Formal, syntactic, semantic and textual features of English shell nouns

---

PHD THESIS

MIGUEL ÁNGEL BENÍTEZ CASTRO

SUPERVISED BY

DR. SALVADOR VALERA HERNÁNDEZ

University of Granada

DR. PAUL THOMPSON

University of Birmingham

PROGRAMA OFICIAL DE DOCTORADO EN LENGUAS, TEXTOS Y CONTEXTOS

UNIVERSIDAD DE GRANADA, 2013

Editor: Editorial de la Universidad de Granada
Autor: Miguel Ángel Benítez Castro
D.L.: GR 866-2014
ISBN: 978-84-9028-920-4

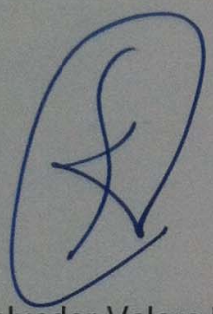
Formal, syntactic, semantic and textual features of English shell nouns

El doctorando **Miguel Ángel Benítez Castro** y los directores de la tesis **Dr. Salvador Valera Hernández (Universidad de Granada)** y **Dr. Paul Thompson (Universidad de Birmingham)** garantizamos, al firmar esta tesis doctoral, que el trabajo ha sido realizado por el doctorando bajo la dirección de los directores de la tesis y hasta donde nuestro conocimiento alcanza, en la realización del trabajo, se han respetado los derechos de otros autores a ser citados, cuando se han utilizado sus resultados o publicaciones.

Granada, 21/10/2013

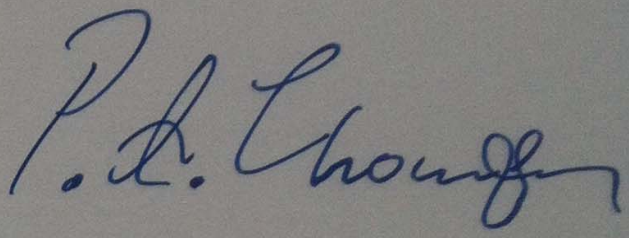
Directores de la Tesis

Fdo.:



Dr. Salvador Valera Hernández

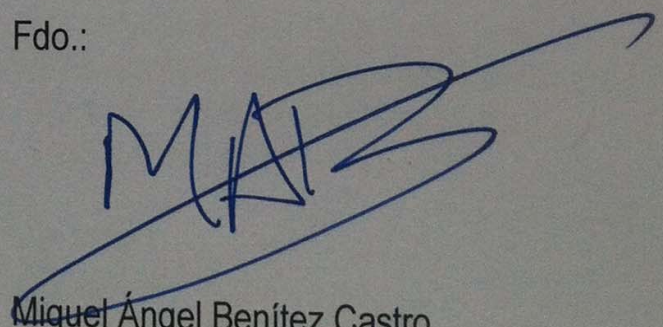
Fdo.:



Dr. Paul Thompson

Doctorando

Fdo.:



Miguel Ángel Benítez Castro

CONTENTS

TABLES

FIGURES

APPENDICES

ACKNOWLEDGEMENTS

First and foremost, I would like to express my deepest gratitude to my supervisors, Dr. Salvador Valera Hernández and Dr. Paul Thompson, for their invaluable guidance and their unwavering support throughout the different stages leading to this PhD thesis. The sound and constructive advice and generous help from these two excellent scholars have enabled me to attain a degree of academic maturity, rigour and autonomy which will stand me in good stead for the future.

Special thanks are also due to the members of my TEAM (in capital letters), among whom I can mention Leanne Bartley, Dr. Ana Díaz Negrillo, and Dr. Jesús Fernández Domínguez. To quote my dear friend Leanne Bartley back in 2009, "I do not think that it's the norm to find a team nor friends so warm". I couldn't agree more with her. I feel immensely proud to have shared countless laughs and chats (research and non-research related) with each and everyone of them over the past four years. Feeling part of such a great team has helped me cope better with the ups and downs linked to the arduous PhD journey.

My dear friend Leanne Bartley deserves special mention. I would like to thank her for being always there, for sharing the good and bad moments with me, for being all ears and acting as my personal psychologist whenever I needed to vent my frustration in one of those 'down' days; in short, for her tremendous support, help and encouragement over the past four years. THANK YOU.

There are not enough words to describe the unconditional and wholehearted support that I have also received from Quica. Thank you for always offering a helping hand, thank you for always bringing a smile when I needed it most, thank you for choosing the right words to cheer me up, thank you for lightening one of those gloomy days.

I am also indebted to the staff at the Centre for Advanced Research in English (University of Birmingham), where I was allowed to spend six of the most fruitful, inspiring and enriching months of my PhD research. My attendance at several MA courses on corpus linguistics and discourse analysis, as well as the regular meetings I had with Dr. Paul Thompson, one of my supervisors, had a profound influence on my PhD research and on the kind of research that I intend to pursue in the future.

Thanks are also due to the Department of English and German Philologies at the University of Granada for the kind welcome and assistance that I received following the transfer of my research grant from the University of Jaén to the University of Granada in January 2012.

I am also extremely thankful to Prof. Dr. Susan Hunston (University of Birmingham, UK) and Prof. Dr. Pavel Štekauer (Pavol Jozef Šafárik

University, Košice, Slovakia), who kindkly agreed to review my thesis and write the two external referee reports required for the award of an International Doctorate.

Lastly, I wish to express my most heartfelt gratitude to my family: thank you for being always there, thank you for your support, thank you for your encouragement, thank you for understanding my constant absence.

This thesis was possible thanks to a four-year research scholarship funded by the Spanish Ministry of Education (Beca de Formación de Profesorado Universitario, FPU, AP2008-030204: August 2009-August 2013).

ABBREVIATIONS

BAWE	*British Academic Written English Corpus*
BNC	*British National Corpus*
BoE	*Bank of English*
CCD	*Collins Cobuild Advanced Learners' English Dictionary*
CDA	Critical Discourse Analysis
CIC	*Cambridge International Corpus*
fpmw	Frequency per million words
ICLE	*International Corpus of Learner English*
L1	First Language
L2	Second Language
LDCE	*Longman Dictionary of Contemporary English*
LOB	*Lancaster-Oslo Bergen Corpus*
LOCNESS	*Louvain Corpus of Native English Essays*
LSWEC	*Longman Spoken and Written English Corpus*
MICASE	*Michigan Corpus of Academic Spoken English*
OALD	*Oxford Advanced Learner's Dictionary*
OCD	*Oxford Collocations Dictionary*
OED	*Oxford English Dictionary*
POS	Part-of-speech
SFG	Systemic-Functional Grammar
SFL	Systemic-Functional Linguistics
T2K-SWAL	*TOEFL 2000 Spoken and Written Academic Language Corpus*
TG	Transformational Grammar

1 INTRODUCTION

1.1 INTRODUCTION

Numerous studies have cast light on the role of abstract nouns such as *objective* or *problem* in the rhetorical organisation of certain broad written genres, namely academic prose and newspaper language. This raises the question of whether the encapsulating and reifying discourse function of these nouns, labelled here as 'shell nouns' (cf. 1.2), is genre-specific or a wide-ranging phenomenon of language.

Closely linked to the description of the role of shell nouns are the corpus employed and the type of analysis adopted, ranging from the fully automated analysis of a large corpus such as the *Bank of English* to the manual analysis of a small corpus. The former relies on a set of predefined queries modelled on the syntactic patterns assumed to prevail in the contextual use of these nouns, i.e. noun-clause and noun-*be*-clause. Manual analyses are often centred on genre-specific corpora. These analyses appear to be mainly concerned with the identification of the rhetorical features of these nouns.

Thus, there seems to be a need for a study that employs a small though well-balanced general corpus in order to provide an all-encompassing linguistic perspective on the use of these units. This thesis aims at the identification of formal, syntactic, semantic and textual features of shell-noun phrases on the basis of a general corpus and of the data extracted manually from it. The goal is to provide an account of shell-noun behaviour based on a manual and contextualised analysis of a representative sample of the English language at large, i.e. the *British National Corpus Sampler*.

This chapter is the introduction to that account. It comprises four sections. Section 1.2 overviews the state of the art of shell nouns. Section 1.3 presents the rationale, method and objectives underpinning this thesis. Lastly, sections 1.4 and 1.5 look at the structure of the thesis and the typographical and terminological conventions used.

1.2 BACKGROUND

Over the past forty years, nouns like *fact*, *idea* or *warning* have received considerable attention from numerous approaches. This is evident in the wide range of terms proposed to highlight various aspects of these units: 'general nouns' (Halliday & Hasan 1976; Mahlberg 2005), 'Vocabulary 3 items' (Winter 1977), 'lexical signalling' (Hoey 1979), 'enumerables' and 'advance labelling' (Tadros 1985; 1994), 'anaphoric nouns' (Francis 1986), 'carrier nouns' (Ivanič 1991), 'advance' and 'retrospective labels' (Francis 1994), 'shell nouns' (Hunston & Francis 2000; Schmid 2000) and 'signalling nouns' (Flowerdew 2003a). Despite differences in the analytical scope of the terms proposed, all of them conceptualise shell-noun behaviour along similar lines. They are summarised below as five frequent properties:

i) Abstraction: Shell-noun description is closely linked to the notion of abstraction, whose identification hinges on formal and semantic criteria. Formal criteria lie at the core of Vendler's (1968) approach to the ontological classification of nominals (Events or Facts), based on their occurrence in fixed structural patterns known as 'containers' (Vendler 1968: 33). One such container is the 'N *is* N' pattern, where the shell noun occurs as a complement of copulative *be*, as in example (1) below.

(1) 'That he died/His death' is **a fact** (Vendler 1968: 73)

Lyons' (1977, II: 442–7) description of abstraction is more semantic in nature, as no connection is established between the ontological status of nominal units and typical sentential patterns. His threefold semantic classification of nouns comprises 'first-order entities' (prototypical concrete nouns; e.g. *table*, *spoon*), 'second-order entities' (nominalised processes and events; e.g. *activity*, *destruction*) and 'third-order entities' (facts and propositions, e.g. *issue*, *problem*). Schmid (2000: 68) argues that, in a scale of prototypicality, third-order entities are the '[...] core of the class of shell nouns'.

ii) Open- and closed-class status, dictionary and context-dependent meaning: Another distinguishing feature of these units concerns the often observed indeterminacy between open-class and closed-class items (e.g. Halliday & Hasan 1976: 275; Winter 1977: 2; Francis 1986: 3; Ivanič 1991: 103). Their open-class status stems from their 'dictionary' meaning, while their closed-class status relates to their 'specific' or context-dependent meaning (Ivanič 1991: 95). Shell

nouns are thus argued to carry an inherent 'dictionary' meaning, remaining constant, and one which, like pronouns, varies depending on the context where the noun appears. Such a variable meaning may be explained by reference to Lyons' (1977, II: 668) concept of 'impure textual deixis' and Fraurud's (1992: 4) notion of 'situation reference'. Both terms capture the idea that, in some cases of anaphoric reference, the pronoun is not co-referential with a first-order entity, but with a longer discourse segment. Example (2) illustrates the more concrete or 'object' (Fraurud 1992: 3) type of anaphora, whilst (3a) and (3b) illustrate the more impure or situation-dependent anaphora. Examples (3a) and (3b) contain the same information, the difference lying in the use of a referential demonstrative pronoun in (3b) and a referential shell-noun phrase in (3a). The use of a shell noun in (3a) endows the writer with a means of evaluating the underlined stretch of discourse. It is this 'characterisation' potential of shell nouns (Schmid 2000: 13) that distinguishes shell nouns from pronouns. Thus, whilst similar to pronouns in their context-specific meaning, the evaluative potential of shell-nouns is lexical or open-ended in nature.

(2) 'The boys went home. They were tired' (Fraurud 1992: 3)

(3) (a)'[...] foul-smelling algae, the product of exceptionally high temperatures and high levels of sea pollution, which led to a huge drop in bookings. Fortunately this problem does not seem to have recurred this summer' (Schmid 2000: 124)
(b)'[...] foul-smelling algae, the product of exceptionally high temperatures and high levels of sea pollution, which led to a huge drop in bookings. Fortunately this does not seem to have recurred this summer'

iii) Long antecedent: In the literature, shell-noun behaviour is generally associated with the encapsulation of clausal, sentential or extended discourse segments, as in (4) below. Thus, in principle, single-word encapsulation is no guarantee of '[...] shell-nounhood' (Schmid 2000: 13). Still, some references (Ivanič 1991: 109; Flowerdew 2003a: 336; Gray 2010: 179) contemplate the possibility of noun-phrase antecedents for shell-noun phrases, as in (5):

(4) '[...] In reply to that question a golfing colleague of mine offered two reasons. The first was that beginners usually start with handed-down clubs, which are usually right-handed. The second was that, for technical reasons, left-handed individuals make good right-handed golfers.' (Francis 1994: 84)

(5) 'It is interesting to read about the items electors mentioned as having, in their view, specially affected the election. […] rash Labour promises – cost of new pension scheme – bribery of electorate [...] strikes' (Ivanič 1991: 109)

iv) Anaphora and intersentential realisation: Anaphoric uses, as in (3), occupy a prominent position in the literature (e.g. Halliday & Hasan 1976; Francis 1986; Conte 1996; Charles 2003; Moreno 2004; Gray 2010). This is inspired by the numerous references on (mainly) pronominal anaphora resolution (e.g. Chomsky 1981; Fox 1987; Fraurud 1988; Asher 1993). Other types of encapsulation are less conspicuous in shell-noun research (e.g. Winter 1977, Hunston & Francis 2000 on cataphoric uses; Francis 1994, Schmid 2000 on anaphoric and cataphoric uses; Ivanič 1991, Flowerdew 2003a on endophoric and exophoric uses).

Given the emphasis on retrospective uses, it stands to reason that most research attention is also on the intersentential realisation of shell-noun instances, as in (3) and (6). Comparatively few studies mention reference either within the boundaries of the sentence (e.g. Hunston & Francis 2000; Biber 2006; Charles 2007, as in (7)) or both within and outside these boundaries (e.g. Ivanič 1991; Winter 1992; Schmid 2000; Flowerdew 2003a; Caldwell 2009).

(6) '[…] the Soviet Union has "shot its bolt", and that only the unreconstructed Cold Warriors are losing any sleep about the Russian menace. James Reston has readily and complacently echoed this assessment in his criticism of the Reagan équipe' (Francis 1986: 27)

(7) '[...] the best the White House has been able to conjure up is the tired accusation that they are liberals and lackeys of special interest groups.' (Schmid 2000: 135)

v) Deictically specific shell-noun phrases and the N-cl, N-*be*-cl patterns: As regards formal structure, shell-noun behaviour is typically linked to definite and demonstrative instances, as in (3), (5), (6) and (7). Few references allow indefinite cases (e.g. Ivanič 1991: 111; Partington 1998: 92–3; Aktas & Cortes 2008: 10), as in (8):

(8) 'In a move to tighten control of a far-reaching empire and to improve the group's own image, Maurice and Charles Saatchi, credited with building up the company, have stepped down from the day-to-day running of the group' (Partington 1998: 94)

> The literature also places considerable emphasis on two post-nuclear structures, i.e. *that* and *to*-infinitive 'noun complement clauses' (Biber et al. 1999: 645), as shown in N-cl and N-*be*-cl patterns, as in (9) and (10). Prepositional phrases, as in (11), tend to be disregarded in most shell-noun descriptions available. Exceptions include Winter (1992: 157), Flowerdew (2003a: 337) and Caldwell (2009: 176).

(9) 'The Association will give a warning that poll tax bills in some Conservative districts will exceed government guidelines […]' (Schmid 2000: 135)
(10) 'The first action was to place the vessel under cover and remove the deckhouse' (Schmid 2000: 263)
(11) 'its function of providing mechanical strength' (Flowerdew 2003a: 337)

Most generalisations about shell-noun use draw on findings from small and genre-specific corpora. Academic discourse is paramount there, as most research is geared to the analysis of academic sub-genres such as textbooks, research articles or essays (e.g. Tadros 1985; Flowerdew 2003a; Moreno 2004; Charles 2007). Newspaper discourse is also central to shell-noun descriptions, especially where data retrieval rests on large general English corpora such as the *Bank of English* (henceforth, *BoE*; e.g. Hunston & Francis 2000, Schmid 2000). The ever-growing *BoE* offers size (reaching 225 million words when Schmid 2000 conducted his analysis), but fails to provide a balanced representation of genres, because 70% of the corpus is accounted for by journalistic prose. Use of other general English corpora like the *Lancaster-Oslo Bergen Corpus* (henceforth, *LOB*) or the *British National Corpus* (henceforth, *BNC*) offers a better balance. However, when they are used for this topic, the research focus is limited to qualitative findings (e.g. Ivanič 1991) or to quantitative mode-related data (i.e. spoken vs. written; e.g. Aijmer 2007, Yamasaki 2008).

Data analysis in the literature is often based on retrieval of predefined patterns, particularly N-cl and N-*be*-cl, which, as stated above, are prototypically associated with shell-noun use (e.g. Francis 1993; Hunston & Francis 2000; Schmid 2000 and 2007). Manual approaches to the analysis of corpus data identify patterns other than noun complement clauses (e.g. N-*of*, N-*which*), but in those cases the scope is restricted to particular sub-genres and nouns (e.g. Francis 1986 on the monthly journal *Encounter*, Flowerdew 2003a on biology textbooks and lectures; Hoey 1993 on the noun *reason*; Lorés 2006 on *thing* and *idea*).

1.3 THE THESIS

1.3.1 *Rationale*

The overview in 1.2 presents shell nouns as a widely researched area, but also as one with certain gaps:

i) Considerable attention is devoted to academic and journalistic prose, and little to other genres.
ii) Most research presents a partial description of shell-noun use, failing to offer a systematic and thorough account of formal, syntactic, semantic and textual features.
iii) Data analysis is often limited to specific patterns (e.g. N-cl and N-*be*-cl) and encapsulating relations (e.g. anaphoric uses).
iv) Automated analytical approaches limited to certain patterns prevail over manual ones.
v) Small-scale manual analyses only use genre-specific corpora.

With this in mind, there arises a need for a study where:

i) Shell-noun use is described on the basis of a small but representative sample of the English language at large.
ii) Formal, syntactic, semantic and textual levels of linguistic analysis are investigated.
iii) All patterns and uses are included, thereby foregrounding a manual corpus-driven approach to data analysis.

1.3.2 *Method*

This thesis uses a fully manual and corpus-driven method for the analysis of shell-noun data. Complete automaticity allows quick processing of large amounts of data from a large corpus. On the downside, the research scope is limited by the restrictive nature of predefined automated corpus queries, as this may force the exclusion of certain linguistic features. In line with Sinclair (2004: 23), this study advocates an open approach to data: one which does not impose preconceptions (e.g. shell nouns and noun complement clauses), and one which, in short, 'trust[s] the text'. As in Mahlberg's (2005: 31–8) 'corpus theoretical approach' to general nouns (e.g. *thing*, *fact*, *people*, *world*), only two assumptions about shell-noun use guide the analysis: one such assumption relates shell nouns to semantically unspecific second- and third-order entities (e.g. *action*, *event*, *idea*, *point*), and the other lays emphasis on the contextualised interpretation of these

units. The latter implies that, regardless of the form of the noun phrase (definite or indefinite), the interpretation of a shell noun is always influenced by the surrounding context. In addition to Sinclair (2004) and Mahlberg (2005), this thesis is also underpinned by Hoey's (2005) theory of lexical priming, as shell nouns are not explored from just one linguistic perspective, but from a range of them: formal/structural, syntactic, semantic and textual.

The thesis uses a small corpus of contemporary written and spoken British English, the *BNC Sampler.* Containing around 2 million words, the *Sampler*, a 2% sample of the entire *BNC*, is smaller but considerably better balanced than the *BNC*. Unlike the *BNC*, where written and spoken genres occur in a 90%-10% distribution, written and spoken language in the *Sampler* are evenly distributed (50%-50%).

The study sample comprises 1447 concordance lines corresponding to 60 lemmas. The units are extracted from the top, middle and bottom frequency ranges of a list of 922 shell-like units. The list brings together 15 lists of shell-like units found in the literature, the most extensive being Schmid's (2000) with 670 units. A random set of 40 concordances (20 for the singular and 20 for the plural) is analysed for each of the 60 lemmas.

Each concordance is tagged according to nine variables, reflecting the multifaceted approach applied here:

i) Genre of the text (e.g. academic prose, conversation, etc.).
ii) Semantic or experiential structure of the noun phrase (e.g. Deictic, Epithet, etc.).
iii) Formal structure (e.g. definite article, prepositional phrase, etc.).
iv) Syntactic function (e.g. direct object, subject).
v) Participant type (e.g. Goal, Attribute).
vi) Theme-Rheme.
vii) Direction of encapsulation (e.g. intersentential anaphora, intrasentential cataphora, etc.).
viii) Antecedent (e.g. extended discourse, sentence, etc.).
ix) Semantic type of shell noun (e.g. Factual, Mental, etc.).

1.3.3 *Objectives*

This thesis has three general and four specific objectives. The general objectives are:

i) To examine the lexico-grammatical, syntactico-semantic and textual criteria involved in the use of shell nouns from a synchronic perspective.

ii) To make progress on the identification of how different criteria interrelate with each other.
iii) To investigate the connection between the criteria and shell-noun meanings.

The specific objectives are:

i) To retrieve and analyse manually a sample of 1447 concordances for 60 lemmas extracted from the *BNC Sampler*.
ii) To identify criteria for the distinction between shell and non-shell uses.
iii) To tag shell-noun instances according to nine lexico-grammatical, syntactico-semantic and textual variables.
iv) To assess the extent to which these units contribute to the organisation of different types of spoken and written discourse.

1.4 CONTENTS

This thesis is divided into seven chapters, including this introduction (Chapter 1). Chapter 2 elaborates on the theoretical background to shell-noun description, with special emphasis on definitions, identification criteria and genre-related uses. Chapter 3 lays the theoretical foundations for the analytical approach of this thesis. It casts light on the research gaps that guide the choice of the corpus, the variables and the method used for the analysis of the study sample. Chapter 4 explains the decisions on the study corpus and sample, and describes the manual corpus theoretical approach of this thesis. It also distinguishes between shell and non-shell uses. Chapter 5 details the principles underlying the nine analytical variables used. This chapter combines methodological procedures with qualitative observations on some of the variables. Chapter 6 presents and discusses quantitative results for the variables both in isolation and in relation to one another. Chapter 7 summarises the contents of the thesis and proposes possible research avenues.

1.5 TYPOGRAPHICAL AND TERMINOLOGICAL CONVENTIONS

This thesis uses the following typographical conventions:

i) SMALL CAPITALS are used for section headings.
ii) Initial capitals are used with Systemic-Functional terms (e.g. Deictic, Actor, Theme, as in Halliday & Matthiessen 2004) and with

semantic features of shell-noun uses (e.g. Factual, Mental, Agentive, Dynamic).

iii) *Italics* are used for examples in the running text, names of corpora, software packages, dictionaries, projects and their acronyms.

iv) Single quotations marks (‘ ’) are used for the meanings and senses of words and sentences, quotations and direct speech, as well as for the metalinguistic use of terms (e.g. the term ‘shell noun’, the concept of ‘factuality’) or when a term is first presented and defined.

v) Double quotation marks (“ ”) are used for quotations within quotations.

vi) For the sake of consistency, in all the examples separated from the running text, the shell noun is indicated in **boldface** and its lexical realisation is underlined. Single quotation marks (‘ ’) are used for examples obtained from the literature. Both literature and study examples are followed by their source. In the latter case, source codes include the name of the corpus, the corpus text which the example is extracted from and its genre category (e.g. *BNC Sampler*: CF6, W:newsp:other:report).

vii) Citation of bibliographical references in the main text and in the section of References is in accordance with the style sheet of the journal *English Language and Linguistics*. Following its style sheet, if more than one article is cited from a single edited volume, a short reference to the volume appears in the article entries and the full detals of the volume apear in a separate entry.

This thesis uses the following terminological conventions:

i) In line with Schmid (2000; see 3.2.5.2), ‘shell noun’, ‘shell-noun phrase’, ‘shell-noun use’ and ‘shell use’ are used interchangeably.

ii) No distinction is made between ‘co-text’ and ‘context’. Both terms apply to any discourse segment enabling the interpretation of a particular shell noun (cf., however, Halliday 1978: 133 and Brown & Yule 1983: 46–7, where the terms are distinguished).

iii) Following Stirling & Huddleston (2002), ‘antecedent’ is used to refer to anaphoric and cataphoric encapsulation (see 3.2.2).

iv) ‘Encapsulation’ is used instead of ‘reference’ to describe the link between shell noun and antecedent (Sinclair 1993, 2004; see 3.2.2).

2 LITERATURE REVIEW: SHELL-NOUN DEFINITIONS AND GENRE VARIATION

2.1 Introduction

This chapter is intended as a summary of the bibliographical background to the study of shell nouns. With this goal in mind, it has been structured around two major sections: definitions and identification criteria (section 2.2), and shell nouns and genre variation (section 2.3).

The first section is an overview of the characterisation of these units in the literature. Specifically, it deals with the attention given to shell nouns in general grammars of English (2.2.1) and with research on the definition of their formal and semantico-pragmatic boundaries (2.2.2). All in all, section 2.2 is crucial to the study conducted in this thesis, inasmuch as it casts light on the similarities and differences evident in many of the definitions reported in the literature, while also emphasising the need for a multifaceted approach that integrates the issues raised in such definitions.

Section 2.3 reviews references where insights are offered about possible variations in the use of these items in different contexts. The section comprises research on written discourse (2.3.1), written and spoken discourse (2.3.2) and spoken discourse (2.3.3). Written academic discourse (2.3.1.2) proves to be the genre that has so far received most attention in the literature. This appears to underline the need for a study where shell nouns are described on the basis of a wide range of genres of the English language.

2.2 Shell nouns: Definitions and identification criteria

The class of nouns which, in line with Hunston & Francis (2000) and Schmid (2000), shall be referred to here as 'shell nouns' has been the subject of considerable academic debate over the past four decades. Numerous terms, definitions and identification criteria have been suggested to account for '[t]he property of shell-nounhood' (Schmid 2000: 13). Despite the manifold approaches applied to their description, widespread agreement exists on the scope of such a property, found primarily among semantically

incomplete abstract nouns (e.g. *fact*, *assumption*, *possibility*) requiring information from the surrounding co-text. Example (12) below is a typical instance of the use of these nouns, with a following *that*-clause specifying the meaning of the head noun *theory.* This section starts by reviewing the definitions and identification criteria offered by general grammars of English (2.2.1). The focus will then be directed to research specifically devoted to shell-like units (2.2.2).

(12) 'The discovery of twin pandas in Sichuan province has refuted the theory that only one of any pair of giant panda twins could survive' (Hunston & Francis 2000: 186)

2.2.1 *General grammars*

2.2.1.1 *Early descriptive grammars*

This section examines the treatment of shell nouns in the main grammars of English dating back to the late 19th and early 20th centuries (Sweet 1891–8, I; Poutsma 1904–29, I; Jespersen 1909–49, III and VII; Kruisinga 1931–2, II and III; Curme 1947). The earliest of these grammars, Sweet (1891–8), defines a new era in the study of English, where the focus is shifted from the 18th century prescriptivism to an increasing interest in the systematic and scientific description of language. This said, prescriptivist attitudes are still evident in the reliance of these grammars on present and past literary language, which is considered as the linguistic norm. Diachronic (or historical) insights, therefore, play a paramount role in these references, especially in the provision of examples. Kruisinga (1931–2) and Zandvoort (1972) are the only authors who decide to eschew historical considerations by focusing only on contemporary language.

Sweet (1891–8) offers only a cursory treatment of shell nouns. His grammar describes abstract nouns in terms of their relation to attributes (e.g. *redness*, *tiredness*) and phenomena (e.g. *action*, *reading*) (Sweet 1891, I: 61). Formally, they often appear in the shape of morphological derivations from verbs and adjectives, as in *converse>conversation* or *strong>strength*. In addition to his definition of abstract nouns, especially relevant to the unspecific meaning of shell nouns is Sweet's (1891, I: 155) definition of 'sentence' as '[…] a word or group of words whose form makes us expect it to express a full meaning'. The use of the verb *expect* is explained in terms of the central importance placed on context for the more or less complete interpretation of the meaning of a sentence. Sweet's contribution to the structural description of shell nouns lies in his discussion of 'noun-clause(s)' (Sweet 1891, I: 171). These are claimed to have the ability to perform

several functions (subject, predicate, direct object and apposition), only two of which represent typical shell-noun patterns (predicate noun-clause, as in (13) and apposition, as in (14)). Poutsma (1905, I: 404), where (as in Sweet 1891, I) only passing mention is made of shell nouns, also presents these two patterns. Poutsma (1905, I: 404), however, differs from Sweet (1891, I) in the use of the term 'subordinate statement' instead of 'noun-clause' and 'attributive adnominal adjunct' instead of 'apposition'.

(13) '[…] this is what I mean – my opinion is that he is mistaken' (Sweet 1891, I: 171)
(14) '[…] the wish that he may succeed is very general' (Sweet 1891, I: 171)

Jespersen's (1909–49) grammar gives a more detailed account of shell nouns than either of the aforementioned references. These nouns are discussed in the context of Jespersen's (1927, III: 23–36) 'content-clauses', a term he coined to refer to Sweet's (1891, I) 'noun-clauses' and Poutsma's (1905, I) 'subordinate statements'. The choice of the term lies in Jespersen's (1927, III: 24) reluctance to use either 'noun-clause' or '*that*-clause', on the grounds of the lack of formal identity between nouns and clauses and the occurrence of many instances where *that* is deleted ('I believe *he is ill*'). Content-clauses perform functions typical of nouns, such as subject and direct object. There are instances, however, where the abstract nouns *fact* or *circumstance* are added to the *that*-clause to make it sound more natural in subject or object position. In Jespersen's (1927, III: 24, 26) words, these items (as in (15) and (16)) '[…] prop up the clause', a use which is '[…] especially frequent in modern scientific prose'. Although *fact* and *circumstance* are the default options for this prop-up function, mention is also made of other non-derived and deverbal nouns similarly followed by a *that*-clause (e.g. *belief*, *hope*, *wish*, *desire*, *idea*, *notion*, *sentiment*, *doctrine*). As regards the phrasal status of the *that*-clause, Jespersen (1927, III: 27) states that two analyses might be proposed, one where the clause is treated as an object and another one where the clause is treated as an apposition. He ends up adopting the latter analysis, in line with Sweet (1891, I).

(15) 'But the fact that it was thought necessary to disguise these exactions under the names of benevolence and loans sufficiently proves that the authority of the great constitutional rule was recognised' (Jespersen 1927, III: 24)
(16) 'I'm afraid you overlook the circumstance that you've been requested to leave my house' (Jespersen 1927, III: 26)

Interestingly enough, Jespersen's (1949, VII) last volume of his grammar contains a passing reference to a use of shell nouns that corpus linguistics will later on reveal as typical of informal spoken discourse (cf. for example Schmid 2001 and Aijmer 2007 in 2.3.2.1 and 2.3.3). Under the term 'prosiopesis', Jespersen (1949, VII: 415) includes instances where the definite article is dropped from certain nouns occurring in subject position (e.g. *fact*, *question*, *thing*, *trouble*, *truth*). Example (17) illustrates what Schmid (2000: 329–37) will later on describe as the focusing function of shell nouns (see 2.2.2.2):

(17) '[…] But look here – question is, are our characters good enough just now[…]?' (Jespersen 1949, VII: 415)

Kruisinga's (1931–2) contribution to shell-noun description is apparent in different sections of two of his three volumes on English syntax. Following Sweet (1891, I), abstract nouns are described as those expressing attributes or phenomena (e.g. *grandeur*, *arrival*) (Kruisinga 1932, II: 24). Together with nouns of materials (e.g. *iron*, *gold*), abstract items are said to fall into the category of 'non-class-nouns', i.e. nouns which do not accept the plural inflection. The division between class and non-class nouns is not clear-cut, as there are many instances where the same word is said to occur as a countable class noun in some sentences but as an uncountable abstract item in others (see (18a) and (18b)).

(18) (a)'Legend […] has it that her body became so miraculously heavy that they could not lift her' [uncountable abstract noun] (Kruisinga 1932, II: 25)
(b)'[…] celestial ministrants had wrapped him in a white robe of supernatural beauty and had led him into the Divine Presence to receive the granting of his petition of the indulgence. How far the legend is fictitious need not be debated' [class-noun] (Kruisinga 1932, II: 25)

The intersentential reference of a shell noun in (18b) is further illustrated in the context of Kruisinga's (1932, II: 202–15, 239–41) distinction between deictic and anaphoric uses of demonstrative pronouns (and determiners) and the definite article. Deictic uses (as in (19)) are frequently anticipatory or cataphoric, while anaphoric uses, as in (18b), are retrospective. The noun *thing* is so common as part of a deictic expression that Kruisinga (1932, II: 329) argues that it '[...] may be looked upon as a pronoun', its only function being to allow an adjective to appear in a nominal context (as in (20)) (cf. Jespersen's 1927, III prop-up function).

(19) 'We don't understand girls, but we ask **this question** of those who do: Is it possible that Miss Sally was impressed by the splendid arm with the name tattooed on it?' (Kruisinga 1932, II: 206)

(20) '**The most remarkable thing about the Lapp** is that he can live in Lapland […]' (Kruisinga 1932, II: 329)

Kruisinga's (1932, III: 364–84) third volume on syntax discusses, among other things, five types of subordinate clauses, i.e. subject clauses, object clauses, predicate clauses, apposition clauses and attributive clauses. Only object, predicate and apposition clauses are relevant to shell-noun description. Unlike the above-mentioned grammarians, Kruisinga (1932, III) distinguishes between nouns introducing object clauses and those introducing apposition clauses. The former are nouns derived from verbs and adjectives followed by *that*-clauses (e.g. *discovery*, *fear*, *certainty*, *confidence* < *discover that*, *fear that*, *certain that*, *confident that*) (Kruisinga 1932, III: 369), while the latter correspond to non-derived nouns (e.g. *fact*, *news*, *idea*, *thing*) (Kruisinga 1932, III: 374). As regards predicate clauses, Kruisinga (1932, III: 372) presents examples like (20) above, where the *that*-clause acts as subject complement. Mention is also made of instances like (21), where the conjunction is deleted and replaced with a comma. This is another manifestation of Jespersen's (1949, VII: 415) 'prosiopesis' (see (17) above).

(21) '**The reason** was, Shirley's head ran on other things than money and position' (Kruisinga 1932, III: 373)

It is worth noting that Kruisinga's (1932, II and III) framework is followed closely by Zandvoort (1972). The only difference with the former lies in the adoption of Jespersen's (1927, VII) 'content-clauses' to refer to Kruisinga's (1932, III) 'object' and 'apposition' clauses (Zandvoort 1972: 222).

This section concludes with Curme (1947), where nothing new is added to what has been said so far about shell nouns. These units feature in his discussion of adjective clauses, which fall into attributive substantive clauses and attributive adjective clauses (i.e. relative clauses) (Curme 1947: 162–3). Shell nouns occur in the former, where the *that*-clause has '[…] the force of an attributive appositive noun, or substantive' (Curme 1947: 162).

2.2.1.2 *Contemporary descriptive grammars: Focus on form*

This section is concerned with the attention given to shell nouns in more modern grammatical descriptions of English (Chomsky 1970; Quirk et al. 1985; Greenbaum 1996; Huddleston & Pullum 2002). With the exception of Chomsky (1970), where there is a definite preference for made-up

examples, the other references rely (to a greater or lesser degree) on examples drawn from general English corpora. The corpora, however, are used merely as one of several sources of examples (other sources being the linguist's or other speakers' intuitions, for example), and they inform rather than drive the description. It is also important to mention that the description in these references is primarily formal in that little or no emphasis is laid on the connection between form and the contextual use of language, as is shown in 2.2.1.3.

Chomsky (1970) is not a book but a paper in a monograph on various aspects of English Transformational Grammar (henceforth, TG). The paper looks at two types of nominalisations: gerundive nominals and derived nominals. Chomsky's (1970: 188) main argument is that the transformations leading to the former are less restrictive than those leading to the latter. As such, based on (22a), (22b) (a gerundive nominal) is grammatical, while (22c) (a derived nominal) is ungrammatical. Example (22c) would only prove grammatical in a different construction, as in (23).

(22) (a)'John is certain (likely) to win the prize' (Chomsky 1970: 188)
(b)'John's being certain (likely) to win the prize' (Chomsky 1970: 188)
(c)'*John's certainty (likelihood) to win the prize' (Chomsky 1970: 188)

(23) 'John's certainty that Bill will win the prize' (Chomsky 1970: 189)

Grammaticality or the lack thereof in these examples is explained on the basis of the 'lexicalist hypothesis' (Chomsky 1970: 191–2). This hypothesis suggests that the acceptability of noun complement clauses depends on whether a complement is inherent in the lexical item undergoing nominalisation. For example, the deep structure for (22a) is an extraposed construction (i.e. *it is certain that John will win the prize*) where *to win the prize* is not a complement of the adjective *certain*. As such, a nominalisation where the *to*-infinitive clause is a complement of *certainty* is not acceptable (as in (22c)). Example (23) is regarded as acceptable on the grounds that *certainty* originates from a use of *certain* where the adjective is followed by a *that*-complement clause (i.e. *John is certain that Bill will win the prize*).

Several other instances of shell-noun patterns in the article lend further support to the lexicalist hypothesis (Chomsky 1970: 197–8). For example, the ungrammaticality of (24a) is due to the misinterpretation of the adjoining *that*-clause as an optional modifier rather than a complement. Hence, in (24b) and (24c), the semantic gap in the noun *excuse* is entirely filled by the *that*-complement clause, which implies that, should another *that*-clause be added, the example would be semantically and grammatically unacceptable, as in (24a).

(24) (a)'*The excuse that John had left was that Bill should stay' (Chomsky 1970: 198)
(b)'The excuse that John had left' (Chomsky 1970: 197)
(c)'The excuse was that John had left' (Chomsky 1970: 197)

More recently, Aarts (2001: 111, 122, 137) uses Chomskyan theory to offer a similar explanation for the analysis of intraclausal shell-noun patterns. Postnominal clauses, as in (23) or (24b), are argued to function as subcategorised complements. The concept of 'subcategorisation' explains the downgrading that clauses undergo from sentential to phrasal constituents (Aarts 2001: 122). Overall, the relevance of TG to shell-noun description lies in the introduction of the term 'complement' to describe all instances of head-clause shell-noun patterns (cf. 2.2.1.1).

In Quirk et al'.s (1985) grammar, shell nouns are accounted for in the course of their discussion of nominal postmodification (1985: 1260–2, 1272–4). *That*-clauses are claimed to occur as appositives of general abstract nouns such as *fact*, *idea* or *proposition.* Their appositional status is explained in terms of the semantic identity resulting from the possible introduction of copular *be* between noun and clause (e.g. *the belief is that*, *the news was that*, etc.). The noun is often a singular deverbal or deadjectival nominalisation preceded by the definite article (e.g. *believe>the belief that*, *possible>the possibility that*). However, (25) and (26) show two exceptions to this rule, (25) containing an indefinite article and (26) a plural head.

(25) 'A message that he would be late arrived by special delivery' (Quirk et al. 1985: 1261)
(26) 'The reason probably lies in the facts that the Intelligence Service is rather despised, that the individual members change rapidly and are therefore inexperienced […]' (Quirk et al. 1985: 1261)

In addition to appositive *that*-clauses, the description includes apposition realised by *to*-infinitive and *of* V-*ing*. In this respect, a distinction is made between nouns followed by the former appositive only (e.g. *agreement*, *proposal*, *decision*, as in (27)), nouns followed by the latter only (e.g. *risk*, *prospect*, *hope*, as in (28)) and nouns followed by either (e.g. *change*, *possibility*, *intention*, as in (29)) (Quirk et al. 1985: 1272–4). The preference for one pattern over the other is claimed to rest on modality. As such, nouns implying human control over events (e.g. *invitation*, *willingness*, *refusal*) prefer *to*-infinitive, while nouns involving human judgement prefer *of V-ing* (Quirk et al. 1985: 1272–3).

(27) 'Anna's willingness to do […] the job' (Quirk et al. 1985: 1273)

(28) 'There is actually no hope of (them/their) winning the war' (Quirk et al. 1985: 1273)
(29) 'Their chance to go/of going abroad was lost' (Quirk et al. 1985: 1272)

The only reference made in this grammar to the intersentential use of shell nouns is a passing comment on 'general hypernyms' such as *problem*, *situation*, *event*, *subject* and *thing* (Quirk et al. 1985: 1442). These are argued to resemble pronouns and other substitutes in their ability to endow a text with cohesion. An example is given below, with *situation* referring anaphorically to the previous sentence:

(30) 'Thousands were out of work; there was hunger, anger, and unrest. The situation required careful handling' (Quirk et al. 1985: 1442)

Greenbaum's (1996) grammar does not differ much from Quirk et al'.s (1985) in the attention given to shell nouns (1996: 219, 330–1, 357). A seeming contradiction, however, is observed in the use of both 'appositive postmodifier' and 'complement' to refer to post-nuclear *that*-clauses. From Greenbaum's (1996: 357) explanation, it may be inferred that the complement analysis applies only to deverbal nouns in reported speech sentences (as in (31)), while the appositive analysis would apply to any other instances of the construction (as in (32)).

(31) 'This reinforces the earlier statement, that man is blind to what he cannot see' (Greenbaum 1996: 357)
(32) 'It's really shorthand for the view that well-being depends on more than the absence of disease' (Greenbaum 1996: 219)

The last grammar in this section, Huddleston & Pullum (2002), follows the others in the emphasis laid on the complement clause pattern. Drawing on Jespersen (1927, VII: 23–36), the term 'content clause' is used to refer to the subordinate clause in this construction (Huddleston 2002: 1016–17). TG also lies behind the choice of 'complement' to label the function of such a clause.

Payne & Huddleston (2002: 448) and Huddleston (2002: 1017) argue against the treatment of these clauses as appositives, on the grounds that in a sentence like (33) the omission of the head noun alters the meaning of the construction. *That they cheated was quite outrageous* would imply that the cheating did take place. This change of meaning would not apply to typical instances of nominal appositives, as in *my friend Mary is sad*, where meaning would remain constant should one of the two elements (i.e. *my friend* or *Mary*) be omitted. Further evidence against the appositive analysis appears in such sentences as (34), where the introduction of *be* between

the noun and the clause would result in an ungrammatical sequence. It is worth mentioning here that, in a data-driven study, Francis (1993: 151) finds substantial corpus evidence which similarly questions the appositive analysis. She observes that for many nouns of feeling (e.g. *annoyance*, *astonishment*, *anxiety*, *concern*), the *that*-clause does not express the identity expected from appositive noun phrases, but the cause of the feeling (as in (35)).

(33) 'The suggestion that they cheated was quite outrageous' (Payne & Huddleston 2002: 448)
(34) 'Their insistence that the meetings should be held at lunch-time angered the staff' (Huddleston 2002: 1017)
(35) '[...] He felt a surge of happiness that she could not hurt him any more' (Francis 1993: 151)

2.2.1.3 *Contemporary descriptive grammars: Focus on function*

The grammars in this section are inspired by a common goal to describe language based on the functions to which it is put in different contexts. Lexis and grammatical structure are no longer treated as separate components, but as a single system of choices: the lexicogrammar. Such a system is not immutable; it varies according to the purpose and function for which people use language. The first half of this section (2.2.1.3.1) looks at the treatment of shell nouns in Hallidayan Systemic-Functional Grammar (henceforth, SFG or SF, for Systemic-Functional). The second half (2.2.1.3.2) discusses corpus-driven grammars of English.

As was the case in 2.2.1.2, in SFG (2.2.1.3.1), computerised corpora (if used) are only sources of examples, but they do not drive the description. Corpus-driven grammars (2.2.1.3.2), however, base linguistic description on '[...] the recurrent patterns and the frequency distributions [...]' (Tognini-Bonelli 2001: 87) emerging from the analysis of computerised corpora.

2.2.1.3.1 *Systemic-Functional Grammar*

Halliday & Matthiessen (2004: 467–80) explore shell nouns in the context of embedded clauses. They argue that, just as a verb can project a clause (e.g. *they assert that…*), so can a noun be said to be followed by a projected qualifier (e.g. *their assertion that…*). Such projecting nouns are essential for '[…] the representation of arguments […] in newspaper reports and scientific discourse' (Halliday & Matthiessen 2004: 468). These nouns fall into two main groups, those derived from projecting verbal and mental processes or verbs (e.g. *state that>statement that*, *know that>knowledge that*), and 'fact nouns', i.e. those which are either non-derived (e.g. *fact*,

case, *principle*) or derived from words other than verbal and mental processes (e.g. *possible>possibility*, *likely>likelihood*, *certain>certainty*; Halliday & Matthiessen 2004: 469). Encompassing this twofold distinction is a more general one between 'propositions' and 'proposals'. Propositions contain 'stating nouns', followed by projected *that*-clauses or *of* V-*ing* (e.g. *claim*, *argument*, *proposition*, *assumption*), as well as 'questioning nouns', followed by *if-/whether-/wh*-clauses (e.g. *question*, *query*, *doubt*). Proposals, by contrast, include 'offering nouns' (e.g. *offer*, *suggestion*, *proposal*) and 'commanding nouns' (e.g. *order*, *instruction*, *demand*), both followed mainly by *to*-infinitive clauses.

Halliday & Matthiessen (2004: 470–80) devote considerable space to a discussion of fact nouns and clauses. These are claimed to prevail in relational (i.e. copulative) environments (Halliday & Matthiessen 2004: 472–5). In this respect, only nouns with overt evaluative nuances typically occur as Attributes (e.g. *pity*, *shame*, *relief*, *tragedy*, *nuisance*). Neutral fact nouns such as *fact* or *principle* may also appear as Attributes, but they are often premodified by adjectives (as in (36)). In identifying relational clauses, fact nouns function as subjects assessing the information in the *that*-clause. This assessment is often highlighted through evaluative adjectives, as shown in (37).

(36) 'Until 1940 it was an observable fact that there were composers whose music was highly prized in some countries and entirely neglected by their neighbours […]' (Halliday & Matthiessen 2004: 473)

(37) 'The thorniest problem for next week's conference is to settle the relationships between them and the rest of the country' (Halliday & Matthiessen 2004: 473)

It should be noted that, whilst their account of shell nouns is restricted to intrasentential patterns, Halliday & Matthiessen (2004: 468, 471) make a passing reference to intersentential anaphoric uses, as in (38). These are claimed to contribute to the cohesion of discourse.

(38) 'The Labour Party opposed Thor missiles, because, he said, they were out of date and vulnerable and would attract enemy action. That argument did not apply to the Polaris submarine' (Halliday & Matthiessen 2004: 468)

Particularly relevant to SFG explorations of shell nouns is also their close association with 'grammatical metaphors' (Halliday & Matthiessen 2004: 586–658), the Hallidayan term used to account for instances of nominalisation. Martin (1992) offers a classification of grammatical metaphors that is significantly related to shell nouns. Three types of

metaphor are included in his description: experiential, logical and textual metaphors. 'Experiential metaphors' (Martin 1992: 409–11) correspond to nominalisations of actions (e.g. *use something>the use of something*) and qualities (e.g. *inadequate>inadequacy*). 'Logical metaphors' (Martin 1992: 408–9) are concerned with the representation of conjunctions (e.g. *because*) and conjuncts (e.g. *therefore*) as prepositions (e.g. *due to*), verbs (e.g. *cause*) or nouns (e.g. *reason*). Logical metaphors are thus illustrative of Winter's (1977) Vocabulary 3 items (see 2.2.2.1.2), as they foster '[…] reasoning within rather than between clauses […]' (Schleppegrell 2004: 177). To provide an example, in (39) below, the logical metaphor *cause* conveys the meaning of the complex preposition *because of* and the subordinating conjunction *because* in considerably more transparent versions. The paraphrase with *because* is deliberately contrived to show the interclausal relations established by conjunctions, as opposed to the intraclausal ones apparent in prepositions (e.g. *because of*) or nouns (e.g. *cause*):

(39) 'The cause of the restructuring of the Australian economy towards a manufacturing basis was the Second World War'
<
'Because of the Second World War the Australian economy was restructured towards a manufacturing basis' (Martin 1992: 409)
<
The Australian economy was restructured towards a manufacturing basis, because the Second World War broke out. (My paraphrase)

The signalling function evident in some logical metaphors leads Martin (1992: 416–17) to suggest a textual type of grammatical metaphor. The motivation behind this category lies in the assumption that nouns like *reason*, *example*, *point* or *factor* '[…] organise text, not field' (Martin 1992: 416). From his explanation, it may be inferred that whilst logical metaphors contribute to in-clause reasoning, the overlapping category of textual metaphors is aimed at developing the overall structure of a text. Example (40) below underlines such a text-organising function, with *a number of reasons* anticipating a stretch of discourse, and *for example* introducing one of the lexical realisations of the paragraph-initial textual metaphor. Another difference between both metaphors involves the possibility that textual metaphors allow for the expression of the speaker's evaluation of the meanings being made, as in the highly attitudinal and oral-like example provided in (41).

(40) 'I think Governments are necessary at different levels for a number of reasons. For example, they make laws, without which people would be killing themselves […]' (Martin 1992: 416)
(41) 'That point is just silly!' (Martin 1992: 417)

The last reference to be presented in this section is Downing & Locke (2006). In this introductory textbook to English grammar, Hallidayan theory appears in combination with more structural perspectives (as in 2.2.1.2). As in Halliday & Matthiessen (2004), the main focus here is on the complement clause pattern (Downing & Locke 2006: 457–60). In line with some of the references in 2.2.1.2, 'content clause' is substituted for 'projection' and 'noun complement clause' for 'projecting nouns'. Drawing on Quirk et al. (1985), the head noun in this construction is often a definite deverbal or deadjectival noun (e.g. *knowledge*, *belief*, *awareness*, *probability*; Downing & Locke 2006: 457). The use of these nouns may endow the complement clause with a particular stance (cf. Biber et al. 1999 in 2.2.1.3.2), manifested in the semantic division of these units into nouns of cognition and reasoning (e.g. *knowledge*, *belief*), speech-act nouns (e.g. *suggestion*, *proposal*) and personal assessment nouns (e.g. *possibility*, *doubt*; Downing & Locke 2006: 458).

2.2.1.3.2 *Corpus-driven grammars*

Sinclair et al. (1990), the first major corpus-driven grammar of English, derives linguistic description from the analysis of the 20 million-word *Birmingham Collection of English Texts*, the precursor of the *BoE.* Unlike all the grammars discussed so far (2.2.1.1-2.2.1.3.1), Sinclair et al. (1990) gives equal weight to both intra- and intersentential patterns of shell nouns. As regards the former, the analysis reveals a group of nouns followed by *to*-infinitive clauses and another one followed by reported *that*-clauses. Those with a *to*-infinitive are nouns derived from verbs or adjectives also followed by *to*-infinitives (e.g. *fail to>failure to*, *able to>ability to*) (Sinclair et al. 1990: 134). Those with a *that*-clause are primarily related to reporting verbs (e.g. *feel that>feeling that*, *state that>statement that*; Sinclair et al. 1990: 338). Some non-derived nouns expressing facts or beliefs are also followed by *that*-clauses (e.g. *advantage*, *benefit*, *danger*).

With respect to the intersentential use of shell nouns, the corpus shows a range of nouns with the ability to refer back '[…] to whole sections of spoken or written text' (Sinclair et al. 1990: 389), as in (42).

(42) '"Martin, what are you going to do?" – "That's a good question, Larry"' (Sinclair et al. 1990: 389)

Semantically, these nouns represent 'verbal actions' (e.g. *account*, *accusation*, *advice*), 'ideas' (e.g. *analysis*, *assessment*, *assumption*), and 'pieces of writing' (e.g. *paragraph*, *phrase*, *example*; Sinclair et al. 1990: 389–91). There is another category, however, which subsumes a whole range of items not belonging to any of the former three categories (Sinclair et al. 1990: 390): nouns like *action*, *circumstances*, *development*, *factor*, *fact* and *aspect* are all said to refer back to 'actions and events', even when some of these nouns (e.g. *factor*, *fact*, *aspect*) are clearly factual in meaning.

In addition to anaphoric nouns, Sinclair et al. (1990: 429) mention 'prefacing structures', a group of subject definite noun phrases which are used to emphasise or to label what the speaker is about to say. Emphasis is most frequent with *fact*, *point* and *thing*, as in (43), while labelling is evident with such nouns as *rule*, *answer*, *conclusion* or *problem*, as in (44).

(43) 'The thing is, how are we to get her out?' (Sinclair et al. 1990: 430)
(44) 'The inevitable conclusion is that man is not responsible for what he does' (Sinclair et al. 1990: 430)

Almost a decade after the publication of Sinclair et al. (1990), Francis et al. (1998) produced a grammar of noun and adjective patterns. This grammar describes the phraseology linked to nouns and adjectives. The patterns are retrieved from the 350 million-word *BoE*, and are presented in relation to semantic groups of units. The grammar gives a complete list of the units specific to every single semantic group within each pattern, thus proving invaluable for the semantic categorisation of shell nouns. Underlying this presentation of examples is the assumption that form and meaning are inseparable, to the extent that words with a similar meaning share the same pattern. To provide an example of the amount of semantic detail given, a typical shell-noun pattern like N-*that* (i.e. noun + *that*-clause) features in six meaning groups, outlined below (Francis et al. 1998: 108–13):

i) The 'suggestion group', referring to written or spoken types of language (e.g. *accusation*, *denial*, *testimony*),
ii) The 'belief group', referring to mental processes (e.g. *acceptance*, *awareness*, *realisation*),
iii) The 'happiness group', referring to emotions (e.g. *amazement*, *gratitude*, *pleasure*),
iv) The 'sign group', implying that something serves as evidence to prove something else (e.g. *clue*, *indication*, *proof*),
v) The 'possibility group', referring to degrees of likelihood (e.g. *chance*, *hope*, *odds*), and
vi) Nouns with other meanings (e.g. *advantage*, *benefit*, *problem*)

Biber et al. (1999), like Sinclair et al. (1990) and Francis et al. (1998), uses a corpus as the basis for linguistic description. It differs in the wealth of genre-related information offered to support its claims. This grammar draws on the 40 million-word *Longman Spoken and Written English Corpus* (henceforth, *LSWEC*), which covers four broad genres[1]: conversation, fiction, newspaper language and academic prose. Shell-noun description in this reference is restricted to four types of 'noun complement clause': *that*-clauses, *to*-infinitive clauses, *wh*-clauses and *of* V-*ing* clauses (Biber et al. 1999: 645–56). Overall, complement clauses headed by nouns are considerably less frequent than those headed by verbs or adjectives (Biber et al. 1999: 647). Even so, they are found to be prevalent in academic prose and rare in conversation. The nouns heading these constructions are said to belong to a closed set of abstract lexical items at the speakers' or writers' disposal to express their certainty of the complement clause or their attitude towards it (Biber et al. 1999: 647).

That-complement clauses reveal a strong preference for deverbal or deadjectival definite and singular noun phrases, those of which are used mainly in academic prose to express four types of stance, i.e. 'linguistic communication' (e.g. *claim*, *report*), 'cognitive reasoning' (e.g. *assumption*, *hypothesis*) and 'personal belief' (e.g. *belief*, *hope*) (Biber et al. 1999: 648). *To*-infinitive clauses, common in journalistic prose, are observed to collocate with nouns implying '[…] human goals, opportunities or actions' (Biber et al. 1999: 653), as illustrated by *opportunity*, *decision*, *capacity*, *bid* or *battle*. *Of* V-*ing* clauses, in turn, show no semantic co-occurrence tendencies, some nouns taking only this complementation pattern (e.g. *cost*, *task*) and others occurring also in one of the other structures (e.g. *idea*, *hope*, taking either *of* V-*ing* or *that*) (Biber et al. 1999: 653–5). Finally, *wh*-clauses feature as the least frequent complementation pattern, restricted primarily to the head noun *question* (i.e. *the question if/whether*…; Biber et al. 1999: 656).

Carter and McCarthy (2006) utilise the 700 million-word *Cambridge International Corpus* (henceforth, *CIC*) in their grammar of spoken and written English. Their discussion of noun complement clauses is more concise though similar in every detail to Biber et al'.s (1999). Therefore, it is not presented here (Carter and McCarthy 2006: 329–30). It is worth mentioning, however, their analysis of *thing* and *stuff* (Carter and McCarthy 2006: 147–9). These nouns, especially *thing*, are highly frequent in spoken conversation, where they are used to focus listeners' attention on what follows (as in (45)) and to describe phenomena that the speaker finds

[1] Biber et al. (1999) label them as 'registers'. See 4.2.1 for the reasons behind the use of 'genre' in this thesis.

difficult to label (as in (46); cf. Jespersen 1927, III and Sinclair et al. 1990 above). It is also noted that the occurrence of *thing* in focusing constructions (as in (45)) is often linked to a *problem* meaning (Carter and McCarthy 2006: 148).

(45) 'Yeah, but, you know, **thing** is, she's left it rather late' (Carter and McCarthy 2006: 148)
(46) 'I think **the whole Euro thing** has got completely out of control' (Carter and McCarthy 2006: 148)

2.2.2 *Specific research*

Following the overview of general grammars in 2.2.1, this section turns to a discussion of the literature specifically devoted to the formal, syntactic, textual and semantico-pragmatic features of shell-like units. The description of these units rests on nine different terms, i.e. 'general nouns' (Halliday & Hasan 1976; Mahlberg 2005), 'Vocabulary 3 items' (Winter 1977), 'lexical signalling' (Hoey 1979), 'enumerables and advance labelling' (Tadros 1985 and 1994), 'anaphoric nouns' (Francis 1986), 'carrier nouns' (Ivanič 1991), 'advance and retrospective labels' (Francis 1994), 'shell nouns' (Hunston & Francis 2000; Schmid 2000) and 'signalling nouns' (Flowerdew 2003a). The following description of these terms is structured around three subsections: 2.2.2.1 deals with the references where definitions start off highlighting the discursive function of shell-like units to subsequently present a set of formal identification criteria; 2.2.2.2 includes the studies where the identification of formal patterns precedes the functional interpretation of shell-noun uses; 2.2.2.3 focuses on research where definitions of shell-like units are driven by pedagogical concerns.

2.2.2.1 *From function to form*

2.2.2.1.1 *General nouns*

Back in the 1970s, Halliday & Hasan (1976: 274–82) reported on the cohesive function of a set of nouns functioning as '[…] superordinate members of major lexical sets […]' (Halliday & Hasan 1976: 275). Some examples of these nouns, then termed 'general nouns', are as follows:

(47) *people*, *person* [human]
(48) *thing*, *object* [inanimate concrete noun]
(49) *business*, *affair* [inanimate abstract]
(50) *question*, *idea* [fact] (Halliday & Hasan 1976: 274)

General nouns, sparsely discussed in their chapter on lexical cohesion, constitute an intermediate category between open and closed-class items (e.g. *table* and *this* respectively). Their indeterminate word-class status is motivated by their referential function in discourse, straddling the boundaries between grammatical referential cohesion (performed by pronouns) and lexical reiterative cohesion (involving the use of repetition and synonymy) (Halliday & Hasan 1976: 278–9). Their frequent co-occurrence with the definite article *the* and the demonstrative determiners *this*/*that* accounts for their similarity to anaphors realised by demonstrative and personal pronouns (Halliday & Hasan 1976: 275). Example (51) below may illustrate this point: *the stuff* could be replaced with the personal pronoun *it* with no drastic change in meaning:

(51) 'Leave **the stuff** there' (Halliday & Hasan 1976: 275)

Being so close to pronouns, however, general nouns are said to resemble prototypical lexical items in the possibility to carry an interpersonal or attitudinal meaning. The evaluation, linked to familiarity, may be conveyed either through the head noun alone (e.g. *idiot*, *devil*, *dear*) or in combination with attitudinal modifiers (e.g. *the stupid thing*). In this respect, Halliday & Hasan (1976: 276–7) emphasise that only attitudinal modifiers (e.g. *stupid*, *lucky*) are prevalent in these nouns (cf. Ivanič 1991: 106; Francis 1994: 95; Schmid 2000: 318 and Flowerdew 2003a: 335 below, where other types of modifiers are also found to accompany shell nouns).

Whilst it is argued that general words (comprising both nouns like *thing* or *person* and verbs like *make* or *do*) are '[…] limited in number […]', their borderline status is said to prevent the compilation of a complete list of potential instances (Halliday & Hasan 1976: 280).

Bolinger (1977: 5, 50–1) uses the terms 'low-content nouns' and 'classifiers' to account for Halliday & Hasan's (1976) general nouns. He argues that nouns like *region*, *thing*, *creature*, *action* or *device* are similar to pronouns in their co-referentiality and in their lack of prosodic stress, the latter being further linked to the semantically unspecific nature of both (Bolinger 1977: 5, 50–1). Thus, in example (52), the low-content noun *the scheme* simply echoes the meaning of *planned*, adding little new information to the sentence. Replacement of *the scheme* with *the conspiracy* would endow the sentence with more descriptive detail, as it would now be clear that the plan was secretly made. In Bolinger's (1977: 51) view, therefore, *the conspiracy* is informative and prosodically accented, while *the scheme* is low-content and unaccented.

(52) 'They planned to assassinate the King, but called off **the scheme**' (Bolinger 1977: 51)

In a chapter on corpus-based research into lexical cohesion, Partington (1998: 90–6) uses evidence from the *BoE* to question some of Halliday & Hasan's (1976) claims about the behaviour of these units. A distinction is made between general nouns referring to specific entities (those labelled by Lyons 1977, II: 442 'first-order entities', e.g. *man*, *object* or *place*; see 3.2.1) and those referring to longer stretches of discourse, such as *matter* or *question* (Partington 1998: 91). It is further argued that corpus evidence reveals a substantial number of cases of so-called general nouns being used with determiners other than *the* and *this* (Partington 1998: 91). Especially prominent in this respect are *a* and *such*, exemplified in (53) and (54) below. Finally, the connection that Halliday & Hasan (1976) establish between general nouns and anaphoric reference is, in Partington's (1998: 92) words, '[…] too restrictive', in view of the existence of such examples as (53) below, where the general noun has cataphoric reference.

(53) 'In **a move to tighten control of a far-reaching empire and to improve the group's own image**, Maurice and Charles Saatchi, credited with building up the company, have stepped down from the day-to-day running of the group' (Partington 1998: 94)

(54) 'They abandoned the felling of tropical trees to get at the plants established in their branching, but burned a forest to ensure a monopoly in orchids. **Such things** hit you in the pit of the stomach' (Partington 1998: 92)

The corpus approach adopted by Partington (1998) is exploited to the full by Mahlberg (2005) in a major study on the textual behaviour of general nouns. Three assumptions about these units are taken as the starting point for the analysis, namely that they are frequent nouns, that they perform local textual functions and that they are indeed nouns (Mahlberg 2005: 37). The frequency assumption manifests itself in the sampling procedure followed: 20 nouns are selected from the top frequency ranges of nouns in the *BoE* and the *BNC* (Mahlberg 2005: 51), and 100 concordance lines are then analysed for each unit drawing on the *BoE.* The second assumption is explored through the establishment of functional groups in the use of time nouns (*time*, *times*, *year*, *years*, *day*), people nouns (*man*, *woman*, *men*, *women*, *people*, *family*) and the more heterogeneous group of world nouns (*life*, *world*, *way*, *part*, *end*, *place*, *things*, *business*, *thing*). Functional groups stem from the meaning similarities that different lexicogrammatical patterns of these nouns show when used in natural discourse. Functional distinctions are made based on word-forms, as evidenced from the aforementioned list.

Examples (55) and (56) below exemplify this approach. Singular *time* in (55) means 'time passes', whilst plural *times* in (56) means 'measurement'. The conclusion is that different word-forms occur with different lexicogrammatical patterns and meanings.

(55) 'Magda we can't; we'll lose **time**' (Mahlberg 2005: 67)
(56) '[…] three **times** the size of Britain' (Mahlberg 2005: 66)

Finally, as regards the third assumption, i.e. that general nouns are nouns, Mahlberg (2005: 177) posits that Halliday & Hasan's (1976) claim about the borderline word-class status of these units is only '[…] of secondary importance in a corpus linguistic approach'. Any so-called traditional word-class is '[…] bound to have fuzzy edges' (Mahlberg 2005: 161) when explored in context.

Mahlberg's (2005) approach thus brings to the forefront the importance of examining the cohesive function of general nouns based on the specific functions (or meanings) that they reveal in individual texts.

Of all the textual functions performed by general nouns, one appears to play a key role in their use in context: the support function. This function is the subject of Mahlberg's (2003) paper. A noun is claimed to display this behaviour where little or no semantic contribution (i.e. little new meaning) is evident from its use, as its main purpose is to adapt speakers' or writers' communicative needs to the form of their message (Mahlberg 2003: 100). For example, in (57), *a move* adds little new meaning to the message: the writer's intention is not to emphasise that the adjoining relative clause is *a move*, but to introduce the point about to be made in a more compact form.

(57) 'Dressler, from Germany, has been mentioned as a possible bidder in the French market – **a move** to which the French Government is opposed – while Germany's Commerzbank has been rumoured to be interested in buying up Flemings in the UK' (Mahlberg 2003: 101)

Three reasons are suggested for the use of this function: laying emphasis, adding information in passing and providing an introduction (Mahlberg 2003: 102–5). Example (58), where *a man* is intended to highlight the information that follows, illustrates the first reason. The second reason is shown in (59), where the noun phrase is presented as Given (clause-initial), in spite of also introducing new evaluative information through the prepositional phrase. This ties in with the presuppositions argued to arise from the use of abstract shell nouns in subject position (see Schmid 2001: 1545 in 2.3.2.1). The third reason is most evident in *the thing is*, as in (60), a phrase often used in spoken discourse to initiate a new turn (cf. Sinclair et al. 1990: 429 in 2.2.1.3.2 and Schmid 2000: 329–37 in 2.2.2.2).

(58) 'It would doubtless be too much to expect Spurs fans to suddenly express a sweetness for Alan Sugar, a man who's been subjected to more abuse and hate mail than the average child molester' (Mahlberg 2003: 102)
(59) 'And, of course, where there's ladies (First or otherwise), there's George Hamilton. The man with the chicken tikka complexion pitches up in London this Saturday for a gig […]' (Mahlberg 2003: 103)
(60) '[…] The thing is, you can be huge in Europe […]' (Mahlberg 2003: 105)

The support function of these units is also remarked on by Mihatsch (2009). In her view, words like *thing, matter* and *affair* are nouns whose meanings have, over the centuries, gone through a process of increasing grammaticalisation (Mihatsch 2009: 84). This implies that their lexical specificity has gradually become blurred, and, as a result, they have come to function as 'placeholder nouns' (Mihatsch 2009: 83). Placeholder nouns are often drawn on to maintain the flow of discourse whenever the speaker is unable to recall a more specific lexical alternative (Mihatsch 2009: 85). Their use may also be motivated by the need to introduce an adjective in discourse ('adjective support', e.g. *the nice thing/the bad thing is that...*; Mihatsch 2009: 879).

Use of more specific items like *tree*, *chair* or *embezzlement* instead of *thing* (for all three) or *affair* (for the latter) makes their identification more straightforward and less context-dependent. Placeholder nouns, by contrast, are purely deictic, inasmuch as they indicate knowledge shared by speaker and hearer (Mihatsch 2009: 86). Thus, the occurrence of *that thing on the news* in a particular discourse situation would prove meaningful provided that both speaker and hearer share the information that, for example, someone has been accused of embezzlement. Failure to follow the news regularly would result in the hearer's inability to work out the meaning of *thing*. Similarly, the understanding of *that thing* to refer to an object deictically implied (e.g. *a chair*, *a knife*) is possible only if the hearer is present in the discourse situation. This shows that context is of paramount importance for the understanding of such highly unspecific nouns as *thing* or *matter*.

2.2.2.1.2 *Vocabulary 3 items and lexical signalling*

The discourse-organising function of general nouns (only touched on in the above references) lies at the core of Winter's (1977) threefold division of cohesive lexis into 'Vocabulary 1', 'Vocabulary 2' and 'Vocabulary 3 items' (cf. Martin 1992 in 2.2.1.3.1).

Winter (1977) argues that the processing and understanding of relations between clauses and sentences is enabled thanks to three types of clause-relational words. 'Vocabulary 1' and 'Vocabulary 2' comprise closed-class

items, i.e. subordinators (e.g. *after*, *provided that*, *whereas*) and sentence connectors or conjuncts (e.g. *however*, *generally*, *in fact*). 'Vocabulary 3', by contrast, consists of nouns (e.g. *distinction*, *requirement*), verbs (e.g. *achieve*, *affirm*) and adjectives (e.g. *analogous*, *common*), most of which can paraphrase subordinators, sentence connectors or both. To provide an example, the noun *method* is a Vocabulary 3 item that paraphrases Vocabulary 1 *by* [...] *-ing* and Vocabulary 2 *thereby* (Winter 1977: 23).

Vocabulary 3 items are, like general nouns above, open word-classes with semantic properties of closed word-classes. The former status is explained on the grounds of the modification possibilities (pre- and postmodification) granted by prototypical open-system items (e.g. *striking* in *a striking example*) (Winter 1977: 23–4). Their closed-system semantics stems from their function as cataphoric 'signposts' (Winter 1977: 2), as their use is often influenced by their role as '[…] anticipators for the next part of their *paragraphs* […]' (Winter 1977: 9, italics as in the original). This implies that the meaning of Vocabulary 3 items always finds its 'lexical realisation' (Winter 1977: 8) in the following '[…] clause or group of clauses […]' (Winter 1977: 7). The predictive potential of these words is regarded as '[…] almost inevitable' (Winter 1977: 36). An example of the use of a Vocabulary 3 item is the following, where the noun *requirement* predicts the modal verb *must* as well as the sentence connector *otherwise*.

(61) '(1) One **requirement** for the success of the course is obvious. (2) The student **must** like the course; **otherwise** he will not follow it with enthusiasm' (Winter 1977: 21)

Winter (1977: 27) agrees with Halliday & Hasan (1976: 280) in considering Vocabulary 3 as '[…] a small and fairly stable vocabulary […]'. As a result, a list of 104 potential instances of Vocabulary 3 items is proposed (Winter 1977: 20). The list comprises 60 nouns (e.g. *situation*, *consequence*, *feature*) and 44 adjectives and verbs (e.g. *correct*, *real*, *reciprocate*, *specify*).

Winter's (1977) interest in the contextual analysis of English clauses becomes the main focus of his 1982 book on the grammar of English clauses and sentences. The thrust of Winter's argument is that clauses are marginally informative on their own, thereby always needing some kind of lexical realisation or specification by surrounding clauses (Winter 1982: 40). This is so much so that, in Winter's words, research on English grammar ought to be primarily concerned with the identification of the relation '[…] between incomplete and complete clause(s)' (Winter 1982: 44). Contextual incompleteness is not only specific to pronouns, but represents '[…] a general semantic characteristic of many lexical items in the clause' (Winter

1982: 40). These lexical items are nouns, adjectives and verbs falling within the scope of Vocabulary 3 items. Example (62) below illustrates this point. *Something* is an indefinite pronoun that is lexically specified by the adjoining clause. Replacement of *something* with a noun phrase, e.g. *a significant event*, would endow the first clause with similar incompleteness, while also constituting a more subjective selection. Winter (1982: 44) thus claims that '[…] every lexical selection is evaluative or subjective […]', to the extent that, unlike pronouns, the choice of a lexical item hinges on the speaker's or writer's personal interpretation and opinion on the surrounding co-text.

(62) 'Something of significance did, however, happen in the middle of all this: a television camera was admitted into the chamber of the House of Commons for the first time […]' (Winter 1982: 40)

A decade later, Winter (1992) revised his notion of Vocabulary 3 items in the light of Francis' (1986) and Ivanič's (1991) research on anaphoric nouns and carrier nouns respectively (see 2.2.2.1.4). The conclusion is that Vocabulary 3 items belong to the larger group of metalinguistic lexical items (Winter 1992: 140). By 'metalinguistic', reference is made to those lexical items which '[…] "talk about" the nature of the clause or sentence as a message in the text itself, and […] do not refer to concrete things in the outside world' (Winter 1992: 133). Both metalanguage nouns and Vocabulary 3 items are unspecific items lexically realised by more specific discourse elements (Winter 1992: 140). However, the latter spell out the kind of clause relation holding between two clauses, while the former simply ascribe a specific meaning to one or more clauses, irrespective of their relation (Winter 1992: 156). *Result* in (63) represents a prototypical Vocabulary 3 item, one which, like many others (e.g. *difference*, *reason*, *way*, *means*), could be replaced with a Vocabulary 2 item (i.e. conjunct) such as *therefore*. By contrast, *idea* in (64) represents an item of the metalanguage (e.g. *assumption*, *assessment*, *theory*, *news*), one which just offers the speaker's or writer's interpretation of the surrounding co-text.

(63) 'In some places there was no rain at all. The result was the same – the farmers lost their always precarious crops of maize […]' (Winter 1992: 134)

(64) '[…] we had no idea we'd get such an overwhelming response' (Winter 1992: 133)

Winter's discussion of the non-specificity of metalanguistic items involves a crucial distinction between two kinds of specifics, i.e. 'specifics of identity' and 'specifics by clause' (Winter 1992: 154–5). By 'specifics' is meant the surrounding linguistic material one has to draw on in order to fully

understand and interpret the meaning of the unspecific metalinguistic item. 'Specifics of identity' are realised by premodifiers and postmodifiers whose only role is to restrict the reference of the head noun without necessarily identifying it. When applied to shell nouns, these specifics indicate what the head noun relates to, but not what the head noun is. A perfect example of these specifics at work is that of relative clauses. In (65) below, the relative clause *that threaten public health and well-being* does not tell the reader what some of these problems are; it simply restricts the reference of the head noun (and, to a certain extent, also evaluates it). An example of 'specifics by clause' is (66), where the actual nature of these *conditions* (i.e. what these conditions are) is not made explicit by the postmodifier *at home*, but by the three sentences that follow. Specifics by clause thus comprise one or more sentences and, if present, enable an understanding of the contextual or text-related meaning of the unspecific noun.

(65) 'Now we are hearing from concerned citizens in all parts of the country who want to know what they can do to hold local officials accountable for tackling **population-related problems that threaten public health and well-being**' (Winter 1992: 155)

(66) 'His marriage was in tatters. **Conditions at home** were terrible. The house was in a shambles. His wife drank. The children screamed all day, and his mother-in-law had moved in to restore order' (Winter 1992: 156)

As regards the direction of the link between unspecific and specific, cataphoric (anticipatory) reference, foregrounded in Winter (1977), is now considered alongside anaphoric (retrospective) reference (Winter 1992: 154). This widening of the referential scope is also apparent in Winter (1982: 32), where it is suggested that the study of signalling in discourse should lay emphasis on '[…] the signalling role of any word in the clause, whether it signals backwards in its sentence or beyond its sentence to a preceding sentence, or whether it signals forward within its sentence or beyond its sentence to a sentence which follows it'.

Winter's (1992: 159) article concludes with a list of 131 metalanguage nouns retrieved from an unspecified 1.3 million-word corpus. Two patterns underlie the automated queries run in the corpus, i.e noun + *that*/*to*-infinitive clause and subject + predicator + *that*/*to*-infinitive clause.

Winter's (1977) clause-relational approach to the analysis of discourse is further developed by Hoey (1979, 1983 and 1994). Hoey's main concern is with Vocabulary 3 items, particularly with the way in which such lexical signals are used in the organisation of certain discourse patterns. In all three studies, Hoey expands the research scope from the level of the paragraph or below, as in Winter (1977, 1982 and 1992), to '[…] whole discourses' (Hoey 1994: 34). One discourse pattern figuring prominently in

Hoey's research is the 'Situation-Problem-Response-Result-Evaluation' pattern (Hoey 1983: 61). Example (67) illustrates the Situation-Problem section of a real text. The occurrence of the lexical signal *problems* immediately after the Situation (i.e. *Helicopters are very convenient*) spells out clearly the nature of what follows. In this respect, Hoey (1983: 52) argues that there is no one-to-one relationship between real-world problems and 'linguistic problem(s)', inasmuch as the reason for the use of the lexical signal *problem* rests on the writer's own interpretation of linguistic information as a problem. The Problem-Solution pattern has more recently been examined through corpus techniques, as epitomised by Flowerdew (2008) (see 2.3.1.2.2).

(67) 'Helicopters are very convenient for dropping freight by parachute, but this system has **its problems**. Somehow the landing impact has to be cushioned to give a soft landing. The movement to be absorbed depends on the weight and the speed at which the charge falls. Unfortunately most normal spring systems bounce the load as it lands, sometimes turning it over […]' (Hoey 1983: 68)

It is important to point out that Hoey's research into lexical signalling does not consider the occurrence of lexical signals as a prerequisite for discourse cohesion. It is claimed that the sequential ordering of sentences in real discourse may suffice to enable the readers' understanding of textual structure (Hoey 1983: 60). This ties in with de Beaugrande & Dressler's (1981: 4, 36, 200), Brown & Yule's (1983: 65) and McCarthy's (1984, cited in Carter 1998: 86) claim that, regardless of the occurrence of cohesive elements or not (e.g. conjunctions, conjuncts, nouns, etc.), readers and listeners will always strive to make some sense of the structure of the stretch of discourse they are being exposed to. Example (68) illustrates how the absence of a lexical signal such as *example* does not hinder an interpretation of the second sentence as an example for the information contained in the first one.

(68) 'Second, to avoid those heavy shadows under the eyes with bounce flash, try keeping your distance with a medium telephoto for portraits. I took the portrait of the little girl on the left with a 100mm lens on my Minolta from about 3m, with a Sunpak 4205G hammer gun aimed at a wall on the left and only slightly upward' (Carter 1998: 86)

2.2.2.1.3 *Prediction in discourse*

Winter's (1977) initial emphasis on the cataphoric function of Vocabulary 3 items was echoed by Tadros (1985 and 1994). Tadros relies on the basic

assumption that expository writing is essentially predictive, where 'prediction' is defined as '[…] a prospective rhetorical device which commits the writer at one point in the text to a future discourse act' (Tadros 1994: 70). Prediction is argued to stand in contrast with 'anticipation', which involves the reader's expectations about the structure and content of discourse (Tadros 1985: 6). Ideally, what the reader anticipates should coincide with what the text predicts.

Six categories of prediction are presented: enumeration, advance labelling, reporting, recapitulation, hypotheticality and question. Of these categories, only the former four are of immediate relevance to the study of shell nouns:

i) 'Enumeration' (Tadros 1994: 71–3) implies the use of certain sub-technical nouns (e.g. *advantages*, *reasons*, *aspects*) and discourse reference nouns (e.g. *examples*, *definitions*, *classifications*) as predictors of lists of items, categories, ideas, arguments, etc. These nouns, occurring in sequences like example (69) below, are termed 'enumerables'. Tadros (1994: 72) cites 47 items where this function is apparent.

(69) '[…] there are a number of ways by which risks can be reduced' (Tadros 1994: 72).

ii) 'Advance labelling' (Tadros 1994: 73–4) does not predict enumeration, but simply the writer's commitment '[…] to perform a discourse act'. Three types of advance labelling are proposed, one where the label is followed by linear text (70), one where the label is followed by non-linear text (71), and one where it is followed by both (72):

(70) 'This analysis leads us to make the important distinction between real income and money income […]' (Tadros 1994: 73)
(71) 'We can show this in a simple diagram as follows: […]' (Tadros 1994: 74)
(72) 'Consider now the following cost schedule of a firm […]' (Tadros 1994: 74)

iii) 'Reporting' (Tadros 1994: 74–6) entails the writer's detachment from propositions by citing other sources. This detachment leads in turn to the writer's subsequent evaluation of what has been reported. Even though this category is primarily associated with reporting verbs such as *show*, *prove*, *claim* or *suggest*, the possibility for inclusion of nominalised metadiscursive nouns is contemplated, as in (73):

(73) 'Halliday's (1970) discussion of language structure and function is pitched at a different level […]' (Tadros 1994: 76)

iv) 'Recapitulation' (Tadros 1994: 76) summarises what has been said in order to introduce new information. This category is often associated with combinations of verbs such as *mention*, *consider* or *note* and nouns such as *section*, *chapter* or *paragraph*:

(74) '[…] It was mentioned […] in the preceding section' (Tadros 1994: 76)

2.2.2.1.4 *Anaphoric nouns, labels and carrier nouns*

The studies reported so far reveal a rather partial account of shell-like units. Their description is often subsidiary to a more general concern with discourse structure, as is evident in sections 2.2.2.1.2 and 2.2.2.1.3. In relation to general nouns (2.2.2.1.1), their treatment is, for the most part, superficial (with the exception of Mahlberg 2005) and, in any case, the presence of concrete units in their ranks (e.g. *man*, *person*, *object*) indicates only a partial overlap with prototypical shell nouns. The three terms to be presented below (i.e. 'anaphoric nouns', 'labels' and 'carrier nouns') offer a more inclusive definition of the overall formal and functional boundaries of shell nouns.

Francis (1986) proposes the term 'Anaphoric nouns' or 'A-nouns' to account for the metadiscursive use of certain nouns. This implies that any noun '[…] which can be used to *talk about* the ongoing discourse' (Francis 1986: 3, italics as in the original) qualifies as an A-noun. Based on this criterion, anaphoric nouns are broken down into two major semantic classes (Francis 1986: 11-19), i.e. 'purely metadiscursive nouns' (illocutionary nouns, verbal activity nouns, cognition nouns and text nouns; e.g. *accusation*, *exposition*, *assumption* and *paragraph*) and so-called '"ownerless" nouns' (e.g. *aspect*, *matter*, *subject*, *problem*). The latter share with typical metadiscursive nouns their referential non-specificity, but not their potential for labelling types of language; they are claimed to '[…] exist in the world outside discourse' (Francis 1986: 17). Further details on Francis' (1986) semantic typology are offered in 3.2.5. Francis' (1986: 11–18) discussion of semantic categories of A-nouns also involves the presentation of a list of 234 units. Such a list, however, is far from comprehensive in that '[…] any noun which can be used metadiscursively can function as an A-noun within a discourse' (Francis 1986: 7). Additional features relevant to the description and identification of these units are the following:

i) Almost as important as the semantic (i.e. metadiscursive) criterion is the '"pro-form" criterion' (Francis 1986: 27), whereby the interpretation of any A-noun necessarily depends on a preceding stretch of discourse.
ii) Following Halliday & Hasan (1976), reference in A-noun phrases is realised only through definite determiners, the most frequent being *the*, *this*, *these*, *that*, *those* and *such* (Francis 1986: 27).
iii) A-nouns are argued to '[…] face two ways' (Francis 1986: 38), in that they condense a previous stretch of discourse which is then introduced as a new entity (with all the nominal modification possibilities that this implies). Hence, an A-noun like *this assessment* in (75) nominalises the previous co-text and, in so doing, subsequent reference to the information it encapsulates is made easier thanks to its nominal status. The use of a noun phrase instead of a pronoun also allows the writer to evaluate the previous discourse (Francis 1986: 48).

(75) '[…] the Soviet Union has 'shot its bolt', and that only the unreconstructed Cold Warriors are losing any sleep about the Russian menace. James Reston has readily and complacently echoed **this assessment** in his criticism of the Reagan équipe' (Francis 1986: 27)

iv) Evaluation may be conveyed through the head noun alone or in combination with modifiers. A-nouns fall into two evaluative sub-groups (Francis 1986: 49), i.e. attitudinally neutral ones (e.g. *comparison*, *issue*, *approach*) and evaluative ones (e.g. *insight*, *realisation*, *eloquence*: positive evaluation; *distortion*, *exaggeration*, *fabrication*: negative evaluation).
v) One of the reasons for choosing an A-noun instead of a pronoun is its potential for modification (Francis 1986: 55–63). A-nouns are often preceded by attitudinal modifiers (e.g. *carefully chosen*, *inaccurate*, *down-to-earth*). Modification may also be 'propositional', 'organisational' and 'comparative'. 'Propositional content' modifiers assign the head noun to a particular objective subclass (e.g. *this monist/anthropocentric vision*). 'Organisational' modifiers comprise items such as *another*, *other*, *similar* and *the same*, those of which help to reinforce the connection between noun and antecedent. Finally, 'comparative' modifiers are similar to organisational modifiers in their connective role, but differ in their evaluative meaning (e.g. *an even more decisive argument*, *a more plausible explanation*).

vi) While the emphasis is on anaphora, passing reference is made to the existence of cataphoric uses of A-nouns (Francis 1986: 104), as in (76) below:

(76) 'There are **a number of extraordinary contradictions** in all this […]' (Francis 1986: 104)

Cataphora finds a place in Francis' description of shell-like nouns in 1988, where the term 'label' is used to differentiate between 'retrospective' (or anaphoric) and 'advance' (or cataphoric) units (Francis 1988: 326–7). Such a characterisation of these nouns is inspired by Tadros' (1985 and 1994) concept of 'advance labelling', one of the six categories comprising her account of prediction in discourse (see 2.2.2.1.3). In view of the pedagogical nature of Francis (1988), a fuller discussion of its relevance for shell noun categorisation is deferred until section 2.2.2.3.

Francis (1994) draws on the advance-retrospective label distinction to provide an update on the theory underlying her 1986 monograph. Francis (1994: 83) defines a 'label' as '[…] an inherently unspecific nominal element whose specific meaning in the discourse needs to be precisely spelled out […]' (Francis 1994: 83). Her definition rests on Winter's (1992) description of Vocabulary 3 items (among other metalanguage nouns) as unspecific discourse items (see 2.2.2.1.2 above).

The scope of the paper is arguably limited in two respects. On the one hand, as in Francis (1986), the link between the head noun and its antecedent is restricted to intersentential boundaries. No mention is therefore made of instances where the link appears within the same sentence (cf. Ivanič 1991; Schmid 2000 and Flowerdew 2003a, where both inter- and intrasentential uses are included). On the other hand, whilst both anaphora and cataphora are allowed for, considerably more attention is paid to retrospective labels, on account of their reportedly greater frequency and wider lexical range (Francis 1994: 95).

The form and function of labels does not differ much from the A-noun features reported above. Some basic aspects of labels are summarised in the following:

i) Specific deictics (e.g. *the*, *this*, *that*, *such*), along with the optional presence of modifiers and qualifiers, endow labels with functions specific to pro-forms (Francis 1994: 85).
ii) Reference exists only between labels and stretches of discourse, leading Francis (1994: 85) to state categorically that a '[…] major criterion for identifying an anaphorically cohesive group as a

retrospective label is that there is no single nominal group to which it refers'.

iii) They have a '[…] topic-shifting and topic-linking function […]' (Francis 1994: 86), in the sense that their introduction as given elements of the clause provides a link back to the previous discourse, while also serving as topic initiators. In (77), the noun *move* offers the writer the possibility to summarise the government's action and, as such, it acts as a springboard for the writer's evaluation of this action:

(77) '[…] The Polish government is on the verge of outlawing abortion, which has been free on demand since 1956. This move in itself is deplorable, but is made far worse by the fact that contraception is virtually unobtainable […]' (Francis 1994: 87)

iv) They are often used with a 'fuzzy reference' (Francis 1994: 88). This is a result of the difficulties often encountered when locating the specific stretch of discourse which the label lexicalises. Although such indistinctness is usually unconscious, there are cases where its use may be motivated by a wish to persuade or manipulate the addressee.

v) Francis' (1986: 11–19) distinction between metadiscursive and ownerless A-nouns is now replaced with a twofold semantic division into 'metalinguistic' and 'non-metalinguistic' labels. Metalinguistic labels name '[…] a stretch of discourse as being a particular type of language […]' (Francis 1994: 89). Therefore, nouns like *accident*, *occasion*, *process* or *topic* are non-metalinguistic (i.e. ownerless above), while nouns like *explanation*, *distinction*, *theory* and *paragraph* are metalinguistic, insofar as each falls into one of the four metalinguistic categories proposed by Francis (1994: 90): illocutionary nouns, language activity nouns (verbal activity nouns above), mental process nouns (cognition nouns above) and text nouns. Further details on the semantic taxonomy are given in 3.2.5.

It is relevant to note that Francis (1994) differs from its 1986 precursor in its more explicit orientation towards Hallidayan SFG. This is evident in the passing reference made to the metafunctional properties of labels. For example, *this move* in (77) above has an ideational meaning conveyed through its role as Carrier in an attributive relational clause, an interpersonal meaning implicit in the official and almost unchangeable nature of a political move, and a textual meaning stemming from its status as given information (Francis 1994: 88). Indeed, despite their obvious text-organising function as

given elements of the clause, labels may carry additional evaluative nuances. Such implicit interpersonal meaning leads Francis (1994: 93) to the creation of a category of 'evaluative retrospective labels'. Though excluding advance labels, this category contains attitudinally neutral head nouns (e.g. *statement*, *belief*) and clearly attitudinally marked nouns (e.g. *nonsense*, *squabble*).

The three Hallidayan metafunctions are further explored through their presence in the premodifiers of retrospective labels (Francis 1994: 95–100). Modification in retrospective labels is thus of three types:

i) 'Ideational' modification ('propositional' above) further restricts the reference of the head noun, either through classification or definition (e.g. *this new confectionary concept*).
ii) 'Interpersonal' modification ('attitudinal' above) helps convey the writer's attitude in attitudinally-neutral labels presented as Given (e.g. *this far-sighted recommendation*).
iii) 'Textual' modification ('organisational' above) fulfils a discourse-organising function whereby the modifier is presented as New and the rest of the nominal group as Given (e.g. *a similar/different/another blunder*).

Francis (1994: 89–93) provides a list of 234 labels, the same number as in Francis (1986). The lists, however, are not identical, which follows from the corpus used in each case (see 2.3.1.1). Compilation of an exhaustive list of labels is in Francis (1994) also claimed to prove unfeasible, but the reason given is not based on the metadiscursive criterion, as in Francis (1986: 7) (see above). In Francis (1994: 88), the impracticality of such a list stems from the idea that '[…] any noun can be the head noun of a label if it is unspecific and requires lexical realisation in its immediate context […]'. The influence of Winter (1992) on Francis' (1994) definition of labels is also evident in this claim.

The last term to be discussed in this section is Ivanič's (1991) 'carrier nouns'. These are described as countable abstract nouns which '[…] carry a specific meaning within their context in addition to their dictionary meaning' (Ivanič 1991: 95). Accordingly, nouns such as *example*, *aspect*, *advantage*, *purpose*, *question* and *feature* are all said to have a meaning which remains 'constant' and an additional 'variable meaning' (Ivanič 1991: 96) whose identification depends on their context of appearance. The variability in their meaning is linked to the assumption that carrier nouns are '[…] not subject-specific' (Ivanič 1991: 96). Ivanič's (1991: 97) use of the term 'carrier' hinges on the identification of this category on the basis of Halliday's (1985)

SFG. As such, carrier nouns resemble the Carrier participant in their frequent occurrence in relational process clauses (cf. 2.2.1.3.1).

The closed-system nature of these nouns is attributed to their similarity to pronouns, in that both have a variable meaning interpreted in terms of either the surrounding textual context, i.e. endophoric reference, or the reader's background knowledge, i.e. exophoric reference (Ivanič 1991: 103). With regard to the former, Ivanič (1991: 104–5) acknowledges the existence of both anaphoric and cataphoric carrier nouns, and provides an example where both types interact:

(78) '[…] Up to a point this is a useful analogy, but it breaks down for several reasons. […]' (Ivanič 1991: 104)

This sentence contains two carrier nouns, one pointing backwards and acting as Given (*analogy*), and the other referring to the following stretch of discourse (*reasons*).

Ivanič's (1991) study is not solely concerned with the intersentential function of carrier nouns, as is the case in Halliday & Hasan (1976), Winter (1977), Francis (1986 and 1994) and Tadros (1985 and 1994). Special consideration is also given to their intrasentential function (Ivanič 1991: 101–3), as manifested in their occurrence within Vendler's (1968) container sentences (see 2.2.2.2). Vendler's container nouns are closely linked to a prototypical syntactic pattern, where 'N *is* nominalisation and nominalisation *is* N'. This means that a container noun (N) can combine with a complement (i.e. nominalisation) in the form of a *that*-clause, a *to*-infinitive clause, a *wh*-question clause or a deverbal noun. For example, in (79) below, the carrier noun *purpose* is lexically realised within the same clause by the *to*-infinitive clause acting as subject complement:

(79) 'The purpose of the following section is to provide an elementary account of the magnetic properties of ferrites' (Ivanič 1991: 101)

The referential potential of carrier nouns should be understood in conjunction with their purely nominal properties. In this respect, it is claimed that the syntactic versatility of nouns is considerably higher than that of pronouns (Ivanič 1991: 107), inasmuch as their lexical meaning enables them to take different positions in the information structure of the clause. More important, however, is the fact that, unlike pronouns, carrier nouns can be accompanied by determiners and modifiers (Ivanič 1991: 108). Modification of carrier nouns is often accomplished through modifiers restricting the reference of the head noun (as in (80)) and through evaluative modifiers (as in (81)). This claim contrasts with Halliday &

Hasan's (1976: 276–7) remark on the dominance of only attitudinal modifiers with general nouns.

(80) '[…] certain **food** problems' (Ivanič 1991: 106)
(81) 'These **stern** measures […]' (Ivanič 1991: 108)

Whilst providing many textual examples to substantiate the description of carrier nouns, Ivanič's (1991) article does not include a complete list of instances, which is explained on the basis of the large variety of countable abstract nouns to be found in discourse (Ivanič 1991: 99). The list offered comprises only 33 examples.

2.2.2.2 *From pattern to function*

The references discussed so far take as their starting point a functional description of shell nouns, thanks to which a set of formal features are subsequently identified:

i) General nouns are explained in relation to their cohesive use as anaphoric lexical substitutes (Halliday & Hasan 1976), as well as the semantic and pragmatic functions they perform when used in context (Mahlberg 2003 and 2005).
ii) Vocabulary 3 items (e.g. Winter 1977) and lexical signals (e.g. Hoey 1983) are identified on the basis of their potential for specifying clause relations and discourse structure.
iii) Anaphoric nouns (Francis 1986) are closely connected with the semantic criterion of metadiscursivity, while labels exist in terms of their unspecific meaning (Francis 1994).
iv) Carrier nouns (Ivanič's 1991) are initially defined in terms of their meaning (constant vs. variable meaning) and inherent formal properties (countable abstract nouns).

In the references reviewed below (Vendler 1968; Hunston & Francis 2000; Schmid 2000), the identification of shell-noun phrases rests largely on a set of predefined syntactic patterns. With the exception of Vendler (1968), examples occurring in these patterns are retrieved automatically from large corpora. Compared to the research reported in the previous section, where the emphasis is often laid on the intersentential lexicalisation of shell-noun phrases, the studies below focus particularly on encapsulation within the same sentence. Noun complement clauses (Biber et al. 1999: 644–56, see 2.2.1.3.2), i.e. noun-clause/noun-*be*-clause, occupy a prominent role in this

research. It is in two of these references, i.e. Hunston & Francis (2000) and Schmid (2000), that the term 'shell noun' is first introduced in the literature.

Vendler (1968) explores English nominalisations from a TG approach (cf. Chomsky 1970 in 2.2.1.2). In Vendler's framework, nominalisations comprise morphologically derived deverbal and deadjectival nouns, and by extension any structure that enables '[…] the incorporation of a sentence into another' (Vendler 1968: 27). This means that any subordinate nominal clause (e.g. *that*-clauses, *-ing* clauses, *to*-infinitive clauses, etc.) would be regarded as a nominal or a nominalised form. The relevance of such a concept of nominalisation to shell nouns lies in Vendler's proposal of 'container sentences' or 'containers'. In simple terms, these are sentence patterns with nominal gaps. The completion of these gaps crucially depends on '[…] the co-occurrence restrictions between families of nominals and families of containers' (Vendler 1968: 34). Examples (82) and (83) represent two container sentences, the former admitting nominals like *his death*, *that he died* and *his having died*, and the latter only admitting *his death*.

(82) 'O surprised me' (Vendler 1968: 33)
(83) 'O occurred at noon' (Vendler 1968: 33)

The examples of container shell-like nouns featuring in Vendler's (1968) book are often of the 'O *is* N' type, where 'O' stands for a subordinate nominal *that*- or *to*-infinitive clause, and 'N' for an abstract noun. Examples (84) and (85) illustrate this pattern. The class of nouns that may occupy the N gap in these sentences is, according to Vendler (1968: 63), '[…] not easy to define'. One feature these nouns seem to share is the occurrence of a premodifying adjective, as in (84), i.e. *difficult*.

(84) 'To build a battleship is a difficult task for any shipyard' (Vendler 1968: 63)
(85) 'That he died is a fact' (Vendler 1968: 73)

Passing mention is also made of the complement clause pattern. In this respect, Vendler (1968: 37) argues that only in certain container sentences is the insertion of an abstract noun allowed before the conjunction introducing a clausal nominal. For example, the insertion of the noun phrase *the fact* would be acceptable in (86) but not in (87).

(86) 'I know the fact that he died' (Vendler 1968: 37)
(87) '*I think the fact that he died' (Vendler 1968: 37)

From the above, it is evident that Vendler (1968) only scratches the surface of all that is involved in shell-noun behaviour. His account of nominals is

restricted to the level of the sentence, which means that no attention is given to the wider context. This, in turn, implies a focus on specific patterns (e.g. O *is* N), and some disregard of patterns that a more context-sensitive analysis might reveal. A major contribution of Vendler (1968) is the connection between grammatical patterns and lexis, where acceptability judgements draw on specific container sentences and nominal gaps. This shows an interest in the lexicogrammatical environment favoured by specific nominals. It stands to reason, for example, that a fully factual noun like *fact* or *point* cannot occur in combination with an eventive verb like *occur*. With the advent of corpus linguistics, intuitions of this kind started to be tested against contextualised examples extracted from computerised samples of real language, and not against invented examples (as in TG). Hunston & Francis' (2000) *Pattern Grammar* is paramount in this respect (cf. Sinclair et al. 1990 and Francis et al. 1998 in 2.2.1.3.2). Grammar, in this framework, is not a repository of empty structural patterns waiting to be filled by any lexical item. On the contrary, a 'pattern' is closely associated with the specific phraseology that words appear as when used in context. The identification of a pattern is thus dependent on whether '[…] a combination of words occurs relatively frequently, [whether] it is dependent on a particular word choice, and [whether] there is a clear meaning associated with it' (Hunston & Francis 2000: 37).

Four pages are devoted in *Pattern Grammar* to what Hunston & Francis (2000: 185–8) regard as a new word-class, i.e. 'shell nouns'. The notion of word-class in this grammatical approach has to be understood based on the idea that '[…] words do not "have" classes as something intrinsic to them' (Hunston & Francis 2000: 197). Instead, word-classes can only be established once the use of individual lexical items is examined in context. It is this context-sensitive exploration of the patterns of groups of lexical items that led them to come up with this subclass of nouns.

Shell nouns are defined as '[…] nouns which require lexicalisation in their immediate context' (Hunston & Francis 2000: 185). Drawing on frequency data from a large corpus (*BoE*, see 2.3.2.1), shell nouns are found to dominate the pattern where the head noun is immediately followed by a *that*-clause (see (88)). Despite the significance of this pattern, it is also noted that, on the basis of Francis' (1994) paper, shell nouns may occur as advance or retrospective labels across sentence boundaries (see (89)).

(88) 'He had **an unshakable premonition** that he would die' (Hunston & Francis 2000: 186)

(89) '[…] Why, then, should there be such a preponderance of right-handed golfers, which extends, I am informed, to club level? In reply to **that question** a golfing colleague of mine offered **two reasons**. The first was

> that beginners usually start with handed-down clubs, which are usually right-handed […]' (Hunston & Francis 2000: 187)

A threefold semantic classification of shell nouns is developed on the basis of their frequent occurrence in such a pattern (Hunston & Francis 2000: 186–7), i.e. 'nouns referring to something that is written or spoken' (e.g. *assurance*, *remark*, *prediction*), 'nouns referring to something that is thought or believed' (e.g. *inference*, *hope*, *recollection*) and 'nouns not fitting into either of these two groups' (e.g. *fact*). All in all, 194 cases of shell nouns are presented.

Research on shell nouns finds its most comprehensive treatment to date in Schmid (2000). Schmid (2000: 4) defines them as '[…] abstract nouns that have […] the potential for being used as conceptual shells for complex, proposition-like pieces of information'. Given that shell nouns can only be understood in connection with the propositional content they encapsulate, Schmid's (2000: 8) study is especially concerned with 'shell-content complexes', rather than with shell nouns alone. In (90) below, the shell noun *aim* and the propositional content starting with *to make* constitute a shell-content complex:

(90) 'The Government's aim is to make GP's more financially accountable, in charge of their own budgets, as well as to extend the choice of the patient. […]' (Schmid 2000: 7)

Schmid (2000: 13) lays emphasis on the fact that shell nouns can only be understood from a functional perspective. As such, they do not exhibit any inherent formal properties, as prototypical nouns do (cf. Ivanic's 1991 claim that carrier nouns are countable abstract nouns). Instead, their 'shell-nounhood' (Schmid 2000: 13) is crucially dependent on their use in context. In this respect, three key functions are ascribed to shell nouns: 'characterisation', 'temporary concept-formation' and 'linking'. By 'characterisation', it is meant that, despite the anaphoric potential of both pronouns and shell nouns, only the latter possess the ability to name a stretch of discourse in a certain way, for example, as a *fact* or a *problem* (Schmid 2000: 15). 'Temporary concept-formation' rests on the potential shown by shell nouns for turning a complex piece of information such as (91) into a concept:

(91) '[…] That it is now very much a regular winter visitor […]'
>
'[…] The fact that it is now very much a regular winter visitor […]' (Schmid 2000: 367)

This concept is, nevertheless, of a temporary nature, since its meaning varies according to the surrounding discourse. Such temporariness appears to tie in with Ivanič's (1991: 95) distinction between the constant and variable meanings of carrier nouns.

The third function, 'linking', underlines the frequent use of shell nouns as anaphoric textual linkers:

(92) '[…] Since 1902 archaeologists have usually attributed such findings to cannibalism, though few have explained their reasons for **this explanation** adequately' (Schmid 2000: 343–4)

In view of the above emphasis on the functional nature of these nouns, it is paradoxical, however, that purely syntactic criteria are applied to the retrieval of examples from the *BoE*. The following lexicogrammatical patterns represent the four prototypical syntactic environments where shell nouns are found to occur (Schmid 2000: 22):

i) Shell noun + *that-/to-/wh-* clause (N-cl):

(93) 'Mr Bush said Iraq's leaders had to face **the fact** that the rest of the world was against them'

ii) Shell noun + *be* + *that-/to-/wh-* clause (N-*be*-cl)*:*

(94) '**The advantage** is that there is a huge audience that can hear other things you may have to say'

iii) Specific deictic (*the*, *this*, *that*, *other*, *same*, *such*) + (premodifier) + shell noun (*th*-N):

(95) '[…] I hope it is unnecessary to say that **this accusation** is also completely unjustified'

iv) Referring item (*this*, *that*, *it*) + *be* + shell noun (*th-be*-N):

(96) '[…] That was **a great achievement** […]'

A list of 670 shell nouns results from this pattern-based search, thus rendering Schmid's (2000: 383–407) the most exhaustive list of shell nouns compiled to date. The impossibility of providing a complete list of these nouns is recognised, inasmuch as many other potential members may well appear '[…] in suitable contexts […]' (Schmid 2000: 14).

The adoption of these syntactic criteria stems from the use of automated corpus retrieval methods, preferred over manual data collection on account of the sheer size of the corpus used (250 million words). Such methods involve transforming the above patterns into a set of computer-readable queries (Schmid 2000: 44), as in NN+*that*/CS, where 'NN' stands for 'noun' and 'CS' stands for 'subordinating conjunction'. Of the four patterns above, however, only N-cl and N-*be*-cl were queried, on the grounds that nouns occurring in these patterns tend to feature as shell nouns. The occurrence of concrete nouns like *book* or *blackboard* in the patterns *th*-N and *th*-*be*-N (e.g. *this book* and *this is a blackboard*) precluded their initial retrieval from the corpus. Instead, for these two patterns examples were retrieved on the basis of the 670 shell nouns previously identified in N-cl and N-*be*-cl. Instances where only *th*-N and *th*-*be*-N were present in the data were excluded from the count, despite their possible occurrence in N-cl and N-*be*-cl (Schmid 2000: 42).

The lack of flexibility inherent in the automaticity of this method necessarily results in 'systematic misses', as acknowledged by the author (Schmid 2000: 51–3). A major source of exclusion of prototypical instances of shell nouns (estimated by the author to rate 30-40% of excluded examples) concerns the insertion of linguistic material between the head noun and its lexical realisation. This is often the case where a prepositional phrase occurs as postmodifier of a shell noun:

(97) 'The next part of the project is to go back and to identify where these products come from' (Schmid 2000: 52)

The elimination of all *of*-prepositional phrases from the quantitative analysis is explained in terms of their occurrence with a relatively small group of nouns (Schmid 2000: 26)[2]. Further omissions of potential shell nouns from the analysis include the plural forms of likely candidates, as they do not usually appear in the patterns N-cl and N-*be*-cl (Schmid 2000: 53). This property, however, stands in contrast to Ivanič's (1991: 93) finding that carrier nouns '[…] are common in the plural'.

Notwithstanding some of the drawbacks of the retrieval method used, in the second part of his book Schmid (2000) provides one of the most comprehensive and fine-grained semantic classifications of shell nouns to be found in the literature. The classification relies on six superordinate semantic features to label the six semantic categories of shell nouns:

[2] Shell nouns, however, are reported to occur with many other prepositions in addition to *of* (Flowerdew 2006: 350).

i) Factual (e.g. *fact*, *reason*, *aspect*),
ii) Mental (e.g. *belief*, *doubt*, *wish*),
iii) Linguistic (e.g. *statement*, *argument*, *claim*),
iv) Modal (e.g. *possibility*, *permission*, *ability*),
v) Eventive (e.g. *action*, *effort*, *success*), and
vi) Circumstantial (e.g. *situation*, *area*, *way*)

A more detailed explanation of these categories will be offered in 3.2.5. Mention should be made here of Schmid (2007), a paper on the connections and mismatches between the above semantic types and the meanings of *to*-infinitive and *that*-clauses in the N-cl and N-*be*-cl patterns (see 3.2.5.2 for details).

In the last part of his book, Schmid (2000) brings together the syntactic and the functional description of shell nouns in a discussion of the potentiality for each of the four syntactic patterns to act in close association with any of the three functions presented above. To start with, the focus is directed to the semantic effect of characterisation (Schmid 2000: 309–28). This function is most noticeable in the *th*-*be*-N pattern, where overtly attitudinal nouns (e.g. *problem*, *trouble*, *pity*, *tragedy*) prevail. The N-*be*-cl pattern is also remarkable in this respect, though attitudinally neutral nouns are also common (e.g. *fact*, *point*, *idea*). N-cl and *th*-N are, by contrast, less amenable to characterisation and, as such, this function tends to be implied rather than clearly conveyed.

It should also be noted that the function of characterisation is often performed by premodifying adjective phrases. Schmid (2000: 318–19) classifies those adjectives found to premodify shell nouns into five groups:

i) 'descriptive' adjectives,
ii) 'evaluative' adjectives,
iii) 'classifying' adjectives,
iv) 'restrictive' adjectives, and
v) 'cohesive' adjectives.

The difference between the first two types is often hard to ascertain, in that apparently 'descriptive' adjectives such as *big*, *huge* or *small* (normally expressing qualities) may in some contexts carry attitudinal nuances similar to those found in such prototypical 'evaluative' adjectives as *horrendous* or *tremendous*. 'Classifying' adjectives enable nouns to be allocated to objective classes, as in *scientific*, *medical*, *fascist*, etc. 'Restrictive' adjectives such as *main*, *only* or *real* narrow down the reference of the head noun. Finally, 'cohesive' adjectives like *next*, *other* or *different* serve a discourse-organising role.

Characterisation is in turn followed by an examination of the pragmatic, rhetorical and textual functions of shell nouns (Schmid 2000: 329–59): 'focusing and topicalising', 'linking' and 'signposting'. The function of 'focusing and topicalising' (Schmid 2000: 329–37) concerns the organisation of clause constituents into given or known information (i.e. the Topic) and new or unknown information (i.e. the Focus). The *th*-*be*-N pattern (as in (98)) reveals the unmarked or default order, with the subject pronoun acting as Topic (known information) and the shell-noun phrase representing the Focus (new information). The *th*-N pattern, by contrast, places the shell-noun phrase in Topic (or Given) position (see (99)). This is supported by Schmid's (2000: 330) claim that '[…] by far the majority of shell-noun phrases in the pattern *th*-N occur in clause-initial position or after one or two clause-initial adverbials'.

(98) '[…] With all the physiotherapy and swimming she built herself up and has been managing to walk using a frame. That is a **marvellous achievement**' (Schmid 2000: 345)

(99) '[…] the goal of the welfare state, within a society in which economic competition under capitalism dominates, must be to effect gradual reform. **This Fabian approach** argues for […]' (Schmid 2000: 344)

More interesting, from a pragmatic point of view, is the N-*be*-cl pattern. The shell noun in this pattern is often highly general in meaning (e.g. *thing*, *point*, *question*), adding little information to the clause. Its function, as in (100) below, is to focus the listener's attention on the propositional content of the complement clause by highlighting its importance. This is demonstrated by the possibility to paraphrase the subject with a *wh*-cleft sentence, as in (101). The reasons behind the use of focusing in N-*be*-cl are threefold (Schmid 2000: 333–4). Firstly, it is a useful means for bringing to the forefront information that the speaker considers important. Secondly, it enables the speaker to draw a contrast between his/her own current claims and previous ones, thereby implying some evaluation too. Finally, it may manifest itself as a hesitation device, giving speakers time to better express their ideas.[3]

[3] Delahunty (2012: 63–4, 73) further explains the focusinng function on the grounds that the highly unspecific meaning of *thing*, along with the use of the definite article, introduce the presupposition that a specific denotatum exists and that such a denotatum contrasts with and displaces the immediately preceding proposition. This contrast, however, does not entail a change of topic, but only '[...] a change in the expected trajectory of the current topic [...]' (Delahunty 2012: 64). For example, if someone says *so the thing is, it can be done* after a series of arguments against writing in the bath with more traditional means (a typeweriter,

(100) 'Why […] why get rid of people with experience? I think **the thing** is that people have seen technology change dramatically within the business' (Schmid 2000: 332)

(101) 'What is important is that people have seen technology change dramatically within the business' (Schmid 2000: 333)

Data for the N-*be*-cl pattern also include a range of semantically more specific nouns, such as *difference*, *aim*, *answer*, *problem*, *fear* or *feeling*. The use of such nouns is illustrative of the combined effect of focusing and topicalising (Schmid 2000: 334–7). In examples like (102), the shell-content complex both brings into focus the content of the complement clause and assigns an anticipatory peak of prominence to the head noun acting as Topic (cf. Sinclair et al. 1990: 429, in 2.2.1.3.2).

(102) '[…] **My own feeling** is that alpha emitters, like plutonium, are going to prove an important pathway through fathers to their unborn children' (Schmid 2001: 202).

The function of 'linking' (Schmid 2000: 339–48) relates primarily to the use of the *th*-N and *th-be*-N patterns as anaphoric reference items. Drawing on cognitive theories of anaphora (e.g. Prince 1981), it is argued that the subject position of the referring item in *th-be*-N helps to encapsulate information which is already mentally active thanks to the antecedent in the preceding clause or within the same paragraph. For example, in (98) above the shell noun-phrase *a marvellous achievement* refers back to the subject pronoun *that*, which in turn refers back to the information contained in the previous adjoining sentence. A similar situation applies to the *th*-N pattern, particularly with regard to the proximal demonstratives *this/these*, as in (99).

Whilst linking emphasises the anaphoric potential of shell nouns, 'signposting' (Schmid 2000: 349–59) stresses their use as cataphoric markers (cf. Winter 1977: 2 in 2.2.2.1.2, where the notion of signposting is also used). Schmid (2000: 350) claims that this function manifests itself through two main types. Firstly, there are many instances where the shell noun features as an anaphoric signal that summarises the preceding discourse and acts as a topic change marker (cf. Francis 1986: 38 and 1994: 86 in 2.2.2.1.4):

---

a pen, etc.), the conclusion is that, by using a PowerBook, writing in the bath is feasible (Delahunty 2012: 67). *The thing is*, therefore, entails an element of conclusion that does not change the topic of discourse (the possibility of using a PowerBook anywhere), but merely directs it towards a concluding statement.

(103) '[…] the supply of children far exceeds the pool of people prepared to provide them with a home and a family life. The problem was highlighted this week when a newspaper in Oxford published […]' (Schmid 2000: 350)

Secondly, there are cases of cataphoric signposts by which writers organise the internal structure of their texts to guide readers through their argumentation (cf. Tadros 1985 and 1994 in 2.2.2.1.3). This is often associated with plural shell-noun uses:

(104) '[…] Three unresolved issues, currently exercising the government's best minds, all have features in common […]' (Schmid 2000: 356)

Schmid (2000: 360–76) devotes a final chapter to the cognitive functions of shell nouns: 'conceptual partitioning', 'reification' and 'integration'. The first function, 'conceptual partitioning' (Schmid 2000: 360–3), concerns the ability of the N-cl and N-*be*-cl patterns to transform events or abstract relations into temporary conceptual entities. For example, in a sentence like (105), the underlined complement clause represents an event condensed into a single temporary discourse entity, one which at this point in discourse is an *aim*, but which may change as discourse unfolds (hence its temporary nature; e.g. *their aim, this event, this problem*).

(105) 'Their aim is to meet President Saddam in Baghdad' (Schmid 2000: 68).

A direct result of partitioning resides in the function of reification. 'Reification' (Schmid 2000: 363–9) explains how the packaging of complex pieces of information into nouns involves a series of significant advantages not found in clauses. For example, unlike clauses, nouns can act as anchors for the expression of the writer's or speaker's attitudes, as topic initiators in clauses, as anaphoric linkers and as signposts (Schmid 2000: 369). The role of shell nouns as anaphoric signals and signposts underlies the last cognitive function, that of 'integration' (Schmid 2000: 370–6). It is stated that the encapsulation resulting from the use of shell nouns entails a semantic integration whereby cognitively complex pieces of information are incorporated into single semantic units. This function is clearly at play in instances where a plural shell noun appears in the *th*-N pattern, as in these cases the semantic integration relies on long and complex stretches of discourse:

(106) 'The company said yesterday that it would sell or close its 12 remaining abbattoirs, was cutting chicken production from over three million birds a

> week to two million, and had abandoned property trading. **These measures** resulted in an extraordinary charge of pounds 92 million, which wiped out the year's profits' (Schmid 2000: 370)

2.2.2.3 *From description to pedagogy*

The review of the related literature in the previous sections (2.2.2.1 and 2.2.2.2) has contemplated purely descriptive research. The study of shell nouns, however, has also been inspired by a focus on the teaching and learning of these units, as evident in the references discussed below (Francis 1988; Carter 1998; Flowerdew 2002 and 2003a and Hinkel 2004).

Francis (1988) looks at ways of helping foreign learners to improve their use of lexical cohesion in their written production. A small-scale study is also reported on the use of cohesive devices by foreign and native writers.

Lexical cohesion is argued to comprise two main subtypes: 'labels' and 'equivalents'. For a description of the behaviour of advance and retrospective labels (Francis 1988: 326–8), see Francis (1994) in 2.2.2.1.4. 'Equivalents' (Francis 1988: 328–30) account for those instances where, unlike labels, a noun refers back to a single noun phrase. Equivalents fall into three subtypes: 'true equivalents', 'instantial equivalents' and 'processed equivalents'. Synonymy lies at the centre of so-called 'true equivalents', as observed in noun pairs like *notion* and *idea*. 'Instantial equivalents' are synonyms which are not based on the language system, but are created by the writer/speaker in a specific context. *Situation* and *dilemma* in (107) are examples of this type of equivalence. 'Processed equivalents' resemble the former in their context or discourse-based synonymy. They are more complex, however, as their discourse-based change of status echoes real-life processes. This is the case in (108), where the change of status from *militants*, through *culprits*, to *defendants* reflects the real-life judicial process of presumed crime (i.e. *militants*), arrest (i.e. *culprits*) and trial (i.e. *defendants*). Instantial and processed equivalence are claimed to outweigh true equivalence, as meaning does not remain static and is more often than not negotiated and modified as discourse unfolds (Francis 1988: 330).

(107) '[…] I called the ROV and explained **my situation** over the phone. A man said to bring the keys to ROV in Sin Ming Drive and that they would handle it. The following afternoon, I drove to ROV and explained **my dilemma** to a woman there […]' (Francis 1988: 327)

(108) 'The last upheaval in October when **militants belonging to a hitherto unknown 'Party of God'** (Hezbollah) assassinated in Benghazi a member of […] **The seven alleged culprits** had to go through the traditional ordeal of self-criticism on television […] **All the defendants**

> obligingly admitted that "they deserved to die" […]' (Francis 1988: 329–30)

Francis (1988: 334–7) concludes with a set of exercises aimed at raising students' awareness of the use of retrospective labelling. Four main types of exercise are proposed, ranging from those where the students are asked to delimit the stretch of discourse a retrospective label refers to, through those where labels have to be supplied (with and without alternatives), to those where the goal is to provide follow-up sentences for a label used at the beginning of a specific text.

Carter (1998) gives an overview of research on vocabulary from an applied linguistic perspective. One of the chapters (Carter 1998: 79–115) is devoted to the interface between lexis and discourse, reflected in the three main strands of research described in section 2.2.2.1, i.e. Halliday & Hasan's (1976) notion of lexical cohesion, lexical signalling (e.g. Winter 1977 and Hoey 1994) and Francis' (1986) anaphoric nouns. Regarding anaphoric nouns, Carter (1998: 90) suggests that certain items '[…] are more core than others', as they are attitudinally neutral and common to a range of text types. Consequently, items such as *means*, *move*, *issue*, *question* and *fact* are supposedly easier to use than other more specific or evaluative items such as *proviso*, *persiflage* or *absurdity*. In line with Francis (1988: 330) above, Carter (1998: 91) further observes that '[W]ord meaning in discourse is regularly instantial'. Drawing also on Ivanič (1991: 95), this implies that meaning does not always remain constant, but may change once the same lexical item is used in different texts.

Flowerdew's (2002) paper summarises the literature on shell nouns from Halliday & Hasan (1976) up to Winter (1992), with the aim of outlining the main ideas in the form of a succinct pedagogical grammar of what he terms 'signalling nouns'. His 8-point synthesis (Flowerdew 2002: 152–4) may serve as the basis for a more in-depth study of these units:

i) Signalling nouns are semantically unspecific items whose meaning is specified in context.
ii) Signalling is most commonly performed by nouns, but verbs and adjectives can also play a role there (see Winter 1977 in 2.2.2.1.2).
iii) The semantic specification of signalling nouns may be realised across clauses, within the clause or outside the text.
iv) The across-clause function may be anaphoric or cataphoric.
v) Signalling may refer to individual clauses or longer stretches of discourse.

vi) Evaluation is at times implicit in either the shell noun alone (e.g. *distortion*, *fabrication*) or in its modifiers (e.g. *down-to-earth approach*).
vii) In cases of external semantic specification of the shell noun, reference may be either exophoric, based on general background knowledge, or homophoric, based on specific knowledge shared by the interlocutors.

Flowerdew's review of the major aspects covered in the literature on shell nouns is in turn followed by another article where his theory of signalling nouns takes centre stage. Flowerdew (2003a) is an attempt to provide '[…] a pedagogically appropriate description of […] "signalling nouns" […]' (Flowerdew 2003a: 329). The term is applied to '[…] any abstract noun […]' whose meaning can be specified '[…] by reference to its context'. The usefulness of the study lies mainly in better defining the essential features of these nouns (e.g. *attitude*, *assistance*, *difficulty*), with a view to raising learners' awareness of their pervasiveness in academic language (Flowerdew 2003a: 331).

The discussion is structured around three discourse contexts enabling the semantic specification of signalling nouns, i.e. across clauses, within the clause and exophoric. The across-clause function (Flowerdew 2003a: 333–6) may include cases of anaphoric reference, cataphoric reference or both, as in (109) below:

(109) '[…] As a reaction to this extreme **view** a rival **idea** has grown up, the organismal theory, which proposes that the whole organism is the basic entity and the cells merely incidental sub-units' (Flowerdew 2003a: 335)

The in-clause function (Flowerdew 2003a: 336–8) is evidenced by the use of signalling nouns as subjects of relational clauses, where the semantic specification is conveyed as a *that*-clause (110), a *to*-infinitive clause (111) or a deverbal noun (112):

(110) '**The reason why they're green** is that they have chlorophyll' (Flowerdew 2003a: 336)
(111) '[…] **the function** is to produce seeds which will then grow into new plants' (Flowerdew 2003a: 336)
(112) '**Another important structural characteristic of monosaccharides** is the occurrence of isomerism' (Flowerdew 2003a: 336)

Finally, the exophoric function (cf. Ivanič 1991: 103, where it is also considered) entails the use of background knowledge. More often than not,

however, exophoric signals need not be semantically unpacked to facilitate understanding, as in (113):

(113) '[…] This modification of Whitakker's original five kingdom system has certain **advantages** but it will not be used here' (Flowerdew 2003a: 339)

Signalling nouns in any of their three discourse functions are often accompanied by modifiers. These help to narrow down the reference of the head noun, sometimes being '[…] of more semantic importance than the signal […]' (Flowerdew 2003a: 336). Several examples are provided of adjectival premodification, as in (114), clausal postmodification, as in (115), but, more significantly, unlike Schmid (2000: 26), prepositional postmodification, as in (116), is included as one of the possible types of lexical realisation of the head noun:

(114) '[…] During the last forty years or so **such functional studies** have become more and more chemical […]' (Flowerdew 2003a: 335)
(115) 'a **theory that is generally accepted** that mitochondria found in animal and plant cells were originally bacteria […]' (Flowerdew 2003a: 337)
(116) 'its **function** of providing mechanical strength' (Flowerdew 2003a: 337)

As a result of a manual analysis based on the above criteria, Flowerdew (2003a: 341–2) identifies 166 and 112 signalling nouns in his written and spoken biology corpora respectively (cf. section 2.3.2.2). It is, however, admitted that a wider range of items would be likely to appear in larger, less domain-specific corpora (Flowerdew 2003a: 341).

In view of the large number of signalling nouns retrieved from the corpus, special emphasis is placed on their major role in the organisation of scientific discourse. Flowerdew's (2003a: 343–4) contention is that learners should be made aware of these nouns through their study in context. This may be accomplished through Francis' (1988) set of controlled activities (see above). Such activities may nonetheless be further supplemented with hands-on analysis of unsorted authentic concordances.

Whilst Flowerdew (2002) and (2003a) are specifically devoted to signalling nouns, Hinkel's (2004) approach to these nouns falls within the scope of a more general study on techniques for the teaching of vocabulary and grammar in ESL writing. Hinkel (2004: 100) shows how L2 writers of academic English tend to overuse highly frequent simple nouns such as *people*, *man*, *woman*, *stuff* and *thing*, at the expense of more specific lexical

substitutes. It is also noted that L2 writers sometimes draw on highly context-specific and attitudinal nouns like *miracle* or *magic*, which should be avoided in academic prose (Hinkel 2004: 129).

Hinkel's (2004) characterisation of shell nouns is dependent on their enumerative function. The category of 'enumerative "catch-all" nouns' discussed by Hinkel (2004: 135–6, 284–5) is based on Tadros' (1994: 71–3) enumerables. These are described as '[…] lexically simple nouns […]' (Hinkel 2004: 135) referring to stretches of discourse or individual nouns. It is on account of their encapsulating function that they are also termed 'catch-all nouns', as their enumerative meanings enable them to refer to previous or following textual points. In addition to enumeration, these nouns may also act as nominal substitutes contributing to the creation of cohesive chains. In this respect, therefore, enumerative nouns resemble Halliday & Hasan's (1976) general nouns. Example (117) shows a cohesive chain relying on three enumerative nouns (i.e. *issues-problems-challenges*):

(117) 'The author mentions pollution, water shortage, and loss of soil **issues** concerning the threat of overpopulation. In his article, he […] does not mention the health and nutrition **problems**. The health **challenges** are created when there are too many people in the world' (Hinkel 2004: 284)

Hinkel's (2004) brief description of enumerative nouns is illustrated by a list of 63 common instances, out of which 34 nouns are said to be highly prevalent in discourse (e.g. *approach*, *facet*, *reason*, *task*) (Hinkel 2004: 284). It is also suggested that the group of enumerative nouns in English may not comprise more than a hundred instances (Hinkel 2004: 285). This claim may prove to be contentious, in light of the research reported so far.

As regards the pedagogical implications of this approach, it is argued that the explicit teaching of these nouns may help to foster students' awareness of the differences between informal conversation and the highly nominalised written academic registers (Hinkel 2004: 136). The students will consequently be better prepared to tackle the lexical dearth often found in their writing.

2.3 SHELL NOUNS AND GENRE VARIATION

The study of shell nouns has not been based solely on the formal and functional identification of this category. Of particular relevance is also their use in context and, as such, the related literature offers insights into the influence that genre and, specifically, mode (i.e. written and spoken discourse) exert on the occurrence of these lexical items in discourse. Except for Halliday & Hasan (1976) and Winter (1977 and 1982), all the other references in this section make explicit mention of the use of corpora in their research. Section 2.3 relies on such corpora as its basic organising principle. Three main subsections will be considered on these grounds, one on written corpora (2.3.1), one on corpora of written and spoken discourse (2.3.2) and a final one on spoken discourse (2.3.3). Each subsection further encompasses various degrees of genre specificity, ranging from general to genre-specific corpora (e.g. corpora of academic discourse). The reason why the last section concerns spoken discourse lies in the minimal research (only two studies) conducted on this mode alone.

Before moving on to section 2.3.1, it is worth noting Halliday & Hasan's (1976) and Winter's (1977 and 1982) remarks on the genre-specificity of general nouns and Vocabulary 3 items respectively. Halliday & Hasan (1976: 274) suggest that, on account of the highly attitudinal nature of general nouns (e.g. *you crazy fool*), their cohesive function is paramount in spoken language. Winter (1977 and 1982), by contrast, underlines the discourse-organising function of Vocabulary 3 items in written discourse. Most of the examples provided, however, are of a journalistic nature (i.e. newspapers), so his claims concerning written discourse might be skewed towards media language.

2.3.1 *Written discourse*

2.3.1.1 *General prose and journalistic prose*

Of the references included in section 2.3.1, only Ivanič (1991) draws on a general corpus of written discourse. Her use of the 1 million-word *LOB* corpus is nevertheless not inspired by quantitative concerns. As such, no detailed account is provided of the genre-specific frequency of carrier nouns, inasmuch as the corpus is employed as a source of examples to support her theory. Two tentative conclusions reached from the analysis of the examples point to their substantial presence in academic discourse (Ivanič 1991: 113), as well as their common occurrence as exophoric markers in news reporting (Ivanič 1991: 106). The former conclusion is explained in terms of the ability shown by carrier nouns to encapsulate and

summarise long stretches of discourse. The latter stems from the frequent assumption in news reporting that readers are familiar with previous articles, as in (118):

(118) 'Mrs Thatcher's Cabinet has offered to compromise national pride and even modify **her own principles**, as she initially articulated them' (Ivanič 1991: 106).

Journalistic language is precisely at the core of Francis' (1986 and 1994) discussion of A-nouns and labels respectively. The 1986 monograph draws on expository articles from the monthly journal *Encounter* and the 1994 article is based on the *BoE*, specifically on articles from the *The Times* sub-corpus. Both references share with Ivanič (1991) their qualitative rather than quantitative nature, as no numerical data substantiate her claims. Whilst it is argued that the type of cohesion enabled by labels prevails in the discourse of the press, a need is recognised for further studies on the frequency of labels across various spoken and written genres (Francis 1994: 100).

2.3.1.2 *Academic prose*

Academic prose, the other genre identified by Ivanič (1991: 113) as being relevant to shell nouns, features prominently in the genre-specific literature reviewed. A threefold distinction shall be drawn here between research on professional writing (2.3.1.2.1), on comparisons between professional and discipline-specific student writing (2.3.1.2.2) and on argumentative learner writing (2.3.1.2.3).

2.3.1.2.1 *Professional writing*

The scope of the studies reported in this section is not limited to academic prose in general, but covers a range of different academic sub-genres, i.e. textbooks (Tadros 1985 and 1994), research articles (Moreno 2004; Gray 2010 and Kanté 2010a) and theses (Charles 2003 and 2007). Tadros (1994) (see 2.2.2.1.3) examines her predictive categories on the basis of 7 academic textbooks, 3 on economics, 3 on linguistics and 1 on law. Tadros (1985), by contrast, explores only 1 textbook on economics. Expository writing is best manifested in this sub-genre and, as such, textbooks prove of paramount importance for the students' acquisition of the text reporting and writer detachment techniques found in the writing of theses or research papers (Tadros 1994: 81).

Moreno (2004) focuses on 36 economics and business research articles in English and Spanish (72 articles in total) and explores the extent to which

language and discipline influence the use of retrospective labelling in premise-conclusion metatext. In line with Francis' (1994: 86) statement of the '[…] topic-shifting and topic-linking function […]' of retrospective labels (see 2.2.2.1.4), this type of metatext names the statement from which a conclusion follows, as in (119):

(119) '[…] These results would appear to indicate Q […]' (Moreno 2004: 323)

Moreno (2004: 326–7) classifies retrospective labels into 'explicit', 'fuzzy' and 'implicit' labels. 'Explicit' labels may be semantically specified in previous stretches of discourse, as in *these circumstances* or *these results*. 'Fuzzy' labels are more unspecific, as only a demonstrative pronoun is used (e.g. *this* and *all this*). Lastly, 'implicit' labels leave out the referential item from which the conclusion is to be drawn, as most typically found in conjuncts like *as a consequence* or *therefore*.

Explicit labels are further broken down into Francis' (1994) threefold distinction between metalinguistic labels, non-metalinguistic labels and evaluative labels (Moreno 2004: 328–30). Moreno's (2004) metalinguistic labels, however, include two categories not present in Francis' (1994) categories: 'visual unit noun' (e.g. *table 1*, *figure 6*) and 'research-related noun' (e.g. *the statistics*, *this research literature*). The former subcategory relates to Tadros' (1994: 73–4) non-linear advance labelling.

The type of retrospective labelling involved in this study comprises several general degrees of formal realisation (Moreno 2004: 330–1), ranging from least to most complex. Pronouns and modified pronouns (e.g. *this*, *all this*) feature as the least complex. These are followed by combinations of a determiner and a lexical item (e.g. *our analysis*), a determiner plus a modified lexical item (e.g. *the preceding subsection*) and a modified lexical item (e.g. *evidence presented in the previous section*). The latter type illustrates how, as opposed to what is claimed in the literature reviewed in section 2.2.2, the co-occurrence of a specific deictic is not a necessary requirement for shell-noun membership.

The modification of retrospective labels is, like Francis (1994: 95–100), also categorised in terms of Halliday's (1994) three metafunctions. Three types of modifiers are therefore examined: ideational modifiers (e.g. *large*), interpersonal modifiers (e.g. *significant*) and textual modifiers (e.g. *similar*). Particularly noteworthy is Moreno's division of ideational modification into five categories inspired by the nature of the corpus used: 'dimension' (e.g. *relatively large*), 'quantity' (e.g. *three*), 'relative location' (e.g. *preceding*), 'precise location' (e.g. *in table 2*) and 'class' (e.g. *empirical*).

The results obtained from the application of the above criteria reveal a greater frequency of implicit labels in both Spanish and English (Moreno 2004: 334). More instances of fuzzy labels, however, occur in the Spanish corpus. Semantically, economics and business articles in both languages appear to favour research-related metalinguistic labels, though differences arise in visual unit nouns and non-metalinguistic nouns, with Spanish preferring the former and English preferring the latter (Moreno 2004: 334–5). Finally, the modification of the labels found in both corpora proves to be mainly ideational. The English articles are nonetheless said to be more attitudinal, as retrospective evaluative labels (e.g. *such an error*, *these inconsistencies*) are more frequent there.

Gray (2010) also draws on research articles, but the aim here is to look at the use of sentence-initial *this*/*these* as pronouns and determiners, as in (120) and (121) respectively:

(120) '[…] **This** should not come as a surprise […]' (Gray 2010: 169)
(121) '[…] **This trend of rising interregional income disparities** was the main force behind the dramatic rise in global inequality […]' (Gray 2010: 170)

On the basis of a 480,989-word corpus of 40 published research articles in education and sociology (20 for each discipline), the analysis concentrates on the types of verbs following pronominal *this*/*these* (i.e. copular and non-copular) and on the types of nouns occurring with the demonstrative determiners. Interestingly, attention is also paid to the nature of the structures preceding both pronominal and determinative uses. Antecedents are reported to fall into those based on 'local discourse' and those based on 'global discourse' (Gray 2010: 173–4). 'Local antecedents' materialise in noun phrases (simple: with no postmodification; complex: with postmodification; see (122)) and sentences or clauses (see (123)).

(122) '[…] developed a domain-specific measure to assess students' beliefs about knowledge in history and mathematics. **These specific domains** were selected to address beliefs [...]' (Gray 2010: 174)
(123) 'For example, mid-career entrants were more than three times as likely as their first-career counterparts to move from one school to another. **This comparison** suggests that, as a group, they were fickle or unstable' (Gray 2010: 174)

'Global antecedents', on the other hand, contain longer stretches of discourse, and comprise cases of 'extended preceding discourse' and cases of 'overall discourse'. The former corresponds to instances where a specific segment of text can be identified and where two or more sentences

are involved (see (124)). Extended preceding is also applicable to examples where the exact boundaries of the antecedent are difficult to delimit, but where reference is still made to information contained in the preceding text. Finally, overall discourse constitutes reference to the entire text or the study as a whole (see (125)).

(124) 'We use only nonviolent protest by African-Americans, including public demonstrations and marches, sit-ins, rallies, freedom rides, boycotts, and other protest actions. We exclude riots, melees, and racial confrontations that lacked a clear protest quality as well as routine institutional actions […]. This means our counts differ from the total of "movement-initiated" events' (Gray 2010: 173)

(125) 'Schema-based transfer instruction (SBTI) explicitly teaches transfer features that change problems in superficial ways to make them appear novel even though they still require known solution strategies. This study assessed the effects of an expanded version of SBTI […]' (Gray 2010: 173)

Gray's (2010) account of shell nouns lies in their inclusion within a fourfold taxonomy of nouns modified by the demonstrative determiner: 'concrete nouns', 'species nouns', 'abstract nouns' and 'shell nouns' (Gray 2010: 172). 'Concrete nouns' represent tangible entities from the world outside the text, as in *newspaper*, *teacher* or *student*. 'Species nouns' such as *type*, *kind* and *form* collocate with an *of*-phrase which restricts the reference of the head noun. Finally, 'abstract nouns' and 'shell nouns' are distinguished on the basis of Schmid's (2000) list. Only those instances appearing in Schmid's (2000) 670-unit list are coded as shell nouns.

The analysis of data shows that, despite a marked preference for determinative uses (76.15%), pronominal instances of *this*/*these* are not absent from the study corpus (23.85%; Gray 2010: 174). Disciplinary variation here is minimal, with a slightly higher frequency of pronouns in sociology and a slightly higher frequency of determiners in education. With regard to the taxonomy of nouns, abstract and shell nouns clearly stand out over concrete and species nouns (Gray 2010: 177). Disciplinary variation is once again marginally significant, with education tending to use more shell nouns and sociology featuring more examples of abstract nouns.

Three semantic types of shell nouns stem from the examination of the study corpus (Gray 2010: 178): 'results/findings' (e.g. *finding*, *result*, *difference*), 'verbal/mental' (e.g. *hypothesis*, *argument*, *assumption*) and 'other' (e.g. *approach*, *possibility*, *role*). Abstract nouns, however, do not lend themselves to such a semantic classification, on the grounds of their wide-ranging variety. Gray's discussion of the noun taxonomy is further complemented by data illustrating the extent to which the four members of

the taxonomy rely on pre- and postmodification (Gray 2010: 178). The evidence indicates a scanty use of modification in the study corpus, with concrete nouns showing barely any, and abstract and shell nouns featuring only a few more instances of premodification than postmodification.

In relation to the types of antecedents (Gray 2010: 179), pronouns show a marked preference for clausal antecedents (70%), whilst abstract nouns are evenly distributed across all four types. Shell nouns appear to favour clausal and extended antecedents (35% and 32% respectively), but more significantly, they also feature 33% of instances of phrasal antecedents (26% corresponding to complex noun phrases and 7% to simple noun phrases). The remarkable frequency of longer antecedents agrees with the frequent claim about the encapsulating potential of shell nouns. The occurrence of phrasal antecedents, however, seems to imply that, in line with Ivanič (1991: 110), shell nouns may also refer to individual phrases.

Another academic genre that has also received attention in the study of shell nouns is that of theses. In this respect, Charles (2003) and (2007) revolve around the stance implications of shell nouns, as evidenced in a corpus of 8 MPhil theses in politics and international relations (about 200,000 words), and 8 PhD theses in materials science (about 300,000 words), all by native speakers of English. Charles (2003) deals with sentence-initial retrospective labels preceded by *this*, as in (126):

(126) '[…] This mystery is what makes the model so fascinating […]' (Charles 2003: 315)

The overall occurrence of these labels is slightly higher in the politics corpus (76.8 tokens vs. 60 tokens per 100,000 words; Charles 2003: 316). Such a minor difference, however, proves more noticeable when considering head noun types, with politics containing a much wider range of head nouns than materials science (56.8 types vs. 33.0 types per 100,000 words). With regard to Francis' (1994) dichotomy between metalinguistic and non-metalinguistic labels, the findings suggest that metalinguistic labels are twice as frequent in Politics as in the Materials corpus (Charles 2003: 317). This is indicative of the text-based nature of politics, as opposed to the empirical and experimental nature of materials science.

It is argued that the choice of labels is not random, but often underpinned by the writer's perceptions. The discourse-organising function of shell nouns is therefore also a means of conveying a stance that tends to be imposed by the writer's familiarity with the concerns of a given academic discipline (Charles 2003: 318). For example, in (127) and (128), the retrospective labels *this process* and *this claim* show the writer's discipline-specific stance,

as involved in the encapsulation of discourse whose labelling would most likely differ if applied by a non-specialist:

(127) 'As the applied field moves beyond its maximum and starts to fall the screening currents within the sample respond [….]. This process continues as the field drops further until the state shown in (c)' (Charles 2003: 318)
(128) 'Policies aimed at raising the stakes faced by a state are likely to produce convergence [….] This claim is made on the basis of the conceptual analysis offered here' (Charles 2003: 318)

Stance is further examined in terms of the twofold distinction proposed by Biber et al. (1999: 972, 974) between 'epistemic' and 'attitudinal' stance (Charles 2003: 320). 'Epistemic stance' entails an assessment of the validity of a proposition, while 'attitudinal stance' highlights the writer's feelings towards the information presented. The retrospective labels in both corpora show a clear preference for epistemic stance, as in *claim*, *theory*, *explanation* and *assumption*. Attitudinal stance (e.g. *limitation*, *atrocity*, *dilemma* and *disjointedness*), though low in both corpora, is slightly more common in the politics corpus.

Charles (2007) takes a different perspective on stance, as the focus is now shifted to the N-*that* pattern. Drawing on Francis et al. (1998: 108–13; see 2.2.1.3.2), shell nouns in this pattern are assigned to five semantic groups (Charles 2007: 208): 'idea' (e.g. *idea*, *assumption*, *belief*), 'argument' (e.g. *argument*, *contention*, *point*), 'evidence' (e.g. *evidence*, *indication*, *observation*), 'possibility' (e.g. *possibility*, *probability*, *chance*) and 'other' (e.g. *fact*, *case*, *concern*). Significant differences arise in the extent of use of these types in both corpora, politics favouring argument and idea nouns, and materials science preferring evidence nouns. This seems to support the assumption that social sciences rely on understanding and interpretation, as in (129), while hard sciences depend on experimental evidence to prove or reject a hypothesis, as in ((130); Charles 2007: 208–9):

(129) 'It is my contention that pressure from parts of the Labour Party and Movement was not central to the decision taken to withdraw' (Charles 2007: 208)
(130) 'Despite the limited set this provides convincing evidence that growth was halted by quenching and then re-initiated' (Charles 2007: 209)

In addition to the semantics of head nouns, the article also explores complement clauses in terms of the type of source from which propositions originate (Charles 2007: 210–12). Three sources emerged from the analysis

of the study corpus: 'self source', 'research source' and 'non-research source'. 'Self source' lies in the writer's research, as in (131) (i.e. *this*), 'research source' stems from other academic references, as in (132) (i.e. *recent academic work*), and 'non-research source' is based on people or entities not present in the research process, as in (133) (i.e. people *on the ground*):

(131) '[…] This lends strong support to **the supposition** that it is the ratio of the applied current […]' (Charles 2007: 210)
(132) 'Recent academic work takes **the attitude** that the middle power idea must still be dealt with […]' (Charles 2007: 210)
(133) 'This disjointedness led to trouble on the ground, as it produced **the belief** that UNPROFOR was an ally of the Muslims [...]' (Charles 2007: 211)

The frequency data suggest a marked contrast between both corpora (Charles 2007: 211), especially with respect to non-research sources, which are absent from materials science but rank highest in politics. As to research sources and self sources, the former feature more frequently in politics, and the latter figure more prominently in materials science. The implication here seems to be that politics builds on other sources as an integral part of the presentation of the writer's arguments. Materials science, by contrast, draws conclusions on the basis of the experimental evidence obtained and, as such, there may be less need to constantly refer to other sources.

The study concludes with a combined analysis of semantic types and sources (Charles 2007: 212). Two clear tendencies appear in the data: one for the combination in politics of argument and idea nouns with research sources and the other for the combination in materials science of possibility and idea nouns with self source.

Charles' (2003 and 2007) interest in the use of noun complement clauses as stance-marking devices is also the focus of Kanté's (2010a and 2010b) research. Kanté's (2010a) paper draws on data from two markedly different corpora, one comprising 16 research papers from the *Journal of English Linguistics* (100,000 words) and the other a 100,000-word sample of a larger diachronic corpus of court proceedings (*the Old Bailey Proceedings corpus*, 25 million words)[4]. The main thrust of Kanté's (2010a) argument lies

[4] Diachronic evidence is also used in Schmid (1997), where the noun *idea* is investigated in terms of its metaphorical conceptualisation from the 16th to the 20th centuries (from object-like meanings only, as in *have an idea*, to metaphorical associations with visible objects, as in *bright idea*, artefacts, as in *wrap up an idea*, etc.).

in the genre-specific modal meaning implicit in the use of shell nouns in noun complement *that*-clauses. The discussion revolves around three types of stance-marking shell nouns (Kanté 2010a: 131): 'epistemic' (knowledge) nouns (e.g. *allegation*, *assumption*, *doubt*), 'deontic' (obligation, necessity) nouns (e.g. *order*, *regulations*, *rule*) and 'attitudinal-evaluative' nouns (e.g. *advantage*, *flaw*, *problem*).

The findings reveal a pronounced tendency for noun complement clauses to occur in linguistics research papers (72% vs. 28% in the courtroom corpus; Kanté 2010a: 130; cf. also Biber et al. 1999: 648, where academic prose contains the largest number of noun complement clauses in their corpus). As regards the semantic classification, epistemic nouns are slightly more frequent in the linguistics corpus, while deontic nouns figure only in the courtroom corpus and attitudinal-evaluative nouns only in the linguistics corpus (Kanté 2010a: 131).The deontic orientation of the courtroom corpus relates to the authoritative nature of the legal institution, while the more epistemic and attitudinal orientation of the linguistics corpus is more in line with the '[…] persuasive and non-authoritative function' of research papers (Kanté 2010a: 134).

Kanté (2010b) delves further into modality-marking in shell nouns in a paper comparing the choice of the indicative or the subjunctive mood in complement *that*-clauses in French and English. The study uses the written section of the *BNC* (90 million words) and *Franztext*, a 210 million-word diachronic corpus of literary and academic prose. The results for both languages (Kanté 2010b: 287) reveal a semantic continuum whereby the closer a noun is to the epistemic end, the more likely it is to choose the indicative mood in the *that*-clause, and the closer it is to the deontic end, the more likely it is to choose the subjunctive mood (or *should* in English) in the *that*-clause. Example (134) illustrates the former and (135) the latter. Kanté (2010b) does not offer any insights into the genre-specific distribution of different semantic types.

(134) 'If he suspected Pascoe, Rain had to concede the probability that he was right' (Kanté 2010b: 285)

(135) 'In trying to find answers, managers – whether they were heads or governors – were caught between the need to make progress and to prepare for broad changes which had been widely publicized by central government and to prepare for the necessity that they should lead changes' (Kanté 2010b: 286)

2.3.1.2.2 *Professional and discipline-specific student writing*

Overall, the references presented so far (Tadros 1985 and 1994; Charles 2003 and 2007; Moreno 2004; Gray 2010; and Kanté 2010a) account for various aspects underlying the use of shell nouns in advanced native academic writing (published and unpublished). This section describes four studies that pave the way for the learner writing references reported in 2.3.1.2.3: Aktas & Cortes (2008), Flowerdew (2008), Caldwell (2009), Lin (2012) and Nesi & Moreton (2012). They share with the research reported so far their focus on discipline and genre-specific writing. The difference lies in their comparison between:

i) Professional writing and discipline-specific undergraduate writing (Flowerdew 2008 and Caldwell 2009),
ii) Professional writing and discipline-specific graduate writing (Aktas & Cortes 2008 and Lin 2012), and
iii) Native and non-native discipline-specific undergraduate writing (Nesi & Moreton 2012).

Flowerdew (2008) explores how the signalling of the Problem-Solution discourse pattern (cf. Hoey 1983 in 2.2.2.1.2) differs between two corpora, one of professional writing and the other of undergraduate writing (totalling 225,000 words each). The former contains 60 professional recommendation reports commissioned by several environmental agencies, and the latter comprises 80 recommendation reports on a range of issues produced by 2nd and 3rd year undergraduate engineering students in an ESP course at a Hong Kong university. The analysis rests on Martin's (2000) Appraisal system, whereby lexis is claimed to fall into 'inscribed' and 'evoking' options (Flowerdew 2008: 33–4). 'Inscribed' items are those whose inherent meaning is evaluative; for example, *problem* is intrinsically negative while *solution* is intrinsically positive. 'Evoking' items are only evaluative through their contextual associations. Nouns like *noise*, *construction* and *traffic*, for instance, are attitudinally negative in the professional corpus, where the aim is to offer suggestions for the protection of the environment. The signalling of the Problem-Solution pattern through these items is markedly different in both corpora (Flowerdew 2008: 49–50). Professional writers rely on a variety of inscribed and evoking items to signal the components of the pattern (e.g. *mitigation*, *measures*, *impact*, *discharge*), while students show an over-reliance on inscribed items (e.g. *problem*, *need*, *result*, *solution*). This may be due, on the one hand, to the more specialised nature of the professional corpus, resulting in the use of more concrete and discipline-specific nouns to refer to problems or solutions. On the other hand, however,

students' repetition of key inscribed signals may be explained by a more limited vocabulary range than that of professional writers (Flowerdew 2008: 50).

Following the identification of inscribed vs. evoking items, the analysis turns to the possible connections between clause relations (e.g. reason-result, means-result, condition-consequence) and lexical signals in the Problem and Solution stages of the pattern. The most remarkable finding in this respect concerns the close association between the noun *problem* and the reason-result relation, as in (136), where *works at the tunnel portal* represents the reason and *a noise problem* the result (Flowerdew 2008: 57). While the student corpus uses this clause relation to a similar extent as the professional corpus, the range of verbs in the pattern (e.g. *cause*, *create*, *pose*) is more limited and errors, as in (137), are bound to occur (Flowerdew 2008: 113). On the whole, Flowerdew's study underlines the importance of examining discourse patterns from '[…] a multifaceted approach' (Flowerdew 2008: 19), one which combines the phraseological analysis typical of corpus-based methodologies (collocations and colligations) with a more functional analysis of examples of specific genres.

(136) '[…] works at the tunnel portal will create a noise problem' (Flowerdew 2008: 58)
(137) '*The problem seems to be arised out of the fact that […]' (Flowerdew 2008: 100)

Compared to Flowerdew (2008), where the research scope is limited to the signalling of a specific discourse pattern, Aktas & Cortes (2008), Caldwell (2009), Lin (2012) and Nesi & Moreton (2012), while also based on small genre-specific corpora, offer more general insights into the overall use of shell nouns. Aktas & Cortes (2008) look at shell nouns in two corpora, one of published research articles (166 texts, 721,553 words) and another of research articles by non-native MA and PhD students (28 texts, 66,459 words). The corpora represent six academic disciplines: art and design, biology, computer science, economics, environmental engineering, and physics and astronomy. The analysis of data, however, is solely concerned with the dichotomy between native and non-native writing. As a result, no attempt is made to identify the extent of use of shell nouns across the disciplines contained in both corpora.

The identification of shell nouns rests on Hinkel's (2004: 284) list of 34 'highly prevalent' enumerative nouns in writing (see 2.2.2.3). Out of these nouns, the selection is further restricted to the 6 most frequent items in the published writing corpus: *effect*, *result*, *fact*, *system*, *process* and *problem*. The aim is to survey the lexicogrammatical and functional properties

displayed in both corpora. The analysis of such properties draws on Schmid's (2000) 4 syntactic patterns and 3 functional categories (see section 2.2.2.2).

The results reflect a general greater frequency of Hinkel's (2004) nouns in the student corpus (549 vs. 274 in 100,000 words) (Aktas & Cortes 2008: 7). This finding should nonetheless be treated with caution, insofar as the high frequency of some of the nouns in the student corpus is by only 1 or 2 students out of 28. This implies that professional writing shows fewer tokens, but the frequency of types is '[…] better distributed across authors' (Aktas & Cortes 2008: 9).

As regards Schmid's (2000) patterns, the data contain instances of only 2 of his 4 possible combinations: N-cl and *th*-N. The shell noun *fact* illustrates a considerable difference with respect to these 2 patterns. It is observed to dominate in the cataphoric N-cl pattern (i.e. *the fact that...*) in the published writing corpus, but in the anaphoric *th*-N pattern (i.e. *this fact is...*) in the student corpus (Aktas & Cortes 2008: 10). Interestingly, in addition to N-cl and *th*-N, 5 new patterns unaccounted for in Schmid (2000) appear in the corpora (Aktas & Cortes 2008: 10):

i) Definite article + Noun (*the*-N):

(138) '[…] the lensing properties of **the system**'

ii) Definite article + Noun + *of*-prepositional phrase (*the*-N-*of*):

(139) '[…] **the process of** administering the service was different […]'

iii) Indefinite article + Noun (*a*/*an*-N):

(140) '[…] **an effect** we interpret as retail price fixity'

iv) Indefinite article + Noun + *of*-prepositional phrase (*a*/*an*-N-*of*):

(141) '[…] one way to begin **a process of** empowerment might be […]'

v) Definite article + *same* + Noun (*the* same-N):

(142) '**The same result** was obtained with ramets […]'

Two of these new patterns are frequent in both corpora: *the*-N and *the*-N-*of* (Aktas & Cortes 2008: 10–11). The former is most common with the noun *system*, while the latter is often found with the noun *effect.*

The lexicogrammatical analysis is linked to a functional analysis based on Schmid's (2000) three functions of characterisation, temporary concept-formation and linking (Aktas & Cortes 2008: 11–13). Characterisation (as in (143)) is reported to apply to N-cl with the noun *fact* in the published writing corpus and to the *the*-N-*of* pattern with the other five nouns in the student corpus. Temporary concept-formation (as in (144)) seems to prevail in the nouns *effect*, *results* and *process* with the *the*-N-*of* pattern in both corpora. Finally, linking (as in (145)) relies mainly on *th*-N, with published writers using the nouns *effect*, *results*, *system* and *process*, and students using *fact* and *problem*.

(143) '[…] The problem of this technique, [is that] it does not work if […]' (Aktas & Cortes 2008: 11)
(144) 'The process of administering the survey was slightly different […]' (Aktas & Cortes 2008: 12)
(145) '[…] This fact suggests that only part of chemisorbed ammonia is […]' (Aktas & Cortes 2008: 13)

Caldwell (2009) analyses professional native writing and undergraduate native (L1) and non-native (L2) writing respectively in order to ascertain whether shell nouns have any influence on the lexical vagueness typical of student writing. The study sample, totalling about 80,000 words (39 texts), comprises two corpora of writing by native and non-native undergraduate students on a Cognitive Psychology course at a South-African university. The written production contained in these two corpora corresponds to an assessed essay that all students enrolled in this course were required to submit. The professional writing corpus is made up of published articles from prestigious Cognitive Psychology journals.

Caldwell's (2009) analytical approach stands in marked contrast with the pattern-based research discussed in 2.2.2.2. She emphasises that Vendler's (1968) and Schmid's (2000) syntactic identification of shell nouns is overly restrictive, as no account is taken of abstract nouns which, whilst infrequent in Schmid's (2000) set of automated predefined patterns (N-cl and N-*be*-cl), may nonetheless be found to encapsulate stretches of discourse (Caldwell 2009: 66). It is further argued that many so-called prototypical shell nouns occur frequently in patterns disregarded by Schmid (2000).

Ivanič's (1991) list of 33 carrier nouns is utilised to investigate their overall distribution across the three corpora. The data show that carrier nouns are significantly more prevalent in the professional corpus than in either of the two undergraduate corpora (Caldwell 2009: 71). Further differences between the three corpora arise after the examination of a range of syntactic patterns. In this respect, a decision is made to explore the patterns

in Schmid (2000), as well as a set of those from his group of 'systematic misses' (Schmid 2000: 51–3; see 2.2.2.2). As regards Schmid's (2000) patterns, only *these*-N and *th-be*-N differ between the three corpora. The former pattern (e.g. *these findings…*) is significantly more frequent in the L1 corpus, while the latter (e.g. *this is the crucial issue*) is only present in the professional and the L1 corpus (Caldwell 2009: 88–9). Its absence from the L2 corpus is explained in terms of its considerable complexity for non-native learners, as its clear characterising potential (see 2.2.2.2, Schmid 2000: 309–28) forces the writer to use overtly attitudinal shell-noun phrases (e.g. *crucial issue*, *pity*, *tragedy*, etc.), and this may require an advanced level of competence in the language. In relation to the patterns excluded from Schmid (2000), mention is made of the following (Caldwell 2009: 81–94):

i) (Determiner) + (Modifier) + Noun + *which* (e.g. *the system which*…)
ii) (Determiner) + (Modifier) + Noun + *in which* (e.g. *the context in which we*…)
iii) (Determiner) + (Modifier) + Noun + zero relativiser (e.g. *the way we respond to*…)
iv) (Determiner) + (Modifier) + Noun + *of*-phrase (e.g. *the nature of the responses to*…)
v) (Determiner) + (Modifier) + Noun + *of* + *wh*-clause (e.g. *a question of whether*…)

Two of the relative clause patterns, i.e. i) and iii), reveal a tendency to occur frequently in L1 and L2 student writing (Caldwell 2009: 82, 90). The more substantial occurrence of the zero relativiser in both undergraduate corpora may be due to the more informal and colloquial tone adopted by novice students in their academic writing (Caldwell 2009: 90). The other relative clause pattern, ii), shows the opposite tendency, standing out only in professional and L1 writing (Caldwell 2009: 94). With respect to the two *of*-phrase patterns, iv) and v), no statistically significant difference is observed across the three corpora, but the results indicate that pattern iv) is '[…] clearly the most prolific of all the patterns extracted as sources of potential shell nouns' (Caldwell 2009: 93).

The first part of Caldwell's (2009) study is, as evident from the above, fully automated, as the chief aim is to determine the distribution of a range of syntactic patterns across the three corpora. This is in turn followed by a more manual and context-sensitive analysis of text. The analysis consists in the identification and coding of all definite noun phrases in the first 200 words of the texts in the three corpora. The coding scheme comprises six categories (Caldwell 2009: 117–18):

i) ESP (expansion/specification) is applied to instances where the noun is postmodified. ESP includes any postmodifying or complementing structures, ranging from prepositional phrases (as in (146)), through relative clauses, to noun complement clauses.

(146) '[…] the analogical abilities (ESP1) *of young children* have challenged […]' (Caldwell 2009: 117, typography as in the original)

ii) A (anaphoric) concerns cases where the specification of a given noun is found in the preceding text:

(147) '[…] that [reasoning by analogy is a developmentally sophisticated skill] (A1). (A1) This notion arose […]' (Caldwell 2009: 117, typography as in the original)

iii) TXT (text reference) entails metadiscursive reference to the paper/assignment itself:

(148) 'We begin (TXT) this article with […]' (Caldwell 2009: 117, typography as in the original)

iv) GEN (generic) applies to nouns for which no textual co-reference exists on account of their relation to general classes of individuals:

(149) '[…] an interplay between (GEN) the world-information available in (GEN) the environment' (Caldwell 2009: 118, typography as in the original)

v) NFS (no further specification) is used to annotate those instances for which no specification appears in the surrounding co-text, and which do not fall into any of the above categories:

(150) '(NFS) The language is a barrier toward Respondent A as he may not fully understand (NFS) the task [...]' (Caldwell 2009: 118, typography as in the original)

Of these patterns, ESP occupies a prominent role in professional writing, compared to the two undergraduate corpora (Caldwell 2009: 121). The conclusion here seems to be that packaging at the level of the noun phrase is a feature common to professional academic writing and, as such, novice writers are expected to learn to produce texts which show increasingly more nominal formality and complexity. This does not mean that students do not make use of complex noun phrases; they do, but their noun phrases are

often too vague. Such vagueness results from the occurrence of unspecific shell nouns at both head and postmodifier levels (Caldwell 2009: 134). Example (151) illustrates this. This particular instance contains two abstract definite noun phrases: *responses* and *tasks*. The definite article in *the tasks* misleads the reader into believing that the nature of these tasks is already known or active in the text. Nevertheless, none of the nouns in this phrase, i.e. *responses* and *tasks*, have been previously specified, forcing the reader to keep two unspecific referential expressions in short-term memory. Example (146) above, however, proves easier to process, inasmuch as only one of the two nouns is unspecific: *abilities*. The reader will therefore need only further information to understand *abilities*, but not *young children*, which is already highly specific and informative.

(151) '(ESP4) the different responses (ESP4) *to* (NFS) *the tasks*' (Caldwell 2009: 132, typography as in the original)

Caldwell's (2009) study thus proves that vagueness in academic writing does not follow from the mere occurrence of abstract nouns, but from '[…] particular ways of *using* "abstract nouns" […]' (Caldwell 2009: 177, italics as in the original).

Lin (2012) looks at shell-noun use in applied linguistics abstracts by professional and non-native post-graduate and graduate scholars. Her corpus comprises 40 abstracts, 20 from six prestigious applied linguistics journals (3,223 words) and 20 from the proceedings of a conference intended for post-graduate and graduate students (4,039 words). Following Aktas & Cortes (2008), sampling is restricted to Hinkel's (2004) list of 34 shell nouns. Of these nouns, 8 are significantly more frequent in the professional corpus: *result*, *effect*, *change*, *type*, *task*, *finding*, *method* and *measure* (Lin 2012: 231). An examination of shell-noun use in the signalling of rhetorical moves in the abstracts (Background-Purpose-Method-Results-Conclusion) reveals further that professional texts include a greater number of shell nouns than non-native ones (Lin 2012: 238). This indicates that student writers are not fully aware of the useful encapsulating function served by shell nouns in such a concise academic text type.

Nesi & Moreton (2012) differs from the above-mentioned references in leaving professional writing out of the comparison. In this paper, the aim is to compare shell-noun use in discipline-specific non-native writing with similar writing produced by native undergraduate speakers. The analysis uses the *British Academic Written English Corpus* (henceforth, *BAWE*), a 6.5 million-word collection of high-quality assessed native and non-native undergraduate writing distributed across 4 broad disciplines (Arts and Humanities, Social Sciences, Life Sciences and Physical Sciences) and 4

levels of study (levels 1, 2 and 3: undergraduate; level 4: masters). Nesi & Moreton (2012) is inspired by Aktas & Cortes (2008) and, as such, it contrasts the findings in both.

Based on Hinkel's (2004: 284) list of 34 shell nouns (as in Aktas & Cortes 2008 and Lin 2012), no significant difference arises in their frequency of occurrence in native and non-native writing (Nesi & Moreton 2012: 131). The scope of the analysis is subsequently narrowed to the 9 most frequent nouns in *BAWE* and in Aktas & Cortes' (2008) published writing corpus: *change*, *effect*, *fact*, *factor*, *method*, *process*, *problem*, *result* and *system* (Nesi & Moreton 2012: 133). An examination of the 8 formal patterns identified in Aktas and Cortés (2008: 10) shows N-cl, *th*-N and *a/the*-N as the most frequent alternatives both in native and non-native writing (Nesi & Moreton 2012: 135).

In search of possible differences between both groups of students, the analysis then turns to the noun *fact*. The evidence indicates that *th*-N is more frequent in non-native writing, while N-cl is more common in native production (Nesi & Moreton 2012: 136). This being the case, however, N-cl is proportionally the most frequent pattern for both types of writing, failing to confirm Aktas & Cortes' (2008: 10) finding about non-native writers' preference for *th*-N. Further inspection shows that *th*-N is more frequent in unmarked subject position in non-native writing (as in (152)); native students, by contrast, prefer post-verbal positions (as in (153); Nesi & Moreton 2012: 137). This seems to suggest that native writing will tend to place highly unspecific nouns such as *fact* in Rheme positions, leaving Theme for more specific or informative elements.

(152) 'Nepal is literate, whether in the UK it is not a concern (appendix 2). This fact has a major impact over the human resources […]' (Nesi & Moreton 2012: 138)

(153) 'Emphasis was put on their undesirable behaviour, for example crime, and prostitution, despite the fact that much of this was not carried out by immigrants' (Nesi & Moreton 2012: 137)

2.3.1.2.3 *Learner writing*

Most research on shell nouns reports on native standards of use, but, as section 2.2.2.3 suggests, non-native learners of English require explicit pedagogic strategies for the proper understanding of such powerful discourse-organising items. In view of such a need, care must be taken to develop an appropriate pedagogy that builds on native-like uses of shell nouns, but also on their use in non-native writing. The identification of problems and areas for improvement in learner writing has therefore been considered essential for the adoption of appropriate remedial techniques.

This section addresses such problems, as identified by Petch-Tyson (2000), Flowerdew (2006), Flowerdew (2010) and Hasselgård (2012). Unlike the references in 2.3.1.2.2, based on discipline-specific academic writing, the study corpus in the following references draws on undergraduate argumentative writing.

Petch-Tyson (2000) examines the use of demonstrative pronouns and determiners in a 100,000-word sample of non-native (L2) and native (L1) writing. The 50,000-word L2 sample is obtained from the Dutch, French, Finnish and Swedish sections of the *International Corpus of Learner English* (henceforth, *ICLE*) and the 50,000-word L1 sample from the American English component of the *Louvain Corpus of Native English Essays* (henceforth, *LOCNESS*). Both corpora contain argumentative essays written by undergraduate students.

Petch-Tyson's (2000) primary research focus is on Fraurud's (1992) concept of 'situation reference' (see 3.2.2). This kind of reference is often established between demonstrative expressions and '[…] higher-order, non-nominal antecedents' (Petch-Tyson 2000: 45), as often claimed in the literature on shell nouns. The analysis involves a sample of 150 demonstrative expressions from each corpus. Some of these expressions are then tested against the whole corpus for confirmation of the results obtained from the small-scale study (Petch-Tyson 2000: 51).

The findings suggest that situation reference is only frequent among native writers, non-native students preferring reference to other noun phrases (Petch-Tyson 2000: 54). Certain differences emerge from the non-native sub-corpora, the most revealing being the larger proportion of demonstrative noun phrases (i.e. *th*-N) in the French texts (Petch-Tyson 2000: 54). Their frequency, however, is no indication of native or near-native use. French and, by extension, all non-native texts use demonstrative nominal anaphors to nominalise verbal processes previously mentioned in discourse, while L1 writers use them as metadiscursive labels applied to previous stretches of discourse (Petch-Tyson 2000: 58). The L2 use of these nouns is therefore cohesive by near-repetition (as in (154): *go sideways>this step sideways*) and the L1 use is cohesive by a more evaluative encapsulation, the one associated with shell-noun phrases (as in (155): *this tactic<it is important to stress* […]).

(154) 'No, I think we should keep on going forward, but also a little bit sideways in order to come closer to Nature. But this step sideways cannot be taken too quickly' (Petch-Tyson 2000: 58)

(155) 'It is important to stress the terrible accidents that occur with other power plants when arguing for nuclear power. This tactic is very good when one looks at the entire issue at hand' (Petch-Tyson 2000: 59)

Flowerdew (2006) proposes a taxonomy of signalling noun errors detected in a 110,000-word sample drawn from a 390,587-word corpus of assessed argumentative writing by Cantonese-speaking first-year undergraduates (included in *ICLE*). The taxonomy comprises four main categories: 'colligation errors', 'incorrect signalling nouns', 'collocation errors' and 'missing signalling nouns'. 'Colligation errors' (Flowerdew 2006: 350–2) result from wrong combinations of lexical and grammatical words and tend to recur in the prepositions acting as heads in postmodifying prepositional phrases, as in (156). 'Incorrect signalling nouns' (Flowerdew 2006: 352–5) constitute cases of semantic misselection of the signalling noun, as in (157), as well as cases of derivational confusions, as in (158). 'Collocation errors' (Flowerdew 2006: 356–7) arise from wrong combinations of lexical items and typically involve instances of lexical misselection of verbs or premodifying adjectives, as in (159). Lastly, 'omissions' (Flowerdew 2006: 355–6) entail cases where meaning is not complete unless a signalling item is provided, as in (160).

(156) 'the major **argument in** supporting the development of country parks' (Flowerdew 2006: 352)
(157) 'It may further worsen our economic **pollution** also after the economic downturn […]' (Flowerdew 2006: 352)
(158) 'People against this points out that this would hurt unemployment and the **prospection** of local university graduates […]' (Flowerdew 2006: 353)
(159) '[…] After this **law** has been **opened** to the public […]' (Flowerdew 2006: 356)
(160) '[…] it may also contains some adverse effects on young people's **physically and mentally**' (Flowerdew 2006: 355)

The frequency data (Flowerdew 2006: 357) show colligation errors as a remarkably numerous category (989 tokens), followed by incorrect signalling nouns (283 tokens), collocation errors (153 tokens) and omissions (26 tokens). The clear dominance of colligation in Flowerdew's taxonomy suggests that students' use of shell nouns is considerably more problematic at the syntagmatic axis than at the paradigmatic axis (Flowerdew 2006: 359). The results are also indicative of two trends relating the students' grades to the number of signalling nouns used and the number of signalling noun errors committed. Thus, it is claimed that, the lower the grade, the fewer signalling nouns are used and the more errors are made (Flowerdew 2006: 357–8).

Flowerdew (2010) broadens the scope of its predecessor by comparing an 111,558-word sample from the same Cantonese-speaking corpus and the American component of *LOCNESS* (110,537 words in total). As in Petch-

Tyson (2000), Flowerdew's (2010) aim is to analyse the use of signalling nouns in native and non-native English. The study sets out to address a range of research questions concerning the overall frequencies of signalling nouns and their rhetorical functions (i.e. across-clause vs. in-clause), the use and range of specific signalling noun types, and the frequency of different in-clause patterns (Flowerdew 2010: 39–41). In the latter respect, Flowerdew (2010: 40–1) identifies 8 in-clause patterns, only 3 of which appear in Schmid's (2000) framework (i.e. N-*to* clause, N-*that* clause, N-*be*-*that* clause; cf. Aktas & Cortes 2008 and Caldwell 2009 in 2.3.1.2.2). The remaining 5 patterns are the following:

i) Signalling Noun + *of* + Noun Phrase (SN-*of*-NP):

(161) 'a consideration **of the sanctity and value of human life**'

ii) Signalling Noun + *of* + *-ing* form (SN-*of*-V-*ing*):

(162) 'the idea **of banning of smoking**'

iii) Signalling Noun + *be* + Identifier (SN-*be*-Identifier):

(163) '**an everyday example of this** is Charles Barkley compared to Madonna'

iv) Identifier + *be* + Signalling Noun (Identifier-*be*-SN):

(164) 'A great deal of people are arguing if students using credit cards is a **good thing**'

v) Signalling Noun + *such as* + Noun Phrase/*-ing* form (SN-*such as* NP/V-*ing*):

(165) '**the drawback of using the card**, such as getting into uncontrollable purchasing habits […]'

The analysis reveals a significant occurrence of signalling nouns in the native corpus (also demonstrated by Petch-Tyson 2000), with an individual text-based mean frequency of 50.4 nouns, as opposed to 18.5 nouns in the learner corpus (Flowerdew 2010: 43). A similar variation is also apparent in the extent of use of the different rhetorical functions, with a preference for across-clause anaphoric nouns and in-clause nouns in native writing, and a greater frequency of across-clause cataphoric nouns in learner writing (Flowerdew 2010: 44). The reliance of non-native writing on the cataphoric

function might lie in either teachers' emphasis on the structure of introductions in essay writing (where cataphoric nouns dominate) or in the text format of the essays, often eliciting comparisons and opinions (e.g. compare and contrast…; discuss the advantages and disadvantages of…; Flowerdew 2010: 48–9).

Concerning the use of specific nouns, the corpora show a superficial similarity in the occurrence of 9 signalling nouns among the 20 most frequent types in both kinds of writing: *argument*, *problem*, *way*, *right*, *issue*, *result*, *example*, *time* and *effect* (Flowerdew 2010: 45). Such a similarity, however, proves misleading when looking at the frequency range of individual types across the two corpora, with native writing containing considerably more types among lower token frequency ranges than non-native writing. This appears to indicate that, unlike L1 writing, learner writing makes use of a small range of signalling noun types which are frequently repeated (Flowerdew 2010: 51).

Finally, the corpora are also observed to differ in the use of certain in-clause patterns (Flowerdew 2010: 47). Consequently, while SN-*of*-NP tops the list of native patterns, Cantonese-speaking learners of English are more likely to use Identifier-*be*-SN. Learners therefore choose the patterns that resemble those found in their L1, and they also avoid patterns featuring the amount of condensation involved in cases of nominalisation and grammatical metaphor, as in (161) above (Flowerdew 2010: 50).

Hasselgård (2012) resembles Petch-Tyson (2000) and Flowerdew (2010) in its comparison between *ICLE* non-native writing and *LOCNESS* native writing. Her interest, however, lies in the analysis of the recurring word combinations or clusters of five '[…] somewhat randomly chosen […]' shell nouns: *fact*, *idea*, *question*, *problem* and *issue* (Hasselgård 2012: 22). For the analysis of L2 writing, the corpus selection is restricted to the Norwegian, German and French sections of *ICLE* (amounting to 661,051 words). The data for L1 writing are obtained from the entire *LOCNESS* corpus (326,089 words in total).

The study shows that French writers' use of these nouns is closer to L1 standards than is the case in the other L2 corpora (cf. also Petch-Tyson 2000 above). This is reflected, for example, in the more frequent occurrence of the five units in the French corpus (Hasselgård 2012: 25), as well as in the similar distribution of the N-*of* and N-*that* patterns for *idea* and *problem* in the *LOCNESS* and French-*ICLE* texts (Hasselgård 2012: 34, 43). In certain other respects, however, all L2 writers exhibit similar tendencies. This is exemplified by *issue*, a noun that is generally underused by all L2 writers (Hasselgård 2012: 26). Underuse of certain nouns and patterns is at times specific to a particular L2 sub-corpus; for instance, *in fact* and *matter of fact* feature frequently in the French sub-corpus (while *the fact that* tops

the list of L1 combinations), and *it is a fact that* is only common in the Norwegian sub-corpus (Hasselgård 2012: 28, 33).

Hasselgård (2012: 51) accounts for the close association observed between French-*ICLE* and *LOCNESS* use of these nouns in terms of the high frequency of the N-*that* and N-*of* patterns in French (e.g. *l'idée de/que*). As far as these nouns are concerned, English therefore is more similar to Romance languages than to the Germanic-language backgrounds represented in the two other L2 corpora.

2.3.2 *Written and spoken discourse*

2.3.2.1 *General*

Use of a written and spoken corpus-based methodology in the investigation of shell nouns is of marginal significance compared to the wealth of studies on genre-specific written language (2.3.1). This section discusses references where a large written and spoken general English corpus is used. One such corpus, the *BoE*, is especially prominent in this research (Schmid 2000; Schmid 2001 and Mahlberg 2005). Only one reference, Yamasaki (2008), has been found to make use of the written and spoken components of the *BNC*.

Whilst using large general English corpora, genre concerns are only secondary to the main arguments advanced in these studies. As such, information on the genre-specificity of shell nouns is sparse and often does not extend beyond generalisations about overall written or spoken discourse.

Schmid (2000) and (2001) draw on the 225 million-word British section of the *BoE* in order to provide a purely descriptive and quantitative account of shell nouns, either generally (Schmid 2000), or specifically (Schmid 2001). Schmid's choice of the *BoE*, instead of other available corpora like the *BNC*, is motivated by the so-called 'From-Corpus-to-Cognition Principle' (Schmid 2000: 38). The rationale here is that the frequency of any linguistic pattern in a given corpus constitutes compelling evidence that such a pattern is part of the overall linguistic system. Accordingly, the real importance of a pattern can only be assessed, provided that a sufficiently large corpus is used. The automated analysis applied to such a large corpus as the *BoE* is intended to retrieve only the most frequent and '[…] linguistically preferred […]' (Schmid 2000: 40) examples of a category, in this case, of shell nouns. This means that less frequent instances of a category are systematically excluded from the analysis, on the grounds that they are not as cognitively salient as frequent ones. Schmid's principle may therefore be summarised as

'Frequency in text instantiates entrenchment in the cognitive system' (Schmid 2000: 39)[5].

In view of such stress on frequency, the *BoE* proves necessary for Schmid's (2000) quantification of shell nouns. It should be noted that drawing on this corpus helps to endow the research with large amounts of data, but fails to allow for '[…] the balance of sources […]' (Schmid 2000: 43–4). This lack of balance resides in the sheer dominance of written media language: around 70% of the words in the study corpus (i.e. 152 million words). The corpus is thus not broadly representative of English, but '[…] is clearly skewed towards the type of language typically used by the media'. (Schmid 2000: 43)

Schmid's (2000: 379) study concludes with the recognition of the need to go deeper into shell noun variation across various spoken and written text types. Several hypotheses are suggested in this respect. For example, it is claimed that the extent of use of shell nouns appears to correlate with the degree of abstraction of a text. Formal expository writing is hence more likely to feature more instances of shell nouns than colloquial spoken English, where only a limited set of highly frequent short nouns (e.g. *thing*, *fact*, *reason*) is bound to occur (Schmid 2000: 379–80). Special attention is also given to the connection between genre and the semantic, cognitive and pragmatic and rhetorical functions of shell nouns (Schmid 2000: 380; see section 2.2.2.2). As regards the semantic function of characterisation, the most typical realisation may be found in spoken language, particularly with the highly attitudinal nouns (e.g. *problem*, *tragedy*) often appearing in the *th-be*-N pattern. Emphatic focusing (e.g. *the thing is*) is similarly prevalent in spoken language. Signposting, by contrast, is closely associated with

[5] Schmid (2010: 125) questions his earlier claim about the influence of frequency on the cognitive salience of linguistic behaviour by arguing that '[...] so far we have understood neither the nature of frequency itself nor its relation to entrenchment [...]'. In his view, frequency ought to be studied in both absolute and relative terms, insofar as much cognitive entrenchment is 'cotextual' rather than 'cotext-free' (Schmid 2010: 120). Cotextual entrenchment is consistent with Hoey's (2005) theory of lexical priming (see 3.2.6), which advocates the study of linguistic phenomena in terms of their cognitive dependence and activation of other structures, functions, etc. Schmid (2010: 120) suggests that cotextual cognitive entrenchment is affected by co-text free entrenchment. This implies that the cognitive salience of a particular linguistic expression hinges on both the overall frequency of the lemma in the linguistic system and its association with specific patterns, meanings and uses. Therefore, Schmid (2010) seems to show awareness of the idea that perhaps shell-noun use is not only a matter of frequent lemmas in the N-cl and N-*be*-cl patterns, but of how both more and less frequent lemmas behave according to a range of patterns, meanings and uses.

expository writing. Finally, integration is reported to be frequent in both spoken and written abstract discourse.

Schmid (2001) elaborates on the pragmatic implications of one of the above functions, i.e. focusing and topicalising. The research focus of this paper is on the spoken and written use of the N-*be*-cl pattern as an often unconscious manipulation tool. The data obtained from the *BoE* show many instances of mental shell nouns (e.g. *guess*, *fear*, *feeling*, *hunch*) preceded by a possessive determiner, as in (166). In such cases, the shell noun acts as a subjective or tentative marker, implying the speaker's uncertainty about the propositional content in the complement clause (Schmid 2001: 1541). There are instances, however, where the definite article is substituted for a possessive determiner. The definite article differs from possessive determiners in triggering a '[…] complete backgrounding of the speaker' (Schmid 2001: 1542), leading readers or listeners to accept that the proposition is indeed what the noun implies; for example, a fear, as in (167).

(166) '[…] And I think **my feeling** is that I mean I I'm not god I don't know what the true answer is […]' (Schmid 2001: 1541)

(167) '[…] But **the greatest fear** is that the new arrangement will deplete the numbers of Britain's researchers severely and permanently' (Schmid 2001: 1543)

What is described as a fear is therefore only '[…] a bluff' (Schmid 2001: 1545), a mere presupposition whereby the speaker or writer presents his/her own opinion as given information (cf. Mahlberg 2003: 103 in 2.2.2.1.1). Readers or listeners are unconsciously deceived into taking the writer or speaker's labelling at face value. This explains why disagreement may refer to the complement clause, but not to the sentence-initial shell noun (Schmid 2001: 1548). In view of the evaluative subtlety and formality of this construction, it is more frequent in written language, whilst cases like (166) are claimed to be more typical of spoken language (Schmid 2001: 1541–2).

Mahlberg (2005), in her book on general nouns, also uses the British English component of the *BoE*, but in an updated 300 million-word version (see also Mahlberg 2003). As mentioned above (2.2.2.1.1), frequency information from the *BoE* and the *BNC* is used in the selection of the 20 study units. The analysis, however, is restricted only to the *BoE*. Mahlberg (2005: 42, 164) is aware of the unbalanced nature of the corpus, to the extent that most of her findings are explicitly claimed to apply to written journalistic language. Genre considerations are not only limited by the corpus used, but also by the scope of the study itself, as demonstrated by Mahlberg's (2005: 130) assertion that '[A] detailed discussion of the context-dependency of meaning with regard to register or genre factors is beyond

the scope of this study'. Future research is nonetheless claimed to benefit from a more genre-specific analysis of the textual functions that general nouns perform in different text types (Mahlberg 2005: 183).

Yamasaki (2008) differs from the above references in the use of a smaller though more balanced corpus, the 100 million-word *BNC* (see 4.2.2 for details). Whilst more balanced in the overall distribution of genres than the *BoE*, the *BNC*, however, is also skewed towards written discourse (90% written vs. 10% spoken). Yamasaki's (2008) paper looks at the evaluative function performed by anaphoric shell nouns in two of Schmid's (2000) patterns: *th*-N (e.g. *this problem*) and *th-be*-N (e.g. *this is a problem*). On the basis of a list of 73 retrospective labels from Francis (1994), a search of the entire *BNC* reveals that the former pattern is implicitly evaluative while the latter is explicitly so (Yamasaki 2008: 79–81). Such a finding echoes Schmid's (2000: 309) observation about the implicit characterisation typical of *th*-N, as opposed to the more overt and emphatic characterisation of *th-be*-N. A subsequent analysis of 4 labels (*change*, *shift*, *failure* and *mistake*) shows that attitudinally neutral nouns (*change* and *shift*) are often preceded by evaluative adjectives when occurring in *th-be*-N (as in (168)), while clearly attitudinal nouns (*failure*, *mistake*) usually occur with no premodification in the same pattern (as in (169); Yamasaki 2008: 81–8). The evaluative potential of the entire noun phrase is therefore essential to *th-be*-N, and less so to *th*-N, where modification is often of an objective or classifying nature (e.g. *technological*, *theoretical*, *cultural*, *epistemic*).

(168) 'It is a fundamental change in the way we do business' (Yamasaki 2008: 83)

(169) 'Visitors to the villa arrive at the front, to be greeted by a huge work by Marino Marini, and if they are not careful see only the front. This is a mistake' (Yamasaki 2008: 84)

At the end of the paper, Yamasaki (2008: 88–93) explores the lexicogrammatical behaviour of *this/that problem* and *this/that attitude* in the written and spoken sections of the *BNC*. For the written component, the analysis includes only the text types of books and periodicals, excluding written to be spoken and written miscellaneous. Overall, these noun phrases are uncommon in sentence-initial position, which entails a stronger preference for object or complement positions. This said, the written sub-corpus contains a larger proportion of sentence-initial instances than the spoken one. An examination of their collocations in both sub-corpora further indicates that spoken discourse shows a marked tendency for these nouns to occur as the objects of possessive *have/have got* (as in (170)). This appears to confirm Biber et al'.s (1999: 235–6, 1067) claim that only

pronouns (and not full noun phrases) are prevalent in subject position in spoken language (Yamasaki 2008: 88).

(170) '[…] No if you […] by a different make of […], some of them stick out further and they foul it. We've had **this problem** before' (Yamasaki 2008: 91)

2.3.2.2 *Academic discourse*

The study of shell nouns across spoken and written academic genres receives a shallow treatment in the literature reviewed. Consequently, as opposed to the long section on academic prose (2.3.1.2), the arguments in 2.3.2.2 are supported by only three references: Biber (2006), Flowerdew (2003a) and Flowerdew (2003b), where the former relies on a corpus of spoken and written university genres, and the latter two on a corpus of biology.

Biber (2006), drawing on the *TOEFL 2000 Spoken and Written Academic Language Corpus* (2.7 million words; henceforth, *T2K-SWAL*), examines the occurrence of three realisations of stance (modal verbs, stance adverbs and stance complement clauses) in two spoken and two written university genres (classroom teaching, class management talk, textbooks and syllabuses). Two of such genres are purely academic (classroom teaching and textbooks), while the two others are institutional varieties (class management talk and syllabuses). Stance is perceived to be most prevalent in spoken discourse, with a preference for modal verbs over the other markers (Biber 2006: 103). With regard to complement clauses, the pattern stance verb-*that/to*-infinite clause, as in (171) below, is reported to predominate over that of stance adjective/noun-*that/to*-infinitive clause, as in (172) and (173) (Biber 2006: 109; cf. Biber et al. 1999 discussed in 2.2.1.3.2).

(171) 'We **recognize** that it's a real error […]' (Biber 2006: 108)
(172) 'Even fairly recent specimens are **unlikely** to have any of the radioactive isotope left to be measured' (Biber 2006: 110)
(173) 'I started out with **the assumption** that consciousness is complete with the ontological proof' (Biber 2006: 109)

The relevance of this article for the study of shell nouns lies in the observation that noun complement clauses introduced by *that* (as in (173)) are only frequent in classroom teaching and textbooks. As such, only the more academic genres are bound to feature nouns such as *argument*, *idea* or *claim*, whose use is aimed at identifying '[…] the status of the information presented in the *that*-clause […]' (Biber 2006: 109). *To*-infinitive clauses

controlled by adjectives and nouns are found to prevail in the written genres. Adjectives (as in (173)) are more frequent than nouns (Biber 2006: 110).

Flowerdew's (2003a and 2003b) approach to shell nouns is more comprehensive than Biber's (2006), but the corpus in the former is much smaller and more discipline and genre-specific than that in the latter. In both papers (Flowerdew 2003a and 2003b), the author uses a corpus consisting of undergraduate lectures on biology (92,939 words) and another one comprising relevant sections of the prescribed textbook for the course (90,482 words).

Flowerdew's (2003a) exploration of signalling nouns (see 2.2.2.3) is supplemented by a number of genre-related insights (Flowerdew 2003a: 340–1). As shown by the frequency data, signalling nouns are considerably more common in the book corpus, with 20 items per 1,000 words, as opposed to only 9 items in the lecture corpus. In a list of the six most frequent nouns in the corpora (*function*, *way*, *result*, *case*, *effect*, *kind*), differences arise as to the frequency of certain items. For example, whilst *result* is frequent in the book, no occurrence is attested in the lecture corpus. A noun like *kind* is, by contrast, more prevalent in lectures. Variation is also evident in the preference for certain rhetorical functions (in-clause, across-clause and exophoric) by specific items. In this respect, the noun *function* tends to appear in the across-clause usage, while *result* is more frequent in the in-clause function.

Springing from his general characterisation of signalling nouns, Flowerdew (2003b) brings to light the potential register-related variation in signalling noun usage, based on Halliday's (1978) three contextual parameters of field, tenor and mode. The article is noteworthy in this regard, in that none of the references reported in section 2.3 takes explicit account of Halliday's register parameters. Regarding field (Flowerdew 2003b: 39–42), the biology corpora are claimed to share a common goal to provide novice students with a discipline-specific body of knowledge. With this aim in mind, it is noted that the use of signalling nouns is determined by the rhetorical acts specific to a given field of study. As such, elementary biology is frequently linked to signalling nouns conveying classification (e.g. *basis*, *category*, *division*, *group*), description of structure and function (e.g. *appearance*, *constituent*, *activation*, *conversion*) and cause-effect relations (e.g. *reason*, *result*, *difference*; Flowerdew 2003b: 40–1).

Concerning tenor (Flowerdew 2003b: 42–4), the expository nature of the study corpus entails a role relation between expert and novice. Given that elementary biology is concerned with the presentation of basic facts, attitudinal signalling nouns are rare in the corpora. Evaluative modification is, by contrast, slightly more frequent, especially with regard to such

modifiers as *essential*, *fundamental* or *major*, all highlighting the importance of the stretch of discourse encapsulated by the signalling noun (Flowerdew 2003b: 43). Finally, mode (Flowerdew 2003b: 44) is, as evidenced by Flowerdew (2003a), the most crucial parameter, inasmuch as signalling noun variation is most substantial when written discourse is contrasted with spoken discourse.

To the best of my knowledge, Flowerdew is currently testing some of his earlier findings (Flowerdew 2003a; Flowerdew 2003b; Flowerdew 2006; and Flowerdew 2010) on the basis of a 600,000-word corpus of academic discourse comprising three text types, two written and one spoken: research articles, textbook chapters and lectures. In a recent conference presentation (Flowerdew & Flowerdew 2013), a description is provided of the intraclausal patterns of the signalling nouns in the aforementioned corpus. Their focus on these patterns is explained by the limited attention that they have received in the literature, as opposed to the '[…] fair amount of attention […]' given to interclausal relations. From the conference abstract, it may be gathered that their findings point to some tendencies for certain intraclausal patterns to favour Given or New positions in the sentence.

2.3.3 *Spoken discourse*

Sections 2.3.1 and 2.3.2 cast light on the dominance of written discourse in the research to date. Comparisons between written and spoken English are, as shown in 2.3.2, either too general or too specific in their scope. When a general English corpus is utilised, little or no attention is paid to genre-specific variation, and when a specific type of discourse is examined (in this case, academic discourse), the corpus is often limited in its coverage of genres and disciplines. This section looks at research whose focus is on spoken discourse. As stated in the introduction to section 2.3, spoken language is left for the end on account of the scant attention given to this mode alone. Of the two references included here, one deals with spoken academic discourse (Lorés 2006) and the other with spoken general English (Aijmer 2007).

Lorés (2006) explores the associations between reference and formal pattern in the use of *thing*(*s*) and *idea*(*s*) in 10 Biology and Health Sciences lectures (about 101,674 words) from the 1.8 million-word *Michigan Corpus of Academic Spoken English* (henceforth, *MICASE*). The results indicate a strong preference for the cataphoric use of both nouns (as in (174)), with almost half of the instances for each unit corresponding to this reference type (43.32% for *thing*(*s*) and 50.75% for *idea*(*s*); Lorés 2006: 323–4). Anaphora and exophora (as in (175) and (176)), though less frequent, show an even distribution in both lemmas (29.37%, 27.29% for *thing*(*s*) and

28.03%, 21.21% for *idea*(*s*)). As regards reference-pattern associations, the top three combinations for *thing*(*s*) (62% of examples) are exophora (27.29%), anaphoric *th*-N (18.39%) and cataphoric N-relative clause (16.32%). *Idea*(*s*)' top three patterns (69% of examples) comprise a substantial majority of N-*that*/-*of* (33.33%), followed by exophora (21.21%) and anaphoric *th*-*be*-N (15.15%). Lorés (2006: 327) explains the prevalent use of cataphora in the study corpus on the grounds of its usefulness for the expression of hesitation in spoken discourse. Cataphoric nouns like *thing(s)* and *idea(s)* may thus give speakers time to better form their ideas.

(174) 'He did figure out two really important things. He said organisms can evolve and he also said that one organism is ancestral to another' (Lorés 2006: 321)
(175) 'It only came when we started changing the food that people started to care what was happening in the plant world and genetically modified crops, and so these things have come a long way before they got any public attention' (Lorés 2006: 321)
(176) '[…] it was Western ideas that influenced Darwin' (Lorés 2006: 323)

Aijmer (2007) uses the spoken section of the *BNC* to investigate the pragmatic implications of expressions such as *the fact is that*, *the thing is that* or *the evidence is that*. Her central argument is that only nouns like *fact*, *truth*, *thing* and *trouble* may act as pragmatic markers in a similar way to typical instances of this category (e.g. *you know*, *I mean*). Nouns like *suggestion*, *consensus*, *tragedy* or *difficulty*, by contrast, help to characterise the following *that*-clause, but lack the focusing function inherent in more general and unspecific nouns (Aijmer 2007: 35). Hence, in line with Schmid (2000: 329–37, 2.2.2.2), only the former fulfil a focalising function. Such a function leads to the increased grammaticalisation and formal reduction of the pattern, as in (177), where the definite article is dropped, or (178), where *that* is deleted instead (cf. Jespersen 1949 in 2.2.1.1).

(177) 'Thing is that the furniture trade was really booming then […]' (Aijmer 2007: 34)
(178) '[…] The trouble is there's a hundred and one things to do before we go' (Aijmer 2007: 35)

The interpretation of examples like the ones above entails an understanding of *thing is* and *the trouble is* as prefabricated units or 'fixed phrases' (Aijmer 2007: 40). These phrases (epitomised by *the fact is that*) are treated as main clauses structurally but as subordinate from a pragmatic perspective

(Aijmer 2007: 38). This implies that, pragmatically, only the *that*-clause is truly informative, while the clause-initial noun phrase would correspond to what Biber et al. (1999: 1073–6) call an 'utterance launcher'. From this utterance-initiating role, a range of interesting conversational implicatures arise (Aijmer 2007: 40–3). For example, in (179) *the fact is* has a disclaiming or adversative function similar to *however*, while in (180), a meaning of justification or elaboration is at issue:

(179) '[…] So you can consult people in […] about what's happening in […] and wider afield. **The fact is**, there's been a complete lack of consultation' (Aijmer 2007: 42)
(180) '[…] She is really a grand person and **the fact is** that she really is very often right' (Aijmer 2007: 42)

2.4 CONCLUSION

Chapter 2 surveys the theoretical grounding of this thesis, with special emphasis on the wide-ranging definitions of shell units and their genre-specific behaviour.

Section 2.2.1 shows how awareness of these nouns dates back to the first half of the 20th century. The early descriptive grammarians make passing reference to deverbal and deadjectival nouns (e.g. *discovery, certainty*) in need of semantic completion by a following *that*-clause. Such a clause is variously treated as an appositive (e.g. Sweet 1891–8) or an object (e.g. Kruisinga 1931–2). In grammatical descriptions from the late 20th and early 21st centuries, the focus is still primarily on intrasentential patterns. Especially noteworthy here is Chomsky's (1970) treatment of post-noun clauses as complements (see also e.g. Biber et al. 1999 and Huddleston & Pullum 2002), laying emphasis on the idea that both nouns and verbs may take complements, and that in most cases of nominalised shell nouns, the following clause is the rankshifted verbal complement. Special mention should also be made of the functional approach to shell-noun description in SFG (e.g. Halliday & Matthiessen 2004) and in corpus-driven grammars (e.g. Sinclair et al. 1990), where, despite the dominance of intrasentential patterns, attention is also given to the wider discursive relevance of these nouns.

Section 2.2.2 describes the wealth of terms and definitions contained in the literature on shell-like units. The classification of these terms rests on their defining criteria. General nouns (e.g. Halliday & Hasan 1976), Vocabulary 3 items (e.g. Winter 1977), enumerables (e.g. Tadros 1985), anaphoric nouns (Francis 1986), labels (Francis 1994) and carrier nouns (Ivanič 1991) all stress in different ways the unspecific, metadiscursive and discourse-organising character of shell nouns (2.2.2.1). Formal features,

especially their frequent combination with specific deictics and complement clauses, follow the initial functional criteria. Another group of references (Vendler 1968; Hunston & Francis 2000 and Schmid 2000) base their functional interpretation of shell-noun uses on a limited set of predefined structural patterns (mainly N-cl and N-*be*-cl) (2.2.2.2). Many potential shell-noun patterns are excluded due to the limitations of the automated retrieval of textual evidence from large corpora. Shell-noun research is also motivated by teaching and learning concerns (Francis 1988; Carter 1998; Flowerdew 2002 and 2003a and Hinkel 2004) (2.2.2.3). Flowerdew (2003a), with its consideration of both endophoric and exophoric reference as well as its inclusion of intrasentential patterns other than noun complement clauses, is particularly relevant to this thesis.

Section 2.3 reveals that most research focuses on written discourse and, more specifically, on two broad genres, i.e. academic and journalistic prose. General English corpora are of marginal significance, and, when used, the corpus is only a source of examples (e.g. Ivanič 1991) or is heavily skewed towards a particular genre (e.g. *BoE*, Schmid 2000). Of the references discussed in 2.3, Aktas & Cortes (2008), Caldwell (2009) and Flowerdew (2010) (along with Flowerdew 2003a) need highlighting for their examination of a range of patterns other than those typically linked to shell-noun use (e.g. Schmid's 2000 four patterns).

The following chapter is also theoretical, but its scope is narrower and more related to the the study in this thesis, in that it discusses a range of issues of paramount importance for the methodological decisions reported in chapters 4 and 5.

3 LITERATURE REVIEW: VARIABLES AND ANALYTICAL PROCEDURES

3.1 INTRODUCTION

Chapter 3 acts as a nexus between the literature review and the methodological decisions presented in chapter 4. The goal here is to provide theoretical standpoints behind the analytical approach of this thesis. The chapter comprises two main sections: key issues in shell-noun description (3.2) and corpus methodology and analytical procedures in shell-noun research (3.3).

Section 3.2 looks at six formal, semantico-pragmatic and textual issues of crucial importance for a comprehensive approach to shell-noun description. One such issue is that of abstraction (3.2.1). This subsection raises the question of whether the abstract nature of shell nouns is similar to that of other so-called abstract nouns. Section 3.2.2 explores the often-claimed association between shell nouns and the encapsulation of preceding discourse segments. Section 3.2.3 deals with the formal and semantic structure of shell-noun phrases, as reported in the literature. Attention is then turned to the treatment given to the syntactico-semantic and textual functions of these units (3.2.4). Section 3.2.5 is devoted to some of the most frequent semantic classifications of shell nouns available in the literature. Lastly, section 3.2.6 looks at how a multifaceted approach to shell-noun description might be inspired by Hoey's (2005) lexical priming, a corpus-driven approach intended to '[...] describe everything that we know about a word' (Mahlberg 2005: 27), including its meaning, form and function in discourse.

Section 3.3, shorter in length, surveys the corpus methodology employed in the literature reviewed. Two aspects will be considered: the types of corpora used and the analytical procedures applied. The chapter closes with a section of recapitulation (3.4).

3.2 KEY ISSUES IN SHELL-NOUN DESCRIPTION

3.2.1 *Abstraction*

A key factor frequently invoked in the description of shell nouns is that of abstraction. Abstraction has been linked traditionally to all instances of deverbal nominalisation, ranging from action nouns like *kiss* or *reading* to less concrete examples like *belief* or *grandeur* (cf. for example Sweet 1891, I: 61 and Kruisinga 1932, II: 24 in 2.2.1.1). Langacker (1987: 90–1) argues that all nominalisations, regardless of their more or less concrete nature (e.g. *explode>explosion*, *to hope>their hope*), are used to transform the processual component of verbs into 'abstract regions'. Many of such regions are formally realised by mass or uncountable nouns, to the extent that '[..] a considerable degree of overlap […]' (Quirk et al. 1985: 247) is claimed to exist between abstraction and the lack of the plural inflection. Exceptions, however, occur where the abstract noun is used to refer to a particular instance of the phenomenon or quality at issue (Quirk et al. 1985: 286 and Sinclair et al. 1990: 10–11), as in examples (181a) and (181b) below.

(181) (a)'Many parents were alarmed to find themselves in **open conflict** with the church' (Sinclair et al. 1990: 10)
(b)'Russia had been successful in **previous conflicts**' (Sinclair et al. 1990: 11)

Ivanič (1991: 98–101), as stated in 2.2.2.1.4, draws on these countable uses as the main criterion for the identification of carrier nouns. Hence, only countable abstract nouns like *purpose*, *issue* or *aspect* are prototypical instances of the category. Some of these examples have both countable and uncountable uses (e.g. *these difficulties*, *without difficulty*; Ivanič 1991: 99). Some items, by contrast, are always countable, especially those carrier nouns whose meanings stem from metaphorical extensions of concrete nouns, as in *component*, *element*, *area* and *feature* (Ivanič 1991: 100). An example of such a contrast is (182) below, where *feature* is first used as a concrete noun ((182a)) and then as a metaphorical carrier noun ((182b)):

(182) (a)'The nose was **the predominant feature of his face**' (Ivanič 1991: 100)
(b)'Perhaps **the awkward feature** is that the plane can only be used to a limited extent […]' (Ivanič 1991: 100)

Based on the above, abstract nouns appear as a rather heterogeneous class of mostly uncontable non-concrete untouchable entities. Fraurud (1992: 7) calls into question their lack of concreteness by suggesting the

possibility of an abstract-concrete cline within abstract nouns. As such, nouns like *kiss* or *arrival* are arguably more concrete and visible than more abstract items like *plan* or *want*. Asher (1993) builds on this idea in a major study on abstract objects in discourse. In his framework, only elements denoting propositons and thoughts (e.g. *fact*, *issue*, *idea*, *assumption*) are 'purely abstract objects' (Asher 1993: 2). Events and states (e.g. *action*, *struggle*, *fight*), by contrast, are considered as 'world immanent objects' (Asher 1993: 2) on account of their connection with time and space. Distinctions between eventive and factual objects are often influenced by co-occurrence restrictions. For example, (183) below is unacceptable on the basis of the occurrence of a factual propositional noun (*fact*) in combination with an overtly eventive verb (*occur*). Further mismatches are those between verbal processes and thoughts or propositions, as in (184) (Asher 1993: 31–2). This being the case, however, the choice of a particular abstract object (i.e. shell noun) to label a clause responds to our individual need to categorise linguistic material according to our pragmatic needs (Asher 1993: 60).

(183) '*The fact that Mary fell down occurred in the park at noon' (Asher 1993: 31)
(184) '*John announced the thought/claim that he was married' (Asher 1993: 31)

Returning to Fraurud's (1992: 7) claim above, Asher (1993: 212) agrees that discourse entities occur in a cline that ranges from '[…] "semi-concrete" eventualities to abstract entities'. Consten et al. (2007), in a study on so-called 'complex anaphors' (their own term for shell nouns), propose a similar cline comprising various degrees of abstraction. Propositions and facts represent the highest degrees, followed by states, processes and events as the least abstract categories (Consten et al. 2007: 86). Interestingly, anaphoric encapsulation in discourse is argued to proceed only from least to most abstract, rather than the other way round (Consten et al. 2007: 93). This is evident in (185), where the oddness of *event* at the end of the sequence would be explained by the natural ontological order of entities in the preceding discourse, i.e. *process/state- fact-possibility*.

(185) 'The earth turns about the sun. This process/This state will presumably last for 7-10 years. This fact is well known since the Middle Ages. Researchers of the Vatican were not allowed to examine this possibility. *This event […]' (Consten et al. 2007: 93)

Martin (1997: 30–1) categorises abstraction in a way that differs markedly from the clines above, as no mention is made of such terms as 'event', 'fact'

or 'proposition'. This SF framework classifies nouns under three levels, i.e. 'concrete' (e.g. *apple*, *anvil*), 'abstract' (e.g. *inflation*, *fact*) and 'metaphoric' (e.g. *collision*, *strength*). Abstract nouns are subdivided into four further levels, i.e. technical, institutional, semiotic and generic, with only the latter two containing instances of shell nouns. 'Technical' and 'institutional' nouns contain highly specialised vocabulary (e.g. *gene*, *metafunction*) and lexis related to bureaucracy (e.g. *regulation*, *bureau*). 'Semiotic' nouns include such general items as *fact*, *idea*, *word*, *concept* or *notion*, those of which express mental and verbal (linguistic) meanings. Lastly, 'generic' nouns represent unspecific abstract meanings other than those encoded by the previous category (e.g. *manner*, *way*, *cause*).

As regards 'metaphoric nouns', these fall into nominalised processes or verbs (e.g. *collision<collide*, *perception<perceive*) and nominalised qualities or adjectives (e.g. *strength<strong*, *bravery<brave*). Martin's (1997) approach, therefore, fails to equate abstraction with deverbal or deadjectival nominalisation, as is often the case in the literature, granting it a separate category instead.

Of paramount importance to the notion of abstraction followed in this thesis is the threefold distinction that Lyons (1977, II: 442–7) proposes between 'first-order', 'second-order' and 'third-order entities'. The class of 'first-order entities' comprises tangible, real-world elements, such as people, animals and objects (e.g. *waiter*, *lion*, *pencil*), those of which have a location in time and space. 'Second-order entities' are nouns which do not exist (unlike first-order entities), but occur or take place (e.g. *crime*, *move*, *fight*). Finally, 'third-order entities' convey purely abstract meanings, as shown by ideas, propositions and facts (e.g. *theory*, *claim*, *aspect*). These items share with second-order ones their nominalised origin (e.g. *to work>work, to assess>assessment*), but differ in their being unobservable and outside the spatiotemporal dimension.

Schmid's (2000) exploration of shell nouns rests on the aforementioned distinction as its chief identification criterion. Only nouns denoting second- (e.g. *act*, *change*) and third-order entities (e.g. *problem*, *reason*) are treated as shell nouns (Schmid 2000: 63–73). It is important, however, to distinguish between second- and third-order shell nouns from other non-shell abstract entities. Schmid (1999: 223) points out that, although such nouns as *love*, *democracy* or *inflation* are abstract in their denotation of intangible concepts, their meaning remains roughly constant in different contexts of use. Shell nouns like *problem*, *idea* or *event*, by contrast, '[…] depend almost completely on the linguistic and situational context' (Schmid 1999: 223), insofar as the understanding of a particular problem, idea or event is specific to a particular utterance situation. From Schmid's (1999) argument, it appears that second- and third-order entities fall in a cline that

comprises highly unspecific entities in need of lexical realisation (i.e. shell nouns) and abstract entities which, whilst not denoting tangible units, are almost as semantically bounded as typical concrete entities (cf. Martin 1997 above). Such semi-concrete entities are best exemplified by technical or discipline-specific terms like *subordination*, *extraposition* or *morphological derivation* (Schmid 1999: 222).

In Schmid (2000), second- and third-order shell nouns are termed 'events' and 'abstract relations', and are further split into three subcategories each (Schmid 2000: 65–7)[6]. 'Events' comprise activities, processes and states, while 'abstract relations' consist of facts, ideas and utterances. Events are characterised in terms of the semantic features Dynamic and Agentive, with 'activities' showing both, 'processes' featuring only Dynamic and 'states' lacking both. Concerning abstract relations, states of affairs are involved in 'facts' (e.g. *fact*, *aspect*, *point*), objects of thought in 'ideas' (e.g. *thought*, *assumption*) and linguistic expressions of ideas in 'utterances' (e.g. *claim*, *argument*). As shown in 3.2.5 and in 5.3.5, the analysis of particular shell units often defies a straightforward ontological classification, inasmuch as it is not lemmas, but their context-dependent uses, that carry specific meanings.

With this classification in mind, Schmid's (2000: 67–8) data reflect a tendency for eventive shell contents to occur with abstract shell nouns, but not the other way round. Abstract shell contents are therefore expected to be preceded by abstract shell nouns, but not by eventive nouns as, according to Schmid, only abstract nouns can modify the conceptual status of clauses. Example (186) below shows an eventive shell content introduced by an abstract shell noun (i.e. *aim*). Example (187), by contrast, illustrates the combination of an abstract shell noun (i.e. *irony*) and an abstract shell content. The strangeness of a paraphrase containing an eventive noun like *process* is evident here.

(186) 'Their aim is to meet President Saddam in Baghdad' (Schmid 2000: 68)
(187) 'The irony of the anti-hunting councillors' action is that this motion will not have any significant effect on hunting in Hampshire' (Schmid 2000: 68)

[6] In Paradis (2004: 58), shell nouns are only third-order entities. Following Schmid (2000), however, in this thesis, second- and third-order entities are treated as shell nouns, inasmuch as eventive nouns like *action*, *venture* or *crime* may be as unspecific and context-dependent as more abstract third-order nouns like *problem*, *aspect* or *thing*. Thus, the belief here is that both second- and third-order entities may carry the context-specific gaps typical of shell-noun uses.

The conceptual shift apparent in examples like (186) underlies the function of temporary concept-formation discussed in 2.2.2.2. The reification of discourse, however, is not always linked to the formation of concepts. Abstraction is at times exploited for mere stylistic purposes, turning congruent representations of reality into abstract entities. This is where the Hallidayan concept of 'grammatical metaphor' comes into play (see 2.2.1.3.1). Schmid (2000: 70–3) argues that, in examples like (188) below, the possibility to paraphrase the shell-content complex (i.e. *the firm belief that*…) into a verbal construction (i.e. *he firmly believes that*…) brings to light a stylistic abstractness that differs from that found in cases like (189), where such a paraphrase is not possible for want of a related verb. The difficulty in finding a satisfactory paraphrase for (189) indicates a conceptual shift (from event to abstract relation), but no grammatical metaphor.

(188) '[…] his Buddhist mind-training exercises led him to hold **the firm belief** that the mind has unrecognized powers' (Schmid 2000: 71)

(189) 'A very minor flaw lay in **the fact** that Tom Wood, alone among his travelling companions, had met Markham face to face' (Schmid 2000: 71)

To conclude this review, it should be emphasised that, whilst Lyons' (1977) threefold ontological classification is useful for distinguishing between shell nouns and other non-shell concrete and abstract entities, the contextual examination of units reveals certain degrees of indeterminacy in the identification of shell-noun uses. The contention here is that the notion of second- and third-order entities should not be applied to the letter in shell-noun research. Real discourse contains many instances where, despite the existence of an observable concrete or semi-concrete product, the noun phrase itself still requires lexicalisation by another discourse segment. This may be observed in example (190) below:

(190) 'At the end of the eighties a rogue characterized the decadent gluttony: 'Buy nicer and nicer, fly further and further, shag faster and faster'. During several months **this graffito** ornamented the walls of an old comfort station at Hamburg's Yuppie district […]' (Consten et al. 2007: 82)

Consten et al. (2007: 83) claim that *this graffito* cannot be treated as a complex anaphor because, strictly speaking, it is neither an abstract nor an eventive entity, but an observable first-order entity associated with the presence of paint on a wall. Discarding this example seems contradictory in the light of Consten et al'.s (2007: 83) second identification criterion, namely that only noun phrases referring back to at least a clause are to be treated

as complex anaphors. In this particular instance, *this graffito*, whilst not a prototypical second- or third-order entity, clearly refers back or encapsulates the preceding quote. This implies that, although, strictly speaking, *graffito* here is an observable entity occurring in time and space, such an entity represents the product of a prior act of writing, and this product behaves in this particular context as an unspecific lexical item in need of lexicalisation.

3.2.2 *Encapsulation*

An often-repeated claim about shell nouns is their use to refer to long discourse segments. The literature distinguishes between this kind of reference and more typical instances of pronoun-based anaphora, where a pronoun (e.g. *he*) relates back to a first-order entity (e.g. *the plumber*, *John*). Several terms are put forward to account for the longer antecedents underlying the use of shell nouns, i.e. 'impure textual deixis' (Lyons 1977), 'discourse deixis' (Levinson 1983 and Stirling & Huddleston 2002), 'situation reference' (Fraurud 1992), 'encapsulation' (Sinclair 1993; Conte 1996 and Sinclair 2004), 'abstract entity anaphora' (Asher 1993), 'indirect anaphora' (Botley 2006) and 'complex anaphora' (Consten et al. 2007). Many of these references are concerned primarily with pronominal anaphoric or retrospective reference, as in (191). Such a focus on anaphora offers an explanation for the prominence of anaphoric shell nouns in the literature reviewed in chapter 2.

(191) 'Sam suspects that either Fred is at a party or he is at a bar. Jenny thinks that too' (Asher 1993: 49)

Lyons (1977, II: 658–64, 673) notes that pronouns may perform either a deictic or a referential function. While the interpretation of 'deixis' draws on the extralinguistic context of communication, 'reference' depends on the occurrence of an antecedent in the surrounding co-text. As such, in (192), the pronoun *he* may be referential if taken to imply *John*, or deictic, if the form is stressed and no linguistic referential expression occurs in the co-text.

(192) 'John looked up when he came in' (Lyons 1977, II: 661)

Straddling deixis and referential anaphora is the category of 'textual deixis' (Lyons 1977, II: 667–8). 'Textual deictics' are referring expressions (usually demonstrative pronouns) interpreted on the basis of formal linguistic entities in the surrounding co-text. In (193), for example, *it* does not entail the real-world referent of *rhinoceros* (i.e. the animal), but the actual lexeme. Therefore, textual deictics do not point to the discursive referents of real-

world entities (as anaphoric reference would do), but to language segments (e.g. words, expressions, clauses, sentences). Falling within the scope of textual deixis is a less straightforward subtype. In sentences like (194), the noun *lie* does not objectively label a previous stretch of discourse: it offers an interpretation of its pragmatic implication. In these instances, textual deixis is described as 'impure', as the referential expression does not relate to a surface linguistic form, but to the pragmatic and propositional meaning underlying such a form.

(193) 'That's a rhinoceros (and Y responds)
A what?
Spell it for me' (Lyons 1977, II: 667)
(194) 'I've never even seen him (and Y responds)
That's a lie' (Lyons 1977, II: 668)

Levinson (1983: 85–8) follows the deictic analysis of these expressions and introduces the term 'discourse deixis'. 'Discourse deixis' applies to cases of metalinguistic reference to preceding or subsequent stretches of discourse, as in (195) and (196) respectively. Reference is metalinguistic on account of the property of 'token reflexivity' (Levinson 1983: 86), whereby discourse segments are characterised by speakers or writers according to different types of language (e.g. *story*, *lie*, *myth*). This characterisation may be more or less objective, in which case, as in Lyons (1977), a distinction is made between 'pure' and 'impure' discourse deixis. Levinson (1983: 87) illustrates the latter subtype with the same example as in Lyons (1977; see (194)). It is worth mentioning that, unlike Lyons (1977), Levinson (1983: 87) also equates discourse deixis with clause-initial conjunctive adverbials (e.g. *therefore*, *however*, *besides*), insofar as they spell out the relationship between the current sentence and the preceding discourse.

(195) 'That was the funniest story I've ever heard' (Levinson 1983: 80)
(196) 'I bet you haven't heard this story' (Levinson 1983: 80)

Stirling & Huddleston (2002: 1460–1) employ the term 'discourse deixis' in a similar way to Levinson (1983). This said, they separate instances where a pronoun is preceded by a stretch of discourse from instances where both a pronoun and a metalinguistic noun refer back to a discourse segment. The former represents 'anaphora', as in (197), and the latter is treated as 'discourse deixis', as in (198). In their framework, 'anaphora' and 'antecedent' appear as general terms subsuming retrospective and anticipatory (i.e. cataphoric) reference (Stirling & Huddleston 2002: 1455). Following Stirling & Huddleston (2002), this thesis applies the term 'antecedent' to anaphoric and cataphoric encapsulation.

(197) 'A: Kim has been falsifying the accounts. B: **That**'s terrible' (Stirling & Huddleston 2002: 1461)
(198) 'A: Kim has been falsifying the accounts. B: **That**'s a **lie**' (Stirling & Huddleston 2002: 1461)

Fraurud (1992: 4) coins the term 'situation reference' for any formal realisation of anaphoric reference to 'eventualities' (i.e. events, processes and states) and 'factualities' (i.e. facts and propositions). Fraurud (1992: 39) is aware of the considerable emphasis that related research has placed on singular neuter pronouns (e.g. *it*, *this*, *that*), while disregarding instances of situation reference realised by full definite or demonstrative noun phrases. In support of the latter, she provides examples where pronominal and nominal anaphors replace both sentential antecedents, as in (199), and shorter noun-based ones, as in (200). Unlike most related research, account is thus taken of pronominal and nominal referring expressions and of nominal and non-nominal antecedents (Fraurud 1992: 45). At the end of the paper, however, it is suggested that a more thorough exploration of nominal anaphors (i.e. shell nouns) and nominal antecedents would prove beneficial to an overall account of the use of situation reference in discourse (Fraurud 1992: 60–1).

(199) 'Today a woman was assaulted and her baby kidnapped. **The incident/It** took place in the disreputable X district' (Fraurud 1992: 45)
(200) 'Today a kidnapping and an assault were reported. **The incidents/They** took place in the disreputable X district' (Fraurud 1992: 45)

Asher's (1993) 'abstract entity anaphora' resembles Fraurud (1992) in the connection established between referential expressions and different ontological types of antecedent (e.g. events, propositions, etc.; see 3.2.1). Similarly, no distinction is made between pure or impure anaphora, as in Lyons (1977) or Levinson (1983). Instead, a cline of abstraction is said to apply to the use of these anaphors in discourse, with facts and propositions being clearly abstract and events being closer to the concrete end of the continuum (Asher 1993: 212).

Abstract entity anaphora, whilst being most often introduced by the pronouns *this*, *that* and *it*, may also occur in the shape of a noun phrase (Asher 1993: 225). As regards types of antecedent, Asher (1993: 260) observes that on many occasions the interpretation of an abstract anaphor draws on information that is not contained in single or adjoining clauses, but in a range of non-adjacent clauses. This underlines the influence that abstract entity anaphora has on the overall structure of discourse (Asher 1993: 261). While the occurrence of a long antecedent is an important

prerequisite for abstract anaphora, passing mention is made of the existence of nominal antecedents, as in (201). In these instances, the analysis of abstract anaphora is said not to differ from that of first-order concrete expressions, as in (202) (Asher 1993: 229). Compared to this assumption, Fraurud (1992: 60) comments that further investigation is needed to detect the possible formal and semantic differences between both kinds of anaphora.

(201) 'The claim that Susan got a C on the test was surprising. John did not believe **it**' (Asher 1993: 229)
(202) 'The man in the black raincoat looked suspicious. Frank did not trust **him**' (Asher 1993: 229)

Botley's (2006) use of the term 'indirect anaphora' involves Francis' (1986 and 1994) nominal 'labelling' (see 2.2.2.1.4) and the two pronoun-based concepts of 'situation reference' (Fraurud 1992) and 'textual' or 'discourse deixis' (Lyons 1977 and Levinson 1983). The adjective 'indirect' accounts for the nature of the antecedents, more complex and longer than those of more straightforward instances of relations between pronouns and concrete first-order entities (i.e. 'direct anaphora', as in (202)). This study highlights the many difficulties posed by the identification of antecedent boundaries, with almost 30% of the data proving problematic or unclear (Botley 2006: 102). Examples (203) and (204) illustrate such problems, insofar as in both cases the noun phrase (i.e. *this suspense* and *this time*) does not encapsulate a clearly delimited stretch of discourse. The noun phrase cannot be taken as exophoric, since the noun has a context-specific meaning, one which stems from the overall situation implied in the preceding co-text (i.e. *the suspense* caused by not knowing who committed the crime and *by the time* when the shooting and the escape down the river bank took place).

(203) '[…] Ay, truly, I believe you; I believe poor Harry is killed; and I believe his murderer (for what purpose, God alone can tell) is still lurking in his victim's room. Well, let our name be vengeance. Call Bradshaw". The footman came at the summons, very white and nervous. "Pull yourself together, Bradshaw," said the lawyer. "**This suspense**, I know, is telling upon all of you […]"' (Botley 2006: 91)
(204) '"I watched the shooting from the roof of my house", said Rauf, one of four Afghans who led this reporter into rebel territory. "I also saw a number of men slip away from the jirga and escape down the river bank". By **this time**, 400-500 government troops were going through the town' (Botley 2006: 97)

Kolhatkar et al. (2013: 114), a computational linguistic study on the annotation of *this*-N anaphoric shell nouns, similarly points to the challenges involved in the accurate identification of antecedents, arguing that '[...] antecedent annotation is a complex task, as it involves deeply understanding the discourse and interpreting it'. Following a preliminary analysis of about 15 examples of 10 highly frequent shell nouns in a corpus of *New York Times* articles, Kolhatkar et al. (2013: 113) conclude that antecedent annotation proves particularly difficult with circumstantial and eventive shell nouns (e.g. *time*, *way*, *action*, *attempt*), as the antecedents are often long and unclear (no examples are given of these unclear cases; see, however, (204) above for an unclear circumstantial instance in Botley 2006). At a later stage, their annotation of about 500 concordances for 6 factual, mental and modal nouns (i.e. *fact*, *reason*, *issue*, *decision* and *possibility*) reveals a close correspondence between their expert annotation and the suggested antecedents by two non-expert judges (72% of agreement; Kolhatkar et al. 2013: 119). This shows that despite the troublesome nature of antecedent annotation, readers and listeners in real discourse situations tend to make an effort to understand shell nouns, irrespective of the specificity of antecedent boundaries (hence the agreement between expert and non-expert annotators).

Consten et al'.s (2007) concept of 'complex anaphora' is also considered as a subtype of indirect anaphora. The term applies to Francis' (1986 and 1994) labels and Schmid's (2000) shell nouns. Anaphora in these cases is argued to be 'complex' on account of the evaluative potential involved in the use of nominal referring expressions to encapsulate discourse segments (Consten et al. 2007: 82). Evaluation is apparent in the ability of these nouns to change the ontological status of complex antecedents, as in (205), where the underlined stretch undergoes a semantic shift from event into fact once the noun *fact* itself is used as a referring expression.

(205) 'The Americans tried to invade the building but were forced back by shots from the top floor. Rumsfeld had to explain the consequences resulting from **this fact** during a press conference in the afternoon' (Consten et al. 2007: 91)

Sinclair's (1993 and 2004) 'encapsulation' is the last term to be explained in this section. The discussion of this approach has been postponed based on the use of the term in this thesis to describe the relation between shell noun and antecedent. The choice of the term rests on the need to avoid the terminological controversy implied in such concepts as reference and deixis.

Sinclair (1993: 8–9) distinguishes between two types of cohesion, a 'point-to-point' one and 'encapsulation'. The former concerns cases of pronominal

reference to other first-order entities in the text, while the latter entails reference to linguistically complex antecedents (i.e. clauses and sentences). 'Encapsulation' is contrasted with 'prospection', the former being backward-looking (anaphoric) and the latter forward-looking (cataphoric). Prospection is inspired by Tadros' (1985 and 1994, see 2.2.2.1.3) notion of 'advance labelling', both sharing the assumption that certain expressions commit the writer to the provision of specific information in a subsequent discourse segment (Sinclair 1993: 12–13). This is evident in (206), where *his message* presupposes the following sentence.

(206) '[…] Last week he addressed British industrialists, and his message was typically forthright: "In two years' time, the United Kingdom will find itself part of a single market and, in effect, a single population of 320 million people […]' (Sinclair 1993: 13)

In Sinclair (2004), the prospective function of sentences is brought to the forefront of linguistic analysis. His argument is that every sentence in discourse encapsulates all previous sentences, to the extent that '[T]he whole text is present in each sentence' (Sinclair 2004: 14). With this in mind, it is claimed that little or no use lies in the identification of anaphoric antecedents, since readers or listeners do not constantly refer back to the previous co-text in their interpretation of current discourse (Sinclair 2004: 14). All previous sentences become part of their knowledge or experience of the text and, as such, it is only what follows (i.e. prospection) that is of real interest (Sinclair 2004: 69). Each sentence, therefore, '[…] encapsulates what has gone before in a single act of reference' (Sinclair 2004: 15) and, by so doing, points forward to the following sentence. Brown & Yule (1983: 46), though not using the terms 'encapsulation' or 'prospection', similarly posit that any sentence in discourse is always dependent on the whole of the previous text for an understanding of its semantico-pragmatic meaning. In their view, '[…] every sentence forms part of a developing, cumulative instruction […]' enabling readers or listeners to make sense of subsequent sentences (Brown & Yule 1983: 134).

Conte (1996) employs 'encapsulation' with a more explicit link to shell-noun use. Her paper resembles Francis (1986, see 2.2.2.1.4) in its emphasis on anaphoric encapsulation, but differs in its reluctance to use the term 'anaphoric noun'. In this respect, it is argued that some general nouns have a potential for anaphoric reference, but they are not anaphoric per se (Conte 1996: 3). Anaphoric encapsulation results in the phenomenon of 'hypostasis', where, subsequent to nominalisation, a stretch of discourse is integrated as a single entity in the flow of discourse (Conte 1996: 4). Closely associated with Lyons' (1977) or Levinson's (1983) impure textual or

discourse deixis is Conte's (1996: 6) 'pragmatic hypostasis', which explains how, more often than not, nouns do not objectively encapsulate the linguistic status of the antecedent, but provide a particular subjective interpretation of the illocutionary force involved, as in (207).

(207) '[…] With only a 10-seat majority, Kohl must now keep everybody happy, so CDU general secretary Peter Hintze immediately announced that the party "did not accept the employers' proposals". If the CDU keeps **that promise**, the central test […]' (Conte 1996: 6)

This brief review shows how the encapsulation that underlies shell-noun use appears under various terms and definitions. While certain differences exist, all of these terms draw attention to the complex relation that holds between Lyons' (1977) second- and third-order entities and their discursive antecedents. Although most such antecedents involve discourse segments, some of the above references make passing reference to the possibility of lexical realisation by noun phrases (e.g. Fraurud 1992 and Asher 1993). Caldwell (2009: 46) questions the rigidity inherent in the overall criterion of encapsulation by claiming that even so-called first-order entities may at times be used to refer to longer stretches of discourse. Hence, it is argued that Schmid's (2000: 15) function of temporary concept-formation (see 2.2.2.2) is applicable not only to shell nouns but, by extension, to any other noun whose pragmatic and semantic import is similarly temporary. This is shown in (208), where the first-order entity *this cat* has a temporary or context-specific meaning that stems from its encapsulation of the underlined segment. The contextual interpretation of this noun phrase is thus coloured by what has already been said about the animal.

(208) 'Two small piercing eyes. The attitude of a temperamental landmine. Ten years old and with timing that put my alarm clock to shame. **This cat** is what stood between me and my first cup of coffee every morning […]' (Caldwell 2009: 46)

To conclude this section, attention will be given to the treatment that encapsulation receives in the literature on shell nouns reviewed in chapter 2.

As regards its direction, anaphoric encapsulation is at the core of most research (e.g. Halliday & Hasan 1976; Francis 1986; Charles 2003; Moreno 2004; Yamasaki 2008; Gray 2010; Kolhatkar et al. 2013). This is unsurprising in light of the emphasis that the above-mentioned references place on (mainly pronominal) anaphora (e.g. Fraurud 1992; Asher 1993). The prominence of anaphoric encapsulation, however, could prove misleading on the grounds that most such references focus principally on demonstrative noun phrases (particularly *this* and *these* + noun). For

example, Botley's (2006: 85) finding that '[…] cases of cataphora are much rarer than cases of anaphora […]' should not be generalised to all instances of indirect anaphora, insofar as his study is only concerned with demonstratives *this*, *that*, *these* and *those.* The dominance of anaphoric uses among demonstrative determiners is, based on Biber et al. (1999: 273), not specific to shell nouns, but a general tendency displayed by the English noun phrase. This suggests, therefore, that the regular attention devoted to anaphoric encapsulation in the literature is only the result of the kinds of determiners examined.

Comparatively few references focus only on the cataphoric function of shell-like units, as shown for example in Winter (1977), Tadros (1985 and 1994), Hunston & Francis (2000) and Charles (2007). The interplay between anaphora and cataphora is explored in Francis (1994), Schmid (2000) and Mahlberg (2005). Finally, both endophoric (anaphoric and cataphoric) and exophoric uses are mentioned in Ivanič (1991), Partington (1998) and Flowerdew (2003a). One caveat here is that, even when an all-encompassing approach to encapsulation is adopted, no quantitative data are offered about the frequency of different types.

In relation to the link between antecedent and noun, the literature consulted appears to be primarily concerned with the intersentential functions of shell-like items (e.g. Halliday & Hasan 1976; Tadros 1985; Francis 1986; Charles 2003; Moreno 2004). That this is indeed the case is not unexpected based on the prevalence of anaphora in most related research. Encapsulation within the boundaries of the sentence (e.g. Hunston & Francis 2000; Biber 2006; Charles 2007) or both within and outside these boundaries (e.g. Ivanič 1991; Winter 1992; Schmid 2000; Flowerdew 2003a; Caldwell 2009) features less frequently in the literature. The treatment of intrasentential encapsulation, however, is, as noted in 3.2.3, most often limited to *that* noun complement clauses.

Finally, concerning antecedent type, encapsulation of long discourse segments appears to be the norm for shell-noun use. Only three of the references reported in chapter 2 contemplate the possibility of noun phrase antecedents for these units (Ivanič 1991: 109; Flowerdew 2003a: 336 and Gray 2010: 179). In the examples provided, the nominal antecedent has the form of a deverbal nominalisation (see (209), (210) below and also (200) and (201) above). Example (211) shows a non-derived unit, which, whilst not being a deverbal nominalisation, is still a third-order entity (i.e. two academic disciplines). Hence, from the few examples in the literature, a plausible hypothesis might be that second- or third-order nominalised and non-nominalised entities may act as shell-noun antecedents in the absence of longer discourse segments.

(209) 'It is interesting to read about the items electors mentioned as having, in their view, specially affected the election. Of the total, 32% thought rash Labour promises – cost of new pension scheme – bribery of electorate had a bad effect […]' (Ivanič 1991: 109)
(210) 'Another important structural characteristic of monosaccharides is the occurrence of isomerism' (Flowerdew 2003a: 336)
(211) '[…] developed a domain-specific measure to assess students' beliefs about knowledge in history and mathematics. These specific domains were selected to address beliefs [...]' (Gray 2010: 174)

3.2.3 *The shell-noun phrase: formal and semantic structure*

Section 3.2.2 presents anaphora as the most prominent encapsulating direction cited in the literature. Such an emphasis seems to tie in with the usual argument that only 'specific Deictics' (Halliday & Matthiessen 2004: 312–17) correlate closely with shell-noun use. 'Specific Deictics' comprise the definite article, demonstrative determiners (*this*, *that*, *these*, *those*) and possessive determiners (e.g. *my*, *your*, *his*).

The occurrence of these determiners in shell-noun phrases could in turn be linked to the length and complexity of antecedents. Ariel (1990) lends support to this claim. This study explores the connection between the form of the noun phrase and the degree of cognitive accessibility of its antecedent expression. The result is a 15-point scale of accessibility ranging from least accessible (proper noun, long/short definite description) through partially accessible (distal demonstrative, proximal demonstrative) to most accessible (pronouns) (Ariel 1990: 73). Only pronouns, on account of their attenuated form, entail an antecedent that is immediately relevant to the discourse situation. Full noun phrases (especiallly definite), by contrast, require an antecedent that is not active in the reader's or listener's mind, i.e. one which is not near the referring expression and, at times, one which crosses sentence boundaries (Ariel 1990: 57). Shell-noun phrases would thus resemble pronouns in their context-specific meaning, but not in their descriptive potential, which is better suited to distant and longer antecedents according to Ariel (1990). Brown & Yule (1983: 174) offer a similar suggestion, with definite noun phrases being used to refer to 'displaced entities' and pronouns identifying 'current entities'.

Compared to the amount of attention given to definite shell-noun phrases (including *the* and the demonstratives), the literature reveals a cursory treatment of indefinite units (Quirk et al. 1985: 1261; Ivanič 1991: 111; Francis 1993: 154; Partington 1998: 92–3 and Aktas and Cortes 2008: 10). This may once again be due to the dominant position held by anaphora in most related research. Anaphora is typically associated with definiteness, while first-mention is the most frequent use of indefiniteness. In Prince

(1981: 235), for example, indefinite noun phrases like *a guy I work with* or *a bus* are treated as 'brand-new' discourse entities, insofar as they cannot be assumed to exist in the hearer's or reader's mental discourse model. Hawkins (1978: 172–3), Brown & Yule (1983: 188) and Fraurud (1990: 404) show that such a tendency is not always apparent in discourse. These references contemplate the possibility of indefinite noun phrases with an anaphoric function, as well as of definite phrases with a first-mention use. The former is evident, for example, in (212), where *a student* represents a member of the larger group of students introduced in the preceding sentence.

(212) 'Some students were standing outside the factory gate. Bill kept his eye on them. After a while a student came up to him and asked him his name' (Hawkins 1978: 174)

As regards first-mention definiteness, one of the cases discussed is that of noun complement clauses. In these instances, the use of the definite article is motivated by the occurrence of a post-nuclear clause (Hawkins 1978: 140–4). The absence of the complement clause would indicate an anaphoric interpretation of the noun, as shown in (213b) (Hawkins 1978: 140). Hawkins (1978: 143) explains the first-mention definiteness of these examples in terms of a transformation from the underlying structure *that* X *is a fact*, where the shell noun is a first-mention indefinite noun phrase in complement position (e.g. *that the Prime minister is going to resign is a fact*). The use of corpus techniques by Francis (1993: 154) '[…] does not confirm that the definite article is the norm' in these constructions. Hawkins' (1978: 140–4) theory is thus refuted in the light of many nouns for which the indefinite article appears to be the only option in this pattern, as in (214) (cf. also Quirk et al. 1985: 1261 and Partington 1998: 92–3 in 2.2.1.2 and 2.2.2.1.1 for similar examples).

(213) (a)'London has been buzzing with the rumour that the Prime Minister is going to resign' (Hawkins 1978: 102)
(b)'London has been buzzing with the rumour' (Hawkins 1978: 102)

(214) 'But today they'll have made a pact that each must be courageous for the other's sake' (Francis 1993: 154)

Indefinite shell-noun phrases may occur not only in noun complement constructions but also, as argued by Ivanič (1991: 110–1), in other intra- or intersentential instances of (anaphoric or cataphoric) encapsulation. In her framework, prototypical carrier nouns are leftmost in a context-dependency cline ranging from most context-dependent uses, as in countable definite

nouns with endophoric reference, to most self-contained uses, represented by uncountable nouns. The two intermediate categories are countable definite nouns with exophoric reference and countable indefinite nouns.

Of particular interest is Ivanič's (1991: 111) view of indefinite noun phrases and uncountable nouns as potential instances of carrier nouns. In this respect, it is suggested that '[…] the context seems to play a role in their interpretation even when they are accompanied by indefinite reference, or when they occur as uncountables' (Ivanič 1991: 112). Thus, in examples like (215) and (216), the uncountable noun (i.e. *difficulty*) and the indefinite noun phrase (i.e. *an explanation*) appear to require the reader to look for some sort of context-dependent meaning in the surrounding co-text. In the case of (215), one could argue that the lack of such a quality in *without difficulty* is linked to the action of *going further and giving relief on an individual's income*. In (216), although no immediate explanation is given, the reader expects the subsequent discourse to offer clues as to the interpretation of this carrier noun. Fiction, the genre of this example, makes frequent use of these indefinite entities. For the sake of suspense, the interpretation of these units is often not linked to clearly delimited stretches of discourse, but to information lying scattered throughout the previous or subsequent discourse (cf. also (203) and (204) above). In addition to indefinite and uncountable nouns, even the meaning of highly general plural nouns without a determiner may be context-dependent. This is shown in (217), where, according to Sinclair (1993: 11), the noun *things* performs a deictic act that encapsulates the underlined segment.

(215) 'If the chancellor really wanted to get more people into the saving and investing habit he could, without **difficulty**, go further and give relief on the first slice of an individual's income from his investments' (Ivanič 1991: 99)

(216) '"Where are we going, Mam?" inquired little Cadwallader when his child mind grasped the fact that green earth was falling away at a rate of knots. "Where indeed?" Megan Thomas spoke sharply to the conductor, demanding **an explanation**. But non-plussed, the conductor was' (Ivanič 1991: 111)

(217) 'The British have become less insular in some respects (cuisine is one of them) but they are still disastrously monoglot. This is the first generation in history to delude itself into thinking that because one particular language, English, seems to be very widely understood, no other language need be learned. The foreign language requirement in the UK's National Curriculum will help to change **things** […]' (Sinclair 1993: 11)

A plausible cognitive argument for the shell-noun status of examples like the aforementioned lies in Garnham & Oakhill's (1990: 380) claim that textual comprehension is enabled by the mental model that readers create of texts as they are being processed. The understanding of the current sentence or clause is thus dependent on the readers' integration of previous information into their mental discourse model (cf. Sinclair 1993 and 2004 in 3.2.2). On many occasions, however, the absence of explicit linguistic information in the co-text forces readers to search for an interpretation in their overall knowledge repository (Garnham & Oakhill 1990: 392). With this in mind, Brown & Yule (1983: 201) posit that '[…] the distinction between endophoric and exophoric co-reference becomes much harder to draw', insofar as readers always strive to make sense of text, irrespective of whether their interpretation is motivated by linguistic information in the text itself or by their knowledge of the world.

From what has been said so far, it is evident that definite or specific determiners are important indicators of shell-noun use, but also that they are by no means unique. Similarly, as regards the structures of postmodification, the literature places considerable emphasis on noun complement clauses (i.e. N-*that*/-*to*-cl and N-*be*-*that*/*to*-cl), as these are typical '[…] conceptual shells for complex, proposition-like pieces of information' (Schmid 2000: 4). Prepositional phrases, paramount among the types of noun phrase postmodification according to Biber et al. (1999: 634), tend to be disregarded in most shell-noun descriptions available. Some exceptions include Winter (1992: 157), Flowerdew (2003a: 337 and 2006: 358), Aktas & Cortes (2008: 10) and Caldwell (2009: 176). In these references, mention is made of shell-noun patterns with prepositional phrases headed by *of*, this being by far the most frequent postmodifying preposition in English (Biber et al. 1999: 635).

The importance of this preposition is such that Renouf & Sinclair (1991) conduct a corpus study on the possible co-occurrence tendencies between *of*-phrases and specific kinds of nouns. Their analysis rests on the concept of 'collocational framework', i.e. a sequence of grammatical words with an empty slot corresponding to a lexical item (Renouf & Sinclair 1991: 128). Two examples of such frameworks are *a/an* + *?* + *of* and *many* + *?* + *of*. Semantically, these two frameworks show a preference for 'abstract support nouns', which include shell nouns like *act*, *example*, *idea*, *aspect* or *impression* (Renouf & Sinclair 1991: 133, 138). Their support function hinges upon their delimitation and characterisation of the following noun phrase (e.g. *an example of this function*, *an aspect of this category*; cf. Mahlberg 2003 in 2.2.2.1.1). On many occasions, these collocational frameworks belong to specific phraseologies. For example, *an idea of* is often preceded by the verbs *give*, *get* or *have*, and followed by *what*, *how* or

a determiner (Renouf & Sinclair 1991: 140). According to Hunston (2008: 283–4), prepositions are crucial for the identification of certain phraseologies. This is borne out by the evidence obtained from a search for the pattern *the*-N-*that* in the *New Scientist* sub-corpus within the *BoE*. The most frequent items preceding the pattern are prepositions: *of*, *in*, *to* and *on*. An example is *the assumption that*, which is most often introduced by the preposition *on*, i.e. *on the assumption that*.

The discussion so far has revolved around the formal structure of shell-noun phrases. The literature, however, also makes passing reference to several semantic types of shell-noun premodifiers (Francis 1986; Ivanič 1991; Winter 1992; Francis 1994; Schmid 2000 and Flowerdew 2003a). These fall into two basic categories: modifiers with a subjective and attitudinal meaning (e.g. *inaccurate*, *useful*, *tremendous*, *wonderful*) and modifiers that objectively describe or classify the head noun (e.g. *big*, *small*, *scientific*, *medical*). The former are variously described as:

i) 'attitudinal modifiers' (Francis 1986: 55–6) (e.g. *down-to-earth approach*),
ii) 'modifiers with an evaluative role' (Ivanič 1991: 108) (e.g. *stern measures*),
iii) 'interpersonal modifiers' (Francis 1994: 95–8 and Moreno 2004: 335) (e.g. *hotly debated question*), and
iv) 'evaluative adjectives' (Schmid 2000: 318) (e.g. *terrible problem*).

The latter are described as:

i) 'propositional content modifiers' (Francis 1986: 59) (e.g. *monist vision*),
ii) 'modifiers restricting the range of reference of the head noun' (Ivanič 1991: 108) (e.g. *food problems*),
iii) 'specifics of identity' (Winter 1992: 154–5) (e.g. *population-related problems*),
iv) 'ideational modifiers' (Francis 1994: 95–8 and Moreno 2004: 335) (e.g. *new confectionary concept*, *empirical evidence*),
v) 'descriptive adjectives' (Schmid 2000: 318) (e.g. *clandestine atempt*),
vi) 'classifying adjectives' (Schmid 2000: 318) (e.g. *anthroposophical thinking*), and
vii) 'modifiers of more semantic importance than the signal itself' (Flowerdew 2003a: 336) (e.g. *functional studies*)

In addition to the two main categories presented above, the semantic modification of shell nouns is also apparent in three further types of modifiers. One of these types comprises words like *other*, *another*, *similar*, *same* or *different.* These are textual markers whose chief function is to spell out the clause relations underlying discourse structure (e.g. similarity, difference, addition, etc.). Only four references consider these modifiers, variously termed 'organisational' (Francis 1986: 60), 'textual' (Francis 1994: 98 and Moreno 2004: 333), and 'cohesive' (Schmid 2000: 318). Francis (1986: 62 and 1994: 99) introduced another category, that of 'comparative adjectives or epithets', which resemble textual modifiers in their connective role but differ in their overt attitudinal meaning (e.g. *an even more decisive argument*). The last semantic type is that of 'restrictive adjectives', mentioned only by Schmid (2000: 318). This category comprises adjectives like *main*, *only* or *real*, meant to restrict the reference of the noun.

Table 3.1 below outlines all the above-mentioned semantic categories and also illustrates their sparse coverage in the literature:

Table 3.1 Semantic types of shell-noun premodifiers

Francis (1986)	**Ivanič** (1991)	Winter (1992)	Francis (1994)	Schmid (2000)	Flowerdew (2003)	Moreno (2004)
Attitudinal Propositional	Evaluative Restrictive (range of reference)	Specifics of identity	Interpersonal Ideational	Evaluative Descriptive, Classifying	Semantically important	Interpersonal Ideational
Organisational			Textual	Cohesive		Textual
Comparative adjectives			Comparative epithets			
				Restrictive		

3.2.4 *The shell-noun phrase: syntactico-semantic functions and textual positions*

Research has produced valuable insights into the form of shell-noun phrases and their encapsulating function, but little is said about their syntactico-semantic role in clausal structure. Syntactically, the pre-eminence of Schmid's (2000) four shell-noun patterns in the literature leads to the generalisation that these units occur as subjects in N-*be*-cl and *th*-N (e.g. *the problem is that*, *this problem shows that*), as complements in *th-be*-N (e.g. *this is an important problem*) and as objects in N-cl (e.g. *they make the assumption that*). The typical occurrence of *th*-N in subject position may stem from the anaphoric referential status of definite and demonstrative nouns (see 3.2.2), coupled with their frequent encapsulation of given or known information (Yamasaki 2008: 80).

Leaving shell-noun patterns aside, English noun phrases are, according to Biber et al. (1999: 235), considerably more likely to appear as objects or complements of prepositions than as subjects. Aarts (1971, reprinted in 2004) finds an explanation for this in the structural complexity of the noun phrase. The examination of a 72,000-word sample from *the Survey of English Usage Corpus* reveals that the subject position is most often occupied by structurally light items, i.e. pronouns, names and nouns without pre- or postmodification (Aarts 2004: 36). This is especially the case in fiction and informal conversation. When it comes to structurally heavy items (i.e. nouns with pre- or postmodification), however, the corpus shows a marked preference for post-verbal positions (e.g. direct object, complement, etc.; Aarts 2004: 43). Scientific writing and formal spoken and written texts feature the largest number of post-verbal heavy noun phrases.

De Haan (1991), looking at postmodifying clauses, presents similar findings. Noun phrases with postmodifying clauses do not favour subject positions, as the subject slot is often occupied by given or familiar information which is in no need of further specification by a following clause (de Haan 1991: 55–6). Only noun phrases postmodified by non-restrictive relative clauses are found to occur frequently in subject slots. This is due to the non-specifying contribution of such clauses to the meaning of the noun. While relative-clause modification is not uncommon in subject position, noun phrases with appositive or complement clauses are predominantly sentence-final (de Haan 1991: 60). This is explained in terms of their frequent occurrence as complements of post-verbal prepositional phrases, as in (218).

(218) 'He pointed to the fact that this had not been mentioned before' (de Haan 1991: 60)

Compared to the slot-filling approach to syntax of the three aforementioned references, Francis (1991) investigates whether the syntactic distribution of nouns is equally frequent across individual items or not. The study uses the *Cobuild Corpus* to analyse 100 concordance lines for each of the items in a group of 38 nouns (e.g. *accident*, *artist*, *context*, *darkness*, *furniture*, etc.). The data reveal differences in the syntactic preferences of individual nouns (Francis 1991: 146–8). *Context*, for example, is chiefly an adjunct (e.g. *in the context of*), while *accident* tends to be adjunct or complement, as in (219a) and (219b). Interestingly enough, most such differences are linked to specific semantic senses and phraseologies (Francis 1991: 149–51). For example, the 'effect' meaning of the noun *impact* is associated with object position and with the phraseology *have/make* + *impact* + *on/upon*, as in (220a). The 'hitting' meaning of *impact*, however, is usually realised by an

adjunct, as in (220b). Francis (1991) brings to light the need to look at syntactic functions not as empty slots where any noun may occur, but as part of the grammar of individual lexical items.

(219) (a)'He vanished. Killed **in a motor accident** when we were in Cardiff […]' (Francis 1991: 148)
(b)'It must have been **an accident**, he felt certain of that […]' (Francis 1991: 149)
(220) (a)'[…] had an immediate and startling **impact on production** […]' (Francis 1991: 150)
(b) 'Hill 402 seemed to crumble with **the impact of enemy artillery fire**' (Francis 1991: 151)

As regards the semantic function of shell nouns, the sparse information provided in the literature appears to indicate that relational processes (e.g. *be*, *become*, *appear*) dominate their use. For example, Vendler's (1968: 73) connection between abstract nouns and the 'O *is* N' pattern (see 2.2.2.2) is indicative of such a tendency. Similarly, Ivanič's (1991: 97) choice of the term 'carrier' is not random, as it emphasises the role of these nouns '[…] in the structure of relational process clauses'. Lastly, Halliday & Matthiessen (2004: 472) observe that so-called fact nouns (e.g. *fact*, *sign*, *aspect*; see 2.2.1.3.1) are primarily used in intensive relational clauses, as evident in the extraposed construction in (221). Existential processes are a less frequent but not less significant choice for these nouns, as in (222) (Halliday & Matthiessen 2004: 473).

(221) 'In that article, it's **no coincidence** that I have a big fight with Twain and Eliot […]' (Halliday & Matthiessen 2004: 473)
(222) 'There is **evidence** that the Russians were just as surprised as anyone else at the suddenness and violence of them […]' (Halliday & Matthiessen 2004: 473)

Closely related to the syntactico-semantic behaviour of shell nouns is their textual association with either Theme or Rheme positions in clausal structure. In simple terms, 'Theme' is clause-initial and constitutes the starting point for the message, while 'Rheme' comprises the rest of the message (Halliday & Matthiessen 2004: 64). There is a tendency for Theme to occur as subject and to provide given or known information and for Rheme to introduce new or unknown information (see 5.3.3.3 for further details on the Theme system). From the literature reviewed in chapter 2, the conclusion appears to be that, through the encapsulation of a discourse segment, shell nouns act as both given and new discourse markers (cf. Francis 1986: 38 and 1994: 86; Schmid 2000: 350). Their givenness stems

from their reference to information which is already present in discourse, while their newness is a product of their characterisation and evaluation of such information through a single discourse entity. This double textual function leads Consten et al. (2007: 99) to argue that '[…] the strict dichotomy of thematic versus rhematic entities in texts has to be abandoned […]', at least in the analysis of shell nouns, as '[…] they are thematical […] and rhematical at the same time' (Consten et al. 2007: 96).

In line with these claims, Ravelli (2003: 57–60) suggests that shell nouns feature frequently as part of hyper-Themes[7]. These are introductory sentences anticipating the Themes to follow in a given paragraph. An example is (223), where *a further similarity*, in Theme position, establishes a connection between the similarities discussed in previous paragraphs and the similarity to be explained in the current paragraph. Ravelli (2003: 60) observes that abstract nouns like *similarity* are crucial to the organisation of discourse, as they '[…] function as the lynch-pin of the hyper-Theme, facilitating forward and backward connections'.

(223) 'A further similarity is that both systems have relied ultimately on force' (Ravelli 2003: 57)

Mauranen (1993: 108–11) comments that the degree of specificity of a given shell noun in a hyper-Theme is often matched by that of subsequent Themes. In (224), for example, *the results*, does not place any restriction on the Theme to follow because it is highly unspecific in meaning. Example (225) is more specific, as the Rheme of the first sentence does not refer to any argument, but to one in particular, i.e. Stiglitz's. This argument, or rather, its difference from that in the first sentence, is explained in the second sentence. These two examples reveal two kinds of '[…] prospective specificity […]' (Mauranen 1993: 109), one more neutral or unspecific than the other.

(224) 'When we take uncertainty into account, the results change dramatically. Investors would now require that the expected after-tax real rate of return on an investment at least equal the after-tax real risk-free interest rate […]' (Mauranen 1993: 108)

[7] Thompson (forthcoming: 360) provides evidence from the *BAWE* corpus that suggests that, at least, as far as undergraduate essay writing is concerned, '[T]he elements in argumentation [...]' or shell nouns like *conclusion*, *problem* and *argument* are strongly primed for paragraph-initial positions. Conversely, what he calls '[...] relations in argumentation' or conjunctive adverbials such as *thus*, *therefore* and *hence* show a closer association with paragraph-final sentences.

(225) 'This argument that the taxation income is non-distorting is entirely different from that of Stiglitz. Stiglitz's argument, developed in a certainty setting, relied on the possibility of 100 percent debt finance for marginal investments' (Mauranen 1993: 109)

3.2.5 *Semantic types of shell noun*

3.2.5.1 *Classifications*

Shell nouns may be classified as a range of semantic types. Despite differences in the scope of the classifications available, two categories prevail. One of such categories comprises nouns representing mental states and processes and their results (e.g. *consideration*, *speculation*, *theory*, *insight*). These nouns are variously described as:

i) 'Cognition nouns' (Francis 1986: 14–15),
ii) 'Mental-process nouns' (Francis 1994: 92–3),
iii) 'Belief nouns' (Francis et al. 1998: 109–10),
iv) 'Nouns referring to something thought or believed' (Hunston & Francis 2000: 187), and
v) 'Mental nouns' (Schmid 2000: 184–230).

The other category is that of nouns with a verbal or linguistic meaning. The realisation of such a meaning is most evident in nouns derived from illocutionary or speech act verbs, such as *announcement* (<*they announced that*), *prediction* (<*they predicted that*) or *suggestion* (<*they suggested that*). The category, however, also contains nouns which, whilst not derived from speech act verbs, illustrate various types of language providing propositional or informative content (e.g. *myth*, *definition*, *news*, *story*, *metaphor*). Various terms are used in the literature to describe the overall category of verbal or linguistic nouns. As shown below, only Francis (1986 and 1994) proposes two separate semantic types for the two aforementioned subcategories: 'illocutionary' and 'verbal' or 'language-activity' nouns. All the other terms either subsume both types under a single category or contemplate two subcategories. The latter is the case in Schmid (2000), where the class of linguistic nouns comprises the two subtypes of 'illocutionary uses' and 'propositional uses'. The terms proposed in the literature are the following:

i) 'Illocutionary nouns' and 'verbal activity nouns' (Francis 1986: 11–14),
ii) 'Illocutionary nouns' and 'language-activity nouns' (Francis 1994: 90–1),

iii) 'Suggestion and Answer nouns' (Francis et al. 1998: 108–9, 237),
iv) 'Nouns referring to something that is written or spoken' (Hunston & Francis 2000: 186–7),
v) 'Linguistic nouns', with the two subtypes of 'propositional uses' and 'illocutionary uses' (Schmid 2000: 131–83).

It is important to note that the boundaries between verbal and mental nouns are fuzzy in many respects. Francis (1994: 92) and Schmid (2000: 136–8), for example, claim that there are many nouns like *assumption*, *guess* or *assessment* that could be taken to denote both mental activities and their verbal outcomes. These ambiguous units would, according to Francis (1994: 92), occur in the middle of a cline that ranges from purely illocutionary nouns like *claim* or *statement* to purely cognitive nouns like *belief* and *idea*. As a rough-and-ready rule, Schmid (2000: 137–8) suggests that typical illocutionary nouns may collocate with such verbs as *perform*, *make* or *give* (e.g. *give a speech*, *make a point*), while typical mental nouns would primarily collocate with *have* or *hold* (e.g. *have an idea*, *hold a belief*). Schmid (2000: 138), however, recognises that even such a test may prove fallible in the light of such collocations as *form a plan*, where a dynamic verb appears with a clear instance of a mental shell noun.

Within the spectrum of verbal or linguistic uses of shell nouns, there is one other category that is contemplated only by Francis (1986: 16–17 and 1994: 93): that of 'text nouns'. The items contained in this category (e.g. *chapter*, *paragraph*, *passage*, *quotation* or *sentence*) are semantically neutral terms for different levels of discourse structure. Drawing on Lyons (1977, II: 667–8, see 3.2.2), nouns of this kind would fall within the scope of pure textual deixis, as their contribution to discourse is purely formal or syntactic rather than semantic.

There are many shell nouns which defy classification as either mental or linguistic (e.g. *fact*, *move*, *attempt*, *effort*, *chance*). As a result, these units are often put into such heterogeneous or ragbag categories as 'ownerless nouns' (Francis 1986: 17–18), 'non-metalinguistic labels' (Francis 1994: 89) and 'nouns that do not fit in any of the other groups' (Hunston & Francis 2000: 187–8). Schmid (2000) addresses such indeterminacy by proposing four further semantic types, i.e. 'factual', 'modal', 'eventive' and 'circumstantial' shell nouns.

'Factual' nouns are used to encapsulate various kinds of facts and states of affairs (e.g. *thing*, *reason*, *evidence*, *difference*, *aspect*, *problem*, *advantage*; Schmid 2000: 92–130). 'Modal' nouns express the speakers' or writer's assessment of the likelihood or necessity of the information in the encapsulated discourse segment (e.g. *possibility*, *probability*, *certainty*, *truth*, *permission*, *obligation*, *ability*; Schmid 2000: 231–60). 'Eventive' nouns rest

on Lyons' (1977, II) second-order entities (see 3.2.1), providing labels for activities, processes and states (e.g. *event*, *change*, *attempt*, *custom*, *achievement*, *crime*, *fault*; Schmid 2000: 261–74). Schmid's (2000) last semantic type is that of 'circumstantial' nouns, where such notions as time, place, manner or condition are involved (e.g. *position*, *area*, *moment*, *method*, *proviso*; Schmid 2000: 275–91).

Two other references where a subclassification is given of so-called ownerless or non-metalinguistic nouns (Francis 1986 and 1994) are Halliday & Hasan (1976: 274) and Francis et al. (1998: 108–20; see 2.2.1.3.2 and 2.2.2.1.1). Their categories are considerably more numerous than Schmid's (2000), with such labels as 'fact' (e.g. *question*), 'place' (e.g. *place*), 'action' (e.g. *move*), 'inanimate abstract' (e.g. *business*) or 'inanimate inconcrete mass' (e.g. *stuff*) in Halliday & Hasan (1976), and such fine-grained labels as 'desire' (e.g. *intention*), 'arrangement' (e.g. *deal*), 'ability' (e.g. *capability*), 'request' (e.g. *appeal*) or 'happiness' (e.g. *amazement*) throughout Francis et al'.s (1998) grammar. Schmid's (2000) framework provides a similar degree of detail. The difference lies in that, while in Francis et al. (1998) categories like 'request' or 'ability' are treated as main semantic types, in Schmid (2000) such fine-grained labels are subsumed under the more general semantic features of Linguistic and Modal, respectively.

Schmid (2000), therefore, classifies shell nouns into six main semantic features (i.e. Mental, Verbal, Factual, Modal, Eventive and Circumstantial), those of which are then subdivided into groups and families. Groups contain families of nouns sharing a specific semantic feature. For example, the nouns *goal*, *wish*, *motivation* and *determination* are all part of the mental group of psychological state uses and of the sub-group of volitional uses, as they all share the semantic feature Volitional (Schmid 2000: 209–25). These examples, however, differ in how this '"wanting" that future events […] take place' (Schmid 2000: 209) is actualised. As such, a noun like *goal* belongs to the family of 'Aim' nouns (e.g. *goal*, *aim*, *objective*, *idea*), all of which share the semantic feature Conclusive, insofar as the state of volition is intended with an endpoint or conclusion in mind (e.g. *his goal is to get there*). By contrast, a noun like *wish* would fall into the family of 'Desire' nouns (e.g. *desire*, *intention*, *willingness*, *longing*), where an Emotive or personal nuance is added to the state of wanting something to happen (e.g. *his desire is to get there*).

Schmid (2000: 85–6) uses his semantic typology to distinguish three degrees of shell-noun typicality, i.e. 'prime', 'good' and 'less good' shell nouns. 'Prime shell nouns' constitute the most prototypical instances of the category, as illustrated by most factual nouns (e.g. *fact*, *reason*) and some mental (e.g. *idea*, *concept*) and linguistic (e.g. *message*, *rumour*) nouns.

Many of them share a lack of morphological transparency, reflected in the absence of synchronic verbal or adjectival bases implicit in their internal structure. 'Good shell nouns' are shown in the classes of linguistic (e.g. *order*, *proposal*), mental (e.g. *belief*, *assumption*) and modal (e.g. *likelihood*, *certainty*) shell nouns. Unlike prime shell nouns, morphological decomposition into related verbs and adjectives is allowed in most cases (e.g. *proposal<propose*, *likelihood<likely*). Lastly, 'less good shell nouns' are exemplified by the eventive (e.g. *move*, *measure*) and circumstantial (e.g. *way*, *procedure*) types. Many of these nouns represent events, rather than abstract relations, which explains why, unlike typically abstract prime or good shell nouns, abstract shell contents cannot occur with them. Furthermore, these nouns are found to occur much less frequently in the patterns N-cl and N-*be*-cl.

With all of the above in mind, Schmid (2000) is revealed as the most thorough semantic classification of shell nouns to date. The strengths of this taxonomy include its identification of meanings often subsumed under single ragbag categories and its hierarchy of semantic features, ranging from most general or coarse-grained (e.g. Factual) to most specific (e.g. Conclusive). For these reasons, Schmid's (2000) classification is chosen in this thesis as the basis for the semantic categorisation of the units sampled.

Table 3.2 below presents an overview of the areas of overlap of all the semantic classifications and categories referred to throughout this section.

Table 3.2 Semantic types of shell noun

Halliday & Hasan (1976)	Francis (1986)	Francis (1994)	Francis et al. (1998)	Hunston & Francis (2000)	Schmid (2000)
	Illocutionary nouns Verbal-activity nouns	Illocutionary nouns Language-activity nouns	Suggestion, Answer	Nouns referring to sth. written/spoken	Linguistic
	Cognition nouns	Mental-process nouns	Belief, Happiness, Desire […]	Nouns referring to sth. thought/believed	Mental
	Text nouns	Text nouns			
Fact, Inanimate abstract			Sign, Reason, Other […]		Factual
	Ownerless nouns	Non-metalinguistic labels	Possibility, Ability […]	Nouns which do not fit into any	Modal
Action			Attempt, Challenge, Tragedy […]		Eventive
Place			Time, Place, Way […]		Circumstantial

3.2.5.2 *Shell-noun uses*

Shell nouns are assigned semantic features based on their use in context (Schmid 2000: 87). As such, labels like Linguistic or Mental are not treated as semantic types, but as semantic uses. Such an emphasis on use leads Schmid (2000) to allow for the existence of nouns with multiple membership. *Point*, for example, may be factual in focusing constructions (e.g. *the point is…*), linguistic, when implying an assertive illocutionary act (e.g. *she made a point<she pointed out…*), and circumstantial, when followed by a *wh*-clause (e.g. *we progressed to the point where they have become…*) (Schmid 2000: 93–102, 140–4, 279–82).

Whilst not denying the important context-sensitive component of Schmid's (2000) taxonomy, one caveat is that many semantic senses of nouns are missed out as a result of the strong emphasis on the N-cl and N-*be*-cl patterns, as well as on '[…] the frequent and […] linguistically preferred and cognitively more entrenched, shell nouns […]' (Schmid 2000: 40). An example of overlooked semantic senses is (226), contained in the sample used for this thesis. The meaning of *point* here is that of 'purpose' or 'aim', a semantic sense missed by Schmid (2000). This is probably due to the frequent co-occurrence of this sense with a following *of*-phrase, one of the patterns excluded from Schmid's (2000) automated analysis. In this particular instance, what the *point* is *of* is implied in the underlined clause (i.e. *the point of going back into the city*), but the actual *purpose* or *point* of such an action is not specified in discourse (*dying of starvation* does not explain why they would want to go there, only the likely result of their going back there).

(226) […] they didn't want to go over to the Serians because they'd be killed, they didn't want <u>to go back into the city</u> because they weren't allowed there and any way what was **the point**, they'd only die of starvation in there <pause>[…] (*BNC Sampler*: FLU, S:sermon)

Schmid's (2000) contextualised semantic approach is mainly influenced by the structural patterns in which shell nouns occur. In this respect, Schmid (2007) goes further into the kinds of meanings emerging from the N-*that*, N-*to*, N-*be*-*that* and N-*be*-*to* patterns. Inspired by Goldberg's (2006) Construction Grammar, this paper explores the co-occurrence restrictions between nouns and *that*- or *to*-infinitive clauses, as well as the '[…] emergent meaning components […]' (Schmid 2007: 315) that certain shell nouns acquire in combination with either of these two clause types. Three of Schmid's (2000) semantic types are often linked to only one of these clauses. Factual nouns are compatible only with *that*-clauses (e.g. *the fact*

that, *the thing is that*), while eventive and circumstantial nouns co-occur with *to*-infinitive clauses (e.g. *an attempt to*, *the time to*; Schmid 2007: 325). Linguistic, mental and modal nouns accept both *that*- and *to*-infinitive, the former associated with propositional, conceptual and epistemic uses (e.g. *the news is that*, *the idea that*, *the possibility that*) and the latter with directive, volitional, deontic and dynamic uses (e.g. *the order was to*, *the plan to*, *the obligation to*, *the ability to*; Schmid 2007: 325).

While these co-occurrence restrictions reflect a general tendency in shell-noun use, they do not fully account for the many instances where a specific semantic feature arises that is not part of the meaning of the lexeme in isolation. This is demonstrated in relation to three 'test cases' (Schmid 2007: 326–8). One such case concerns nouns showing a marked semantic difference linked to the occurrence of a *that*- or a *to*-infinitive complement. For example, the noun *answer* is most typically a linguistic noun, as in (227a). However, when a *to*-infinitive clause is added, as in (227b), the meaning of the noun changes into that of 'solution', thereby adopting a clear eventive and volitional meaning. The second test case corresponds to highly unspecific circumstantial nouns (e.g. *time*, *place*, *way*) acquiring a modal meaning when combined with a *to*-infinitive clause. This is shown in (228) below, where, as indicated by the paraphrase between brackets, the noun *time* acquires an extra semantic nuance of dynamic modality contributed by the following infinitive clause. The third test case looks at the various meanings of the infinitive clause in relation to specific nouns. For instance, while in (229) the infinitive adds a meaning of deontic obligation (i.e. *what we have to do is raise…*), in (230) a meaning of volition is entailed (i.e. *what we want to do is raise…*). From these three test cases, Schmid (2007: 328) highlights the complexity of N-cl and N-*be*-cl constructions, noting that '[…] the constructional meaning must be the result of an intricate interaction of noun and complement meanings'.

(227) (a)'The answer is that we haven't got any money' (Schmid 2007: 326)
(b)'The answer was to raise money for books in the library' (Schmid 2007: 326)

(228) 'The best time to encourage your older child to start caring for a new baby is before the birth' ('the time when it is best possible to…') (Schmid 2007: 327)

(229) 'The task is to raise money for new books in the library' (Schmid 2007: 327)

(230) 'The idea is to raise money for new books in the library' (Schmid 2007: 327)

At the end of the article, Schmid (2007: 329–33) offers evidence for the constructional meaning of the N-*be*-*to* pattern. His automated retrieval of

examples from the *BoE* reveals that the pattern is often used in conjunction with a meaning of 'intended result' and 'obligation' (Schmid 2007: 330). The former meaning is by far the most prominent, as shown in the occurrence of *aim* as the most frequent noun in the construction. Synonyms rank high too, like *intention*, *objective* or *purpose*, followed by obligation nouns like *job* or *task*. In view of the strong association of the pattern with nouns like *aim* and *job*, it is unsurprising that nouns with none of these meanings in isolation end up assimilating the semantic features of the N-*be-to* construction (Schmid 2007: 331). For example, *idea* becomes synonymous with *aim* when combined with an infinitive clause, as in (230).

Schmid's (2000 and 2007) research shows the identification of shell-noun meanings to be crucially dependent on their associated complementation structures. From such a standpoint, shell nouns are not watertight self-contained entities, but lexical items with a variety of semantic senses emerging from their use in context. This said, however, it should be noted that, whilst *that*- and *to*-infinitive clauses are useful indicators of shell meanings, they are not unique. Meaning, and by extension, formal, pragmatic and textual aspects of shell nouns closely relate to their use in a range of structures (not only *that*- and *to*-infinitive clauses) and discourse situations. Therefore, following Schmid (2000: 13, 88), the emphasis in this thesis is on 'shell-noun uses'. Bearing this in mind, the terms 'shell noun', 'shell-noun phrase', 'shell-noun use' and 'shell use', are, as in Schmid (2000), utilised here interchangeably in relation to any shell-like context-specific sense of English second- or third-order abstract nouns. For a full discussion of the criteria applied in this thesis to separate shell from non-shell uses, reference is made to section 4.4.2 in the following chapter.

3.2.6 *Towards a multifaceted approach: Hoey's (2005) lexical priming*

The previous sections discuss five key issues of the description of shell nouns. From the literature reviewed in chapter 2 and from 3.2.1 to 3.2.5 in the current chapter, the conclusion appears to be that no study has yet taken explicit and systematic account of lexicogrammatical, semantico-pragmatic and textual areas of linguistic enquiry. One such multifaceted approach to the study of shell nouns could be inspired by Hoey's (2005) theory of 'lexical priming'.

Hoey's (2005) theory rests on the assumption that the use of any word is affected by the writer's or speaker's previous experience of the word. Every single word is thus '[…] primed for use in discourse […]' (Hoey 2005: 13) based on the amount and type of prior exposure to such an item. Lexical priming comprises 10 hypotheses covering formal, semantico-pragmatic

and textual areas of linguistic analysis (Hoey 2005: 13). These may be summarised in the following 5 points:

i) Every word collocates with particular lexical items (i.e. 'collocations') showing specific semantic associations.
ii) Every word is linked to specific pragmatic functions.
iii) Every word is likely to occur only in certain grammatical patterns, in particular positions in the clause (i.e. Theme or Rheme) and with certain syntactic functions. These are the word's 'colligations'.
iv) Collocational, semantic and colligational primings are associated with individual words. Synonymy and co-hyponymy, therefore, do not lead to priming similarities. Polysemy produces priming differences, with individual word senses having specific collocations, semantic associations and colligations.
v) Textually, every word favours particular types of cohesive and semantic relations, as well as certain positions in the overall structure of discourse. These are a word's 'textual collocations', 'textual semantic associations' and 'textual colligations'.

Hoey (2005) tests his hypotheses primarily on data from a 95 million-word corpus of newspaper articles from *The Guardian*. The choice of such a genre-specific corpus is guided by the idea that any '[...] claims about priming have to be domain and genre-specific' (Hoey 2005: 133). Corpus investigation is of paramount importance in the theory of lexical priming, where, in the absence of data for every individual's lexical primings, the corpus provides valuable evidence on the kinds of data language users encounter in their everyday experience of language (Hoey 2005: 14).

Hoey's (2005) exploration of the above-mentioned hypotheses makes frequent reference to the shell noun *consequence*. As regards hypothesis i), *consequence* is found to collocate most frequently (59%) with premodifying adjectives expressing the logic of the process that the noun is describing (Hoey 2005: 24–6). These logical adjectives fall into 'necessity' (e.g. *inevitable*, *inescapable*), 'directness' (e.g. *direct*, *ultimate*) and 'naturalness' or 'expectedness' adjectives (e.g. *natural*, *predictable*). As an example of hypothesis ii), Hoey (2005: 28) mentions the noun *reason*. The corpus evidence used by Hoey shows that, on many occasions (34%), the noun is used to deny the existence, knowledge or importance of a reason, as in (231):

(231) 'Mahathir sees no reason to tinker with success' (Hoey 2005: 28)

Hypothesis iii) reveals a strong tendency for *consequence* to occur as part of an adverbial adjunct (43%), as in (232). When compared to the nouns *question*, *preference*, *aversion* and *use*, *consequence* is shown to be positively primed for adjunct and subject complement positions (24%) (e.g. *this is a consequence of…*), but negatively so for object positions (4%) (e.g. *it had important consequences*; Hoey 2005: 46–7). In relation to Theme and Rheme, almost half of the instances in the corpus (43%) occur as Themes, most of which are adverbial adjuncts (as in (232); Hoey 2005: 49–50). Grammatical colligations are also apparent in the determiners used in certain syntactic functions. For example, in subject position, *consequence*, *question*, *preference*, *aversion* and *use* are more common with definite than with indefinite determiners (Hoey 2005: 56–7). However, *consequence* and *use* feature more indefinite determiners than the other nouns (both 33%), the former preferring *one* (61%) and the latter *a*/an (67%). With regard to definite determiners, *consequence*, *question* and *use* are positively primed for *the* (99%, 96%, 64%), while *preference* and *aversion* prefer a possessive determiner (70%, 73%). The subject position of *consequence* is also associated with a particular pattern, i.e. *consequence* + *be* + *that*-clause/nominal group (Hoey 2005: 57). Example (233) below illustrates the pattern with a noun-phrase complement (cf. Ivanič 1991; Flowerdew 2003a and Gray 2010 in 3.2.2, where the possibility of a noun-phrase antecedent is also considered).

(232) 'But as a consequence of past neglect, this 'recovery' is different' (Hoey 2005: 52)
(233) 'But the consequence could be the retention of large numbers of alternative syllabuses in the subject' (Hoey 2005: 57)

The fourth hypothesis (iv) casts light on the priming differences between *consequence* and *result*. Looking at the premodifying adjectives of each noun, it is shown that *consequence* has a negative semantic association (3% positive, 15% negative), while *result* has a positive one (22% positive, 8% negative) (e.g. *adverse consequence* vs. *encouraging result*; Hoey 2005: 70). Another difference concerns the closer connection between *consequence* and indefinite determiners (33%) than that revealed by *result* (6%), which is more clearly primed for definiteness (94% vs. 67% for *consequence*; Hoey 2005: 71). Priming differences are also evident in polysemous words. In search of support for this claim, Hoey (2005: 83–101) looks at two senses of the nouns *consequence* and *reason.* The findings indicate that the primings linked to the most frequent senses differ markedly from those of the least frequent meanings. For example, while the 'result' sense of *consequence* is primed for Theme, as in (233), the 'importance'

meaning prefers Rheme (e.g. *your opinion is of little consequence*; Hoey 2005: 87). In the case of *reason*, an example is the association between the 'cause' meaning and the definite article (e.g. *the reason I called…*) and that between the zero article and the 'rational faculty' meaning (e.g. *listen to reason*; Hoey 2005: 89).

For the fifth hypothesis (v), no example of shell noun is provided. The conclusions reached, however, are especially relevant to the encapsulating and referential functions of shell nouns. Hoey (2005: 115–22) posits that words are more or less primed to occur in cohesive chains and to favour particular types of cohesion (i.e. 'textual collocations'). For instance, a noun like *planet* is likely to feature in *Guardian* articles as part of lexical chains comprising repetitions and hyponyms (e.g. *planet – Uranus – Saturn – planets – Pluto*; Hoey 2005: 122). Proper nouns, by contrast, are more primed to occur in cohesive chains with pronouns (e.g. *Mr Blair – his – Mr Blair's – he*). Whilst most such primings are linked to concrete or proper nouns, it is argued that even words with '[…] weak denotations (e.g. *ridiculous*, *make*, *action*) […]' (Hoey 2005: 118) may be primed for cohesive chains in special domains and genres. With this in mind, it appears that even such highly unspecific items as shell nouns may be associated with specific kinds of cohesion.

Textual primings of a semantic kind are also evident in discourse, the so-called 'textual semantic associations'. The argument is that every word will tend to prefer or reject certain clause relations and textual patterns (e.g. contrast, cause-effect, problem-solution; Hoey 2005: 122–4). Such an idea evolves from Winter's (1977, 1982 and 1992) and Hoey's (1979, 1983 and 1994) research on Vocabulary 3 items and signalling in discourse (see 2.2.2.1.2). Hoey's (2005: 123) claim is substantiated by corpus evidence for *sixty* and *ago*, which are shown to prevail in contrast relations. Hunston (2001: 28–30) offers a shell-noun instance of one such textual semantic association. *On the downside* in (234) is part of a contrast relation, inasmuch as it represents a link between the positive information in the previous sentence and the negative information to follow. Hunston (2001: 31) notes that these associations rest not on individual lexical items, but on phraseologies. Hence, the contrast relation in (234) is most often tied to *on the downside*, just as the contrast involved in some uses of *may* is often linked to the sequence *may not be a*.

(234) 'What follows is his selection of science writings that move chronologically from the Renaissance to the present. On the downside, we are, inevitably, limited to snapshots of each of the author's writings' (Hunston 2001: 30)

In addition to cohesive and semantic textual primings, words may also show a preference for certain positions in the overall structure of discourse. Hoey (2005: 129–51) proposes the term 'textual colligations' with a view to endowing the concepts of Theme and Rheme with further detail. The analysis looks at the occurrence of words in the first sentence of a text, in the first sentence of a paragraph that is not the first in the text and in any sentence that does not open a paragraph or text. Attention is also given to whether the word occurs in the first half of a sentence or in the second half. The application of this framework to the noun *consequence* reveals that, while *consequence* prefers Theme position, it tends to avoid paragraph-initial and text-initial positions (Hoey 2005: 130). The plural *consequences*, however, is less likely to occur as Theme, but is primed for paragraph-initial positions. Hoey & O'Donnell (2008) offer a detailed examination of textual colligations in an article from *The Guardian*. A shell-noun example in this paper is that of the comparison between the position of the noun *move* in the aforementioned newspaper article and its textual distribution in the *BoE*. The findings (Hoey & O'Donnell 2008: 303) point to the dominance of *a move* in text-initial sentences and of *the move* in paragraph-initial ones[8].

From the above hypotheses, Hoey's (2005) theory of lexical priming may be said to prove helpful in comprehensive research. Its consideration of a range of levels of linguistic description is the driving force behind the formal, semantic and textual variables analysed in this thesis. An example of a study combining a range of variables in the analysis of shell nouns is Hoey (1993). As a precursor to the theory of lexical priming, that paper examines 493 concordance lines for the noun *reason* retrieved from the *Birmingham Corpus of English* (the current *BoE*). The following summarises some of Hoey's (1993) most important findings:

i) 59% of the causal meaning of *reason* corresponds to clause-initial and subject complement positions (35%, 24%), as in (235) and (236). The judgement or rationality meaning of *reason*, as in (237), however, is only frequent in other positions (83%; Hoey 1993: 70).

(235) 'For some reason, the portrait looked like […]' (Hoey 1993: 70)
(236) 'That was the very reason I wanted to play a woman' (Hoey 1993: 70)
(237) 'But cool reason says there is no continuity' (Hoey 1993: 71)

[8] In a study on textual colligations in essays from the *BAWE* corpus, Thompson (forthcoming: 370) shows that textual priming does not affect only individual words. Textual colligations are relevant to combinations of individual words (i.e. words and their collocates), P-frames or collocational frameworks such as *one of the* * (see Renouf & Sinclair 1991 in 3.2.3) and their specific semantic, pragmatic and rhetorical associations.

ii) In clause-initial position, causal *reason* is more frequent as subject (60%) than as any other function (40%). This said, however, the frequency of *reason* as part of a clause-initial adverbial adjunct, as in (235), is higher for this noun than for other nouns in the language (Hoey 1993: 71–3).
iii) Certain correlations are revealed between the form of the noun phrase and its function (Hoey 1993: 77). Causal *reason* as subject will tend to be postmodified by *for*-x and by *Ø*-clause (e.g. *the first reason for caution is…*, *the main reason they won was…*). As a complement, causal *reason* is often followed by a *why*-clause (e.g. *this was the reason why…*). As an object, a preference exists for a postmodifying *to*-infinitive clause (e.g. *he sees no reason to…*).
iv) Causal *reason* in complement and object positions followed by *why*- or *to*-infinitive clauses is often associated with a denial of the knowledge or existence of the reason (e.g. *I see no reason why…*, *this was no reason to…*; Hoey 1993: 78)

3.3 CORPUS METHODOLOGY AND ANALYTICAL PROCEDURES IN SHELL-NOUN RESEARCH

This section addresses the corpus methodology followed in the literature on shell nouns. Two issues will be dealt with here, one being the corpora employed and, the other, the procedures used for the analysis. As regards the former issue, chapter 2 (section 2.3) offers an extensive discussion of the findings from the analysis of a range of general and genre-specific corpora. The following is a concise summary of such a section, geared only to the corpora utilised in such research.

Most of the generalisations about the use of shell nouns stem from the analysis of corpora which are both small and genre-specific. One broad genre, academic discourse (both written and spoken), occupies a prominent role in most related research. Four subtypes of academic corpora may be distinguished, i.e. discipline-specific professional writing (as found in textbooks, research articles and theses: e.g. Tadros 1985; Moreno 2004; Charles 2007; Gray 2010), undergraduate L2 learner writing (e.g. Francis 1988; Flowerdew 2006; Hasselgård 2012), L1 (usually professional) vs. L2 (undergraduate or graduate) writing (e.g. Aktas and Cortés 2008; Caldwell 2009) and lectures (e.g. Flowerdew 2003a; Lorés 2006). Another genre featuring in the literature is journalistic prose, epitomised by Francis' (1986 and 1994) use of articles from *The Times* and the monthly journal *Encounter*. It should also be noted that some genre-specific research has relied on diachronic corpora, as in Kanté (2010a), which compares a

synchronic corpus of research papers on linguistics with a diachronic corpus of judicial proceedings.

Compared to the dominance of genre-specific references, very few studies to date have made use of general English corpora. The three corpora employed in those studies are the *LOB Corpus* (Ivanič 1991), the *BoE* (through its different compilation stages: Hoey 1993; Hunston & Francis 2000; Schmid 2000; Mahlberg 2005) and the *BNC* (Aijmer 2007; Yamasaki 2008). *LOB*, containing approximately 1 million words of written English, is the smallest of the three corpora. Ivanič (1991) utilises this corpus as a source of examples, but she does not provide any frequency data about genre-related distribution of units and patterns. The *BoE*, amounting to 450 million words by the time Mahlberg (2005) went into print, is the largest of the general English corpora available. However, as Schmid (2000: 43) recognises, this corpus provides size, but not balance, inasmuch as newspaper language makes up about 70% of the whole corpus. Lastly, the *BNC*, with about 100 million words, is not as large as the *BoE,* but in terms of genre distribution, it is better balanced than the *BoE* (notwithstanding the 90% vs. 10% proportion of written and spoken modes). The two studies where the *BNC* is used, nevertheless, do not report any genre-specific findings, insofar as the focus is placed solely on mode (i.e. spoken language: Aijmer 2007, and written vs. spoken language: Yamasaki 2008).

From the above it follows that the choice of a genre-balanced corpus may play a significant role in the results obtained. Use of a small genre-specific corpus will reveal insights into shell-noun behaviour in a specific type of written or spoken text, but, at the same time, unless findings are compared with those from a better balanced corpus, no generalisations ought to be made about such behaviour in general English.

Just as relevant to a description of shell-noun use is also the procedure applied to the analysis of the corpus. The literature consulted reveals three types of analysis, i.e. automated large-scale corpus analysis, semi-automated large- or small-scale corpus analysis and manual small-scale corpus analysis. 'Large-' and 'small-scale corpus analyses' form part of Bednarek's (2009: 20–2) three-pronged approach to corpus-assisted discourse analysis. The third method, that of 'individual text analysis' (Bednarek 2009: 22–3), applies to the intensive examination of individual texts, as advocated by critical discourse analysis.

'Automated large-scale corpus analyses' (e.g. Francis 1993; Biber et al. 1999; Hunston & Francis 2000; Schmid 2000 and 2007) are those conducted on large corpora (most commonly, the *BoE*), such that examples are retrieved from the corpus based on predefined queries, those of which typically represent the N-cl and N-*be*-cl patterns. While offering abundant evidence on the most frequent nouns in these patterns, the rigidity of corpus

queries results in the exclusion of many potential shell-noun instances. This is most evident in the frequent disregard of prepositional patterns.

A lower degree of automation appears in 'semi-automated large or small-scale corpus studies', where predefined querying is followed by manual analysis. Two references illustrate this approach, i.e. Yamasaki (2008) and Caldwell (2009), the former using the *BNC* (a large corpus) and the latter using three small corpora of undergraduate L1, L2 and professional writing on cognitive psychology. Yamasaki (2008) looks at the *th*-N (e.g. *this problem*) and *th-be*-N (e.g. *this is a problem*) patterns in terms of their evaluative potential. The decision to focus on both follows the retrieval and subsequent manual analysis of a set of random concordances for 73 shell nouns. Queries for both patterns are then run on the corpus based on 5 nouns, i.e. *change*, *shift*, *failure*, *mistake* and *problem*. A subsequent manual analysis identifies the type of premodification and generic distribution (written or spoken English) of the nouns selected. Caldwell (2009: 78) conducts what she calls 'syntactically-motivated searches' of both typical shell-noun patterns and patterns excluded from fully automated analyses (e.g. N-*of*), with a view to comparing their occurrence across non-professional native and non-native writing, and professional writing. A manual analysis then establishes the kind of reference that definite shell-noun phrases show in the first 200 words of every text that makes up the three corpora. From the methodology followed in Yamasaki (2008) and Caldwell (2009), their semi-automaticity is explained to the extent that predefined queries of patterns are used as a starting point for some sort of subsequent experimental manual analysis of contextualised data.

Automation is minimal in 'manual small-scale corpus analysis'. In general terms, no predefined patterns are used and small genre-specific samples of language are preferred over large corpora. The scope of the references consulted (e.g. Francis 1986; Hoey 1993; Flowerdew 2003a; Mahlberg 2005; Lorés 2006), however, proves either too general or too specific. Francis (1986) and Flowerdew (2003a), whilst looking at a wide range of nouns, are primarily concerned with the discursive and semantic aspects of these units. The data (presented as generalisations) tend to be unaccompanied by quantitative results. Besides, whilst the former looks at newspaper language, the latter examines a corpus of undergraduate biology lectures and textbooks. Mahlberg (2005) uses the *BoE* to describe the '[…] local textual functions' (Mahlberg 2005: 59) of 20 general nouns. Such functions represent the meanings that general nouns acquire when used in context (e.g. *time* meaning 'measurement', 'history', 'life', etc.). Some general nouns qualify also as shell nouns (e.g. *thing*, *fact*, *system*, *problem*), whilst many others (e.g. *year*, *number*, *man*, *people*) are like shell nouns only on the basis of their unspecific meaning, but not in terms of their reference,

inasmuch as only shell nouns may refer to stretches of discourse. Finally, Hoey (1993) and Lorés (2006), unlike the former, provide more quantitative detail, but their findings are restricted to only the noun *reason* in Hoey (1993) and *thing(s)* and *idea(s)* in Lorés (2006).

A fully manual analysis of corpus data is important in certain areas of discourse analysis, especially those where the aim is to further our understanding of the overall linguistic and discourse behaviour of some units. This approach has limitations in that the close reading of long passages of discourse forces the researcher to reduce considerably the number of examples to analyse. A fully automated analysis has the advantage that it allows the processing of large amounts of data and is less taxing for the researcher, but at the same time it is less '[…] revealing with regard to textual patterns' (Mahlberg 2005: 58). Drawing on Bednarek's (2009)'s three-pronged approach to corpus research, the study in this thesis would thus fall into manual small-scale corpus analyses, in that a '[…] manual, context-sensitive analysis […]' takes precedence over the pattern-based automated analyses of lexical cohesion favoured in the literature on shell nouns (Bednarek 2009: 21–2). With about 2 million words, the corpus employed (i.e. the *BNC Sampler*) is not as small as those found in small-scale corpus analyses (approximately 70,000-80,000 words), but one which, in Flowerdew's (2011) terms, might be called a 'large small corpus', i.e. large enough to be representative of the language but small enough for the purposes of a manual approach to the analysis of lexical cohesion.

3.4 CONCLUSION

This chapter has reviewed the analytical decisions driving the identification of shell nouns in the literature. The following summarises those decisions with a direct bearing on the research methodology followed in this thesis (see chapters 4 and 5):

i) Drawing on Schmid (2000), only semantically abstract second- and third-order entities (Lyons 1977) qualify as shell nouns. This being the case, however, second-order entities like *arrival* or *fight* and third-order units like *isolation* and *loneliness* should not be treated as shell nouns, insofar as they are semantically bounded and in no need of further specification by the surrounding co-text (Schmid 1999). With this in mind, Lyons' (1977) ontological categories ought to be treated as a cline, with certain abstract units closer to the specificity of first-order concrete entities, and others, whilst more concrete in nature, requiring the context-sensitive interpretation typical of shell nouns (see, for example, *this graffito* in (190) above).

ii) 'Encapsulation' (Sinclair 1993 and 2004) proves to be the most appropriate term to describe the relation between shell noun and antecedent, as it avoids the terminological confusion apparent in such other terms as 'reference' or 'deixis'. In view of the almost exclusive concern with anaphoric encapsulation in the literature, it is worth investigating the extent of use of cataphora and exophora, especially in the light of Sinclair's (1993 and 2004) claim about the prospective or forward-looking nature of written and spoken discourse. Besides, whilst most research looks at long antecedents, the question is raised whether shorter noun-based antecedents are equally frequent in shell-noun use.

iii) In the literature, it is often repeated that shell-noun use correlates only with specific determiners, particularly *the* and the demonstratives *this* and *these*. This follows from the emphasis on anaphoric encapsulation, as *the*, *this* and *these* are frequently employed to refer to preceding linguistic information. Nevertheless, on the basis of Ivanič's (1991: 99, 111) and Sinclair's (1993: 11) tentative evidence, the question arises whether the strict separation between specific and non-specific determiners becomes blurred when it comes to the understanding of shell-noun phrases. Thus, '[…] the context seems to play a role […]' in shell-noun interpretation (Ivanič 1991: 112; my emphasis), irrespective of whether the noun is preceded by a definite, indefinite or zero determiner.

iv) Shell-nouns are often said to occur in N-cl and N-*be*-cl patterns, i.e. noun complement clauses. While this may be the case, more attention should be given to other structures of nominal postmodification, particularly prepositional phrases, especially as, according to Biber et al. (1999: 606), these prevail in English noun phrases.

v) Regarding syntactico-semantic and textual functions, the literature makes only passing reference to their relation to shell-noun use. Following Francis (1991) and Hoey (1993 and 2005), syntactic, semantic and textual positions may not be generalised to the entire class of English noun phrases, as they are specific to individual items. It would thus be interesting to examine whether shell-noun use in general is primed for certain syntactico-semantic and textual functions and whether any differences arise from individual items.

vi) With respect to the semantic categorisation of shell nouns, Schmid's (2000) classification is the most comprehensive. Its structuring of categories around different levels of granularity, as well as its inclusion of four further categories to account for Francis'

> (1986) ownerless nouns make Schmid (2000) the most useful semantic taxonomy of shell nouns. It should be noted that, whilst this taxonomy relies on the context-sensitive identification of shell-noun senses, many senses are missed due to Schmid's (2000) primary focus on noun complement clauses and on the most frequent meanings. It is thus worth looking at both frequent and less frequent senses emerging from all formal patterns of shell-noun use.

Chapters 2 and 3 have shown that shell nouns represent a widely researched area, and also that a research niche appears to exist for a study where all linguistic levels are considered (formal, syntactic, semantic and textual) and where a corpus-driven approach to data is foregrounded. Given the genre-specificity of most related research, such an approach would need to be based on a small but representative (and well-balanced) sample of language. The following chapter presents the methodology employed in this thesis, with special emphasis on the corpus, the sample and the analytical procedure.

4 THE CORPUS, THE SAMPLE AND THE ANALYSIS

4.1 INTRODUCTION

Chapter 4 is one of two chapters dealing with the method used in this thesis, specifically with the method used for the identification of shell instances. The chapter comprises three main sections: the corpus (4.2), the sample (4.3) and the analysis (4.4).

The first section is about corpus selection. Firstly, it looks at representativeness and genre-related balance in English general corpora (4.2.1). It then explains the choice of the corpus, which, as noted at the end of chapter 3, should ideally be manageable but representative of the English language at large (4.2.2). Section 4.3 shifts the focus to the study sample. It comprises two subsections: one lists and describes the decisions made during data collection, i.e. extraction of a shell-noun frequency list (4.3.1), and the other lists and describes the decisions made during data analysis, i.e. the sampling procedures employed for the analysis of the list described in 4.3.1 (4.3.2).

Section 4.4 turns to the analysis itself. It starts by describing the corpus theoretical approach (Mahlberg 2005) underpinning the study (4.4.1). The variables used for the analysis (nine in total) are outlined there. Subsequent to the rationale behind the analysis, section 4.4.2 explores up to twelve categories of excluded non-shell instances. The aim of the latter section is to clarify the distinction between shell and clearly non-shell uses prior to the discussion of the nine variables in chapter 5.

4.2 THE CORPUS

4.2.1 *Representativeness, balance and genre typology in corpora*

Chapters 2 and 3 have brought to light the distinct genre-specific nature of most research on shell nouns. The corpora used in many of such studies are geared to certain broad written genres, namely academic prose and newspaper language. This raises the question of whether the encapsulating

discourse function of these nouns is genre-specific or not. With this in mind, this thesis sets out to describe shell nouns on the basis of a wider range of genres. As noted at the end of chapter 3, this goal is closely linked to the need for a study where a more context-sensitive analytical approach is foregrounded. In order to meet such goals or needs, it is important to reach a compromise between a manageable corpus, and one which offers a well-balanced and representative coverage of a range of English genres.

The first decision was to use a 'sample corpus' instead of a 'monitor corpus' (McEnery & Hardie 2012: 6–10). The *BoE* is a prime example of the latter. At almost half a billion words nowadays (McEnery & Hardie: 2012: 7), the *BoE* is a dynamic corpus of English made up of complete texts obtained mainly from newspapers. The genre bias of this corpus is justified by Hunston (2002: 31) in terms of the free and easy access to newspaper language, enabling the regular update of the corpus contents. Drawing on Sinclair's (1991: 18) claim that '[…] a corpus should be as large as possible […]', the *BoE* offers size at the expense of the desirability of annually updating the corpus with material from a range of sources (Hunston 2002: 31). Such a range or balance of sources lies at the core of the design of 'sample' or 'snapshot corpora' (McEnery & Hardie 2012: 8–9).

Unlike monitor corpora, these are compiled on the basis of a predefined sampling frame in order to ensure that the corpus is well-balanced and representative within that frame. Their 'snapshot' nature stems from their capture of the language used at a particular point in time (e.g. the 1960s in the *Brown Corpus*, the 1970s through the 1990s in the *BNC*, etc.). Unlike the *BoE*, these corpora are never updated. One other feature that distinguishes sample from monitor corpora is the selection of portions of texts rather than entire texts (as is the case with the *BoE*). This is done with the aim of avoiding over-representation of certain text types and, thus, to '[…] maintain, as far as possible, balance and representativeness' (McEnery & Hardie 2012: 152). An example of these corpora is the 100 million-word *BNC*. With texts in this corpus dating back to the latter part of the 20th century (i.e. 1974–93) (Hoffmann et al. 2008: 29), it might be argued that, whilst better balanced and more representative than the *BoE*, the *BNC* is rather outdated. However, as noted by Hoffmann et al. (2008: 45), except for the failure of the *BNC* to account for the proliferation of internet words during the last decade (e.g. *blog*, *website*, *to google*, etc.), many other areas of English lexis and grammar have remained unchanged and will remain so for a long period of time. Hence, despite its static nature, the *BNC* may still be said to contain language samples of present-day English.

The issue of representativeness in corpora generates considerable controversy in the corpus linguistics literature. This is due to the lack of hard and fast criteria as to what constitutes a representative corpus (Hoffmann et

al. 2008: 17), as well as the consideration of balance and representativeness as '[…] matters of degree' (McEnery & Hardie 2012: 10). Should representativeness be understood as comprising the most typical genres a language user is exposed to in everyday situations, then a representative corpus would be one with a large proportion of spontaneous spoken conversation and TV or radio programmes (Hoffmann et al. 2008: 16). However, if representativeness is to be understood as covering the overall use of the language, regardless of users' degree of everyday exposure to specific genres, then the corpus must contain as wide a variety of text types as enabled by the initial sampling frame (Hoffmann et al. 2008: 16).

The categorisation of corpus texts is also a contentious issue in corpus research. Lee (2001) finds a general lack of agreement in the criteria behind the scope of such terms as 'style', 'register' and 'genre'. Following a review of the various interpretations of these terms in the literature, Lee (2001: 45–6) posits that 'style' refers to a writer's or speaker's personal use of language, while 'register' concerns the lexicogrammatical field (i.e. topic), tenor (i.e. formal, informal) and mode (i.e. written or spoken) choices linked to general situations of language use. 'Genres' are the products of registers, insofar as registers lead to genres, i.e. culturally established text groups (Lee 2001: 46–7). For instance, the legal register represents a configuration of lexical and grammatical choices instantiated in such genres as courtroom debates or wills (Lee 2001: 46). Similarly, formal registers are realised in such broad genres as academic prose or official documents.

Lee (2001: 47) further refines the concept of 'genre' by arguing that it applies to '[…] groups of texts collected and compiled for corpora or corpus-based studies'. It is then claimed that seven attributes are involved in the identification of genres (Lee 2001: 49):

i) 'Domain', or the area of activity or knowledge where a genre is used (e.g. art, science, religion),
ii) 'Medium', or the channel through which a genre is used (e.g. spoken, written),
iii) 'Content', or what the genre is about (e.g. syntax, nuclear fission),
iv) 'Form', or the internal structure of the genre (e.g. research paper: introduction, methods, results and discussion, conclusion),
v) 'Function', or what the genre is used to achieve (e.g. informative, persuasive, instructive),
vi) 'Type', including such rhetorical categories as narrative, description or exposition, and
vii) 'Language', or the lexicogrammatical features associated with a particular genre.

These attributes classify genres into three levels of delicacy: 'super-genres', 'genres' and 'sub-genres' (Lee 2001: 48–9). 'Super-genres' contain such broad categories as fiction or newspapers. 'Genres', as basic-level categories, involve such labels as prose, poetry or drama for fiction, and broadsheets or tabloids for newspapers. 'Sub-genres', as the most delicate categories, include labels like western or romance for prose fiction and editorial, sports or commerce for newspapers.

An examination of textual categorisation in English corpora reveals a lack of consensus on the classification criteria adopted in each case (Lee 2001: 49–52). For example, the 1 million-word *LOB* Corpus blends together genres (e.g. press: reportage), sub-genres (e.g. mystery and detective fiction) and domains (e.g. religion) as main categories within its sampling frame. The *BNC*, however, draws only on general domain categories such as context-governed leisure or business for its spoken component and imaginative or informative pure science for its written component. Such broad domain labels are useful as superordinate categories subsuming a wide range of texts, but they do not account for specific types of texts. This is most evident in the category of informative texts, where no account is taken of such different genres as newspapers, non-academic (popularised) articles, research papers, reports or leaflets (Lee 2001: 53).

Lee (2001: 65) addresses the fuzziness of the *BNC* categories by providing a list of 70 genre tags incorporated in subsequent releases of the corpus. 46 of these correspond to written language (W) and 24 relate to spoken language (S). The tags give the maximum amount of textual detail, encoding information on genre and sub-genre. If narrowed to the second or third components of each tag (e.g. W:newsp and W:ac), the original list is reduced to 34 super-genre categories (18 for written and 16 for spoken). Appendix 1 presents Lee's (2001) list of 70 categories, as well as the reduced list of 34 super-genres. Descriptive information is added to the former with a view to spelling out the scope of application of each tag.

4.2.2 *Corpus selection*

Following the decision to use a sample corpus, a choice was made as to which corpus offers the most representative and balanced coverage of English genres. Four possibilities were investigated, a large corpus (i.e. the *BNC*) and three small ones (i.e. *BNC Baby*, *BNC Sampler* and *ICE -GB*).

The *BNC* contains about 100 million words of written and spoken British English (4049 texts in total) from the latter part of the 20th century. The corpus is heavily skewed towards written language, with a 90%-10% distribution of written and spoken English respectively. As stated above, the sampling frame followed originally draws only on domain information. As

such, the written component is divided into imaginative and informative texts, the latter further broken down into natural and pure sciences, applied science, social science, world affairs, commerce and finance, arts, belief and thought and leisure. The spoken sub-corpus consists of demographically sampled texts (i.e. spontaneous spoken conversations) and context-governed texts. The latter represent instances of spoken language used in four different domains, i.e. educational and informative, business, institutional and leisure (for further detail on the *BNC*, see Burnard 2007).

The *BNC Sampler* is a 2% sample of the entire *BNC*. The corpus contains about 2 million words and 184 texts in total, with a 50%-50% distribution of spoken and written language (1 million words each). The small-scale nature of this corpus allowed the development of a more detailed and accurate tagset than the one found originally in the larger corpus. Using the *BNC* sampling frame, the *BNC Sampler* was designed with the aim of mirroring the structure of the *BNC* as much as possible, while also providing a balance of written vs. spoken modes (50%-50%) lacking in the parent corpus. According to Lee (2001: 53–4), the *BNC Sampler* is representative and well-balanced in terms of domain, but not in terms of genre. With domain as the key sampling criterion, many *BNC* genres are either not present or marginally so in the *BNC Sampler*. For example, no university lectures are included, and only texts from the *Guardian* occur in the category of national broadsheets (for further detail on the *BNC Sampler*, see Burnard 2008a).

The *BNC Baby* is a 4 million-word sample of the entire *BNC*. It consists of 182 texts distributed over four 1 million-word sub-corpora covering four *BNC* super-genres, i.e. fiction, newspapers, academic prose and spoken conversation. Whilst differing from the *Sampler* in the more substantial evidence on four key super-genres, the *Baby* is considerably less representative of the *BNC* than the *Sampler*. For example, the *Baby's* spoken sub-corpus accounts only for conversation, excluding all the context-governed language that forms part of both the *BNC* and the *Sampler* (for further detail on the *BNC Baby*, see Burnard 2008b).

The *ICE-GB* Corpus does not belong to the *BNC* family. It is the British English component of the *International Corpus of English*. The corpus is the smallest of the four options considered, comprising about 1 million words and 500 texts of spoken and written British English produced from 1990 to 1993. The larger number of texts in this corpus is explained by its sampling procedure, with each text amounting to approximately 2000 words. The sampling procedure in the *BNC* family, however, allows for varying lengths. This results in fewer texts (as evident in the *Sampler* and the *Baby*). Written and spoken language in the corpus is not as evenly distributed as in the

Sampler, insofar as 60% of the corpus consists of spoken texts, while only 40% is made up of written samples. The textual categorisation in *ICE-GB* is, according to Lee (2001: 49–50), rather heterogeneous. It comprises three levels of granularity (sorted from least to most fine-grained): medium or interaction type (i.e. dialogue, monologue, mixed, non-printed and printed), super-genre or function (i.e. private, public, unscripted, scripted, non-professional writing, correspondence, academic writing, non-academic writing, reportage, instructional writing, persuasive writing and creative writing) and genres or sub-genres (e.g. face-to-face conversations, broadcast talks, humanities, novels) (for further detail on the *ICE-GB*, see Nelson et al. 2002).

Prior to corpus selection, and in order to better ascertain the balance and representativeness of these four corpora, a calculation was made of the proportion or percentages of Lee's (2001) genre categories in each corpus. Genre data for the *BNC* and the *BNC Baby* are offered in Burnard (2007: web page) and Burnard (2008b: 7) respectively. For the *Sampler*, however, only domain information is given, which means that genre data had to be obtained and calculated manually. The only way that this could be done was by extracting a sentence or two from each of the 184 .txt files comprising the XML version of the *Sampler*. The textual evidence obtained from each file was then fed into *BNCweb*, a web-based interface for the *BNC* created at Lancaster University (Hoffmann et al. 2008). Each of the files in the interface specifies the genre category of the text on the basis of Lee's (2001) tags. Only when each sentence from the *Sampler* was matched with the corresponding *BNC* file, was it possible to establish the genre category of the *Sampler* file. The number of words in each file was obtained from the bibliographical details provided in Burnard (2008a: 6–37). With regard to the *ICE-GB*, the text categories in this corpus were matched with Lee's (2001) using the bibliographical information in Nelson et al. (2002: 309–31). For example, *ICE-GB* categories such as academic writing humanities or news reportage correspond to the *BNC* tags W:ac:humanities_arts and W:newsp:brdsht_nat:report.

Table 4.1 below brings together the percentages for each of Lee's (2001) 70 genre categories in the four corpora. The percentages draw on the number of words for each genre category divided by the total number of words in each corpus. The choice of a tabular format is motivated by the sheer number of categories, which, if displayed in a graph, would be hardly visible.

Table 4.1 Proportions of the four corpora according to Lee's (2001) 70 genre categories. The shaded cells indicate lack of textual evidence for a given category

	BNC	*SAMPLER*	*BABY*	*ICE-GB*
S:brdcast:discuss	0.77	1.66		4.00
S:brdcast:documentary	0.04	0.50		4.00
S:brdcast:news	0.27			4.00
S:classroom	0.44	2.80		
S:consult	0.13			
S:conv	4.30	24.77	25.40	20.00
S:courtroom	0.13	1.21		4.00
S:demonstratn	0.03			2.00
S:interview	0.13			2.00
S:interview:oral_history	0.84	3.28		
S:lect:commerce	0.02			4.00
S:lect:humanities_arts	0.05			
S:lect:nat_science	0.02			
S:lect:polit_law_edu	0.05			
S:lect:soc_science	0.16			
S:meeting	1.41	8.59		2.00
S:parliament	0.10			2.00
S:pub_debate	0.29	0.75		
S:sermon	0.08	0.22		
S:speech:scripted	0.20	1.29		2.00
S:speech:unscripted	0.48	2.99		6.00
S:sportslive	0.03			4.00
S:tutorial	0.15			
S:unclassified	0.43	1.63		
W:ac:humanities_arts	3.41	1.34	5.62	2.00
W:ac:medicine	1.46		2.24	
W:ac:nat_science	1.14	1.65	5.38	2.00
W:ac:polit_law_edu	4.78	0.53	4.89	
W:ac:soc_science	4.87	1.07	5.24	2.00
W:ac:tech_engin	0.70	0.78	1.94	2.00
W:admin	0.23	0.31		2.00
W:advert	0.56	0.31		
W:biography	3.62	2.50		
W:commerce	3.87	3.33		
W:email	0.22			
W:essay:school	0.15			
W:essay:univ	0.06	0.73		4.00
W:fict:drama	0.05	1.19		
W:fict:poetry	0.23	1.12		
W:fict:prose	16.30	8.44	25.23	4.00
W:hansard	1.19			
W:institut_doc	0.56	0.92		
W:instructional	0.45			
W:letters:personal	0.05			3.00
W:letters:prof	0.07			3.00
W:misc	9.39	6.87		
W:news_script	1.27	0.15		

W:newsp:brdsht_nat:arts	0.36		0.91	
W:newsp:brdsht_nat:commerce	0.44		1.60	
W:newsp:brdsht_nat:editorial	0.10		0.22	2.00
W:newsp:brdsht_nat:misc	1.06		3.03	
W:newsp:brdsht_nat:report	0.68	6.03	1.20	4.00
W:newsp:brdsht_nat:science	0.07		0.46	
W:newsp:brdsht_nat:social	0.08		0.86	
W:newsp:brdsht_nat:sports	0.31		0.92	
W:newsp:other:arts	0.24		1.09	
W:newsp:other:commerce	0.43	1.04	2.23	
W:newsp:other:report	2.78	1.14	5.81	
W:newsp:other:science	0.06		0.34	
W:newsp:other:social	1.17		2.36	
W:newsp:other:sports	1.05	2.11		
W:newsp:tabloid	0.75		3.03	
W:non_ac:humanities_arts	3.81	2.28		2.00
W:non_ac:medicine	0.51			
W:non_ac:nat_science	2.58			2.00
W:non_ac:polit_law_edu	4.60	2.31		
W:non_ac:soc_science	3.77	0.61		2.00
W:non_ac:tech_engin	1.24	0.42		2.00
W:pop_lore	7.58	2.75		2.00
W:religion	1.15	0.36		

A comparison of *Sampler*, *Baby* and *ICE-GB* data reveals that, notwithstanding differences in proportions, the *Sampler* provides the most thorough coverage of the *BNC* genres. Whilst the *Sampler* fails to include 32 of the 70 *BNC* genres, the figure in the *ICE-GB* and the *Baby* rises to 37 and 48 respectively. This would translate into 12.45% of the *BNC* unaccounted for in the *Sampler*, 44.64% in *ICE-GB* and 54.51% in the *BNC Baby*. Of the three small corpora, the *Sampler* is thus the corpus that would, according to this, give the highest degree of genre representativeness of the *BNC* and, possibly, of English in general.

Figures 4.1 through 4.4 illustrate the distribution (in percentages) of Lee's (2001) 34 super-genres in each corpus. A trend line is added with a view to bringing to light any skewing in the proportion of certain genres.

Figure 4.1 The distribution of super-genres in the *BNC*

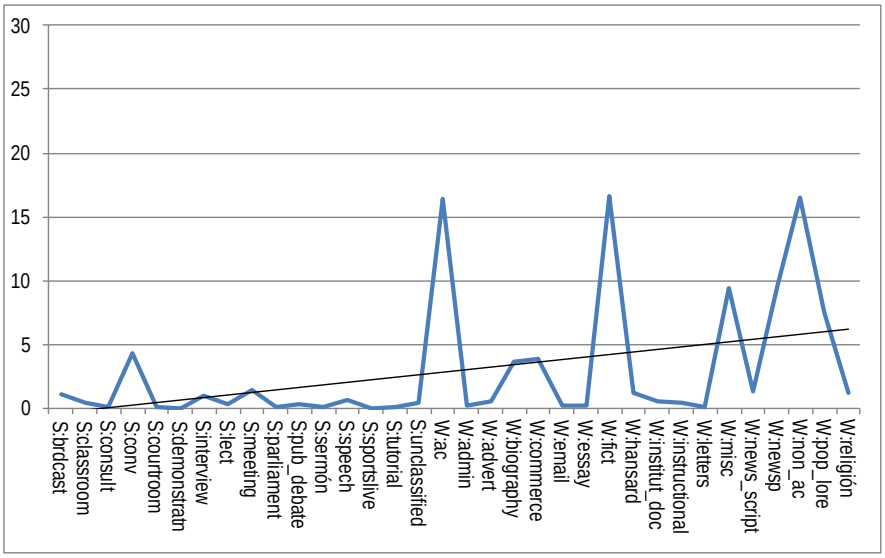


Figure 4.2 The distribution of super-genres in the *BNC Sampler*

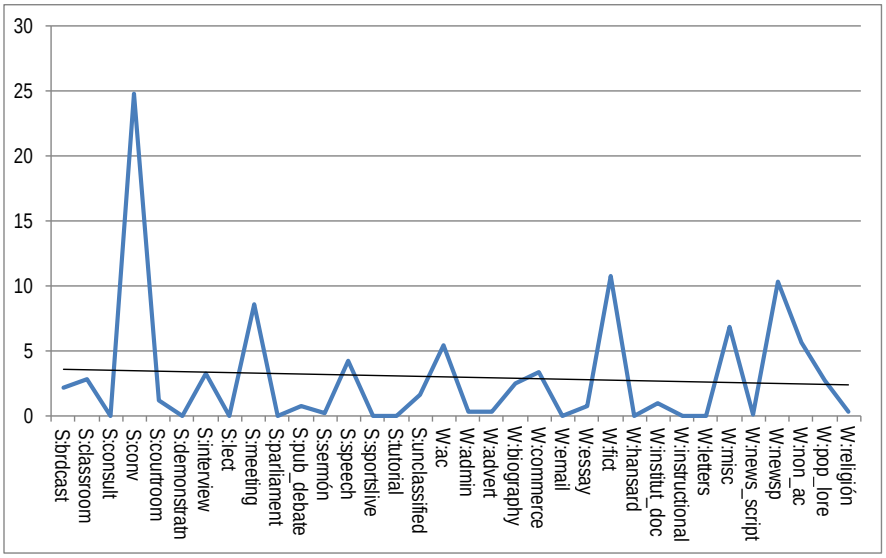


Figure 4.3 The distribution of super-genres in the *BNC Baby*

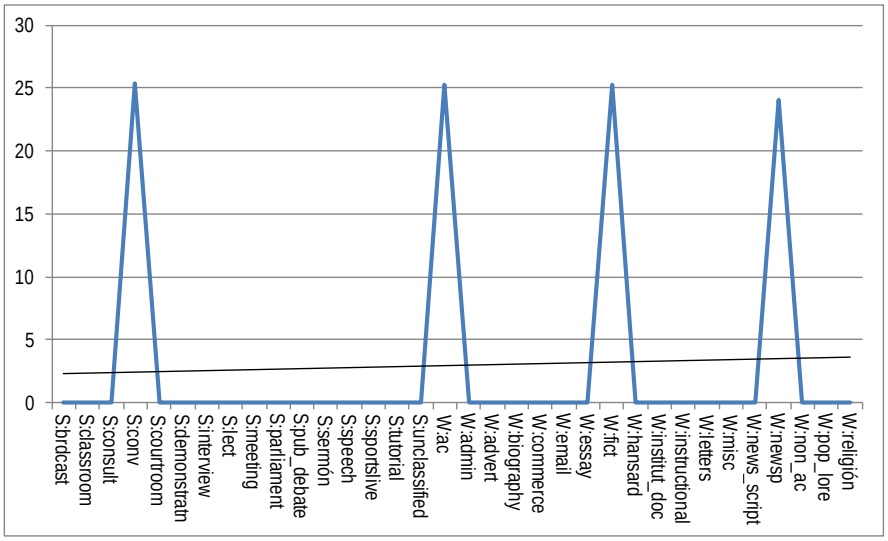
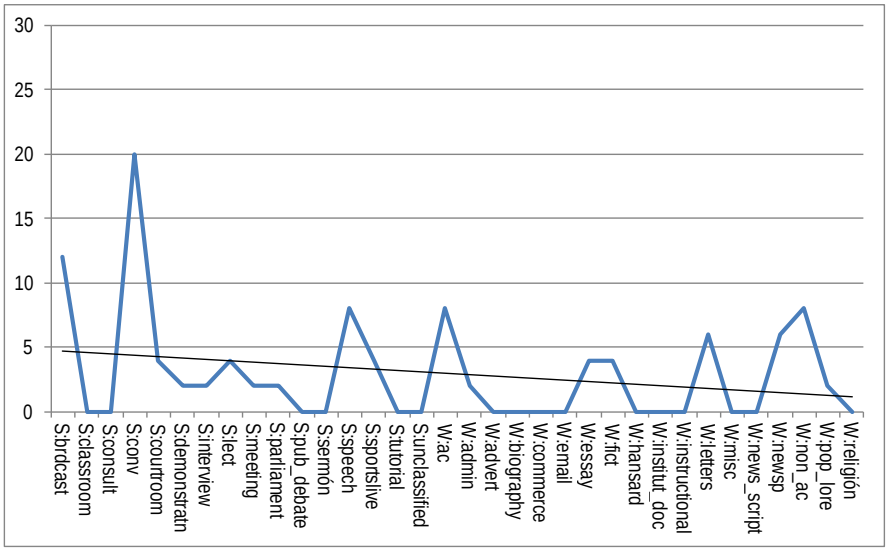


Figure 4.4 The distribution of super-genres in *ICE-GB*



Once again, of the three small corpora, the *Sampler* covers most of the *BNC* super-genres, with only 10 categories missing from this corpus. In the *ICE-GB* and the *Baby*, by contrast, this number rises to 16 and 30 respectively. In terms of linear distribution, the rising trend line in Figure 4.1 illustrates the markedly higher proportion of written language in the *BNC* (90%). Four super-genres peak high in this corpus: academic and non-academic prose, fiction and miscellanea.

Figure 4.2 reveals that, except for conversation (standing at 24.77%), a more even spread of the graph line is observed in the *Sampler* than in the *BNC*. Most genres do not peak higher than 5%, and those that do (meetings, fiction, miscellanea and newspapers) stay within the 5%-10% range.

Figure 4.3 shows the dominance of four super-genres in the *BNC Baby*: conversation, academic prose, fiction and newspapers. Lastly, Figure 4.4 suggests a more patchy coverage of Lee's (2001) super-genres in the *ICE-GB* than in the *Sampler*. Whilst the area from S:courtroom to S:parliament is more even than that in the *Sampler*, the *ICE-GB* written mode contains more areas for which no textual evidence exists. The *ICE-GB* downward trend line points to the higher proportion of spoken texts in this corpus (60%). Although a downward line is also apparent in the *Sampler*, a greater linear inclination is evident in the *ICE-GB*.

Taking the four figures into consideration, the *Sampler* reveals itself as the corpus covering most genres and offering the best balance and representativeness. It was, thus, decided to use the *BNC Sampler* for the study in this thesis. With most references based either on small genre-specific corpora or large monitor corpora like the *BoE*, the combination in the *Sampler* of small size and coverage of a wide range of genres is well-suited to the multifaceted manual analysis of shell nouns proposed in this thesis. One criticism that might be levelled at this decision is that such a small corpus as the *Sampler* (2 million words) cannot provide enough data about shell-noun use. Bearing in mind Lee's (2001: 53–4) misgivings about the representativeness of the *Sampler*, it might also be argued that the sparse textual representation of certain genres cannot lead to unfalsifiable conclusions about the genre-based distribution of shell nouns. Whilst not denying the truth of these claims, it is important to emphasise that, of the corpora tested, the *Sampler* offers the most even distribution of genres and modes (50% written, 50% spoken). In this respect, the *Sampler* fares better than the much larger but less balanced *BNC* and *BoE*.

4.3 THE SAMPLE

4.3.1 *The frequency list*

Following corpus selection, the literature was consulted for lists of shell nouns in order to compile a list of potential shell nouns as complete as possible on which to base the analysis. Of the references reviewed here, the following include more or less extensive lists of these items:

Table 4.2 A checklist of references including inventories of shell nouns

Lists of shell nouns in the literature
Winter (1977)
Hoey (1979)
Tadros (1985)
Francis (1986)
Ivanič (1991)
Winter (1992)
Francis (1993)
Tadros (1994)
Francis (1994)
Francis et al. (1998)
Hunston & Francis (2000)
Schmid (2000)
Flowerdew (2003a)
Hinkel (2004)
Flowerdew (2006)

Of these, Schmid (2000), as mentioned in previous chapters, gives the most comprehensive list to date, with 670 units.

Subsequent to the combination of all lists into a single list and to the deletion of repeated lemmas, the resulting study list contains 978 units. Such a list was then sorted in terms of the token frequency or the number of individual occurrences of each unit in the *BNC Sampler*. Frequency data are lemmatised in two ways, i.e. in the combination of singular and plural inflections under one lexical item (e.g. *area*, *areas>area*) and in the fact that, at this stage, no account is taken of the amount of shell-noun uses per item. This means that the token frequency of nouns like *room* or *field* encompasses such concrete first-order uses as *spacious room* or *grassy field* along with third-order uses like *no room for doubt* or *research field*.

Extraction of the frequency list drew on the 2005 XML edition of the *Sampler*. Following conversion of the 184 .xml files into .txt format, these were fed into *AntConc 3.2.4w* (Anthony 2011). *AntConc* is a freeware tool that enables corpus research through collocational, word-list, keyword-list, cluster and concordance analysis. Initially, the frequency list for all words in the corpus was extracted through the Word List tool. A preliminary examination of the list revealed that, whilst for certain noun-only lemmas (e.g. *thing*, *assumption*), the frequency figures could be typed into the shell-noun list straight away, for many other instances, more than one word-class could be assigned to a single lemma (e.g. *point*: verb, noun; *place*: verb, noun). Therefore, in order to ensure that the shell-noun frequency list only contains noun tokens, *AntConc* was searched for noun uses of each lemma.

The *BNC Sampler* contains part-of-speech (henceforth, POS) annotation through the *CLAWS7* tagset developed at Lancaster University (Burnard

2008a). Two of the tags in *CLAWS7* accompany most shell nouns in the list: <w NN1> (singular common noun: *problem*, *issue*, *point*) and <w NN2> (plural common noun: *problems*, *issues*, *points*). Three other noun tags occur only with certain shell nouns: <w NN>, <w NNT1> and <w NNT2>. The former in this sequence is attached in the corpus to common nouns which are morphologically neutral for number (e.g. *data*, *works*), while the latter two, used with singular and plural temporal nouns (e.g. *day*, *week*), occur in the shell-noun list only with the noun *time*. Additional noun tags like <w ND1> (noun of direction, e.g. *north*), <w NNA> (noun of title, e.g. *Mr.*), <w NP> (proper noun: e.g. *Andes*), <w NPD> (weekday noun: e.g. *Sunday*) and <w NPM> (month noun: e.g. *October*) are not used with shell nouns.

The number of singular and plural tokens for each lemma in the list was searched in *AntConc* primarily through <w NN1> and <w NN2>, as well as through <w NN>, <w NNT1> and <w NNT2> for those instances which were found to occur with none of the above tags. Prior to the search, the Words box in *AntConc* was checked so only the queried item (e.g. *trouble*), and no unwanted combinations (e.g. *troublemaker*), made it into the final count. The searches were conducted by typing in the singular or plural tag followed by the inflected form of the noun. For example, a query of the sequence <w NN1>*assumption* returns 23 hits, while <w NN2>*assumptions* returns 15 hits. The addition of both figures amounts to 38, the overall token frequency for the lemma *assumption*. The frequency list for this study is sorted on the basis of such added frequencies. The resulting list contains 922 lemmas, i.e. 56 of the original 978 units are unaccounted for in the *BNC Sampler*.

A sample of *AntConc* frequencies was then compared with those obtained from the *CQPweb* online interface for the *Sampler* (Hardie 2012). The aim was to check for any frequency differences arising from the use of two corpus tools. A random search for 4 noun lemmas from the top of the list (*system*, *place*, *point* and *answer*), 4 from the middle (*objective*, *surprise*, *recognition* and *excuse*) and 4 from the bottom (*drawback, regret, quest* and *absurdity*) revealed that in 7 out of the 12 units examined (*system*, *surprise*, *recognition*, *drawback*, *regret*, *quest* and *absurdity*), token frequency was the same in both tools. In 3 of the remaining instances (*place*, *point* and *objective*), a difference of less than 1% emerged. In the 2 units left (*answer* and *excuse*), the two tools showed a difference between 1% and 2%.

Thus, the marginal discrepancies found in the figures obtained do not alter the overall frequency distribution of the units examined[9]. Table 4.3 below

[9] A possible explanation for the divergence may lie in the erroneous POS annotation of cases of formal identity, as the 5 cases where differences arise belong to more than one word-class (*place*: noun, verb; *point*: noun, verb; *objective*: adjective, noun; *excuse*: noun, verb). Admittedly, both tools give the same figure for other similar cases (*surprise*, *regret* and *quest*).

presents the exact data for the nouns explored. As explained in 4.3.2, use of *AntConc* was discarded for the extraction and analysis of examples, because it does not provide a facility for random sorting of examples. For such a purpose, the more useful *CQPweb* interface was used.

Table 4.3 Token frequency discrepancies between *AntConc* and *CQPweb* expressed as percentages[10]

	AntConc (raw data)	*CQPweb* (raw data)	Difference (%)
System	1262	1262	0
Place	911	900	0.61
Point	1064	1044	0.95
Answer	243	238	1.03
Objective	74	73	0.69
Surprise	75	75	0
Recognition	39	39	0
Excuse	28	27	1.81
Drawback	12	12	0
Regret	9	9	0
Quest	7	7	0
Absurdity	3	3	0

4.3.2 *Sampling*

The 922-lemma frequency list thus obtained for the study was then divided into three frequency ranges for a stratified sample: high (1), middle (2) and low (3). Given that 922 divided by 3 gives 307.3, frequency ranges 1 and 2 (i.e. high and middle) contain 307 lemmas each, while range 3 (i.e. low) contains 308 lemmas. Each frequency range was in turn subdivided into three sub-ranges (i.e. upper: A, middle: B and lower: C). Once again, 307.3 divided by 3 gives a decimal figure (102.43). Hence, ranges 1A, 1B, 2A, 2B, 3A and 3B include 102 units, while ranges 1C, 2C contain 103 lemmas and 3C,104. The data analysed in this thesis comprise 1447 concordances corresponding to 60 shell-noun lemmas. Lemmas were chosen from the three main frequency ranges (1, 2, 3) and then from the upper and middle sub-ranges within each main range (i.e. A and B). The lower sub-ranges (i.e. C) were discarded due to the scant corpus evidence for 3C, where token frequencies range from 3 to 1 (i.e. hapax legomena) (cf., however, 3A and

[10] The calculation of such differences draws on the addition of the raw data from *AntConc* and *CQPweb*. The figure resulting from this addition is then used to calculate the percentage for each corpus tool. The subtraction of one percentage from the other gives the difference between both tools.

3B, with frequencies ranging from 14 to 7 hits). This thesis thus relies on the first 10 lemmas from ranges 1A, 1B, 2A, 2B, 3A and 3B, (60 in total). Table 4.4 below lists the 60 lemmas and their distribution in the frequency list (see Appendix 2 for the complete frequency list):

Table 4.4 Lemmas and frequency distribution

Range	Sub-range	Lemmas	Total
1	1A	time, thing, way, work, system, problem, area, point, part, word	10
	1B	example, application, project, detail, practice, evidence, experience, sense, answer, chance	10
	1C		0
2	2A	surprise, vision, crime, failure, objective, assessment, capacity, challenge, joke, leave	10
	2B	phenomenon, philosophy, dimension, myth, recommendation, suspicion, characteristic, finding, venture, warning	10
	2C		0
3	3A	motivation, opposite, prejudice, scandal, terror, triumph, anger, contradiction, correction, endeavour	10
	3B	endorsement, facet, foreboding, impetus, irony, misfortune, proviso, quest, recollection, testimony	10
	3C		0

Three lemmas within these 10-unit groups were discarded from the analysis due to the *Sampler*'s failure to account for any shell-noun uses. Such exclusions do not affect the number of sampled lemmas from each range. In these instances, the non-shell lemma was skipped, and sampling continued for the following units until a total of 10 was reached.

One such word is *number* (in range 1A), a noun occurring only in Flowerdew (2006)'s list of top 100 signalling nouns from an academic corpus. A search of the *BNC Sampler* returned instances with the meaning 'a sign that represents an exact amount or quantity' (Fox & Combley 2009, *Longman Dictionary of Contemporary English*, 2009; henceforth, *LDCE*), as in (238a), 'an amount of something that can be counted' (*LDCE*), as in (238b), and 'a number used to show the position of something in an ordered set or list', as in (238c) (*LDCE*). These instances were excluded because, whilst they are as unspecific as *fact* or *argument*, they do not refer to language (either directly, as with *sentence* and *paragraph* or indirectly, as with *claim* or *lie*), but to actual numbers or amounts, and are therefore closer to first-order than to third-order (more abstract) entities.

(238) (a) […] the total **number** of transferable votes cast, gives 0.39 (*BNC Sampler*: EW4, W:non_ac:polit_law_edu)
(b) Issuing bonds in the form of Eurobonds has a **number** of advantages (*BNC Sampler*: HY1, W:commerce)
(c) **Number** eighty-nine, a cluster ring again (*BNC Sampler*: G5A, S:unclassified)

The two other excluded lemmas are *vocabulary* and *wit*, both occurring within 3A. With regard to *vocabulary* (accounted for only in Francis 1986), the corpus returns highly general instances with meanings encompassing 'all the words that someone knows or uses' (*LDCE*), as in (239a) and 'all the words in a particular language' (*LDCE*), as in (239b). Although the word itself refers to language, in these particular instances *vocabulary* implies the general capacity or repository of words and, as such, no encapsulation of another discourse segment is intended. The third excluded item, *wit*, is only considered by Schmid (2000). Its shell use stems from its occurrence in *have the wit to do something* (i.e. *have the wit* [*necessary*] *to do something*). No such pattern occurs in the instances returned by the *Sampler*. Only cases with the general meaning of 'the ability to say things that are clever and amusing' (*LDCE*), as in (240a), and 'your ability to think quickly and make the right decisions' (*LDCE*), as in (240b), (with no connection to what this ability might be needed for) are accounted for in the corpus.

(239) (a) […] determining the precise sense of a word in an author's **vocabulary** (*BNC Sampler*: F98, W:ac:humanities_arts)
(b) […] the word first entered the Russian **vocabulary** (*BNC Sampler*: FB4, W:ac:humanities_arts)
(240) (a) […] facing the single-entendres that were Gav's best approximation of **wit**, and which inevitably followed any sexual adventure of mine […] (*BNC Sampler*: G0A, W:fict:prose)
(b) So what it means is, you've got to have **wits** about you when you go shopping (*BNC Sampler*: FUT, S:speech:unscripted)

Following sampling, the amount of evidence to be analysed for each of the 60 lemmas was decided. In view of the multifaceted concordance analysis proposed in this thesis (see 4.4.1), only 40 concordance lines were considered per lemma for the sake of manageability. This decision was inspired by Mahlberg's (2005) method, where a detailed analysis similarly leads to the selection of a limited set of concordance lines. While the amount of evidence for each lemma is larger in Mahlberg (2005) (100 vs. 40 concordances), considerably fewer units are included in her analysis (20 vs. 60). The latter is due to one of the minimal assumptions on which she bases her study of general nouns, i.e. that they are highly frequent nouns (e.g.

people, *time*, *fact*). This explains why her sampling includes only units from the top frequency ranges in the *BoE* and why, based on the amount of evidence per lemma, only 20 units are considered (see 2.2.2.1.1 for further details). Sinclair (1991: 29, 41–2) and Tognini-Bonelli (2001: 92–8) were also drawn on for the decision to split sampling into 20 concordances for the singular and 20 for the plural forms of each noun. The aim was to obtain, whenever possible, similar proportions of singular and plural word-forms in order to allow for any (patterning, syntactic, etc.) differences to arise.

As shown in Table 4.5 below, only 18 out of the 60 lemmas (i.e. 30%) feature equal proportions for both inflected forms. These coincide with nouns from the higher frequency ranges (1A and 2A). The remaining 42 lemmas (70%) contain uneven proportions, with 30% accounting for nouns for which fewer than 20 plural shell-noun forms occur in the corpus (i.e. not enough plural), 20% for lemmas which contain fewer than 20 shell instances for both singular and plural (i.e. not enough singular and plural), 15% for instances with only singular evidence and with fewer than 20 shell concordances for this word-form (i.e. only singular + not enough singular), 3.33% for units with fewer than 20 singular shell forms (i.e. not enough singular) and 1.67% for units with only singular evidence but with 20 concordances (i.e. only singular). Although initially the possibility was raised of completing the missing evidence with examples extracted from the *BNC*, the 50%-50% mode distribution obtained from the *Sampler* would be significantly affected by the 90%-10% distribution of the parent corpus. Therefore, for the sake of the balance and representativeness intended in this study, only the shell-noun evidence retrieved from the *BNC Sampler* was included in the analysis database, even if the resulting evidence is not as proportionate as regards number as was originally envisaged.

Table 4.5 Singular vs. plural proportions of the 60 lemmas selected. The number of lemmas per inflection is shown between brackets

	Lemmas	%
20 singular-20 plural	example, application, project, detail, practice, experience, answer, chance, time, thing, way, crime, objective, joke, problem, point, area, word (18)	30
Only singular	evidence (1)	1.67
Only singular + not enough singular	endorsement, impetus, irony, quest, testimony, terror, anger, correction, leave (9)	15
Not enough singular	characteristic, finding (2)	3.33
Not enough plural	sense, phenomenon, philosophy, dimension, myth, recommendation, suspicion, venture, warning, surprise, vision, failure, assessment, capacity, challenge, work, system, part (18)	30
Not enough singular and plural	facet, foreboding, misfortune, proviso, recollection, motivation, opposite, prejudice, scandal, triumph, contradiction, endeavour (12)	20

Following definition of the scope of the analysis, attention was turned to the choice of the concordancer for the purposes of this thesis. In this respect, two options were contemplated: *AntConc* and *CQPweb*. Although *AntConc* appears as the natural choice in view of its use for the extraction of the 922-unit frequency list, it does not offer a facility for random arrangement of concordances. The reason why this is important lies in the small sampling adopted for the analysis of lemmas. In order to compensate for the small size of the evidence (40 concordances), the concordancer had to offer a sample of hits representative of the overall occurrence of the lemma in the corpus.

AntConc arranges data in alphabetical order by text codes (A7V, A8J, A8W, etc., as in Figure 4.5). The only possible way of forcing a random order on data is by getting *AntConc* to sort the node word based on a range of words to the right or left of the queried item[11]. Whilst such a procedure returns what would appear as a fully random order, it is not such, as shown in the screenshot in Figure 4.6. The search in this Figure draws on the fifth word to the left and the fifth to the right of the node word *problem*. The screenshot illustrates how the occurrence of the words *Administration* and *Administrator* 5 words to the left in 7 concordance lines from text EAP somehow reduces the reliability of the intended random order. Hence, the contention is that, even if a higher range is applied (8, 9 words to the left or right), the mere recurrence of a given item in one text would skew, to a certain extent, the assumed representativeness of the returned hits.

CQPweb addresses such a limitation through the 'Show in random order' button (also present in *BNCweb*). As in *AntConc*, concordances in *CQPweb* are, by default, presented in 'corpus order', i.e. in alphabetical order by text codes (see Figure 4.7). The 'random order' button (located at the top of the concordances), however, ensures that '[…] whatever the number of examples […], it will always be a random selection of the full set' (Hoffmann et al. 2008: 166). This is shown in Figure 4.8, which presents a random arrangement of the concordances in Figure 4.7. It is important to note that whilst a range of texts are randomly selected, the same order is returned no matter how many times a given search is repeated (Hoffmann et al. 2008: 53). This being the case, *CQPweb* still offers a more varied distribution of corpus evidence than is apparent in *AntConc*. It was therefore decided to use *CQPweb* for retrieval and analysis of examples of shell nouns.

[11] Such a tool is also available in *CQPweb*.

Figure 4.5 A screenshot of a set of *AntConc* concordance lines for *problem* (non-random)

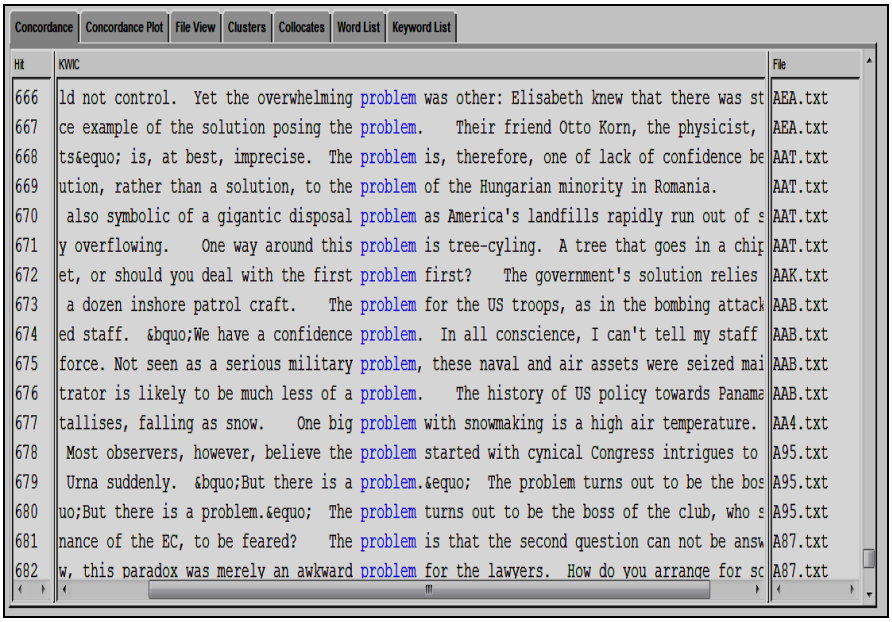


Figure 4.6 A screenshot of a set of *AntConc* concordance lines for *problem* (sorted 5 words to the left and 5 to the right)

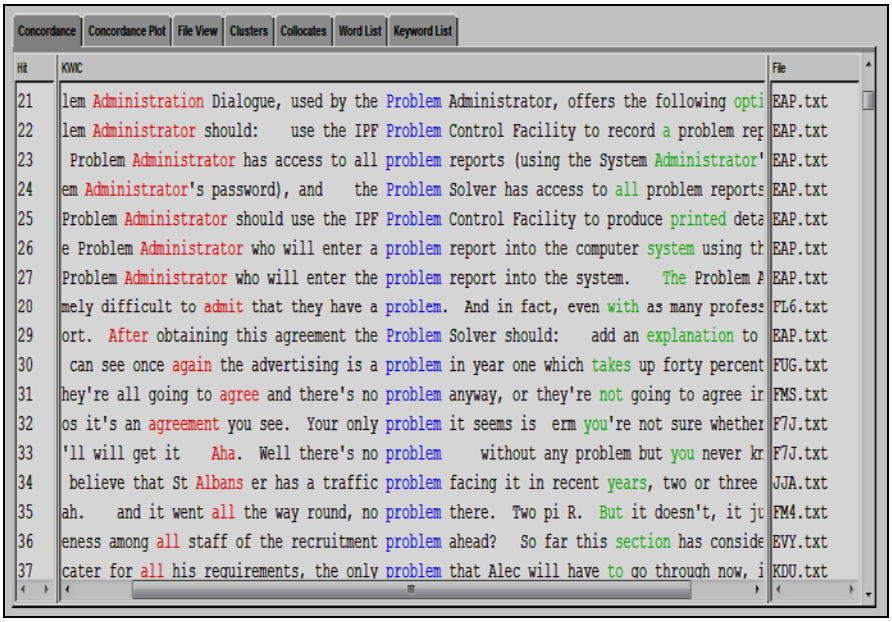
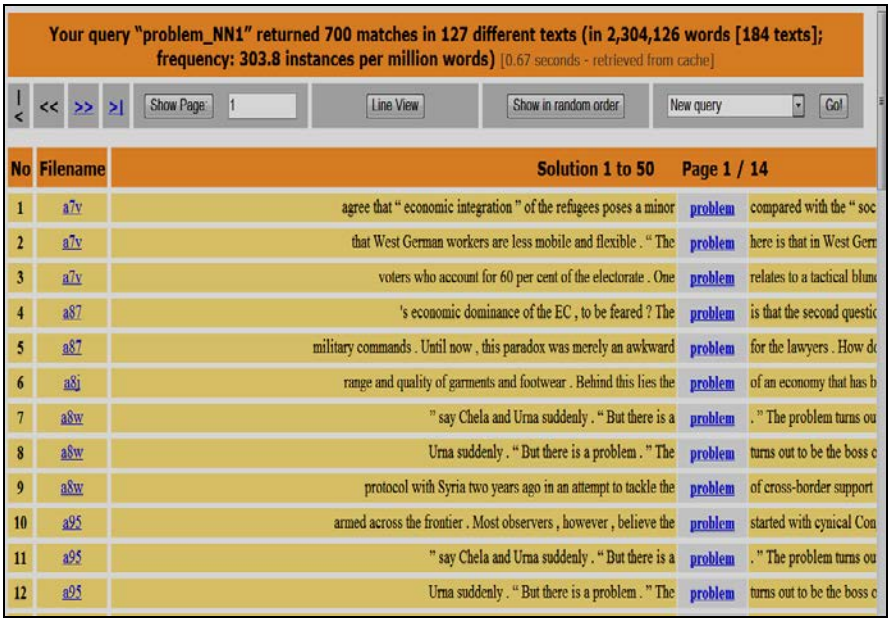


Figure 4.7 A screenshot of a set of *CQPweb* concordance lines for *problem* (corpus order)

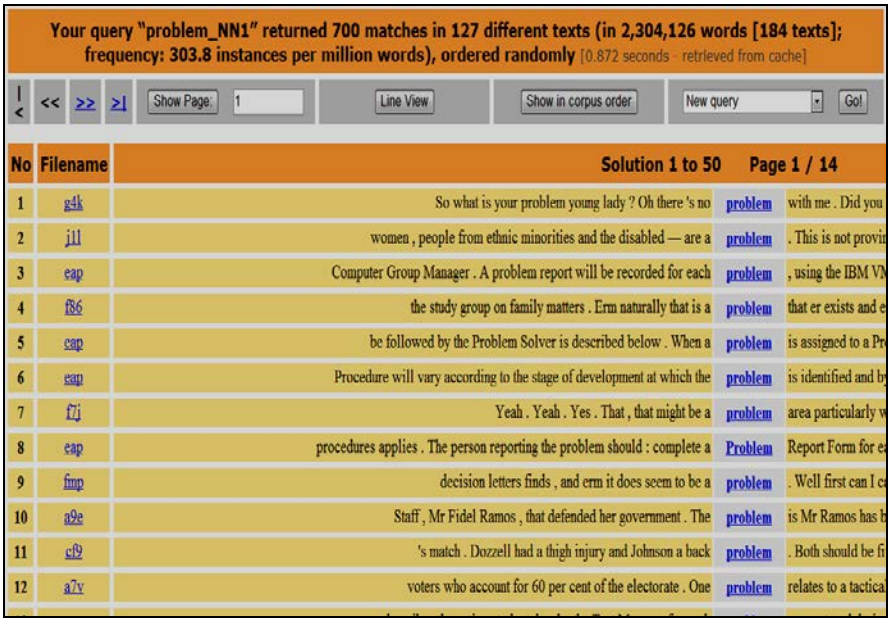


Your query "problem_NN1" returned 700 matches in 127 different texts (in 2,304,126 words [184 texts]; frequency: 303.8 instances per million words) [0.67 seconds - retrieved from cache]

|< << >> >| Show Page: 1 Line View Show in random order New query Go!

No	Filename	Solution 1 to 50 Page 1 / 14		
1	a7v	agree that " economic integration " of the refugees poses a minor	problem	compared with the " soc
2	a7v	that West German workers are less mobile and flexible . " The	problem	here is that in West Gerr
3	a7v	voters who account for 60 per cent of the electorate . One	problem	relates to a tactical blunc
4	a87	's economic dominance of the EC , to be feared ? The	problem	is that the second questic
5	a87	military commands . Until now , this paradox was merely an awkward	problem	for the lawyers . How dc
6	a8j	range and quality of garments and footwear . Behind this lies the	problem	of an economy that has b
7	a8w	" say Chela and Urna suddenly . " But there is a	problem	. " The problem turns ou
8	a8w	Urna suddenly . " But there is a problem . " The	problem	turns out to be the boss c
9	a8w	protocol with Syria two years ago in an attempt to tackle the	problem	of cross-border support
10	a95	armed across the frontier . Most observers , however , believe the	problem	started with cynical Con
11	a95	" say Chela and Urna suddenly . " But there is a	problem	. " The problem turns ou
12	a95	Urna suddenly . " But there is a problem . " The	problem	turns out to be the boss c

Figure 4.8 A screenshot of a set of *CQPweb* concordance lines for *problem* (random order)



Your query "problem_NN1" returned 700 matches in 127 different texts (in 2,304,126 words [184 texts]; frequency: 303.8 instances per million words), ordered randomly [0.872 seconds - retrieved from cache]

|< << >> >| Show Page: 1 Line View Show in corpus order New query Go!

No	Filename	Solution 1 to 50 Page 1 / 14		
1	g4k	So what is your problem young lady ? Oh there 's no	problem	with me . Did you
2	j11	women , people from ethnic minorities and the disabled — are a	problem	. This is not provir
3	eap	Computer Group Manager . A problem report will be recorded for each	problem	, using the IBM VM
4	f86	the study group on family matters . Erm naturally that is a	problem	that er exists and e
5	eap	be followed by the Problem Solver is described below . When a	problem	is assigned to a Pr
6	eap	Procedure will vary according to the stage of development at which the	problem	is identified and by
7	f7j	Yeah . Yeah . Yes . That , that might be a	problem	area particularly w
8	eap	procedures applies . The person reporting the problem should : complete a	Problem	Report Form for e
9	fmp	decision letters finds , and erm it does seem to be a	problem	. Well first can I c
10	a9e	Staff , Mr Fidel Ramos , that defended her government . The	problem	is Mr Ramos has b
11	cf9	's match . Dozzell had a thigh injury and Johnson a back	problem	. Both should be fi
12	a7v	voters who account for 60 per cent of the electorate . One	problem	relates to a tactica

Following the choice of the concordancer, it was decided that, for better balance, the analysis of the evidence for each word-form (i.e. singular vs.

plural) would include no more than 2 concordance lines per text. Thus, in the instance that a given text occurs more than twice among the first 20 random concordances for a particular word-form, once 2 of the occurrences are analysed, subsequent ones are skipped until a total of 20 shell-noun instances is reached. The reason for this decision lies in that, in many cases, one and the same shell noun is repeated more than once to refer to the same or to similar discourse segments. Should all repetitions be included as part of the set of 20 shell-noun singular or plural examples, the evidence for that particular word-form would be skewed towards a particular text and, more precisely, towards a restricted range of encapsulations in that text. Thus, in Figure 4.9, 5 out of the first 20 random concordances for singular *project* belong to text BMJ. The noun *project* in this text is primarily linked to the different stages of the project whose final aim was to build the Channel Tunnel between France and England. Inclusion of all these examples would boost text BMJ to 25% of the evidence for singular *project*.

Figure 4.9 A screenshot of the first 20 random concordances for singular *project*

8	bmj	the passage of the bill the various participants in the Channel Tunnel	project	continued their preparations . In August 1986 the construction
9	h47	O Nine Thousand ? I S O Nine Thousand and One through	project	quality plans . Those which fulfil the requirements of I S O
10	bmj	it seems unlikely that the international banking consortium will abandon the	project	given the large amounts of irretrievable money already com
11	gt9	in 1934 , but he continued to assist Waismann in this Schlick-directed	project	. Schlick 's murder in 1936 terminated Waismann 's employm
12	bmj	authorities was held soon after the January 1986 announcement of the Tunnel	project	in order to consider its likely impact on the region , and
13	h47	of project that you get . The prestige project , the complex	project	or the quick , slick project . In terms of your organization
14	h47	are which procedures is the quality plan . So for any particular	project	we want a quality plan which will tell us what the client
15	jnm	so that it could also conduct the management of the the Eurofighter	project	. erm One of our buyers , Spain , er was not
16	h47	The prestige project , the complex project or the quick , slick	project	. In terms of your organization , I think we 've got
17	bmj	Channel Tunnel Rail Link In the early stages of the Channel Tunnel	project	British Rail believed that the existing rail network in the South
18	h47	a function engineer . In addition to their other responsibilities , the	project	coordinator is responsible for a number activities . He is respo
19	bmj	Committee was set up in 1868 to consider a twin bored tunnel	project	developed by Thome de Gamond and the British engineers Jo

The small amount of corpus evidence for 21 out of the 60 lemmas under study (i.e. 35%) forced the inclusion, in these particular instances, of more than 2 concordance lines for a given text. This was done with a view to reaching, whenever possible, the total of 20 shell-noun concordances for each word-form. Table 4.6 below displays the extent (in %) of textual over-representation found among the 25 word-forms falling within the scope of the aforementioned 21 lemmas. The 6 over-representation categories comprise instances for which one text accounts for less than 50% of the 20 singular or plural word-form concordances (i.e. 1 text < 50% singular or plural), cases where the figure rises to 50% or more (i.e. 1 text 50% or >

50% singular or plural) and cases where two texts make up 50% or more of the concordances (i.e. 2 texts 50% or > 50% singular or plural). The two most frequent categories are singular nouns with less, as well as those with more than 50% of the evidence from one text (i.e. 28% and 20% respectively).

Table 4.6 Over-represented word-forms. The number of lemmas per inflection is shown between brackets

	Word-forms	%
1 text < 50% (singular)	phenomenon, philosophy, dimension, warning, assessment, capacity, opposite (7)	28
1 text < 50% (plural)	characteristics, findings, crimes (3)	12
1 text 50 % or > 50% (singular)	foreboding, recollection, motivation, anger, leave (5)	20
1 text 50 % or > 50% (plural)	phenomena, dimensions, visions, jokes (4)	16
2 texts 50% or > 50% (singular)	myth, recommendation (2)	8
2 texts 50% or > 50% (plural)	myths, recommendations, challenges, systems (4)	16

4.4 DATA ANALYSIS

4.4.1 *A corpus theoretical approach to shell-noun description*

Not all concordances returned by the corpus represent shell-noun uses. For every single node word, it was therefore necessary to ascertain whether inclusion in the group of shell-noun uses would make sense in the light of the evidence obtained. This was accomplished through manual and context-sensitive analysis of every concordance line. This kind of intensive textual analysis requires moving from the limited scope of the concordance line to the wider context (Mahlberg 2005: 58).

A click on any node word in *CQPweb* allows access to the extended context for any concordance. While lexical realisation for a shell noun often appeared in the surrounding text, the initial context given by the corpus in many instances had to be extended to large non-contiguous stretches of discourse in order to identify the encapsulated discourse segment. The close reading of individual corpus texts needed for this kind of analysis is, according to Botley (2006: 102), in apparent contradiction with typical concordance-based corpus methods. Example (241) illustrates the amount of textual evidence that had to be considered for the interpretation of certain shell-noun instances. In the example, *his vision* does not refer to adjacent information. An understanding of what *his vision* might involve can only be gained from the underlined segments, which are (as evident from the numbered lines) far from the node word itself.

(241) 106 After all Koresh's apocalyptic vision, his talk of the seven seals that only the lamb of God can unlock, sounded like refinements of what they already believed. [...] david koresh:
174 Someone is going to rule whether the big world likes it or not.
175 Thou shalt break them with the rod of iron, thou shalt dash them in pieces like a potter's vessel. [...] Well David said that we were at the end, everyone believed that they were that the time has come that we was at the end of the world. [...]
388 In order to make his own sort of erm prophetic vision come true.
389 Erm he decided to stage a fire in which it would make it appear as though er this was a result of some sort of erm armed confrontation between law enforcement and his group. (*BNC Sampler*: HE3, S:brdcast: documentary)

The detailed analysis of corpus data in the present study is linked to a corpus theoretical framework. Drawing on Mahlberg (2005: 31–8), a 'corpus theoretical approach' takes issue with a focus on predefined patterns and corpus queries (e.g. N-cl and N-*be*-cl) in favour of a more corpus-driven treatment of data. In contrast to the extreme version of corpus-driven research (e.g. Tognini-Bonelli 2001: 67–71; Sinclair 2004: 23), data analysis is conducted with certain theoretical preconceptions in mind. Thus, as noted by McEnery & Hardie (2012: 161), '[…] it is arguably impossible to approach corpus evidence with no preconceptions about language'. Such preconceptions are often shaped by theoretical notions and categories felt to be useful for the purposes of a particular corpus-driven linguistic analysis. In addition to the reliance on relevant theories, a corpus-theoretical approach also starts with a range of '[…] minimal assumptions […]' (Mahlberg 2005: 35–8) about the linguistic phenomenon under study. Three minimal assumptions guiding Mahlberg's (2005: 37–8) analysis of general nouns are that they are highly frequent, that they perform local textual functions and that they are nouns.

The corpus theoretical perspective used in this thesis rests on an eclectic theoretical background (cf. Schmid 2000: 20 and Mahlberg 2005: 33, where an eclectic approach is advocated too). Such eclecticism is motivated by the multifaceted formal, textual and semantico-pragmatic analysis proposed. The analysis of shell-noun instances revolves around 9 variables adopted mainly from contemporary descriptive grammars of English (Quirk et al. 1985) and from SFG (Halliday & Matthiessen 2004). The following outlines the 9 variables under scrutiny, deferring a more detailed discussion of each variable until chapter 5:

i) Genre of the text: The genre categories used in this study (e.g. W:ac, i.e. written academic; S: meeting, i.e. spoken meeting) are those taken from *BNCweb* (see 4.2.2).
ii) Experiential structure: This variable concerns the semantic structure of the noun phrase, as explained by SFG (Halliday & Matthiessen 2004). It comprises Deictic, post-Deictic, Epithet, Classifier and Qualifier. Hallidayan grammar offers a more systematic classification of semantic premodifiers than that found in shell-noun research (see 3.2.3 and Table 3.1).
iii) Formal structure: The labels used for this variable (e.g. definite article, relative clause, noun complement clause, etc.) are adopted from Quirk et al'.s (1985) grammar. Hallidayan grammar was not used for this variable on account of the greater amount of descriptive detail found in Quirk et al. (1985).
iv) Syntactic function: The framework followed here (e.g. direct object, subject, adverbial adjunct, object complement, etc.) is Quirk et al'.s (1985). Once again, the more detailed description given in Quirk et al. (1985) stands in contrast with Halliday & Matthiessen's (2004) distinction between only four types of syntactic function (subject, predicator, complement and adjunct). In SFG, the function of complement subsumes Quirk et al'.s (1985) direct object, indirect object, subject complement and object complement.
v) Participant type: The analysis of the semantic roles played by nouns follows Halliday & Matthiessen's (2004) system of Transitivity (e.g. Actor, Goal, Identifier, Carrier, Circumstance, etc.).
vi) Theme vs. Rheme: This variable draws on Halliday & Matthiessen's (2004) system of Theme.
vii) Direction of encapsulation: This variable comprises the two main types of endophoric reference (anaphora and cataphora), their realisation (intersentential, intrasentential), as well as exophoric reference. The analysis of this variable draws on the literature on encapsulation reported in 3.2.2.
viii) Antecedent: Following Stirling & Huddleston (2002: 1455), the term 'antecedent' is used here to refer to discourse encapsulated both anaphorically and cataphorically (see 3.2.2). The labels used in this variable are adopted from Gray (2010), where a distinction is made between 'Global Extended Discourse' (where the antecedent is either difficult to delimit or crosses sentence boundaries) and 'Local Discourse', comprising Noun Phrase (Simple: no postmodification and Complex: with postmodification) and Sentence/Clause (see 2.3.1.2.1).

ix) Semantic type: This variable concerns the meaning carried by a shell noun in context. The semantic classification followed here is the one by Schmid (2000): factual, mental, linguistic, circumstantial, modal and eventive shell nouns (see 3.2.5).

The corpus theoretical rationale of this study is thus inspired by SFG for the semantic and textual variables (i.e. experiential structure, participant type, and Theme vs. Rheme) and by Quirk et al'.s (1985) grammar for the more formal or structural ones (i.e. structural pattern and syntactic function). The remaining four variables, i.e. genre, direction of encapsulation, antecedent and semantic type, are more heterogeneous in their theoretical underpinnings. In view of such mixed categories, the argument might be advanced that the present study does not use one reference for its theoretical foundations. Had a single theoretical background been used, the scope of the variables would have been restricted to either more formal (i.e. Quirk et al. 1985) or more semantico-pragmatic and textual aspects (i.e. SFG). By drawing on an eclectic approach, a range of variables is explored, and each variable is encoded with higher descriptive detail.

To conclude this section, it must be pointed out that, apart from the theoretical framework underlying the variables, the corpus theoretical approach is also evident in two minimal assumptions, both of which are implied at the end of the previous chapter. One such assumption is that shell uses are only apparent in semantically unspecific second- and third-order entities (e.g. *action*, *event*, *idea*, *point*). Nouns like *love*, *arrival* or *democracy*, while abstract in nature, do not qualify as shell nouns on account of their being more semantically bounded and so less dependent on lexical realisation than typical shell-noun instances (see Schmid 1999: 223 in 3.2.1). Hence, while it may be asked what an *action* or an *idea* is or involves in a particular context, it is not common to ask the same question about nouns like *love*, *arrival* or *democracy*. These are general concepts for which few or no semantic gaps exist: most people would know what *love*, an *arrival* or *democracy* are (despite more positive or negative personal associations with the concepts), but greater difficulty would stem from knowing exactly what an *action*, *idea* or *point* refer to in the absence of context.

The second assumption is, in fact, related to context. Following Ivanič (1991: 111), it is assumed that, regardless of the determiner accompanying shell nouns (specific, non-specific or zero), the surrounding context will, on most occasions, influence their interpretation. As discourse unfolds, readers draw on the mental model created by the text so far, as well as on their extra-textual models of world knowledge in order to make sense of any discourse entity (see Brown & Yule 1983: 201 and Garnham & Oakhill 1990:

380 in 3.2.3). If close reading of corpus texts reveals no contextual information to spell out the meaning of a shell noun, the assumption is that the reader is being instructed to rely on extra-textual knowledge and, as such, the example is treated as exophoric. Examples (242) and (243) are two cases in point. In (242), the noun phrase *an unambiguous policy recommendation* has generic reference, as it implies *any potential recommendation on policy that might be considered unambiguous* (i.e. *one such recommendation*). As no such recommendation is spelled out in either the preceding or subsequent discourse, *recommendation* is here treated as exophoric. Example (243), whilst also containing an indefinite article, is different, as an understanding of what *a clear vision* could involve is entailed in the text through the underlined discourse segment (i.e. *the BBC's vision would be to introduce a wide range of high quality programmes, greater efficiency,* etc.). In this case, the shell-noun phrase is endophoric.

(242) Clearly this might not be a task that can be undertaken with any great degree of precision. Yet if we are to pronounce on the desirability or otherwise of individual monopoly situations, or proposed mergers, this is clearly what is required to provide an unambiguous policy recommendation. (*BNC Sampler*: HXN, W:commerce)

(243) 'The BBC must therefore have a clear vision if it is to retain its role as the cornerstone of British broadcasting and continue to command respect and admiration in Britain and throughout the world.' In January John Birt laid out that vision of a wide range of high quality programmes, greater efficiency and accountability with value for money for licence payers. 'We, the ten members of the board of management, believe that John Birt is the best person to lead the BBC and he has our unanimous support (*BNC Sampler*: CF6, W:newsp:other:report)

The analysis disregards the often-repeated assumptions that shell nouns are most frequent in N-cl and N-*be*-cl patterns, that only specific determiners (especially definite and demonstrative) correlate with their use or that only long stretches of discourse may be referred to. The analysis of corpus data in this thesis is expected to either corroborate or refute these two minimal assumptions.

4.4.2 *Non-shell instances: types of excluded items*

This section presents a range of non-shell uses found in the study sample. It precedes the detailed discussion of the 9 variables in chapter 5 because an understanding of what shell nouns are needs to rest also on what they are not. This section and chapter 5 (both qualitative in nature) should pave the way for the quantitative analysis of the 1447 shell-noun instances

retrieved from the *Sampler* in chapter 6. The categories of excluded items are listed in the following 12 subsections.

4.4.2.1 *Premodifying nouns*

The assumed shell noun acts as premodifier of another head noun. While the head noun is the nucleus of the referential expression, the nominal premodifier merely restricts the meaning of the following noun. This is shown in (244) and (245) below, where the query lemma appears in boldface. In the instance that a potential shell-noun phrase occurs as a complement of a preposition, a shell-noun analysis will usually apply, as in (246). Whilst not performing a clause-level function, the noun is often found to encapsulate preceding or subsequent discourse segments. Nominal premodifiers are therefore similar to adjectives in being non-referential, while nominal prepositional complements are still referring expressions, and are thereby treated as shell nouns.

(244) […] take some important initiatives in terms of helping the pensioners in the city, **crime** prevention, and other issues, so we think we've done a reasonable job. (*BNC Sampler*: HDT, S:speech:scripted)

(245) Democrats in Congress led the condemnation of the **surprise** arrival of two senior Bush Administration aides in Beijing. (*BNC Sampler*: A9M, W:newsp:brdsht_nat:report)

(246) 'But, Prentice, it's not as though you even believe in Christianity or anything like that. Shit, I can't work out what it is you do believe in … God?' I shifted uncomfortably in the thin seat. 'I don't know; not God, not as such, not as a man, something in human form, or even in an actual thing, just … just a field … a force — '[…] I keep getting this feeling it's already there, like in quantum physics, where matter is mostly space, and space, even the vacuum, seethes with creation and annihilation all the time, and nothing is absolute, and two particles at opposite ends of the universe react together as soon as one's interfered with; all that stuff. It's like it's there and it's staring us in the face but I just can't … can't access it'. […] Are you tryin to get back at your dad because of **this stupit** [sic] **religious thing**? (*BNC Sampler*: G0A, W:fict:prose)

4.4.2.2 *Repetitions*

The shell-noun phrase is repeated in the same sentence, often due to hesitation or rephrasing in spoken discourse. In these cases, only the first occurrence of the noun is kept in the analysis (see (247)). Repetition also features in instances where two identical concordance lines are returned by the corpus. This is shown with certain concordances from texts A95 and

A9E, two newspaper files which, for some reason, share the same texts. The analysis in these cases includes only one of the two repeated lines (see (248)).

(247) And that 's gonna create opportunities but also er the possibilities of joint ventures be it on a UK, Europe or the world basis er joint ventures will I mean a lot more cooperation between er th the large players with regards to product development what have you. (*BNC Sampler*: FUG, S:unclassified)

(248) Mr Ray's last warning was both unheeded and unnecessary. (*BNC Sampler*: A95, A9E, W:newsp:brdsht_nat:report)

Examples of partial repetition are also found in the corpus. Unlike complete repetition, as in (247) and (248), partially repeated examples are included in the shell-noun count. In these instances, the noun phrase in the concordance is a shorter or more distilled[12] version of a preceding more complex and more informative noun phrase or of a range of non-adjacent discourse segments. The simpler noun phrase is treated as encapsulating the preceding information, provided that the more complex noun phrase or discourse segment allows understanding of the unspecific shell unit.

This is exemplified in (249) and (250) below. In (249), whilst the syntactico-semantic function of the noun phrase is unclear, its contextualised meaning seems to draw on the preceding underlined segments. From the immediately preceding stretch, it may be inferred that the application or request is to obtain planning probably to build a road through an area. By referring further back in the text, it will be discovered that the planning application dates back to 1964 and that it applies to a specific site and area. In this example, therefore, the context-dependent meaning of the shell-noun phrase hinges on information scattered throughout the previous discourse. Such information is gradually condensed until *the nineteen sixty four application* is reached, a noun phrase which, at this point in discourse, is meaningful to the hearer. In example (250), information as to what *an application* is for may be obtained from the underlined partially repeated noun phrase. The request or application in this case is to get training, the nature of which is not clear from the surrounding co-text. This being the case, however, at least it is known what the purpose

[12] Halliday & Martin (1993: 30) use the concept of 'distillation' for the process whereby the semantic opaqueness of many so-called technical terms (e.g. *coelomates*, *arthropods*) stems from a gradual condensation of more congruent or transparent paraphrases (e.g. *an invertebrate animal with an external skeleton*, for *arthropod*). A fully distilled form is the end result of the process of grammatical metaphor (see 2.2.1.3.1 and 4.4.2.9).

of the application is, and the example could be said to partially encapsulate or refer back to the underlined complex noun phrase.

(249) We know from the appearances that you were present at that enquiry <-|-> <unclear> <-|->.
<-|-> I did.
If we then look at <pause> paragraph five, that's the first page of appendix two, <pause> <cough> <pause> we see here <voice quality: reading>The appeal site has a long planning history. It lies within an area designated as <unclear> planning permission since nineteen sixty four<end of voice quality>.
Yes.
I don't know whether it would be helpful, sir, if Mrs <gap desc='name' reason='anonymization'> could indicate on the plan recently handed in to you
Yes.
where that nineteen sixty four planning permission <pause>
Right.
lies. […]
Perha-- perhaps if I <unclear> the appeal site <unclear> wedge with the erm <pause> <unclear> intending to build a road through it.
The nineteen sixty four application <unclear>.
Er let me just say that aloud <-|-> <unclear> <-|-> (*BNC Sampler*: FMP, S:pub_debate)

(250) Okay erm yeah a-- again related to special needs is the number of training applications
Yeah yeah.
erm <-|-> which I think we can assume <-|->
<-|-> <unclear> <-|->
is reasonably standard.
<-|-> Well <-|-> I mean the arguments I've had is that they're not.
What that in some areas somebody wouldn't get an application whereas that they would in other areas? (*BNC Sampler*: H5D, S:meeting)

In other cases, similar instances of gradual condensation of discourse information may not be taken as encapsulating the longer noun phrase, as its modifiers do not help to spell out with any degree of certainty what the simpler shell noun implies. The complex noun phrase only provides specifics of identity (see Winter 1992: 154–5 in 2.2.2.1.2), insofar as its modifiers simply restrict the referential scope of the head noun, without specifying what the noun entails. More informative details or specifics by clause (i.e. what the shell noun is encapsulating) have to be searched for elsewhere. This is apparent in example (251) below: *such projects* refers back to the complex noun phrase *ambitious research projects undertaken*

for methodical demonstration purposes. The simple phrase, however, could not be said to encapsulate the complex one, in that the modifying *-ed* participle clause merely specifies the types of projects discussed here (i.e. only the projects undertaken for demonstration purposes). More precise information on what such projects involve is provided in the following stretch of discourse.

(251) At the other end of the scale there are ambitious research projects undertaken for methodical demonstration purposes. […] Two examples of such projects are discussed in a recent paper by Manfred Thaller. In France Gian Piero Zarri has developed a system known as RESEDA to allow Artificial Intelligence techniques to be applied to historical source material (concerning the Hundred Years War). […] Again, in 1981–7 the Volkswagenstiftung funded on a large scale the ARCOS project, a system which photographs a three-dimensional archaeological object, provides a drawing and converts it into data which lend themselves easily to methods like Cluster Analysis. (*BNC Sampler*: F98, W:ac:humanities_arts)

4.4.2.3 *Incorrect POS annotation*

The node word is incorrectly POS annotated. In 39 concordance lines in total, the node word is erroneously tagged as a noun (i.e. NN1 or NN2), when in fact it is either a verb, as in (252) or an adjective, as in (253).

(252) Does that answer your question? (*BNC Sampler*: FM7, S:unclassified)
(253) […] he derived from him a characteristic liking for low tones […] (*BNC Sampler*: CN4, W:pop_lore)

4.4.2.4 *Unintended items*

As a result of hesitations and disfluencies in spoken discourse, the corpus may return assumed shell-noun instances which, when explored in context, reveal themselves as unintended lexical items. This is evident as soon as the speaker (through the use of repair strategies) utters the intended lexical item. Examples (254) and (255) illustrate this. In (254), it is not the noun *works*, but *workers* that the speaker is trying to use. Similarly, in (255), it is not the noun *part* but *party* that is actually intended in this discourse situation.

(254) Some works at the <-|-> station <-|->
<-|-> Some workers.
some workers at the station took Scott inside. (*BNC Sampler*: KB3, S:conv)

(255) […] the beginning of individual voting of trade union members within the part--Party. (*BNC Sampler*: HDT, S:speech:scripted)

4.4.2.5 *Non-referential naming expressions*

The assumed shell-noun is part of a non-referential naming expression. The excluded instances range from nouns occurring in book and newspaper titles (e.g. *How I see **Philosophy**?*, *1968*; *the Irish **Times***), through names of competitions and exhibitions (e.g. *The Forbra Gold **Challenge** Cup*, *Reverie **Myth** and Sensuality*) to brand and flower names (e.g. *a **Triumph** motorbike*, *'Snowball' and red '**Triumph**' also look good*).

4.4.2.6 *First-order entities: closed sets and value ranges*

The assumed shell noun is a first-order concrete entity which is either part of a closed set of items or whose meaning comprises a range of values or quantities. The former applies to *sense*, when referring to any of the human powers of sight, hearing, taste, smell and touch. The latter is observed with *dimension*, when meaning the measurement of something (e.g. *make the length and breadth **dimension** about 23 cm*), *capacity*, when denoting an amount of space (e.g. *the storage **capacity** is about 230 megabytes*) and *point*, when implying a score (e.g. *The nine gleaned from this encounter puts Essex ahead of the field on 30 **points***). *Time* would also fall within the scope of the latter group when the meaning is that of 'the thing that is measured in minutes, hours, days, years etc using clocks' (*LDCE*) (e.g. *a personal best **time** of 1.24.41*) and 'a particular point in time shown on a clock in hours and minutes' (*LDCE*) (e.g. *Do we have any idea what **time** we'll be finished*?). Only when *time* carries the meaning of 'an occasion when something happens or someone does something' (*LDCE*) may a shell use be assigned to the noun. This is in line with Schmid's (2000: 276) claim that '[…] circumstantial shell nouns do not shell "circumstances" but events'.

Hence, for *time* to be considered a shell noun, it needs to encapsulate neither a quantifiable amount of time (e.g. *four*, *five minutes*) nor a temporal point (e.g. *10 a.m*), but an event reported within the universe of discourse. Example (256) is a case in point. In this instance, *at that time* is not linked to a specific time, but to the general situation or event described in the underlined segment. Such a time is therefore the occasion when the arrest and the questioning took place. A similar example is (257). Drawing on Botley (2006: 97), in this particular instance, *this time* does not refer to a clearly delimited discourse segment, but to '[...] the situation described previously', i.e. *the time when Prentice screeches after learning about someone else's pregnancy.* Given that the understanding of *time* in both

(256) and (257) is not exophoric, but relies on the general situation described in the preceding co-text, their use is treated as '[...] a class of situation reference' (Botley 2006: 97; see also Fraurud 1992 in 3.2.2).

(256) No <pause> but on this occasion we didn't have a spare holster to put the weapon into. […]
I then went forward to the man on the floor er, handcuffed him and stepped back. […]
Did you tell him you were arresting him?
Yes I did, yes.
When did you arrest him?
In the meantime, as I recall, er <pause> I <pause> I asked him who he was and I asked his wife who he was because he had asked what we were doing in his house <pause> and when he said his house, I thought then for the first time that perhaps this wasn't <gap desc='name' reason='anonymization'> who er <pause> we had on the floor. But <pause> I have to say that at that time I still didn't know, I had a good idea, <pause> that it wasn't the man we were looking for but therefore he was still arrested and he was arrested on sus-- well for harbouring an escapee.
I--i--if you can remember, can you tell us the words you used throughout the <unclear> or the gist of the words? (*BNC Sampler*: JJW, S:courtroom)

(257) Married? No! 'But isn't this …' My voice had risen a good half-octave and my hands were waggling around on the end of my arms as though I was trying to shake off bits of Sellotape. '… rather soon?' I finished, lamely. 'Well, yes,' mum said, sipping her cappuccino. 'It is'. She smiled brightly. 'I mean, not that she's pregnant or anything, but' 'Pregnant! I screeched. The very idea! The thought of the two of them fucking was bad enough; Lewis impregnating that gorgeous creature was infinitely worse. 'Prentice!' Mother whispered urgently, leaning closer and glancing round again. This time we were getting a few funny looks from other customers. (*BNC Sampler*: G0A, W:fict:prose)

The meaning of *point* in (258) is similar to that of *time* in the former two examples, i.e. 'an exact moment, time, or stage in the development of something' (*LDCE*). In this case, it is possible to eliminate *the point* with no change in meaning (i.e. *until they were almost ripe*). This indicates that the noun here performs a mere support function within this discourse context (cf. Mahlberg 2003 in 2.2.2.1.1), inasmuch as it acts as an anchor enabling the quicker processing of the postmodifying clause through a single noun-based discourse entity.

(258) Once again this autumn, I lost the race with the squirrels to harvest the hazel nuts. I checked daily, in September, until **the point** where they were almost ripe. (*BNC Sampler*: C9C, W:pop_lore)

4.4.2.7 *Visible and tangible first-order entities*

The assumed shell noun is a visible and more or less tangible first-order entity. This is apparent with *thing*, when its meaning is that of 'object' (e.g. *fixings, fasteners, that type of* ***thing***), *work*, when used to refer to artistic output (e.g. *a* ***work*** *of art*) and *point*, when implying a particular place (e.g. *please use the regular crossing* ***points*** *of our capital*). In these cases, only certain senses qualify as shell nouns, thereby forcing the analysis to leave first-order concrete uses out of the count. An explanation of the rationale behind this decision will now be given for the nouns *area, application* and *part*, whose classification proved difficult (see 5.3.5 for the shell-noun uses of *thing*, *work* and *point*).

With regard to *area*, drawing on Schmid (2000: 279–82), example (259) below represents a shell-noun use, while (260) is a first-order non-shell use. Based on the argument that circumstantial shell nouns are found in combination with events (see 4.4.2.6 above), instances like (259) are treated as shell uses. Therefore, *the area where you play football* is *an area where something happens* and, as such, it is shell in nature. Schmid (2000: 280–1) claims that the N-*where* pattern is closely linked to shell uses among circumstantial place nouns (e.g. *place*, *area*, *region*), insofar as the combination of a place noun plus a following *where* clause results in a similar kind of experiential identity to that found among typical N-cl shell instances. For example, just as *the fact that she is nice* is equivalent to *the fact is that she is nice*, so is *the area where you play football* similar to *this area is where you play football* (despite the different determiner).

(259) Why don't you use the green one?
Well I like the ride on one because it's easier.
The, the <pause> green one it takes hours to cut <pause> used to take about three hours for me to cut all the grass with that one.
But first, but when we first came here I spent three hours on a Sunday morning, and I only did **that area where you play football**. […] <-|-> If <-|-> you'd started the motor and used it it would have helped. (*BNC Sampler*: KCH, S:conv)

(260) The main colour in your scheme should reflect this mood. Restrict this colour to **large areas** such as walls, carpet or curtaining and build up from here, introducing one or two further colours or tones of one shade, and adding interest with patterns and texture. (*BNC Sampler*: GUB, W:misc)

Schmid (2000: 281) further notes that the *th*-N pattern is also frequent with place nouns, and lists adjectives typically collocating with *this* and *that place*, *area* and *region* in the newspaper section of the *BoE* (Schmid 2000: 282). Among the adjectives collocating with *area*, the list includes instances like *geographical* or *tumbledown*, which would be associated with first-order *area*, as well as *vital* or *crucial*, which are more in line with an abstract use of the noun. No concordance lines are offered in support of the shell or non-shell analysis of *th*-N cases. Without any extended textual evidence, it is difficult to ascertain the extent to which a shell-noun analysis might apply to this pattern. In this respect, Schmid (2000: 282) hints that certain uses of place nouns in *th*-N are non-shell on account of their being deictic rather than anaphoric markers. This is evident, for example, with *this place*, when reference is made to a house or flat in the real world.

With this in mind, for the sake of consistency, it was decided to automatically discard any instances of *area* where the noun is observed to denote 'a particular part of a country, town etc'. (*LDCE*) or 'a part of a house, office, garden etc that is used for a particular purpose' (*LDCE*). Following Paradis (2004: 59), the contention here is that physical locations are first-order concrete-entities. In example (259), it is believed that the *where* clause does not act as the lexical realisation of *area*. Its function is merely to provide specifics of identity restricting the reference of a first-order place in the extralinguistic world (i.e. *what area?*, *not any area/place*, *but the area/place where/in which the action of playing football takes place*)[13]. Similarly, in (260), *large areas* is not shell in nature, as it relates to tangible parts of a house, i.e. *walls*, *carpet* and *curtaining*. Underlying these examples is not a shell-noun use, but one which is more akin to that of general nouns (cf. Halliday & Hasan 1976; Mahlberg 2003 and 2005 in 2.2.2.1.1).

Given that, in (259) and (260), *area* does not entail or refer to second- or third-order information (i.e. it is not metadiscursive), but to real-world entities, its role is simply to act as a superordinate term for a more concrete first-order entity, this being either within the text itself, as in (260) (*walls*, *carpet>house areas*) or deictically implied, as in (259). It is such superordinate reference to other concrete entities that makes these uses more general rather than shell. On these grounds, examples of *area* were included in the analysis provided that their meaning is that of 'a particular

[13] *Point* in (258) also occurs in the N-*where* pattern. In this case, however, the noun is metadiscursive and shell in nature, as no reference is made to an extralinguistic first-order entity, but to a particular stage in the harvest of hazelnuts (i.e. the stage where they are almost ripe). Considering that the noun here introduces a state and that no link is established with a single concrete entity in the outside world, *point* is treated as a second-order entity shell noun.

subject, range of activities, or group of related subjects' (*LDCE*). This is shown in (261) and (262), where *area* refers to a type of planning activity (i.e. *greenbelt protection*) and to a subject of debate (i.e. *the hard data like annual mileage*), respectively. Thus, the shell use of these two examples lies in the eventive (i.e. *area of planning activity)* and linguistic domains (i.e. *area of debate*) that the use of *area* restricts in each case (for more information on the semantic categorisation of these uses, see 5.3.5).

(261) There is however one er <pause> overriding point here Mr <gap desc='name' reason='anonymization'>.
That is, you recognize, do you not, that greenbelt protection includes the concept of the initial negative presumption that is not contained in any other area of planning <unclear>? (*BNC Sampler*: FMP, S:pub_debate)

(262) I mean even the simple things, the hard data like annual mileage, is clearly an area of debate because you and Cynthia would argue different points you know. (*BNC Sampler*: H5D, S:meeting)

Another noun that deserves special mention is *application*. One of the senses of this noun is that of 'the practical purpose for which a machine, idea etc can be used, or a situation when this is used' (*LDCE*). This sense is in (263), where a particular technology is said to have a range of practical purposes such as airline reservations or warehousing.

(263) The sorts of application where we believe this technology is appropriate might be in such things as credit card authorization, inventory management, airline reservations, warehousing, customer service and command and control. (*BNC Sampler*: HDF, S:speech:unscripted)

Another sense is that of 'a piece of computer software which does a particular job' (*LDCE*). The last part of the definition seems to relate to the aforementioned 'practical purposes' meaning. A search in the on-line *Oxford English Dictionary* (henceforth, *OED*) gives the following definition of the computing sense: 'A function performed by a computer to meet a specific user requirement; (now usually) a program or piece of software designed to perform such a function [...]'. From this definition, it appears that the 'purpose' or 'function' meaning preceded that of the 'software-related' one. Therefore, despite the more concrete 'software' meaning, the purpose that an application was designed to serve is still very much at the semantic core of the noun itself. In (264), for example, it might be claimed that the purpose of one such store-based application is to enable all shops to have an *RS six thousand* installed. Similarly, in (265), whilst *Wordperfect InForms* is a piece of computer software, its practical use is evident in the following

prepositional phrase: to create and distribute fill-in forms. As the computing first-order sense of *application* stems from a more abstract one, examples like (264) and (265) were initially left in the shell-noun count.

(264) One of <gap desc='name' reason='anonymization'> partners is <gap desc='name' reason='anonymization'> Stores who are implementing a strategic store-based application which involves having an R S six thousand in every single shop with a SQL Server on every single box. With three hundred and sixty SQL Servers, there's no way they're going to have a DBA in every store. (*BNC Sampler*: HDG, S:speech:unscripted)

(265) However, not everything is being bundled into Office, the company's other great hope is Wordperfect InForms, an application for creating, and electronically distributing fill-in forms. (*BNC Sampler*: CL8, W:non_ac:tech_engin)

Additional corpus evidence later revealed that the computing sense of *application* tends to favour a material environment, combining either with material processes (e.g. *make*, *use*, *develop*, *build*, *write*, *send*, *ship*, *work with*, *inter-operate with an application*; *an application runs*) or referring to other first-order entities (e.g. *an object developed in C*, *a product*). This kind of environment points to a more material and, thus, first-order interpretation of the computing sense of *application.* Whilst application software is not as tangible as hardware, an application is a sort of virtual device that allows the user to perform a certain function. As the function is subsidiary to the physical programme itself, a first-order analysis would apply, leading to the exclusion of computing instances of *application* from the count. Cases of *application* meaning 'practical purpose or use', as in (263) were, by contrast, included in the analysis (for further details on the semantic categorisation of *application*, see 5.3.5).

As regards *part*, following Schmid (2000: 118–9), a decision was made to include only those instances where the *of*-phrase in the N-*of* pattern (i.e. *part of*) contains a second- or third-order entity and where the whole noun phrase refers to second- or third-order abstract information. In the instance that either the N-*of* or *th*-N (e.g. *the*/*this*/*that part*) patterns of *part* relate to sections of first-order real world entities, then a shell-noun analysis will not apply. This is because in examples like (266) and (267), *part* denotes sections of physical concrete entities (rockets and earrings) and, as such, no metadiscursive meaning is entailed.

(266) They, they brought in a heap of people to disassemble these rockets and inspectors as-- er examined every, every part of the rockets. (*BNC Sampler*: H47, S:speech:unscripted)

(267) <-|-> she said <-|-> yeah, she, erm, the reason that women who does the earrings couldn't make them for me, was that she buys all the parts, well the flower earrings she makes, she buys **all these parts** and she ex-- assembles them which ever parts you want. (*BNC Sampler*: KB8, S:conv)

In (268) and (269) *part* restricts a third-order entity (i.e. abstract *truth*) and a second-order one (i.e. eventive *defence*) respectively. In (268), *part of the truth* refers only to the underlined segment as, if only *the truth* had been used, reference would have been made to the whole truth (i.e. the official biased version, as well as the more objective scholarly version). In that instance, the encapsulation would have comprised both the preceding (i.e. official version) and the following (i.e. scholarly version) stretches of discourse (i.e. anaphoric and cataphoric encapsulation). Therefore, as reference is only to part of the overall truth (i.e. only the official version), only the underlined segment is here treated as the lexical realisation of the entire shell-noun phrase. Similarly, in (269), one possible part of the overall action of *defence* is that of *attack*. Hence, reference is made not to all kinds of *defence*, but to a specific aspect of this second-order entity.

(268) Russian colonial policies in the seventeenth and eighteenth centuries were aimed at maximizing the income of the state by the subjection of aboriginal peoples to fur tribute. It was fashionable at one time to see this process crudely in terms of mere military subjugation. Then the pendulum swung too far in the other direction, towards an assertion that the Russians avoided the unpleasant aspects of colonialism as exemplified in Spanish, Portuguese or British experience. Neither position, however, contains more than **part of the truth**. Fortunately we are now able to see the complexity of the processes more clearly through the painstaking researches of Soviet scholars. Russian servicemen were sometimes, but not always, brutal exploiters; they often felt frightened, vulnerable and far from home. Native peoples were not simply savages unable to pit their wits against a superior enemy, nor were they the Russians' dupes. Weakened by warfare, imported diseases and the excessive demands of their overlords, they were obliged in the end to submit. (*BNC Sampler*. FB4, W:ac:humanities_arts)

(269) 'Please, *maman*,' put in the prince in his most wheedling tone. 'Please let me join Edward'. Edward? Elizabeth Woodville's voice was unusually stern — her defences were down and attack was **the better part of defence**. 'To whom do you refer?' 'The king my brother,' the prince replied cheerfully. (*BNC Sampler*: *CCD*, W:fict:prose)

4.4.2.8 *Generic self-contained uses*

The use of the noun is most 'self-contained' (Ivanič 1991: 110) and generic, inasmuch as no reliance on any contextual information (endophoric or exophoric) is required for its interpretation (cf. Martin 1992: 103). This is typically associated with noun uses carrying no determiners. Examples (270) and (271) are two cases in point. In (270), the noun *vision* has the uncountable sense of 'the knowledge and imagination that are needed in planning for the future with a clear purpose' (*LDCE*). As it expresses a general human quality with no context-dependent meaning, the example is excluded from the shell-noun count. Example (271) is similarly excluded from the analysis on the basis of its uncountable meaning of 'the study of the nature and meaning of existence, truth, good and evil etc' (*LDCE*). If reference had been made to a particular set of ideas rather than to the discipline of philosophy, then the example would have been left in the analysis database.

(270) Acting chairman Derek Round, the former chief executive of Colchester and East Essex Co-operative Society, said: 'We are planning to appoint a mix of hands-on people with vision and commitment who have knowledge and expertise in all aspects of visual arts, including education as well as skills in marketing and business-related subjects. (*BNC Sampler*: CF6, W:newsp:other)

(271) From shorthand notes of conversations with Wittgenstein, supplemented by dictations and typescripts, Waismann wrote lectures and articles on Wittgenstein's conception of mathematics, his view of logic, and his treatment of identity and probability. The culmination was to be a systematic presentation of these ideas on logic, language, and philosophy. (*BNC Sampler*: GT9, W:biography)

However, there are many instances for which, despite the lack of a determiner, a context-dependent meaning is apparent (see Ivanič 1991: 111 and Sinclair 1993: 11 in 3.2.3). In these cases, a shell-noun analysis would apply, thereby leading to their inclusion in the count. For example, in (272), if *suspicion* were analysed out of context, it would be difficult to see what lies behind this use of the noun. Despite the absence of a specific determiner, a certain amount of contextual information is drawn on for its understanding. Such information cannot be clearly pinpointed or delimited, as it is only implied in the preceding co-text by such elements as *were brutally murdered*, *Thomas was apparently away in Reigate*, *she did not do it* and *her husband's inconsistencies and contradictions.* By the time the reader gets to the noun *suspicion*, the mental model created by the text leads to the understanding that it is not any suspicion that is at stake here.

The suspicion is that Thomas probably murdered his children, based on the evidence that the certainty of his being away in Reigate was only apparent, that his wife pleaded not guilty, and that, prior to the murders, he had been accused of several other crimes (fraud and arson). With all that negative textual prosody, *suspicion* is thus context-dependent and qualifies as a shell-noun instance. In this example, what is at issue is not a clear and definite relationship of encapsulation between a delimited discourse segment and a shell noun, but something more in line with Fraurud's (1992) notion of situation reference (see 3.2.2 and examples (256) and (257) in 4.4.2.6), inasmuch as the beliefs underlying this suspicion are only entailed at various points in the preceding discourse situation.

(272) Bacon, however, soon came into financial difficulties and he tried to foil a £90 debt owed to a local stonemason by forging a receipt. On discovery of the fraud, though, he destroyed the receipt to avoid incrimination. One night during 1856 his house was burnt to the ground and he was tried before Lincoln Assizes on a charge of arson of which he was acquitted. Around November 1856, Bacon and his family moved to Walworth in London and on 29 December their two children, Edwin Fuller Bacon, aged 2 ½, and Sarah Ann Bacon, aged 11 months, were brutally murdered by having their throats cut. As Thomas was apparently away in Reigate, Marina, his wife, aged 26, who had a history of mental illness, was charged with the murder. She pleaded that she did not do it and her husband's inconsistencies and contradictions aroused **suspicion**. Marina then blamed her husband for the act and this revived interest in the sudden death of Bacon's mother, Ann, at Stamford in May 1855. (*BNC Sampler*: CBB, W:non_ac:humanities_arts)

A similar example is that of the noun *crime*. The instances where *crime* is used with the uncountable sense of 'illegal activities in general' (*LDCE*) were often excluded from the analysis. This is shown in (273) below, where no specific kinds of crime are implied in the surrounding co-text. In other cases, however, the general meaning of uncountable *crime* is narrowed by context-specific information. In (274), for example, *so much crime* enables the writer to intensify the strength of the subsequent statement (i.e. the underlined stretch). As such, it is not crime in general that the writer is particularly worried about, but about crime against the elderly. The noun phrase here establishes a prosody that is taken up by the immediate specifics in the following sentence, which serve as emphatic examples to illustrate the point being made.

(273) Police dog handler Barry Barlow is seeking a mutt with 'a bit of sparkle' to be his new partner against **crime**. Pc Barlow's present dog, Ben, is

due to retire later this year and the officer from Thorpe-le-Soken is trying to find a replacement. (*BNC Sampler*: CF9, W:newsp:other:sports)

(274) Even the grocery man would go in, leave the goods on the table, take his money and leave the change. How different today. I wonder why there is so much crime or is it the world we live in which makes people so greedy? How sad to see our old people not just robbed but beaten up so badly as well. Sadly, I know there is less work today and many have turned to drugs and drink. (*BNC Sampler*: CF9, W:newsp:other:sports)

Plural *things* is another case in point. In (275), for instance, *things* is highly general and unspecific in meaning, as no details are found in the surrounding discourse to support a context-specific semantic interpretation of the noun. *Things* in this context would have the meaning of 'life in general and the way it is affecting people' (*LDCE*). In (276), by contrast, *things* is semantically restricted by the situation described in the conversation. If someone joined the conversation at that point, the news that one of the interlocutors could not sleep well because of things in her mind would motivate a question as to the reasons for such a state of mind. In this particular instance, such reasons are scattered throughout the previous discourse, and *things* is treated as a shell noun which, in Sinclair's (1993: 11) terms, performs a deictic act that encapsulates the underlined segment (see example (217) in 3.2.3).

(275) And the fact he can't just accept things, he has to question things.
And we were talking about ME, and how
Oh he's a very intense person.
how Dave, I was saying about Dave and how he's sort of taken, taken a completely different attitude to things.
More relaxed. (*BNC Sampler*: KCH, S:conv)

(276) Well I don't know, I feel happier now your dad's had that bit of a do with that bloke, cos I feel that <pause> we can go to him
Yeah.
and, and wi-- you know, hopefully get paid, there shouldn't be any problem. <pause dur='6'> Cos that erm <pause> that Sen, that <pause> accountant obviously isn't up to much. He promised your dad, oh it'll all be sorted within twenty four, forty eight hours and that was how many, three
Yeah.
four days ago?
I don't think he knows everything ab-- about that or what?
Erm <pause> he, he told him this morning on the phone, he's looked into it, he's the director <pause> and he apologizes.
[…]I didn't realize things were like this.
That's why <pause> you sent them the conditions!

> Well, that's what Andrew said. If, if they'd have <pause> if they'd have talked to Andrew first, Andrew would have told him why we'd sent it.
> Andrew can't <unclear>, he's just sort of passed that, that on. Why has it gone to somebody who <unclear> ?
> Well apparently <pause> it had gone to Andrew, and, Andrew had <pause> cos Andrew isn't in the office all the time, and he'd left it in the office, and <pause> I think they faxed it <pause> by th-- Andrew was gonna talk to somebody about it, but of course, by the time he'd got there, it'd already been done and this had happened so I'm gonna ring erm <pause> Job Centre when I get home.
> I, they'll say, I know what they're gonna say. Like, you know <pause> there's nothing they can do. […]
> <pause dur='6'> <cough> <pause dur='7'> Although I seem, I slept alright last night, but I think when you've got **things** on your mind<pause>no I didn't sleep alright, I was too warm, I tossed and turned most of night! (*BNC Sampler*: KB9, S:conv)

4.4.2.9 *Nominalised shell and non-shell nouns*

The assumed shell noun is, in Levi's (1978: 169) words, an 'act nominalisation', i.e. a verb-based nominalisation whose meaning is paraphrased as 'an act of X-ing' (e.g. *parental refusal = 'act of parents refusing'*). Two such instances from the corpus are (277a) and (278a). In (277a), the noun phrase in boldface implies *the process/act of assessing (or calculating) the second votes*, while in (278a) it entails *the process/act of correcting errors*. Whilst the same nouns occur in (277b) and (278b), their meaning is that of 'product nominalisations' (Levi 1978: 169), as their paraphrase is 'that which is produced by (the act of) X-ing' (e.g. *human error* = '*that which is produced by (the act of) humans erring*'). Given that product nominalisations result from prior (mental or eventive) acts, specifics identifying such results are expected to be in the surrounding discourse. In (277b) such specifics are in the underlined complement clause, while in (278b) no such specifics are given endophorically.

(277) (a) These bonuses, however, are not like those capriciously awarded under the STV: they are far smaller and they are shared out proportionally. Once the national allotment of seats has been fixed, a second stage in **the assessment of the second votes** gives each party its share of seats in each of the ten lander. (*BNC Sampler*. EW4, W:non_ac:polit_law_edu)

(b) He said that his assessment, **the telephonist's assessment of the call** was that that he was genuine. (*BNC Sampler*: JJV, S:courtroom)

(278) (a) **Correction of errors identified by users** will be controlled by the Computer Group Manager. (*BNC Sampler*: EAP, W:admin)
(b) However, it may be helpful if the system can sometimes suggest a **correction for a miskeyed word.** (*BNC Sampler*: H0S, W:misc)

In addition to act and product, Levi (1978: 169) also mentions 'agent nominalisations', whereby the verbal subject becomes a nominal or adjectival premodifier in the resulting noun phrase (e.g. *film cutter* = '*x such that x cuts film*'). This is shown in (279a) below. In this case, *market* does not specify the content of *failure*, as the meaning is not *markets are a failure*, but rather, *in the event that markets failed*. In (279b), by contrast, the former interpretation would apply, i.e. *the entire gassing operation was a failure*. The noun *failure* is thus being used to characterise a second-order eventive discourse entity, i.e. *gassing operation*.

(279) (a) Yet there are many ways in which **market failure** can arise, and State aid in those circumstances can form a legitimate means of corrective intervention. (*BNC Sampler*: G0C, W:commerce)
(b) Last week's report would call the entire gassing operation **a failure**. (*BNC Sampler*: HE3, S:brdcast:documentary)

Halliday (1998: 213) argues that examples like (279a) are closer to the verbal rather than to the nominal end of the linguistic continuum. Unlike prototypical instances of grammatical metaphor, where verbal and nominal features combine into a final noun phrase (i.e. 'semantic junction'), in examples like (279a) (e.g. *engine*/*heart*/*crop failure*), no complete semantic junction is apparent, in that the noun still preserves much of the nature of the verb it derives from. Underlying this argument is the fact that, although (279b) is product-like (and thus, more nominal), (279a) would clearly fall within the scope of act nominalisations, which are more verbal and in less need of lexical realisation by the surrounding discourse.

With this in mind, it appears that the shell or non-shell status of nominalised processes is a matter of degree. Deverbal nouns like *warning*, *answer* or *recommendation* are the clearest instances of the shell category, in that they derive from verbs often accompanied by projected clauses (e.g. *I warned them that*…, *I answered that*…, *I assumed that*…). When turned into nouns, the result is commonly a product nominalisation with inherent semantic gaps lexically realised by the following inherited projected clause, as in (280) (cf. Chomsky 1970 in 2.2.1.2).

(280) He said that during his recent visit to Moscow, President Gorbachev had given **a warning** that the failure of East Germany's reform process would not only destabilise the GDR but also have 'serious

repercussions' for perestroika in the Soviet Union. (*BNC Sampler*: A9M, W:newsp:brdsht_nat:report)

There are many cases, however, where the nominalised unit derives from a verb for which projection is not allowed. This is evident with *assessment* and *correction* above. As one can only assess and correct something (but not assess or correct that something is…), both verbs are more likely to undergo an act nominalisation rather than a product one (i.e. *the assessment or correction of something>the act/process of correcting something*). Telling a shell from a non-shell use is, in these cases, entirely dependent on the co-text. Only if the noun is found to contain semantic gaps in need of lexical realisation is the example counted as a shell noun. Otherwise, if no context-specific meaning is apparent, the example is treated as an act nominalisation and excluded from the count (e.g. (277a), (278a), (279a)).

Another example of a noun derived from a non-projecting verb is *endorsement*. According to Ashby (2005; *Oxford Advanced Learner's* Dictionary; henceforth, *OALD*), *endorsement* is defined as 'a public statement or action showing that you support somebody or something'. Only examples drawing attention to the information contained in such a statement or action are included in the count. If the meaning is that of 'confirmation, ratification, approving testimony' (*OED*), *endorsement* is treated as the result of agreeing with something, and the example is thus excluded from the count. This is shown in (281) and (282) below. These two examples do not highlight the actual words involved in the act of endorsing, but the support that follows from those words, i.e. the fact that these countries support or agree with the Treaty. Hence, *endorsement* in these examples is a self-contained and bounded discourse entity with little or no need for further specification from the co-text.

(281) Agreement on revisions to the Rome Treaty by the end of 1991 would allow a further 12 months for the delicate business of securing **the endorsement of all 12 national parliaments**. (*BNC Sampler*: A9M, W:newp:brdsht_nat:report)

(282) But all the signs are that if Britain still withholds **its endorsement**, the other 11 will conclude an agreement among themselves outside the formal EC framework. (*BNC Sampler*: A9M, W:newsp:brdsht_nat:report)

Examples (283) and (284), by contrast, are treated as shell on account of the greater emphasis on the statement, rather than on the ensuing support. In (283), *his endorsement of perestroika* follows a number of reporting clauses (i.e. *he said…*) offering Mr. Ashkenazy's actual words in showing or stating his support for perestroika. A similar explanation would apply to

(284), where *his generous endorsement of the work that the Board has done* constitutes the speaker's evaluation of someone else's words. A reader or listener first encountering these two examples might ask 'what made you say that X has endorsed Y?', and the writer or speaker may well answer 'his/her words or what he/she said when showing support for Y'.

(283) VLADIMIR Ashkenazy, one of the world's greatest pianists, said yesterday he felt far more at home in the Soviet Union now than when he departed under a cloud half a lifetime ago. Returning to Moscow at the age of 52, for the first time since he left to live in London 26 years ago, Ashkenazy said: 'In my modest way, I thought I could endorse what's happening in my country'. Although he was never a political defector, he said he probably would not have come back before President Gorbachev launched his perestroika policy. In 1963 'there were some very unpleasant moments', he told journalists at the Soviet Culture Foundation. For many years in Moscow, I felt I couldn't leave the country'. But now, he said, 'I see open faces. People aren't frightened. I feel at ease — much better than I did then'. […] The Soviet world of culture is not alone in regarding Ashkenazy as an immensely welcome guest. His endorsement of perestroika will be music as sweet in the ears of Mr Gorbachev and his supporters, as the brilliant notes everyone confidently expects to hear from him. (*BNC Sampler*: A87, W:newsp:brdsht_nat:report)

(284) Moderator could the convenor tell us why the Board wishes to be discharged its remit on health and healing when it's been so successful in bringing this to the notice of the whole church?
<pause>
<unclear> grateful to Mr <gap desc='name' reason='anonymization'> for his very erm generous en--endorsement of the work that the Board has done on health and healing. (*BNC Sampler*: F86, S:meeting)

Mention should also be made of eventive objects (Quirk et al. 1985: 750–2), i.e. combinations of a semantically general verb like *have*, *make* or *do* and a nominalised unit as object. Schmid (2000: 25) suggests that these examples are not '[…] good examples of shell nouns […]', insofar as the whole expression is equivalent to a single verb. The contention here is that, so long as the example is observed to draw on contextual specifics, a shell-noun analysis will apply. This is apparent in (285) below. The noun in boldface occurs in the introductory section of an academic paper. In line with Tadros (1985 and 1994 in 2.2.2.1.3), here *assessment* functions as an advance label committing the writer to a subsequent discourse act. The realisation of such an act is shown in the underlined segments, where the evaluative stance adopted by the writer (e.g. *major shift*, *largely procedural*,

the main problem is, etc.) is in tune with the inherent meaning of the noun *assessment* (i.e. make a judgement about something).

(285) This is followed by a section outlining current European Community policy. Finally, an assessment is made of the effectiveness of current policy, and with the creation of the SEM, the important implications arising from the Treaty of Rome are considered in relation to the future direction of British policy and proposals for change. [...] In the case of restrictive practices the new proposals represent a major shift towards a European type approach. However, in the case of mergers the proposed changes are largely procedural, and no significant reorientation of present policy is envisaged. There are at present no specific proposals for changes in the treatment of dominant firm monopolies. [...] Despite the early successes of restrictive practices legislation in combating collusive agreements, there is now a belief that within the present business climate the legislation possesses fundamental weaknesses. This belief was articulated within the official review of restrictive practices policy (Department of Trade and Industry 1988c). The main problem is that the policy approach that has evolved is no longer seen as efficient. In particular the deterrent effect is weak since the penalties for operating an illegal cartel are inadequate [...] (*BNC Sampler*: HXN, W:commerce)

Examples like (286) and (287) are excluded on the grounds of their non-shell use. While (285) is closer to a product nominalisation in the subsequent provision of informative specifics, (286) and (287) are clear act nominalisations. In (286), *assessment* does not encapsulate any preceding or subsequent evaluative specifics, as its meaning is that of a process, i.e. *the framework within which the suitability of any publication could be assessed*. Given the nature of the surrounding co-text, the noun is generic in meaning, which means that there is no expectation for a following discourse act that the noun might be taken to encapsulate. Similarly, in (287), no particular *joke* is intended, which leads to a verbal understanding of the whole expression as *he was joking about himself* or *he was laughing at himself*. Falling within the scope of excluded eventive objects are also expressions such as *put into practice* (<*to practice*), *have application* (<*apply*), *make sense of* (<*understand*) and *take part in* (<*participate*), all of which may be replaced by single verbs, but show no inherent semantic gaps in their use in context.

(286) While the government does not foresee the need for pre-censorship of publications, films and video cassettes, it is the intention of the government to enact appropriate legislation to provide the framework within which an assessment could be made of the

suitability of any publication, film or video cassette for public and private consumption with regard to obscenity and offending of religious convictions. (*BNC Sampler*: EBJ, W:misc)

(287) […] I, don't actually know an awful lot about Gaugin, but if, if I knew nothing about him at all, I would of thought he was having a bit of joke of himself with this, but er, being the person that he was […] (*BNC Sampler*: F71, S:speech:unscripted)

From all of the above, the conclusion appears to be that shell uses typically correlate with product nominalisations. Act-like nominalisations may also occur as shells, but only when their use involves context-specific semantic gaps. On certain occasions, however, an act nominalisation is included in the count despite a lack of clear semantic gaps. This applies when the potential shell noun occurs as subject or subject complement in a relational environment. In the absence of more informative specifics in the surrounding discourse, reference is held to apply between the potential deverbal shell noun and the other noun. Whilst clearly verbal in meaning, the occurrence of these nouns in relational environments imposes a contextual interpretation of one noun in terms of the other. These instances are accordingly left in the shell-noun count. Examples (288a) and (289a) illustrate this kind of context-dependent meaning. In these two examples, *endorsement* does not entail the words behind a public statement in support of something, as in (283) and (284), but rather, the act of endorsing something, as evident from the paraphrases in (288b) and (289b). As the act of endorsing is here understood in connection with another act, the meaning of *endorsement* is treated as context-dependent and a shell analysis applies.

(288) (a) For your company to participate in this scheme, contact the Recycling Officer. (Inclusion in the leaflet does not imply endorsement of the goods & services provided). (*BNC Sampler*: G2R, W:institut_doc)

(b) Including the name of your company in the leaflet does not force you to endorse the goods and services provided.

(289) (a) A leader of Socialist Left, a radical pressure group, said the confirmation in office of the old team was 'a serious mistake, because it signals the endorsement of an economic policy criticised by a lot of people'. (*BNC Sampler*: A95, W:newsp:brdsht_nat:report)

(b) The fact that the old team has been confirmed means that the government still endorses an economic policy criticised by a lot of people.

Similar examples may be found in definitions. In these cases, one of the two nouns, the defined term, is semantically more distilled than the other, the

defining noun, which, whilst more complex in form, is more transparent in meaning (cf. Halliday & Martin 1993: 30 in footnote 10, 4.4.2.2). A hypernym-hyponym relationship underlies these examples, with one of the two nominals being the hyponym of the pair and the other the hypernym.

This is shown in (290) and (291) below. In (290), *collection failure* is the more specific member of the two, as it assigns a term to the more self-explanatory phenomenon of *the absence of the sought item or topic from the collection*. *Collection failure* is thus introduced as a generic concept, one which, at this stage, is not used to encapsulate a particular failure. On these grounds, prototypical shell-noun encapsulation would not follow from what has been said. What this example shows is semantic distillation, inasmuch as a shell-noun phrase with a specific and precise meaning (i.e. *collection failure*) distils what is taken to be the more congruent or transparent representation of this phenomenon. In the case of *assessment* in (291), distillation takes place in a different direction. The shell-noun phrase acts as the hyponym, thus rephrasing a distilled specific item (i.e. *quality control*) into more congruent and clearer words. *Assessment* in this instance is an act nominalisation, inasmuch as it entails *the process whereby inspection results are assessed. Assessment*, therefore, contains no inherent semantic gaps in need of lexical realisation. Its inclusion in the count is, once again, dependent on the contextual interpretation of one act nominalisation (i.e. one process) on the basis of another.

(290) The session failure rate would have been considerably lower, but it is very difficult to estimate this from log data alone. Note that failure does not include **collection failure** — the absence of the sought item or topic from the collection; one of the functions of a good catalogue is to inform users as quickly as possible that the library does not hold what they are looking for. (*BNC Sampler*: H0S, W:misc)

(291) So you're saying that **an assessment of the inspection results** is quality control?
Well you've gotta have a definition or a standard to inspect against really, I think Geoff, as well. Otherwise you can't say it's good, bad or indifferent. (*BNC Sampler*: H47, S:speech:unscripted)

There are two additional instances which are not nominalisations, but share the same kind of analysis as examples (290) and (291) above. Like (291), (292) shows in boldface a more congruent representation of the underlined term. It differs from the former example in that, whilst *assessment* is used in (291) as a self-contained act nominalisation, *words* appears to need further specification from the surrounding co-text. In this particular instance, the reader might ask: 'What words? Can you give an example?'. In (291), by contrast, a question such as 'what is your assessment of the inspection

results?' would sound decidedly odd, insofar as *assessment* there does not imply the product of a prior act of assessing (the most typical shell sense), but the act itself. With this in mind, *words* in (292) has a twofold reference: the relation of context-dependent distillation between defining and defined term, as in (290) and (291), as well as the more typical shell-like encapsulation of other elements of discourse. As no examples are given of such *words* or *hapax legomena* in the text, encapsulation here is taken to be exophoric. With regard to example (293), the lexical realisation of *system* seems to be split between the subject (i.e. *Conductive Education*) and the postmodifying clause (i.e. *aimed at stopping a disability from becoming a handicap*). Flowerdew (2003a: 336–7) argues that the pattern where this noun occurs is typical of formal written definitions, which tend to comprise a term in subject position and both a class and a characteristic in complement position. In these cases, the signalling or shell noun is lexically realised '[…] between the subject and its own postmodification' (Flowerdew 2003a: 336). In (293), *a preventive system* (i.e. the class) is realised by the term in subject position (i.e. *Conductive Education*) and by the characteristic, represented by the postmodifying clause (i.e. *aimed at stopping*).

(292) Biblical scholars have long noticed, and based arguments on, the occurrence of hapax legomena, or words which occur once only in a given text or author. But if one is to allow counting at all in a literary context, it seems perverse to insist that counting should stop at the number one and go no further. (*BNC Sampler*: F98, W:ac:humanities_arts)

(293) Conductive Education is a preventive system aimed at stopping a disability from becoming a handicap. Motor and intellectual development are inseparable. (*BNC Sampler*: APJ, W:misc)

4.4.2.10 *Prepositional phrases with a subjunct or conjunct role*

The assumed shell noun forms part of a prepositional phrase with an adverbial subjunct or conjunct function (Quirk et al. 1985: 566–612, 631–47). The subjunct function is evident in cases of intensifying prepositional phrases, which amplify or tone down an element of the utterance or the whole utterance. The conjunct function, by contrast, spells out the relation between any two discourse segments.

Examples (294) to (296) illustrate the former function. *In a sense*, according to the *OED*, presents '[…] a particular acceptation or interpretation (of a word, phrase, etc.)', and '[…] sometimes come[s] to mean "in some degree", "in no respect", "on every account", etc'. The occurrence of this phrase in (294) is tied to the 'in some degree' interpretation, inasmuch as the whole prepositional phrase, in the form of a

sentence-initial wide orientation subjunct (Quirk et al. 1985: 568–72), helps to reduce the assertiveness of the adjoining proposition, thereby making it more tentative. As regards examples (295) and (296), *in a way* and *in some ways* are similar to (294) in their initial position, as well as in their subjunct-like meaning. According to Sinclair (2006; *Collins Cobuild Advanced Learners' English Dictionary*; henceforth, *CCD*), *in a way*, as in (295) is used '[…] to reduce the force of a statement', and *in some ways*, as in (296) indicates '[…] the degree or extent to which a statement is true'. Once again, therefore, the whole prepositional phrase serves to tone down the propositional meaning of the adjoining sentence.

(294) Like the Boar scene, the scene in the hut includes farther-reaching reminiscence, in the big chunk of 'They listen to money' from Act One scene <gap desc='figure'>one [9] But the main role of reminiscence in Grimes is reserved for the final interlude and the monologue which follows it. In a sense, just as out and out dementia is less interesting than neurosis and eccentricity, so the music of these episodes, with their rather obvious distortions and wrong attributions of earlier themes, is less gripping than the Passacaglia, where Grimes was still somewhat in control of his faculties. The scene is nevertheless necessary and, if well sung and acted, very moving. (*BNC Sampler*: J55, W:non_ac:humanities_arts)

(295) <-|-> And I'm not working next <-|-> year anyway cos I've done it for two years on the run and I'm not doing it next year.
I've done it five, this is the first time in five years I've had New Year's Eve off. […]
<-|-> In a way I think <-|-> Gwynn had one too many New Year's Eve. (*BNC Sampler*: KCA, S:conv)

(296) […] it is a recursive procedure for transforming IF b P to b, x-normal form, where P is an x-IF/ALT program without uninitialised variables. Indeed in some ways the proof is rather simpler than Theorem 1, since it does not need such a complex structure of nested recursions. (*BNC Sampler*: G3N, W:ac:tech_engin)

Examples (297) and (298) are different in that the prepositional phrase is not a clause-initial adverbial merely narrowing the truth scope of a particular proposition. In these two instances, the information focus is not so much on the pragmatic force of the entire prepositional phrase, as in (294) through (296), as on the complementing noun phrase itself.

(297) <pause> Yeah lactic acid <pause>
Is poisonous.
is poisonous in one <-|-> sense <-|-> .
<-|-> Gives you <-|-> stitch.

What Yes.
Neil, say it again.
Gives you stitch and cramps.
Yes, gives you stitch and cramps. (*BNC Sampler*: FLY, S:classroom)

(298) Rather, it is that there exist people willing to pay prices lower than Pm, but which are in excess of the marginal costs of providing additional units. The existence of monopoly denies them that opportunity, and this is manifest in the inevitable reduction in total surplus. It is in this sense that monopoly is said to be economically inefficient, and to misallocate resources through the restriction of output. (*BNC Sampler*: HXN, W:commerce)

The noun-based orientation of these expressions is borne out by the fact that, unlike examples (294) through (296), *in this* sense and (less frequently) *in one sense* allow the prepositional phrase to be the focus of a cleft sentence. With this test, Quirk et al. (1985: 504) tell intraclausal adverbial adjuncts, whose interpretation depends on the main verb, from sentential adverbials, which modify the whole sentence, as in (294) through (296). Winter (1977: 16) uses the same test for a similar distinction between 'anaphoric adjuncts' (e.g. *in such circumstances*, *for this reason*) and 'sentence adjuncts' (e.g. *accordingly*, *in contrast*). This said, a search in *BNCweb* for the sequence *it {be} in this sense that*, as in (298), returns 31 hits, while a search for *it is {be} in a sense that*, *it {be} in a way that* and *it {be} in some ways that* returns no hits. Incidentally, the same query in *google.co.uk* gives 79300 hits for *it is in this sense that*, only 2 hits for *it is in a sense that* (these being, however, relative clauses, i.e. *it is in a sense that/which*), 7 million hits for *it is in a way that* (most of which are once again relative clauses, i.e. *it is in a way that/which*) and no hits for *it is in some ways that*.

In the light of the abundant evidence for the cleft pattern containing *this*, it appears that the moment a specific determiner (e.g. *this*, *that*, *the*) is used instead of an unspecific one (e.g. *a/an*, *some*, *zero*), the noun phrase in the formerly intensifying expression becomes more referential and context-bound. This shows that the emphasis is shifted from the whole prepositional phrase to the noun itself, thereby leading the noun to acquire the context-specific semantic gaps typical of shell items. Such semantic gaps are evident in examples (297) and (298). If the underlined segments were omitted in both, any speaker or writer in such discourse situations might ask: 'could you please tell us in what sense is lactic acid poisonous?' and 'in what sense is monopoly economically inefficient?'. It should be noted that, whilst the nouns in both (297) and (298) are shell in their referential and context-specific meaning, only (298) admits a cleft pattern. A search in *BNCweb* and *google.co.uk* returns no examples for *it is in one sense that*.

This being the case, *one sense* is treated as shell on account of the semantic gaps that the speaker (here, the teacher) expects listeners (here, students) to fill (i.e. *gives you stitch and cramps*).

Examples (299) and (300) resemble the latter two in the occurrence of a specific deictic (i.e. *this* and *that*) and in the context-specific gaps inherent in the noun phrase itself. In (299), *in this way* functions as a manner adjunct specifying what his thinking involves, i.e. the boy, the jersey, Ellen, the sea, etc. In (300), *in all that way* is a respect adjunct showing the basis on which *he's quite correct* needs to be interpreted, i.e. *he's quite correct in relation to his argument that you can't get fish from prize thinking fish because the fish itself is a living thing.* Both instances would therefore fall into Winter's (1977: 16) anaphoric adjuncts, insofar as the complementing noun phrase contains semantic gaps lexically realised by preceding discourse segments.

(299) As his mind rambles from the boy to the jersey to Ellen (who knitted it), to the sea and the Borough (via the fish Grimes is going to get rich by catching), back to Ellen and the future life they will never lead because Grimes cannot escape his past — as his mind rambles **in this way**, so the music probes our sympathies by constant allusion back to the variations in the interlude. (*BNC Sampler*: J55, W:non_ac:humanities_arts)

(300) you can't get fish <unclear> from <pause> prize thinking <-|-> fish <-|-> <-|-> No.
can you?
No.
But I know they blamed him for that.
And he says you can because he says it's important for it to be a living thing that's got to change with the time.
Mm.
In all that way he's quite <-|-> correct <-|-> (*BNC Sampler*: KC0, S:conv)

At times, the occurrence of a specific deictic is no guarantee of shell status. This is especially the case when such a deictic is *the*. In example (301a), *at the same time* is semantically equivalent to *however* or *yet*, which implies that the entire prepositional phrase is an adverbial conjunct. Whilst conjuncts realise the semantic relations between discourse segments, their connective role depends on the prepositional phrase rather than on the complementing noun phrase. In most such cases, the prepositional phrase may be replaced by another Vocabulary 1 (i.e. conjunction) or Vocabulary 2 item or expression (i.e. adverbial conjunct; see Winter 1977 in 2.2.2.1.2). This lends further support to the argument that the emphasis is on the prepositional phrase. Such a lack of semantic gaps in the noun phrase is also evident in other conjuncts like *for example* and *in other words* in (302)

and (303a). Both expressions are analogous to other reformulatory or appositive conjuncts, e.g. *for instance*, *that is to say*, *namely* and *more accurately*. Examples (301) through (303) share the fact that the noun phrase does not admit any other determiner. This proves that the prepositional phrase is fossilised in its conjunct meaning and that the referential status of the noun phrase becomes markedly weakened. Substituting *this* for *the* and *other* in *at the same time* and *in other words* would transform the conjunct phrase into a shell-like anaphoric adjunct, with the complementing noun phrase (i.e. *at this same time*, *in these words*) now performing a referential function in discourse (see (301b) and (303b), queried in *BNCweb*, but not included here because they do not occur in the *Sampler*).

(301) (a) I was certainly to do with <pause> feeling that I would have to be a little bit slimmer, that I would have to be acceptable to other people, that I would have to change my shape <pause> and at the same time, I could not resist consuming large amounts of food. (*BNC Sampler*: FL6, S:brdcast:discussn)
(b) It was this very quality of non-conformity for which Leslie had expressed admiration in June 1941 when Wavell was relieved of his Middle East command. At this same time Leslie had voiced his discontent at the 'ponderous inefficiency' and 'heartless bureaucracy' of the conventional army [...] (*BNC*: AMC, W:biography)

(302) However, this is no excuse for inaction, for a number of systems have been described which are of considerable assistance; for example, the Telford system which relies largely on professional judgement can be used to good effect. (*BNC Sampler*: EVY, W:commerce)

(303) (a) Of the considerations governing the selection of constituency candidates by one of the major Irish parties the most cogent could be called the 'seniority factor'. In other words the candidates most likely to be nominated if they so wish will be those who were nominated and became TDs at the previous general election. (*BNC Sampler*: EW4, W:non_ac:polit_law_edu)
(b) A clue is contained in Paul's command, 'Children, obey your parents, in the Lord', which is bracketed with the quotation of the fifth commandment, 'Honour your father and mother' (Eph. 6:1–2). In these words, Paul turns from the principle of the learning process to the command to glorify. (*BNC*: CGE, W:religion)

Overall, when it comes to the analysis of subjunct- and conjunct-like prepositional phrases, a shell analysis applies only where the complementing noun contains a specific deictic (most often, *this*) and where context-specific semantic gaps underlie their use, as in (297) through (300) and (301b) and (303b). By contrast, if a clear subjunct or conjunct function

prevails (often with non-specific determiners), the example is then treated as non-shell and is discarded, because it is the prepositional phrase (not the noun) that is at issue here, as in (294) through (296), (301a) and (303a).

4.4.2.11 *Shell and non-shell uses in set phrases*

The assumed shell noun forms part of a non-referential set phrase. As in (301a), (302) and (303a), varying degrees of formal fossilisation underlie these expressions. Their meaning is often more metaphorical, idiomatic and less compositional than that of conjunct phrases. Two cases in point are (304) and (305). The occurrence of *way* and *surprise* in the set phrases in boldface does not entail a shell analysis, as the noun phrase is not referential, but is part of an expression semantically equivalent to a less metaphorical lexical item or expression (i.e. *pave the* **way** *for>enable*, *something catches you by* **surprise***>something happens unexpectedly*). Similar examples as *be on your* **way** (<*head somewhere*), *something is* **at work** (<*something affects something else*), *by any* **chance** (<*perhaps*) and *take your* **leave** *of somebody* (<*to say goodbye to somebody*) are also excluded.

(304) REMITTANCE Man's early season setback has **paved the way for** Katabatic to regain the Queen Mother Champion Chase crown at Cheltenham today. (*BNC Sampler*: CF9, W:newp:other:sports)

(305) I can't understand how I did it you know. <pause>
Well you didn't realize it was going over did you?
No. <pause>
It **catches you by surprise**.
I think I must have had my weight to the front of the bike, you know, coming up on the braking. <pause> (*BNC Sampler*: KCL, S:conv)

This said, there are metaphorical set phrases where the noun may be argued to contain shell-like context-dependent gaps. Three such examples are discussed in the following. In (306), *suspicion* is included in the count despite its lack of a determiner and its occurrence in a metaphorical expression. This is explained on the grounds that saying that *the finger of suspicion is pointed at X* is equivalent to saying that *the suspicion is that X is responsible for Y.* In this particular instance, it may be inferred that the suspicion is that *methyl mercury that is accumulated in the soil is to blame for the high incidence of mental retardation and other congenital conditions*. The encapsulation is thus both anaphoric and cataphoric. In (307), *ways out* is a metaphorical expression meaning 'a way to escape a difficult or bad situation' (*LDCE*). As three such ways are offered straight away, *ways* is treated as a context-dependent shell.

(306) In the old gold mining regions of Min Sherice and Gojas scientists are finding an unusually high incidence of mental retardation, and other congenital conditions, and the finger of **suspicion** is pointed at methyl mercury that 's accumulated in the soil and in the ground water. (*BNC Sampler*: HE4, S:brdcast:documentary)

(307) Those who foresaw the inevitability of this could take **one of three ways out**: commit suicide like the Itelmens; resist and fight like the Chukchis; or learn to get along with the newcomers. Only those who chose the last option survived. (*BNC Sampler*: FB4, W:ac:humanities_arts)

The expression *take my word for it* in (308) is used to ask somebody to '[...] believe you because you are telling the truth' (*CCD*). What this person said (i.e. *his word*) on that occasion is implied in the situation described in the underlined discourse segment. His word or message involved a chicken that had gone off, a complaint about that chicken and the act of taking the label back in order to obtain a refund. The actual words in uttering that message are not explicit, but the underlined specifics offer useful clues as to the content of that message and, as such, the example is left in the shell-noun count. A search in *BNCweb* for the sequence {*take*} + *possessive determiner* + *word for it* reveals that variation is apparent in the possessive determiner slot (i.e. *my*, *his*, *your*, etc.). This implies that, in spite of its apparent rigidity, the construction is still used meaningfully within any given textual context. Further, the existence of examples where a *that*-complement clause follows *for it* somehow blurs the independent syntactic slot that *for it* might be claimed to occupy. This is also supported by the weak prominence that *for it* receives in pronunciation. Example (309), retrieved from *BNCweb*[14], is a case in point. In this instance, the writer's word or message is that *a hill-walk is still an exhilarating day out.*

(308) And that was forty eight pence for one grapefruit because it was pink.
Doesn't even look pink, does it?
No. Well I think you complain. I don't think you have to take it back, I didn't take that chicken back to Safeway's if you remember.
What chicken was that?
The one that smelt. <pause>
Did you complain?
Yes. Don't you remember that?
I remember <-|-> the chicken. But I don't remember <|->

[14] No examples of *that* appositive clauses following *for it* are given in the *Sampler*. This example is therefore not included in this study.

<-|-> Some months ago.<-|-> I took the label back. Didn't take, didn't bother with the chicken. And she didn't, she scarcely even looked at the label. She took my word for it. Unfortunately because it, it was a special sell today job with a reduction off others I didn't get much money back. (*BNC Sampler*: KBK, S:conv)

(309) Perhaps it's the moaning you can do at the top, or the relief of getting down and into the warmth. Whatever the reason, you will have to take my word for it that a hill-walk is still an exhilarating day out even when you can't see a thing. (*BNC*: AS3, W:misc)

4.4.2.12 *Incomplete discourse*

The unplanned nature of spontaneous spoken discourse often results in hesitations, pauses and false starts, which may hinder the contextual interpretation of a shell-noun item. The instances where the surrounding discourse is unclear or incomplete are included in the count, with the caveat that only certain of the 9 study variables (see 4.4.1) are assigned a specific value. Those variables for which the incomplete nature of the corpus written transcript forbids their analysis are annotated as unclear, i.e. UNC. Three such examples are given below.

In (310), the occurrence of the <unclear> tag before and after *your chances* makes it difficult to ascertain its syntactic function, its participant type, its direction of encapsulation and its antecedent. From the corpus transcript, it may only be inferred that there is a possibility of something happening, but it is not known what such a possibility entails.

(310) Cos they all have a go
You don't all have a go on this one.
We do when we're stuck on <-|-> a <unclear> <-|->
<-|-> Can we have one of these?
<unclear>
You do get stuck on a regular, then your chances <unclear>
What is it? What is it? What number?
It's a green one.
Ready steady <pause> go. (*BNC Sampler*: KD0, S:conv)

Example (311) illustrates a false start, insofar as the speaker introduces something worrying, but no subsequent specifics are given to realise the encapsulation. The speaker's utterance is interrupted by another interlocutor's observation about air bricks, these being a possible solution to the lack of ventilation in the speaker's loft. Whilst *I can't understand how they managed to get* might be treated as potential specifics for *thing*, this is unlikely, because this is a new problem derived from the air bricks solution, not the problem or worrying situation originally intended by the speaker (cf.

Hoey 1983 in 2.2.2.1.2). With this in mind, both direction of encapsulation and antecedent are treated as unclear. Notwithstanding the absence of a verb, the expectation is that a relational process should follow the noun phrase (i.e. *the only thing that worries me is that...the...*). In view of such an implied verb, syntactically and semantically, the noun phrase is treated as syntactic subject and as semantic Identified.

(311) I said to them <pause> er, I've got no ventilation there, I said, I'd need something. He said, oh that's easily done. Well erm <pause> erm <pause> the chappie who did my door said, oh, there's no problem there you can easily put <-|-> ventilation <-|->
<-|-> Ah well.
in there.
Oh.
But it does need it.
I don't know
And the only thing that erm <pause> <-|-> worries me <-|->
<-|-> See we've <-|-> got air bricks up around <-|-> ours.
<-|-> Yes.
We got two.
Yeah. You got <unclear> I can't understand how they managed to get <-|-> <unclear> <-|->
<-|-> Well don't they put air bricks <-|-> in places?
Well they should do, I would <-|-> have thought. (*BNC Sampler*: KST, S:conv)

Example (312) is unclear with regard to direction of encapsulation and antecedent. The occurrence of up to 5 <unclear> tags in such a small discourse segment obscures the interpretation of *their crime*. A search for the word *crime* and *children* in text JJA reveals no other examples of any of these nouns. Hence, it is not possible to establish whether children are the victims or the instigators of crime or what the nature of such a crime is.

(312) I hardly dare mention traffic calming but it does seem to be the flavour of of the month at the moment but <unclear> to see that the existing work <unclear> doubled by <unclear> and I know it's going to cost the county erm this <unclear> but if you don't want to talk about the children and their crime it is er traffic calming <unclear> it certainly is a <unclear> and I do believe. (*BNC Sampler*: JJA, S:meeting)

4.5 CONCLUSION

This chapter describes three aspects of the research process: corpus selection, sampling and the method for data analysis. It has been demonstrated that of the small general English corpora available, the *BNC Sampler* offers the highest degree of balance and representativeness. Whilst being a microscopic version of its parent corpus, the *Sampler*'s 50%-50% coverage of written and spoken language is better suited for this thesis than the larger but less well-balanced *BNC*. As the objective is a detailed analysis of shell units, it was deemed necessary to reduce the research scope to 60 lemmas and 40 concordance lines per lemma.

Such a decision was influenced by a corpus theoretical approach to shell-noun description. Inspired by Mahlberg's (2005) seminal study on general nouns (partiallly overlapping with shell nouns), this approach involves the analysis of a small sample of corpus data on the basis of only two assumptions about shell-noun behaviour, namely that they are second- and third-order entities and that their interpretation is context-bound. Bearing these two assumptions in mind, it was decided that, prior to the discussion of the analytical decisions for each of the 9 variables considered here (see chapter 5), it was important to separate shell from non-shell nouns. Only if such a distinction is clear, may the decisions in the following chapter be better understood. The 12 categories of excluded instances presented in 4.4.2 may be summarised as follows:

i) Premodifying nouns: Premodifying nouns are left out of the analysis database because they are non-referential. Only where the noun occurs at clause level or as complement in a prepositional phrase is it as a referential shell unit.
ii) Complete repetition: Only partially repeated nouns are included, provided that the simpler noun stems from a more complex and informative noun phrase or discourse segment.
iii) Erroneous word-class tag: The node word, adjectival or verbal in form, is erroneously tagged as a noun.
iv) Wrong item: The speaker uses a shell unit by mistake.
v) Naming expressions: The noun forms part of a non-referential naming expression such as a book or newspaper title.
vi) First-order entity: closed sets and value ranges: The noun labels either measurements (e.g. *dimension*) or closed sets of units (e.g. *human senses*). Circumstantial nouns such as *time* or *point* are treated as shell only if they are found to encapsulate eventive information.

vii) First-order entity: visible and tangible unit: In general terms, nouns denoting real-world physical entities are treated as non-shell, insofar as they are neither abstract (second- and third-order) nor meta-discursive in meaning. Therefore, nouns like *area*, *application* and *part* are shell only in the abstract senses of each word.
viii) Generic uses: Those noun uses denoting general qualities or concepts lacking in any context-specific semantic gaps are excluded from the analysis database. Only if the close reading of the surrounding co-text reveals certain contextual specifics does a shell analysis apply.
ix) Nominalisations: Only product-like nominalisations are likely to contain shell context-bound gaps. Act nominalisations are treated as non-shell unless the example is dependent on the surrounding co-text for its interpretation.
x) Conjuncts and subjuncts: These are discarded, because the attention is directed not to the complement noun, but to the entire prepositional phrase. Only those instances where the noun itself is referential in meaning (typically associated with *this*) are included in the count.
xi) Set phrases: Only in those cases where the noun in a metaphorical set phrase is observed to contain context-bound gaps does a shell analysis apply.
xii) Incomplete discourse: The fuzzy nature of spoken discourse at times forces the annotation of certain variables (most often, direction of encapsulation and antecedent) as unclear.

In all the above categories, context plays a major role. The second assumption is therefore crucial for the identification of shell nouns in this thesis. This implies that, regardless of the form or pattern of the noun phrase, any second- or third-order entity is treated as shell provided that a context-bound interpretation is required for the lexical realisation of inherent semantic gaps. As illustrated by the examples given in this chapter, such semantic gaps are not linked to lemmas, but to senses of these lemmas (see 3.2.5.2).

5 THE NINE VARIABLES

5.1 INTRODUCTION

Section 4.4.1 outlined the nine analytical variables used in this study. Chapter 5 delves into each of these variables with a view to presenting a detailed account of all the categories, principles and decisions underlying the identification of shell-noun instances in the study sample. As in chapter 4 (especially in section 4.4), the focus of this chapter is on the qualitative results obtained by the analysis of examples. The chapter has two main sections: the analysis database (5.2) and the variables (5.3).

Section 5.2 describes the layout and coding of the analysis database. Section 5.3 is the core of this chapter, because it examines the rationale behind the nine variables. The first subsection (5.3.1) briefly discusses genre categorisation in the analysis database. Section 5.3.2 then turns to the semantic and formal structure of the noun phrase. Following the description of each variable in 5.3.2.1 and 5.3.2.2, 5.3.2.3 looks at mismatches between formal and semantic structure. In 5.3.3, the focus shifts from phrase-level patterns to clause-level functions. It deals with SFG participant types, syntactic function and Theme-Rheme. This section also discusses some of the problems found in the application of Hallidayan Transitivity to second- and third-order shell entities (5.3.3.1.1 through 5.3.3.1.5). Section 5.3.4 moves from clause-level functions to the encapsulating potential of shell-noun phrases. It comprises two subsections: 5.3.4.1 is devoted to encapsulating directions, and 5.3.4.2 to antecedent types. Lastly, section 5.3.5 is about the semantic categorisation of the shell units in the study sample, with emphasis on units (5.3.5.1) and senses (5.3.5.2) unaccounted for in Schmid's (2000) study.

5.2 THE ANALYSIS DATABASE

As noted in 4.4.1, this study takes account of nine variables spanning formal, semantico-pragmatic and textual levels of linguistic description: genre, experiential structure, formal structure, syntactic function, participant type,

Theme vs. Rheme, direction of encapsulation, antecedent and semantic type. Each variable is displayed in separate columns in a *Microsoft Excel* spreadsheet, which is referred to below as the database. Overall, the database comprises 15 columns (from left to right):

i) The numbered list of word-form examples for each lemma (20 singular and 20 plural). For instance, the 20 examples of singular *project* are numbered as *project (1), project (2),* etc., whilst plural *project* appears as *projects (1), projects (2)*, etc.
ii) The second column specifies the *BNC Sampler* text code from which each example is drawn (e.g. H4A, FUT, AEA). This is followed by another column with the concordance line, plus any additional text needed for the interpretation of the shell unit.
iii) The following four columns specify the genre, text type, domain and medium of each text.
iv) The remaining variables are displayed in eight columns: experiential structure, formal structure, syntactic function, participant type, Theme vs. Rheme, direction of encapsulation, antecedent and semantic type.

All variables except the genre-related ones (i.e. genre, text type, domain and medium) and semantic type are analysed on the basis of 187 individual codes[15] (see Appendix 3 for the list of codes). Three of the variables (syntactic function, participant type and Theme vs. Rheme), are assigned single codes (e.g. SCL or DO, standing for subject complement and direct object; PHE or IDR (rel.in), standing for Phenomenon and relational intensive Identifier; and T, *T, R, standing for Theme, marked Theme and Rheme). The 4 remaining variables (experiential structure, formal structure, direction of encapsulation and antecedent) are more complex in their categorisation, thereby forcing the use of combined codes (e.g. DC.NSP.PT.NSL^TG^QF, standing for non-specific partial Deictic^Thing^Qualifier). An example of a fully coded shell unit is given in (313) and in Table 5.1 below:

(313) Apart from the visual appearance of curtain fabric, and suitability for its purpose, **the other important characteristic to consider** is its drapability.

[15] For text type, domain, medium and semantic type, analytical categories are given in full (e.g. Context-governed for text type or Mental Volitional Detached 'Purpose' for semantic type). Genre, by contrast, uses Lee's (2001) 70 abbreviated genre categories (e.g. S:meeting for meeting or W:fict:prose for fictional prose).

Table 5.1 The complete analysis of a shell unit

Genre	Text type	Domain	Medium
W:misc (i.e. Written miscellanea)	Written books and periodicals	Informative: Leisure	Book

Experiential Structure	Formal Structure
DC.SP.DM.DV^PDC.EL.ID^EP.IP^TG^QF (i.e. Specific demonstrative determinative Deictic^Elaborating identity post-Deictic^interpersonal Epithet^Thing^Qualifier)	DF.AR^NUM.GO^AJ^H^TI.CL (i.e Definite article^general ordinal numeral^adjective^head^*to*-infinitive clause)

Syntactic Function	Participant Type	Theme vs. Rheme
SB (i.e. Subject)	IDD (rel.in) (i.e. Relational intensive Identified)	R (i.e. Rheme)

Direction of Encapsulation	Antecedent
AF.INTRA&CF.INTRA&AF.INTRA&CF.INTRA (i.e. Intrasentential anaphora&intrasentential cataphora&intrasentential anaphora&intrasentential cataphora)	LC.SNP(SI)&LC.CL(SI)&LC.CNP&LC.SNP (i.e. Local simple noun phrase: specifics of identity [i.e. *curtain fabric*] & local clause: specifics of identity [i.e. *to consider*] & local complex noun phrase [i.e. *visual appearance of curtain fabric*, *suitability for its purpose*] & local simple noun phrase [i.e. *its drapability*])

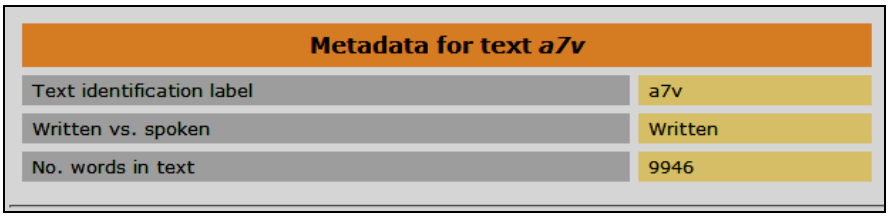
Semantic Type
Factual Attributive Part-whole 'Aspect'

5.3 THE VARIABLES

5.3.1 *Genre*

In view of the lack of genre information in the *BNC Sampler*, it was decided to turn to *BNCweb* for the genre category of each text. As explained in 4.2.2, *BNCweb* classifies texts according to Lee's (2001) 70 genre categories (see Appendix 1), whilst the *CQPweb Sampler* interface gives only mode information (i.e. written vs. spoken). Figures 5.1 and 5.2 show the amount of metadata detail that each corpus interface gives for a particular text (i.e. A7V). From the wide range of textual metadata offered by *BNCweb*, three categories were chosen for inclusion in the spreadsheet: genre, text type, domain and medium. Sociolinguistic categories such as age, sex or level of difficulty are not considered here.

Figure 5.1 Metadata for text A7V in *CQPweb* (*BNC Sampler*)



Metadata for text *a7v*	
Text identification label	a7v
Written vs. spoken	Written
No. words in text	9946

Figure 5.2 Metadata for text A7V in *BNCweb* (*BNC*)

BNC header information for file A7V	
Title:	The Guardian, electronic edition of 1989-11-08: Foreign news pages. Sample containing about 8802 words from a periodical (domain: world affairs)
Spoken or Written:	Written
Number of Words (tagged items):	8,870
Average sentence length (<w>-tags per <s>-unit):	20.0226
Derived text type:	Newspapers
Genre:	W:newsp:brdsht_nat:report
Text type:	Written books and periodicals
Publication date:	1985-1993
Age of Author:	unknown
Domicile of Author:	unknown
Sex of Author:	unknown
Type of Author:	Multiple
Age of Audience:	Adult
Text Domain:	Informative: World affairs
Perceived level of difficulty:	Medium
Medium of Text:	Periodical
Place of publication:	UK: South (south of Bristol Channel-Wash line)
Text Sample:	unknown
Estimated circulation size:	High
Target audience sex:	Mixed

5.3.2 *Experiential and formal structure*

5.3.2.1 *Experiential structure*

The semantic or experiential structure of shell-noun phrases is analysed on the basis of Halliday & Matthiessen's (2004: 312–35) framework. In SFG, noun phrases (or, in their terminology, 'noun groups') are built around a nucleus or Thing (the head noun), which may be accompanied by up to six semantic constituents, five preceding the Thing (Deictic, post-Deictic,

Numerative, Epithet and Classifier) and one following it (Qualifier). A description of each component is given below:

i) Deictics: Deictics are the elements known traditionally as 'determiners'. In SFG, a twofold distinction is made between 'specific' and 'non-specific' Deictics. The former comprises demonstrative and possessive determiners (e.g. *this*, *that*, *my, your*), whilst the latter ranges from the indefinite article *a/an*, through items like *each*, *every*, *either* or *all*, to the assertive, non-assertive and negative determiners *some*, *any* and *no*. It should be noted that in SFG, the definite article *the* is a demonstrative specific Deictic, along with *this*, *that*, *these* and *those*.

ii) Post-Deictics: In addition to a Deictic element, noun groups may contain a post-Deictic to highlight the familiar or unfamiliar status of an item in the text or the way it compares to other items. As such, post-Deictics narrow the reference of the head noun further. Post-Deictics fall into six main groups, which are then subdivided into ten sub-categories. The six main categories are 'elaborating' (e.g. *same*, *identical*, *particular*, *various*), 'extending' (e.g. *complete*, *entire*, *whole*), 'enhancing' (e.g. *above*, *aforementioned*, *similar*, *different*), 'modality: modalisation' (e.g. *certain*, *possible*, *habitual*, *normal*), 'modality: modulation' (e.g. *necessary*, *required*, *intended*, *desired*) and 'report' (e.g. *alleged*, *so-called*, *hypothetical*, *purported*). The noun phrase in example (314) contains one such post-Deictic, i.e. *next*. Whilst *next* is treated as a spatial and temporal post-Deictic in SFG (i.e. PDC.EN.SPA-TM), on account of its anticipating the following discourse segment, it is analysed as a general ordinal in Quirk et al. (1985: 262) (along with other SFG post-Deictics like *last*, *following* or *other/another*; i.e. NUM.GO).

(314) [...] if you look at the next examples, forty four <unclear> we are told here were in favour of the <unclear> strategy and thirty six per cent opposed. (*BNC Sampler*: JJA, S:meeting)

Many of Halliday & Matthiessen's (2004) post-Deictics correspond to a group of adjectives whose occurrence is restricted to pre-head positions. This is why both Quirk et al. (1985: 428–32) and Pullum & Huddleston (2002: 555–9) refer to them as 'attributive-only adjectives'. Quirk et al. (1985: 428–32) distinguish between two main classes of attributive-only adjectives: 'intensifying adjectives' and 'restrictive adjectives', the former further broken down into 'emphasizers', 'amplifiers' and 'downtoners'. Emphasizers stress

qualities of nouns (e.g. *a **clear** failure*, *__plain__ nonsense*, *the **simple** truth*), while amplifiers widen the semantic scope of nouns (e.g. *a **complete** victory*, *__great__ destruction*, *a **complete** fool*) and downtoners do the opposite (e.g. *a **feeble** joke*, *a __slight__ effort*). Telling an emphasizer from an amplifier sometimes depends on whether the adjective allows predicative position with no marked change in meaning. In that case, the adjective is probably an amplifier, as in *complete victory* (i.e. *the victory is complete*). Restrictive adjectives, rather than intensify the meaning of the head noun, restrict its reference (e.g. *main*, *chief*, *certain*, *particular*, *only*). For example, in *the **main** problem* or *a __particular__ problem*, the focus is not on any problem, but on a specific one.

Pullum & Huddleston's (2002: 555–9) classification of attributive-only adjectives, with seven categories, is more detailed than Quirk et al.'s (1985). It consists of 'degree and quantifying attributives' (e.g. *a **complete** fool*, *a __real__ help*), 'temporal and locational attributives' (e.g. *his **former** wife*, *its **ultimate** demise*), 'associative attributes' (e.g. *a __lunar__ landing*, *__musical__ analysis*), 'process-oriented attributives' (e.g. *a __big__ eater*, *a __slow__ learner*), 'modal attributives' (e.g. *an __apparent__ discrepancy*, *the __likely__ benefit*), 'particularising attributes' (e.g. *a __particular__ area*, *the __chief__ reason*) and 'expressive attributives' (e.g. *my __dear__ mother*, *the __wretched__ car*).

Quirk et al.'s (1985) distinction between emphasizers, amplifiers, downtoners and restrictive adjectives is used to endow SFG post-Deictics with greater detail. When a post-Deictic is observed to also fall into any of these four categories, the code EZ (i.e. emphasizer), AM (i.e amplifier) or RT (i.e. restrictive)[16] is added to the post-Deictic tag. This is because a failure to account for these categories would miss the intensifying or restrictive role of quite a few so-called post-Deictics. Pullum & Huddleston's (2002) classification could have been used for the same purpose, but their categories overlap with SFG's or Quirk et al.'s (1985). This is shown in the correspondence between Pullum & Huddleston's (2002) degree and particularising attributes and Quirk et al.'s (1985) emphasizers, amplifiers and restrictive adjectives. The five other categories have SFG equivalents. Temporal and locational attributives relate to space-time post-Deictics (as in (314)), and modal attributes correspond to the two SFG modality post-Deictics. Finally, associative, process-oriented and expressive attributes fall into

[16] No examples of downtoners were found in the corpus evidence analysed.

three experiential SFG premodifiers: Classifiers, experiential Epithets and interpersonal Epithets.

Examples (315) to (317) illustrate the use of an emphasizer, an amplifier and a restrictive adjective in three shell units. *Express* in (315) is an attributive-only adjective meaning 'clear and definite' (*LDCE*). According to the *OED*, it is used to make '[...] definitely formulated' statements. In view of its absence from Halliday & Matthiessen's (2004) list of post-Deictics, a decision was made to classify *express* as a report locution post-Deictic (e.g. *alleged*, *so-called*) on account of its linguistic meaning (as shown in the *OED*). It was then further decided to treat it as an emphasizer, based on its attributive-only position and its stress on the definite nature of the recommendation. The adjective is thus tagged as PDC.RP.LN.EZ. *Total* in (316) is analysed as an extending post-Deictic based on its similarity to *complete*, *entire* or *whole.* Its treatment as an amplifier rests on the fact that, whilst it is rare to find *these failures were total*, the meaning of the adjective ('complete, or as great as is possible', *LDCE*) does not just emphasize the nature of the failures, but points to the degree in which such failures are apparent (not partial but complete failures). The adjective is thus tagged as PDC.ET.AM. *Particular* in (317) is a typical instance of an elaborating exemplification post-Deictic (e.g. *certain*, *given*, *various*, *different*, etc.) and of a restrictive adjective. Its restrictive meaning is evident in its definition: 'a particular thing or person is the one that you are talking about, and not any other' (*LDCE*). It is tagged as PDC.EL.EM.RT.

(315) Decisions to terminate legal assistance under rule twenty seven are different from other decisions made within the union, because they are invariably made on the express written recommendation of a specialist in law. (*BNC Sampler*: HLW, S:speech:scripted)

(316) There are short sessions which are total failures — there is a total mismatch between the user and the system. (*BNC Sampler*: H0S, W:misc)

(317) The particular challenge for donors has been how best to support the fragile but vital process <pause> of transition to a pluralist democracy and market economies in the region. (*BNC Sampler*: JNG, S:meeting)

iii) Numeratives: Two sub-classes of Numeratives are described in SFG: 'quantitative' (e.g. *one*, *two*, *few*, *much*) and 'ordinative' (e.g. *first*, *second*, *preceding*, *subsequent*), corresponding to the traditional distinction between cardinal and ordinal numerals. Each

is in turn subdivided into 'definite' (e.g. *one, two, a couple of*, *first*, *second*) and 'indefinite' (e.g. *little*, *a bit of*, *several*, *many*, *more*).

iv) Epithets: Epithets are used to attribute qualities to head nouns. SFG differentiates between 'experiential' and 'interpersonal' Epithets, the former expressing objective qualities (e.g. *old, long, tall*) and the latter conveying the speaker's or writer's subjective attitudes (e.g. *splendid, fantastic, stupid*). Halliday & Matthiessen (2004: 319) are aware of the difficulty in distinguishing one from the other, arguing that even typical experiential Epithets like *little* or *old* may come to function as interpersonal Epithets given a particular discourse situation. In their view, '[...] expressions of attitude tend to be strung prosodically throughout the clause [...]', which implies that '[...] very few words [...] serve only an atttitudinal function'.

Some tests are offered to distinguish between experiential and interpersonal Epithets. One such test rests on the assumption that experiential Epithets objectively define the head noun (i.e. they are '[...] potentially defining [...]', Halliday & Matthiessen 2004: 319), while interpersonal Epithets assign a quality that is not inherent in the noun itself. Although such a test may be easily applied to first-order entities (as in *the long train* vs. *the mighty train*), when it comes to the analysis of second- and third-order unspecific shells, it is not so much a matter of a quality being inherent in the noun, but of a quality being either more or less objective depending on the surrounding discourse. In examples (318) and (319), *new* and *short* are factual qualities indicating that an argument (*that mothers have to go out to work*) is not recent and that an answer is not detailed enough. This is why they are both treated as experiential Epithets.

(318) I know mothers have to go out to work to help with the bills but this is no new thing. (*BNC Sampler*: CF9, W:newsp:other:sports)

(319) CONFUSED shareholders have been seeking help on the complex question of Eurotunnel warrants and what they are worth. Not much, is the short answer. (*BNC Sampler*: CEL, W:newsp:other:commerce)

In (320) and (321), *remarkable*, *superficial* and *subjective* function as interpersonal Epithets. The evaluative meaning of *very remarkable* is evident from its occurrence at the end of an impassioned courtroom plea. However, the interpersonal treatment of *superficial* and *subjective* could be open to question. A case may be made for the inherent objective meaning of both adjectives (i.e. assessments with the inherent qualities of superficiality and subjectivity), but the belief here is that the kind of negative semantic prosody entailed in the italicised segments (*any fool*, *less sanguine*,

inevitably, *prejudice*, *irrationality*) somehow imposes an interpersonal interpretation on both adjectives.

(320) Indeed yesterday you may well seen on on the news, heard on the radios, seen in the papers, forget such attacks. t's also not about macho, gun <unclear> , gung ho policeman as has been suggested <unclear>. What it is about, members of the jury, is a very professional police force doing the best they could in the circumstances. <pause> And what do they try to do, they were trying, above all else, to protect you and I, the general public. And this is a police force, you will remember, Mr <gap desc='name' reason='anonymization'> telling you <pause> <unclear> very remarkable characteristic. (*BNC Sampler*: JJV, S:courtroom)

(321) Of course *any fool* can write down 1,2,3,4 and even 5,6,7,8 opposite the names of candidates on a ballot paper. Whether he would put those figures in the same places tomorrow is perhaps open to question. To the extent that rank-ordering means giving preferences to candidates of one party rather than another it is both *easy and reasonable*. *One may be less sanguine* about the average voter 's capacity to rank candidates , whether of one preferred party or more than one , according to their relative merit and aptitude. That exercise, *inevitably* based on superficial and subjective assessments, allows *too much scope for prejudice and irrationality*. (*BNC Sampler*: EW4, W:non_ac_polit_law_edu)

Further to the above test, Halliday & Matthiessen (2004: 319) claim that there is a tendency for interpersonal Epithets to occur before experiential ones. *Considerable* in example (322) is a case in point:

(322) I agree that Gloriana_was a considerable challenge. (*BNC Sampler*: J55, W:non_ac:humanities_arts)

LDCE defines it as 'fairly large, especially large enough to have an effect or be important'. It is synonymous with such other adjectives as *big*, *great* or *significant*. A search for the sequence *considerable _{ADJ}* in *BNCweb* reveals that *considerable* often occurs before objective classifying adjectives, such as *financial*, *political*, *social*, *economic*, *local* and *regional*. A search for *_{ADJ} considerable* shows that in 57.3% of hits, *considerable* is preceded by the assertive determiner *some.* This is in turn followed by *any*, *this* and *that*. The reason why such a query returns a mixture of determiners (e.g. *some*, *any*, *this*, *that*) and adjectives (e.g. *exciting*, *obscure*, *perceived*, *consequent*) is unknown. However, the argument remains that *considerable* tends to precede adjectives denoting

inherent objective qualities and that it is often preceded by Deictics, and only rarely by other Epithets.

In view of the emphasising meaning of the adjective, the possibility was raised of treating the adjective as an intensifying post-Deictic (emphasizer or amplifier). However, *considerable* does not express an upward intensification from an assumed norm (as in *great destruction*, *utter folly* or *total irresponsibility*), but something that is important enough to be 'worthy of consideration or regard' (*OED*). Its role, therefore, is not so much to amplify the challenge as to highlight the difficulty of the task or challenge at hand. Nevertheless, it cannot be analysed as an emphasizer, as emphasizers are attributive only and, in this particular instance, one might also say something like: *the challenge of staging Gloriana was considerable*[17]. Be that as it may, whether examples such as *considerable*, *significant* or *important* are treated as interpersonal Epithets rather than as intensifying post-Deictics does not invalidate the idea that both intensifying and interpersonal Epithets have a similar evaluative import. They only differ in that, whilst intensifying post-Deictics involve different degrees through which speakers and writers produce and perceive discourse entities (e.g. *total*, *complete*, *clear*), interpersonal Epithets lay emphasis on how they actually feel about such discourse entities (e.g. *awful*, *horrifying*, *funny*). The line separating intensifying post-Deictics from Epithets is fuzzy in many respects. For the purposes of this study, a decision was made to classify as intensifying post-Deictics either those instances occurring in Halliday & Matthiessen's (2004), Quirk et al.'s (1985) and Pullum & Huddleston (2002)'s lists of post-Deictics and attributive-only adjectives or any other adjective synonymous with those in any of the aforementioned lists.

Examples (323) and (324) show how an intensifying post-Deictic and an Epithet analysis may apply to one and the same adjective (in this case, *clear*), thereby serving as further evidence for the blurred boundary between both categories. In (323), the meaning of *clear* is that of '[...] obvious and impossible to be mistaken about' (*CCD*). It could be replaced with such other adjectives as *evident*, *obvious* or *patent*, which fall into report, idea post-Deictics. Its semantic contribution to the head noun is therefore not descriptive,

[17] A *BNCweb* query of *_{N} {be} considerable* (i.e. any form of a noun + any form of the verb *be* + *considerable*) returns 236 hits for such a sequence (e.g. *the obstacles/advantages/effects are considerable*), thereby proving a greater mobility than that of clear emphasising adjectives like *definite* (only 16 hits), *sheer* (only 5 hits) and *outright* (no hits).

as it does not assign the quality of clarity to *suspicion*. Rather, it lays emphasis on the existence of one such suspicion and, as such, *clear* in (323) is tagged as PDC.RP.IA.EZ (i.e. report, idea post-Deictic: emphasizer). In (324), *clear* conveys a different meaning. Substituting *obvious*, *evident* or *patent* for *clear* would not make much sense in this particular discourse situation. *Clear* here is more in line with 'something that is [...] easy to understand, see, or hear' (*CCD*). As such, its meaning is closer to one of the senses of *definite*, i.e. 'clearly known, seen, or stated' (*LDCE*). Hence, what is at issue here is the quality of clarity, as it applies to a specific idea. Therefore, the adjective is treated as an experiential Epithet.

(323) Ideally, all the above qualifications would need to be taken into account in any attempt to prescribe policy. The predictions of economic theory are not sufficiently clear-cut to permit us to proscribe monopoly outright. Theory does point to a clear suspicion that a lack of competition can, most certainly, lead to inefficiencies, but it also identifies possible benefits from the attainment of lower-cost production. (*BNC Sampler*: HXN, W:commerce)

(324) 'The BBC must therefore have a clear vision if it is to retain its role as the cornerstone of British broadcasting and continue to command respect and admiration in Britain and throughout the world. 'In January John Birt laid out that vision of a wide range of high quality programmes, greater efficiency and accountability with value for money for licence payers. (*BNC Sampler*: CF6, W:newsp:other:report)

v) Classifiers: Classifiers bear the closest relation to the head noun, both semantically and structurally. They are used to assign the noun to particular objective sub-classes, differing from Epithets in their failure to accept intensification (e.g. *slightly*, *very*, *highly*) and comparison, as well as in their occurring closest to the head noun (following Deictics, post-Deictics and Epithets). Unlike Epithets, Classifiers comprise both adjectives and nouns (e.g. ***parental*** *rights*, ***tonal*** *excellence*, ***foundation*** *stone*, ***restraint*** *chairs*).

Examples (325) and (326) contain two shell-noun phrases with Classifiers. In (325), *scientific* is analysed as a Classifier on account of its failure to allow for the comparative form in this particular sense (i.e. 'about or related to science, or using its methods', *LDCE*)[18]. Thus, a phenomenon may be scientific or non-scientific, but it cannot be more or less so. Additional evidence for its

[18] The *related to* paraphrase is a useful test for the identification of Classifiers.

classifying role is its occurrence after *curious*, treated here as an interpersonal Epithet. Example (326) contains two Classifiers, one nominal (i.e. *Community*) and the other adjectival (i.e. *economic*). In both cases, the Classifier accepts a *related to* paraphrase (i.e. *economic*: objectives related to the economy; *Community*: related or applying to the European Community). Besides, while *broad* can be intensified and compared (e.g. *very broad*, *broader objectives*), *Community* and *economic* are more static, not allowing any of these transformations (e.g. **very/more economic objectives*). Thus, *broad* is analysed as an experiential Epithet and *Community* and *economic* as Classifiers.

(325) ROS: (Cutting his fingernails) Another curious scientific phenomenon is the fact that the fingernails grow after death, as does the beard. (*BNC Sampler*: FU6, W:fict:drama)

(326) One problem will undoubtedly be the growing demands of the European Parliament, backed by Chancellor Kohl and President Mitterrand, to be given a bigger say with the Council of Ministers in setting broad Community economic objectives and in holding a future EC central bank to account. (*BNC Sampler*: A9E, W:newsp:brdsht_nat:report)

vi) Qualifiers: Qualifiers comprise any postmodifying structure, ranging from phrasal postmodification (prepositional and appositive noun phrases) to clausal postmodification (e.g. relative clauses, *-ed* clauses, noun complement clauses). The relative clause and the prepositional phrase in (327) and (328) respectively are two examples of Qualifiers.

(327) Even so, in countless places around Europe, people behave and think in ways which are totally distinctive. (*BNC Sampler*: A8W, W:newsp:brdsht_nat:report)

(328) A useful system for handling practical program equivalence questions must be able to deal with programs containing loops. (*BNC Sampler*: G3N, W:ac:tech_engin)

5.3.2.2 *Formal structure*

From the above, it is clear that SFG's semantic description of the noun phrase provides useful insights into the meaning distinctions applying to various kinds of premodifiers (i.e. post-Deictics, Epithets and Classifiers). However, when it comes to the analysis of Deictics and Qualifiers, SFG only scratches the surface of all that is involved in these areas. This explains why formal structure is also included in the analysis. For such a variable,

the analysis uses the categories in Quirk et al. (1985: 245–331; 1238–352). Given that a full description of their framework would take up much space, the following outlines some of the formal realisations of the three phrase-level functions of determiner, premodifier and postmodifier:

i) Determiners: Quirk et al. (1985: 253–65) distinguish between central, pre- and postdeterminers, the latter two occurring before and after central determiners respectively. 'Central determiners' comprise, for example, the definite and indefinite articles, the possessive determiners (e.g. *my*, *your*), the demonstrative determiners (e.g. *this*, *that*) and the assertive and non-assertive determiners (i.e. *some*, *any*). 'Predeterminers' consist of such items as *all* (universal determiner), *both*, *half*, multipliers (e.g. *double*, *twice*), fractions (e.g. *one-third*, *one-fifth*), *such* and *what*. Finally, 'postdeterminers' involve cardinal and ordinal numerals (e.g. *two*, *three*, *second*, *third*) and quantifiers, which are divided into closed-class (e.g. *many*, *few*, *much*, *several*, *a little*) and open-class ones (e.g. *plenty of*, *a lot of*, *a number of*, *a great deal of*).

ii) Premodifying structures: Head nouns may be preceded by adjectives (e.g. *the* ***deplorable*** *example*), nouns (e.g. *a* ***membership*** *application*) or genitive noun phrases (e.g. ***the queen's*** *forebodings*) (Quirk et al. 1985: 1321–46). In the former two cases, the coding in this study takes account of whether the premodifying noun or adjective occurs as a single item or as part of a larger noun or adjective phrase. For example, the premodifier in ***tunnel*** *myths* and *such* ***terrible*** *things* is tagged as N (i.e. noun) and AJ (i.e. adjective), respectively. However, in *the* ***Channel Tunnel*** *project* and in *two* ***very comprehensive*** *answers*, *Channel Tunnel* and *very comprehensive* are annotated as NP (i.e. noun phrase) and AJP (i.e. adjective phrase). The tag for instances of genitive noun phrases is DF.DV.GV.NP (i.e. definite determinative genitive noun phrase). It is important to note that genitive noun phrases are treated in SFG as possessive determinative Deictics, on a par with possessive determiners like *my*, *his* or *their*.

iii) Postmodifying structures: Postmodification may be realised by both clausal and phrasal constituents (Quirk et al. 1985: 1244–321). Clausal structures comprise restrictive and non-restrictive relative clauses (as in (327)), appositive clauses (both *that*- and *to*-infinitive) (as in (329a) and (329b)) and non-finite clauses (*-ing*, *-ed* and *to*-infinitive, as in (330a) and (330b)). Phrasal structures include prepositional phrases (as in (328))), adjectives (as in (331)) and nominal appositives (as in (332)). In the case of prepositional

phrases, the coding in this study indicates the preposition introducing the structure, followed by the formal realisation of its complement (i.e. *n*: noun or *-ing*). As such, in (328) above, the prepositional phrase is annotated as PP(*for -ing*). As regards nominal apposition, the coding rests on Quirk et al.'s (1985: 1300–21) categorisation of this postmodifying structure. Thus, in (332), *the Muldergate affair in the late 70s* is tagged as FL.ST.NR.IT.AP (i.e. full strict non-restrictive identification appositive). It is treated as an identification appositive on account of the possibility of replacing the initial dash with *that is*, *namely* or *i.e*. In addition, as one of the two nouns may be omitted with no change in meaning, and as the appositive phrase is a noun, the example is further analysed as full and strict.

(329) (a) A poll showed that Mr Da Silva won it, burying the myth that as an ex-factory worker with only a few years schooling he is unqualified for the presidency. (*BNC Sampler*: A9M, W:newsp:brdsht_nat:report)
(b) Summit offers chance to improve British image in EC (*BNC Sampler*: A95, W:newsp:brdsht_nat:report)

(330) (a) This is a collaborative venture, involving three publishers (OUP, Longman, Chambers), the British Library, and two universities (Oxford and Lancaster). (*BNC Sampler*: F98, W:ac:humanities_arts)
(b) Overcoming a natural resistance to change is a challenge faced by many companies that want to progress. (*BNC Sampler*: J24 W:misc)

(331) Perhaps lent the final impetus necessary for successful reception by the proximity of the area of stimulus to that of cognition, this hair-pulling signal finally seemed to awaken Gavin to the perception that there might be something else going on [...] (*BNC Sampler*: G0A, W:fict:prose)

(332) In the last government scandal — the Muldergate affair in the late 70s — Mr Botha, was forced to fire a judicial inquiry because the judge refused to cover up corruption. (*BNC Sampler*: A9E, W:newsp:brdsht_nat:report)

5.3.2.3 *Semantic Thing and formal head: some mismatches*

To conclude section 5.3.2, special mention should be made of the mismatch existing between semantic Thing and structural head in certain constructions (Halliday & Matthiessen 2004: 333). In one such structure, the formal head is filled by a non-specific Deictic (e.g. *each*, *none*, *some*, *any*) or a Numerative (definite: e.g. *one*, *two*, *three*; indefinite: e.g. *many*, *much*, *few*), followed by an *of*-prepositional phrase containing the actual semantic head or Thing. Quirk et al.'s (1985: 380) grammar accounts for this construction on the basis of an indefinite pronoun followed by what they

consider a 'partitive *of*-phrase'. Examples (333) and (334) are two shell-noun instances of this structure. In (333), *many* is, in SFG's terms, an indefinite quantitative Numerative, which is not the semantic nucleus of the noun group (it being *projects*). On these grounds, the semantic tag for this example is NUM.IN.QV^DC.SP.DM.DV^TG. Structurally, the noun phrase is tagged as QT.PN^PV.PP(*of* DM.DT(*THESE*)^H), with *many* treated as a quantifying pronoun, followed by a partitive prepositional phrase whose head is the shell unit under scrutiny (i.e. *projects*). The analysis of (334) is essentially similar to that of (333), the only difference lying in the treatment of *some*. Semantically, *some* is a partial, selective, non-specific Deictic (DC.NSP.PT.SL), whilst syntactically, it is an assertive pronoun (AS.PN).

(333) The latest roads White Paper also includes further improvements to the M20, M23 and M25 but, despite government streamlining of the planning and construction process, it is unlikely that many of these projects will be completed before the year 2000. (*BNC Sampler*: BMJ, W:misc)

(334) Let me just mention some of the ways in which you, the volunteers, are getting together, enriching your skills and supporting each other generally in support of the seventy fifth. (*BNC Sampler*: JNG, S:meeting)

Difficulties in the identification of heads or Things also arise in a group of constructions that Keizer (2007: 109–84) terms 'pseudo-partitive'. Anyone presented with a noun followed by an *of*-prepositional phrase would argue that the first noun is the head of the noun phrase. This is often the case, but there are instances where the semantic, pragmatic and syntactic focus is on the rightmost noun (i.e. N_2). In these cases, the noun phrase preceding *of* (i.e. N_1) is treated as a complex quantifier or qualifier. Examples (335) and (336) are two cases in point. In (335), the emphasis is not on the number of such practices (e.g. three, four, etc.), but on the fact that there are *many* (i.e. *a substantial number of*) and *various* (i.e. *a variety of*) practices. The underlined segment specifies examples of such practices, rather than their actual number. As such, *a substantial number and variety of* is treated as an indefinite quantitative Numerative (NUM.IN.QV) in experiential structure, and as a quantifier (QT) in formal structure. However, in example (336) the focus is on the number of such applications. This is borne out by the occurrence of the definite article (cf. the indefinite article in (335)) and by the singular agreement of *is* in *is reasonably standard* with the singular N_1 (i.e. *number*). Unlike *number* in (335), where it is part of a pre-head quantifying expression, in (336), it is '[...] used referentially' (Keizer 2007: 112). This implies that subsequent reference to the noun phrase will be made to N_1 rather than to N_2. Accordingly, *the number of* is not part of the shell-noun

phrase *training applications* and, as such, its tagging in experiential structure and formal structure is as follows: CS^TG (Classifier: *training*, Thing), (ING)N^H (*-ing* noun + head). In this instance, the existence of N_1 is acknowledged in the variable syntactic function, where the tag CP.*OF*(<SB) indicates that the shell-noun phrase is the complement of the preposition *of* which postmodifies the noun phrase *the number,* whose overall syntactic function is that of subject.

(335) The Commission's record on attacking cartels and concerted practices is impressive, for it has vigorously pursued and successfully secured the termination of a substantial number and variety of concerted practices. In particular, it has taken a strong line against price fixing (e.g. dyestuffs , glass containers) and market sharing or quantity agreements (e.g. cement) [...] (*BNC Sampler*: HXN, W:dcommerce)

(336) Okay erm yeah a-- again related to special needs is the number of training applications
Yeah yeah.
erm <-|-> which I think we can assume <-|->
<-|-> <unclear> <-|->
is reasonably standard. (*BNC Sampler*: H5D, S:meeting)

Pseudo-partitives are not only restricted to expressions like *a number of* or *a great deal of*. They are also evident among such other expressions as *a bit of*, *an item of*, *a cup of*, *a type of*, *a kind of* and *a sort of*. The three former examples would, according to Quirk et al. (1985: 249–50), qualify as 'quantity partitives', while the latter three would fall into 'quality partitives'. From their claim that such instances as *a nice kind of coffee* or *two lumps of sugar* may be reclassified as *a nice coffee* and *two sugars*, it may be inferred that, in their view, it is N_2 that constitutes the head of the construction. In Halliday & Matthiessen (2004: 333), N_2 is similarly treated as the semantic nucleus or Thing of the construction, with N_1 forming part of either a quantity or a quality 'extended Numerative'. The use of such a term, however, could prove confusing, as it brings together cases of indefinite Numeratives like *a number of* or *a bit of* (which may be replaced by single-word quantifiers like *some*, *many* or *little*, as in (335)) with more specific quantifying and qualifying partitives like *a cup of* and *a kind of*. For a better descriptive category for the semantic analysis of these structures, it was decided to turn to Matthiessen's (1995: 655–7) 'Facets'.

Such a concept features as one of the sub-classes of quality extended Numeratives in Halliday & Matthiessen (2004: 333), its use being restricted to the physical parts of various first-order entities, as in *the front of the house* or *three sides of a square*. In Matthiessen (1995: 655–7), by contrast, Facet applies to a range of nouns that roughly correspond to Quirk et al.'s

(1985: 249–50) quantity and quality partitives. They comprise five sub-types: 'symbolization' (e.g. *picture*, *symbol*, *sign*), 'exemplification' (e.g. *kind*, *type*, *class*), 'aggregate' (e.g. *set*, *collection*, *group*), 'measure' (e.g. *cup*, *glass*, *jug*) and 'partitive' (e.g. *bit*, *piece*, *top*). Thus, a decision was made to tag any pseudo-partitives of the type of *piece of* and *kind of* as quantity or quality partitives (QL.PV, QV.PV) in formal structure, and as facets (FC) in experiential structure. For instance, in (337), *sort of* is tagged as FC.EL.EM (i.e. Facet elaborating exemplification) in experiential structure and as QL.PV (i.e. quality partitive) in formal structure. *Examples* is in both cases the nucleus of the noun phrase, i.e. Thing and head.

(337) And what do y-- what sort of examples can you give about discipline then, <unclear> harsh discipline? What what what sort of things went on? (*BNC Sampler*. FYJ, S:interview:oral_history)

As Keizer (2007: 151) argues, whilst the distinction between pseudo-partitive and referential expressions is well-founded in theory, in practice '[...] the classification of authentic examples may not always be easy'. The corpus evidence used in this study reveals a cline that ranges from clear instances of partitive and facet expressions, through instances which are less clearly so, to clear examples of non-partitive referential expressions. With regard to partitive cases, these are found to share a number of characteristics:

i) N_2 is the semantic and syntactic head of the construction.
ii) N_1 is Facet and quantity or quality partitive.
iii) N_2 tends to be uncountable (only applying to quantity partitives).
iv) The initial determiner accompanies N_2.
v) N_1 may be deleted with no major change in meaning.
vi) The main verb imposes semantic and formal restrictions on N_2, not N_1.

The three examples below contain quality partitives. At first glance, of the three instances, *all manner of* in (338) appears to be the clearest partitive, because it may be easily omitted or replaced by a quantifier like *many* or *several* (i.e. *all other crimes, many/several other crimes*). Besides, the occurrence of the non-specific Deictic *all* seems to be closely linked to a partitive interpretation (cf. also (335) and (337)). Matthiessen (1995: 657) explains such a tendency on the grounds that, the more specific the assumed Facet is, the more likely it is to be interpreted as an N_1-headed construction. Thus, (339) and (340), preceded by demonstrative specific *this*, seem more referential (or N_1-headed) than partititve (or N_2-headed).

However, the surrounding co-text, shows that the reader's attention is not directed to a *backcloth* or a *circle*, but to the existence of *evidence* and *terror*, which are lexically realised by the underlined segments. Hence, the belief here is that *with this backcloth of evidence* entails *with all this evidence*, while *this circle of terror* entails *this terror*. Further support for the partitive analysis of the latter lies in the semantic selection restrictions imposed by the verb *continue*. It makes sense to say that a frightening situation (i.e. *terror*) continues, but it is at least odd to say that a circle continues (it can either exist or not, but it cannot continue). Given the symbolic nature of *backcloth* and *circle*, both examples are semantically annotated as elaborating symbolization Facets (i.e. FC.EL.SM). By contrast, *All manner of* is treated as an elaborating exemplification Facet (i.e. FC.EL.EM), based on its similarity to *kinds, classes* or *types*.

(338) They say she procured her daughter's marriage to the king by sorcery and has committed all manner of other crimes.' (*BNC Sampler*: CCD, W:fict:prose)

(339) [...] there were weaknesses in the control and deployment of nursing staff. [...] apparently there had been little work done on the effects of skill mix on patient services and that 'a higher priority should be given by management to achieving the best value for money by the adoption of methods of allocating staffing resources more closely related to the needs of patients and ward objectives [...] With this 'backcloth of evidence', it was no surprise that the Personnel Director of the Management Board of the DHSS asked regional general managers to discuss the issues with regional nursing officers and district colleagues. (*BNC Sampler*: EVY, W:commerce)

(340) How long will this circle of terror continue?/Hunter and hunted, beater and beaten,/when will this madness end?' (*BNC Sampler*: A8W, W:newsp:brdsht_nat:report)

Quantity partitives, as shown in the examples below, behave similarly to quality ones: (341) and (342) best represent the category; (343) seems more doubtful.

(341) One last word of warning, if your pond does freeze over at any stage, don't smash the ice as the shock waves created will also kill the fish. (*BNC Sampler*: C9C, W:pop_lore)

(342) Well there is a list of things, list of points here, 4C does not apply to scan <unclear> 5, does not apply to scan T G I questionnaire. 9 Research approval forms for relevant to T G I space. 10 does not apply to T G I Survey. (*BNC Sampler*: J97, S:meeting)

(343) Indeed, your latest victim seems to have been a veritable bundle of surprises — not a maidservant but a duchess, not dead but alive —

> and, to top all, affianced to an eleven-year-old and not a virgin! (*BNC Sampler*: CCD, W:fict:prose)

In (341), *one last word* merely quantifies a warning and lays emphasis on the fact that one such warning is about to be presented. In (342), *a list of* may be said to be equivalent to *several* (i.e. *there is a list of points here>there are several points here*). Further evidence in support of this analysis lies in the possibility of turning *there is a list of points* into *there are a list of points*, with the plural existential verb agreeing with N_2 instead of with N_1. A *BNCweb* search for *there are a list of* returns 1 hit (vs. 13 hits for *there is a list of*), and the same search in *google.co.uk* returns 3 million hits (vs. 6 million hits for *there is a list of*). The fact that such a structure is used indicates that, more often than not, the assumed referential status of *list* becomes blurred in spontaneous discourse, thereby favouring a partitive analysis.

Regarding (343), the contention here is that *a veritable bundle of* is equivalent to *many* or *a lot of* (i.e. *your victim seems to have been/contained a lot of/many surprises*). Such an analysis draws on Keizer's (2007: 138) claim that a pseudo-partitive interpretation is highly likely in those cases where N_1 is preceded by an intensifying or downtoning adjective, as exemplified by *respectable* in *a **respectable** handful of private individuals* or *significant* in ***significant** lengths of time* (see also *substantial* in (335)). *Veritable* in (343) is one of these adjectives, as it is '[...] used to emphasize a description of someone or something' (*LDCE*). In this example, *veritable* does not premodify *surprises*, but *bundle*, as it emphasises the amount of such surprises, rather than the surprises themselves. On these grounds, *a veritable bundle of* is treated as a quantity partitive and as an extending aggregate Facet (i.e. FC.ET.AGG). The same Facet analysis would apply to *list of* in (342), while for (341) it makes more sense to treat *word of* as an extending partitive Facet (i.e. FC.ET.PV).

As stated above, the database also contains cases of non-partitive referential expressions. These examples share the following features:

i) The omission of N_1 causes a change in meaning (though to varying degrees).
ii) The semantico-pragmatic focus is placed on N_1.
iii) N_1 tends to be premodified by a descriptive adjective (i.e. an Epithet) (Keizer 2007: 139–40)
iv) The verb tends to agree with N_1 in number.
v) N_1 is treated as the Thing and head of the whole construction.

Examples (344) through (346) are three cases in point. In the three examples, the assumed partitive is preceded by an experiential Epithet (i.e. *non-exclusive*, *widespread*, *late*). Keizer (2007: 139–40) notes that in these instances, the construction is left-headed, with the '[...] N_1 denoting a specific object or entity complemented or modified by the following PP' (e.g. *half-filled* in *a* ***half-filled*** *mug of coffee* and *parallel* in ***parallel*** *sets of masculine and feminine numbers*). Further to this, the omission of N_1 in each case would somehow modify the intended meaning. For example, in (345), it is in spite of the phenomenon being widespread that there is no agreed explanation (not in spite of the phenomenon itself). Similarly, in (346), it is not the original myth that Plutarch presents, but a late form or his own version of such a myth. Additional support for the referential analysis lies in the number agreement of the verb with N_1, as shown in example (344) (i.e. *a list is provided*; cf. *list* in (342)). Referential cases like (344) through (346) follow the tagging policy of (336). As explained above, no mention is made of N_1 in experiential and formal structure, which means that no Facet or partitive analysis applies. The existence of N_1 is only identified in syntactic function, where the shell-noun phrase is annotated as being a complement of a preposition within a larger noun phrase with a specific syntactic function. In example (345), for example, the syntactic function assigned to *the phenomenon* is CP.*OF*(<CP.*IN SPITE OF*<AB.AT.CC) (i.e. complement of the preposition *of* < complement of the complex preposition *in spite of* < adverbial adjunct of concession).

(344) A non-exclusive list of examples of such abuses is provided [...] (*BNC Sampler*: CCD, W:fict:prose)
(345) Nonetheless, in spite of the widespread nature of the phenomenon, there is no agreed explanation of the causes [...] (*BNC Sampler*: FR2, W:ac:soc_science)
(346) Plutarch clearly presents a late form of the myth with some Greek influences by providing a very useful story outline. (*BNC Sampler*: EVR, W:non_ac: humanities_arts)

5.3.3 *Participant type, syntactic function and Theme vs. Rheme*

5.3.3.1 *Participant type: Hallidayan Transitivity*

This study uses Halliday & Matthiessen's (2004: 168–305) system of Transitivity for the analysis of the semantic roles of shell-noun phrases. In SFG, clauses are semantically structured around three components, i.e. the process, the participants involved in such a process and any additional circumstances. Six process types are accounted for in Halliday & Matthiessen (2004: 168–305):

i) Material processes: These are verbs used to present actions (i.e. doing) and events (i.e. happening) (e.g. *create*, *make*, *develop*, *grow*). The associated participants are Actor, Goal, Scope, Recipient and Client.
ii) Mental processes: These are verbs implying psychological states or processes (e.g. *like*, *see*, *fear*, *think*). The associated participants are Senser and Phenomenon.
iii) Relational processes: These roughly correspond to linking or copular verbs, which are used in the characterisation or identification of subjects. They fall into three main types, each of which is further divided into attributive and identifying: intensive (e.g. *be*, *become*, *seem*, *look*), circumstantial (e.g. *cause*, *last*, *depend on*, *concern*) and possessive (e.g. *have*, *lack*, *comprise*, *contain*). The associated participants are Carrier-Attribute (for the attributive subtypes) and Identified-Identifier (for the identifying subtypes).
iv) Behavioural processes: These verbs are mental or verbal in meaning, but material in their permitting the progressive form (i.e. *watch*, *listen*, *chatter*, *argue*). The associated participants are Behaver and Behaviour.
v) Verbal processes: These verbs are used to present acts of saying (e.g. *say*, *tell*, *ask*, *order*, *convince*). The associated participants are Sayer, Verbiage, Receiver and Target.
vi) Existential processes: This group comprises all those verbs used in the presentative construction *there is* (e.g. *exist*, *remain*, *follow*, *grow*). The participant associated with these processes is Existent.

As to Circumstances, SFG considers 9 main types, which are further subdivided into 22 subtypes. The nine main Circumstances are 'extent' (e.g. *for ten years*), 'location' (e.g. *in London*), 'manner' (e.g. *in a dignified way*), 'cause' (e.g. *thanks to him*), 'contingency' (e.g. *in spite of this*), 'accompaniment' (e.g. *with him*), 'role' (i.e. *in the shape of a dolphin*), 'matter' (e.g. *concerning this event*) and 'angle' (e.g. *according to him*).

It is important to stress that, whilst Hallidayan theory is useful for the semantic analysis of verbs and their associated elements, it contains inconsistencies. They are principally the result of the lists of examples given for each process type. Verbs are classified into one or other category with no explanation as to what sense of each verb is intended. For example, while the verb *provide* is classified as both material and identifying possessive relational, synonymous verbs like *give* or *present* are treated only as material (Halliday & Matthiessen 2004: 188, 246). Similarly, *talk* is classed as both behavioural and verbal, even when contradicting one of the tests applied to the identification of verbal processes, i.e. the possibility of

projection (*say that*, *tell somebody that* vs. **talk that*) (Halliday & Matthiessen 2004: 251, 255). A lack of consistency is also apparent in the variety of coding procedures followed. As O'Donnell et al. (2008: 47) note, '[...] SFL does not provide a single process type classification of any clause'. This is because, when faced with the analysis of Transitivity, two options arise: to draw on Hallidayan syntactic tests for the identification of process types, or to focus on the concept underlying the verb (O'Donnell et al. 2008: 56–7). For example, if syntactic criteria are drawn on, a verb such as *watch* is treated as behavioural, because it allows the progressive form and cannot project, but if conceptual criteria are employed, the verb implies sensory perception and is thereby analysed as mental.

The approach used here is in line with the latter option. Shell-noun phrases are assigned participant roles based on the context-sensitive interpretation of each verb or process type. Verbs are therefore classed as one process or another depending on the particular senses displayed in specific discourse situations. The identification of such senses rests on the surrounding co-text, and on the definitions given in any of the following dictionaries: *CCD*, *OALD*, *LDCE* and *OED*. Two examples of this type of analysis are (347) and (348). The meaning of *can be found* in (347) is 'exist in a place' (*LDCE*), which, according to *CCD*, is associated with the passive inflection of *find*. Given the equivalence of *further details can be found* to *further details exist* or *there are further details*, the shell-noun phrase is tagged as Existent (i.e. EXI). *Find* in (348) is semantically different, as it means '[...] discover, see, or get something that you have been searching for' (*LDCE*). *Finding a way of doing something* implies exploring, figuring out or identifying such a way. Considering the mental nature of such a sense of *find*, *ways of doing that* is here tagged as Phenomenon (i.e. PHE).

(347) Further details can be found in {11}. (*BNC Sampler*: J2J, W:ac:nat_science)

(348) [...] that is a luxury I think erm to be able to do it, if you want to write and earn money, which I do as well as, as socialize while I'm working erm you've got to find ways of, of, of doing that and erm just writing books and poetry just wouldn't I, I, I couldn't survive on that, I really do have to find other ways of, of earning money which is enjoying more [...] (*BNC Sampler*: F71, S:speech:unscripted)

The following subsections detail some of the decisions that had to be made in the analysis of each process type.

5.3.3.1.1 *Material or relational?*

If a first-order entity fills the subject or object slot of a material-like verb, a material analysis will apply. This is evident in (349) through (351). *The project* in example (349) is the object or semantic Goal of *abandon*, whose meaning here is clearly material (i.e. 'to stop doing something because there are too many problems [...]', *LDCE*). Its subject (*international banking consortium*), is neither a process nor a metadiscursive abstract entity, but a first-order entity meaning 'a group of companies [...] who are working together to do something' (*LDCE*). The subject in (350) is similarly first-order in meaning (*we*), with *our all-year capacity* thereby analysed as Goal.

However in (351) the shell-noun phrase is not Goal but Actor. According to *CCD*, *discredit* is applied to someone or something when something '[...] cause[s] them to lose people's respect or trust'. In this instance, *the scandal* initiates or triggers the loss of respect, which is actualised through the leader of this political movement (i.e. a first-order entity). It is thus similar to two examples given by Matthiessen (1995: 240): *the flood widened the river* and *they beat him blue,* where the objects are analysed as Goals. In both cases, an external entity (external force: *flood* or people: *they*) causes a change in a first-order entity (i.e. *making the river wider, leaving him bruised* or *damaging someone's reputation*, as in (351)).

(349) However, it seems unlikely that the international banking consortium will abandon **the project** given the large amounts of irretrievable money already committed. (*BNC Sampler*: BMJ, W:misc)

(350) The suggestion is not that Britain be turned into a glasshouse containing a mass of Disneylands and Mafia-run casinos; merely that we take a serious and professional look at extending **our all-year capacity** to give people what they want. (*BNC Sampler*: G0C, W:commerce)

(351) The leader of the Afrikaner Weerstandsbeweging (Afrikaner Resistance Movement) has been largely discredited in far-right circles by **the recent scandal** over his relations with a glamorous English-language newspaper columnist. (*BNC Sampler*: A7V, W:newsp:brdsht_nat:report)

If a material-like process is found to link two second- or third-order entities, the verb is analysed as a relational circumstantial process with a causative meaning. The shell-noun phrase in these instances is thereby treated as a circumstantial Identified or Identifier. This analysis follows Halliday's (1993: 61–5) claim that a verb should be coded as relational circumstantial provided that a link is established between two process-based grammatical metaphors. In examples such as (352) through (354), a *cause X to X* paraphrase shows how one process is causally related to the other. In (352), *decrease* implies that, should these organisms continue to grow, they might

not survive (i.e. *their growth would cause these entities' chances of survival to become smaller*). In (353), *dispel* conveys the idea that what people suspected would be proved wrong as soon as Mr Horn arrived in Bucharest (i.e. *his arrival would cause these suspicions to disappear*). Finally, in (354), *restrict* implies that what one group does may limit what another group does as part of their tradition (i.e. *what one group does may cause the practices of another group to become more limited*). In the three cases, the shell-noun phrase in boldface is tagged as relational circumstantial Identifier (i.e. IDR (rel.ci)).

(352) Crabs select the larger individuals from amongst first-year dog-whelks (Feare, 1970a) so that rapid growth at this stage may decrease **the individual's chances of survival**. (*BNC Sampler*: FU0, W:ac:nat_science)
(353) The arrival of Mr Horn in Bucharest was intended to dispel **some of these suspicions**. (*BNC Sampler*: AAT, W:newsp:brdsht_nat:report)
(354) Indigenous religions shall be tolerated, and even encouraged in our polytetheastic society. However, where the activities of one group restrict **the** freedoms, **practices** and conduct **of another**, such as events in Amritsar we shall clamp down — Sheikh/Sikh and ye shall be fined. (*BNC Sampler*: AP6, W:misc)

5.3.3.1.2 *Verbal or relational?*

Typical instances of verbal processes are those where a conscious participant (i.e. the Sayer) communicates something (i.e. the Verbiage). This is shown in (355) and (356), where the shell noun functions as Verbiage.

(355) Maybe if I tell her **one of my pooh jokes** she'll throw us out! (*BNC Sampler*: CHR, W:misc)
(356) [...] we're just going to mention **one or two other points on erm our customer base**. (*BNC Sampler*: HDF, S:speech:unscripted)

There are also examples where the subject of a verbal-like process is not a person, but a non-conscious (first-, second- or third-order) entity. This is illustrated in (357) through (359), with *reports* and *myth(s)* as subjects of the projecting verbal processes *reveal, explain* and *emphasise* (i.e. *reveal*, *explain*, *emphasise that...*).

(357) Official news reports of the Assembly debate failed to reveal **details of the controversial content of the law**, especially its much-criticised Article 5. (*BNC Sampler*: AAT, W:newsp:brdsht_nat:report)

(358) The creation myths explain how the world emerged from chaos, a world which to the Egyptian comprised a flat earth, a flat sky above it supported by air and an underworld below the earth through which the sun travelled at night. (*BNC Sampler*: EVR, W:non_ac:humanities_arts)

(359) Another myth preserved in magical papyri emphasised the power of the divine name and was a tale about Isis and Re, usually called 'The Secret Name of Re'. (*BNC Sampler*: EVR, W:non_ac:humanities_arts)

Halliday & Matthiessen (2004) would argue that a verbal analysis applies to each of these examples. In their view, verbal processes involve '[...] any kind of symbolic exchange of meaning' (Halliday & Matthiessen 2004: 253), with such examples as *my watch says it's half past ten* or *the notice tells you to keep quiet* treated as verbal. They add, however, that if the subject is non-personal, the present simple selection of the verb resembles that of relational processes (Halliday & Matthiessen 2004: 255). Such a claim is exemplified with *the study says that such a diversified village structure produces a dualistic pattern of migration* (cf. **the study is saying that...*), with *the study* featuring as Sayer and the *that*-clause as Verbiage. It seems contradictory that other potential verbal processes such as s*uggest*, *mean* or *indicate* are treated as identifying relational intensive (Halliday & Matthiessen 2004: 238).

The 1994 edition of Halliday's SFG makes the point that such verbs as *imply*, *indicate*, *show*, *demonstrate*, *signify* or *suggest* '[...] may function either as "saying" or as "being (a sign of)"' (Halliday 1994: 142). The reason behind this classification lies in the consciousness of the subject: if it is realised by a conscious being, the clause is analysed as verbal (as in *the minister implied a change of policy*); conversely, if the subject is non-conscious, the clause is coded as relational intensive (as in *the minister's response implied a change of policy*). Martin et al. (1997: 126) hold the same view, with verbs like *show*, *indicate* or *suggest* analysed as both verbal (as in *she showed us that the substance was potassium*) and relational (as in *the result showed that the substance was potassium*). The analysis in this study follows this view. Therefore, the shell-noun phrase *details* in (357) is tagged as relational intensive Identifier (i.e. IDR(rel.in)), whilst *myths* and *myth* in (358) and (359) are both relational intensive Identified (i.e. IDD(rel.in)).

5.3.3.1.3 *Mental or relational?*

Mental clauses are typically those with a conscious participant as subject. Examples (360) and (361) are two cases in point, with the shell-noun phrases in boldface coded as Phenomenon.

(360) 'I've always …' Fergus lowered his voice as he bent to kiss her again, face in her hair and then his mouth on hers. Rory missed the exact words. (*BNC Sampler*: G0A, W:fict.prose)
(361) 'Maybe it's like your particle; inevitably uncertain. Soon as you understand one part of what it means, you lose any chance of understanding the rest.' (*BNC Sampler*: G0A, W:fict:prose)

However, as in 5.3.3.1.2, there are also cases of mental verbs with non-conscious subjects. This is shown in examples (362) through (364), where *book*, *study* and *finding* function as subjects.

(362) For example, a book about human rights in the context in which we live can be considered subversive because it disagrees with the vision of those in authority. ' (*BNC Sampler*: EBK, W:misc)
(363) Cronin's study of catalogue non-users sought to determine the characteristics and attitudes of non-users [...] (*BNC Sampler*: H0S, W:misc)
(364) The second finding looks at the impact that CNN has had on the viewing of news. (*BNC Sampler*: J1L, W:misc)

Halliday & Matthiessen's (2004: 203) view as to the nature of these examples is that, whilst the consciousness of the Senser is desirable, it is not required for a mental interpretation. In their opinion, such instances as *the film imagines that the FBI imported a free-lance black operative [...]* ought to be treated as mental, on a par with more typical instances with a conscious participant as subject. However, Martin et al. (1997: 120–1) suggest that, only if the Senser is a conscious being, can a mental analysis apply (e.g. *violent TV programmes affect/influence our children*). Otherwise, the example is treated as relational (e.g. *temperature affects/influences humidity*). As in 5.3.3.1.2, this study adheres to the latter view.

In (362), the idea conveyed by the use of *disagree with* is that *the book does not match/is not consistent with that vision* or that *the vision given in that book does not resemble that of those in authority*. Based on its similarity to verbs like *match* or *resemble*, classified by Halliday & Matthiessen (2004: 243) as identifying relational circumstantial comparative processes, the shell-noun phrase headed by *vision* is thereby coded as relational circumstantial Identifier (i.e. IDR(rel.ci)). In (363), *determine* carries the meaning 'to find out the results about something' (*LDCE*). Had the researchers behind the study occurred as subject, the example would have been coded as mental. Nevertheless, as it is the analytical process (i.e. *study*) that is brought to the forefront here, the example appears to be semantically equivalent to identifying relational intensive verbs like *show*, *define* or *reveal* (i.e. *the study sought to reveal/show/define/cast light on*

these characteristics). The shell noun *characteristics of non-users* is thus tagged as relational intensive Identifier (i.e. IDR(rel.in)). Finally, as regards (364), a likely paraphrase of its meaning could be: *the second finding is about/concerns/deals with/goes into that impact*. As all of these verbs feature in Halliday & Matthiessen's (2004: 243) class of attributive relational circumstantial matter processes, the shell-noun in boldface is coded as relational circumstantial Carrier (i.e. CAR(rel.ci)).

5.3.3.1.4 *Existential or material?*

Halliday & Matthiessen's (2004: 258) discussion of existential processes includes verbs which are also classified as intransitive material (e.g. *emerge, occur*, *happen*, *take place*) (Halliday & Matthiessen 2004: 187). Although their meaning in both uses is 'come into existence', it is argued that material senses show a preference for progressive forms, while existential senses are more likely to occur as simple forms (Halliday & Matthiessen 2004: 185). Besides, only the latter accept *there* as grammatical subject. Hence, in *the spiritual life is developing*, the verb is treated as material on the grounds of the *-ing* form and also of the oddity of a *there* paraphrase (i.e. **there is developing the spiritual life*). An example like *a similar pattern emerges*, however, is coded as existential on account of the possibility of a *there* transformation (i.e. *there emerges a similar pattern*). With respect to Quirk et al. (1985: 1408–9), existential clauses are claimed to include intransitive and presentative verbs covering such meanings as motion (e.g. *arrive*, *enter*), inception (e.g. *emerge*, *spring up*) and stance (e.g. *live*, *remain*).

Drawing on both grammars, a decision was made to apply an existential analysis to those instances where an entity is introduced for the first time in discourse, and where a *there* paraphrase is deemed acceptable. The presentative meaning of existential clauses is found to be linked to non-specific Deictics (e.g. *a/an*, *another*) occurring in noun phrases with cataphoric reference. In most such cases, the question *what happened?* identifies the Existent. Examples (365) through (367) illustrate these uses. For example, in (365) *along comes another challenge* has cataphoric reference and accepts a *there* paraphrase (i.e. *there comes another challenge*), thereby suggesting its existential status. Similarly, the acceptability of *there occurred a classic example* and *there grew a modest myth* in (366) and (367), along with the possibility of querying the subject through *what happened?*, reveals once again the existential meaning of these two examples.

(365) JUST when you thought that marathons, duathlons, biathlons and triathlons were the pinnacle of endurance tests, along comes another awesome challenge. (*BNC Sampler*: CF9, W:newsp:other:sports)
(366) It is in this area that a classic example of planners failing to see what contribution women are making occurred. (*BNC Sampler*: H8W, W:essay:univ)
(367) A late nineteenth-century Hepplewhite revival led to an 1897 reprint of the Guide by Batsford, and a modest myth grew up around George Hepplewhite, who was stated to have been apprenticed to Gillows of Lancaster and London [...] (*BNC Sampler*: GT9, W:biography)

Many existential processes may also occur as material, provided that a *there* construction does not make sense in context. This is normally due to the use or the potential for the use of the progressive form. In these cases, the verb is often non-presentative, which means that, unlike in examples like (365) through (367), reference tends to be anaphoric. The subject in these instances is treated as Actor, following Halliday & Matthiessen's (2004: 184–5) analysis of intransitive material processes. These clauses may be queried through both *what happened?* and *what happened to X?*. Examples (368) through (370) show three such material processes. *Assessment* in (368) is coded as Actor, based on the progressive form of *emerge* (i.e. *new details and a frank assessment are emerging*), as well as on the oddity of a *there* construction (i.e. **there are emerging new details and a frank assessment*). The anaphora test would not apply in this instance, as *assessment* encapsulates a subsequent discourse segment. Reference to a preceding segment is evident in (369) and (370), with the verb in the former being in the progressive form (i.e. *were opening*).

(368) New details of the summit conversations are emerging in the course of White House briefings, including a surprisingly frank assessment by the Soviet leader of the looming economic crisis. Mr Gorbachev told the Americans that the Soviet people would judge his perestroika by how quickly he could get food and consumer goods into the shops. (*BNC Sampler*: A8W, W:newsp:brdsht_nat:report)
(369) During nineteen ninety one the Board has been delighted to open new areas of work in Inverness where our first designated place and associated hostel was opened on a most happened--happy day by Sir Russell <gap desc='name' reason='anonymization'>. In Elderslie near Paisley <pause> where Lady <gap desc='name' reason='anonymization'> the wife of last year's Lord High Commissioner opened our fourth senile dementia unit. In Haddington in, in East Lothian where Tyne Park House now offers a mental health service within that county. But while new projects were opening, the Board was

forced to depart from its Shetland Islands project [...] (*BNC Sampler*: F86, S:meeting)

(370) While small, dead and still-bleeding little bodies lay like exposed carrion round the harbour, the weeping fathers hovered over them to stop the crows and flies finishing off what Gorbrandt's men had started. Feelings of grief far outweighed thoughts of revenge. [...] The invaders' next mission was to capture our treasure, stored in the crypt of this church. [...] In a booming voice infused with all the wrath of the Old Testament deity the pastor gave the answer: 'So that **such terrible things** never happen again!' (*BNC Sampler*: AEA, W:fict:prose)

5.3.3.1.5 *Behavioural processes*

In line with Martin et al. (1997: 127), the object of a so-called behavioural process is treated as Phenomenon or Verbiage according to whether the verb is conceptually mental or verbal. This implies that, in spite of allowing the progressive form as material verbs do, and in spite of its failure to project clauses, the verb still conveys a communicative or mental process, and, as such, it is either verbal or mental in meaning. In Halliday & Matthiessen (2004: 251), this analysis is restricted to the verb *watch*, whose object is said to be a Phenomenon. In their view, most other behavioural processes are intransitive (e.g. *breathe*, *grumble*, *chatter*). Those that have an object are similar to material Scope-taking verbs like *take a walk* or *have a shower*, in that the object is not really a participant, but the nominalised verb itself (e.g. *give a yawn*, *sing a song*), which they term Behaviour.

In the four examples below, (371) through (374), the shell-noun phrases in boldface are coded as behavioural Verbiage (i.e. VER(beh), (371)-(372)) and behavioural Phenomenon respectively (i.e. PHE(beh), (373)-(374)). *Describe* and *express* are two processes which do not project *that*-clauses but are clearly verbal or communicative in meaning, the former involving the act of saying '[...] what something or someone is like by giving details about them' (*LDCE*) and the latter used 'to tell or show what you are feeling or thinking by using words, looks, or actions' (*LDCE*) (in (372), it is the use of words that is at issue, i.e. *he said he was surprised that...*). *Experience* and *examine*, in (373) and (374), entail a mental process: passive in *experiencing visions* (i.e. someone feels or receives those visions) and active in *examining prejudices* (i.e. someone considers those prejudices carefully).

(371) Anna had explained, and it had given her the opportunity to describe **details her daughter found miraculous**. (*BNC Sampler*: AEA, W:fict:prose)

(372) On a one-day visit from Washington, where he now resides, the US ambassador expressed his exasperated surprise that the Lebanese Forces had not joined the 'legitimate' government. (*BNC Sampler*: A8W, W:newsp:brdsht_nat:report)
(373) [...] places where visions, transcendental states of prophecy, may be experienced. (*BNC Sampler*: CBB, W:non_ac:humanities_arts)
(374) The decline in membership, the economic recession, has damaged all the trade unions and has forced us all to examine all prejudices, to look above the sectional interests and wherever possible do all we can to protect our movement. (*BNC Sampler*: HLW, S:speech:scripted)

Special mention should be made of *have*, a verb often classed as relational possessive, but which, in the light of examples like (375) or (376), could also be coded as behavioural mental. According to the *OED*, one of the senses of *have* is '[T]o be possessed or affected with (something physical or mental); to be subjected to; to experience; to enjoy or suffer', as in *have an accident, have a good holiday* or *have difficulties*. As neither (375) nor (376) is about the possession of *dreadful experiences* or *brilliant times*, but about experiencing, suffering (as in (375)) or enjoying (as in (376)) such things, the shell-noun phrase in each case is tagged as behavioural Phenomenon.

(375) If er you've had dreadful experiences with er guests staying at your place, give me a ring [...] (*BNC Sampler*: HM4, S:brdcast:discussn)
(376) But I, I can understand why she'd want to marry, why she'd want to do that, cos she hasn't ha-- well had an awful lot of <pause> brilliant times in her life since she left Nigeria. (*BNC Sampler*: KC7, S:conv)

The use of *have* in the following examples shows a related sense: '[T]o hold or entertain in the mind (a feeling, opinion, etc.); to entertain, hold, cherish' (*OED*), as in *have an idea*, *have a doubt*, *have a feeling*. According to *LDCE* and *OALD*, in this sense *have* is interchangeable with *have got*, and, unlike the two aforementioned examples, it may not be used in the progressive tenses. As such, while it is possible to say *I've got a suspicion* or *I've got certain prejudices*, it is unacceptable to say **I've got dreadful experiences* or **I've got many brilliant times*. Conversely, one may find such instances as *I'm having a dreadful experience with those guests* or *I'm having a brilliant time here*, but it is rare to find **I'm having a suspicion* or **I'm having many prejudices*. Drawing on Halliday & Matthiessen's (2004: 245) claim that even abstractions such as *idea* can occur as objects of relational possessive verbs, the shell-noun phrases in (377) and (378) are coded as relational possessive Attributes (i.e. ATT(rel.po)).

(377) I've a sneaking suspicion he bins any letter that criticizes his government, before he even reads it. (*BNC Sampler*: H4A, S:meeting)
(378) All wood-burning pundits have their prejudices and the commonest one is against elm. (*BNC Sampler*: H13, W:fict:prose)

5.3.3.2 *Syntactic function*

Quirk et al.'s (1985) categories are used for the analysis of the syntactic functions of shell-noun phrases. In broad terms, nouns may fulfil five main functions, four of them at clause level and one at phrase level:

i) Subject (i.e. SB), as in (369) or (370) above.
ii) Object, which comprises direct objects (DO), as in (371) or (372) above, indirect objects (IO), as in (379), and prepositional objects (PO), as in (380).

(379) This was 'to give greater flexibility to party work, and to remove unnecessary parallelism in the activity of the regional and city committees'. (*BNC Sampler*: A8J, W:newsp:brdsht_nat:report)
(380) [...] I think you've got to think about the kind of words you use. (*BNC Sampler*: JJS, S:classroom)

iii) Complement, containing subject complements (SCL), as in (381), and object complements (OCL), as in (382).

(381) It is a terrible irony that although they are invisible to the planners <pause> all of us who've travelled in developing countries cannot fail to have noticed these children out of school. (*BNC Sampler*: JNG, S:meeting)
(382) [...] this chapter considers population and employment as contiguous rather than necessarily related phenomena. (*BNC Sampler*: FR2, W:ac:soc_science)

iv) Adverbial, as in (383) and (384), where the shell-noun phrase in each functions as a time and manner adverbial adjunct respectively (AB.AT.TM, AB.AT.MAN).

(383) Erm every time you use a tape <pause> you fill in <pause> this is tape side B <pause> tape one side B <pause> tape two side A <pause> tape two side B. (*BNC Sampler*: KCL, S:conv)
(384) [...] i order from MY PeN, i verse the way i like [...] (*BNC Sampler*: F9M, W:fict:poetry)

v) Prepositional complement, where the noun follows a preposition. In these cases, the coding indicates the preposition the noun is a complement of, as well as the clause-level function served by the entire structure. It is important to note that, semantically, the annotation considers only the participant type of the whole structure. For example, the shell-noun phrase in (385) is the complement of the preposition *of*, which heads a prepositional phrase acting as complement of the noun *rejection*, whose clause-level function is subject (CP.*OF*(<SB)). In terms of participant type, the example is annotated as relational intensive Identified (IDD (rel.in)). The coding of (386) is more complex, in that the shell-noun phrase is the complement of *in*, heading a prepositional phrase which complements *experts*, and which, in turn, acts as the complement of the preposition *of*, heading a prepositional phrase which complements *approach*, the head of the entire noun phrase. *The approach of experts in the area of aesthetic and artistic judgement* is the complement of *from*, heading a prepositional phrase whose clause-level function is a means adverbial adjunct. The syntactic tag for this example is CP.*IN*(<CP.*OF*<CP.*FROM*<AB.AT.MS). Semantically, the example is coded as a Circumstance of manner and means (CR.MAN.MS).

(385) Meanwhile, Turkey's Foreign Minister, Mr Ali Bozer, said yesterday that the European Commission's rejection of Turkey's application to join the EC until after 1992 did 'not fully correspond to Turkey's aspirations'. (*BNC Sampler*: AA4, W:newsp:brdsht_nat:report)

(386) [...] any method should include room for different tastes (Jacques, 1980), and that much could be learnt from the approach of experts in the area of aesthetic and artistic judgement (Carlson, 1977). (*BNC Sampler*: FR2, W:ac:soc_science)

5.3.3.3 *Theme vs. Rheme*

This variable draws on Halliday & Matthiessen's (2004: 64–105) Theme system. Roughly speaking, the system comprises three main categories: Theme (T), marked Theme (*T) and Rheme (R). The Theme corresponds to any clause-initial constituent realised by one participant, one circumstance or one process. In declarative clauses, the Theme slot is typically occupied by a nominal participant (noun or pronoun). In interrogative clauses, by contrast, it is a finite verbal operator plus subject (e.g. *do you*, *is he*, *can they*) or a *wh*-interrogative (e.g. *who*, *where*, *how many*) that acts as the starting point for the message. As regards imperative clauses, it is the process or verb that functions as Theme (e.g. *turn it down*, *don't do that*).

Any constituent following the typical thematic choice for each clause type is treated as Rheme.

If a circumstance, object or complement occurs as the clause-initial constituent in any of the aforementioned clause types, the Theme selection is coded as marked and any element following it is part of the Rheme. If a conjunction (e.g. *but*, *when*, *so that*), an adverbial conjunct (e.g. *therefore*, *nevertheless*, *actually*) or an adverbial disjunct (e.g. *surely*, *unfortunately*, *happily*) is used in Theme position, any participant, circumstance or process following such an element is still treated as Theme, or, more specifically, 'topical Theme'. Clause-initial conjunctions and conjunctive or disjunctive adverbials are coded in SFG as 'textual' and 'interpersonal Themes'. The analysis in the present study makes no distinction between textual, interpersonal and topical Themes, thereby annotating shell-noun phrases in Theme position simply as Theme (T).

Examples (387) and (388) contain two shell-noun phrases in unmarked Theme positions (see (384) and (386) for two examples of Rheme). In (387), *employment numbers in these areas* follows the conjunctive textual Theme *and*, while in (388), *one of the big ways* is both clause and sentence-initial.

(387) Agriculture, with only 1 man needed per 200 acres on arable land, and industries like car-manufacturing with its increasing robotic aids, are unlikely candidates, and employment numbers in these areas are bound to fall. (*BNC Sampler*: G0C, W:commerce)

(388) One of the big ways is, is spiritualism and things that are not real, things that are sort of er blind in the mind of the unbeliever as it were, in other words material things probably and things that have under revelation <pause> (*BNC Sampler*: KBX, S:conv)

The shell nouns in (389) and (390) are marked Themes. In (389), *objectives* is part of a locative circumstance introducing a coordinate clause. In (390), *project* occurs in a clause- and sentence-initial purpose circumstance following a conjunctive textual Theme.

(389) I refer specifically to the recreational page from the Estuaries Consultation Document from English Nature and in one of their er proposals and objectives they say encouraging self regulation and observance of Code of Practice by local clubs and groups to avoid clon-- conflict with and or disturbance to other users including nature conservation interests. (*BNC Sampler*: J3W, S:meeting)

(390) So for any particular project we want a quality plan which will tell us what the client brief is [...] (*BNC Sampler*: H47, S:speech:unscripted)

Particular note should be made of the application of the Theme system to examples drawn from real or fictional spoken conversation. If a given shell

noun is part of the Rheme of a preceding spoken turn, or occurs alone as a reaction to a previous comment (a statement or a question), the example is tagged as Rheme. The former is evident in (391), with *a crime* being the prepositional object and the matter circumstance of the verb *tell about*. The latter is shown in (392) and (393). In (392), *very small* projects is the answer to a previous question with the verb *manage*. By *very small projects*, therefore, the hearer understands *I manage very small projects*, with the shell-noun phrase analysed as direct object, Goal and Rheme. In (393), *good point* may be taken as the subject complement or Attribute of a missing intensive relational process (i.e. *That is a good point*). Based on the typical connection established between Theme and Given information and Rheme and New information, *good point*, being one of the speaker's reactions to another spoken turn, appears to provide new information, thereby lending further support to the Rheme analysis.

(391) Fireman, we're on to firemen now, what, why do policemen, what do policemen go to the houses for? They go to tell about
A crime. (*BNC Sampler*: FM7, S:unclassified)
(392) <pause> You manage projects don't you?
Very small projects and some quite large ones. (*BNC Sampler*: H47, S:speech:unscripted)
(393) I think we should stick together. He might be violent.
Good point. I'll come with you. (*BNC Sampler*: FU6, W:fict:drama)

5.3.4 *Direction of encapsulation and antecedent*

5.3.4.1 *Direction of encapsulation*

Examples are coded in terms of three main encapsulating relations: anaphora, cataphora and exophora. Endophoric encapsulation is subdivided into intra- and intersentential realisations. Five primary codes are thus used in the analysis of this variable: AF.INTRA (intrasentential anaphora, as in (394)), AF.INTER (intersentential anaphora, as in (395)), CF.INTRA (intrasentential cataphora, as in (396)), CF.INTER (intersentential cataphora, as in (397)) and EXO (exophora, as in (398)). These occur alone or in combination.

(394) The latest roads White Paper also includes <u>further improvements to the M20, M23 and M25</u> but, despite government streamlining of the planning and construction process, it is unlikely that many of these projects will be completed before the year 2000. (*BNC Sampler*: BMJ, W:misc)

(395) Participants were particularly inspired by the lead given by Namibia in defining communication policy (see pages 4 and 5 of this newsletter) and in placing it at the service of justice, freedom and democratic growth. Namibia's example has encouraged the Eritreans to press ahead with their own policy and to respond to the widespread need for media training. (*BNC Sampler*: EBK, W:misc)

(396) Applications for voluntary repatriation virtually dried up. (*BNC Sampler*: A9V, W:newsp:brdsht_nat:report)

(397) Like other intellectuals of his time, such as Isaac Newton, he retained his enthusiasm for investigating curious phenomena. He went on to prove that earthquakes were the result of electrical disturbances of the air of which the movement of the Earth was a secondary effect, that sponges were not living creatures and that the deluge took place in the Autumn. (*BNC Sampler*: CBB, W:non_ac:humanities_arts)

(398) [...] such sites become 'cosmic power points on the surface of the earth, special places where the mind can expand into new levels of consciousness, places where visions, transcendental states of prophecy, may be experienced'. (*BNC Sampler*: CBB, W:non_ac:humanities_arts)

The identification of intersentential realisations draws on sentence boundaries as indicated by full stops, question marks and exclamation marks. Commas separate clauses within one and the same sentence, and, as such, are always intrasentential. As regards colons and semicolons, only if the example suggests an appositional reading does an intrasentential analysis apply; otherwise, if judged to be equivalent to a full stop, the example is treated as intersentential. Examples (399) and (400) are two cases in point. The colon introducing the underlined segment in (399) is equivalent to *such as*, thereby favouring the CF.INTRA analysis of *practical applications*. By contrast, the colon introducing the shell content in (400) is not as easily replaced with appositive markers (e.g. *such as* or *that is to say*) as in (399), and, on these grounds, *evidence* is treated as CF.INTER. Similarly, the anaphoric realisation of *these phenomena* in (401) is coded as intersentential, as the brackets seem to introduce information that would otherwise be separated by a full stop.

(399) Nevertheless, there are a number of practical applications for the laws described in this paper: proving programs equivalent to one another, transforming programs to make them more efficient, and transforming programs to a restricted syntax for special applications. (*BNC Sampler*: G3N, W:ac:tech_engin)

(400) This fact, and the decay-rate restrictions discussed below, lie behind the paucity of experimental evidence on Ikeda instabilities: only in two all-optical systems (both pulsed rather than continuous-wave) has 2tR

oscillation and chaos been observed to date. (*BNC Sampler*: J2H, W:ac:nat_science)

(401) When Ikeda in 1979 performed a stability analysis starting from the full Maxwell-Bloch equations, with propagation effects included {22}, he discovered an apparently distinct family of instabilities, and indeed chaotic behaviour, which will be the dominant topic of the remainder of this section (as Lugiato {26} has pointed out, these phenomena are not actually distinct from the Bonifacio-Lugiato instabilities, and can be analysed by similar methods). The simplest approach to the Ikeda instability makes the limit I'00 in the Maxwell-Bloch equations (7.2). It is then possible (at least in the ring resonator) to integrate the full equation along the characteristic, z = ct + constant, to obtain an explicit relationship between E (0, t) and E (L, j + L/c). Further simplification ensues if we now let y and take the dispersive or Kerr limit finite. (*BNC Sampler*: J2H, W:ac:nat_science)

If specifics lie within one spoken turn, and no strong punctuation mark is used in the transcribed text (e.g. full stop, colon, semicolon), the encapsulating direction is treated as intrasentential. This is shown in (402) and (403), where *this particular venture* and *a joke* encapsulate information scattered throughout the same spoken turn.

(402) <pause> Er and I well remember er <pause> even as early as nineteen nineteen <pause> the election which took place immediately after world war one, I remember being <pause> er sort of dragged round the streets, <pause> you know, er <pause> I think it was enjoyable, I don't know <pause> er by my mother, er attending these street meetings, <pause> er <pause> I I I very well remember it because I had a a a new coat, a new coat for for er for this particular venture, <pause> and er the two things you know are fairly deep in my memory. (*BNC Sampler*: FYJ, S:interview:oral_history)

(403) How're you stuck?
I, I'm not
you're not stuck, you're pretending, was that a joke?, <-|-> Richard come on love <-|-> (*BNC Sampler*: KB8, S:conv)

As explained in 4.4.1 and 4.4.2.8, in this thesis endophoric encapsulation is not restricted to specific Deictics (especially, the definite article and demonstrative determiners), as is often the case in the literature. Based on the contextual assumption underlying the corpus theoretical approach of this study (see 4.4.1), endophora applies whenever the interpretation of a shell-noun phrase hinges on information contained in the surrounding co-text. Examples (404) through (407) illustrate the four main types of endophoric encapsulation with non-specific shell-noun phrases found in the

study sample: zero article/undetermined, negative, indefinite and non-assertive noun phrases.

(404) So if for example a dishonest car dealer said bargain, thirty thousand miles only, when really it had done ninety thousand miles then that would be a criminal offence under the Trade Descriptions Act. Because he's claiming the car is something which it clearly is not. Or if a jeweller said, you know, the rings in my window are made of gold when really they were not made of gold at all, then that would be a false trade description.On the streets of Nottingham, just before last Christmas, there was some street traders selling coats er which they described as sheepskin.We suspected they were not made of sheepskin at all er and we took some of them away and analyzed them to find out what exactly are they made of. And they turned out to be made of manmade fibres. So those gentlemen were not only infringing the Sale of Goods Act they were committing a criminal offence as well by infringing the Trade Descriptions Act so they ended up being prosecuted by Trading Standards. But I, I do feel the Trade Descriptions Act is a is a very important er piece of consumer protection designed of course to protect not only you and me as consumer but also designed to protect traders against unfair trading practices. (*BNC Sampler*: FUT, S:speech:unscripted)

(405) Gen Noreiga said last Friday that Panama considered itself at war with the US so long as American aggression against Panamanians continued. The reaction of a Pentagon spokesman to the remark was first an incredulous 'What?' then laughter. The implication was that Gen Noreiga's bravado could in no legal sense be taken as a declaration of hostilities. The mechanisms for registering and resolving a conflict via the United Nations, which would apply in this situation, had not been gone through by the general. (*BNC Sampler*: AAB, W:newsp:brdsht_nat:report)

(406) He said, 'That boy at the counter said you got a stage ticket.' John Russell opened his hand on his lap. 'This?' 'That's it. You give it to me and you can take the next stage.' 'I have to take this one,' Russell said. 'No, you want to is all. But it would be better if you waited. You can get drunk tonight. How does that sound?' 'I have to take this one,' John Russell said. 'I have to take it and I want to take it.' 'Leave him alone,' the ex-soldier said then. 'You come late, you find your own way.' Frank Braden looked at him. 'What did you say? ' 'I said why don't you leave him alone.' His tone changed. All of a sudden it sounded friendlier, more reasonable. 'He wants to take this stage, let him take it,' the ex-soldier said. You heard that ching sound again as Frank Braden shifted around to face the ex-soldier. He stared at him and said, 'I guess I'll use your ticket instead. The ex-soldier hadn't moved, his big hands resting on his knees, his feet still propped on the canvas bag. 'You just walk in,' he said, 'and take somebody else's seat?' Braden's pointed hat brim

moved up and down. 'That's the way it is.' The ex-soldier glanced at John Russell, then over at me. 'Somebody's pulling a joke on somebody,' he said. Russell didn't say anything. He had made a cigarette and now he lit it, looking at Braden as he blew the smoke up in the air. 'You think I come in here to kid?' Braden asked the ex-soldier. (*BNC Sampler*: J2G, W:fict:prose)

(407) Recently on the cover of Direct, was a young black person who was er, a depiction of somebody unemployed in Britain today. Unfortunately, the photograph could have been er, mistaken as a stereotypical photograph of somebody from Los Angeles, or any other part of the world where there is problems in inner city area. Please, Tom when er placing photographs in Direct, please try to make them a little more appropriate. Thank you.
Now turn to page one O four, one O five, one O six and one O seven, and one O eight. Tom could you respond to those points please. <pause dur='13'> [...]
On the other point that was raised in respect of er Direct. I think our delegate will appreciate that er, it is a difficult issue, we do try as much as we possibly can to project all aspects of our membership in a fair fashion, and we go out of our way er, to be fair in respect of er certain er areas. I would simply conclude that while we're always open to any sort of criticism, we're always open to any kind of recommendation that might be made, erm to us in respect of er Direct. (*BNC Sampler*: HDT, S:speech:scripted)

In (404), the lack of a determiner at first sight seems to impose a generic interpretation of *unfair trading practices*. A closer reading of the previous co-text reveals that participants listening from the beginning of the speech are, at this precise moment, aware of what *unfair practices* may involve (e.g. dishonest car dealers, fake sheepskin). On these grounds, the contention here is that the shell-noun phrase, whilst lacking a determiner, is meaningful to participants in this discourse situation, as it activates information present in their minds. In (405), the occurrence of a negative Deictic would normally be associated with exophora (e.g. *there are no problems*, *there are no ideas*). In this particular context, however, *no legal sense* seems to anticipate the underlined segment. Without the underlined specifics, a reader may wonder why Noriega's declaration of war cannot be legally understood as such. The answer is in the underlined specifics, where a legal understanding or a legal sense would have involved going through the mechanisms for registering and resolving a conflict via the United Nations, which general Noriega did not do. In (406), whilst *somebody is pulling a joke on somebody* might be taken as a generic statement out of context, the indefinite shell-noun *a joke* is found to refer back to the situation presented in the underlined segment upon closer reading, i.e. someone trying to trick

somebody else into giving him his ticket. Lastly, in (407), the underlined sentence is a recommendation that someone in the audience is making about the stereotypical implications arising from some of the covers of Direct. At a later stage, Tom, the editor of the magazine, acknowledges the illocutionary force of such a statement by using the non-assertive noun phrase *any kind of recommendation that might be made to us in respect of Direct*. Again, whilst out of context *any kind of recommendation* suggests a generic reading (i.e. *not any in particular*), the extended context shows how the shell-noun phrase indicates awareness of one such recommendation.

In the examples presented so far, encapsulation is unidirectional, as the shell noun points backwards, forwards or out of the text. There are instances where encapsulation is bidirectional, especially when the shell noun is preceded by post-Deictics like *another*, *other*, *further*, *additional*, *similar* and *different*, and by indefinite quantitative Numeratives like *more* (cf. also Francis 1986: 90 and 1994: 98, where the bidirectionality of these modifiers is also considered). The three examples below illustrate this use. In (408), *some more ways to picture it* is treated as AF.INTER&EXO, on the grounds that Russell refuses at this point to hear any more ways of taking advantage of his life in a new land (hence the use of EXO). If he had wanted to hear more ways, a bidirectional anaphoric and cataphoric analysis would have applied. In (409) and (410), as specifics are given both anaphorically and cataphorically, *another awesome challenge* and *similar problems* are coded as AF.INTER&CF.INTER.

(408) 'But it's a big something to think about,' Mr. Mendez said. 'Going to Contention. Going there to live among white men. To live as a white man on land a white man has given you. [...] 'Sure,' Mr. Mendez said. 'You could sell it. Buy a horse and a new gun with some of the money. See people at San Carlos hungry and given them the rest. Then you got nothing.' Russell shrugged. 'Maybe so.' 'Or you sell only the herd and grow corn on the land and make tizwin, enough to keep you drunk for seven years.' 'Even that,' Russell said. 'Or you can work the herd and watch it grow,' Mr. Mendez said. 'You can marry and raise a family. You can live there the rest of your life.' He waited a little. 'You want some more ways to picture it?' 'I have too many ways now,' Russell said. But he didn't sound worried about it. That didn't satisfy Mr. Mendez. He was trying to convince him of something and kept at it. He said then, 'I hear it's a good house.' Russel nodded. 'If living there is worth it to you.' 'Man' Mendez said, like something good was staring at Russell and he didn't know enough to take it. 'What do you want?' (*BNC Sampler*: J2G, W:fict:prose)

(409) JUST when you thought that marathons, duathlons, biathlons and triathlons were the pinnacle of endurance tests, along comes another awesome challenge. The quadriathlon, still in its infancy, is basically a

triathlon with a gruelling canoe leg thrown in for good measure, and East Anglia has one of the leading exponents in this new sport. (*BNC Sampler*: CF9, W:newsp:other:sports)

(410) These papers were a series of articles published in 1787 and 1788 to persuade the citizens of New York to ratify the U.S. Constitution; most of the papers were known to have been written by Hamilton or by Madison, but the authorship of twelve of them was disputed. [...] A partly similar problem is presented by the Aristotelian corpus, which contains two ethical treatises of parallel structure, the Nicomachean Ethics and the Eudemian Ethics. The problem here is not one of authorship attribution, since most scholars regard them both as genuinely Aristotelian; the puzzle is that three books make a double appearance in the manuscript tradition, once as books five, six and seven of the Nicomachean Ethics, once as books four, five and six of the Eudemian Ethics. [...] **Similar problems** have been studied by similar methods in the very different context of the law-courts. It is sometimes a question whether a particular confession is the unaided work of an accused or is a fabrication of a police officer. [...] The Californian court trying Patricia Hearst for bank robbery was asked to hear stylometric evidence to help determine whether she composed the propaganda for the Symbionese Liberation Army which was part of the case against her. (*BNC Sampler*: F98, W:ac:humanities_arts)

Bidirectional encapsulation is also shown where no post-Deictic occurs. This applies especially to spoken discourse situations, where the question often arises whether the information provided anaphorically is enough for an adequate understanding of the use of a particular shell noun. For example, in (411), the noun phrase in boldface represents the primary informative focus of this trial, i.e. ascertaining the advice that solicitors give clients on commercial conveyancing. The segments in italics reveal how the lawyer is constantly pressing the judge to accept evidence from expert witnesses. This being the case, the underlined segments still offer insights into what this practice entails. *The accepted and standard practice*, the main information focus, is mentally present throughout this discourse situation, forcing listeners to pay attention to every detail in order to obtain a general idea of the implications behind this practice. This means that the noun is neither clearly referring to anaphoric information nor to cataphoric information, but to specifics scattered throughout the previous and following discourse segment, hence its coding as AF.INTER&CF.INTER. In (412), the reader who gets to the noun in boldface already knows about the McLaren girl's experience. All the previous chapters have spelt out the nature of this experience, so the noun is mainly anaphoric. However, at this point, the subject of her experience is brought up in a dialogue within the novel. Although the noun as such is already mentally present for the reader, it is

apparently not so for some of the interlocutors in this conversation, as one of them asks (perhaps maliciously) what happened (see question in italics). Therefore, the noun acts as an anchor between what the reader knows from the narrator and what the McLaren girl herself is going to say about her experience in this particular conversation. On these grounds, the example is coded as both anaphoric and cataphoric, i.e. AF.INTER&CF.INTER.

(411) My Lord if I could question it. It is not, er this evidence does not go to a matter of law er er and the duty but it <unclear> matter of practice and my Lord what this case is dealing with is about what if, what is or *should be the practice of a solicitors engaged in commercial conveyancing as to the advice that is given to clients* <pause> and er my Lord the er <pause> commercial conveyancing is obviously a matter which particularly <-|-> concerns <unclear>
<-|-> Can I, can I <-|-> take it this way. Analysing your <pause> opening to me <pause> you were saying that there's expert evidence <pause> that <pause> solicitors instructed in the purchase of a property must <pause> ask about financial arrangements and advise about them. <pause> Secondly it is the solicitors <pause> must advise about clause twenty two <pause> and thirdly <pause> well that's about it really, isn't it <-|->. [...]
Even if there is no directly relevant practice of the profession*, the evidence of other practitioners as to what they do in similar cases and why, is likely to be of benefit to the court*. A knowledge of the working of a solicitor's office, particularly ero-- of those departments handling non-contentious business, cannot be automatically imputed to the judge or to council <pause> and he may as well make <unclear> *it is not uncommon for an expert witness to give evidence of what he would have done in a particular situation after consideration* and er I resign on that because in my submission er the issues in this case are clearly issues of mixed fact and law [...]
[...]My Lord it will be helped in that situation in my submission by evidence from <pause> which is dealing with what is the accepted and standard practice in that solicitors in handling these forms of transactions <pause> and what advice is given <unclear> my Lord undoubtedly that must be of benefit to the courts.
[...]That <pause> the negligence or rather the breach of the implied term, was failing to ensure as is the common practice amongst solicitors that the proposed source of finance had in fact agreed both finance for the transaction and the term thereon. [...] (*BNC Sampler*: JJV, S:courtroom)

(412) I was out in front of the Hatch & Hodges office at the time, directly across the street, and I got a clear look at the girl even with all the people around. She was seventeen or eighteen and certainly pretty. [...] But she looked good anyway. Even after living with Apaches over a month and after all the things they must have done to her. Somebody

> said the girl had been taken by Chiricahuas on a raid and held four or five weeks before a patrol out of Fort Thomas surprised their rancheria and found her. She had stayed at Thomas a while and now this officer was to put her on a stage for home. [...] 'It's a rosary.' 'I don't know why,' Mrs. Favor said, 'I thought they were Indian beads.' Her voice was soft and sort of lazy sounding, the kind of voice that most of the time you aren't sure if the person is kidding or being serious. 'You might say they are Indian beads,' the girl said. 'I made them.' 'During **your experience**?' Dr. Favor said, 'Audra,' very low, meaning for her to keep quiet. [...] Braden, I noticed, was looking at the McLaren girl too. *'What happened?'* he said. [...] 'Did they treat you all right?' 'As well as you could expect, I guess.' 'I suppose they kept you with the women.' 'Well, we were on the move most of the time.' [...] 'Did they-bother you?' 'Well,' the McLaren girl said, 'I guess the whole thing was kind of a bother, but I hadn't thought of it that way. One of the women cut my hair off. (*BNC Sampler*: J2G, W:fict:prose)

In addition to bidirectional encapsulation, there are also instances of intra- and intersentential realisations of one encapsulating direction. Examples (413) and (414) below illustrate intra- and intersentential cataphora (CF.INTRA&CF.INTER). In (413), the appositive noun phrase *the French multinational telecommunications* is not the only specifics for *an example*. Prior to the occurrence of the shell noun in bold, a situation is described, one whereby capitalism and imperialism force countries to export their products abroad, because national citizens cannot afford to buy them. French multinational telecommunication companies are an example of such a state of affairs, but, the appositive noun phrase does not clarify the point to be made by using this example alone. The point is specified in the intersentential segment, where French telecommunication industries are said to be doing much better abroad than in France. Similarly, in (414), whilst the main problem occurs intrasententially, this is elaborated intersententially.

(413) For example, when you talk of, imperialism, when you look at, when we, we've study it from a <unclear>, erm, quite a few, so it means that <unclear> that capitalism <unclear> supply [...] and, but the problem is, that, although it creates supply it laboured people who worked, yeah [...] cannot buy it because it's too expensive for them. So capitalism has to export it over its boundary and that's one of the company, what you could call it period, for example, erm [...] So for example, I don't know, erm, <unclear>, let's take **an example**, <unclear>, <unclear>, the French, erm multinational, tele-- <pause>telecommunications [...] Multinationals, they have erm, <unclear>, doing much better abroad than it is doing in France. [...] Because the French cannot always buy

what <-|-> <unclear> <-|-> [...] That means imperialism means in a way they're the French market it's well known, and, makes itself sell at much a <unclear> than in its, in its own country, so you can say that for example Thompson has a monopoly tower, most European and <unclear> <-|-> have a monopoly tower <-|-> (*BNC Sampler*: KCV, S:conv)

(414) Now you've got a **real problem** here because you just <pause> er <unclear> you know <unclear> that this is the only room available. There's no other committee room available on that Friday. Er admittedly they'll all be available on the Saturday <pause> but perhaps we would need, you know, we've perhaps a little bit more <laugh>more work to be done. (*BNC Sampler*: F7J, S:meeting)

Examples (415) and (416) show the most common realisation of intra- and intersentential anaphora (AF.INTRA&AF.INTER) in the sample. In these cases, the shell noun in bold refers back to a pronoun (e.g. *this*, *it*), which, in turn, encapsulates information lying elsewhere. The pronouns are deictic in function, as they just draw attention to the truly informative intersentential specifics. Example (417), whilst also coded as AF.INTRA&AF.INTER, differs from (415) and (416) in that, unlike non-personal pronouns (e.g. *it*, *this*, *that*), *you* does not refer back to previous or subsequent discourse segments, but to a person. Thus, *a mass of prejudice* is kept as a shell use, in that, by attributing a quality to an entity (a conscious one here), it serves the typical function of characterisation (Schmid 2000: 15; see 2.2.2.2) associated with shell nouns. However, *You*, on account of its being a person pronoun, does not offer any information as to what this prejudice is about, hence its intersentential encapsulation of informative specifics.

(415) What it is about, members of the jury, is a very professional police force doing the best they could in the circumstances. <pause> And what do they try to do, they were trying, above all else, to protect you and I, the general public. And this is a police force, you will remember, Mr <gap desc='name' reason='anonymization'> telling you <pause> <unclear> **very remarkable characteristic**. Never in it's entire history has it shot a gun in anger at a human being. (*BNC Sampler*: JJV, S:courtroom)

(416) We'll try and be a bit more adventurous this year and go a bit further! <laugh> <pause> But I can't sit in the car very long. It's terrible! It's a **problem**. (*BNC Sampler*: KST, S:conv)

(417) That must be east, then. I think we can assume that.
I'm assuming nothing.
No, it's all right. That's the sun. East.
(Looks up) Where?
I watched it come up.

> No … it was light all the time, you see, and you opened your eyes very, very slowly. If you'd been facing back there you'd be swearing that was east.
> (Standing up) You're a mass of prejudice.
> I've been taken in before. (*BNC Sampler*: FU6, W:fict:drama)

In (418) and (419), the shell-noun phrase does not relate to a pronoun, but to another shell noun for which informative specifics are given intersententially. These cases reveal a direct relationship of experiential identity between both shell nouns (e.g. *these ideas are an important challenge*, *this story is a fairly complete myth*), and an indirect one between the shell noun under analysis and the lexical realisation of the other shell noun. In these cases, (SN2) is added to the encapsulating relation and antecedent of the second shell noun. Thus, *challenge* in (418) is tagged as AF.INTRA&AF.INTER&AF.INTER(SN2) in terms of encapsulating direction, and as LC.SNP(*THEY*)&LC.SNP&GB.ET.PR(SN2) in terms of antecedent. This means that *challenge* is intrasententially realised by the pronoun *they,* which relates intersententially to the shell noun *these ideas*, it being lexically realised by an extended discourse segment. In (419), *myth* is similarly tagged as AF.INTRA&AF.INTER(SN2), as *myth* refers back to *this story*, which directly encapsulates the events recounted in the preceding discourse segment.

(418) The implication of all this is that large numbers of firms are unnecessary to achieve economic efficiency, and so it is equally possible for oligopolies to attain an efficient allocation of resources. Thus, on this view, attention should be focused upon the freeing of conditions of exit, rather than solely upon encouraging actual competition from increasing numbers of firms. It is not possible to enter into this debate here, for as yet these ideas have not gained universal acceptance. However, we do need to note that they are posing an important challenge to the conventional wisdom on the circumstances of resource misallocation. (*BNC Sampler*: HXN, W:commerce)

(419) [...] Re was at first pleased but soon realised that Hathor's delight in killing might lead to the destruction of all mankind, which he had not intended. He hastily ordered large amounts of red ochre to be brought from Elephantine and to be ground to powder by the High Priest of Heliopolis and stirred into seven thousand jugs of beer. The liquid looks very like human blood and was poured at night on to the fields. When Hathor arrived in the morning to continue her killing, she was completely deceived and began to drink. She <gap desc='figure'>soon became very drunk and forgot about mankind, so they were saved from destruction. This story forms a fairly complete myth of early origin, but in the New Kingdom texts it is followed by various additions, overlaid

with some degree of satire, which explain how certain things happened. (*BNC Sampler*: EVR, W:non_ac:humanities_arts)

5.3.4.2 *Antecedent*

As stated in 4.4.1, the identification of antecedents in this thesis rests on Gray's (2010: 173) taxonomy of antecedent types. They fall into two main types (local and global discourse) and four subtypes (noun phrase: simple and complex, sentence/clause, extended preceding discourse and overall discourse). Application of Gray's (2010) taxonomy to the *CQPweb* data in this study indicated the need for further specificity among the existing categories, and for the addition of antecedent types unaccounted for in the original classification. The motivation behind the enhanced antecedent taxonomy proposed here lies in the narrow scope of Gray's (2010) data, which apply only to research articles and anaphoric sentence initial *this/these*. The following describes all the antecedent types used in this study. They fall into two subsections: 5.3.4.2.1 for antecedents outside the shell-noun phrase (i.e. extra-syntagmatic) and 5.3.4.2.2 for antecedents inside the shell-noun phrase (i.e. intra-syntagmatic). Section 5.3.4.2.3 discusses the treatment of uninformative or identity specifics in this study.

5.3.4.2.1 *Extra-syntagmatic antecedents*

i) Overall Discourse (OV): The only instances of OV in the study sample are (420) and (421). *Myths* in (420) is part of a title introducing the nature of the contents to follow. *Jokes* in (421) occurs in a preface to a book of jokes. In both cases, therefore, the noun encapsulates a whole text, rather than only a portion or segment, as GB.ET.PR below does.

(420) MYTHS AND LEGENDS
THE GREAT UNIVERSITY OF STAMFORD
The Ancient University Myth
IN 863BC, King Bladud, the Trojan king of ancient Albion (or Britain) and the father of King Lear, founded the first university in the world at Stamford. (*BNC Sampler*: CBB, W:non_ac:humanities_arts)

(421) Mr Blackadder tolled me this buk is ful of jokes that are abowt as funny as getting your bottom caut in a bacon slicer. (*BNC Sampler*: CHR, W:misc)

ii) Global Extended Preceding Discourse (GB.ET.PR): As in Gray (2010), GB.ET.PR applies where shell-noun specifics cross sentence boundaries or where specifics, whilst present in the text,

may not be clearly delimited. Examples (418) and (419) illustrate clearly delimited discourse segments, whilst (411) and (412) show specifics scattered throughout the surrounding co-text.

Examples (422) and (423) deserve special mention. In (422), it is shown how GB.ET.PR applies to specifics which, whilst nominal in form (e.g. *an end to recession*, *low unemployment*, etc.), span sentence boundaries and thus favour a GB.ET.PR analysis. Although GB.ET.PR is primarily linked to intersentential realisations, examples (423) and (402) are two exceptions, as GB.ET.PR applies to realisations within a spoken turn (no full stop separation). In these instances, the occurrence of <pause> and <unclear> tags hampers the establishment of clear antecedent boundaries, hence the suitability of GB.ET.PR.

(422) I'm here because I believe that we share many objectives and because I know that there's more that unites us than divides us. I'm sure that when you adopted your working together slogan you were not thinking that it included employers. But why shouldn't it? Surely we all want to achieve the same objectives, an end to recession, low unemployment, a prosperous and fair economy, good education, training and health care. In short, a stable and caring society in which everyone can live in reasonable comfort and security and to which everyone contributesaccording to their means. (*BNC Sampler*: HLW, S:speech:scripted)

(423) So so what do you get for the thirty four <-|-> pounds <-|->[...] <-|-> Well <-|-> , this is the thing, we have a caretaker<pause> but er <pause> fo-- for his services, but I've been at the council and I have got nowhere, I have <unclear> from the corporation, there's an awful lot of people complaining but they'll do nothing about it! (*BNC Sampler*: FX5, S:brdcast:discussn)

iii) Local Sentence (LC.SC) and Local Clause (LC.CL): Unlike Gray (2010), where sentences and clauses are conflated under the sentence/clause category, the analysis in this thesis distinguishes both.

Sentential antecedents apply to whole sentences and to main clauses in complex sentences. The former is shown in (424), and the latter in (425). The antecedent in (425) is coded as sentential on account of its main clause status within a conditional sentence (i.e. *if you were in there and you've got a closed door, there could be a build of fumes*).

(424) Evil in Germany was so heinous, so unmistakable in the late 1930s, that all the world would assuredly join forces to wipe it out. Here was a nice

example of the solution posing the problem. (*BNC Sampler*: AEA, W:fict:prose)

(425) Like say, if you were in there like and you've got a closed door
<-|-> Mummy, <unclear> <-|->
<-|-> there could be a build of fumes.
When we going?
Do you see what I mean?
Yeah.
You know, that, that's the only thing <pause> <-|-> er<-|-> (*BNC Sampler*: KCN, S:conv)

Clausal antecedents comprise subordinate and coordinate clauses, as in (426) and (427). Verbless clauses, as in (428), are also coded as LC.CL. Appositive clauses following the head noun, as in (429), are tagged as LC.CL(AP).

(426) And our philosophy as a truly worldwide supplier of accounting software is that we should increase our presence in different parts of the world as our sales and marketing activity there merits it. (*BNC Sampler*: HDF, S:speech:unscripted)

(427) This in turn creates its own difficulties. At a local level sponsors so often are involved in the club on a more direct way but it isn't always seen as a problem and can be enormously beneficial to both parties. (*BNC Sampler*: J3W, S:meeting)

(428) 'Out of sight, out of mind, is the practice with youths of his age,' said the voice of experience. (*BNC Sampler*: CCD, W:fict:prose)

(429) Nothing in European life, even in Albania, retains that capacity to shock, even to shock people who have never left West Cumberland in their lives. (*BNC Sampler*: A8W, W:newsp:brdsht_nat:report)

If a particular shell noun is realised by several clauses belonging to one and the same sentence, LC.CL applies. Otherwise, if several clauses are shelled by a particular noun and these clauses span sentence boundaries, then the example is treated as GB.ET.PR, as stated above. Two examples of this use of LC.CL are (430) and (431), where *findings* and *points of accord* encapsulate a range of *that-* and *to*-infinitive subordinate clauses within one long sentence.

(430) An analysis and assessment of the findings of previous catalogue use studies may shed some light on the users' apparent change of heart towards the library catalogue since the advent of online systems. The overall results have been brought together in three major reviews by Krikelas, Hafter and Markey. These studies indicate that: (a) 25% to 50% of library users use the catalogue; (b) students account for the greatest proportion of the user population; (c) the catalogue is used

predominantly for known-item searching and use increases with users' educational level; (d) public library users do more subject searching than academic library users. (*BNC Sampler*: H0S, W:misc)

(431) Points of accord. By Reuter in Strasbourg. THE European Community leaders attending the summit here agreed to: Help Eastern European countries embracing democracy, and declare support for the reform policies of President Mikhail Gorbachev; Endorse reunification of the two Germanys Accelerate the EC's drive towards economic and monetary union and meet in the second half of 1990 to consider creating a single currency and a European central bank; Adopt, with the exception of Britain, a charter guaranteeing workers' rights in the single European market to be created in 1992. (*BNC Sampler*: A9M, W:newsp:brdsht_nat:report)

Examples where clausal antecedents span two or more spoken turns by one and the same speaker are treated as LC.CL, provided that all the speaker's turns form one sentence. For example, in (432), *having carpet down in one corridor* and *painting the walls* are two clauses linked to *such as?*, which in turn relates to the sentence where *some good things* belongs. Thus, since the underlined segment is clausal in form and may be easily integrated into the sentence where the shell noun occurs (i.e. *They're doing some good things at weekend, such as having carpet down in one corridor and painting the walls*), it makes sense to code the antecedent for *some good things* as LC.CL.

(432) <unclear> talking about school!
Yeah.
[...]School's banned!
They're doing some good things at weekend.
Oh aye? Such as?
Having carpet down in one <-|-> corridor!
<-|-> <laugh> <-|-> Nice.
And they're painting the walls! (*BNC Sampler*: KBG, S:conv)

iv) Local Simple Noun Phrase, Complex Noun Phrase and Pronoun (LC.SNP, LC.CNP, LC.SNP(PN)): As in Gray (2010), the difference between LC.SNP and LC.CNP lies in the use of postmodification in the latter (as in (435) and (436)), and the lack thereof in the former (as in (433) and (434)).

(433) Profitability of the model and model layout sector has declined due to increased foreign competition but market share has been maintained.

The directors anticipate improved performance in this area next year. (*BNC Sampler*: FEJ, W:commerce)

(434) The Prime Minister emphasised that even she shared the goal of monetary union, eschewing on this occasion the expression 'economic and monetary reform' which she has preferred in the past. 'We have been in favour of this objective since 1962,' she said. (*BNC Sampler*: A9E, W:newsp:brdsht_nat:report)

(435) After his wife's departure in 1789, Henry retired to the village of Bolas Magna in Shropshire to escape both the scandal and his heavy gambling debts. (*BNC Sampler*: CBB, W:non_ac:humanities_arts)

(436) Historically, one of the earliest frequency-locking phenomena to be studied in laser physics was the locking of three or more longitudinal modes. When many modes are involved, this phenomenon results in very short pulses — typically picoseconds — which have grabbed most of the attention because of their many applications. (*BNC Sampler*: J2H, W:ac:nat_science)

Juxtaposed or coordinated nominal antecedents are coded as LC.SNP or LC.CNP provided that they are all either simple or complex (as in (437), where they are both LC.SNP), and as LC.SNP^CNP when combining simple and complex noun phrases (as in (438), where specifics cover LC.SNPs like *spiritualism* and LC.CNPs like *things that are not real*).

(437) Their husbands objected at first, but soon stopped when the money started coming in and anyway a quarter of the women were on their own, who had been abandoned or divorced. Problems encountered included rape and wife beating. (*BNC Sampler*: H8W, W:essay:univ)

(438) <laugh> <pause> Well getting back to these fallen angels in what way must try and lead mankind today? <pause>
One of the big ways is, is spiritualism and things that are not real, things that are sort of er blind in the mind of the unbeliever as it were, in other words material things probably and things that have under revelation <pause> (*BNC Sampler*: KBX, S:conv)

If specifics occur in appositive noun phrases, the code used is LC.SNP or LC.CNP(RT.AP) if restrictive, and LC.SNP or LC.CNP(NR.AP) if non-restrictive. Examples (439) and (440) are two cases in point, with *murder* coded as LC.SNP(RT.AP), and *television and attractions*, as LC.SNP(NR.AP).

(439) The total of forty-two sins ranged from serious crimes like murder to minor wrongdoings like listening to gossip. (*BNC Sampler*: EVR, W:non_ac:humanities_arts)

(440) Entertainment, we divide our entertainment into two parts, television and attractions. (*BNC Sampler*: HYF, S:meeting)

Examples (441) and (442) illustrate the use of pronominal antecedents (see also (415) and (416) for similar instances of AF.INTRA&AF.INTER encapsulation). *It* and *this* are coded as LC.SNP(*IT*) and LC.SNP(*THIS*) respectively. *Not much* in (443) is coded as LC.SNP[QT.PN], as it is a quantifier pronoun, not a personal or a demonstrative pronoun (as in (441) and (442)). This is the only example of the type in the sample.

(441) How many voted <pause> to put, smack the Conservatives, and don't worry about politics I'm just using it as an example, in the face. (*BNC Sampler*: H4A, S:meeting)
(442) However, the labour market has absorbed far more IT Advanced Courses students in 1984 than were even available to employers in 1983. This was a positive finding. (*BNC Sampler*: H0H, W:non_ac:polit_law_edu)
(443) CONFUSED shareholders have been seeking help on the complex question of Eurotunnel warrants and what they are worth. Not much, is the short answer. (*BNC Sampler*: CEL, W:newsp:other:commerce)

v) Numeral (LC.SNP[NUM]): The only occurrence of this antecedent is in example (444):

(444) Right, now, what erm what would the answer be here
That would be thirty-four <pause> four <pause> er point seven one eight. (*BNC Sampler*: JJS, S:classroom)

vi) Local Adjective (LC.AJ): Adjectival antecedents are exemplified in (445) and (446). In (445), acrylic fabric is said to resemble wool in its being soft and warm (i.e. its characteristics are those of being soft and warm). In (446), the only word to describe her underclothes is an adjective, *deadly serviceable*.

(445) Acrylic fibres are made into fabric which is soft and warm, and consequently has similar characteristics to wool. (*BNC Sampler*: GUB, W:misc)
(446) Deadly serviceable, the only word for her underclothes. (*BNC Sampler*: AEA, W:fict:prose)

vii) Local Verb (LC.VR): The only instance of this antecedent type is (447), where *a good word* refers back to *nicked*, the past participle of *nick*.

(447) Judge, judge. Any more?
Nicked, well that's a good word, I like that, go on
Nicked, you're nicked. <pause> (*BNC Sampler*: FM7, S:unclassified)

viii) Local Prepositional Phrase (LC.PP): This antecedent category is only found with *way*, as in (448) and (449). In both instances, the noun encapsulates adjuncts or circumstances. In (448), *one of three strategic ways* shells several manner adjuncts, as is evident in the *by*-phrases. In (449), *this way* does not shell thoughts, but the sequence whereby these thoughts come to mind (i.e. *from one thing to another*).

(448) Finally, Stanley and Farrington (1981, 78) also conclude that public transport is only one<gap desc='figure'><gap desc='table'><gap desc='table'>component of accessibility and argue that accessibility deprivation can be alleviated in one of three strategic ways: By improving the conditions of the population; by redesigning the provision of facilities required by the population; and by providing public transport [orig: explicity] explicitly tailored to the population's needs. (*BNC Sampler*: FR2, W:ac:soc_science)

(449) As his mind rambles from the boy to the jersey to Ellen (who knitted it), to the sea and the Borough (via the fish Grimes is going to get rich by catching), back to Ellen and the future life they will never lead because Grimes cannot escape his past — as his mind rambles in this way, so the music probes our sympathies by constant allusion back to the variations in the interlude. (*BNC Sampler*: J55, W:non_ac:humanities_arts)

ix) Local Interjection (LC.IJ): The only instance of this antecedent category is (450):

(450) Now what is the opposite of glum ?
Wally — Cheerful .
Teacher — Quite right.
And the opposite of woe?
Wally — GIDDYUP ! (*BNC Sampler*: CHR, W:misc)

x) Local Reaction Signal (LC.RN.SG): This antecedent category accounts for *yes*/*no* answers to questions (Quirk et al. 1985: 444). Only two instances occur in the study sample, (451) and (452).

(451) In the British mind the distinction between constituency MdBs and list MdBs will seem sharp: must they not divide into two classes, first and

second? The West German answer is emphatically 'No'. (*BNC Sampler*: EW4, W:non_ac:polit_law_edu)

(452) Would you make any conscious decision to move slightly down market or less up market?
I myself wouldn't use that phrase. I think that er, er, what I would say is that er, we would follow the markets erm, where they lead us.
And the answer's yes. (*BNC Sampler*: HYF, S:meeting)

xi) Local Formula (LC.FR): The two examples below, both from the same text, are the only occurrences of this antecedent type in the study sample. Thus, LC.FR applies to scientific or mathematical rules, which, on account of their being letter and number combinations, do not fit into any of the other antecedent categories.

(453) For such techniques we will probably need to discover a number of algebraic laws involving WHILE. We have not needed any of these so far, because finite programs contain no loops. Five examples are given below, each of which is easily derived from our existing systems. (Each requires an application of Infinitary rule 1 and induction.) (W1) WHILE = <gap desc='formula'>(W2) WHILE <gap desc='formula'>(W3) WHILE = <gap desc='formula'>(W4) WHILE true <gap desc='formula'>(W5) WHILE b SEQ (P, Q) <gap desc='formula'>if no variable appearing in b is input or assigned to in Q. (*BNC Sampler*: G3N, W:ac:tech_engin)

(454) Other laws can be proved in much the same way (often rather more easily). Some examples are given below. a) [FORMULA]; b) [FORMULA]; if [FORMULA]; true and no variable in any bi is altered by P. c) [FORMULA]; provided U1 declares all global variables and channels used by P, and [FORMULA]; declares none of them. (*BNC Sampler*: G3N, W:ac:tech_engin)

xii) Non-linear text: Local Figure and Local Table (LC.FG and LC.TB): In line with Tadros (1994: 73), this category includes instances of non-linear cataphoric or 'advance labelling', where, as in (455) and (456), specifics are given as figures or tables, rather than in linear discourse segments.

(455) Nonetheless as Moseley (1980) has pointed out, there are many areas of overlap between urban and rural deprivation, particularly in the inner urban and outer rural areas, as Figure 6.2 shows. (*BNC Sampler*: FR2, W:ac:soc_science)

(456) Table 3.4 provides details of new issues of government stock between 1 October and 31 December 1991. (*BNC Sampler*: HY1, W:commerce)

Finally, mention should be made of two variants of some of the above antecedent categories. One such variant applies where a local antecedent occurs two or more sentences away from a shell-noun phrase. For want of a better tag, Local and Global Extended Preceding tags are blended into the new category of Local Extended Preceding, applying in the study sample to nominal, clausal and sentential antecedents (LC.ET.PR.SNP, LC.ET.PR.CNP, LC.ET.PR.CL and LC.ET.PR.SC). Examples (457) through (459) are three cases in point (LC.ET.PR.CNP in (457), LC.ET.PR.CL in (458) and LC.ET.PR.SC in (459)).

(457) SYRIA yesterday denounced the assassination of the Lebanese President, Rene Muawad, 17 days after he took office, and pointed the finger of suspicion at General Michel Aoun. Gen Aoun controls Lebanon's Christian enclave and opposes the presence of Syrian troops in the country. He rejected the authority of President Muawad, who was elected with Syrian backing. 'The hand of betrayal and treason has reached the martyr President Rene Muawad,' the official Syrian news agency, Sana, said. 'This ugly crime against Lebanon and its legitimacy followed a series of threats by the rebellious officer (Aoun). (*BNC Sampler*: A8J, W:newsp:brdsht_nat:report)

(458) In addition to providing information to the Science and Engineering Research Council on the initial destinations of IT Advanced Course students completing courses in 1984, a secondary objective was to test a method of monitoring the output from the courses that may be adopted in future years. [...] By involving course organisers in the distribution of destinations questionnaires and in the collection of completed questionnaires for despatch to IMS for analysis and interpretation, it was felt that this mechanism would assist the flow of information. In the parallel exercise to information engineering research students (see Part 0) a different method of data collection was used, with questionnaires being sent directly to named students in departments, together with a reply-paid envelope addressed to IMS. The difference in response rates achieved is of more than passing interest. 3.2 Preliminary Information In early summer 1984, IMS wrote to the organisers of the 77 conversion and specialist information technology Advanced Courses running in 1983 – 84 (see Appendix I). The course organisers were informed about the study and its objectives and were asked for their cooperation in data collection. (*BNC Sampler*: H0H, W:non_ac:polit_law_edu)

(459) Sir Thomas remonstrated with him, as was his duty, reminding him of his high calling and how he will be required, in accordance with royal custom, to make a match advantageous to the realm — that is to say with a foreign princess.' Joan felt a trifle faint and was grateful for the fact that she was sitting down. But her hands were trembling uncontrollably, giving away to the duchess the measure of her anxiety.

> 'And how did the prince make answer, my lady?' she asked. 'He refused to hearken to Sir Thomas's words of wisdom, disdaining to discuss the matter further. (*BNC Sampler*: CCD, W:fict:prose)

Another variant is shown in examples (460) through (463), where shell-noun specifics are given in quotes. In (460), LC.SNP^CNP[QU] accounts for the quoted simple and complex noun phrases shelled by *a clear vision*. In (461) and (462), *Alienor's parting words* and *Lavery's assessment* are coded as LC.CL[QU] and LC.SC[QU] respectively, given the direct speech clausal specifics in the former (*she had said X*), and the sentential specifics in the latter. In (463), where the quoted specifics comprise more than one sentence, the category used is GB.ET.PR[QU]. It is important to note that, if the quoted antecedent is part of the syntax of a particular sentence, as in (463) (*the study found that X*), intrasentential encapsulation applies (CF.INTRA in (463)). However, in examples like (462), intersentential encapsulation is more appropriate (in this case, CF.INTER), as shell-noun specifics are not syntactically attached to the previous sentence.

(460) 'The BBC must therefore have a clear vision if it is to retain its role as the cornerstone of British broadcasting and continue to command respect and admiration in Britain and throughout the world. 'In January John Birt laid out that vision of a wide range of high quality programmes, greater efficiency and accountability with value for money for licence payers. (*BNC Sampler*: CF6, W:newsp:other:report)

(461) She shivered of a sudden, and recalled Alianor's parting words. 'Not exactly a time to go visiting, lass!' that lady had said with a sigh. (*BNC Sampler*: CCD, W:fict:prose)

(462) This is not the only model to produce a close fit and in conclusion it seems that Lavery's (1975, 198) assessment of the use of models in forecasting recreation demand is still very relevant for the mid to late 1980s: 'Although our understanding of the whole and individual elements of the recreation system is very imperfect, we do know the main activities which generate the greatest growth and demand for space and we can identify current and potential conflicts with other land use activities'. (*BNC Sampler*: FR2, W:ac:soc_science)

(463) Cronin's study of catalogue non-users sought to determine the characteristics and attitudes of non-users and found that, 'Non-use does not appear to be a function of personal dissatisfaction with any aspect of the library services or staff. The library and its catalogue do not warrant high priority rating simply because they are not central to the day-to-day survival of the non user'. (*BNC Sampler*: H0S, W:misc)

5.3.4.2.2 *Intra-syntagmatic antecedents*

i) Local Complement (LC.CT): LC.CT covers instances where informative specifics occur in appositive *of*-phrases. The evidence analysed suggests three identifying features:

- The noun phrase is often deictically specific, as in (464) and (465). Exceptions are (466) and (467), where the noun is indefinite. *A time of recession* in (466) may be paraphrased as *a time/occasion when recession is happening*. The possibility of eliminating *time* with no change in meaning (*during recession*, *during the current recession*) further indicates that, in its introduction of the more informative element (*recession*), *time* serves a mere support function (see Mahlberg 2003 in 2.2.2.1.1). In (467), *a sense of their worth as women* involves giving these women a feeling of how worthy or important they are. The actual feeling or sense is, thus, that of their own worth, hence the LC.CT analysis.

(464) Nonetheless, you draw this distinction, despite the fact that the characteristic of both those areas is the characteristic of openness. (*BNC Sampler*: FMP, S:pub_debate)
(465) Given Theorem 2 above, the following theorem shows that we have achieved our objective of completely characterising the semantics of finite programs. (*BNC Sampler*: G3N, W:ac:tech_engin)
(466) Although I have to confess that at present, during a time of recession, it is increasingly difficult to attract the backing of galleries for such shows [...] (*BNC Sampler*: CN4, W:pop_lore)
(467) And I know a lot of things too with the girls who came to us, we tried to erm give them er a sense of their worth as a women and not to constantly be oppressed and to accept erm what their boyfriends did or said, and so on. (*BNC Sampler*: FY8, S:interview:oral_history)

- The *of*-phrase condenses information conveyed in a previous discourse segment. For example, in (468), out of context, *the myth of Bladud's University* suggests that the myth concerns a university and someone called Bladud. It is only by reference to the segment in italics that the distillation of this instance becomes clear, *Bladud's University* implying *Bladud's establishment of a university in Britain*. The shell-noun phrase in boldface may thus be paraphrased as *the existence of Bladud's University is fictitious or a myth*. In (469), *Sharp's vision of a monastic metropolis* picks up on a point made earlier about the large concentration of monastic academic halls in Stamford (see stretch in italics). Hence, the content of Sharp's idea

or image is the existence of a monastic metropolis in Stamford back in the Middle Ages. Unlike (468) and (469), in (470), the *of*-phrase does not provide informative specifics, insofar as *the myth of Osiris* informs only about the protagonist of the story, but not about its content. In (468) and (469), however, whilst condensed, the *of*-phrase specifies both content (*the existence of a University* and *of a monastic metropolis*) and protagonist (*the existence of a University founded by Bladud*). Thus, *of Osiris* in (470) is treated as specifics of identity, as it only restricts the reference of the head noun. Informative specifics in this case may be found in a previous discourse segment (GB.ET.PR).

(468) The story of King Bladud was first chronicled by Geoffrey of Monmouth (c. 1100–1154) in his Historia Regum Britanniae, a collection of mythology and history relating to the early history of Britain, which he claimed he had translated from a lost book of Breton legends. Geoffrey, however, makes no reference to *Bladud establishing a university in Britain*. John Hardyng, a mid fifteenth-century poet and chronicler, is the first to mention the story and quotes as his source the sixth-century Merlin of Caledonia, more commonly known as Merlin the Wizard. There is no evidence, though, for the existence of 'Merlin' and writings attributed to him are spurious. This implies that the myth of Bladud's University was invented by Hardyng, and was inspired by the 1333 secession. Blore in 1813 suggested: The tale was then copied by later historians and poets such as John Ross, the late fifteenth-century Warwick antiquary, John Higgins (working c. 1570–1602), Michael Drayton (1563–1631), and local historians such as Richard Butcher and Francis Peck. (*BNC Sampler*: CBB, W:non_ac:humanities_arts)

(469) *As these colleges became overcrowded, the large monasteries set up their own academic halls, including Sempringham Hall and Durham Hall in St. Peter's Street, Peterborough Hall in All Saints' Street and Vaudey Hall (owned by Vallis Dei Abbey near Grimsthorpe) in St. Mary's Street. Each monastic order established a religious foundation in the town resulting in a proliferation of churches* until as Samuel Sharp says: '*Stamford…must have been a kind of monastic metropolis.*' [...] Even if there were monastic or friary-owned schools in the town they would still not constitute a university, as each school would be administered separately and would be dominated by the beliefs of the monastic order. Samuel Sharp's vision of a 'monastic metropolis' is an exaggerated fantasy and the reason for the large number of churches in the town was not monastic but the result of early self-government and the absence of any single overlord other than the king. (*BNC Sampler*: CBB, W:non_ac:humanities_arts)

(470) Three main themes are contained in the myth of Osiris, which are political, agricultural and ritual and which serve to explain and justify

particular historical events and the growth of certain beliefs and practices. (*BNC Sampler*: EVR, W:non_ac:humanities_arts)

- The information contained in appositive *of*-phrases is often eventive or dynamic (i.e. something happening). Non-appositive *of*-phrases, by contrast, are semantically more static or abstract. Examples (471) through (474) illustrate both uses. In (471), the interference of curtain fabric with the doors is a problematic event, hence its treatment as appositive LC.CT. In (472), however, the reader knows that there are intricate problems linked to a particular field of study (*structural geology*), but not what these problems are. The *of*-phrase is thus treated as specifics of identity (LC.CT(SI)). In (473), the suspicion is that an undesirable interaction between roof ventilation and sprinklers is happening, hence its analysis as appositive LC.CT. In (474), however, the reader is told about the subject of Earl John's suspicion (*enthusiasm*, *religious fervour* and *triumphalism*), but not what this suspicion actually involves. In other words, the reader is informed about the suspected abstract entities, but not about the implications behind the suspicion. On these grounds, the *of*-phrase is treated as specifics of identity (LC.CT(SI)).

(471) Doors that swing inwards present the problem of curtain fabric interfering with the operation of the doors. (*BNC Sampler*: GUB, W:misc)

(472) On his appointment in Scotland he formed part of a group of geologists who, in the period before World War I, were internationally without parallel for their original insights into metamorphic and igneous rocks and the intricate problems of structural geology. (*BNC Sampler*: GT9, W:biography)

(473) It would appear that any suspicion of undesirable interaction between roof ventilation and sprinklers would be unjustified in the case of a zoned sprinkler system. (*BNC Sampler*: G0K, W:institut_doc)

(474) He took a far more realistic approach to the revival and certainly did not share their conviction that the Oxford Movement would lead to the reunion of the Church of England with the Church of Rome. Although hailed by the leaders of the Catholic revival as their patron, 'the good Earl John' (as he was known by them) retained something of the recusant Catholic families' suspicion of 'enthusiasm', religious fervour, and triumphalism. (*BNC Sampler*: GT9, W:biography)

In addition to appositive *of*-phrases, LC.CT also applies where a nominalised prepositional complement stems from a verbal prepositional object. The prepositional complement in these instances may be paraphrased as an appositive *to*-infinitive clause

in the typical N-*to* or N-*be-to* shell-noun patterns. This is shown in (475) and (476), where the shell-noun phrases in boldface allow paraphrases as *an application to obtain funds* and *their challenge* (i.e. *attempt*) *to get promoted to the East Super League*.

(475) The criteria for deciding whether the banks would support **an application** for funds led many banks to the conclusion that, if a case warranted support, the government's guarantee was superfluous [...] (*BNC Sampler*: G0C, W:commerce)

(476) ESSEX **champions Braintree's challenge** for promotion to the East Super League started well in the first of three play-off matches against Pelicans, from King's Lynn. (*BNC Sampler*: CF9, W:newsp:other:sports)

ii) Local Premodifier (LC.PRM): Flowerdew (2003a: 338) argues that premodifying Classifiers like *division* in *the division process* specify '[...] the signal [...] in part'. This is also the case with many of the Classifier instances in the study sample, as in (477) and (478). *Community* and *Channel Tunnel* restrict the semantic scope of the head noun, but do not fully specify what the *dimension* and *project* involve. More informative specifics occur elsewhere. Thus, in (477), the underlined segment reveals that *mergers with a Community dimension* are those where *the parties have an aggregate turnover in excess of ECU 5bn*. In (478), the underlined specifics show the stages and objectives underlying the construction of a tunnel between England and France, i.e. *the Channel Tunnel*. Premodifying Classifiers in these two examples are treated as specifics of identity (LC.PRM(SI)), as they indicate what the head noun relates to, but not what it actually involves.

(477) This now enables the Commission to investigate and control those 'concentrations' (i.e. mergers and takeovers) which have a 'Community dimension', while those mergers not having **a Community dimension** remain subject to domestic policies. Mergers with a Community dimension are defined as those where the parties have an aggregate turnover in excess of ECU 5 bn, and where at least two of those parties have a Community turnover greater than ECU 250 m. (*BNC Sampler*: HXN, W:commerce)

(478) However, the history of the Channel Tunnel is generally said to have begun with the 1820 proposal of French mining engineer Albert Mathieu-Favier for a road tunnel lit by gas and ventilated through chimneys emerging above sea level. Napoleon is believed to have been warmly in favour even though the tunnel was not designed for military purposes. [...] In 1957 a Channel Tunnel Study Group with British, French and American membership undertook extensive survey work

and considered the relative merits of a range of fixed link designs, finally opting in 1960 for a twin bored tunnel with accompanying service tunnel. [...] In March the formal concession was signed Eurotunnel for the construction of two 7.6m diameter running tunnels and a 4.8m diameter service tunnel between Britain and France, and the operation of shuttle train service for the carriage of road vehicles. [...] During the passage of the bill the various participants in **the Channel Tunnel project** continued their preparations. (*BNC Sampler*: BMJ, W:misc)

Despite the close association between Classifiers and specifics of identity, there are cases where LC.PRM is treated as informative specifics. This applies particularly to indefinite and plural nouns with generic reference. Their generic orientation precludes the existence of more informative specifics, and turns the restrictive modifier (as in (477) and (478)) into the only informative anchor for the understanding of the shell noun. For example, in (479), *a membership application* entails *an application for membership* or *an application to obtain membership*, showing how a postnuclear complement is here condensed into a premodifying Classifier. In (480), the Classifier is analysed as informative specifics on the grounds that *flood* is an example of the '[...] bad, dangerous, or annoying [...]' (*LDCE*) things that tend to be predicted by *warnings*. The acceptability of a paraphrase such as *warnings that a flood is underway or about to happen* is further proof of the LC.PRM antecedent in this particular example.

(479) With the system correctly set up, the whole process can be completely, completed automatically by the use of technology and it could be set up speedily upon receipt of a **membership application**. (*BNC Sampler*: HLW, S:speech:scripted)

(480) FORECASTS of the highest tides for decades have prompted an Essex council to distribute flood warning leaflets to people living in low-lying areas. [...] 'We certainly don't want to alarm anyone, but it is important for people to know what they can do in advance to prepare their homes against damage.' If necessary **flood warnings** will be issued by the National Rivers Authority and police will try to warn residents using local radio, television and loud hailer vans touring risk areas. (*BNC Sampler*: CF5, W:newsp:other:report)

5.3.4.2.3 *Specifics of identity*

Drawing on Winter (1992: 154; see 2.2.2.1.2), pre- and postmodifiers restricting the reference of the head noun, but failing to provide truly informative specifics, are coded as Specifics of Identity. As stated in

2.2.2.1.2, these specifics indicate what the noun relates to, but not what it actually involves. Identity specifics comprise most premodifying adjectival and nominal Classifiers, relative clauses, participle *-ing* and *-ed* clauses, non-appositive *to*-infinitive clauses, and prepositional phrases other than appositive *of*-phrases. The following three examples illustrate the use of participle *-ed* clauses (*faced by many companies*, as in (481)), complement prepositional phrases (*about the quality*, as in (482)) and both Classifiers and prepositional complements (*socio-demographic*, *of movers*, as in (483)). Each example is followed by a tag combining the annotation for direction of encapsulation and antecedent. Identity specifics are coded by attaching (SI) to the their realisation (e.g. LC.CL(SI), LC.PRM(SI), LC.CT(SI), etc.)

(481) Overcoming a natural resistance to change is a challenge faced by many companies that want to progress. (*BNC Sampler*: J24, W:misc)
[CF.INTRA&AF.INTRA-LC.CL(SI)&LC.CL]

(482) Dislike of party lists is often fortified by dark suspicions about the quality of the candidates they supposedly include. These, it is argued, will be mediocrities, no more than 'good party men', 'loyal apparatchiks'. (*BNC Sampler*: EW4, W:non_ac:polit_law_edu)
[CF.INTRA&CF.INTER-LC.CT(SI)&LC.SC]

(483) The migration reversal is as pervasive across the socio-demographic characteristics of movers as it is across regions. (*BNC Sampler*: W:ac:soc_science)
[AF.INTRA&CF.INTRA&EXO-LC.PRM(SI)&LC.CT(SI)&NA]

In relation to LC.CT identity specifics, prepositional phrases are treated as complements if featuring in any of the noun patterns in Francis et al.'s (1998) grammar or in the *Oxford Collocations Dictionary* (McIntosh et al. 2009; henceforth, *OCD*). This is illustrated in example (484), where the *to*-phrase following *testimony* is coded as LC.CT(SI). LC.CT(SI) also applies to nominalised verbal arguments, as in (485), where the *by*- and *of*-phrases following *assessment* correspond to the subject and object of the verb *assess*. All other prepositional identity specifics are coded as postmodifiers, LC.PM(SI), as in (486), where *since the NHS reforms* introduces an optional temporal restriction on *experiences*.

(484) External aspects of religion such as temples, statues, reliefs and paintings of the many gods and goddesses, sacred object, writings and burial customs survive as a unique testimony to that religious experience but its inner meaning and significance are far more intangible. (*BNC Sampler*: EVR, W:non_ac:humanities_arts)

(485) New details of the summit conversations are emerging in the course of White House briefings, including a surprisingly frank assessment by the Soviet leader of the looming economic crisis. Mr Gorbachev told

the Americans that the Soviet people would judge his perestroika by how quickly he could get food and consumer goods into the shops. (*BNC Sampler*: A8W, W:newsp:brdsht_nat:report)

(486) And one of my experiences since the NHS reforms has been working in hospitals and hopefully defending our members' interests. (*BNC Sampler*: H4A, S:meeting)

Although relative and participle clauses are primarily linked to identity specifics, there are cases where they occur as lexical realisations of the shell noun. For example, in (487), the clause following *in which* spells out one of the two opposite or countervailing visions presented in this passage, thereby favouring its treatment as informative specifics. In (488), the relative clause following *philosophy* reveals the content of such a view or attitude (classing people as units of labour) and, on these grounds, it is coded as informative LC.CL specifics. A similar example is (489), where the participle *-ing* clause details the aim of the project (to compare dog-whelk shells from exposed and sheltered shores).

(487) I was drawn to contrary poles. I yearned towards the mystical earnestness which saw through the outer facing of existence in a oneness and blinding intensity which went direct to some essence of being. But I could appreciate a countervailing vision of tolerant scepticism in which the surface texture of life was a source of amiability and pleasure. (*BNC Sampler*: FU7, W:non_ac:soc_science)

(488) We hope that sooner or later the philosophy that classes people as mere units of labour will be consigned to the dustbin of history which it deserves. (*BNC Sampler*: HLW, S:speech:scripted)

(489) On arrival at Dale Fort Field Centre (overlooking Milford Haven) in 1963, I found John Barrett's staff using a project on their A-level biology courses involving the comparison of dog-whelk shells from exposed and sheltered shores. (*BNC Sampler*: FU0, W:ac:nat_science)

In the examples presented so far, identity specifics are realised by intra-syntagmatic elements, as originally conceived by Winter (1992). Interestingly, the analysis in this thesis also reveals instances of extra-syntagmatic (SI) realisations. These fall into three main groups:

i) Relational meanings: If a reader is presented with the (a) versions of examples (490) through (492), he/she may be left wondering what *variegated forms of Geranium macrorrhizum* are examples of, what something is a detail of and what the evidence removed is of. The complete passages in (b) show that *variegated forms of Geranium macrorrhizum* are examples of *varieties whose foliage is*

misshapen, that *every detailed she recalled* is of *what had taken place* and that *the evidence removed* (i.e. *the bodies*) is of *five murders*. *Example*, *detail*, *evidence*, *characteristic* and *facet* are all examples of 'relational nouns' (Keizer 2007: 64). Their relational meaning stems from their use with *of*-phrases, which denote the reference points whereby the head noun is to be interpreted. In the absence of the *of*-phrase, the attributive component inherent in these nouns is syntagmatically not explicit, but contextually implied. Such contextual implication is evident in the following three examples, where the italicised segments bring to light the attributive or partitive element missing from the intra-syntagmatic realisation of the shell-noun phrase. These extra-syntagmatic identity specifics are realised by a complex noun phrase in (490), by a clause in (491) and by a simple noun phrase in (492).

(490) (a) Examples include variegated forms of Geranium macrorrhizum and Prunus lusitanica.
(b) Avoid *varieties whose foliage is misshapen*. Examples include variegated forms of Geranium macrorrhizum and Prunus lusitanica. (*BNC Sampler*. C9C, W:pop_lore)

(491) (a) She vividly recalled every detail.
(b) Joan was recalling the day, not much more than a year ago, when she had found herself alone in a locked room with the innkeeper. She remembered the moment when he had locked the door and she had known herself trapped — she could still hear the sound of Rose Trivet crying her heart out in another part of the house. [...] But Joan was harking back to that awful day. Still she could hear her own cries, as she had begged to be released — but nay, he had flung her down on the stone floor and leaned over her. She had twisted and turned, struggling to free herself to no avail, until a well-aimed kick had won her a temporary respite. With an obscene oath he had released her. 'You little slut!' he had said through clenched teeth, his face ugly with fury — and some other emotion she had not then recognized. 'You'll pay for that, you trollop!' [...] She gave thought to it now, dwelling for the first time, without fear, without flinching, on *what had taken place*. She vividly recalled every detail. (*BNC Sampler*: CCD, W:fict:prose)

(492) (a) [...] his bizarre method of removing the evidence [...]
(b) What made the whole case so remarkable was his confession to *five other murders*, his bizarre method of removing the evidence — by dissolving the bodies in acid — and his claims that he indulged in vampirism. (*BNC Sampler*: CBB, W:non_ac:humanities_arts)

In addition to Factual Attributive and Partitive nouns like *aspect*, *characteristic* or *example* (Schmid 2000: 116), Schmid's (2000: 218–19) family of Mental Volitional Detached 'Purpose' nouns (e.g. *idea*, *purpose*, *function*) is also markedly relational. These nouns shell the aim that an Experiencer attributes to a particular entity. The occurrence of the entity is thus required for the identification of the intended aim. For example, in (493) and (494), full understanding of *the final objective* and *these applications* needs not only the informative specifics in each case (what the objective is and what the applications are), but also what *extinguishing the fire* is an objective of and what things like *proving programs equivalent to one another* or *transforming programs* are uses, functions or applications of. The *of* component is here implied in the italicised simple noun phrase in (493) (LC.SNP(SI)) and the complex noun phrase in (494) (LC.CNP(SI))

(493) It is realistic to require, therefore, that *the automatic system* shall have achieved substantial control i.e. to have prevented further upward spread, by three minutes, before the flames rise out of the original level, or at least have not ignited the goods in the level above it. At this stage in the fire's development, sideways spread is not a problem and can easily be constrained by a good protective system. **The final objective** is to extinguish the fire, and this should occur within eight to ten minutes of ignition. (*BNC Sampler*: G0K, W:institut_doc)

(494) Nevertheless, there are a number of practical applications for *the laws described in this paper*: proving programs equivalent to one another, transforming programs to make them more efficient, and transforming programs to a restricted syntax for special applications. In the three following subsections we examine their potential for **these applications**. (*BNC Sampler*: G3N, W:ac:tech_engin)

ii) Discourse patterns: The italicised segments in (495) and (496) are coded as identity specifics by virtue of their occurrence in such frequent discourse patterns as Problem-Solution and Question-Answer (see Winter 1977, 1982 and Hoey 1983 in 2.2.2.1.2). If presented only with *the negative answer* or *automatic fire fighting installations are the only answer*, a reader or listener would be justified in asking what *negative* (i.e. *no*) is an answer to and what *automatic fire fighting installations* are a solution to. The principle here is that full understanding of the contextual significance of *answer* can only be gained by reference to what the answer is (i.e. informative specifics) and to what problem or question it relates to

(i.e. identity specifics). The identity specifics in (495) and (496) are coded as LC.CL(SI) and LC.SC(SI).

(495) Tish refers to all those who will ask us *'have you seen…?'* Then, when faced with the negative answer, 'Really … ?' From a train window the views are democratically arrayed so that basilica and hedgerow, back street and castle have equal viewing time and space. (*BNC Sampler*: CN4, W:pop_lore)

(496) *However, the internal layout of high-bay warehouses, the high degree of contents-density, limited access and the likelihood of rapid fire spread reduces the possibility of success by conventional fire fighting methods*. Automatic fire fighting installations capable of achieving total control and extinguishment appear to be the only answer. (*BNC Sampler*: G0K, W:institut_doc)

iii) Names: If the shell noun refers to the name of a company, system, etc., the name is treated as identity specifics, as it merely labels the noun without specifying its content. This is evident in examples (497) and (498), where *Model Marketing Company* and *the Telford system which relies largely on professional judgement* are coded as LC.SNP(SI) and LC.CNP(SI). The purpose of the joint venture and the procedures in the system are found in an intrasentential *to*-infinitive clause and in an intersentential discourse segment respectively (LC.CL in (497) and GB.ET.PR in (498))

(497) The company also became a party to a joint venture, *Model Marketing Company*, set up to market model layout products in the Far East and Australasia. (*BNC Sampler*: FEJ, W:commerce)
(CF.INTRA=>LC.SNP(NR.AP))

(498) However, this is no excuse for inaction, for a number of systems have been described which are of considerable assistance; for example, *the Telford system which relies largely on professional judgement* can be used to good effect (1). A 'bottom up' approach is used to construct a manpower profile. Each patient area manager is asked to make an assessment of manpower needs and required to justify this assessment to other nurses. The assessment relies largely on a combination of operational experience and professional judgement. A recent report by the DHSS Operational Research Service reviews methods for calculating nurse demand and recommends an evolutionary — start simple and refine as necessary — approach which has built-in evaluation and cross-checking elements to allow judgemental inputs to be scrutinized and upon which the process of informing, questioning and understanding can be based (1). Once completed the exercise enables comparison to be made between hospitals/specialties and thus the

> inequities in financial distribution to be identified. (*BNC Sampler*: EVY, W:commerce)

5.3.5 *Semantic type of shell noun*

As stated in 3.2.5.1, this thesis follows Schmid's (2000) semantic classification. Its fine-grained and comprehensive treatment of shell-noun meanings contrasts with the all-embracing nature of other classifications (e.g. Francis 1986 and 1994; Hunston & Francis 2000). In 3.2.5.2, it was noted that, whilst not denying the valuable contribution of Schmid's (2000) taxonomy to shell-noun description, its failure to account for potential shell senses stems from the author's primary focus on the most frequent shell nouns in the N-cl and N-*be*-cl patterns. Although the semantic tags used here are in the main Schmid's (2000), the approach of this thesis required important decisions to be made in two respects:

i) Lemmas unaccounted for in Schmid (2000) were accommodated in existing, modified or new categories.
ii) Lemmas included in Schmid's (2000) list but for which certain senses are unaccounted for were coded on the basis of existing, modified or new categories.

The above points are developed in sections 5.3.5.1 and 5.3.5.2. Prior to this, however, it should be mentioned that, as with Transitivity in 5.3.3.1, the contextualised analysis of data in this section is assisted by the definitions provided in any of the following dictionaries: *CCD*, *OALD*, *LDCE* and *OED*. Further assistance is provided by classifications of nouns in Francis (1986), Francis (1994), Francis et al. (1998) and Hunston & Francis (2000), and by corpus evidence from *BNCweb* in cases of assimilation of the semantic frame associated with particular lexico-grammatical patterns.

5.3.5.1 *Lemmas unaccounted for in Schmid's (2000) list*

Eleven of the 60 lemmas analysed in this study are not in Schmid's (2000) list: *correction*, *detail*, *dimension*, *endorsement*, *foreboding*, *leave*, *opposite*, *sense*, *system*, *word* and *work*.

i) Correction: Francis (1986: 12) labels *correction* as an illocutionary noun, illustrating it with example (499). One of the two instances of this lemma in the study sample is (500). For want of a better category, *correction* is coded as Linguistic Illocutionary Assertive

Argumentative 'Amendment' (Schmid 2000: 157) on the grounds of its similarity to *amendment*.

(499) '[...] that there had been a misunderstanding, and that only 1,000 Jews had been killed. **This 'correction'** was, of course, quite misleading' (Francis 1986: 52)
(500) However, it may be helpful if the system can sometimes suggest a **correction for a miskeyed word**. (*BNC Sampler*: H0S, W:misc)

ii) Detail: The evidence for this lemma falls into two of Schmid's semantic categories: Factual Attributive Part-Whole 'Aspect' and Linguistic Propositional 'News' (Schmid 2000: 116–18, 140–4). The former is evident in example (501), where *detail* carries the sense of 'a minor point or aspect of something, as opposed to the central ones' (*CCD*). By contrast, in (502) *detail* is synonymous with *information*, one of the nouns included in Schmid's 'News' class.

(501) More subtle is the way the minor second also becomes symbolic of the Storm, again starting as **a purely graphic detail**. (*BNC Sampler*: J55, W:non_ac:humanities_arts)
(502) Official news reports of the Assembly debate failed to reveal **details of the controversial content of the law**, especially its much-criticised Article 5. (*BNC Sampler*: AAT, W:newsp:brdsht_nat:report)

iii) Dimension: Example (503) illustrates the expected shell sense of *dimension*, that of 'aspect' or 'a part of a situation or a quality involved in it' (*LDCE*). Thus, similarly to the 'aspect' meaning of *detail* above, instances like (503) are coded as Factual Attributive Part-Whole 'Aspect'.

(503) An unemployed member now could be a full member in the near future, possibly in firms where we have previously had no members. There is **another dimension to this motion that needs to be considered**. If we continue to ignore these people, especially the young, we may be fighting them in the near fu--future as they are actively being targeted by the far right British National Party and other fascist organizations that are using this sense of desertion. (*BNC Sampler*: HLW, S:speech:scripted)

Examples (504) and (505) represent two semantic rarities, insofar as *dimension* acquires a meaning that is not inherent in the lemma itself, but one linked to a specific phraseology and context of use. Example (504) is coded as Specific Circumstantial Manner 'Way' (Schmid 2000: 284–9), based on its semantic similarity to *way*,

approach and *method*. This is confirmed by *BNCweb*, where a search of '*the simplest _{N} is to*' returns 32 hits, 68.77% of which (22 tokens) occur with *method*, *way*, *approach*, *means* and *procedure*. Example (505) is also treated as Circumstantial, but this is coded as General Circumstantial 'Situation' (Schmid 2000: 277–9). This interpretation is borne out by the context, where *the majority of voters in Northern Ireland will not have an institutionalised Irish dimension* seems to imply: *these voters will not tolerate a situation/scenario/context involving an institutional attachment to the Republic of Ireland*.

(504) Not surprisingly, therefore, most approaches have been based on the cultural dimension with the addition of ecological and occupational factors. The simplest cultural dimension is to use local authority administrative areas but these often lag behind significant changes in population composition, and although Robertson 's (1961) classification already outlined, defined 75 per cent of the areas as rural-urban, she was also able to comment that this reflected not only the extent of the outward movement from towns, but also the inadequacy of the census definition of rural population, namely all persons living in administrative Rural Districts. (*BNC Sampler*: FR2, W:ac:soc_science)

(505) Thus the percentage of the poll gained by UUUC was again over 50% in the third election within fifteen months and the result of the election suggests again that the majority of voters in Northern Ireland will not have an institutionalised Irish dimension and that they are profoundly suspicious of any party which is prepared to share power with those whose explicit or implicit aim seems to be to 'destroy the state'. (*BNC Sampler*: H7C, W:non_ac:polit_law_edu)

iv) Endorsement: All instances of *endorsement* (as in (506)) are tagged as Linguistic Illocutionary Assertive Argumentative Defending (Schmid 2000: 157). Its definition as 'a public statement or action showing that you support somebody/something' (*OALD*) supports this interpretation, as it is synonymous with other Defending lemmas like *affirmation* or *defence*. The illocutionary nature of this lemma is confirmed by Francis (1986: 12), who includes *endorsement* in her list of illocutionary nouns.

(506) Moderator could the convenor tell us why the Board wishes to be discharged its remit on health and healing when it's been so successful in bringing this to the notice of the whole church? <pause>
<unclear> grateful to Mr <gap desc="name" reason="anonymization"> for his very erm generous en-- endorsement of the work that the

Board has done on health and healing. (*BNC Sampler*: F86, S:meeting)

v) Foreboding: This lemma is treated as Mental Emotive Event-Related 'Fear' (Schmid 2000: 228–30). Its definition as 'a strong feeling that something bad is going to happen soon' (*LDCE*) ties in with Schmid's (2000: 228) claim that Event-Related 'Fear' nouns (e.g. *fear*, *concern*, *worry*) represent states of emotion '[...] caused by IDEAS representing future events' (capitals as in the original). The future orientation of *foreboding* is illustrated in example (507). The fear in this particular case concerns the sense of impending doom linked to a pre-war context (note the sense of futurity in the lexis employed: *imminent*, *is about to be shattered*, *it will never be the same*).

(507) 'War's imminent!' [...] We may not see this place again for a long time.' [...]It's awful; I feel somehow that our life, our happiness, our everything is about to be shattered.' [...] And then after a long pause, he sighed, 'But it will never be quite the same.' [...] The sense of foreboding was invasive. It tinged their love-making. (*BNC Sampler*: AEA, W:fict:prose)

vi) Leave: *Leave* is coded as a Modal Deontic Possible 'Permission' noun (Schmid 2000: 245–7) based on its similarity to such other items as *permission* or *licence* (see (508)).

(508) Likewise therefore this appeal is not one which entitles any party to call evidence unless there are exceptional circumstances and leave is given (*BNC Sampler*: FCF, W:ac:polit_law_edu)

vii) Opposite: For want of a better semantic tag, all instances of this noun are assigned to Schmid's (2000: 113) Factual Comparative 'Difference' class, comprising such nouns as *difference*, *alternative*, *distinction*, *contrast* and *discrepancy* (see (509)).

(509) Perhaps only because Green is the opposite of Red many people said Green when asked what party they would support in the forthcoming elections. (*BNC Sampler*: AAT, W:newsp:brdsht_nat:report)

viii) Sense: The evidence for this lemma in the study sample corresponds to four semantic tags, i.e. Mental Creditive 'Belief-Feeling' (as in (510)), Mental Creditive 'Belief-Understanding' (as in (511)), Mental Volitional Conditional 'Determination' (as in (512)) and Factual Causal 'Reason' (as in (513)).

(510) At the same time, she felt a sense of guilt, of disloyalty, that she could even for an instant contemplate the idea of deserting her mother and turning her back on the place she called home. (*BNC Sampler*: CCD, W:fict:prose)

(511) The existence of monopoly denies them that opportunity, and this is manifest in the inevitable reduction in total surplus. It is in this sense that monopoly is said to be economically inefficient, and to misallocate resources through the restriction of output. (*BNC Sampler*: HXN, W:commerce)

(512) If somebody else was trying to book it, one of those Welshy types, they haven't got the sense to <pause> <-|-> ask <unclear> <-|-> (*BNC Sampler*: KD8, S:conv)

(513) Now I can't see any sense in throwing <pause> whatsit, the javelin. (*BNC Sampler*: KC4, S:conv)

The former two meanings draw on Schmid's (2000: 195–203) Mental Creditive 'Belief' category, the most general of his group of psychological-state uses. This category comprises such nouns as *idea*, *belief*, *hope*, *understanding* or *knowledge*. In this study, the tag is slightly modified by the addition of 'Feeling' in (510) and 'Understanding' in (511). This is in order to account for the difference between *sense* as 'a feeling about something important' (*OALD*) (*a feeling of guilt*) and *sense* as 'a way of understanding something' (*OALD*) (*from this understanding*, *from this way of understanding*). The last two semantic categories are found only with two examples: (512) and (513), respectively. Schmid's (2000: 222–4) Mental Volitional Conditional 'Determination' class contains nouns such as *determination*, *confidence*, *courage*, *readiness* or *heart*, all of which imply psychological states which may contribute to the achievement of an aim. Thus, in (512), *having the sense to ask* is roughly equivalent to saying *having the sense necessary to ask*. *Sense* in (513) is classed as Factual Causal 'Reason' (Schmid 2000: 102–6) on the grounds of its similarity to *can't see any reason in throwing*. The 'reason' meaning is evident in the definition given by *CCD*: 'If you say that there is no sense or little sense in doing something, you mean that it is not a sensible thing to do because nothing useful would be gained by doing it'. The definition given by the *OCD* is even more closely related to the 'reason', where *sense* in this construction is equivalent to 'a sensible or practical reason'.

ix) System: As a shell noun, *system* means 'an organized set of ideas, methods, or ways of working' (*LDCE*). On the basis of its semantic affinity with such other items as *approach*, *method* and *procedure*,

system is classified under Schmid's (2000: 284–9) Specific Circumstantial Manner 'Way' nouns (see (514)).

(514) The West German system is at the same time a plurality or "first-past-the-post" system, a list system and an additional member system: in short a mixed system, which may be conveniently designated in the following pages by the initials WGMS. (*BNC Sampler*: EW4, W:non_ac:polit_law_edu)

x) Word: As mentioned in 3.2.5.1, only Francis (1986: 16–17; 1994: 93) allows for the existence of 'text nouns', whose function is to assign formal labels to the structural components of discourse (ranging from *term*, through *passage*, *quotation* and *paragraph*, to *page* and *text*). According to Lyons (1977, II: 667–8, see 3.2.2), nouns of this kind would represent typical examples of pure textual deixis, as their cohesive contribution is more formal or syntactic than semantic. Thus, in cases like (515), reference is made primarily to the linguistic form of *nicked*, this being a word. Whilst evaluation is not absent from these items (as evident from the interpersonal Epithet *good*), emphasis is on their pointing or deictic use. Thus, a new semantic category is created in this thesis for instances like (515), that of Linguistic Metalinguistic Textual Deixis 'Word'. It should be noted that, whilst singular *word* is excluded from Francis' (1986 and 1994) list of text nouns, *term*, largely synonymous with *word*, is not. In cases like (516), plural *words* is tagged as Linguistic Metalinguistic Textual Deixis 'Text', as the noun refers to a whole text (the lyrics of a song).

(515) Nicked, well that's a good word, I like that, go on (*BNC Sampler*: FM7, S:unclassified)

(516) Give us an A please.
<voice quality: singing> <-|-> La la la la la la la la la la la <-|-> la la la la la la la la la la. <end of voice quality>
Are the words in English to that? (*BNC Sampler*: FXR, S:meeting)

Among the *word* examples, there are instances where the noun is not intended as a mere formal label for a discourse segment ((515) and (516)). In examples such as (517) and (518), the focus seems to shift to what is being said, i.e. to the content of the message. Hence, based on their similarity to *news*, *information* or *message*, these instances are coded as Linguistic Propositional 'News' (Schmid 2000: 140–4). This means that in (517) and (518) attention is not drawn to the fact that the Queen's and Alianor's messages

are expressed in a text or in spoken words, but to the content of the messages themselves.

(517) 'I am grateful, sir,' Edward replied with a smile. 'I had not till now received word from the queen my mother.' [...] The letter from their mother was a joy and an inspiration to both princes. But so graphically did her words convey her love and concern for them that, reading it aloud, both were reduced to tears. 'Have faith in the future, my princes,' she had concluded, 'and be of good cheer — you are your father's sons, and he who laughed always at adversity would have been proud of you. 'God send you good keeping. 'Written this day at Westminster by the hand of your loving mother 'ELIZABETH WOODVILLE' (*BNC Sampler*: CCD, W:fict:prose)

(518) She shivered of a sudden, and recalled Alianor's parting words. 'Not exactly a time to go visiting, lass!' that lady had said with a sigh. (*BNC Sampler*: CCD, W:fict:prose)

Lastly, (519) and (520) are two marginal categories. In (519), *words* is tagged as Linguistic Illocutionary Assertive Retrodictive 'Report' (Schmid 2000: 160–1) based on its link to such other nouns as *report*, *story*, *account* or *tale*. The Retrodictive component of these units lies in their sharing '[...] the specific feature that the propositional content of the reported utterance represents a previous event' (Schmid 2000: 160). The past orientation of *words* in (519) is shown in the historical significance of the account or version given by two survivors of a massacre that took place a long time ago. *A word* in (520) implies a short conversation. This use of *word*, frequent in conversation, is coded as Linguistic Illocutionary Assertive Interactive 'Statement'. In Schmid (200: 155), Linguistic Illocutionary Assertive 'Statement' nouns comprise units like *assertion*, *observation* and *statement*, which represent '[...] acts of stating or saying something'. A conversation similarly entails an act of saying something, but unlike a *statement* or an *observation*, a conversation is dialogic and interactive. For this reason, Interactive is added to Schmid's (2000: 155) tag in (520).

(519) 'And if you were to look in here,' and he pointed to his side, where an ancient oak chest, bound with brass braces and secured with a giant lock and key stood, 'you would find Møn's Book of Legends, in which the massacre is chronicled in words by two contemporary survivors. (*BNC Sampler*: AEA, W:fict:prose)

(520) jolly lucky <pause> whether they er, whether they all turn up on Friday to er, have a word with Eileen I don't know <pause> (*BNC Sampler*: KC2, S:conv)

xi) Work: Most shell-noun uses of *work* (singular and plural) carry the meaning of an activity or action intended to achieve something through effort. This is shown in the definitions given in *LDCE* and *OED*: 'when you use physical or mental effort in order to achieve something', 'action involving effort or exertion directed to a definite end'. Therefore, based on its similarity to such other shell nouns as *event*, *act* or *action*, but also in line with the effort implied in this activity, the shell uses of *work* are tagged as General Eventive Neutral 'Event-Effort'. This category differs from Schmid's (2000: 262) slightly, as it adds the example noun 'Effort' to 'Event'. As 'Effort' is part of the semantic frame of *work*, the question arises as to why this lemma is not assigned to Schmid's (2000: 266) class of Specific Eventive Purposive 'Attempt' nouns (e.g. *attempt*, *effort*, *campaign*, *move*). The reason is that, whilst 'Purpose' lies at the core of 'Attempt' nouns, the purposive component of *work* is only subsidiary to the activity as such. In other words, it is what the work implies that matters most. The effort involved in its performance is a secondary nuance expressing the manner through which the activity itself is done (with a certain effort). This is evident in example (521), where the focus is primarily on the activity itself (*the tooling for the development aircraft and roll equipment*).

(521) Yes well if we take for example er <pause> probably the largest of the packages which we have priced converted to a fixed price basis so far which was <pause> package number three, er the tooling for the development aircraft and roll equipment which is erm <pause> er forty eight million pounds, that was priced in August nineteen ninety two when **eighty percent of that work** had been done. (*BNC Sampler*: JNM, S:meeting)

Example (522) is the only occurrence of *work* with a different semantic annotation. Its sense hinges on the 'local textual function' (Mahlberg 2005: 3) the noun is used to perform in a specific context. *Work* in this context is synonymous with other nouns like *duty*, *role*, *task*, *mission* or *mandate*, all of which belong to Schmid's (2000: 247–9) Modal Deontic Probable 'Job' class. Although all these nouns denote actions, and could also qualify as eventive nouns, they are modally oriented, as the emphasis is not so much on the activity itself as on the requirement or desirability to perform such an activity. In Schmid's (2000: 248) words, '[...] "Job" nouns refer to ACTIVITIES that [AGENTS] are supposed, meant, or even more or less required, but not absolutely obliged or forced to do'. The underlined segment in (522) reveals how such work (advising on

matters of health and healing), represents a responsibility, a remit, a commission that the Board accepted at some point. Now that the Board is no longer willing to offer the service, members are faced with the dilemma of where such a mission ought to go.

(522) Moderator could the convenor tell us why the Board wishes to be discharged its remit on health and healing when it's been so successful in bringing this to the notice of the whole church? <pause> [...]
As I said in pre-- presenting the Board's report, when this matter was first given to us, we were invited to advise the general assembly on matters relating to health and healing, and that we did. At that time the general assembly didn't remove the remit from us and so we were emboldened by being allowed to continue erm to run some training conferences for ministers and others in the church who were involved in sharing this ministry of healing within the nation. However th-- th-- the fit of that training work for ministers and others within the church has sat less and less easily with the Board's commission to advise the church, and that's our social interests commission, to advise the church erm on matters of social, ethical or moral importance within its remit. And so because of that sense of ill ease of fit, we decided that perhaps it would be appropriate for the assembly council to look at where this work ought to most comfortably go. (*BNC Sampler*: F86, S:meeting)

5.3.5.2 *Senses not accounted for in Schmid's (2000) study*

The study sample reveals 29 lemmas, which are in Schmid (2000), but are only partially covered in terms of their semantic potential. In the following, new senses are presented and subsumed under existing or new categories. In some cases, senses accounted for in Schmid (2000) are reclassified as categories other than his. The 29 units discussed below are *anger*, *answer*, *application*, *area*, *assessment*, *capacity*, *challenge*, *chance*, *contradiction*, *crime*, *evidence*, *experience*, *failure*, *impetus*, *irony*, *joke*, *misfortune*, *part*, *phenomenon*, *point*, *practice*, *scandal*, *surprise*, *terror*, *testimony*, *thing*, *time*, *vision* and *way*.

i) Answer: Schmid (2000: 157) only accounts for the linguistic use of this lemma. This is illustrated in (523), where *answer* is coded as belonging to the Linguistic Illocutionary Assertive Argumentative Reactive 'Answer' family (e.g. *answer*, *contradiction*, *counterclaim*, *reaction*, *reply*).

(523) CONFUSED shareholders have been seeking help on the complex question of Eurotunnel warrants and what they are worth. Not much, is the short answer. (*BNC Sampler*: CEL, W:newsp:other:commerce)

Three of the 40 instances for this noun are synonymous with *solution* (see (524)). *Solution* is classified by Schmid (2000: 219–20) as a Mental Volitional Resultative 'Solution' noun (e.g. *solution*, *key*, *remedy*, *cure*). Schmid (2000: 220) explains the Mental Resultative nature of these nouns in terms of their shelling '[...] PSYCHOLOGICAL STATES as results of acts of thinking'. From this perspective, solutions or remedies are arrived at by Experiencers in their search for ways to address problems.

(524) [...] Christ teaches that marriage is special and divorce is not the way he would wish for his people and then this sentence <pause> no-one would wish to condens-- condemn someone to stay in a marriage which is beyond redeeming. We are not trying to say that divorce is never the answer. What we are trying to argue for is <pause> to try to prevent the cost, the pain, the hurt that is experienced after a divorce by alerting people to the fact that a divorce won't solve all their problems [...] (*BNC Sampler*: F86, S:meeting)

ii) Application: Many of the uses of *application* in the sample (31 out of 40) correspond to Schmid's (2000: 173–4) Linguistic Illocutionary Directive 'Invitation' nouns (e.g. *invitation*, *appeal*, *application*, *petition*; see (525)). In Schmid's (2000: 170) framework, directive uses place the illocutionary force on the act of asking the hearer to do something for the speaker. Unlike cases such as *demand*, *order* or *injunction*, the members of the 'Invitation' family do not imply an element of authority and compulsion on the part of the speaker (Schmid 2000: 174). Speakers in these cases are aware that their petition might not be accepted.

(525) On 19 December 1991 the local authority made their application for an interim care order and I have said what the order was. (*BNC Sampler*: FCF, W:ac:polit_law_edu)

In example (526), *applications* does not entail a petition, but 'the practical purpose for which a machine, idea etc can be used, or a situation when this is used' (*LDCE*), a sense unaccounted for in Schmid (2000). Schmid (2000: 218–9) treats items like *purpose* and *function* as Mental Volitional Detached 'Purpose' nouns, based on the idea that the function or purpose of a concrete or abstract entity depends on what an Experiencer intends it to be good for. Thus, *applications* in (526) is used for the many purposes that researchers attach to short pulses.

(526) When many modes are involved, this phenomenon results in very short pulses — typically picoseconds — which have grabbed most of the attention because of their many applications. (*BNC Sampler*: J2H, W:ac:nat_science)

iii) Area: This lemma is thoroughly discussed in 4.4.2.7. Here, attention is only given to the new semantic category created to accommodate its shell-noun uses. In order to account for the two domains of application of abstract *area* ('a particular subject or topic' and 'a particular part of a larger, more general situation or activity', *CCD*), a distinction is made between Specific Abstract Circumstantial Locative 'Place' (Mental domain) and Specific Abstract Circumstantial Locative 'Place' (Eventive domain). The former is evident in (527) and (528), where *area* is synonymous with subject of study (as in (527)) and with content that someone is supposed to learn (as in (528)). The latter is illustrated in (529) and (530), where *area* implies a type of activity or work.

(527) the two IT subject areas which, for SERC students, had the highest proportions in employment and the lowest seeking work, were microelectronics and IKBS/MMI/AI. Computing had the third highest percentage of SERC students in work. (*BNC Sampler*: H0H, W:non_ac:polit_law_edu)

(528) Teresa says, the only one what presents any difficulty is usually the medical aspects one.
Well why should that present anything different?
Because it covers such a wide area. It says the English paper you only need, you only need seventy five percent <pause> for your <pause> erm <pause> it's seventy five percent pass rate. (*BNC Sampler*: KB1, S:conv)

(529) Over the past twelve months it has been my great privilege to become more closely involved with those with disabilities, I have seen the enormous efforts being made to encourage participation at a very high competitive level as well as tremendous strides being made at many clubs to bring disabled people into sailing at all levels. This is an area where we can make much more progress and I'm delighted to see the advances that have already taken place. (*BNC Sampler*: J3W, S:meeting)

(530) you can't hear me, right, ok, I'll try and speak up a little bit louder, erm my work, I work erm here and I work abroad and I live in Durham, er I'm a writer and erm one of the things that I'm most interested in is using art as an inspiration for my work as a, as a writer and I'd like to be able to show you by this talk, how I do that, erm, I have a few publications, my work's performed by a local group actually based in Newcastle and I'm a member of er a group called Another Story and my work's been put to

music by a composer and er a play and the play and sing that, that the pieces of [...] erm another area of my work is <pause> as I er writing workshops [...] (*BNC Sampler*: F71, S:speech:unscripted)

iv) Assessment: The coding of *assessment* (e.g. (531) below[19]) in this thesis does not agree with Schmid's (2000). Schmid (2000: 163–4) assigns it to the small family of Linguistic Illocutionary Assertive Verdictive nouns (*judgement*, *assessment* and *verdict* being the only members), whilst its treatment here is as a Mental noun falling into the family of Mental Creditive Resultative 'Realisation' units (Schmid 2000: 206–7; e.g. *realisation*, *discovery*, *reasoning*, *diagnosis*, *deduction*). The mental analysis adopted here is in line with such other references as Francis (1986: 15, 1994: 92), Francis et al. (1998: 110, 136, 242) and Hunston & Francis (2000: 187). Despite arguing in favour of a linguistic analysis, Schmid (2000: 163), nonetheless, points out that *judgement* and *assessment* '[...] include strong implications as to the cognitive processes that precede acts of judging and assessing'.

(531) 'He's a selfish little swine who behaved appallingly at Catterick,' was her assessment of Quick Reaction's display in a Hunter Chase three days earlier, 'so I ran him here as a punishment!' (*BNC Sampler*: CF9, W:newsp:other:sports)

This apparent contradiction between the category applied and the semantic implications of the lemma ties in with '[...] the systematic ambiguity between speech- and thought-reporting nouns [...]' (Schmid 2000: 149). Francis (1986: 9) suggests that such an ambiguity is explained by the existence of a cline between cognitive and linguistic uses, with clearly illocutionary nouns (e.g. *claim*, *statement*) at one end, clearly cognitive nouns (e.g. *belief*, *idea*) at the other, and nouns straddling cognitive processes and illocutionary acts in between (e.g. *acceptance*, *conclusion*, *judgement*). This being the case, Francis (1986: 14, 1994: 92) opts for the mental analysis of nouns like *assessment*, *reasoning*, *thinking, supposition* and *inference* based on their reference to '[...] cognitive states and processes and the results thereof'. Illocutionary nouns (e.g. *accusation*, *advice*, *answer*, *rejection*), by contrast, label illocutionary acts, with no implication of prior cognitive processing (Francis 1986: 11; 1994: 90). Drawing on Francis (1986,

[19] See 4.4.2.9 for further examples of *assessment* and for the distinction between nominalised shell and non-shell uses.

1994), it is here believed that, in deverbal nominalisations, the Mental feature ought to be applied where the encapsulated discourse segment results from a previous act or process of cognition (e.g. *assessment>assess*, *interpretation>interpret*, *analysis>analyse*, *realisation>realise*). Deverbal nominalisations like *answer*, *suggestion*, *application*, *recommendation* and *warning*, however, are best treated as Linguistic Illocutionary nouns, as their lexical realisation stems from on-the-spot linguistic acts.

The choice of the Mental Resultative 'Realisation' family for *assessment* is motivated by dictionary definitions for this lemma and by Schmid's (2000) inclusion of semantically-related lemmas in this family. *LDCE* defines *assessment* as 'a process in which you make a judgment about a person or situation, or the judgment you make'. *OALD* similarly defines it as 'an opinion or a judgement about somebody/something that has been thought about very carefully', or as 'the act of judging or forming an opinion about somebody/something'. Compared to *reflection*, one of the members in Schmid's (2000: 206–7) Mental Resultative family, dictionaries definitions resemble those for *assessment*: 'careful thought, or an idea or opinion based on this' (*LDCE*), 'careful thought about something, sometimes over a long period of time' (*OALD*), 'your written or spoken thoughts about a particular subject or topic'. In both cases, the nouns imply acts of thinking, as well as the ideas resulting from these acts (cf. also *diagnosis*, *interpretation* or *reasoning*). Therefore, just as units like *diagnosis*, *interpretation*, *reasoning* and *reflection* '[...] typically describe different kinds of mental processes [...] which can lead up to certain BELIEFS' (Schmid 2000: 206), so does *assessment* entail both a mental process and the result thereof, lending support to its coding as a Mental Creditive Resultative 'Realisation' noun.

v) Capacity: All but 4 instances of this lemma belong to Schmid's (2000: 252–4) family of Modal Dynamic (+)Possible Subject-Oriented 'Ability' nouns (e.g. *ability*, *power*, *failure*, *capacity*), where '[...] the possibility of an EVENT is portrayed as being dependent on qualities attributed to an agent' (Schmid 2000: 252). Examples (532) and (533) are two cases in point.

(532) I I think there is certainly **capacity erm for Skelton to expand**, or I wouldn't be sitting here supporting erm <pause> the people I represent. (*BNC Sampler*: FMP, S:pub_debate)

(533) But Gen Noriega's capacity to inspire personal loyalty, as well as the power of his purse, is well known to US Intelligence. (*BNC Sampler*: AAK, W:newsp:brdsht_nat:report)

In his account of the typical 'Ability' sense, Schmid (2000) fails to include a sense of *capacity* featuring four times in the study sample. In examples like (534) and (535), *capacity* stands for 'someone's job, position, or duty' (*LDCE*). Based on its semantic affinity with *role*, *duty* and *job*, these uses are coded as Modal Deontic Probable 'Job' (Schmid 2000: 247–9). This family of Deontic uses comprises nouns shelling activities that agents are supposed but not forced to do (Schmid 2000: 248; cf. the greater compulsion of nouns in the Deontic Necessary 'Necessity' family, e.g. *obligation*, *requirement*, *imperative*, *pressure*, Schmid 2000: 249). The examples considered by Schmid (2000: 247–9) are all linked to the N-*to* and N-*be-to* patterns (e.g. *a mission to*, *his job is to*, etc.). The antecedents in (534) and (535), whilst being nominal in form, may be paraphrased using N-*be-to* clauses. Thus, the role of Chief Racing Coach in (534) implies that his role, job or duty is to act as a Chief Racing Coach. Similarly, in (535), broker and dealer are two roles or duties that market makers are supposed to perform (their role is to act as both brokers and dealers). *Part*, in (536), is also coded as Modal Deontic Probable 'Job', this being the only occurrence with this sense. All the other instances are tagged as Factual Partitive Part-whole 'Part', the only meaning included in Schmid's (2000) study. In this case, the part or role that Osiris played on earth is specified in the underlined segment (to be a just and wise king, to organise the agricultural life of his people, etc.).

(534) Your Royal Highness, Ladies and Gentlemen my name's Rod <gap desc='name' reason='anonymization'> I normally masquerade as the Chief Racing Coach for the Royal Yachting Association but I'm not here in that capacity this afternoon but the coordinator rather a grand title for the <unclear> Year of Youth Sailing and I've been asked to give you a short ten minutes or so briefing on where we're up to with th-- this project this year [...] (*BNC Sampler*: J3W, S:meeting)

(535) The most significant of the changes was the ending of single capacity with its sharp demarcation between the function of broker and that of jobber and the creation of a number of gilt-edged market makers (GEMMs) able to act in the dual capacity of broker and dealer. (*BNC Sampler*: HY1, W:commerce)

(536) The earthly rule of Osiris was that of a just and wise king, who organised the agricultural, religious and secular life of his people, and who also concerned himself with peaceful foreign conquest, ably

assisted by Isis, by his vizier Thoth, and by his officials Anubis and Wepwawet. This happy state of affairs was soon to be destroyed, for Seth, the younger brother of Osiris, was jealous of Osiris's power and prestige and determined to seize the throne for himself. [...] The resurrected Osiris had **no further part to play on earth**, and as the immortal dead king he became the ruler of the dead. (*BNC Sampler*: EVR, W:non_ac:humanities_arts)

vi) Challenge: Schmid (2000: 247–9) accounts only for Modal Deontic uses of this lemma. In his view, examples like (537) ought to be coded as Modal Deontic Probable 'Job', based on their semantic affinity with such other nouns as *job*, *duty*, *role* or *task*. This interpretation is confirmed by a search of the sequence *your _NN1 {be} to* in *BNCweb*, with *task*, *job*, *mission*, *role*, *work* and *commission* topping the list. Such a finding is coherent with the reported dominance of 'obligation' and 'intended result' meanings in the N-*be-to* pattern (Schmid 2007: 329–33, in 3.2.5.2). In addition to N-*to* and N-*be-to* occurrences, the study sample also shows cases where the Deontic interpretation hinges on lexical signals in the preceding co-text. This is shown in (538), where the 'Job' sense is implied in the sense of necessity conveyed in the italicised segments.

(537) There were daily chances of grace. Roads were hazardous practical examinations in virtue. **Your challenge** was to encounter a dying man or child, the victim of an accident, and to have the presence of mind to rescue his soul, in a terrible elevation to sacred power, by the formula of baptism. In such a way a child's promptness might save a soul from the long blankness of Limbo, or a child's inattention or failure of nerve condemn it to years of the deadliest waiting until the day of Judgement, held still in the amber of the centuries. (*BNC Sampler*: EX7, W:non_ac:soc_science)

(538) What rights do these children have when *their primary need* is to survive in often harsh environments and perhaps even help to support their families. […] Yes *children should be* at school […] that *income generating programmes for children need consideration* [...] *Such children must be assisted* in gaining education and skills built around their existing work activities. [...] **The challenge of meeting all children's needs and of helping them achieve their rights** can only be met by concerted efforts of governments of multi-lateral agencies and bi-lateral donors, as well as by the path-breaking initiatives of NGOs such as Save The Children. (*BNC Sampler*: JNG, S:meeting)

In examples like (539) and (540), a different kind of meaning is at issue. *Challenge* in these cases is not modally deontic, since its sense is that of 'a statement or an action that shows that somebody refuses to accept something and questions whether it is right, legal, etc.' (*OALD*). The subversive action in (539) involves the forming of South Africa's first trade union for police and prison warders, whilst that in (540) implies the writing and addressing of an open letter of defiance to President Ceausescu. The eventive orientation of these uses of *challenge* matches other units included in Schmid's (2000: 266–8) family of Specific Eventive Purposive 'Attempt' nouns. Replacing *challenge* with *attempt*, *effort*, *campaign* or *move* preserves the intended semantic effect of (539) and (540) (e.g. *in a renewed attempt/campaign/move against white authority; their move was clearly the most serious*), hence the suitability of this semantic category for 'subversive action' uses of *challenge.*

(539) THE rebel Coloured policeman, Mr Gregory Rockman, has formed South Africa's first trade union for police and prison warders in a **renewed challenge to white authority**. (*BNC Sampler*: A7V, W:newsp:brdsht_nat:report)

(540) Most notable in the first category have been six veterans of the Romanian Communist Party who at the beginning of the year addressed an open letter of defiance to President Ceausescu, telling him that 'the very idea of socialism for which we have fought is discredited by your policies'. [...] While **their challenge** was clearly the most serious to have been presented to President Ceausescu, and was made with the full knowledge of important facts, it was also made by people who have long since left the political scene, mainly having retired, or been retired [...]. (*BNC Sampler*: AAK, W:newsp:brdsht_nat:report)

Lastly, instances like (541) and (542) seem to highlight the difficult or problematic nature of particular situations. This interpretation ties in with one of the senses of *challenge* as 'something that tests strength, skill, or ability, especially in a way that is interesting' (*LDCE*) or 'something new and difficult which requires great effort and determination' (*CCD*). Further support for the 'problem' sense comes from Francis et al. (1998: 226, 244), where *challenge* features alongside *problem*, *difficulty*, *matter* and *nightmare* in families indicating '[...] that something is difficult or easy', as well as '[...] something difficult that needs to be done or dealt with'. Unlike cases like (537) and (538), where the focus is primarily on the need to fulfil a task, in instances such as (541) and (542) the emphasis is on how difficult or problematic a task or a situation is. The

occurrence of the verb *face* in the preceding co-text in (541) and (542) (*faces the most potent challenge*, *a major new challenge the governments have had to face*) is further proof of the 'problem' interpretation, as, according to *CCD*, *face* is often used with things which are '[...] difficult or unpleasant'. This is confirmed by *BNCweb*, where *face* is associated with negative entities such as *problems*, *difficulties*, *charges*, *death*, *starvation* and *criticism.* On these grounds, the problem, complication or predicament that President Ortega faces in (541) is that of the terrible economic situation in his country, specifically, his being called to account for such a disaster. In (542), the problematic state of affairs created by the collapse of communism and the disintegration of the Soviet Union poses the dilemma or problem of not knowing exactly how to support the transition to democracy. Examples like (541) are coded as Factual Attitudinal Impeding 'Problem' (Schmid 2000: 121–5), as attention is drawn to the difficulties or problems linked to particular situations or state of affairs. By contrast, (542) is better suited for the family of Eventive Attitudinal Manner Deontic 'Trouble' nouns (Schmid 2000: 270–2). In these cases, the adjoining *to*-infinitive clause adds a meaning of 'knowing how to do something' absent from typical Factual uses of nouns like *problem*, *difficulty* or *dilemma.*

(541) President Ortega faces the most potent challenge to his leadership since the Sandinistas took full control in 1981. The challenge comes not from the US, not from the contra rebels, nor from hostile neighbouring countries, but from within. In short, the government is being called to account for the economic disaster that is present-day Nicaragua. (*BNC Sampler*: A95, W:newsp:brdsht_nat:report)

(542) The collapse of communism in the former Eastern block and the disintegration of the former Soviet Union represent a major new challenge the governments have had to face. The particular challenge for donors has been how best to support the fragile but vital process <pause> of transition to a pluralist democracy and market economies in the region. (*BNC Sampler*: JNG, S:meeting)

vii) Chance: Of the 40 instances for this lemma, 36 are synonymous with *opportunity* and *possibility*. The former interpretation is evident in (543), where, in terms of Schmid's (2000: 254–8) family of Modal Dynamic Possible Neutral 'Opportunity' nouns (e.g. *chance*, *opportunity*, *possibility*, *occasion*), *chance* stands for the '[...] situation or circumstances where it is possible to do something' (Schmid 2000: 254). In this particular instance, the match in Rugby

Park is the situation that enabled this team to go within one game of their previous best Premier League run of ten matches undefeated.

(543) Now to Rugby Park, where Kilmarnock, after a superb start to the season, had the chance to go within one game of their previous best Premier League run of ten matches undefeated. (*BNC Sampler*: J1N, W:news_script)

According to Schmid (2000: 257), this sense of *chance* is primarily linked to the N-*to* pattern, as confirmed by *BNCweb*, where *opportunity to* and *chance to* reveal 4902 and 3837 tokens as opposed to only 9 and 38 tokens for *possibility to* and *probability to*. Example (544) is an exception to this phraseological tendency, as an 'opportunity' interpretation arises despite the occurrence of *chance* with an *of*-phrase. In this instance, Gary's *chance of overtime* implies his opportunity to work overtime, thereby earning more money than usual, which explains his wife's support: *if you can get it*, *get it* (i.e. *if you have this opportunity*, *just take it*).

(544) I might give you a ring later and say can you have Aaron
Gary's gonna find out, he's got <-|-> chance of overtime so I said well if you can get it, get it <-|-> and I'll <-|-> drop Aaron off in the morning. (*BNC Sampler*: KD1, S:conv)

With regard to the 'possibility' sense, Schmid's (2000: 236–40) emphasis on the high frequency of N-*that* and N-*be-that* patterns with Modal Epistemic Possible 'Possibility' nouns (e.g. *chance, possibility*, *danger*, *option*, *uncertainty*) does not match the evidence for *chance* in the study sample. Of the 20 instances of 'possibility' *chance* 19 occur with an appositive *of*-phrase, and 17 are plural in form (cf., however, 'opportunity' *chance*, where 15 of the 16 examples of this use occur in the singular).

It should be remembered at this point that Schmid (2000) excludes plural and *of*-phrase shell-noun instances, hence the focus on the N-*that* and N-*be-that* patterns of 'Possibility' nouns. However, a *BNCweb* search for the sequences *{chance/N} _CJT* (*chance/chances* + conjunction *that*), *{chance/N} of* (*chance/chances* + *of*) and *{chance/N} _TO0* (*chance/chances* + *to-infinitive*) reveals the N-*that* pattern as the least frequent, with only 445 occurrences, as opposed to 4389 for *chance(s) of* and 3920 for *chance(s) to*. Therefore, it is little wonder that N-*of* and N-*to* prevail in the evidence from the study sample. The 'possibility' sense of *chance of* occurrences (as in (545) and (546)) is confirmed by *CCD*,

where it is stated that '[I]f there is a chance of something happening, it is possible that it will happen'. *BNCweb* evidence is further proof of this meaning, as a search for *possibility of*, *chance of*, *probability of* and *opportunity of* returns 3829 and 3117 tokens for the former two and 567 and 668 for the latter two. Similarly, in the plural, *chances of* and *possibilities of* prevail over *probabilities of* and *opportunities of* (1269, 503, 75, 117 tokens). An exception to the dominance of N-*of* in cases of 'possibility' *chance* is example (547), where possibility is entailed in the hypothetical or tentative modal *could* in the underlined clauses.

(545) Bilici Constantin, a cheery film man in his early 40s, hurried back from France after an absence of 10 years and arrived just as Ceausescu fell. He and a few of his colleagues were inspired by the early struggles of Solidarity in Poland a decade ago. In Ceausescu's Romania there was no chance of pressing their claim too far. They just wanted to redress some glaring injustices in their terms of employment. (*BNC Sampler*. AAT, W:newsp:brdsht_nat:report)

(546) For the average hernia operation which is done remember as cold surgery, not as an emergency, the chances of dying under the anaesthetic are vanishingly small. <pause> (*BNC Sampler*: FLY, S:classroom)

(547) Well what I did, there were a competition, win a thousand pounds worth of Air Miles <pause> and I, I entered every single one of them because of Irene. You know, she's been invited to that wedding in Africa? And I thought <pause> well she's got two chances that way, either <pause> I could win, or she could win, or <pause> neither of us 'll win. (*BNC Sampler*. KCX, S:conv)

One other meaning of *chance* related to the 'Possibility' family is shown in (548), where *take chances* involves doing something '[...] although there is a large risk of danger or failure' (*CCD*). *Risk* itself is a member of Schmid's (2000: 238) 'Possibility' family, where it characterises possible future events as undesirable along with, e.g. *danger*. In (548), the risk that Richard of Gloucester is unwilling to take is the over-exposure of the future king and his brother to the public: if this occurred, their life would be in danger (see the segment in italics: *I'd not give a fig for the young king's life*). In this thesis, the tag Modal Epistemic Possible 'Possibility-Risk' is used to distinguish between instances of 'Possibility' and 'Risk' (cf., however, Schmid 2000: 236, where no such distinction is made).

(548) 'I expected to see the prince my husband at Mass this morning, my lady, she said. Could it be that he and the king his brother hear Mass at St.

> Peter ad Vincula, the chapel outside the White Tower?' 'Both the king and his brother, I understand, hear Mass at a small private chapel close to their apartments,' came the answer. 'The protector desires that, prior to his coronation, the king should not be seen too frequently in public.' More disappointed than she was prepared to admit, Joan could not entirely conceal her chagrin. 'Regular attendance at the chapel of St. John hardly constitutes over-exposure to the public,' Joan demurred. 'Does not the king himself have a say in such matters?' [...] Clearly though she is disappointed at not seeing him — and is concerned for his welfare. And in that she is not alone. There is much uneasiness and speculation — all of it, please God, unfounded! The king and his brother are seen from time to time, if infrequently — and appear lively and in good health. Nay, one must give Richard of Gloucester the benefit of the doubt — he is taking no chances this side of the coronation. Various factions are at work, each seeking to promote its own interests — I'd not give a fig for the young king's life if he fell in with one of them! One cannot blame the protector for being zealous in his guardianship of his nephews. (*BNC Sampler*: CCD, W:fict:prose)

Lastly, 2 of the 40 instances of *chance* show a sense of 'Probability' linked to Schmid's (2000: 240–2) Modal Epistemic Probable 'Probability' family (e.g. *chance*, *likelihood*, *probability*). This is shown in (549) and (550), where the occurrence of *will* and of nouns like *law* or *probability* in the same sentence suggests a higher degree of certainty than that found in cases like (235).

(549) The equanimity of your average tosser of coins depends upon the law, or rather a tendency, or let us say a probability, or at any rate a mathematically calculable chance, which ensures that he will not upset himself by losing too much nor upset his opponent by winning too often. (*BNC Sampler*: FU6, W:fict:drama)

(550) <pause> The chances are we won't be going! (*BNC Sampler*: KCN, S:conv)

viii) Contradiction: In Schmid's (2000: 157) categorisation, *contradiction* is a Linguistic Illocutionary Assertive Argumentative Aggressive noun. This family comprises units like *counterclaim*, *criticism*, *critique*, *protest*, *provocation* and *contradiction*. According to Schmid (2000: 157), these '[...]characterize assertions as acts of attacking the positions of other discourse participants' (hence the Aggressive feature). However, the linguistic illocutionary uses retrieved from the *BNC Sampler* do not represent cases of assertions uttered in order to attack someone else's position, or, as defined in the *OED*, 'the action of contradicting or declaring to be

untrue or erroneous; affirming the contrary; assertion of the direct opposite; denial'. The meaning conveyed is 'a statement containing propositions one of which denies or is logically at variance with the other; also a contradictory proposition' (*OED*) or 'a statement or phrase which is self-contradictory on the face of it' (*OED*). Examples (551) and (552) are two cases in point. The underlined segments in them represent two self-contradictory assertions: the contrast between noise and silence and between job seeking among specialist IT and conversion students. A new semantic category is created to accommodate these uses, that of Linguistic Illocutionary Assertive Difference 'Contradiction'. This category rests on the assumption that *contradiction* in these examples is not a difference as such or a statement uttered against another statement, but simply a statement of difference or containing difference or contrast.

(551) The scent of rotting leaves, beech nuts and bracken combined in a heady mélange Elisabeth found intoxicating. She, who was the victim of the harsh tumult of dissonance, observed how the noiselessness of the forest was made evident by the occasional snap of a twig or dry thud of a pine-cone falling to the ground. Noise creates silence! The seeming contradiction reminded her of another: 'the solution posed the question'. (*BNC Sampler*: AEA, W:fict:prose)

(552) Students studying on specialist IT programmes were less likely than conversion students to be in employment or to be seeking work (Table 4.3). This apparent contradiction is explained by the fact that, overall, almost one in five IT specialists were continuing their studies, usually aiming for PhD qualifications. (*BNC Sampler*: H0H, W:non_ac:polit_law_edu)

In addition to Linguistic uses, examples like (553) and (554) illustrate a Factual reading. In these cases, *contradiction* implies 'a state or condition of opposition in things compared; variance; inconsistency, contrariety' (*OED*) or 'a lack of agreement between facts, opinions, actions, etc.' (*OALD*). In (553), *the contradiction between political renewal and economic crisis* condenses the sense of contrast implied in the underlined discourse segment into a single noun phrase. The sentence in italics, introduced by the adversative conjunction *but*, signals the contrast between Mr Prokhorov's comment on the current economic downturn in the Soviet Union and his optimism about Soviet politics. A reader presented with the sentence where the noun phrase in boldface occurs would be justified in asking what this contradictory state of

affairs between political renewal and economic crisis involves. Only by reference to the underlined co-text can the reader understand the significance of this contradictory situation. In (554), the inconsistency or lack of agreement relates to television commercials supporting abortion, which, according to the Catholic Church, oppose its moral teachings. In other words, these commercials represent a contradiction or inconsistency in the light of religious dogma. These uses of *contradiction* have been assigned to Schmid's (2000: 113–16) family of Factual Comparative 'Difference' nouns, based on their semantic affinity with 'Difference' nouns like *contrast*, *discrepancy* and *inconsistency*.

(553) SOME people believed that the Soviet Union had nothing in particular to celebrate, observed Mr Boris Prokhorov, of the official Soviet news agency Tass, yesterday. Commenting on the 72nd anniversary of the Russian Revolution, in the openly frank vein that is the hallmark of glasnost, Mr Prokhorov went on: 'Well, one may think so, looking at the industrial down-swing, inflation, transport malfunctions, coal miners' strikes and semi-barren shelves of shops. Yes, all this exists. These are the realities of the truly not easy days of our perestroika.*' But for Mr Prokhorov, there is another side to things*.'The first genuinely free elections were held in the Soviet Union,' he pointed out. 'The first Congress of People's Deputies was held and became a political sensation. The first 'real' Supreme Soviet began to function.' And, he added, the broadcasts of parliamentary sessions attracted more viewers than films based on Agatha Christie novels. Besides that, Soviet troops had withdrawn from Afghanistan and a whole range of nuclear missiles had been destroyed. Mr Prokhorov said that the answer to **the contradiction between political renewal and economic crisis** lay in Marxism, which 'teaches us that the superstructure is more mobile than the base, and that it is the first to react to a change of conditions.' (*BNC Sampler*: A7V, W:newsp:brdsht_nat:report)

(554) Bishop Leo Maher, head of the San Diego diocese, banned Mrs Lucy Killea from communion after the 67-year-old grandmother began screening television commercials promoting her support of women's right to choose abortion. Bishop Maher denounced them as 'in **complete contradiction of the moral teachings of the Catholic Church**.' (*BNC Sampler*: A8W, W:newsp:brdsht_nat:report)

ix) Crime: *Crime* features only in Schmid's (2000: 273–4) family of Eventive Attitudinal Negative 'Mistake' nouns (e.g. *mistake*, *fault*, *error*, *sin*, *offence*). However, only one of the *crime* instances in the study sample illustrates this sense. In (555), *crime* is used with the

meaning of 'an act that you think is immoral or is a big mistake' (*OALD*), the act here being falling in love.

(555) I will give the crown to my brother, and leave England. I must follow my heart. You tell me that it's a crime to fall in love. You tell me that it's wrong to be happy. How strange this country is!' (*BNC Sampler*: GV9, W:fict:prose)

The remaining instances do not imply a mistake, but the more typical sense of 'an illegal act or activity that can be punished by law' (*OALD*). This is evident in (556), where *a state crime* encapsulates the act of killing someone. A new category is proposed in this thesis to accommodate these uses: Specific Eventive Negative 'Crime'. Whilst, at first glance, Schmid's (2000: 262) family of General Eventive Neutral 'Event' nouns (e.g. *event*, *act*, *action*) seems appropriate for the 'act' sense of *crime*, there arises a need for a category including nouns portraying specific types of general events unrelated to any of Schmid's (2000: 266–70) family of Specific Eventive senses (i.e. 'Attempt', e.g. *attempt*, *effort*, *campaign*, *move*; 'Tradition', e.g. *tradition*, *habit*, *convention*, *custom*; 'Option', e.g. *option*, *alternative*, *choice*, *preference*). The illegal nature of a crime explains the inclusion of Negative in its semantic categorisation.

(556) Herri Batasuna has called the killing a state crime and blamed the Socialist government. (*BNC Sampler*: A8J, W:newsp:brdsht_nat:report)

x) Evidence: *Evidence* is included in Schmid's (2000: 110–13) family of Factual Causal Mental 'Evidence' nouns (e.g. *evidence*, *sign*, *proof*, *indication*). The author characterises the semantic frame of these units as one where '[...] an observable state of affairs or fact is conceived of as a reason for a mental state [...]' (Schmid 2000: 110). The former element corresponds to what he terms a 'SIGN', whilst the latter constitutes the resulting 'BELIEF'. In the N-cl examples he gives, the lexical realisation of 'Evidence' nouns is attributed to the following *that* noun complement clause, i.e. the BELIEF. However, it is here argued that the discourse segment that provides information as to what the evidence involves is the SIGN and not the BELIEF. This interpretation is supported by the following dictionary definitions of the noun *evidence*: 'Facts or signs that show clearly that something exists or is true' (*LDCE*), 'Evidence is anything that you see, experience, read, or are told that causes you to believe that something is true or has really happened' (*CCD*).

Hence, in reading examples like (557), the reader would be justified in asking what *a lot of evidence* refers to. This implies that it is what causes the claim about the marked effect of petrol prices (i.e. the SIGN) that provides truly informative specifics as to what *evidence* is. In this particular instance, the SIGN component is absent from the frame, which means that *a lot of evidence* is exophoric. The resulting BELIEF (realised by the *that* complement clause) is here treated as specifics of identity on the grounds of its indirect contribution to the relationship of experiential identity between *evidence* and the SIGN. Thus, in terms of encapsulating direction and antecedent, the *that* clause is coded as CF.INTRA^LC.CL(SI.RU), 'RU' indicating the resultative meaning of the BELIEF component. Unlike (557), the SIGN component in (558) is endophoric, this being realised by the underlined segments.

(557) For example, there has been a lot of evidence that petrol prices have a marked effect on rates of rural recreation (Shucksmith, 1980b), since the demand relationship is not only stable, but is also highly price-elastic. (*BNC Sampler*: FR2, W:ac:soc_science)

(558) In July 1985 the Comptroller and Auditor General's Report was published and it concluded that there were weaknesses in the control and deployment of nursing staff. (1) A 9% increase, in real terms, in the number of nurses employed in England between 1976 and 1983 was the justification for the National Audit Office Enquiry. [...] The report was published in July 1986 and concluded that apparently there had been little work done on the effects of skill mix on patient services and that 'a higher priority should be given by management to achieving the best value for money by the adoption of methods of allocating staffing resources more closely related to the needs of patients and ward objectives. In its report, Control of Nursing Manpower, published in February 1986, the Committee of Public Accounts acknowledged that progress had been made by nurse managers but went on to recommend that the 'DHSS should ensure that all general managers are made aware of all the possibilities for economies identified in the C & AG's report' (2). With this 'backcloth of evidence', it was no surprise that the Personnel Director of the Management Board of the DHSS asked regional general managers to discuss the issues with regional nursing officers and district colleagues. (*BNC Sampler*: EVY, W:commerce)

In addition to the typical factual sense of *evidence*, (559) shows another sense missed by Schmid (2000). In this case, *evidence* entails 'information that is given in a court of law in order to prove that someone is guilty or not guilty' (*LDCE*). Based on its semantic

affinity with nouns like *information* and *argument*, *evidence* in (559) is coded as Linguistic Propositional 'News' (Schmid 2000: 140–4).

(559) He erm <pause> he wasn't giving **evidence** at the enquiry, but erm it was at his suggestion that I was asked to meet with the Chief Planning Officer prior to the <unclear> to talk about alternative sites. <pause> (*BNC Sampler*: FMP, S:pub_debate)

xi) Experience: Schmid (2000: 203–6) classifies *experience* as a Mental Creditive Attitudinal 'View' noun in terms of its correspondence to such other items as *idea*, *opinion* or *perception*. This use relates closely to the N-*be*-cl pattern, as in (560), where *experience* indicates that, based on the writer's life, his/her opinion or view is that field sports help to make people more sociable. However, none of the instances retrieved from the *BNC Sampler* feature the N-*be*-cl pattern, thereby carrying senses other than the attitudinal one identified by Schmid (2000).

(560) 'After nearly 60 years on this earth, 46 of them spent in the pleasurable pursuit of fish, **my experience** is that people who follow field sports, particularly angling and particularly young anglers, are far less antisocial than those who never go near the countryside or its pursuits.' (Schmid 2000: 204)

Quite an experience in (561) does not involve an opinion, but 'something that happens to you or something you do, especially when this has an effect on what you feel or think' (*LDCE*). In view of the subject-oriented and eventive nature of this sense, instances like (561) are coded as Mental Event-Related 'Experience'. Thus, the psychological effect entailed in this use stems from something that has taken place or something that the Experiencer has seen (as in (561)). Mental Event-Related 'Experience' is not a category in Schmid's (2000) taxonomy. It is partially inspired by the family of Mental Emotive Event-Related 'Fear' nouns (Schmid 2000: 228–30; e.g. *fear*, *concern*, *worry*, *anxiety*). The emotive component in *experience* may be conveyed through premodifying interpersonal Epithets (e.g. *good*, *awful*, *terrifying experiences*), but it is not itself part of the semantic frame of the lemma[20], hence the exclusion of 'Emotive' from its semantic tag. This category is also used with two instances of *time*, one of them being (262), where *time* implies 'a

[20] Cf., however, *fear*, *worry*, *anxiety*, *apprehension*

period of time when you have good, bad, difficult etc experiences' (*LDCE*).

(561) Er, and all in all, it was quite er it's quite it was quite an experience, er er to have seen er this this er this police er er baton charge, [pause] er and er we were fortunate enough in in being able to to to get out of the way. (*BNC Sampler*: FYJ, S:interview:oral_history)

(562) But I, I can understand why she'd want to marry, why she'd want to do that, cos she hasn't ha-- well had an awful lot of <pause> brilliant times in her life since she left Nigeria. (*BNC Sampler*: KC7, S:conv)

In examples (563) and (564), the focus is not so much on a particular event affecting somebody, as on the knowledge that somebody has gained from doing or seeing something. *Experience* in (563) seems to imply the 'knowledge that you gain about life and the world by being in different situations and meeting different people' (*LDCE*), the knowledge or experience in this case being gained after living for a while in a deprived area. The meaning of (564) is more in line with the 'knowledge or skill that you gain from doing a job or activity' (*LDCE*). In this instance, his employment in the leisure industry has allowed him to gain knowledge and skill in such areas as the marketing of sports clothing and keep-fit equipment. For want of a better semantic category, examples like (563) and (564) are classed as Mental Creditive 'Belief' nouns (Schmid 2000: 195–203), based on their semantic affinity with *knowledge* (knowledge after seeing or experiencing something and knowledge about how to do something).

(563) What did you find best about living there.
Er well I think the experience that I gained of the the kind of erm living situation erm because I you know <unclear> we underwent the same kind of difficulties. Erm e-- even like things like emptying the bins. Er I know there were complaints that people used to throw rubbish out of the windows. Well it did take a bit of organization to get yourself to the end of the walk, erm with your your weekly rubbish or your daily rubbish, to get that put away. Er and you know to really see what kind of erm design, what design can do to the actual living situation. Erm I think too, the closeness to erm our neighbours and to to the people was a very good experience. And also just the the whole erm general feeling of being with the people in the flats. (*BNC Sampler*: FY8, S:interview:oral_history)

(564) He was born in Toronto, Canada in 1940 and has extensive experience in the leisure industry, particularly in the marketing of

sports clothing and keep-fit equipment. (*BNC Sampler*: FEJ, W:commerce)

xii) Failure: Schmid (2000) only presents Modal uses of this lemma. Example (565) is a case in point, where the impossibility of obtaining a sponsor for the national championships is attributed to the Squash Rackets Association. The inability of an entity to make an event happen is at the core of the family of Modal Dynamic (-)Possible[21] Subject-Oriented 'Ability' nouns (Schmid 2000: 252–4, e.g. *inability*, *failure*, *incapacity*). These nouns indicate their Subject-Oriented nature through possessive determiners, premodifying genitive phrases or *of*-phrases (as in (565)). However, there are instances where the agent is not explicitly marked in the noun phrase. This is illustrated in (566), where the Modal interpretation of *failure* is enabled by the information in the previous sentence. *The refusal of the contras to demobilise* introduces an element of difficulty or challenge that, according to Mr Arias, may thwart Sandinista attempts to disarm and relocate the contras, as a result of which, the polls may be cancelled.

(565) Surprisingly he must take part in the qualifying section of the Daily Express Newspapers and Pursers inaugural British Squash Players' Championship to be played between March 23 and 28 at Bromley, Beckenham and Cheshunt. This event sprung up following the failure of the Squash Rackets Association to secure a sponsor for this year's national championships and a resultant boycott from most of the leading professionals. (*BNC Sampler*: CF9, W:newsp:other:sports)

(566) The other main subject of discussion will be the refusal of the contras to demobilise under the terms of the Tela accord signed by the presidents last August. Although the Sandinistas have repeatedly pledged democratic elections in February, Mr Arias was said to be fearful that a failure to disarm, relocate or disband the contras could provoke a Sandinista cross-border sweep against the rebels and cancellation of the polls. (*BNC Sampler*: A9M, W:newsp:brdsht_nat:report)

Three additional senses not recorded by Schmid (2000) are illustrated in (567) through (569). *Failure* in these examples is used with the meaning 'someone or something that is not successful' (*LDCE*). If the unsuccessful thing is eventive in nature, as in (567) (*the entire gassing operation*), the example is coded as Eventive Attitudinal Evaluative Negative 'Tragedy', based on its semantic

[21] Cf. also the (+)Possible variant, with such nouns as *ability*, *capacity* and *potential*

affinity with nouns like *disaster*, *catastrophe* or *debacle*[22]. However, in cases like (568) and (569) *failure* encapsulates a person and a thing respectively (*you* and *computer programs*). This use of *failure* is motivated by the need to describe how a negative state of affairs applies to a particular entity. In other words, being a failure is a state of affairs that impinges on an individual and an object. These two examples are tagged as Factual Attitudinal Evaluative Negative 'Tragedy' (Schmid 2000: 129), this being the Factual equivalent of the Eventive 'Tragedy' family.

(567) Over loud speakers the FBI told the Davidians it was not an assault, just an attempt to force them out. But the Davidians stayed put. Last week's report would call the entire gassing operation a **failure**. (*BNC Sampler*: HE3, S:brdcast:documentary)

(568) The thing is er, you can end up feeling like a bit of a, a **failure**, almost like a traitor to womankind if you do take courses that are traditionally female. (*BNC Sampler*: FLK, S:brdcast:discussn)

(569) Neglect of this obvious truth led to computer programs (say, for language analysis) being regarded as **failures** if they could not record 100% success. (*BNC Sampler*: F98, W:ac:humanities_arts)

Lastly, *failure* in (570) conveys the sense of 'an act of not doing something which should be done or which people expect you to do' (*LDCE*). Specifically, it suggests that the supporters' act of not transferring votes on elimination caused these votes to be non-transferable. The eventive and negative polarity nature of this use of *failure* led to its inclusion in a new semantic category, that of General Eventive Neutral Negative Polarity 'Failure'. The tag draws on Schmid's (2000: 262–6) family of General Eventive Neutral 'Event' nouns (e.g. *event*, *act, action*), but it adds the Negative Polarity component of *failure*.

(570) As far as the 1975 election is concerned VUPP had a greater proportion of seats than was warranted by its first preference vote for two reasons. First, other parties in the UUUC put up rather fewer candidates in certain constituencies than quotas of first preference votes would have suggested and there was a subsequent transfer of votes from those parties in VUPP's favour. Second, **failure of supporters of other parties** to transfer votes on elimination caused these to be non-transferable and made it possible for VUPP candidates to be elected

[22] Eventive Attitudinal Evaluative Negative 'Tragedy' is a new category proposed in this thesis (see the discussion on *scandal* and *misfortune* below for the reasons behind this decision).

'without reaching the quota' in the last stage of a count. (*BNC Sampler*: H7C, W:non_ac:polit_law_edu)

xiii) Impetus: *Impetus* occurs in Schmid's (2000: 222–4) class of Mental Volitional Conditional 'Determination' nouns (e.g. *determination*, *motivation*, *nerve*, *courage*). This category accommodates nouns with a specific phraseology, i.e. *have the* NOUN *to do something*. The NOUN slot is occupied by psychological states enabling the accomplishment of a particular aim (e.g. *have the gall to ask about something*, *have the grace to admit something*, etc.). However, in none of the seven instances of *impetus* from the *BNC Sampler* does this phraseology occur. The sense of 'an influence that makes something happen or makes it happen more quickly' (*LDCE*) or 'something that encourages a process or activity to develop more quickly' (*OALD*) underlies these examples. These definitions suggest a causal semantic frame, whereby something (the influence) causes the quicker development of something else. This is evident in (571) and (572). In (571), the admiral's visit (i.e. his coming) is the influence that, according to the admiral, will invigorate the relationship between both countries. *Impetus* in (572) implies that the proliferation and development of research on laser instabilities might be triggered or enhanced by the development of a true wave-all optical passive resonator system. In other words, only such a development could trigger or be a starting point for further research on laser instabilities. In view of the causality involved in these cases, and for want of a better category, a decision was made to code them as Factual Causal 'Reason' (Schmid 2000: 102–6; e.g. *reason*, *thing*, *cause*, *ground*).

(571) But the visit amounted to an urgent appeal from Washington to stop what has seemed an inexorable downward slide in its relations with Beijing since the bloody crackdown on anti-government protests in June. The admiral said he had come 'to bring new impetus and vigour into our bilateral relationship'. (*BNC Sampler*: A9M, W:newsp:brdsht_nat:report)

(572) These experiments are, however, still pulsed, and it would be extremely valuable if a true continuous wave all-optical passive resonator system could be developed, to give the sort of experimental impetus that the xenon and helium-xenon lasers have given to laser instabilities. (*BNC Sampler*: J2H, W:ac:nat_science)

xiv) Irony: In Schmid's (2000: 127–9) classification, *irony* appears only as 'a situation that is unusual or amusing because something

strange happens, or the opposite of what is expected happens or is true' (*LDCE*). One such example is (573), where, following Schmid (2000: 127–9), *irony* is coded as a Factual Attitudinal Descriptive 'Irony' noun. Nouns like *irony*, *coincidence*, *peculiarity* or *absurdity* are used to highlight the noteworthy nature of a fact or state of affairs. *Irony* in (573) casts light on the apparent contradiction between developments in urban planning and the large number of out-of-school children in developing countries.

(573) It is a **terrible irony** that although they are invisible to the planners <pause> all of us who've travelled in developing countries cannot fail to have noticed these children out of school. (*BNC Sampler*: JNG, S:meeting)

Schmid (2000) overlooks the linguistic sense of *irony*, i.e. 'when you use words that are the opposite of what you really mean, often in order to be amusing' (*LDCE*). This is shown in (574), where Britten's use of *irony* is illustrated by the quoted passages from one of his operas. For want of a better tag, these uses are coded in terms of one of Schmid's (2000: 144–5) three Metalinguistic families: Linguistic Propositional Metalinguistic 'Adage' (e.g. *joke*, *nonsense*, *adage*, *allegory*). Classification into the 'Rumour' and 'Myth' families (Schmid 2000: 145–7; e.g. *talk*, *rumour*, *gossip*; *myth*, *teaching*, *stereotype*) was considered inappropriate on the grounds that, whilst 'Adage', 'Rumour' and 'Myth' are all Metalinguistic in their reference to types of language, unlike members of the two former families, 'Adage' nouns do not seem to imply people's widespread familiarity with the encapsulated discourse segments (cf. *irony* vs. a *rumour*, *a myth*).

(574) If anything Britten outdoes his models by the pointedness of his treatment, in which **irony** is constantly heightened into a more specific symbolism. In the church scene the words of the Anglican matins are made to allude heavily to Grimes's troubles at almost every point: 'we have erred and strayed from thy ways' just as Ellen notices the tear in the apprentice's coat; '0 Lord open Thou our lips' with Ellen's 'John, what are you trying to hide?'; and finally 'Amen' as answer to Ellen's 'We've failed', which Peter takes up literally in his most crucial phrase 'So be it, and God have mercy upon me!' (As Philip Brett has now shown, this was originally even more loaded, with 'He descended into Hell' answered by Peter's 'To Hell then, and God have mercy upon me'. Britten took this out, however, although he kept the obvious reference to it in Grimes's last soliloquy, 'To Hell with all your mercy'.) (*BNC Sampler*: J55, W:non_ac:humanities_arts)

xv) Joke: Most uses of this lemma involve '[...] a funny story' (*OALD*), as in (575) and (576). Schmid (2000: 144–5) assigns this sense to the family of Linguistic Propositional Metalinguistic 'Adage' nouns (e.g. *joke*, *nonsense*, *adage*, *preface*, *proverb* and *allegory*).

(575) Nor show a blindworm how to blink, Nor teach an old racoon Chinese. The juiciest orange feels the squeeze; Who spends his portion will be broke; Who has no milk can make no cheese — So say I and so say the folk. He makes no blot who has no ink, Nor gathers honey who keeps no bees.The ship that does not float will sink; Who'd travel far must cross the seas. Lone wolves are seldom seen in threes; [...] Dear friends! If adages like these Should seem banal, or just a joke, Remember fish don't grow on trees — So say I and so say the folk. (*BNC Sampler*: G11, W:fict:poetry)

(576) JOKES NOT TO TELL AT THE ZOO [...] Why did the fox cross the road? To eat the squashed rabbit! A bee and a butterfly settled on a flower together. 'Funny how your stripes look like a rugby jersey, isn't it?' said the butterfly. (*BNC Sampler*: CHR, W:misc)

Whilst considering only metalinguistic *joke*, Schmid (2000: 145) suggests that factual interpretations seem to arise from *th-be*-N occurrences of the lemma. Examples (577) and (578) are two cases in point, as *joke* refers to 'a person, thing or situation that is ridiculous or annoying and cannot be taken seriously' (*OALD*). The ridiculous thing in (577) is the amount of subsidy, while in (578), it is the situation described. With respect to the latter example, *the next one* refers to a client and to a funny situation, but not to a metalinguistic joke. Thus, the intended meaning is not that of *I am going to tell a joke*, but that of *this situation and what happened to this person was funny*. On these grounds, (577) and (578) seem to fit the family of Factual Attitudinal Descriptive 'Irony' nouns (Schmid 2000: 127–9, e.g. *coincidence*, *irony*, *paradox*, *folly*, *oddity*, *absurdity*), where facts or situations are described as remarkable due to certain qualities (in this case, the quality of being ridiculous or funny).

(577) After the conference in conversation with the section secretary and myself, he stated that the fiver subsidy was a joke. National Power were buying coal from Australia for half the price of the subsidized figure. (*BNC Sampler*: HDT, S:speech:scripted)

(578) the next one was a joke. Bloke come and bought a Teletext with his friend, come all the way from Falmouth, he pulled on the front, paid me his money, saw the telly working, put the remote in his pocket, my son carried the telly down, put it in the back of the car <pause> the bloke

shut the boot, he drove off the front, I picked up the phone <pause> he er and he just swung round off the front, the boot come open <pause> out come the telly!
Oh goodness.
So he just got out, picked it up, shoved it back in and drove away. (*BNC Sampler*: KC1, S:conv)

In cases like (579) and (580), *joke* is not a funny story or a funny situation, but a trick or 'something that you do to make somebody believe something which is not true, or to annoy somebody' (*OALD*). The eventive and purposive sense in these examples (i.e. something done to annoy somebody) ties in well with Schmid's (2000: 266–8) family of Specific Eventive Purposive 'Attempt' nouns, where semantically similar units like *ruse* or *trick* occur.

(579) She gets a smoke cos her, her brother smokes stuff and sometimes he gives it to her and she blows it in her dog's face. And then she tried to put it in the washing machine for a joke and the dog was gonna get in. <pause> The dog's so stupid. (*BNC Sampler*: KPG, S:conv)

(580) The Duke had a passion for practical jokes-wetting people with hoses and putting itching powder in their beds — and he was a devoted animal lover, always surrounded by dogs and preferring the ugliest because no-one else would be kind to it. (*BNC Sampler*: CBB, W:non_ac:humanities_arts)

xvi) Phenomenon: Of the 36 instances of this lemma, 34 fall into Schmid's (2000: 93–101) family of Factual General Neutral 'Thing' nouns (e.g. *fact*, *point*, *case*, *business*, *phenomenon*). One such example is (581), where *phenomenon* implies 'something that happens or exists in society, science or nature, especially something that is studied because it is difficult to understand' (*LDCE*), this *something* corresponding here to *turbulence in lasers and other optical systems*.

(581) Turbulence in lasers and other optical systems is a newly recognised rather than new phenomenon. (*BNC Sampler*: J2H, W:ac:nat_science)

However, two of the concordances for this lemma carry the sense of 'something or someone that is very unusual because of a rare quality or ability that they have' (*LDCE*). This meaning is not in Schmid's (2000) taxonomy. The semantic tag applied in these cases is that of Factual Attitudinal Descriptive 'Irony' (Schmid 2000: 127–9), a category comprising nouns like *coincidence*, *irony*, *accident*, *paradox*, *curiosity* and *oddity*. According to Schmid (2000:

127), these nouns describe '[...] facts as being remarkable or noteworthy because of certain qualities'. Whilst, at first glance, *Japan* in (582) is not a fact but a first-order entity (i.e. a country), the surrounding context seems to impose an abstract reading. It is not the geographical beauty of the island nation of Japan that is considered remarkable, but its economy or, better, its outstanding economic performance.

(582) As to whether change is necessary or not one cannot help pay some regard of that phenomenon of the post war world, Japan: If time travel was a fact and it was possible to transport someone from the middle of the social scale of Victorian Britain to the present time, he would think a revolution had taken place; but the basic ground rules of social life and commerce would shortly become comprehensible to him. The equivalent Japanese, transported to modern Tokyo would think he was on a different planet! (*BNC Sampler*: G0C, W:commerce)

xvii) Point: *Point* is the most versatile of the lemmas studied here. It covers eight semantic tags straddling Factual, Linguistic, Mental, Eventive and Circumstantial uses. The two most frequent uses fall into the Linguistic Propositional 'News' and Factual Attributive Part-Whole 'Aspect' categories (Schmid 2000: 140–4, 116–18), both accounting for 26 of the 40 examples for this lemma. The former is shown in (583) and (584), where *point* is equivalent to 'a thing that somebody says or writes giving their opinion or stating a fact' (*OALD*), the expressed opinion here being the underlined discourse segment. Linguistic instances are often the objects of *make* and *raise*, both synonymous with *mention*. However, in cases like (584), the linguistic interpretation is borne out by the occurrence of other linguistic nouns in the antecedent: *a proposal that the European Parliament have the right to initiate legislation*, *a suggestion that the Council of Ministers meet in public*. The 'Aspect' sense manifests itself in (585), where *point* implies 'a particular quality or feature that somebody/something has' (*OALD*). In this instance, the partitive meaning draws on the writer's suggestion that, prior to cutting the fabric, attention should be given to some of its qualities or aspects.

(583) A, a point on programming too er I make really because of the discussion at the General Purposes Committee last week. Erm, we have looked again at the programme for building this road. There's certainly no way we can shorten it and in fact erm s-- programme timing is very dependent on the ease or difficulty with which we actually get the land.

Er, so I just make that point because of the debate last week. Thank you. (*BNC Sampler*: J44, S:meeting)

(584) Mr Baron's 10 points yesterday were not fleshed out either, but they go further than MEPs have yet done. They include a proposal that the European Parliament have the right to initiate legislation, albeit perhaps jointly with the Commission; and what is called 'a joint decision-making by Parliament in any new areas introduced under the new treaties'. [...] There is also a suggestion that the Council of Ministers meet in public when it is acting as a legislative body, a suggestion which some parliamentarians first attribute to Mr Tony Benn in his days as Energy Secretary. (*BNC Sampler*: A93, W:newsp:brdsht_nat:report)

(585) 8 CUTTING OUT AND HANDLING FABRICS Look at your fabric carefully before starting to cut out, as there are several points to observe. Give yourself plenty of room, particularly when dealing with floor-length curtains. Work on the floor if necessary. Strictly speaking all fabrics should be cut on the straight grain. True straight grain occurs when the weft (crosswise) and warp (lengthwise) threads are at right angles to each other. [...] The fabric will be on grain if its width is square with the end of the table. [...] Problems occur on some printed fabrics when the pattern has been printed off grain. With a small print, follow the design instead of the weave so that the curtain looks right visually. [...]Unfortunately, the problem increases with multiple-width curtains. Either strike a balance between the pattern and grain or, if the problem is very bad, return the fabric before cutting. (*BNC Sampler*: GUB, W:misc)

Point may also refer to 'the main or most important idea in something that is said or done' (*OALD*) or 'the most important fact or idea' (*LDCE*). Examples (586) and (587) illustrate this. Their occurrence in the N-*be-that* pattern is indicative of their focusing function (Schmid 200: 95–6), reducing the semantic potential of the noun in favour of its emphasis on the following *that*-clause. For example, in (586), the shell-noun phrase and its lexical realisation are roughly equivalent to *what is important/what really matters is that estuary and coastal zone management is now a reality*. On these grounds, and in line with Schmid (2000: 93–101), these examples are assigned to the family of Factual General Neutral 'Thing' nouns.

(586) Now I have considerable respect for the House of Commons' Environment Committee but I think on this occasion they must have been tired and emotional when they wrote that.
<clears throat>The point of issue is that estuary and coastal zone management is a reality and what seems to be happening is all sorts of different plans are popping up on various estuaries and Geoff knows

them well now because he's been involved with the one on the Stour and Orwell and what is happening is that these are starting at local level so the first contact with sailors may be from a local authority or someone at local level. (*BNC Sampler*: J3W, S:meeting)

(587) So er, she phoned the corporation and she asked the man, to, to say she was leaving on the Saturday morning <pause> and <pause> were they, were they gonna come and <pause> er ba-- er what wi-- , er, bar up the windows?
Mhm.
And their, well their answer's says are they broken? She said no. Well nothing we can do! But the point is David, they are broken, now! So, they've had to get men in, to board the windows up <pause> then if somebody else gets the house, they've got to get these men to come and take the boards down again, then they glazier has to put the window in. This, to me, is a waste of money! (*BNC Sampler*: FX5, S:brdcast:discussn)

Attention should also be drawn to the '[...] additional covert meanings [...]' (Schmid 2000: 94) underlying focusing constructions. Schmid (2000: 121–5) accounts for these senses by including cases of N-*be*-cl *point* with 'problem' nuances in the family of Factual Attitudinal Impeding 'Problem' nouns. However, 'Problem' occurrences are not coded in this study as Factual 'Problem' nouns, but as Factual Neutral 'Thing' nouns. This is borne out by the fact that 'Problem' does not belong to the semantic frame for the lemma itself, but arises from its use in a Situation-Problem-Solution discourse situation (cf. Hoey 1983 in 2.2.2.1.2). Thus, as with *thing* below, the additional 'Problem' nuances are between square brackets (i.e. Factual General Neutral 'Thing' [Factual Attitudinal Impeding 'Problem']). Example (587) illustrates one of these problem-oriented uses of *point*. In this discourse context, a situation is presented, a woman inquiring as to when her windows are going to be barred up. Following the situation, the adversative conjunction *but* in *but the point* introduces the problem, namely that the windows were not broken before but they are now. This leads to a solution, getting men in to board the windows up, which leads to another problem: the solution is a waste of money.

In addition to Linguistic and Factual uses, the evidence for *point* also contains instances of Mental and Circumstantial senses. For example, *talking points* in (588) is coded as a Mental Conceptual 'Idea' noun (Schmid 2000: 189–94). The reason for this lies in the definition of *talking point* as 'a subject that is talked about or discussed by many people' (*OALD*). Thus, *one of the talking points*

in (588) may be rephrased as *one of the subjects/issues/ideas to talk about this morning*. In Schmid's (2000: 189) framework, nouns like *point*, *idea*, *topic* or *issue* are treated as mental entities or concepts, hence the classification of example (588) into the Mental Conceptual 'Idea' family. *Point* in (589) is also Mental in meaning, but its specific sense is not considered by Schmid (2000). The use of the lemma in this instance implies 'the purpose or aim of something' (*OALD*), the purpose here relating to the charging for every television set. Based on the semantic affinity with *purpose*, (589) is coded as a Mental Volitional Detached 'Purpose' use (Schmid 2000: 218–19; see also *application* above).

(588) Three British hostages are on their way home from Iraq, after being freed from a jail in Baghdad. The release of Paul <gap desc='name' reason='anonymization'>, Michael <gap desc='name' reason='anonymization'> and Simon <gap desc='name' reason='anonymization'> follows a mercy mission by former Prime Minister, Sir Edward Heath. Iraq is denying that the move is a bid to have sanctions lifted. Well it's one of the talking points this morning, if you'd like to give me a call. Has a deal been struck do you think, what's in it for Saddam, what's in it for Britain? (*BNC Sampler*: HM4, S:brdcast:discussn)

(589) Now, I must say that I see there is a point of charging for every television set, erm, I don't have any er, er objection to that, but one rider, that senior citizens should be helped. (*BNC Sampler*: HM4, S:brdcast:discussn)

The last two senses concern Circumstantial and Eventive uses. With regard to the former, *the point* in (590) is tagged as a Specific Circumstantial Temporal 'Time' noun (Schmid 2000: 282–4) based of its semantic affinity with *time*, *stage* or *moment*. However, in (591) *point* is clearly eventive in meaning and, as such, it is classed as General Eventive Neutral 'Event' item (Schmid 2000: 262–6). Along with the Mental 'Purpose' sense above, the eventive meaning of *point* is not included in Schmid (2000) either. This is not surprising, as it is quite likely that the speaker used a noun that he did not mean to use. He was probably looking for a better Eventive noun, but ended up using a highly general shell noun instead, i.e. *point*. The Eventive interpretation stems from its occurrence with the verb *do*: facts, ideas or arguments are not done (actions are).

(590) Once again this autumn, I lost the race with the squirrels to harvest the hazel nuts. I checked daily, in September, until the point where they were almost ripe. (*BNC Sampler*: C9C, W:pop_lore)

(591) [...] and one of the actual points that were starting to do now <pause> is we're actually going to start reporting tu-- , temperature and humidity on the first, second, and third floors within GP House. (*BNC Sampler*: FLS, S:meeting)

xviii) Practice: This noun features only in Schmid's (2000: 284–8) Specific Circumstantial Manner 'Way' uses. This meaning is shown in (592), where *this practice* encapsulates *pairing. Pairing* is a method or procedure allowing Members of Parliament to abstain from voting by coming '[...] to an arrangement not to vote as (one of) a pair' (*OED*). Similarly, *industrial practices* in (593) refers to industrial procedures or methods such as building a car in 19 hours thanks to semi-automatic aids.

(592) We've all seen debates on TV, er, only an handful of MPs are present. Where are all the rest. Paired off. [...] He explained the Party line, in favour. He then explained his own position, which was quite different. Strongly against, but he could see the need for some front bench MPs to pair. New MPs tend not to pair and will attend most sittings, he said. [...] At my first important debate, I paired up with five of the Tory buggers and then I turned up. Pairing allows Tory MPs to sit on their boards of directors, manipulate your pension funds, to rally support for their party, and enhance their salaries. We should not be supporting this practice. (*BNC Sampler*: HDT, S:speech:scripted)

(593) Meanwhile, beyond our shores, the world's industrial practices and capacities advanced. In the late 1960's, a Volkswagen car took 19 man-hours to build, thanks largely to semi-automatic aids. The most efficiently-built British car of comparable size took 85 man hours. (*BNC Sampler*: G0C, W:commerce)

However, there are cases where the procedural meaning of *practice* is combined with a sense of 'habit or custom' (*OALD*). In examples like (594) and (595)*, practice* relates to old traditions and customs, i.e. events or activities that '[...] have been carried out in the same way for a long time' (Schmid 2000: 268). In Schmid's (2000: 268) framework, *tradition*, *habit*, *convention*, *custom* and *ritual* are all classified as Specific Eventive Habitual Temporal Manner 'Tradition' nouns. Following Schmid (2000), instances like (594) and (595) are coded in the same way.

(594) It is probable that in very early times if his powers weakened a king would be ritually killed. This practice was abandoned before the Dynastic Period but was symbolically retained in the Heb-sed or jubilee festival. (*BNC Sampler*: EVR, W:non_ac:humanities_arts)

(595) Goryushkin's conclusion is that in the process of peasant immigration and colonization, not only did the incoming millions bring about innovations in agricultural methods, animal husbandry, cottage industry and the social customs of the Siberian peasantry, but they themselves were also deeply influenced by the **traditional local practices of the old Siberians.** (*BNC Sampler*: FB4, W:ac:humanities_arts)

In (596), *practice* entails neither a procedure nor a custom, but the 'repeated exercise in or performance of an activity so as to acquire, improve, or maintain proficiency in it' (*OED*). The repeated exercise in (596) involves adding (i.e. this student has had a lot of repeated exercise, and this exercise involves adding). Considering that this sense of *practice* entails doing something with a particular aim or purpose in mind (i.e. improving someone's skill), (596) fits into Schmid's (2000: 266–8) Specific Eventive Purposive 'Attempt' nouns (e.g. *attempt*, *struggle*, *endeavour, effort*). As these nouns '[...] shell the ACTIVITIES undertaken in the pursuit of aims [...]' (Schmid 2000: 266) and, for want of a better tag, examples like (596) are coded thus.

(596) Cos five add seven <pause> seven add five would come to twelve. Now <-|-> <unclear> <-|->
<-|-> Twenty four.
add up to twenty four okay and <-|-> right okay <-|->
<-|-> And there's thirty six there.
So you had **a lot of practice** <u>adding there</u>.
Now what I want you to do <pause> I want to just have a look at it now and I want you to <pause> play with it over the holiday on your own. (*BNC Sampler*: FUH, S:classroom)

The purposive meaning of (596) is absent from (597). In this particular instance, *practice* implies 'the actual application or use of an idea, belief, or method, as opposed to the theory or principles of it' (*OED*). Thus, *practice* in (597) does not carry a sense of repeated exercise with a goal in mind (i.e. to improve someone's skill), but the act of testing a theory on a particular occasion. Such an act is here implied in the underlined segments, where the protagonist tries to do as she was told by her friend. On these grounds, the example is coded as General Eventive Neutral 'Event' (Schmid 2000: 262–6), a category subsuming such general nouns as *change*, *event*, *act* and *action*.

(597) If 'they' tried to kiss you with your mouth open you must keep your teeth tight shut. (Well who could want such sloppy kisses anyway thought I)

> And if you lay full length with them on the beach it was a 'sign' you would go all the way. [...] It was two years before I tested Joanna 's theory in practice. Janina's family took me to Switzerland for a holiday. We went dancing in discos and an Italian boy, Giacomo, kissed me at the end of a dance. On Joanna 's instructions I set my teeth resolutely. When I sat down I told Janina 's mother , expecting her to tutut approvingly at my astute virtue. Instead she threw back her head in laughter and said, 'Next time he tries, open your mouth.[...] Joseph must have been a country boy . He took mountains in his stride. Up and up we went and when I could walk no longer he lay me down and covered me with kisses and love bites. (*BNC Sampler*: FU7, W:non_ac:soc_science)

xix) Scandal and misfortune: Schmid (2000: 129) classifies *scandal* and *misfortune* under Factual Attitudinal Evaluative Negative 'Tragedy' nouns (e.g. *pity*, *tragedy*, *disaster*, *offence*, *disgrace*). He posits that Factual uses of these attitudinally negative nouns occur only with *that*-clauses in N-cl and N-*be*-cl patterns. Subsequently, prior to discussing attitudinal Eventive uses (Schmid 2000: 270–4), he argues that nouns like *scandal*, *misfortune*, *tragedy* or *disaster* may receive Eventive interpretations in *th*-N and *th*-*be*-N patterns. This being the case, no explicit category is introduced to accommodate the Eventive uses of the nouns in the Factual 'Tragedy' family. Eventive Attitudinal nouns in Schmid's (2000: 270–4) framework fall into three families: Eventive Attitudinal Manner Deontic 'Trouble' (e.g. *problem*, *trouble*, *difficulty*), Eventive Attitudinal Positive 'Success' (e.g. *success*, *achievement*, *improvement*) and Eventive Attitudinal Negative 'Mistake' (e.g. *mistake*, *crime*, *fault*). None of these three categories capture the sense of negative situation entailed in nouns like *misfortune*, *scandal* or *disgrace*. For this reason, a new semantic category was created to accommodate Eventive senses of Factual 'Tragedy' nouns: Eventive Attitudinal Negative 'Tragedy'. Examples (598) and (599) illustrate the Factual use of *scandal* and *misfortune*, this corresponding to a general state of affairs in (598) and to a mental entity (i.e. a feeling) in (599). In (600) and (601), *scandal* and *misfortune* are Eventive in meaning, as the scandal is an underhand action by the government (i.e. the running of professional death squads), whilst the misfortune is caused by a recent event in the princes' lives (i.e. the change of abode and servants).

(598) <-|-> Now <-|->, I'm a pensioner, and I think it's a scandal what we get <pause> for a pension after all the years that you, you fight for you, you

put away for your old age, and before you know it you've nothing! (*BNC Sampler*: FX5, S:brdcast:discussn)

(599) And what was the attitude over and above, other than obviously coming in and out of work, what was the attitude of the local neighbours and whatnot towards <-|-> them <-|->?
<-|-> Well <-|-> er th-- er this is one of the misfortunes er is it not, you see, this bad feeling.This bad feeling er lived on into old age. (*BNC Sampler*: FYJ, S:interview:oral_history)

(600) THE man responsible for uncovering South Africa's hit-squad scandal, the condemned security branch policeman Almond Nofomela, is being brought to court tomorrow by the authorities who are seemingly intent on hurrying him on to his postponed appointment with the hangman. Mr Nofomela is to appear in a Natal magistrate's court to be charged with the murder of the Durban civil rights lawyer, Griffiths Mxenge. It was Mr Nofomela's confession to the killing — the day before he was due to be executed for the murder of a white farmer — which led to the disclosures that the force has been running professional death squads. (*BNC Sampler*: A9V, W:newsp:brdsht_nat:report)

(601) The change of abode and servants was not the princes' only misfortune. (*BNC Sampler*: CCD, W:fict:prose)

xx) Surprise, anger and terror: These three units belong to a group of shell nouns denoting feelings or emotions resulting from a particular situation. Therefore, the typical relationship of experiential identity between shell noun and lexical realisation (A is B, B is A) does not hold, as the encapsulated segment is not what the feeling is, but rather the cause of that feeling. According to Schmid (2000: 226–30), the nature of the causing situation may be either Fact-Related or Event-Related. On these grounds, Schmid proposes two categories of Emotive uses: Mental Emotive Fact-Related 'Surprise' (e.g. *surprise*, *relief, pleasure*, *shock*, *regret*, *pain*, *terror*) and Mental Emotive Event-Related 'Fear' (e.g. *fear*, *concern*, *worry*, *anxiety*, *apprehension*, *dread*, *premonition*). The 'Fear' family is well suited for such lemmas as *foreboding* in 5.3.5.1, where the anxiety or fear is caused by the uncertainty surrounding an event yet to take place (Schmid 2000: 228). However, members of the 'Surprise' family like *surprise*, *anger* or *terror* do not carry a sense of futurity in their instantiation in discourse. Thus, in order to accommodate Event-Related uses of these nouns, a new category was created, that of Mental Emotive Event-Related 'Surprise' uses. Thus, instances of *surprise*, *anger* and *terror* are classified as either Mental Emotive Fact-Related 'Surprise' or Mental Emotive Event-Related 'Surprise', according to whether the cause of the feeling is factual or eventive.

The distinction between one subtype and the other drew on the following coding policy[23]:

- Encapsulated *that*-clauses seem to be associated with Fact-related interpretations, e.g. (602), whilst any other structure causing the feeling is often Even-Related, e.g. (603). A possible paraphrase for (602) is *the fact that parliament should have approved a conservative plan is no surprise*. Whilst the approval of the plan was a specific event, its reporting and evaluation in this newspaper article (*no surprise*) is factually-oriented. This is further supported by the strong link that Schmid (2007; see 3.2.5.2) establishes between *that* complement clauses and factual shell nouns. A similar factual paraphrase for (603) seems unlikely (*my surprise was the fact of discovering a mid-19th century edition in San Diego*). In this particular instance, it makes more sense to relate the feeling of surprise to the act of discovering, hence the suitability of the Eventive interpretation.

(602) It is no surprise that parliament should have approved a conservative plan. (*BNC Sampler*: AAK, W:newsp:brdsht_nat:report)

(603) It was at this time that we read Morgan Forster's article in the Listener about George Crabbe and we were touched and interested enough to seek out a copy of Crabbe's poems and I well remember my surprise in discovering a mid-19th-century edition in San Diego, I think. (*BNC Sampler*: J55, W:non_ac:humanities_arts)

- The shelled content of Fact-Related feelings is often relational, as in (604) and (605), whilst that of Event-Related feelings tends to be material, as in (606) and (607). Hence, in (604) and (605), the surprise is caused by the fact that the beef surprise was dog food and by the fact that the victim was not a maidservant but a duchess (amongst other factual details). By contrast, *surprise* and *anger* in (606) and (607) are not caused by relational states of affairs, but by actual events (the moment Middlesbrough took the lead and the rise of Mr Rene Muawad to power).

(604) The beef surprise was lovely, but what's the surprise?
It was dog food, sir. (*BNC Sampler*: CHR, W:misc)

[23] Many of the examples given in the following are for the noun *surprise*. This is due to the more substantial evidence for this lemma in the study sample (25 concordances) compared with, e.g. *anger* (8 instances) or *surprise* (5).

(605) Indeed , your latest victim seems to have been a veritable bundle of surprises — not a maidservant but a duchess , not dead but alive — and , to top all , affianced to an eleven-year-old and not a virgin! (*BNC Sampler*: CCD, W:fict:prose)

(606) At the other end Thompson's flag kick was headed into the net by Linighan wide of the far post but a linesman had his flag up very early and the effort did not count. It was no great surprise when Middlesbrough took the lead. Williams fouled Hendrie just outside the penalty area on the Middlesbrough left. (*BNC Sampler*: CF9, W:newsp:other:sports)

(607) LEBANON'S new President, Mr Rene Muawad, yesterday worked to weld old militia foes into a cabinet to govern his divided country while fellow Christians, demonstrating against him, shut half of Beirut. Supporters of General Michel Aoun, staged noisy street protests in the Christian enclave. About 1,000 marched on the French embassy to denounce support for the Arab peace plan which brought Mr Muawad to power on Sunday. 'France has deceived us,' one poster said. Another charged that Lebanon's Christians had been let down by President Franois Mitterrand. [...] The anger expressed in East Beirut reflected a sense of betrayal in the Christian camp. (*BNC Sampler*: A7V, W:newsp:brdsht_nat:report)

- Feelings caused by linguistic information describing events (e.g. *account*, *story*, *news*, etc.) are interpreted as Fact-Related, in that the person has not experienced the events first-hand, but reacts to someone's account of the events. In other words, the events are filtered by someone's report and thus the events become factual entities. For example, in (608) and (609), children and UN officials are terrified and enraged by the events described in a terrible story and by the report or news of forcible Vietnamese repatriations.

(608) 'Gorbrandt wanted Møn for the amber he knew could be picked up on our shores. And so that he might ensure that our island would be his for all time, he instructed his hordes to slaughter all the children on Møn. 'Gorbrandt was a coarse killer, but a subtle man. He reckoned that the islanders would be so devastated by the deaths of their children that they would be incapable of taking up arms against the invaders and, later, would be easily subjugated. [...] 'News of the massacre of the children reached Sandweg a full two hours before the invaders. By the time that the invaders poured across our flatlands this church had been manned and victualled and the women of the parish safely concealed in outlying farms, under the protection of what you, today, would refer to as 'teenagers'. The youths had been trained especially to protect the women. And, of course, the women took their children with them. [...]

The children turned to face one another for the first time since the pastor had started to tell the story of the massacre. He had discharged **their terror** by his final sentence. (*BNC Sampler*: JNG, S:meeting)

(609) OFFICIALS at the office of the United Nations High Commissioner for Refugees in Geneva reacted with **resigned anger** to the news of yesterday's forcible repatriation of 51 Vietnamese from Hong Kong. (*BNC Sampler*: A9V, W:newsp:brdsht_nat:report)

- Feelings caused by someone's realisation/knowledge about what is happening in their immediate context are coded as Event-Related, as the event and the knowledge about the event are experienced by one and the same person (cf., however, (608) and (609), where feelings are caused by second-hand reports). Thus, in (610), Pugwash's *surprise* stems from his realisation of what was going on (he was surprised at seeing that his enemies had lowered their hands, that drinks were being served, etc.). Similarly, in (611), Re is angered by his finding out that something treacherous was happening behind his back (men scheming against him).

(610) And now, to **his** horror and **surprise**, Pugwash realised that his enemies had lowered their hands, that drinks were being served to them and that the soldiers were transferring their attentions to himself, the Mate, Willy and Barnabas. (*BNC Sampler*: ALS, W:fict:prose)

(611) He became aware that men were scheming against him and therefore summoned a council of gods. The primordial deity, Nun, advised Re to use his powerful Eye, the sun itself and possessor of its own complex mythology, to exact vengeance on the evildoers, and furthermore to send the Eye in the person of Hathor. The goddess Hathor is thus given an uncharacteristic fierce role, but one which she fulfilled with great zeal. She found those who had fled from **Re's anger** in the desert and killed very many of them, thus gaining the name of Sekhmet, the lioness goddess of war, with whom she is here identified. (*BNC Sampler*: EVR, W:non_ac:humanities_arts)

xxi) Terror: *Terror* features only in Schmid's (2000: 226–8) family of Mental Emotive Fact-Related 'Surprise' nouns. However, in (612) and (613) *terror* does not imply the feeling itself, but 'violent action or the threat of violent action that is intended to cause fear, usually for political purposes' (*OALD*). The kind of violent action involved in (612) and (613) is shown in the information interspersed throughout the surrounding co-text (e.g. brutal repression of protesters, firing

on fellow citizens, loss of life, factories under guard, etc.[24]). Uses of *terror* with this sense are classed as Specific Eventive Purposive 'Attempt' (Schmid 2000: 266–8). This category comprises units shelling '[...] the ACTIVITIES undertaken in the pursuit of aims [...]' (Schmid 2000: 266; e.g. *attempt*, *campaign*, *struggle*). In the case of *terror*, the aim behind the activity or action is to cause fear for political purposes.

(612) In Bonn, Chancellor Helmut Kohl was said to be 'deeply shocked over the brutal repression of the justified and peaceful protests'.[...] condemning the 'brutal action of the Romanian state power against peaceful citizens demonstrating to assert their elementary human rights'. Thousands of East German troops urged their comrades in the Romanian forces not to fire on their fellow citizens peacefully demanding their human rights. Demonstrators picketing the Romanian embassy in East Berlin urged the people of other East European capitals to hold simultaneous rallies on Saturday afternoon 'against the bloody terror of the corrupt Romanian dictator'. (*BNC Sampler*: AAK, W:newsp:brdsht_nat:report)

(613) By Foreign Staff OUTRAGE and concern at the violence in Romania during the weekend came from all parts of the world yesterday, with the Soviet Union and the US in the lead. [...]European foreign ministers, meeting in Brussels, condemned 'in the strongest possible terms' the measures taken by the Romanian security forces. [...]he said 'if indeed some loss of life has occurred, I can only express my very profound regret'. [...] It reported that Romanian state institutions and factories were under intensified guard, and that the frontiers were closed to tourists. [...]The Polish Parliament stood for a minute's silence after approving unanimously a resolution which accused the Romanian authorities of 'exceptional brutality' leading to children being shot at and expressed solidarity with the 'victims of terror.' The World Council of Churches, the World Alliance of Reformed Churches, the Lutheran World Federation and the Conference of European Churches said in a joint telegram they were 'deeply disturbed by the disquieting intervention of the army against civilian people in Timisoara.' (*BNC Sampler*: AAB, W:newsp:brdsht_nat:report)

xxii) Testimony: The only sense accounted for in Schmid (2000) is that of 'a formal written or spoken statement saying what you know to be true, usually in court' (*OALD*). This is shown in (614) and (615), both coded as Linguistic Illocutionary Assertive Public 'Proclamation' (Schmid 2000: 155–6; e.g. *announcement*,

[24] See 4.4.2.8 for further examples of uncountable generic self-contained shell-noun uses

proclamation, *statement*). The Public feature of this family conveys the official and well-planned nature of these nouns.

(614) THIS mid seventeenth-century story is recorded in the second edition of Richard Butcher's Survey of Stamford of 1717 and is taken from a testimony written by Samuel Wallis, a Stamford shoemaker, upon whom the 'miracle' was performed. Wallis was critically afflicted with consumption and William Foster, in a letter published in Peck's Annals , claimed that Wallis had been sick for thirteen years. On Whitsunday 1658 a stranger called at his house and begged for some small beer. He then instructed Wallis to take a herbal remedy of red sage leaves for twelve days after which time he would be cured. The miracle occurred just as predicted. Some people considered this stranger to be a devil changed into the illusion of an angel of light, or a witch. (*BNC Sampler*: CBB, W:non_ac:humanities_arts)

(615) The presence of hundreds of journalists for the game ensured the kind of publicity that Miami preferred not to have. Extraordinary precautions have been taken to defuse tempers in case of acquittal. Key testimony has been broadcast live on television and radio to help keep the city fully abreast of the proceedings and to show the wheels of justice turning. Mr Lorenzo contends he fired in self-defence because he was about to be run down by the motorcycle, which was being chased by a police car. Mr Lloyd was carrying marijuana and $1,514 in cash, and his blood contained alcohol , marijuana, and traces of cocaine. (*BNC Sampler*: A8W, W:newsp:brdsht_nat:report)

However, in such instances as (616) and (617), *testimony* carries the sense of 'a thing that shows that something else exists or is true' (*OALD*). On the basis of its semantic affinity with units like *evidence*, *proof* or *indication*, the most appropriate category for this sense is that of Factual Causal Mental 'Evidence' nouns (Schmid 2000: 110–13).

(616) Chief executive Sid Hopkins said the results represented the biggest year-on-year improvement achieved by the group. 'It is testimony to the determined and vigorous actions taken during the past two years. (*BNC Sampler*: CF8, W:newsp:other:report)

(617) Donors will help, Britain's commitment of five hundred and fifty million pounds for the special programme of assistance to Africa since nineteen eighty, is testimony to that <pause> but African governments must own their reform programmes. (*BNC Sampler*: JNG, S:meeting)

xxiii) Thing: Schmid's (2000) categorisation of this highly unspecific noun comprises three families: Factual General Neutral 'Thing', Factual Causal 'Reason' and Factual Attitudinal Impeding 'Problem'

(Schmid 2000: 93–101; 102–6; 121–5). *Thing* instances often occur in N-*be*-cl focusing constructions (e.g. *the thing is*, *the thing is that*), which, according to Schmid (2000: 94), '[...] express additional covert meanings on top of the 'pure' concept-forming effect'. These covert meanings explain the classification of *thing* under the 'Reason' and 'Problem' families[25]. In view of the highly unspecific nature evident in the factual uses of this lemma, a decision was made to disregard this threefold distinction, and to code all instances with 'Problem' or 'Reason' nuances in terms of the Factual General Neutral 'Thing' category. This decision is supported by the nature of focusing constructions, whose chief function is to highlight the following *that*-clause. Any 'Problem' or 'Reason' nuances that might arise from these constructions are secondary to that focusing function. In these cases, any covert meanings are indicated between square brackets (as with *point* above).

Example (618) illustrates the Factual Neutral use of *thing*, with no added nuances. By contrast, examples (619) through (621) convey additional meanings, (619) and (620) being coded as Factual General Neutral 'Thing' [Factual Attitudinal Impeding 'Problem'], and (621) as Factual General Neutral 'Thing' [Modal Deontic Necessary 'Necessity']. The 'Problem' nuances of the former two examples are lexically signalled. In (619), *though* introduces the contrast between a positive state of affairs (the existence of five good teams) and a problem (the thing or problem being that whilst five teams are good, all the rest are rubbish). Example (620) underlines the 'Problem' meaning through the premodifying Epithet *difficult*. The impeding situation or problem here is not knowing when to stop. In (621), the Modal 'Necessity' interpretation stems from the surrounding discourse context. The *to*-infinitive clause colours the factuality of *thing* with modal deontic nuances, as *the important thing* conveys the speaker's plea for people to listen to genuine concerns and to address political problems. With this in mind, (621) might be rephrased as *what people really need to do is to listen to the genuine concerns and to sort out these problems*.

(618) Mr Prokhorov should have the last word. 'The main thing is perhaps the fact that we are now having a fresh look, without blinkers, at

[25] Delahunty (2012: 66-9) takes issue with Schmid's (2000) classification of *thing* instances, stating that focusing constructions may imply not only reasons or problems, but also solutions, premises and conclusions (see footnote 3 in 2.2.2.2 for details about Delahunty's 2012 description of focusing constructions).

ourselves and our country,' he said. (*BNC Sampler*: A7V, W:newsp:brdsht_nat:report)

(619) why do Scotland always <-|-> score more goals?
<-|-> I think the goals <-|-> must be bigger or something, but they always get loads <-|-> they always <-|->
<-|-> they do <-|->
<-|-> and the goals <-|->
<-|-> always high scorers <-|->
are far more fu-- , fun to watch because they always fly in from miles out and if you look at the Scottish <-|-> results <-|->
<-|-> the thing <-|-> is in Scotland though, there's five good teams and the rest are rubbish
yeah (*BNC Sampler*: KD6, S:conv)

(620) I read a book by the American painter Charles Reid who said that rather than paint a dark background all round a daisy it is sometimes better to put just one dark drop of colour behind the vase of the petal to get the most impact. The most difficult thing is to know when to stop and when I feel as though the painting is nearly there I usually stop , make myself a cup of coffee , and contemplate what I have achieved from the comfort of a kitchen stool. (*BNC Sampler*: CN4, W:pop_lore)

(621) [...] in one of their er proposals and objectives they say encouraging self regulation and observance of Code of Practice by local clubs and groups to avoid clon-- conflict with and or disturbance to other users including nature conservation interests. Now if that is really saying to the water users that it's down to you, there are genuine concerns of conservation and I think they are, there is a real problem in some areas. The important thing is for people to listen to the genuine concerns, sort out the, the real issues from the power politics and then hopefully go forward in, in a, in a manner of conciliation. (*BNC Sampler*: J3W, S:meeting)

All Factual instances of *thing* occur with relational processes, especially in intensive Identified positions (e.g. *the thing is*...). However, in (622) and (623) *thing* depends on verbal and material processes, respectively. In these cases, it seems reasonable to apply a linguistic and eventive categorisation. This is supported by the oddness of such factual paraphrases as *I was about to ask the same fact/point/case*[26] and *The first fact/point/case/problem that Daniel did was to remove the record*[27]. On these grounds, *thing* in (622) is tagged as Linguistic Propositional 'News', and as General Eventive Neutral 'Event' in (623). Whilst it would seem appropriate

[26] Cf., however, *I was about to ask about the same fact/point* or *I was about to ask the same question.*

[27] Cf., however, *the first action he took was to remove the record.*

to treat linguistic *thing* as an illocutionary noun based on its affinity with *question*, this meaning is context-driven and not part of the semantic frame of the lemma. Therefore, it made more sense to use the default or most general linguistic tag, that of such nouns as *information*, *message* or *news*.

(622) Do you want to tell us what interest there was.
I, I was about to <-|-> ask the same thing, yeah <-|-> . (*BNC Sampler*: DCH, S:meeting)

(623) The first thing that Daniel did when he came back into the room was to remove the record and switch off the gramophone. (*BNC Sampler*: AEA, W:fict:prose)

xxiv) Vision: Schmid (2000: 210–12) categorises this lemma in terms of its Mental Volitional Conclusive 'Aim' use. According to Schmid (2000: 210), 'Aim' nouns (e.g. *point*, *idea*, *hope*, *goal*, *target*, *ideal*) are closely related to the N-*be-to* pattern, where they '[...] shell things that EXPERIENCERS want to achieve by means of certain activities' (cf. also Schmid 2007: 329–33 in 3.2.5.2 for the frequent association between *aim* and the N-*be-to* pattern). One such example is (624), where substitution of *aim* or *goal* for *vision* does not alter the meaning. Example (625) is also tagged as Mental 'Aim', but in this case the volitional interpretation does not stem from the N-*be-to* pattern. It is instead conveyed by the lexical specifics in the preceding discourse, which elaborate on Haine's goals and ambitions (turning Aden into a strategic centre, transforming a ramshackle town into an important city, etc.).

(624) I hope that two two eight can be referred so that that discussion can be taken into account in determining the needs of members in the new world, and therefore the design of the new union. Now, on the Report itself, and just a few quick points. The idea of course the vision is to create a better union, not one that is exactly like us because er, difficult for me to say so, but we have got the occasional fault here and there. (*BNC Sampler*: HLW, S:speech:scripted)

(625) Following his failure to buy or seize Socotra, Haines convinced the government of Bombay that Aden could be made into both a strategic and a great commercial centre. [...] After he had beaten off three attacks to retake Aden, Haines was able to concentrate on his plans for turning a ramshackle town where 600 people, nearly half Jews, lived in squalor into the main entrepôt for Arabia and East Africa. [...] At the same time, offering low rents to those who would agree to build in stone, he laid out a town which after seven years had a population of 25,000, served by a free port. [...] Haines was a strange mixture, a man with a romantic

vision for the future of the first conquest of Queen Victoria's reign, a skilful politician amongst Arabs but not amongst his fellow-countrymen, with a sailor's practicality in such matters as building a town but not in administration. (*BNC Sampler*: GT9, W:biography)

Whilst still in the Mental domain, instances like (626) and (627) are conceptual rather than volitional (i.e. what something is, rather than what you want something to be). These uses accord with the definition of *vision* as 'an idea or picture in your imagination' (*OALD*). The 'idea' or 'image' sense of *vision* is entailed in (626) in the verb *imagine*, whose occurrence in the first line of the extract suggests a conceptual interpretation of the following encapsulated segment. Similarly, the connection between *vision* and *idea* in (627) lies in the occurrence of *concept* and *image* in the same sentence. Instances of *vision* with this sense are thus coded as Mental Conceptual 'Idea' (Schmid 2000: 189–94; e.g. *idea*, *notion*, *stereotype*, *hypothesis*, *concept*, *image*).

(626) As 1989 ends, it is tempting to imagine that Europe is emerging into a wholly new identity, governed by peace not war, common aspirations rather than sectional concerns. In this Europe there is a Benetton in every high street, Badoit and Czech Budweiser in every fridge, an Armani jacket in every wardrobe, Beaujolais Nouveau on every table, cable and satellite television channels in many languages in every living room, an Umberto Eco novel on every bookshelf, a Volvo in every garage, where CDs of The Orchestra of the Age of Enlightenment lie casually next to Eurythmics, and where nipping across to Paris for the day is as natural as doing a day's business in London. It's an attractive vision, up to a point, and it has a certain reality. (*BNC Sampler*: A8W, W:newsp:brdsht_nat:report)

(627) It is, of course, both a name and a concept which readily stimulates knee jerk responses, stereotyped visions and hackneyed images in most people's minds — images which will almost invariably feature great frozen wildernesses, blinding blizzards, steel-shattering frosts, and, of course , legions of fur-wrapped, fettered convicts and political prisoners — 'exiled to Siberia', in the chilling cliché — by the autocratic Russian state. (*BNC Sampler*: FB4, W:ac:humanities_arts)

Another sense of *vision* not recorded in Schmid (2000) concerns neither intended goals nor concepts or ideas, but things perceived through the sense of sight. This is obvious in (628), where *vision* implies the perception of 'a person of great beauty or who shows the quality mentioned' (*OALD*). Example (629) is also visually oriented, but less so than (628), as the sense here is that of 'a

dream or similar experience, especially of a religious kind' (*OALD*). Both senses are subsumed under a new semantic category: Mental Perceptive Visual nouns. None of Schmid's (2000) mental categories account for first-hand direct or indirect visual experiences, hence the suitability of this new tag.

(628) 'His ample figure and full-moon face, with its fringe of curls, were always a pleasant vision, and he had a persuasive manner that was hard to resist [...] (*BNC Sampler*: GT9, W:biography)

(629) [...] it's in first Corinthians, it talks about for, I think I'm in a vision of the third heaven, which was actually a vision, it wasn't a reality it's like a revelation was a vision, I don't think it was the reality of heaven was it? (*BNC Sampler*: KBX, S:conv)

Lastly, underlying instances like (630) and (631) is a contextual sense unaccounted for in any of the dictionaries consulted. In these two examples, *vision* shows a strong semantic affinity with *view*, *standpoint* or *opinion*, lemmas falling into Schmid's (2000: 203–6) family of Mental Creditive Attitudinal 'View' units. Thus, *vision* in (630) implies the way of thinking, perspective or viewpoint of those in authority, whilst in (631) the Attitudinal interpretation is suggested by the occurrence of evaluative expressions in the underlined segment (*I feel perhaps*, *I do just wonder*).

(630) For example, a book about human rights in the context in which we live can be considered subversive because it disagrees with the vision of those in authority. (*BNC Sampler*: EBK, W:misc)

(631) You see my visions for qualitative are slightly different to the four months at the moment, and I feel perhaps could exchange Q P16 for the control procedures to be a longer procedure but to have everything covering the project plan in progress monitoring through to quality control procedures for just the quality. So if you look at something like Q P5 where you have got this point <unclear> qualitated, in fact a lot of the points don't actually apply to qualitated because many of the points relate to erm, D P erm, <unclear> qualitative field as opposed to survey. So I do just wonder if I couldn't pick out point say 1 and 2 of Q P5 and 6 which are about booking and booking forms and confirming the commission of qualitative through to whatever is required during the run off of it and then add the quality control check and finally you have got this, the final review. (*BNC Sampler*: J97, S:meeting)

xxv) Way: Of the 40 instances of *way*, 34 are coded as Specific Circumstantial Manner 'Way' (Schmid 2000: 284–9; e.g. *way*,

approach, *method*, *technique*, *practice*), (632) and (633) being two cases in point.

(632) [...] how are you gonna grab me? what kind of words are you gonna use to grab me
bodies, victims, flames what do those words do to me as I'm reading
I think I might, I dunno I've gotta think of ways to change it so there's more than one line taken
<laugh>You're gonna make the actual statistics worse okay yeah you can do that <laugh>but it's the vocabulary that's not strong enough. (*BNC Sampler*: JJS, S:classroom)

(633) Did they di-- do you think they <pause> I was going to say enjoy there but I do I don't think that'd be the right word, but d-- did they like or did they prefer to work piece work or was it something that they loathed or I mean there was more money to be made at it wasn't <-|-> there <-|-> ? [...] Well it was a system that er <pause> wasn't liked but it was operated because, in my opinion anyway, because the employers er had seen it as a way of getting more work out of you. (*BNC Sampler*: G63, S:interview:oral_history)

Examples (634) and (635) deserve special attention, as they are '[...] systematically ambiguous between manner readings and dynamic modal readings' (Schmid 2000: 286).

(634) Th-- th-- In them days you hadn't got er what is commonly known now as hire purchase, nowadays, you know, that that mostly started after the war.
<-|-> That would've been <-|->
<-|-> <unclear> <-|-> the only way of Maybe not the only way, but the the easiest way for people to er get debt, if you want to call it that, was er to use this Co book. And of course erm for some people it became a way of living, <pause> <-|-> <unclear> and just <-|-> (*BNC Sampler*: G63, S:interview:oral_history)

(635) If their giro didn't come, erm the only way of of contacting DHSS, was either to go down, or to walk all round until they could find a telephone to do it. (*BNC Sampler*: FY8, S:interview:oral_history)

Schmid (2000: 256–7) argues that the Modal orientation of these cases stems from their combination with postmodifying *to*-infinitive clauses, which introduce an element of possibility in the circumstantial semantic frame of these uses. Hence, just as *the opportunity to* implies the '[...] situation or circumstances where it is possible to do something', and is therefore coded as Modal Dynamic Possible Neutral 'Opportunity' (Schmid 2000: 254–8), so

is *the easiest way to* or *the only way of* amenable to classification into the same category. This is due to their semantic equivalence to *the only/easiest way in which it is possible to*.

A distinction needs to be made here between examples like (636), where 'Opportunity' features as the primary sense, and (634) and (635), where it is subsidiary to the 'Manner' sense. The latter is explained on the grounds of the strong connection between N-*be-to* and the Circumstantial 'Way' family (Schmid 2000: 286). Thus, in (636), there is little doubt that the prevailing meaning is that of *chance* or *opportunity*, as evident in the acceptability of such paraphrase as *the visit offers China's leadership a face saving opportunity/chance to patch up its rift with Washington* and the oddity of *the visit offers China's leadership a manner/way to patch up its rift*. The latter paraphrase shows how the visit is not the manner of patching up the rift with Washington, but the opportunity which will hopefully lead to such a solution. Example (636) is thus coded as Modal Dynamic Possible Neutral 'Opportunity'. Another lemma that, according to Schmid (2000: 254), may also occur in Modal Dynamic 'Opportunity' uses is *time*, as in (637), where a possible paraphrase might be *I've got so much to do and little/almost no chance/opportunity to do it*.

(636) Admiral Scowcroft's visit at the head of a senior delegation including the Deputy Secretary of State, Mr Lawrence Eagleburger, seemed to offer China's leadership a face saving way to patch up its rift with Washington, with whom China's relations have been at their lowest ebb since Mr Nixon's trail-blazing 1972 trip to Beijing. (*BNC Sampler*: A9M, W:newsp:brdsht_nat:report)

(637) So we're gonna catch every traffic light on red now <pause> there was something I wanted to go into Wisbech for but I can't think for the life of me what it was, I must get stuck into this Christmas shopping again, really must <pause> so much to do and such little time to do it in <pause> oh he's, he's put that car down five hundred pounds, that price of that car was, that he got for sale on it <-|->seven, seven, nine, five [...] (*BNC Sampler*: KC8, S:conv)

In (634) and (635), it is the 'manner' meaning that lies at the core of their interpretation, whilst 'opportunity' is only entailed in the postmodifying *to*-clause and *of*-phrase respectively. Hence, the focus is not on the opportunity to do something, but on how contacting DHSS or getting debt may be best achieved. In other words, the chance or opportunity to get debt or contact DHSS is best achieved by using a book or finding a telephone. In order to

account for the primary 'manner' sense of these instances, and following Schmid (2000: 286), use is made of the Specific Circumstantial Manner 'Way' family. However, in this thesis, the additional Modal nuances are indicated between square brackets (e.g. Specific Circumstantial Manner 'Way' [Modal Dynamic Possible Neutral]).

Lastly, mention should be made of (638), where *way* does not imply 'manner' or 'opportunity', but *Elias Hrawi's attempt/effort/move to unite and bring peace to Lebanon*. In other words, General Michel Aoun is making things difficult for Elias Hrawi, who is trying hard to unite and bring peace to Lebanon. This interpretation is suggested by a search of *{block/V} * _{N} to* in *CQPweb*, returning such instances as *blocked efforts to*, *blocked moves to*, *block any moves to*, *blocked an attempt to* and *block our attempts to.* On these grounds, (638) is coded as belonging to Schmid's (2000: 266–8) family of Specific Eventive Purposive 'Attempt' nouns (e.g. *attempt*, *effort*, *campaign*, *move*, *struggle*, *initiative*, *fight*). This example is particularly interesting in that it reveals a metaphorical extension from a concrete event (blocking someone's way or path) to an abstract domain, where it acquires the meaning of *attempt*, *move* or *effort*.

(638) By Reuter in Beirut PRESIDENT Elias Hrawi has threatened to resign if he fails to topple General Michel Aoun, whom he accused of blocking his way to unite and bring peace to Lebanon. (*BNC Sampler*: A9E, W:newsp:brdsht_nat:report

This section includes a table showing all the semantic tags used with the 60 lemmas in this thesis. Table 5.2 below consists of four columns listing the six main semantic classes of shell nouns, the families under each class, the lemmas carrying each sense, and an example sentence for each family. The number of instances for each class, family and lemma is between brackets. Overall, the evidence in the analysis database falls into 70 semantic families, 51 of these taken directly from Schmid's (2000) study and 19 created on the basis of existing categories (see the shaded cells). The three most frequent semantic classes are Factual, Mental and Linguistic, accounting for 67.79% or 981 of the 1447 shell-noun instances in the study sample. As regards the Lemma column, red marks units that do not occur in Schmid's (2000) shell-noun list (see 5.3.5.1), and green marks senses not recorded in his taxonomy (see 5.3.5.2). The remaining units correspond to senses recorded in his analysis.

Table 5.2 Semantic classification of the 60 lemmas in the study sample

Class	Family	Lemmas	Example
Factual (345)	Factual Attitudinal Descriptive 'Irony' (9)	*irony*(4), *joke*(3), *phenomenon*(2)	Erm, the irony of the sort of split responsibility is that er one of the things that we have to obviously bear in mind with any future in-- erm extension of residents' parking schemes is the cost
	Factual Attitudinal Evaluative Negative 'Tragedy' (8)	*failure*(4), *misfortune*(1), *scandal*(3)	I think it's a scandal what we get <pause> for a pension after all the years that you, you fight for you, you put away for your old age
	Factual Attitudinal Impeding 'Problem' (50)	*challenge*(10), *problem*(40)	Where the problem arose was when there was no record of anything having been sent or what had been sent or anything.
	Factual Attributive Part-whole 'Aspect' (74)	*characteristic*(29), *detail*(8), *dimension*(22), *facet*(7), *point*(8)	The hotel industry, the entertainment business, and gambling are all facets of tourist facilities which ought to be taken into consideration.
	Factual Causal Mental 'Evidence' (45)	*evidence*(13), *finding*(28), *testimony*(4)	However, the labour market has absorbed far more IT Advanced Courses students in 1984 than were even available to employers in 1983. This was a positive finding.
	Factual Causal 'Reason' (7)	*sense*(1), *impetus*(6)	This laser provided an early impetus for studies of instabilities by tending to produce noisy, spiked output even under quasi-steady excitation.
	Factual Comparative 'Difference' (14)	*contradiction*(4), *opposite*(10)	The Americans are clearly hoping that it will be possible to bring the Palestinians and Israelis around the negotiating table without confronting the contradictions and achieve an historic breakthrough
	Factual General Neutral 'Thing' (60)	*phenomenon*(34), *point*(3), *thing*(23)	The best thing about Christmas is buying well thought out gifts that will give pleasure and use for many years to come
	Factual General Neutral 'Thing' [Factual Attitudinal Impeding 'Problem'] (6)	*point*(3), *thing*(3)	<-|-> the thing <-|-> is in Scotland though, there's five good teams and the rest are rubbish
	Factual General Neutral 'Thing' [Modal Deontic Necessary] (1)	*thing*(1)	The important thing is for people to listen to the genuine concerns, sort out the, the real issues from the power politics and then hopefully go forward in, in a, in a manner of conciliation.
	Factual Partitive Class-Member 'Example' (40)	*example*(40)	The workshop of a garage is a typical example of a service trade.

	Factual Partitive Part-whole 'Part' (31)	*part*(31)	In art history a large part of electronic data processing takes the form of cataloguing works of art […]
Mental (336)	Mental Conceptual 'Idea' (19)	*myth*(7), *point*(3), *vision*(9)	Have Post-war changes in educational policy and provision brought Britain closer to realising the vision of the meritocratic society?
	Mental Creditive Attitudinal 'View' (35)	*philosophy*(18), *prejudice*(11), *vision*(6)	Journalists with no axe to grind, no message to put across, no prejudice to bear, positively seeking after the truth.
	Mental Creditive 'Belief' (32)	*experience*(7), *suspicion*(25)	We have a suspicion that very little reaction has taken place because in fact very few bubbles have risen to the surface.
	Mental Creditive 'Belief-Feeling' (7)	*sense* (7)	[…] we tried to erm give them er a sense of their worth as a women […]
	Mental Creditive 'Belief-Understanding' (13)	*sense*(13)	It is just different from what it would be had Gloriana been intended as a Grand Opera in the Aida sense.
	Mental Creditive Resultative 'Realisation' (31)	*assessment*(26), *recollection*(5)	They were considered, both of them considered, both of them considered erm but it is my recollection that erm been the inference that Mr had made the telephone calls himself.
	Mental Emotive Event-Related 'Fear' (7)	*foreboding*(7)	She was assailed by apprehension; by a sense of foreboding, of doom, of indescribable evil…
	Mental Emotive Event-Related 'Surprise' (17)	*anger*(3), *surprise*(10), *terror*(4)	Peter Scudamore's late decision to partner Granville Again was no surprise.
	Mental Emotive Fact-Related 'Surprise' (19)	*anger*(2), *surprise*(15), *terror*(2)	When you know what we should be saying to John Major and employers is that it isn't any surprise that er disablement er has increased over the last thirteen, fourteen years.
	Mental Event-Related 'Experience' (35)	*experience*(33), *time*(2)	If er you 've had dreadful experiences with er guests staying at your place, give me a ring on and share your horrid experiences with me.
	Mental Perceptive Visual (9)	*vision*(9)	[…] special places where the mind can expand into new levels of consciousness, places where visions, transcendental states of prophecy, may be experienced […]
	Mental Volitional Causal 'Motivation' (5)	*motivation*(5)	The 'motivation' for the session we watched was a backcloth of a ski slope.
	Mental Volitional Conclusive 'Aim' (45)	*objective*(40), *vision*(5)	In arriving at the correct mix of full-time to part-time staff the objective should be to determine the best fit between the staff required and the staff available.
	Mental Volitional Conditional 'Determination' (4)	*sense*(1), *motivation*(3)	This is coupled with a team of people with extensive international marketing and commercial acumen and the drive and motivation to continue our

			commitment to international success.
	Mental Volitional Detached 'Purpose' (11)	*application*(9), *point*(2)	The second possible practical application of algebraic laws is for transforming programs to improve their efficiency in some way.
	Mental Volitional Manner 'Plan' (44)	*philosophy*(4), *project*(40)	They raise money for village development projects, clean public buildings and help with the catering for conferences and state functions […]
	Mental Volitional Resultative 'Solution' (3)	*answer*(3)	Automatic fire fighting installations capable of achieving total control and extinguishment appear to be the only answer.
Linguistic (300)	Linguistic Illocutionary Assertive Argumentative 'Amendment' (2)	*correction*(2)	However, it may be helpful if the system can sometimes suggest a correction for a miskeyed word.
	Linguistic Illocutionary Assertive Argumentative Defending 'Endorsement' (4)	*endorsement*(4)	A leader of Socialist Left, a radical pressure group, said the confirmation in office of the old team was 'a serious mistake, because it signals the endorsement of an economic policy criticised by a lot of people'.
	Linguistic Illocutionary Assertive Argumentative Reactive 'Answer' (37)	*answer*(37)	Not much, is the short answer.
	Linguistic Illocutionary Assertive Difference 'Contradiction' (4)	*contradiction*(4)	She pleaded that she did not do it and her husband's inconsistencies and contradictions aroused suspicion.
	Linguistic Illocutionary Assertive Interactive 'Statement' (2)	*word*(2)	Do you wanna have a wee word with granny?
	Linguistic Illocutionary Assertive Public 'Proclamation' (3)	*testimony*(3)	Key testimony has been broadcast live on television and radio to help keep the city fully abreast of the proceedings and to show the wheels of justice turning.
	Linguistic Illocutionary Assertive Retrodictive 'Report' (2)	*word*(2)	And she didn't, she scarcely even looked at the label. She took my word for it.
	Linguistic Illocutionary Commissive 'Threat' (29)	*warning*(29)	Recent warnings from politicians of all parties that the influx can not continue indefinitely are therefore seen as being based primarily on fears […]
	Linguistic Illocutionary Directive 'Advice' (32)	*recommendation*(32)	I welcome the recommendation in this report that income generating programmes for children need consideration <pause> but with great care in their design.
	Linguistic Illocutionary Directive 'Invitation' (31)	*application*(31)	Had the Borough Council objected to the application for 60 housing units, then Pubmaster might have appealed against their decision […]

	Linguistic Metalinguistic Textual Deixis 'Text' (3)	*word*(3)	You feel that the music carries the words along on the back of its urge to secure release.
	Linguistic Metalinguistic Textual Deixis 'Word' (19)	*word*(19)	Tenuous, was the word that came to mind.
	Linguistic Propositional Metalinguistic 'Adage' (34)	*irony*(3), *joke*(31)	You see, by starting with it you emphasise the wholly hopeful, optimistic expectations the words imply and then, by ending with it, you can stress the double-edged sword quality: the irony.
	Linguistic Propositional Metalinguistic 'Myth' (23)	*myth*(23)	Plutarch clearly presents a late form of the myth with some Greek influences by providing a very useful story outline.
	Linguistic Propositional 'News' (75)	*detail*(32), *evidence*(7), *point*(18), *thing*(4), *word*(14)	Er, so I just make that point because of the debate last week.
	Eventive Attitudinal Manner Deontic 'Trouble' (8)	*challenge*(8)	Overcoming a natural resistance to change is a challenge faced by many companies that want to progress.
	Eventive Attitudinal Negative 'Mistake' (2)	*crime*(1), *failure*(1)	But the clear intelligence failure which let Gen Noriega escape, to organise what could prove to be a prolonged guerrilla resistance, threatens to turn this latest military adventure into a political embarrassment.
	Eventive Attitudinal Negative 'Tragedy' (19)	*failure*(3), *misfortune*(6), *scandal*(10)	Are mama and my sisters truly in sanctuary — or could it be that some grave misfortune has overtaken them also?
	Eventive Attitudinal Positive 'Success' (8)	*triumph*(8)	Yet once again there was a medical triumph. Dr Laurent came in late one cold night with the results of lung resection and a complete cure from seemingly hopeless tuberculosis.
Eventive (188)	General Eventive Neutral 'Event' (11)	*point*(1), *practice*(1), *thing*(9)	[…] the first thing she does, quite unsurprising, is screams […]
	General Eventive Neutral 'Event-Effort' (26)	*work*(26)	[…] they enjoyed showing the 'greats' who had little rooms adjoining each other on that top floor, the good work that had been done in 'the acute exanthemata.'
	General Eventive Neutral Negative Polarity 'Failure' (3)	*failure*(3)	Failure to persuade medical staff of the adverse consequences on patient care should be reported to the next level of management.
	Specific Eventive Habitual Temporal Manner 'Tradition' (6)	*practice*(6)	'Out of sight, out of mind, is the practice with youths of his age,' said the voice of experience.

	Specific Eventive Negative 'Crime' (38)	*crime*(38)	They had committed the 'counter-revolutionary crime' of celebrating the Dalai Lama 's Nobel Peace Prize.
	Specific Eventive Purposive 'Attempt' (67)	*challenge*(8), *endeavour*(9), *joke*(6), *practice*(7), *quest*(5), *terror*(2), *venture*(29), *way*(1)	But I'm quite sure his earlier associations within the family, er you know helped him in er in his endeavour to become a county councillor after the er after the second world war.
Circumstantial (186)	General Circumstantial 'Situation' (1)	*dimension*(1)	[…] the majority of voters in Northern Ireland will not have an institutionalised Irish dimension […]
	Specific Abstract Circumstantial Locative 'Place' (Eventive domain) (16)	*area*(16)	That is, you recognize, do you not, that greenbelt protection includes the concept of the initial negative presumption that is not contained in any other area of planning <unclear>?
	Specific Abstract Circumstantial Locative 'Place' (Mental domain) (24)	*area*(24)	I mean even the simple things, the hard data like annual mileage, is clearly an area of debate […]
	Specific Circumstantial Conditional 'Condition' (7)	*proviso*(7)	The objective of the operation was to arrest Mr with the proviso that it was to be done with the safety of the public to be paramount and that the safety of the police officers and Mr was also to be considered.
	Specific Circumstantial Locative 'Place' (1)	*point*(1)	Erm <pause> but there comes a point where th-- , the government, the line that the government draws for means testing's always <pause> very low.
	Specific Circumstantial Manner 'Way' (95)	*dimension*(1), *practice*(26), *system*(34), *way*(34)	In continental Europe political parties in power have been known to institute electoral systems calculated to give them unfair advantage.
	Specific Circumstantial Manner 'Way'[Modal Dynamic Possible Neutral] (4)	*way*(4)	Several, for instance, are deciduous, and the only way to have berries at Christmas would be to emigrate to Chile.
	Specific Circumstantial Temporal 'Time' (38)	*point*(1), *time*(37)	Although I have to confess that at present, during a time of recession, it is increasingly difficult to attract the backing of galleries for such shows […]
Modal (91)	Modal Deontic Possible 'Permission' (4)	*leave*(4)	[…] on an appeal to the High Court from a decision of justices under the Children Act 1989 fresh evidence could be adduced only with leave in exceptional circumstances […]
	Modal Deontic Probable 'Job' (14)	*capacity*(4), *challenge*(8), *part*(1),	The appropriate body within a member state (e.g. DTI and/or OFT within UK) is then consulted and assists in an advisory capacity [...]

	work(1)	
Modal Dynamic (-)Possible Subject-Oriented 'Ability' (12)	*failure*(12)	Then loneliness and **their failure** to appreciate our culinary efforts must have depressed the visitors [...]
Modal Dynamic (+)Possible Subject-Oriented 'Ability' (19)	*capacity*(19)	One may be less sanguine about **the average voter's capacity** to rank candidates, whether of one preferred party or more than one, according to their relative merit and aptitude.
Modal Dynamic Possible Neutral 'Opportunity' (18)	*chance*(16), *time*(1), *way*(1)	Summit offers **chance** to improve British image in EC
Modal Epistemic Possible 'Possibility' (20)	*chance*(20)	Stu , I mean I live in Glasgow , **the chances** of me leaving Glasgow are pro , pre pretty slim but yeah [...]
Modal Epistemic Possible 'Possibility-Risk' (2)	*chance*(2)	Probably took **a chance with it**.
Modal Epistemic Probable 'Probability' (2)	*chance*(2)	**The chances** are we won't be going!

5.4 CONCLUSION

Chapter 5 describes the analytical process of this study, particularly in relation to the coding of shell-noun instances. Whilst the nine variables under scrutiny draw on existing models and categorisations, the corpus-theoretical and manual approach followed in this thesis (see 4.4.1) required modifications to tailor existing frameworks to the nature of the data at hand. The following summarises some of the most relevant decisions in this respect:

i) Halliday & Matthiessen's (2004) post-Deictic categories are enhanced with Quirk et al.'s (1985) emphasising, amplifying and restrictive adjectives.
ii) N-*of* instances occur in a cline that ranges from partitive expressions, where N_2 is head, to non-partitive ones, with N_1 as head (see 5.3.2.3).
iii) Unlike Halliday & Matthiessen (2004), where Sayers and Experiencers are applicable to conscious and non-conscious subjects, shell nouns are coded as Sayers or Experiencers in this thesis only if the subject is human. Otherwise, if the subject is performed by a second- or third-order abstract entity, the shell noun is a relational participant, in line with Martin et al. (1997).
iv) Endophoric encapsulation covers specific and non-specific Deictics. Context is a major criterion in the identification of shell-noun specifics.
v) Encapsulation is uni- and bidirectional. The latter applies not only to shell nouns preceded by post-Deictics (e.g. *similar*, *different*, *other*), but also to instances where anaphoric specifics on their own are deemed insufficient for a complete understanding of the shell noun.
vi) Gray's (2010) antecedent taxonomy is enriched with 12 new categories and 2 variants (pronoun, numeral, adjective, verb, prepositional phrase, interjection, reaction signal, formula, figure, table, complement and premodifier; local extended preceding and quotation).
vii) Specifics of identity occur inside the noun phrase (as in Winter 1992) and elsewhere.

As in 4.5, it is worth emphasising once again the importance of a sense-oriented approach to the analysis conducted in this thesis. Only the close reading of extended discourse segments, aided by dictionary and corpus evidence, can enhance the accuracy of findings. This is particularly evident in the semantic categorisation of shell instances, where extended reading,

dictionary and corpus evidence allow the identification of numerous shell senses unaccounted for in Schmid's (2000) automated study.

Following the methodological and qualitative description offered in chapters 4 and 5, the next chapter explores how methodological and qualitative decisions translate into quantitative or numerical findings.

6 RESULTS

6.1 INTRODUCTION

This chapter presents the quantitative findings for the nine variables under scrutiny. It comprises two major sections, 6.2 and 6.3. The former constitutes the core of this chapter, as it brings together the results for each variable under four main subsections. Section 6.2.1 is the most general, as it deals with the genre distribution of shell units. Sections 6.2.2 and 6.2.3 shift the focus from text type to the form and behaviour of shell-noun phrases at clause level. The former looks at the experiential and formal structure of shell-noun phrases, whilst the latter is concerned with their syntactico-semantic functions and textual positions (syntactic function, participant type and Theme-Rheme). Section 6.2.4 moves from phrase and clause levels to the wider discourse context, and delves into encapsulating relations, antecedents, and the relationship between both.

Each of these sections is in turn subdivided into four subsections. The first subsection in each case presents the general findings for each variable, whilst the following two narrow the scope to the results for mode and genre and, subsequently, to shell-noun distribution across lemmas and semantic types. The last of these subsections discusses findings in relation to the literature. Prior to this final subsection, 6.2.3 and 6.2.4 contain an additional subsection on the connection with some of the variables in 6.2.1 and 6.2.2 (6.2.3.4 and 6.2.4.4). Following the detailed presentation of findings in 6.2, 6.3 brings this chapter to a close by synthesising results into eight points.

All data in this chapter are given in percentages. Line graphs are chosen for the visual presentation of results where the number of categories does not impede visibility. Scales are set at 50% if lines rank below 50%, and at 100%, if they rank above 50% in order to enhance visibility where lines cluster around low frequency ranges. If categories are too numerous, data are presented in tabular format. The length of some of the lists forces at times the inclusion of only the top 50% categories in the main text. In such cases, the complete lists are given in the Appendix.

6.2 RESULTS AND DISCUSSION

6.2.1 *Genre of the text*

6.2.1.1 *Overall distribution*

The study sample shows a higher percentage of shell-noun use in written genres: 64.55% (934 out of the total 1447 concordances analysed) occur in written genres, while 35.45% (513 concordances) occur in spoken genres.

In Table 6.1, the study sample is shown to contain shell-noun instances in 38 out of Lee's (2001) 70 genre categories and in 24 out of the 34 super-genre categories. The top ten genre and super-genre categories (shaded in grey) account for 64.27% and 83.83% of the evidence.

Table 6.1 Distribution of the evidence in percentages across genres and super-genres

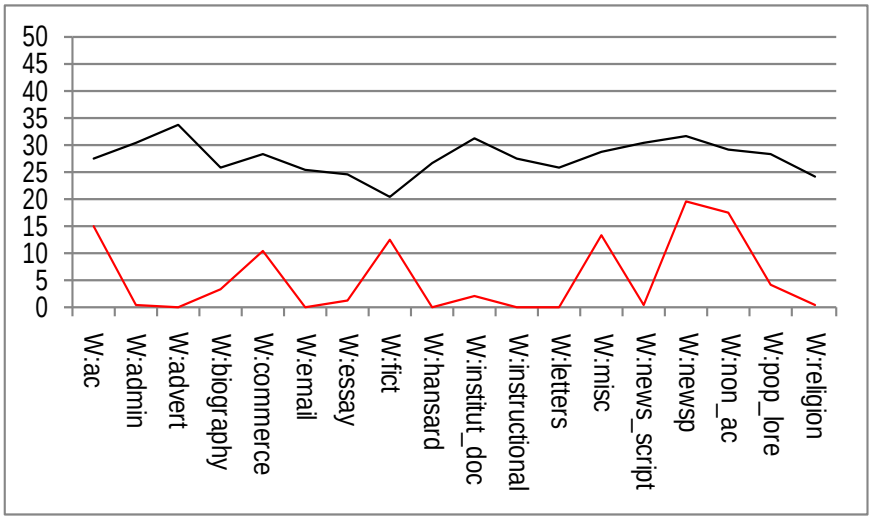
Genres	%	Super-genres	%
S:meeting	9.95	W:newsp	12.72
W:misc	8.64	W:non_ac	11.33
W:newsp:brdsht_nat:report	8.36	S:meeting	9.95
S:conv	6.91	W:ac	9.68
W:commerce	6.77	W:misc	8.64
W:fict:prose	6.63	W:fict	8.15
W:non_ac:humanities_arts	6.29	S:speech	6.98
S:speech:unscripted	3.94	S:conv	6.91
W:non_ac:polit_law_edu	3.73	W:commerce	6.77
S:speech:scripted	3.04	S:brdcast	2.70
W:ac:humanities_arts	2.90	W:pop_lore	2.56
W:ac:nat_science	2.63	S:interview	2.42
W:pop_lore	2.56	W:biography	2.00
S:interview:oral_history	2.42	S:classroom	1.80
W:ac:soc_science	2.07	S:courtroom	1.59
S:brdcast:discussn	2.00	S:pub_debate	1.59
W:biography	2.00	S:unclassified	1.31
W:newsp:other:sports	1.94	W:institut_doc	1.31
S:classroom	1.80	W:essay:univ	0.76
S:courtroom	1.59	S:sermon	0.21
S:pub_debate	1.59	W:admin	0.21
W:newsp:other:commerce	1.38	W:news_script	0.21
S:unclassified	1.31	W:religion	0.14
W:ac:tech_engin	1.31	W:advert	0.07
W:institut_doc	1.31		
W:fict:drama	1.24		
W:newsp:other:report	1.04		
W:non_ac:soc_science	1.04		
W:ac:polit_law_edu	0.76		
W:essay:univ	0.76		

S:brdcast:documentary	0.69
W:fict:poetry	0.28
W:non_ac:tech_engin	0.28
S:sermon	0.21
W:admin	0.21
W:news_script	0.21
W:religion	0.14
W:advert	0.07

Six of the top ten categories in each case fall into the written mode (40.43% for genres and 57.29% for super-genres), while only four belong to the spoken mode (23.84% for genres and 26.54% for super-genres).

With the above results in mind, the proportion of shell-noun use was calculated as percentages based on the amount of evidence for each mode (934 concordances, written; 513 concordances, spoken) and the resulting percentages were compared with genre-related noun use in *BNCweb*[28]. Percentages for the latter were obtained after dividing the number of nouns (_{N}) in *BNCweb* by the number of words for each super-genre. Figures 6.1 and 6.2 below present the distribution of shell nouns (study sample) and nouns (*BNC*) across written and spoken super-genres.

Figure 6.1 Distribution of nouns (*BNC*, black) and shell nouns (study sample, red) in written super-genres



[28] The *BNC Sampler* (*CQPweb*) was not used for comparison, because data there are given only for the mode-related distribution of corpus searches (i.e. written vs. spoken; see 4.2.1).

Figure 6.2 Distribution of nouns (*BNC*, black) and shell nouns (study sample, red) in spoken super-genres

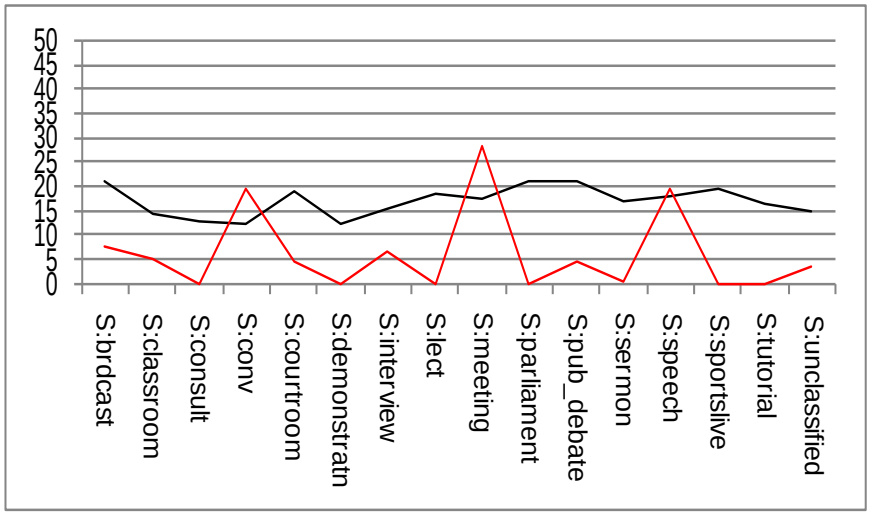


Figure 6.1 shows that, whilst the study sample does not overlap with the *BNC*, the patterns of both corpora differ. This is shown in the greater prevalence of certain super-genres in the study sample. W:fict is especially prominent in this respect, as, according to *BNCweb*, noun use is fairly rare in this written super-genre (hence, the falling line in the *BNC* and the rising one in the study sample). Several other super-genres are also particularly frequent in the study sample: W:ac, W:commerce, W:misc, W:newsp and W:non_ac. Figure 6.2 is more revealing than Figure 6.1 in this respect, as three super-genres (S:conv, S:meeting and S:speech) peak slightly above the line for general noun use in the *BNC*.

Following the above comparison, the genre-related proportions of shell-noun use in the study sample were recalculated on the basis of the total number of words for each super-genre in the *BNC Sampler* (see 4.2.2 for how genre data were extracted from the *Sampler*). The aim was to assess the randomised shell-noun evidence from *CQPweb* in terms of the overall size of each super-genre category. Tables 6.2 and 6.3 below present the *Sampler*-normalised data (second column) alongside the mode-related results from the study sample (first column). The choice of a tabular format is explained by the markedly low figures in the *BNC Sampler* column. Graphical representation of these data (as in Figures 6.1 and 6.2) would have shown a hardly visible line for the *Sampler*-normalised proportions.

Table 6.2 Mode-related distribution of shell-noun use in the study sample and *Sampler*-normalised frequencies (Written)

Study Sample	%	BNC Sampler	%
W:newsp	19.70	W:commerce	0.1474
W:non_ac	17.56	W:non_ac	0.1462
W:ac	14.99	W:ac	0.1307
W:misc	13.38	W:news_script	0.1033
W:fict	12.63	W:institut_doc	0.1032
W:commerce	10.49	W:misc	0.0913
W:pop_lore	3.96	W:newsp	0.0895
W:biography	3.10	W:essay	0.0757
W:institut_doc	2.03	W:pop_lore	0.0675
W:essay	1.18	W:biography	0.0581
W:admin	0.32	W:fict	0.0551
W:news_script	0.32	W:admin	0.0481
W:religion	0.21	W:religion	0.0278
W:advert	0.11	W:advert	0.0160

Table 6.3 Mode-related distribution of shell-noun use in the study sample and *Sampler*-normalised frequencies (Spoken)

Study Sample	%	BNC Sampler	%
S:meeting	28.07	S:pub_debate	0.1528
S:speech	19.69	S:speech	0.1185
S:conv	19.49	S:courtroom	0.0957
S:brdcast	7.60	S:brdcast	0.0904
S:interview	6.82	S:meeting	0.0841
S:classroom	5.07	S:sermon	0.0690
S:courtroom	4.48	S:unclassified	0.0584
S:pub_debate	4.48	S:interview	0.0536
S:unclassified	3.70	S:classroom	0.0465
S:sermon	0.58	S:conv	0.0202

The side-by-side presentation of results in both tables shows more differences than similarities in genre-related distribution. While certain super-genres share the same rank in both sets, most are ranked differently, making either a greater or smaller contribution to the *Sampler*-normalised data than that found in the study sample.

Thus, in Table 6.2, W:non_ac, W:ac, W:religion and W:advert rank equal in both data sets. Conversely, 4 written super-genres feature more prominently in the *Sampler*-normalised data: W:commerce, W:news_script, W:institut_doc and W:essay. This is especially noticeable in relation to W:commerce and W:news_script. There are also 6 categories where the contribution is more prominent in the study sample and less so in the *Sampler*-normalised data: W:newsp, W:misc, W:fict, W:pop_lore, W:biography and W:admin. This is most noticeable with W:newsp and W:fict.

In Table 6.3, 2 categories (S:speech and S:brdcast) rank equal in both data sets, but 4 rank higher in the *Sampler*-normalised data: S:courtroom, S:pub_debate, S:unclassified and S:sermon. This is particularly evident with S:courtroom and S:pub_debate. A smaller contribution is apparent with S:meeting and S:conv, and less markedly with S:interview and S:classroom.

6.2.1.2 *Distribution across lemmas*

The foregoing discussion brings to light a number of variations in the genre-related distribution of shell-noun use. With this mind, it is now worth looking at the units involved in such use for any lexically-driven explanations for the results of Table 6.1. In the following, the focus shifts from the mode-related distribution of the 60 sampled lemmas (Tables 6.4, 6.6a and 6.6b) to their genre-specific frequencies (Tables 6.7a through 6.7b).

Table 6.4 Mode-related distribution of the 60 sampled lemmas. Percentages are based on the written and spoken components of the study sample (513 and 934 concordances)

Written	%	Spoken	%
Phenomenon	3.75	Point	5.85
Crime	3.21	Thing	5.65
Detail	3.21	Time	5.26
Objective	3.10	Answer	4.29
Myth	3.00	Area	4.09
Example	2.89	Chance	4.09
Challenge	2.68	Experience	4.09
Practice	2.68	Problem	3.70
Project	2.68	Application	3.51
Venture	2.68	Word	3.51
Way	2.68	Joke	3.12
Joke	2.57	Part	2.92
Dimension	2.46	Practice	2.92
Warning	2.46	Project	2.92
Application	2.36	Recommendation	2.92
Characteristic	2.36	Way	2.92
Finding	2.36	Example	2.53
Word	2.36	System	2.53
Problem	2.25	Vision	2.34
System	2.25	Objective	2.14
Assessment	2.14	Crime	1.95
Failure	2.14	Detail	1.95
Area	2.03	Work	1.95
Chance	2.03	Sense	1.75
Experience	2.03	Challenge	1.75
Surprise	2.03	Evidence	1.56
Answer	1.93	Philosophy	1.56
Capacity	1.93	Suspicion	1.56

Part	1.82	Characteristic	1.36
Recommendation	1.82	Assessment	1.17
Suspicion	1.82	Finding	1.17
Vision	1.82	Surprise	1.17
Work	1.82	Warning	1.17
Philosophy	1.50	Capacity	0.97
Sense	1.39	Endeavour	0.78
Time	1.39	Motivation	0.78
Evidence	1.28	Recollection	0.78
Thing	1.18	Scandal	0.78
Point	1.07	Venture	0.78
Prejudice	1.07	Failure	0.58
Opposite	0.96	Irony	0.58
Scandal	0.96	Myth	0.39
Triumph	0.86	Contradiction	0.19
Contradiction	0.75	Dimension	0.19
Foreboding	0.75	Endorsement	0.19
Terror	0.75	Facet	0.19
Facet	0.64	Misfortune	0.19
Impetus	0.64	Opposite	0.19
Misfortune	0.64	Phenomenon	0.19
Proviso	0.64	Prejudice	0.19
Testimony	0.64	Proviso	0.19
Anger	0.54	Terror	0.19
Endeavour	0.54	Testimony	0.19
Quest	0.54		
Irony	0.43		
Leave	0.43		
Motivation	0.43		
Endorsement	0.32		
Correction	0.21		
Recollection	0.11		

Table 6.4 shows that the written component of the sample contains a slightly wider lexical range than the spoken one, where 7 of the 60 lemmas are unaccounted for. The top ten lemmas in each case reveal a greater concentration of shell-noun use in the spoken section (44.05% spoken vs. 29.87% written). Thus, shell-noun use in written English appears to draw on a variety of types spread evenly across the study sample. Conversely, spoken English uses slightly fewer types, the highest-ranking ones being repeated more frequently.

When compared to the lemmas in the written list, it emerges that whilst all 60 units are similarly unspecific in their reference, high-ranking written nouns seem to be semantically more specialised than their counterparts in the spoken list (e.g. *crime*, *myth*, *project*, *venture* vs. *point*, *thing*, *time*, *area*). Table 6.5 below shows that semantic specificity is related to frequency. This

table shows the token frequency with which the top ten units in each mode occur in the *BNC Sampler*.

Table 6.5 Token frequencies for the top ten units in each mode (*BNC Sampler*). The shaded cells represent the ten highest-ranking units in the written mode

Lemmas (Token frequency)			
Time	3746	Practice	254
Thing	2616	Experience	248
Problem	1132	Answer	238
Area	1117	Example	177
Point	1044	Crime	74
Word	934	Objective	73
Project	263	Challenge	70
Chance	260	Phenomenon	45
Application	256	Venture	42
Detail	256	Myth	40

Except for *project*, *detail* and *practice*, the remaining seven written lemmas occur at the bottom of the list. This indicates that, while high-ranking units in the spoken section are also highly frequent in the corpus (especially *time*, *thing*, *problem*, *area*, *point* and *word*, with token frequencies ranging from 3746 to 934), high-ranking written units take lower frequency ranges (from 263 to 40).

Before moving on to the genre-specific data in Tables 6.7a through 6.7b, it is worth establishing whether, according to the *BNC Sampler*, the top ten units in each mode also show dissimilar values in spoken and written English. Tables 6.6a and 6.6b present the figures for the top ten written and spoken units. Each table compares the attested frequencies of use based on the amount of written and spoken evidence in the study sample and in the *BNC Sampler* (*CQPweb*). Figures are expressed differently in each case, i.e. percentages in the former and frequencies per million words (henceforth, fpmw) in the latter. *CQPweb* uses fpmw on account of the greater amount of evidence in the *BNC Sampler*, i.e. 1 million words for spoken and written English (cf. however, the evidence from the study sample: 934 written and 513 spoken).

Table 6.6a Mode-related frequencies for the top ten written units in the study sample: Study sample (%) vs. *BNC Sampler* (fpmw). The shaded cells in Tables 6.6a and 6.6b correspond to units for which the mode-related priming in the study sample does not match the distribution observed in the *BNC Sampler*

	Top ten units (Written)		Top ten units (Spoken)	
	Study sample	BNC Sampler	Study sample	BNC Sampler
Phenomenon	3.75	37.80	0.19	0.88
Crime	3.21	38.66	1.95	25.44
Detail	3.21	150.33	1.95	71.05
Objective	3.10	38.66	2.14	24.56
Myth	3.00	32.64	0.39	1.75
Example	2.89	93.65	2.53	59.65
Challenge	2.68	45.53	1.75	14.91
Practice	2.68	138.31	2.92	81.58
Project	2.68	128.86	2.92	99.12
Venture	2.68	31.79	0.78	4.39
Way	2.68	740.5	2.92	1004.34

Table 6.6b Mode-related frequencies for the top ten spoken units in the study sample: Study sample (%) vs. *BNC Sampler* (fpmw)

	Top ten units (Written)		Top ten units (Spoken)	
	Study sample	BNC Sampler	Study sample	BNC Sampler
Point	1.07	362.52	5.85	549.59
Thing	1.18	420.08	5.65	1865.7
Time	1.39	1410.57	5.26	1845.52
Answer	1.93	90.2	4.29	116.66
Area	2.03	613.36	4.09	353.49
Chance	2.03	143.46	4.09	81.58
Experience	2.03	119.41	4.09	95.61
Problem	2.25	538.63	3.70	442.96
Application	2.36	101.37	3.51	121.05
Word	2.36	347.06	3.51	464.89

Table 6.6a reveals that the study sample's preference for the written use of these lemmas agrees with their overall distribution in the *BNC Sampler*, where they are also primed for written use. The only exception is *way*, occurring more frequently in the spoken mode. Table 6.6b shows four lemmas whose priming in the study sample does not match the attested preference for written English in the *BNC Sampler*: *area*, *chance*, *experience*, *problem*. Of the four lemmas, *area* is most obviously primed for written use in the *Sampler*, with a difference of 260 fpm between written and

spoken English (cf., by contrast, *chance*, *experience* and *problem*, with differences ranging from 62 to 96 fpm).

Following the mode-related data above, Tables 6.7a through 6.7b bring together all the lemmas in the ten most frequent super-genres in the study sample (see Table 6.1). Although shell-noun use occurs in 24 super-genres, as mentioned above, the first ten categories are highly representative of the overall evidence (i.e. 1447 concordances), accounting for 83.83% of the total. i.e.

Table 6.7a The 60 lemmas and their distribution across the top ten super-genres. Percentages in Tables 6.7a and 6.7b are obtained on the basis of the number of shell-noun instances within each super-genre[29]

W:newsp	%	W:non_ac	%	S:meeting	%	W:ac	%	W:misc	%
Warning	9.78	Myth	14.63	Recommendation	7.64	Phenomenon	17.86	Joke	10.40
Crime	5.98	Example	4.27	Point	6.94	Characteristic	5.71	Challenge	8.00
Challenge	4.35	Assessment	4.27	Project	6.25	System	5.00	Project	7.20
Chance	4.35	Finding	3.66	Answer	5.56	Dimension	5.00	Objective	6.40
Problem	3.80	Experience	3.66	Problem	4.86	Application	5.00	Characteristic	5.60
Objective	3.80	Detail	3.66	Example	4.86	Finding	4.29	Venture	5.60
Application	3.26	Characteristic	3.05	Part	4.86	Example	4.29	Finding	4.80
Scandal	3.26	Prejudice	3.05	Application	4.86	Project	4.29	Assessment	4.80
Suspicion	2.72	Part	3.05	Work	4.86	Way	4.29	Recommendation	4.00
Project	2.72	Capacity	3.05	Area	4.86	Part	3.57	Problem	3.20
Failure	2.72	Practice	3.05	Challenge	4.86	Area	3.57	Area	3.20
Surprise	2.72	Word	3.05	System	3.47	Evidence	3.57	Failure	3.20
Dimension	2.17	Work	2.44	Thing	3.47	Work	2.86	Opposite	2.40
Contradiction	2.17	System	2.44	Word	3.47	Impetus	2.14	Point	2.40
Venture	2.17	Philosophy	1.83	Detail	2.78	Detail	2.14	Detail	2.40
Terror	2.17	Surprise	1.83	Experience	2.08	Philosophy	2.14	Vision	2.40
Vision	2.17	Area	1.83	Objective	2.08	Objective	2.14	Practice	2.40
Way	2.17	Way	1.83	Crime	2.08	Sense	2.14	Example	1.60
Triumph	1.63	Time	1.83	Irony	1.39	Practice	2.14	Correction	1.60
Finding	1.63	Crime	1.83	Finding	1.39	Problem	1.43	Suspicion	1.60
Example	1.63	Dimension	1.83	Surprise	1.39	Answer	1.43	Way	1.60
Work	1.63	Evidence	1.83	Practice	1.39	Leave	1.43	Time	1.60
Detail	1.63	Testimony	1.83	Way	1.39	Crime	1.43	Crime	1.60
Assessment	1.63	Phenomenon	1.83	Vision	1.39	Endeavour	1.43	Experience	1.60
Capacity	1.63	Challenge	1.83	Chance	1.39	Failure	1.43	Motivation	1.60
Time	1.63	Failure	1.83	Evidence	0.69	Misfortune	0.71	Dimension	0.80
Anger	1.63	Vision	1.83	Testimony	0.69	Facet	0.71	Evidence	0.80
Area	1.63	Scandal	1.22	Endorsement	0.69	Opposite	0.71	Phenomenon	0.80
Practice	1.63	Answer	1.22	Contradiction	0.69	Word	0.71	Work	0.80
Impetus	1.09	Contradiction	1.22	Warning	0.69	Point	0.71	Application	0.80
Opposite	1.09	Suspicion	1.22	Joke	0.69	Vision	0.71	Myth	0.80

[29] 184 in W:newsp, 164 in W:non_ac, 144 in S:meeting, 140 in W:ac, 125 in W:misc, 118 in W:fict, 101 in S:speech, 100 in S:conv, 98 in W:commerce and 39 in S:brdcast.

Thing	1.09	Chance	1.22	Philosophy	0.69	Prejudice	0.71	Philosophy	0.80
Part	1.09	Quest	1.22	Suspicion	0.69	Assessment	0.71	Prejudice	0.80
Endorsement	1.09	Venture	1.22	Sense	0.69	Motivation	0.71	Surprise	0.80
Point	1.09	Joke	1.22	Assessment	0.69	Capacity	0.71	Chance	0.80
Sense	1.09	Sense	1.22	Recollection	0.69	Chance	0.71	System	0.80
Endeavour	1.09	Misfortune	0.61	Capacity	0.69	Proviso	0.71		
Joke	1.09	Triumph	0.61	Time	0.69	Venture	0.71		
Quest	1.09	Facet	0.61	Endeavour	0.69				
Testimony	1.09	Point	0.61	Venture	0.69				
Answer	1.09	Application	0.61						
Word	1.09	Irony	0.61						
Characteristic	0.54	Recollection	0.61						
Facet	0.54	Anger	0.61						
Evidence	0.54	Terror	0.61						
Phenomenon	0.54	Objective	0.61						
Recommendation	0.54	Project	0.61						
Myth	0.54	Proviso	0.61						
Prejudice	0.54	Endeavour	0.61						
Experience	0.54								
Proviso	0.54								
System	0.54								

Table 6.7b The 60 lemmas and their distribution across the top ten super-genres

W:fict	%	S:speech	%	S:conv	%	W:commerce	%	S:brdcast	%
Word	7.63	Area	8.91	Thing	15.00	Recommendation	9.18	Experience	15.4
Crime	6.78	System	5.94	Joke	13.00	Practice	8.16	Point	10.3
Foreboding	5.93	Project	4.95	Time	12.00	System	7.14	Philosophy	7.69
Joke	5.93	Time	4.95	Chance	10.00	Venture	7.14	Failure	7.69
Thing	5.08	Answer	4.95	Vision	6.00	Objective	6.12	Chance	5.13
Surprise	5.08	Philosophy	3.96	Point	5.00	Application	6.12	Time	5.13
Answer	4.24	Objective	3.96	Way	4.00	Detail	5.10	Suspicion	5.13
Detail	4.24	Practice	3.96	Word	4.00	Capacity	5.10	Warning	5.13
Capacity	3.39	Application	3.96	Problem	3.00	Dimension	4.08	Thing	5.13
Way	3.39	Point	3.96	Detail	3.00	Experience	4.08	Answer	2.56
Chance	3.39	Problem	2.97	Experience	3.00	Area	4.08	Application	2.56
Misfortune	2.54	Capacity	2.97	Example	2.00	Assessment	3.06	Characteristic	2.56
Phenomenon	2.54	Way	2.97	Work	2.00	Challenge	3.06	Endeavour	2.56
Part	2.54	Chance	2.97	Surprise	2.00	Failure	3.06	Objective	2.56
Warning	2.54	Vision	2.97	Practice	2.00	Answer	2.04	Part	2.56
Suspicion	2.54	Finding	1.98	Sense	2.00	Example	2.04	Problem	2.56
Experience	2.54	Thing	1.98	Myth	2.00	Facet	2.04	Recommendation	2.56
Time	2.54	Example	1.98	Area	2.00	Project	2.04	Scandal	2.56
Irony	2.54	Recommendation	1.98	Scandal	1.00	Sense	2.04	Crime	2.56
Dimension	1.69	Word	1.98	Part	1.00	Surprise	2.04	Evidence	2.56
Opposite	1.69	Detail	1.98	Answer	1.00	Suspicion	2.04	Finding	2.56
Vision	1.69	Experience	1.98	Warning	1.00	Way	2.04	Vision	2.56
Prejudice	1.69	Challenge	1.98	Terror	1.00	Evidence	1.02		
Sense	1.69	Scandal	1.98	Objective	1.00	Motivation	1.02		
Triumph	0.85	Joke	1.98	System	1.00	Phenomenon	1.02		

Problem	0.85	Evidence	1.98	Venture	1.00	Philosophy	1.02
Characteristic	0.85	Sense	1.98			Problem	1.02
Finding	0.85	Irony	0.99			Proviso	1.02
Impetus	0.85	Dimension	0.99			Time	1.02
Example	0.85	Facet	0.99			Work	1.02
Work	0.85	Opposite	0.99				
Contradiction	0.85	Phenomenon	0.99				
Testimony	0.85	Part	0.99				
Point	0.85	Work	0.99				
Philosophy	0.85	Warning	0.99				
Anger	0.85	Prejudice	0.99				
Terror	0.85	Assessment	0.99				
Leave	0.85	Recollection	0.99				
Failure	0.85	Surprise	0.99				
Proviso	0.85						
Practice	0.85						
Quest	0.85						
Venture	0.85						

Tables 6.7a and 6.7b show that the 60 lemmas considered in this thesis are not distributed evenly across the ten most frequent super-genres. It follows that, as stated in the literature (see 2.3), shell-noun use is highly context- and genre-sensitive. A cursory glance at the top of each list reveals different distribution patterns. These seem to accord with the various configurations of entities (concrete and abstract) that underlie the 'universe of discourse' associated with particular super-genres and with instantiations of such general categories in specific texts (van Dijk 1977: 26; cf. also Charles 2003: 318 in 2.3.1.2.1, where shell-noun use is linked to the concerns of each academic discipline).

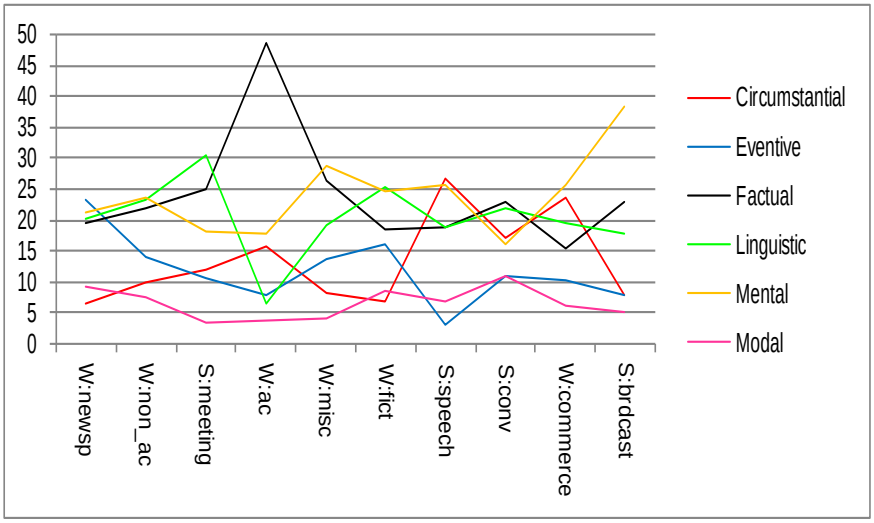
In the case of W:newsp, the high frequency of such units as *warning*, *challenge*, *problem*, *failure*, *suspicion* or *venture* relates to the state of political uncertainty and controversy following the fall of Berlin's wall in the late 1980s, a recurring issue in many of the W:newsp instances retrieved from the corpus. The prominence of *myth* in W:non_ac is text- rather than genre-related, inasmuch as 24 out of the 30 concordances for this lemma come from two books on Egyptian myths and on local myths and legends about a particular town. S:meeting appears to favour such nouns as *recommendation*, *point*, *project*, *answer* or *problem*, which is hardly surprising in view of the debate and confrontation often involved in these situations (where problems are raised, recommendations and answers are requested and given, points are made and projects are discussed). W:ac is also self-explanatory, with some of the high-ranking nouns (e.g. *phenomenon*, *characteristic*, *system*, *example*, *way*) falling into Flowerdew's (2006: 354) top 100 signalling nouns from an academic corpus. W:misc and W:fict are more text-related; the high-ranking nouns in W:misc come from a

book of jokes, a book on the construction of the Channel Tunnel and a booklet advertising Guiness Brewing Worldwide, while the most frequent lemmas for fiction occur in three novels and one play.

S:speech contains samples of speeches delivered at political, trade union and business meetings, which explains why such nouns as *area* (e.g. *area of debate*, *area of planning*), *system* (as a way of doing something which either does or does not serve its purpose), *project*, *answer* or *philosophy* (the set of principles guiding an institution) top the list of lemmas in this super-genre. As regards S:conv, the dominance of *thing* is not unexpected based on its frequency as a focusing discourse marker in *the thing is* (*that*) pattern (cf. for example Schmid 2000: 94; Carter & McCarthy 2006: 147–9; Aijmer 2007 in chapter 2). *Time*, third in the list, is, according to *BNCweb*, the second most frequent noun lemma in S:conv (0.19%; the first being *thing*, with 0.20%). However, *joke*, *chance* and *vision*, semantically more specific, are considerably less frequent in S:conv in *BNCweb*, with 0.0061%, 0.0092% and 0.0010%. Therefore their prominence in the study sample is likely to be text-related, especially in the case of *vision*, where in 4 of the 6 S:conv instances, the lemma occurs in a conversation on religious philosophy. The top lemmas in W:commerce are in line with the purpose of economic or financial treatises and articles, where official recommendations are sought and given, systems of best practice and new business ventures are discussed, and objectives are formulated. Finally, the high rank of *experience* and *point* in S:brdcast can be explained in terms of the nature of TV and radio discussions and phone-in shows, where people are invited to contribute with their own comments on a given topic or issue. The other lemmas are text- rather than genre-related.

Drawing on Tables 6.7a through 6.7b above, Figure 6.3 represents genre-related use in the six semantic types of shell noun (circumstantial, eventive, factual, linguistic, mental and modal shell nouns) in relation to the top ten super-genres in the study sample.

Figure 6.3 Genre-related use in the six semantic types of shell noun



At first sight, the most remarkable finding concerns the close association between W:ac and factual nouns (48.57%; e.g. *phenomenon*, *characteristic*, *dimension*, *finding*, *example*, *part*), these being rare in W:commerce (15.31%). Concerning the other categories, circumstantial nouns feature most frequently in S:speech (26.73%; e.g. *system*, *area*, *time*, *practice*, *way*) and W:commerce (23.47%; e.g. *practice*, *system*, *area*, *way*, *proviso*), and only marginally in W:fict (6.78%) and W:newsp (6.52%). By contrast, eventive nouns seem to associate mainly with W:newsp, where the reporting of current affairs occupies a crucial role (23.37%; e.g. *crime*, *challenge*, *scandal*, *venture*, *triumph*). This is in turn followed by W:fict (16.10%; *crime*, *misfortune*, *triumph*, *work*, *quest*) and W:non_ac (14.02%; e.g. *work*, *crime*, *challenge*, *scandal*, *quest*). Eventive meanings are infrequent in S:speech (2.97%). Linguistic nouns stand out in S:meeting (30.56%; e.g. *recommendation*, *point*, *answer*, *application*, *warning*), followed by W:fiction (25.42%; e.g. *word*, *joke*, *answer*, *warning*, *irony*). Their use is limited in W:ac (6.43%). Mental nouns prevail in S:brdcast (38.46%; e.g. *experience*, *philosophy*, *suspicion*, *objective*, *vision*), followed by W:misc (28.80%; e.g. *project*, *objective*, *assessment*, *vision*, *suspicion*) and are least frequent in S:conv (16.00%). Lastly, modal nouns, minimally represented in the study sample, appear to associate with S:conv (11.00%; *chance*) and W:newsp (9.24%; *chance*, *failure*, *capacity*) and are not frequent in S:meeting (3.47%) and W:ac (3.57%).

6.2.1.3 *Discussion*

The dominance of written English in shell noun use is confirmed by data for general noun use in the *BNC Sampler*, where 66.30% of nouns occur in the written mode and 33.69% in the spoken mode (cf. 64.55% and 35.45% for shell nouns in the study sample). Therefore, shell nouns do not show a different mode distribution to that of nouns in general (cf. also Biber et al. 1999: 231).

Although Table 6.1 confirms the often-repeated claim that shell nouns are common in written journalistic and academic discourse, other written and spoken super-genres seem to be at least as important. They are W:non_ac, S:meeting, W:misc, W:fict, S:speech, S:conv and W:commerce. The prominence of shell-noun use in academic and newspaper discourse is unsurprising, as this is also common to the entire noun word-class (Biber et al. 1999: 65 and Biber & Gray 2011: 228). In these references, academic and newspaper discourse are in turn followed by fiction and conversation.

The occurrence of S:meeting and S:speech among the top ten super-genres is borne out by the continuum between spoken and written discourse (Lyons 1977, I: 69; Halliday 1978: 224; Carter & McCarthy 2006: 164; and Biber & Conrad 2009: 261). As noted by Lyons (1977, I: 69), genres like academic lectures are both spoken and written: they use the spoken mode for their delivery, and written lexicogrammatical resources for expression of their contents. Thus, meetings and speeches, both linked to formal contexts, will draw on the lexical density of written discourse for the structuring of ideas. Written discourse is lexically dense on the grounds of its strong reliance on nominalisation, which packages what would otherwise be expressed as a string of clauses (as in the spoken mode) into '[...] an edifice of words and phrases [...]' (Halliday & Martin 1993: 39; cf. also Halliday & Matthiessen 2004: 654). This affects the textual organisation of written discourse, as the use of single entities (nouns) creates a range of possibilities for the distribution and manipulation of information both intra- and inter-sententially (see 2.2.1.3.1; cf. also Halliday & Martin 1993: 39 and Ravelli 2003: 49).

With regard to the results in Figure 6.2, the prevalence of S:conv among spoken shell nouns contrasts with Biber et al.'s (1999: 65) data for nouns in general, as these are claimed to be '[...] by far least common in conversation'. This claim is also confirmed in the *BNC*, with S:conv (of all genres) also showing the lowest percentage of noun use (12.51%). Table 6.3 confirms Biber et al.'s (1999) and the *BNC* evidence, S:conv featuring marginally in the *Sampler*-normalised data.

Table 6.2 reveals further divergence in W:fict and W:newsp, both being more primed for shell nouns and less so for general noun use. The lower

rank of W:newsp could be explained by the unrepresentative coverage of newspaper discourse in the *BNC Sampler*, with evidence only from the foreign news section of one broadsheet (*The Guardian*) and none from tabloids (Lee 2001: 64). According to Biber et al. (1999: 237), Aarts (2004: 43) and Carter & McCarthy (2006: 169), the small proportion of fiction and conversation in the written and spoken modes is due to the association between these two genres and pronouns and unmodified or structurally light noun phrases. The reason for such a similarity lies in the frequent use that written fiction makes of conversational language (in dialogues and introspection).

As to Table 6.4, the greater lexical range apparent in written shell nouns matches Flowerdew's (2010: 51; see 2.3.1.2.3) conclusion on native vs. non-native argumentative writing, where the former resembles the written use of shell nouns evident in Table 6.4 and the latter is more in line with their spoken use. More generally, this finding is also consistent with Halliday (1978: 224), Biber et al. (1999: 1049), Carter & McCarthy (2006: 169) and Biber & Conrad (2009: 94–5, 114–15, 262), where it is argued that written English is lexically more dense and varied than spoken English. However, Spoken English is syntactically more intricate than written English, as vocabulary is more limited and general in meaning, and information is spread over long strings of clauses instead of being concentrated in complex noun phrases (as in written English).

Finally, Table 6.5 lends support to Schmid's (2000: 379–80; see 2.3.2.1) 'general expectation' that, as opposed to spoken English, where short and highly frequent shell nouns are bound to prevail (e.g. *thing*, *fact*, *reason*, *point*), written English makes more use of morphologically complex and semantically specific units (e.g. *declaration*, *pronouncement*, *affirmation*, *preconception*). On these grounds, the more frequent a shell noun is in the language, the more likely it is to be used in spoken discourse. This is due to the '[...] "online" nature of spoken communication' (Carter & McCarthy 2006: 169): it forces the speaker to draw on shell nouns which, whilst not as informative as more specific and less frequent candidates, are similarly effective in the encapsulation and reification of complex discourse segments.

6.2.2 *Experiential and formal structure*

6.2.2.1 *Overall distribution*

This section examines the formal and semantic patterns underlying shell-noun use. Table 6.8 brings together the most prevalent patterns in each case. In terms of experiential structure, the 60 sampled lemmas fall into 289 patterns, 10 of which (those in Table 6.8) comprise 50.03% of the overall

evidence. Regarding formal structure, the number of patterns stands at 513 when most fine-grained and 285 when combined or most coarse-grained (e.g. PP *of*, *in*, *at*, *for*>PP). Thirty-three patterns account for 50.93% of the detailed list and 16 for 54.39% of the combined list.

Table 6.8 Top 50% for experiential (detailed) and formal patterns (detailed and combined)[30] See Appendices 4.1 through 4.3 for the complete lists

Experiential patterns	%	Formal patterns	%	Formal patterns (combined)	%
DC.SP.DM.DV^TG	10.44	DF.AR^H	6.36	DF.AR^H	6.36
DC.SP.DM.DV^TG^QF	9.40	H	5.25	DF.AR^H^PP	6.08
TG^QF	6.77	DF.AR^H^PP(*of* n)	3.80	H	5.25
TG	5.25	AJ^H	3.59	H^PP	4.91
DC.NSP.PT.NSL^TG^QF	3.87	IN.AR^H	3.04	AJ^H	4.63
DC.NSP.PT.NSL^TG	3.46	H^PP(*of* n)	2.63	DM.DT^H	3.73
CS^TG	3.11	DF.AR^AJ^H	2.42	IN.AR^AJ^H	3.11
DC.SP.PS.DV^TG	3.11	IN.AR^AJ^H	2.42	IN.AR^H	3.11
DC.SP.PS.DV^TG^QF	3.11	PS.DT^H	2.28	DF.AR^AJ^H	2.97
DC.NSP.PT.NSL^CS^TG	1.52	DM.DT(*THIS*)^H	1.59	DF.AR^AJ^H^PP	2.70
		DF.AR^H^RT.RV.CL	1.45	PS.DT^H	2.35
		DF.AR^AJ^H^PP(*of* n)	1.24	IN.AR^H^PP	2.21
		DM.DT(*THESE*)^H	1.17	IN.AR^AJ^H^PP	2.00
		QT^H	1.17	DF.AR^H^RV.CL	1.66
		IN.AR^H^PP(*of* n)	1.11	N^H	1.66
		N^H	0.97	PS.DT^H^PP	1.66
		DM.DT(*THAT*)^H	0.90		
		NUM.CD^H	0.90		
		PS.DT^AJ^H	0.90		
		DF.DV.GV.NP^H	0.83		
		AS.DT^H	0.69		
		NUM.GO^H	0.62		
		DF.AR^H^PP(*of* -ing)	0.55		
		DF.DV.GV.NP^H^PP(*of* n)	0.55		
		H^RT.RV.CL	0.55		
		IN.AR^N^H	0.55		
		AJ^AJ^H	0.48		
		DF.AR^N^H	0.48		
		H^PP(*for* n)	0.48		
		IN.AR^AJ^H^PP(*of* n)	0.48		
		IN.AR^H^AP.TI.CL	0.48		
		NAS.DT^H	0.48		
		NG.DT^AJ^H	0.48		

[30] In tables presenting only the top 50%, the cut-off point is approximate, often rising above 50% in order to allow for the existence of two or more patterns with the same frequency.

Semantically, the two most frequent experiential patterns (DC.SP.DM.DV^TG and DC.SP.DM.DV^TG^QF) contain a specific demonstrative Deictic, with the Thing being either unmodified or modified by a Qualifier. The specific Deictic and the Qualifier in experiential structure are most frequently realised by the definite article and by prepositional phrases (e.g. ***the** challenge* and ***the** challenge **of this abrasive colour***). In the latter, the preposition *of* seems to prevail in the shell-noun phrases in the study sample. The undetermined zero-article head ranks high in the list of semantic and formal patterns (TG/H, e.g. *projects* and TG/H^QF/PP, e.g. *projects with very large sample groups*). Experiential premodifiers are fairly uncommon among the top 50% semantic patterns: only Classifiers feature in 2 of the 10 patterns (4.63%; CS^TG, e.g. ***restrictive** practices*, ***natural** phenomena*; DC.NSP.PT.NSL^CS^TG, e.g. *a **community** dimension, a **business** venture*).

Table 6.9 is more specific than Table 6.8. It focuses on Deictic and determiner use. The bottom of each list contains cases of combined Deictic or determiner use (most often, partitive in nature, as in NUM.CD^DF.AR for *one of the* or *two of the*, and AS.PN^DF.AR, for *some of the*; see also cases like UV.PDT^DF.AR, as in *all the*, where no partitive *of*-phrase occurs).

Table 6.9 Deictics and determiners in the study sample

Deictics	%	Determiners	%
DC.SP.DM.DV	33.31	DF.AR	27.02
Ø (TG or CS/EP + TG)	20.25	Ø (H or AJ/N + H)	22.11
DC.NSP.PT.NSL	17.28	IN.AR	16.10
DC.SP.PS.DV	10.09	PS.DT	7.05
NUM.IN.QV	3.66	DM.DT	5.74
PDC	3.18	QT	3.87
DC.NSP.TL.NG	2.76	DF.DV.GV.NP	3.04
NUM.DF.QV	2.28	NUM.CD	2.35
DC.NSP.PT.SL	2.07	NG.DT	2.00
DC.NSP.PT.SL^DC.SP.DM.DV	1.45	AS.DT	1.59
DC.NSP.TL.POS	0.90	NAS.DT	1.59
DC.NSP.TL.POS^DC.SP.DM.DV	0.69	NUM.GO	1.31
DC.SP.DM.IV	0.35	NUM.CD^DF.AR	0.97
DC.SP.PS.IV	0.35	PDT	0.62
NUM.IN.QV^DC.SP.DM.DV	0.35	UV.DT	0.48
DC.NSP.PT.SL^DC.SP.PS.DV	0.28	UV.PDT	0.41
FC.ET.AGG	0.14	UV.PDT^DF.AR	0.41
NUM.IN.QV^DC.NSP.PT.NSL	0.14	IV.DT	0.35
DC.NSP.TL.POS^DC.SP.PS.DV	0.07	RV.DT	0.35
DC.SP.DM.DV^DC.NSP.PT.NSL	0.07	NUM.CD^NUM.GO	0.28
DC.SP.DM.DV^DC.NSP.PT.SL	0.07	NUM.CD^PS.DT	0.28
DC.SP.DM.DV^DC.SP.PS.DV	0.07	NUM.CD^NUM.CD	0.21
FC.ET.PV^DC.NSP.TL.POS^DC.SP.PS.DV	0.07	QT.PN^DF.AR	0.21
FC.ET.PV^DC.SP.DM.DV	0.07	AS.PN^DF.AR	0.14

NUM.IN.QV^DC.SP.PS.DV	0.07	NAS.PN^DF.AR	0.14
		QT^IN.AR	0.14
		QV.PV	0.14
		UV.PDT^DM.DT	0.14
		AS.PN^DF.DV.GV.NP	0.07
		AS.PN^DM.DT	0.07
		DF.AR^DF.DV.GV.NP	0.07
		DF.AR^DM.DT	0.07
		MR.GV	0.07
		NAS.PN^DM.DT	0.07
		NUM.CD^DM.DT	0.07
		QT.PN^DM.DT	0.07
		QT^DF.AR	0.07
		QV.PV^DM.DT	0.07
		QV.PV^UV.PDT^PS.DT	0.07
		UV.PDT^PS.DT	0.07
		UV.PN^DF.AR	0.07
		UV.PN^DM.DT	0.07

Table 6.9 indicates that, while specific demonstrative Deictics (most commonly *the*) are highly frequent with shell-noun behaviour (33.31%), non-specific partial non-selective Deictics (usually, *a/an*) and even absence of a Deictic are as relevant to shell-noun use in the study sample (17.28% and 20.25%). If all specific and non-specific Deictic categories are subsumed under DC.SP (specific Deictic) and DC.NSP (non-specific Deictic), the above 33.31% for DC.SP.DM.DV rises to 44.09% for DC.SP, showing that almost half of the Deictics used in the study sample are specific in nature. Unlike the definite article and possessive determiners (27.02% and 7.05%), demonstrative determiners rank low among specific Deictics (5.74%). *This* is the most frequent choice (2.83%), followed by *that* (1.38%), *these* (1.31%) and *those* (0.21%).

Regarding experiential premodifiers, the evidence seems to disfavour premodification in the shape of post-Deictics, Epithets or Classifiers. Only 37.39% of the shell-noun instances contain experiential premodification, while 62.61% lacks premodification. Table 6.10 lists the types of semantic modification in the study sample (48 in total).

Table 6.10 Types of experiential premodification in the study sample. Percentages are based on the total of instances with experiential premodification (541 concordances)

Semantic premodification (%)			
CS	32.16	PDC.EL.ID&EP.IP	0.37
EP.IP	18.11	PDC.EL.ID.EZ	0.37
EP.EX	11.65	PDC.RP.IA.EZ	0.37
PDC.EL.ID	6.84	CS&EP.EX	0.18
PDC.EL.ID.RT	4.99	PDC.EL.EM&PDC.EL.ID	0.18
PDC.EL.EM.RT	3.14	PDC.EL.ID&CS	0.18
EP.EX&CS	2.96	PDC.EL.ID&EP.EX&CS	0.18
PDC.EN.SPA-TM	2.77	PDC.EN.CV&PDC.EN.CV	0.18
CS&CS	2.40	PDC.EN.SPA-TM&PDC.EL.ID	0.18
PDC.ET.AM	1.48	PDC.ET.RT	0.18
PDC.EL.EM	1.11	PDC.MD.OB	0.18
PDC.EN.CV	1.11	PDC.MD.PB&EP.EX	0.18
PDC.MD.PB	0.92	PDC.MD.PB.EZ	0.18
PDC.EL.EM.EZ	0.74	PDC.MD.RD	0.18
PDC.MD.US	0.74	PDC.MD.US&CS	0.18
EP.IP&EP.IP	0.55	PDC.RP.IA	0.18
PDC.EN.SPA-TM&CS	0.55	PDC.RP.IA&PDC.MD.US	0.18
PDC.ET.EZ	0.55	PDC.RP.IA.EZ&CS	0.18
CS&CS&CS	0.37	PDC.RP.LN	0.18
EP.EX&CS&CS	0.37	PDC.RP.LN.EZ&CS	0.18
EP.EX&EP.EX	0.37	PDC.ET.EZ&EP.IP	0.18
EP.IP&CS	0.37	PDC.MD.US.EZ	0.18
EP.IP&EP.EX	0.37	PDC.RP.IA&CS	0.18
PDC.EL.EM.AM	0.37	CS&EP.IP	0.18

Classifiers prevail among experiential premodifiers (32.16%; as in (639)). They are followed by Epithets (29.76%). Of the two Epithet subtypes, interpersonal Epithets are more frequent than experiential ones (18.11% for EP.IP, as in (640a) vs. 11.65% for EP.EX, as in (640b)). Use of post-Deictics appears to be concentrated in PDC.EL.ID (12.55%, as in (641a)), PD.EL.EM (5.54%, as in (641b)), PDC.EN.SPA-TM (3.5%, as in (641c)) and PDC.ET (2.39%, as in (641d))[31].

(639) planning application, experimental evidence and 40m chance

(640) (a) the strange phenomenon, a positive finding and one of the direst warnings
(b) colourful visions, serious crimes and new capacity to produce

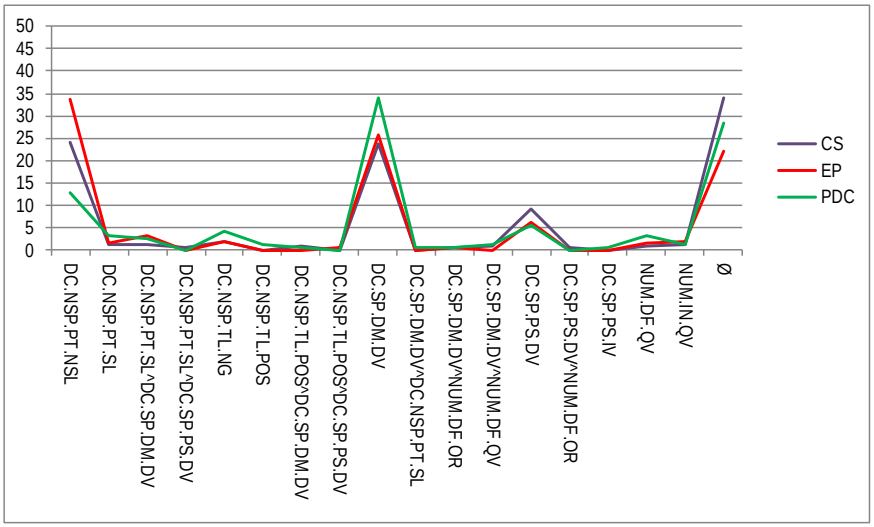
(641) (a) another awesome chalenge, one or two other points and the only answer
(b) this particular venture, certain provisos and a specific warning

[31] The four percentages given here include AM, EZ and RT subtypes, as well as combined patterns, e.g. PDC.EL.EM&PDC.EL.ID.

(c) the **next** example, the **current** project and the **earlier** myth
(d) the **entire** project, **overall** philosophy and **total** triumph

Appendix 5 presents the experiential patterns where Classifiers, Epithets and post-Deictics are used. Figure 6.4 narrows the scope of data in these tables by focusing on Deictic and Numerative use before Classifiers, Epithets and post-Deictics.

Figure 6.4 Use of Deictics and Numeratives in patterns with Classifiers, Epithets and post-Deictics. Percentages are based on the total number of pattern tokens with one or more examples of each semantic category



Although the three lines in this figure share a similar and almost overlapping distribution, slight differences emerge. For example, patterns with non-specific partial non-selective Deictics (most commonly *a/an*) appear to be associated with Epithets instead of with Classifiers or post-Deictics. Post-Deictics, by contrast, are more frequent among patterns with specific demonstrative Deictics (most commonly *the*), while Classifiers seem to be associated with patterns with no Deictic. In the case of specific possessive Deictics and non-specific negative Deictics, the association with Classifiers in the former and post-Deictics in the latter is minimal.

As to postmodification structures, 48.51% of the overall evidence contains postmodification, while 51.49% does not. Table 6.11 brings together the types of postmodifying structures (33 in total) in the study sample. It also includes a list of the prepositions occurring in PP structures.

Table 6.11 Postmodification structures and prepositions in PP. Percentages are based on the total number of postmodification tokens (702) for the former and on the total number of prepositional tokens (453) for the latter

Postmodification structures	%	Prepositions	%
PP	59.83	OF	59.82
RT.RV.CL	13.68	FOR	9.05
AP.TI.CL	6.27	IN	6.40
PL.ED.CL	3.28	TO	4.64
AP.THAT.CL	2.99	ON	3.53
TI.CL	2.42	ABOUT	3.09
PT.WK.NR.AP	1.28	WITH	2.87
NR.RV.CL	0.85	FROM	1.32
PL.ING.CL	0.85	BETWEEN	1.10
PP^AP.TI.CL	0.85	AGAINST	0.88
PP^RT.RV.CL	0.85	AS TO	0.66
AJ	0.71	AT	0.66
FL.ST.NR.AP	0.71	BEHIND	0.66
PP^AP.THAT.CL	0.71	OF^IN	0.66
FL.ST.RT.AP	0.57	OVER	0.66
PT.ST.NR.AP	0.57	AS	0.44
PP^NR.RV.CL	0.43	AMONG	0.22
PP^PT.WK.NR.AP	0.43	AT^IN	0.22
RT.RV.CL^PP	0.43	BY	0.22
AV	0.28	BY^OF	0.22
PP^PT.ST.NR.AP	0.28	CONCERNING	0.22
AV^PT.WK.NR.AP	0.14	DURING	0.22
FL.ST.NR.AP^PL.ED.CL	0.14	IN TERMS OF	0.22
FL.WK.NR.AP	0.14	INTO	0.22
PL.ED.CL^AP.THAT.CL	0.14	LIKE	0.22
PL.ING.CL^PP	0.14	OF^ABOUT	0.22
PP^FL.ST.NR.AP	0.14	OF^AT	0.22
PP^PL.ED.CL	0.14	OF^FOR	0.22
PP^PL.ING.CL	0.14	OF^ON	0.22
PT.ST.NR.AP^NR.RV.CL	0.14	OF^OVER	0.22
PT.ST.NR.AP^RT.RV.CL	0.14	OTHER THAN	0.22
RT.RV.CL^PT.ST.NR.AP	0.14	SINCE	0.22
TI.CL^PP	0.14		

Table 6.11 indicates a dominance of prepositional phrases (especially *of*-phrases) among postmodification structures (59.83%, as in (642)). These are followed by restrictive relative clauses (13.68%, as in (643)), appositive *to*-infinitive clauses (6.27%, as in (644)), participle *-ed* clauses (3.28%, as in (645)) and appositive *that*-clauses (2.99%, as in (646)).

(642) examples **of products in this class**, alarming details **about Bulgaria's ailing economy**

(643) differing work practices which I try to erode, the experiences that we go through
(644) a chance to speak, his way to unite and bring peace to Lebanon
(645) natural capacities bestowed as a gift, a challenge faced by many companies that want to progress
(646) a suspicion that a secret cartel is in operation, recent warnings from politicians of all parties that the influx cannot continue indefinitely

Table 6.12 merges the 33 postmodification types in Table 6.11 into 21 more general categories, namely RT.RV.CL and NR.RV.CL>RV.CL, AP.TI.CL and AP.THAT.CL>AP.CL, PL.ING.CL and PL.ED.CL>PL.CL and FL/PT.NR.AP>NR.AP. The top three structures, i.e. prepositional phrases, relative clauses and appositive clauses, make up 83.62% of all general or combined postmodification types.

Table 6.12 Postmodification structures (combined)

Postmodification structures (combined)			
PP	59.83	RT.AP	0.57
RV.CL	14.53	RV.CL^PP	0.43
AP.CL	9.26	AV	0.28
PL.CL	4.13	PP^PL.CL	0.28
NR.AP	2.71	NR.AP^RV.CL	0.28
TI.CL	2.42	AV^NR.AP	0.14
PP^AP.CL	1.57	NR.AP^PL.CL	0.14
PP^RV.CL	1.28	PL.CL^AP.CL	0.14
PP^NR.AP	0.85	PL.CL^PP	0.14
AJ	0.71	RV.CL^NR.AP	0.14
		TI.CL^PP	0.14

6.2.2.2 *Mode and genre distribution*

The overall distribution of experiential and formal patterns presented above will now be explored in relation to mode and genre. Tables 6.13 through 6.16 below show the top 50% experiential and formal patterns in the written and spoken sections of the study sample (934 and 513 concordances; see 6.2.1.1). The realisation of written instances comprises 219 experiential patterns and 366 formal patterns. Thirteen pattern types make up 53.64% of the former, while 31 correspond to 50.64% of the latter (see Tables 6.13 and 6.14). Spoken instances fall into fewer experiential and formal patterns, 148 and 237 respectively. Tables 6.13 and 6.14 contain 50.88% of the list of experiential patterns (9 patterns) and 54% of the the list of formal patterns (32 patterns).

Table 6.13 Mode distribution of experiential patterns (Top 50%). See Appendix 6.1 for the complete list

Written	%	Spoken	%
DC.SP.DM.DV^TG^QF	9.21	DC.SP.DM.DV^TG	12.87
DC.SP.DM.DV^TG	9.10	DC.SP.DM.DV^TG^QF	9.75
TG^QF	7.28	TG^QF	5.85
TG	6.00	DC.NSP.PT.NSL^TG	5.65
DC.SP.PS.DV^TG^QF	3.85	DC.NSP.PT.NSL^TG^QF	5.07
CS^TG	3.75	TG	3.70
DC.NSP.PT.NSL^TG^QF	3.21	DC.SP.PS.DV^TG	3.51
DC.SP.PS.DV^TG	2.89	NUM.IN.QV^TG	2.53
DC.NSP.PT.NSL^TG	2.25	CS^TG	1.95
CS^TG^QF	1.61		
DC.NSP.PT.NSL^CS^TG	1.50		
DC.NSP.PT.NSL^CS^TG^QF	1.50		
DC.SP.DM.DV^CS^TG	1.50		

Table 6.14 Mode distribution of formal patterns (Top 50%). See Appendix 6.2 for the complete list

Written	%	Spoken	%
H	6.00	DF.AR^H	8.77
DF.AR^H	5.03	IN.AR^H	5.26
AJ^H	4.39	H	3.70
DF.AR^H^PP(of n)	4.07	DF.AR^H^PP(of n)	3.12
H^PP(of n)	2.89	IN.AR^AJ^H	2.92
DF.AR^AJ^H	2.36	DF.AR^H^RT.RV.CL	2.73
IN.AR^AJ^H	2.14	PS.DT^H	2.53
PS.DT^H	2.14	DF.AR^AJ^H	2.34
DM.DT(THIS)^H	1.82	DM.DT(THAT)^H	2.34
IN.AR^H	1.82	AJ^H	2.14
DF.AR^AJ^H^PP(of n)	1.71	H^PP(of n)	2.14
DM.DT(THESE)^H	1.50	QT^H	1.95
IN.AR^H^PP(of n)	1.18	DM.DT(THIS)^H	1.17
N^H	1.18	DF.DV.GV.NP^H	0.97
PS.DT^AJ^H	1.07	IN.AR^H^PP(of n)	0.97
NUM.CD^H	0.86	NUM.CD^H	0.97
DF.AR^H^RT.RV.CL	0.75	AS.DT^H	0.78
DF.AR^N^H	0.75	DF.AR^AJ^H^RT.RV.CL	0.78
DF.DV.GV.NP^H	0.75	NAS.DT^H	0.78
DF.DV.GV.NP^H^PP(of n)	0.75	DF.AR^H^PP(to n)	0.58
H^RT.RV.CL	0.75	DF.AR^NP^H	0.58
NG.DT^AJ^H	0.75	DF.AR^QL.PV^H^RT.RV.CL	0.58
NUM.GO^H	0.75	DM.DT(THESE)^H	0.58
QT^H	0.75	H^RT.RV.CL(when)	0.58
AS.DT^H	0.64	IN.AR^H^AP.TI.CL	0.58
DF.AR^H^PP(of -ing)	0.64	IN.AR^H^PP(of -ing)	0.58
DF.DV.GV.NP^AJ^H	0.64	IN.AR^H^RT.RV.CL(where)	0.58
H^PP(for n)	0.64	N^H	0.58
IN.AR^AJ^H^PP(of n)	0.64	NP^H	0.58

IN.AR^N^H	0.64	NUM.CD^PV.PP(of DF.AR^AJ^H^RT.RV.CL)	0.58
NG.DT^H	0.64	NUM.CD^PV.PP(of DF.AR^H^RT.RV.CL)	0.58
		PS.DT^AJ^H	0.58

Table 6.13 reveals a coincidence with the top three patterns in Table 6.8. Whilst use of DC.SP.DM.DV^TG^QF is roughly equal in both modes (9.10% and 9.75%), DC.SP.DM.DV^TG and TG^QF seem to be associated slightly more frequently with spoken and written English respectively (12.87% vs. 9.10%; 7.28% vs. 5.85%). The differences concerning TG, CS^TG, and DC.NSP.PT.NSL^TG are also worth mentioning. The former two feature more frequently in written English (6.00% vs. 3.70%; 3.75% vs. 1.95%), whilst the latter is more frequent in the spoken mode (5.65% vs. 2.25%). Interestingly, 5 of the 13 written patterns contain semantic premodification in the shape of Classifiers, while their occurrence is restricted to only one of the top 50% spoken patterns: CS^TG.

The differences regarding formal patterns are more revealing. The list of written patterns in Table 6.14 is topped by H, followed by DF.AR^H and AJ^H. By contrast, in the spoken mode H comes third, preceded by DF.AR^H and IN.AR^H. The third most frequent pattern in Table 6.8, i.e. DF.AR^H^PP(*of* n), ranks fourth in both modes. Divergences are most marked with DF.AR^H and IN.AR^H: they appear to be associated with spoken English, showing a difference with the written mode of 3.74% in the case of DF.AR^H and 3.44% in the case of IN.AR^H. Other noticeable differences concern H, AJ^H and DF.AR^H^RT.RV.CL, the former two being more frequent in the written mode and the latter in the spoken mode. In relation to DF.AR^H^RT.RV.CL, it is worth noting that whilst restrictive relative clauses occur in 6 of the top 50% spoken patterns, they feature in only 2 of the written patterns. This appears to indicate that experiential Qualifiers are realised by restrictive relative clauses in spoken English and by prepositional phrases in written English. Finally, DM.DT(THAT)^H is strongly primed for spoken English (2.34% vs. 0.11% in written English).

Turning now to genre, Tables 6.15a through 6.15c and 6.16a through 6.16c display the top 50% experiential and formal patterns for the top ten super-genres in the study sample (see 6.2.1.2). As in the tables above, the cut-off point for most genre categories rises above 50% to keep together groups of cut-off instances with the same frequency[32]. Each group of tables

[32] In four cases, however, the number of patterns presented does not reach 50%, as all remaining pattern types following the cut-off point occur only once in the super-genre. This applies to W:ac, S:speech, W:commerce and S:brdcast, where the cut-off point within formal structure stands at 45.71% (16 out of 92 patterns), 45.54% (15 out of 70 patterns), 21.20% (14 out of 73 patterns) and 35.90% (3 out of 28 patterns; see Tables 6.16a through 6.16c)

is followed by a figure showing the genre distribution of the top ten experiential and formal patterns in Table 6.8. The larger number of patterns in Figure 6.6 (13 instead of 10) is intended to bring to light differences in the distribution of the four demonstrative determiners in the DM.DT^H pattern.

Table 6.15a Top ten super-genres and their experiential patterns (Top 50%). See Appendix 7.1 for the complete lists

W:newsp	%	W:non_ac	%	S:meeting	%	W:ac	%
DC.SP.DM.DV^TG^QF	8.70	DC.SP.DM.DV^TG	12.80	DC.SP.DM.DV^TG	15.28	DC.SP.DM.DV^TG	13.57
TG^QF	8.70	TG	9.15	DC.SP.DM.DV^TG^QF	12.50	DC.SP.DM.DV^TG^QF	7.86
CS^TG	5.43	DC.SP.DM.DV^TG^QF	7.93	DC.SP.PS.DV^TG	5.56	TG^QF	5.00
DC.SP.DM.DV^TG	5.43	DC.SP.PS.DV^TG^QF	6.71	DC.NSP.PT.NSL^TG^QF	4.86	DC.SP.PS.DV^TG^QF	4.29
DC.SP.PS.DV^TG^QF	4.89	TG^QF	6.10	TG^QF	4.17	CS^TG	2.86
TG	4.89	DC.NSP.PT.NSL^TG^QF	4.27	DC.NSP.PT.NSL^TG	2.78	DC.NSP.PT.NSL^EP.EX^TG^QF	2.86
DC.NSP.PT.NSL^CS^TG	2.72	CS^TG	3.66	TG	2.78	DC.SP.PS.DV^TG	2.86
DC.SP.PS.DV^CS^TG	2.72			DC.NSP.PT.SL^DC.SP.DM.DV^TG^QF	2.08	NUM.IN.QV^TG^QF	2.86
DC.SP.PS.DV^TG	2.72			DC.SP.DM.DV^PDC.EN.SPA-TM^TG	2.08	TG	2.86
DC.SP.DM.DV^CS^TG^QF	2.17			DC.SP.DM.DV^TG^QF^QF	2.08	CS^TG^QF	2.14
DC.NSP.PT.NSL^CS^TG^QF	1.63					DC.NSP.PT.NSL^TG^QF	2.14
DC.NSP.PT.NSL^EP.IP^TG^QF	1.63					PDC.EL.ID^TG	2.14
DC.NSP.PT.NSL^TG^QF	1.63						
DC.SP.DM.DV^CS^TG	1.63						
DC.SP.DM.DV^EP.IP^TG	1.63						
DC.SP.DM.DV^TG^QF^QF	1.63						
EP.EX^TG	1.63						
NUM.DF.QV^TG	1.63						

Table 6.15b Top ten super-genres and their experiential patterns (Top 50%)

W:misc	%	W:fict	%	S:speech	%	S:conv	%
DC.SP.DM.DV^TG^QF	14.40	DC.NSP.PT.NSL^TG	11.02	DC.SP.DM.DV^TG	13.86	DC.NSP.PT.NSL^TG	13
DC.SP.DM.DV^TG	8.00	TG	11.02	DC.SP.DM.DV^TG^QF	10.89	DC.SP.DM.DV^TG	9
TG^QF	6.40	DC.SP.DM.DV^TG	6.78	TG^QF	5.94	DC.NSP.PT.NSL^TG^QF	7
DC.NSP.PT.NSL^TG^QF	4.80	DC.SP.DM.DV^TG^QF	6.78	DC.NSP.PT.NSL^TG	4.95	TG^QF	6
TG	4.80	DC.SP.PS.DV^TG	6.78	DC.SP.PS.DV^TG	3.96	DC.SP.DM.DV^TG^QF	5
CS^TG	3.20	TG^QF	5.93	TG	3.96	NUM.IN.QV^TG	5
DC.NSP.PT.NSL^CS^TG^QF	3.20	DC.NSP.PT.NSL^TG^QF	3.39	CS^TG	2.97	TG	5
CS^TG^QF	2.40			DC.NSP.PT.NSL^TG^QF	2.97		
DC.SP.DM.DV^CS^TG^QF	2.40			DC.SP.DM.DV^EP.IP^TG^QF	2.97		
DC.SP.PS.DV^TG	2.40			DC.SP.PS.DV^TG^QF	2.97		
NUM.IN.QV^TG^QF	2.40			NUM.IN.QV^TG	2.97		

Table 6.15c Top ten super-genres and their experiential patterns (Top 50%)

W:commerce	%	S:brdcast	%
TG^QF	12.24	DC.SP.DM.DV^TG	23.08
DC.SP.DM.DV^TG	11.22	DC.NSP.PT.NSL^TG	10.26
DC.SP.DM.DV^TG^QF	11.22	CS^TG	5.13
CS^TG	6.12	DC.NSP.PT.NSL^TG^QF	5.13
DC.NSP.PT.NSL^CS^TG	4.08	DC.NSP.TL.NG^TG^QF	5.13
CS^TG^QF	3.06	TG^QF	5.13
DC.NSP.PT.NSL^TG^QF	3.06		
DC.SP.PS.DV^CS^TG	3.06		
TG^QF^QF	3.06		

The highest-ranking experiential patterns in 7 out of 10 super-genres contain a specific demonstrative Deictic (DC.SP.DM.DV^TG). Non-specific partial non-selective Deictics in DC.NSP.PT.NSL^TG top the list of W:fict and S:conv, while TG^QF ranks highest in W:commerce. Patterns without Deictics, especially TG^QF and TG, rank second or third in most super-genres (as shown also in Table 6.8; cf., however, S:meeting and S:conv, where DC.SP.PS.DV^TG and DC.NSP.PT.NSL^TG^QF occur as the third most frequent patterns in each case).

Figure 6.5 Genre distribution of the top ten experiential patterns

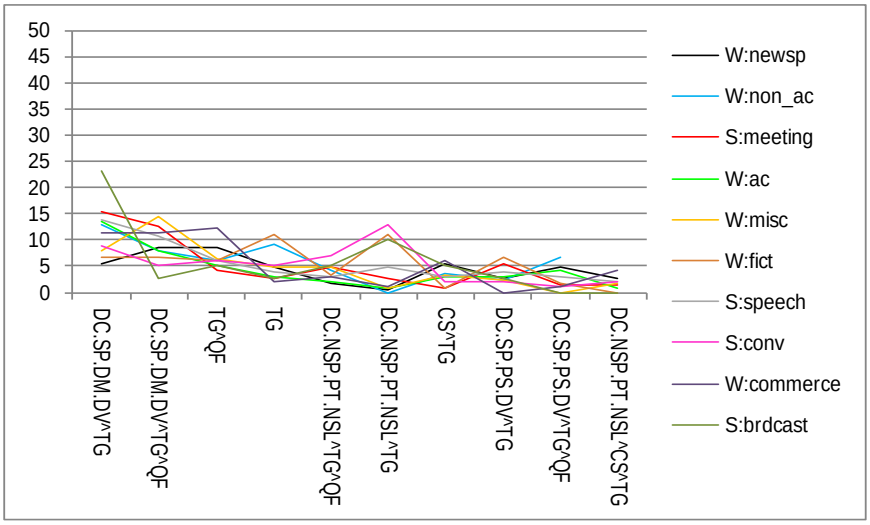


Figure 6.5 shows several other associations among the ten most frequent experiential patterns. The most remarkable of these concern DC.SP.DM.DV^TG (S:brdcast), TG^QF (W:commerce), TG (W:fict and W:non_ac) and DC.NSP.PT.NSL (S:conv, W:fict and S:brdcast; cf. DC.NSP.PT.NSL^TG^QF, where S:conv also ranks highest). There are weaker associations between DC.SP.DM.DV^TG^QF and W:misc and S:meeting, CS^TG and W:commerce and W:newsp, DC.SP.PS.DV^TG and W:fict and S:meeting, and DC.SP.PS.DV^TG^QF and W:newsp and W:ac.

Table 6.16a Top ten super-genres and their formal patterns (Top 50%). See Appendix 7.2 for the complete lists

W:newsp	%	W:non_ac	%	S:meeting	%	W:ac	%
AJ^H	5.43	H	9.15	DF.AR^H	9.03	AJ^H	5.71
H	4.89	DF.AR^H	8.54	DF.AR^H^PP(of n)	5.56	DF.AR^H	5.00
DF.AR^H	4.35	DF.AR^H^PP(of n)	5.49	PS.DT^H	4.17	DM.DT(THESE)^H	5.00
DF.AR^H^ PP(of n)	3.80	H^PP(of n)	3.66	DF.AR^AJ^H	3.47	DF.AR^H^PP(of n)	2.86
H^PP(of n)	3.80	AJ^H	3.05	DM.DT(THAT)^H	3.47	DM.DT(THIS)^H	2.86
DF.AR^AJ^H	3.26	DF.AR^AJ^H^PP(of n)	3.05	AJ^H	2.78	H	2.86
N^H	2.72	IN.AR^H^PP(of n)	3.05	DF.AR^H^RT.RV.CL	2.78	IN.AR^AJ^H	2.86
IN.AR^AJ^H	2.17	DM.DT(THIS)^H	2.44	H	2.78	PS.DT^H	2.86
NG.DT^AJ^H	2.17	DF.AR^N^H	1.83	IN.AR^AJ^H	2.78	DF.AR^AJ^H	2.14
NUM.CD^H	2.17	DF.DV.GV.NP^H^ PP(of n)	1.83	DM.DT(THIS)^H	2.08	DF.AR^AJ^H^PP(of n)	2.14
DF.DV.GV.NP^H	1.63	IN.AR^AJ^H	1.83	IN.AR^H	2.08	DF.DV.GV.NP^H^ PP(of n)	2.14
DF.DV.GV.NP^H^ AP.TI.CL	1.63	PS.DT^H	1.83	AS.DT^H	1.39	H^RT.RV.CL	2.14

H^PP(for n)	1.63	AJ^AJ^H	1.22	DF.AR^H^PP(on n)	1.39	IN.AR^AJ^H^PP(of n)	2.14
IN.AR^AJ^H^AP.TI.CL	1.63	AJ^H^PL.ED.CL	1.22	DF.DV.GV.NP^H	1.39	NUM.GO^H	2.14
IN.AR^N^H	1.63	DF.AR^AJ^H	1.22	H^PP(of n)	1.39	AS.DT^H	1.43
PS.DT^AJ^H	1.63	DF.AR^AJP^H	1.22	IN.AR^H^PP(of n)	1.39	NUM.CD^H	1.43
DF.AR^AJ^H^PP(of n)	1.09	DF.AR^H^PP(of n)^AP.TI.CL	1.22	NAS.DT^H	1.39		
DF.AR^N^H	1.09	DF.AR^H^RT.RV.CL	1.22	NUM.CD^PV.PP(of DF.AR^AJ^H^RT.RV.CL)	1.39		
DF.DV.GV.NP^N^H	1.09	DM.DT(THIS)^AJ^H	1.22	NUM.CD^PV.PP(of DF.AR^H^RT.RV.CL)	1.39		
DM.DT(THIS)^H	1.09	IN.AR^AJ^H^PP(to n)	1.22	QT^H	1.39		
IN.AR^H^AP.TI.CL	1.09	IN.AR^AJP^H	1.22				
NAS.DT^AJ^H	1.09	N^H	1.22				
NG.DT^H	1.09	PS.DT^AJ^H	1.22				
PS.DT^H	1.09						
PS.DT^H^PP(of n)	1.09						
QT^H	1.09						

Table 6.16b Top ten super-genres and their formal patterns (Top 50%)

W:misc	%	W:fict	%	S:speech	%	S:conv	%
DF.AR^H^PP(of n)	6.40	H	11.02	DF.AR^H	10.89	IN.AR^H	13.00
AJ^H	4.80	IN.AR^H	11.02	DF.AR^H^RT.RV.CL	3.96	DF.AR^H	7.00
H	4.80	PS.DT^H	5.93	H	3.96	H	5.00
DF.AR^H	4.00	AJ^H	3.39	IN.AR^H	3.96	DF.AR^H^RT.RV.CL	4.00
DF.AR^AJ^H	3.20	DF.AR^H	3.39	QT^H	2.97	DF.AR^AJ^H	3.00
DF.AR^AJ^H^PP(of n)	3.20	DF.AR^AJ^H	2.54	AS.DT^H	1.98	H^PP(of n)	3.00
DM.DT(THESE)^H	2.40	DF.AR^H^PP(of n)	2.54	DF.AR^AJ^H	1.98	QT^H	3.00
H^TI.CL	2.40	DF.AR^H^RT.RV.CL	2.54	DF.AR^H^PP(of n)	1.98	AJ^H	2.00
IN.AR^AJ^H	2.40	DF.DV.GV.NP^AJ^H	2.54	DF.DV.GV.NP^H	1.98	IN.AR^AJ^H	2.00
AJ^H^PP(in n)	1.60	PDT^H	2.54	DM.DT(THIS)^H	1.98	IN.AR^H^AP.TI.CL	2.00
AJ^H^PP(of n)	1.60	H^PP(of n)	1.69	H^PP(of n)	1.98	IN.AR^H^PP(with n)	2.00
DF.AR^H^PL.ED.CL	1.60	IN.AR^H^PP(of n)	1.69	H^RT.RV.CL(when)	1.98	N^H	2.00
DF.AR^H^PP(of -ing)	1.60	NG.DT^H	1.69	IN.AR^H^PP(of n)	1.98	PS.DT^H	2.00
DF.DV.GV.NP^H	1.60	PS.DT^AJ^H	1.69	NP^H	1.98	QT^AJ^H	2.00
DM.DT(THIS)^H	1.60	QT^H	1.69	PS.DT^H	1.98	QT^H^RT.RV.CL	2.00
H^PP(of n)	1.60	UV.DT^H	1.69			UV.PDT^DF.AR^H	2.00
H^RT.RV.CL	1.60						
IN.AR^AJ^H^PL.ED.CL	1.60						
IN.AR^H^PP(of n)	1.60						
IN.AR^H^RT.RV.CL	1.60						
N^H	1.60						
NUM.GO^H	1.60						
PS.DT^AJ^H	1.60						

PS.DT^AJ^H^PP(in n)	1.60

Table 6.16c Top ten super-genres and their formal patterns (Top 50%)

W:commerce	%	S:brdcast	%
DF.AR^H	2.72	DF.AR^H	17.95
H^PP(of n)	2.72	IN.AR^H	10.26
AJ^H	2.17	AJ^H	7.69
DF.AR^H^PP(of n)	2.17		
DM.DT(THIS)^H	1.63		
AJ^AJ^H	1.09		
DF.AR^H^PP(of -ing)	1.09		
DF.DV.GV.NP^AJ^H	1.09		
DM.DT(THESE)^H	1.09		
H	1.09		
H^PP(for n)	1.09		
IN.AR^AJ^H	1.09		
IN.AR^AJ^N^H	1.09		
N^H	1.09		

Tables 6.16a through 6.16c show four prominent formal patterns in the top three of most super-genres: DF.AR^H, DF.AR^H^PP(*of* n), H and AJ^H. Other patterns seem strongly primed for specific genre categories. For example, PS.DT^H ranks third in S:meeting and W:fict (e.g. *my project*, *my visions*; *your experience*, *his surprise*), and DM.DT(*THESE*)^H occupies the same position in W:ac (e.g. *these phenomena*, *these findings*). As expected from Figure 6.5, IN.AR^H prevails in W:fict, S:conv and S:brdcast (e.g. *a crime*, *a warning*; *a point*, *a joke*; *a philosophy*, *a scandal*).

Figure 6.6 Genre distribution of the top ten formal patterns

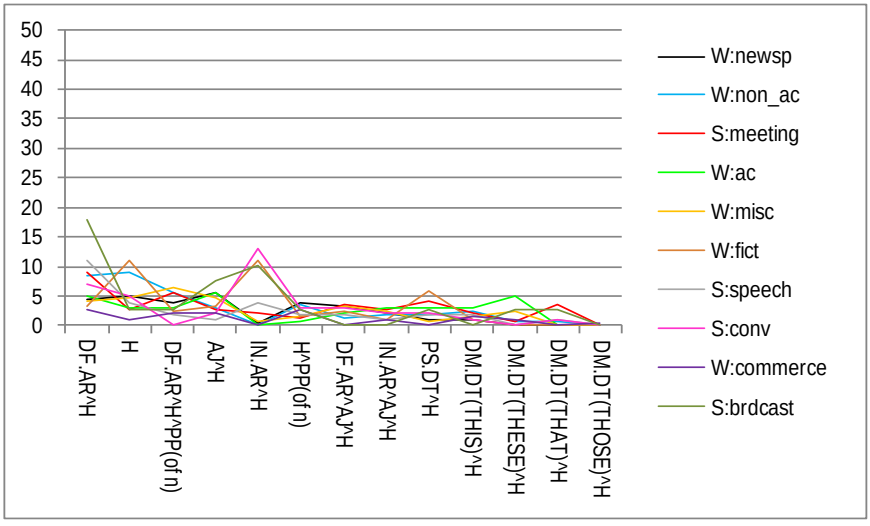


Figure 6.6 reveals preferences which roughly coincide with some of those in Figure 6.5. The most obvious ones are those regarding DF.AR^H (S:brdcast), H (W:fict and W:non_ac), AJ^H (S:brdcast), IN.AR^H (S:conv, W:fict and S:brdcast) and PS.DT^H (W:fict and S:meeting). Concerning the demonstrative determiners, the link shown in Table 6.14 between DM.DT(*THIS*)/(*THESE*)^H and written English and DM.DT(*THAT*)^H and spoken English is here reflected in the stronger priming between the former two and W:ac, and the latter and S:meeting and S:brdcast.

With the above results in mind, in the following the focus is on the mode and genre distribution of the ten most frequent determiners, semantic premodifiers and formal postmodifying structures.The choice of the top ten is motivated by the need to make the figures below more visible, given the large number of categories in each dataset (see Tables 6.9 through 6.11 above). This being the case, the ten most frequent categories are highly representative of each dataset, comprising 90.88% of all determiner types, 86.50% of semantic premodifiers and 92.30% of postmodifying structures.

Figures 6.7 and 6.8 below illustrate the extent of use of determiner types in relation to mode and super-genre.

Figure 6.7 Mode distribution of the top ten determiners. Percentages in Figures 6.7 and 6.8 draw on the total amount of evidence for each mode and each super-genre

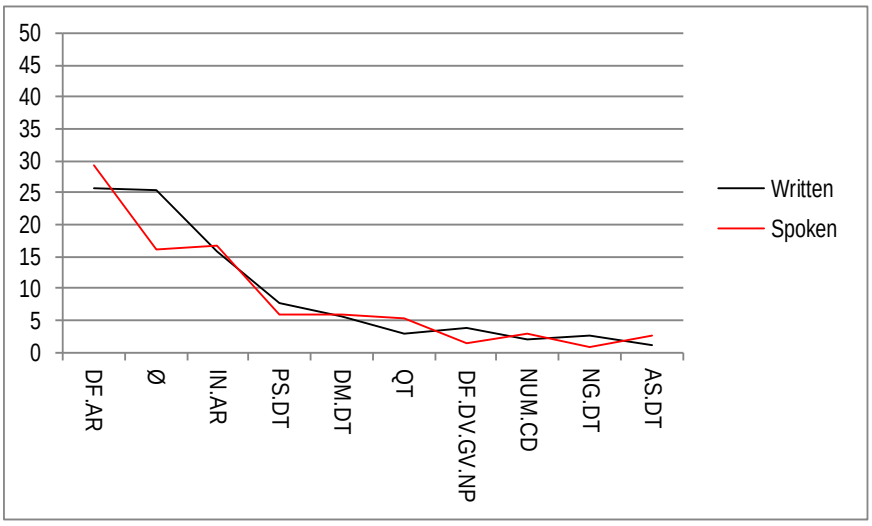
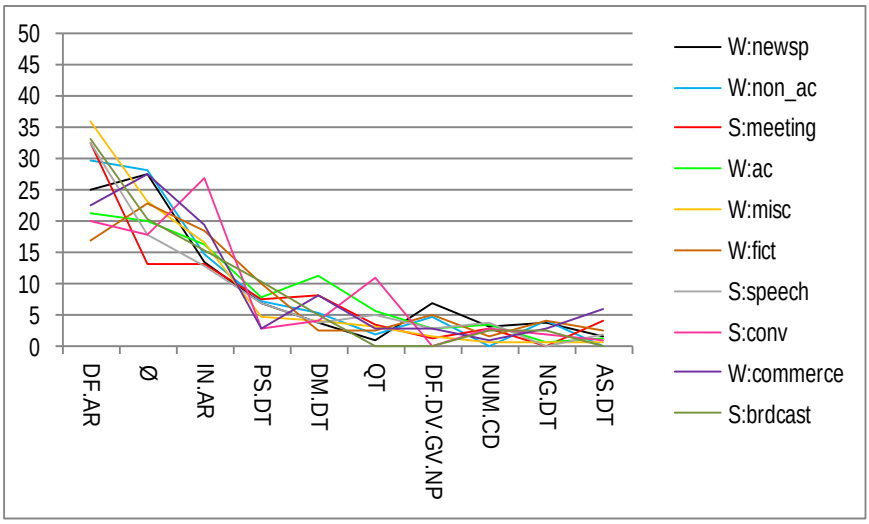


Figure 6.8 Genre distribution of the top ten determiners



The clearest connection in Figure 6.7 is between undetermined heads and written English (cf. also Table 6.14). The distribution of the other categories is more even: spoken English is more primed for the definite article, the indefinite article, demonstrative determiners, quantifiers, cardinal numerals and assertive determiners, and written English shows a slight association with possessive determiners, genitive phrases and negative determiners. Concerning genre, Figure 6.8 reveals three noteworthy findings: the high frequency of the indefinite article and quantifiers in S:conv and of demonstrative determiners in W:ac (see Figure 6.6). Also worth mentioning are the association between the definite article and W:misc, S:brdcast, S:speech and S:meeting, the zero article and W:non_ac, W:newsp and W:commerce (hence its prevalence in the written mode), genitive phrases and W:newsp, W:fict and W:non_ac, and assertive determiners and W:commerce and S:meeting.

As regards semantic premodification, Figures 6.9 and 6.10 below show that the proportion of shell-noun instances with semantic premodifiers (in percentages) is higher in the written mode. This applies particularly to W:newsp, W:ac and W:commerce. Lack of premodification is most common in W:fict, S:conv, S:meeting and S:brdcast.

Figure 6.9 Mode distribution of experiential premodification

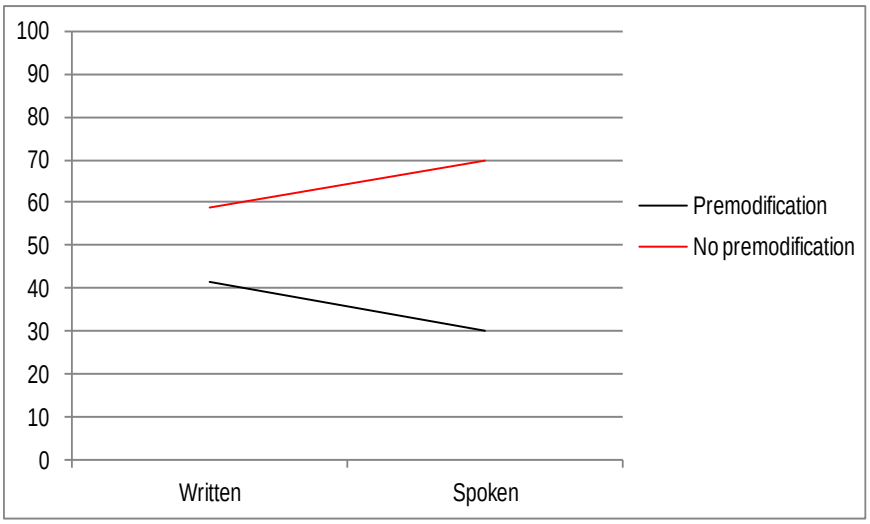
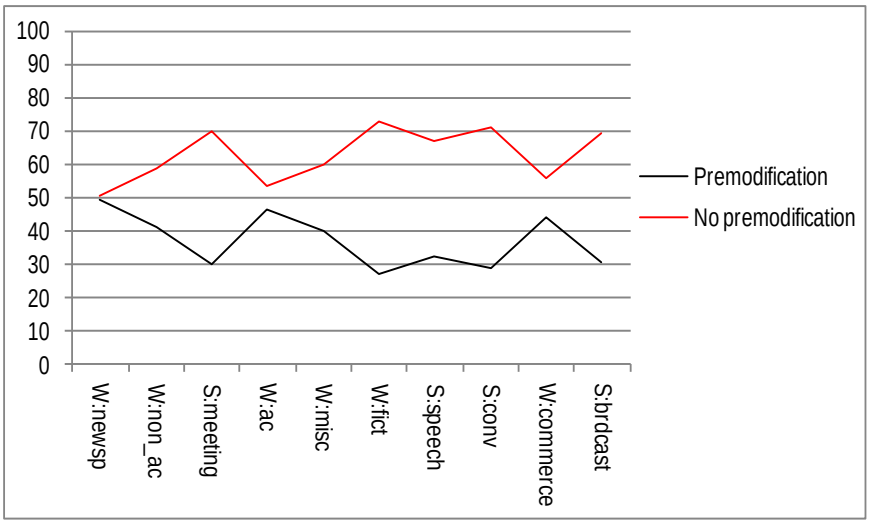


Figure 6.10 Genre distribution of experiential premodification



Figures 6.11 and 6.12 illustrate the distribution of the ten most frequent types of semantic premodification in respect of mode and super-genre.

Figure 6.11 Mode distribution of the top ten experiential premodifiers. Percentages in Figures 6.11 and 6.12 are calculated on the basis of the number of premodified instances in each mode and each super-genre

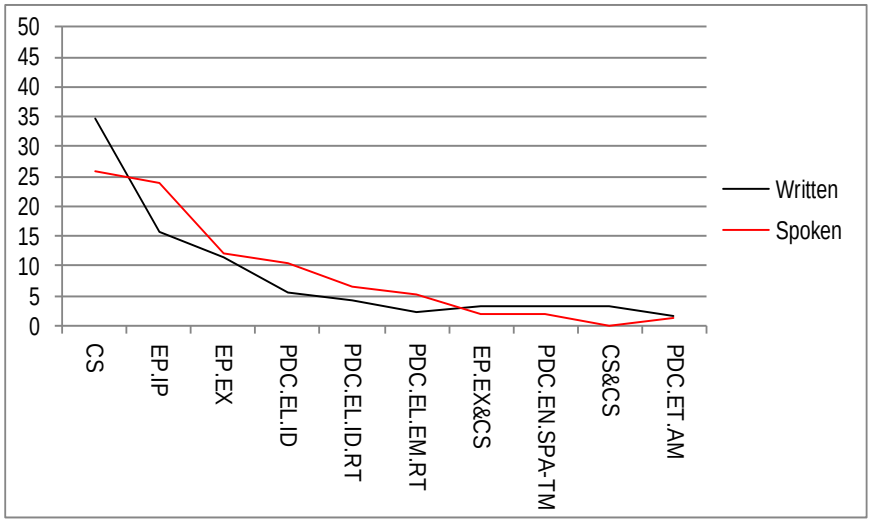
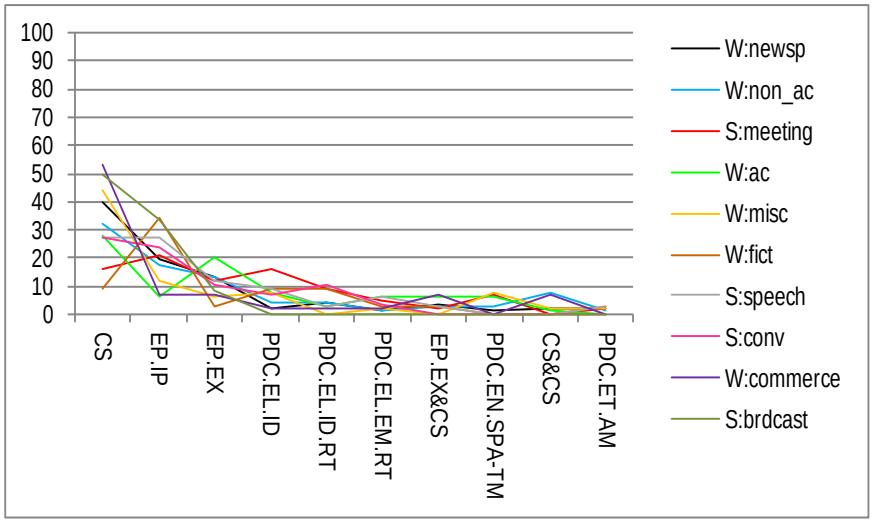


Figure 6.12 Genre distribution of the top ten experiential premodifiers



The most noticeable mode differences in Figure 6.11 concern Classifiers, interpersonal Epithets and elaborating identity post-Deictics. Use of Classifiers (alone or in combination, i.e. EP.EX&CS and CS&CS) prevails in written English, while interpersonal Epithets and elaborating identity post-Deictics do in spoken English. Except for space-time and extending amplifier post-Deictics, the three other types are primed for spoken use. In experiential Epithets and extending amplifier post-Deictics, the difference

between both modes is marginal (0.86% with experiential Epithets and 0.26% with extending amplifier post-Deictics).

In terms of genre distribution, Figure 6.12 displays a number of preferences, the most evident found among Classifiers, interpersonal Epithets, experiential Epithets and elaborating identity post-Deictics. Classifiers are most common in W:commerce (as in (647a)), S:brdcast (as in (647b)), W:misc (as in (647c)) and W:newsp (as in (647d)), hence their prevalence in written English. W:fict, negatively primed for Classifiers, ranks highest with interpersonal Epithets (as in (648a)), followed by S:brdcast (as in (648b)), S:speech (as in (648c)) and S:conv (as in (648d)), which explains their dominance in spoken English. Their use is minimal in W:ac and W:commerce. The associations observed with interpersonal Epithets reverse with experiential Epithets, as W:ac ranks highest (as in (649)), whilst W:fict ranks lowest. Elaborating identity post-Deictics are most clearly primed for S:meeting, as in (650a) (cf. also restrictive identity post-Deictics, where S:meeting is also prevalent, along with S:conv and W:fict, as in (650b) through (650d)). Only slight associations are shown between space-time enhancing post-Deictics, W:misc, S:meeting and W:ac (as in (651a) through (651c)), between combinations of experiential Epithets and Classifiers, W:ac and W:commerce (as in (652a) through (652b)), and between combinations of Classifiers, W:non_ac and W:commerce (as in (653a) through (653b)).

(647) (a) their **employment** practices, the goverment's **fiscal** philosophies
(b) the **Taylor Report** recommendations, **financial** objectives
(c) **Okapi** projects, an **Irish** joke
(d) **political** opposites, a **state** crime

(648) (a) such **terrible** things, no **small** surprise
(b) the **most inopportune** times, **dreadful** experiences
(c) the **strange** phenomenon, the **relevant** points
(d) some **good** things, a couple of **good** works

(649) a **culturally progressive** phenomenon, a **broader-based** system

(650) (a) various **different** areas, all **other** parts of the Board's work
(b) the **same** thing, one **sole** objective
(c) the **only** thing, the **only** problem
(d) his **chief** characteristic, the princes' **only** misfortune

(651) (a) the **following** joke
(b) the **next** examples
(c) the **latter** phenomena

(652) (a) the **hated exile** system
(b) any **categoric policy** recommendation

(653) (a) Stukeley's **historical** and **religious** philosophy
(b) EC **trading** objectives

As to formal postmodification, Figure 6.13 shows a similar proportion of instances containing postmodifying structures in both modes, these being only slightly more frequent in written English. Clearer differences emerge from the genre-related data in Figure 6.14. Frequent occurrence of postmodification stands out in W:misc, S:speech and W:commerce, and infrequent occurrence predominates in S:brdcast, W:fict and S:conv. Interestingly, the latter three super-genres also feature among those with the least amount of premodification (see Figure 6.10).

Figure 6.13 Mode distribution of formal postmodification

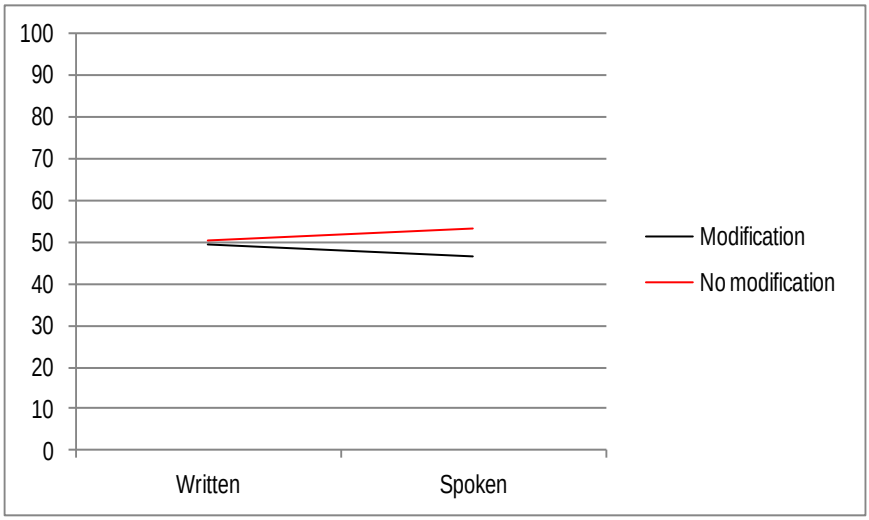
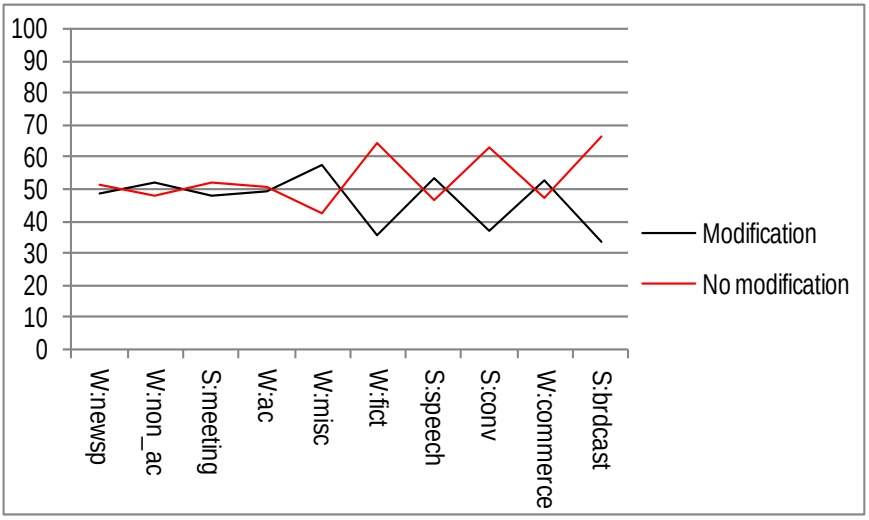


Figure 6.14 Genre distribution of formal postmodification



Figures 6.15 and 6.16 look at the frequency of the ten highest-ranking postmodifying structures in relation to mode and genre.

Figure 6.15 Mode distribution of the top ten postmodifying structures. Percentages in Figures 6.15 and 6.16 are based on the total number of postmodified items in each mode and each super-genre

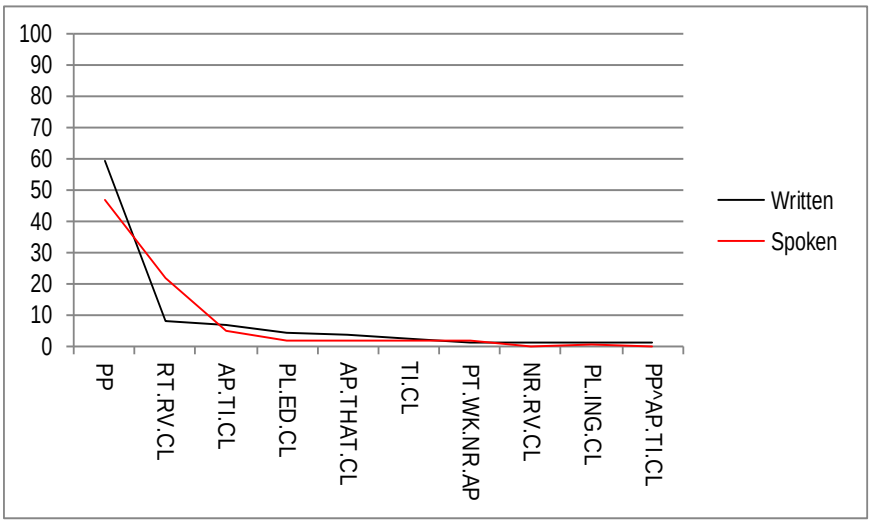
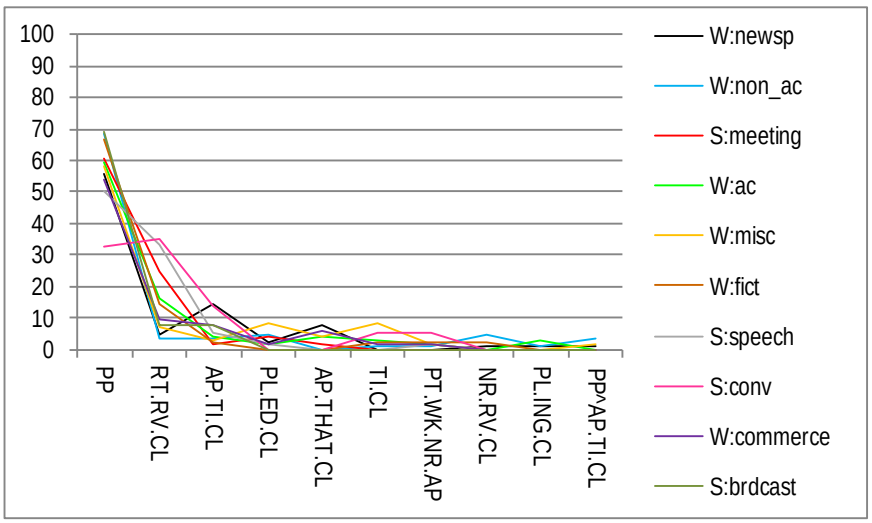


Figure 6.16 Genre distribution of the top ten postmodifying structures



The greatest mode differences in Figure 6.15 lie in the top two postmodifying structures, i.e. prepositional phrases and restrictive relative clauses, the former being more frequent in written English and the latter in

spoken English. This lends support to the above observation about the frequent occurrence of prepositional phrases among written patterns and of relative clauses among spoken patterns (see Table 6.14). The remaining structures are either overlapping or only slightly more frequent in one of the two modes (e.g. participle *-ed* clauses and appositive *that*-clauses in written English).

Regarding genre distribution, of the top two structures, only restrictive relative clauses reveal a distinct priming. Except for S:conv, ranking lowest in prepositional use, most of the other super-genres cluster around the 50%-60% ranges. However, S:brdcast, W:non_ac and W:fict rank slightly higher (as in (654a) through (654c)). Restrictive relative clauses occur most frequently in three spoken super-genres (S:conv, S:speech and S:meeting, as in (655a) through (655c)), and least so in W:non_ac and W:newsp (cf., however, non-restrictive relative clauses, where the only examples of this construction occur in W:non_ac, W:fict and W:newsp). Concerning the other genre categories, appositive *to*-infinitive and *that*-clauses reveal a stronger association with W:newsp, as in (656a) and (656b). Appositive *to*-infinitive clauses are closely followed by S:conv, and appositive *that*-clauses by W:commerce, W:misc and W:ac (S:conv contains no instances of *that*-clauses). In the case of participle clauses, *-ed* clauses are primed for W:misc (as in (657a)), while *-ing* clauses are only minimally so for W:ac, W:non_ac and W:newsp (as in (657b) through (657d)). Lastly, *to*-infinitive clauses prevail in W:misc and S:conv (as in (658a) through (658b)), while partial weak non-restrictive appositives are strongly primed only for S:conv (as in (659)).

(654) (a) the chances of me leaving Glasgow
(b) the myth of Osiris
(c) the opposite of people

(655) (a) the visions that they had
(b) the relevant points that others may have missed
(c) differing work practices which I try to erode

(656) (a) Gen Noriega's capacity to inspire personal loyalty
(b) his exasperated surprise that the Lebanese Forces had not joined the 'legitimate' government

(657) (a) a preventive system aimed at stopping a disability from becoming a handicap
(b) a collaborative venture involving three publishers
(c) part of the work being carried out by IMS
(d) one of the direst warnings, pleading for immediate though unspecified action

(658) (a) the other important characteristic to consider
(b) things to say

(659) an objective, you've got to get it down a hole

6.2.2.3 *Distribution across lemmas and semantic types*

This section draws on the results in 6.2.2.1 in order to identify possible priming differences in the distribution of lemmas across determiners, semantic premodifiers and postmodifying structures. As in Figures 6.7 through 6.16 above, data are provided only for the top ten categories in each case. In the following, frequency-sorted lists of lemmas are given for the top 10 units in each category. For the complete lists (with the 60 lemmas), reference is made to Appendix 8. Each group of tables is accompanied by a figure illustrating behaviour related to the six semantic types of shell nouns.

Tables 6.17a through 6.17c and Figure 6.17 below present the results for the ten most frequent determiners.

Table 6.17a Top ten lemmas for the top ten determiners. Percentages in all the tables in section 6.2.2.3 are based on the evidence obtained for each lemma (usually 40 concordances, but see Table 4.5 in 4.3.2 for instances with fewer concordances). See Appendix 8.1 for the complete lists

DF.AR	%	Ø	%	IN.AR	%	PS.DT	%	DM.DT	%
Opposite	90.00	Foreboding	85.71	Correction	100.00	Recollection	80.00	Phenomenon	27.78
Answer	65.00	Leave	75.00	Venture	44.83	Quest	60.00	Terror	25.00
Challenge	64.71	Anger	60.00	Testimony	42.86	Endorsement	50.00	Area	20.00
Motivation	62.50	Detail	57.50	Joke	40.00	Endeavour	44.44	Finding	17.86
Scandal	61.54	Irony	57.14	Assessment	38.46	Leave	25.00	Misfortune	14.29
Characteristic	55.17	Terror	50.00	Dimension	37.50	Experience	22.50	Proviso	14.29
Contradiction	50.00	Evidence	45.00	Triumph	37.50	Philosophy	18.18	Contradiction	12.50
Myth	43.33	Part	43.75	Endeavour	33.33	Prejudice	18.18	Motivation	12.50
Finding	42.86	Testimony	42.86	Impetus	33.33	System	17.65	Part	12.50
Proviso	42.86	Suspicion	40.00	Example	32.50	Capacity	17.39	Triumph	12.50

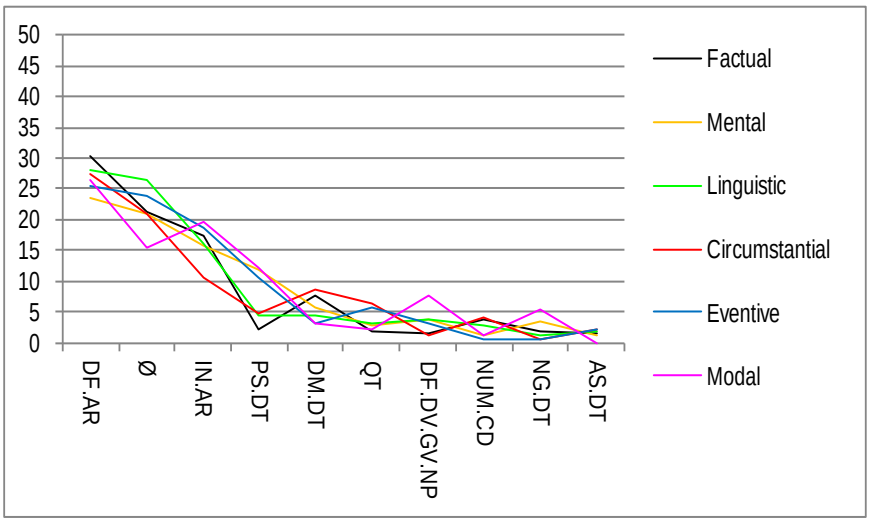
Table 6.17b Top ten lemmas for the top ten determiners

QT	%	DF.DV.GV.NP	%	NUM.CD	%	NG.DT	%	AS.DT	%
Time	15.00	Anger	20.00	Facet	28.57	Surprise	32.00	Facet	14.29
Work	14.81	Quest	20.00	Proviso	14.29	Chance	10.00	Misfortune	14.29
Detail	12.50	Capacity	17.39	Way	10.00	Evidence	10.00	Endeavour	11.11
Way	12.50	Assessment	15.38	Part	9.38	Prejudice	9.09	Suspicion	8.00
Evidence	10.00	Foreboding	14.29	Sense	9.09	Experience	5.00	Part	6.25
Point	10.00	Misfortune	14.29	Example	7.50	Sense	4.55	Area	5.00
Practice	10.00	Testimony	14.29	Thing	7.50	Dimension	4.17	Example	5.00
Prejudice	9.09	Contradiction	12.50	Warning	6.90	Warning	3.45	Point	5.00
Surprise	8.00	Triumph	12.50	Answer	5.00	Myth	3.33	Word	5.00
Scandal	7.69	Vision	10.34	Area	5.00	Part	3.13	Sense	4.55

Table 6.17c Top ten lemmas for the top ten determiners

DM.DT(THIS)	%	DM.DT(THAT)	%	DM.DT(THESE)	%	DM.DT(THOSE)	%
Terror	25.00	Misfortune	14.29	Phenomenon	13.89	Assessment	3.85
Proviso	14.29	Motivation	12.50	Finding	10.71	Part	3.13
Area	12.50	Capacity	8.70	Application	5.00	Experience	2.50
Contradiction	12.50	Part	6.25	Joke	5.00	Anger	0.00
Triumph	12.50	Area	5.00	Point	5.00	Answer	0.00
Phenomenon	11.11	Evidence	5.00	Venture	3.45	Application	0.00
Project	7.50	Point	5.00	Recommendation	3.13	Area	0.00
Finding	7.14	Time	5.00	Area	2.50	Capacity	0.00
Evidence	5.00	Vision	3.45	Experience	2.50	Challenge	0.00
Objective	5.00	Recommendation	3.13	Project	2.50	Chance	0.00

Figure 6.17 General semantic types and the top ten determiners. Percentages in all the figures in section 6.2.2.3 are based on the total number of concordances for each general semantic type (see Table 5.2 in 5.3.5)



Determiner use in Figure 6.17 and in Tables 6.17a through 6.17c reveals slight associations with one or the other semantic type. The definite article is more frequent among factual and linguistic nouns (e.g. *opposite*, *characteristic*, *finding*; *answer*, *contradiction*, *myth*) and least so among mental nouns. *Opposite* is strongly primed for this determiner. At first glance, the zero article is positively primed for uncountable lemmas (e.g. *foreboding*, *leave*, *anger*, *terror*, *evidence*). The occurrence of *foreboding*, *anger*, *terror* and *suspicion* in the top ten seems to further indicate an association with mental nouns. However, Figure 6.17 suggests that the lack of a determiner is, in general, more frequent among linguistic and eventive nouns (e.g. *detail*, *irony*, *application*, *warning*; *crime*, *work*, *venture*, *quest*), and least so among modal nouns. The indefinite article is slightly more prevalent among

modal, eventive and factual nouns (e.g. *failure*, *chance*; *venture*, *triumph*, *endeavour*; *testimony*, *dimension*, *impetus*) and relatively rare among circumstantial nouns. The strong priming shown by *correction* (100%) is due to the scant evidence for this lemma (only 2 concordances).

Possessive determiners appear to be associated with modal, mental and eventive nouns (e.g. *capacity*; *recollection*, *experience*, *philosophy*; *quest*, *endeavour*, *crime*) and to be negatively primed factual nouns. This is not surprising in view of the close link between these nouns and an individual, in that use of *capacity*, *experience* or *endeavour* is often intended to refer to someone's capacity, experience or endeavour. Demonstrative determiners seem primed for circumstantial and factual nouns (e.g. *area*, *proviso*; *phenomenon*, *finding*, *part*) and only minimally for modal and eventive nouns. Concerning demonstrative subtypes, the first position occupied by *terror* and *misfortune* in the *this* and *that* columns of Table 6.17c might be explained by the need to emphasise the negative nature of *terror* (a feeling) and *misfortune* (an event). Use of the definite article would render both nouns more neutral and somehow less threatening than is apparent with *this* or *that*. Circumstantial nouns are also common with quantifiers, even if in this case they are closely followed by eventive nouns (e.g. *time*, *way*, *practice*; *work*, *practice*, *scandal*).

Like possessive determiners, genitive noun phrases are most frequent among modal nouns (e.g. *capacity*, *chance*, *failure*). Mental and linguistic nouns rank second and third, but they cover a wider range of units than is the case with modal nouns (e.g. *anger*, *assessment*, *foreboding*; *testimony*, *contradiction*, *warning*). Circumstantial and factual nouns are negatively primed for this type of determiner (cf. also possessive determiners). Cardinal numerals are common among circumstantial and factual nouns (e.g. *proviso*, *way*, *area*; *facet*, *part*, *example*) and rare among eventive and modal units.The negative determiner *no* seems to stand out among modal and mental nouns (e.g. *chance*; *surprise*, *prejudice*, *experience*) and not to occur with units with circumstantial or eventive meanings. Lastly, the assertive determiner *some* is slightly more frequent among circumstantial, eventive and linguistic nouns (e.g. *area*, *way*; *misfortune*, *endeavour*; *point*, *word*). It does not occur in modal nouns.

Tables 6.18a through 6.18c and Figures 6.18 and 6.19 display the results for the top ten semantic premodifiers in the study sample.

Table 6.18a Top ten lemmas for the top ten experiential premodifiers. Percentages are calculated in terms of the overall evidence for each lemma, including instances without premodification (see Table 6.18c). See Appendix 8.2 for the complete lists

CS	%	EP.IP	%	EP.EX	%	PDC.EL.ID	%	PDC.EL.ID.RT	%
Venture	62.07	Testimony	42.86	Phenomenon	22.22	Dimension	20.83	Misfortune	14.29
Practice	37.50	Endorsement	25.00	Anger	20.00	Facet	14.29	Proviso	14.29
System	35.29	Vision	17.24	Answer	12.50	Misfortune	14.29	Endeavour	11.11
Scandal	30.77	Impetus	16.67	Project	12.50	Part	12.50	Objective	10.00
Project	30.00	Experience	15.00	Venture	10.34	Way	12.50	Thing	10.00
Dimension	29.17	Thing	15.00	Evidence	10.00	Endeavour	11.11	Finding	7.14
Characteristic	27.59	Facet	14.29	Experience	10.00	Area	10.00	Part	6.25
Capacity	26.09	Irony	14.29	Challenge	8.82	Thing	10.00	Area	5.00
Application	25.00	Misfortune	14.29	Capacity	8.70	Example	5.00	Way	5.00
Triumph	25.00	Proviso	14.29	Dimension	8.33	Objective	5.00	Dimension	4.17

Table 6.18b Top ten lemmas for the top ten experiential premodifiers

PDC.EL.EM.RT	%	EP.EX&CS	%	PDC.EN.SPA-TM	%	CS&CS	%	PDC.ET.AM	%
Proviso	14.29	Dimension	12.50	Impetus	16.67	Prejudice	18.18	Contradiction	12.50
Area	7.50	Endeavour	11.11	Proviso	14.29	Philosophy	9.09	Triumph	12.50
Characteristic	6.90	Practice	10.00	Joke	5.00	Dimension	8.33	Philosophy	4.55
Surprise	4.00	Recommendation	9.38	Sense	4.55	Venture	6.90	Capacity	4.35
Assessment	3.85	Philosophy	4.55	Assessment	3.85	System	2.94	Failure	4.35
Finding	3.57	Assessment	3.85	Myth	3.33	Area	2.50	Myth	3.33
Venture	3.45	Work	3.70	Warning	3.33	Example	2.50	Detail	2.50
Warning	3.33	Myth	3.33	Phenomenon	2.78	Objective	2.50	Project	2.50
Challenge	2.94	System	2.94	Area	2.50	Practice	2.50	Anger	0.00
Application	2.50	Anger	0.00	Example	2.50	Anger	0.00	Answer	0.00

Table 6.18c Top and bottom ten lemmas, and the lack of premodification

Ø (%) (top ten)	%	Ø (%) (bottom ten)	%
Foreboding	100.00	Project	45.00
Leave	100.00	Endeavour	44.44
Motivation	100.00	System	44.12
Recollection	100.00	Testimony	42.86
Opposite	88.89	Characteristic	41.38
Detail	87.50	Practice	32.50
Time	87.50	Proviso	28.57
Failure	86.96	Impetus	16.67
Chance	85.00	Dimension	8.33
Suspicion	84.00	Venture	6.90

Figure 6.18 General semantic types and the occurrence of experiential premodification

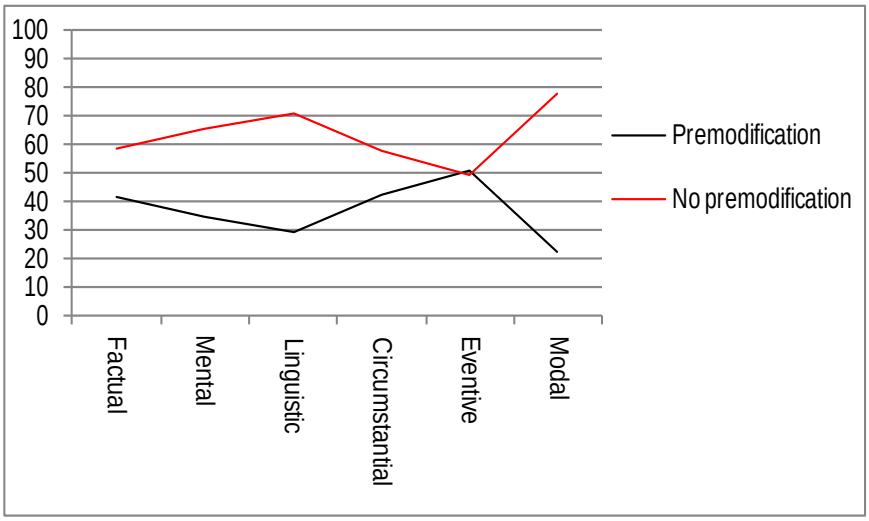
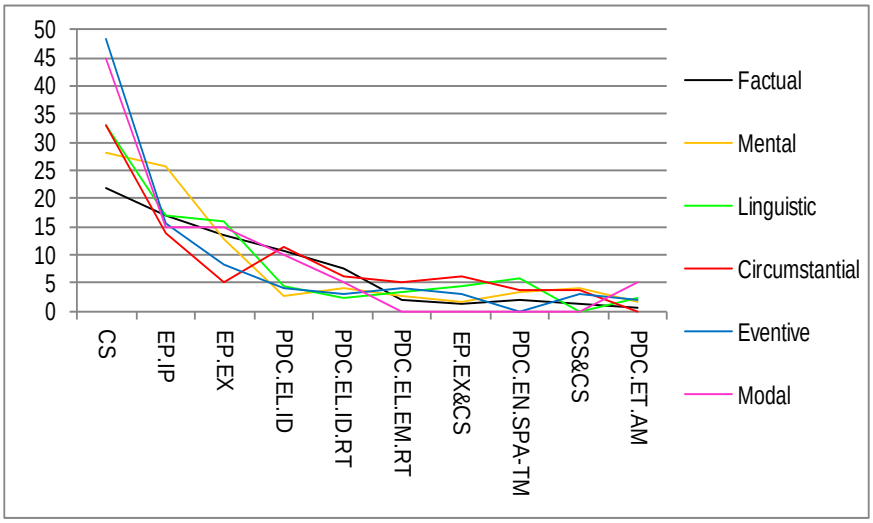


Figure 6.19 General semantic types and the top ten experiential premodifiers. The data in this figure rest on the amount of semantic premodification in each semantic type (see Figure 6.18)[33]



[33] This was intended to compensate for the small amount of semantic premodification in the study sample (37.99% vs. 62.61%, see above), and to make the results easier to interpret.

Semantic premodification, as shown in Figure 6.18, prevails among eventive, circumstantial and factual nouns, but ranks low among modal and linguistic nouns (see Table 6.18c for the ten lemmas with the least frequent semantic premodification, i.e. from *foreboding* to *suspicion*, and the ten lemmas where it is most frequent, i.e. from *project* to *venture*). Specifically, Figure 6.19 brings to light a number of preferences in the instantiation of individual semantic premodifiers across shell-noun semantic types. Paramount among these preferences are those concerning Classifiers and interpersonal Epithets. Classifiers occur more frequently in eventive and modal nouns (as in (660a) and (660b)), and least so in factual nouns (cf. CS&CS, showing a preference for mental nouns, as in (660c)). The dominance of eventive nouns in Classifier use may be due exclusively to the strong priming of *venture* for this type of semantic premodifier (62.07%). As to interpersonal Epithets, these stand out among mental nouns, as in (661a). The strong priming of *testimony* for interpersonal Epithets (42.86%) may be explained by the small amount of evidence for this lemma (only 7 concordances and 3 instances of EP.IP; see (661b)).

(660) (a) a business venture, an ecumenical venture, a teaching practice, religious practices, the latest political scandal, the last government scandal
(b) an advisory capacity, industrial and technological capacities, a mathematically calculable chance, a 40m chance, collection failure, the clear intelligence failure
(c) their Christian European prejudices, Home Office and Department of Environent philosophy

(661) (a) a pleasant vision, an attractive vision, a good experience, a bitterly disappointing experience, a lovely surprise, his exasperated surprise
(b) a sad testimony to a town in decline, key testimony, the pastor's impotent testimony

Regarding the other premodifiers, experiential Epithets seem to associate with lemmas with linguistic, modal and factual meanings, as in (662a) through (662c). They are rare with circumstantial nouns. The combination of experiential Epithet and Classifier is more frequent among circumstantial and linguistic nouns (as in (663a) and (663b)), and it is absent from modal nouns.

(662) (a) the uncompromising answer, the short answer, the correct word, one new word, a veiled warning, recent warnings
(b) new capacity to produce, the new capacity to create stability

(c) a temporary phenomenon, a culturally progressive phenomenon, this romantic dimension, a new dimension, the simple example of shoes, all these random examples

(663) (a) traditional management practices
(b) any categoric policy recommendation

Among the four post-Deictic types in the top ten units, circumstantial nouns are frequent in all but PDC.ET.AM, where no circumstantial instances occur (i.e. PDC.EL.ID, PDC.EL.ID.RT, PDC.EL.EM.RT and PDC.EN.SPA-TM, as in (664a) through (664d)). As is evident in the examples, the high frequency of circumstantial nouns is motivated by two lemmas in particular, *proviso* and *area*. Factual and modal nouns also occur frequently with PDC.EL.ID and PDC.EL.ID.RT (as in (665a) and (665b)), while eventive and linguistic nouns show stronger priming for PDC.EL.EM.RT and PDC.SPA-TM (as in (666a) and (666b)). PDC.ET.AM is frequent with modal nouns, as in (667). The least frequent semantic types in each case are mental nouns with PDC.EL.ID, linguistic nouns with PDC.EL.ID.RT, modal nouns with PDC.EL.EM.RT, eventive and modal nouns with PDC.EN.SPA-TM and circumstantial and factual nouns with PDC.ET.AM.

(664) (a) the other way in which you can get more close to live Amnesty, any other area of planning
(b) the only proviso that it is only permissible to those who can do without it, the main area of concern, the only way to have berries at Christmas
(c) certain provisos, that particular area
(d) this last proviso, the following areas

(665) (a) a further dimension, one other facet of rural community change, the main findings of the initial destinations survey, the only thing
(b) another chance, the same chances of being appointed

(666) (a) this particular venture
(b) all of the above four points, no advance warning

(667) our total capacity to produce, total failures

Tables 6.19a through 6.19c and Figures 6.20 and 6.21 focus on the top ten postmodifying structures in the study sample.

Table 6.19a Top ten lemmas for the top ten postmodifying structures. Percentages are calculated in terms of the overall evidence for each lemma (with and without postmodification; see Table 6.19c for the latter). See Appendix 8.3 for the complete lists

PP	%	RT.RV.CL	%	AP.TI.CL	%	PL.ED.CL	%	AP.THAT.CL	%
Endorsement	100.00	Time	40.00	Capacity	52.17	Anger	20.00	Proviso	28.57
Part	81.25	Thing	25.00	Endeavour	44.44	Testimony	14.29	Suspicion	28.00
Opposite	80.00	Challenge	20.59	Failure	39.13	Challenge	11.76	Evidence	15.00
Motivation	75.00	System	20.59	Chance	25.00	System	5.88	Warning	13.79
Assessment	65.38	Area	20.00	Leave	25.00	Objective	5.00	Philosophy	9.09
Quest	60.00	Impetus	16.67	Misfortune	14.29	Problem	5.00	Sense	4.55
Recollection	60.00	Point	12.50	Motivation	12.50	Capacity	4.35	Surprise	4.00
Characteristic	58.62	Way	12.50	Application	5.00	Assessment	3.85	Myth	3.33
Facet	57.14	Word	12.50	Way	5.00	Work	3.70	Anger	0.00
Testimony	57.14	Problem	10.00	Sense	4.55	Characteristic	3.45	Answer	0.00

Table 6.19b Top ten lemmas for the top ten postmodifying structures

TI.CL	%	PT.WK.NR.AP	%	NR.RV.CL	%	PL.ING.CL	%	PP^AP.TI.CL	%
Way	12.50	Proviso	14.29	Experience	5.00	Scandal	7.69	Failure	21.74
Prejudice	9.09	Prejudice	9.09	System	2.94	Work	3.70	Capacity	4.35
Evidence	5.00	Objective	5.00	Chance	2.50	Venture	3.45	Anger	0.00
Joke	5.00	Surprise	4.00	Crime	2.50	Warning	3.45	Answer	0.00
Work	3.70	Chance	2.50	Example	2.50	Practice	2.50	Application	0.00
Characteristic	3.45	Example	2.50	Anger	0.00	Project	2.50	Area	0.00
Part	3.13	Joke	2.50	Answer	0.00	Anger	0.00	Assessment	0.00
Phenomenon	2.78	Thing	2.50	Application	0.00	Answer	0.00	Challenge	0.00
Application	2.50	Anger	0.00	Area	0.00	Application	0.00	Chance	0.00
Point	2.50	Answer	0.00	Assessment	0.00	Area	0.00	Characteristic	0.00

Table 6.19c Top and bottom ten lemmas, and the lack of postmodification

Ø (%) (top ten)	%	Ø (%) (bottom ten)	%
Foreboding	100.00	Characteristic	27.59
Irony	85.71	Way	27.50
Misfortune	85.71	Assessment	26.92
Joke	85.00	Chance	22.50
Surprise	84.00	Capacity	21.74
Prejudice	81.82	Opposite	20.00
Phenomenon	80.56	Facet	14.29
Anger	80.00	Motivation	12.50
Dimension	79.17	Part	6.25
Answer	75.00	Endorsement	0.00

Figure 6.20 General semantic types and the occurrence of postmodification

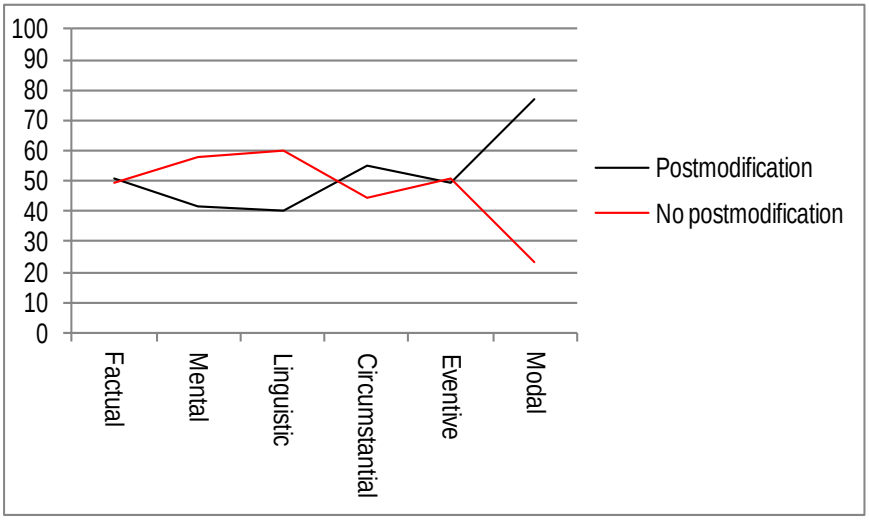


Figure 6.21 General semantic types and the top ten postmodifying structures. Percentages are based on the number of occurrences with postmodification in each semantic type

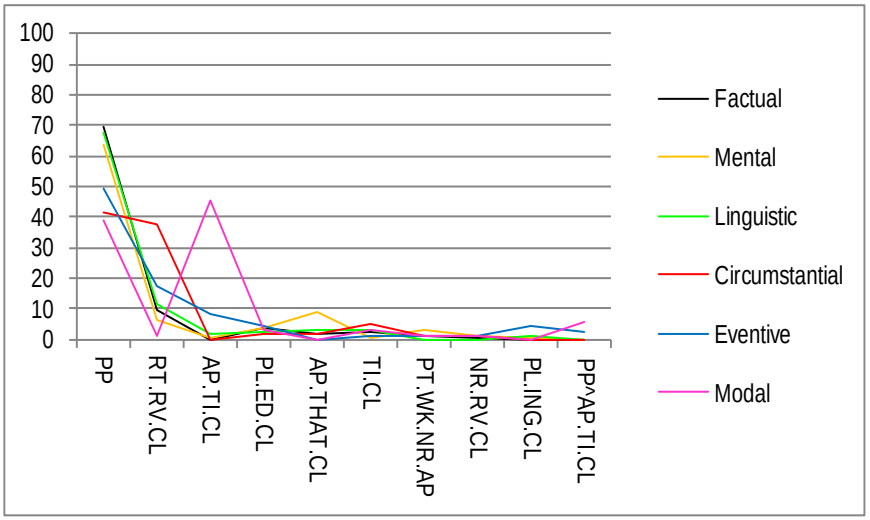


Figure 6.20 indicates a greater use of structural postmodification among modal and circumstantial nouns, while the lack thereof is most evident among linguistic and mental units (*irony*, *joke*, *answer*, *foreboding*, *surprise*, *prejudice* and *anger* in Table 6.19c).

The ten bottom lemmas in Table 19c appear to require postnuclear structures most. Most prevalent among these are *characteristic*, *opposite*, *facet* and *part*, which Keizer (2007: 64) terms 'relational nouns'. Their

relational meaning stems from their frequent association with other discourse entities expressed in following *of*-phrases. Thus, when reference is made to a *characteristic*, the *opposite*, a *facet* and a *part*, it is usually in relation to something else (e.g. *a prominent characteristic of the STV*, *facets of the divinity*). Certain other nouns in the list (i.e. *asssessment* and *endorsement*) are also frequent with *of*-phrases, but these are derived nouns, i.e. nominalised entities requiring the transformation of verbal arguments into prepositional *of*-phrases (e.g. *an assessment of its prospects for the future*, *his endorsement of perestroika* < *its prospects for the future are assessed* and *perestroika is endorsed*; see 4.4.2.9). In the case of *way*, *chance*, *capacity* and *motivation*, use of postnuclear structures is split between appositive *to*-infinitive clauses (most common with *chance* and *capacity*) and prepositional phrases.

Figure 6.21 shows the distribution of postmodifying structures across semantic types. The most remarkable findings concern prepositional phrases, restrictive relative clauses and appositive *to*-infinitive clauses. Prepositional phrases are strongly primed for factual, linguistic and mental nouns (e.g. *part*, *opposite*, *characteristic*, *facet*, *testimony*; *endorsement*, *correction*, *detail*, *word*; *motivation*, *assessment*, *recollection*, *sense*), and negatively so for modal nouns. Restrictive relative clauses are dominant among circumstantial nouns (e.g. *time*, *system*, *area*, *way*), and rare among modal nouns. *Time* is strongly primed for this structure (40.00%) on the grounds of the frequent instantiation of its shell meaning (i.e. 'occasion') through *the time that* or *the time when something happened.* Appositive *to*-infinitive clauses show an association with modal nouns (e.g. *capacity*, *failure*, *chance*, *leave*), and are absent from factual and circumstantial nouns and almost so from mental nouns (see also PP^AP.TI.CL, whose use is restricted to two nouns, *failure* and *capacity*, the latter always modal and the former featuring modal and eventive senses). The occurrence of *endeavour* as the second most frequent lemma (44.44%) is in line with the second position that eventive nouns occupy in Figure 6.21 (e.g. *endeavour*, *venture*). The association of appositive *to*-infinitive clauses with modal and eventive nouns is reversed with appositive *that*-clauses, where mental and linguistic nouns prevail (e.g. *suspicion*, *philosophy*, *sense*, *surprise*; *warning*, *myth*).

With respect to the other structures, participle *-ed* clauses are slightly more frequent among eventive and factual nouns (e.g. *work*, *venture*; *problem*, *characteristic*). Participle *-ing* clauses are primed for eventive lemmas, and absent or almost so from the other semantic types (e.g. *scandal*, *work*, *venture*). Lastly, *to*-infinitive clauses, partial weak non-restrictive appositives and non-restrictive relative clauses show a minimal

association with circumstantial, mental and modal nouns (e.g. *way*, *time*; *prejudice*, *surprise*; *chance*).

6.2.2.4 *Discussion*

The results presented above do not seem to confirm the often-claimed link between shell-noun behaviour and specific Deictics, inasmuch as slightly fewer than half of the instances in the study sample occur with these Deictics (44.09%, see Table 6.9). The definite article ranks highest (27.02%), in accordance with the dominance reported in Halliday & Hasan (1976: 275) and Francis (1986: 27). By contrast, demonstrative determiners are not as prevalent as claimed in the literature: they rank fifth in Table 6.9 (5.74%) and are outranked by the zero and indefinite articles (22.11%, 16.10%). This runs against Schmid's (2000: 25) claim that '[...] indefinite noun phrases do not create as strong conceptual boundaries as the definite noun phrases in which shell nouns tend to occur'. In terms of distribution, the prevalence of the indefinite article and quantifiers in conversation and that of demonstrative determiners in academic prose appear as the most remarkable connections. Also worth emphasising is the greater frequency shown by the definite article, demonstrative determiners and cardinal numerals with factual and circumstantial nouns, and by possessive determiners with modal, mental and eventive nouns. The latter, as argued in 6.2.2.3, may be due to the human component of such nouns as *capacity*, *recollection*, *quest*, *endeavour* or *experience*, where the possessive determiner often corresponds to a human subject (e.g. *they endeavour to*, *they recollect, they are able to*, etc.). The factual and circumstantial uses of the definite article, demonstrative determiners and cardinal numerals, however, rest precisely on the lack of such a human component in non-derived factual and circumstantial instances like *opposite*, *characteristic*, *area* or *proviso*. These, being non-derived, are also more object-like in nature and more open to quantification through cardinal numerals (e.g. *there are three characteristics*, *two areas*, etc.).

A comparison of some of the above findings with Biber et al.'s (1999) shows considerable similarities with nouns in general. The definite article is, as expected, the most frequent specific determiner, followed by possessive and demonstrative determiners, the latter being '[...] far less frequent than the definite article' (Biber et al. 1999: 270). This is also evident in the study sample, where 27.02% of determiners correspond to the definite article, whilst 7.05% and 5.74% correspond to possessive and demonstrative determiners. These proportions resemble those found with noun use in BNC*web* (23.69% for nouns with the definite article, 9.68% for the indefinite

article, 5.14% for possessive determiners and 3.85% for demonstrative determiners).

Concerning mode distribution, according to Biber et al. (1999: 267), the definite article is most frequent in written English, where it occurs twice as frequently as the indefinite article. *BNCweb* lends support to this finding, as the definite article comprises 23.58% of noun use in written English, compared with only 9.42% for the indefinite article. Paradoxically, the mode differences for the definite and the indefinite articles are minimal in the sample, both being slightly more common in spoken genres (29.43% vs. 25.69% for the definite article; 16.76% vs. 15.84% for the indefinite article).

In terms of genre distribution, Biber et al. (1999: 270) state that possessive and demonstrative determiners prevail in fiction and academic prose. Figure 6.8 confirms this, although use of possessive determiners is very close both in fiction and broadcast programmes (10.17% vs. 10.26%). The pre-eminence of these two super-genres is unsurprising given their frequent priming for mental nouns (see Figure 6.3; as pointed out above, mental nouns are also strongly primed for possessive determiners, e.g. *my recollection of the last meeting*, *her experience*, *our philosophy*). With respect to conversation, Biber et al. (1999: 267) discover only a minor difference in the use of both articles, *the* featuring slighly more frequently than *a/an*. This is supported by *BNCweb* data, where 21.84% of noun use is with the definite article and 15.27% with the indefinite article. However, the sample shows a strong association with the indefinite article in conversation (27% for the indefinite article vs. 20% for the definite article).

Turning now to semantic premodification, the evidence in the study sample is consistent with Biber et al.'s (1999: 589) finding about the dominance of premodifiers in written genres, particularly in newspapers and academic prose (see Figures 6.9 and 6.10). As shown in Figure 6.18, premodifiers are most common with eventive, factual and mental nouns (50.53%, 41.16%, 34.82%), and comparatively less frequent with modal nouns (21.98%). Following Schmid (2000: 322), where modal nouns feature as '[...] bad carriers for evaluations', this might be due to the inherent ability of eventive, factual and mental nouns to describe events and states of affairs, as opposed to the focus of modal nouns on the range of (epistemic and deontic) perspectives from which such descriptions can be made (hence their uncommon priming for descriptive or evaluative adjectives).

In relation to the three main semantic premodifiers, there is a strong association between non-specific Deictics and Epithets, and specific demonstrative Deictics and post-Deictics (see Figure 6.4). Drawing on Yamasaki (2008: 79–81) and Schmid (2000: 309), Epithets favour non-specific Deictics, given that, as shown in Figure 6.45, these occur most frequently with shell instances in attributive positions, as in *it may be a*

lovely surprise. According to the two references above, the attributive position is the most explicitly evaluative, as it favours the inclusion of evaluative or descriptive detail through premodification or through nouns which are attitudinal or evaluative themselves, as in *that is a problem that exists.* The priming of post-Deictics for specific Deictics (mainly *the*) is also syntactically motivated. Many of the post-Deictics in the sample are restrictive in nature (e.g. *only*, *particular*, *same*, *main*, *chief*), laying emphasis on subject shell nouns prior to their intrasentential cataphoric lexicalisation. From Schmid's (2000: 334–5) perspective, the shell nouns in these cases perform a focusing function which draws attention to the importance of the adjacent information. As also noted by Schmid (2000: 334), attitudinally neutral nouns prevail in these, e.g. *thing*, *facet*, *part*, *way*, *objective* or *area*, and are particularly primed for elaborating identity (restrictive and non-restrictive) post-Deictics in subject positions (subject shell nouns being primed for the definite article; see Figure 6.42), as in *the only way of contacting DHSS was either to go down or to walk all round* or *the main thing is perhaps the fact that we are now having a fresh look, without blinkers, at ourselves and our country*. Factual and circumstantial nouns are prominent in these cases (see Figure 6.19).

Concerning the genre distribution of premodifying elements, the most relevant data are those regarding the higher frequency of Classifiers in written genres (more specifically, W:commerce, W:misc and W:newsp), of interpersonal Epithets in W:fict, S:speech and S:conv, and of experiential Epithets in W:ac. This is in line with Biber et al. (1999: 510–1), who shows that evaluative descriptors (i.e. interpersonal Epithets) predominate in fiction and conversation, whilst Classifiers prefer academic prose and newspapers. It should be noted that academic prose is relatively infrequent with Classifiers in the study sample, favouring instead experiential Epithets. At any rate, whilst subjective premodifiers are more primed for genres with a strong emotive component, where personal opinion is freely expressed (fiction, conversation, speeches), objective premodifiers (subsuming Classifiers and experiential Epithets) are more likely to occur in genres where evaluation and opinion are often sacrificed for the sake of descriptive objectivity (e.g. economic treatises and textbooks, academic prose, newspapers). It is precisely the emotive connotations of interpersonal Epithets that may also explain their association with mental nouns like *vision*, *experience*, *terror*, *surprise*, *suspicion* (cf. also Figure 6.3, where mental and linguistic nouns rank highest with fiction).

Finally, with regard to postmodifying structures, the dominant position that noun complement clauses occupy in the literature is not confirmed here. This agrees with the references based on manual methods (Winter 1992: 157; Flowerdew 2003a: 337 and 2006: 358; Aktas & Cortes 2008: 10 and

Caldwell 2009: 176). Thus, Caldwell (2009: 176) argues that *of*-phrases are '[...] by far the most prolific of all the patterns extracted as potential shell noun "hosts"'. This is also in the sample's data, where prepositional phrases make up 59.83% of all postmodifying structures, and *of*-phrases account for 59.82% of all prepositional instances. This is not specific to shell nouns: according to Biber et al. (1999: 606, 635), prepositional phrases comprise 65-80% of all postmodifying structures, and *of is* the most frequent alternative (60-65%). Noun complement clauses, by contrast, '[...] are only moderately common' (Biber et al. 1999: 647), because it is adjectives and not nouns that seem to prevail with complement clauses (cf. also Biber 2006: 103, where it is argued that both adjectives and nouns '[...] are generally rare controlling a complement clause').

The genre distribution of prepositional phrases is similar in all super-genres except in conversation, where they rank lowest (cf. also Biber et al. 1999: 606). Academic prose, prevalent with prepositional phrases according to Biber et al. (1999: 606), ranks fifth among the top ten genres in the study sample after S:brdcast, W:non_ac, W:fict and S:meeting. Restrictive relative clauses are second in the analysis database (13.68%), and reveal a strong association with S:conv, S:speech and S:meeting. Again, this is against Biber et al. (1999: 606): in their corpus, relative clauses prevail in fiction. An explanation for this divergence lies in the nature of the units with the strongest priming for this structure (see Table 6.19a). Nouns like *time*, *thing*, *system*, *area*, *point* or *way* are frequent in spoken super-genres (see Tables 6.7a and 6.7b), and are often used in combination with relative clauses, as these help to restrict their highly general semantic scope prior to their full lexical realisation in the preceding or subsequent discourse (e.g. *the last time you had your gas boiler* serviced, *the only thing that worries me*, *the area that as a writer I'm particulary drawn to*). Appositive or noun complement clauses in the sample show a clearer association with W:newsp, followed by S:conv in the case of *to*-infinitive clauses and by W:commerce, W:misc and W:ac in the case of *that*-clauses. The prevalence of academic prose and newspapers with noun complement clauses is also captured by Biber et al. (1999: 647), where *to*-clauses rank highest with newspapers, and *that*-clauses, with academic prose. Remarkably, whilst conversation disfavours appositive clauses in Biber et al. (1999: 647), use of *to*-infinitive instances in the study sample is almost equally distributed in newspapers and conversation. This appears to be lexically related to *chance*, which occurs among the highest-ranking conversation shell nouns (see Table 6.7b), as well as among the units with the strongest priming for this structure (see Table 6.19a; e.g. *I might even get a chance to try one of these days*).

In terms of lemma distribution, the association between prepositional phrases and factual, linguistic and mental nouns matches the marked influence of *of*-prepositional phrases on relational factual nouns and nominalised mental and linguistic instances (cf. Keizer 2007: 64 in 6.2.2.3; e.g. *part*, *opposite*, *asessment*, *recollection*). The connection between *that*-complement clauses and mental and linguistic nouns, and between *to*-infinitive-complement clauses and modal and eventive nouns are also logical. Schmid (2007), as discussed 3.2.5.2, offers revealing insights into the close correspondence between *that*-clauses and factual, mental and linguistic nouns, and *to*-infinitive clauses and eventive and modal nouns.

6.2.3 *Syntactico-semantic functions and textual positions*

6.2.3.1 *Overall distribution*

This and the following section (6.2.4) shift the focus from the internal structure of shell-noun phrases to their role in discourse. Section 6.2.3 is concerned with the syntactic, semantic and textual functions performed by shell nouns at clause level and 6.2.4 looks at the wider discourse context (encapsulation and antecedent). Three variables are at issue here: syntactic function, participant type and Theme-Rheme.

Tables 6.20 through 6.22 present the results for syntactic function. Table 6.20 is the most fine-grained, because it distinguishes between clause- and phrase-level functions, the latter being instances where the noun is complement of a preposition. Given the size of the list obtained (180 categories), Table 6.20 includes only the top ten functions, which account for 74.02% of the overall evidence (i.e. 1071 instances). Unlike Table 6.8, which shows the top 50%, the choice of the top ten in this case is motivated by the concentration of almost half of the instances in only two functions: DO and SB (47.62%). Table 6.21 is the most general, as it disregards the clause-phrase level distinction by merging all instances of CP into the syntactic functions of the prepositional phrases where shell nouns are phrase-level constituents. For example, CP.*IN*(<AB.AT.SPA), CP.*AT*(<AB.AT.SPA), CP.*ON*(<AB.AT.SPA) are conflated under AB.AT.SPA. Table 6.22 lists all instances of CP in the sample.

Table 6.20 Top ten syntactic functions in the study sample (clause- and phrase-level). See Appendix 9 for the complete list

Syntactic function (specific: top ten) %			
DO	27.16	NA	1.94
SB	20.46	CP.*OF*(<DO)	1.66
SCL	12.23	CP.*IN*(<AB.AT.SPA)	1.38
SB(not)	4.22	CP.*OF*(<SB)	1.17
PO	2.76	CP.*IN*(<AB.AT.RS)	1.04

Table 6.21 Syntactic functions in the study sample (complete list: only clause-level)

Syntactic function (general) %			
DO	31.17	AB.AT.RE	0.28
SB	23.50	AP.NR&SCL	0.28
SCL	14.51	AB.AT.CG.CI	0.28
SB(not)	4.63	AP.NR&SB	0.21
AB.AT.RS	4.01	OCL	0.21
PO	3.11	AB.DJ.CO	0.14
AB.AT.TM	2.70	AB.DJ.SY.ML.CM	0.14
NA	2.42	AP.NR&AB.AT.AG	0.14
AB.AT.SPA	2.35	PD.AT.MAT	0.14
AB.AT.MAN	1.52	SB.AT	0.14
AB.AT.MS	1.45	IO	0.14
AB.AT.CAU	1.17	AB.AT.RU	0.07
AB.AT.PU	1.04	AB.DJ.CM.CL	0.07
AB.AT.AG	0.83	AP.NR&AB.AT.SPA	0.07
PSCL	0.76	AP.NR&AB.AT.TM	0.07
POCL	0.62	AP.NR&DO&AB.AT.RS	0.07
AB.AT.CC	0.48	AP.NR&IO	0.07
AP.NR&DO	0.41	SCL&DO	0.07
AP.NR&PO	0.35	SJ.VI	0.07
AB.AT.MAN.RL	0.28	AB.AT.AC	0.07

Table 6.22 Realisations of the prepositional complements in the study sample. Percentages are based on the total number of CP tokens in the sample (i.e. 425)

Prepositional complements			
<CP.*OF*>	22.59	<CP.*FOLLOWING*>	0.47
<CP.*IN*>	18.59	<CP.*INCLUDING*>	0.47
<CP.*WITH*>	10.59	<CP.*LIKE*>	0.47
<CP.*FOR*>	7.76	<CP.*THROUGH*>	0.47
<CP.*TO*>	6.12	<CP.*UNTIL*>	0.47
<CP.*ON*>	5.18	<CP.*UPON*>	0.47
<CP.*BY*>	4.47	<CP.*WITHIN*>	0.47
<CP.*AS*>	4.00	<CP.*WITHOUT*>	0.47
<CP.*FROM*>	4.00	<CP.*ACROSS*>	0.24
<CP.*AT*>	3.29	<CP.*BEHIND*>	0.24
<CP.*ABOUT*>	1.65	<CP.*CONTRARY TO*>	0.24
<CP.*INTO*>	1.18	<CP.*DUE TO*>	0.24
<CP.*BECAUSE OF*>	0.94	<CP.*IN RESPECT OF*>	0.24

<CP.*DESPITE*>	0.94	<CP.*OVER*>	0.24
<CP.*AFTER*>	0.71	<CP.*PLUS*>	0.24
<CP.*BETWEEN*>	0.71	<CP.*UNLIKE*>	0.24
<CP.*DURING*>	0.71	<CP.*WITH REGARD TO*>	0.24
<CP.*AGAINST*>	0.47	<CP.*CONCERNING*>	0.24

Overall, clause-level predominate over phrase-level positions. The former comprise 70.63% of the evidence, and the latter 29.37%. Concerning realisation, 51.77% of all cases of CP correspond to three prepositions: *of*, *in* and *with* (cf. Table 6.11 for similar frequencies of postmodifying prepositions in shell-noun phrases). Concerning clause-level functions, direct object, subject and subject complement make up 59.85% of the evidence in Table 6.20 and 69.18% in Table 6.21. Table 6.21 reveals a preference for adverbial adjuncts of respect (4.01%), time (2.70%) and space (2.35%).

Table 6.23 and Figures 6.22a through 6.22b go into greater detail about the distribution of clause-level functions through the examination of participant types. As with Table 6.21, Table 6.23 gives percentages for Transitivity participants and circumstances, irrespective of clause- or phrase-level positions. Figures 6.22a and 6.22b bring together all categories in Table 6.23 under their corresponding processses. The difference between both figures lies in the distinction that the former makes between the three subtypes of relational processes (intensive, circumstantial and possessive).

Table 6.23 Participant types in the study sample

Participant type (detailed) (%)			
IDR (rel.in)	9.61	CR.CAU.PU	1.04
ATT (rel.in)	8.36	CR.MAN.QL	1.04
GOA	7.39	ATT (rel.ci)	0.97
IDD (rel.in)	7.33	CAR (rel.po)	0.62
VER	6.15	CR.RL.GS	0.62
EXI	5.46	CR.CG.CC	0.48
CAR (rel.in)	5.39	IDD (rel.po)	0.48
CR.MAT	5.11	CR.CG.CI	0.28
PHE	4.98	CR.MAN.CV	0.28
ATT (rel.po)	4.49	TR	0.28
IDR (rel.ci)	4.15	CM.AT	0.21
SCO	3.73	CR.RL.PCT	0.21
CR.LOC.TM	2.76	CR.MAN.DG	0.14
PHE (beh)	2.70	SCO (men)	0.14
NA	2.42	ATR (rel.in)	0.07
IDD (rel.ci)	2.28	ATT(rel.in)&GOA	0.07
ACT	1.73	CR.AC.AD	0.07
CR.LOC.PC	1.66	CR.AN.VI	0.07
CAR (rel.ci)	1.45	CR.CAU.BF	0.07

CR.MAN.MS	1.45	GOA&PHE	0.07
CR.CAU.RE	1.38	INI	0.07
IDR (rel.po)	1.31	PHE&CR.MAT	0.07
VER (beh)	1.31	SEN (met)	0.07

Figure 6.22a Process types in the study sample (with relational subtypes)

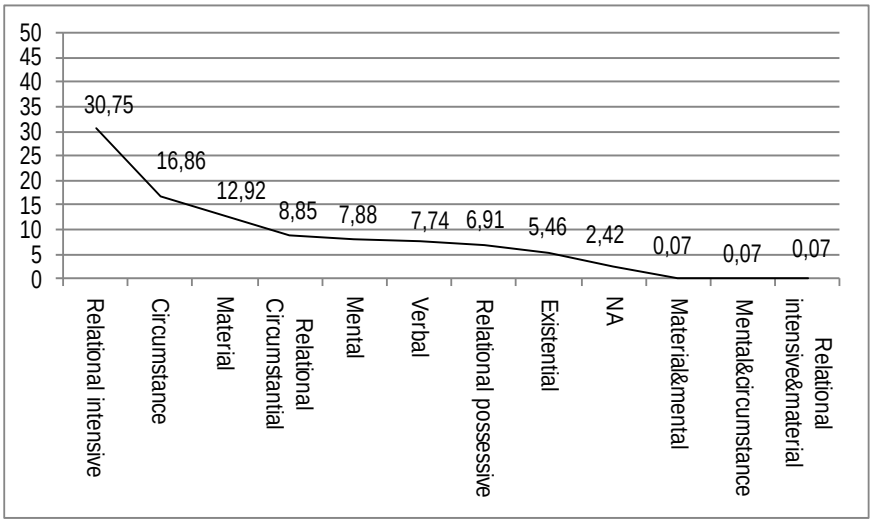
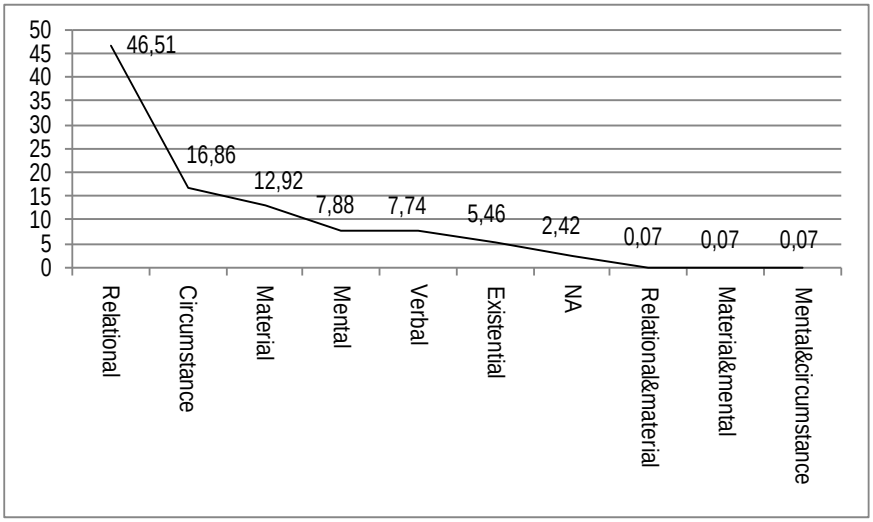


Figure 6.22b Process types in the study sample (general)



Of the 46 Transitivity categories in Table 6.23, the first ten make up 64.27% of the overall shell-noun evidence. The three most frequent participants are relational intensive Identifier, relational intensive Attribute and Goal (as in (668a) through (668c)), all occurring after the verb, and comprising 25.36%

of the cases. Regarding circumstances, matter, time and place feature most frequently in the sample (5.11%, 2.76%, 1.66%; as in (669a) through (669c)). This ranking coincides with the prevalence of respect, time and space adjuncts in syntax, but the frequencies diverge on account of differences in the coding practices for each variable. For example, while *about the problem* in *think about the problem* is a matter circumstance in Hallidayan Transitivity, *the problem* is a prepositional object in Quirk et al. (1985)[34]. Similarly, in an example like *the decay-rate restrictions discussed below lie* ***behind the paucity of experimental*** evidence, the prepositional phrase in boldface is a space adjunct in Quirk et al. (1985), but a relational circumstantial Attribute in SFG.

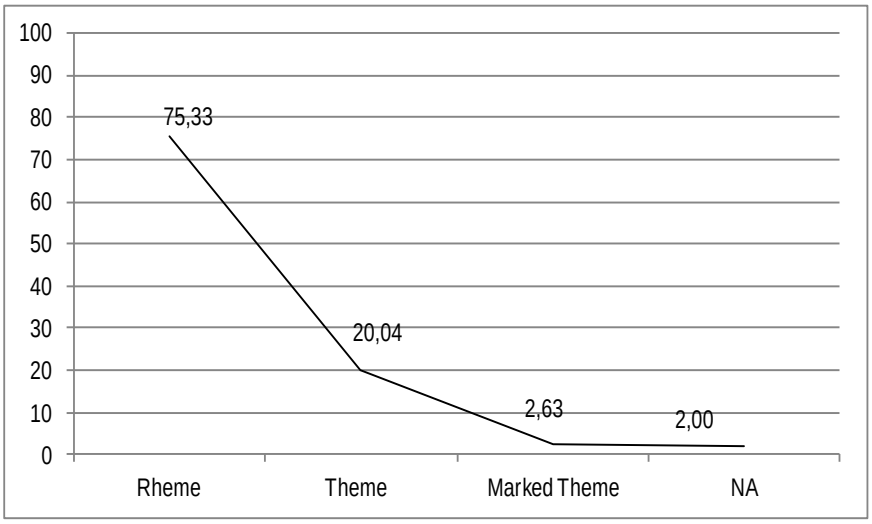
(668) (a) it was **our second chance**
(b) that was **a good experience**
(c) extending **our all-year capacity to give people what they want**
(669) (a) ask you to be warned **about two things**
(b) I wasn't very happy about it **at the time**
(c) **in this myth**, Isis was a woman who longed to increase her magic powers

The prominence of relational participants in Table 6.23 hinges on the overall dominance of relational processes in the study sample, with 46.51% of instances occurring with this process type (see Figure 6.22b). Specifically, relational intensive processes are the most frequent alternative, followed by relational circumstantial and relational possessive processes (30.75%, 8.85%, 7.74%; see Figure 6.22a). Interestingly, circumstances rank second in both figures, and show that in 63.37% of the evidence, shell-noun use is strongly primed for relational participants and, occasionally, for circumstantial roles. Material, mental, verbal and existential participants account for 34% of shell-noun instances, followed by uncoded cases (NA) and marginal instances of coordination.

Finally, with regard to the Theme system, Figure 6.23 shows the majority of examples in Rheme position, i.e. 75.33%. However, shell-noun use in the sample is negatively primed for Theme and marked Theme: both accounting for only 22.67% of the evidence.

[34] This is due to the possibility of a passive paraphrase like *the problem needs to be thought about*, with the original prepositional complement as subject (cf., however, *I was not actually consulted* ***on these applications***, where such a paraphrase is odd, and thus, the prepositional phrase in boldface is a respect adjunct, i.e. *in relation to these applications*).

Figure 6.23 Shell-noun use in Theme/Rheme positions



6.2.3.2 *Mode and genre distribution*

This section describes the behaviour of the three syntactico-semantic and textual variables with regard to mode and genre. Figures 6.24 and 6.25 focus on the distribution of shell-noun instances in terms of syntactic functions. Data (in percentages) are given for the top five functions in Table 6.20 and for cases where the shell noun acts as a phrase-level prepositional complement (i.e. CP).

Figure 6.24 Mode distribution of the top six syntactic functions (clause and phrase-level)

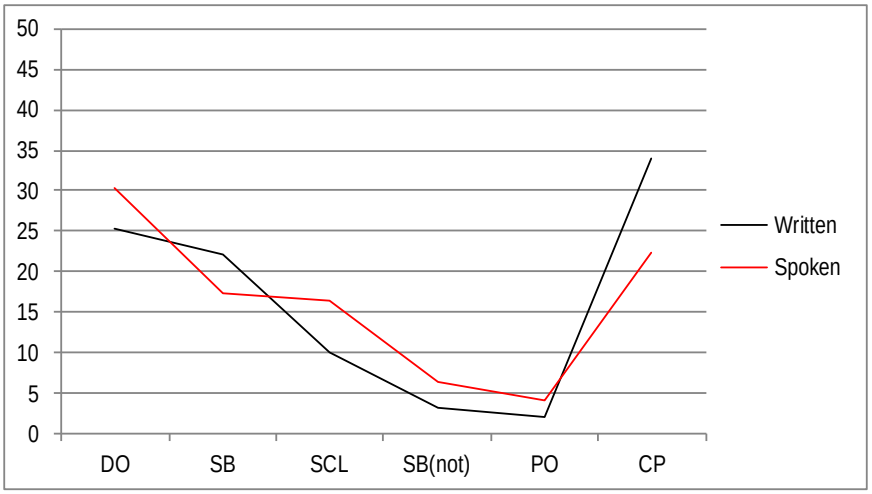
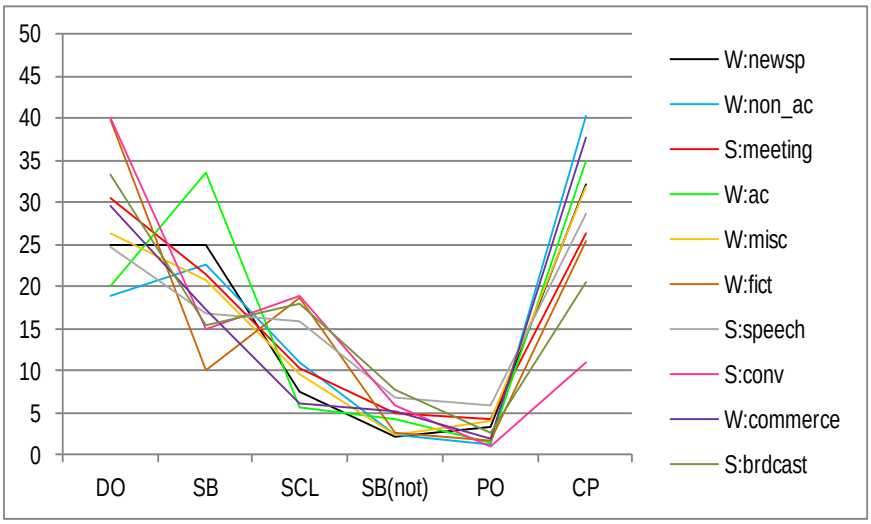


Figure 6.25 Genre distribution of the top six syntactic functions (clause and phrase-level)



All but two functions in Figure 6.24 show an association with the spoken mode. Shell-noun instances as direct objects seem more primed for spoken English, while subject shell units are more frequent in written English. Subject complements, notional subjects and prepositional objects are more common in spoken language, whilst prepositional complements are primed for written language (cf. Figure 6.15, where prepositional phrases as postmodifying structures prevail in written English). Mode differences are most marked with subject complements and prepositional complements: 6.41% in the former case and 11.62% in the latter.

Turning to Figure 6.25, S:conv, W:fict and S:brdcast rank highest in the use of direct object and subject complement shell nouns. These three super-genres are infrequent with subject nouns (especially in the case of W:fict). W:ac ranks highest among subjects, followed by W:newsp (cf., however, direct objects and subject complements, where W:ac is among the least frequent super-genres). The distribution of notional subjects and prepositional objects is more even: most super-genres cluster around the 2%–7% ranges. This being the case, notional subjects reveal a stronger priming for S:brdcast, S:speech and S:conv, while prepositional objects are more frequent in S:speech, S:meeting and W:misc. Lastly, the connection between phrase-level prepositional complements and the written mode manifests itself in four written super-genres: W:non_ac, W:commerce, W:ac and W:misc. S:conv ranks lowest in the use of phrase-level shell nouns (cf. Figures 6.14 and 6.16, where structural postmodification of shell-nouns, particularly by prepositional phrases, is relatively infrequent in S:conv).

Figures 6.26 through 6.29 display the mode and genre distributions for the ten most frequent participant types and the ten most frequent circumstances in Table 6.23.

Figure 6.26 Mode distribution of the top ten participant types

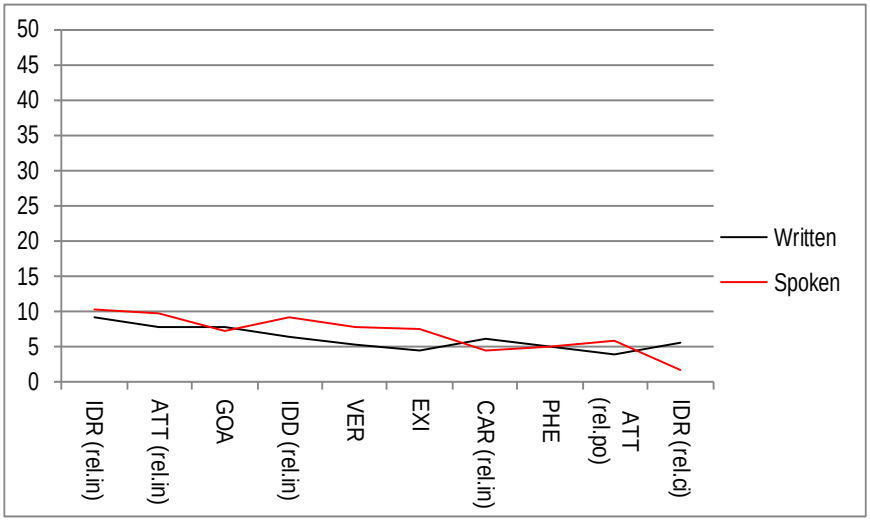


Figure 6.27 Mode distribution of the top ten circumstances

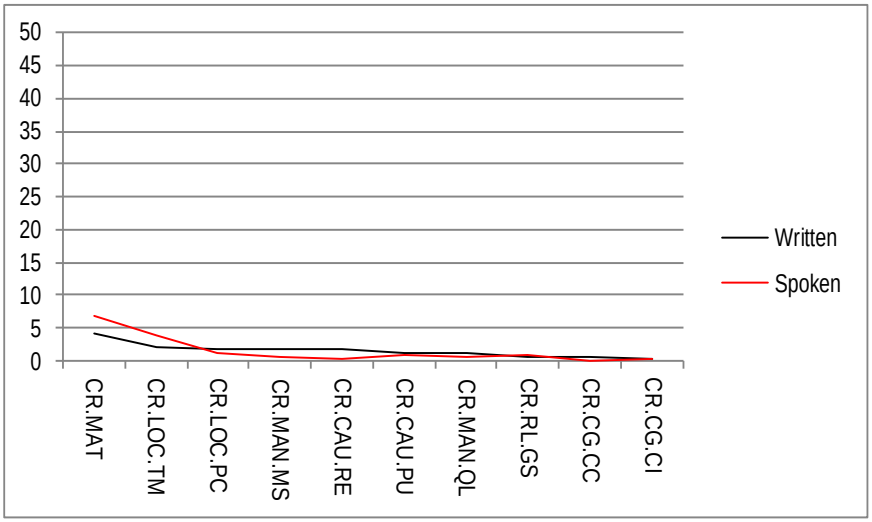


Figure 6.28 Genre distribution of the top ten participant types

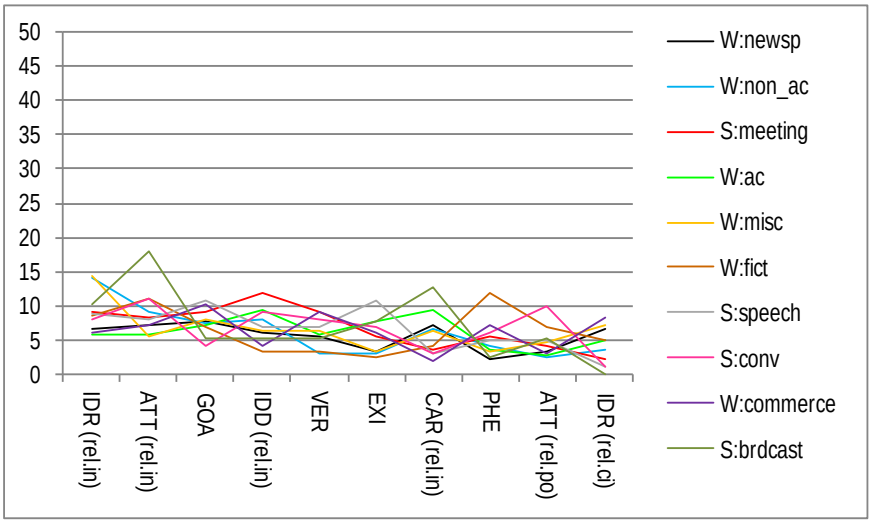
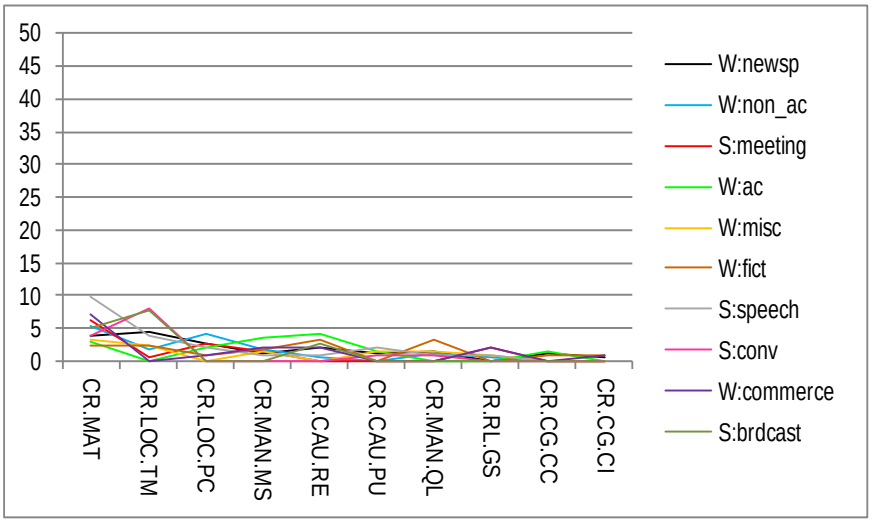


Figure 6.29 Genre distribution of the top ten circumstances



The most noticeable mode differences in Figure 6.26 concern circumstantial Identifiers, primed for written genres (4.01% difference between both modes), and Existent, intensive Identified and Verbiage, all more frequent in spoken English (their mode difference is 3.02%, 2.84% and 2.25%). Less clear is the connection between spoken language, intensive Identifier, and intensive and possessive Attribute, and that between written language and intensive Carrier. Goal and Phenomenon are frequent in both modes. As regards circumstances (Figure 6.27), matter and time are weakly primed for

spoken language, while place, manner:means and cause:reason are slightly more frequent in written English. The remaining categories almost overlap in their mode distribution.

The most relevant data in Figure 6.28 regard intensive Attribute, intensive Identified, Existent, intensive Carrier, Phenomenon and possessive Attribute. Both intensive Attribute and Carrier prevail in S:brdcast, as in (670a) and (670b). Interestingly, whilst W:ac ranks lowest in the former, it occurs as the second most frequent super-genre in relation to the latter (see (670c)). Intensive Identified stands out in S:meeting, followed by W:ac and S:conv (as in (671a) through (671c)). The high frequency of W:ac with intensive Identified and Carrier, two subject participant types, explains the strong connection between syntactic subject and this super-genre in Figure 6.25. Existent occurs most frequently in S:speech, followed by W:ac and S:brdcast (as in (672a) through (672c)), and rarely in W:fict. By contrast, W:fict is most prominent with Phenomenon (as in (673)), where it is first, and with possessive Attribute, where it is second (preceded by S:conv; as in (374a) through (374b)).

(670) (a) it's a scandal what we get for a pension
(b) the chances of me leaving Glasgow are pretty slim
(c) their clan system was in decline
(671) (a) one of the problems that we've been facing in the past is the amount of stuff that's on the floor
(b) the simplest cultural dimension is to use local authority administrative areas
(c) the other thing is that I haven't got any leave
(672) (a) the ultimate objective remains
(b) further details can be found in {11}
(c) there is a point of charging for every television set
(673) she vividly recalled every detail
(674) (a) all wood-burning pundits have their prejudices
(b) they haven't got the sense to ask

The remaining participant types reveal a more balanced use across genres. Intensive Identifier is most common in W:misc and W:non_ac (as in (675a) through (675b)), and is rare in W:ac. Goal instances predominate in S:speech, W:commerce and S:meeting (as in (676a) through (676c)), but are rare in S:conv. Verbiage shows an association with W:commerce, S:meeting and S:conv (as in (677a) through (677c)), but not with W:fict and W:non_ac. Finally, circumstantial Identifier is most frequent with W:commerce, W:misc and W:newsp (as in (678a) through (678c)), and least so with S:brdcast.

(675) (a) the education market is an area of their business they are expecting to expand in the UK
(b) this story forms a fairly complete myth of early origin
(676) (a) the philosophy that classes people as mere units of labour will be consigned to the dustbin of history
(b) the director of nursing may devise the system for allocating annual leave
(c) some additional parts have been written
(677) (a) before any categoric policy recommendation could be made
(b) Cynthia raised that point yesterday
(c) he tells jokes
(678) (a) staff availability should be determined by the work to be done
(b) the activities of one group restrict the freedoms, practices and conduct of another
(c) the arrival of Mr Horn in Bucharest was intended to dispel some of these suspicions

With respect to circumstances, Figure 6.29 shows weaker associations. The most noticeable concern matter, time, place, means, reason, manner:quality and role:guise. Matter is common in S:speech, W:commerce and S:meeting (as in (679a) through (679c)), and infrequent in W:fict and W:ac. Time peaks higher in S:conv and S:brdcast (as in (680a) through (680b)), while no instances occur in W:commerce and W:ac. Use of place circumstances, by contrast, is absent from S:conv, S:brdcast and W:misc, but stands out in W:non_ac, as in (681). Means and reason are primed for W:ac (as in (682) and (683)), and least so for S:conv. Lastly, manner:quality and role:guise show a stronger association with W:fict and W:commerce, as in (684) and (685).

(679) (a) we go out of our way to be fair in respect of certain areas
(b) it is in this sense that monopoly is said to be economically inefficient
(c) who could talk about experiences
(680) (a) I went with her a few times she frightened the life out of me
(b) after my own experiences, I wouldn't advice someone not to go to their GP
(681) this story is taken from a testimony written by Samuel Wallis
(682) to enrich the domain of research with fundamentally new findings
(683) because of the way fresh bound variable is created for every input, no variable that contains a value relevant to the program is overwritten until this final assignment
(684) I verse the way I like
(685) able to act in the dual capacity of broker and dealer

Figures 6.30 through 6.33 draw on the model of Figures 6.22a and 6.22b, with all participants merged into their corresponding process types and all circumstances conflated under a single category. As in Figures 6.22a and 6.22b, 6.30 and 6.32 distinguish between the three subtypes of relational processes, while 6.31 and 6.33 show combined results.

Figure 6.30 Mode distribution of process types (with relational subtypes)

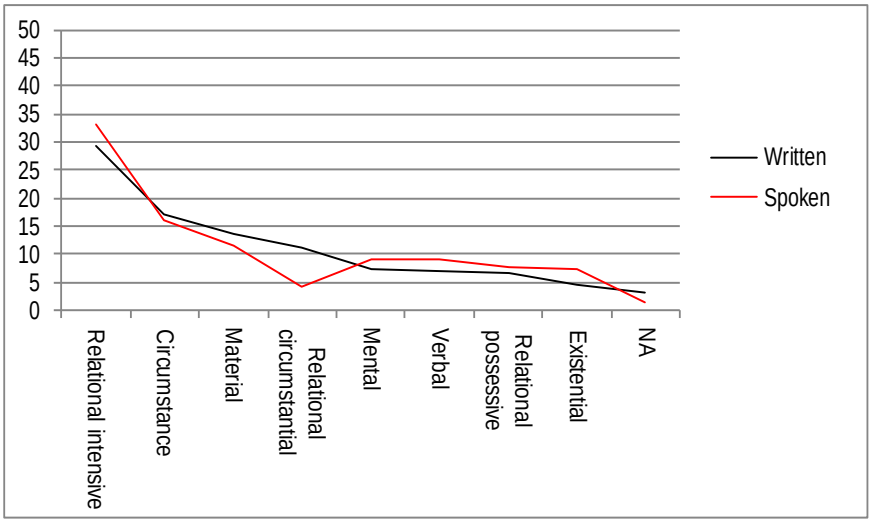


Figure 6.31 Mode distribution of process types (general)

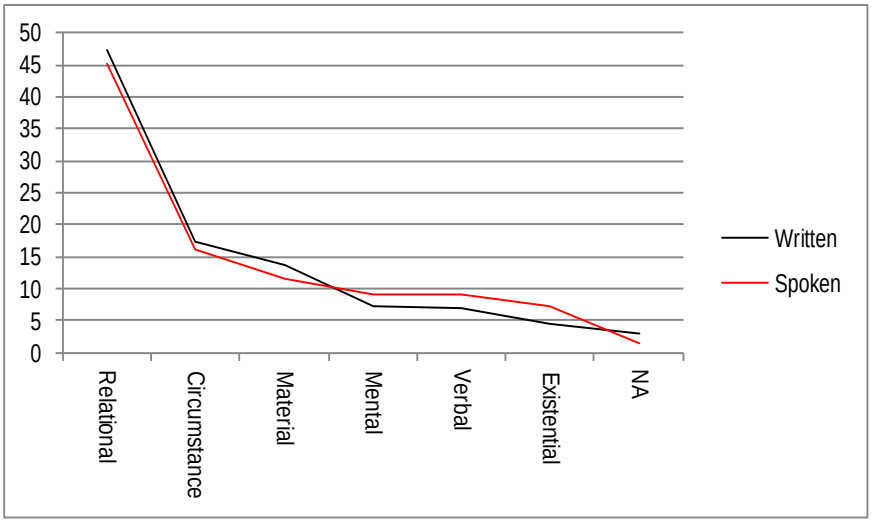


Figure 6.32 Genre distribution of process types (relational subtypes)

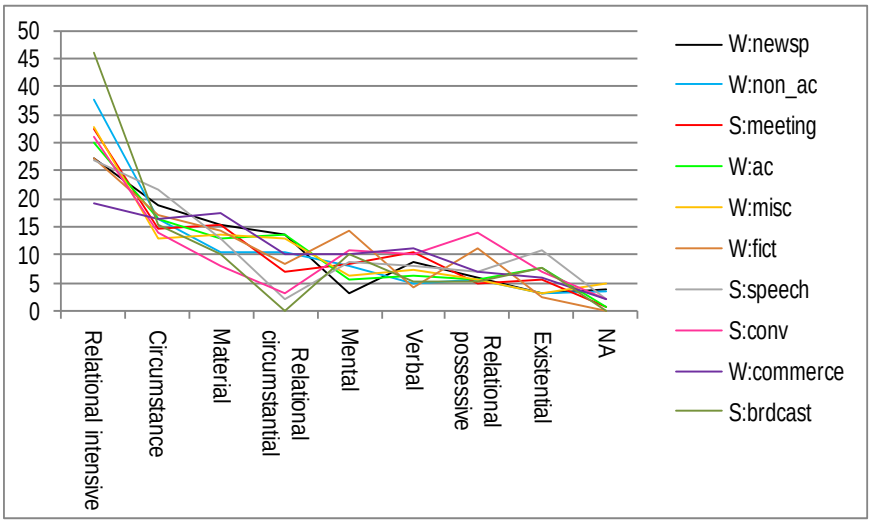
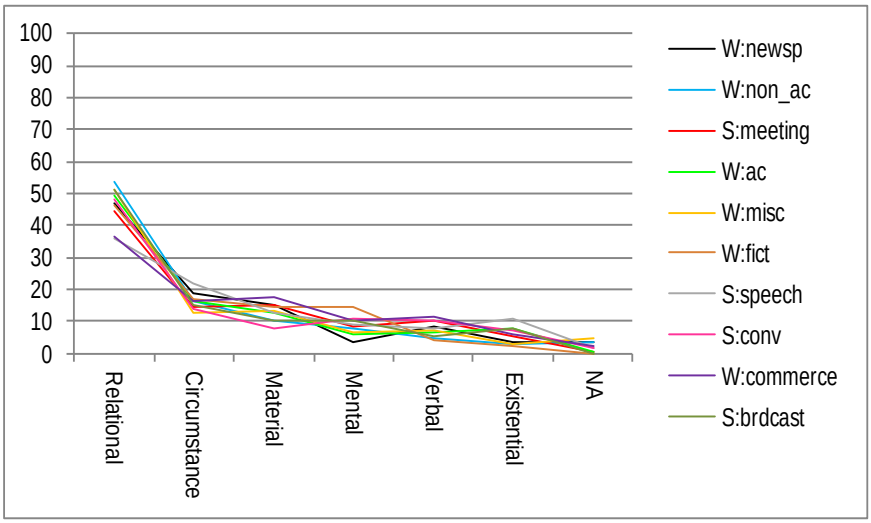


Figure 6.33 Genre distribution of process types (general)



In Figures 6.30 and 6.31 the most marked difference between both modes lies in relational circumstantial processes, primed strongly for written English. Relational intensive and possessive processes are more frequent in the spoken mode. This is also the case with mental, verbal and existential processes. Use of circumstances is roughly similar in both modes, this being slightly more frequent in written language. Uncoded instances (NA) are more prevalent in written English. This is logical considering their association with shell-noun use in titles or headlines, often lacking a

process. Shell-noun use among relational processes in general is almost equal in both modes according to Figure 6.31: written English is only slightly more frequent than written English.

In terms of genre distribution, relational intensive processes in Figure 6.32 rank highest in S:brdcast and W:non_ac, and lowest in W:commerce. W:commerce prevails among material and verbal processes. In the former, W:commerce is followed by W:newsp and S:meeting; in the latter, it is close to S:meeting and S:conv. Use of material and verbal processes is most limited among S:conv and W:fict, which rank highest with relational possessive and mental processes. Shell-nouns with mental processes are rare in W:newsp. This super-genre is most frequent with circumstances and relational circumstantial and material processes. Relational circumstantial processes in W:newsp rank almost equal to W:ac and W:misc. As to circumstances, W:newsp is the second most frequent super-genre, preceded by S:speech. S:speech also prevails with existential processes. Relational processes in Figure 6.33 show most genres clustered in the 45%–53% range. The three most frequent categories are W:non_ac, S:brdcast and W:ac, and the least so are W:commerce and S:speech.

Figures 6.34 and 6.35 bring together the more fine-grained information offered by syntactic function and participant type under the Theme system.

Figure 6.34 Mode distribution of Theme/Rheme

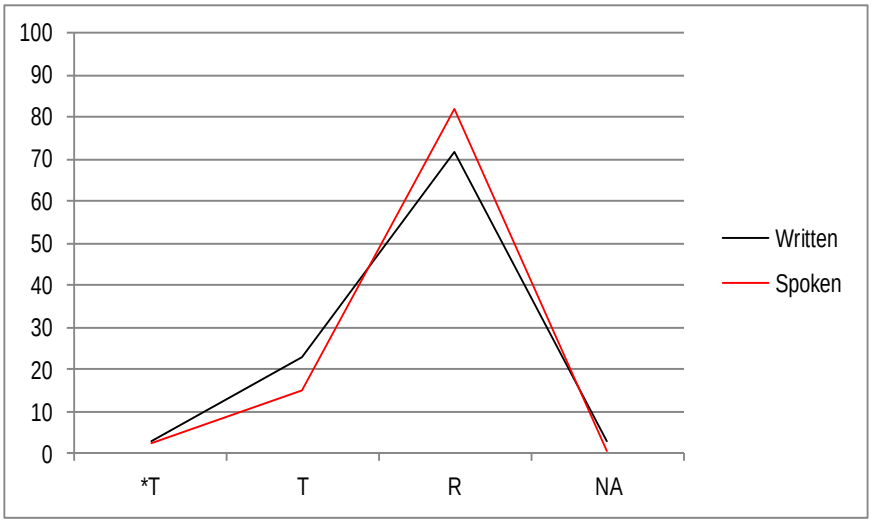


Figure 6.35 Genre distribution of Theme/Rheme

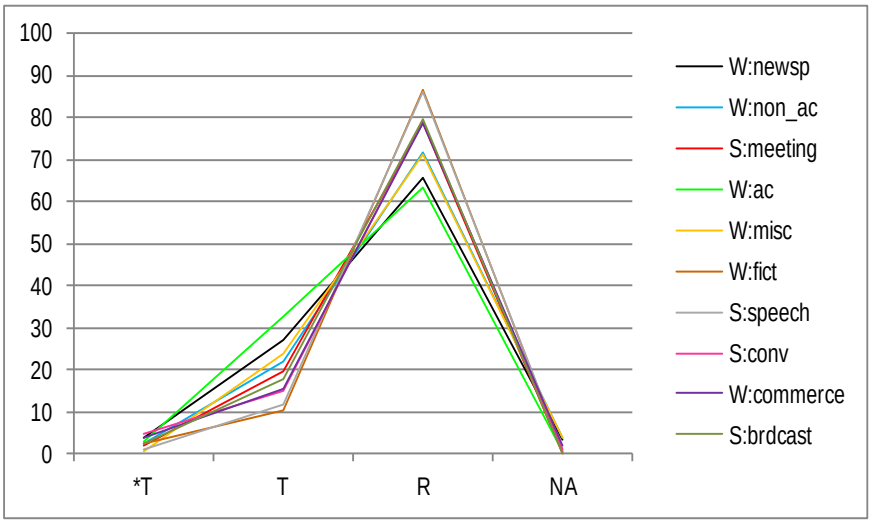


Figure 6.34 shows opposite lines: Theme is frequent in written English and Rheme is in spoken English. The mode difference regarding marked Theme is minimal (written language is only 0.44% more frequent than spoken language). Figure 6.35 indicates that the greater frequency of Theme in written English is best evidenced by two super-genres: W:ac and W:newsp (cf. Figure 6.25 for the prevalence of subjects in W:ac and W:newsp). W:fict and S:speech, which rank lowest among Theme shell-nouns, rank highest among Rheme units. The opposite applies to W:ac and W:newsp.

6.2.3.3 *Distribution across lemmas and semantic types*

Following the results for the general and mode/genre-specific distribution of syntactico-semantic and textual functions, the focus now turns to their behaviour across lemmas and semantic types of shell nouns. Tables 6.24a through 6.24b and Figure 6.36 below display the extent of use (in percentages) of the top five clause-level syntactic functions and of phrase-level instances (as in Figures 6.24 and 6.25 above). As in 6.2.2.3, all lists of lemmas in this section comprise only the top ten units.

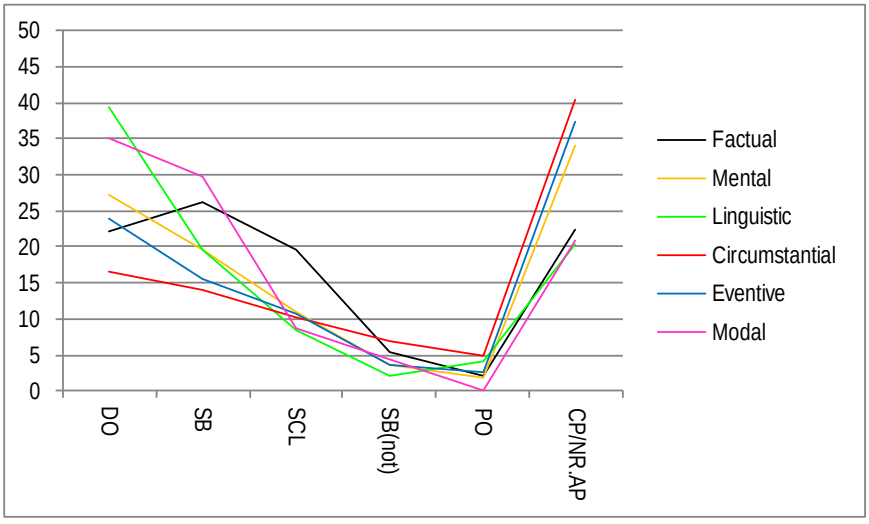
Table 6.24a Top ten lemmas for the top six syntactic functions (clause- and phrase-level). See Appendix 10.1 for the complete lists

DO	%	SB	%	SCL	%	SB(not)	%	PO	%
Correction	100.00	Facet	42.86	Opposite	70.00	Evidence	30.00	Scandal	15.38
Impetus	100.00	Misfortune	42.86	Surprise	44.00	Facet	14.29	Detail	7.50
Chance	52.50	Recollection	40.00	Testimony	42.86	Contradiction	12.50	Experience	7.50
Dimension	50.00	Contradiction	37.50	Recollection	40.00	Motivation	12.50	Recommendation	6.25
Endorsement	50.00	Motivation	37.50	Misfortune	28.57	Point	12.50	System	5.88
Scandal	46.15	Prejudice	36.36	Part	28.13	Time	12.50	Application	5.00
Suspicion	44.00	Problem	35.00	Motivation	25.00	Triumph	12.50	Area	5.00
Recommendation	43.75	Thing	35.00	Venture	24.14	Suspicion	12.00	Crime	5.00
Detail	42.50	Failure	34.78	Area	22.50	Way	10.00	Evidence	5.00
Joke	42.50	Assessment	34.62	Joke	22.50	Surprise	8.00	Practice	5.00

Table 6.24b Top ten lemmas for the top six syntactic functions (clause- and phrase-level)

Phrase level (CP) (%)			
Foreboding	85.71	Quest	60.00
Terror	75.00	Sense	59.09
Proviso	71.43	Irony	57.14
Triumph	62.50	Endeavour	55.56
Anger	60.00	Venture	55.17

Figure 6.36 General semantic types and the top six syntactic functions (clause- and phrase-level)



The above tables and figure point to a number of connections. Direct object, for example, ranks highest among linguistic nouns, followed closely by modal instances (e.g. *correction*, *endorsement*, *recommendation*, *detail*,

joke; *chance*)[35]. It is important to note that the overall prevalence of modal nouns relates closely to *chance* (52.50%; cf. *leave*, *capacity*, *failure*, with frequencies ranging from 25% to 8%). Use of subject shell noun-phrases is most frequent with modal and factual lemmas (e.g. *failure, chance, capacity*; *facet, problem, thing, part, example*). Factual units, ranking second with subjects, prevail with subject complements (e.g. *opposite*, *testimony*, *point*, *example*, *problem*). *Opposite*, topping the list, is most clearly primed for this syntactic function (70%). Regardless of semantic type, subject complement is particularly common with nouns showing attitudinal nuances (e.g. *surprise, misfortune*, *motivation*, *joke*) and with highly semantically unspecific nouns like *part*, *area* or *point*, where evaluation is often conveyed through premodifying Epithets (e.g. *the better part of defence*, *one of the key areas*, *a good point*).

With regard to notional subjects and prepositional objects, semantic preferences are more evenly distributed. Notional subjects are more frequent among circumstantial and factual nouns (e.g. *time*, *way*; *evidence*, *facet*, *point*), whilst prepositional objects are primed for circumstantial and linguistic nouns (e.g. *system*, *area*, *practice*; *detail*, *recommendation*, *application*). Lastly, phrase-level instances are associated with circumstantial, eventive and mental nouns (e.g. *proviso*, *area*, *time*; *triumph*, *quest*, *endeavour*, *venture*; *foreboding*, *terror*, *anger*, *sense*). The frequency of circumstantial instances among prepositional complements is not unexpected given the common occurrence of some lemmas in circumstantial or adjunct prepositional phrases (e.g. *with the proviso that*, *in this/that area*, *at this time, by that time*, etc.). Similarly, uncountable instances like *foreboding*, *terror* or *anger* rank highest among mental nouns in phrase-level positions. They are often preceded by partitive noun phrases like *a sense/feeling of foreboding*, *terror* or *anger*. *Sense* is also frequent on account of its frequent phraseology in *in this/that sense.*

Tables 6.25a through 6.26b and Figures 6.37 and 6.38 bring together the results for the ten most frequent participant types and the ten most frequent circumstances (as in 6.2.3.2).

[35] The strong priming shown with *correction* and *impetus* (both 100%) is due to the small amount of evidence for these two lemmas (two instances for the former and six for the latter).

Table 6.25a Top ten lemmas for the top ten participant types. See Appendix 10.2 for the complete lists

IDR (rel.in)	%	ATT (rel.in)	%	GOA	%	IDD (rel.in)	%	VER	%
Opposite	70.00	Surprise	44.00	Correction	50.00	Recollection	40.00	Correction	50.00
Endorsement	50.00	Failure	30.43	Project	30.00	Thing	30.00	Scandal	30.77
Misfortune	42.86	Testimony	28.57	System	23.53	Motivation	25.00	Recommendation	28.13
Recollection	40.00	Joke	27.50	Challenge	20.59	Anger	20.00	Detail	25.00
Facet	28.57	Endorsement	25.00	Opposite	20.00	Example	20.00	Point	25.00
Testimony	28.57	Endeavour	22.22	Practice	15.00	Objective	20.00	Warning	20.69
Motivation	25.00	Venture	20.69	Finding	14.29	Assessment	19.23	Evidence	20.00
Characteristic	24.14	Phenomenon	19.44	Testimony	14.29	Point	17.50	Joke	20.00
Part	21.88	Vision	17.24	Application	12.50	Facet	14.29	Assessment	19.23
Finding	21.43	Scandal	15.38	Contradiction	12.50	Irony	14.29	Application	15.00

Table 6.25b Top ten lemmas for the top ten participant types

EXI	%	CAR (rel.in)	%	PHE	%	ATT (rel.po)	%	IDR (rel.ci)	%
Evidence	30.00	Endorsement	25.00	Foreboding	28.57	Chance	17.50	Impetus	50.00
Facet	14.29	Quest	20.00	Contradiction	25.00	Foreboding	14.29	Suspicion	24.00
Contradiction	12.50	Recollection	20.00	Experience	22.50	Misfortune	14.29	Problem	17.50
Motivation	12.50	Prejudice	18.18	Word	22.50	Capacity	13.04	Facet	14.29
Point	12.50	Chance	17.50	Philosophy	18.18	Dimension	12.50	Testimony	14.29
Problem	12.50	Vision	17.24	Prejudice	18.18	Objective	12.50	Venture	10.34
Terror	12.50	Facet	14.29	Answer	17.50	Suspicion	12.00	Vision	10.34
Time	12.50	Finding	14.29	Vision	17.24	Characteristic	10.34	Myth	10.00
Triumph	12.50	Foreboding	14.29	Recommendation	15.63	Problem	10.00	Prejudice	9.09
Surprise	12.00	Phenomenon	13.89	Scandal	15.38	Prejudice	9.09	Dimension	8.33

Table 6.26a Top ten lemmas for the top ten circumstances. See Appendix 10.3 for the complete lists

CR.MAT	%	CR.LOC.TM	%	CR.LOC.PC	%	CR.MAN.MS	%	CR.CAU.RE	%
Sense	50.00	Time	65.00	Anger	20.00	Leave	25.00	Anger	20.00
Challenge	17.65	Quest	20.00	Quest	20.00	Proviso	14.29	Irony	14.29
Recommendation	15.63	Scandal	7.69	Area	15.00	Triumph	12.50	Crime	12.50
Capacity	13.04	Experience	7.50	Facet	14.29	Way	12.50	Terror	12.50
Crime	12.50	Work	7.41	Testimony	14.29	Word	7.50	Evidence	5.00
Endeavour	11.11	Evidence	5.00	Triumph	12.50	Venture	6.90	Philosophy	4.55
Finding	10.71	Assessment	3.85	Myth	10.00	Finding	3.57	Capacity	4.35
Evidence	10.00	Warning	3.45	Venture	6.90	Warning	3.45	Failure	4.35
Practice	10.00	Recommendation	3.13	Capacity	4.35	System	2.94	Work	3.70
Experience	7.50	Application	2.50	Suspicion	4.00	Area	2.50	Vision	3.45

Table 6.26b Top ten lemmas for the top ten circumstances

CR.CAU.PU	%	CR.MAN.QL	%	CR.RL.GS	%	CR.CG.CC	%	CR.CG.CI	%
Endeavour	11.11	Irony	28.57	Capacity	13.04	Warning	10.34	Proviso	57.14
Challenge	5.88	Leave	25.00	Example	5.00	Suspicion	4.00	Anger	0.00
Area	5.00	Anger	20.00	Venture	3.45	Finding	3.57	Answer	0.00
Detail	5.00	Way	15.00	Warning	3.45	Phenomenon	2.78	Application	0.00
Joke	5.00	Suspicion	12.00	Project	2.50	Experience	2.50	Area	0.00
Project	5.00	Objective	2.50	Word	2.50	Anger	0.00	Assessment	0.00
Surprise	4.00	Practice	2.50	Anger	0.00	Answer	0.00	Capacity	0.00
Venture	3.45	Answer	0.00	Answer	0.00	Application	0.00	Challenge	0.00
Answer	2.50	Application	0.00	Application	0.00	Area	0.00	Chance	0.00
Thing	2.50	Area	0.00	Area	0.00	Assessment	0.00	Characteristic	0.00

Figure 6.37 General semantic types and the top ten participant types

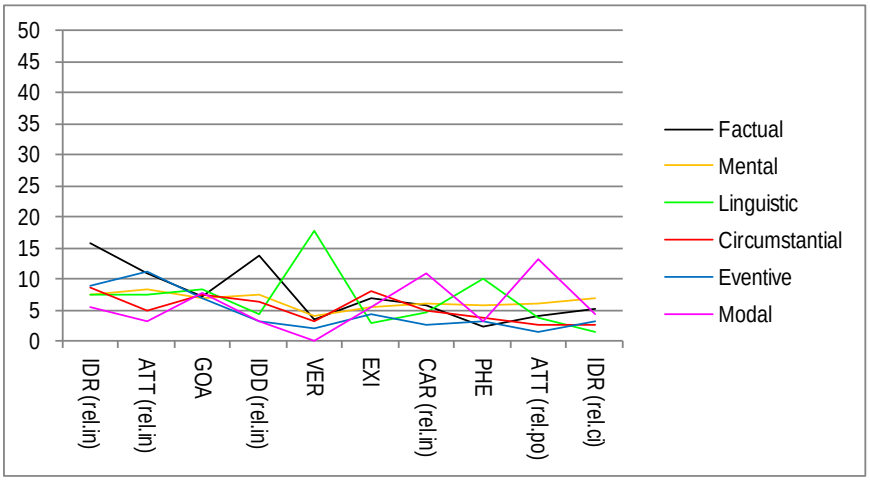
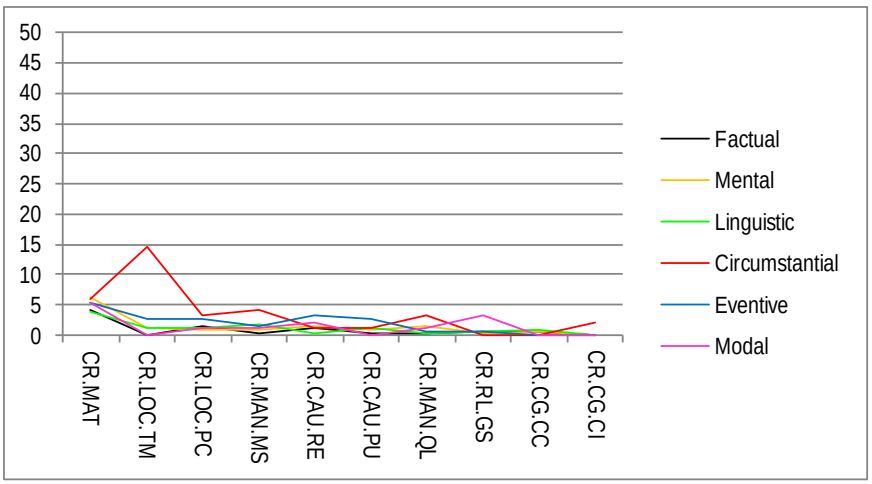


Figure 6.38 General semantic types and the top ten circumstances



Among the top ten participant types, the clearest semantic associations in Figure 6.37 are shown by intensive Identifier and Identified, Verbiage, intensive Carrier, Phenomenon and possessive Attribute. In the case of intensive Identifier and Identified, Factual nouns prevail (e.g. *opposite*, *facet*, *characteristic*, *part*, *finding* for IDR (rel.in), as in (686a); *thing*, *example*, *point*, *facet* for IDD (rel.in), as in (687a)). The strong connection between *opposite* and subject complement in Table 6.24a is now found to relate semantically to intensive Identifier (as in (686b)), sharing the same 70% proportion with its syntactic counterpart. Concerning intensive Identified, the occurrence of *thing* or *point* among the top ten lemmas in Table 6.25a is explained by their widespread use in focusing constructions, as in (687b) through (687c) (see Schmid 2000: 329–37 in 2.2.2.2). Verbiage and Phenomenon favour linguistic nouns, but this is most evident with Verbiage (e.g. *correction*, *recommendation*, *detail*, *point*, *warning*, *joke*, *application* for VER, as in (688); *contradiction*, *word*, *answer*, *recommendation* for PHE, as in (689a)). Linguistic nouns as Phenomenon are closely followed by mental nouns (e.g. *foreboding*, *experience*, *prejudice*, *vision*, as in (689b)). The association between Verbiage and linguistic nouns such as *recommendation*, *point*, *warning* and *application* rests on their frequent association with verbs like *give* and *make* (e.g. *make a recommendation*, *give a warning*). These, according to Halliday & Matthiessen (2004: 256), are '"empty" verbs' followed by nouns conveying 'the *name of the saying*' (emphasis as in the original), which thereby qualify as Verbiage. Modal nouns, rare among the four aforementioned participants (i.e. intensive Identifier and Identified, Verbiage and Phenomenon), prevail with intensive Carrier and possessive Attribute. Use of modal nouns is in both cases restricted to *chance* and *capacity*, the former topping the list of possessive Attributes (see (690) for CAR(rel.in) and (691) for ATT(rel.po)). Mental and factual nouns feature as the second most frequent semantic types in both participant types (e.g. *endorsement*, *recollection*, *prejudice*, *facet*, *finding*, *phenomenon* for CAR (rel.in); *foreboding*, *suspicion*, *prejudice*, *dimension*, *characteristic*, *problem* for ATT (rel.po)).

(686) (a) another variable to be considered would be social characteristics
(b) glazing is the exact opposite to scumbling

(687) (a) one other facet of rural community is provided by the fourth element of Pahl's classification
(b) the other thing is that I haven't got any leave
(c) the point of issue is that estuary and coastal zone management is a reality

(688) I was telling jokes

(689) (a) we should know the answer to this

(b) the vision that I had of a three tier cake came to me as I was striding home

(690) inherent in any permanent way job is the capacity for something to go wrong

(691) they have no chance to notice or see anything

The least clear associations in Figure 6.37 concern intensive Attribute, Goal, Existent and circumstantial Identifier. Intensive Attribute is most common with eventive, factual and mental nouns (e.g. *endeavour*, *venture*, as in (692a); *testimony*, *phenomenon*, as in (692b); *surprise*, *vision*, as in (692c)). As pointed out with regard to subject complements, intensive Attributes are similarly primed for nouns with attitudinal or evaluative meanings. Nouns like *surprise* involve a positive attitude, whilst nouns like *failure*, *joke* or *scandal* are negatively evaluative. Use of Goal shell nouns is almost equally distributed across the six semantic types. Goal is only slightly more frequent with linguistic, modal, circumstantial and factual nouns (e.g. *correction*, *application*, as in (693a); *chance*, as in (693b); *system*, *practice*, as in (693c); *opposite*, *finding*, as in (693d)). Existent, as with notional subject in Figure 6.36, is most frequent with circumstantial and factual nouns (e.g. *time*, *way*, as in (694a); *evidence*, *facet*, *point*, *problem*, as in (694b)). In this case, priming is markedly frequent with *evidence* (30%), which makes sense given the common occurrence of this noun in contexts where emphasis is laid on the need to prove something through the existence or lack of evidence (especially in academic prose). Finally, circumstantial Identifier is associated with mental, factual and modal nouns (e.g. *suspicion*, *vision*, *prejudice*, as in (695a); *impetus*, *problem*, *facet*, *dimension*, as in (695b); *chance*, *capacity*, as in (695c)).

(692) (a) that's not a business venture
(b) the ATS were a wartime phenomenon
(c) Peter Scudamore's late decision to partner Granville Again was no surprise

(693) (a) it has submitted a planning application to double the size of its current premises
(b) I don't think we missed any chances
(c) the system not having been adopted by all education authorities
(d) they wanted to write up their findings

(694) (a) there are times when that has to be done
(b) there is evidence to make the conventional figure inappropriate

(695) (a) her husband's inconsistencies and contradictions aroused suspicion
(b) this created a tremendous problem
(c) rapid growth at this stage may decrease the individual's chances of survival

As to the top ten circumstances, the most remarkable association in Figure 6.38 is between time circumstances and circumstantial nouns, as in (696). This is due to pre-eminence of *time* among time circumstances (65%). Circumstantial nouns also peak slightly higher with place, manner:means, manner:quality and contingency:condition circumstances (e.g. *area*, as in (697); *proviso, way*, *system*, *area*, as in (698); *way*, as in (699); *proviso*, as in (700)). Priming is most marked in the case of *proviso* (57.14%). Circumstantial nouns are less prominent with cause:reason and cause:purpose, where eventive nouns are more frequent (e.g. *crime*, *work*, as in (701); *endeavour*, *venture*, as in (702)). Lastly, use of role:guise stands out with modal nouns (e.g. *capacity*, as in (703)), while matter is only marginally more common with mental, circumstantial, modal and eventive nouns (e.g. *sense*, as in (704a); *practice*, as in (704b); *capacity*, as in (704c); *crime*, *endeavour*, as in (704d)). Priming within this circumstance is most obvious with *sense*, 50% of its use occurring as matter circumstances or respect adjuncts in *in one/this/that sense* instances.

(696) by the time you've got all the dishes done, the glasses have all drained off
(697) it is in the area of State aid that there is likely to be greater conflict
(698) landscapes can now be seen in a variety of ways
(699) has made its views evident only in oblique ways
(700) he agreed with one proviso
(701) their own khan chopped off heads for crimes which merited mere birching
(702) the computer is used in an endeavour to replace intuition with quantification
(703) the appropriate body within a member state assists in an advisory capacity
(704) (a) lactic acid is poisonous in one sense
(b) taping Hugh my father on shipyard riveting practices
(c) candidates would have to be chosen with regard to their capacity to attract votes
(d) the farmers were charged with the crime of growing opium

Tables 6.27a and 6.27b and Figures 6.39 and 6.40 shift the focus from specific participant types and circumstances to processes and circumstances.

Table 6.27a Process types and their top ten lemmas. See Appendix 10.4 for the complete lists

Relational	%	Relational intensive	%	Relational circumstantial	%	Relational possessive	%	Circumstance	%
Endorsement	100.00	Endorsement	100.00	Impetus	50.00	Foreboding	57.14	Proviso	71.43
Impetus	100.00	Recollection	100.00	Failure	26.09	Impetus	50.00	Time	67.50
Recollection	100.00	Motivation	75.00	Contradiction	25.00	Dimension	37.50	Anger	60.00
Characteristic	86.21	Opposite	70.00	Problem	25.00	Chance	32.50	Sense	54.55
Dimension	79.17	Characteristic	62.07	Suspicion	24.00	Sense	18.18	Leave	50.00
Failure	78.26	Facet	57.14	Myth	23.33	Objective	17.50	Irony	42.86
Motivation	75.00	Testimony	57.14	Experience	17.50	Misfortune	14.29	Quest	40.00
Problem	72.50	Part	53.13	Venture	17.24	Characteristic	13.79	Way	37.50
Facet	71.43	Thing	52.50	Work	14.81	Capacity	13.04	Capacity	34.78
Foreboding	71.43	Surprise	52.00	Facet	14.29	Problem	12.50	Crime	30.00

Table 6.27b Process types and their top ten lemmas

Material	%	Mental	%	Verbal	%	Existential	%
Correction	50.00	Foreboding	28.57	Correction	50.00	Evidence	30.00
Leave	50.00	Prejudice	27.27	Scandal	30.77	Facet	14.29
Terror	50.00	Contradiction	25.00	Recommendation	28.13	Contradiction	12.50
Work	40.74	Experience	22.50	Detail	25.00	Motivation	12.50
Project	37.50	Word	22.50	Point	25.00	Point	12.50
Opposite	30.00	Philosophy	18.18	Warning	20.69	Problem	12.50
Challenge	29.41	Answer	17.50	Application	20.00	Terror	12.50
Misfortune	28.57	Vision	17.24	Evidence	20.00	Time	12.50
System	26.47	Surprise	16.00	Joke	20.00	Triumph	12.50
Practice	25.00	Recommendation	15.63	Assessment	19.23	Surprise	12.00

Figure 6.39 General semantic types and process types (relational subtypes)

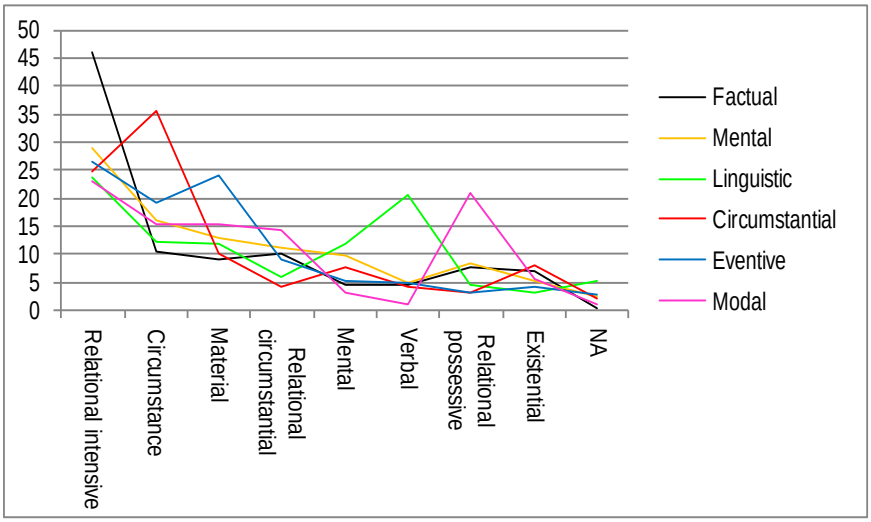
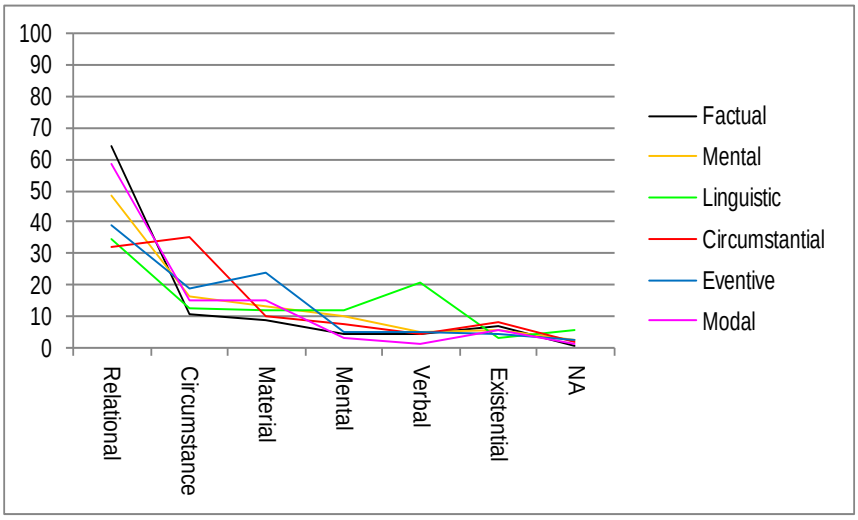


Figure 6.40 General semantic types and process types (general)



In terms of relational processes, Figure 6.39 shows that the realisation of relational intensive participants is primarily associated with factual nouns (e.g. *opposite*, *characteristic*, *facet*, *part*, *thing*), while relational possessive instances are mostly modal in meaning (e.g. *chance*, *capacity*). Whilst mental and factual nouns rank low with regard to semantic types (Figure 6.39), in relation to lemmas (Table 6.27a), they are more numerous and widespread in relational possessive uses than modal nouns (e.g. *foreboding*, *sense*, *objective*; *impetus*, *dimension*, *characteristic*, *problem*). Modal nouns also rank higher in relational circumstantial roles, where they are closely followed by mental, factual and eventive instances (e.g. *failure*; *suspicion*, *experience*; *impetus*, *problem*, *facet; venture*, *work*). Overall, as shown in Figure 6.40 and Table 6.27a, relational participants are most common among factual, modal and mental nouns (e.g. *endorsement*, *impetus*, *characteristic*, *failure*; *chance*, *capacity*; *recollection*, *motivation*, *foreboding*). Of the remaining process types, only material and verbal show definite preferences. The former is most frequent with eventive, modal and mental units (e.g. *work*, *misfortune*; *leave*, *chance*; *terror*, *project*), and the latter is strongly primed for linguistic nouns (e.g. *correction*, *recommendation*, *detail*, *warning*, *application*). Existential uses, as shown in Figure 6.37, are slightly more frequent among circumstantial and factual nouns (e.g. *time*, *way*; *evidence*, *facet*). Lastly, circumstances prevail among circumstantial and eventive nouns (e.g. *proviso*, *time*, *way*; *quest*, *crime*, *venture*).

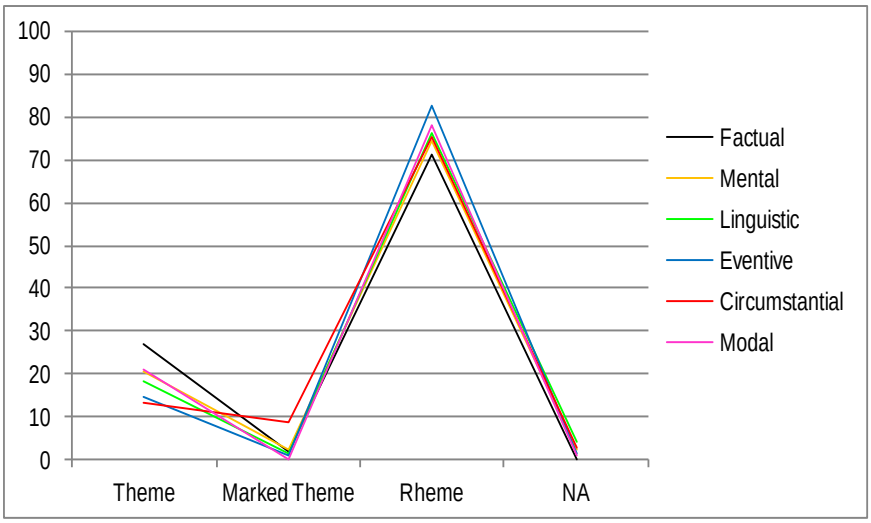
Turning finally to the Theme system, Table 6.28 and Figure 6.41 indicate a stronger connection between Theme and factual, mental and modal nouns (e.g. *example*, *thing*, *problem*; *recollection*, *motivation*, *assessment*; *failure*), between marked Theme and circumstantial nouns (e.g. *time*, *proviso*, *area*),

and between Rheme and eventive, modal and linguistic units (e.g. *crime*, *work*, *practice*; *capacity*, *chance*; *correction*, *recommendation, word*).

Table 6.28 Top ten lemmas for Theme/Rheme. See Appendix 10.5 for the complete lists

Marked Theme	%	Theme	%	Rheme	%
Time	22.50	Contradiction	50.00	Correction	100.00
Quest	20.00	Misfortune	42.86	Impetus	100.00
Irony	14.29	Recollection	40.00	Sense	100.00
Proviso	14.29	Motivation	37.50	Surprise	96.00
Philosophy	9.09	Example	35.00	Crime	95.00
Scandal	7.69	Thing	35.00	Evidence	90.00
Area	7.50	Failure	34.78	Opposite	90.00
Warning	6.90	Assessment	34.62	Work	89.29
Phenomenon	5.56	Phenomenon	33.33	Suspicion	88.00
Evidence	5.00	Problem	32.50	Practice	87.50

Figure 6.41 General semantic types and Theme/Rheme



6.2.3.4 *Syntactico-semantic function and other variables*

Before moving on to the wider discourse context in 6.2.4, this section adds further detail to the findings on the clause-level behaviour of shell-noun instances by examining the link between syntactic function and formal structure, and between participant type and semantic premodification.

Table 6.29 and Figures 6.42 through 6.44 explore the connection between syntax and noun-phrase structure. The combination of syntactic function and formal structure, both in their most fine-grained detail (see Tables 6.8

and 6.20), returns 895 patterns. Table 6.29 contains only the top 170 combinations, which account for 49.90% of data.

Table 6.29 Syntactic functions and their formal realisations (Top 50%). See Appendix 11 for the complete list

Syntactic function and formal structure (Top 50%)			
SB-DF.AR^H	2.35	AB.AT.TM-DM.DT(THIS)^H	0.14
DO-DF.AR^H	1.80	DO-AJ^H^PP(about)	0.14
SCL-IN.AR^H	1.73	DO-AJ^H^PP(in)	0.14
DO-H	1.45	DO-AS.DT^AJ^H	0.14
DO-DF.AR^H^PP(of)	1.24	DO-DF.AR^N^H	0.14
SB-DF.AR^H^PP(of)	1.24	DO-DF.AR^H^AP.THAT.CL	0.14
DO-AJ^H	1.04	DO-DM.DT(THAT)^H^PP(of)	0.14
SCL-IN.AR^AJ^H	1.04	DO-H^PP(from)	0.14
DO-H^PP(of)	0.90	DO-IN.AR^AJ^AJ^H	0.14
SB-DF.AR^AJ^H	0.90	DO-IN.AR^AJ^H^AP.THAT.CL	0.14
SB-H	0.90	DO-IN.AR^AJ^H^AP.TI.CL	0.14
DO-IN.AR^AJ^H	0.83	DO-IN.AR^H^AP.THAT.CL	0.14
SB-AJ^H	0.83	DO-IN.AR^H^PP(for)	0.14
SB-PS.DT^H	0.83	DO-IN.AR^H^PP(with)	0.14
DO-IN.AR^H	0.69	DO-IN.AR^H^PT.WK.NR.IT.AP	0.14
SCL-DF.AR^H	0.69	DO-NG.DT^AJ^H	0.14
SCL-H^PP(of)	0.69	DO-NG.DT^H	0.14
SB-DM.DT(THESE)^H	0.62	DO-NG.DT^H^PP(of)	0.14
DO-DF.AR^AJ^H	0.55	DO-NUM.CD^H^PP(of)	0.14
SCL-DF.AR^H^PP(of)	0.55	DO-PDT^H	0.14
DO-IN.AR^H^PP(of)	0.48	DO-PS.DT^H^PP(of)	0.14
DO-PS.DT^H	0.48	DO-QT^AJ^H	0.14
SB-H^PP(of)	0.48	DO-QT^H^PP(of)	0.14
DO-QT^H	0.41	DO-UV.DT^H	0.14
DO-PS.DT^AJ^H	0.41	DO-UV.PDT^DF.AR^H	0.14
SB-DF.AR^N^H	0.41	SB-AS.DT^AJ^H	0.14
CP.OF(<DO)-H	0.41	SB-DF.AR^AJ^H^PP(about)	0.14
DO-DF.AR^AJ^H^PP(of)	0.35	SB-DF.AR^AJ^H^PP(of)	0.14
DO-DF.AR^H^RT.RV.CL	0.35	SB-DF.AR^H^PP(to)	0.14
DO-DM.DT(THIS)^H	0.35	SB-DF.AR^H^PP(for)	0.14
SCL-AJ^H	0.35	SB-DF.AR^NUM.CD^AJ^H^PP(of)	0.14
SCL-DF.AR^AJ^H	0.35	SB-DF.AR^NUM.OR^H^RT.RV.CL	0.14
SCL-DF.AR^H^RT.RV.CL	0.35	SB-DF.DV.GV.NP^AJ^H	0.14
SCL-IN.AR^AJ^H^PP(of)	0.35	SB-DM.DT(THIS)^AJ^H	0.14
DO-DF.AR^H^AP.TI.CL	0.28	SB-H^AP.TI.CL	0.14
DO-DM.DT(THAT)^H	0.28	SB-H^PP(for)	0.14
DO-H^AP.TI.CL	0.28	SB-NUM.CD^H^PP(of)	0.14
DO-NAS.DT^H	0.28	SB-NUM.CD^PV.PP(of DF.AR^AJ^H^RT.RV.CL)	0.14
DO-NUM.GO^H	0.28	SB-NUM.CD^PV.PP(of DF.AR^H^RT.RV.CL)	0.14
SB-DF.AR^H^RT.RV.CL	0.28	SB-NUM.GO^H^PP(of)	0.14
SB-DF.DV.GV.NP^H^PP(of)	0.28	SB-PS.DT^AJ^H	0.14
SB-IN.AR^H	0.28	SB-PS.DT^H^PP(for)	0.14

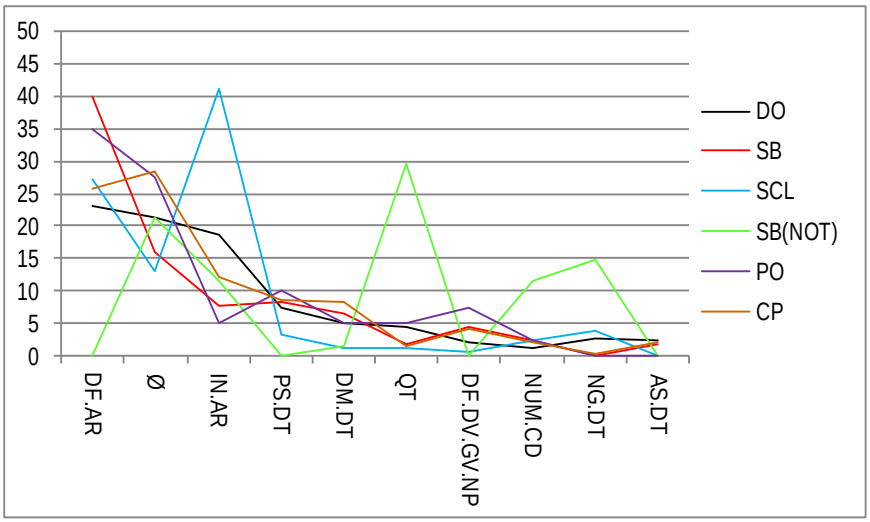
SB-IN.AR^H^PP(of)	0.28	SB-RV.DT^H	0.14
SB-IN.AR^AJ^H	0.28	SB-RV.DT^H^PP(of)	0.14
SCL-IN.AR^H^PP(of)	0.28	SCL-DF.AR^AJ^H^PP(for)	0.14
SCL-NG.DT^H	0.28	SCL-DF.AR^QL.PV^H^RT.RV.CL	0.14
SB(not)-QT^H	0.28	SCL-H^PP(to)	0.14
SB(not)-QT^H^PP(of)	0.28	SCL-IN.AR^AJ^H^PL.ED.CL	0.14
NA-H	0.28	SCL-IN.AR^AJ^H^PP(to)	0.14
CP.OF(<DO)-AJ^H	0.28	SCL-IN.AR^H^RT.RV.CL(where)	0.14
CP.OF(<SB)-H	0.28	SCL-IN.AR^N^AJ^H	0.14
DO-DF.AR^AJ^H^RT.RV.CL	0.21	SCL-IN.AR^N^H	0.14
DO-DF.DV.GV.NP^H	0.21	SCL-NUM.CD^AJ^H	0.14
DO-DM.DT(THESE)^H	0.21	SCL-NUM.CD^PV.PP(of DF.AR^H)	0.14
DO-H^PP(for)	0.21	SCL-PS.DT^H^PP(of)	0.14
DO-H^PP(to)	0.21	SCL-QT^H	0.14
DO-H^RT.RV.CL	0.21	SB(not)-AJ^H^PP(of)	0.14
DO-IN.AR^AJ^H^PP(for)	0.21	SB(not)-QT^H^RT.RV.CL	0.14
DO-IN.AR^H^PP(about)	0.21	SB(not)-NG.DT^H^PP(of)	0.14
DO-IN.AR^H^AP.TI.CL	0.21	SB(not)-NUM.CD^H^PP(of)	0.14
DO-IN.AR^N^H	0.21	PO-DF.AR^AJ^H^PP(of)	0.14
DO-PS.DT^AJ^H^AP.TI.CL	0.21	PO-PS.DT^H	0.14
DO-PS.DT^H^PP(for)	0.21	NA-AJ^H	0.14
SB-AS.DT^H	0.21	NA-DF.AR^H	0.14
SB-QT^H	0.21	NA-DF.AR^H^PP(of)	0.14
SB-DF.AR^H^PL.ED.CL	0.21	NA-H^TI.CL	0.14
SB-DF.DV.GV.NP^H	0.21	CP.AS(<POCL)-IN.AR^H^PP(of)	0.14
SB-DM.DT(THIS)^H	0.21	CP.AT(<AB.AT.TM)-DM.DT(THAT)^H	0.14
SB-IN.AR^AJ^H^PP(of)	0.21	CP.FOR(<AB.AT.PU)-H	0.14
SB-IN.AR^H^AP.TI.CL	0.21	CP.IN(<AB.AT.MS)-NUM.CD^PV.PP(of NUM.CD^AJ^H)	0.14
SB-N^H	0.21	CP.IN(<AB.AT.RS)-AS.DT^H	0.14
SB-NUM.CD^H	0.21	CP.IN(<AB.AT.RS)-NUM.CD^H	0.14
SB-NUM.GO^H	0.21	CP.IN(<AB.AT.SPA)-AS.DT^H	0.14
SB-PS.DT^H^PP(of)	0.21	CP.IN(<AB.AT.SPA)-DF.AR^H	0.14
SCL-DF.AR^AJ^H^PP(of)	0.21	CP.IN(<AB.AT.SPA)-DF.AR^H^PP(of)	0.14
SCL-NG.DT^AJ^H	0.21	CP.OF(<AB.AT.RS)-AJ^H	0.14
SB(not)-H^RT.RV.CL(when)	0.21	CP.OF(<DO)-DF.AR^AJ^H^PP(of)	0.14
PO-DF.AR^H^PP(of)	0.21	CP.OF(<DO)-H^PP(for)	0.14
PO-H^PP(of)	0.21	CP.OF(<SB)-DF.AR^H^PP(of)	0.14
PO-H^TI.CL	0.21	CP.ON(<AB.AT.RS)-DF.AR^H^RT.RV.CL	0.14
NA-H^PP(of)	0.21	CP.TO(<AB.AT.RS)-AJ^H	0.14
NA-N^H	0.21	CP.WITH(<AB.AT.MAN)-AJ^H	0.14
CP.AS(<PSCL)-H	0.21	CP.WITH(<AB.AT.MS)-AJ^H	0.14
CP.AT(<AB.AT.TM)-DF.AR^H	0.21	CP.WITH(<AB.AT.RS)-DF.AR^H^PP(of)	0.14
CP.OF(<DO)-DF.AR^H	0.21	CP.WITH(<SCL)-DF.AR^AJ^H	0.14

A glance at the top three combinations for the five most frequent syntactic functions reveals distinct preferences. Both subject and direct object share a frequent priming for DF.AR^H (2.35% for the former and 1.89% for the

latter) and DF.AR^H^PP(*of*) (both 1.24%). H, however, makes it into the top three for direct object (1.45%), while DF.AR^AJ^H ranks third with subject (0.90%). The definite nature of subject and direct object contrasts with subject complement, where its top two patterns contain the indefinite article (i.e. IN.AR^H, 1.73%; IN.AR^AJ^H, 1.04%; cf., however, DF.AR^H, ranking third, with 0.69%). Notional subject seems to occur with quantifiers, as evident from its top two patterns (i.e. QT^H and QT H^PP(*of*), both showing 0.28%; H^RT.RV.CL(*when*), with 0.21%, ranks third). Lastly, prepositional object shares the same frequency (0.21%) with its top three patterns (i.e. DF.AR^H^PP(*of*), H^PP(*of*) and H^TI.CL).

Figure 6.42 below narrows the focus to the ten highest-ranking determiners in the study sample (as in the figures in 6.2.2.2 and 6.2.2.3).

Figure 6.42 Syntactic function and determiner use



The clearest associations in Figure 6.42 are found with subject, subject complement and notional subject. Subject noun phrases are strongly primed for the definite article (as in (705)), while subject complement instances are markedly high with the indefinite article (followed by the definite article; see (706a) and (706b)). Notional subjects, by contrast, prevail with quantifiers, zero determiners, negative determiners and cardinal numerals (as in (707a) through (707d)).

(705) **the problem** is Mr Ramos has become a focus of discontent among younger officers
(706) (a) that's **a good word**
(b) Janina's family, Methodism and history turned out to be **the answers**
(707) (a) I wonder why there is **so much crime**

(b) there might be problems
(c) there appears to be no standard creation myth
(d) there's one sole objective

Concerning the other functions, use of direct object shell nouns is similarly distributed across the definite article, the zero determiner and the indefinite article, ranking slighly higher with the definite article (see (708a) through (708c)). Prepositional objects also rank second with the zero article, but are more frequent with the definite article than direct objects (see (709a) through (709b)). Phrase-level examples, like direct objects, do not differ much between the definite and the zero articles, even if priming is higher for the latter (see (710a) through (710b)).

(708) (a) Henry retired to the village of Bolas Magna in Shropshire to escape both the scandal and his heavy gambling debts
(b) anything that can boost overseas ventures for UK groups
(c) the BBC must therefore have a clear vision
(709) (a) looking for evidence of official misdeeds
(b) when you look at the scandal of the share allocations
(710) (a) despite warnings, Mrs Thatcher stressed her leadership role in Europe
(b) one may well speculate on possible causes for the phenomenon

Figures 6.43 and 6.44 look at postmodification structures across syntactic functions. Figure 6.43 shows that postmodification among shell-noun phrases prevails among notional subjects. Postmodification is also frequent with prepositional objects and direct objects, especially in the former (cf., however, the almost equal proportion of modified and unmodified instances of direct object). Lack of postmodification is most frequent with phrase-level, subject and subject complement shell nouns.The difference between modication and the lack thereof is marginal in the latter case.

Figure 6.43 Syntactic function and the occurrence of postmodification

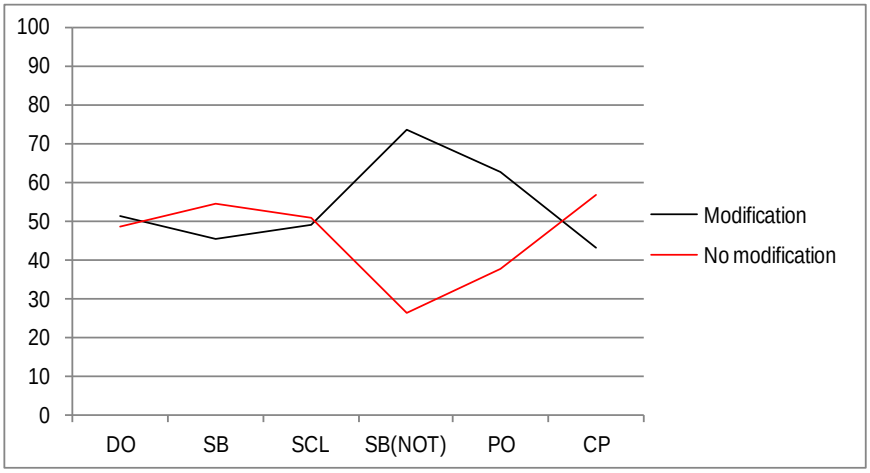
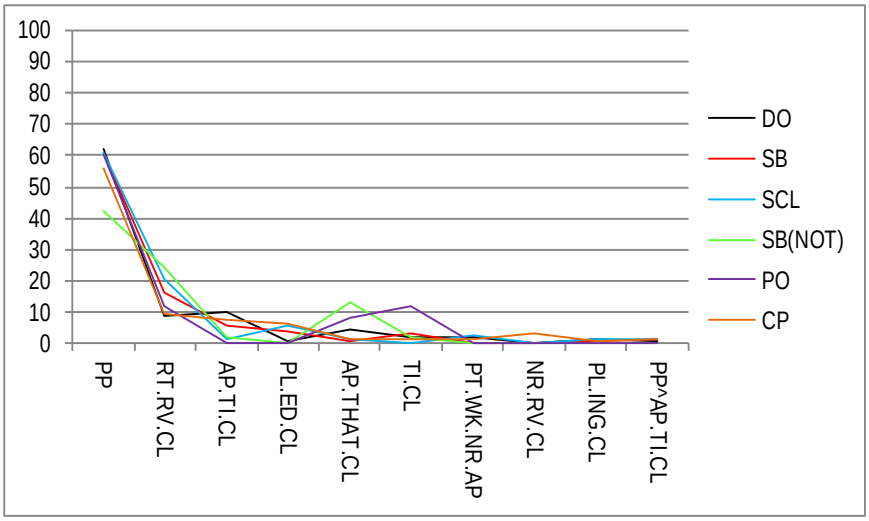


Figure 6.44, as in 6.2.2.2 and 6.2.2.3, is restricted to the top ten postmodifying structures in the sample.

Figure 6.44 Syntactic function and postmodifying structures. Percentages are calculated on the amount of postmodification for each syntactic function (see Figure 6.43)



The most noteworthy results of Figure 6.44 are those regarding appositive *that*-clauses and *to*-infinitive clauses, the former topped by notional subjects and prepositional objects (as in (711a) and (711b)), and the latter occurring most frequently with prepositional objects (as in (712)).

(711) (a) there is a strong suspicion that the billions he has pumped in may not be quite enough
(b) Franklin subscribe to the Apple philosophy that if you can encourage children to use your products, they will continue to use them when they are adults
(712) we continually look for ways to break brewing barriers and set new industry standards

The association is weaker in the case of restrictive relative clauses, appositive *to*-infinitive clauses, participle *-ed* clauses and non-restrictive relative clauses. Restrictive relative clauses are more primed for notional subjects, subject complements and subjects (as in (713a) through (713c)), while appositive *to*-infinitive clauses seem to associate with direct objects, prepositional complements and subjects (as in (714a) through (714c)). Participle *-ed* clauses are slightly more frequent with prepositional complements, subject complements and subjects (as in (715a) through (715c)), while non-restrictive relative clauses occur only with prepositional complements, as in (716). Use of the remaining postmodifying structures reveals a similar distribution across syntactic functions. Prepositional phrases prevail among postmodifying structures, and are used primarily in direct objects, subjects, subject complements and prepositional objects (as in (717a) through (717d)), but they rank lowest among prepositional complements and notional subjects.

(713) (a) there's been a couple of times that he's really lost his temper with her
(b) these are the sorts of problems we get
(c) the visions that they had were not as reality
(714) (a) gastropods have the capacity to store sperm for short periods
(b) numerous reports this year of Ipswich Town's failure to get to grips with the increased demand for tickets
(c) Gen Noriega's capacity to inspire personal loyalty is well known to US intelligence
(715) (a) decided upon the extreme course of assassinating the prime minister − a crime carried out, with pistols, coolly and deliberately
(b) these to her were natural capacities bestowed as a gift
(c) the challenges presented by the developing world do not diminish
(716) yet in all these random examples, which are each dimensions of the new Europe
(717) (a) the local authority made their application for an interim care order
(b) two examples of such projects are discussed in a recent paper
(c) these have become key areas of the department's work
(d) he refused to hearken to Sir Thomas's words of wisdom

Figures 6.45 through 6.48 turn to Transitivity and its association with the semantic premodification of shell-noun phrases. As in 6.2.2.2 and 6.2.2.3, the results draw only on the ten most common types of premodifiers in the study sample. Following the model of 6.2.3.2 and 6.2.3.3, Figures 6.45 and 6.46 are concerned with the ten highest-ranking participant types and circumstances, and Figures 6.47 and 6.48 look at processes and circumstances in general.

Figure 6.45 Participant type and experiential premodification. The percentages shown in Figures 6.45 through 6.48 derive from the distribution of participant types on the basis of the total number of instances for each semantic premodifier (see Table 6.10 in 6.2.2.1 for the overall frequencies of all types of semantic premodification)

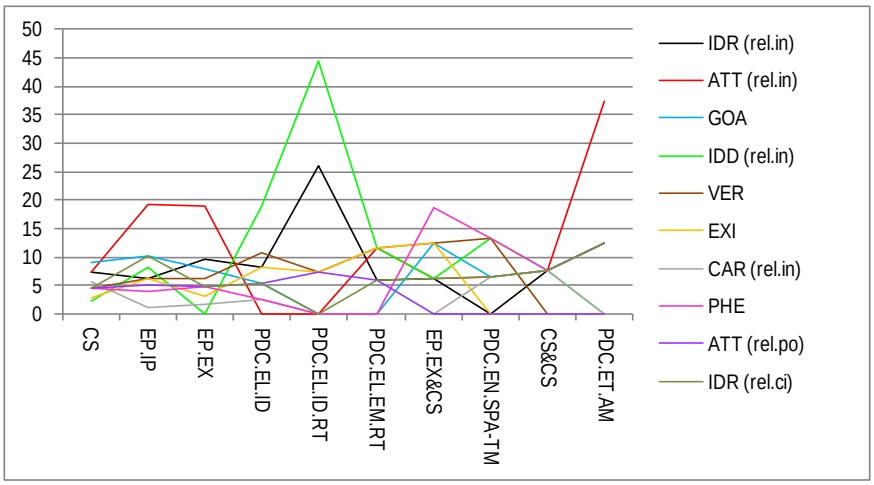


Figure 6.46 Circumstance and experiential premodification

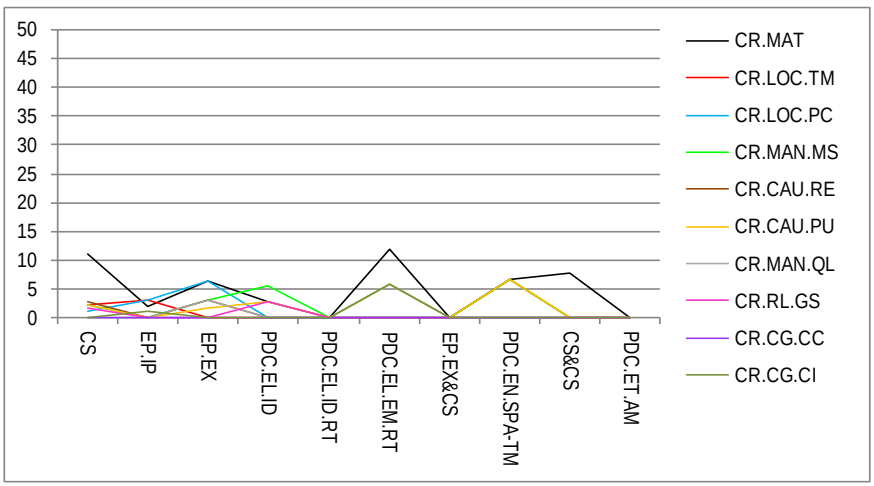


Figure 6.47 Process type (relational subtypes) and experiential premodification

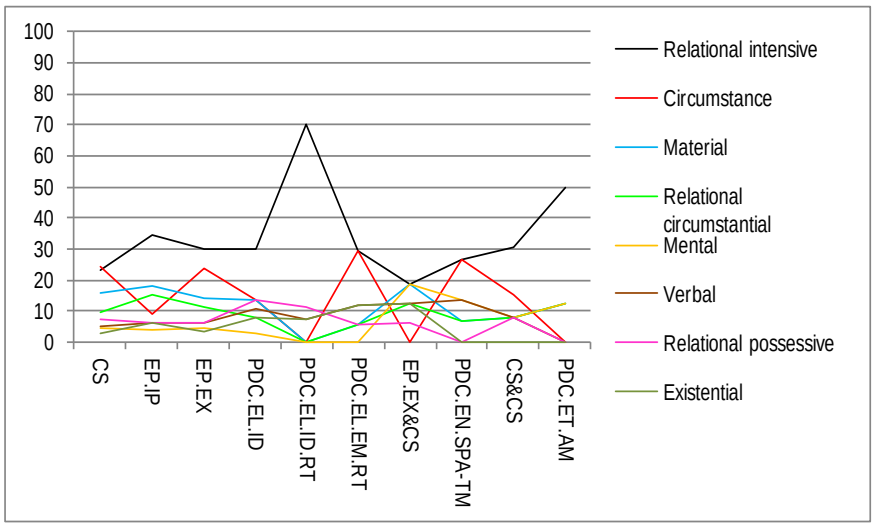
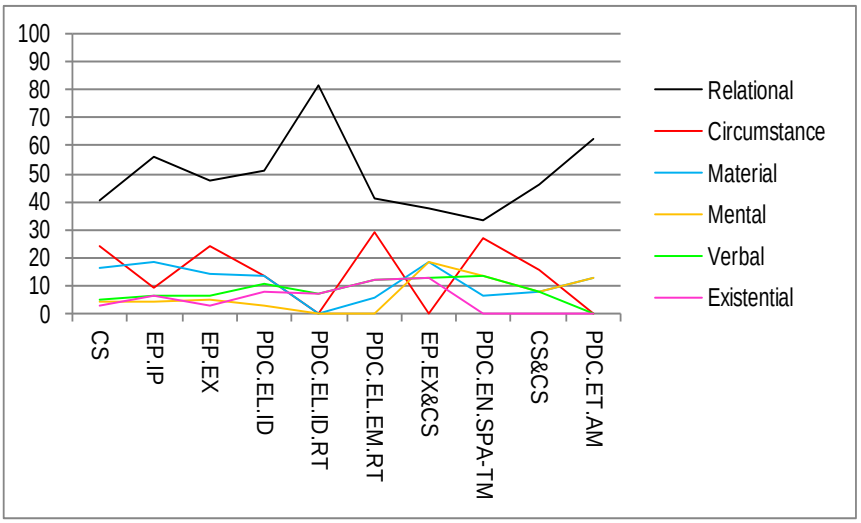


Figure 6.48 Process type (general) and experiential premodification



The main results in Figure 6.45 concern interpersonal and experiential Epithets, and elaborating identity and extending amplifier post-Deictics. Both kinds of Epithets prevail in intensive attributive positions (see (718a) for EP.IP and (718b) for EP.EX). At this point, it is important to recall the pre-eminence of Epithets with non-specific Deictics shown in Figure 6.4. When considered in conjunction with the dominance of the indefinite article with subject complements in Figure 6.42, the association shown of Epithets with intensive Attributes ties in with their frequent occurrence in indefinite

noun phrases. Intensive Attributes are also markedly frequent with extending amplifier post-Deictics (as in (718c)), which, like interpersonal Epithets, carry evaluative nuances. With regard to general and restrictive types of elaborating identity post-Deictics, priming is particularly strong with shell-nouns in intensive Identified positions, the connection being most evident with restrictive instances (see (719a) for PDC.EL.ID and (719b) for PDC.EL.ID.RT). Intensive Identifiers are also prominent with restrictive cases, as in (720). In relation to the remaining premodifying types, only combinations of experiential Epithets and Classifiers show a clearer preference for the participant Phenomenon, as in (721). All the other premodifiers are fairly evenly distributed across participants.

(718) (a) Gloriana was a considerable challenge
(b) unemployment among refugees was only a temporary phenomenon
(c) it was a moment of total triumph
(719) (a) the other way in which you can get more close to live Amnesty is to affiliate as a school group
(b) the main part about it is to remember what you've read
(720) automatic fire fighting installations capable of achieving total control and extinguishment appear to be the only answer
(721) we only have to look at some of the long-term deals, the single union deals and changed working practices to see that

Use of circumstances in Figure 6.46 is clearest with Classifiers (both single and combined) and restrictive elaborating exemplification post-Deictics, where matter circumstances prevail (see (722a) and (722b)). In the case of experiential Epithets and general elaborating identity post-Deictics, a slight association is shown with place and matter by the former, as in (723) and (724), and manner:means by the latter, as in (725). The remaining premodifiers reveal an equal or almost equal proportion of a range of circumstances (interpersonal Epithets and space-time post-Deictics) or a lack thereof (restrictive elaborating identity and extending amplifier post-Deictics, and combinations of experiential Epithets and Classifiers).

(722) (a) ministerial judgements made on the basis of administrative recommendations
(b) the county council was not actually consulted on these particular applications
(723) there may be a very good cause to have a proper course in how to get on with other people, in these sort of cooperative ventures
(724) respond quickly to new circumstances and challenges
(725) public access evolved in two different ways

Turning lastly to Figures 6.47 and 6.48, the most remarkable connections are, as expected from Figure 6.45, those between relational processes (especially, intensive ones) and restrictive elaborating identity and extending amplifier post-Deictics. Circumstances follow relational intensive participants closely in the case of Classifiers, experiential Epithets, and restrictive elaborating exemplification and space-time post-Deictics. Use of the remaining process types is largely concentrated in the 5%–20% frequency range.

6.2.3.5 *Discussion*

The evidence regarding syntactic function in 6.2.3.1 is generally in line with the syntactic behaviour of English nouns. Biber et al. (1999: 235) claim that nouns predominate as objects and prepositional complements, and are less frequent as subjects. The sample contains 29.37% of prepositional complements, 27.16% of direct objects and 20.46% of subjects, which confirms the pattern. Subjects, in view of their close association with given information, are often realised by pronouns (Biber et al. 1999: 236): they are shorter in form and refer to previous information (Ariel 1990: 57; see 3.2.3).

The high prevalence of subject shell nouns in academic and journalistic prose (see Figure 6.25) is confirmed in Biber et al.'s (1999: 236) data, where 75% and 80% of the subjects in these two super-genres are realised by nouns. In fiction, however, only 35% of the subject instances are nouns. The infrequent use of nominal subjects in Biber et al.'s (1999: 236) fiction sub-corpus (35%) contrasts with the more frequent occurrence of pronominal subjects (65%). This might explain the strong connection shown in Figure 6.25 between object shell nouns and fiction and conversation, two super-genres where pronouns are more frequent than in journalistic and academic prose (Biber et al. 1999: 235), and where subject slots are thus more likely to be occupied by pronouns.

By contrast, academic prose, structurally elaborated and abstract at the level of the noun phrase (Biber & Gray 2010: 18), is better suited to bringing second and third-order entity shell nouns to Theme and subject positions (see Figure 6.35 for the higher frequency of Theme shell nouns in academic and journalistic prose). This would apply to newspaper discourse too, where attitudinally loaded nouns like *challenge*, *problem*, *failure*, *warning* or *vision* are often made subjects in order to summarise and bring to the forefront the range of implications underlying a particular news story. It should be noted that, whilst conversation is negatively primed for subject shell nouns in general, it associates strongly with subject shell nouns like *thing*, *point* or *problem*. According to Schmid (2000: 333–4, 334–7), these nouns favour subject positions in focusing constructions, where the aim is not so much to

encapsulate, as to lay emphasis on the point about to be made (see (726)). Conversely, academic and journalistic prose make use of semantically more specific subject shell nouns (e.g. *phenomenon*, *finding*, *example*, *challenge*, *chance* or *problem*; see (727)) which label the preceding discourse, and in so doing, 'face forwards' (Francis 1986: 38), enabling expository written discourse '[...] to move forward by logical and coherent steps [...]' (Halliday 1993: 64).

(726) Yeah I, I, I think that er <pause> the, the, the M S Society have er er say that we've got eighty thousand members in the United Kingdom, Great Britain, Northern Ireland <pause> **only problem** is they appear to have eighty thousand different conditions <laugh> (*BNC Sampler*: KC3, S:conv)

(727) A staggering 52% — mostly females — said they expect rates to be higher in 12 months' time with a consequent knock-on effect on mortgages. This is making people hesitate to go out and spend money. The same survey, looking at a three-month period, told a different story. Seventy per cent then thought bank base rates would either fall or plateau. [...] **The Pearl findings** coincide with the latest Home Owners Confidence Monitor unveiled by the Britannia Building Society. (*BNC Sampler*: CEL, W:newsp:other:commerce)

It thus seems clear that shell nouns in subject or Theme positions are restricted to written expository prose and to emphatic spoken instances. Even so, Rheme is still the default option for the shell nouns in the sample. A possible explanation for the outstanding dominance of Rheme is provided by Lyons (1977, II: 510–11) and Martin (1992: 107), who claim that the chances of Theme occurrences are determined by the centrality of participants. Lyons (1977, II: 510–11) even suggests that only first-order entities (especially people) are suited for Theme positions. On these grounds, it might be argued that the pre-eminence of Rheme shell nouns is due to their abtract nature and to their cohesive effect. Thus, despite the powerful discourse-organising and evaluative functions of shell nouns, their semantic contribution is not as important as that of first-order participants, which may explain why they are often linked to Rheme positions. Only academic prose and newspaper stories seem to cope better with Theme positions. This is logical, in view of the use of nouns in academic prose to freeze '[...] a world swarming with activity [...] so that it can become the object of scientific scrutiny' (Melrose 2003: 427), and, in journalistic prose to glean valuable insights from news stories, and in so doing, to slip in interpretations that the reader is expected to accept.

The results in Figure 6.43 match those obtained by de Haan (1991) and Aarts (2004) (see 3.2.4), as postmodification in the study sample prevails in

post-verbal positions and is relatively infrequent with subject shell nouns. Again, this is not surprising considering the expected priming of pronouns and structurally light nouns for subject positions (see Biber et al. 1999: 236 above and Aarts 2004: 43 in 3.2.4). Figure 6.42 is also consistent with noun behaviour in general, as it corroborates Biber et al.'s (1999: 269) observation that the definite article prefers subject and prepositional complement positions. The definite nature of subjects is explained on the grounds of their association with anaphoric uses [36]. Whilst lacking corroborating evidence, general noun behaviour may also be taken to account for the marked association between subject complement and the indefinite article, and between notional subjects and the zero article, the indefinite article, quantifiers, cardinal numerals and negative determiners. The association between the indefinite article and subject complement stands to reason based on the attributive function of many shell-noun instances (as in (728); see Table 6.23, where intensive Attributes rank second). Similarly, existential or notional subjects are expected to occur in zero article, indefinite article, quantifier, cardinal numeral and negative determiner instances as (729a) through (729e).

(728) it does seem to be a problem
(729) (a) there were sound examples of rare stone
(b) there is a suspicion that a secret cartel is in operation
(c) there are several points to observe
(d) there's two violin practices today
(e) there was no word on price

With regard to Transitivity, the dominance of relational participants in the sample (46.51%, see Figure 6.22b) contrasts with Matthiessen's (1999: 14) findings based on a 14,500-word sample of written English. Whilst highly exploratory and tentative, Matthiessen's article provides tendencies that might reflect the frequency of Transitivity selections in a more representative sample of the language. His results differ from those presented here in that they point to the dominance of material processes (51%), followed by relational (23%), verbal (10%), mental (9%) and existential (2%). Although verbal, mental and existential participants are also rare in the sample, relational processes prevail over material ones. Matthiessen (1995: 245) and (2006: 107) similarly comment on the pre-eminence of material processes in English. Still, these three references acknowledge that different genres will show different distributions, material processes predominating in such text types as instruction manuals and

[36] Cf., however, Fraurud 1990: 404 in 3.2.3, where this view is questioned

novels, and relational processes in expository prose (Matthiessen 1995: 360; 1999: 14; 2006: 107).

A possible explanation for the marked relational nature of shell nouns lies in Halliday & Matthiessen's (2004: 480) claim that their class of fact nouns is best suited for relational contexts, and occasionally, for mental or verbal ones, but never for material processes, insofar as '[...] facts cannot do things, or have things done to them'. Fact nouns in SFG comprise instances straddling Schmid's (2000) factual and modal nouns (e.g. *case*, *point*, *evidence*, *chance*, *possibility*, *obligation*) (Halliday & Matthiessen 2004: 469), which, interestingly, rank highest with relational processes in Figure 6.40. With this in mind, it is logical that factual nouns rank lowest with material processes in the study sample. Material processes are primed for eventive shell nouns, most likely due to their being second-order entities, which, according to Fraurud (1992: 7; see 3.2.1), are more concrete in nature than prototypical third-order factual instances (see, for example, (730a) through (730c)). The occurrence of mental and linguistic nouns as material participants relates to shell nouns that represent the result of a previous mental or verbal process, as in (731a) through (731b).

(730) (a) commit a crime
(b) do that work
(c) other misfortunes might have occurred
(731) (a) launching a major research project
(b) this motion is not an attempt to preempt applications

Just as noteworthy as the correspondence between relational processes and factual and modal nouns, and between material processes and eventive nouns is the link between verbal processes and linguistic nouns, between mental processes and linguistic and mental nouns, and between circumstances and circumstantial nouns. These connections suggest associations between process categories and semantic types of shell nouns. Whilst all shell nouns are most primed for relational participants, a cline is apparent in their instantiation with the other process types, such that the more abstract (and third-order) the shell noun is, the more likely it is to occur in relational contexts, and the closer it is to concrete meanings, the stronger the priming for material processes.

6.2.4 *Direction of encapsulation and antecedent*

This section is about the discursive behaviour of shell nouns in terms of their relationship with an antecedent. Thus, it brings together all the findings on encapsulating directions, types of antecedent and the link between both.

6.2.4.1 *Overall distribution*

Figure 6.49 below represents the extent of use of the main encapsulation types in the sample. Cataphora appears as the most frequent option: 40.29% of shell nouns have their antecedent in a following discourse segment. Anaphora and exophora follow: 30.75% of cases correspond to the former and 24.05% to the latter. Combinations of any of the three main encapsulation types are not frequent (4.49%). The rest are unclear examples (0.41%). Overall, 75.53% of the evidence comprises cases where the antecedent occurs in the text (i.e. cataphora, anaphora, combined), and 24.05% is made up of extra-textual or exophoric instances.

Figure 6.49 General encapsulating directions

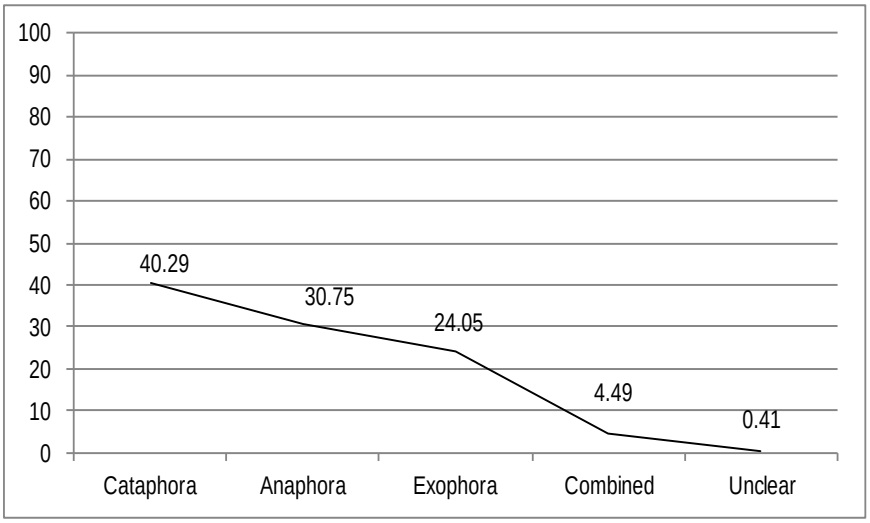


Table 6.30 is narrower in scope than Figure 6.49, as it specifies the intra- or intersentential realisation of endophoric encapsulation. According to Table 6.30, the sample contains 26 different types of encapsulating relations, the top ten making up 96.61% of all shell-noun instances.

Table 6.30 Specific encapsulating directions

Direction of encapsulation (%)			
EXO	24.05	AF.INTRA&AF.INTER(SN2)	0.35
CF.INTRA	23.29	AF.INTRA&CF.INTRA	0.35
AF.INTER	17.28	AF.INTRA&CF.INTER	0.28
CF.INTER	16.03	AF.INTRA&AF.INTER&CF.INTER	0.21
AF.INTRA	8.57	CF.INTRA&EXO	0.14
AF.INTRA&AF.INTER	3.52	CF.INTRA^AF.INTER(SN2)	0.14
AF.INTER&CF.INTER	1.45	AF.INTER&AF.INTER(SN2)	0.07
AF.INTRA&AF.INTRA	0.90	AF.INTER&CF.INTRA&CF.INTER	0.07
AF.INTER&CF.INTRA	0.76	AF.INTRA&AF.INTER&AF.INTER(SN2)	0.07
CF.INTRA&CF.INTER	0.76	AF.INTRA&AF.INTER&CF.INTRA	0.07
AF.INTER&EXO	0.62	CF.INTER&CF.INTER(SN2)	0.07
AF.INTRA&EXO	0.41	CF.INTRA&CF.INTER(SN2)	0.07
UNC	0.41	CF.INTRA&CF.INTRA	0.07

Extra-textual realisation or exophora ranks highest in the list (24.05%). The prevalence of cataphora in Figure 6.49 is now observed to relate to intrasentential examples, featuring second in the list (23.29%). This said, almost half of the evidence in the study sample (47.34%) consists of exophoric and intrasentential cataphoric instances. Intersentential anaphora ranks third (17.28%), followed by intersentential cataphora (16.03%). Intrasentential anaphora, unlike its cataphoric counterpart, is relatively infrequent (8.57%). The remaining encapsulating categories comprise combinations of intra- and intersentential realisations of a single encapsulating direction, as well as combinations of both anaphora and cataphora, the latter making up the 4.49% of combined instances in Figure 6.49. Five frequent combinations occur in the top ten of Table 6.30. Intra- and intersentential anaphoric combinations are the most frequent (3.52%), followed by instances lexically realised by intersentential anaphoric and cataphoric antecedents (1.45%). The use of double intrasentential anaphora (0.90%), intersentential anaphora and intrasentential cataphora (0.76%), and intra- and intersentential cataphora is more limited (0.76%). Examples (732) through (741) illustrate the top ten encapsulating relations.

(732) [EXO] And I think you've gotta address my doubts because if I have doubts how am I going to <unclear> stand in front of members and say <pause> hang on a minute, I can give you **all of these answers**, I can tell you what unison's going to be like and this is what's been decided, now come back to me and give me your fears, give me your doubts and I'll go and take them on to represent you. (*BNC Sampler*: F7J, S:meeting)

(733) [CF.INTRA] A dogwhelk's tolerance of reduced salinity is probably low, despite Pelseneer's (1935) statement that they can survive 9.5 days in fresh water (Moore 1938b) and Agersborg's (1929) **finding of living**

animals in freshwater pools at high water mark in Norway. (*BNC Sampler*: FU0, W:ac:nat_science)

(734) [AF.INTER] [...] the American airline took the unprecedented step of revealing a bomb threat. [...] the airline's head office in St Paul, Minnesota, revealed that a man with a Middle Eastern accent had telephoned the company and threatened Flight 51. [...]A Northwest spokesman in St Paul, Mr Douglas Miller, said that the latest threat was unusual as a specific flight and day had been mentioned by the caller, who warned of reprisals for the jailing of two Palestinians for a series of bombings. [...] The deputy chief of France's air and border police, Mr Pierre Quilici, said the warning had been passed on by the Americans several days ago. (*BNC Sampler*: AAT, W:newsp:brdsht_nat:report)

(735) [CF.INTER] Finally members of the jury by way of opening tonight, ask you to be warned about two things. Firstly <pause> it is inevitable and natural and quite proper that you should have a sympathy, in particular for Mrs <gap desc="name" reason="anonymization"> and her children and <gap desc="name" reason="anonymization">. None of us <unclear> it is a matter of extreme regret that they were terrified in the way they claim they were, but that sympathy must not be allowed to cloud your judgement on the issue about whether the police were reasonable. <pause> Secondly members of the jury, please don't do what the press are always doing which is act on the basis of hyper. If the police had a crystal ball <pause> and a gypsy woman they could bring in and ask her to look and see what was in the flat, we wouldn't be here today would we? But they don't have a crystal ball, all they have is their judgement and they do what's best in the circumstances. (*BNC Sampler*: JJV, S:courtroom)

(736) [AF INTRA] Peter Scudamore's late decision to partner Granville Again was no surprise. (*BNC Sampler*: CF9, W:newsp:other:sports)

(737) [AF.INTRA&AF.INTER] 'Someone once said that playing patience was the nearest thing to being dead.'
'Really, Miss Danziger! Well I never; it's not like you to be offensive!'
I must get out of here or I shall say other things. It's best to be alone when the noises get this loud. [...]
'Was that a joke, Miss Danziger?'
'I don't think so. It was an American poet who made the observation. He was a solemn man.' (*BNC Sampler*: AEA, W:fict:prose)

(738) [AF.INTER&CF.INTER] In recent years tourism has also made an increasing impact on farming, particularly in the more scenically attractive and marginal farming areas, and in Scotland, at least 20 per cent of farms indulge in farm tourism, and farm tourism accounts for at least 10 per cent of net income (Denman, 1978). Similar findings have been reported by Davies (1983) for the 'less favoured areas' (the uplands) of England and Wales. In these areas 20 per cent of the farms sampled indulged in farm tourism, but most farmers saw it as a supplement rather than as an alternative to farm income, although

Davies argued that it could be a much <gap desc="table"><gap desc="table">more profitable use of the farms' resources. (*BNC Sampler*: FR2, W:ac:soc_science)

(739) [AF.INTRA&AF.INTRA] How you attack those markets can be different in the sense that in er London the penetration of erm er mineral water into the target market is very high and therefore it's already an accepted practice. Whereas in Yorkshire the penetration and usage is very low perhaps the lowest in the country erm where there is still very much a heritage of, Well what's wrong with what's in t' tap lad? (*BNC Sampler*: FUG, S:unclassified)

(740) [AF.INTER&CF.INTRA] Okay but [...]
<-|-> If we going out <unclear> <-|->
<-|-> you'd have to <-|-> you'd have to come and get me and we'd <-|-> go straight off <-|-> . [...]
The other thing erm is that I haven't got any leave. (*BNC Sampler*: KE3, S:conv)

(741) [CF.INTRA&CF.INTER] And of course this created this created a a tr-- a tremendous problem, because er these few people that went to the er odd pits were in need of er a very strong police escort. <pause> And there were hundreds of police who were <pause> drafted into the into the town, <pause> er billeted on the er local pubs, er and so on. And d-- it was the duty of these police to protect these er er these people, these scabs as they were called, <pause> er and escort them from their homes to the pit, and see them back. [...] when er a couple of these er er people were being escorted back towards the centre of the town, after they'd done er a day's work <pause> and there were lots and lots of er er er people about, men and women, who were shouting and jeering, at er at er at these at these people, who had been er who had er er violated the union decision and gone to work [...] (*BNC Sampler*: FYJ, S:interview:oral_history)

Tables 6.31 through 6.32b show the tags for direction of encapsulation blended with those for antecedent. The combined tags contain the most fine-grained coding for each variable, including antecedent types and specifics of identity[37]. The sample contains 193 different combinations, 16 of which account for 51% of the evidence. Table 6.31 presents only these 16 combinations.

[37] Note that the findings in Figure 6.49 and Table 6.30 above leave (SI) codes out of the count, because specifics of identity narrow the reference of the head noun but do not detail what the head noun is or involves (i.e. are not proper antecedents). Thus, the count in Figure 6.49 and Table 6.30 for a combined tag as CF.INTRA^AF.INTER&LC.CT(SI)^GB.ET.PR, draws only on the AF.INTER part of the tag, because CF.INTRA corresponds to specifics of identity, realised in this case by a complement prepositional phrase.

Table 6.31 Specific encapsulating directions and their antecedents (Top 50%). See Appendix 12 for the complete list

Encapsulation and antecedent (Top 50%)			
EXO^NA	8.09	CF.INTRA^EXO&LC.CL(SI)^NA	2.35
AF.INTER&GB.ET.PR	5.74	CF.INTRA^CF.INTER&LC.CT(SI)^GB.ET.PR	2.28
CF.INTRA&LC.CT	4.77	CF.INTRA^AF.INTER&LC.CT(SI)^GB.ET.PR	1.87
CF.INTRA&LC.CL	4.49	AF.INTER&LC.CL	1.52
CF.INTRA^EXO&LC.CT(SI)^NA	4.35	CF.INTRA^EXO&LC.PM(SI)^NA	1.52
CF.INTER&GB.ET.PR	3.73	AF.INTRA^AF.INTER&LC.PRM(SI)^GB.ET.PR	1.31
CF.INTRA&LC.CL(AP)	3.73	AF.INTRA&LC.SNP	1.17
AF.INTRA^EXO&LC.PRM(SI)^NA	2.97	AF.INTRA&LC.CNP	1.11

The list is topped by EXO^NA and AF.INTER&GB.ET.PR, followed by CF.INTRA&LC.CT and CF.INTRA&LC.CL. This means that global extended discourse is the most frequent option for instances of intersentential anaphora, whilst complement prepositional phrases and clauses (non-appositive or non-complement: 4.49%, and appositive or complement: 3.73%) rank highest with intrasentential cataphora. The prominence of exophora, intrasentential cataphora and intersentential anaphora in Table 6.31 (19.28%, 12.99% and 10.44%, after exclusion of (SI)) contrasts with the limited use of intersentential cataphora and intrasentential anaphora (6.01% and 2.28%). This is in line with what Table 6.30 shows. Intersentential cataphora, like its anaphoric equivalent, prevails with global extended discourse, while intrasentential anaphora is more primed for simple or complex nominal antecedents.

With regard to use of (SI), 11.19% of the exophoric tags in Table 6.31 occur with specifics of identity (5.87% for cataphoric prepositional complements or postmodifiers, 2.97% for anaphoric premodifiers, and 2.35% for cataphoric intrasentential clauses). Intersentential anaphora features only 3.18% of (SI) instances, which fall into cataphoric prepositional complements (1.87%) and anaphoric premodifiers (1.31%). Intersentential cataphora contains 2.28% of (SI) use, realised, in this case, only by cataphoric prepositional complements. There are not any cases of (SI) in Table 6.31 for intrasentential cataphora and anaphora.

Following the data of Table 6.31, Tables 6.32a and 6.32b display the distribution of the 193 combined tags across anaphora, cataphora, exophora and combinations. Given the size of the lists, the results for anaphora, cataphora and exophora are limited to the top 50% tags in each case (9 out of 140 tags, i.e. 50.34% for anaphora; 7 out of 154 tags, i.e. 52.14% for cataphora; and 2 out of 38 tags, 51.72% for exophora). The data for combinations comprise 30.77% of the whole list (7 out of 52), as all remaining tags occur once in the sample (1.54%).

Table 6.32a Distribution of specific encapsulating directions and antecedents (combined) across the four general encapsulating directions (Top 50%). Percentages are based on the total number of instances for each encapsulating relation. See Appendix 13 for the complete lists

Anaphora (Top 50%)	%	Cataphora (Top 50%)	%
AF.INTER&GB.ET.PR	18.65	CF.INTRA&LC.CT	11.84
CF.INTRA^AF.INTER&LC.CT(SI)^GB.ET.PR	6.07	CF.INTRA&LC.CL	11.15
AF.INTER&LC.CL	4.94	CF.INTER&GB.ET.PR	9.26
AF.INTRA^AF.INTER&LC.PRM(SI)^GB.ET.PR	4.27	CF.INTRA&LC.CL(AP)	9.26
AF.INTRA&LC.SNP	3.82	CF.INTRA^CF.INTER&LC.CT(SI)^GB.ET.PR	5.66
AF.INTRA&LC.CNP	3.60	AF.INTRA^CF.INTER&LC.PRM(SI)^GB.ET.PR	2.57
AF.INTER&LC.SC	3.15	CF.INTRA^CF.INTER&LC.CL(SI)^GB.ET.PR	2.40
AF.INTRA&LC.PRM	2.92		
CF.INTRA^AF.INTRA&LC.CT(SI)^LC.SNP	2.92		

Table 6.32b Distribution of specific encapsulating directions and antecedents (combined) across the four general encapsulating directions (Top 50%)

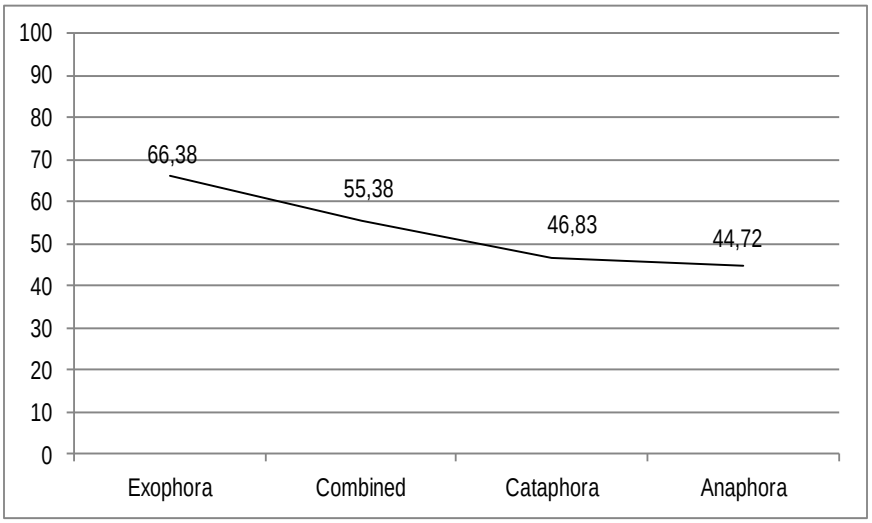
Exophora (Top 50%)	%	Combined (Top 30.77%)	%
EXO^NA	33.62	AF.INTER^CF.INTER&GB.ET.PR	7.69
CF.INTRA^EXO&LC.CT(SI)^NA	18.10	AF.INTER^EXO&GB.ET.PR^NA	4.62
		CF.INTRA^AF.INTER^CF.INTER&LC.CT(SI)^GB.ET.PR	4.62
		CF.INTRA^AF.INTER^CF.INTER&LC.PM(SI)^GB.ET.PR	4.62
		AF.INTER^CF.INTER&LC.SNP^LC.SC	3.08
		AF.INTRA^EXO&LC.CL^NA	3.08
		AF.INTRA^EXO&LC.SNP^NA	3.08

As expected from the data in Table 6.31, anaphoric uses in Table 6.32a are topped by intersentential realisations through global extended antecedents (18.65%). Showing a 12.58% difference with the highest-ranking anaphoric tag, there are instances of global extended intersentential realisations with prepositional (SI). Intrasentential anaphoric uses are, as is also evident in Table 6.31, most frequent with simple and complex noun phrases (3.82%, 3.60%). Turning to cataphora, intrasentential realisations through prepositional complements and clauses prevail (11.84% for LC.CT and 20.41% for LC.CL, out of which 11.15% are non-complement or non-appositive clauses and 9.26% are appositive or noun complement clauses). Intersentential global extended antecedents (excluding (SI)) are more frequent in the top 50% for anaphora than they are with cataphora (22.92% vs. 19.89%). Use of exophora is, as shown in Table 6.32b, most common on its own and in combination with cataphoric prepositional specifics of identity. Concerning combinations, 16.93% of the 30.77% in Table 6.32b (excluding (SI)) is made up of intersentential anaphora and cataphora realised by global extended antecedents. The remaining 13.84% consists in intra- or intersentential anaphoric and exophoric encapsulation, as well as

realisations of the AF.INTER^CF.INTER combination through simple noun phrases and sentences.

Table 6.31 suggested a higher frequency of specifics of identity among exophoric tags. As results in that table draw only on the top 50% combined tags, such a hypothesis needs to be examined on the basis of the overall evidence for each main encapsulating relation. Figure 6.50 looks at the proportion of tags including (SI) in each kind of encapsulation. Interestingly, the figure confirms the high frequency of specifics of identity among exophoric uses, followed by combinations, cataphora and anaphora. Whilst the occurrence of (SI) is half or almost half as frequent as the lack thereof for combined, cataphoric and anaphoric uses, its higher frequency for exophoric uses might be explained on the grounds that, in the absence of endophoric informative specifics, identity specifics in the shape of internal modifiers provide provisional interpretative 'anchors' (Fraurud 1990: 424). These enable the readers' or listeners' search for their background knowledge for the interpretation of the exophoric noun.

Figure 6.50 Use of (SI) tags in the four general encapsulating directions



Examples of (SI) use for the four main encapsulating relations in Tables 6.32a and 6.32b are given in (742) through (745) below. All examples contain a prepositional complement (SI) (i.e. *of a military alliance between two states*, *of dealing with these risks when they are unwanted*, *of its prospects for the future*, and *of the Board's work*), but (742) is intersential global extended anaphoric, (743) is global extended cataphoric, (744) is exophoric, and (745) is global extended anaphoric and cataphoric.

(742) [CF.INTRA^AF.INTER&LC.CT(SI)^GB.ET.PR] The Teleuts made an agreement (shert) with Muscovite voevody, but never consented to pay yasak during the whole seventeenth century, even when offered the inducement of paying a nominal tenth of the usual quantity. They did give presents (pominki), but these were voluntary. It seems that this was their means of asserting their continuing independence of Moscow [...] Nor did the Teleuts hand over hostages to the Russians, despite pressure being placed on them to do so. [...] But Umanskii sees a significant difference: feudal subjects of the seigneur are usually from the same nation as him. The Teleuts ,however, had been sovereign in their own domains for a long period before the Russians arrived on the scene. Rather than being a case of feudal vassaldom, this has more of the characteristics of a military alliance between two states, an unstable one, it is true, and between partners of very unequal power. (*BNC Sampler*: FB4, W:ac:humanities_arts)

(743) [CF.INTRA^CF.INTER&LC.CT(SI)^GB.ET.PR] There are a number of ways of dealing with these risks when they are unwanted as opposed to being deliberately chosen in the hope of making capital gains. Market makers can aim to run only modest positions in any group of similar stocks or they can run large but offsetting positions, i.e. being long on certain stocks and short on others such that capital gains and losses can be expected to cancel each other out. A further means of risk management is provided by the derivative markets in futures and options in which GEMMs can broadly offset spot positions by dealing in bond futures or options. (*BNC Sampler*: HY1, W:commerce)

(744) [CF.INTRA^EXO&LC.CT(SI)^NA] The booklet Angola: The Possible Peace provides a concise analysis of the country's history and an assessment of its prospects for the future. It asks what Angola's place in the regional political order is likely to be. (*BNC Sampler*: EBK, W:misc)

(745) [CF.INTRA^AF.INTER^CF.INTER&LC.CT(SI)^GB.ET.PR] [...] although other parts of the Board's work, and the Board's investicl-- invested capital have performed well from a financial point of view, the anticipated deficit on thirty eight eventide homes has been sustained. [...] We see this underfunding of eventide care as limiting the choice available to people who have come to that point in their lives when they're making decisions about the sort of care that they will next require. [...] All other parts of the Board's work have been budgeted within acceptable limits. And I should say to the general assembly that the Board's four homes for people with senile dementia are differently funded and any deficits there are made up by grants from health boards and social work departments which are prohibited from making up the deficit in eventide care. During nineteen ninety one the Board has been delighted to open new areas of work in Inverness where our first designated place and associated hostel was opened on a most happened-- happy day by Sir Russell <gap desc="name"

reason="anonymization">. In Elderslie near Paisley <pause> where Lady <gap desc="name" reason="anonymization"> the wife of last year's Lord High Commissioner opened our fourth senile dementia unit. (*BNC Sampler*: F86, S:meeting)

Table 6.33 focuses on the extent of use of individual antecedents for the intra- and intersentential realisations of anaphora and cataphora. The count records the total number of antecedent types found with AF.INTRA, AF.INTER, CF.INTRA and CF.INTER, regardless of whether these occur alone (e.g. AF.INTRA) or in combination (e.g. AF.INTRA&AF.INTER). Thus, for CF.INTRA^AF.INTRA^AF.INTER&LC.CT(SI)&LC.SNP(*THIS*)&LC.CNP, the demonstrative antecedent LC.SNP(*THIS*) receives one token within the AF.INTRA column, and LC.CNP is given another token within the AF.INTER column. Similarly, in combined cases like AF.INTER^CF.INTER&GB.ET.PR, one token is assigned to GB.ET.PR in AF.INTER and one in CF.INTER.

Table 6.33 Distribution of antecedents across intra- and intersentential anaphora and cataphora. Percentages are based on the total number of antecedent tokens occurring with each encapsulating relation

AF.INTRA	%	AF.INTER	%	CF.INTRA	%	CF.INTER	%
LC.SNP	27.43	GB.ET.PR	59.33	LC.CL	31.37	GB.ET.PR	66.91
LC.CNP	15.93	LC.CL	10.58	LC.CT	20.64	LC.SC	10.55
LC.CL	13.27	LC.SC	8.36	LC.CL(AP)	19.03	LC.CL	6.18
LC.SNP(IT)	11.50	LC.SNP	6.13	LC.CNP	8.58	LC.CNP	2.18
LC.SNP(THAT)	9.73	LC.CNP	5.29	LC.SNP	5.36	LC.SNP	2.18
LC.PRM	6.19	LC.ET.PR.CL	2.51	LC.CL(NR.AP)	2.95	LC.ET.PR.SC	1.45
LC.SNP(THIS)	5.75	LC.ET.PR.SC	1.11	LC.CL[QU]	1.61	LC.SC[QU]	1.45
LC.SNP(THESE)	2.21	LC.ET.PR.SNP	1.11	LC.SNP(NR.AP)	1.61	LC.CL[QU]	1.09
LC.AJ	1.33	LC.SC[QU]	1.11	LC.SNP^CNP	1.34	LC.TB	1.09
LC.SNP^CNP	1.33	GB.ET.PR[QU]	0.84	LC.CNP(NR.AP)	1.07	OV	1.09
GB.ET.PR	0.44	LC.CL[QU]	0.84	LC.SC	1.07	LC.ET.PR.CL	0.73
LC.AJP	0.44	LC.ET.PR.CNP	0.84	LC.SNP(RT.AP)	0.80	LC.FG	0.73
LC.CL[QU]	0.44	LC.ET.PR.SC[QU]	0.84	GB.ET.PR[QU]	0.54	LC.FR	0.73
LC.CNP^CL	0.44	LC.AJ	0.28	LC.PM	0.54	GB.ET.PR[QU]	0.36
LC.PP	0.44	LC.CT	0.28	LC.PP	0.54	LC.AJ	0.36
LC.SC	0.44	LC.ET.PR.SNP[NUM]	0.28	LC.SNP^CNP(NR.AP)	0.54	LC.CT	0.36
LC.SC[QU]	0.44	LC.ET.PR.TB	0.28	GB.ET.PR	0.27	LC.ET.PR.SC[QU]	0.36
LC.SNP(THEM)	0.44			LC.AJ	0.27	LC.ET.PR.TB	0.36
LC.SNP(THEY)	0.44			LC.CNP(RT.AP)	0.27	LC.IJ	0.36
LC.SNP(THOSE)	0.44			LC.PP(NR.AP)	0.27	LC.SNP[NUM]	0.36
LC.SNP[QT.PN]	0.44			LC.RN.SG	0.27	LC.SNP^CNP	0.36
LC.VR	0.44			LC.RN.SG[QU]	0.27	LC.SNP^CNP[QU]	0.36
				LC.SNP(THAT)	0.27	LC.SNP^LC.AJ	0.36
				LC.SNP(THESE)	0.27		
				UNC(NA)	0.27		

The results in Table 6.33 are in line with the data of Tables 6.32a and 6.32b. Intrasentential anaphora is most frequent with simple and complex noun phrase antecedents (both accounting for 43.36%), as well as with non-appositive clauses (13.27%). Personal and demonstrative pronouns follow closely, the most prominent being *it* and *that* (11.50% and 9.73%; cf. *this* and *these*, 5.75% and 2.21%). By contrast, intersentential anaphora is strongly primed for global extended antecedents (59.33%). Clauses, sentences and noun phrases show a weaker association with this encapsulating direction: their frequencies range from 10% to 5.29%. Non-appositive clauses rank highest with intrasentential cataphora (31.37%), followed by prepositional complements and appositive or complement clauses (20.64% and 19.03%). Nouns rank fourth and fifth and are slighly more frequent as complex phrases than as simple ones[38] (8.58% and 5.36%). Lastly, intersentential cataphora shows a strong priming for global extended antecedents, even more so than intersentential anaphora (66.91% vs. 59.33%). Clauses, sentences and noun phrases, as in intersentential anaphora, also follow extended antecedents. The only difference between both lies in the more frequent occurrence of sentential antecedents with intersentential cataphora (10.55% vs. 8.36%) and of clausal ones with intersentential anaphora (10.58% vs. 6.18%). The following examples are illustrative of the top three antecedents in each encapsulating relation in Table 6.33:

(746) [AF.INTRA&LC.SNP] Although rape and looting were technically **crimes** according to both armies rules of conduct, the commanders considered it a necessity to 'cleanse' their enemies of their women. (*BNC Sampler*: H8W, W:essay:univ)

(747) [AF.INTRA&LC.CNP] Furthermore, while the Constitution guarantees freedom of expression and freedom of the media, it does not make explicit provision for free and equal access to the media for all citizens — **an equally desirable objective**. (*BNC Sampler*: EBK, W:misc)

(748) [AF.INTRA&LC.CL] And then she tried to put it in the washing machine for a **joke** and the dog was gonna get in. (*BNC Sampler*: KPG, S:conv)

(749) [AF.INTER&GB.ET.PR] What happened was that the water first backed up behind the lipped edge of the top step, then over-flowed, and spilled with increasing force, in a sort of hydro-chain-reaction, down each subsequent step to the bottom of the channel. There followed a period of comparative quiet, while the water built up again behind the top step and those beneath. You might guess it was my dad who first pointed out **this odd (and classically Chaotic) phenomenon** and brought it to the attention of us kids. (*BNC Sampler*: G0A, W:fict:prose)

[38] Cf., however, AF.INTRA and AF.INTER, where simple noun phrases predominate over complex ones, this being particularly marked with AF.INTRA.

(750) [AF.INTER&LC.CL] At the same time it is recommended that sufficient potassium permanganate crystals be added to the water to turn it violet. What this latter achieves I am not quite sure, but I have yet to hear of a gardener who has experienced any problems with free lime after following this recommendation. (*BNC Sampler*: GV1, W:misc)

(751) [AF.INTER&LC.SC] Do you want to tell us what interest there was. I, I was about to <-|-> ask the same thing, yeah <-|->. (*BNC Sampler*: DCH, S:meeting)

(752) [CF.INTRA&LC.CL] And our philosophy as a truly worldwide supplier of accounting software is that we should increase our presence in different parts of the world as our sales and marketing activity there merits it. (*BNC Sampler*: HDF, S:speech:unscripted)

(753) [CF.INTRA&LC.CT] It would appear that any suspicion of undesirable interaction between roof ventilation and sprinklers would be unjustified in the case of a zoned sprinkler system. (*BNC Sampler*: G0K, W:institut_doc)

(754) [CF.INTRA&LC.CL(AP)] He said that during his recent visit to Moscow, President Gorbachev had given a warning that the failure of East Germany's reform process would not only destabilise the GDR but also have "serious repercussions" for perestroika in the Soviet Union. (*BNC Sampler*: A9M, W:newsp:brdsht_nat:report)

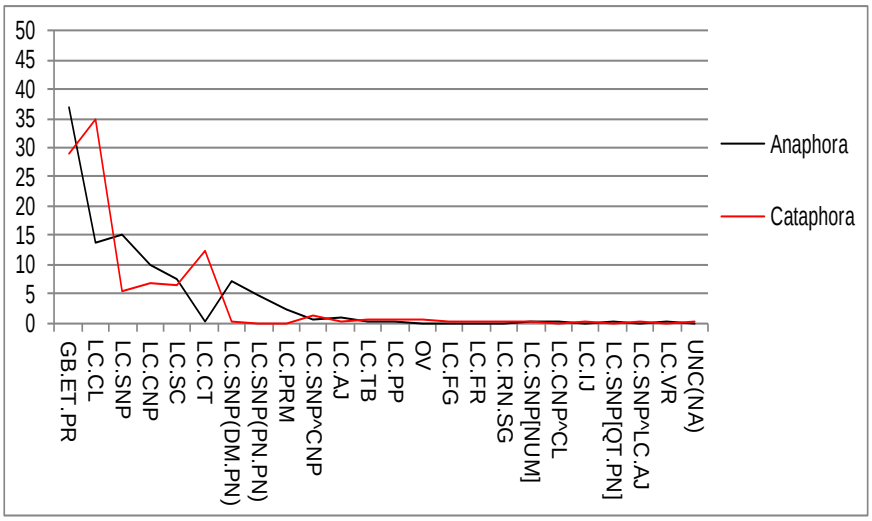
(755) [CF.INTER&GB.ET.PR] Nevertheless, prepare yourself for some surprises when you cross the borders. We have wrongly formed a grim and forbidding picture in our minds but instead you will find beautiful scenery, historic cities and friendly and welcoming people. The holiday resorts in our programme have a wide appeal. In Poland, Zakopane, in the mountains, offers wonderful scenery and Gdynia on the Baltic Coast has all the attractions of the beach plus lovely countryside and interesting cities to visit. In Czechoslovakia , we now have two resorts, we return to our popular holiday centre just outside the historic city of Brno and also to Roznov near the Beskydy mountains. In Hungary, Lake Balaton is larger than Lakes Geneva or Garda and is a well developed and long established holiday region. (*BNC Sampler*: CAA, W:advert)

(756) [CF.INTER&LC.SC] The most passionate prejudice was reserved for the Jews. Resentment of their success in the Leeds business world acted to explain anyone else's failure, without questioning individual acumen or the inherent virtue of capitalism. (*BNC Sampler*: FU7, W:non_ac:soc_science)

(757) [CF.INTER&LC.CL] Er you need only accept a credit note if the problem is your own fault in some way. For example if you've changed your mind and decide you didn't want the goods or if you had damaged the goods and the shop offer you a credit note, I think you really have to accept that with good grace, cos it is really your fault. (*BNC Sampler*: FUT, S:speech:unscripted)

Figure 6.51 merges all intra- and intersentential results of Table 6.33 under anaphora and cataphora. Unlike that table, antecedents are given in their most general form, conflating all subtypes into wider categories. For example, appositive, quoted and local extended preceding clauses (i.e. LC.CL(AP), LC.CL[QU] and LC.ET.PR.CL) appear under LC.CL. Similarly, all cases of pronominal antecedents appear under LC.SNP(DM.PN) (if a demonstrative pronoun) or LC.SNP(PN.PN) (if a personal pronoun).

Figure 6.51 Distribution of antecedents across anaphora and cataphora



Anaphora stands out compared with cataphora with global extended, nominal (simple and complex), sentential and pronominal antecedents. By contrast, cataphoric uses rank higher with clausal and prepositional complement antecedents. The most substantial difference in Figure 6.51 concerns clausal antecedents, where 35.03% are cataphoric and 13.85% anaphoric. Also worth noting is the absence of prepositional complement antecedents with anaphora (cf. 12.35% for cataphora) and the marginal and almost unnoticeable occurrence of pronominal antecedents with cataphora (only 0.31% for demonstrative pronouns[39]).

Table 6.34 presents the overall proportion of antecedent use in the study sample, regardless of the distinction between anaphora and cataphora.

[39] Cf. anaphora, where demonstrative pronouns account for 7.01% of antecedents, while personal pronouns comprise 4.79%.

Table 6.34 Antecedent types in the study sample. Percentages are based on the total number of antecedent tokens in the sample (1233). The list of antecedents follows the one in Figure 6.51

Antecedent type (overall) (%)			
GB.ET.PR	32.85	LC.PP	0.32
LC.CL	24.98	OV	0.24
LC.SNP	9.98	LC.FG	0.16
LC.CNP	8.19	LC.FR	0.16
LC.SC	6.89	LC.RN.SG	0.16
LC.CT	6.57	LC.SNP[NUM]	0.16
LC.SNP(DM.PN)	3.49	LC.CNP^CL	0.08
LC.SNP(PN.PN)	2.27	LC.IJ	0.08
LC.PRM	1.14	LC.SNP[QT.PN]	0.08
LC.SNP^CNP	0.97	LC.SNP^LC.AJ	0.08
LC.AJ	0.57	LC.VR	0.08
LC.TB	0.41	UNC(NA)	0.08

64.96% of antecedent tokens in the study sample are stretches of discourse, ranging from subordinate or co-ordinate clauses (LC.CL, 24.98%) through complete sentences (LC.SC, 6.89%), to more or less definite discourse segments spanning sentence boundaries (GB.ET.PR, 32.85%). The remaining 35.04% is for single-word, intra-phrasal and genre-specific antecedents (e.g. figures, formulas and tables in academic writing). The most frequent cases are simple and complex noun phrases (9.98% and 8.19%), and prepositional complements (6.57%).

6.2.4.2 *Mode and genre distribution*

This section examines the extent of use of encapsulating relations and antecedents across modes and super-genres. Figures 6.52 and 6.53 are the most general and capture the distribution of the primary encapsulating relations, regardless of their intra- or intersentential realisations (as in Figure 6.49).

Figure 6.52 Mode distribution of the general encapsulating directions

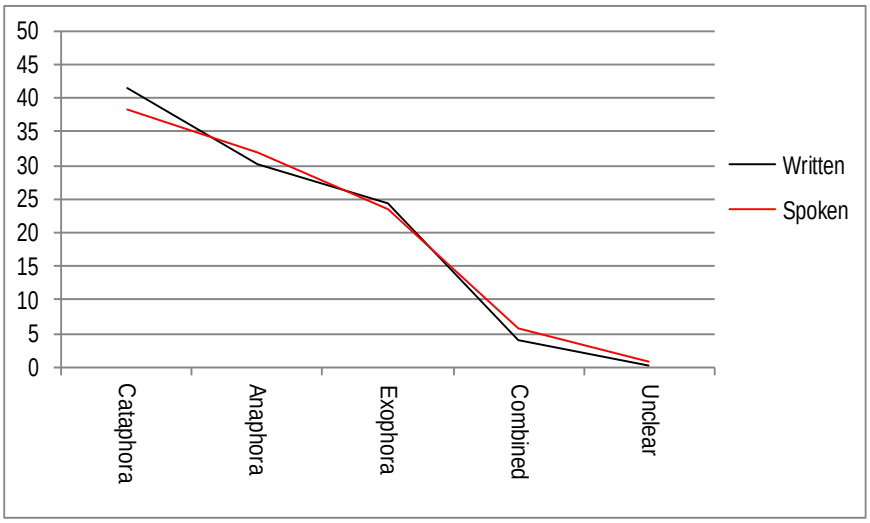


Figure 6.53 Genre distribution of the general encapsulating directions

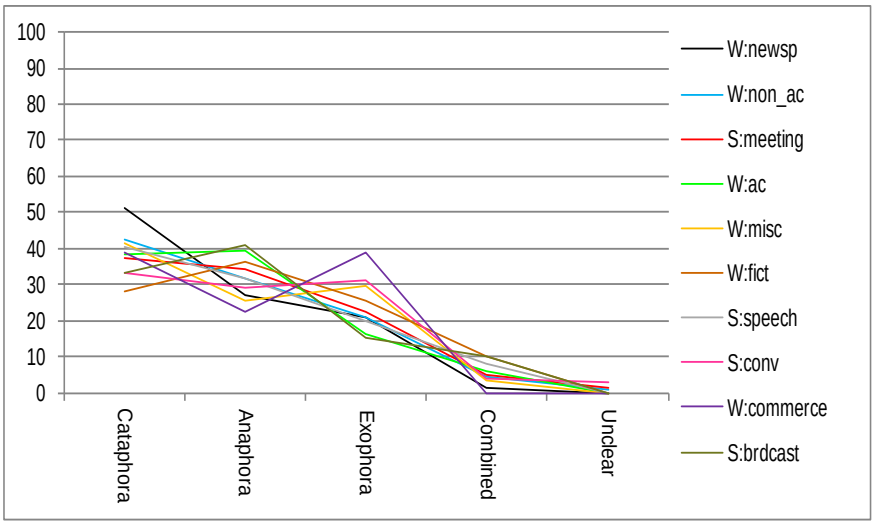


Figure 6.52 does not show any major differences between written and spoken English. Cataphora is slightly more frequent in written language (41.54% vs. 38.21%), whilst anaphora and combinations reveal a minor association with spoken language (31.97% vs. 30.09% for anaphora; 5.65% vs. 3.85% for combinations). A small difference is observed with exophora, which is more frequent in written than in spoken English (24.41% vs. 23.39%).

The connection between encapsulating relations and super-genres is more revealing. Cataphora in Figure 6.53 prevails in W:newsp (51.09%), followed by W:non_ac, W:misc and S:speech (42.68%, 41.60%, 40.50%). It is rare with S:conv and W:fict. W:fict is the third most frequent super-genre in anaphoric uses (36.44%). It is preceded by S:brdcast and W:ac (41.03% and 39.29%). These two super-genres rank lowest with exophora, where W:commerce, S:conv and W:misc prevail (38.78%, 31.00% and 29.60%; cf. the lower rank of S:conv in cataphora and anaphora). The pre-eminence of W:commerce in exophora contrasts with its weaker contribution to anaphoric and combined uses, where it is the least frequent alternative (this being closely so with W:newsp in the case of combined instances). Combinations are topped by S:brdcast and W:fict (10.26% and 10.17%), which also occur in the top three for anaphoric uses. Finally, unclear instances are most prevalent in S:conv, as is logical considering the difficulty of spelling out the contextual significance of shell nouns in often incomplete conversational exchanges.

Figures 6.54 and 6.55 present the ten highest-ranking specific encapsulating relations. As shown in Table 6.30, these cover 96.61% of the evidence in the sample.

Figure 6.54 Mode distribution of the top ten specific encapsulating directions

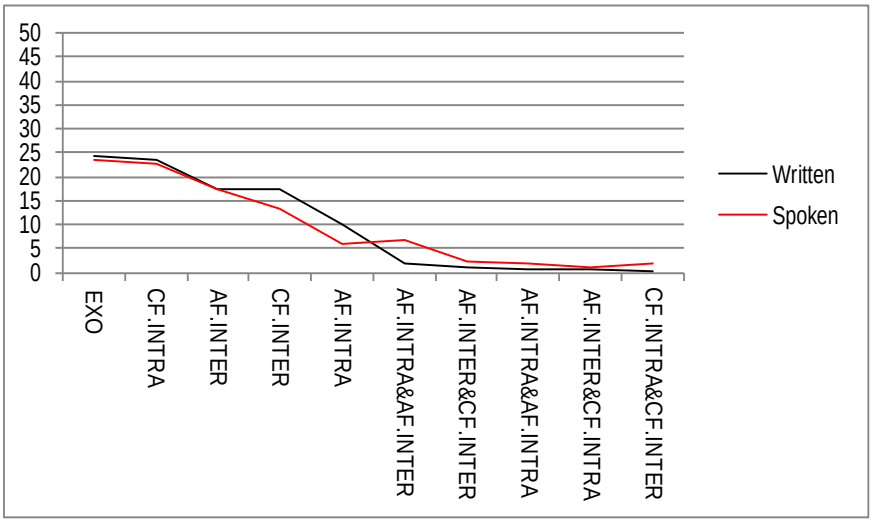


Figure 6.55 Genre distribution of the top ten specific encapsulating directions

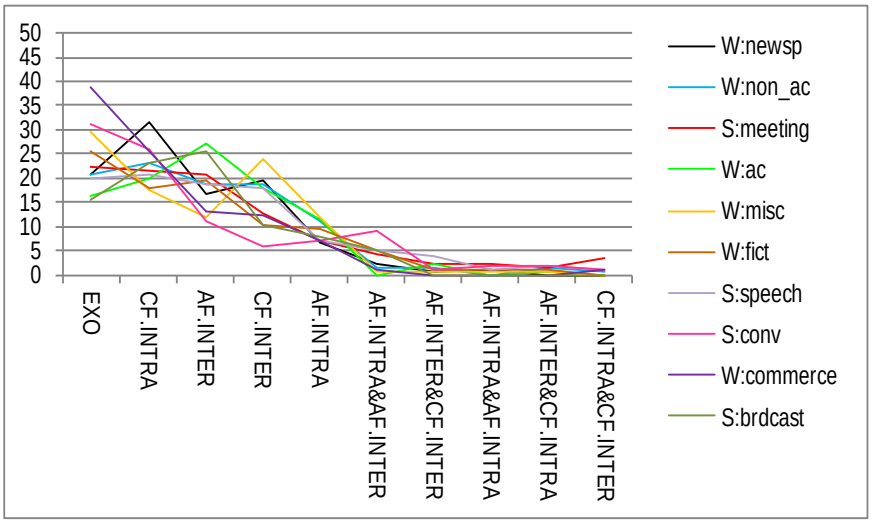


Figure 6.54 shows that the most noticeable associations are between CF.INTER, AF.INTRA and written English (17.56% vs. 13.45%; 9.96% vs. 6.04%), and AF.INTRA&AF.INTER and spoken English (6.63% vs. 1.82%). In all other cases, the distribution is similar in both modes, this being slightly higher with spoken English in all combinations other than AF.INTRA&AF.INTER, and with written English in EXO.

With regard to Figure 6.55, the dominance of W:newsp in cataphora is now found to relate to its prevalence in CF.INTRA (31.52%), where it is closely followed by S:conv (26.00%) and W:commerce (25.51%). W:newsp is also primed for CF.INTER, where it ranks second (19.57%), preceded by W:misc (24.00%), and followed by W:non_ac (18.90%), W:ac (17.86%) and S:speech (17.82%). W:ac is particularly prominent with CF.INTER and most frequent among AF.INTER, almost on a par with S:brdcast (27.14% and 25.64%). The intrasentential equivalent of anaphora, i.e. AF.INTRA, also reveals a strong priming for W:ac (11.43%), which is the second most frequent super-genre in this case (W:misc ranks first, 12.00%). However, AF.INTRA&AF.INTER does not occur in W:ac, as this encapsulating relation prevails only in spoken super-genres, the most common being S:conv (9.00%). S:conv is also frequent with EXO and CF.INTRA, where it ranks second (31.00% and 26.00%). Weaker associations appear for AF.INTER&CF.INTER and CF.INTRA&CF.INTER: the former is mainy restricted to S:speech, W:ac and S:meeting (3.96%, 2.14%, 2.08%), and the latter is primarily associated with S:meeting (3.47%).

Ancedent use is displayed in Figures 6.56 and 6.57. Data are given only for the top ten general antecedent categories in Table 6.34, accounting for 97.32% of the study sample.

Figure 6.56 Mode distribution of the top ten antecedent types. The count draws on the number of individual antecedent tokens in each mode and each super-genre

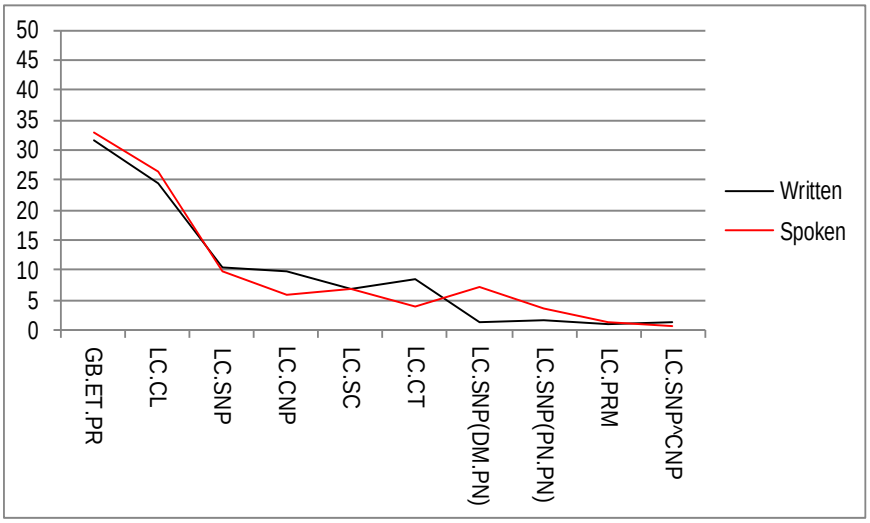
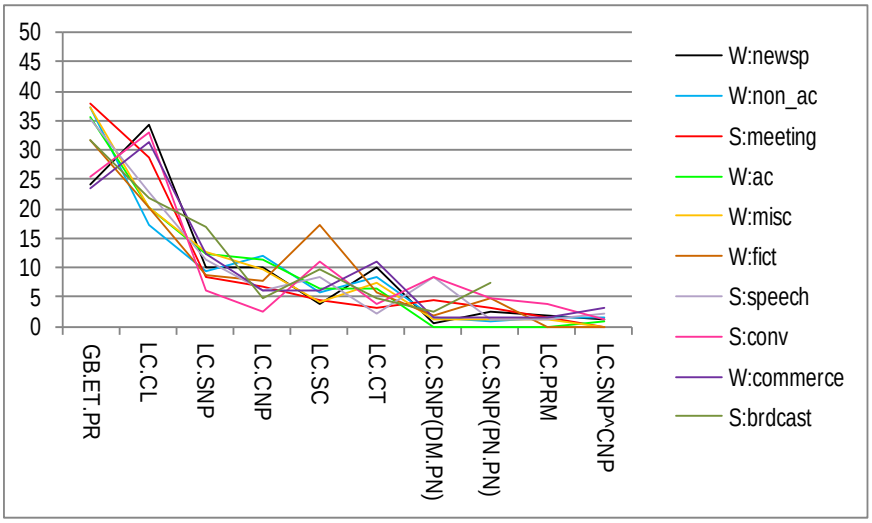


Figure 6.57 Genre distribution of the top ten antecedent types



In Figure 6.56, only four antecedent categories show noteworthy mode differences. Complex noun phrases and prepositional complements are more frequent in written English (9.64% vs. 5.83%; 8.32% vs. 3.89%), whilst pronouns prevail in spoken discourse, this being most marked with demonstrative pronouns (7.13% vs. 1.32%; cf. 3.46% vs. 1.59% for personal pronouns). The differences regarding extended discourse and clauses are minimal (26.57% spoken vs. 24.44% written; 32.83% spoken vs. 31.57% written), while the use of the other antecedents is almost equally distributed in both modes.

Concerning Figure 6.57, mention must first be made of clausal antecedents. The reported prevalence of W:newsp, S:conv and W:commerce in the use of both intrasentential cataphora (Figure 6.55) and appositive or noun-complement clauses (Figure 6.16) matches the results for clausal antecedents in Figure 6.57, where these three genres prevail, followed closely by S:meeting (34.18% for W:newsp, 32.93% for S:conv, 31.25% for W:commerce, and 28.79% for S:meeting). Interestingly, S:conv, W:newsp and W:commerce, whilst especially common with clauses, are the three least frequent super-genres in the use of extended antecedents. These are most frequent in S:meeting, W:misc, W:non_ac, W:ac and S:speech (37.88%, 37.23%, 37.14%, 35.77%, 35.42%).

Prepositional complement antecedents, like clausal ones, show a stronger association with W:commerce and W:newsp, followed by W:non_ac and W:misc (10.94%, 10.13%, 8.57% and 7.44%). W:non_ac features first with complex noun phrases, W:ac and W:newsp ranking closely behind (12.14%, 11.38% and 10.13%). By contrast, simple noun phrases are more frequent in S:brdcast, W:misc, W:commercce and W:ac (17.07%, 12.77%, 12.50% and 12.20%). S:conv, marginally frequent with simple and complex noun phrases, is frequent with demonstrative pronouns and premodifiers (8.54% and 3.66%). The use of demonstrative antecedents is mainly concentrated in S:conv and S:speech (8.54% and 8.33%), while personal pronouns are primed for S:brdcast, S:conv and W:fict (7.32%, 4.88% and 4.81%). The prominence of S:conv with demonstrative and personal pronouns is logical considering its close connection with AF.INTRA&AF.INTER, where AF.INTRA is often realised by a pronoun, as in ***that****'s the only thing* or ***it****'s only a warning.* W:fiction, ranks third with personal pronoun antecedents and stands out with sentential ones (17.31%; S:conv and S:brdcast rank second and third, 10.98%, 9.76%).

6.2.4.3 *Distribution across lemmas and semantic types*

This section examines the behaviour of encapsulating relations and antecedents across lemmas and semantic types of shell nouns. As in 6.2.2.3 and 6.2.3.3, the lists of lemmas in the main text comprise only the top ten units in each case.

Table 6.35 and Figure 6.58 show the four main types of encapsulation.

Table 6.35 Top ten lemmas for the general encapsulating directions. See Appendix 14.1 for the complete lists

Anaphora	%	Cataphora	%	Exophora	%	Combined	%
Endorsement	100.00	Quest	100.00	Correction	100.00	Opposite	30.00
Leave	75.00	Proviso	85.71	Prejudice	63.64	Misfortune	28.57
Testimony	71.43	Chance	77.50	Detail	50.00	Terror	25.00
Phenomenon	69.44	Capacity	73.91	Evidence	50.00	Thing	15.00
Impetus	66.67	Sense	72.73	Recommendation	50.00	Dimension	12.50
Contradiction	62.50	Failure	69.57	Crime	45.00	Motivation	12.50
Anger	60.00	Surprise	64.00	Venture	41.38	Assessment	11.54
Terror	50.00	Suspicion	64.00	System	41.18	Point	10.00
Triumph	50.00	Motivation	62.50	Work	40.74	Part	9.38
Finding	46.43	Recollection	60.00	Answer	40.00	Area	7.50

Figure 6.58 General semantic types and the general encapsulating directions

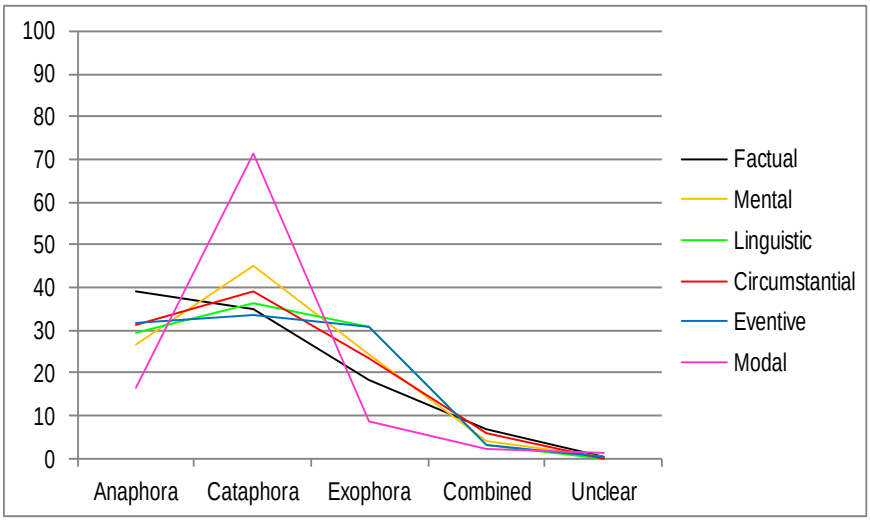


Figure 6.58 suggests the marked influence of cataphora on modal nouns (71.43%; e.g. *chance*, *capacity*, *failure*), followed by mental units (44.94%; e.g. *sense*, *surprise*, *suspicion*, *motivation*, *recollection*). Anaphora is rare with modal and mental meanings, and ranks highest with factual nouns

(39.13%, e.g. *testimony*, *phenomenon*, *impetus*, *finding*, *part*). These are closely followed by eventive and circumstantial nouns (31.91%, e.g. *triumph*, *venture*, *work*; 31.18%, e.g. *area*, *practice*, *time*). Linguistic and eventive nouns occur most frequently with exophora (31.00%, e.g. *correction*, *detail*, *recommendation*, *answer*; 30.85%, e.g. *crime*, *venture*, *work*, *scandal*), where factual and modal nouns rank lowest. Lastly, combinations suggest an association with factual and circumstantial units (e.g. 6.96%, e.g. *opposite*, *thing*, *dimension*, *point*, *part*; 5.91%, e.g. *area*, *way*, *system*), but a negative priming with modal nouns.

Tables 6.36a through 6.36b and Figure 6.59 elaborate the above by narrowing their scope to the ten highest-ranking specific encapsulating relations, as in Figures 6.54 and 6.55. Exophora, whose results are presented in Table 6.35 and Figure 6.58, is excluded from the tables and figures below.

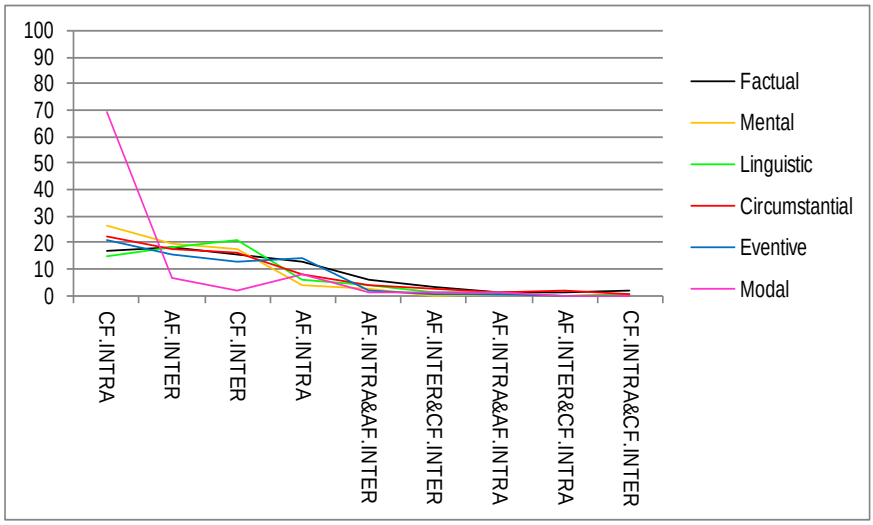
Table 6.36a Top ten lemmas for the top ten specific encapsulating directions. See Appendix 14.2 for the complete lists

CF.INTRA	%	AF.INTER	%	CF.INTER	%	AF.INTRA	%
Quest	80.00	Contradiction	50.00	Triumph	37.50	Leave	75.00
Chance	77.50	Endorsement	50.00	Assessment	34.62	Impetus	66.67
Capacity	69.57	Phenomenon	44.44	Myth	33.33	Testimony	42.86
Failure	69.57	Foreboding	42.86	Joke	30.00	Opposite	30.00
Recollection	60.00	Anger	40.00	Opposite	30.00	Facet	28.57
Proviso	57.14	Terror	37.50	Way	30.00	Failure	26.09
Suspicion	56.00	Project	32.50	Facet	28.57	Endorsement	25.00
Surprise	52.00	Finding	32.14	Foreboding	28.57	Triumph	25.00
Endeavour	44.44	Myth	30.00	Irony	28.57	Phenomenon	22.22
Time	42.50	Application	27.50	Proviso	28.57	Anger	20.00

Table 6.36b Top ten lemmas for the top ten specific encapsulating directions

AF.INTRA& AF.INTER	%	AF.INTER& CF.INTER	%	AF.INTRA& AF.INTRA	%	AF.INTER& CF.INTRA	%	CF.INTRA& CF.INTER	%
Recollection	20.00	Terror	12.50	Endorsement	25.00	Motivation	12.50	Problem	7.50
Testimony	14.29	Opposite	10.00	Failure	4.35	Area	5.00	Point	5.00
Contradiction	12.50	Dimension	8.33	Surprise	4.00	Thing	5.00	Vision	3.45
Part	12.50	Part	6.25	Part	3.13	Assessment	3.85	Myth	3.33
Word	12.50	System	5.88	Area	2.50	Vision	3.45	Part	3.13
Joke	10.00	Point	5.00	Example	2.50	Phenomenon	2.78	System	2.94
Problem	10.00	Practice	5.00	Joke	2.50	Objective	2.50	Example	2.50
Prejudice	9.09	Assessment	3.85	Point	2.50	Point	2.50	Objective	2.50
Surprise	8.00	Finding	3.57	Practice	2.50	Way	2.50	Anger	0.00
Scandal	7.69	Myth	3.33	Problem	2.50	Anger	0.00	Answer	0.00

Figure 6.59 General semantic types and the top ten specific encapsulating directions



The reported dominance of modal nouns with cataphoric uses is, as shown in Figue 6.59, accounted for by their prevalence with intrasentential cataphora. 69.23% of modal instances is made up of CF.INTRA uses, and this is the most remarkable connection (e.g. *chance*, *capacity*, *failure*). Mental and circumstantial nouns, with frequencies of 26.59% and 22.58%, rank second and third (e.g. *recollection*, *suspicion*, *surprise*, *sense*; *proviso*, *time*). Linguistic nouns show the weakest priming with CF.INTRA but are most frequent with CF.INTER, where they are closely followed by mental and circumstantial nouns (21.00%, e.g. *myth*, *joke*, *irony*; 17.26%, e.g. *assessment*, *foreboding*, *sense*; and 16.13%; e.g. *way*, *proviso*, *system*). AF.INTER is slightly more frequent with mental, factual and linguistic nouns (19.35%, e.g. *foreboding*, *anger*, *terror*, *project*, *philosophy*; 18.26%, e.g. *phenomenon*, *finding*, *point*, *characteristic*; and 18.00%, e.g. *contradiction*, *endorsement*, *myth*, *application*, *warning*). Mental and linguistic nouns, whilst topping AF.INTER, are scarcely used with AF.INTRA, where eventive and factual nouns prevail (14.36%, e.g. *triumph*, *practice*, *misfortune*; 13.04%, e.g. *impetus*, *testimony*, *opposite*, *facet*, *phenomenon*, *example*). Factual nouns also prevail with AF.INTRA&AF.INTER, where they rank first, followed by circumstantial and linguistic nouns (6.09%, e.g. *testimony*, *contradiction*, *part*, *problem*, *thing*, *finding*; 3.76%, e.g. *way*, *practice*, *time*; and 3.67%, e.g. *word*, *joke*, *warning*).

The remaining four encapsulating relations show a similar association with factual units, with the caveat that differences between semantic types are negligible. In the case of AF.INTER&CF.INTER, use is almost equally

distributed over factual and circumstantial nouns (2.90%, e.g. *opposite*, *dimension*, *part*, *point*, *finding*; 2.69%, e.g. *system*, *practice*, *way*). An even distribution is also apparent with AF.INTRA&AF.INTRA, where use is by factual, modal, circumstantial and linguistic nouns (e.g. 1.16%, e.g. *part*, *example*, *point*, *problem*; 1.10%, e.g. *failure*; 1.08%, e.g. *area*, *practice*; 1.00%, e.g. *endorsement*, *joke*, *word*). AF.INTER&CF.INTRA occurs only with circumstantial, mental and factual nouns (1.61%, e.g. *area*, *way*; 0.89%, e.g. *motivation*, *assessment*, *vision*, *objective*; 0.87%, e.g. *thing*, *phenomenon*, *point*). Lastly, CF.INTRA&CF.INTER is mainly primed for factual nouns, followed by mental, circumstantial and linguistic nouns (2.03%, e.g. *problem*, *point*, *part*, *example*; 0.60%, e.g. *vision*, *objective*; 0.54%, e.g. *system*; 0.33%, e.g. *myth*).

Tables 6.37a through 6.37b and Figure 6.60 show the ten most frequent general antecedents, as in Figures 6.56 and 6.57.

Table 6.37a Top ten lemmas for the top ten antecedents. Percentages are based on the total number of antecedent tokens for each lemma. See Appendix 14.3 for the complete lists

GB.ET.PR	%	LC.CL	%	LC.SNP	%	LC.CNP	%	LC.SC	%
System	70.83	Leave	100.00	Opposite	50.00	Testimony	50.00	Foreboding	60.00
Detail	70.00	Proviso	85.71	Facet	40.00	Work	43.75	Prejudice	40.00
Terror	70.00	Capacity	66.67	Area	38.46	Anger	40.00	Contradiction	33.33
Project	69.23	Endeavour	62.50	Word	28.13	Endorsement	40.00	Sense	33.33
Myth	67.74	Failure	62.50	Failure	25.00	Impetus	33.33	Dimension	26.32
Experience	65.38	Recollection	60.00	Philosophy	22.22	Phenomenon	31.43	Warning	17.24
Irony	60.00	Objective	58.33	Phenomenon	20.00	Motivation	30.00	Finding	16.67
Point	52.38	Recommendation	56.25	Prejudice	20.00	Scandal	22.22	Opposite	14.29
Evidence	50.00	Suspicion	52.17	Challenge	18.75	Characteristic	20.00	Triumph	14.29
Practice	48.48	Surprise	51.85	Example	18.18	Facet	20.00	Joke	13.51

Table 6.37b Top ten lemmas for the top ten antecedents

LC.CT	%	LC.SNP(DM.PN)	%	LC.SNP(PN.PN)	%	LC.PRM	%	LC.SNP^CNP	%
Quest	80.00	Recollection	20.00	Endorsement	20.00	Practice	12.12	Vision	16.00
Chance	50.00	Word	18.75	Contradiction	16.67	Application	11.11	Example	6.82
Application	27.78	Misfortune	14.29	Testimony	12.50	Dimension	5.26	Answer	4.00
Sense	27.78	Thing	11.36	Scandal	11.11	Capacity	4.76	Problem	2.94
Vision	24.00	Joke	10.81	Part	8.82	Answer	4.00	Application	2.78
Recollection	20.00	Venture	10.00	Problem	8.82	Project	3.85	Way	2.63
Time	16.13	Finding	8.33	Assessment	8.33	Warning	3.45	Area	2.56
Capacity	14.29	Area	7.69	Surprise	7.41	Time	3.23	Anger	0.00
Crime	13.04	Surprise	7.41	Warning	6.90	Anger	0.00	Assessment	0.00
Suspicion	13.04	Point	7.14	Practice	6.06	Area	0.00	Capacity	0.00

Figure 6.60 General semantic types and the top ten antecedents

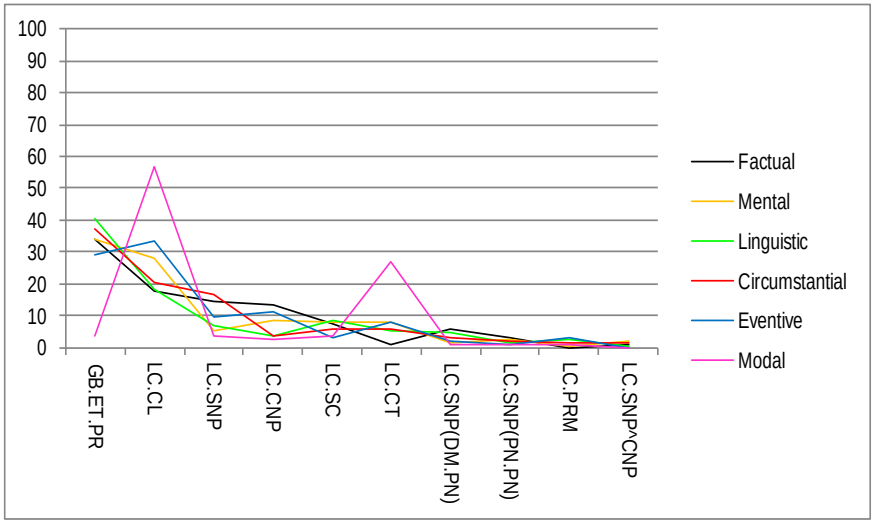


Figure 6.60 shows that the preponderance of intrasentential cataphora among modal nouns corresponds to their prevalence with clausal and prepositional complement antecedents (56.47%, e.g. *capacity*, *failure*, *chance*; 27.06%, *chance*, *capacity*). Priming in the latter case is especially marked with *chance*, as in *no chance of a football stadium* or *a chance of having one* (50% of LC.CT use). Eventive and mental nouns rank second and third with both LC.CL and LC.CT (33.33% vs. 28.11% for LC.CL, e.g. *endeavour*, *venture*, *crime*, *recollection*, *objective*, *suspicion*, *surprise*; 8.15% vs. 7.83% for LC.CT, e.g. *quest*, *crime*, *challenge*, *sense*, *vision*, *recollection*, *suspicion*). The evidence regarding clausal antecedents is consistent with the high frequency shown by appositive *to*-infinitive clauses with modal and eventive nouns, and by appositive *that*-clauses with mental nouns (see Figure 6.21).

Eventive nouns are also frequent with simple and complex noun phrases and premodifiers. Premodifiers rank highest with eventive nouns, followed closely by linguistic and circumstantial nouns (2.96%, e.g. *practice*, as in *a **teaching** practice*; 2.61%, e.g. *application*, *answer*, *warning*, as in *a **membership** application*; 1.24%, e.g. *time*, as in *__dinner__ time*). Concerning noun-phrase antecedents, LC.SNP is most frequent with circumstantial, factual and eventive nouns (16.77%, e.g. *area*, *system*, *way*; 14.37%, e.g. *opposite*, *facet*, *phenomenon*, *example*; 9.63%, e.g. *crime*, *work*, *venture*). By contrast, LC.CNP prevails with factual, eventive and mental nouns (13.49%, e.g. *testimony*, *impetus*, *phenomenon*, *characteristic*, *facet*; 11.11%, e.g. *work*, *scandal*, *triumph*; 8.54%, e.g. *anger*, *motivation*,

assessment; cf. also LC.SNP^CNP, slightly more frequent with mental, circumstantial and factual nouns, 1.78%, 1.24% and 1.17%).

It should be noted that the strong connection between noun-phrase antecedents and factual and eventive nouns ties in with their frequent use with intrasentential anaphora (see Figure 6.59; e.g. *unemployment among refugees was still only a **temporary phenomenon***, *the tunnel project can succeed entirely as a **private venture***). Factual nouns prevail among uses of AF.INTRA&AF.INTER and rank first with both pronominal antecedents. In the case of demonstrative pronouns, factual, linguistic and circumstantial nouns prevail (5.57%, e.g. *thing*, *finding*, *point*, *example*, *part*; 4.78%, e.g. *word*, *joke*, *answer*; 3.11%, e.g. *area*, *time*, *way*), while with personal pronouns, the top three comprises factual, mental and circumstantial nouns (3.23%, e.g. *contradiction*, *testimony*, *part*, *problem*; 2.49%, e.g. *assessment*, *surprise*, *philosophy*, *vision*; 1.85%, e.g. *practice*, *way*).

Lastly, in relation to extended and sentential antecedents, linguistic nouns are primed for both (40.43%, e.g. *detail*, *myth*, *irony*, *point*; 8.70%, e.g. *warning*, *joke*, *word*, *answer*). In the former case, linguistic nouns are followed by circumstantial, mental and factual (37.27%, e.g. *system*, *practice*, *way*; 33.81%, e.g. *terror*, *project*, *experience*, *assessment*, *philosophy*; 33.72%, e.g. *point*, *evidence*, *characteristic*, *part*), while in the latter, mental, factual and circumstantial rank closely behind (8.19%, e.g. *foreboding*, *prejudice*, *sense*, *surprise*; 7.62%, e.g. *dimension*, *finding*, *opposite*, *thing*; 5.59%, e.g. *system*, *way*, *time*).

6.2.4.4 *Direction of encapsulation and other variables*

This section is about the relation between encapsulating direction and two other variables: formal structure and Theme-Rheme. The aims are to glean insights into the formal realisation of the encapsulating relations evidenced by shell nouns, and to determine their influence on the clause-initial or clause-final position of these units.

Tables 6.38a and 6.38b display the evidence concerning formal structure. The tags in both tables combine the coding for formal structure and direction of encapsulation. Table 6.38a is the most specific, as it details the intra- or intersentential realisation of the four main encapsulating relations (i.e., anaphora, cataphora, exophora and combinations). Table 6.38b looks at these four encapsulating relations in general. As in previous sections, the size of the lists forces the inclusion of only the top 50% patterns in the main text. Table 6.38a comprises 115 patterns accounting for 54.32% of the study sample (688 patterns), and Table 6.38b contains 64 patterns representing 51.83% of the evidence (579 patterns).

Table 6.38a Formal patterns and specific encapsulating directions (Top 50%). See Appendix 15.1 for the complete list

Formal structure & Direction of encapsulation (Detailed: Top 50 %)			
DF.AR^H-AF.INTER	2.35	H^PP(of)-AF.INTRA&AF.INTER	0.35
H-EXO	2.28	IN.AR^H^PP(of)-AF.INTRA	0.35
DF.AR^H^PP(of)-CF.INTRA	2.00	DF.AR^H^RV.CL-CF.INTRA	0.35
AJ^H-EXO	1.66	H-CF.INTRA	0.35
DM.DT(THIS)^H-AF.INTER	1.31	IN.AR^AJ^H-CF.INTRA	0.35
DF.AR^H-CF.INTRA	1.24	PS.DT^H^AP.TI.CL-CF.INTRA	0.35
DF.AR^H-EXO	1.17	N^H-AF.INTER	0.28
DF.AR^H^PP(of)-CF.INTER	0.97	DF.DV.GV.NP^AJ^H-CF.INTER	0.28
H-CF.INTER	0.97	IN.AR^AJ^N^H-EXO	0.28
AJ^H-CF.INTER	0.90	AJ^H^PP(in)-EXO	0.28
DF.AR^H-CF.INTER	0.90	H^PP(for)-CF.INTRA	0.28
H-AF.INTER	0.90	H^PP(of)-AF.INTRA	0.28
IN.AR^AJ^H-AF.INTRA	0.83	IN.AR^AJ^H^PP(to)-AF.INTRA	0.28
DM.DT(THESE)^H-AF.INTER	0.83	PS.DT^H^PP(for)-CF.INTRA	0.28
PS.DT^H-EXO	0.83	DF.AR^H^PP(to)-EXO	0.28
DF.AR^H^PP(of)-EXO	0.76	H^PP(of)-AF.INTER	0.28
IN.AR^H-AF.INTRA&AF.INTER	0.76	QT^H^PP(of)-CF.INTER	0.28
DF.AR^AJ^H-AF.INTER	0.69	DF.AR^H^RV.CL-CF.INTER	0.28
H^PP(of)-EXO	0.69	PS.DT^H-CF.INTRA	0.28
IN.AR^H-EXO	0.69	AJ^AJ^H-EXO	0.28
DM.DT(THAT)^H-AF.INTER	0.69	DF.AR^H-AF.INTRA	0.28
NUM.CD^H-CF.INTER	0.69	DF.AR^H^AP.THAT.CL-CF.INTRA	0.28
DF.AR^H^PP(of)-AF.INTER	0.62	DM.DT(THIS)^AJ^H-AF.INTER	0.28
DF.AR^AJ^H-CF.INTRA	0.55	H^RV.CL-EXO	0.28
H^PP(of)-CF.INTER	0.55	IN.AR^H-AF.INTER	0.28
H^PP(of)-CF.INTRA	0.55	IN.AR^H-AF.INTRA&AF.INTRA	0.28
IN.AR^H-AF.INTRA	0.55	NAS.DT^H-EXO	0.28
PS.DT^H-AF.INTER	0.55	NG.DT^H-CF.INTRA	0.28
QT^H-EXO	0.48	PDT^H-AF.INTER	0.28
IN.AR^AJ^H-CF.INTER	0.48	PS.DT^AJ^H-CF.INTER	0.28
DF.AR^AJ^H^PP(of)-EXO	0.48	PS.DT^N^H-EXO	0.21
IN.AR^H^PP(of)-CF.INTRA	0.48	IN.AR^AJ^H^PP(of)-CF.INTRA	0.21
DF.AR^AJ^H^PP(of)-CF.INTER	0.48	IN.AR^AJ^H-AF.INTER	0.21
H^AP.TI.CL-CF.INTRA	0.48	H^PP(to)-EXO	0.21
AJ^H-AF.INTER	0.48	IN.AR^AJ^H^PP(of)-AF.INTRA	0.21
AJ^H-CF.INTRA	0.48	H^PP(for)-EXO	0.21
DF.AR^AJ^H-EXO	0.48	IN.AR^AJ^H^PP(for)-AF.INTRA	0.21
DF.AR^H^RV.CL-EXO	0.48	NUM.CD^H^PP(of)-CF.INTRA	0.21
H-AF.INTRA	0.48	IN.AR^H^PP(of)-EXO	0.21
IN.AR^H^AP.TI.CL-CF.INTRA	0.48	IN.AR^AJ^H^PP(of)-CF.INTER	0.21
QT^H-CF.INTER	0.48	PS.DT^AJ^H^AP.TI.CL-CF.INTRA	0.21
N^H-EXO	0.41	AJ^H-AF.INTRA	0.21
DF.AR^AJ^H-CF.INTER	0.41	AS.DT^H-EXO	0.21
DF.AR^AJ^H^PP(of)-CF.INTRA	0.41	DF.AR^AJ^H-AF.INTRA	0.21
DF.AR^N^H-AF.INTER	0.41	DF.AR^NUM.OR^H^RV.CL-CF.INTRA	0.21
DF.AR^H^AP.TI.CL-CF.INTRA	0.41	H^RV.CL-CF.INTER	0.21

IN.AR^AJ^H-EXO	0.41	H^TI.CL-CF.INTER	0.21
AS.DT^H-CF.INTER	0.41	H^TI.CL-EXO	0.21
N^H-AF.INTRA	0.35	IN.AR^AJ^H-AF.INTRA&AF.INTER	0.21
N^H-CF.INTER	0.35	IN.AR^AJ^H-AF.INTRA&AF.INTRA	0.21
DF.DV.GV.NP^H^PP(of)-CF.INTRA	0.35	IN.AR^AJ^H^AP.TI.CL-CF.INTRA	0.21
DF.DV.GV.NP^H-EXO	0.35	IN.AR^H-CF.INTRA	0.21
DF.DV.GV.NP^H^AP.TI.CL-CF.INTRA	0.35	IN.AR^H^AP.THAT.CL-CF.INTRA	0.21
DF.DV.GV.NP^H-AF.INTER	0.35	IN.AR^H^RV.CL-AF.INTRA	0.21
IN.AR^AJ^H^AP.THAT.CL-CF.INTRA	0.35	IN.AR^N^H-EXO	0.21
IN.AR^N^H-AF.INTRA	0.35	PS.DT^H-CF.INTER	0.21
DF.AR^H^PP(of)^AP.TI.CL-CF.INTRA	0.35	PS.DT^AJ^H-CF.INTRA	0.21
		UV.DT^H-EXO	0.21

Table 6.38b Formal patterns and general encapsulating directions (Top 50%). See Appendix 15.2 for the complete list

Formal structure & Direction of encapsulation (AF, CF, EXO: 50%)			
DF.AR^H^PP(of)-CF	3.04	IN.AR^N^H-AF	0.55
DF.AR^H-AF	2.76	DF.AR^N^H-AF	0.55
H-EXO	2.28	QT^H-CF	0.55
DF.AR^H-CF	2.21	DF.DV.GV.NP^H^PP(of)-CF	0.48
IN.AR^H-AF	2.00	DF.AR^AJ^H^PP(of)-EXO	0.48
AJ^H-EXO	1.66	N^H-CF	0.48
IN.AR^AJ^H-AF	1.45	PS.DT^AJ^H-CF	0.48
DM.DT(THIS)^H-AF	1.45	PS.DT^H-CF	0.48
H-AF	1.38	QT^H-EXO	0.48
AJ^H-CF	1.38	DF.AR^AJ^H-EXO	0.48
H-CF	1.38	DF.AR^H^RV.CL-EXO	0.48
H^PP(of)-CF	1.24	AS.DT^H-CF	0.48
DF.AR^H-EXO	1.17	H^AP.TI.CL-CF	0.48
DF.AR^AJ^H-AF	1.11	IN.AR^H^AP.TI.CL-CF	0.48
DF.AR^AJ^H-CF	1.04	IN.AR^AJ^H^PP(of)-CF	0.41
IN.AR^AJ^H-CF	0.97	H^PP(of n)-AF	0.41
DF.AR^H^PP(of)-AF	0.90	DF.AR^H^RV.CL-AF	0.41
DF.AR^AJ^H^PP(of)-CF	0.90	IN.AR^AJ^H-EXO	0.41
DM.DT(THESE)^H-AF	0.90	DF.AR^H^AP.TI.CL-CF	0.41
PS.DT^H-EXO	0.83	H^PP(for)-CF	0.35
DF.AR^H^PP(of)-EXO	0.76	NUM.CD^H^PP(of)-CF	0.35
AJ^H-AF	0.76	IN.AR^AJ^H^PP(of)-AF	0.35
PS.DT^H-AF	0.76	QT^H^PP(of)-CF	0.35
NUM.CD^H-CF	0.76	PS.DT^H^PP(for)-CF	0.35
DM.DT(THAT)^H-AF	0.76	IN.AR^AJ^H^PP(to)-AF	0.35
H^PP(of n)-EXO	0.69	DF.AR^H^PP(of)^AP.TI.CL-CF	0.35
IN.AR^H-EXO	0.69	DM.DT(THIS)^AJ^H-AF	0.35
IN.AR^H^PP(of)-CF	0.62	DF.DV.GV.NP^H-EXO	0.35
H^PP(of)-AF	0.62	IN.AR^AJ^H^AP.THAT.CL-CF	0.35
N^H-AF	0.62	DF.DV.GV.NP^H^AP.TI.CL-CF	0.35

DF.AR^H^RV.CL-CF	0.62	DF.DV.GV.NP^H-AF	0.35
IN.AR^H^PP(of)-AF	0.55	PS.DT^H^AP.TI.CL-CF	0.35

The top ten patterns in each table show a number of associations. For example, DF.AR^H is most frequently used anaphorically (2.76% for AF and 2.35% for AF.INTER) and only slightly less so, cataphorically (2.21% for CF and 1.24% for CF.INTRA). By contrast, DF.AR^H^PP(*of*) is markedly frequent with cataphora (3.04% for CF, 2.00% for CF.INTRA) and only slightly so with anaphora (0.90% for AF, 0.62% for AF.INTER). H and AJ^H are frequent with exophora (2.28% for the former and 1.66% for the latter). In the case of AJ^H, exophora is closely followed by cataphora, (1.38% for CF, 0.90% for CF.INTER), whilst with H, anaphora and cataphora rank second, both with the same frequency (1.38% for AF and CF, 0.97% for CF.INTER, 0.90% for AF.INTER). Anaphora is most frequent with DM.DT(THIS)^H, where it is the only encapsulating relation (1.45% for AF, 1.31% for AF.INTER), and with IN.AR^H (2.00% for AF, 0.76% for AF.INTRA&AF.INTER, 0.55% for AF.INTRA), where exophora ranks second (0.69%).

Figures 6.61 and 6.62 summarises the above to the ten highest-ranking determiners in the study sample (as in 6.2.3.4). Figure 6.61, as in 6.2.4.2 and 6.2.4.3, presents the ten most frequent specific encapsulating relations, while Figure 6.62 focuses on the main relations, irrespective of their realisation. The data in each case represent the proportion of encapsulating relations for each determiner.

Figure 6.61 Determiner use and specific encapsulating directions. Percentages are based on the total number of shell instances with each determiner

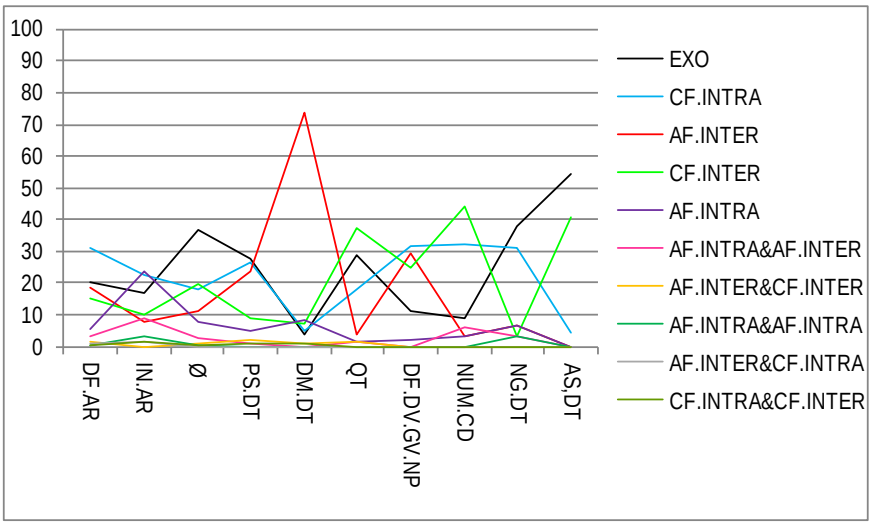
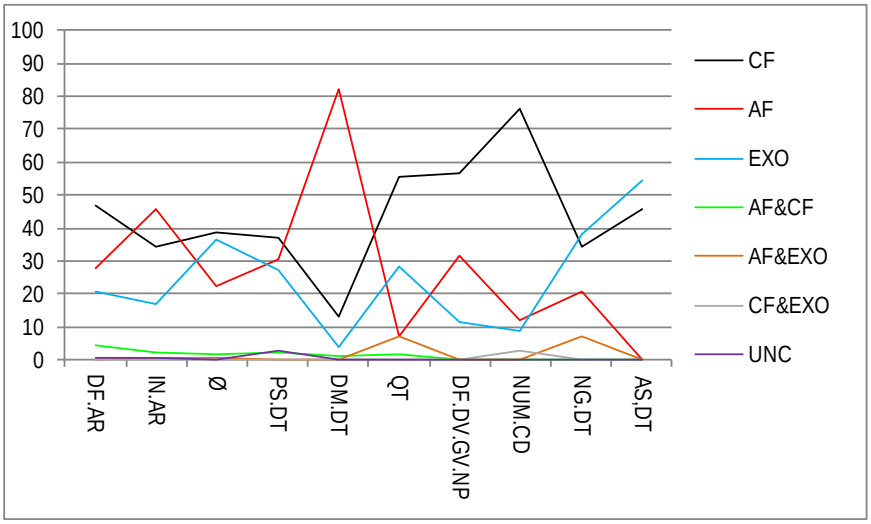


Figure 6.62 Determiner use and the general encapsulating directions



Four connections stand out in Figure 6.62: the marked frequency of anaphora with demonstrative determiners (81.93%) and of cataphora with cardinal numerals (76.47%), with genitive noun phrases (56.82%) and with quantifiers (55.36%). This is primarily associated with the AF.INTER use of demonstrative determiners (73.49%), the prevalence of CF.INTER and CF.INTRA in the case of quantifiers and cardinal numerals (44.12% vs. 32.35%; 37.50% vs. 17.86%), and of CF.INTRA and CF.INTER with genitive noun phrases[40] (31.82%, 25.00%).

With respect to the top four determiners, the definite article is strongly primed for cataphora (particularly, the intrasentential subtype), followed by anaphora and exophora (46.80% for CF and 31.20% for CF.INTRA, 27.88% for AF and 20.46% for EXO). Anaphora, especially AF.INTRA, occurs as the prevailing relation with the indefinite article, cataphora and exophora ranking second and third (45.73% for AF and 23.50% for AF.INTRA, 34.19% for CF and 16.67% for EXO). The high frequency of indefinite anaphoric shell nouns seems to tie in with the close link between indefinite articles and subject complements in Figure 6.42. Therefore, many of the indefinite shell-noun subject complements often find their lexical realisation in anaphoric subject noun phrases, which explains the greater frequency of AF.INTRA (e.g. *this is a problem*, *this is a surprise*). Zero-article shell nouns are almost equally common with cataphora and exophora, and are relatively infrequent with anaphora (38.75%, 36.56%, 22.19%). Of the top four determiners, exophora ranks highest with the zero article, thereby

[40] Cf. also the prominence of AF.INTER, ranking second in Figure 6.61: 29.55%.

explaining its prevalence with H and AJ^H in Tables 6.38a and 6.38b. Use of possessive determiners is frequent in cataphora, anaphora and exophora, their frequencies differing only slightly (37.25%, 30.39%, 27.45%). Lastly, negative and assertive determiners are primed for exophora and cataphora, most clearly so with assertive instances, where exophora ranks highest (37.93% vs. 34.48% for NG.DT and 54.55% vs. 45.45% for AS.DT).

Figures 6.63 through 6.66 explore the association between postmodifying structures and encapsulating relations. Figures 6.63 and 6.64 are the most general, as they distinguish only between shell-noun instances with and without postmodification. In this respect, only anaphora and cataphora differ substantially between modified and unmodified units. Anaphora, specifically its intersentential realisation, prevails with unmodified cases (40.26% vs. 20.79% for AF; 25.50% vs. 8.54% for AF.INTER), whilst cataphora, especially its intrasentential realisation, prevails among postmodified nouns (52.13% vs. 29.26% for CF; 35.75% vs. 11.81% for CF.INTRA). The differences regarding exophora and combinations are minimal.

Figure 6.63 Specific encapsulating directions and the occurrence of postmodification

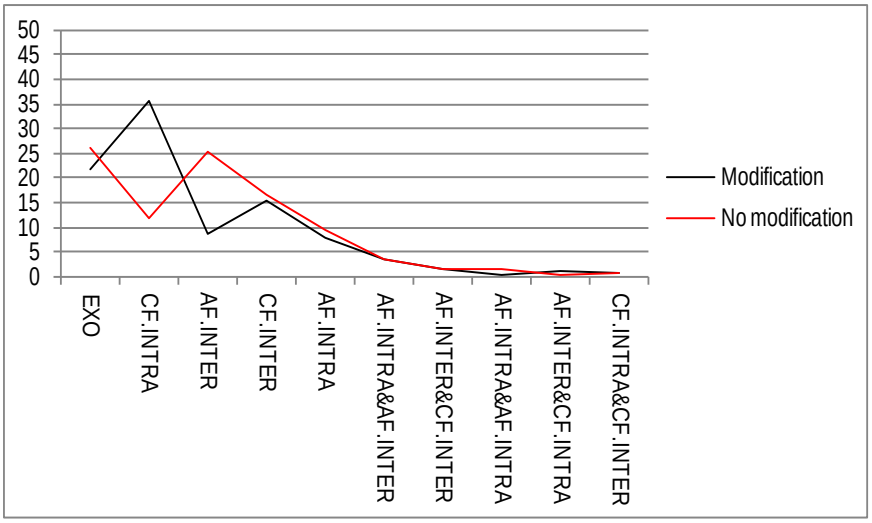
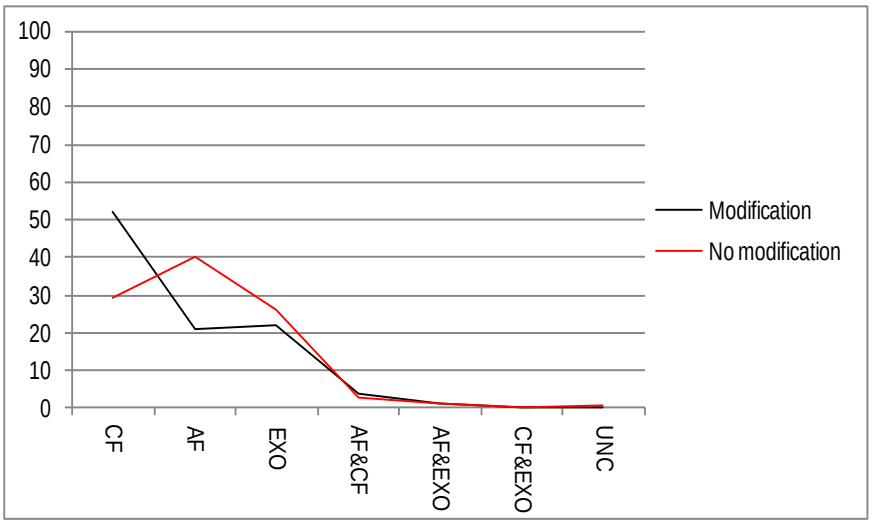


Figure 6.64 General encapsulating directions and the occurrence of postmodification



Figures 6.65 and 6.66 turn to the ten most common postmodifying structures in the study sample (as in 6.2.3.4).

Figure 6.65 Postmodifying structures and specific encapsulating directions. Percentages are calculated on the basis of the total number of shell-noun instances for each postmodifying structure

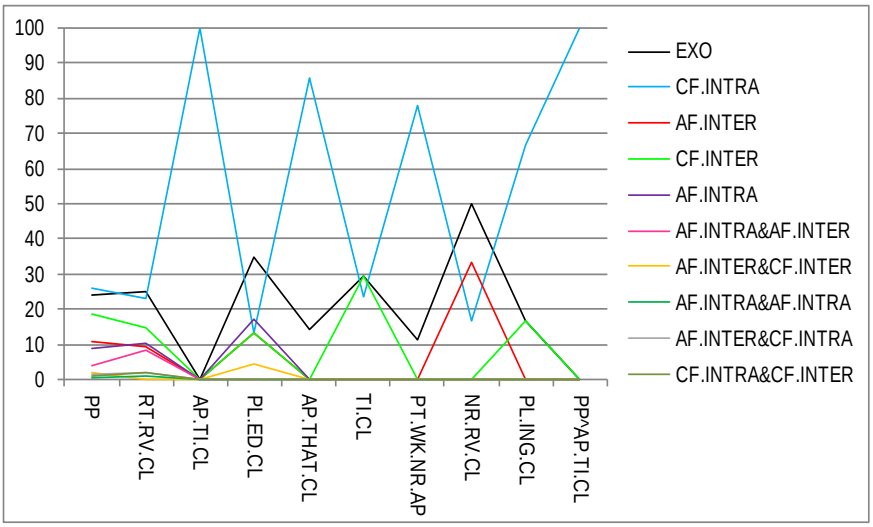
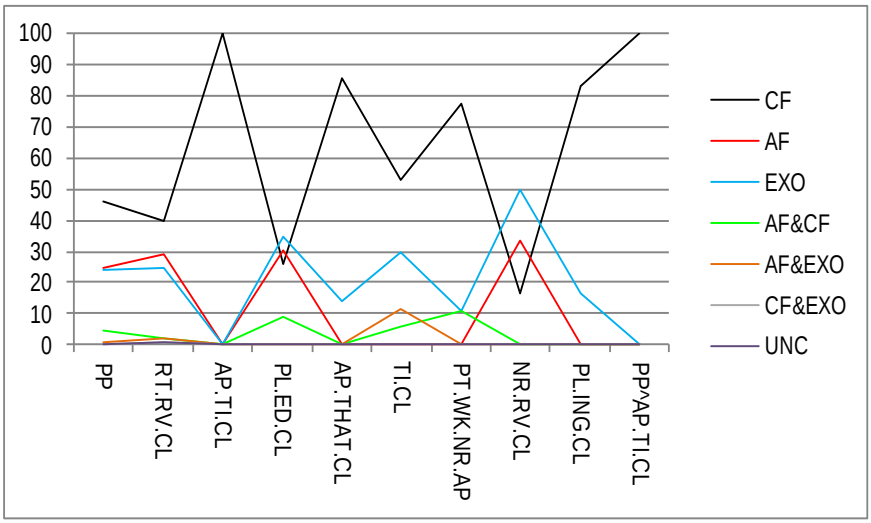


Figure 6.66 Postmodifying structures and general encapsulating directions



According to Figure 6.66, six of the top ten structures are fully or almost fully primed for cataphoric uses: AP.TI.CL, AP.THAT.CL, TI.CL, PT.WK.NR.AP, PL.ING.CL and PP^AP.TI.CL. This is logical with appositive and noun complement *to*-infinitive and *that*-clauses, where the adjoining clause acts as the intrasentential realisation of the head noun (hence the dominance of CF.INTRA in Figure 6.65). Partial weak non-restrictive appositives, whilst not imposed by the head noun, also prove useful for the intrasentential realisation of shell nouns through clauses, as in example (758). As to *to*-infinitive and *-ing* participle clauses, their prevalence with intrasentential cataphora may be explained by their common occurrence in examples such as (759) and (760). The head noun in these cases is often a newly introduced discourse entity whose referential scope is narrowed by the postmodifying clause (acting as specifics of identity), and whose encapsulation is prospectively realised within the same sentence (e.g. (760)) or elsewhere (e.g. (759)). As revealed by Figure 6.65, the priming for intrasentential cataphora is stronger with participle *-ing* clauses, since, with *to*-infinitive clauses, EXO, CF.INTER and CF.INTRA share an almost equal distribution (29.41%, 29.41%%, 23.53%).

(758) And I thought <pause> well she's got **two chances** that way, <u>either <pause> I could win, or she could win, or <pause> neither of us'll win</u>. (*BNC Sampler*: KCX, S:conv).

(759) I have **an application to make to your Lordship** er in due course when your Lordship feels it is appropriate for an application to be made.[...] <u>Er, in my respectful submission er some <unclear> of inadmissible evidence put before a jury are simply too important and too</u>

central <pause> for a jury to be directed to disregard and then safely <pause> er to the <unclear> which we known that the jury will disregard them. If for instance this was a criminal trial as to which will rule this to inadmissibility exactly the same. In my respectful submission there could be A, no question that the jury would be discharged and B, in the submission er some er judicial displeasure of an experienced police officer volunteering what any police officer must know is A inadmissible and B not an answer to the question. Now my Lord I'm <pause> I make the application with no <pause> with no joy at all but the <unclear> for a day <pause> this matter has proceeded on the basis that yes these two men might have known each other, but that's as far as it goes. (*BNC Sampler*: JJV, S:courtroom).

(760) Part of the work being carried out by IMS for the Science and Engineering Research Council involves a short follow-up study in Spring 1985, to check on the short term destinations of research students, their salary levels, and the extent to which those looking for work in September 1984 managed to find it (see 1.3). (*BNC Sampler*: H0H, W:non_ac:polit_law_edu).

In relation to the four remaining structures, prepositional phrases and restrictive relative clauses show a similar ssociation with cataphora and anaphora, the former being more marked with prepositional phrases, and the latter with restrictive relative clauses (46.10% vs. 39.58% for CF; 24.88% vs. 29.17% for AF). Whilst CF.INTRA and CF.INTER (along with EXO) prevail in both, they occur more frequently with prepositional phrases (26.10% vs. 22.92% for CF.INTRA; 18.78% vs. 14.58% for CF.INTER). Lastly, participle *-ed* clauses and non-restrictive relative clauses are more primed for exophora (34.78% vs. 50.00%).

After examining the influence of formal structure on the distribution of encapsulating relations, Figures 6.67 and 6.68 investigate the connection between these relations and the Theme system.

Figure 6.67 Specific encapsulating directions and Theme/Rheme. Percentages rest on the number of instances for each of the components of the Theme system, i.e. marked Theme, Theme and Rheme

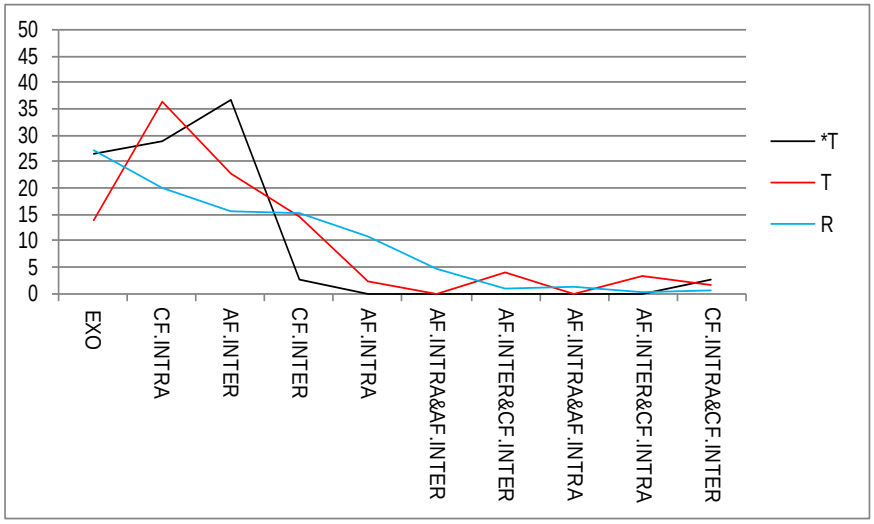


Figure 6.68 General encapsulating directions and Theme/Rheme

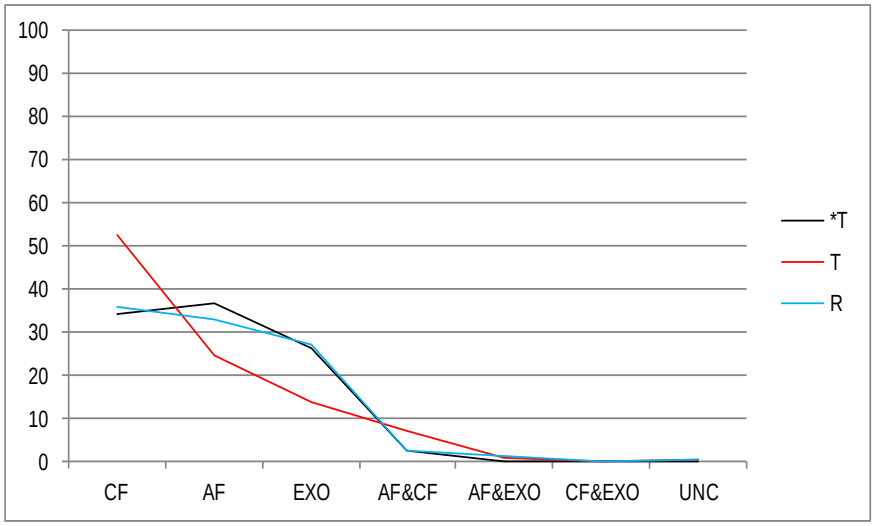


Figure 6.68 shows that cataphoric uses prevail in Theme positions (52.76% for T, 35.87% for R and 34.21% for *T). This is particularly noticeable with intrasentential cataphora (36.55% for T, 28.94% for *T and 20.09% for R), as the frequencies for the intersentential subtype are closely similar for Theme and Rheme positions (15.04% for R, 14.48% for T and 2.63% for *T). By contrast, anaphora is negatively primed for Theme, and occurs more frequently with marked Theme and Rheme (36.84% for *T, 32.84% for R

and 24.83% for T). As shown in Figure 6.67, this association relates to the dominance of Rheme with intrasentential anaphora alone and in combination with its intersentential equivalent (10.73%, R, vs. 2.06%, T, for AF.INTRA; 4.67%, R, vs. 0%, T, for AF.INTRA&AF.INTER). Converely, intersentential anaphora reveals a strong priming for marked Theme and Theme (36.84% for *T, 22.75% for T and 15.59% for R). Theme is also prevalent with combinations of anaphora and cataphora (7.24%, R, 2.63%, *T, 2.39%, R, for AF&CF; 3.79%, T, 1.01%, R, for AF.INTER&CF.INTER).

6.2.4.5 *Discussion*

In the light of the dominance of cataphora in Figure 6.49 (40.29%), it may be argued that Francis' (1994: 89) claim that '[...] retrospective labels are far commoner than advance labels [...]' may apply only to the corpus and the formal structure investigated. The prominent position occupied by anaphora in the literature (see 3.2.2) seems thus to rest on the special attention given to definite and demonstrative noun phrases, which prevail as anaphoric expressions in Biber et al. (1999: 237). This prominence may also explain the marked emphasis on academic prose in shell-noun research (see 2.3.1.2): according to Biber et al. (1999: 237, 274), anaphoric definite and demonstrative noun phrases are strongly primed for this super-genre. Not surprisingly, of the four super-genres considered in Biber et al. (1999), academic prose also makes more use of definite and demonstrative shell-noun phrases than fiction or conversation in the sample (see Figure 6.8).

This super-genre also shows the largest proportion of anaphoric shell nouns (see Figures 6.53 and 6.55). The close correspondence between academic prose and anaphoric encapsulation seems to hinge on the high frequency of factual nouns such as *finding*, *dimension*, *characteristic* or *phenomenon* with this encapsulating direction (see Figure 6.58). According to Figure 6.3, nouns of this type are highly frequent in academic prose, thereby explaining the connection between this super-genre, specific Deictics and anaphora.

Whilst most shell-noun research is primarily concerned with intersentential anaphora and with noun complement clauses, intrasentential patterns other than the latter are just as relevant or even more so. Intrasentential cataphora is most prominent in this respect, ranking closely behind exophora in Table 6.30. Flowerdew's (2003a: 337) observation about the prepositional realisation of signalling nouns is here reflected in the second position held by prepositional complements (LC.CT) in the CF.INTRA column in Table 6.33, showing that appositive prepositional phrases, as in *the practice* ***of exiling malefactors and malcontents to Siberia,*** may prove particularly useful as encapsulating devices in certain written genres

(W:commerce, W:newsp, W:non_ac, W:misc, W:ac). This preference for written genres is logical given the higher occurrence of prepositional postmodification in all written genres except in conversation (cf. Biber et al. 1999: 606 in 6.2.2.4). Intrasentential cataphora in Table 6.33 also reveals that, whilst appositive clauses (LC.CL(AP)) are frequent encapsulating realisations (ranking third), non-appositive clauses (LC.CL) prevail. This is most likely due to their association with the N-*be*-cl pattern, which explains why conversation tops LC.CL use in Figure 6.57. Conversation, as stated in 6.2.3.4, is strongly primed for subject factual shell nouns in focusing constructions like *the thing is that*, *the point is that* or *the problem is that*.

Concerning intrasentential anaphora, its association with simple and complex noun phrases in Table 6.33 appears to relate to the higher occurrence of AF.INTRA cases in W:ac, W:non_ac and W:misc (see Figure 6.55). The former two are most prevalent with complex noun phrases. By contrast, simple noun phrases are primed for S:brdcast and show a weaker association with W:misc and W:ac. The connection between complex nominal antecedents and academic prose (popularised and non-popularised) ties in with Gray's (2010: 179; see 2.3.1.2.1) findings from a corpus of research articles, where complex nominal antecedents prevail over simple ones (26% vs. 7%). Again, this is logical considering the greater frequency of nouns and postmodifying structures in academic prose (Biber et al. 1999: 235, 607). The frequency of simple noun phrase and pronoun antecedents with spoken super-genres is also in line with general noun use (see Figure 6.57; cf. Biber et al. 1999: 235, where pronouns and simple noun phrases are found to prevail in conversation). Conversation, primed for pronominal antecedents, is more prominent with exophora (see Figure 6.53). Thus, whilst the use of antecedents is not absent form conversation, the spontaneous and often unconnected nature of this super-genre hampers the recovery of antecedents in many cases. Only the interlocutors in a particular conversational exchange have all the background information required for the disambiguation of exophoric shell nouns. Biber et al. (1999: 266) corroborate the observed correspondence between exophoric uses and conversation, finding that situational reference (i.e. exophora) accounts for more than 50% of definite noun phrases in this super-genre. Carter & McCarthy (2006: 245) also comment on the rarity of cataphoric nouns in informal spoken language, as evidenced in Figure 6.53, where conversation ranks second to last with cataphora (followed by fiction).

Comparison of the two intersentential encapsulating relations reveals a larger concentration of written genres in cataphora than is apparent in anaphora, where written and spoken genres share a similar distribution (Figure 6.54). If Carter & McCarthy's (2006: 245) claim that '[...] anaphoric references are by far the more common type of endophoric reference' is

anything to go by, it means that, although backward reference or anaphora may indeed be more frequent than cataphora in English at large, this does not imply that cataphora is uncommon. Based on Figures 6.54 and 6.55, intersentential cataphora, whilst rare with spoken genres, is still frequent with such carefully planned and edited super-genres as journalistic and academic prose (popularised and non-popularised). This appears to indicate that prospective encapsulation requires a great deal of planning in order to ensure that the the writer's discourse act is duly fulfilled in a subsequent discourse segment. Winter's (1977) and Tadros' (1985, 1994) focus on the prospective nature of Vocabulary 3 and advance labels makes sense in view of their analysis of highly structured journalistic and textbook prose, where the occurrence of such instances as (761) and (762) is motivated by the need to maintain the reader's interest in what is yet to come. It is worth noting that S:speech, whilst spoken in nature, occurs among the top five super-genres for CF.INTER in Figure 6.55. This suggests that the prospective potential of carefully planned written genres is also evident in formal spoken super-genres like speeches (despite their spoken delivery, they draw on written language to a large extent; see example (763)).

(761) One problem relates to a tactical blunder by Swapo. It allowed a splinter party, Swapo-Democrats, to appropriate and register the symbol which Swapo used for nearly 30 years, a hand holding a flaming torch. Swapo went into the election with a symbol which was only weeks old: a man with a raised fist. The dangers of confusion for illiterate voters is compounded by two factors. Firstly, the Swapo-Democrat emblem is placed just above the Swapo emblem on the ballot paper, meaning that it will be seen first. Secondly, the raised fist in the new Swapo emblem is inconspicuous in the reduced size needed to squeeze Swapo's new symbol into the box on the ballot paper. (*BNC Sampler*: A7V, W:newsp:brdsht_nat:report).

(762) From a practical point of view there are three ways of approaching the valuation of the dependency. The first is quite simply to take each item of expenditure by the deceased on the dependants and add them up, and this will provide the annual dependency figure. A second way of doing this is to take the deceased 's net income and deduct his estimated expenditure on himself. [...]The third way of calculating the multiplicand and now the one that finds most favour with the courts is to deduct a percentage from the deceased 's net income figure to represent what he would have spent exclusively on himself. (*BNC Sampler*: J6W, W:ac:polit_law_edu).

(763) Now on this slide we have a hexagon, there are six probably important facets in terms of standards which we should follow or adopt in the development of the product. [...] If I start at the top with

> the user interface, then clearly we should use or employ a graphical user interface. [...] Well we should use X-networking standards, and at the moment if a network complies to T C P I P as the communications protocol then we can work with that.[...] we should support true electronic data interchange.[...] On operating systems we wanted to be as truly hardware independent as we can, so we have chosen to operate potentially under any Posix-compliant operating system [...] As I identified a minute or two ago, our approach to computerizing accounting applications fits hand in glove with the concept of relational database.[...] we've written our product in C to retain that independence within there. (*BNC Sampler*: HDF, S:speech:unscripted).

Turning to the results reported in 6.2.4.4, and as noted by Hawkins (1978: 172–3), Brown & Yule (1983: 188) and Fraurud (1990: 404) (see 3.2.3), definiteness does not necessarily correlate with anaphora, and neither does indefiniteness relate only to cataphora. Figure 6.61 shows that the definite article ranks highest with intrasentential cataphora, while the indefinite article ranks highest with both intrasentential cataphora and intrasentential anaphora. Cataphoric definiteness seems connected with the strong association with subject positions by definite noun phrases (see Figure 6.42), which also explains the high occurrence of Theme positions with intrasentential cataphora in Figure 6.67. In these cases, the definite shell-noun phrase occurs in N-cl, N-*of* and N-*be*-cl patterns (e.g. *the assumption that*, *the problem of*, *the point is that*), where prospective encapsulation follows from the restriction imposed by the postnuclear complement in N-cl and N-*of*, and from the focusing and topicalizing functions of N-*be*-cl (cf. Schmid 2000: 329–37 in 2.2.2.2). As regards anaphoric indefiniteness, there is a clear association between the dominance of the indefinite article in subject complement positions (see Figure 6.42), its priming for intrasentential cataphora in Figure 6.61, and the prevalence of Rheme positions with intrasentential cataphora in Figure 6.67 (e.g. *The operetta was a bitterly disappointing experience for Britten*).

Finally, and concerning the results in Figure 6.63, a plausible explanation for the higher occurrence of simple noun phrases (i.e. without postmodification) with intersentential anaphora lies in their frequent instantiation in Theme positions (see Figure 6.67). Based on the discussion in 6.2.3.4, and in line with de Haan (1991) and Aarts (2004), Theme positions are negatively primed for structural postmodification, and instead complex noun phrases are associated with Rheme positions. This matches the strong priming of simple noun phrases for intersentential anaphora.

6.3 CONCLUSION

This section summarises the findings in 6.2 as a profile of shell-noun use in the study sample. Following Flowerdew's (2002; see 2.2.2.3) 8-point synthesis of signalling noun behaviour, the results for the nine variables included in this study are similarly brought together under eight points:

i) Shell-noun use is largely concentrated in written super-genres, particularly in W:newsp, W:non_ac and W:ac. A comparison of spoken noun use in *BNCweb* with spoken shell-noun use reveals S:conv, S:meeting and S:speech as particularly prominent with shell nouns. The highest-ranking units in the written mode are semantically more specific and more evenly distributed than those in the spoken mode. Semantically, the most remarkable connection is between academic prose and factual nouns.
ii) The three most frequent determiners are the definite article, the zero article and the indefinite article. Definite shell nouns are more frequent in spoken super-genres, especially S:brdcast, S:speech and S:meeting. They are strongly primed for subject and Theme positions, where W:ac and W:newsp prevail. Intrasentential cataphora and exophora are the most frequent encapsulating directions. Zero article is most common in written genres, especially in W:non_ac, W:newsp and W:commerce. Syntactically, it is primed for prepositional complements and prepositional objects, whilst textually, it is frequently exophoric. Indefinite shell-noun phrases show a marked association with S:conv and W:fict, and with subject complement positions. In terms of semantic preference, the definite article occurs frequently with factual and linguistic nouns, the zero article, with linguistic and eventive nouns, and the indefinite article, with modal and eventive nouns.
iii) Semantic premodification is relatively infrequent. There is a strong association between Epithets and attributive positions, and this explains their frequent association with indefinite noun phrases. Specifically, interpersonal Epithets predominate in spoken super-genres (S:brdcast, S:speech, S:conv) and in fiction. Mental nouns are associated with this semantic premodifier. Experiential Epithets are clearly associated with W:ac, and with linguistic and modal nouns. Classifiers are associated with expository prose (W:commerce, W:misc and W:newsp). Eventive and modal nouns are most frequent with this premodifier. Identity post-Deictics are frequent in S:meeting, and show a marked priming for intensive

Identified positions in focusing constructions (e.g. *the only thing is*, *the main problem is*), hence their frequent use with factual nouns.

iv) Prepositional postmodification by *of*-phrases is highly prevalent among postmodifying structures. Genre distribution in this case shows no marked primings, as prepositional phrases feature similarly in all super-genres. Restrictive relative clauses rank second among postmodifying structures and are nonetheless primed for spoken genres, particularly S:conv, S:speech and S:meeting. Factual, mental and linguistic nouns are primed for prepositional postmodification, whilst circumstantial nouns prevail with restrictive relative clauses.

v) Syntactically, direct object, subject and subject complement positions reveal the highest concentration of shell-noun use. Subject instances prevail in W:ac and W:newsp, while direct object and subject complement cases prevail in S:conv and W:fict. Subject shell nouns are most frequent with modal and factual meanings, the latter also being frequent with subject complement uses. Object shell nouns are primed for linguistic, modal and mental nouns. Modal and linguistic nouns occur among the highest-ranking semantic classes for Rheme positions. Theme is frequently occupied by factual nouns, as is also the case with subject.

vi) With regard to Transitivity, relational participants (especially, intensive Identifier, Attribute and Identified) stand out across super-genres, followed by circumstances. Relational contexts show a strong preference for factual nouns, while circumstances are primed for circumstantial nouns. The semantic type of shell noun is associated with the process type (i.e. relational and factual, material and eventive, verbal and linguistic, circumstance and circumstantial).

vii) Cataphora is the most frequent encapsulating relation, followed closely by anaphora and exophora. Cataphora is most prevalent in W:newsp, and anaphora, in S:brdcast and W:ac. Exophora seems better suited for W:commerce and S:conv. Concerning semantic preferences, cataphora is prominent with modal nouns, anaphora with factual nouns, and exophora with linguistic and eventive nouns.

viii) In terms of lexical realisation, intrasentential cataphora tops the list of encapsulating relations, followed by intersentential anaphora and cataphora, and intrasentential anaphora. Clausal and prepositional complement antecedents prevail with intrasentential cataphora, while noun phrase antecedents prevail with intrasentential anaphora. Intersentential anaphora and cataphora are both strongly primed for global extended antecedents. The most noteworthy

genre and lemma-related findings for antecedent use concern clauses, noun phrases and pronouns. Clauses are primed for W:newsp, S:conv and W:commerce, and for modal, eventive and mental nouns. Noun phrases, especially complex ones, are best suited for expositoy prose (W:ac and W:non_ac) and factual nouns. Pronominal antecedents are also linked to factual nouns, but in this case spoken super-genres, especially S:conv, predominate.

7 CONCLUSIONS

7.1 INTRODUCTION

Chapter 7 concludes this thesis with an appraisal of the study and the proposal of future research avenues. Section 7.2 looks at the strengths of this study in relation to previous research. Section 7.3 turns to the question of what shell nouns are, and is intended to cast light on the nature of these units based on the qualitative and quantitative findings from the study. Section 7.4 describes possible research areas in need of further exploration.

7.2 THE STUDY AND ITS CONTRIBUTION TO SHELL-NOUN DESCRIPTION

More than a decade ago, Schmid (2000: 379) concluded his seminal study on shell nouns suggesting two future research avenues. In the first, shell-noun description was expected to benefit from corpus methods accounting for as many patterns as possible. This could only be achieved by investigating shell-noun use '[...] less mechanically [...]' or, in other words, by drawing less on automated corpus queries and more on the thorough analysis of corpus data (Schmid 2000: 379). In the second, emphasis was laid on the need for further research on the genre- and mode-related distribution of these units, which would involve using better balanced corpora than the *BoE.*

Subsequent to Schmid (2000), several studies broadened the research scope to patterns other than N-cl and N-*be*-cl (e.g. Flowerdew 2003a; Aktas & Cortes 2008; Caldwell 2009). This being the case, most related research to date has revolved around academic discourse. Even Flowerdew (2003a: 331), purporting to provide a '[...] comprehensive pedagogically useful description [...]' of these units (Flowerdew 2003a: 331), uses discipline-specific academic data to support his findings. Whilst not denying the value of the contribution of these units to the structuring of academic texts, Table 6.1 shows that shell-noun use may also be relevant to certain other written and spoken super-genres such as newspapers, formal meetings, fiction, speeches and conversations. Be that as it may, the dominance of shell-

noun use in academic discourse should not be generalised to all shell instances. Even a small sample like the one used in this thesis reveals that genre-related differences are semantically motivated. This is particularly evident in the case of academic prose, where lemmas with factual meanings prevail (cf. Figure 6.3).

The research presented in this thesis explores both of Schmid's (2000: 379) suggested avenues. It uses a small though well-balanced corpus of English and, unlike previous studies, genre information is explicitly coded in the analysis database. In addition, the restrictive nature of automated corpus queries is here addressed through a fully manual and comprehensive approach to corpus data analysis. This thesis thus subscribes to Mahlberg's (2005: 188) view that '[...] detailed analysis and human interpretation is – at present – the only way to developing corpus linguistic theories on textual phenomena'.

Indeed, in the case of shell nouns, where context plays such a key role in the form, function and meaning of these units, the close reading of discourse passages spanning sentences, paragraphs and even pages becomes indispensable. This is shown in the amount of detail aimed at in many of the examples in this thesis, where the limited scope of the concordance line is extended to long discourse segments in order to offer an accurate representation of readers' or listeners' context-specific 'discourse models' (Prince 1981: 235). Such an analysis questions the validity of purely quantitative concordance-based approaches in cases where the focus is not only on meaning and form within the limits of the concordance, but also on how meaning and form emerge from and relate to the function that a particular discourse entity performs at a specific point in discourse. This is one of the intended contributions of this thesis to shell-noun description: to the best of my knowledge, this is the first study where formal, syntactic, semantic and textual features of these units are all systematically investigated. Notwithstanding the small-scale nature of the study sample, the fine-grained annotation scheme used here reveals a larger number of features than automated methods do. Thus, as in Mahlberg (2005: 180), it is here believed that '[...] the quantitative limitations' are outweighed by '[T]he nature of the results and their implications [...]'. Section 7.3 looks at some of these implications in the light of the results obtained.

7.3 WHAT ARE SHELL NOUNS?

The quantitative and qualitative analysis of this thesis shows that shell nouns are, first and foremost, nouns. Whilst this may seem a truism, the affinity of shell nouns with nouns in general is not sufficiently discussed in

the literature. This is especially so in view of claims such as Hunston & Francis' (2000: 185) that shell nouns represent a separate class of nouns. The quantitative evidence in chapter 6 points to a number of similarities between the class of nouns in general and shell nouns:

i) The definite article prevails in both, followed by the indefinite article, possessive determiners and demonstrative determiners (see Table 6.9 and 6.2.2.4).
ii) Possessive determiners predominate in fiction, and demonstrative determiners do in academic prose (see Figure 6.8 and 6.2.2.4).
iii) Premodifiers are common in written genres, particularly in journalistic and academic prose (see Figure 6.9 and 6.10, and 6.2.2.4).
iv) Evaluative adjectives are frequent in fiction and conversation, whilst classifying and descriptive adjectives prevail in journalistic and academic prose (see Figure 6.12 and 6.2.2.4).
v) Prepositional postmodification (especially, *of*) is the most frequent postmodifying structure in both (see Table 6.11 and 6.2.2.4).
vi) Syntactically, noun and shell-noun use is more frequent in object and prepositional complement position than in subject position (see Table 6.20 and 6.2.3.5).
vii) Subject nouns are strongly primed for academic and journalistic prose, whilst object nouns prevail in fiction (see Figure 6.25 and 6.2.3.5).
viii) Structural postmodification is most frequent in post-verbal positions, whilst the lack thereof is associated with subject positions (see Figure 6.43 and 6.2.3.5).
ix) Definite noun phrases occur frequently as subjects, and indefinite nouns do as subject complements (see Figure 6.42 and 6.2.3.5).
x) Anaphoric uses stand out in academic prose, and exophoric uses, in conversation (see Figure 6.53 and 6.2.4.5)

Shell instances behave like nouns in their association with the same types of determiners, premodifiers, postmodifying structures and clause-level positions. This said, the question arises as to what makes a noun a shell noun. Although the answer still remains elusive, the following discusses key findings from this study in the light of previous definitions:

i) Lack of semantic specificity and immediate lexical realisation:

- '[...] a label [...] is an **inherently unspecific** nominal element whose specific meaning in the discourse needs to be **precisely spelled out**'. (Francis 1994: 83; my emphasis)
- '[...] these are nouns which require lexicalisation in **their immediate context**' (Hunston & Francis 2000: 185; my emphasis)

In the literature, shell-noun use is associated with the inherent lack of semantic specificity in second and third-order abstract entities. The belief here is that lack of specificity is not a monolithic concept or property, but one that varies according to meaning and context. The meaning of shell units is instantiated in highly frequent words like *thing*, *way*, *area* or *point* and in infrequent ones like *contradiction*, *endeavour*, *foreboding* or *recollection*. This thesis shows a connection between frequency and shell uses, such that high frequency involves greater dependence on contextual specifics, whilst low frequency leads to more self-contained non-shell uses. Low-frequency ranges in the study sample are occupied by generic and nominalised lemmas (e.g. *terror*, *anger*, *impetus*, *contradiction*, *endorsement*, *endeavour*). As discussed in 4.4.2.8 and 4.4.2.9, these challenge the distinction between shell and non-shell uses, as well as the value of the connection between frequency and shell use, so much so that unless further conditions are researched, no firm conclusion can be drawn regarding frequency and shells.

Context is a major disambiguating factor in the identification of shell instances in this thesis. The analysis adopts the perspective of the reader or listener in a specific discourse situation in order to gauge the contextual significance of a particular noun. It is important to stress that shell-noun specifics cannot always be '[...] precisely spelled out', and neither is their lexical realisation always in '[...] the immediate context', as stated by Francis (1994: 83) and Hunston & Francis (2000: 185) in the above quotations. Such an emphasis on precise and immediate specifics stems from the often-repeated claim about the definite and demonstrative nature of shell nouns, as well as their use in N-cl and N-*be*-cl patterns.

Thus, according to the literature, examples like (764) and (765) would represent prototypical shell uses, as the noun phrase in both is definite, and contextual specifics may be clearly delimited. However, example (766), described in chapter 5 as (613), is fuzzy in its contextual significance, as informative specifics are neither clear nor immediate. This being the case, *terror* is still context-

dependent, since it is not terror or violent actions in general that inspired this statement, but the specific terror used by the Romanian government over the weekend (i.e. violence, loss of life, factories under guard, brutality, children shot at, military intervention, etc.).

(764) Lethal Force tested the suspicion that men who could have been captured or wanted to surrender had been shot dead [...] (*BNC Sampler*: J1L, W:misc)

(765) Mr Baker said the main area of concern for over the next few years was 'serious downturn' in the economies of export markets in Europe and Japan. (*BNC Sampler*: CF8, W:newsp:other:report)

(766) By Foreign Staff OUTRAGE and concern at the violence in Romania during the weekend came from all parts of the world yesterday, with the Soviet Union and the US in the lead. [...] European foreign ministers, meeting in Brussels, condemned 'in the strongest possible terms' the measures taken by the Romanian security forces. [...] he said 'if indeed some loss of life has occurred, I can only express my very profound regret'. [...] It reported that Romanian state institutions and factories were under intensified guard, and that the frontiers were closed to tourists. [...]The Polish Parliament stood for a minute's silence after approving unanimously a resolution which accused the Romanian authorities of 'exceptional brutality' leading to children being shot at and expressed solidarity with the 'victims of terror.' The World Council of Churches, the World Alliance of Reformed Churches, the Lutheran World Federation and the Conference of European Churches said in a joint telegram they were 'deeply disturbed by the disquieting intervention of the army against civilian people in Timisoara.' (*BNC Sampler*: AAB, W:newsp:brdsht_nat:report)

Therefore, lack of semantic specificity is not necessarily linked to definite or demonstrative abstract entities. It may occur with any formal realisation of the noun phrase provided that the noun is contextually meaningful. With this in mind, Ivanič's (1991: 112, my emphasis) observation that, in the study of carrier nouns, '[...] the context seems to play a role in their interpretation even when they are accompanied by indefinite reference, or when they occur as uncountables' could be less tentatively rephrased as 'the context always plays a role in the interpretation of shell-noun instances, regardless of the determiner used'. The importance of context is such that even concrete first-order entities may acquire shell uses in cases where the focus is on the metadiscursive content of the noun, this being propositional or eventive in nature. This is evident in example (767), presented in chapter 3 as (190). Consten et al.

(2007: 82) treat *this graffito* as a non-shell instance on the grounds of its denotation of a visible first-order entity (i.e. paint on a wall). However, analysed in context, a shell-noun interpretation arises, as the focus is on the product and on the propositional content or information of such a product (i.e. the underlined segment).

(767) 'At the end of the eighties a rogue characterized the decadent gluttony: 'Buy nicer and nicer, fly further and further, shag faster and faster'. During several months **this graffito** ornamented the walls of an old comfort station at Hamburg's Yuppie district [...]' (Consten et al. 2007: 82)

ii) Open- and closed-class status and the pro-form criterion:

- '[...] a general noun is itself a **borderline case** between a **lexical item** (member of an open set) and a **grammatical item** (member of a closed system)' (Halliday & Hasan 1976: 274; my emphasis)
- '[...] an A-noun [...] must be functioning as a **pro-form** and as such be an **anaphorically cohesive device**, referring metadiscursively to a stretch of discourse [...]' (Francis 1986: 4; my emphasis)
- 'The head nouns of retrospective labels are almost always preceded by a specific deictic like *the, this, that* or *such*, and may have other modifiers and qualifiers too. The whole nominal group functions very much like a **pro-form or reference item**.' (Francis 1994: 85; my emphasis)

Francis (1986 and 1994) draw on Halliday & Hasan (1976: 274)'s definition of general nouns (e.g. *thing*, *fact*, *person*, *creature*) as half lexical and half grammatical to describe anaphoric nouns and labels, respectively. Following Halliday & Hasan (1976), shell-like units are often likened to pronouns and pro-forms (e.g. *this*, *they*, *it*, *so*) on the basis of their definiteness, frequent anaphoricity and semantic unspecificity. Ivanič (1991: 107–8) argues that carrier nouns resemble pronouns in their semantic unspecificity, but also that they are nouns in every other respect (i.e. more informative, evaluative potential and pre- and postmodification). Mahlberg (2005: 177) settles the question of the categorial indeterminacy of general nouns by positing that a corpus linguistic approach to the study of these units is blind to any watertight grammatical distinctions.

Indeterminacy is at the core of corpus linguistic findings, as nouns (and, by extension, any other category in the conventional system of word-classes), are shown to contain fuzzy edges when explored

in the light of corpus data. Thus, the research focus should be on '[...] meaning as use' (Mahlberg 2005: 177), or on how meaning is formally instantiated in specific text types. This explains Mahlberg's (2005) identification of the specific meaning categories linked to the use of general nouns in context (i.e. local textual functions; e.g. time orientation, measurement, people in contrast to government, humans/men, etc.).

In this thesis, the assumed categorial indeterminacy in shell-noun use is not subscribed to. As stated above, shell nouns behave like nouns in their form, syntactic function and lexical meaning and, as such, they are treated as nouns. What makes these nouns special is their abstract meaning and their need for contextual specifics to clarify their contribution to discourse. In view of the formal, syntactic and semantic associations of shell units in general with the word-class noun, following Mahlberg (2005) and Hoey (2005), the real interest lies not in their nominal or pronominal categorisation, but in how the context-dependent meanings of second- and third-order abstract shells in specific text types influence their overall formal, syntactic and textual behaviour.

iii) Shell nouns and their primings: From the above, it is clear that an accurate and comprehensive definition of shell nouns cannot rest on formal and syntactic criteria, as is often the case in the literature. This being the case, only the semantic and textual variables in this thesis seem to offer revealing insights into '[...] the property of shell-nounhood [...]' (Schmid 2000: 13). The widespread use of these units with relational processes (see Figure 6.22b) and long antecedents (clausal, sentential and extended: 64.96%; see Table 6.34) is particularly noteworthy in this respect, failing to reflect the expected association between concrete first-order nouns, material processes and short antecedents. On these grounds, shell nouns are shells by virtue of their frequent occurrence with relational processes and their encapsulation of discourse segments. In this thesis, the latter confirms the often-repeated claim about shell-noun reference to stretches of discourse by detailed quantitative data comprising a wide range of antecedent types. The fine-grained antecedent taxonomy in this thesis (see 5.3.4.2) shows that, despite the close link with long antecedents, 35.04% of shell instances accounts for phrasal antecedents (mainly noun phrases and prepositional complements).

A definition of shell-noun use on the basis of participant type and antecedent is still not sufficiently accurate and comprehensive, as it

fails to capture the influence that meaning and text type may have on the formal, syntactic and semantic behaviour of shell instances. In order to illustrate such an influence, Table 7.1 brings together the five highest-ranking categories for the variables genre, formal structure, semantic structure, syntactic function, participant, Theme/Rheme, encapsulating direction and antecedent (the two leftmost columns), alongside the top three super-genres and semantic types for each category (remaining columns). The six rightmost columns explore the link between structure (formal and semantic), direction of encapsulation and other variables further (see chapter 6 for the full results). The following draws on the categories in red to outline some of the most remarkable shell-noun primings in the study sample (see chapter 6 for full details):

- Academic prose and formal meetings are strongly primed for the use of factual and linguistic nouns respectively. Semantic preferences relate closely to the typical communicative functions of each super-genre (e.g. providing factual information, offering suggestions for improvement, etc.; see 6.2.1.3).
- Definite shell nouns are often linked to subject positions and intrasentential cataphoric realisations. Undetermined instances are markedly primed for linguistic meanings, prepositional complement positions and exophoric uses. Indefinite nouns prevail in conversation and in subject complement positions. Demonstrative determiners are highly frequent in academic prose, and rarely occur with uses other than intersentential anaphora.
- Prepositional postmodification stands out with factual, linguistic and mental nouns. Relative clauses show a strong priming for circumstantial nouns and spoken super-genres (especially, conversations and speeches). Noun complement *to*-infinitive clauses are most frequent with modal meanings, and prevail in journalistic prose and conversations. Noun complement *that* clauses are strongly primed for mental nouns, as well as for notional subject and prepositional object positions.
- As regards semantic premodifiers, the most noticeable associations are those between Classifiers, eventive and modal meanings, between interpersonal Epithets, fiction, broadcast programmes, mental nouns and intensive attributive positions, between experiential Epithets,

academic prose and intensive attributive positions, and between elaborating identity post-Deictics, formal meetings and intensive Identified positions.

- Syntactically, object shell nouns are most frequent in conversation and fiction, whilst subject and Theme uses prevail in academic and journalistic prose. The former are strongly primed for linguistic and modal meanings, and the latter for modal and factual meanings.
- In relation to participant type, the results reveal a close association between Transitivity components and semantic types of shell noun. Relational participants are strongly primed for factual nouns, circumstances for circumstantial nouns, material participants for eventive nouns, and verbal participants for linguistic nouns (see 6.2.3.5 for details).
- Turning to direction of encapsulation, exophoric uses are strongly primed for economic treatises and conversations, linguistic and eventive nouns, and Rheme and marked Theme positions. Intrasentential cataphora occurs most frequently in newspapers, and is strongly associated with modal meanings and Theme positions. Intersentential anaphora is most frequent in academic prose and broadcast programmes, as well as in Theme positions. Intersentential cataphora, by contrast, prevails in leaflets, user manuals and other miscellaneous texts, and is almost equally distributed in Rheme and Theme positions. Lastly, intrasentential anaphora is most clearly associated with eventive and factual nouns in Rheme positions.
- The strongest associations concerning antecedent types are those between clauses and modal nouns, between simple noun phrases and broadcast programmes, between complex noun phrases and academic discourse (professional and popularised), and between sentences and fiction.

Table 7.1 Shell-noun use in the study sample (overview). The categories in red are those for which the linear patterns of figures in chapter 6 stand out. Yellow shading indicates cases where shell-noun use in the study sample coincides with noun behaviour in general

	Top 5 categories	Super-genre (Top 3)			Semantic type (Top 3)			Syntactico-semantic function and Theme/Rheme (Top 3)			Direction of encapsulation (Top 3)		
Super-genre	W:newsp				Eventive	Mental	Linguistic						
	W:non_ac				Mental	Linguistic	Factual						
	S:meeting				Linguistic	Factual	Mental						
	W:ac				Factual	Mental	Circumstantial						
	W:misc				Mental	Factual	Linguistic						
Formal structure (Determiner)	DF.AR	W:misc	S:brdcast	S:speech	Factual	Linguistic	Circumstantial	SB	PO	SCL	CF.INTRA	EXO	AF. INTER
	Ø	W:non_ac	W:commerce	W:newsp	Linguistic	Eventive	Factual	CP	PO	DO	EXO	CF.INTER	CF. INTRA
	IN.AR	S:conv	W:commerce	W:fict	Modal	Eventive	Factual	SCL	DO	CP	AF.INTRA	CF.INTRA	EXO
	PS.DT	S:brdcast	W:fict	W:ac	Modal	Mental	Eventive	PO	CP	SB	EXO	CF.INTRA	AF. INTER
	DM.DT	W:ac	S:meeting	W:commerce	Circumstantial	Factual	Mental	CP	SB	PO	AF.INTER	AF.INTRA	CF. INTER
Formal structure (Postmodifier)	PP	S:brdcast	W:non_ac	W:fict	Factual	Linguistic	Mental	DO	SCL	PO	CF.INTRA	EXO	CF. INTER
	RT.RV.CL	S:conv	S:speech	S:meeting	Circumstantial	Eventive	Linguistic	SB(NOT)	SCL	SB	EXO	CF.INTRA	CF. INTER
	AP.TI.CL	W:newsp	S:conv	S:brdcast	Modal	Eventive	Linguistic	DO	CP	SB	CF.INTRA		
	PL.ED.CL	W:misc	W:non_ac	S:meeting	Eventive	Factual	Mental	CP	SCL	SB	EXO	AF.INTRA	CF. INTRA
	AP.THAT.CL	W:newsp	W:commerce	W:misc	Mental	Linguistic	Circumstantial	SB(NOT)	PO	DO	CF.INTRA	EXO	
Semantic structure	CS	W:commerce	S:brdcast	W:misc	Eventive	Modal	Circumstantial	GOA	ATT(rel.in)	IDR (rel.in)			
	EP.IP	W:fict	S:brdcast	S:speech	Mental	Linguistic	Factual	ATT (rel.in)	GOA	IDR (rel.ci)			

(Premodifier)	EP.EX	W:ac	W:non_ac	W:newsp	Linguistic	Modal	Factual	ATT (rel.in)	IDR (rel.in)	GOA	
	PDC.EL.ID	S:meeting	S:speech	W:fict	Circumstantial	Factual	Modal	IDD (rel.in)	VER	IDR (rel.in)	
	PD.EL.ID.RT	S:conv	S:meeting	W:fict	Factual	Circumstantial	Modal	IDD (rel.in)	IDR (rel.in)	VER	
Syntactic function	DO	S:conv	W:fict	S:brdcast	Linguistic	Modal	Mental				
	SB	W:ac	W:newsp	W:non_ac	Modal	Factual	Mental				
	SCL	S:conv	W:fict	S:brdcast	Factual	Mental	Eventive				
	SB(not)	S:brdcast	S:speech	S:conv	Circumstantial	Factual	Modal				
	PO	S:speech	S:meeting	W:misc	Circumstantial	Linguistic	Eventive				
Participant type	IDR (rel.in)	W:misc	W:non_ac	S:brdcast	Factual	Eventive	Circumstantial				
	ATT (rel.in)	S:brdcast	S:conv	W:fict	Eventive	Factual	Mental				
	GOA	S:speech	W:commerce	S:meeting	Linguistic	Modal	Circumstantial				
	IDD (rel.in)	S:meeting	W:ac	S:conv	Factual	Mental	Circumstantial				
	VER	W:commerce	S:meeting	S:conv	Linguistic	Mental	Factual				
Process type	Relational	W:non_ac	S:brdcast	W:ac	Factual	Modal	Mental				
	Circumstance	S:speech	W:newsp	W:fict	Circumstantial	Eventive	Mental				
	Material	W:commerce	W:newsp	S:meeting	Eventive	Modal	Mental				
	Mental	W:fict	S:conv	S:brdcast	Linguistic	Mental	Circumstantial				
	Verbal	W:commerce	S:meeting	S:conv	Linguistic	Eventive	Mental				
Theme/Rheme	R	W:fict	S:speech	W:commerce	Eventive	Modal	Linguistic				
	T	W:ac	W:newsp	W:misc	Factual	Modal	Mental				
	*T	S:conv	W:commerce	W:newsp	Circumstantial	Mental	Factual				
Direction of encapsulation	CF	W:newsp	W:non_ac	W:misc	Modal	Mental	Circumstantial				
	AF	S:brdcast	W:ac	W:fict	Factual	Eventive	Circumstantial				

(General)	EXO	W:commerce	S:conv	W:misc	Linguistic	Eventive	Mental			
	CB	S:brdcast	W:fict	S:speech	Factual	Circumstantial	Mental			
Direction of encapsulation (Specific)	EXO	W:commerce	S:conv	W:misc	Linguistic	Eventive	Mental	R	*T	T
	CF.INTRA	W:newsp	S:conv	W:commerce	Modal	Mental	Circumstantial	T	*T	R
	AF.INTER	W:ac	S:brdcast	S:meeting	Mental	Factual	Linguistic	*T	T	R
	CF.INTER	W:misc	W:newsp	W:non_ac	Linguistic	Mental	Circumstantial	R	T	*T
	AF.INTRA	W:misc	W:ac	W:non_ac	Eventive	Factual	Modal	R	T	
Antecedent	GB.ET.PR	S:meeting	W:misc	W:non_ac	Linguistic	Circumstantial	Factual			
	LC.CL	W:newsp	S:conv	W:commerce	Modal	Eventive	Mental			
	LC.SNP	S:brdcast	W:misc	W:ac	Circumstantial	Factual	Eventive			
	LC.CNP	W:non_ac	W:ac	W:newsp	Factual	Eventive	Mental			
	LC.SC	W:fict	S:conv	S:brdcast	Linguistic	Mental	Factual			

All in all, although shell nouns share certain features (e.g. their abstract meaning and their relational roles), their behaviour in real discourse situations reveals a range of semantic and genre-specific primings. These preclude any generalisations about their use, in that their formal, syntactico-semantic and textual features are shaped by their context-specific senses in particular text types. Thus, on the surface, shell nouns do not differ from nouns in general. It is only when explored in relation to particular contexts and meanings that differences emerge. Such differences indicate that shell-noun use is a matter of degree, as also observed by Ivanič (1991:109) and Schmid (2000: 14). However, any quantification of such a degree should not rest on morphological, semantic and syntactic criteria alone, as in Schmid (2000: 85–6), where shell nouns are classed as prime, good or less good depending on their semantic type (e.g. prime: factual nouns, less good or peripheral: eventive and circumstantial nouns), their morphological structure (e.g. prime: non-derived factual nouns, good: mental or linguistic nominalisations, less good: eventive nominalisations) and their formal pattern (e.g. prime and good: N-cl and N-be-cl, less good: other patterns) (see 3.2.5.1 for further details).

In line with Ivanič (1991: 109–12), the thesis presented here is that shell nouns occur in a cline of context-dependency: one which classes nouns not only in terms of their form and meaning, but also in terms of the extent to which context-dependent uses arise in particular text types. Context-dependency challenges preconceptions about shell uses in favour of comprehensive approaches to data analysis. In this respect, the contention is that shell nouns are not shells due to gaps inherent in the semantic structure of the lemma, as Schmid (2000: 76) claims. Instead, 'shell-nounhood' (Schmid 2000: 13) stems from contextually-imposed gaps. The examples presented in this thesis show how these gaps may be associated with both deictically specific noun phrases and with seemingly non-shell units like indefinite and uncountable nouns.

This thesis aims at identifying and quantifying all possible shell uses, but the size of the sample prevents any firm conclusions about the context-dependent cline of shell nouns. As stated in (i) above, further evidence is needed to ascertain the potential for shell-nounhood among low-frequency non-projecting nominalisations (e.g. *assessment*, *correction*, *examination*, etc.). In this respect, the best that can be claimed of this thesis is that it paves the way for further research into how shell-nounhood materialises in Schmid's (2000: 86) 'less good' or 'peripheral' shell units.

7.4 WAYS FORWARD

Despite the strengths of the multifaceted approach presented here, this thesis only scratches the surface of the range of semantic and genre-related primings involved in shell-noun use. Thus, the conclusions drawn in relation to meaning and genre are only tentative, as they are based on a limited set of lemmas (only 60) and data (1447 concordances). Further evidence is needed to determine whether the meaning and genre primings observed here relate only to the data considered or whether they indeed indicate assocations between the formal, syntactico-semantic and textual behaviour of shell units, and particular genres and meanings.

In what follows, a range of future research avenues are suggested to address the limitations of this thesis and, as a result, to pursue further as yet uncharted and under-researched areas of enquiry. These fall into two subsections, 7.4.1 on research areas related to shell-noun use according to genre, discourse structure and social purposes, and 7.4.2 on the limits and coverage of the shell or signalling function.

7.4.1 *Genres and sub-genres, discourse structure and social implications*

This first set of suggestions comprises four possible research areas:

i) More data are needed to confirm or reject the genre-related and semantic primings found in this thesis. This would involve broadening the analytical scope to more lemmas and more evidence for each lemma (e.g. 100 concordances, as in Mahlberg 2005, instead of only 40).
ii) Substantial research has been devoted to shell-noun use in a number of academic disciplines and sub-genres, but the focus is still primarily on specific formal patterns and their functions (e.g. *th*-N, N-cl). There is a need for multifaceted analyses of shell-noun use in particular academic sub-genres and disciplines, where formal, syntactic, semantic and textual features of shell nouns are investigated systematically and thoroughly. Such a need would apply not only to academic discourse but, by extension, to any other super-genre.
iii) The influence of discourse structure on shell-noun behaviour needs further research. To the best of my knowledge, the scant research available on such an influence merely identifies the range of shell

nouns occurring in specific moves[41] (e.g. Background, Purpose, Method, etc. in journal abstracts as in Lin 2012; see 2.3.1.2.2) and rhetorical acts (e.g. description of structure, measurement and quantification, exemplification, etc. as in Flowerdew 2003b; see 2.3.2.2). Little is said about how the formal, syntactico-semantic and textual behaviour of shell nouns varies across moves and acts.

iv) The claim is often made that the characterising function of shell nouns may also entail manipulative purposes (e.g. Conte 1996: 6; Schmid 2000: 8; Schmid 2001). In using a particular shell noun, the writer or speaker may be persuading the reader or listener to take a particular statement or claim at face value, which '[...] politicians and other people with debating experience [...]' (Schmid 2000: 8) are proficient in. In an article on the presuppositional meaning of shell nouns, Schmid (2001; see 2.3.2.1) finds that the N-*be*-cl pattern is particularly well suited for manipulative purposes (e.g. *the fear is that*, *the problem is that*), as, by using the definite article, the writer or listener introduces shell contents as general truths (cf. *the fear is that* vs. *my fear is that*).

Whilst the manipulative potential of shell units is acknowedged in the literature, this is often no more than a passing comment and, in those cases where it is explicitly investigated (as in Schmid 2001), attention is only given to specific patterns (e.g. N-*be*-cl). On these grounds, it is here believed that shell-noun description would benefit from more explicit attention to the ideological and manipulative connotations linked to particular shell meanings and patterns (formal and semantic). Such a research goal is consistent with the principles underlying Critical Discourse Analysis (henceforth, CDA), as epitomised by Fairclough (1995, 2000), van Dijk (1997), van Leeuwen (2006, 2008) and Wodak & Meyer (2009). Combining CDA with the multifaceted approach to discourse analysis of this thesis could provide revealing insights into how highly ideological genres such as political speeches, debates and journalistic prose make use of shell nouns to get their messages across.

[41] See Swales (1990) and Bhatia (1993) for the theoretical and practical underpinnings of move analysis in genre theory (especially in academic and professional contexts).

7.4.2 *Formal limits of the shell or signalling function*

The second set of suggestions comprises two research areas:

i) Schmid (1999: 222-3; see 3.2.1) raises the issue of the shell-noun status of nominalised second- and third-order abstract entities. His explanation seems to suggest the existence of a cline ranging from typically context-dependent (and thus, shell) units like *warning*, *claim* or *assumption* to semantically abstract self-contained (and thus, non-shell) entities like *inflation*, *love* or *derivation*. In 4.4.2.9, the analysis of the study sample points to a similar conclusion. Nouns like *warning* or *recommendation* feature as typical instances of the property of shell-nounhood, as their inherent semantic gaps are inherited from the clausal complements of their verbal counterparts (i.e. *warning that* < *warn somebody that*, *recommendation that* < *recommend that*). In cases like *assessment* and *endorsement*, where their corresponding verbs do not allow projection of *that* or *to*-infinitive clauses, shell-nounhood is a matter of degree. Only if the nominalisation is product-like, rather than act-like, does a shell-noun interpretation arise. However, there are instances where even act nominalisations may suggest shell interpretations (see 4.4.2.9 for details). The surrounding co-text is essential for the distinction between shell- and process-like non-shell uses. Future research should explore the interface between derivational morphology and shell-noun use in order to quantify and cast light on the degree to which various non-projecting nominalisations (e.g. *assessment*, *endorsement*, *correction*) allow shell interpretations. Such a goal could be accomplished by looking at the link between a range of word-formation processes (e.g. affixation, conversion, etc.) and context-dependent shell-uses.

ii) Francis (1986: 104) ends her monograph on anaphoric nouns by arguing that '[A]ny item, or almost any item, can be used to achieve anaphoric or cataphoric cohesion, so long as the context allows such an interpretation'. This statement raises the question of whether the property of shell-nounhood relates only to semantically abstract second- and third-order nouns or whether it applies to any noun in need of contextual specifics. Examples (767) above and (768), presented in 3.2.2 as (208), are two cases in point. Caldwell (2009: 46) argues that *this cat* in (768) does not imply a merely deictic use of *cat* (i.e. this is a cat) but, rather, its meaning is here influenced or couloured by the writer's perception of the animal. Thus, the focus is not so much on the first-order entity, as on the

representation of the entity at a particular point in the universe of discourse (i.e. not any cat, but this cat with two small piercing eyes, the attitude of a landmine, etc.).

(768) 'Two small piercing eyes. The attitude of a temperamental landmine. Ten years old and with timing that put my alarm clock to shame. This cat is what stood between me and my first cup of coffee every morning […]' (Caldwell 2009: 46)

Francis (1986: 104) even goes so far as to suggest the existence not only of anaphoric nouns, but also of 'anaphoric verbs' (e.g. *aware of what was happening*), 'anaphoric adjectives' (e.g. *even more extraordinary is the notion that*) and 'anaphoric adverbs' (e.g. *more disastrously*; no further context is given).

Later on, Partington (1998: 101–4), in his discussion of labels and general nouns, includes 'general verbs' (e.g. *happen*, *occur*), which are used to replace more specific actions or events. Such an emphasis on anaphoric or cataphoric uses of words other than nouns is inspired by Winter's (1977: 20) class of Vocabulary 3 items (see 2.2.2.1.2 for details), comprising clause-relational nouns like *attribute*, *justification* and *method*, verbs like *affirm*, *confirm* and *repeat*, and adjectives like *hypothetical*, *similar* and *converse*. Whilst non-nominal Vocabulary 3 items have been investigated in terms of their signalling of clause relations and discourse patterns (Hoey 1979, 1983, 1994; Winter 1977, 1982, 1992; e.g. Cause-Consequence, Problem-Solution, etc.), the research focus has often been on the in-depth analysis of individual texts.

Echoing Hoey's words in 1993, '[...] Winter's (1977) Vocabulary 3 items are more complex in their functioning than either he or I bargained for [...] we have only just begun the proper description of signalling in discourse', there is a clear need for a study where metadiscursive signalling is thoroughly investigated on the basis of multifaceted analyses of corpus evidence, including formal, semantic, syntactic and textual variables.

APPENDICES

APPENDIX 1 LEE'S (2001) *BNC* GENRE AND SUPER-GENRE CATEGORIES

Genre categories	Descriptive information
W:ac:humanities_arts	academic prose: humanities
W:ac:medicine	academic prose: medicine
W:ac:nat_science	academic prose: natural sciences
W:ac:polit_law_edu	academic prose: politics law education
W:ac:soc_science	academic prose: social and behavioural sciences
W:ac:tech_engin	academic prose: technology computing engineering
W:admin	administrative and regulatory texts
W:advert	print advertisements
W:biography	biographies and autobiographies
W:commerce	commerce and finance, economics
W:email	e-mail sports discussion list
W:essay:school	school essays
W:essay:univ	university essays
W:fict:drama	plays (drama scripts)
W:fict:poetry	Poems
W:fict:prose	novels and short stories
W:hansard	parliamentary proceedings
W:institut_doc	official and governmental documents
W:instructional	Instructional and DIY texts
W:letters:personal	personal letters
W:letters:prof	professional and business letters
W:misc	miscellaneous texts
W:news_script	TV autocue data
W:newsp:brdsht_nat:arts	broadsheet national newspapers: arts and cultural material
W:newsp:brdsht_nat:commerce	broadsheet national newspapers: commerce and finance
W:newsp:brdsht_nat:editorial	broadsheet national newspapers: letters to the editor
W:newsp:brdsht_nat:misc	broadsheet national newspapers: miscellaneous material
W:newsp:brdsht_nat:report	broadsheet national newspapers: home and foreign news reportage
W:newsp:brdsht_nat:science	broadsheet national newspapers: science material
W:newsp:brdsht_nat:social	broadsheet national newspapers: material on lifestyle and leisure
W:newsp:brdsht_nat:sports	broadsheet national newspapers: sports material
W:newsp:other:arts	regional and local newspapers: arts
W:newsp:other:commerce	regional and local newspapers: commerce and finance
W:newsp:other:report	regional and local newspapers: home and foreign news reportage
W:newsp:other:science	regional and local newspapers: science material
W:newsp:other:social	regional and local newspapers: material on lifestyle and leisure
W:newsp:other:sports	regional and local newspapers: sports material
W:newsp:tabloid	tabloid newspapers
W:non_ac:humanities_arts	non-academic/non-fiction: humanities
W:non_ac:medicine	non-academic: medical/health matters
W:non_ac:nat_science	non-academic: natural sciences
W:non_ac:polit_law_edu	non-academic: politics law education
W:non_ac:soc_science	non-academic: social and behavioural sciences
W:non_ac:tech_engin	non-academic: technology, computing, engineering
W:pop_lore	popular magazines
W:religion	religious texts, excluding philosophy
S:brdcast:discussn	TV or radio discussions
S:brdcast:documentary	TV documentaries
S:brdcast:news	TV or radio news broadcasts
S:classroom	non-tertiary classroom discourse
S:consult	medical and legal consultations
S:conv	spontaneous conversations

S:courtroom	legal presentations or debates
S:demonstratn	'live' demonstrations
S:interview	job interviews and other types
S:interview:oral_history	oral history interviews/narratives, some broadcast
S:lect:commerce	lectures on economics commerce and finance
S:lect:humanities_arts	lectures on humanities and arts subjects
S:lect:nat_science	lectures on the natural sciences
S:lect:polit_law_edu	lectures on politics, law or education
S:lect:soc_science	lectures on the social and behavioural sciences
S:meeting	business or committee meetings
S:parliament	transcribed parliamentary speeches
S:pub_debate	public debates, discussions, meetings
S:sermon	religious sermons
S:speech:scripted	planned speech, whether dialogue or monologue
S:speech:unscripted	more or less unprepared speech, whether dialogue or monologue
S:sportslive	'live' sports commentaries and discussions
S:tutorial	university-level tutorials
S:unclassified	miscellaneous spoken genres

Super-genres	
W:ac	W:religión
W:admin	S:brdcast
W:advert	S:classroom
W:biography	S:consult
W:commerce	S:conv
W:email	S:courtroom
W:essay	S:demonstratn
W:fict	S:interview
W:hansard	S:lect
W:institut_doc	S:meeting
W:instructional	S:parliament
W:letters	S:pub_debate
W:misc	S:sermon
W:news_script	S:speech
W:newsp	S:sportslive
W:non_ac	S:tutorial
W:pop_lore	S:unclassified

APPENDIX 2 *BNC SAMPLER* FREQUENCY LIST DIVIDED INTO THE THREE MAIN FREQUENCY RANGES. THE SHADED CELLS REPRESENT THE 60 SAMPLED LEMMAS

1 (High)	Freq.	2 (Middle)	Freq.	3 (Low)	Freq.
Time	3721	Surprise	75	Motivation	14
Thing	2613	Vision	75	Opposite	14
Way	1981	Crime	74	Prejudice	14
Number	1566	Failure	74	Scandal	14
Work	1330	Objective	74	Terror	14
System	1262	Assessment	73	Triumph	14
Problem	1138	Capacity	73	Vocabulary	14
Area	1126	Challenge	72	Wit	14
Point	1064	Joke	72	Anger	13
Part	982	Leave	72	Contradiction	13
Word	932	Paragraph	72	Correction	13
Place	911	Sequence	72	Endeavour	13
Form	806	Exception	71	Fitnesss	13
Room	791	Fault	71	Grouping	13
Case	764	Initiative	70	Identification	13
Job	742	Reaction	70	Instalment	13
Fact	715	Beauty	68	Norm	13
Question	691	Content	68	Prescription	13
Name	690	Incident	68	Readiness	13
Programme	682	Intention	68	Realisation	13
Bit	672	Pleasure	68	Rejection	13
Sort	669	Priority	68	Reluctance	13
Report	615	Enterprise	66	Scenario	13
Information	609	Joy	66	Breakthrough	12
Change	599	Tale	65	Conception	12
Idea	588	Expansion	64	Conspiracy	12
Line	574	Phrase	64	Drawback	12
Paper	572	Reply	64	Fiction	12
Business	538	Bid	63	Frustration	12
State	535	Burden	63	Legacy	12
Use	527	Emphasis	63	Oath	12
Level	522	Mixture	63	Puzzle	12
Reason	515	Struggle	63	Rationale	12
Policy	512	Belief	62	Resentment	12
Table	509	Presentation	62	Reversal	12
Development	507	Resolution	62	Similarity	12
Interest	486	Licence	61	Verdict	12
Figure	483	Declaration	60	Anomaly	11
Effect	471	Extension	60	Anticipation	11
Result	461	Impression	60	Attribute	11
Need	453	Proof	60	Coincidence	11
Course (Of Action)	448	Convention	59	Configuration	11
Rate	448	Custom	59	Correlation	11
Moment	441	Disaster	59	Curiosity	11
Issue	432	Zone	59	Deadline	11
Value	432	Intelligence	58	Deed	11
Law	429	Target	58	Gossip	11
Process	425	Teaching	58	Motive	11
Quality	424	Topic	58	Persuasion	11
Account	414	Assurance	57	Praise	11
Order	410	Explanation	57	Preoccupation	11

Period	409	Mechanism	57	Temptation	11
Product	407	Phase	57	Bitterness	10
Kind	403	Premise	57	Destiny	10
Stage	403	Classification	56	Disappointment	10
Type	402	Mess	56	Divergence	10
Action	392	Observation	56	Grievance	10
Plan	388	Contrast	55	Illusion	10
Position	387	Implication	55	Indicator	10
Power	379	Permission	55	Longing	10
Term	379	Protest	55	Segment	10
Right	377	Regard	55	Sentiment	10
Study	375	Formula	54	Setback	10
Piece	374	Instance	54	Thesis	10
View	372	Objection	54	Contingency	9
Force	366	Concentration	53	Discrepancy	9
Space	359	Connection	53	Disgrace	9
Region	357	Drama	53	Distortion	9
Control	355	Mission	53	Nuisance	9
Page	350	Resistance	53	Presumption	9
Section	344	Criticism	52	Regret	9
Subject	343	Grace	52	Sadness	9
Support	343	Indication	52	Sensation	9
Situation	342	Prize	52	Standpoint	9
Size	336	Acceptance	51	Terminology	9
Story	336	Agenda	51	Twist	9
Advance	334	Error	51	Viewpoint	9
Authority	329	Expectation	51	Whisper	9
Matter	328	Experiment	51	Compliment	8
Ground	327	Outcome	51	Dilemma	8
Site	326	Sentence	51	Hypothesis	8
Stuff	324	Complaint	50	Ideology	8
Operation	323	Conversion	49	Imagery	8
Age	322	Desire	49	Likeness	8
News	313	Fight	49	Manifestation	8
Procedure	306	Luck	49	Manoeuvre	8
Decision	305	Promise	49	Outrage	8
Benefit	301	Secret	49	Permit	8
Pattern	287	Fate	48	Projection	8
Condition	282	Meaning	48	Rage	8
Activity	279	Potential	48	Reverse	8
Opportunity	276	Preparation	48	Ruling	8
Picture	276	Symbol	48	Scepticism	8
Range	276	Announcement	47	Sphere	8
Art	275	Favourite	47	Urge	8
Difference	275	Guarantee	47	Vow	8
Event	275	Shock	47	Ambiguity	7
Field	275	Sin	47	Brief	7
Basis	272	Wish	47	Condemnation	7
Class	272	Completion	45	Consolation	7
Approach	271	Frame	45	Decree	7
Act	266	Legend	45	Deviation	7
Example	263	Phenomenon	45	Endorsement	7
Application	259	Philosophy	45	Facet	7
Project	259	Dimension	44	Foreboding	7
Detail	257	Myth	44	Impetus	7
Practice	254	Recommendation	44	Irony	7

Evidence	253	Suspicion	44	Misfortune	7
Experience	252	Characteristic	43	Proviso	7
Sense	246	Finding	43	Quest	7
Answer	243	Venture	43	Recollection	7
Chance	242	Warning	43	Testimony	7
Item	242	Achievement	42	Unknown	7
Source	240	Comfort	42	Converse	6
Rule	239	Comparison	42	Deception	6
Statement	232	Summary	42	Disgust	6
Role	231	Thinking	42	Equal	6
Scheme	227	Wonder	42	Folly	6
Growth	224	Evaluation	41	Franchise	6
Facility	220	Shame	41	Fury	6
Style	220	Calculation	40	Gamble	6
Division	219	Disspute	40	Inability	6
Method	219	Passage	40	Inconsistency	6
Step	218	Qualification	40	Inference	6
Choice	216	Query	40	Instinct	6
Model	214	Recognition	39	Narrative	6
Preference	212	Assumption	38	Novelty	6
Agreement	210	Composition	38	Outburst	6
Movement	210	Coverage	38	Paradox	6
Data	209	Confusion	37	Propensity	6
Effort	208	Habit	37	Refinement	6
Note	207	Obligation	37	Unhappiness	6
Base	205	Reading	37	Yearning	6
Motion	203	Signal	37	Accusation	5
Heart	201	Weakness	37	Bribe	5
Will	201	Assignment	36	Foresight	5
Pressure	200	Interpretation	36	Generalisation	5
Campaign	196	Lie	36	Gist	5
Trouble	196	Worry	36	Impatience	5
Charge	195	Forecast	35	Impossibility	5
Requirement	192	Innovation	35	Intuition	5
Heading	191	Mention	35	Jealousy	5
Provision	191	Passion	35	Maxim	5
Demand	187	Specification	35	Notification	5
Test	186	Speculation	35	Portrayal	5
Purpose	185	Flexibility	34	Precondition	5
Asset	183	Limitation	34	Pretence	5
Branch	183	Pact	34	Provocation	5
Success	183	Pride	34	Rationalisation	5
Responsibility	182	Reward	34	Resemblance	5
Truth	182	Cycle	33	Retort	5
Comment	181	Doctrine	33	Saying	5
Call	180	Nerve	33	Stamina	5
Duty	178	Scope	33	Stereotype	5
Commission	176	Talent	33	Acknowledgement	4
Help	176	Attraction	32	Adjunct	4
Difficulty	175	Criterion	32	Conjecture	4
Example	175	Grief	32	Denunciation	4
Match	175	Reservation	32	Dread	4
Factor	174	Restriction	32	Excerpt	4
Skill	174	Rush	32	Flaw	4
Talk	173	Ambition	31	Go-Ahead	4
Attempt	170	Cheek	31	Impulse	4

Feature	170	Command	31	Indictment	4
Deal	167	Courage	31	Obsession	4
Shape	167	Determination	31	Oddity	4
Alternative	166	Distinction	31	Proverb	4
Key	165	Examination	31	Reasoning	4
Arrangement	163	Invitation	31	Reckoning	4
Property	162	Offence	31	Stipulation	4
Solution	162	Perspective	31	Thrill	4
Survey	162	Profile	31	Uniformity	4
Chapter	161	Satisfaction	31	Vocation	4
Language	161	Uncertainty	31	Zeal	4
Variety	159	Estimate	30	Absurdity	3
Behaviour	158	Nightmare	30	Adaptation	3
Advantage	157	Precaution	30	Analogy	3
Contract	157	Significance	30	Assertion	3
Feeling	157	Tip	30	Betting	3
Technology	157	Blow	29	Blunder	3
Impact	156	Coup	29	Calamity	3
Rise	156	Illustration	29	Characterisation	3
Link	155	Nonsense	29	Compulsion	3
Task	155	Pity	29	Corollary	3
Addition	150	Plot	29	Cue	3
Concern	150	Rumour	29	Curse	3
Means	149	Stand	29	Decency	3
Danger	148	Wisdom	29	Disbelief	3
Response	148	Amendment	28	Disposition	3
Series	148	Bargain	28	Enigma	3
Consideration	147	Charm	28	Exaggeration	3
Hope	145	Excuse	28	Exposition	3
Thought	143	Notion	28	Fabrication	3
Claim	141	Petition	28	Fallacy	3
Defence	141	Revelation	28	Gall	3
Expression	141	Capability	27	Guff	3
Treatment	141	Concession	27	Inducement	3
Discussion	140	Conviction	27	Inevitability	3
Knowledge	140	Delight	27	Inkling	3
Aspect	139	Occurrence	27	Irritation	3
Goal	139	Probability	27	Knack	3
Circumstance	138	Remark	27	Motto	3
Proposal	138	Discovery	26	Ploy	3
Image	137	Necessity	26	Polemics	3
Measure	137	Proposition	26	Preface	3
Possibility	137	Awareness	25	Pretext	3
Sign	137	Bonus	25	Prophecy	3
Accident	136	Complexity	25	Protestation	3
Exchange	136	Constraint	25	Reassurance	3
Risk	136	Disclosure	25	Repudiation	3
Extent	135	Mystery	25	Superstition	3
Occasion	135	Privilege	25	Unwillingnesss	3
Reference	135	Usage	25	Upshot	3
Argument	134	Disadvantage	24	Yen	3
Theory	133	Merit	24	Affirmation	2
Association	132	Odds	24	Affront	2
Noise	132	Reflection	24	Apprehension	2
Opinion	132	Refusal	24	Astonishment	2
Affair	131	Tragedy	24	Axiom	2

Discussion	131	Blessing	23	Boast	2
Message	131	Courtesy	23	Catastrophe	2
Move	130	Deduction	23	Caveat	2
Analysis	129	Hurdle	23	Cliché	2
Attack	129	Incentive	23	Credo	2
Element	129	Parallel	23	Critique	2
Fear	129	Tendency	23	Crusade	2
Option	128	Adoption	22	Debacle	2
Suggestion	126	Consensus	22	Delusion	2
Confidence	125	Happiness	22	Denial	2
Relationship	125	Logic	22	Diagnosis	2
Strategy	125	Recipe	22	Digression	2
Technique	125	Trick	22	Entreaty	2
Advice	123	Compensation	21	Euphemism	2
Function	122	Compromise	21	Farce	2
Context	119	Constituent	21	Inclination	2
Doubt	119	Disagreement	21	Indignation	2
Offer	119	Episode	21	Lust	2
Appeal	116	Inspiration	21	Marvel	2
Contribution	116	Pledge	21	Metaphor	2
Ability	114	Stress	21	Miscalculation	2
Energy	114	Appreciation	20	Misconception	2
Definition	112	Certainty	20	Misinterpretation	2
Equation	111	Explosion	20	Payoff	2
Selection	110	Obstacle	20	Phraseology	2
Freedom	109	Routine	20	Preconception	2
Principle	109	Slogan	20	Predicament	2
Cause	108	Symptom	20	Premonition	2
Text	108	Virtue	20	Prognosiss	2
Debate	107	Anxiety	19	Recapitulation	2
Exercise	107	Apology	19	Remorse	2
Threat	107	Authorisation	19	Ruse	2
Lesson	106	Confirmation	19	Snag	2
Version	106	Encouragement	19	Surmisal/Surmise	2
Background	105	Hint	19	Tenet	2
Race	105	Ideal	19	Uniqueness	2
Strength	102	Likelihood	19	Abstraction	1
Appearance	101	Mandate	19	Activation	1
Concept	101	Prediction	19	Adage	1
Description	101	Punishment	19	Adaption	1
Progress	101	Ritual	19	Admonition	1
Shift	101	Tactic	19	Amazement	1
Prospect	100	Aspiration	18	Anachronism	1
Notice	98	Clue	18	Annoyance	1
Influence	96	Cure	18	Aphorism	1
Spot	96	Essence	18	Audacity	1
Commitment	95	Insistence	18	Blasphemy	1
Improvement	95	Quotation	18	Carte Blanche	1
Mistake	95	Remedy	18	Clamour	1
Battle	93	Reminder	18	Clut	1
Device	93	Repetition	18	Coefficient	1
Importance	93	Shortfall	18	Concoction	1
Understanding	93	Split	18	Conundrum	1
Representation	92	Stance	18	Countermeasure	1
Aim	90	Willingness	18	Crux	1
Conclusion	90	Bet	17	Dictum	1

Faith	88	Complication	17	Dogma	1
Parameter	88	Confession	17	Edict	1
Trial	88	Diversity	17	Ethos	1
Writing	88	Fantasy	17	Falsification	1
Category	87	Formation	17	Gambit	1
Instruction	87	Handicap	17	Gratitude	1
Introduction	87	Hurry	17	Grumble	1
Limit	87	Insight	17	Gumption	1
Tradition	87	Justification	17	Hunch	1
Relief	86	Misunderstanding	17	Imperative	1
Combination	85	Particular	17	Intimation	1
Drive	85	Sorrow	17	Itch	1
Angle	83	Allegation	16	Lament	1
Attitude	83	Catch	16	Misapprehension	1
Requesst	83	Contention	16	Misjudgement	1
Revolution	83	Era	16	Misreading	1
Trend	83	Guess	16	Nous	1
Undertaking	83	Miracle	16	Orthodoxy	1
Demonstration	82	Modification	16	Paranoia	1
Dream	82	Optimism	16	Pipedream	1
Manner	82	Plea	16	Pretension	1
Theme	82	Specialty	16	Proclamation	1
Judgement	81	Controversy	15	Quibble	1
Component	80	Extract	15	Ramification	1
Pain	80	Formulation	15	Resolve	1
Respect	79	Perception	15	Scramble	1
Object	78	Pussh	15	Sequel	1
Foundation	77	Queue	15	Squabble	1
Reality	76	Admission	14	Stricture	1
Appointment	75	Directive	14	Supposition	1
Consequence	75	Guilt	14	Synonymy	1
Investigation	75	Intent	14	Travesty	1
				Truism	1

APPENDIX 3 CODES USED IN THE TAGGING OF EXAMPLES

Tagging codes			
*T	Marked Theme	LT	Letter
AB	Adverbial	MAN	Manner
AC	Accompaniment	MAT	Matter
ACT	Actor	MD	Modality
AD	Additive	ML	Metalinguistic
AF	Anaphoric	MR	Measure
AG	Agent	MS	Means
AGG	Aggregate	n	Noun (after prep.)
AJ	Adjective	N	Noun
AJP	Adjective Phrase	NA	Non-Applicable
AM	Amplifier	NAS	Non-Assertive
AN	Angle	NG	Negative
AP	Appositive	NP	Noun Phrase
AR	Article	NR	Non-restrictive
AS	Assertive	NSL	Non-selective
AT	Adjunct	NSP	Non-specific
ATR	Attributor	NUM	Numeral/Number
ATT (rel.ci)	Relational Circumstantial Attribute	OB	Obligation
ATT (rel.in)	Relational Intensive Attribute	OCL	Object Complement
ATT (rel.po)	Relational Possessive Attribute	OR	Ordinative/Ordinal
AV	Adverb	OV	Overall
BF	Behalf	PA	Particularization
CAJ	Complement Adjective	PB	Probability
CAR (rel.ci)	Relational Circumstantial Carrier	PC	Place
CAR (rel.in)	Relational Intensive Carrier	PCT	Product
CAR (rel.po)	Relational Possessive Carrier	PD	Predication
CAU	Cause	PDC	Post-Deictic
CC	Concession	PDT	Pre-Determiner
CD	Cardinal	PHE	Phenomenon
CF	Cataphoric	PHE (beh)	Phenomenon (Behavioural)
CG	Contingency	PHE (met)	Phenomenon (Metaphorical)
CI	Condition	PL	Participle
CL	Clause	PM	Postmodifier
CM	Comment	PN	Pronoun
CN	Caption	PO	Prepositional Object
CO	Content	POCL	Prepositional Object Complement
CP	Prepositional Complement	POS	Positive
CR	Circumstance	PP	Prepositional Phrase
CS	Classifier	PR	Preceding
CT	Complement	PRM	Premodifier
CV	Comparative/Comparison	PS	Possessive
DC	Deictic	PSCL	Prepositional Subject Complement
DF	Definite	PT	Partial
DG	Degree	PU	Purpose
DJ	Disjunct	PV	Partitive
DM	Demonstrative	QF	Qualifier
DN	Designation	QL	Quality
DO	Direct Object	QT	Quantifier
DS	Descriptive	QU	Quote
DT	Determiner	QV	Quantitative/Quantity
DV	Determinative	R	Rheme
EC	Emphatic	RD	Readiness

ED	ED-form
EL	Elaborating
EM	Exemplification
EN	Enhancing
EP	Epithet
ET	Extended/Extending
EX	Experiential
EXI	Existent
EXO	Exophoric
EZ	Emphasizer
FC	Facet
FG	Figure
FL	Full
FR	Formula
GB	Global
GO	General Ordinal
GOA	Goal
GS	Guise
GV	Genitive
H	Head
HD	Heading/Headline
IA	Idea
ID	Identity
IDD (rel.ci)	Relational Circumstantial Identified
IDD (rel.in)	Relational Intensive Identified
IDD (rel.po)	Relational Possessive Identified
IDR (rel.ci)	Relational Circumstantial Identifier
IDR (rel.in)	Relational Intensive Identifier
IDR (rel.po)	Relational Possessive Identifier
IF	Intensifier
IJ	Interjection
IM	Image
IN	Indefinite
INI	Initiator
INTER	Intersentential
INTRA	Intrasentential
IP	Interpersonal
IT	Identification
IV	Interrogative
LC	Local
LN	Locution
LOC	Location
RE	Reason
RF	Reformulation
RL	Role
RN	Reaction
RP	Report
RS	Respect
RT	Restrictive
RU	Result
RV	Relative
SAY	Sayer
SB	Subject
SB(not)	Notional Subject
SC	Sentence
SCL	Subject Complement
SCO	Scope
SEN (met)	Metaphorical Sense
SG	Signal
SI	Specifics of Identity
SJ	Subjunct
SL	Selective
SM	Symbol/Symbolization
SN2	Shell Noun 2
SNP	Simple Noun Phrase
SP	Specific
SPA	Space
ST	Strict
SY	Style
T	Theme
TB	Table
TG	Thing
TI	To-Infinitive
TL	Total
TM	Time
TR	Target
UNC	Unclear
US	Usuality
UV	Universal
VER	Verbiage
VI	Viewpoint
VR	Verb
WK	Weak

APPENDIX 4 EXPERIENTIAL AND FORMAL PATTERNS (COMPLETE LISTS)

4.1 *Experiential patterns*

Experiential patterns (%)			
DC.SP.DM.DV^TG	10.44	DC.NSP.PT.SL^DC.SP.PS.DV^CS^TG	0.07
DC.SP.DM.DV^TG^QF	9.40	DC.NSP.PT.SL^EP.EX^CS^TG	0.07
TG^QF	6.77	DC.NSP.PT.SL^EP.EX^TG	0.07
TG	5.25	DC.NSP.PT.SL^EP.IP^TG^QF	0.07
DC.NSP.PT.NSL^TG^QF	3.87	DC.NSP.PT.SL^FC.EL.EM^TG^QF	0.07
DC.NSP.PT.NSL^TG	3.46	DC.NSP.PT.SL^FC.ET.PV^TG	0.07
CS^TG	3.11	DC.NSP.PT.SL^NUM.DF.OR^FC.ET.PV^TG	0.07
DC.SP.PS.DV^TG	3.11	DC.NSP.PT.SL^NUM.IN.QV^TG	0.07
DC.SP.PS.DV^TG^QF	3.11	DC.NSP.PT.SL^PDC.EL.EM.RT^TG	0.07
DC.NSP.PT.NSL^CS^TG	1.52	DC.NSP.PT.SL^PDC.EL.EM.RT^TG^QF	0.07
DC.SP.DM.DV^CS^TG	1.45	DC.NSP.PT.SL^PDC.EL.ID.RT^TG	0.07
NUM.IN.QV^TG	1.45	DC.NSP.PT.SL^TG[…]^QF	0.07
DC.SP.DM.DV^TG^QF^QF	1.31	DC.NSP.PT.SL^TG^QF^QF	0.07
DC.NSP.PT.NSL^EP.IP^TG	1.24	DC.NSP.TL.NG^EP.EX^CS^TG	0.07
DC.NSP.PT.NSL^CS^TG^QF	1.17	DC.NSP.TL.NG^EP.IP^TG	0.07
CS^TG^QF	1.11	DC.NSP.TL.NG^PDC.EL.EM.AM^TG	0.07
DC.NSP.PT.NSL^EP.IP^TG^QF	1.11	DC.NSP.TL.NG^PDC.EL.EM.RT^TG	0.07
NUM.IN.QV^TG^QF	1.04	DC.NSP.TL.NG^PDC.EL.ID^TG	0.07
DC.NSP.TL.NG^TG^QF	0.97	DC.NSP.TL.NG^PDC.EN.SPA-TM^TG	0.07
DC.SP.DM.DV^CS^TG^QF	0.97	DC.NSP.TL.NG^PDC.RP.LN^TG^QF	0.07
DC.SP.DM.DV^EP.IP^TG^QF	0.97	DC.NSP.TL.POS^DC.SP.DM.DV^EP.EX^TG^QF	0.07
DC.SP.PS.DV^CS^TG	0.90	DC.NSP.TL.POS^DC.SP.DM.DV^EP.IP^TG^QF	0.07
DC.NSP.TL.NG^TG	0.83	DC.NSP.TL.POS^DC.SP.DM.DV^PDC.EN.SPA-TM^NUM.DF.QV^TG	0.07
NUM.DF.QV^TG	0.76	DC.NSP.TL.POS^DC.SP.PS.DV^EP.IP^TG	0.07
DC.SP.DM.DV^EP.IP^TG	0.69	DC.NSP.TL.POS^FC.EL.EM^PDC.EL.ID^TG	0.07
PDC.EL.ID^TG	0.69	DC.NSP.TL.POS^PDC.EL.ID^TG^QF	0.07
DC.NSP.PT.NSL^EP.EX^TG	0.62	DC.SP.DM.DV^[…]TG^[…]QF	0.07
DC.SP.DM.DV^PDC.EL.ID.RT^TG^QF	0.62	DC.SP.DM.DV^CS^[…]EP.EX^TG^QF	0.07
DC.SP.DM.DV^PDC.EN.SPA-TM^TG	0.62	DC.SP.DM.DV^CS^TG^QF^QF	0.07
EP.EX^TG	0.62	DC.SP.DM.DV^DC.NSP.PT.NSL^TG	0.07
DC.SP.DM.DV^EP.EX^TG	0.55	DC.SP.DM.DV^DC.NSP.PT.SL^PDC.EL.ID^TG^QF	0.07
NUM.DF.QV^TG^QF	0.55	DC.SP.DM.DV^EP.EX^[…]CS^CS^TG	0.07
DC.NSP.PT.NSL^EP.EX^TG^QF	0.48	DC.SP.DM.DV^EP.IP^[…]DC.SP.PS.DV^TG	0.07
DC.NSP.PT.SL^TG	0.48	DC.SP.DM.DV^EP.IP^CS^TG	0.07
DC.NSP.PT.SL^TG^QF	0.48	DC.SP.DM.DV^EP.IP^EP.EX^TG	0.07
DC.SP.DM.DV^EP.EX^TG^QF	0.48	DC.SP.DM.DV^FC.EL.EM^CS^[…]^TG^QF	0.07
EP.IP^TG	0.48	DC.SP.DM.DV^FC.EL.EM^CS^TG^QF	0.07
EP.IP^TG^QF	0.48	DC.SP.DM.DV^FC.EL.EM^EP.EX^TG	0.07
DC.NSP.PT.SL^DC.SP.DM.DV^TG^QF	0.41	DC.SP.DM.DV^FC.EL.EM^EP.IP^TG^QF	0.07
DC.SP.DM.DV^PDC.EL.ID.RT^TG	0.41	DC.SP.DM.DV^FC.EL.EM^TG	0.07
PDC.EL.ID^TG^QF	0.41	DC.SP.DM.DV^NUM.DF.OR^CS^TG^QF	0.07
TG^QF^QF	0.41	DC.SP.DM.DV^NUM.DF.OR^DC.SP.DM.DV^TG	0.07
DC.NSP.TL.POS^TG	0.35	DC.SP.DM.DV^NUM.DF.OR^PDC.MD.PB^EP.EX^TG^QF	0.07

DC.SP.DM.DV^[…]TG^QF	0.35	DC.SP.DM.DV^NUM.DF.OR^TG	0.07
DC.SP.DM.DV^FC.EL.EM^TG^QF	0.35	DC.SP.DM.DV^NUM.DF.OR^TG^QF^QF	0.07
DC.SP.DM.DV^NUM.DF.OR^TG^QF	0.35	DC.SP.DM.DV^NUM.DF.QV^CS^CS^TG^QF	0.07
DC.SP.PS.DV^EP.IP^TG	0.35	DC.SP.DM.DV^NUM.DF.QV^CS^TG^QF	0.07
EP.EX^TG^QF	0.35	DC.SP.DM.DV^NUM.DF.QV^PDC.EL.EM^TG^QF	0.07
PDC.EN.CV^TG	0.35	DC.SP.DM.DV^NUM.DF.QV^PDC.EL.ID.RT^TG^QF	0.07
CS^CS^TG	0.28	DC.SP.DM.DV^NUM.DF.QV^TG	0.07
DC.NSP.PT.NSL^EP.EX^CS^TG	0.28	DC.SP.DM.DV^PDC.EL.EM.EZ^TG^QF	0.07
DC.NSP.PT.SL^DC.SP.DM.DV^TG	0.28	DC.SP.DM.DV^PDC.EL.EM^TG^QF	0.07
DC.NSP.TL.POS^DC.SP.DM.DV^TG	0.28	DC.SP.DM.DV^PDC.EL.ID^EP.IP^TG^QF	0.07
DC.NSP.TL.POS^TG^QF	0.28	DC.SP.DM.DV^PDC.EL.ID^TG	0.07
DC.SP.PS.DV^PDC.EL.ID.RT^TG	0.28	DC.SP.DM.DV^PDC.EL.ID^TG^QF	0.07
NUM.IN.QV^TG[…]^QF	0.28	DC.SP.DM.DV^PDC.EN.SPA-TM^CS^TG^QF	0.07
PDC.ET.AM^TG	0.28	DC.SP.DM.DV^PDC.EN.SPA-TM^PDC.EL.ID^[…]TG^QF^[…]^QF	0.07
TG^[…]^QF	0.28	DC.SP.DM.DV^PDC.EN.SPA-TM^TG^QF	0.07
DC.NSP.PT.NSL^CS^CS^TG	0.21	DC.SP.DM.DV^PDC.ET.AM^TG	0.07
DC.NSP.PT.NSL^TG[…]^QF	0.21	DC.SP.DM.DV^PDC.ET.EZ^TG^QF	0.07
DC.NSP.PT.NSL^TG^QF^QF	0.21	DC.SP.DM.DV^PDC.ET.RT^TG	0.07
DC.NSP.PT.SL^DC.SP.DM.DV^EP.IP^TG^QF	0.21	DC.SP.DM.DV^PDC.MD.OB^TG	0.07
DC.NSP.TL.NG^CS^TG	0.21	DC.SP.DM.DV^PDC.MD.US^CS^TG^QF	0.07
DC.NSP.TL.POS^DC.SP.DM.DV^TG^QF	0.21	DC.SP.DM.DV^PDC.MD.US^TG^QF	0.07
DC.SP.DM.DV^PDC.EL.EM.RT^TG	0.21	DC.SP.DM.DV^PDC.RP.IA.EZ^CS^TG^QF	0.07
DC.SP.PS.DV^NUM.DF.OR^TG	0.21	DC.SP.DM.DV^PDC.RP.IA^CS^TG	0.07
DC.SP.PS.DV^NUM.DF.OR^TG^QF	0.21	DC.SP.DM.DV^PDC.RP.IA^PDC.MD.US^TG^QF	0.07
DC.SP.PS.DV^TG[…]^QF	0.21	DC.SP.DM.DV^PDC.RP.LN.EZ^CS^TG^QF	0.07
NUM.IN.QV^DC.SP.DM.DV^TG^QF	0.21	DC.SP.DM.DV^UNC^TG^[…]^QF	0.07
PDC.EL.EM.RT^TG	0.21	DC.SP.DM.IV^[…]TG	0.07
DC.NSP.PT.NSL^EP.EX^CS^TG^QF	0.14	DC.SP.DM.IV^FC.EL.EM^TG^[…]^QF	0.07
DC.NSP.PT.NSL^EP.EX^TG[…]^QF	0.14	DC.SP.DM.IV^TG^QF	0.07
DC.NSP.PT.NSL^EP.IP^TG^QF^QF	0.14	DC.SP.PS.DV^[…]^EP.IP^EP.IP^TG	0.07
DC.NSP.PT.NSL^FC.EL.EM^TG	0.14	DC.SP.PS.DV^[…]TG^QF	0.07
DC.NSP.PT.NSL^PDC.EL.EM.RT^TG^QF	0.14	DC.SP.PS.DV^CS^[…]TG	0.07
DC.NSP.PT.NSL^PDC.EL.ID^TG	0.14	DC.SP.PS.DV^CS^CS^CS^TG	0.07
DC.NSP.PT.NSL^PDC.MD.US^TG^QF	0.14	DC.SP.PS.DV^CS^CS^TG	0.07
DC.NSP.PT.SL^CS^TG	0.14	DC.SP.PS.DV^CS^CS^TG^QF	0.07
DC.NSP.PT.SL^DC.SP.PS.DV^TG^QF	0.14	DC.SP.PS.DV^EP.EX^TG	0.07
DC.NSP.PT.SL^PDC.EL.ID^TG^QF	0.14	DC.SP.PS.DV^EP.EX^TG^QF	0.07
DC.NSP.TL.NG^EP.EX^TG	0.14	DC.SP.PS.DV^EP.IP^TG^[…]QF	0.07
DC.NSP.TL.NG^PDC.EL.ID^TG^QF	0.14	DC.SP.PS.DV^FC.EL.EM^CS^TG	0.07
DC.NSP.TL.POS^FC.EL.EM^TG	0.14	DC.SP.PS.DV^NUM.DF.OR^CS^TG	0.07
DC.SP.DM.DV^CS^CS^TG	0.14	DC.SP.PS.DV^NUM.IN.QV^TG	0.07
DC.SP.DM.DV^EP.EX^CS^TG	0.14	DC.SP.PS.DV^PDC.EL.ID.EZ^TG	0.07
DC.SP.DM.DV^EP.EX^CS^TG^QF	0.14	DC.SP.PS.DV^PDC.EL.ID^TG	0.07
DC.SP.DM.DV^FC.EL.SM^TG	0.14	DC.SP.PS.DV^PDC.EN.SPA-TM^TG	0.07
DC.SP.DM.DV^PDC.EL.EM.EZ^TG	0.14	DC.SP.PS.DV^PDC.ET.AM^TG^QF	0.07
DC.SP.DM.DV^PDC.EL.EM.RT^TG^QF	0.14	DC.SP.PS.DV^PDC.MD.RD^TG	0.07
DC.SP.DM.DV^PDC.EL.ID.RT^TG^QF^QF	0.14	DC.SP.PS.IV^PDC.EL.ID.RT^TG	0.07
DC.SP.DM.DV^PDC.ET.EZ^TG	0.14	EP.EX^[…]TG	0.07
DC.SP.DM.DV^PDC.MD.PB^TG	0.14	EP.EX^CS^CS^TG	0.07

DC.SP.DM.DV^TG[…]^QF	0.14	EP.EX^CS^TG^[…]QF	0.07
DC.SP.DM.IV^TG	0.14	EP.EX^CS^TG^QF	0.07
DC.SP.PS.DV^[…]^TG	0.14	EP.EX^CS^TG^QF^QF	0.07
DC.SP.PS.DV^CS^TG^QF	0.14	EP.EX^EP.EX^TG	0.07
DC.SP.PS.DV^EP.IP^TG^QF	0.14	EP.EX^TG^[…]QF	0.07
DC.SP.PS.DV^NUM.DF.QV^TG	0.14	EP.IP^DC.NSP.PT.NSL^TG	0.07
DC.SP.PS.IV^TG	0.14	EP.IP^EP.EX^TG	0.07
DC.SP.PS.IV^TG^QF	0.14	EP.IP^EP.IP^TG	0.07
EP.EX^TG^QF^QF	0.14	EP.IP^TG^[…]^QF	0.07
NUM.DF.QV^TG^QF^QF	0.14	FC.ET.AGG^TG	0.07
NUM.IN.QV^CS^TG	0.14	FC.ET.AGG^TG^QF	0.07
NUM.IN.QV^DC.NSP.PT.NSL^TG	0.14	FC.ET.PV^DC.NSP.TL.POS^DC.SP.PS.DV^TG	0.07
NUM.IN.QV^DC.SP.DM.DV^TG	0.14	FC.ET.PV^DC.SP.DM.DV^TG	0.07
NUM.IN.QV^PDC.EL.ID^TG^QF	0.14	IF^DC.NSP.PT.NSL^EP.EX^TG	0.07
NUM.IN.QV^TG^QF^QF	0.14	IF^DC.NSP.PT.NSL^TG	0.07
PDC.EL.EM.RT^TG^QF	0.14	IF^NUM.IN.QV^TG^QF	0.07
CS^CS^CS^TG	0.07	NUM.DF.QV^CS^TG	0.07
CS^CS^TG^QF	0.07	NUM.DF.QV^EP.EX^TG	0.07
CS^TG^QF^QF	0.07	NUM.DF.QV^EP.IP^TG	0.07
DC.NSP.PT.NSL^[…]^TG	0.07	NUM.DF.QV^EP.IP^TG^QF	0.07
DC.NSP.PT.NSL^[…]CS^TG	0.07	NUM.DF.QV^NUM.DF.QV^CS^TG	0.07
DC.NSP.PT.NSL^[…]TG^QF	0.07	NUM.DF.QV^NUM.DF.QV^PDC.EL.ID^TG	0.07
DC.NSP.PT.NSL^CS^[…]^EP.IP^TG	0.07	NUM.DF.QV^NUM.DF.QV^TG^QF^QF	0.07
DC.NSP.PT.NSL^CS^[…]TG^QF	0.07	NUM.DF.QV^PDC.EL.ID.RT^TG	0.07
DC.NSP.PT.NSL^CS^TG^QF^QF	0.07	NUM.DF.QV^PDC.EL.ID^TG	0.07
DC.NSP.PT.NSL^EP.EX^EP.EX^TG	0.07	NUM.DF.QV^PDC.EL.ID^TG^QF	0.07
DC.NSP.PT.NSL^EP.IP^CS^TG^QF	0.07	NUM.DF.QV^PDC.MD.PB^TG^QF	0.07
DC.NSP.PT.NSL^EP.IP^EP.IP^TG^QF	0.07	NUM.DF.QV^TG^[…]^QF	0.07
DC.NSP.PT.NSL^NUM.IN.QV^TG^QF	0.07	NUM.IN.QV.IV^TG	0.07
DC.NSP.PT.NSL^PDC.EL.EM.AM^TG	0.07	NUM.IN.QV^CS^TG^QF	0.07
DC.NSP.PT.NSL^PDC.EL.EM.EZ^TG	0.07	NUM.IN.QV^DC.SP.PS.DV^TG^QF	0.07
DC.NSP.PT.NSL^PDC.EL.EM.RT^TG	0.07	NUM.IN.QV^EP.EX^TG	0.07
DC.NSP.PT.NSL^PDC.EL.EM^TG	0.07	NUM.IN.QV^EP.EX^TG^QF^QF	0.07
DC.NSP.PT.NSL^PDC.EL.ID.EZ^TG^QF	0.07	NUM.IN.QV^EP.IP^TG	0.07
DC.NSP.PT.NSL^PDC.EL.ID.RT^TG	0.07	NUM.IN.QV^EP.IP^TG^QF	0.07
DC.NSP.PT.NSL^PDC.EL.ID^TG^QF	0.07	PDC.EL.EM.RT^[…]TG	0.07
DC.NSP.PT.NSL^PDC.EN.CV^TG^QF^QF	0.07	PDC.EL.EM^PDC.EL.ID^TG	0.07
DC.NSP.PT.NSL^PDC.EN.SPA-TM^TG^QF	0.07	PDC.EL.EM^TG	0.07
DC.NSP.PT.NSL^PDC.ET.AM^TG^QF	0.07	PDC.EL.EM^TG^QF	0.07
DC.NSP.PT.NSL^PDC.MD.PB^TG	0.07	PDC.EL.ID.RT^TG	0.07
DC.NSP.PT.NSL^PDC.MD.PB^TG^QF	0.07	PDC.EL.ID^CS^TG^QF	0.07
DC.NSP.PT.NSL^PDC.MD.US.EZ^TG	0.07	PDC.EL.ID^EP.EX^CS^TG	0.07
DC.NSP.PT.NSL^PDC.RP.IA.EZ^TG^QF	0.07	PDC.EL.ID^EP.IP^TG	0.07
DC.NSP.PT.NSL^PDC.RP.IA^TG	0.07	PDC.EL.ID^TG^QF^QF	0.07
DC.NSP.PT.SL^DC.SP.DM.DV^[…]EP.EX^CS^TG	0.07	PDC.EN.CV^PDC.EN.CV^TG	0.07
DC.NSP.PT.SL^DC.SP.DM.DV^CS^TG	0.07	PDC.EN.SPA-TM^CS^TG	0.07
DC.NSP.PT.SL^DC.SP.DM.DV^EP.EX^TG^QF	0.07	PDC.EN.SPA-TM^TG^QF	0.07
DC.NSP.PT.SL^DC.SP.DM.DV^EP.IP^TG	0.07	PDC.ET.AM^TG^QF	0.07
DC.NSP.PT.SL^DC.SP.DM.DV^PDC.EL.EM^TG^QF	0.07	PDC.ET.EZ^EP.IP^TG	0.07
DC.NSP.PT.SL^DC.SP.DM.DV^PDC.EL.ID^TG^QF	0.07	PDC.MD.US^TG	0.07
DC.NSP.PT.SL^DC.SP.DM.DV^	0.07	PDC.RP.IA.EZ^TG	0.07

PDC.EN.SPA-TM^CS^TG^QF			
DC.NSP.PT.SL^DC.SP.DM.DV^PDC.MD.PB.EZ^TG^QF	0.07	TG^QF^[…]^QF	0.07
DC.NSP.PT.SL^DC.SP.PS.DV^[…]TG	0.07		

4.2 *Formal patterns (detailed)*

Formal patterns (%)			
DF.AR^H	6.36	DF.AR^UNC^H^[…]^RT.RV.CL	0.07
H	5.25	DF.DV.GV.NP^[…]H	0.07
DF.AR^H^PP(*of* n)	3.80	DF.DV.GV.NP^AJ^[…]H	0.07
AJ^H	3.59	DF.DV.GV.NP^AJ^AJ^H	0.07
IN.AR^H	3.04	DF.DV.GV.NP^H^[…]AP.TI.CL	0.07
H^PP(*of* n)	2.63	DF.DV.GV.NP^H^PP(*of* -ing)	0.07
DF.AR^AJ^H	2.42	DF.DV.GV.NP^H^PP(*to* n)	0.07
IN.AR^AJ^H	2.42	DF.DV.GV.NP^NUM.GO^H	0.07
PS.DT^H	2.28	DM.DT(*THAT*)^AJ^H	0.07
DM.DT(*THIS*)^H	1.59	DM.DT(*THAT*)^H^AP.TI.CL	0.07
DF.AR^H^RT.RV.CL	1.45	DM.DT(*THAT*)^H^PP(*of* n)^FL.ST.NR.IT.AP	0.07
DF.AR^AJ^H^PP(*of* n)	1.24	DM.DT(*THAT*)^QL.PV^H	0.07
DM.DT(*THESE*)^H	1.17	DM.DT(*THESE*)^AJ^H	0.07
QT^H	1.17	DM.DT(*THESE*)^QL.PV^AJ^H	0.07
IN.AR^H^PP(*of* n)	1.11	DM.DT(*THIS*)^(ING)N^H	0.07
N^H	0.97	DM.DT(*THIS*)^AJ^AJ^H	0.07
DM.DT(*THAT*)^H	0.90	DM.DT(*THIS*)^AJ^AJP^H	0.07
NUM.CD^H	0.90	DM.DT(*THIS*)^AJ^H^PP(*against* n)	0.07
PS.DT^AJ^H	0.90	DM.DT(*THIS*)^H^PP(*in* n)	0.07
DF.DV.GV.NP^H	0.83	DM.DT(*THIS*)^H^PP(*of* n)^PP(*over* n)	0.07
AS.DT^H	0.69	DM.DT(*THIS*)^N^H	0.07
NUM.GO^H	0.62	DM.DT(*THIS*)^NUM.GO^H	0.07
DF.AR^H^PP(*of* -ing)	0.55	DM.DT(*THIS*)^QL.PV^AJ^H^RT.RV.CL	0.07
DF.DV.GV.NP^H^PP(*of* n)	0.55	DM.DT(*THOSE*)^AJ^H	0.07
H^RT.RV.CL	0.55	DM.DT(*THOSE*)^H	0.07
IN.AR^N^H	0.55	DM.DT(*THOSE*)^H^PP(*of* n)^PL.ING.CL	0.07
AJ^AJ^H	0.48	H^(*for* n AP.TI.CL)	0.07
DF.AR^N^H	0.48	H^[…]^PP(*of* n)	0.07
H^PP(*for* n)	0.48	H^[…]AP.TI.CL	0.07
IN.AR^AJ^H^PP(*of* n)	0.48	H^[…]PP(*in* n)	0.07
IN.AR^H^AP.TI.CL	0.48	H^[…]PP(*of* -ing)	0.07
NAS.DT^H	0.48	H^AP.*THAT*.CL	0.07
NG.DT^AJ^H	0.48	H^CV.AJP	0.07
DF.AR^H^PL.ED.CL	0.41	H^FL.ST.RT.EM.AP	0.07
DF.DV.GV.NP^AJ^H	0.41	H^PP(*about* n -ing)	0.07
H^AP.TI.CL	0.41	H^PP(*about* n)	0.07
H^TI.CL	0.41	H^PP(*against* n)	0.07
IN.AR^AJP^H	0.41	H^PP(*among* n)^AP.*THAT*.CL	0.07
NG.DT^H	0.41	H^PP(*at* n)	0.07
NP^H	0.41	H^PP(*at* n)^PP(*in* n)	0.07
PS.DT^H^PP(*for* n)	0.41	H^PP(*by* n)	0.07
AJ^H^PP(*in* n)	0.35	H^PP(*for* n)^PL.ED.CL	0.07
DF.AR^AJ^H^RT.RV.CL	0.35	H^PP(*for* wh)	0.07
DF.AR^H^AP.TI.CL	0.35	H^PP(*of* n -ing)	0.07
DF.AR^H^PP(*of* n)^AP.TI.CL	0.35	H^PP(*of* n)^[…]^PT.WK.NR.IT.AP	0.07
DF.AR^NP^H	0.35	H^PP(*of* n)^AP.TI.CL	0.07
DM.DT(*THIS*)^AJ^H	0.35	H^PP(*of* n)^PP(*of* n)	0.07

H^PP(*to* n)	0.35	H^PP(*of* n)^RT.RV.CL	0.07
IN.AR^AJ^H^PP(*to* n)	0.35	H^PP(*on* n)	0.07
PDT^H	0.35	H^PP(*other than* n)	0.07
PS.DT^H^AP.TI.CL	0.35	IF.AV^IN.AR^AJ^H	0.07
AJ^H^PP(*of* n)	0.28	IF.AV^IN.AR^H	0.07
AS.DT^AJ^H	0.28	IF.AV^QT^H^PP(*of* n)	0.07
DF.AR^AJ^H^PP(*of* -ing)	0.28	IN.AR^(ING)N^H	0.07
DF.AR^H^AP.*THAT*.CL	0.28	IN.AR^(ING)N^H^AP.TI.CL	0.07
DF.AR^H^PP(*to* n)	0.28	IN.AR^(ING)N^H^PP(*at* n)	0.07
DF.AR^QL.PV^H^RT.RV.CL	0.28	IN.AR^[…]^H	0.07
DF.DV.GV.NP^H^AP.TI.CL	0.28	IN.AR^[…]AJ^H	0.07
H^RT.RV.CL(*when*)	0.28	IN.AR^[…]H^PP(*of* n)	0.07
IN.AR^H^RT.RV.CL	0.28	IN.AR^AJ^AJ^H^AJP	0.07
NUM.CD^AJ^H	0.28	IN.AR^AJ^H[…]^AP.*THAT*.CL	0.07
PS.DT^H^PP(*of* n)	0.28	IN.AR^AJ^H^[…]PP(*on* n)	0.07
PS.DT^N^H	0.28	IN.AR^AJ^H^FL.ST.NR.IT.AP^PL.ED.CL	0.07
QT^AJ^H	0.28	IN.AR^AJ^H^NR.AP.*THAT*.CL	0.07
QT^H^PP(*of* -ing)	0.28	IN.AR^AJ^H^NR.PL.ED.CL	0.07
UV.DT^H	0.28	IN.AR^AJ^H^NR.PL.ING.CL	0.07
AJ^H^PP(*about* n)	0.21	IN.AR^AJ^H^NR.RV.CL	0.07
AJP^H	0.21	IN.AR^AJ^H^PP(*for* -ing)	0.07
CV.QT^H	0.21	IN.AR^AJ^H^PP(*of* ?)	0.07
DF.AR^AJ^AJ^H	0.21	IN.AR^AJ^H^PP(*of* n)^PP(*in* which)	0.07
DF.AR^AJ^AJ^H^PP(*of* n)	0.21	IN.AR^AJ^H^PP(*over* n)	0.07
DF.AR^AJ^H^PP(*for* n)	0.21	IN.AR^AJ^H^PP(*with* n)	0.07
DF.AR^AJP^H	0.21	IN.AR^AJ^H^RT.RV.CL	0.07
DF.AR^H^PP(*for* n)	0.21	IN.AR^AJ^N^H^PP(*on* n)	0.07
DF.AR^H^PP(*of* n -ing)	0.21	IN.AR^AJP^AJ^H	0.07
DF.AR^NUM.OR^H^RT.RV.CL	0.21	IN.AR^AJP^H^NR.PL.ED.CL^AP.*THAT*.CL	0.07
DF.DV.GV.NP^N^H	0.21	IN.AR^AJP^H^NR.RV.CL	0.07
DM.DT(*THAT*)^H^PP(*of* n)	0.21	IN.AR^AJP^H^PP(*by* n)^PP(*of* n)	0.07
H^PP(*in* n)	0.21	IN.AR^AJP^H^PP(*of* n)	0.07
IN.AR^AJ^H^AP.*THAT*.CL	0.21	IN.AR^AJP^N^H	0.07
IN.AR^AJ^H^AP.TI.CL	0.21	IN.AR^CV.AJ^AJ^H^AP.TI.CL	0.07
IN.AR^AJ^H^PL.ED.CL	0.21	IN.AR^CV.AJ^H^PP(*in* which)	0.07
IN.AR^AJ^H^PP(*for* n)	0.21	IN.AR^CV.AJ^H^PP(*of* n)	0.07
IN.AR^H^AP.*THAT*.CL	0.21	IN.AR^CV.AJP^N^H	0.07
IN.AR^H^PL.ED.CL	0.21	IN.AR^H^[…]PL.ING.CL	0.07
IN.AR^H^PP(*for* n)	0.21	IN.AR^H^[…]PP(*about* n)	0.07
IN.AR^H^PP(*of* -ing)	0.21	IN.AR^H^[…]PP(*of* n)	0.07
IN.AR^H^RT.RV.CL(*where*)	0.21	IN.AR^H^PP(*like* n)	0.07
IN.AR^NP^H	0.21	IN.AR^H^PP(*of* -ing)^RT.RV.CL	0.07
NAS.DT^AJ^H	0.21	IN.AR^H^PP(*of* n)^NR.RV.CL	0.07
NUM.CD^PV.PP(*of* DF.AR^AJ^H^RT.RV.CL)	0.21	IN.AR^H^PP(*of* n)^RT.RV.CL	0.07
NUM.CD^PV.PP(*of* DF.AR^H^RT.RV.CL)	0.21	IN.AR^H^PP(*of* wh)	0.07
(ING)N^H	0.14	IN.AR^H^PP(*to* n)	0.07
AJ^[…]H	0.14	IN.AR^H^PT.ST.NR.IT.AP	0.07
AJ^H^PL.ED.CL	0.14	IN.AR^H^TI.CL	0.07
AJ^H^RT.RV.CL	0.14	IN.AR^N^[…]H^PP(*of* n)	0.07
AS.DT^QL.PV^H	0.14	IN.AR^N^AJ^H	0.07
DF.AR^[…]H^PP(*of* n)	0.14	IN.AR^N^H^PP(*against* n)	0.07
DF.AR^AJ^H^PP(*about* n)	0.14	IN.AR^N^H^PP(*of* which)	0.07
DF.AR^CV.AJ^H^PP(*of* n)	0.14	IN.AR^NP^[…]^AJP^H	0.07
DF.AR^H^AV	0.14	IN.AR^NP^H^RT.RV.CL(*whereby*)	0.07
DF.AR^H^FL.ST.RT.AP	0.14	IN.AR^NUM.GO^H	0.07

DF.AR^H^PP(*of* n)^AP.*THAT*.CL	0.14	IV.DT^[…]H	0.07
DF.AR^H^PP(*of* n)^PP(*in* n)	0.14	IV.DT^H^RT.TV.CL	0.07
DF.AR^H^PP(*on* n)	0.14	IV.DT^QL.PV^H^[…]^PP(*about* n)	0.07
DF.AR^H^RT.RV.CL^PP(*of* n)	0.14	MR.GV^H^PP(*with* n)	0.07
DF.AR^NUM.CD^AJ^H^PP(*of* n)	0.14	N^H^AJP	0.07
DF.AR^NUM.CD^H	0.14	N^H^FL.ST.NR.RF.AP	0.07
DF.AR^NUM.GO^H	0.14	N^H^PT.ST.NR.EM.AP^RT.RV.CL	0.07
DF.DV.GV.NP^H^PP(*for* n)	0.14	N^N^H	0.07
DF.DV.GV.NP^NUM.CD^H	0.14	NAS.DT^AJ^H^PP(*of* n -ing)	0.07
DM.DT(*THIS*)^H^PP(*of* n)	0.14	NAS.DT^AJ^N^H	0.07
DM.DT(*THIS*)^QL.PV^H	0.14	NAS.DT^CV.QT^H	0.07
H^PL.ED.CL	0.14	NAS.DT^H[…]^RT.RV.CL	0.07
H^PP(*from* n)	0.14	NAS.DT^H^AP.*THAT*.CL	0.07
H^PP(*of* -ing)	0.14	NAS.DT^H^PP(*as to* n)	0.07
H^PP(*with* n)	0.14	NAS.DT^H^PP(*between* n)	0.07
IN.AR^AJ^AJ^H	0.14	NAS.DT^H^PP(*in* -ing)	0.07
IN.AR^AJ^H^PP(*between* n)	0.14	NAS.DT^H^PP(*of* n)	0.07
IN.AR^AJ^H^PP(*for* which)	0.14	NAS.DT^H^PP(*of* n)^PP(*at* n)	0.07
IN.AR^AJ^H^PP(*of* n -ing)	0.14	NAS.DT^H^RT.RV.CL(*when*)	0.07
IN.AR^AJ^H^PT.WK.NR.IT.AP	0.14	NAS.DT^NUM.GO^H^PP(*of* n)	0.07
IN.AR^AJ^N^H	0.14	NAS.DT^QL.PV^H^RT.RV.CL	0.07
IN.AR^H^PP(*about* n)	0.14	NAS.PN^PV.PP(*of* DF.AR^H^PL.ED.CL)	0.07
IN.AR^H^PP(*with* n)	0.14	NAS.PN^PV.PP(*of* DF.AR^H^PP *in* n)	0.07
IN.AR^H^PT.WK.NR.IT.AP	0.14	NAS.PN^PV.PP(*of* DM.DT(*THESE*)^H)	0.07
IV.DT^H	0.14	NG.DT^AJ^AJ^H	0.07
NG.DT^H^PP(*of* n)	0.14	NG.DT^AJ^H^RT.RV.CL	0.07
NUM.CD^H^PP(*of* -ing)	0.14	NG.DT^H^AP.*THAT*.CL	0.07
NUM.CD^H^PP(*of* n)	0.14	NG.DT^H^AP.TI.CL	0.07
NUM.CD^PV.PP(*of* DF.AR^H)	0.14	NG.DT^H^PP(*for* n)	0.07
NUM.CD^PV.PP(*of* NUM.CD^AJ^H)	0.14	NG.DT^H^PP(*for* which)	0.07
NUM.GO^H^PP(*of* -ing)	0.14	NG.DT^H^PP(*of* -ing)	0.07
NUM.GO^H^PP(*of* n)	0.14	NG.DT^H^PP(*of* n -ing)	0.07
PDT^H^RT.RV.CL(*as*)	0.14	NG.DT^H^PP(*on* n)	0.07
PS.DT^AJ^H^AP.TI.CL	0.14	NG.DT^H^PP(*with* n)	0.07
PS.DT^AJ^H^PP(*in* n)	0.14	NG.DT^H^TI.CL	0.07
PS.DT^H^PP(*as* n)	0.14	NG.DT^N^H	0.07
PS.DT^H^PP(*in* -ing)	0.14	NG.DT^NUM.GO^H	0.07
PS.DT^H^PP(*of* -ing)	0.14	NG.DT^NUM.GO^H^TI.CL	0.07
QT^H^RT.RV.CL	0.14	NP^H^RT.RV.CL	0.07
QT^H^TI.CL	0.14	NP^NP^H^PP(*as to* n)	0.07
QT^IN.AR^H	0.14	NUM.CD(QT)^H	0.07
RV.DT^H	0.14	NUM.CD(QT)^H^[…]PP(*about* n)	0.07
UV.PDT^DF.AR^H	0.14	NUM.CD(QT)^NUM.GO^H^PP(*on* n)	0.07
(ING)AJ^H	0.07	NUM.CD^AJ^H^PT.ST.NR.IT.AP	0.07
(ING)N^(ING)N^H	0.07	NUM.CD^AJ^H^RT.RV.CL	0.07
(ING)N^H^PP(*for* n)	0.07	NUM.CD^AJP^H	0.07
AJ^AJ^AJ^H	0.07	NUM.CD^AJP^H^PP(*in terms of* n)	0.07
AJ^AJP^H	0.07	NUM.CD^H^[…]^PP(*of* n)	0.07
AJ^H^[…]^AP.*THAT*.CL	0.07	NUM.CD^H^CV.AJP	0.07
AJ^H^[…]PP(*of* n)	0.07	NUM.CD^H^FL.ST.NR.IT.AP	0.07
AJ^H^AP.TI.CL	0.07	NUM.CD^H^PP(*of* n)^NR.RV.CL	0.07
AJ^H^FL.ST.RT.EM.AP	0.07	NUM.CD^H^PP(*of* n)^RT.RV.CL	0.07
AJ^H^NR.RV.CL(*where*)	0.07	NUM.CD^H^PP(*of* wh)	0.07
AJ^H^PP(*against* n)	0.07	NUM.CD^H^PT.WK.NR.IT.AP	0.07
AJ^H^PP(*for* n)	0.07	NUM.CD^N^H	0.07
AJ^H^PP(*from* n)^AP.*THAT*.CL	0.07	NUM.CD^NUM.GO^H^PL.ED.CL	0.07

AJ^H^PP(*in* n)^PT.WK.NR.PA.AP	0.07	NUM.CD^NUM.GO^H^PP(*of* n)	0.07
AJ^H^PP(*in* which)	0.07	NUM.CD^NUM.GO^H^RT.RV.CL(*where*)	0.07
AJ^H^PP(*of* n -ing)	0.07	NUM.CD^NUM.GO^QV.PV^H	0.07
AJ^H^PP(*on* n)	0.07	NUM.CD^PV.PP(*of* DF.AR^(ING)N^H)	0.07
AJ^H^PP(*to* n)	0.07	NUM.CD^PV.PP(*of* DF.AR^AJ^AJ^H^TI.CL)	0.07
AJ^H^PP(*with* n -ing)	0.07	NUM.CD^PV.PP(*of* DF.AR^AJ^H)	0.07
AJ^H^PP(*with* n)	0.07	NUM.CD^PV.PP(*of* DF.AR^AJ^H^NR.PL.ING.CL)	0.07
AJ^H^PT.ST.NR.EM.AP	0.07	NUM.CD^PV.PP(*of* DF.AR^AJ^H^ PP(*about* n))	0.07
AJ^H^PT.WK.NR.EM.AP	0.07	NUM.CD^PV.PP(*of* DF.AR^NUM.GO^H^RT.RV.CL)	0.07
AJ^N^AJ^H	0.07	NUM.CD^PV.PP(*of* DM.DT(*THOSE*)^AJ^H^RT.RV.CL)	0.07
AJ^N^H	0.07	NUM.CD^PV.PP(*of* NUM.CD^H^AV^PT.WK.NR.IT.AP)	0.07
AJ^N^H^[…]RT.RV.CL	0.07	NUM.CD^PV.PP(*of* PS.DT^[…]H)	0.07
AJ^N^H^PT.ST.NR.EM.AP^NR.RV.CL	0.07	NUM.CD^PV.PP(*of* PS.DT^H^ PP(*behind* -ing))	0.07
AJ^N^H^RT.RV.CL	0.07	NUM.CD^PV.PP(*of* PS.DT^H^PP(*since* n))	0.07
AJP^IN.AR^H	0.07	NUM.CD^PV.PP(*of* PS.DT^N^H)	0.07
AS.DT^AJP^AJP^H^PP(*in* n)	0.07	NUM.CD^QV.PV^H	0.07
AS.DT^CV.QT^H^TI.CL	0.07	NUM.GO^AJ^AJ^H	0.07
AS.DT^H^PP(*for* n)	0.07	NUM.GO^AJ^H	0.07
AS.DT^H^PP(*in* n)	0.07	NUM.GO^AJ^H^PL.ED.CL	0.07
AS.DT^H^PP(*of* n)	0.07	NUM.GO^H^PL.ED.CL	0.07
AS.DT^H^PP(*to* n)	0.07	NUM.GO^H^PP(*on* n)	0.07
AS.DT^N^H	0.07	NUM.GO^H^PP(*to* n)^RT.RV.CL	0.07
AS.PN^PV.PP(of DF.AR[…]^AJ^(ING)N^H)	0.07	PDT(CV.RU)^IN.AR^AJ^H	0.07
AS.PN^PV.PP(of DF.AR^H^PP(in which))	0.07	PDT^AJ^H	0.07
AS.PN^PV.PP(of DF.DV.GV.NP^H^PP of n)	0.07	PDT^IN.AR^H	0.07
AS.PN^PV.PP(of DM.DT(THESE)^H)	0.07	PS.DT^(ING)N^H	0.07
CV.QT.PN^PV.PP(of DF.AR^H^PP of n)	0.07	PS.DT^[…]^AJ^AJ^H	0.07
CV.QT^CV.AJP^H^PP(in n)	0.07	PS.DT^[…]^H	0.07
CV.QT^H[…]^RT.RV.CL	0.07	PS.DT^[…]H^PP(*on* n)	0.07
CV.QT^H^PP(on n)	0.07	PS.DT^AJ^AJ^AJ^H	0.07
DF.AR^(ING)N^H	0.07	PS.DT^AJ^AJ^H^PT.WK.NR.PA.AP	0.07
DF.AR^[…]H^[…]PP(*of* n)	0.07	PS.DT^AJ^H^[…]AP.TI.CL	0.07
DF.AR^[…]H^AP.TI.CL	0.07	PS.DT^AJ^H^AP.*THAT*.CL	0.07
DF.AR^[…]H^PP(*of* wh)	0.07	PS.DT^AJP^H^PP(*of* n)	0.07
DF.AR^[…]H^RT.RV.CL	0.07	PS.DT^H[…]^PP(*of* wh)	0.07
DF.AR^AJ^[…]AJ^AJ^H	0.07	PS.DT^H^[…]PP(*for* -ing)	0.07
DF.AR^AJ^AJ^H^PL.ING.CL	0.07	PS.DT^H^AP.*THAT*.CL	0.07
DF.AR^AJ^AJ^H^PP(*in* -ing)	0.07	PS.DT^H^NR.RV.CL	0.07
DF.AR^AJ^H^(*for* n TI.CL)	0.07	PS.DT^H^PL.ED.CL	0.07
DF.AR^AJ^H^AJP	0.07	PS.DT^H^PP(*at* -ing)	0.07
DF.AR^AJ^H^AP.*THAT*.CL	0.07	PS.DT^H^PP(*behind* -ing)	0.07
DF.AR^AJ^H^AP.TI.CL	0.07	PS.DT^H^PP(*during* n)	0.07
DF.AR^AJ^H^PP(*between* n)	0.07	PS.DT^H^PP(*in* n)	0.07
DF.AR^AJ^H^PP(*from within* n)	0.07	PS.DT^H^PP(*on* n)	0.07
DF.AR^AJ^H^PP(*in* n)	0.07	PS.DT^NUM.GO^H	0.07
DF.AR^AJ^H^PP(*of* n)^PP(*for* n)	0.07	PS.DT^NUM.OR^AJ^H	0.07
DF.AR^AJ^H^PP(*of* n)^RT.RV.CL	0.07	PS.DT^NUM.OR^H	0.07
DF.AR^AJ^H^PP(*over* n)	0.07	PS.DT^NUM.OR^H^PP(*into* n)	0.07
DF.AR^AJ^H^PP(*to* n)	0.07	PS.DT^NUM.OR^H^PP(*of* n)	0.07
DF.AR^AJ^H^PT.WK.NR.AP	0.07	PS.DT^NUM.OR^H^PP(*over* n)	0.07

DF.AR^AJ^H^RT.RV.CL^PP(*with* n)	0.07	PS.DT^QL.PV^AJ^H	0.07
DF.AR^AJ^H^TI.CL	0.07	PS.DT^QT^H	0.07
DF.AR^AJ^N^H	0.07	QT.IV^H	0.07
DF.AR^AJ^N^H^NR.RV.CL	0.07	QT.PN^PV.PP(*of* DF.AR^H^PP *of* n)	0.07
DF.AR^AJ^N^H^RT.RV.CL	0.07	QT.PN^PV.PP(*of* DF.AR^H^RT.RV.CL)	0.07
DF.AR^AJ^NUM.GO^[…]H^PP(*of* n)^[…]^PT.ST.NR.IT.AP	0.07	QT.PN^PV.PP(*of* DM.DT(*THESE*)^H)	0.07
DF.AR^AJP^H^FL.WK.NR.IT.AP	0.07	QT^AJ^H^PP(*for* n)^PT.WK.NR.IT.AP	0.07
DF.AR^AJP^H^PP(*of* n)	0.07	QT^AJ^H^PP(*on* n)	0.07
DF.AR^CV.AJ^H	0.07	QT^DF.AR^H	0.07
DF.AR^CV.AJP^PP(*of* […] DF.DV.GV.NP^H)	0.07	QT^DS.GV^H	0.07
DF.AR^H[…]^PP(*of* n)	0.07	QT^H[…]^PP(*of* n)	0.07
DF.AR^H^(*for* n AP.TI.CL)	0.07	QT^H^[…]^RT.RV.CL	0.07
DF.AR^H^[…]PP(*to* n)	0.07	QT^H^AP.*THAT*.CL	0.07
DF.AR^H^PL.ING.CL	0.07	QT^H^FL.ST.NR.EM.AP	0.07
DF.AR^H^PP(*behind* n -ing)	0.07	QT^H^FL.ST.NR.IT.AP	0.07
DF.AR^H^PP(*between* n)	0.07	QT^H^PL.ING.CL	0.07
DF.AR^H^PP(*for* -ing)	0.07	QT^H^PL.ING.CL^PP(*as to* wh)	0.07
DF.AR^H^PP(*from* n)	0.07	QT^H^PP(*concerning* n)	0.07
DF.AR^H^PP(*in* n)^AP.*THAT*.CL	0.07	QT^H^PP(*from* n)	0.07
DF.AR^H^PP(*of* n)^NR.RV.CL	0.07	QT^H^PP(*of* n)	0.07
DF.AR^H^PP(*of* n)^PP(*about* n)	0.07	QT^H^PP(*on* n)	0.07
DF.AR^H^PP(*of* n)^PP(*on* n)	0.07	QT^H^PP(*on* n)^PT.ST.NR.EM.AP	0.07
DF.AR^H^PP(*of* ps.dt -ing)	0.07	QT^NUM.GO^H^PP(*in*which)	0.07
DF.AR^H^PP(*with* n)	0.07	QV.PV^DM.DT(*THAT*)^H	0.07
DF.AR^H^PP(*with* which)	0.07	QV.PV^H	0.07
DF.AR^H^RT.RV.CL(*when*)	0.07	QV.PV^H^PT.WK.NR.IT.AP	0.07
DF.AR^H^RT.RV.CL(*where*)	0.07	QV.PV^UV.PDT^PS.DT^H	0.07
DF.AR^H^RT.RV.CL^PT.ST.NR.IT.AP	0.07	RV.DT^AJ^H	0.07
DF.AR^H^TI.CL	0.07	RV.DT^H^PP(*of* -ing)	0.07
DF.AR^N^H^AP.*THAT*.CL	0.07	RV.DT^H^PP(*of* n)	0.07
DF.AR^N^H^PP(*of* n)	0.07	UV.DT^H^PP(*of* n)	0.07
DF.AR^N^H^PT.ST.NR.ATT.AP	0.07	UV.DT^H^RT.RV.CL	0.07
DF.AR^NP^[…]AJ^H^RT.RV.CL	0.07	UV.DT^QL.PV^H	0.07
DF.AR^NP^AJ^H	0.07	UV.PDT^DF.AR^AJ^H^PP(*about* n)	0.07
DF.AR^NUM.CD^AJ^H^RT.RV.CL	0.07	UV.PDT^DF.AR^H^PL.ED.CL	0.07
DF.AR^NUM.CD^N^N^H^RT.RV.CL	0.07	UV.PDT^DF.AR^H^PP(*on* n)	0.07
DF.AR^NUM.CD^NUM.GO^H^RT.RV.CL	0.07	UV.PDT^DF.AR^H^RT.RV.CL	0.07
DF.AR^NUM.GO^AJ^H^TI.CL	0.07	UV.PDT^DM.DT(*THAT*)^H	0.07
DF.AR^NUM.GO^H^PP(*in* which)	0.07	UV.PDT^DM.DT(*THESE*)^AJ^H^NR.RV.CL	0.07
DF.AR^NUM.GO^H^RT.RV.CL	0.07	UV.PDT^H	0.07
DF.AR^NUM.GO^N^H^FL.ST.NR.IT.AP	0.07	UV.PDT^H^PP(*for* n)	0.07
DF.AR^NUM.OR^AJ^AJ^H^PP(*of* n)	0.07	UV.PDT^H^PP(*of* n)	0.07
DF.AR^NUM.OR^H	0.07	UV.PDT^NUM.GO^H^PP(*of* n)	0.07
DF.AR^NUM.OR^H^PP(*of* n)	0.07	UV.PDT^PS.DT^AJ^H	0.07
DF.AR^NUM.OR^H^TI.CL^PP(*with* n)	0.07	UV.PDT^QL.PV^H	0.07
DF.AR^NUM.OR^PV.PP(*of* DM.DT(*THESE*)^H)	0.07	UV.PDT^QL.PV^NUM.GO^H	0.07
DF.AR^QL.PV^AJ^[…]^H^PP(*of* n)	0.07	UV.PN^PV.PP(*of* DF.AR^AJ^NUM.CD^H)	0.07
DF.AR^QL.PV^AJ^H^RT.RV.CL	0.07	UV.PN^PV.PP(*of* DM.DT(*THESE*)^H)	0.07
DF.AR^QL.PV^H^RT.RV.CL(*where*)	0.07		

4.3 *Formal patterns (combined)*

Formal patterns (%)			
DF.AR^H	6.36	DF.AR^AJ^PP(*of* DF.DV.GV.NP^H)	0.07
DF.AR^H^PP	6.08	DF.AR^H^(*for* n AP.TI.CL)	0.07
H	5.25	DF.AR^H^PP^RV.CL	0.07
H^PP	4.91	DF.AR^H^RV.CL^PT.ST.NR.AP	0.07
AJ^H	4.63	DF.AR^H^TI.CL	0.07
DM.DT^H	3.73	DF.AR^N^AJ^H	0.07
IN.AR^AJ^H	3.11	DF.AR^N^AJ^H^RV.CL	0.07
IN.AR^H	3.11	DF.AR^N^H^AP.CL	0.07
DF.AR^AJ^H	2.97	DF.AR^N^H^PP	0.07
DF.AR^AJ^H^PP	2.70	DF.AR^N^H^PT.ST.NR.AP	0.07
PS.DT^H	2.35	DF.AR^NUM.CD^AJ^H^RV.CL	0.07
IN.AR^H^PP	2.21	DF.AR^NUM.CD^N^H^RV.CL	0.07
IN.AR^AJ^H^PP	2.00	DF.AR^NUM.CD^NUM.GO^H^RV.CL	0.07
DF.AR^H^RV.CL	1.66	DF.AR^NUM.GO^AJ^H^TI.CL	0.07
N^H	1.66	DF.AR^NUM.GO^H^PP	0.07
PS.DT^H^PP	1.66	DF.AR^NUM.GO^H^RV.CL	0.07
AJ^H^PP	1.45	DF.AR^NUM.GO^N^H^FL.ST.NR.AP	0.07
QT^H	1.45	DF.AR^NUM.OR^AJ^H^PP	0.07
PS.DT^AJ^H	1.04	DF.AR^NUM.OR^H	0.07
DF.AR^N^H	0.90	DF.AR^NUM.OR^H^PP	0.07
DF.DV.GV.NP^H	0.90	DF.AR^NUM.OR^H^TI.CL^PP	0.07
NUM.CD^H	0.90	DF.AR^NUM.OR^PV.PP(*of* DM.DT^H)	0.07
DF.DV.GV.NP^H^PP	0.83	DF.AR^QL.PV^AJ^H^PP	0.07
H^RV.CL	0.83	DF.AR^QL.PV^AJ^H^RV.CL	0.07
IN.AR^N^H	0.83	DF.AR^UNC^H^RV.CL	0.07
QT^H^PP	0.76	DF.DV.GV.NP^NUM.GO^H	0.07
AS.DT^H	0.69	DM.DT^AJ^H^PP	0.07
DF.AR^H^AP.CL	0.69	DM.DT^H^AP.CL	0.07
DM.DT^AJ^H	0.69	DM.DT^H^PP^FL.ST.NR.AP	0.07
IN.AR^H^AP.CL	0.69	DM.DT^H^PP^PL.CL	0.07
IN.AR^AJ^H^AP.CL	0.62	DM.DT^H^PP^PP	0.07
NUM.GO^H	0.62	DM.DT^NUM.GO^H	0.07
DF.AR^H^PP^AP.CL	0.55	DM.DT^QL.PV^AJ^H	0.07
DF.DV.GV.NP^AJ^H	0.55	DM.DT^QL.PV^AJ^H^RV.CL	0.07
NG.DT^AJ^H	0.55	H^(*for* n AP.CL)	0.07
NG.DT^H^PP	0.55	H^AJ	0.07
DF.AR^H^PL.CL	0.48	H^AP.CL	0.07
H^AP.CL	0.48	H^FL.ST.RT. AP	0.07
IN.AR^H^RV.CL	0.48	H^PP^PL.CL	0.07
NAS.DT^H	0.48	H^PP^PT.WK.NR.AP	0.07
DM.DT^H^PP	0.41	H^PP^RV.CL	0.07
H^TI.CL	0.41	IF.AV^IN.AR^AJ^H	0.07
NG.DT^H	0.41	IF.AV^IN.AR^H	0.07
NUM.CD^H^PP	0.41	IF.AV^QT^H^PP	0.07
PS.DT^H^AP.CL	0.41	IN.AR^AJ^H^AJ	0.07
DF.AR^AJ^H^RV.CL	0.35	IN.AR^AJ^H^FL.ST.NR.AP^PL.CL	0.07
DF.AR^QL.PV^H^RV.CL	0.35	IN.AR^AJ^H^PL.CL^AP.CL	0.07
DF.DV.GV.NP^H^AP.CL	0.35	IN.AR^AJ^N^H^PP	0.07
IN.AR^AJ^H^PL.CL	0.35	IN.AR^H^PT.ST.NR.AP	0.07
NUM.CD^AJ^H	0.35	IN.AR^H^TI.CL	0.07
NUM.GO^H^PP	0.35	IN.AR^N^H^AP.CL	0.07
PDT^H	0.35	IN.AR^N^H^RV.CL	0.07

PS.DT^N^H	0.35	IN.AR^NUM.GO^H	0.07
AS.DT^AJ^H	0.28	IV.DT^H^RV.CL	0.07
AS.DT^H^PP	0.28	IV.DT^QL.PV^H^PP	0.07
DF.AR^H^PP^PP	0.28	MR.GV^H^PP	0.07
IN.AR^AJ^N^H	0.28	N^H^AJ	0.07
IN.AR^H^PL.CL	0.28	N^H^FL.ST.NR.AP	0.07
IN.AR^N^H^PP	0.28	N^H^PT.ST.NR.AP^RV.CL	0.07
NAS.DT^H^PP	0.28	N^H^RV.CL	0.07
PS.DT^AJ^H^AP.CL	0.28	NAS.DT^AJ^H^PP	0.07
QT^AJ^H	0.28	NAS.DT^AJ^N^H	0.07
QT^H^RV.CL	0.28	NAS.DT^CV.QT^H	0.07
UV.DT^H	0.28	NAS.DT^H^AP.CL	0.07
AJ^H^RV.CL	0.21	NAS.DT^H^PP^PP	0.07
DF.AR^NUM.OR^H^RV.CL	0.21	NAS.DT^NUM.GO^H^PP	0.07
DF.DV.GV.NP^N^H	0.21	NAS.DT^QL.PV^H^RV.CL	0.07
DM.DT^QL.PV^H	0.21	NAS.PN^PV.PP(*of* DF.AR^H^PL.CL)	0.07
IN.AR^AJ^H^RV.CL	0.21	NAS.PN^PV.PP(*of* DF.AR^H^PP)	0.07
IN.AR^H^PP^RV.CL	0.21	NAS.PN^PV.PP(*of* DM.DT^H)	0.07
IV.DT^H	0.21	NG.DT^AJ^H^RV.CL	0.07
NAS.DT^AJ^H	0.21	NG.DT^H^TI.CL	0.07
NUM.CD^PV.PP(*of* DF.AR^AJ^H^RV.CL)	0.21	NG.DT^N^H	0.07
NUM.CD^PV.PP(*of* DF.AR^H^RV.CL)	0.21	NG.DT^NUM.GO^H	0.07
PS.DT^AJ^H^PP	0.21	NG.DT^NUM.GO^H^TI.CL	0.07
PS.DT^NUM.OR^H^PP	0.21	NUM.CD^AJ^H^PP	0.07
AJ^H^AP.CL	0.14	NUM.CD^AJ^H^PT.ST.NR.AP	0.07
AJ^H^PL.CL	0.14	NUM.CD^AJ^H^RV.CL	0.07
AJ^N^H^RV.CL	0.14	NUM.CD^H^AJ	0.07
AS.DT^QL.PV^H	0.14	NUM.CD^H^FL.ST.NR.AP	0.07
DF.AR^AJ^H^AP.CL	0.14	NUM.CD^H^PT.WK.NR.AP	0.07
DF.AR^AJ^N^H^RV.CL	0.14	NUM.CD^N^H	0.07
DF.AR^H^AV	0.14	NUM.CD^NUM.GO^H^PL.CL	0.07
DF.AR^H^FL.ST.RT.AP	0.14	NUM.CD^NUM.GO^H^PP	0.07
DF.AR^H^RV.CL^PP	0.14	NUM.CD^NUM.GO^H^RV.CL	0.07
DF.AR^NUM.CD^AJ^H^PP	0.14	NUM.CD^NUM.GO^QV.PV^H	0.07
DF.AR^NUM.CD^H	0.14	NUM.CD^PV.PP(*of* DF.AR^AJ^H)	0.07
DF.AR^NUM.GO^H	0.14	NUM.CD^PV.PP(*of* DF.AR^AJ^H^PL.CL)	0.07
DF.DV.GV.NP^NUM.CD^H	0.14	NUM.CD^PV.PP(*of* DF.AR^AJ^H^PP)	0.07
DM.DT^N^H	0.14	NUM.CD^PV.PP(*of* DF.AR^AJ^H^TI.CL)	0.07
H^PL.CL	0.14	NUM.CD^PV.PP(*of* DF.AR^N^H)	0.07
H^PP^AP.CL	0.14	NUM.CD^PV.PP(*of* DF.AR^NUM.GO^H^RV.CL)	0.07
H^PP^PP	0.14	NUM.CD^PV.PP(*of* DM.DT^AJ^H^RV.CL)	0.07
IN.AR^AJ^H^PP^PP	0.14	NUM.CD^PV.PP(of NUM.CD^H^AV^PT.WK.NR.AP)	0.07
IN.AR^AJ^H^PT.WK.NR.AP	0.14	NUM.CD^PV.PP(*of* PS.DT^H)	0.07
IN.AR^H^PT.WK.NR.AP	0.14	NUM.CD^PV.PP(*of* PS.DT^N^H)	0.07
IN.AR^N^AJ^H	0.14	NUM.CD^QV.PV^H	0.07
N^H^PP	0.14	NUM.GO^AJ^H^PL.CL	0.07
NAS.DT^H^RV.CL	0.14	NUM.GO^H^PL.CL	0.07
NG.DT^H^AP.CL	0.14	NUM.GO^H^PP^RV.CL	0.07
NUM.CD^H^PP^RV.CL	0.14	PDT(CV.RU)^IN.AR^AJ^H	0.07
NUM.CD^PV.PP(*of* DF.AR^H)	0.14	PDT^AJ^H	0.07

NUM.CD^PV.PP(*of* NUM.CD^AJ^H)	0.14
NUM.CD^PV.PP(*of* PS.DT^H^PP)	0.14
NUM.GO^AJ^H	0.14
PDT^H^RV.CL	0.14
QT.PN^PV.PP(*of* DF.AR^H^PP)	0.14
QT^AJ^H^PP(*in*)	0.14
QT^H^FL.ST.NR.AP	0.14
QT^H^TI.CL	0.14
QT^IN.AR^H	0.14
QT^NUM.GO^H^PP	0.14
RV.DT^H	0.14
RV.DT^H^PP	0.14
UV.PDT^DF.AR^H	0.14
UV.PDT^H^PP	0.14
AJ^H^FL.ST.RT.AP	0.07
AJ^H^PP^AP.CL	0.07
AJ^H^PP^PT.WK.NR.AP	0.07
AJ^H^PT.ST.NR.AP	0.07
AJ^H^PT.WK.NR.AP	0.07
AJ^IN.AR^H	0.07
AJ^N^AJ^H	0.07
AJ^N^H	0.07
AJ^N^H^PT.ST.NR.AP^RV.CL	0.07
AS.DT^AJ^H^PP	0.07
AS.DT^CV.QT^H^TI.CL	0.07
AS.DT^N^H	0.07
AS.PN^PV.PP(*of* DF.AR^AJ^N^H)	0.07
AS.PN^PV.PP(*of* DF.AR^H^PP)	0.07
AS.PN^PV.PP(*of* DF.DV.GV.NP^H^PP)	0.07
AS.PN^PV.PP(*of* DM.DT^H)	0.07
DF.AR^AJ^H^(*for* n TI.CL)	0.07
DF.AR^AJ^H^AJ	0.07
DF.AR^AJ^H^FL.WK.NR.AP	0.07
DF.AR^AJ^H^PL.CL	0.07
DF.AR^AJ^H^PP^PP	0.07
DF.AR^AJ^H^PP^RV.CL	0.07
DF.AR^AJ^H^PT.WK.NR.AP	0.07
DF.AR^AJ^H^RT.RV.CL^PP	0.07
DF.AR^AJ^H^TI.CL	0.07
DF.AR^AJ^N^H	0.07
DF.AR^AJ^NUM.GO^H^PP^PT.ST.NR.AP	0.07
PDT^IN.AR^H	0.07
PS.DT^AJ^H^PT.WK.NR.AP	0.07
PS.DT^H^PL.CL	0.07
PS.DT^H^RV.CL	0.07
PS.DT^NUM.GO^H	0.07
PS.DT^NUM.OR^AJ^H	0.07
PS.DT^NUM.OR^H	0.07
PS.DT^QL.PV^AJ^H	0.07
PS.DT^QT^H	0.07
QT.IV^H	0.07
QT.PN^PV.PP(*of* DF.AR^H^RV.CL)	0.07
QT.PN^PV.PP(*of* DM.DT^H)	0.07
QT^AJ^H^PP^PT.WK.NR.AP	0.07
QT^DF.AR^H	0.07
QT^DS.GV^H	0.07
QT^H^AP.*THAT*.CL	0.07
QT^H^PL.CL	0.07
QT^H^PL.CL^PP	0.07
QT^H^PP^PT.ST.NR.AP	0.07
QV.PV^DM.DT^H	0.07
QV.PV^H	0.07
QV.PV^H^PT.WK.NR.AP	0.07
QV.PV^UV.PDT^PS.DT^H	0.07
RV.DT^AJ^H	0.07
UV.DT^H^PP	0.07
UV.DT^H^RV.CL	0.07
UV.DT^QL.PV^H	0.07
UV.PDT^DF.AR^AJ^H^PP	0.07
UV.PDT^DF.AR^H^PL.CL	0.07
UV.PDT^DF.AR^H^PP	0.07
UV.PDT^DF.AR^H^RV.CL	0.07
UV.PDT^DM.DT^AJ^H^RV.CL	0.07
UV.PDT^DM.DT^H	0.07
UV.PDT^H	0.07
UV.PDT^NUM.GO^H^PP	0.07
UV.PDT^PS.DT^AJ^H	0.07
UV.PDT^QL.PV^H	0.07
UV.PDT^QL.PV^NUM.GO^H	0.07
UV.PN^PV.PP(*of* DF.AR^AJ^NUM.CD^H)	0.07
UV.PN^PV.PP(*of* DM.DT^H)	0.07

APPENDIX 5 EXPERIENTIAL PATTERNS WITH CLASSIFIERS, EPITHETS AND POST-DEICTICS. PERCENTAGES ARE BASED ON THE TOTAL NUMBER OF PATTERN TOKENS WITH ONE OR MORE EXAMPLES OF EACH SEMANTIC CATEGORY (220 FOR CLASSIFIERS, 195 FOR EPITHETS AND 162 FOR POST-DEICTICS)

Classifier patterns	%	Epithet patterns	%
CS^TG	20.45	DC.NSP.PT.NSL^EP.IP^TG	9.23
DC.NSP.PT.NSL^CS^TG	10.00	DC.NSP.PT.NSL^EP.IP^TG^QF	8.21
DC.SP.DM.DV^CS^TG	9.55	DC.SP.DM.DV^EP.IP^TG^QF	7.18
DC.NSP.PT.NSL^CS^TG^QF	7.73	DC.SP.DM.DV^EP.IP^TG	5.13
CS^TG^QF	7.27	DC.NSP.PT.NSL^EP.EX^TG	4.62
DC.SP.DM.DV^CS^TG^QF	6.36	EP.EX^TG	4.62
DC.SP.PS.DV^CS^TG	5.91	DC.SP.DM.DV^EP.EX^TG	4.10
CS^CS^TG	1.82	DC.NSP.PT.NSL^EP.EX^TG^QF	3.59
DC.NSP.PT.NSL^EP.EX^CS^TG	1.82	DC.SP.DM.DV^EP.EX^TG^QF	3.59
DC.NSP.PT.NSL^CS^CS^TG	1.36	EP.IP^TG	3.59
DC.NSP.TL.NG^CS^TG	1.36	EP.IP^TG^QF	3.59
DC.NSP.PT.NSL^EP.EX^CS^TG^QF	0.91	DC.SP.PS.DV^EP.IP^TG	2.56
DC.NSP.PT.SL^CS^TG	0.91	EP.EX^TG^QF	2.56
DC.SP.DM.DV^CS^CS^TG	0.91	DC.NSP.PT.NSL^EP.EX^CS^TG	2.05
DC.SP.DM.DV^EP.EX^CS^TG	0.91	DC.NSP.PT.SL^DC.SP.DM.DV^EP.IP^TG^QF	1.54
DC.SP.DM.DV^EP.EX^CS^TG^QF	0.91	DC.NSP.PT.NSL^EP.EX^CS^TG^QF	1.03
DC.SP.PS.DV^CS^TG^QF	0.91	DC.NSP.PT.NSL^EP.EX^TG[…]^QF	1.03
NUM.IN.QV^CS^TG	0.91	DC.NSP.PT.NSL^EP.IP^TG^QF^QF	1.03
CS^CS^CS^TG	0.45	DC.NSP.TL.NG^EP.EX^TG	1.03
CS^CS^TG^QF	0.45	DC.SP.DM.DV^EP.EX^CS^TG	1.03
CS^TG^QF^QF	0.45	DC.SP.DM.DV^EP.EX^CS^TG^QF	1.03
DC.NSP.PT.NSL^[…]CS^TG	0.45	DC.SP.PS.DV^EP.IP^TG^QF	1.03
DC.NSP.PT.NSL^CS^[…]^EP.IP^TG	0.45	EP.EX^TG^QF^QF	1.03
DC.NSP.PT.NSL^CS^[…]TG^QF	0.45	DC.NSP.PT.NSL^CS^[…]^EP.IP^TG	0.51
DC.NSP.PT.NSL^CS^TG^QF^QF	0.45	DC.NSP.PT.NSL^EP.EX^EP.EX^TG	0.51
DC.NSP.PT.NSL^EP.IP^CS^TG^QF	0.45	DC.NSP.PT.NSL^EP.IP^CS^TG^QF	0.51
DC.NSP.PT.SL^DC.SP.DM.DV^[…]EP.EX^CS^TG	0.45	DC.NSP.PT.NSL^EP.IP^EP.IP^TG^QF	0.51
DC.NSP.PT.SL^DC.SP.DM.DV^CS^TG	0.45	DC.NSP.PT.SL^DC.SP.DM.DV^[…]EP.EX^CS^TG	0.51
DC.NSP.PT.SL^DC.SP.DM.DV^PDC.EN.SPA-TM^CS^TG^QF	0.45	DC.NSP.PT.SL^DC.SP.DM.DV^EP.EX^TG^QF	0.51
DC.NSP.PT.SL^DC.SP.PS.DV^CS^TG	0.45	DC.NSP.PT.SL^DC.SP.DM.DV^EP.IP^TG	0.51
DC.NSP.PT.SL^EP.EX^CS^TG	0.45	DC.NSP.PT.SL^EP.EX^CS^TG	0.51
DC.NSP.TL.NG^EP.EX^CS^TG	0.45	DC.NSP.PT.SL^EP.EX^TG	0.51
DC.SP.DM.DV^CS^[…]EP.EX^TG^QF	0.45	DC.NSP.PT.SL^EP.IP^TG^QF	0.51
DC.SP.DM.DV^CS^TG^QF^QF	0.45	DC.NSP.TL.NG^EP.EX^CS^TG	0.51
DC.SP.DM.DV^EP.EX^[…]CS^CS^TG	0.45	DC.NSP.TL.NG^EP.IP^TG	0.51
DC.SP.DM.DV^EP.IP^CS^TG	0.45	DC.NSP.TL.POS^DC.SP.DM.DV^EP.EX^TG^QF	0.51
DC.SP.DM.DV^FC.EL.EM^CS^[…]	0.45	DC.NSP.TL.POS^DC.SP.DM.DV^EP.IP	0.51

^TG^QF	
DC.SP.DM.DV^FC.EL.EM^CS^TG^QF	0.45
DC.SP.DM.DV^NUM.DF.OR^CS^TG^QF	0.45
DC.SP.DM.DV^NUM.DF.QV^CS^CS^TG^QF	0.45
DC.SP.DM.DV^NUM.DF.QV^CS^TG^QF	0.45
DC.SP.DM.DV^PDC.EN.SPA-TM^CS^TG^QF	0.45
DC.SP.DM.DV^PDC.MD.US^CS^TG^QF	0.45
DC.SP.DM.DV^PDC.RP.IA.EZ^CS^TG^QF	0.45
DC.SP.DM.DV^PDC.RP.IA^CS^TG	0.45
DC.SP.DM.DV^PDC.RP.LN.EZ^CS^TG^QF	0.45
DC.SP.PS.DV^CS^[…]TG	0.45
DC.SP.PS.DV^CS^CS^CS^TG	0.45
DC.SP.PS.DV^CS^CS^TG	0.45
DC.SP.PS.DV^CS^CS^TG^QF	0.45
DC.SP.PS.DV^FC.EL.EM^CS^TG	0.45
DC.SP.PS.DV^NUM.DF.OR^CS^TG	0.45
EP.EX^CS^CS^TG	0.45
EP.EX^CS^TG^[…]QF	0.45
EP.EX^CS^TG^QF	0.45
EP.EX^CS^TG^QF^QF	0.45
NUM.DF.QV^CS^TG	0.45
NUM.DF.QV^NUM.DF.QV^CS^TG	0.45
NUM.IN.QV^CS^TG^QF	0.45
PDC.EL.ID^CS^TG^QF	0.45
PDC.EL.ID^EP.EX^CS^TG	0.45
PDC.EN.SPA-TM^CS^TG	0.45

^TG^QF	
DC.NSP.TL.POS^DC.SP.PS.DV^EP.IP^TG	0.51
DC.SP.DM.DV^CS^[…]EP.EX^TG^QF	0.51
DC.SP.DM.DV^EP.EX^[…]CS^CS^TG	0.51
DC.SP.DM.DV^EP.IP^[…]DC.SP.PS.DV^TG	0.51
DC.SP.DM.DV^EP.IP^CS^TG	0.51
DC.SP.DM.DV^EP.IP^EP.EX^TG	0.51
DC.SP.DM.DV^FC.EL.EM^EP.EX^TG	0.51
DC.SP.DM.DV^FC.EL.EM^EP.IP^TG^QF	0.51
DC.SP.DM.DV^NUM.DF.OR^PDC.MD.PB^EP.EX^TG^QF	0.51
DC.SP.DM.DV^PDC.EL.ID^EP.IP^TG^QF	0.51
DC.SP.PS.DV^[…]^EP.IP^EP.IP^TG	0.51
DC.SP.PS.DV^EP.EX^TG	0.51
DC.SP.PS.DV^EP.EX^TG^QF	0.51
DC.SP.PS.DV^EP.IP^TG^[…]QF	0.51
EP.EX^[…]TG	0.51
EP.EX^CS^CS^TG	0.51
EP.EX^CS^TG^[…]QF	0.51
EP.EX^CS^TG^QF	0.51
EP.EX^CS^TG^QF^QF	0.51
EP.EX^EP.EX^TG	0.51
EP.EX^TG^[…]QF	0.51
EP.IP^DC.NSP.PT.NSL^TG	0.51
EP.IP^EP.EX^TG	0.51
EP.IP^EP.IP^TG	0.51
EP.IP^TG^[…]^QF	0.51
IF^DC.NSP.PT.NSL^EP.EX^TG	0.51
NUM.DF.QV^EP.EX^TG	0.51
NUM.DF.QV^EP.IP^TG	0.51
NUM.DF.QV^EP.IP^TG^QF	0.51
NUM.IN.QV^EP.EX^TG	0.51
NUM.IN.QV^EP.EX^TG^QF^QF	0.51
NUM.IN.QV^EP.IP^TG	0.51
NUM.IN.QV^EP.IP^TG^QF	0.51
PDC.EL.ID^EP.EX^CS^TG	0.51
PDC.EL.ID^EP.IP^TG	0.51
PDC.ET.EZ^EP.IP^TG	0.51

post-Deictic patterns (%)			
PDC.EL.ID^TG	6.17	DC.NSP.TL.POS^PDC.EL.ID^TG^QF	0.62
DC.SP.DM.DV^PDC.EL.ID.RT^TG^QF	5.56	DC.SP.DM.DV^DC.NSP.PT.SL^PDC.EL.ID^TG^QF	0.62
DC.SP.DM.DV^PDC.EN.SPA-TM^TG	5.56	DC.SP.DM.DV^NUM.DF.OR^PDC.MD.PB^EP.EX^TG^QF	0.62
DC.SP.DM.DV^PDC.EL.ID.RT^TG	3.70	DC.SP.DM.DV^NUM.DF.QV^PDC.EL.EM^TG^QF	0.62
PDC.EL.ID^TG^QF	3.70	DC.SP.DM.DV^NUM.DF.QV^PDC.EL.ID.RT^TG^QF	0.62
PDC.EN.CV^TG	3.09	DC.SP.DM.DV^PDC.EL.EM.EZ^TG^QF	0.62
DC.SP.PS.DV^PDC.EL.ID.RT^TG	2.47	DC.SP.DM.DV^PDC.EL.EM^TG^QF	0.62
PDC.ET.AM^TG	2.47	DC.SP.DM.DV^PDC.EL.ID^EP.IP^TG^QF	0.62
DC.SP.DM.DV^PDC.EL.EM.RT^TG	1.85	DC.SP.DM.DV^PDC.EL.ID^TG	0.62
PDC.EL.EM.RT^TG	1.85	DC.SP.DM.DV^PDC.EL.ID^TG^QF	0.62
DC.NSP.PT.NSL^PDC.EL.EM.RT^TG^QF	1.23	DC.SP.DM.DV^PDC.EN.SPA-TM^CS^TG^QF	0.62
DC.NSP.PT.NSL^PDC.EL.ID^TG	1.23	DC.SP.DM.DV^PDC.EN.SPA-TM^PDC.EL.ID^[…]TG^QF^[…]^QF	0.62
DC.NSP.PT.NSL^PDC.MD.US^TG^QF	1.23	DC.SP.DM.DV^PDC.EN.SPA-TM^TG^QF	0.62
DC.NSP.PT.SL^PDC.EL.ID^TG^QF	1.23	DC.SP.DM.DV^PDC.ET.AM^TG	0.62
DC.NSP.TL.NG^PDC.EL.ID^TG^QF	1.23	DC.SP.DM.DV^PDC.ET.EZ^TG^QF	0.62
DC.SP.DM.DV^PDC.EL.EM.EZ^TG	1.23	DC.SP.DM.DV^PDC.ET.RT^TG	0.62
DC.SP.DM.DV^PDC.EL.EM.RT^TG^QF	1.23	DC.SP.DM.DV^PDC.MD.OB^TG	0.62
DC.SP.DM.DV^PDC.EL.ID.RT^TG^QF^QF	1.23	DC.SP.DM.DV^PDC.MD.US^CS^TG^QF	0.62
DC.SP.DM.DV^PDC.ET.EZ^TG	1.23	DC.SP.DM.DV^PDC.MD.US^TG^QF	0.62
DC.SP.DM.DV^PDC.MD.PB^TG	1.23	DC.SP.DM.DV^PDC.RP.IA.EZ^CS^TG^QF	0.62
NUM.IN.QV^PDC.EL.ID^TG^QF	1.23	DC.SP.DM.DV^PDC.RP.IA^CS^TG	0.62
PDC.EL.EM.RT^TG^QF	1.23	DC.SP.DM.DV^PDC.RP.IA^PDC.MD.US^TG^QF	0.62
DC.NSP.PT.NSL^PDC.EL.EM.AM^TG	0.62	DC.SP.DM.DV^PDC.RP.LN.EZ^CS^TG^QF	0.62
DC.NSP.PT.NSL^PDC.EL.EM.EZ^TG	0.62	DC.SP.PS.DV^PDC.EL.ID.EZ^TG	0.62
DC.NSP.PT.NSL^PDC.EL.EM.RT^TG	0.62	DC.SP.PS.DV^PDC.EL.ID^TG	0.62
DC.NSP.PT.NSL^PDC.EL.EM^TG	0.62	DC.SP.PS.DV^PDC.EN.SPA-TM^TG	0.62
DC.NSP.PT.NSL^PDC.EL.ID.EZ^TG^QF	0.62	DC.SP.PS.DV^PDC.ET.AM^TG^QF	0.62
DC.NSP.PT.NSL^PDC.EL.ID.RT^TG	0.62	DC.SP.PS.DV^PDC.MD.RD^TG	0.62
DC.NSP.PT.NSL^PDC.EL.ID^TG^QF	0.62	DC.SP.PS.IV^PDC.EL.ID.RT^TG	0.62
DC.NSP.PT.NSL^PDC.EN.CV^TG^QF^QF	0.62	NUM.DF.QV^NUM.DF.QV^PDC.EL.ID^TG	0.62
DC.NSP.PT.NSL^	0.62	NUM.DF.QV^PDC.EL.ID.RT^TG	0.62

PDC.EN.SPA-TM^TG^QF			
DC.NSP.PT.NSL^PDC.ET.AM^TG^QF	0.62	NUM.DF.QV^PDC.EL.ID^TG	0.62
DC.NSP.PT.NSL^PDC.MD.PB^TG	0.62	NUM.DF.QV^PDC.EL.ID^TG^QF	0.62
DC.NSP.PT.NSL^PDC.MD.PB^TG^QF	0.62	NUM.DF.QV^PDC.MD.PB^TG^QF	0.62
DC.NSP.PT.NSL^PDC.MD.US.EZ^TG	0.62	PDC.EL.EM.RT^[...]TG	0.62
DC.NSP.PT.NSL^PDC.RP.IA.EZ^TG^QF	0.62	PDC.EL.EM^PDC.EL.ID^TG	0.62
DC.NSP.PT.NSL^PDC.RP.IA^TG	0.62	PDC.EL.EM^TG	0.62
DC.NSP.PT.SL^DC.SP.DM.DV^PDC.EL.EM^TG^QF	0.62	PDC.EL.EM^TG^QF	0.62
DC.NSP.PT.SL^DC.SP.DM.DV^PDC.EL.ID^TG^QF	0.62	PDC.EL.ID.RT^TG	0.62
DC.NSP.PT.SL^DC.SP.DM.DV^PDC.EN.SPA-TM^CS^TG^QF	0.62	PDC.EL.ID^CS^TG^QF	0.62
DC.NSP.PT.SL^DC.SP.DM.DV^PDC.MD.PB.EZ^TG^QF	0.62	PDC.EL.ID^EP.EX^CS^TG	0.62
DC.NSP.PT.SL^PDC.EL.EM.RT^TG	0.62	PDC.EL.ID^EP.IP^TG	0.62
DC.NSP.PT.SL^PDC.EL.EM.RT^TG^QF	0.62	PDC.EL.ID^TG^QF^QF	0.62
DC.NSP.PT.SL^PDC.EL.ID.RT^TG	0.62	PDC.EN.CV^PDC.EN.CV^TG	0.62
DC.NSP.TL.NG^PDC.EL.EM.AM^TG	0.62	PDC.EN.SPA-TM^CS^TG	0.62
DC.NSP.TL.NG^PDC.EL.EM.RT^TG	0.62	PDC.EN.SPA-TM^TG^QF	0.62
DC.NSP.TL.NG^PDC.EL.ID^TG	0.62	PDC.ET.AM^TG^QF	0.62
DC.NSP.TL.NG^PDC.EN.SPA-TM^TG	0.62	PDC.ET.EZ^EP.IP^TG	0.62
DC.NSP.TL.NG^PDC.RP.LN^TG^QF	0.62	PDC.MD.US^TG	0.62
DC.NSP.TL.POS^DC.SP.DM.DV^PDC.EN.SPA-TM^NUM.DF.QV^TG	0.62	PDC.RP.IA.EZ^TG	0.62
DC.NSP.TL.POS^FC.EL.EM^PDC.EL.ID^TG	0.62		

APPENDIX 6 MODE DISTRIBUTION OF EXPERIENTIAL AND FORMAL PATTERNS (COMPLETE LISTS)

6.1 *Experiential patterns*

Written	%	Spoken	%
DC.SP.DM.DV^TG^QF	9.21	DC.SP.DM.DV^TG	12.87
DC.SP.DM.DV^TG	9.10	DC.SP.DM.DV^TG^QF	9.75
TG^QF	7.28	TG^QF	5.85
TG	6.00	DC.NSP.PT.NSL^TG	5.65
DC.SP.PS.DV^TG^QF	3.85	DC.NSP.PT.NSL^TG^QF	5.07
CS^TG	3.75	TG	3.70
DC.NSP.PT.NSL^TG^QF	3.21	DC.SP.PS.DV^TG	3.51
DC.SP.PS.DV^TG	2.89	NUM.IN.QV^TG	2.53
DC.NSP.PT.NSL^TG	2.25	CS^TG	1.95
CS^TG^QF	1.61	DC.SP.PS.DV^TG^QF	1.75
DC.NSP.PT.NSL^CS^TG	1.50	DC.NSP.PT.NSL^CS^TG	1.56
DC.NSP.PT.NSL^CS^TG^QF	1.50	DC.NSP.PT.NSL^EP.IP^TG	1.36
DC.SP.DM.DV^CS^TG	1.50	DC.SP.DM.DV^CS^TG	1.36
DC.SP.DM.DV^TG^QF^QF	1.39	DC.SP.DM.DV^EP.IP^TG^QF	1.36
DC.NSP.PT.NSL^EP.IP^TG^QF	1.28	DC.NSP.TL.NG^TG^QF	1.17
DC.SP.DM.DV^CS^TG^QF	1.28	DC.SP.DM.DV^TG^QF^QF	1.17
DC.NSP.PT.NSL^EP.IP^TG	1.18	NUM.DF.QV^TG^QF	0.97
DC.SP.PS.DV^CS^TG	1.18	DC.NSP.PT.NSL^EP.IP^TG^QF	0.78
NUM.IN.QV^TG^QF	1.18	DC.NSP.PT.SL^DC.SP.DM.DV^TG^QF	0.78
DC.NSP.TL.NG^TG	0.86	DC.NSP.TL.NG^TG	0.78
DC.NSP.TL.NG^TG^QF	0.86	DC.NSP.TL.POS^DC.SP.DM.DV^TG	0.78
NUM.IN.QV^TG	0.86	DC.SP.DM.DV^FC.EL.EM^TG^QF	0.78
DC.NSP.PT.NSL^EP.EX^TG^QF	0.75	NUM.DF.QV^TG	0.78
DC.SP.DM.DV^EP.IP^TG	0.75	NUM.IN.QV^TG^QF	0.78
DC.SP.DM.DV^EP.IP^TG^QF	0.75	TG^[…]^QF	0.78
NUM.DF.QV^TG	0.75	DC.NSP.PT.NSL^CS^TG^QF	0.58
PDC.EL.ID^TG	0.75	DC.NSP.PT.NSL^EP.EX^TG	0.58
DC.NSP.PT.NSL^EP.EX^TG	0.64	DC.NSP.PT.SL^TG	0.58
DC.SP.DM.DV^EP.EX^TG	0.64	DC.SP.DM.DV^EP.IP^TG	0.58
DC.SP.DM.DV^PDC.EL.ID.RT^TG^QF	0.64	DC.SP.DM.DV^NUM.DF.OR^TG^QF	0.58
DC.SP.DM.DV^PDC.EN.SPA-TM^TG	0.64	DC.SP.DM.DV^PDC.EL.EM.RT^TG	0.58
EP.EX^TG	0.64	DC.SP.DM.DV^PDC.EL.ID.RT^TG	0.58
EP.IP^TG	0.64	DC.SP.DM.DV^PDC.EL.ID.RT^TG^QF	0.58
TG^QF^QF	0.64	DC.SP.DM.DV^PDC.EN.SPA-TM^TG	0.58
DC.NSP.PT.SL^TG^QF	0.54	EP.EX^TG	0.58
DC.SP.DM.DV^EP.EX^TG^QF	0.54	PDC.EL.ID^TG	0.58
EP.IP^TG^QF	0.54	PDC.EL.ID^TG^QF	0.58
CS^CS^TG	0.43	DC.NSP.PT.NSL^FC.EL.EM^TG	0.39
DC.NSP.PT.SL^TG	0.43	DC.NSP.PT.SL^DC.SP.DM.DV^EP.IP^TG^QF	0.39
DC.NSP.TL.POS^TG	0.43	DC.NSP.PT.SL^DC.SP.DM.DV^TG	0.39
PDC.EN.CV^TG	0.43	DC.NSP.PT.SL^DC.SP.PS.DV^TG^QF	0.39
DC.NSP.PT.NSL^CS^CS^TG	0.32	DC.NSP.PT.SL^TG^QF	0.39
DC.NSP.PT.NSL^EP.EX^CS^TG	0.32	DC.NSP.TL.POS^DC.SP.DM.DV^TG^QF	0.39
DC.NSP.TL.POS^TG^QF	0.32	DC.SP.DM.DV^[…]TG^QF	0.39
DC.SP.DM.DV^[…]TG^QF	0.32	DC.SP.DM.DV^CS^TG^QF	0.39
DC.SP.DM.DV^PDC.EL.ID.RT^TG	0.32	DC.SP.DM.DV^EP.EX^TG	0.39

DC.SP.PS.DV^EP.IP^TG	0.32	DC.SP.DM.DV^EP.EX^TG^QF	0.39
DC.SP.PS.DV^NUM.DF.OR^TG	0.32	DC.SP.DM.DV^PDC.ET.EZ^TG	0.39
DC.SP.PS.DV^NUM.DF.OR^TG^QF	0.32	DC.SP.PS.DV^CS^TG	0.39
DC.SP.PS.DV^PDC.EL.ID.RT^TG	0.32	DC.SP.PS.DV^EP.IP^TG	0.39
EP.EX^TG^QF	0.32	DC.SP.PS.DV^TG[…]^QF	0.39
NUM.DF.QV^TG^QF	0.32	EP.EX^TG^QF	0.39
NUM.IN.QV^DC.SP.DM.DV^TG^QF	0.32	EP.IP^TG^QF	0.39
PDC.EL.ID^TG^QF	0.32	NUM.IN.QV^DC.NSP.PT.NSL^TG	0.39
PDC.ET.AM^TG	0.32	NUM.IN.QV^TG[…]^QF	0.39
DC.NSP.PT.NSL^EP.EX^CS^TG^QF	0.21	CS^TG^QF	0.19
DC.NSP.PT.NSL^EP.IP^TG^QF^QF	0.21	DC.NSP.PT.NSL^CS^[…]^EP.IP^TG	0.19
DC.NSP.PT.NSL^PDC.EL.EM.RT^TG^QF	0.21	DC.NSP.PT.NSL^CS^[…]TG^QF	0.19
DC.NSP.PT.NSL^PDC.MD.US^TG^QF	0.21	DC.NSP.PT.NSL^EP.EX^CS^TG	0.19
DC.NSP.PT.NSL^TG[…]^QF	0.21	DC.NSP.PT.NSL^EP.EX^TG^[…]QF	0.19
DC.NSP.PT.NSL^TG^QF^QF	0.21	DC.NSP.PT.NSL^EP.IP^EP.IP^TG^QF	0.19
DC.NSP.PT.SL^CS^TG	0.21	DC.NSP.PT.NSL^PDC.EL.EM.EZ^TG	0.19
DC.NSP.PT.SL^DC.SP.DM.DV^TG	0.21	DC.NSP.PT.NSL^PDC.EL.EM.RT^TG	0.19
DC.NSP.PT.SL^DC.SP.DM.DV^TG^QF	0.21	DC.NSP.PT.NSL^PDC.EL.ID^TG	0.19
DC.NSP.TL.NG^CS^TG	0.21	DC.NSP.PT.NSL^PDC.MD.US.EZ^TG	0.19
DC.NSP.TL.NG^EP.EX^TG	0.21	DC.NSP.PT.NSL^PDC.RP.IA^TG	0.19
DC.SP.DM.DV^CS^CS^TG	0.21	DC.NSP.PT.NSL^TG^[…]QF	0.19
DC.SP.DM.DV^EP.EX^CS^TG	0.21	DC.NSP.PT.NSL^TG^QF^QF	0.19
DC.SP.DM.DV^EP.EX^CS^TG^QF	0.21	DC.NSP.PT.SL^DC.SP.DM.DV^[…]EP.EX^CS^TG	0.19
DC.SP.DM.DV^FC.EL.SM^TG	0.21	DC.NSP.PT.SL^DC.SP.DM.DV^CS^TG	0.19
DC.SP.DM.DV^NUM.DF.OR^TG^QF	0.21	DC.NSP.PT.SL^DC.SP.DM.DV^EP.EX^TG^QF	0.19
DC.SP.DM.DV^PDC.EL.ID.RT^TG^QF^QF	0.21	DC.NSP.PT.SL^DC.SP.DM.DV^EP.IP^TG	0.19
DC.SP.DM.DV^PDC.MD.PB^TG	0.21	DC.NSP.PT.SL^DC.SP.DM.DV^PDC.EL.EM^TG^QF	0.19
DC.SP.PS.DV^[…]^TG	0.21	DC.NSP.PT.SL^DC.SP.DM.DV^PDC.EL.ID^TG^QF	0.19
DC.SP.PS.DV^CS^TG^QF	0.21	DC.NSP.PT.SL^DC.SP.DM.DV^PDC.MD.PB.EZ^TG^QF	0.19
DC.SP.PS.DV^NUM.DF.QV^TG	0.21	DC.NSP.PT.SL^DC.SP.PS.DV^[…]TG	0.19
DC.SP.PS.IV^TG	0.21	DC.NSP.PT.SL^EP.EX^TG	0.19
EP.EX^TG^QF^QF	0.21	DC.NSP.PT.SL^FC.EL.EM^TG^QF	0.19
NUM.IN.QV^TG[…]^QF	0.21	DC.NSP.PT.SL^FC.ET.PV^TG	0.19
PDC.EL.EM.RT^TG	0.21	DC.NSP.PT.SL^NUM.IN.QV^TG	0.19
(NG)TG	0.11	DC.NSP.PT.SL^PDC.EL.EM.RT^TG	0.19
CS^CS^CS^TG	0.11	DC.NSP.PT.SL^PDC.EL.ID.RT^TG	0.19
CS^CS^TG^QF	0.11	DC.NSP.PT.SL^PDC.EL.ID^TG^QF	0.19
CS^TG^QF^QF	0.11	DC.NSP.PT.SL^TG[…]^QF	0.19
DC.NSP.PT.NSL^[…]^TG	0.11	DC.NSP.PT.SL^TG^QF^QF	0.19
DC.NSP.PT.NSL^[…]CS^TG	0.11	DC.NSP.TL.NG^CS^TG	0.19
DC.NSP.PT.NSL^[…]TG^QF	0.11	DC.NSP.TL.NG^PDC.EL.ID^TG^QF	0.19
DC.NSP.PT.NSL^CS^TG^QF^QF	0.11	DC.NSP.TL.POS^DC.SP.DM.DV^EP.IP^TG^QF	0.19
DC.NSP.PT.NSL^EP.EX^EP.EX^TG	0.11	DC.NSP.TL.POS^FC.EL.EM^TG	0.19
DC.NSP.PT.NSL^EP.EX^TG[…]^QF	0.11	DC.NSP.TL.POS^PDC.EL.ID^TG^QF	0.19
DC.NSP.PT.NSL^EP.IP^CS^TG^QF	0.11	DC.NSP.TL.POS^TG	0.19
DC.NSP.PT.NSL^NUM.IN.QV^TG^	0.11	DC.NSP.TL.POS^TG^QF	0.19

QF			
DC.NSP.PT.NSL^PDC.EL.EM.AM^TG	0.11	DC.SP.DM.DV^CS^[…]EP.EX^TG^QF	0.19
DC.NSP.PT.NSL^PDC.EL.EM^TG	0.11	DC.SP.DM.DV^DC.NSP.PT.SL^PDC.EL.ID^TG^QF	0.19
DC.NSP.PT.NSL^PDC.EL.ID.EZ^TG^QF	0.11	DC.SP.DM.DV^FC.EL.EM^CS^[…]^TG^QF	0.19
DC.NSP.PT.NSL^PDC.EL.ID.RT^TG	0.11	DC.SP.DM.DV^FC.EL.EM^EP.EX^TG	0.19
DC.NSP.PT.NSL^PDC.EL.ID^TG	0.11	DC.SP.DM.DV^FC.EL.EM^EP.IP^TG^QF	0.19
DC.NSP.PT.NSL^PDC.EL.ID^TG^QF	0.11	DC.SP.DM.DV^FC.EL.EM^TG	0.19
DC.NSP.PT.NSL^PDC.EN.CV^TG^QF^QF	0.11	DC.SP.DM.DV^NUM.DF.QV^PDC.EL.ID.RT^TG^QF	0.19
DC.NSP.PT.NSL^PDC.EN.SPA-TM^TG^QF	0.11	DC.SP.DM.DV^NUM.DF.QV^TG	0.19
DC.NSP.PT.NSL^PDC.ET.AM^TG^QF	0.11	DC.SP.DM.DV^PDC.EL.EM.RT^TG^QF	0.19
DC.NSP.PT.NSL^PDC.MD.PB^TG	0.11	DC.SP.DM.DV^PDC.EL.ID^TG	0.19
DC.NSP.PT.NSL^PDC.MD.PB^TG^QF	0.11	DC.SP.DM.DV^PDC.EL.ID^TG^QF	0.19
DC.NSP.PT.NSL^PDC.RP.IA.EZ^TG^QF	0.11	DC.SP.DM.DV^PDC.RP.IA^CS^TG	0.19
DC.NSP.PT.SL^DC.SP.DM.DV^EP.IP^TG^QF	0.11	DC.SP.DM.DV^PDC.RP.IA^PDC.MD.US^TG^QF	0.19
DC.NSP.PT.SL^DC.SP.DM.DV^PDC.EN.SPA-TM^CS^TG^QF	0.11	DC.SP.DM.DV^PDC.RP.LN.EZ^CS^TG^QF	0.19
DC.NSP.PT.SL^DC.SP.PS.DV^CS^TG	0.11	DC.SP.DM.DV^TG^[…]QF	0.19
DC.NSP.PT.SL^EP.EX^CS^TG	0.11	DC.SP.DM.DV^UNC^TG^[…]^QF	0.19
DC.NSP.PT.SL^EP.IP^TG^QF	0.11	DC.SP.DM.IV^FC.EL.EM^TG^[…]^QF	0.19
DC.NSP.PT.SL^NUM.DF.OR^FC.ET.PV^TG	0.11	DC.SP.DM.IV^TG	0.19
DC.NSP.PT.SL^PDC.EL.EM.RT^TG^QF	0.11	DC.SP.DM.IV^TG^QF	0.19
DC.NSP.PT.SL^PDC.EL.ID^TG^QF	0.11	DC.SP.PS.DV^EP.IP^TG^[…]QF	0.19
DC.NSP.TL.NG^EP.EX^CS^TG	0.11	DC.SP.PS.DV^EP.IP^TG^QF	0.19
DC.NSP.TL.NG^EP.IP^TG	0.11	DC.SP.PS.DV^FC.EL.EM^CS^TG	0.19
DC.NSP.TL.NG^PDC.EL.EM.AM^TG	0.11	DC.SP.PS.DV^PDC.EL.ID.RT^TG	0.19
DC.NSP.TL.NG^PDC.EL.EM.RT^TG	0.11	DC.SP.PS.DV^PDC.ET.AM^TG^QF	0.19
DC.NSP.TL.NG^PDC.EL.ID^TG	0.11	DC.SP.PS.IV^TG^QF	0.19
DC.NSP.TL.NG^PDC.EL.ID^TG^QF	0.11	EP.EX^CS^TG^QF	0.19
DC.NSP.TL.NG^PDC.EN.SPA-TM^TG	0.11	EP.EX^TG^[…]QF	0.19
DC.NSP.TL.NG^PDC.RP.LN^TG^QF	0.11	EP.IP^TG	0.19
DC.NSP.TL.POS^DC.SP.DM.DV^EP.EX^TG^QF	0.11	FC.ET.AGG^TG	0.19
DC.NSP.TL.POS^DC.SP.DM.DV^PDC.EN.SPA-TM^NUM.DF.QV^TG	0.11	FC.ET.PV^DC.NSP.TL.POS^DC.SP.PS.DV^TG	0.19
DC.NSP.TL.POS^DC.SP.DM.DV^TG^QF	0.11	FC.ET.PV^DC.SP.DM.DV^TG	0.19
DC.NSP.TL.POS^DC.SP.PS.DV^EP.IP^TG	0.11	IF^DC.NSP.PT.NSL^EP.EX^TG	0.19
DC.NSP.TL.POS^FC.EL.EM^PDC.EL.ID^TG	0.11	IF^DC.NSP.PT.NSL^TG	0.19
DC.NSP.TL.POS^FC.EL.EM^TG	0.11	IF^NUM.IN.QV^TG^QF	0.19
DC.SP.DM.DV^[…]TG^[…]QF	0.11	NUM.DF.QV^CS^TG	0.19

DC.SP.DM.DV^CS^TG^QF^QF	0.11	NUM.DF.QV^EP.EX^TG	0.19
DC.SP.DM.DV^DC.NSP.PT.NSL^TG	0.11	NUM.DF.QV^EP.IP^TG	0.19
DC.SP.DM.DV^EP.EX^[…]CS^CS^TG	0.11	NUM.DF.QV^EP.IP^TG^QF	0.19
DC.SP.DM.DV^EP.IP^[…] DC.SP.PS.DV^TG	0.11	NUM.DF.QV^PDC.MD.PB^TG^QF	0.19
DC.SP.DM.DV^EP.IP^CS^TG	0.11	NUM.DF.QV^TG^QF^QF	0.19
DC.SP.DM.DV^EP.IP^EP.EX^TG	0.11	NUM.IN.QV.IV^TG	0.19
DC.SP.DM.DV^FC.EL.EM^CS^TG^QF	0.11	NUM.IN.QV^CS^TG	0.19
DC.SP.DM.DV^FC.EL.EM^TG^QF	0.11	NUM.IN.QV^DC.SP.DM.DV^TG	0.19
DC.SP.DM.DV^NUM.DF.OR^CS^TG^QF	0.11	NUM.IN.QV^EP.IP^TG	0.19
DC.SP.DM.DV^NUM.DF.OR^ DC.SP.DM.DV^TG	0.11	NUM.IN.QV^EP.IP^TG^QF	0.19
DC.SP.DM.DV^NUM.DF.OR^ PDC.MD.PB^EP.EX^TG^QF	0.11	NUM.IN.QV^PDC.EL.ID^TG^QF	0.19
DC.SP.DM.DV^NUM.DF.OR^TG	0.11	NUM.IN.QV^TG^QF^QF	0.19
DC.SP.DM.DV^NUM.DF.OR^TG^QF^QF	0.11	PDC.EL.EM.RT^TG	0.19
DC.SP.DM.DV^NUM.DF.QV^CS^CS^TG^QF	0.11	PDC.EL.EM.RT^TG^QF	0.19
DC.SP.DM.DV^NUM.DF.QV^CS^TG^QF	0.11	PDC.EL.EM^PDC.EL.ID^TG	0.19
DC.SP.DM.DV^NUM.DF.QV^ PDC.EL.EM^TG^QF	0.11	PDC.EL.ID.RT^TG	0.19
DC.SP.DM.DV^PDC.EL.EM.EZ^(NG)TG	0.11	PDC.EL.ID^TG^QF^QF	0.19
DC.SP.DM.DV^PDC.EL.EM.EZ^TG	0.11	PDC.EN.CV^TG	0.19
DC.SP.DM.DV^PDC.EL.EM.EZ^TG^QF	0.11	PDC.ET.AM^TG	0.19
DC.SP.DM.DV^PDC.EL.EM.RT^TG^QF	0.11	PDC.ET.EZ^EP.IP^TG	0.19
DC.SP.DM.DV^PDC.EL.EM^TG^QF	0.11	PDC.MD.US^TG	0.19
DC.SP.DM.DV^PDC.EL.ID^EP.IP^TG^QF	0.11		
DC.SP.DM.DV^PDC.EN.SPA-TM^CS^TG^QF	0.11		
DC.SP.DM.DV^PDC.EN.SPA-TM^PDC.EL.ID^[…]TG^QF^[…]^QF	0.11		
DC.SP.DM.DV^PDC.EN.SPA-TM^TG^QF	0.11		
DC.SP.DM.DV^PDC.ET.AM^TG	0.11		
DC.SP.DM.DV^PDC.ET.EZ^TG^QF	0.11		
DC.SP.DM.DV^PDC.ET.RT^TG	0.11		
DC.SP.DM.DV^PDC.MD.OB^TG	0.11		
DC.SP.DM.DV^PDC.MD.US^CS^TG^QF	0.11		
DC.SP.DM.DV^PDC.MD.US^TG^QF	0.11		
DC.SP.DM.DV^PDC.RP.IA.EZ^CS^TG^QF	0.11		
DC.SP.DM.DV^TG[…]^QF	0.11		
DC.SP.DM.IV^[…]TG	0.11		
DC.SP.DM.IV^TG	0.11		
DC.SP.PS.DV^[…]^EP.IP^EP.IP^TG	0.11		
DC.SP.PS.DV^[…]TG^QF	0.11		

DC.SP.PS.DV^CS^[…]TG	0.11
DC.SP.PS.DV^CS^CS^CS^TG	0.11
DC.SP.PS.DV^CS^CS^TG	0.11
DC.SP.PS.DV^CS^CS^TG^QF	0.11
DC.SP.PS.DV^EP.EX^TG	0.11
DC.SP.PS.DV^EP.EX^TG^QF	0.11
DC.SP.PS.DV^EP.IP^TG^QF	0.11
DC.SP.PS.DV^NUM.DF.OR^CS^TG	0.11
DC.SP.PS.DV^NUM.IN.QV^TG	0.11
DC.SP.PS.DV^PDC.EL.ID.EZ^TG	0.11
DC.SP.PS.DV^PDC.EL.ID^TG	0.11
DC.SP.PS.DV^PDC.EN.SPA-TM^TG	0.11
DC.SP.PS.DV^PDC.MD.RD^TG	0.11
DC.SP.PS.DV^TG^[…]QF	0.11
DC.SP.PS.IV^PDC.EL.ID.RT^TG	0.11
DC.SP.PS.IV^TG^QF	0.11
EP.EX^[…]TG	0.11
EP.EX^CS^CS^TG	0.11
EP.EX^CS^TG^[…]QF	0.11
EP.EX^CS^TG^QF^QF	0.11
EP.EX^EP.EX^TG	0.11
EP.IP^DC.NSP.PT.NSL^TG	0.11
EP.IP^EP.EX^TG	0.11
EP.IP^EP.IP^TG	0.11
EP.IP^TG^[…]^QF	0.11
FC.ET.AGG^TG^QF	0.11
NUM.DF.QV^NUM.DF.QV^CS^TG	0.11
NUM.DF.QV^NUM.DF.QV^ PDC.EL.ID^TG	0.11
NUM.DF.QV^NUM.DF.QV^TG^QF^ QF	0.11
NUM.DF.QV^PDC.EL.ID.RT^TG	0.11
NUM.DF.QV^PDC.EL.ID^TG	0.11
NUM.DF.QV^PDC.EL.ID^TG^QF	0.11
NUM.DF.QV^TG^[…]^QF	0.11
NUM.DF.QV^TG^QF^QF	0.11
NUM.IN.QV^CS^TG	0.11
NUM.IN.QV^CS^TG^QF	0.11
NUM.IN.QV^DC.SP.DM.DV^TG	0.11
NUM.IN.QV^DC.SP.PS.DV^TG^QF	0.11
NUM.IN.QV^EP.EX^TG	0.11
NUM.IN.QV^EP.EX^TG^QF^QF	0.11
NUM.IN.QV^PDC.EL.ID^TG^QF	0.11
NUM.IN.QV^TG^QF^QF	0.11
PDC.EL.EM.RT^[…]TG	0.11
PDC.EL.EM.RT^TG^QF	0.11
PDC.EL.EM^TG	0.11
PDC.EL.EM^TG^QF	0.11
PDC.EL.ID^CS^TG^QF	0.11
PDC.EL.ID^EP.EX^CS^TG	0.11
PDC.EL.ID^EP.IP^TG	0.11
PDC.EN.CV^PDC.EN.CV^TG	0.11
PDC.EN.SPA-TM^CS^TG	0.11
PDC.EN.SPA-TM^TG^QF	0.11
PDC.ET.AM^TG^QF	0.11
PDC.RP.IA.EZ^TG	0.11

TG^QF^[...]^QF	0.11

6.2 *Formal patterns*

Written	%	Spoken	%
H	6.00	DF.AR^H	8.77
DF.AR^H	5.03	IN.AR^H	5.26
AJ^H	4.39	H	3.70
DF.AR^H^PP(of n)	4.07	DF.AR^H^PP(of n)	3.12
H^PP(of n)	2.89	IN.AR^AJ^H	2.92
DF.AR^AJ^H	2.36	DF.AR^H^RT.RV.CL	2.73
IN.AR^AJ^H	2.14	PS.DT^H	2.53
PS.DT^H	2.14	DF.AR^AJ^H	2.34
DM.DT(THIS)^H	1.82	DM.DT(THAT)^H	2.34
IN.AR^H	1.82	AJ^H	2.14
DF.AR^AJ^H^PP(of n)	1.71	H^PP(of n)	2.14
DM.DT(THESE)^H	1.50	QT^H	1.95
IN.AR^H^PP(of n)	1.18	DM.DT(THIS)^H	1.17
N^H	1.18	DF.DV.GV.NP^H	0.97
PS.DT^AJ^H	1.07	IN.AR^H^PP(of n)	0.97
NUM.CD^H	0.86	NUM.CD^H	0.97
DF.AR^H^RT.RV.CL	0.75	AS.DT^H	0.78
DF.AR^N^H	0.75	DF.AR^AJ^H^RT.RV.CL	0.78
DF.DV.GV.NP^H	0.75	NAS.DT^H	0.78
DF.DV.GV.NP^H^PP(of n)	0.75	DF.AR^H^PP(to n)	0.58
H^RT.RV.CL	0.75	DF.AR^NP^H	0.58
NG.DT^AJ^H	0.75	DF.AR^QL.PV^H^RT.RV.CL	0.58
NUM.GO^H	0.75	DM.DT(THESE)^H	0.58
QT^H	0.75	H^RT.RV.CL(when)	0.58
AS.DT^H	0.64	IN.AR^H^AP.TI.CL	0.58
DF.AR^H^PP(of -ing)	0.64	IN.AR^H^PP(of -ing)	0.58
DF.DV.GV.NP^AJ^H	0.64	IN.AR^H^RT.RV.CL(where)	0.58
H^PP(for n)	0.64	N^H	0.58
IN.AR^AJ^H^PP(of n)	0.64	NP^H	0.58
IN.AR^N^H	0.64	NUM.CD^PV.PP(of DF.AR^AJ^H^RT.RV.CL)	0.58
NG.DT^H	0.64	NUM.CD^PV.PP(of DF.AR^H^RT.RV.CL)	0.58
AJ^AJ^H	0.54	PS.DT^AJ^H	0.58
AJ^H^PP(in n)	0.54	AJ^AJ^H	0.39
DF.AR^H^PP(of n)^AP.TI.CL	0.54	AJP^H	0.39
H^AP.TI.CL	0.54	AS.DT^AJ^H	0.39
IN.AR^AJ^H^PP(to n)	0.54	AS.DT^QL.PV^H	0.39
IN.AR^AJP^H	0.54	CV.QT^H	0.39
PDT^H	0.54	DF.AR^AJ^H^PP(for n)	0.39
DF.AR^H^AP.TI.CL	0.43	DF.AR^AJ^H^PP(of -ing)	0.39
DF.AR^H^PL.ED.CL	0.43	DF.AR^AJ^H^PP(of n)	0.39
DF.DV.GV.NP^H^AP.TI.CL	0.43	DF.AR^H^PL.ED.CL	0.39
DM.DT(THIS)^AJ^H	0.43	DF.AR^H^PP(of -ing)	0.39
H^TI.CL	0.43	DF.AR^H^PP(on n)	0.39
IN.AR^H^AP.TI.CL	0.43	DF.AR^H^RT.RV.CL^PP(of n)	0.39
PS.DT^H^AP.TI.CL	0.43	DF.AR^NUM.CD^H	0.39
PS.DT^H^PP(for n)	0.43	DF.AR^NUM.GO^H	0.39
UV.DT^H	0.43	DF.AR^NUM.OR^H^RT.RV.CL	0.39
AJ^H^PP(of n)	0.32	H^PP(of -ing)	0.39

DF.AR^AJP^H	0.32	H^PP(to n)	0.39
DF.AR^H^AP.THAT.CL	0.32	H^TI.CL	0.39
DF.AR^H^PP(for n)	0.32	IN.AR^H^PP(with n)	0.39
H^PP(in n)	0.32	IN.AR^H^PT.WK.NR.IT.AP	0.39
H^PP(to n)	0.32	IN.AR^N^H	0.39
IN.AR^AJ^H^AP.TI.CL	0.32	NUM.CD^AJ^H	0.39
IN.AR^AJ^H^PL.ED.CL	0.32	NUM.CD^PV.PP(of DF.AR^H)	0.39
IN.AR^H^PL.ED.CL	0.32	NUM.GO^H	0.39
IN.AR^H^PP(for n)	0.32	NUM.GO^H^PP(of n)	0.39
IN.AR^H^RT.RV.CL	0.32	PDT^H^RT.RV.CL(as)	0.39
NAS.DT^H	0.32	PS.DT^H^PP(for n)	0.39
NP^H	0.32	QT^AJ^H	0.39
PS.DT^H^PP(of n)	0.32	QT^H^RT.RV.CL	0.39
PS.DT^N^H	0.32	QT^IN.AR^H	0.39
QT^H^PP(of -ing)	0.32	UV.PDT^DF.AR^H	0.39
AJ^[…]H	0.21	(ING)N^H	0.19
AJ^H^PL.ED.CL	0.21	(ING)N^H^PP(for n)	0.19
AJ^H^PP(about n)	0.21	AJ^H^[…]PP(of n)	0.19
AS.DT^AJ^H	0.21	AJ^H^AP.TI.CL	0.19
DF.AR^[…]H^PP(of n)	0.21	AJ^H^PP(about n)	0.19
DF.AR^AJ^AJ^H	0.21	AJ^H^PP(of n)	0.19
DF.AR^AJ^AJ^H^PP(of n)	0.21	AJ^H^PP(with n -ing)	0.19
DF.AR^AJ^H^PP(of -ing)	0.21	AJ^H^RT.RV.CL	0.19
DF.AR^CV.AJ^H^PP(of n)	0.21	AJ^N^H^RT.RV.CL	0.19
DF.AR^H^AV	0.21	AS.DT^AJP^AJP^H^PP(in n)	0.19
DF.AR^H^PP(of n -ing)	0.21	AS.DT^H^PP(in n)	0.19
DF.AR^NP^H	0.21	AS.DT^H^PP(of n)	0.19
DF.AR^NUM.CD^AJ^H^PP(of n)	0.21	AS.DT^H^PP(to n)	0.19
DF.DV.GV.NP^H^PP(for n)	0.21	AS.DT^N^H	0.19
DF.DV.GV.NP^N^H	0.21	AS.PN^PV.PP(of DF.AR[…]^AJ^(ING)N^H)	0.19
DF.DV.GV.NP^NUM.CD^H	0.21	AS.PN^PV.PP(of DF.AR^H^PP(in which))	0.19
DM.DT(THAT)^H^PP(of n)	0.21	CV.QT^CV.AJP^H^PP(in n)	0.19
DM.DT(THIS)^QL.PV^H	0.21	CV.QT^H[…]^RT.RV.CL	0.19
H^PL.ED.CL	0.21	DF.AR^(ING)N^H	0.19
H^PP(from n)	0.21	DF.AR^[…]H^PP(of wh)	0.19
H^PP(with n)	0.21	DF.AR^[…]H^RT.RV.CL	0.19
IN.AR^AJ^AJ^H	0.21	DF.AR^AJ^AJ^H	0.19
IN.AR^AJ^H^AP.THAT.CL	0.21	DF.AR^AJ^AJ^H^PP(in -ing)	0.19
IN.AR^AJ^H^PP(between n)	0.21	DF.AR^AJ^AJ^H^PP(of n)	0.19
IN.AR^AJ^H^PP(for n)	0.21	DF.AR^AJ^H^(for n TI.CL)	0.19
IN.AR^AJ^H^PP(for which)	0.21	DF.AR^AJ^H^PP(about n)	0.19
IN.AR^AJ^H^PP(of n -ing)	0.21	DF.AR^AJ^H^PP(from within n)	0.19
IN.AR^AJ^N^H	0.21	DF.AR^AJ^H^PP(in n)	0.19
IN.AR^H^AP.THAT.CL	0.21	DF.AR^AJP^H^FL.WK.NR.IT.AP	0.19
IN.AR^NP^H	0.21	DF.AR^H^(for n AP.TI.CL)	0.19
NAS.DT^AJ^H	0.21	DF.AR^H^[…]PP(to n)	0.19
NUM.CD^AJ^H	0.21	DF.AR^H^AP.THAT.CL	0.19
NUM.CD^H^PP(of -ing)	0.21	DF.AR^H^AP.TI.CL	0.19
NUM.CD^H^PP(of n)	0.21	DF.AR^H^FL.ST.RT.AP	0.19
NUM.CD^PV.PP(of NUM.CD^AJ^H)	0.21	DF.AR^H^PP(behind n -ing)	0.19
PS.DT^AJ^H^PP(in n)	0.21	DF.AR^H^PP(in n)^AP.THAT.CL	0.19
PS.DT^H^PP(of -ing)	0.21	DF.AR^H^PP(of n -ing)	0.19
QT^AJ^H	0.21	DF.AR^H^PP(of n)^AP.THAT.CL	0.19
RV.DT^H	0.21	DF.AR^H^PP(of n)^PP(about n)	0.19

(ING)AJ^H	0.11	DF.AR^H^PP(of n)^PP(in n)	0.19
(ING)N^(ING)N^H	0.11	DF.AR^H^PP(of ps.dt -ing)	0.19
(ING)N^H	0.11	DF.AR^NP^[…]AJ^H^RT.RV.CL	0.19
(NG)H	0.11	DF.AR^NUM.CD^AJ^H^RT.RV.CL	0.19
AJ^AJ^AJ^H	0.11	DF.AR^NUM.CD^NUM.GO^H^ RT.RV.CL	0.19
AJ^AJP^H	0.11	DF.AR^NUM.GO^H^PP(in which)	0.19
AJ^H^[…]^AP.THAT.CL	0.11	DF.AR^NUM.GO^H^RT.RV.CL	0.19
AJ^H^FL.ST.RT.EM.AP	0.11	DF.AR^QL.PV^AJ^[…]^H^PP(of n)	0.19
AJ^H^NR.RV.CL(where)	0.11	DF.AR^QL.PV^H^RT.RV.CL(where)	0.19
AJ^H^PP(against n)	0.11	DF.AR^UNC^H^[…]^RT.RV.CL	0.19
AJ^H^PP(for n)	0.11	DF.DV.GV.NP^H^PP(of n)	0.19
AJ^H^PP(from n)^AP.THAT.CL	0.11	DF.DV.GV.NP^N^H	0.19
AJ^H^PP(in n)^PT.WK.NR.PA.AP	0.11	DM.DT(THAT)^AJ^H	0.19
AJ^H^PP(in which)	0.11	DM.DT(THAT)^H^PP(of n)	0.19
AJ^H^PP(of n -ing)	0.11	DM.DT(THAT)^QL.PV^H	0.19
AJ^H^PP(on n)	0.11	DM.DT(THESE)^AJ^H	0.19
AJ^H^PP(to n)	0.11	DM.DT(THESE)^QL.PV^AJ^H	0.19
AJ^H^PP(with n)	0.11	DM.DT(THIS)^(ING)N^H	0.19
AJ^H^PT.ST.NR.EM.AP	0.11	DM.DT(THIS)^AJ^H	0.19
AJ^H^PT.WK.NR.EM.AP	0.11	DM.DT(THIS)^H^PP(of n)	0.19
AJ^H^RT.RV.CL	0.11	DM.DT(THIS)^N^H	0.19
AJ^N^AJ^H	0.11	DM.DT(THIS)^QL.PV^AJ^H^RT.RV.CL	0.19
AJ^N^H	0.11	H^(for n AP.TI.CL)	0.19
AJ^N^H^[…]RT.RV.CL	0.11	H^[…]^PP(of n)	0.19
AJ^N^H^PT.ST.NR.EM.AP^ NR.RV.CL	0.11	H^[…]AP.TI.CL	0.19
AJP^H	0.11	H^[…]PP(in n)	0.19
AJP^IN.AR^H	0.11	H^[…]PP(of -ing)	0.19
AS.DT^CV.QT^H^TI.CL	0.11	H^AP.THAT.CL	0.19
AS.DT^H^PP(for n)	0.11	H^AP.TI.CL	0.19
AS.PN^PV.PP(of DF.DV.GV.NP^H^PP of n)	0.11	H^CV.AJP	0.19
AS.PN^PV.PP(of DM.DT(THESE)^H)	0.11	H^FL.ST.RT.EM.AP	0.19
CV.QT.PN^PV.PP(of DF.AR^H^ PP of n)	0.11	H^PP(about n -ing)	0.19
CV.QT^H	0.11	H^PP(about n)	0.19
CV.QT^H^PP(on n)	0.11	H^PP(against n)	0.19
DF.AR^[…]H^[…]PP(of n)	0.11	H^PP(for n)	0.19
DF.AR^[…]H^AP.TI.CL	0.11	H^RT.RV.CL	0.19
DF.AR^AJ^(NG)H	0.11	IF.AV^IN.AR^AJ^H	0.19
DF.AR^AJ^[…]AJ^AJ^H	0.11	IF.AV^IN.AR^H	0.19
DF.AR^AJ^AJ^H^PL.ING.CL	0.11	IF.AV^QT^H^PP(of n)	0.19
DF.AR^AJ^H^AJP	0.11	IN.AR^(ING)N^H	0.19
DF.AR^AJ^H^AP.THAT.CL	0.11	IN.AR^AJ^H^[…]PP(on n)	0.19
DF.AR^AJ^H^AP.TI.CL	0.11	IN.AR^AJ^H^AP.THAT.CL	0.19
DF.AR^AJ^H^PP(about n)	0.11	IN.AR^AJ^H^PP(for n)	0.19
DF.AR^AJ^H^PP(between n)	0.11	IN.AR^AJ^H^PP(of ?)	0.19
DF.AR^AJ^H^PP(for n)	0.11	IN.AR^AJ^H^PP(of n)	0.19
DF.AR^AJ^H^PP(of n)^PP(for n)	0.11	IN.AR^AJ^H^PP(with n)	0.19
DF.AR^AJ^H^PP(of n)^RT.RV.CL	0.11	IN.AR^AJ^H^PT.WK.NR.IT.AP	0.19
DF.AR^AJ^H^PP(over n)	0.11	IN.AR^AJP^H	0.19
DF.AR^AJ^H^PP(to n)	0.11	IN.AR^AJP^N^H	0.19
DF.AR^AJ^H^PT.WK.NR.AP	0.11	IN.AR^H^[…]PP(about n)	0.19
DF.AR^AJ^H^RT.RV.CL	0.11	IN.AR^H^AP.THAT.CL	0.19
DF.AR^AJ^H^RT.RV.CL^PP(with n)	0.11	IN.AR^H^PP(about n)	0.19

DF.AR^AJ^H^TI.CL	0.11	IN.AR^H^PP(of n)^NR.RV.CL	0.19
DF.AR^AJ^N^H	0.11	IN.AR^H^PP(of wh)	0.19
DF.AR^AJ^N^H^NR.RV.CL	0.11	IN.AR^H^PT.ST.NR.IT.AP	0.19
DF.AR^AJ^N^H^RT.RV.CL	0.11	IN.AR^H^RT.RV.CL	0.19
DF.AR^AJ^NUM.GO^[…]H^ PP(of n)^[…]^PT.ST.NR.IT.AP	0.11	IN.AR^H^TI.CL	0.19
DF.AR^AJP^H^PP(of n)	0.11	IN.AR^N^[…]H^PP(of n)	0.19
DF.AR^CV.AJ^H	0.11	IN.AR^NP^[…]^AJP^H	0.19
DF.AR^CV.AJP^PP(of […] DF.DV.GV.NP^H)	0.11	IN.AR^NP^H	0.19
DF.AR^H[…]^PP(of n)	0.11	IN.AR^NP^H^RT.RV.CL(whereby)	0.19
DF.AR^H^FL.ST.RT.AP	0.11	IV.DT^H	0.19
DF.AR^H^PL.ING.CL	0.11	IV.DT^H^RT.TV.CL	0.19
DF.AR^H^PP(between n)	0.11	IV.DT^QL.PV^H^[…]^PP(about n)	0.19
DF.AR^H^PP(for -ing)	0.11	MR.GV^H^PP(with n)	0.19
DF.AR^H^PP(from n)	0.11	NAS.DT^AJ^H	0.19
DF.AR^H^PP(of aj)	0.11	NAS.DT^CV.QT^H	0.19
DF.AR^H^PP(of n)^AP.THAT.CL	0.11	NAS.DT^H[…]^RT.RV.CL	0.19
DF.AR^H^PP(of n)^NR.RV.CL	0.11	NAS.DT^H^PP(between n)	0.19
DF.AR^H^PP(of n)^PP(in n)	0.11	NAS.DT^H^PP(in -ing)	0.19
DF.AR^H^PP(of n)^PP(on n)	0.11	NAS.DT^H^PP(of n)^PP(at n)	0.19
DF.AR^H^PP(to n)	0.11	NAS.DT^H^RT.RV.CL(when)	0.19
DF.AR^H^PP(with n)	0.11	NAS.DT^NUM.GO^H^PP(of n)	0.19
DF.AR^H^PP(with which)	0.11	NAS.DT^QL.PV^H^RT.RV.CL	0.19
DF.AR^H^RT.RV.CL(when)	0.11	NG.DT^H^AP.TI.CL	0.19
DF.AR^H^RT.RV.CL(where)	0.11	NG.DT^H^PP(of n -ing)	0.19
DF.AR^H^RT.RV.CL^ PT.ST.NR.IT.AP	0.11	NG.DT^H^PP(of n)	0.19
DF.AR^H^TI.CL	0.11	NG.DT^H^PP(with n)	0.19
DF.AR^N^H^AP.THAT.CL	0.11	NG.DT^N^H	0.19
DF.AR^N^H^PP(of n)	0.11	NUM.CD(QT)^H	0.19
DF.AR^N^H^PT.ST.NR.ATT.AP	0.11	NUM.CD(QT)^H^[…]PP(about n)	0.19
DF.AR^NP^AJ^H	0.11	NUM.CD(QT)^NUM.GO^H^PP(on n)	0.19
DF.AR^NUM.CD^N^N^H^RT.RV.CL	0.11	NUM.CD^AJ^H^PT.ST.NR.IT.AP	0.19
DF.AR^NUM.GO^AJ^H^TI.CL	0.11	NUM.CD^AJP^H	0.19
DF.AR^NUM.GO^N^H^ FL.ST.NR.IT.AP	0.11	NUM.CD^AJP^H^PP(in terms of n)	0.19
DF.AR^NUM.OR^AJ^AJ^H^PP(of n)	0.11	NUM.CD^H^FL.ST.NR.IT.AP	0.19
DF.AR^NUM.OR^H	0.11	NUM.CD^H^PP(of n)^RT.RV.CL	0.19
DF.AR^NUM.OR^H^PP(of n)	0.11	NUM.CD^H^PT.WK.NR.IT.AP	0.19
DF.AR^NUM.OR^H^RT.RV.CL	0.11	NUM.CD^N^H	0.19
DF.AR^NUM.OR^H^TI.CL^PP(with n)	0.11	NUM.CD^NUM.GO^H^PL.ED.CL	0.19
DF.AR^NUM.OR^PV.PP(of DM.DT(THESE)^H)	0.11	NUM.CD^PV.PP(of DF.AR^(ING)N^H)	0.19
DF.AR^QL.PV^AJ^H^RT.RV.CL	0.11	NUM.CD^PV.PP(of DF.AR^AJ^H)	0.19
DF.AR^QL.PV^H^RT.RV.CL	0.11	NUM.CD^PV.PP(of DF.AR^AJ^H^PP(about n))	0.19
DF.DV.GV.NP^[…]H	0.11	NUM.CD^PV.PP(of DF.AR^NUM.GO^H^RT.RV.CL)	0.19
DF.DV.GV.NP^AJ^[…]H	0.11	NUM.CD^PV.PP(of DM.DT(THOSE)^AJ^H^RT.RV.CL)	0.19
DF.DV.GV.NP^AJ^AJ^H	0.11	NUM.CD^PV.PP(of PS.DT^[…]H)	0.19
DF.DV.GV.NP^H^[…]AP.TI.CL	0.11	NUM.CD^PV.PP(of PS.DT^H^PP(behind -ing))	0.19
DF.DV.GV.NP^H^PP(of -ing)	0.11	NUM.CD^PV.PP(of PS.DT^H^ PP(since n))	0.19

DF.DV.GV.NP^H^PP(to n)	0.11	NUM.CD^QV.PV^H	0.19
DF.DV.GV.NP^NUM.GO^H	0.11	NUM.GO^H^PP(of -ing)	0.19
DM.DT(THAT)^H	0.11	NUM.GO^H^PP(to n)^RT.RV.CL	0.19
DM.DT(THAT)^H^AP.TI.CL	0.11	PS.DT^AJ^H^[…]AP.TI.CL	0.19
DM.DT(THAT)^H^ PP(of n)^FL.ST.NR.IT.AP	0.11	PS.DT^AJ^H^AP.TI.CL	0.19
DM.DT(THIS)^AJ^AJ^H	0.11	PS.DT^AJP^H^PP(of n)	0.19
DM.DT(THIS)^AJ^AJP^H	0.11	PS.DT^H[…]^PP(of wh)	0.19
DM.DT(THIS)^AJ^H^PP(against n)	0.11	PS.DT^H^[…]PP(for -ing)	0.19
DM.DT(THIS)^H^PP(in n)	0.11	PS.DT^H^AP.TI.CL	0.19
DM.DT(THIS)^H^PP(of n)	0.11	PS.DT^H^PL.ED.CL	0.19
DM.DT(THIS)^H^PP(of n)^PP(over n)	0.11	PS.DT^H^PP(as n)	0.19
DM.DT(THIS)^NUM.GO^H	0.11	PS.DT^H^PP(behind -ing)	0.19
DM.DT(THOSE)^AJ^H	0.11	PS.DT^H^PP(in -ing)	0.19
DM.DT(THOSE)^H	0.11	PS.DT^H^PP(of n)	0.19
DM.DT(THOSE)^H^PP(of n)^PL.ING.CL	0.11	PS.DT^N^H	0.19
H^PP(among n)^AP.THAT.CL	0.11	PS.DT^QL.PV^AJ^H	0.19
H^PP(at n)	0.11	QT.IV^H	0.19
H^PP(at n)^PP(in n)	0.11	QT^DF.AR^H	0.19
H^PP(by n)	0.11	QT^DS.GV^H	0.19
H^PP(for n)^PL.ED.CL	0.11	QT^H^PL.ING.CL	0.19
H^PP(for wh)	0.11	QT^H^PL.ING.CL^PP(as to wh)	0.19
H^PP(of n -ing)	0.11	QT^H^PP(from n)	0.19
H^PP(of n)^[…]^PT.WK.NR.IT.AP	0.11	QT^H^PP(of -ing)	0.19
H^PP(of n)^AP.TI.CL	0.11	QT^H^TI.CL	0.19
H^PP(of n)^PP(of n)	0.11	QV.PV^DM.DT(THAT)^H	0.19
H^PP(of n)^RT.RV.CL	0.11	QV.PV^H	0.19
H^PP(on n)	0.11	QV.PV^UV.PDT^PS.DT^H	0.19
H^PP(other than n)	0.11	RV.DT^H^PP(of -ing)	0.19
H^RT.RV.CL(when)	0.11	UV.DT^H^RT.RV.CL	0.19
IN.AR^(ING)N^H^AP.TI.CL	0.11	UV.PDT^DF.AR^AJ^H^PP(about n)	0.19
IN.AR^(ING)N^H^PP(at n)	0.11	UV.PDT^DF.AR^H^PL.ED.CL	0.19
IN.AR^[…]^H	0.11	UV.PDT^DF.AR^H^PP(on n)	0.19
IN.AR^[…]AJ^H	0.11	UV.PDT^DM.DT(THAT)^H	0.19
IN.AR^[…]H^PP(of n)	0.11	UV.PDT^H	0.19
IN.AR^AJ^AJ^H^AJP	0.11	UV.PDT^NUM.GO^H^PP(of n)	0.19
IN.AR^AJ^H[…]^AP.THAT.CL	0.11	UV.PDT^QL.PV^H	0.19
IN.AR^AJ^H^FL.ST.NR.IT.AP^ PL.ED.CL	0.11	UV.PN^PV.PP(of DM.DT(THESE)^H)	0.19
IN.AR^AJ^H^NR.AP.THAT.CL	0.11		
IN.AR^AJ^H^NR.PL.ED.CL	0.11		
IN.AR^AJ^H^NR.PL.ING.CL	0.11		
IN.AR^AJ^H^NR.RV.CL	0.11		
IN.AR^AJ^H^PP(for -ing)	0.11		
IN.AR^AJ^H^PP(of n)^PP(in which)	0.11		
IN.AR^AJ^H^PP(over n)	0.11		
IN.AR^AJ^H^PT.WK.NR.IT.AP	0.11		
IN.AR^AJ^H^RT.RV.CL	0.11		
IN.AR^AJ^N^H^PP(on n)	0.11		
IN.AR^AJP^AJ^H	0.11		
IN.AR^AJP^H^NR.PL.ED.CL^ AP.THAT.CL	0.11		
IN.AR^AJP^H^NR.RV.CL	0.11		
IN.AR^AJP^H^PP(by n)^PP(of n)	0.11		
IN.AR^AJP^H^PP(of n)	0.11		

IN.AR^CV.AJ^AJ^H^AP.TI.CL	0.11
IN.AR^CV.AJ^H^PP(in which)	0.11
IN.AR^CV.AJ^H^PP(of n)	0.11
IN.AR^CV.AJP^N^H	0.11
IN.AR^H^[…]PL.ING.CL	0.11
IN.AR^H^[…]PP(of n)	0.11
IN.AR^H^PP(about n)	0.11
IN.AR^H^PP(like n)	0.11
IN.AR^H^PP(of -ing)^RT.RV.CL	0.11
IN.AR^H^PP(of n)^RT.RV.CL	0.11
IN.AR^H^PP(to n)	0.11
IN.AR^N^AJ^H	0.11
IN.AR^N^H^PP(against n)	0.11
IN.AR^N^H^PP(of which)	0.11
IN.AR^NUM.GO^H	0.11
IV.DT^[…]H	0.11
IV.DT^H	0.11
N^H^AJP	0.11
N^H^FL.ST.NR.RF.AP	0.11
N^H^PT.ST.NR.EM.AP^RT.RV.CL	0.11
N^N^H	0.11
NAS.DT^AJ^H^PP(of n -ing)	0.11
NAS.DT^AJ^N^H	0.11
NAS.DT^H^AP.THAT.CL	0.11
NAS.DT^H^PP(as to n)	0.11
NAS.DT^H^PP(of n)	0.11
NAS.PN^PV.PP(of DF.AR^H^PL.ED.CL)	0.11
NAS.PN^PV.PP(of DF.AR^H^ PP in n)	0.11
NAS.PN^PV.PP(of DM.DT(THESE)^H)	0.11
NG.DT^AJ^AJ^H	0.11
NG.DT^AJ^H^RT.RV.CL	0.11
NG.DT^H^AP.THAT.CL	0.11
NG.DT^H^PP(for n)	0.11
NG.DT^H^PP(for which)	0.11
NG.DT^H^PP(of -ing)	0.11
NG.DT^H^PP(of n)	0.11
NG.DT^H^PP(on n)	0.11
NG.DT^H^TI.CL	0.11
NG.DT^NUM.GO^H	0.11
NG.DT^NUM.GO^H^TI.CL	0.11
NP^H^RT.RV.CL	0.11
NP^NP^H^PP(as to n)	0.11
NUM.CD^AJ^H^RT.RV.CL	0.11
NUM.CD^H^[…]^PP(of n)	0.11
NUM.CD^H^CV.AJP	0.11
NUM.CD^H^PP(of n)^NR.RV.CL	0.11
NUM.CD^H^PP(of wh)	0.11
NUM.CD^NUM.GO^H^PP(of n)	0.11
NUM.CD^NUM.GO^H^ RT.RV.CL(where)	0.11
NUM.CD^NUM.GO^QV.PV^H	0.11
NUM.CD^PV.PP(of DF.AR^AJ^AJ^H^TI.CL)	0.11

NUM.CD^PV.PP(of DF.AR^AJ^H^NR.PL.ING.CL)	0.11
NUM.CD^PV.PP(of NUM.CD^H^AV^PT.WK.NR.IT.AP)	0.11
NUM.CD^PV.PP(of PS.DT^N^H)	0.11
NUM.GO^AJ^AJ^H	0.11
NUM.GO^AJ^H	0.11
NUM.GO^AJ^H^PL.ED.CL	0.11
NUM.GO^H^PL.ED.CL	0.11
NUM.GO^H^PP(of -ing)	0.11
NUM.GO^H^PP(on n)	0.11
PDT(CV.RU)^IN.AR^AJ^H	0.11
PDT^AJ^H	0.11
PDT^IN.AR^H	0.11
PS.DT^(ING)N^H	0.11
PS.DT^[…]^AJ^AJ^H	0.11
PS.DT^[…]^H	0.11
PS.DT^[…]H^PP(on n)	0.11
PS.DT^AJ^AJ^AJ^H	0.11
PS.DT^AJ^AJ^H^PT.WK.NR.PA.AP	0.11
PS.DT^AJ^H^AP.THAT.CL	0.11
PS.DT^AJ^H^AP.TI.CL	0.11
PS.DT^H^AP.THAT.CL	0.11
PS.DT^H^NR.RV.CL	0.11
PS.DT^H^PP(as n)	0.11
PS.DT^H^PP(at -ing)	0.11
PS.DT^H^PP(during n)	0.11
PS.DT^H^PP(in -ing)	0.11
PS.DT^H^PP(in n)	0.11
PS.DT^H^PP(on n)	0.11
PS.DT^NUM.GO^H	0.11
PS.DT^NUM.OR^AJ^H	0.11
PS.DT^NUM.OR^H	0.11
PS.DT^NUM.OR^H^PP(into n)	0.11
PS.DT^NUM.OR^H^PP(of n)	0.11
PS.DT^NUM.OR^H^PP(over n)	0.11
PS.DT^QT^H	0.11
QT.PN^PV.PP(of DF.AR^H^PP of n)	0.11
QT.PN^PV.PP(of DF.AR^H^RT.RV.CL)	0.11
QT.PN^PV.PP(of DM.DT(THESE)^H)	0.11
QT^AJ^H^PP(for n)^PT.WK.NR.IT.AP	0.11
QT^AJ^H^PP(on n)	0.11
QT^H[…]^PP(of n)	0.11
QT^H^[…]^RT.RV.CL	0.11
QT^H^AP.THAT.CL	0.11
QT^H^FL.ST.NR.EM.AP	0.11
QT^H^FL.ST.NR.IT.AP	0.11
QT^H^PP(concerning n)	0.11
QT^H^PP(of n)	0.11
QT^H^PP(on n)	0.11
QT^H^PP(on n)^PT.ST.NR.EM.AP	0.11
QT^H^TI.CL	0.11
QT^NUM.GO^H^PP(in which)	0.11
QV.PV^H^PT.WK.NR.IT.AP	0.11

RV.DT^AJ^H	0.11
RV.DT^H^PP(of n)	0.11
UV.DT^H^PP(of n)	0.11
UV.DT^QL.PV^H	0.11
UV.PDT^DF.AR^H^RT.RV.CL	0.11
UV.PDT^DM.DT(THESE)^AJ^H^ NR.RV.CL	0.11
UV.PDT^H^PP(for n)	0.11
UV.PDT^H^PP(of n)	0.11
UV.PDT^PS.DT^AJ^H	0.11
UV.PDT^QL.PV^NUM.GO^H	0.11
UV.PN^PV.PP(of DF.AR^AJ^NUM.CD^H)	0.11

APPENDIX 7 GENRE DISTRIBUTION OF EXPERIENTIAL AND FORMAL PATTERNS (TOP TEN SUPER-GENRES: COMPLETE LISTS)

7.1 *Experiential patterns*

W:newsp	%	W:non_ac	%	S:meeting	%	W:ac	%
DC.SP.DM.DV^TG^QF	8.70	DC.SP.DM.DV^TG	12.80	DC.SP.DM.DV^TG	15.28	DC.SP.DM.DV^TG	13.57
TG^QF	8.70	TG	9.15	DC.SP.DM.DV^TG^QF	12.50	DC.SP.DM.DV^TG^QF	7.86
CS^TG	5.43	DC.SP.DM.DV^TG^QF	7.93	DC.SP.PS.DV^TG	5.56	TG^QF	5.00
DC.SP.DM.DV^TG	5.43	DC.SP.PS.DV^TG^QF	6.71	DC.NSP.PT.NSL^TG^QF	4.86	DC.SP.PS.DV^TG^QF	4.29
DC.SP.PS.DV^TG^QF	4.89	TG^QF	6.10	TG^QF	4.17	CS^TG	2.86
TG	4.89	DC.NSP.PT.NSL^TG^QF	4.27	DC.NSP.PT.NSL^TG	2.78	DC.NSP.PT.NSL^EP.EX^TG^QF	2.86
DC.NSP.PT.NSL^CS^TG	2.72	CS^TG	3.66	TG	2.78	DC.SP.PS.DV^TG	2.86
DC.SP.PS.DV^CS^TG	2.72	CS^TG^QF	3.05	DC.NSP.PT.SL^DC.SP.DM.DV^TG^QF	2.08	NUM.IN.QV^TG^QF	2.86
DC.SP.PS.DV^TG	2.72	DC.SP.DM.DV^CS^TG	3.05	DC.SP.DM.DV^PDC.EN.SPA-TM^TG	2.08	TG	2.86
DC.SP.DM.DV^CS^TG^QF	2.17	DC.SP.PS.DV^TG	2.44	DC.SP.DM.DV^TG^QF^QF	2.08	CS^TG^QF	2.14
DC.NSP.PT.NSL^CS^TG^QF	1.63	DC.NSP.PT.NSL^EP.IP^TG	1.83	DC.NSP.PT.NSL^CS^TG	1.39	DC.NSP.PT.NSL^TG^QF	2.14
DC.NSP.PT.NSL^EP.IP^TG^QF	1.63	DC.NSP.TL.NG^TG^QF	1.83	DC.NSP.PT.NSL^EP.IP^TG	1.39	PDC.EL.ID^TG	2.14
DC.NSP.PT.NSL^TG^QF	1.63	DC.SP.DM.DV^CS^TG^QF	1.83	DC.NSP.TL.NG^TG	1.39	DC.NSP.PT.NSL^CS^TG^QF	1.43
DC.SP.DM.DV^CS^TG	1.63	DC.SP.DM.DV^PDC.EL.ID.RT^TG^QF	1.83	DC.NSP.TL.POS^DC.SP.DM.DV^TG^QF	1.39	DC.NSP.PT.NSL^EP.EX^TG	1.43
DC.SP.DM.DV^EP.IP^TG	1.63	DC.SP.DM.DV^TG^QF^QF	1.83	DC.SP.DM.DV^CS^TG	1.39	DC.NSP.PT.NSL^EP.IP^TG^QF	1.43
DC.SP.DM.DV^TG^QF^QF	1.63	DC.NSP.PT.NSL^CS^CS^TG	1.22	DC.SP.DM.DV^EP.IP^TG	1.39	DC.SP.DM.DV^CS^TG	1.43
EP.EX^TG	1.63	DC.NSP.PT.NSL^CS^TG^QF	1.22	DC.SP.PS.DV^TG^QF	1.39	DC.SP.DM.DV^EP.EX^CS^TG	1.43
NUM.DF.QV^TG	1.63	DC.NSP.PT.NSL^EP.IP^TG^QF	1.22	NUM.DF.QV^TG^QF	1.39	DC.SP.DM.DV^EP.EX^TG	1.43
DC.NSP.PT.NSL^EP.EX^TG	1.09	DC.SP.DM.DV^EP.IP^TG	1.22	NUM.IN.QV^TG	1.39	DC.SP.DM.DV^PDC.EN.SPA-TM^TG	1.43
DC.NSP.TL.NG^CS^TG	1.09	DC.SP.DM.DV^PDC.EN.SPA-TM^TG	1.22	CS^TG	0.69	NUM.DF.QV^TG	1.43
DC.NSP.TL.NG^TG	1.09	EP.EX^TG^QF	1.22	DC.NSP.PT.NSL^CS^TG^QF	0.69	NUM.DF.QV^TG^QF	1.43
DC.SP.DM.DV^CS^CS^TG	1.09	EP.IP^TG^QF	1.22	DC.NSP.PT.NSL^EP.EX^TG	0.69	NUM.IN.QV^DC.SP.DM.DV^TG^QF	1.43
DC.SP.DM.DV^EP.IP^TG^QF	1.09	NUM.IN.QV^TG^QF	1.22	DC.NSP.PT.NSL^EP.EX^TG^[…]QF	0.69	PDC.EN.CV^TG	1.43
DC.SP.PS.DV^	1.09	PDC.EL.ID^TG^QF	1.22	DC.NSP.PT.NSL^	0.69	CS^CS^CS^TG	0.71

NUM.DF.OR^TG				EP.IP^EP.IP^TG^QF			
EP.IP^TG	1.09	CS^CS^TG	0.61	DC.NSP.PT.NSL^ EP.IP^TG^QF	0.69	DC.NSP.PT.NSL^CS^ TG	0.71
NUM.IN.QV^TG	1.09	DC.NSP.PT.NSL^ EP.EX^CS^TG^QF	0.61	DC.NSP.PT.NSL^ PDC.EL.EM.EZ^TG	0.69	DC.NSP.PT.NSL^ EP.EX^EP.EX^TG	0.71
CS^TG^QF	0.54	DC.NSP.PT.NSL^ EP.EX^TG	0.61	DC.NSP.PT.NSL^ PDC.EL.ID^TG	0.69	DC.NSP.PT.NSL^ EP.IP^TG	0.71
CS^TG^QF^QF	0.54	DC.NSP.PT.NSL^ EP.EX^TG^QF	0.61	DC.NSP.PT.NSL^TG^ […]QF	0.69	DC.NSP.PT.NSL^ PDC.EL.EM.AM^TG	0.71
DC.NSP.PT.NSL^ EP.EX^CS^TG	0.54	DC.NSP.PT.NSL^ EP.IP^TG^QF^QF	0.61	DC.NSP.PT.SL^ DC.SP.DM.DV^EP.IP^ TG^QF	0.69	DC.NSP.PT.NSL^ PDC.EL.EM.RT^TG^ QF	0.71
DC.NSP.PT.NSL^ EP.EX^CS^TG^QF	0.54	DC.NSP.PT.NSL^ PDC.EL.EM^TG	0.61	DC.NSP.PT.SL^ DC.SP.DM.DV^ PDC.EL.ID^TG^QF	0.69	DC.NSP.PT.NSL^ PDC.EL.ID.RT^TG	0.71
DC.NSP.PT.NSL^ EP.EX^TG[…]^QF	0.54	DC.NSP.PT.NSL^ PDC.EL.ID.EZ^TG^ QF	0.61	DC.NSP.PT.SL^ DC.SP.DM.DV^ PDC.MD.PB.EZ^TG^ QF	0.69	DC.NSP.PT.NSL^ PDC.EL.ID^TG	0.71
DC.NSP.PT.NSL^ EP.EX^TG^QF	0.54	DC.NSP.PT.NSL^ PDC.EN.CV^TG^QF ^QF	0.61	DC.NSP.PT.SL^DC.SP .PS.DV^[…]TG	0.69	DC.NSP.PT.NSL^ PDC.EN.SPA- TM^TG^QF	0.71
DC.NSP.PT.NSL^ EP.IP^TG	0.54	DC.NSP.PT.NSL^ PDC.ET.AM^TG^QF	0.61	DC.NSP.PT.SL^ DC.SP.PS.DV^TG^QF	0.69	DC.NSP.PT.NSL^TG	0.71
DC.NSP.PT.NSL^ EP.IP^TG^QF^QF	0.54	DC.NSP.TL.NG^EP. EX^CS^TG	0.61	DC.NSP.PT.SL^ FC.ET.PV^TG	0.69	DC.NSP.PT.NSL^ TG[…]^QF	0.71
DC.NSP.PT.NSL^ PDC.EL.EM.RT^TG ^QF	0.54	DC.NSP.TL.NG^ EP.EX^TG	0.61	DC.NSP.PT.SL^ PDC.EL.ID.RT^TG	0.69	DC.NSP.PT.SL^DC.SP .DM.DV^ PDC.EN.SPA- TM^CS^TG^QF	0.71
DC.NSP.PT.NSL^ PDC.MD.PB^TG^ QF	0.54	DC.NSP.TL.NG^ PDC.EL.ID^TG^QF	0.61	DC.NSP.PT.SL^ PDC.EL.ID^TG^QF	0.69	DC.NSP.PT.SL^ PDC.EL.EM.RT^TG^ QF	0.71
DC.NSP.PT.NSL^ TG	0.54	DC.NSP.TL.NG^TG	0.61	DC.NSP.PT.SL^TG	0.69	DC.NSP.PT.SL^ PDC.EL.ID^TG^QF	0.71
DC.NSP.PT.SL^CS ^TG	0.54	DC.SP.DM.DV^ EP.EX^TG	0.61	DC.NSP.PT.SL^TG[…] ^QF	0.69	DC.NSP.PT.SL^TG	0.71
DC.NSP.PT.SL^ DC.SP.DM.DV^ EP.IP^TG^QF	0.54	DC.SP.DM.DV^ EP.EX^TG^QF	0.61	DC.NSP.PT.SL^TG^QF	0.69	DC.NSP.TL.NG^TG^ QF	0.71
DC.NSP.PT.SL^ DC.SP.DM.DV^TG	0.54	DC.SP.DM.DV^EP.IP ^[…]DC.SP.PS.DV^ TG	0.61	DC.NSP.TL.POS^DC.S P.DM.DV^TG	0.69	DC.NSP.TL.POS^ DC.SP.DM.DV^ PDC.EN.SPA- TM^NUM.DF.QV^TG	0.71
DC.NSP.PT.SL^ EP.IP^TG^QF	0.54	DC.SP.DM.DV^ NUM.DF.OR^TG^QF	0.61	DC.NSP.TL.POS^PDC. EL.ID^TG^QF	0.69	DC.SP.DM.DV^CS^TG ^QF	0.71
DC.NSP.PT.SL^TG	0.54	DC.SP.DM.DV^ NUM.DF.QV^CS^CS ^TG^QF	0.61	DC.SP.DM.DV^[…]TG^ QF	0.69	DC.SP.DM.DV^ DC.NSP.PT.NSL^TG	0.71
DC.NSP.PT.SL^TG ^QF	0.54	DC.SP.DM.DV^ NUM.DF.QV^ PDC.EL.EM^TG^QF	0.61	DC.SP.DM.DV^EP.IP^ TG^QF	0.69	DC.SP.DM.DV^EP.EX^ CS^TG^QF	0.71
DC.NSP.TL.NG^ EP.EX^TG	0.54	DC.SP.DM.DV^ PDC.MD.PB^TG	0.61	DC.SP.DM.DV^ FC.EL.EM^EP.EX^TG	0.69	DC.SP.DM.DV^EP.EX^ TG^QF	0.71
DC.NSP.TL.NG^ PDC.EL.EM.AM^TG	0.54	DC.SP.DM.DV^PDC. MD.US^CS^TG^QF	0.61	DC.SP.DM.DV^ NUM.DF.QV^	0.69	DC.SP.DM.DV^ FC.EL.EM^CS^TG^QF	0.71

				PDC.EL.ID.RT^TG^QF			
DC.NSP.TL.NG^PDC.EN.SPA-TM^TG	0.54	DC.SP.PS.DV^[…]TG	0.61	DC.SP.DM.DV^PDC.EL.EM.RT^TG	0.69	DC.SP.DM.DV^FC.EL.EM^TG^QF	0.71
DC.NSP.TL.NG^TG^QF	0.54	DC.SP.PS.DV^CS^CS^TG	0.61	DC.SP.DM.DV^PDC.EL.EM.RT^TG^QF	0.69	DC.SP.DM.DV^NUM.DF.OR^DC.SP.DM.DV^TG	0.71
DC.NSP.TL.POS^DC.SP.DM.DV^EP.EX^TG^QF	0.54	DC.SP.PS.DV^EP.EX^TG	0.61	DC.SP.DM.DV^PDC.EL.ID.RT^TG	0.69	DC.SP.DM.DV^NUM.DF.OR^PDC.MD.PB^EP.EX^TG^QF	0.71
DC.NSP.TL.POS^DC.SP.DM.DV^TG^QF	0.54	DC.SP.PS.DV^EP.IP^TG	0.61	DC.SP.DM.DV^PDC.EL.ID^TG^QF	0.69	DC.SP.DM.DV^NUM.DF.QV^CS^TG^QF	0.71
DC.NSP.TL.POS^DC.SP.PS.DV^EP.IP^TG	0.54	DC.SP.PS.DV^NUM.DF.OR^CS^TG	0.61	DC.SP.DM.DV^TG^[…]QF	0.69	DC.SP.DM.DV^PDC.EL.EM.RT^TG^QF	0.71
DC.NSP.TL.POS^TG	0.54	DC.SP.PS.IV^TG	0.61	DC.SP.PS.DV^CS^TG	0.69	DC.SP.DM.DV^PDC.EL.ID.RT^TG^QF^QF	0.71
DC.NSP.TL.POS^TG^QF	0.54	DC.SP.PS.IV^TG^QF	0.61	DC.SP.PS.DV^EP.IP^TG^QF	0.69	DC.SP.PS.DV^CS^CS^TG^QF	0.71
DC.SP.DM.DV^EP.EX^[…]CS^CS^TG	0.54	EP.EX^TG	0.61	DC.SP.PS.DV^PDC.EL.ID.RT^TG	0.69	DC.SP.PS.DV^CS^TG	0.71
DC.SP.DM.DV^EP.EX^TG	0.54	EP.IP^EP.IP^TG	0.61	EP.EX^CS^TG^QF	0.69	DC.SP.PS.DV^NUM.DF.QV^TG	0.71
DC.SP.DM.DV^EP.EX^TG^QF	0.54	NUM.IN.QV^TG^QF^QF	0.61	EP.EX^TG	0.69	DC.SP.PS.DV^NUM.IN.QV^TG	0.71
DC.SP.DM.DV^FC.EL.SM^TG	0.54	PDC.EL.EM.RT^[…]TG	0.61	EP.EX^TG^[…]QF	0.69	DC.SP.PS.DV^PDC.EL.ID.EZ^TG	0.71
DC.SP.DM.DV^NUM.DF.OR^CS^TG^QF	0.54	PDC.EL.EM^TG	0.61	EP.IP^TG	0.69	EP.EX^CS^TG^[…]QF	0.71
DC.SP.DM.DV^PDC.EL.ID.RT^TG	0.54	TG^QF^QF	0.61	FC.ET.AGG^TG	0.69	EP.EX^TG	0.71
DC.SP.DM.DV^PDC.EL.ID.RT^TG^QF^QF	0.54			FC.ET.PV^DC.SP.DM.DV^TG	0.69	EP.EX^TG^QF	0.71
DC.SP.DM.DV^PDC.ET.AM^TG	0.54			IF^NUM.IN.QV^TG^QF	0.69	EP.IP^TG	0.71
DC.SP.DM.DV^PDC.MD.US^TG^QF	0.54			NUM.DF.QV^TG	0.69	NUM.DF.QV^NUM.DF.QV^CS^TG	0.71
DC.SP.DM.DV^PDC.RP.IA.EZ^CS^TG^QF	0.54			NUM.IN.QV^TG[…]^QF	0.69	NUM.DF.QV^NUM.DF.QV^TG^QF^QF	0.71
DC.SP.PS.DV^[…]TG^QF	0.54			PDC.EL.EM^PDC.EL.ID^TG	0.69	NUM.IN.QV^CS^TG^QF	0.71
DC.SP.PS.DV^EP.IP^TG^QF	0.54			PDC.EL.ID^TG	0.69	NUM.IN.QV^EP.EX^TG	0.71
DC.SP.PS.DV^NUM.DF.QV^TG	0.54			PDC.EL.ID^TG^QF	0.69	NUM.IN.QV^EP.EX^TG^QF^QF	0.71
DC.SP.PS.DV^PDC.EL.ID.RT^TG	0.54			PDC.ET.AM^TG	0.69	NUM.IN.QV^TG	0.71
DC.SP.PS.DV^	0.54			TG^[…]QF	0.69	PDC.EL.EM.RT^TG	0.71

PDC.EL.ID^TG	
DC.SP.PS.IV^ PDC.EL.ID.RT^TG	0.54
EP.EX^CS^CS^TG	0.54
EP.EX^CS^TG^QF^ QF	0.54
EP.EX^TG^QF^QF	0.54
EP.IP^TG^[…]^QF	0.54
EP.IP^TG^QF	0.54
NUM.DF.QV^ PDC.EL.ID^TG^QF	0.54
NUM.DF.QV^TG^ […]^QF	0.54
PDC.EL.ID^CS^TG^ QF	0.54
PDC.EL.ID^EP.IP^ TG	0.54
PDC.EN.SPA- TM^CS^TG	0.54
PDC.ET.AM^TG^ QF	0.54
PDC.RP.IA.EZ^TG TG^QF^QF	0.54

PDC.EL.EM^TG^QF	0.71
PDC.EN.CV^ PDC.EN.CV^TG	0.71

W:misc	%	W:fict	%	S:speech	%	S:conv	%
DC.SP.DM.DV^TG^ QF	14.40	DC.NSP.PT.NSL^TG	11.02	DC.SP.DM.DV^TG	13.86	DC.NSP.PT.NSL^TG	13
DC.SP.DM.DV^TG	8.00	TG	11.02	DC.SP.DM.DV^TG^QF	10.89	DC.SP.DM.DV^TG	9
TG^QF	6.40	DC.SP.DM.DV^TG	6.78	TG^QF	5.94	DC.NSP.PT.NSL^TG^ QF	7
DC.NSP.PT.NSL^ TG^QF	4.80	DC.SP.DM.DV^TG^ QF	6.78	DC.NSP.PT.NSL^TG	4.95	TG^QF	6
TG	4.80	DC.SP.PS.DV^TG	6.78	DC.SP.PS.DV^TG	3.96	DC.SP.DM.DV^TG^ QF	5
CS^TG	3.20	TG^QF	5.93	TG	3.96	NUM.IN.QV^TG	5
DC.NSP.PT.NSL^ CS^TG^QF	3.20	DC.NSP.PT.NSL^TG^ QF	3.39	CS^TG	2.97	TG	5
CS^TG^QF	2.40	DC.NSP.PT.SL^TG^ QF	2.54	DC.NSP.PT.NSL^TG^ QF	2.97	DC.NSP.TL.POS^ DC.SP.DM.DV^TG	3
DC.SP.DM.DV^CS^ TG^QF	2.40	DC.NSP.TL.NG^TG	2.54	DC.SP.DM.DV^EP.IP^ TG^QF	2.97	CS^TG	2
DC.SP.PS.DV^TG	2.40	DC.NSP.PT.NSL^ EP.IP^TG	1.69	DC.SP.PS.DV^TG^QF	2.97	DC.NSP.PT.NSL^CS^ TG	2
NUM.IN.QV^TG^ QF	2.40	DC.NSP.PT.NSL^ EP.IP^TG^QF	1.69	NUM.IN.QV^TG	2.97	DC.NSP.TL.NG^TG^ QF	2
DC.NSP.PT.NSL^ CS^TG	1.60	DC.NSP.TL.POS^TG	1.69	TG	1.98	DC.SP.DM.DV^ PDC.ET.EZ^TG	2
DC.NSP.TL.NG^TG ^QF	1.60	DC.SP.PS.DV^ PDC.EL.ID.RT^TG	1.69	DC.NSP.PT.NSL^CS^ TG^QF	1.98	DC.SP.PS.DV^TG	2
DC.SP.DM.DV^[…] TG^QF	1.60	DC.SP.PS.DV^TG^QF	1.69	DC.NSP.PT.NSL^ EP.IP^TG^QF	1.98	NUM.DF.QV^TG^QF	2
DC.SP.DM.DV^CS^ TG	1.60	NUM.IN.QV^TG	1.69	DC.SP.DM.DV^ FC.EL.EM^TG^QF	1.98	NUM.IN.QV^TG^QF	2
DC.SP.DM.DV^ PDC.EN.SPA-	1.60	(NG)TG	0.85	DC.NSP.PT.SL^ DC.SP.DM.DV^	0.99	DC.NSP.PT.NSL^ EP.EX^TG	1

TM^TG				[…]EP.EX^CS^TG			
DC.SP.DM.DV^TG^QF^QF	1.60	CS^TG	0.85	DC.NSP.PT.SL^DC.SP.DM.DV^EP.IP^TG^QF	0.99	DC.NSP.PT.NSL^EP.IP^TG	1
PDC.EL.ID^TG	1.60	DC.NSP.PT.NSL^CS^TG^QF	0.85	DC.NSP.PT.SL^DC.SP.DM.DV^TG^QF	0.99	DC.NSP.PT.NSL^EP.IP^TG^QF	1
PDC.EN.CV^TG	1.60	DC.NSP.PT.NSL^NUM.IN.QV^TG^QF	0.85	DC.NSP.PT.SL^FC.EL.EM^TG^QF	0.99	DC.NSP.PT.NSL^PDC.MD.US.EZ^TG	1
CS^CS^TG	0.80	DC.NSP.PT.NSL^TG^QF^QF	0.85	DC.NSP.PT.SL^PDC.EL.EM.RT^TG	0.99	DC.NSP.PT.NSL^TG^QF^QF	1
DC.NSP.PT.NSL^[…]TG^QF	0.80	DC.NSP.TL.NG^EP.IP^TG	0.85	DC.NSP.PT.SL^TG	0.99	DC.NSP.PT.SL^DC.SP.DM.DV^EP.IP^TG	1
DC.NSP.PT.NSL^EP.EX^TG^QF	0.80	DC.NSP.TL.NG^PDC.EL.ID^TG	0.85	DC.NSP.TL.POS^TG	0.99	DC.NSP.PT.SL^DC.SP.DM.DV^PDC.EL.EM^TG^QF	1
DC.NSP.PT.NSL^EP.IP^CS^TG^QF	0.80	DC.NSP.TL.NG^TG^QF	0.85	DC.SP.DM.DV^CS^[…]EP.EX^TG^QF	0.99	DC.NSP.PT.SL^DC.SP.DM.DV^TG	1
DC.NSP.PT.NSL^EP.IP^TG	0.80	DC.NSP.TL.POS^FC.EL.EM^PDC.EL.ID^TG	0.85	DC.SP.DM.DV^CS^TG	0.99	DC.NSP.TL.NG^CS^TG	1
DC.NSP.PT.NSL^EP.IP^TG^QF	0.80	DC.NSP.TL.POS^FC.EL.EM^TG	0.85	DC.SP.DM.DV^EP.EX^TG	0.99	DC.NSP.TL.POS^DC.SP.DM.DV^EP.IP^TG^QF	1
DC.NSP.PT.NSL^PDC.EL.ID^TG^QF	0.80	DC.NSP.TL.POS^TG^QF	0.85	DC.SP.DM.DV^EP.EX^TG^QF	0.99	DC.NSP.TL.POS^FC.EL.EM^TG	1
DC.NSP.PT.NSL^PDC.MD.PB^TG	0.80	DC.SP.DM.DV^EP.EX^CS^TG^QF	0.85	DC.SP.DM.DV^NUM.DF.OR^TG^QF	0.99	DC.NSP.TL.POS^TG^QF	1
DC.NSP.PT.NSL^TG	0.80	DC.SP.DM.DV^EP.EX^TG	0.85	DC.SP.DM.DV^PDC.EL.ID.RT^TG	0.99	DC.SP.DM.DV^CS^TG	1
DC.NSP.PT.NSL^TG^QF^QF	0.80	DC.SP.DM.DV^EP.IP^CS^TG	0.85	DC.SP.DM.DV^PDC.RP.LN.EZ^CS^TG^QF	0.99	DC.SP.DM.DV^DC.NSP.PT.SL^PDC.EL.ID^TG^QF	1
DC.NSP.PT.SL^DC.SP.DM.DV^TG	0.80	DC.SP.DM.DV^EP.IP^EP.EX^TG	0.85	DC.SP.DM.DV^UNC^TG^[…]^QF	0.99	DC.SP.DM.DV^FC.EL.EM^EP.IP^TG^QF	1
DC.NSP.PT.SL^DC.SP.DM.DV^TG^QF	0.80	DC.SP.DM.DV^EP.IP^TG	0.85	DC.SP.PS.DV^CS^TG	0.99	DC.SP.DM.DV^FC.EL.EM^TG	1
DC.NSP.PT.SL^DC.SP.PS.DV^CS^TG	0.80	DC.SP.DM.DV^EP.IP^TG^QF	0.85	DC.SP.PS.DV^PDC.ET.AM^TG^QF	0.99	DC.SP.DM.DV^NUM.DF.OR^TG^QF	1
DC.NSP.PT.SL^TG	0.80	DC.SP.DM.DV^NUM.DF.OR^TG^QF	0.85	DC.SP.PS.DV^TG[…]^QF	0.99	DC.SP.DM.DV^PDC.EL.ID.RT^TG	1
DC.NSP.TL.NG^TG	0.80	DC.SP.DM.DV^PDC.EL.EM.EZ^TG^QF	0.85	EP.EX^TG	0.99	DC.SP.DM.DV^PDC.EL.ID.RT^TG^QF	1
DC.NSP.TL.POS^TG^QF	0.80	DC.SP.DM.DV^PDC.EL.ID.RT^TG^QF	0.85	EP.EX^TG^QF	0.99	DC.SP.DM.DV^PDC.EL.ID^TG	1
DC.SP.DM.DV^[…]TG^[…]QF	0.80	DC.SP.DM.DV^PDC.ET.RT^TG	0.85	EP.IP^TG^QF	0.99	DC.SP.PS.DV^TG^QF	1
DC.SP.DM.DV^CS^TG^QF^QF	0.80	DC.SP.DM.DV^PDC.MD.PB^TG	0.85	FC.ET.PV^DC.NSP.TL.POS^DC.SP.PS.DV^TG	0.99	EP.EX^TG	1
DC.SP.DM.DV^	0.80	DC.SP.DM.IV^TG	0.85	NUM.DF.QV^EP.IP^TG	0.99	IF^DC.NSP.PT.NSL^	1

EP.IP^TG^QF				^QF		EP.EX^TG	
DC.SP.DM.DV^ NUM.DF.OR^TG	0.80	DC.SP.PS.DV^[…]^ TG	0.85	NUM.DF.QV^ PDC.MD.PB^TG^QF	0.99	NUM.DF.QV^CS^TG	1
DC.SP.DM.DV^ NUM.DF.OR^TG^ QF^QF	0.80	DC.SP.PS.DV^CS^TG	0.85	NUM.DF.QV^TG^QF^ QF	0.99	NUM.DF.QV^EP.IP^ TG	1
DC.SP.DM.DV^ PDC.EL.EM.EZ^TG	0.80	DC.SP.PS.DV^EP.IP^ TG	0.85	NUM.IN.QV^ DC.NSP.PT.NSL^TG	0.99	NUM.IN.QV.IV^TG	1
DC.SP.DM.DV^ PDC.EL.EM^TG^ QF	0.80	DC.SP.PS.DV^ NUM.DF.OR^TG^QF	0.85	NUM.IN.QV^EP.IP^TG ^QF	0.99	NUM.IN.QV^CS^TG	1
DC.SP.DM.DV^ PDC.EL.ID^EP.IP^ TG^QF	0.80	DC.SP.PS.DV^ PDC.MD.RD^TG	0.85	NUM.IN.QV^ PDC.EL.ID^TG^QF	0.99	NUM.IN.QV^EP.IP^ TG	1
DC.SP.DM.DV^ PDC.EN.SPA-TM^TG^QF	0.80	DC.SP.PS.DV^TG^ […]QF	0.85	NUM.IN.QV^TG^[…]QF	0.99	NUM.IN.QV^TG^QF^ QF	1
DC.SP.DM.DV^ PDC.MD.OB^TG	0.80	EP.EX^EP.EX^TG	0.85	NUM.IN.QV^TG^QF	0.99	PDC.EL.EM.RT^TG^ QF	1
DC.SP.DM.DV^ TG[…]^QF	0.80	EP.IP^ DC.NSP.PT.NSL^TG	0.85	PDC.EL.EM.RT^TG	0.99	PDC.EL.ID.RT^TG	1
DC.SP.PS.DV^CS^ TG	0.80	EP.IP^TG	0.85	PDC.EL.ID^TG^QF	0.99	PDC.ET.EZ^EP.IP^TG	1
DC.SP.PS.DV^CS^ TG^QF	0.80	EP.IP^TG^QF	0.85	PDC.EL.ID^TG^QF^QF	0.99	TG^[…]QF	1
DC.SP.PS.DV^ EP.EX^TG^QF	0.80	FC.ET.AGG^TG^QF	0.85	TG^[…]QF	0.99		
DC.SP.PS.DV^ EP.IP^TG	0.80	NUM.IN.QV^TG^QF	0.85				
DC.SP.PS.DV^ PDC.EN.SPA-TM^TG	0.80	PDC.EL.EM.RT^TG	0.85				
EP.EX^[…]TG	0.80	PDC.EL.ID^EP.EX^ CS^TG	0.85				
EP.IP^TG	0.80	PDC.EL.ID^TG^QF	0.85				
EP.IP^TG^QF	0.80	PDC.ET.AM^TG	0.85				
NUM.DF.QV^ PDC.EL.ID^TG	0.80						
NUM.IN.QV^ DC.SP.DM.DV^TG	0.80						
NUM.IN.QV^ DC.SP.DM.DV^TG^ QF	0.80						
NUM.IN.QV^TG	0.80						
PDC.EL.EM.RT^TG ^QF	0.80						
PDC.ET.AM^TG	0.80						

W:commerce	%	S:brdcast	%
TG^QF	12.24	DC.SP.DM.DV^TG	23.08
DC.SP.DM.DV^TG	11.22	DC.NSP.PT.NSL^TG	10.26
DC.SP.DM.DV^TG^QF	11.22	CS^TG	5.13
CS^TG	6.12	DC.NSP.PT.NSL^TG^QF	5.13
DC.NSP.PT.NSL^CS^TG	4.08	DC.NSP.TL.NG^TG^QF	5.13
CS^TG^QF	3.06	TG^QF	5.13
DC.NSP.PT.NSL^TG^QF	3.06	DC.NSP.PT.SL^DC.SP.DM.DV^CS^TG	2.56
DC.SP.PS.DV^CS^TG	3.06	DC.NSP.PT.SL^TG	2.56
TG^QF^QF	3.06	DC.NSP.TL.NG^TG	2.56
CS^CS^TG	2.04	DC.SP.DM.DV^CS^TG	2.56
DC.NSP.PT.NSL^EP.EX^CS^TG	2.04	DC.SP.DM.DV^EP.IP^TG^QF	2.56
DC.SP.DM.DV^TG^QF^QF	2.04	DC.SP.DM.DV^FC.EL.EM^CS^[…]^TG^QF	2.56
TG	2.04	DC.SP.DM.DV^NUM.DF.QV^TG	2.56
DC.NSP.PT.NSL^[…]^TG	1.02	DC.SP.DM.DV^TG^QF	2.56
DC.NSP.PT.NSL^[…]CS^TG	1.02	DC.SP.DM.DV^TG^QF^QF	2.56
DC.NSP.PT.NSL^CS^CS^TG	1.02	DC.SP.PS.DV^EP.IP^TG	2.56
DC.NSP.PT.NSL^CS^TG^QF	1.02	DC.SP.PS.DV^EP.IP^TG^[…]QF	2.56
DC.NSP.PT.NSL^CS^TG^QF^QF	1.02	DC.SP.PS.DV^FC.EL.EM^CS^TG	2.56
DC.NSP.PT.NSL^EP.EX^TG	1.02	DC.SP.PS.DV^TG	2.56
DC.NSP.PT.NSL^EP.IP^TG^QF	1.02	EP.EX^TG^QF	2.56
DC.NSP.PT.NSL^PDC.MD.US^TG^QF	1.02	EP.IP^TG^QF	2.56
DC.NSP.PT.NSL^PDC.RP.IA.EZ^TG^QF	1.02	NUM.IN.QV^DC.NSP.PT.NSL^TG	2.56
DC.NSP.PT.NSL^TG	1.02	PDC.EN.CV^TG	2.56
DC.NSP.PT.NSL^TG^[…]QF	1.02	TG	2.56
DC.NSP.PT.SL^EP.EX^CS^TG	1.02		
DC.NSP.PT.SL^TG	1.02		
DC.NSP.TL.NG^PDC.EL.EM.RT^TG	1.02		
DC.NSP.TL.NG^PDC.RP.LN^TG^QF	1.02		
DC.NSP.TL.NG^TG	1.02		
DC.SP.DM.DV^[…]TG^QF	1.02		
DC.SP.DM.DV^CS^TG^QF	1.02		
DC.SP.DM.DV^EP.EX^TG^QF	1.02		
DC.SP.DM.DV^EP.IP^TG^QF	1.02		
DC.SP.DM.DV^FC.EL.SM^TG	1.02		
DC.SP.DM.DV^PDC.EN.SPA-TM^PDC.EL.ID^[…]TG^QF^[…]^QF	1.02		
DC.SP.PS.DV^CS^[…]TG	1.02		
DC.SP.PS.DV^CS^TG^QF	1.02		
DC.SP.PS.DV^TG^QF	1.02		
EP.EX^TG^QF^QF	1.02		
EP.IP^EP.EX^TG	1.02		
EP.IP^TG	1.02		
NUM.DF.QV^NUM.DF.QV^PDC.EL.ID^TG	1.02		
NUM.DF.QV^PDC.EL.ID.RT^TG	1.02		
NUM.IN.QV^CS^TG	1.02		
NUM.IN.QV^TG^[…]^QF	1.02		
NUM.IN.QV^TG^QF	1.02		
TG^QF^[…]^QF	1.02		

7.2 Formal patterns

W:newsp	%	W:non_ac	%	S:meeting	%	W:ac	%
AJ^H	5.43	H	9.15	DF.AR^H	9.03	AJ^H	5.71
H	4.89	DF.AR^H	8.54	DF.AR^H^PP(of n)	5.56	DF.AR^H	5.00
DF.AR^H	4.35	DF.AR^H^PP(of n)	5.49	PS.DT^H	4.17	DM.DT(THESE)^H	5.00
DF.AR^H^PP(of n)	3.80	H^PP(of n)	3.66	DF.AR^AJ^H	3.47	DF.AR^H^PP(of n)	2.86
H^PP(of n)	3.80	AJ^H	3.05	DM.DT(THAT)^H	3.47	DM.DT(THIS)^H	2.86
		DF.AR^AJ^H^					
DF.AR^AJ^H	3.26	PP(of n)	3.05	AJ^H	2.78	H	2.86
N^H	2.72	IN.AR^H^PP(of n)	3.05	DF.AR^H^RT.RV.CL	2.78	IN.AR^AJ^H	2.86
IN.AR^AJ^H	2.17	DM.DT(THIS)^H	2.44	H	2.78	PS.DT^H	2.86
NG.DT^AJ^H	2.17	DF.AR^N^H	1.83	IN.AR^AJ^H	2.78	DF.AR^AJ^H	2.14
		DF.DV.GV.NP^H^PP(of				DF.AR^AJ^H^	
NUM.CD^H	2.17	n)	1.83	DM.DT(THIS)^H	2.08	PP(of n)	2.14
						DF.DV.GV.NP^H^	
DF.DV.GV.NP^H	1.63	IN.AR^AJ^H	1.83	IN.AR^H	2.08	PP(of n)	2.14
DF.DV.GV.NP^H^							
AP.TI.CL	1.63	PS.DT^H	1.83	AS.DT^H	1.39	H^RT.RV.CL	2.14
						IN.AR^AJ^H^	
H^PP(for n)	1.63	AJ^AJ^H	1.22	DF.AR^H^PP(on n)	1.39	PP(of n)	2.14
IN.AR^AJ^H^							
AP.TI.CL	1.63	AJ^H^PL.ED.CL	1.22	DF.DV.GV.NP^H	1.39	NUM.GO^H	2.14
IN.AR^N^H	1.63	DF.AR^AJ^H	1.22	H^PP(of n)	1.39	AS.DT^H	1.43
PS.DT^AJ^H	1.63	DF.AR^AJP^H	1.22	IN.AR^H^PP(of n)	1.39	NUM.CD^H	1.43
DF.AR^AJ^H^		DF.AR^H^					
PP(of n)	1.09	PP(of n)^AP.TI.CL	1.22	NAS.DT^H	1.39	AJ^AJ^AJ^H	0.71
				NUM.CD^PV.PP(of			
				DF.AR^AJ^H^			
DF.AR^N^H	1.09	DF.AR^H^RT.RV.CL	1.22	RT.RV.CL)	1.39	AJ^AJP^H	0.71
DF.DV.GV.NP^N^				NUM.CD^PV.PP(of			
H	1.09	DM.DT(THIS)^AJ^H	1.22	DF.AR^H^RT.RV.CL)	1.39	AJ^H^PP(against n)	0.71
		IN.AR^AJ^H^					
DM.DT(THIS)^H	1.09	PP(to n)	1.22	QT^H	1.39	AJ^H^PP(in n)	0.71
IN.AR^H^AP.TI.CL	1.09	IN.AR^AJP^H	1.22	(ING)N^H	0.69	AJ^H^PP(in which)	0.71
NAS.DT^AJ^H	1.09	N^H	1.22	AJ^AJ^H	0.69	AJ^H^PP(to n)	0.71
						AJ^N^H^	
NG.DT^H	1.09	PS.DT^AJ^H	1.22	AJ^H^[…]PP(of n)	0.69	[…]RT.RV.CL	0.71
PS.DT^H	1.09	AJ^[…]H	0.61	AJ^N^H^RT.RV.CL	0.69	AJP^H	0.71
		AJ^H^				CV.QT.PN^PV.PP(of	
PS.DT^H^PP(of n)	1.09	FL.ST.RT.EM.AP	0.61	AS.DT^AJ^H	0.69	DF.AR^H^PP of n)	0.71
		AJ^H^		AS.DT^AJP^AJP^H^			
QT^H	1.09	NR.RV.CL(where)	0.61	PP(in n)	0.69	CV.QT^H	0.71
AJ^H^[…]^							
AP.THAT.CL	0.54	AJ^H^PP(about n)	0.61	AS.DT^H^PP(in n)	0.69	CV.QT^H^PP(on n)	0.71
AJ^H^PP(about n)	0.54	AJ^H^PP(in n)	0.61	AS.DT^H^PP(of n)	0.69	DF.AR^AJ^AJ^H	0.71
				AS.PN^PV.PP(of			
				DF.AR^H^)		DF.AR^AJ^AJ^H^	
AJ^H^PP(for n)	0.54	AJ^H^PP(on n)	0.61	PP(in which))	0.69	PP(of n)	0.71
AJ^H^PP(from		AJ^H^		CV.QT^H[…]^		DF.AR^AJ^H^	
n)^AP.THAT.CL	0.54	PT.ST.NR.EM.AP	0.61	RT.RV.CL	0.69	PP(of n)^RT.RV.CL	0.71
		AJ^H^		DF.AR^[…]H^			
AJ^H^PP(of n)	0.54	PT.WK.NR.EM.AP	0.61	RT.RV.CL	0.69	DF.AR^AJ^N^H	0.71
		DF.AR^AJ^H^					
AJ^N^AJ^H	0.54	PP(of -ing)	0.61	DF.AR^AJ^H^PP(for n)	0.69	DF.AR^H^AP.TI.CL	0.71
AJ^N^H	0.54	DF.AR^AJ^N^H^	0.61	DF.AR^AJ^H^	0.69	DF.AR^H^PP(for n)	0.71

AJ^N^H^ PT.ST.NR.EM.AP^ NR.RV.CL	0.54	NR.RV.CL DF.AR^AJP^H^ PP(of n)	0.61	RT.RV.CL DF.AR^H^[…]PP(to n)	0.69	DF.AR^H^PP(of -ing)	0.71
AS.PN^ PV.PP(of DM.DT(THESE)^H	0.54	DF.AR^CV.AJ^H	0.61	DF.AR^H^PL.ED.CL	0.69	DF.AR^H^RT.RV.CL	0.71
DF.AR^AJ^[…]AJ^ AJ^H	0.54	DF.AR^CV.AJP^ PP(of […] DF.DV.GV.NP^H)	0.61	DF.AR^H^ PP(in n)^AP.THAT.CL	0.69	DF.AR^N^H	0.71
DF.AR^AJ^AJ^H	0.54	DF.AR^H^PL.ING.CL	0.61	DF.AR^H^PP(of -ing)	0.69	DF.AR^NUM.CD^AJ^ H^PP(of n)	0.71
DF.AR^AJ^H^ PP(between n)	0.54	DF.AR^H^ PP(with which)	0.61	DF.AR^H^ PP(of n)^AP.THAT.CL	0.69	DF.AR^NUM.OR^AJ ^AJ^H^PP(of n)	0.71
DF.AR^AJ^H^ PP(of -ing)	0.54	DF.AR^NP^H	0.61	DF.AR^H^ PP(of ps.dt -ing)	0.69	DF.AR^NUM.OR^ PV.PP(of DM.DT(THESE)^H)	0.71
DF.AR^AJ^H^ PP(of n)^PP(for n)	0.54	DF.AR^NUM.CD^AJ^H ^PP(of n)	0.61	DF.AR^H^PP(to n)	0.69	DF.AR^QL.PV^AJ^H ^RT.RV.CL	0.71
DF.AR^AJ^H^ PP(over n)	0.54	DF.AR^NUM.CD^N^N^ H^RT.RV.CL	0.61	DF.AR^H^RT.RV.CL^ PP(of n)	0.69	DF.AR^QL.PV^H^ RT.RV.CL	0.71
DF.AR^AJ^N^H^ RT.RV.CL	0.54	DF.AR^NUM.OR^H^PP (of n)	0.61	DF.AR^NP^H	0.69	DF.DV.GV.NP^ NUM.CD^H	0.71
DF.AR^H^ AP.THAT.CL	0.54	DF.DV.GV.NP^[…]H	0.61	DF.AR^NUM.CD^AJ^H ^RT.RV.CL	0.69	DM.DT(THAT)^H^PP (of n)	0.71
DF.AR^H^AV	0.54	DF.DV.GV.NP^AJ^ AJ^H	0.61	DF.AR^NUM.GO^H	0.69	DM.DT(THIS)^AJ^H	0.71
DF.AR^H^ PL.ED.CL	0.54	DF.DV.GV.NP^H	0.61	DF.AR^NUM.GO^H^PP (in which)	0.69	DM.DT(THIS)^H^ PP(in n)	0.71
DF.AR^H^ PP(between n)	0.54	DF.DV.GV.NP^H^ AP.TI.CL	0.61	DM.DT(THAT)^AJ^H	0.69	DM.DT(THIS)^H^ PP(of n)	0.71
DF.AR^H^ PP(for n)	0.54	DF.DV.GV.NP^H^ PP(for n)	0.61	DM.DT(THESE)^H	0.69	DM.DT(THIS)^ NUM.GO^H	0.71
DF.AR^H^ PP(of -ing)	0.54	DM.DT(THAT)^H	0.61	DM.DT(THESE)^ QL.PV^AJ^H	0.69	H^AP.TI.CL	0.71
DF.AR^H^ PP(of n -ing)	0.54	DM.DT(THESE)^H	0.61	DM.DT(THIS)^(ING)N^ H	0.69	H^PP(in n)	0.71
DF.AR^H^PP(of n)^AP.THAT.CL	0.54	DM.DT(THIS)^H^ PP(of n)^PP(over n)	0.61	H^[…]PP(in n)	0.69	H^PP(of n)	0.71
DF.AR^H^ PP(of n)^AP.TI.CL	0.54	H^AP.TI.CL	0.61	H^CV.AJP	0.69	H^TI.CL	0.71
DF.AR^H^ RT.RV.CL^ PT.ST.NR.IT.AP	0.54	H^PP(at n)	0.61	H^PP(about n -ing)	0.69	IN.AR^AJ^H^ NR.PL.ING.CL	0.71
DF.AR^N^H^ PP(of n)	0.54	H^PP(for wh)	0.61	H^PP(for n)	0.69	IN.AR^AJ^H^ PL.ED.CL	0.71
DF.AR^N^H^ PT.ST.NR.ATT.AP	0.54	H^PP(of n -ing)	0.61	H^PP(to n)	0.69	IN.AR^AJ^H^ PP(for -ing)	0.71
DF.AR^NP^AJ^H	0.54	H^PP(of n)^AP.TI.CL	0.61	IF.AV^QT^H^PP(of n)	0.69	IN.AR^AJ^H^ PP(for n)	0.71
DF.AR^NUM.GO^ N^H^		IN.AR^(ING)N^H^		IN.AR^AJ^H^[…]		IN.AR^AJ^H^	
FL.ST.NR.IT.AP	0.54	PP(at n)	0.61	PP(on n)	0.69	RT.RV.CL	0.71
DF.DV.GV.NP^AJ ^H	0.54	IN.AR^AJ^AJ^H	0.61	IN.AR^AJ^H^ AP.THAT.CL	0.69	IN.AR^AJP^AJ^H	0.71
DF.DV.GV.NP^H^	0.54	IN.AR^AJ^AJ^H^AJP	0.61	IN.AR^AJ^H^PP(for n)	0.69	IN.AR^AJP^H	0.71

PP(for n)							
DF.DV.GV.NP^H^		IN.AR^AJ^H^				IN.AR^CV.AJ^H^	
PP(to n)	0.54	NR.RV.CL	0.61	IN.AR^AJP^H	0.69	PP(in which)	0.71
DF.DV.GV.NP^		IN.AR^AJ^H^		IN.AR^H^[…]		IN.AR^H^[…]	
NUM.CD^H	0.54	PP(of n)	0.61	PP(about n)	0.69	PL.ING.CL	0.71
DF.DV.GV.NP^		IN.AR^AJ^H^					
NUM.GO^H	0.54	PP(of n)^PP(in which)	0.61	IN.AR^H^AP.TI.CL	0.69	IN.AR^H^AP.TI.CL	0.71
		IN.AR^AJP^H^					
DM.DT(THAT)^H^		NR.PL.ED.CL^					
AP.TI.CL	0.54	AP.THAT.CL	0.61	IN.AR^H^PP(of wh)	0.69	IN.AR^H^PP(of n)	0.71
DM.DT(THAT)^H^		IN.AR^AJP^H^					
PP(of n)	0.54	PP(of n)	0.61	IN.AR^H^RT.RV.CL	0.69	IN.AR^H^RT.RV.CL	0.71
DM.DT(THIS)^AJ^				IN.AR^H^			
H^PP(against n)	0.54	IN.AR^H^PL.ED.CL	0.61	RT.RV.CL(where)	0.69	IN.AR^N^H	0.71
DM.DT(THIS)^						IN.AR^N^H^	
QL.PV^H	0.54	IN.AR^H^PP(like n)	0.61	IN.AR^NP^H	0.69	PP(against n)	0.71
DM.DT(THOSE)^				NAS.DT^H[…]^			
AJ^H	0.54	IN.AR^N^AJ^H	0.61	RT.RV.CL	0.69	IN.AR^NUM.GO^H	0.71
		IN.AR^N^H^					
H^AP.TI.CL	0.54	PP(of which)	0.61	NUM.CD^AJ^H	0.69	N^H^AJP	0.71
H^PP(among						NG.DT^H^	
n)^AP.THAT.CL	0.54	NG.DT^AJ^AJ^H	0.61	NUM.CD^H	0.69	AP.THAT.CL	0.71
				NUM.CD^H^		NUM.CD^AJ^H^	
H^PP(in n)	0.54	NG.DT^AJ^H	0.61	FL.ST.NR.IT.AP	0.69	RT.RV.CL	0.71
				NUM.CD^NUM.GO^H^		NUM.CD^H^	
H^PP(on n)	0.54	NG.DT^H	0.61	PL.ED.CL	0.69	PP(of -ing)	0.71
				NUM.CD^PV.PP(of			
				DF.AR^NUM.GO^H^			
H^PP(to n)	0.54	NG.DT^H^PP(for n)	0.61	RT.RV.CL)	0.69	NUM.CD^H^PP(of n)	0.71
				NUM.CD^PV.PP(of		NUM.CD^NUM.GO^	
H^PP(with n)	0.54	NG.DT^H^PP(of n)	0.61	PS.DT^[…]H)	0.69	H^PP(of n)	0.71
						NUM.CD^PV.PP(of	
				NUM.CD^PV.PP(of		DF.AR^AJ^AJ^H^	
H^RT.RV.CL	0.54	NG.DT^H^PP(on n)	0.61	PS.DT^H^PP(since n))	0.69	TI.CL)	0.71
IN.AR^(ING)N^H^		NG.DT^NUM.GO^H^TI.				NUM.CD^PV.PP(of	
AP.TI.CL	0.54	CL	0.61	NUM.CD^QV.PV^H	0.69	NUM.CD^AJ^H)	0.71
						NUM.CD^PV.PP(of	
IN.AR^AJ^H[…]^						NUM.CD^H^AV^	
AP.THAT.CL	0.54	NP^H	0.61	NUM.GO^H^PP(of n)	0.69	PT.WK.NR.IT.AP)	0.71
IN.AR^AJ^H^		NUM.GO^H^					
AP.THAT.CL	0.54	PL.ED.CL	0.61	PS.DT^AJ^H	0.69	PDT^H	0.71
IN.AR^AJ^H^		NUM.GO^H^		PS.DT^AJP^H^			
NR.AP.THAT.CL	0.54	PP(of -ing)	0.61	PP(of n)	0.69	PDT^IN.AR^H	0.71
IN.AR^AJ^H^						PS.DT^AJ^AJ^H^	
PP(between n)	0.54	PDT^H	0.61	PS.DT^H^PP(for n)	0.69	PT.WK.NR.PA.AP	0.71
IN.AR^AJ^H^							
PP(over n)	0.54	PS.DT^H^AP.TI.CL	0.61	PS.DT^H^PP(of n)	0.69	PS.DT^AJ^H	0.71
IN.AR^AJ^H^						PS.DT^H^	
PP(to n)	0.54	PS.DT^H^NR.RV.CL	0.61	PS.DT^N^H	0.69	AP.THAT.CL	0.71
IN.AR^AJP^H^							
PP(by n)^PP(of n)	0.54	PS.DT^H^PP(at -ing)	0.61	QT^H^PP(from n)	0.69	PS.DT^H^PP(for n)	0.71
IN.AR^CV.AJ^AJ^				QV.PV^DM.DT(THAT)^			
H^AP.TI.CL	0.54	PS.DT^H^PP(for n)	0.61	H	0.69	PS.DT^H^PP(of -ing)	0.71
IN.AR^CV.AJP^N^							
H	0.54	PS.DT^H^PP(in -ing)	0.61	QV.PV^H	0.69	PS.DT^N^H	0.71
IN.AR^H	0.54	PS.DT^H^PP(on n)	0.61	UV.PDT^DF.AR^H^	0.69	PS.DT^QT^H	0.71

IN.AR^H^AP.THAT.CL	0.54	PS.DT^NUM.OR^AJ^H	0.61	PL.ED.CL		QT.PN^PV.PP(of	
				UV.PDT^DF.AR^H^PP(on n)	0.69	DF.AR^H^RT.RV.CL)	0.71
IN.AR^NP^H	0.54	QT^H^PP (concerning n)	0.61	UV.PDT^NUM.GO^H^PP(of n)	0.69	QT^AJ^H	0.71
N^H^PT.ST.NR.EM.AP^RT.RV.CL	0.54	QT^H^PP(of -ing)	0.61	UV.PN^PV.PP(of DM.DT(THESE)^H)	0.69	QT^AJ^H^PP(for n)^PT.WK.NR.IT.AP	0.71
NAS.DT^AJ^H^PP(of n -ing)	0.54	QT^H^PP(on n)^PT.ST.NR.EM.AP	0.61			QT^AJ^H^PP(on n)	0.71
NG.DT^H^PP(of -ing)	0.54	RV.DT^H	0.61			QT^H^AP.THAT.CL	0.71
NP^H	0.54	RV.DT^H^PP(of n)	0.61			QT^H^FL.ST.NR.IT.AP	0.71
NUM.CD^H^[…]^PP(of n)	0.54					QT^H^PP(of n)	0.71
NUM.CD^H^PP(of n)	0.54					UV.PN^PV.PP(of DF.AR^AJ^NUM.CD^H)	0.71
NUM.CD^NUM.GO^H^RT.RV.CL(where)	0.54						
NUM.CD^PV.PP(of DF.AR^AJ^H^NR.PL.ING.CL)	0.54						
NUM.GO^AJ^H	0.54						
NUM.GO^AJ^H^PL.ED.CL	0.54						
PS.DT^(ING)N^H	0.54						
PS.DT^[…]H^PP(on n)	0.54						
PS.DT^AJ^H^AP.THAT.CL	0.54						
PS.DT^H^AP.TI.CL	0.54						
PS.DT^H^PP(for n)	0.54						
PS.DT^NUM.OR^H	0.54						
RV.DT^AJ^H	0.54						
UV.DT^H	0.54						
UV.DT^H^PP(of n)	0.54						
UV.PDT^DF.AR^H^^RT.RV.CL	0.54						
UV.PDT^DM.DT(THESE)^AJ^H^NR.RV.CL	0.54						
UV.PDT^PS.DT^AJ^H	0.54						

W:misc	%	W:fict	%	S:speech	%	S:conv	%
DF.AR^H^PP(of n)	6.40	H	11.02	DF.AR^H	10.89	IN.AR^H	13.00

AJ^H	4.80	IN.AR^H	11.02	DF.AR^H^RT.RV.CL	3.96	DF.AR^H	7.00
H	4.80	PS.DT^H	5.93	H	3.96	H	5.00
DF.AR^H	4.00	AJ^H	3.39	IN.AR^H	3.96	DF.AR^H^RT.RV.CL	4.00
DF.AR^AJ^H	3.20	DF.AR^H	3.39	QT^H	2.97	DF.AR^AJ^H	3.00
DF.AR^AJ^H^ PP(of n)	3.20	DF.AR^AJ^H	2.54	AS.DT^H	1.98	H^PP(of n)	3.00
DM.DT(THESE)^ H	2.40	DF.AR^H^PP(of n)	2.54	DF.AR^AJ^H	1.98	QT^H	3.00
H^TI.CL	2.40	DF.AR^H^RT.RV.CL	2.54	DF.AR^H^PP(of n)	1.98	AJ^H	2.00
IN.AR^AJ^H	2.40	DF.DV.GV.NP^AJ^H	2.54	DF.DV.GV.NP^H	1.98	IN.AR^AJ^H	2.00
AJ^H^PP(in n)	1.60	PDT^H	2.54	DM.DT(THIS)^H	1.98	IN.AR^H^AP.TI.CL	2.00
AJ^H^PP(of n)	1.60	H^PP(of n)	1.69	H^PP(of n)	1.98	IN.AR^H^PP(with n)	2.00
DF.AR^H^ PL.ED.CL	1.60	IN.AR^H^PP(of n)	1.69	H^RT.RV.CL(when)	1.98	N^H	2.00
DF.AR^H^ PP(of -ing)	1.60	NG.DT^H	1.69	IN.AR^H^PP(of n)	1.98	PS.DT^H	2.00
DF.DV.GV.NP^H	1.60	PS.DT^AJ^H	1.69	NP^H	1.98	QT^AJ^H	2.00
DM.DT(THIS)^H	1.60	QT^H	1.69	PS.DT^H	1.98	QT^H^RT.RV.CL	2.00
H^PP(of n)	1.60	UV.DT^H	1.69	AJ^H	0.99	UV.PDT^DF.AR^H	2.00
H^RT.RV.CL	1.60	(NG)H	0.85	AJ^H^AP.TI.CL	0.99	AJ^AJ^H	1.00
IN.AR^AJ^H^ PL.ED.CL	1.60	AJ^AJ^H	0.85	AJ^H^PP(of n)	0.99	AJ^H^RT.RV.CL	1.00
IN.AR^H^PP(of n)	1.60	AJ^H^PP(of n -ing)	0.85	AJP^H	0.99	AS.DT^AJ^H	1.00
IN.AR^H^ RT.RV.CL	1.60	AJP^IN.AR^H	0.85	AS.PN^PV.PP(of DF.AR[…]^AJ^(ING)N^ H)	0.99	CV.QT^H	1.00
N^H	1.60	AS.DT^AJ^H	0.85	CV.QT^CV.AJP^H^ PP(in n)	0.99	DF.AR^AJ^H^ RT.RV.CL	1.00
NUM.GO^H	1.60	AS.DT^CV.QT^H^TI.CL	0.85	DF.AR^(ING)N^H	0.99	DF.AR^H^AP.TI.CL	1.00
PS.DT^AJ^H	1.60	AS.DT^H^PP(for n)	0.85	DF.AR^AJ^AJ^H^ PP(of n)	0.99	DF.AR^NP^H	1.00
PS.DT^AJ^H^ PP(in n)	1.60	DF.AR^AJ^AJ^H^ PP(of n)	0.85	DF.AR^AJ^H^ PP(from within n)	0.99	DF.AR^NUM.CD^ NUM.GO^H^ RT.RV.CL	1.00
(ING)N^(ING)N^H	0.80	DF.AR^AJ^H^AJP	0.85	DF.AR^AJ^H^PP(in n)	0.99	DF.AR^NUM.GO^H	1.00
AJ^[…]H	0.80	DF.AR^AJ^H^PP(for n)	0.85	DF.AR^AJ^H^PP(of n)	0.99	DF.AR^NUM.OR^H^ RT.RV.CL	1.00
AS.DT^H	0.80	DF.AR^CV.AJ^H^ PP(of n)	0.85	DF.AR^AJ^H^ RT.RV.CL	0.99	DM.DT(THAT)^H	1.00
DF.AR^[…]H^[…] PP(of n)	0.80	DF.AR^H^ FL.ST.RT.AP	0.85	DF.AR^H^(for n AP.TI.CL)	0.99	DM.DT(THAT)^QL.PV ^H	1.00
DF.AR^[…]H^ AP.TI.CL	0.80	DF.AR^H^PP(with n)	0.85	DF.AR^H^ FL.ST.RT.AP	0.99	DM.DT(THIS)^H	1.00
DF.AR^[…]H^ PP(of n)	0.80	DF.AR^NUM.OR^H^ RT.RV.CL	0.85	DF.AR^H^PP(to n)	0.99	DM.DT(THIS)^QL.PV ^AJ^H^RT.RV.CL	1.00
DF.AR^AJ^H^ RT.RV.CL^ PP(with n)	0.80	DF.DV.GV.NP^H	0.85	DF.AR^NP^[…]AJ^H^ RT.RV.CL	0.99	H^[…]AP.TI.CL	1.00
DF.AR^CV.AJ^H^ PP(of n)	0.80	DF.DV.GV.NP^H^ […]AP.TI.CL	0.85	DF.AR^NUM.GO^H^ RT.RV.CL	0.99	H^AP.TI.CL	1.00
DF.AR^H[…]^ PP(of n)	0.80	DF.DV.GV.NP^H^ PP(of n)	0.85	DF.AR^QL.PV^H^ RT.RV.CL	0.99	H^RT.RV.CL	1.00
DF.AR^H^ AP.THAT.CL	0.80	DM.DT(THIS)^AJ^AJ^H	0.85	DF.AR^QL.PV^H^ RT.RV.CL(where)	0.99	H^TI.CL	1.00
DF.AR^H^ AP.TI.CL	0.80	DM.DT(THIS)^AJ^AJP^ H	0.85	DF.AR^UNC^H^[…]^ RT.RV.CL	0.99	IF.AV^IN.AR^AJ^H	1.00

DF.AR^H^PP(for n)	0.80	DM.DT(THIS)^H	0.85	DF.DV.GV.NP^N^H	0.99	IN.AR^(ING)N^H	1.00
DF.AR^H^PP(of aj)	0.80	H^PP(by n)	0.85	DM.DT(THAT)^H	0.99	IN.AR^AJ^H^PP(with n)	1.00
DF.AR^H^PP(of n -ing)	0.80	H^PP(from n)	0.85	DM.DT(THIS)^H^PP(of n)	0.99	IN.AR^H^PP(of -ing)	1.00
DF.AR^H^PP(of n)^AP.TI.CL	0.80	H^PP(to n)	0.85	H^[…]PP(of -ing)	0.99	IN.AR^H^PP(of n)^NR.RV.CL	1.00
DF.AR^H^PP(of n)^NR.RV.CL	0.80	H^RT.RV.CL	0.85	H^PP(of -ing)	0.99	IN.AR^H^PT.ST.NR.IT.AP	1.00
DF.AR^H^RT.RV.CL	0.80	H^RT.RV.CL(when)	0.85	H^PP(to n)	0.99	IN.AR^H^PT.WK.NR.IT.AP	1.00
DF.AR^N^H	0.80	IN.AR^AJ^H	0.85	IN.AR^AJ^H	0.99	IN.AR^N^H	1.00
DF.AR^N^H^AP.THAT.CL	0.80	IN.AR^AJ^H^PP(for which)	0.85	IN.AR^AJ^H^PP(of ?)	0.99	NAS.DT^H^PP(in -ing)	1.00
DF.AR^NP^H	0.80	IN.AR^AJ^H^PP(of n -ing)	0.85	IN.AR^AJ^H^PP(of n)	0.99	NG.DT^H^PP(of n -ing)	1.00
DF.AR^NUM.GO^AJ^H^TI.CL	0.80	IN.AR^AJP^H^NR.RV.CL	0.85	IN.AR^AJ^H^PT.WK.NR.IT.AP	0.99	IN.AR^H^NG.DT^N^H	1.00
DF.AR^NUM.OR^H	0.80	IN.AR^H^PP(about n)	0.85	IN.AR^H^RT.RV.CL(where)	0.99	NUM.CD(QT)^H	1.00
DF.AR^NUM.OR^H^TI.CL^PP(with n)	0.80	IN.AR^H^PP(of -ing)^RT.RV.CL	0.85	IN.AR^N^H	0.99	NUM.CD^H^PT.WK.NR.IT.AP	1.00
H^PP(with n)	0.80	IV.DT^H	0.85	IN.AR^NP^H^RT.RV.CL(whereby)	0.99	NUM.CD^N^H	1.00
IN.AR^[…]H^PP(of n)	0.80	NAS.DT^H	0.85	N^H	0.99	NUM.CD^PV.PP(of DF.AR^AJ^H)	1.00
IN.AR^AJ^H^NR.PL.ED.CL	0.80	NAS.DT^H^PP(as to n)	0.85	NAS.DT^AJ^H	0.99	NUM.CD^PV.PP(of DF.AR^AJ^H^PP(about n))	1.00
IN.AR^AJ^H^PP(of n)	0.80	NG.DT^AJ^H	0.85	NAS.DT^QL.PV^H^RT.RV.CL	0.99	NUM.CD^PV.PP(of DF.AR^H)	1.00
IN.AR^AJ^H^PP(to n)	0.80	NG.DT^H^PP(for which)	0.85	NUM.CD(QT)^H^[…]PP(about n)	0.99	PS.DT^H^PP(for n)	1.00
IN.AR^AJ^H^PT.WK.NR.IT.AP	0.80	NG.DT^NUM.GO^H	0.85	NUM.CD(QT)^NUM.GO^H^PP(on n)	0.99	QT.IV^H	1.00
IN.AR^AJ^N^H^PP(on n)	0.80	NUM.CD^H^CV.AJP	0.85	NUM.CD^AJ^H^PT.ST.NR.IT.AP	0.99	QT^DS.GV^H	1.00
IN.AR^AJP^H	0.80	NUM.CD^H^PP(of wh)	0.85	NUM.CD^AJP^H^PP(in terms of n)	0.99	QT^H^PL.ING.CL^PP(as to wh)	1.00
IN.AR^CV.AJ^H^PP(of n)	0.80	NUM.GO^AJ^AJ^H	0.85	NUM.CD^H^PP(of n)^RT.RV.CL	0.99	QT^H^TI.CL	1.00
IN.AR^H	0.80	NUM.GO^H^PP(on n)	0.85	NUM.CD^PV.PP(of DF.AR^AJ^H^RT.RV.CL)	0.99	UV.DT^H^RT.RV.CL	1.00
IN.AR^H^PL.ED.CL	0.80	PDT^AJ^H	0.85	NUM.CD^PV.PP(of DF.AR^H^RT.RV.CL)	0.99	UV.PDT^DF.AR^AJ^H^PP(about n)	1.00
IN.AR^H^PP(for n)	0.80	PS.DT^[…]^H	0.85	NUM.GO^H^PP(of n)	0.99	UV.PDT^DM.DT(THAT)^H	1.00
IN.AR^H^PP(of n)^RT.RV.CL	0.80	PS.DT^H^PP(for n)	0.85	NUM.GO^H^PP(to n)^RT.RV.CL	0.99	UV.PDT^QL.PV^H	1.00
N^H^FL.ST.NR.RF.AP	0.80	PS.DT^NUM.OR^H^PP(of n)	0.85	PDT^H^RT.RV.CL(as)	0.99		
NAS.DT^H	0.80	QT^H^PP(of -ing)	0.85	PS.DT^AJ^H^AP.TI.CL	0.99		
NAS.DT^H^	0.80	QV.PV^H^	0.85	PS.DT^H[…]^PP(of wh)	0.99		

AP.THAT.CL		PT.WK.NR.IT.AP			
NAS.PN^PV.PP(of DF.AR^H^PP in n)	0.80	UV.DT^QL.PV^H	0.85	PS.DT^H^PL.ED.CL	0.99
NAS.PN^PV.PP(of DM.DT(THESE)^H)	0.80	UV.PDT^H^PP(for n)	0.85	PS.DT^H^PP(as n)	0.99
NG.DT^H^TI.CL	0.80	UV.PDT^QL.PV^NUM.GO^H	0.85	PS.DT^H^PP(in -ing)	0.99
NUM.CD^AJ^H	0.80			QT^H^PP(of -ing)	0.99
NUM.CD^PV.PP(of PS.DT^N^H)	0.80			QT^IN.AR^H	0.99
PS.DT^H	0.80			QV.PV^UV.PDT^PS.DT^H	0.99
PS.DT^N^H	0.80			UV.PDT^H	0.99
QT.PN^PV.PP(of DF.AR^H^PP of n)	0.80				
QT.PN^PV.PP(of DM.DT(THESE)^H)	0.80				
QT^H	0.80				
QT^H^FL.ST.NR.EM.AP	0.80				
QT^H^PP(on n)	0.80				
QT^H^TI.CL	0.80				
UV.PDT^H^PP(of n)	0.80				

W:commerce		S:brdcast	
DF.AR^H	2.72	DF.AR^H	17.95
H^PP(of n)	2.72	IN.AR^H	10.26
AJ^H	2.17	AJ^H	7.69
DF.AR^H^PP(of n)	2.17	AJ^H^PP(about n)	2.56
DM.DT(THIS)^H	1.63	AJ^H^PP(with n -ing)	2.56
AJ^AJ^H	1.09	DF.AR^AJP^H^FL.WK.NR.IT.AP	2.56
DF.AR^H^PP(of -ing)	1.09	DF.AR^H^PP(of n -ing)	2.56
DF.DV.GV.NP^AJ^H	1.09	DF.AR^H^PP(of n)^PP(about n)	2.56
DM.DT(THESE)^H	1.09	DF.AR^NP^H	2.56
H	1.09	DF.AR^NUM.CD^H	2.56
H^PP(for n)	1.09	DF.AR^QL.PV^AJ^[…]^H^PP(of n)	2.56
IN.AR^AJ^H	1.09	DM.DT(THAT)^H	2.56
IN.AR^AJ^N^H	1.09	DM.DT(THESE)^H	2.56
N^H	1.09	H	2.56
N^H	0.54	H^PP(against n)	2.56
AJ^H^PP(in n)^PT.WK.NR.PA.AP	0.54	H^PP(of n)	2.56
AJ^H^PP(with n)	0.54	IN.AR^H^PP(of -ing)	2.56
AJ^H^RT.RV.CL	0.54	IN.AR^H^RT.RV.CL(where)	2.56
AS.DT^H	0.54	NAS.DT^H	2.56
DF.AR^[…]H^PP(of n)	0.54	NAS.DT^H^PP(between n)	2.56
DF.AR^AJ^H^AP.TI.CL	0.54	NG.DT^H^PP(of n)	2.56
DF.AR^AJ^H^PP(of n)	0.54	NUM.CD^H	2.56
DF.AR^AJ^H^PT.WK.NR.AP	0.54	NUM.CD^PV.PP(of DF.AR^(ING)N^H)	2.56

DF.AR^AJ^NUM.GO^[…]H^ PP(of n)^[…]^PT.ST.NR.IT.AP	0.54
DF.AR^H^AP.THAT.CL	0.54
DF.AR^H^PP(for -ing)	0.54
DF.AR^H^PP(of n)^PP(in n)	0.54
DF.AR^H^PP(to n)	0.54
DF.AR^H^RT.RV.CL(when)	0.54
DF.AR^H^TI.CL	0.54
DF.DV.GV.NP^AJ^[…]H	0.54
DM.DT(THAT)^H^ PP(of n)^FL.ST.NR.IT.AP	0.54
DM.DT(THIS)^QL.PV^H	0.54
DM.DT(THOSE)^H	0.54
H^AP.TI.CL	0.54
H^PL.ED.CL	0.54
H^PP(for n)^PL.ED.CL	0.54
H^PP(from n)	0.54
H^PP(in n)	0.54
H^PP(of n)^[…]^PT.WK.NR.IT.AP	0.54
H^PP(of n)^PP(of n)	0.54
H^PP(of n)^RT.RV.CL	0.54
H^PP(other than n)	0.54
IN.AR^[…]^H	0.54
IN.AR^[…]AJ^H	0.54
IN.AR^AJ^AJ^H	0.54
IN.AR^AJ^H^AP.THAT.CL	0.54
IN.AR^AJ^H^FL.ST.NR.IT.AP^ PL.ED.CL	0.54
IN.AR^AJ^H^PP(for which)	0.54
IN.AR^AJ^H^PP(of n)	0.54
IN.AR^AJ^H^PP(to n)	0.54
IN.AR^AJP^H	0.54
IN.AR^H^[…]PP(of n)	0.54
IN.AR^H^AP.THAT.CL	0.54
IN.AR^H^AP.TI.CL	0.54
IN.AR^H^PP(for n)	0.54
IN.AR^N^H	0.54
IN.AR^NP^H	0.54
N^N^H	0.54
NAS.DT^AJ^N^H	0.54
NAS.DT^H	0.54
NG.DT^AJ^H	0.54
NG.DT^AJ^H^RT.RV.CL	0.54
NG.DT^H	0.54
NP^H^RT.RV.CL	0.54
NUM.CD^AJ^H	0.54
NUM.CD^PV.PP(of NUM.CD^AJ^H)	0.54
PS.DT^AJ^H^AP.TI.CL	0.54
PS.DT^H^PP(during n)	0.54
PS.DT^N^H	0.54
QT^AJ^H	0.54
QT^H^[…]^RT.RV.CL	0.54
QT^H^PP(of -ing)	0.54
PS.DT^AJ^H	2.56
PS.DT^AJ^H^[…]AP.TI.CL	2.56
PS.DT^H	2.56
PS.DT^QL.PV^AJ^H	2.56
QT^IN.AR^H	2.56

APPENDIX 8 LEMMA DISTRIBUTION OF DETERMINERS, EXPERIENTIAL PREMODIFIERS AND POSTMODIFYING STRUCTURES (COMPLETE LISTS)

8.1 *Determiners (top ten)*

DF.AR	%	Ø	%	IN.AR	%	PS.DT	%
Opposite	90.00	Foreboding	85.71	Correction	100.00	Recollection	80.00
Answer	65.00	Leave	75.00	Venture	44.83	Quest	60.00
Challenge	64.71	Anger	60.00	Testimony	42.86	Endorsement	50.00
Motivation	62.50	Detail	57.50	Joke	40.00	Endeavour	44.44
Scandal	61.54	Irony	57.14	Assessment	38.46	Leave	25.00
Characteristic	55.17	Terror	50.00	Dimension	37.50	Experience	22.50
Contradiction	50.00	Evidence	45.00	Triumph	37.50	Philosophy	18.18
Myth	43.33	Part	43.75	Endeavour	33.33	Prejudice	18.18
Finding	42.86	Testimony	42.86	Impetus	33.33	System	17.65
Proviso	42.86	Suspicion	40.00	Example	32.50	Capacity	17.39
Recommendation	37.50	Prejudice	36.36	Sense	31.82	Surprise	16.00
Thing	37.50	Application	35.00	Vision	31.03	Objective	15.00
Impetus	33.33	Practice	35.00	Warning	31.03	Crime	12.50
Objective	32.50	Failure	34.78	Failure	26.09	Motivation	12.50
Word	32.50	Impetus	33.33	Chance	22.50	Terror	12.50
System	32.35	Crime	32.50	Problem	22.50	Triumph	12.50
Point	30.00	Project	32.50	Application	20.00	Word	12.50
Practice	30.00	Work	29.63	Crime	20.00	Finding	10.71
Time	30.00	Facet	28.57	Point	20.00	Characteristic	10.34
Irony	28.57	Proviso	28.57	Suspicion	20.00	Venture	10.34
Project	27.50	Venture	27.59	Phenomenon	19.44	Vision	10.34
Philosophy	27.27	Problem	27.50	System	17.65	Chance	10.00
Sense	27.27	Endorsement	25.00	Surprise	16.00	Dimension	8.33
Capacity	26.09	Warning	24.14	Scandal	15.38	Scandal	7.69
Failure	26.09	Myth	23.33	Project	15.00	Application	7.50
Work	25.93	Example	22.50	Word	15.00	Work	7.41
Application	25.00	Joke	22.50	Challenge	14.71	Recommendation	6.25
Endorsement	25.00	System	20.59	Irony	14.29	Challenge	5.88
Phenomenon	25.00	Answer	20.00	Philosophy	13.64	Practice	5.00
Vision	24.14	Experience	20.00	Myth	13.33	Way	5.00
Area	22.50	Quest	20.00	Capacity	13.04	Sense	4.55
Chance	22.50	Time	20.00	Area	12.50	Failure	4.35
Problem	22.50	Philosophy	18.18	Contradiction	12.50	Assessment	3.85
Dimension	20.83	Objective	17.50	Recommendation	12.50	Warning	3.45
Anger	20.00	Characteristic	17.24	Finding	10.71	Detail	2.50
Evidence	20.00	Vision	17.24	Experience	10.00	Joke	2.50
Experience	20.00	Assessment	15.38	Objective	7.50	Project	2.50
Recollection	20.00	Thing	15.00	Practice	7.50	Time	2.50
Assessment	19.23	Way	15.00	Way	7.50	Anger	0.00
Crime	17.50	Misfortune	14.29	Characteristic	6.90	Answer	0.00
Way	17.50	Phenomenon	13.89	Part	6.25	Area	0.00
Warning	17.24	Capacity	13.04	Time	5.00	Contradiction	0.00
Example	15.00	Area	12.50	Detail	2.50	Correction	0.00
Misfortune	14.29	Chance	12.50	Thing	2.50	Evidence	0.00
Detail	12.50	Contradiction	12.50	Anger	0.00	Example	0.00
Part	12.50	Triumph	12.50	Answer	0.00	Facet	0.00
Terror	12.50	Word	12.50	Endorsement	0.00	Foreboding	0.00
Triumph	12.50	Finding	10.71	Evidence	0.00	Impetus	0.00
Suspicion	12.00	Opposite	10.00	Facet	0.00	Irony	0.00

Endeavour	11.11	Recommendation	9.38	Foreboding	0.00	Misfortune	0.00
Venture	10.34	Challenge	8.82	Leave	0.00	Myth	0.00
Prejudice	9.09	Dimension	8.33	Misfortune	0.00	Opposite	0.00
Joke	7.50	Point	5.00	Motivation	0.00	Part	0.00
Surprise	4.00	Correction	0.00	Opposite	0.00	Phenomenon	0.00
Correction	0.00	Endeavour	0.00	Prejudice	0.00	Point	0.00
Facet	0.00	Motivation	0.00	Proviso	0.00	Problem	0.00
Foreboding	0.00	Recollection	0.00	Quest	0.00	Proviso	0.00
Leave	0.00	Scandal	0.00	Recollection	0.00	Suspicion	0.00
Quest	0.00	Sense	0.00	Terror	0.00	Testimony	0.00
Testimony	0.00	Surprise	0.00	Work	0.00	Thing	0.00

DM.DT	%	QT	%	DF.DV.GV.NP	%	NUM.CD	%
Phenomenon	27.78	Time	15.00	Anger	20.00	Facet	28.57
Terror	25.00	Work	14.81	Quest	20.00	Proviso	14.29
Area	20.00	Detail	12.50	Capacity	17.39	Way	10.00
Finding	17.86	Way	12.50	Assessment	15.38	Part	9.38
Misfortune	14.29	Evidence	10.00	Foreboding	14.29	Sense	9.09
Proviso	14.29	Point	10.00	Misfortune	14.29	Example	7.50
Contradiction	12.50	Practice	10.00	Testimony	14.29	Thing	7.50
Motivation	12.50	Prejudice	9.09	Contradiction	12.50	Warning	6.90
Part	12.50	Surprise	8.00	Triumph	12.50	Answer	5.00
Triumph	12.50	Scandal	7.69	Vision	10.34	Area	5.00
Evidence	10.00	Crime	7.50	Philosophy	9.09	Objective	5.00
Point	10.00	Myth	6.67	Scandal	7.69	Word	5.00
Project	10.00	Recommendation	6.25	Finding	7.14	Characteristic	3.45
Time	10.00	Area	5.00	Warning	6.90	Chance	2.50
Recommendation	9.38	Example	5.00	Recommendation	6.25	Point	2.50
Capacity	8.70	Thing	5.00	Chance	5.00	Practice	2.50
Assessment	7.69	Sense	4.55	Example	5.00	Problem	2.50
Application	7.50	Vision	3.45	Word	5.00	Time	2.50
Experience	7.50	System	2.94	Failure	4.35	Anger	0.00
Joke	7.50	Phenomenon	2.78	Dimension	4.17	Application	0.00
Thing	7.50	Application	2.50	Work	3.70	Assessment	0.00
Venture	6.90	Chance	2.50	Challenge	2.94	Capacity	0.00
Objective	5.00	Experience	2.50	System	2.94	Challenge	0.00
Practice	5.00	Joke	2.50	Answer	2.50	Contradiction	0.00
Way	5.00	Objective	2.50	Application	2.50	Correction	0.00
Philosophy	4.55	Anger	0.00	Experience	2.50	Crime	0.00
Sense	4.55	Answer	0.00	Objective	2.50	Detail	0.00
Dimension	4.17	Assessment	0.00	Point	2.50	Dimension	0.00
Work	3.70	Capacity	0.00	Practice	2.50	Endeavour	0.00
Vision	3.45	Challenge	0.00	Area	0.00	Endorsement	0.00
Myth	3.33	Characteristic	0.00	Characteristic	0.00	Evidence	0.00
Crime	2.50	Contradiction	0.00	Correction	0.00	Experience	0.00
Problem	2.50	Correction	0.00	Crime	0.00	Failure	0.00
Anger	0.00	Dimension	0.00	Detail	0.00	Finding	0.00
Answer	0.00	Endeavour	0.00	Endeavour	0.00	Foreboding	0.00
Challenge	0.00	Endorsement	0.00	Endorsement	0.00	Impetus	0.00
Chance	0.00	Facet	0.00	Evidence	0.00	Irony	0.00
Characteristic	0.00	Failure	0.00	Facet	0.00	Joke	0.00
Correction	0.00	Finding	0.00	Impetus	0.00	Leave	0.00
Detail	0.00	Foreboding	0.00	Irony	0.00	Misfortune	0.00
Endeavour	0.00	Impetus	0.00	Joke	0.00	Motivation	0.00
Endorsement	0.00	Irony	0.00	Leave	0.00	Myth	0.00

Example	0.00	Leave	0.00	Motivation	0.00	Opposite	0.00
Facet	0.00	Misfortune	0.00	Myth	0.00	Phenomenon	0.00
Failure	0.00	Motivation	0.00	Opposite	0.00	Philosophy	0.00
Foreboding	0.00	Opposite	0.00	Part	0.00	Prejudice	0.00
Impetus	0.00	Part	0.00	Phenomenon	0.00	Project	0.00
Irony	0.00	Philosophy	0.00	Prejudice	0.00	Quest	0.00
Leave	0.00	Problem	0.00	Problem	0.00	Recollection	0.00
Opposite	0.00	Project	0.00	Project	0.00	Recommendation	0.00
Prejudice	0.00	Proviso	0.00	Proviso	0.00	Scandal	0.00
Quest	0.00	Quest	0.00	Recollection	0.00	Surprise	0.00
Recollection	0.00	Recollection	0.00	Sense	0.00	Suspicion	0.00
Scandal	0.00	Suspicion	0.00	Surprise	0.00	System	0.00
Surprise	0.00	Terror	0.00	Suspicion	0.00	Terror	0.00
Suspicion	0.00	Testimony	0.00	Terror	0.00	Testimony	0.00
System	0.00	Triumph	0.00	Thing	0.00	Triumph	0.00
Testimony	0.00	Venture	0.00	Time	0.00	Venture	0.00
Warning	0.00	Warning	0.00	Venture	0.00	Vision	0.00
Word	0.00	Word	0.00	Way	0.00	Work	0.00

NG.DT	%	AS.DT	%
Surprise	32.00	Facet	14.29
Chance	10.00	Misfortune	14.29
Evidence	10.00	Endeavour	11.11
Prejudice	9.09	Suspicion	8.00
Experience	5.00	Part	6.25
Sense	4.55	Area	5.00
Dimension	4.17	Example	5.00
Warning	3.45	Point	5.00
Myth	3.33	Word	5.00
Part	3.13	Sense	4.55
System	2.94	Surprise	4.00
Phenomenon	2.78	Work	3.70
Crime	2.50	Answer	2.50
Detail	2.50	Joke	2.50
Problem	2.50	Thing	2.50
Thing	2.50	Way	2.50
Word	2.50	Anger	0.00
Anger	0.00	Application	0.00
Answer	0.00	Assessment	0.00
Application	0.00	Capacity	0.00
Area	0.00	Challenge	0.00
Assessment	0.00	Chance	0.00
Capacity	0.00	Characteristic	0.00
Challenge	0.00	Contradiction	0.00
Characteristic	0.00	Correction	0.00
Contradiction	0.00	Crime	0.00
Correction	0.00	Detail	0.00
Endeavour	0.00	Dimension	0.00
Endorsement	0.00	Endorsement	0.00
Example	0.00	Evidence	0.00
Facet	0.00	Experience	0.00
Failure	0.00	Failure	0.00
Finding	0.00	Finding	0.00
Foreboding	0.00	Foreboding	0.00
Impetus	0.00	Impetus	0.00
Irony	0.00	Irony	0.00

Joke	0.00	Leave	0.00
Leave	0.00	Motivation	0.00
Misfortune	0.00	Myth	0.00
Motivation	0.00	Objective	0.00
Objective	0.00	Opposite	0.00
Opposite	0.00	Phenomenon	0.00
Philosophy	0.00	Philosophy	0.00
Point	0.00	Practice	0.00
Practice	0.00	Prejudice	0.00
Project	0.00	Problem	0.00
Proviso	0.00	Project	0.00
Quest	0.00	Proviso	0.00
Recollection	0.00	Quest	0.00
Recommendation	0.00	Recollection	0.00
Scandal	0.00	Recommendation	0.00
Suspicion	0.00	Scandal	0.00
Terror	0.00	System	0.00
Testimony	0.00	Terror	0.00
Time	0.00	Testimony	0.00
Triumph	0.00	Time	0.00
Venture	0.00	Triumph	0.00
Vision	0.00	Venture	0.00
Way	0.00	Vision	0.00
Work	0.00	Warning	0.00

DM.DT(THIS)	%	DM.DT(THAT)	%	DM.DT(THESE)	%	DM.DT(THOSE)	%
Terror	25.00	Misfortune	14.29	Phenomenon	13.89	Assessment	3.85
Proviso	14.29	Motivation	12.50	Finding	10.71	Part	3.13
Area	12.50	Capacity	8.70	Application	5.00	Experience	2.50
Contradiction	12.50	Part	6.25	Joke	5.00	Anger	0.00
Triumph	12.50	Area	5.00	Point	5.00	Answer	0.00
Phenomenon	11.11	Evidence	5.00	Venture	3.45	Application	0.00
Project	7.50	Point	5.00	Recommendation	3.13	Area	0.00
Finding	7.14	Time	5.00	Area	2.50	Capacity	0.00
Evidence	5.00	Vision	3.45	Experience	2.50	Challenge	0.00
Objective	5.00	Recommendation	3.13	Project	2.50	Chance	0.00
Practice	5.00	Phenomenon	2.78	Anger	0.00	Characteristic	0.00
Thing	5.00	Experience	2.50	Answer	0.00	Contradiction	0.00
Time	5.00	Joke	2.50	Assessment	0.00	Correction	0.00
Philosophy	4.55	Thing	2.50	Capacity	0.00	Crime	0.00
Sense	4.55	Way	2.50	Challenge	0.00	Detail	0.00
Dimension	4.17	Anger	0.00	Chance	0.00	Dimension	0.00
Assessment	3.85	Answer	0.00	Characteristic	0.00	Endeavour	0.00
Work	3.70	Application	0.00	Contradiction	0.00	Endorsement	0.00
Venture	3.45	Assessment	0.00	Correction	0.00	Evidence	0.00
Myth	3.33	Challenge	0.00	Crime	0.00	Example	0.00
Part	3.13	Chance	0.00	Detail	0.00	Facet	0.00
Recommendation	3.13	Characteristic	0.00	Dimension	0.00	Failure	0.00
Application	2.50	Contradiction	0.00	Endeavour	0.00	Finding	0.00
Crime	2.50	Correction	0.00	Endorsement	0.00	Foreboding	0.00
Problem	2.50	Crime	0.00	Evidence	0.00	Impetus	0.00
Way	2.50	Detail	0.00	Example	0.00	Irony	0.00
Anger	0.00	Dimension	0.00	Facet	0.00	Joke	0.00
Answer	0.00	Endeavour	0.00	Failure	0.00	Leave	0.00
Capacity	0.00	Endorsement	0.00	Foreboding	0.00	Misfortune	0.00

Challenge	0.00	Example	0.00	Impetus	0.00	Motivation	0.00
Chance	0.00	Facet	0.00	Irony	0.00	Myth	0.00
Characteristic	0.00	Failure	0.00	Leave	0.00	Objective	0.00
Correction	0.00	Finding	0.00	Misfortune	0.00	Opposite	0.00
Detail	0.00	Foreboding	0.00	Motivation	0.00	Phenomenon	0.00
Endeavour	0.00	Impetus	0.00	Myth	0.00	Philosophy	0.00
Endorsement	0.00	Irony	0.00	Objective	0.00	Point	0.00
Example	0.00	Leave	0.00	Opposite	0.00	Practice	0.00
Experience	0.00	Myth	0.00	Part	0.00	Prejudice	0.00
Facet	0.00	Objective	0.00	Philosophy	0.00	Problem	0.00
Failure	0.00	Opposite	0.00	Practice	0.00	Project	0.00
Foreboding	0.00	Philosophy	0.00	Prejudice	0.00	Proviso	0.00
Impetus	0.00	Practice	0.00	Problem	0.00	Quest	0.00
Irony	0.00	Prejudice	0.00	Proviso	0.00	Recollection	0.00
Joke	0.00	Problem	0.00	Quest	0.00	Recommendation	0.00
Leave	0.00	Project	0.00	Recollection	0.00	Scandal	0.00
Misfortune	0.00	Proviso	0.00	Scandal	0.00	Sense	0.00
Motivation	0.00	Quest	0.00	Sense	0.00	Surprise	0.00
Opposite	0.00	Recollection	0.00	Surprise	0.00	Suspicion	0.00
Point	0.00	Scandal	0.00	Suspicion	0.00	System	0.00
Prejudice	0.00	Sense	0.00	System	0.00	Terror	0.00
Quest	0.00	Surprise	0.00	Terror	0.00	Testimony	0.00
Recollection	0.00	Suspicion	0.00	Testimony	0.00	Thing	0.00
Scandal	0.00	System	0.00	Thing	0.00	Time	0.00
Surprise	0.00	Terror	0.00	Time	0.00	Triumph	0.00
Suspicion	0.00	Testimony	0.00	Triumph	0.00	Venture	0.00
System	0.00	Triumph	0.00	Vision	0.00	Vision	0.00
Testimony	0.00	Venture	0.00	Warning	0.00	Warning	0.00
Vision	0.00	Warning	0.00	Way	0.00	Way	0.00
Warning	0.00	Word	0.00	Word	0.00	Word	0.00
Word	0.00	Work	0.00	Work	0.00	Work	0.00

8.2 *Experiential premodifiers (top ten)*

CS	%	EP.IP	%	EP.EX	%	PDC.EL.ID	%
Venture	62.07	Testimony	42.86	Phenomenon	22.22	Dimension	20.83
Practice	37.50	Endorsement	25.00	Anger	20.00	Facet	14.29
System	35.29	Vision	17.24	Answer	12.50	Misfortune	14.29
Scandal	30.77	Impetus	16.67	Project	12.50	Part	12.50
Project	30.00	Experience	15.00	Venture	10.34	Way	12.50
Dimension	29.17	Thing	15.00	Evidence	10.00	Endeavour	11.11
Characteristic	27.59	Facet	14.29	Experience	10.00	Area	10.00
Capacity	26.09	Irony	14.29	Challenge	8.82	Thing	10.00
Application	25.00	Misfortune	14.29	Capacity	8.70	Example	5.00
Triumph	25.00	Proviso	14.29	Dimension	8.33	Objective	5.00
Work	22.22	Sense	13.64	Scandal	7.69	Sense	4.55
Anger	20.00	Example	12.50	Word	7.50	Myth	3.33
Quest	20.00	Terror	12.50	Vision	6.90	Recommendation	3.13
Philosophy	18.18	Triumph	12.50	Warning	6.67	Chance	2.50
Joke	17.50	Surprise	12.00	Area	5.00	Crime	2.50
Impetus	16.67	Suspicion	12.00	Example	5.00	Detail	2.50
Crime	15.00	Challenge	11.76	Point	5.00	Point	2.50
Irony	14.29	Endeavour	11.11	Problem	5.00	Word	2.50
Myth	13.33	Work	11.11	Philosophy	4.55	Anger	0.00
Warning	13.33	Time	10.00	Sense	4.55	Answer	0.00

Endeavour	11.11	Prejudice	9.09	Work	3.70	Application	0.00
Vision	10.34	Assessment	7.69	Finding	3.57	Assessment	0.00
Evidence	10.00	Area	7.50	Part	3.13	Capacity	0.00
Objective	10.00	Crime	7.50	System	2.94	Challenge	0.00
Recommendation	9.38	Point	7.50	Application	2.50	Characteristic	0.00
Prejudice	9.09	Finding	7.14	Crime	2.50	Contradiction	0.00
Sense	9.09	Venture	6.90	Detail	2.50	Correction	0.00
Phenomenon	8.33	Warning	6.67	Joke	2.50	Endorsement	0.00
Assessment	7.69	Part	6.25	Thing	2.50	Evidence	0.00
Experience	7.50	Answer	5.00	Way	2.50	Experience	0.00
Word	7.50	Chance	5.00	Assessment	0.00	Failure	0.00
Finding	7.14	Joke	5.00	Chance	0.00	Finding	0.00
Answer	5.00	Practice	5.00	Characteristic	0.00	Foreboding	0.00
Chance	5.00	Problem	5.00	Contradiction	0.00	Impetus	0.00
Thing	5.00	Way	5.00	Correction	0.00	Irony	0.00
Way	5.00	Word	5.00	Endeavour	0.00	Joke	0.00
Failure	4.35	Philosophy	4.55	Endorsement	0.00	Leave	0.00
Part	3.13	Characteristic	3.45	Facet	0.00	Motivation	0.00
Challenge	2.94	Myth	3.33	Failure	0.00	Opposite	0.00
Area	2.50	Recommendation	3.13	Foreboding	0.00	Phenomenon	0.00
Detail	2.50	System	2.94	Impetus	0.00	Philosophy	0.00
Point	2.50	Phenomenon	2.78	Irony	0.00	Practice	0.00
Problem	2.50	Application	2.50	Leave	0.00	Prejudice	0.00
Time	2.50	Detail	2.50	Misfortune	0.00	Problem	0.00
Contradiction	0.00	Objective	2.50	Motivation	0.00	Project	0.00
Correction	0.00	Project	2.50	Myth	0.00	Proviso	0.00
Endorsement	0.00	Anger	0.00	Objective	0.00	Quest	0.00
Example	0.00	Capacity	0.00	Opposite	0.00	Recollection	0.00
Facet	0.00	Contradiction	0.00	Practice	0.00	Scandal	0.00
Foreboding	0.00	Correction	0.00	Prejudice	0.00	Surprise	0.00
Leave	0.00	Dimension	0.00	Proviso	0.00	Suspicion	0.00
Misfortune	0.00	Evidence	0.00	Quest	0.00	System	0.00
Motivation	0.00	Failure	0.00	Recollection	0.00	Terror	0.00
Opposite	0.00	Foreboding	0.00	Recommendation	0.00	Testimony	0.00
Proviso	0.00	Leave	0.00	Surprise	0.00	Time	0.00
Recollection	0.00	Motivation	0.00	Suspicion	0.00	Triumph	0.00
Surprise	0.00	Opposite	0.00	Terror	0.00	Venture	0.00
Suspicion	0.00	Quest	0.00	Testimony	0.00	Vision	0.00
Terror	0.00	Recollection	0.00	Time	0.00	Warning	0.00
Testimony	0.00	Scandal	0.00	Triumph	0.00	Work	0.00

PDC.EL.ID.RT	%	PDC.EL.EM.RT	%	EP.EX&CS	%	PDC.EN.SPA-TM	%
Misfortune	14.29	Proviso	14.29	Dimension	12.50	Impetus	16.67
Proviso	14.29	Area	7.50	Endeavour	11.11	Proviso	14.29
Endeavour	11.11	Characteristic	6.90	Practice	10.00	Joke	5.00
Objective	10.00	Surprise	4.00	Recommendation	9.38	Sense	4.55
Thing	10.00	Assessment	3.85	Philosophy	4.55	Assessment	3.85
Finding	7.14	Finding	3.57	Assessment	3.85	Myth	3.33
Part	6.25	Venture	3.45	Work	3.70	Warning	3.33
Area	5.00	Warning	3.33	Myth	3.33	Phenomenon	2.78
Way	5.00	Challenge	2.94	System	2.94	Area	2.50
Dimension	4.17	Application	2.50	Anger	0.00	Example	2.50
Characteristic	3.45	Point	2.50	Answer	0.00	Objective	2.50
Answer	2.50	Practice	2.50	Application	0.00	Point	2.50
Chance	2.50	Project	2.50	Area	0.00	Practice	2.50

Crime	2.50	Thing	2.50	Capacity	0.00	Project	2.50
Point	2.50	Anger	0.00	Challenge	0.00	Anger	0.00
Problem	2.50	Answer	0.00	Chance	0.00	Answer	0.00
Word	2.50	Capacity	0.00	Characteristic	0.00	Application	0.00
Anger	0.00	Chance	0.00	Contradiction	0.00	Capacity	0.00
Application	0.00	Contradiction	0.00	Correction	0.00	Challenge	0.00
Assessment	0.00	Correction	0.00	Crime	0.00	Chance	0.00
Capacity	0.00	Crime	0.00	Detail	0.00	Characteristic	0.00
Challenge	0.00	Detail	0.00	Endorsement	0.00	Contradiction	0.00
Contradiction	0.00	Dimension	0.00	Evidence	0.00	Correction	0.00
Correction	0.00	Endeavour	0.00	Example	0.00	Crime	0.00
Detail	0.00	Endorsement	0.00	Experience	0.00	Detail	0.00
Endorsement	0.00	Evidence	0.00	Facet	0.00	Dimension	0.00
Evidence	0.00	Example	0.00	Failure	0.00	Endeavour	0.00
Example	0.00	Experience	0.00	Finding	0.00	Endorsement	0.00
Experience	0.00	Facet	0.00	Foreboding	0.00	Evidence	0.00
Facet	0.00	Failure	0.00	Impetus	0.00	Experience	0.00
Failure	0.00	Foreboding	0.00	Irony	0.00	Facet	0.00
Foreboding	0.00	Impetus	0.00	Joke	0.00	Failure	0.00
Impetus	0.00	Irony	0.00	Leave	0.00	Finding	0.00
Irony	0.00	Joke	0.00	Misfortune	0.00	Foreboding	0.00
Joke	0.00	Leave	0.00	Motivation	0.00	Irony	0.00
Leave	0.00	Misfortune	0.00	Objective	0.00	Leave	0.00
Motivation	0.00	Motivation	0.00	Opposite	0.00	Misfortune	0.00
Myth	0.00	Myth	0.00	Part	0.00	Motivation	0.00
Opposite	0.00	Objective	0.00	Phenomenon	0.00	Opposite	0.00
Phenomenon	0.00	Opposite	0.00	Point	0.00	Part	0.00
Philosophy	0.00	Part	0.00	Prejudice	0.00	Philosophy	0.00
Practice	0.00	Phenomenon	0.00	Problem	0.00	Prejudice	0.00
Prejudice	0.00	Philosophy	0.00	Project	0.00	Problem	0.00
Project	0.00	Prejudice	0.00	Proviso	0.00	Quest	0.00
Quest	0.00	Problem	0.00	Quest	0.00	Recollection	0.00
Recollection	0.00	Quest	0.00	Recollection	0.00	Recommendation	0.00
Recommendation	0.00	Recollection	0.00	Scandal	0.00	Scandal	0.00
Scandal	0.00	Recommendation	0.00	Sense	0.00	Surprise	0.00
Sense	0.00	Scandal	0.00	Surprise	0.00	Suspicion	0.00
Surprise	0.00	Sense	0.00	Suspicion	0.00	System	0.00
Suspicion	0.00	Suspicion	0.00	Terror	0.00	Terror	0.00
System	0.00	System	0.00	Testimony	0.00	Testimony	0.00
Terror	0.00	Terror	0.00	Thing	0.00	Thing	0.00
Testimony	0.00	Testimony	0.00	Time	0.00	Time	0.00
Time	0.00	Time	0.00	Triumph	0.00	Triumph	0.00
Triumph	0.00	Triumph	0.00	Venture	0.00	Venture	0.00
Venture	0.00	Vision	0.00	Vision	0.00	Vision	0.00
Vision	0.00	Way	0.00	Warning	0.00	Way	0.00
Warning	0.00	Word	0.00	Way	0.00	Word	0.00
Work	0.00	Work	0.00	Word	0.00	Work	0.00

CS&CS	%	PDC.ET.AM	%	Ø (%)	%
Prejudice	18.18	Contradiction	12.50	Foreboding	100.00
Philosophy	9.09	Triumph	12.50	Leave	100.00
Dimension	8.33	Philosophy	4.55	Motivation	100.00
Venture	6.90	Capacity	4.35	Recollection	100.00
System	2.94	Failure	4.35	Opposite	88.89
Area	2.50	Myth	3.33	Detail	87.50
Example	2.50	Detail	2.50	Time	87.50

Objective	2.50	Project	2.50	Failure	86.96
Practice	2.50	Anger	0.00	Chance	85.00
Anger	0.00	Answer	0.00	Suspicion	84.00
Answer	0.00	Application	0.00	Evidence	80.00
Application	0.00	Area	0.00	Surprise	80.00
Assessment	0.00	Assessment	0.00	Problem	77.50
Capacity	0.00	Challenge	0.00	Endorsement	75.00
Challenge	0.00	Chance	0.00	Terror	75.00
Chance	0.00	Characteristic	0.00	Answer	72.50
Characteristic	0.00	Correction	0.00	Word	72.50
Contradiction	0.00	Crime	0.00	Recommendation	71.88
Correction	0.00	Dimension	0.00	Facet	71.43
Crime	0.00	Endeavour	0.00	Crime	70.00
Detail	0.00	Endorsement	0.00	Joke	70.00
Endeavour	0.00	Evidence	0.00	Finding	67.86
Endorsement	0.00	Example	0.00	Challenge	67.65
Evidence	0.00	Experience	0.00	Way	67.50
Experience	0.00	Facet	0.00	Myth	66.67
Facet	0.00	Finding	0.00	Part	65.63
Failure	0.00	Foreboding	0.00	Assessment	65.38
Finding	0.00	Impetus	0.00	Experience	65.00
Foreboding	0.00	Irony	0.00	Point	65.00
Impetus	0.00	Joke	0.00	Prejudice	63.64
Irony	0.00	Leave	0.00	Application	62.50
Joke	0.00	Misfortune	0.00	Example	62.50
Leave	0.00	Motivation	0.00	Anger	60.00
Misfortune	0.00	Objective	0.00	Objective	60.00
Motivation	0.00	Opposite	0.00	Quest	60.00
Myth	0.00	Part	0.00	Warning	60.00
Opposite	0.00	Phenomenon	0.00	Sense	59.09
Part	0.00	Point	0.00	Vision	58.62
Phenomenon	0.00	Practice	0.00	Irony	57.14
Point	0.00	Prejudice	0.00	Misfortune	57.14
Problem	0.00	Problem	0.00	Capacity	56.52
Project	0.00	Proviso	0.00	Work	55.56
Proviso	0.00	Quest	0.00	Area	55.00
Quest	0.00	Recollection	0.00	Philosophy	54.55
Recollection	0.00	Recommendation	0.00	Scandal	53.85
Recommendation	0.00	Scandal	0.00	Contradiction	50.00
Scandal	0.00	Sense	0.00	Correction	50.00
Sense	0.00	Surprise	0.00	Phenomenon	50.00
Surprise	0.00	Suspicion	0.00	Thing	50.00
Suspicion	0.00	System	0.00	Triumph	50.00
Terror	0.00	Terror	0.00	Project	45.00
Testimony	0.00	Testimony	0.00	Endeavour	44.44
Thing	0.00	Thing	0.00	System	44.12
Time	0.00	Time	0.00	Testimony	42.86
Triumph	0.00	Venture	0.00	Characteristic	41.38
Vision	0.00	Vision	0.00	Practice	32.50
Warning	0.00	Warning	0.00	Proviso	28.57
Way	0.00	Way	0.00	Impetus	16.67
Word	0.00	Word	0.00	Dimension	8.33
Work	0.00	Work	0.00	Venture	6.90

8.3 *Postmodifying structures (top ten)*

PP	%	RT.RV.CL	%	AP.TI.CL	%	PL.ED.CL	%
Endorsement	100.00	Time	40.00	Capacity	52.17	Anger	20.00
Part	81.25	Thing	25.00	Endeavour	44.44	Testimony	14.29
Opposite	80.00	Challenge	20.59	Failure	39.13	Challenge	11.76
Motivation	75.00	System	20.59	Chance	25.00	System	5.88
Assessment	65.38	Area	20.00	Leave	25.00	Objective	5.00
Quest	60.00	Impetus	16.67	Misfortune	14.29	Problem	5.00
Recollection	60.00	Point	12.50	Motivation	12.50	Capacity	4.35
Characteristic	58.62	Way	12.50	Application	5.00	Assessment	3.85
Facet	57.14	Word	12.50	Way	5.00	Work	3.70
Testimony	57.14	Problem	10.00	Sense	4.55	Characteristic	3.45
Correction	50.00	Philosophy	9.09	Venture	3.45	Venture	3.45
Example	50.00	Phenomenon	8.33	Anger	0.00	Myth	3.33
Triumph	50.00	Work	7.41	Answer	0.00	Phenomenon	2.78
Chance	47.50	Finding	7.14	Area	0.00	Crime	2.50
Detail	40.00	Vision	6.90	Assessment	0.00	Example	2.50
Contradiction	37.50	Recommendation	6.25	Challenge	0.00	Point	2.50
Word	37.50	Answer	5.00	Characteristic	0.00	Project	2.50
Sense	36.36	Detail	5.00	Contradiction	0.00	Answer	0.00
Challenge	35.29	Evidence	5.00	Correction	0.00	Application	0.00
Application	35.00	Project	5.00	Crime	0.00	Area	0.00
Evidence	35.00	Failure	4.35	Detail	0.00	Chance	0.00
Experience	35.00	Characteristic	3.45	Dimension	0.00	Contradiction	0.00
Way	35.00	Application	2.50	Endorsement	0.00	Correction	0.00
Finding	32.14	Crime	2.50	Evidence	0.00	Detail	0.00
Venture	31.03	Experience	2.50	Example	0.00	Dimension	0.00
Vision	31.03	Joke	2.50	Experience	0.00	Endeavour	0.00
Practice	30.00	Objective	2.50	Facet	0.00	Endorsement	0.00
Suspicion	28.00	Practice	2.50	Finding	0.00	Evidence	0.00
Crime	27.50	Anger	0.00	Foreboding	0.00	Experience	0.00
Objective	27.50	Assessment	0.00	Impetus	0.00	Facet	0.00
Point	27.50	Capacity	0.00	Irony	0.00	Failure	0.00
Area	25.00	Chance	0.00	Joke	0.00	Finding	0.00
Leave	25.00	Contradiction	0.00	Myth	0.00	Foreboding	0.00
Terror	25.00	Correction	0.00	Objective	0.00	Impetus	0.00
Problem	22.50	Dimension	0.00	Opposite	0.00	Irony	0.00
Work	22.22	Endeavour	0.00	Part	0.00	Joke	0.00
Recommendation	21.88	Endorsement	0.00	Phenomenon	0.00	Leave	0.00
Myth	20.00	Example	0.00	Philosophy	0.00	Misfortune	0.00
System	17.65	Facet	0.00	Point	0.00	Motivation	0.00
Answer	17.50	Foreboding	0.00	Practice	0.00	Opposite	0.00
Capacity	17.39	Irony	0.00	Prejudice	0.00	Part	0.00
Dimension	16.67	Leave	0.00	Problem	0.00	Philosophy	0.00
Impetus	16.67	Misfortune	0.00	Project	0.00	Practice	0.00
Scandal	15.38	Motivation	0.00	Proviso	0.00	Prejudice	0.00
Project	15.00	Myth	0.00	Quest	0.00	Proviso	0.00
Irony	14.29	Opposite	0.00	Recollection	0.00	Quest	0.00
Philosophy	13.64	Part	0.00	Recommendation	0.00	Recollection	0.00
Time	12.50	Prejudice	0.00	Scandal	0.00	Recommendation	0.00
Endeavour	11.11	Proviso	0.00	Surprise	0.00	Scandal	0.00
Warning	10.34	Quest	0.00	Suspicion	0.00	Sense	0.00
Surprise	8.00	Recollection	0.00	System	0.00	Surprise	0.00
Joke	5.00	Scandal	0.00	Terror	0.00	Suspicion	0.00
Thing	5.00	Sense	0.00	Testimony	0.00	Terror	0.00

Phenomenon	2.78	Surprise	0.00	Thing	0.00	Thing	0.00
Anger	0.00	Suspicion	0.00	Time	0.00	Time	0.00
Failure	0.00	Terror	0.00	Triumph	0.00	Triumph	0.00
Foreboding	0.00	Testimony	0.00	Vision	0.00	Vision	0.00
Misfortune	0.00	Triumph	0.00	Warning	0.00	Warning	0.00
Prejudice	0.00	Venture	0.00	Word	0.00	Way	0.00
Proviso	0.00	Warning	0.00	Work	0.00	Word	0.00

AP.THAT.CL	%	TI.CL	%	PT.WK.NR.AP	%	NR.RV.CL	%
Proviso	28.57	Way	12.50	Proviso	14.29	Experience	5.00
Suspicion	28.00	Prejudice	9.09	Prejudice	9.09	System	2.94
Evidence	15.00	Evidence	5.00	Objective	5.00	Chance	2.50
Warning	13.79	Joke	5.00	Surprise	4.00	Crime	2.50
Philosophy	9.09	Work	3.70	Chance	2.50	Example	2.50
Sense	4.55	Characteristic	3.45	Example	2.50	Anger	0.00
Surprise	4.00	Part	3.13	Joke	2.50	Answer	0.00
Myth	3.33	Phenomenon	2.78	Thing	2.50	Application	0.00
Anger	0.00	Application	2.50	Anger	0.00	Area	0.00
Answer	0.00	Point	2.50	Answer	0.00	Assessment	0.00
Application	0.00	Thing	2.50	Application	0.00	Capacity	0.00
Area	0.00	Time	2.50	Area	0.00	Challenge	0.00
Assessment	0.00	Anger	0.00	Assessment	0.00	Characteristic	0.00
Capacity	0.00	Answer	0.00	Capacity	0.00	Contradiction	0.00
Challenge	0.00	Area	0.00	Challenge	0.00	Correction	0.00
Chance	0.00	Assessment	0.00	Characteristic	0.00	Detail	0.00
Characteristic	0.00	Capacity	0.00	Contradiction	0.00	Dimension	0.00
Contradiction	0.00	Challenge	0.00	Correction	0.00	Endeavour	0.00
Correction	0.00	Chance	0.00	Crime	0.00	Endorsement	0.00
Crime	0.00	Contradiction	0.00	Detail	0.00	Evidence	0.00
Detail	0.00	Correction	0.00	Dimension	0.00	Facet	0.00
Dimension	0.00	Crime	0.00	Endeavour	0.00	Failure	0.00
Endeavour	0.00	Detail	0.00	Endorsement	0.00	Finding	0.00
Endorsement	0.00	Dimension	0.00	Evidence	0.00	Foreboding	0.00
Example	0.00	Endeavour	0.00	Experience	0.00	Impetus	0.00
Experience	0.00	Endorsement	0.00	Facet	0.00	Irony	0.00
Facet	0.00	Example	0.00	Failure	0.00	Joke	0.00
Failure	0.00	Experience	0.00	Finding	0.00	Leave	0.00
Finding	0.00	Facet	0.00	Foreboding	0.00	Misfortune	0.00
Foreboding	0.00	Failure	0.00	Impetus	0.00	Motivation	0.00
Impetus	0.00	Finding	0.00	Irony	0.00	Myth	0.00
Irony	0.00	Foreboding	0.00	Leave	0.00	Objective	0.00
Joke	0.00	Impetus	0.00	Misfortune	0.00	Opposite	0.00
Leave	0.00	Irony	0.00	Motivation	0.00	Part	0.00
Misfortune	0.00	Leave	0.00	Myth	0.00	Phenomenon	0.00
Motivation	0.00	Misfortune	0.00	Opposite	0.00	Philosophy	0.00
Objective	0.00	Motivation	0.00	Part	0.00	Point	0.00
Opposite	0.00	Myth	0.00	Phenomenon	0.00	Practice	0.00
Part	0.00	Objective	0.00	Philosophy	0.00	Prejudice	0.00
Phenomenon	0.00	Opposite	0.00	Point	0.00	Problem	0.00
Point	0.00	Philosophy	0.00	Practice	0.00	Project	0.00
Practice	0.00	Practice	0.00	Problem	0.00	Proviso	0.00
Prejudice	0.00	Problem	0.00	Project	0.00	Quest	0.00
Problem	0.00	Project	0.00	Quest	0.00	Recollection	0.00
Project	0.00	Proviso	0.00	Recollection	0.00	Recommendation	0.00
Quest	0.00	Quest	0.00	Recommendation	0.00	Scandal	0.00

Recollection	0.00	Recollection	0.00	Scandal	0.00	Sense	0.00
Recommendation	0.00	Recommendation	0.00	Sense	0.00	Surprise	0.00
Scandal	0.00	Scandal	0.00	Suspicion	0.00	Suspicion	0.00
System	0.00	Sense	0.00	System	0.00	Terror	0.00
Terror	0.00	Surprise	0.00	Terror	0.00	Testimony	0.00
Testimony	0.00	Suspicion	0.00	Testimony	0.00	Thing	0.00
Thing	0.00	System	0.00	Time	0.00	Time	0.00
Time	0.00	Terror	0.00	Triumph	0.00	Triumph	0.00
Triumph	0.00	Testimony	0.00	Venture	0.00	Venture	0.00
Venture	0.00	Triumph	0.00	Vision	0.00	Vision	0.00
Vision	0.00	Venture	0.00	Warning	0.00	Warning	0.00
Way	0.00	Vision	0.00	Way	0.00	Way	0.00
Word	0.00	Warning	0.00	Word	0.00	Word	0.00
Work	0.00	Word	0.00	Work	0.00	Work	0.00

PL.ING.CL	%	PP^AP.TI.CL	%	Ø	%
Scandal	7.69	Failure	21.74	Foreboding	100.00
Work	3.70	Capacity	4.35	Irony	85.71
Venture	3.45	Anger	0.00	Misfortune	85.71
Warning	3.45	Answer	0.00	Joke	85.00
Practice	2.50	Application	0.00	Surprise	84.00
Project	2.50	Area	0.00	Prejudice	81.82
Anger	0.00	Assessment	0.00	Phenomenon	80.56
Answer	0.00	Challenge	0.00	Anger	80.00
Application	0.00	Chance	0.00	Dimension	79.17
Area	0.00	Characteristic	0.00	Answer	75.00
Assessment	0.00	Contradiction	0.00	Terror	75.00
Capacity	0.00	Correction	0.00	Myth	70.00
Challenge	0.00	Crime	0.00	Project	70.00
Chance	0.00	Detail	0.00	Philosophy	63.64
Characteristic	0.00	Dimension	0.00	Contradiction	62.50
Contradiction	0.00	Endeavour	0.00	Crime	62.50
Correction	0.00	Endorsement	0.00	Recommendation	62.50
Crime	0.00	Evidence	0.00	Warning	62.07
Detail	0.00	Example	0.00	Scandal	61.54
Dimension	0.00	Experience	0.00	Objective	60.00
Endeavour	0.00	Facet	0.00	Practice	60.00
Endorsement	0.00	Finding	0.00	Problem	60.00
Evidence	0.00	Foreboding	0.00	Thing	60.00
Example	0.00	Impetus	0.00	Finding	57.14
Experience	0.00	Irony	0.00	Proviso	57.14
Facet	0.00	Joke	0.00	Detail	55.00
Failure	0.00	Leave	0.00	System	52.94
Finding	0.00	Misfortune	0.00	Point	52.50
Foreboding	0.00	Motivation	0.00	Venture	51.72
Impetus	0.00	Myth	0.00	Vision	51.72
Irony	0.00	Objective	0.00	Correction	50.00
Joke	0.00	Opposite	0.00	Experience	50.00
Leave	0.00	Part	0.00	Impetus	50.00
Misfortune	0.00	Phenomenon	0.00	Leave	50.00
Motivation	0.00	Philosophy	0.00	Sense	50.00
Myth	0.00	Point	0.00	Work	48.15
Objective	0.00	Practice	0.00	Application	47.50
Opposite	0.00	Prejudice	0.00	Area	45.00
Part	0.00	Problem	0.00	Endeavour	44.44
Phenomenon	0.00	Project	0.00	Time	42.50

Philosophy	0.00	Proviso	0.00	Word	42.50
Point	0.00	Quest	0.00	Evidence	40.00
Prejudice	0.00	Recollection	0.00	Quest	40.00
Problem	0.00	Recommendation	0.00	Recollection	40.00
Proviso	0.00	Scandal	0.00	Suspicion	40.00
Quest	0.00	Sense	0.00	Triumph	37.50
Recollection	0.00	Surprise	0.00	Example	32.50
Recommendation	0.00	Suspicion	0.00	Failure	30.43
Sense	0.00	System	0.00	Challenge	29.41
Surprise	0.00	Terror	0.00	Testimony	28.57
Suspicion	0.00	Testimony	0.00	Characteristic	27.59
System	0.00	Thing	0.00	Way	27.50
Terror	0.00	Time	0.00	Assessment	26.92
Testimony	0.00	Triumph	0.00	Chance	22.50
Thing	0.00	Venture	0.00	Capacity	21.74
Time	0.00	Vision	0.00	Opposite	20.00
Triumph	0.00	Warning	0.00	Facet	14.29
Vision	0.00	Way	0.00	Motivation	12.50
Way	0.00	Word	0.00	Part	6.25
Word	0.00	Work	0.00	Endorsement	0.00

APPENDIX 9 SYNTACTIC FUNCTIONS IN THE STUDY SAMPLE (CLAUSE- AND PHRASE-LEVEL) (COMPLETE LIST)

Syntactic function (specific) %			
DO	27.16	CP.*BY*(<AB.AT.TM)	0.07
SB	20.46	CP.*BY*(<CP.*WITHOUT*<AB.AT.MAN)	0.07
SCL	12.23	CP.*BY*(<SB.AT)	0.07
SB(not)	4.22	CP.*BY*(<SCL)	0.07
PO	2.76	CP.*CONCERNING*(<DO)	0.07
NA	1.94	CP.*CONTRARY TO*(<AB.AT.MAN)	0.07
CP.*OF*(<DO)	1.66	CP.*DUE TO*(<AB.AT.CAU)	0.07
CP.*IN*(<AB.AT.SPA)	1.38	CP.*FOLLOWING*(<AB.AT.CAU)	0.07
CP.*OF*(<SB)	1.17	CP.*FOLLOWING*(<AB.AT.TM)	0.07
CP.*IN*(<AB.AT.RS)	1.04	CP.*FOR*(<PO)	0.07
CP.*AT*(<AB.AT.TM)	0.83	CP.*FOR*(<SB(not))	0.07
CP.*BY*(<AB.AT.AG)	0.76	CP.*FOR*(<SCL)	0.07
CP.*OF*(<SCL)	0.69	CP.*FROM*(<AB.AT.CAU)	0.07
CP.*ON*(<AB.AT.RS)	0.69	CP.*FROM*(<AB.AT.MS)	0.07
CP.*FOR*(<AB.AT.PU)	0.62	CP.*FROM*(<AB.AT.RS)	0.07
CP.*TO*(<AB.AT.RS)	0.62	CP.*FROM*(<CP.*OF*<CP.*WITH*<SB)	0.07
CP.*AS*(<PSCL)	0.55	CP.*FROM*(<AP.NR<AB.AT.AG)	0.07
CP.*WITH*(<AB.AT.MS)	0.55	CP.*FROM*(<SB)	0.07
AB.AT.TM	0.55	CP.*FROM*(<SJ.VI)	0.07
CP.*FOR*(<DO)	0.48	CP.*IN RESPECT OF*(<AB.AT.RS)	0.07
CP.*FROM*(<AB.AT.SPA)	0.48	CP.*IN*(<AB.DJ.CM.CL)	0.07
CP.*IN*(<AB.AT.MS)	0.48	CP.*IN*(<CP.*AS*<POCL)	0.07
CP.*WITH*(<AB.AT.RS)	0.48	CP.*IN*(<CP.*FOR*<AB.AT.PU)	0.07
CP.*AS*(<POCL)	0.41	CP.*IN*(<CP.*OF*<CP.*FROM*<AB.AT.MS)	0.07
CP.*IN*(<AB.AT.MAN)	0.41	CP.*IN*(<SB(not))	0.07
CP.*WITH*(<AB.AT.MAN)	0.41	CP.*INTO*(<AB.AT.RU)	0.07
CP.*WITH*(<SCL)	0.41	CP.*INTO*(<AB.AT.SPA)	0.07
CP.*FOR*(<SB)	0.35	CP.*INTO*(<CP.*FOR*<AB.AT.CAU)	0.07
CP.*IN*(<DO)	0.35	CP.*INTO*(<POCL)	0.07
CP.*IN*(<SB)	0.35	CP.*INTO*(<SCL)	0.07
CP.*ABOUT*(<AB.AT.RS)	0.28	CP.*OF*(<CP.*UPON*<AB.AT.TM)	0.07
CP.*DESPITE*(<AB.AT.CC)	0.28	CP.*OF*(<AB.AT.CAU)	0.07
CP.*FOR*(<AB.AT.CAU)	0.28	CP.*OF*(<AB.AT.TM)	0.07
CP.*IN*(<AB.AT.PU)	0.28	CP.*OF*(<CF.*FOR*<DO)	0.07
CP.*OF*(<AB.AT.RS)	0.28	CP.*OF*(<CP.*ABOUT*<DO)	0.07
CP.*OF*(<SB(not))	0.28	CP.*OF*(<CP.*AS*<CP.*FOLLOWING*<AB.AT.TM)	0.07
CP.*ON*(<DO)	0.28	CP.*OF*(<CP.*AS*<PSCL)	0.07
CP.*TO*(<SCL)	0.28	CP.*OF*(<CP.*AT*<DO)	0.07
CP.*WITH*(<AB.AT.CAU)	0.28	CP.*OF*(<CP.*AT*<SCL)	0.07
CP.*WITH*(<AB.AT.CG.CI)	0.28	CP.*OF*(<CP.*BY*<AB.AT.AG)	0.07
AP.NR(<PO)	0.28	CP.*OF*(<CP.*FOR*<DO)	0.07
CP.*AFTER*(<AB.AT.TM)	0.21	CP.*OF*(<CP.*FOR*<SB)	0.07
CP.*BECAUSE OF*(<AB.AT.CAU)	0.21	CP.*OF*(<CP.*FROM*<DO)	0.07
CP.*DURING*(<AB.AT.TM)	0.21	CP.*OF*(<CP.*FROM*<SCL)	0.07
CP.*IN*(<AB.AT.MAN.RL)	0.21	CP.*OF*(<CP.*IN SPITE OF*<AB.AT.CC)	0.07
CP.*IN*(<AB.AT.TM)	0.21	CP.*OF*(<CP.*IN*<AB.AT.RS)	0.07
CP.*IN*(<SCL)	0.21	CP.*OF*(<CP.*IN*<AB.AT.SPA)	0.07
CP.*OF*(<PO)	0.21	CP.*OF*(<CP.*IN*<SB)	0.07
CP.*TO*(<SB)	0.21	CP.*OF*(<CP.*OF*<DO)	0.07
OCL	0.21	CP.*OF*(<CP.*TO*<AB.AT.RS)	0.07

CP.*BY*(<AB.AT.MS)	0.14	CP.*OF*(<CP.*TO*<CP.*OF*<DO)	0.07
CP.*BY*(<SB)	0.14	CP.*OF*(<CP.*TO*<SB)	0.07
CP.*FOR*(<AB.AT.RE)	0.14	CP.*OF*(<CP.*WITH*<AB.AT.MAN)	0.07
CP.*FOR*(<NA)	0.14	CP.*OF*(<CP.*WITH*<AB.AT.RS)	0.07
CP.*FROM*(<AB.AT.MAN)	0.14	CP.*OF*(<CP.*WITH*<DO)	0.07
CP.*IN*(<NA)	0.14	CP.*OF*(<SB.AT)	0.07
CP.*INCLUDING*(<AP.NR<SB)	0.14	CP.*ON*(<AB.AT.SPA)	0.07
CP.*LIKE*(<PSCL)	0.14	CP.*ON*(<AB.AT.TM)	0.07
CP.*OF*(<CP.*WITH*<SCL)	0.14	CP.*ON*(<PD.AT.MAT)	0.07
CP.*OF*(<NA)	0.14	CP.*OVER*(<AB.AT.CAU)	0.07
CP.*OF*(<AP.NR<DO)	0.14	CP.*PLUS*(<AB.AT.AC)	0.07
CP.*ON*(<SB)	0.14	CP.*TO*(<AB.AT.TM)	0.07
CP.*ON*(<SCL)	0.14	CP.*TO*(<AB.DJ.CO)	0.07
CP.*THROUGH*(<AB.AT.MS)	0.14	CP.*TO*(<CP.*FOR*<SB)	0.07
CP.*TO*(<DO)	0.14	CP.*TO*(<NA)	0.07
CP.*TO*(<IO)	0.14	CP.*TO*(<POCL)	0.07
CP.*UNTIL*(<AB.AT.TM)	0.14	CP.*UNLIKE*(<AB.AT.MAN)	0.07
CP.*WITH*(<AP.NR<SCL)	0.14	CP.*UPON*(<AB.AT.RS)	0.07
CP.*WITH*(<DO)	0.14	CP.*UPON*(<SB)	0.07
CP.*WITHOUT*(<AB.AT.CC)	0.14	CP.*WITH REGARD TO*(<AB.AT.RS)	0.07
AB.AT.MAN	0.14	CP.*WITH*(<AB.AT.TM)	0.07
AB.DJ.SY.ML.CM	0.14	CP.*WITH*(<CP.*AROUND*<AB.AT.SPA)	0.07
AP.NR(<DO)	0.14	CP.*WITH*(<CP.*OF*<SB)	0.07
AP.NR(CP.*OF*<CP.*IN*<AB.AT.TM)	0.07	CP.*WITH*(<CP.*ON*<CP.*CONTRARY TO*<AB.AT.MAN)	0.07
AP.NR(<CP.*OF*<PO)	0.07	CP.*WITH*(<AP.NR<SB)	0.07
CP.*ABOUT*(<DO)	0.07	CP.*WITH*(<SB)	0.07
CP.*ABOUT*(<PD.AT.MAT)	0.07	CP.*WITHIN*(<AB.AT.SPA)	0.07
CP.*ABOUT*(<SCL)	0.07	CP.*WITHIN*(<PO)	0.07
CP.*ACROSS*(<AB.AT.SPA)	0.07	AP.NR(<CP.*AT*<IO)	0.07
CP.*AGAINST*(<AB.AT.RS)	0.07	AP.NR(<CP.*TO*<AB.AT.SPA)	0.07
CP.*AGAINST*(<DO)	0.07	AP.NR(<CP.*OF*<DO&CP.*ON*<AB.AT.RS)	0.07
CP.*AS*(<AB.AT.MAN)	0.07	AP.NR(<CP.*FOR*<DO)	0.07
CP.*AS*(<AB.AT.MAN.RL)	0.07	AP.NR(<CP.*FROM*<SCL)	0.07
CP.*AS*(<CP.*OF*<DO)	0.07	AP.NR(<CP.*ON*<DO)	0.07
CP.*AT*(<DO)	0.07	AB.AT.PU	0.07
CP.*BECAUSE OF*(<AB.AT.RE)	0.07	AB.AT.RS	0.07
CP.*BEHIND*(<AB.AT.SPA)	0.07	AB.DJ.CO	0.07
CP.*BETWEEN*(<AB.AT.RE)	0.07	AP.NR(<SCL)	0.07
CP.*BETWEEN*(<DO)	0.07	AP.NR(<AB.AT.AG)	0.07
CP.*BETWEEN*(<SB)	0.07	SCL&DO	0.07

APPENDIX 10 LEMMA DISTRIBUTION OF SYNTACTIC FUNCTIONS, PARTICIPANT TYPES, CIRCUMSTANCES AND THEME/RHEME (COMPLETE LISTS)

10.1 *Syntactic functions (clause- and phrase-level; top six)*

DO	%	SB	%	SCL	%	SB(not)	%
Correction	100.00	Facet	42.86	Opposite	70.00	Evidence	30.00
Impetus	100.00	Misfortune	42.86	Surprise	44.00	Facet	14.29
Chance	52.50	Recollection	40.00	Testimony	42.86	Contradiction	12.50
Dimension	50.00	Contradiction	37.50	Recollection	40.00	Motivation	12.50
Endorsement	50.00	Motivation	37.50	Misfortune	28.57	Point	12.50
Scandal	46.15	Prejudice	36.36	Part	28.13	Time	12.50
Suspicion	44.00	Problem	35.00	Motivation	25.00	Triumph	12.50
Recommendation	43.75	Thing	35.00	Venture	24.14	Suspicion	12.00
Detail	42.50	Failure	34.78	Area	22.50	Way	10.00
Joke	42.50	Assessment	34.62	Joke	22.50	Surprise	8.00
Answer	40.00	Part	34.38	Failure	21.74	Chance	7.50
Quest	40.00	Example	32.50	Point	20.00	Crime	7.50
Word	40.00	Challenge	32.35	Word	20.00	Example	7.50
Sense	36.36	Vision	31.03	Answer	17.50	Problem	7.50
Application	35.00	Phenomenon	30.56	Example	17.50	Thing	7.50
Experience	32.50	Finding	28.57	Problem	17.50	Myth	6.67
Objective	32.50	Irony	28.57	Thing	17.50	Practice	5.00
System	32.35	Warning	27.59	Characteristic	17.24	Sense	4.55
Surprise	32.00	Answer	27.50	Vision	17.24	Capacity	4.35
Characteristic	31.03	Chance	27.50	Scandal	15.38	Dimension	4.17
Point	30.00	Objective	27.50	Practice	15.00	Work	3.70
Project	30.00	Philosophy	27.27	Challenge	14.71	System	2.94
Vision	27.59	Myth	26.67	Facet	14.29	Answer	2.50
Warning	27.59	Capacity	26.09	Irony	14.29	Application	2.50
Problem	27.50	Endorsement	25.00	Phenomenon	13.89	Area	2.50
Challenge	26.47	Leave	25.00	Warning	13.79	Detail	2.50
Work	25.93	System	23.53	Assessment	11.54	Experience	2.50
Contradiction	25.00	Point	22.50	Endeavour	11.11	Joke	2.50
Leave	25.00	Endeavour	22.22	Finding	10.71	Objective	2.50
Way	25.00	Recommendation	21.88	Experience	10.00	Word	2.50
Assessment	23.08	Characteristic	20.69	Philosophy	9.09	Anger	0.00
Crime	22.50	Anger	20.00	Prejudice	9.09	Assessment	0.00
Thing	22.50	Application	20.00	System	8.82	Challenge	0.00
Capacity	21.74	Way	20.00	Capacity	8.70	Characteristic	0.00
Evidence	20.00	Detail	15.00	Dimension	8.33	Correction	0.00
Practice	20.00	Project	15.00	Crime	7.50	Endeavour	0.00
Part	18.75	Work	14.81	Myth	6.67	Endorsement	0.00
Philosophy	18.18	Proviso	14.29	Time	5.00	Failure	0.00
Prejudice	18.18	Testimony	14.29	Work	3.70	Finding	0.00
Finding	17.86	Venture	13.79	Chance	2.50	Foreboding	0.00
Example	17.50	Dimension	12.50	Detail	2.50	Impetus	0.00
Phenomenon	16.67	Experience	12.50	Objective	2.50	Irony	0.00
Facet	14.29	Terror	12.50	Way	2.50	Leave	0.00
Foreboding	14.29	Word	12.50	Anger	0.00	Misfortune	0.00
Misfortune	14.29	Area	10.00	Application	0.00	Opposite	0.00
Proviso	14.29	Opposite	10.00	Contradiction	0.00	Part	0.00
Motivation	12.50	Practice	10.00	Correction	0.00	Phenomenon	0.00
Terror	12.50	Suspicion	8.00	Endorsement	0.00	Philosophy	0.00
Triumph	12.50	Scandal	7.69	Evidence	0.00	Prejudice	0.00

Endeavour	11.11	Crime	7.50	Foreboding	0.00	Project	0.00
Area	10.00	Joke	7.50	Impetus	0.00	Proviso	0.00
Opposite	10.00	Evidence	5.00	Leave	0.00	Quest	0.00
Time	10.00	Time	2.50	Project	0.00	Recollection	0.00
Failure	8.70	Correction	0.00	Proviso	0.00	Recommendation	0.00
Myth	6.67	Foreboding	0.00	Quest	0.00	Scandal	0.00
Anger	0.00	Impetus	0.00	Recommendation	0.00	Terror	0.00
Irony	0.00	Quest	0.00	Sense	0.00	Testimony	0.00
Recollection	0.00	Sense	0.00	Suspicion	0.00	Venture	0.00
Testimony	0.00	Surprise	0.00	Terror	0.00	Vision	0.00
Venture	0.00	Triumph	0.00	Triumph	0.00	Warning	0.00

PO	%	Phrase level (CP)	%
Scandal	15.38	Foreboding	85.71
Detail	7.50	Terror	75.00
Experience	7.50	Proviso	71.43
Recommendation	6.25	Triumph	62.50
System	5.88	Anger	60.00
Application	5.00	Quest	60.00
Area	5.00	Sense	59.09
Crime	5.00	Irony	57.14
Evidence	5.00	Endeavour	55.56
Practice	5.00	Venture	55.17
Way	5.00	Project	52.50
Word	5.00	Leave	50.00
Philosophy	4.55	Area	47.50
Dimension	4.17	Crime	47.50
Suspicion	4.00	Time	47.50
Work	3.70	Work	44.44
Finding	3.57	Testimony	42.86
Characteristic	3.45	Evidence	40.00
Venture	3.45	Practice	40.00
Vision	3.45	Finding	39.29
Myth	3.33	Capacity	39.13
Answer	2.50	Philosophy	36.36
Example	2.50	Prejudice	36.36
Point	2.50	Phenomenon	36.11
Problem	2.50	Experience	35.00
Thing	2.50	Myth	33.33
Time	2.50	Application	32.50
Anger	0.00	Way	32.50
Assessment	0.00	Suspicion	32.00
Capacity	0.00	Failure	30.43
Challenge	0.00	Objective	30.00
Chance	0.00	Recommendation	28.13
Contradiction	0.00	Characteristic	27.59
Correction	0.00	Detail	27.50
Endeavour	0.00	Assessment	26.92
Endorsement	0.00	Contradiction	25.00
Facet	0.00	Endorsement	25.00
Failure	0.00	System	23.53
Foreboding	0.00	Challenge	23.53
Impetus	0.00	Dimension	20.83
Irony	0.00	Vision	20.69
Joke	0.00	Warning	20.69

Leave	0.00	Example	20.00
Misfortune	0.00	Recollection	20.00
Motivation	0.00	Word	20.00
Objective	0.00	Part	18.75
Opposite	0.00	Joke	17.50
Part	0.00	Surprise	16.00
Phenomenon	0.00	Scandal	15.38
Prejudice	0.00	Facet	14.29
Project	0.00	Misfortune	14.29
Proviso	0.00	Motivation	12.50
Quest	0.00	Thing	12.50
Recollection	0.00	Opposite	10.00
Sense	0.00	Problem	10.00
Surprise	0.00	Answer	7.50
Terror	0.00	Chance	7.50
Testimony	0.00	Point	7.50
Triumph	0.00	Correction	0.00
Warning	0.00	Impetus	0.00

10.2 *Participant types (top ten)*

IDR (rel.in)	%	ATT (rel.in)	%	GOA	%	IDD (rel.in)	%
Opposite	70.00	Surprise	44.00	Correction	50.00	Recollection	40.00
Endorsement	50.00	Failure	30.43	Project	30.00	Thing	30.00
Misfortune	42.86	Testimony	28.57	System	23.53	Motivation	25.00
Recollection	40.00	Joke	27.50	Challenge	20.59	Anger	20.00
Facet	28.57	Endorsement	25.00	Opposite	20.00	Example	20.00
Testimony	28.57	Endeavour	22.22	Practice	15.00	Objective	20.00
Motivation	25.00	Venture	20.69	Finding	14.29	Assessment	19.23
Characteristic	24.14	Phenomenon	19.44	Testimony	14.29	Point	17.50
Part	21.88	Vision	17.24	Application	12.50	Facet	14.29
Finding	21.43	Scandal	15.38	Contradiction	12.50	Irony	14.29
Quest	20.00	Crime	15.00	Motivation	12.50	Proviso	14.29
Answer	17.50	Irony	14.29	Part	12.50	Characteristic	13.79
Area	17.50	Characteristic	13.79	Terror	12.50	Answer	12.50
Example	17.50	Contradiction	12.50	Suspicion	12.00	Contradiction	12.50
Word	17.50	Motivation	12.50	Answer	10.00	Part	12.50
Assessment	15.38	Problem	12.50	Detail	10.00	Problem	12.50
Point	15.00	Triumph	12.50	Word	10.00	Way	12.50
Thing	15.00	Work	11.11	Recommendation	9.38	Challenge	11.76
Failure	13.04	Warning	10.34	Philosophy	9.09	Area	10.00
Detail	12.50	Experience	10.00	Phenomenon	8.33	Dimension	8.33
Practice	12.50	Myth	10.00	Chance	7.50	Phenomenon	8.33
Endeavour	11.11	Part	9.38	Example	7.50	Application	7.50
Philosophy	9.09	Recommendation	9.38	Problem	7.50	Finding	7.14
Prejudice	9.09	Prejudice	9.09	Work	7.41	Myth	6.67
Challenge	8.82	Dimension	8.33	Venture	6.90	Recommendation	6.25
System	8.82	Area	7.50	Warning	6.90	Evidence	5.00
Dimension	8.33	Practice	7.50	Crime	5.00	Philosophy	4.55
Surprise	8.00	Challenge	5.88	Evidence	5.00	Capacity	4.35
Problem	7.50	System	5.88	Objective	5.00	Suspicion	4.00
Venture	6.90	Answer	5.00	Point	5.00	Vision	3.45
Myth	6.67	Point	5.00	Thing	5.00	Chance	2.50
Application	5.00	Project	5.00	Sense	4.55	Crime	2.50
Crime	5.00	Thing	5.00	Capacity	4.35	Detail	2.50

Evidence	5.00	Philosophy	4.55	Dimension	4.17	Experience	2.50
Objective	5.00	Capacity	4.35	Characteristic	3.45	Project	2.50
Time	5.00	Assessment	3.85	Experience	2.50	Time	2.50
Way	5.00	Finding	3.57	Time	2.50	Word	2.50
Capacity	4.35	Detail	2.50	Way	2.50	Correction	0.00
Suspicion	4.00	Time	2.50	Anger	0.00	Endeavour	0.00
Work	3.70	Way	2.50	Area	0.00	Endorsement	0.00
Vision	3.45	Word	2.50	Assessment	0.00	Failure	0.00
Warning	3.45	Anger	0.00	Endeavour	0.00	Foreboding	0.00
Phenomenon	2.78	Application	0.00	Endorsement	0.00	Impetus	0.00
Chance	2.50	Chance	0.00	Facet	0.00	Joke	0.00
Experience	2.50	Correction	0.00	Failure	0.00	Leave	0.00
Project	2.50	Evidence	0.00	Foreboding	0.00	Misfortune	0.00
Anger	0.00	Example	0.00	Impetus	0.00	Opposite	0.00
Contradiction	0.00	Facet	0.00	Irony	0.00	Practice	0.00
Correction	0.00	Foreboding	0.00	Joke	0.00	Prejudice	0.00
Foreboding	0.00	Impetus	0.00	Leave	0.00	Quest	0.00
Impetus	0.00	Leave	0.00	Misfortune	0.00	Scandal	0.00
Irony	0.00	Misfortune	0.00	Myth	0.00	Sense	0.00
Joke	0.00	Objective	0.00	Prejudice	0.00	Surprise	0.00
Leave	0.00	Opposite	0.00	Proviso	0.00	System	0.00
Proviso	0.00	Proviso	0.00	Quest	0.00	Terror	0.00
Recommendation	0.00	Quest	0.00	Recollection	0.00	Testimony	0.00
Scandal	0.00	Recollection	0.00	Scandal	0.00	Triumph	0.00
Sense	0.00	Sense	0.00	Surprise	0.00	Venture	0.00
Terror	0.00	Suspicion	0.00	Triumph	0.00	Warning	0.00
Triumph	0.00	Terror	0.00	Vision	0.00	Work	0.00

VER	%	EXI	%	CAR (rel.in)	%	PHE	%
Correction	50.00	Evidence	30.00	Endorsement	25.00	Foreboding	28.57
Scandal	30.77	Facet	14.29	Quest	20.00	Contradiction	25.00
Recommendation	28.13	Contradiction	12.50	Recollection	20.00	Experience	22.50
Detail	25.00	Motivation	12.50	Prejudice	18.18	Word	22.50
Point	25.00	Point	12.50	Chance	17.50	Philosophy	18.18
Warning	20.69	Problem	12.50	Vision	17.24	Prejudice	18.18
Evidence	20.00	Terror	12.50	Facet	14.29	Answer	17.50
Joke	20.00	Time	12.50	Finding	14.29	Vision	17.24
Assessment	19.23	Triumph	12.50	Foreboding	14.29	Recommendation	15.63
Application	15.00	Surprise	12.00	Phenomenon	13.89	Scandal	15.38
Irony	14.29	Suspicion	12.00	Capacity	13.04	Phenomenon	13.89
Proviso	14.29	Detail	10.00	Motivation	12.50	Application	12.50
Terror	12.50	Example	10.00	Terror	12.50	Crime	12.50
Triumph	12.50	Myth	10.00	Assessment	11.54	Way	12.50
Example	10.00	Way	10.00	Endeavour	11.11	Surprise	12.00
Project	10.00	Philosophy	9.09	Characteristic	10.34	Joke	10.00
Thing	10.00	Chance	7.50	Warning	10.34	Practice	10.00
Word	10.00	Crime	7.50	Part	9.38	Sense	9.09
Failure	8.70	Thing	7.50	System	8.82	Assessment	7.69
Practice	7.50	System	5.88	Joke	7.50	Area	7.50
Part	6.25	Area	5.00	Way	7.50	Detail	7.50
System	5.88	Objective	5.00	Challenge	5.88	Example	7.50
Phenomenon	5.56	Practice	5.00	Application	5.00	Warning	6.90
Area	5.00	Sense	4.55	Area	5.00	System	5.88
Experience	5.00	Capacity	4.35	Detail	5.00	Evidence	5.00
Philosophy	4.55	Failure	4.35	Objective	5.00	Point	5.00

Dimension	4.17	Dimension	4.17	Philosophy	4.55	Time	5.00
Surprise	4.00	Work	3.70	Dimension	4.17	Failure	4.35
Suspicion	4.00	Challenge	2.94	Suspicion	4.00	Dimension	4.17
Finding	3.57	Phenomenon	2.78	Venture	3.45	Suspicion	4.00
Characteristic	3.45	Answer	2.50	Recommendation	3.13	Characteristic	3.45
Vision	3.45	Application	2.50	Answer	2.50	Myth	3.33
Myth	3.33	Experience	2.50	Crime	2.50	Part	3.13
Answer	2.50	Joke	2.50	Experience	2.50	Challenge	2.94
Way	2.50	Project	2.50	Practice	2.50	Chance	2.50
Anger	0.00	Word	2.50	Problem	2.50	Problem	2.50
Capacity	0.00	Anger	0.00	Project	2.50	Project	2.50
Challenge	0.00	Assessment	0.00	Thing	2.50	Thing	2.50
Chance	0.00	Characteristic	0.00	Word	2.50	Anger	0.00
Contradiction	0.00	Correction	0.00	Anger	0.00	Capacity	0.00
Crime	0.00	Endeavour	0.00	Contradiction	0.00	Correction	0.00
Endeavour	0.00	Endorsement	0.00	Correction	0.00	Endeavour	0.00
Endorsement	0.00	Finding	0.00	Evidence	0.00	Endorsement	0.00
Facet	0.00	Foreboding	0.00	Example	0.00	Facet	0.00
Foreboding	0.00	Impetus	0.00	Failure	0.00	Finding	0.00
Impetus	0.00	Irony	0.00	Impetus	0.00	Impetus	0.00
Leave	0.00	Leave	0.00	Irony	0.00	Irony	0.00
Misfortune	0.00	Misfortune	0.00	Leave	0.00	Leave	0.00
Motivation	0.00	Opposite	0.00	Misfortune	0.00	Misfortune	0.00
Objective	0.00	Part	0.00	Myth	0.00	Motivation	0.00
Opposite	0.00	Prejudice	0.00	Opposite	0.00	Objective	0.00
Prejudice	0.00	Proviso	0.00	Point	0.00	Opposite	0.00
Problem	0.00	Quest	0.00	Proviso	0.00	Proviso	0.00
Quest	0.00	Recollection	0.00	Scandal	0.00	Quest	0.00
Recollection	0.00	Recommendation	0.00	Sense	0.00	Recollection	0.00
Sense	0.00	Scandal	0.00	Surprise	0.00	Terror	0.00
Testimony	0.00	Testimony	0.00	Testimony	0.00	Testimony	0.00
Time	0.00	Venture	0.00	Time	0.00	Triumph	0.00
Venture	0.00	Vision	0.00	Triumph	0.00	Venture	0.00
Work	0.00	Warning	0.00	Work	0.00	Work	0.00

ATT (rel.po)	%	IDR (rel.ci)	%
Chance	17.50	Impetus	50.00
Foreboding	14.29	Suspicion	24.00
Misfortune	14.29	Problem	17.50
Capacity	13.04	Facet	14.29
Dimension	12.50	Testimony	14.29
Objective	12.50	Venture	10.34
Suspicion	12.00	Vision	10.34
Characteristic	10.34	Myth	10.00
Problem	10.00	Prejudice	9.09
Prejudice	9.09	Dimension	8.33
Sense	9.09	Surprise	8.00
System	8.82	Scandal	7.69
Experience	7.50	Chance	7.50
Vision	6.90	Practice	7.50
Part	6.25	Project	7.50
Answer	5.00	Application	5.00
Detail	5.00	Area	5.00
Joke	5.00	Experience	5.00
Point	5.00	Philosophy	4.55
Thing	5.00	Sense	4.55

Philosophy	4.55	Capacity	4.35
Failure	4.35	Work	3.70
Assessment	3.85	Characteristic	3.45
Venture	3.45	Part	3.13
Warning	3.45	Challenge	2.94
Recommendation	3.13	System	2.94
Application	2.50	Phenomenon	2.78
Area	2.50	Example	2.50
Project	2.50	Joke	2.50
Time	2.50	Objective	2.50
Way	2.50	Anger	0.00
Word	2.50	Answer	0.00
Anger	0.00	Assessment	0.00
Challenge	0.00	Contradiction	0.00
Contradiction	0.00	Correction	0.00
Correction	0.00	Crime	0.00
Crime	0.00	Detail	0.00
Endeavour	0.00	Endeavour	0.00
Endorsement	0.00	Endorsement	0.00
Evidence	0.00	Evidence	0.00
Example	0.00	Failure	0.00
Facet	0.00	Finding	0.00
Finding	0.00	Foreboding	0.00
Impetus	0.00	Irony	0.00
Irony	0.00	Leave	0.00
Leave	0.00	Misfortune	0.00
Motivation	0.00	Motivation	0.00
Myth	0.00	Opposite	0.00
Opposite	0.00	Point	0.00
Phenomenon	0.00	Proviso	0.00
Practice	0.00	Quest	0.00
Proviso	0.00	Recollection	0.00
Quest	0.00	Recommendation	0.00
Recollection	0.00	Terror	0.00
Scandal	0.00	Thing	0.00
Surprise	0.00	Time	0.00
Terror	0.00	Triumph	0.00
Testimony	0.00	Warning	0.00
Triumph	0.00	Way	0.00
Work	0.00	Word	0.00

10.3 *Circumstances (top ten)*

CR.MAT	%	CR.LOC.TM	%	CR.LOC.PC	%	CR.MAN.MS	%
Sense	50.00	Time	65.00	Anger	20.00	Leave	25.00
Challenge	17.65	Quest	20.00	Quest	20.00	Proviso	14.29
Recommendation	15.63	Scandal	7.69	Area	15.00	Triumph	12.50
Capacity	13.04	Experience	7.50	Facet	14.29	Way	12.50
Crime	12.50	Work	7.41	Testimony	14.29	Word	7.50
Endeavour	11.11	Evidence	5.00	Triumph	12.50	Venture	6.90
Finding	10.71	Assessment	3.85	Myth	10.00	Finding	3.57
Evidence	10.00	Warning	3.45	Venture	6.90	Warning	3.45
Practice	10.00	Recommendation	3.13	Capacity	4.35	System	2.94
Experience	7.50	Application	2.50	Suspicion	4.00	Area	2.50
Objective	7.50	Crime	2.50	Characteristic	3.45	Detail	2.50

Project	7.50	Point	2.50	Part	3.13	Experience	2.50
Thing	7.50	Anger	0.00	Phenomenon	2.78	Objective	2.50
Way	7.50	Answer	0.00	Crime	2.50	Project	2.50
System	5.88	Area	0.00	Example	2.50	Anger	0.00
Example	5.00	Capacity	0.00	Objective	2.50	Answer	0.00
Joke	5.00	Challenge	0.00	Answer	0.00	Application	0.00
Philosophy	4.55	Chance	0.00	Application	0.00	Assessment	0.00
Dimension	4.17	Characteristic	0.00	Assessment	0.00	Capacity	0.00
Venture	3.45	Contradiction	0.00	Challenge	0.00	Challenge	0.00
Phenomenon	2.78	Correction	0.00	Chance	0.00	Chance	0.00
Answer	2.50	Detail	0.00	Contradiction	0.00	Characteristic	0.00
Application	2.50	Dimension	0.00	Correction	0.00	Contradiction	0.00
Area	2.50	Endeavour	0.00	Detail	0.00	Correction	0.00
Chance	2.50	Endorsement	0.00	Dimension	0.00	Crime	0.00
Detail	2.50	Example	0.00	Endeavour	0.00	Dimension	0.00
Point	2.50	Facet	0.00	Endorsement	0.00	Endeavour	0.00
Problem	2.50	Failure	0.00	Evidence	0.00	Endorsement	0.00
Time	2.50	Finding	0.00	Experience	0.00	Evidence	0.00
Word	2.50	Foreboding	0.00	Failure	0.00	Example	0.00
Anger	0.00	Impetus	0.00	Finding	0.00	Facet	0.00
Assessment	0.00	Irony	0.00	Foreboding	0.00	Failure	0.00
Characteristic	0.00	Joke	0.00	Impetus	0.00	Foreboding	0.00
Contradiction	0.00	Leave	0.00	Irony	0.00	Impetus	0.00
Correction	0.00	Misfortune	0.00	Joke	0.00	Irony	0.00
Endorsement	0.00	Motivation	0.00	Leave	0.00	Joke	0.00
Facet	0.00	Myth	0.00	Misfortune	0.00	Misfortune	0.00
Failure	0.00	Objective	0.00	Motivation	0.00	Motivation	0.00
Foreboding	0.00	Opposite	0.00	Opposite	0.00	Myth	0.00
Impetus	0.00	Part	0.00	Philosophy	0.00	Opposite	0.00
Irony	0.00	Phenomenon	0.00	Point	0.00	Part	0.00
Leave	0.00	Philosophy	0.00	Practice	0.00	Phenomenon	0.00
Misfortune	0.00	Practice	0.00	Prejudice	0.00	Philosophy	0.00
Motivation	0.00	Prejudice	0.00	Problem	0.00	Point	0.00
Myth	0.00	Problem	0.00	Project	0.00	Practice	0.00
Opposite	0.00	Project	0.00	Proviso	0.00	Prejudice	0.00
Part	0.00	Proviso	0.00	Recollection	0.00	Problem	0.00
Prejudice	0.00	Recollection	0.00	Recommendation	0.00	Quest	0.00
Proviso	0.00	Sense	0.00	Scandal	0.00	Recollection	0.00
Quest	0.00	Surprise	0.00	Sense	0.00	Recommendation	0.00
Recollection	0.00	Suspicion	0.00	Surprise	0.00	Scandal	0.00
Scandal	0.00	System	0.00	System	0.00	Sense	0.00
Surprise	0.00	Terror	0.00	Terror	0.00	Surprise	0.00
Suspicion	0.00	Testimony	0.00	Thing	0.00	Suspicion	0.00
Terror	0.00	Thing	0.00	Time	0.00	Terror	0.00
Testimony	0.00	Triumph	0.00	Vision	0.00	Testimony	0.00
Triumph	0.00	Venture	0.00	Warning	0.00	Thing	0.00
Vision	0.00	Vision	0.00	Way	0.00	Time	0.00
Warning	0.00	Way	0.00	Word	0.00	Vision	0.00
Work	0.00	Word	0.00	Work	0.00	Work	0.00

CR.CAU.RE	%	CR.CAU.PU	%	CR.MAN.QL	%	CR.RL.GS	%
Anger	20.00	Endeavour	11.11	Irony	28.57	Capacity	13.04
Irony	14.29	Challenge	5.88	Leave	25.00	Example	5.00
Crime	12.50	Area	5.00	Anger	20.00	Venture	3.45
Terror	12.50	Detail	5.00	Way	15.00	Warning	3.45
Evidence	5.00	Joke	5.00	Suspicion	12.00	Project	2.50

Philosophy	4.55	Project	5.00	Objective	2.50	Word	2.50
Capacity	4.35	Surprise	4.00	Practice	2.50	Anger	0.00
Failure	4.35	Venture	3.45	Answer	0.00	Answer	0.00
Work	3.70	Answer	2.50	Application	0.00	Application	0.00
Vision	3.45	Thing	2.50	Area	0.00	Area	0.00
Warning	3.45	Anger	0.00	Assessment	0.00	Assessment	0.00
Application	2.50	Application	0.00	Capacity	0.00	Challenge	0.00
Practice	2.50	Assessment	0.00	Challenge	0.00	Chance	0.00
Problem	2.50	Capacity	0.00	Chance	0.00	Characteristic	0.00
Thing	2.50	Chance	0.00	Characteristic	0.00	Contradiction	0.00
Way	2.50	Characteristic	0.00	Contradiction	0.00	Correction	0.00
Answer	0.00	Contradiction	0.00	Correction	0.00	Crime	0.00
Area	0.00	Correction	0.00	Crime	0.00	Detail	0.00
Assessment	0.00	Crime	0.00	Detail	0.00	Dimension	0.00
Challenge	0.00	Dimension	0.00	Dimension	0.00	Endeavour	0.00
Chance	0.00	Endorsement	0.00	Endeavour	0.00	Endorsement	0.00
Characteristic	0.00	Evidence	0.00	Endorsement	0.00	Evidence	0.00
Contradiction	0.00	Example	0.00	Evidence	0.00	Experience	0.00
Correction	0.00	Experience	0.00	Example	0.00	Facet	0.00
Detail	0.00	Facet	0.00	Experience	0.00	Failure	0.00
Dimension	0.00	Failure	0.00	Facet	0.00	Finding	0.00
Endeavour	0.00	Finding	0.00	Failure	0.00	Foreboding	0.00
Endorsement	0.00	Foreboding	0.00	Finding	0.00	Impetus	0.00
Example	0.00	Impetus	0.00	Foreboding	0.00	Irony	0.00
Experience	0.00	Irony	0.00	Impetus	0.00	Joke	0.00
Facet	0.00	Leave	0.00	Joke	0.00	Leave	0.00
Finding	0.00	Misfortune	0.00	Misfortune	0.00	Misfortune	0.00
Foreboding	0.00	Motivation	0.00	Motivation	0.00	Motivation	0.00
Impetus	0.00	Myth	0.00	Myth	0.00	Myth	0.00
Joke	0.00	Objective	0.00	Opposite	0.00	Objective	0.00
Leave	0.00	Opposite	0.00	Part	0.00	Opposite	0.00
Misfortune	0.00	Part	0.00	Phenomenon	0.00	Part	0.00
Motivation	0.00	Phenomenon	0.00	Philosophy	0.00	Phenomenon	0.00
Myth	0.00	Philosophy	0.00	Point	0.00	Philosophy	0.00
Objective	0.00	Point	0.00	Prejudice	0.00	Point	0.00
Opposite	0.00	Practice	0.00	Problem	0.00	Practice	0.00
Part	0.00	Prejudice	0.00	Project	0.00	Prejudice	0.00
Phenomenon	0.00	Problem	0.00	Proviso	0.00	Problem	0.00
Point	0.00	Proviso	0.00	Quest	0.00	Proviso	0.00
Prejudice	0.00	Quest	0.00	Recollection	0.00	Quest	0.00
Project	0.00	Recollection	0.00	Recommendation	0.00	Recollection	0.00
Proviso	0.00	Recommendation	0.00	Scandal	0.00	Recommendation	0.00
Quest	0.00	Scandal	0.00	Sense	0.00	Scandal	0.00
Recollection	0.00	Sense	0.00	Surprise	0.00	Sense	0.00
Recommendation	0.00	Suspicion	0.00	System	0.00	Surprise	0.00
Scandal	0.00	System	0.00	Terror	0.00	Suspicion	0.00
Sense	0.00	Terror	0.00	Testimony	0.00	System	0.00
Surprise	0.00	Testimony	0.00	Thing	0.00	Terror	0.00
Suspicion	0.00	Time	0.00	Time	0.00	Testimony	0.00
System	0.00	Triumph	0.00	Triumph	0.00	Thing	0.00
Testimony	0.00	Vision	0.00	Venture	0.00	Time	0.00
Time	0.00	Warning	0.00	Vision	0.00	Triumph	0.00
Triumph	0.00	Way	0.00	Warning	0.00	Vision	0.00
Venture	0.00	Word	0.00	Word	0.00	Way	0.00
Word	0.00	Work	0.00	Work	0.00	Work	0.00

CR.CG.CC	%	CR.CG.CI	%
Warning	10.34	Proviso	57.14
Suspicion	4.00	Anger	0.00
Finding	3.57	Answer	0.00
Phenomenon	2.78	Application	0.00
Experience	2.50	Area	0.00
Anger	0.00	Assessment	0.00
Answer	0.00	Capacity	0.00
Application	0.00	Challenge	0.00
Area	0.00	Chance	0.00
Assessment	0.00	Characteristic	0.00
Capacity	0.00	Contradiction	0.00
Challenge	0.00	Correction	0.00
Chance	0.00	Crime	0.00
Characteristic	0.00	Detail	0.00
Contradiction	0.00	Dimension	0.00
Correction	0.00	Endeavour	0.00
Crime	0.00	Endorsement	0.00
Detail	0.00	Evidence	0.00
Dimension	0.00	Example	0.00
Endeavour	0.00	Experience	0.00
Endorsement	0.00	Facet	0.00
Evidence	0.00	Failure	0.00
Example	0.00	Finding	0.00
Facet	0.00	Foreboding	0.00
Failure	0.00	Impetus	0.00
Foreboding	0.00	Irony	0.00
Impetus	0.00	Joke	0.00
Irony	0.00	Leave	0.00
Joke	0.00	Misfortune	0.00
Leave	0.00	Motivation	0.00
Misfortune	0.00	Myth	0.00
Motivation	0.00	Objective	0.00
Myth	0.00	Opposite	0.00
Objective	0.00	Part	0.00
Opposite	0.00	Phenomenon	0.00
Part	0.00	Philosophy	0.00
Philosophy	0.00	Point	0.00
Point	0.00	Practice	0.00
Practice	0.00	Prejudice	0.00
Prejudice	0.00	Problem	0.00
Problem	0.00	Project	0.00
Project	0.00	Quest	0.00
Proviso	0.00	Recollection	0.00
Quest	0.00	Recommendation	0.00
Recollection	0.00	Scandal	0.00
Recommendation	0.00	Sense	0.00
Scandal	0.00	Surprise	0.00
Sense	0.00	Suspicion	0.00
Surprise	0.00	System	0.00
System	0.00	Terror	0.00
Terror	0.00	Testimony	0.00
Testimony	0.00	Thing	0.00
Thing	0.00	Time	0.00
Time	0.00	Triumph	0.00
Triumph	0.00	Venture	0.00

Venture	0.00	Vision	0.00
Vision	0.00	Warning	0.00
Way	0.00	Way	0.00
Word	0.00	Word	0.00
Work	0.00	Work	0.00

10.4 *Process types*

Relational	%	Circumstance	%	Material	%	Mental	%
Endorsement	100.00	Proviso	71.43	Correction	50.00	Foreboding	28.57
Impetus	100.00	Time	67.50	Leave	50.00	Prejudice	27.27
Recollection	100.00	Anger	60.00	Terror	50.00	Contradiction	25.00
Characteristic	86.21	Sense	54.55	Work	40.74	Experience	22.50
Dimension	79.17	Leave	50.00	Project	37.50	Word	22.50
Failure	78.26	Irony	42.86	Opposite	30.00	Philosophy	18.18
Motivation	75.00	Quest	40.00	Challenge	29.41	Answer	17.50
Problem	72.50	Way	37.50	Misfortune	28.57	Vision	17.24
Facet	71.43	Capacity	34.78	System	26.47	Surprise	16.00
Foreboding	71.43	Crime	30.00	Practice	25.00	Recommendation	15.63
Misfortune	71.43	Area	25.00	Scandal	23.08	Scandal	15.38
Testimony	71.43	Finding	25.00	Objective	22.50	Application	15.00
Opposite	70.00	Triumph	25.00	Application	20.00	Phenomenon	13.89
Part	68.75	Venture	24.14	Crime	20.00	Crime	12.50
Chance	67.50	Warning	24.14	Answer	17.50	Way	12.50
Vision	62.07	Challenge	23.53	Part	15.63	Joke	10.00
Surprise	60.00	Endeavour	22.22	Chance	15.00	Practice	10.00
Assessment	57.69	Evidence	20.00	Word	15.00	Sense	9.09
Thing	57.50	Experience	20.00	Finding	14.29	Assessment	7.69
Finding	57.14	Suspicion	20.00	Testimony	14.29	Area	7.50
Endeavour	55.56	Recommendation	18.75	Phenomenon	13.89	Detail	7.50
Venture	55.17	Objective	17.50	Venture	13.79	Example	7.50
Prejudice	54.55	Project	17.50	Capacity	13.04	Warning	6.90
Myth	53.33	Example	15.00	Contradiction	12.50	System	5.88
Phenomenon	52.78	Practice	15.00	Motivation	12.50	Evidence	5.00
Answer	52.50	Word	15.00	Triumph	12.50	Point	5.00
Area	52.50	Facet	14.29	Suspicion	12.00	Time	5.00
Contradiction	50.00	Testimony	14.29	Endeavour	11.11	Failure	4.35
Objective	50.00	Philosophy	13.64	Detail	10.00	Dimension	4.17
Suspicion	48.00	Myth	13.33	Evidence	10.00	Suspicion	4.00
Capacity	47.83	Detail	12.50	Example	10.00	Characteristic	3.45
Example	47.50	Terror	12.50	Recommendation	9.38	Myth	3.33
Irony	42.86	Thing	12.50	Philosophy	9.09	Part	3.13
Experience	42.50	Phenomenon	11.11	Prejudice	9.09	Challenge	2.94
Joke	42.50	Work	11.11	Sense	9.09	Chance	2.50
Point	42.50	Joke	10.00	Assessment	7.69	Problem	2.50
System	41.18	Prejudice	9.09	Experience	7.50	Project	2.50
Quest	40.00	System	8.82	Joke	7.50	Thing	2.50
Challenge	38.24	Surprise	8.00	Problem	7.50	Anger	0.00
Work	37.04	Scandal	7.69	Thing	7.50	Capacity	0.00
Philosophy	36.36	Application	7.50	Vision	6.90	Correction	0.00
Warning	34.48	Vision	6.90	Warning	6.90	Endeavour	0.00
Detail	32.50	Part	6.25	Point	5.00	Endorsement	0.00
Practice	32.50	Answer	5.00	Way	5.00	Facet	0.00
Way	32.50	Chance	5.00	Dimension	4.17	Finding	0.00
Word	32.50	Point	5.00	Characteristic	3.45	Impetus	0.00

Application	30.00	Problem	5.00	Area	2.50	Irony	0.00
Crime	30.00	Failure	4.35	Time	2.50	Leave	0.00
Recommendation	28.13	Dimension	4.17	Anger	0.00	Misfortune	0.00
Project	27.50	Assessment	3.85	Endorsement	0.00	Motivation	0.00
Triumph	25.00	Characteristic	3.45	Facet	0.00	Objective	0.00
Scandal	23.08	Contradiction	0.00	Failure	0.00	Opposite	0.00
Sense	22.73	Correction	0.00	Foreboding	0.00	Proviso	0.00
Anger	20.00	Endorsement	0.00	Impetus	0.00	Quest	0.00
Evidence	15.00	Foreboding	0.00	Irony	0.00	Recollection	0.00
Proviso	14.29	Impetus	0.00	Myth	0.00	Terror	0.00
Terror	12.50	Misfortune	0.00	Proviso	0.00	Testimony	0.00
Time	12.50	Motivation	0.00	Quest	0.00	Triumph	0.00
Correction	0.00	Opposite	0.00	Recollection	0.00	Venture	0.00
Leave	0.00	Recollection	0.00	Surprise	0.00	Work	0.00

Verbal	%	Existential	%
Correction	50.00	Evidence	30.00
Scandal	30.77	Facet	14.29
Recommendation	28.13	Contradiction	12.50
Detail	25.00	Motivation	12.50
Point	25.00	Point	12.50
Warning	20.69	Problem	12.50
Application	20.00	Terror	12.50
Evidence	20.00	Time	12.50
Joke	20.00	Triumph	12.50
Assessment	19.23	Surprise	12.00
Irony	14.29	Suspicion	12.00
Proviso	14.29	Detail	10.00
Terror	12.50	Example	10.00
Triumph	12.50	Myth	10.00
Example	10.00	Way	10.00
Project	10.00	Philosophy	9.09
Thing	10.00	Chance	7.50
Word	10.00	Crime	7.50
Failure	8.70	Thing	7.50
Practice	7.50	System	5.88
Part	6.25	Area	5.00
System	5.88	Objective	5.00
Phenomenon	5.56	Practice	5.00
Area	5.00	Sense	4.55
Experience	5.00	Capacity	4.35
Philosophy	4.55	Failure	4.35
Dimension	4.17	Dimension	4.17
Surprise	4.00	Work	3.70
Suspicion	4.00	Challenge	2.94
Work	3.70	Phenomenon	2.78
Finding	3.57	Answer	2.50
Characteristic	3.45	Application	2.50
Venture	3.45	Experience	2.50
Vision	3.45	Joke	2.50
Myth	3.33	Project	2.50
Answer	2.50	Word	2.50
Way	2.50	Anger	0.00
Anger	0.00	Assessment	0.00
Capacity	0.00	Characteristic	0.00
Challenge	0.00	Correction	0.00

Chance	0.00	Endeavour	0.00
Contradiction	0.00	Endorsement	0.00
Crime	0.00	Finding	0.00
Endeavour	0.00	Foreboding	0.00
Endorsement	0.00	Impetus	0.00
Facet	0.00	Irony	0.00
Foreboding	0.00	Leave	0.00
Impetus	0.00	Misfortune	0.00
Leave	0.00	Opposite	0.00
Misfortune	0.00	Part	0.00
Motivation	0.00	Prejudice	0.00
Objective	0.00	Proviso	0.00
Opposite	0.00	Quest	0.00
Prejudice	0.00	Recollection	0.00
Problem	0.00	Recommendation	0.00
Quest	0.00	Scandal	0.00
Recollection	0.00	Testimony	0.00
Sense	0.00	Venture	0.00
Testimony	0.00	Vision	0.00
Time	0.00	Warning	0.00

Relational intensive	%	Relational circumstantial	%	Relational possessive	%
Endorsement	100.00	Impetus	50.00	Foreboding	57.14
Recollection	100.00	Failure	26.09	Impetus	50.00
Motivation	75.00	Contradiction	25.00	Dimension	37.50
Opposite	70.00	Problem	25.00	Chance	32.50
Characteristic	62.07	Suspicion	24.00	Sense	18.18
Facet	57.14	Myth	23.33	Objective	17.50
Testimony	57.14	Experience	17.50	Misfortune	14.29
Part	53.13	Venture	17.24	Characteristic	13.79
Thing	52.50	Work	14.81	Capacity	13.04
Surprise	52.00	Facet	14.29	Problem	12.50
Assessment	50.00	Irony	14.29	Suspicion	12.00
Failure	47.83	Misfortune	14.29	System	11.76
Finding	46.43	Testimony	14.29	Part	9.38
Endeavour	44.44	Vision	13.79	Prejudice	9.09
Phenomenon	44.44	Chance	12.50	Detail	7.50
Misfortune	42.86	Dimension	12.50	Experience	7.50
Vision	41.38	Project	12.50	Work	7.41
Area	40.00	Triumph	12.50	Venture	6.90
Quest	40.00	Endeavour	11.11	Vision	6.90
Answer	37.50	Finding	10.71	Myth	6.67
Example	37.50	Characteristic	10.34	Recommendation	6.25
Point	37.50	Answer	10.00	Answer	5.00
Prejudice	36.36	Application	10.00	Joke	5.00
Joke	35.00	Area	10.00	Point	5.00
Problem	35.00	Practice	10.00	Thing	5.00
Challenge	32.35	Philosophy	9.09	Way	5.00
Venture	31.03	Prejudice	9.09	Philosophy	4.55
Objective	30.00	Capacity	8.70	Failure	4.35
Dimension	29.17	Phenomenon	8.33	Assessment	3.85
Irony	28.57	Surprise	8.00	Warning	3.45
Way	27.50	Scandal	7.69	Application	2.50
Capacity	26.09	Example	7.50	Area	2.50

Contradiction	25.00	Warning	6.90	Example	2.50
Crime	25.00	Part	6.25	Project	2.50
Word	25.00	Challenge	5.88	Time	2.50
Warning	24.14	System	5.88	Word	2.50
System	23.53	Crime	5.00	Anger	0.00
Myth	23.33	Evidence	5.00	Challenge	0.00
Philosophy	22.73	Word	5.00	Contradiction	0.00
Chance	22.50	Sense	4.55	Correction	0.00
Detail	22.50	Assessment	3.85	Crime	0.00
Practice	22.50	Recommendation	3.13	Endeavour	0.00
Anger	20.00	Detail	2.50	Endorsement	0.00
Recommendation	18.75	Joke	2.50	Evidence	0.00
Application	17.50	Objective	2.50	Facet	0.00
Experience	17.50	Anger	0.00	Finding	0.00
Scandal	15.38	Correction	0.00	Irony	0.00
Work	14.81	Endorsement	0.00	Leave	0.00
Foreboding	14.29	Foreboding	0.00	Motivation	0.00
Proviso	14.29	Leave	0.00	Opposite	0.00
Project	12.50	Motivation	0.00	Phenomenon	0.00
Terror	12.50	Opposite	0.00	Practice	0.00
Triumph	12.50	Point	0.00	Proviso	0.00
Suspicion	12.00	Proviso	0.00	Quest	0.00
Evidence	10.00	Quest	0.00	Recollection	0.00
Time	10.00	Recollection	0.00	Scandal	0.00
Correction	0.00	Terror	0.00	Surprise	0.00
Impetus	0.00	Thing	0.00	Terror	0.00
Leave	0.00	Time	0.00	Testimony	0.00
Sense	0.00	Way	0.00	Triumph	0.00

10.5 *Theme/Rheme*

Marked Theme	%	Theme	%	Rheme	%
Time	22.50	Contradiction	50.00	Correction	100.00
Quest	20.00	Misfortune	42.86	Impetus	100.00
Irony	14.29	Recollection	40.00	Sense	100.00
Proviso	14.29	Motivation	37.50	Surprise	96.00
Philosophy	9.09	Example	35.00	Crime	95.00
Scandal	7.69	Thing	35.00	Evidence	90.00
Area	7.50	Failure	34.78	Opposite	90.00
Warning	6.90	Assessment	34.62	Work	89.29
Phenomenon	5.56	Phenomenon	33.33	Suspicion	88.00
Evidence	5.00	Problem	32.50	Practice	87.50
Example	5.00	Challenge	32.35	Recommendation	87.50
Experience	5.00	Finding	32.14	Word	87.50
Way	5.00	Part	31.25	Capacity	86.96
Surprise	4.00	Project	30.00	Foreboding	85.71
Assessment	3.85	Facet	28.57	Testimony	85.71
Myth	3.33	Vision	27.59	Joke	85.00
Part	3.13	Warning	27.59	Scandal	84.62
Answer	2.50	Prejudice	27.27	Dimension	83.33
Objective	2.50	Anger	25.00	Characteristic	82.76
Practice	2.50	Application	25.00	Venture	82.76
Project	2.50	Endorsement	25.00	Chance	80.00
Word	2.50	Leave	25.00	Experience	80.00
Anger	0.00	Terror	25.00	Quest	80.00

Application	0.00	Myth	23.33	Endeavour	77.78
Capacity	0.00	Philosophy	22.73	Detail	77.50
Challenge	0.00	Answer	22.50	Way	77.50
Chance	0.00	Detail	22.50	Application	75.00
Characteristic	0.00	Objective	22.50	Endorsement	75.00
Contradiction	0.00	Point	22.50	Leave	75.00
Correction	0.00	Endeavour	22.22	Point	75.00
Crime	0.00	System	20.59	Terror	75.00
Detail	0.00	Area	20.00	Time	75.00
Dimension	0.00	Chance	17.50	Triumph	75.00
Endeavour	0.00	Characteristic	17.24	System	73.53
Endorsement	0.00	Dimension	16.67	Prejudice	72.73
Facet	0.00	Experience	15.00	Answer	72.50
Failure	0.00	Way	15.00	Facet	71.43
Finding	0.00	Foreboding	14.29	Irony	71.43
Foreboding	0.00	Irony	14.29	Proviso	71.43
Impetus	0.00	Proviso	14.29	Area	70.00
Joke	0.00	Testimony	14.29	Objective	70.00
Leave	0.00	Venture	13.79	Vision	68.97
Misfortune	0.00	Capacity	13.04	Finding	67.86
Motivation	0.00	Recommendation	12.50	Challenge	67.65
Opposite	0.00	Triumph	12.50	Problem	67.50
Point	0.00	Suspicion	12.00	Part	65.63
Prejudice	0.00	Work	10.71	Failure	65.22
Problem	0.00	Opposite	10.00	Project	65.00
Recollection	0.00	Word	10.00	Thing	65.00
Recommendation	0.00	Scandal	7.69	Motivation	62.50
Sense	0.00	Joke	7.50	Phenomenon	61.11
Suspicion	0.00	Crime	5.00	Example	60.00
System	0.00	Evidence	5.00	Recollection	60.00
Terror	0.00	Practice	5.00	Philosophy	59.09
Testimony	0.00	Time	2.50	Warning	58.62
Thing	0.00	Correction	0.00	Assessment	57.69
Triumph	0.00	Impetus	0.00	Misfortune	57.14
Venture	0.00	Quest	0.00	Myth	56.67
Vision	0.00	Sense	0.00	Anger	50.00
Work	0.00	Surprise	0.00	Contradiction	50.00

APPENDIX 11 SYNTACTIC FUNCTIONS AND THEIR FORMAL REALISATIONS (COMPLETE LIST)

Syntactic function and formal structure (%)			
SB-DF.AR^H	2.35	SB(not)-H^AP.THAT.CL	0.07
DO-DF.AR^H	1.80	SB(not)-H^PP(of)^PP(of)	0.07
SCL-IN.AR^H	1.73	SB(not)-H^TI.CL	0.07
DO-H	1.45	SB(not)-IN.AR^AJ^H	0.07
DO-DF.AR^H^PP(of)	1.24	SB(not)-IN.AR^AJ^H^AP.THAT.CL	0.07
SB-DF.AR^H^PP(of)	1.24	SB(not)-IN.AR^AJ^H^PP(between)	0.07
DO-AJ^H	1.04	SB(not)-IN.AR^AJP^H^NR.PL.ED.CL^AP.THAT.CL	0.07
SCL-IN.AR^AJ^H	1.04	SB(not)-IN.AR^H^AP.THAT.CL	0.07
DO-H^PP(of)	0.90	SB(not)-IN.AR^H^PP(of)	0.07
SB-DF.AR^AJ^H	0.90	SB(not)-IN.AR^H^RT.RV.CL(where)	0.07
SB-H	0.90	SB(not)-N^H	0.07
DO-IN.AR^AJ^H	0.83	SB(not)-NAS.DT^QT^H	0.07
SB-AJ^H	0.83	SB(not)-NAS.DT^H^RT.RV.CL	0.07
SB-PS.DT^H	0.83	SB(not)-NAS.DT^H^AP.THAT.CL	0.07
DO-IN.AR^H	0.69	SB(not)-NAS.DT^H^RT.RV.CL(when)	0.07
SCL-DF.AR^H	0.69	SB(not)-NG.DT^AJ^AJ^H	0.07
SCL-H^PP(of)	0.69	SB(not)-NG.DT^AJ^H	0.07
SB-DM.DT(THESE)^H	0.62	SB(not)-NG.DT^AJ^H^RT.RV.CL	0.07
DO-DF.AR^AJ^H	0.55	SB(not)-NG.DT^H^AP.THAT.CL	0.07
SCL-DF.AR^H^PP(of)	0.55	SB(not)-NG.DT^H^PP(for)	0.07
DO-IN.AR^H^PP(of)	0.48	SB(not)-NG.DT^H^PP(on)	0.07
DO-PS.DT^H	0.48	SB(not)-NG.DT^H^PP(with)	0.07
SB-H^PP(of)	0.48	SB(not)-NUM.CD^AJ^H	0.07
DO-QT^H	0.41	SB(not)-NUM.CD^AJ^H^PT.ST.NR.IT.AP	0.07
DO-PS.DT^AJ^H	0.41	SB(not)-NUM.CD^AJP^H^PP(in terms of)	0.07
SB-DF.AR^N^H	0.41	SB(not)-NUM.CD^H^AJ	0.07
CP.OF(<DO)-H	0.41	SB(not)-NUM.CD^N^H	0.07
DO-DF.AR^AJ^H^PP(of)	0.35	SB(not)-NUM.GO^H^PP(to)^RT.RV.CL	0.07
DO-DF.AR^H^RT.RV.CL	0.35	SB(not)-QT^AJ^H^PP(for)^PT.WK.NR.IT.AP	0.07
DO-DM.DT(THIS)^H	0.35	SB(not)-QT^DS.GV^H	0.07
SCL-AJ^H	0.35	SB(not)-QT^H^AP.THAT.CL	0.07
SCL-DF.AR^AJ^H	0.35	SB(not)-QT^H^PL.ING.CL^PP(as to)	0.07
SCL-DF.AR^H^RT.RV.CL	0.35	SB(not)-QT^H^PP(from)	0.07
SCL-IN.AR^AJ^H^PP(of)	0.35	SB(not)-QT^H^TI.CL	0.07
DO-DF.AR^H^AP.TI.CL	0.28	SB(not)-QT^NUM.GO^H^PP(in)	0.07
DO-DM.DT(THAT)^H	0.28	SB(not)-QV.PV^H	0.07
DO-H^AP.TI.CL	0.28	PO-AJ^N^H	0.07
DO-NAS.DT^H	0.28	PO-AS.PN^PV.PP(of DF.AR^AJ^N^H)	0.07
DO-NUM.GO^H	0.28	PO-DF.AR^H	0.07
SB-DF.AR^H^RT.RV.CL	0.28	PO-DF.AR^H^PP(for)	0.07
SB-DF.DV.GV.NP^H^PP(of)	0.28	PO-DF.AR^H^PP(of)^AP.THAT.CL	0.07
SB-IN.AR^H	0.28	PO-DF.AR^N^H^AP.THAT.CL	0.07
SB-IN.AR^H^PP(of)	0.28	PO-DF.AR^N^H^PT.ST.NR.ATT.AP	0.07
SB-IN.AR^AJ^H	0.28	PO-DF.AR^N^AJ^H^RT.RV.CL	0.07
SCL-IN.AR^H^PP(of)	0.28	PO-DF.AR^N^H	0.07
SCL-NG.DT^H	0.28	PO-DF.AR^NUM.GO^H	0.07
SB(not)-QT^H	0.28	PO-DF.AR^QL.PV^H^RT.RV.CL	0.07
SB(not)-QT^H^PP(of)	0.28	PO-DF.DV.GV.NP^H	0.07
NA-H	0.28	PO-DF.DV.GV.NP^H^PP(of)	0.07
CP.OF(<DO)-AJ^H	0.28	PO-DF.DV.GV.NP^NUM.CD^H	0.07

CP.OF(<SB)-H	0.28	PO-DM.DT(THAT)^H	0.07
DO-DF.AR^AJ^H^RT.RV.CL	0.21	PO-DM.DT(THIS)^H	0.07
DO-DF.DV.GV.NP^H	0.21	PO-H^PP(in)	0.07
DO-DM.DT(THESE)^H	0.21	PO-H^PP(about)	0.07
DO-H^PP(for)	0.21	PO-H^PP(with)	0.07
DO-H^PP(to)	0.21	PO-H^RT.RV.CL(when)	0.07
DO-H^RT.RV.CL	0.21	PO-IN.AR^AJ^H^AP.THAT.CL	0.07
DO-IN.AR^AJ^H^PP(for)	0.21	PO-IN.AR^H	0.07
DO-IN.AR^H^PP(about)	0.21	PO-NUM.CD^H	0.07
DO-IN.AR^H^AP.TI.CL	0.21	PO-PS.DT^H^PP(in)	0.07
DO-IN.AR^N^H	0.21	PO-PS.DT^NUM.OR^AJ^H	0.07
DO-PS.DT^AJ^H^AP.TI.CL	0.21	PO-QT^H	0.07
DO-PS.DT^H^PP(for)	0.21	PO-QT^H^PP(concerning)	0.07
SB-AS.DT^H	0.21	UNC(NA)-N^H^PP(for)	0.07
SB-QT^H	0.21	UNC(NA)-PS.DT^H	0.07
SB-DF.AR^H^PL.ED.CL	0.21	UNC(NA)-UV.PDT^QL.PV^H	0.07
SB-DF.DV.GV.NP^H	0.21	OCL-AJ^H	0.07
SB-DM.DT(THIS)^H	0.21	OCL-IN.AR^H	0.07
SB-IN.AR^AJ^H^PP(of)	0.21	OCL-IN.AR^N^H	0.07
SB-IN.AR^H^AP.TI.CL	0.21	NA-N^N^H	0.07
SB-N^H	0.21	NA-AJ^H^PP(in)	0.07
SB-NUM.CD^H	0.21	NA-DF.AR^N^H	0.07
SB-NUM.GO^H	0.21	NA-DF.AR^NUM.CD^H	0.07
SB-PS.DT^H^PP(of)	0.21	NA-DF.DV.GV.NP^N^H	0.07
SCL-DF.AR^AJ^H^PP(of)	0.21	NA-H^PP(on)	0.07
SCL-NG.DT^AJ^H	0.21	NA-NUM.CD(QT)^H^PP(about)	0.07
SB(not)-H^RT.RV.CL(when)	0.21	AP.NR(<CP.OF<CP.IN<AB.AT.TM)-IN.AR^AJ^H^PP(for)	0.07
PO-DF.AR^H^PP(of)	0.21	AP.NR(<CP.OF<PO)-IN.AR^H^PL.ED.CL	0.07
PO-H^PP(of)	0.21	AP.NR(<DO)-DF.AR^H^PP(of)	0.07
PO-H^TI.CL	0.21	AP.NR(<SCL)-IN.AR^H^RT.RV.CL	0.07
NA-H^PP(of)	0.21	CP.ABOUT(<AB.AT.RS)-H^PP(of)	0.07
NA-N^H	0.21	CP.ABOUT(<AB.AT.RS)-NAS.DT^AJ^H	0.07
CP.AS(<PSCL)-H	0.21	CP.ABOUT(<AB.AT.RS)-NUM.CD^H	0.07
CP.AT(<AB.AT.TM)-DF.AR^H	0.21	CP.ABOUT(<AB.AT.RS)-PS.DT^H	0.07
CP.OF(<DO)-DF.AR^H	0.21	CP.ABOUT(<DO)-DF.AR^AJ^H^PP(in)	0.07
AB.AT.TM-DM.DT(THIS)^H	0.14	CP.ABOUT(<PD.AT.MAT)-DF.AR^H^PP(of)	0.07
DO-AJ^H^PP(about)	0.14	CP.ABOUT(<SCL)-DF.DV.GV.NP^H^AP.TI.CL	0.07
DO-AJ^H^PP(in)	0.14	CP.ACROSS(<AB.AT.SPA)-DF.AR^AJ^H^PP(of)	0.07
DO-AS.DT^AJ^H	0.14	CP.AFTER(<AB.AT.TM)-DF.DV.GV.NP^H	0.07
DO-DF.AR^N^H	0.14	CP.AFTER(<AB.AT.TM)-IN.AR^N^H	0.07
DO-DF.AR^H^AP.THAT.CL	0.14	CP.AFTER(<AB.AT.TM)-PS.DT^H	0.07
DO-DM.DT(THAT)^H^PP(of)	0.14	CP.AGAINST(<AB.AT.RS)-N^H	0.07
DO-H^PP(from)	0.14	CP.AGAINST(<DO)-DF.AR^AJ^H^PP(of)	0.07
DO-IN.AR^AJ^AJ^H	0.14	CP.AS(<AB.AT.MAN)-H	0.07
DO-IN.AR^AJ^H^AP.THAT.CL	0.14	CP.AS(<AB.AT.MAN.RL)-IN.AR^AJ^H	0.07
DO-IN.AR^AJ^H^AP.TI.CL	0.14	CP.AS(<CP.OF<DO)-DF.AR^H^PP(of)	0.07
DO-IN.AR^H^AP.THAT.CL	0.14	CP.AS(<POCL)-AJ^AJ^H	0.07
DO-IN.AR^H^PP(for)	0.14	CP.AS(<POCL)-IN.AR^AJ^H	0.07
DO-IN.AR^H^PP(with)	0.14	CP.AS(<POCL)-IN.AR^H	0.07
DO-IN.AR^H^PT.WK.NR.IT.AP	0.14	CP.AS(<POCL)-PS.DT^AJ^H^PP(in)	0.07
DO-NG.DT^AJ^H	0.14	CP.AS(<PSCL)-IN.AR^AJ^H^PP(to)	0.07
DO-NG.DT^H	0.14	CP.AS(<PSCL)-IN.AR^AJ^H	0.07
DO-NG.DT^H^PP(of)	0.14	CP.AS(<PSCL)-IN.AR^H	0.07

DO-NUM.CD^H^PP(of)	0.14	CP.AS(<PSCL)-IN.AR^H^PP(of)	0.07
DO-PDT^H	0.14	CP.AS(<PSCL)-NAS.DT^H	0.07
DO-PS.DT^H^PP(of)	0.14	CP.AT(<AB.AT.TM)-DF.AR^AJ^H	0.07
DO-QT^AJ^H	0.14	CP.AT(<AB.AT.TM)-DF.AR^AJ^H^FL.WK.NR.IT.AP	0.07
DO-QT^H^PP(of)	0.14	CP.AT(<AB.AT.TM)-DF.AR^H^PP(of)	0.07
DO-UV.DT^H	0.14	CP.AT(<AB.AT.TM)-H^PP(of)	0.07
DO-UV.PDT^DF.AR^H	0.14	CP.AT(<AB.AT.TM)-IN.AR^H^PP(of)	0.07
SB-AS.DT^AJ^H	0.14	CP.AT(<AB.AT.TM)-N^H	0.07
SB-DF.AR^AJ^H^PP(about)	0.14	CP.AT(<AB.AT.TM)-NAS.DT^H	0.07
SB-DF.AR^AJ^H^PP(of)	0.14	CP.AT(<DO)-DF.AR^H^AP.TI.CL	0.07
SB-DF.AR^H^PP(to)	0.14	CP.BECAUSE OF(<AB.AT.CAU)-DF.AR^AJ^H^AP.TI.CL	0.07
SB-DF.AR^H^PP(for)	0.14	CP.BECAUSE OF(<AB.AT.CAU)-DF.AR^H^RT.RV.CL	0.07
SB-DF.AR^NUM.CD^AJ^H^PP(of)	0.14	CP.BECAUSE OF(<AB.AT.CAU)-PS.DT^QT^H	0.07
SB-DF.AR^NUM.OR^H^RT.RV.CL	0.14	CP.BECAUSE OF(<AB.AT.RE)-DM.DT(THIS)^AJ^AJ^H	0.07
SB-DF.DV.GV.NP^AJ^H	0.14	CP.BEHIND(<AB.AT.SPA)-QT^AJ^H^PP(on n)	0.07
SB-DM.DT(THIS)^AJ^H	0.14	CP.BETWEEN(<AB.AT.RE)-AJ^H	0.07
SB-H^AP.TI.CL	0.14	CP.BETWEEN(<DO)-DF.AR^AJ^H^PP(of)	0.07
SB-H^PP(for)	0.14	CP.BETWEEN(<SB)-AJ^H	0.07
SB-NUM.CD^H^PP(of)	0.14	CP.BY(<AB.AT.AG)-AJ^H^PP(about)	0.07
SB-NUM.CD^PV.PP(of DF.AR^AJ^H^RT.RV.CL)	0.14	CP.BY(<AB.AT.AG)-DF.AR^AJ^AJ^H^PP(of)	0.07
SB-NUM.CD^PV.PP(of DF.AR^H^RT.RV.CL)	0.14	CP.BY(<AB.AT.AG)-DF.AR^AJ^H^PP(over)	0.07
SB-NUM.GO^H^PP(of)	0.14	CP.BY(<AB.AT.AG)-DF.AR^H	0.07
SB-PS.DT^AJ^H	0.14	CP.BY(<AB.AT.AG)-DF.AR^H^PP(of)	0.07
SB-PS.DT^H^PP(for)	0.14	CP.BY(<AB.AT.AG)-DF.AR^H^TI.CL	0.07
SB-RV.DT^H	0.14	CP.BY(<AB.AT.AG)-DF.DV.GV.NP^AJ^H	0.07
SB-RV.DT^H^PP(of)	0.14	CP.BY(<AB.AT.AG)-DM.DT(THAT)^H	0.07
SCL-DF.AR^AJ^H^PP(for)	0.14	CP.BY(<AB.AT.AG)-IN.AR^AJ^AJ^H^AJ	0.07
SCL-DF.AR^QL.PV^H^RT.RV.CL	0.14	CP.BY(<AB.AT.AG)-IN.AR^AJ^H^PP(of)	0.07
SCL-H^PP(to)	0.14	CP.BY(<AB.AT.AG)-PS.DT^H^PP(during)	0.07
SCL-IN.AR^AJ^H^PL.ED.CL	0.14	CP.BY(<AB.AT.MS)-DF.AR^AJ^N^H^NR.RV.CL	0.07
SCL-IN.AR^AJ^H^PP(to)	0.14	CP.BY(<AB.AT.MS)-PS.DT^H^PP(on)	0.07
SCL-IN.AR^H^RT.RV.CL(where)	0.14	CP.BY(<AB.AT.TM)-DF.AR^H^RT.RV.CL	0.07
SCL-IN.AR^N^AJ^H	0.14	CP.BY(<CP.WITHOUT<AB.AT.MAN)-PS.DT^H	0.07
SCL-IN.AR^N^H	0.14	CP.BY(<SB)-DF.DV.GV.NP^H^PP(of)	0.07
SCL-NUM.CD^AJ^H	0.14	CP.BY(<SB)-H	0.07
SCL-NUM.CD^PV.PP(of DF.AR^H)	0.14	CP.BY(<SB.AT)-DF.DV.GV.NP^H^AP.TI.CL	0.07
SCL-PS.DT^H^PP(of)	0.14	CP.BY(<SCL)-DF.AR^AJ^AJ^H	0.07
SCL-QT^H	0.14	CP.CONCERNING(<DO)-NAS.PN^PV.PP(of DF.AR^H^PL.ED.CL)	0.07
SB(not)-AJ^H^PP(of)	0.14	CP.CONTRARY TO(<AB.AT.MAN)-DF.AR^H^PP(of)	0.07
SB(not)-QT^H^RT.RV.CL	0.14	CP.DESPITE(<AB.AT.CC)-DF.AR^H^PP(of)	0.07
SB(not)-NG.DT^H^PP(of)	0.14	CP.DESPITE(<AB.AT.CC)-	0.07

		DF.DV.GV.NP^H^PP(of)	
SB(not)-NUM.CD^H^PP(of)	0.14	CP.DESPITE(<AB.AT.CC)-H	0.07
PO-DF.AR^AJ^H^PP(of)	0.14	CP.DESPITE(<AB.AT.CC)-PS.DT^H^PP(in)	0.07
		CP.DUE TO(<AB.AT.CAU)-	
PO-PS.DT^H	0.14	DF.AR^H^PP(of)^AP.TI.CL	0.07
NA-AJ^H	0.14	CP.DURING(<AB.AT.TM)-DF.AR^H^PP(to)	0.07
NA-DF.AR^H	0.14	CP.DURING(<AB.AT.TM)-IN.AR^H^PP(of)	0.07
NA-DF.AR^H^PP(of)	0.14	CP.DURING(<AB.AT.TM)-PS.DT^H	0.07
		CP.FOLLOWING(<AB.AT.CAU)-	
NA-H^TI.CL	0.14	DF.AR^H^PP(of)^AP.TI.CL	0.07
CP.AS(<POCL)-		CP.FOLLOWING(<AB.AT.TM)-	
IN.AR^H^PP(of)	0.14	DF.AR^AJ^H^PP(of)	0.07
CP.AT(<AB.AT.TM)-			
DM.DT(THAT)^H	0.14	CP.FOR(<AB.AT.CAU)-AJ^H^PP(against)	0.07
CP.FOR(<AB.AT.PU)-H	0.14	CP.FOR(<AB.AT.CAU)-DF.AR^H^PP(of)	0.07
CP.IN(<AB.AT.MS)-			
NUM.CD^PV.PP(of			
NUM.CD^AJ^H)	0.14	CP.FOR(<AB.AT.CAU)-H^RT.RV.CL	0.07
CP.IN(<AB.AT.RS)-AS.DT^H	0.14	CP.FOR(<AB.AT.CAU)-IN.AR^H	0.07
CP.IN(<AB.AT.RS)-NUM.CD^H	0.14	CP.FOR(<AB.AT.PU)-AS.DT^H	0.07
CP.IN(<AB.AT.SPA)-AS.DT^H	0.14	CP.FOR(<AB.AT.PU)-QT^H^PP(on)	0.07
CP.IN(<AB.AT.SPA)-DF.AR^H	0.14	CP.FOR(<AB.AT.PU)-DF.AR^AJ^H^PP(of)	0.07
CP.IN(<AB.AT.SPA)-			
DF.AR^H^PP(of)	0.14	CP.FOR(<AB.AT.PU)-IN.AR^H	0.07
CP.OF(<AB.AT.RS)-AJ^H	0.14	CP.FOR(<AB.AT.PU)-NAS.DT^AJ^H	0.07
CP.OF(<DO)-			
DF.AR^AJ^H^PP(of)	0.14	CP.FOR(<AB.AT.PU)-N^H	0.07
CP.OF(<DO)-H^PP(for)	0.14	CP.FOR(<AB.AT.PU)-NUM.GO^H	0.07
CP.OF(<SB)-DF.AR^H^PP(of)	0.14	CP.FOR(<AB.AT.RE)-AJ^H	0.07
CP.ON(<AB.AT.RS)-			
DF.AR^H^RT.RV.CL	0.14	CP.FOR(<AB.AT.RE)-DF.AR^H	0.07
CP.TO(<AB.AT.RS)-AJ^H	0.14	CP.FOR(<DO)-AJ^H^PP(on)	0.07
CP.WITH(<AB.AT.MAN)-AJ^H	0.14	CP.FOR(<DO)-AJ^H^PT.WK.NR.EM.AP	0.07
CP.WITH(<AB.AT.MS)-AJ^H	0.14	CP.FOR(<DO)-DM.DT(THESE)^H	0.07
CP.WITH(<AB.AT.RS)-			
DF.AR^H^PP(of)	0.14	CP.FOR(<DO)-DM.DT(THIS)^AJ^H	0.07
CP.WITH(<SCL)-DF.AR^AJ^H	0.14	CP.FOR(<DO)-H	0.07
AB.AT.MAN-			
DF.AR^H^RT.RV.CL	0.07	CP.FOR(<DO)-PDT^IN.AR^H	0.07
AB.AT.MAN-DM.DT(THAT)^H	0.07	CP.FOR(<DO)-PS.DT^H	0.07
AB.AT.PU-IV.DT^H	0.07	CP.FOR(<NA)-DF.DV.GV.NP^AJ^H	0.07
AB.AT.RS-DF.AR^H^PP(of)	0.07	CP.FOR(<NA)-N^H^AJ	0.07
AB.AT.TM-			
DF.AR^NUM.OR^H^RT.RV.CL	0.07	CP.FOR(<PO)-DF.AR^H	0.07
AB.AT.TM-NUM.CD^H	0.07	CP.FOR(<SB(not))-DM.DT(THESE)^H	0.07
AB.AT.TM-QT^H	0.07	CP.FOR(<SB)-DF.AR^AJ^H	0.07
AB.AT.TM-QT^H^RT.RV.CL	0.07	CP.FOR(<SB)-DF.AR^H	0.07
AB.AT.TM-UV.DT^H	0.07	CP.FOR(<SB)-DM.DT(THIS)^H	0.07
AB.AT.TM-UV.DT^H^RT.RV.CL	0.07	CP.FOR(<SB)-H^PP(of)	0.07
AB.DJ.CO-IN.AR^H^RT.RV.CL	0.07	CP.FOR(<SB)-UV.DT^H	0.07
AB.DJ.SY.ML.CM-			
NUM.CD^NUM.GO^H^			
PL.ED.CL	0.07	CP.FOR(<SCL)-PS.DT^AJ^H^PP(of)	0.07
AB.DJ.SY.ML.CM-			
NUM.CD^NUM.GO^QV.PV^H	0.07	CP.FROM(<AB.AT.CAU)-AJ^H^PP(in)	0.07
DO-AJ^AJ^H	0.07	CP.FROM(<AB.AT.MAN)-DF.DV.GV.NP^H	0.07

DO-AJ^H^PL.ED.CL	0.07	CP.FROM(<AB.AT.MAN)-H^AJ	0.07
DO-AJ^H^PP(for)	0.07	CP.FROM(<AB.AT.MS)-DF.AR^AJ^H	0.07
DO-AJ^H^PP(in)^PT.WK.NR.PA.AP	0.07	CP.FROM(<AB.AT.RS)-AJ^H	0.07
DO-AJ^H^PP(of)	0.07	CP.FROM(<AB.AT.SPA)-AJ^H^FL.ST.RT.EM.AP	0.07
DO-AJ^H^PP(to)	0.07	CP.FROM(<AB.AT.SPA)-AS.DT^H^PP(of)	0.07
DO-AJ^H^PP(with)	0.07	CP.FROM(<AB.AT.SPA)-DF.AR^AJ^H^PP(of)	0.07
DO-AJ^H^RT.RV.CL	0.07	CP.FROM(<AB.AT.SPA)-DF.DV.GV.NP^H	0.07
DO-AJ^N^AJ^H	0.07	CP.FROM(<AB.AT.SPA)-H	0.07
DO-AJ^N^H^RT.RV.CL	0.07	CP.FROM(<AB.AT.SPA)-IN.AR^H^PL.ED.CL	0.07
DO-AS.DT^AJ^AJ^H^PP(in)	0.07	CP.FROM(<AB.AT.SPA)-PS.DT^H	0.07
DO-AS.DT^QT^H^TI.CL	0.07	CP.FROM(<CP.OF<CP.WITH<SB)-DF.AR^N^H	0.07
DO-AS.DT^H	0.07	CP.FROM(<PT.WK.NR.AP<AB.AT.AG)-AJ^H^PP(in)	0.07
DO-AS.DT^H^PP(for)	0.07	CP.FROM(<SB)-AJ^AJ^H	0.07
DO-AS.DT^H^PP(in)	0.07	CP.FROM(<SJ.VI)-IV.DT^H	0.07
DO-AS.DT^H^PP(to)	0.07	CP.IN RESPECT OF(<AB.AT.RS)-AJ^H	0.07
DO-AS.DT^N^H	0.07	CP.IN(<AB.AT.MAN)-AJ^H	0.07
DO-AS.PN^PV.PP(of DF.AR^H^PP(in))	0.07	CP.IN(<AB.AT.MAN)-AS.DT^QL.PV^H	0.07
DO-AS.PN^PV.PP(of DF.DV.GV.NP^H^PP of)	0.07	CP.IN(<AB.AT.MAN)-DM.DT(THIS)^H	0.07
DO-AS.PN^PV.PP(of DM.DT(THESE)^H)	0.07	CP.IN(<AB.AT.MAN)-H	0.07
DO-QT.PN^PV.PP(of DF.AR^H^PP of)	0.07	CP.IN(<AB.AT.MAN)-H^RT.RV.CL	0.07
DO-DF.AR^AJ^AJ^H^PL.ING.CL	0.07	CP.IN(<AB.AT.MAN)-IV.DT^H	0.07
DO-DF.AR^AJ^AJ^H^PP(of)	0.07	CP.IN(<AB.AT.MAN.RL)-DF.AR^AJ^H^PP(of)	0.07
DO-DF.AR^AJ^H^AJ	0.07	CP.IN(<AB.AT.MAN.RL)-DM.DT(THAT)^H	0.07
DO-DF.AR^AJ^H^AP.THAT.CL	0.07	CP.IN(<AB.AT.MAN.RL)-IN.AR^AJ^H	0.07
DO-DF.AR^AJ^H^PP(from)	0.07	CP.IN(<AB.AT.MS)-H^PP(by)	0.07
DO-DF.AR^AJ^N^H	0.07	CP.IN(<AB.AT.MS)-NUM.CD^AJ^H	0.07
DO-DF.AR^AJ^PP(of DF.DV.GV.NP^H)	0.07	CP.IN(<AB.AT.MS)-NUM.CD^H	0.07
DO-DF.AR^H^FL.ST.RT.AP	0.07	CP.IN(<AB.AT.MS)-PS.DT^H	0.07
DO-DF.AR^H^PP(for)	0.07	CP.IN(<AB.AT.MS)-QT^H^FL.ST.NR.IT.AP	0.07
DO-DF.AR^H^PP(from)	0.07	CP.IN(<AB.AT.PU)-DF.AR^AJ^H	0.07
DO-DF.AR^H^PP(in)^AP.THAT.CL	0.07	CP.IN(<AB.AT.PU)-IN.AR^AJ^H^AP.TI.CL	0.07
DO-DF.AR^H^PP(of)^AP.THAT.CL	0.07	CP.IN(<AB.AT.PU)-IN.AR^AJ^H^PP(to)	0.07
DO-DF.AR^H^PP(of)^AP.TI.CL	0.07	CP.IN(<AB.AT.PU)-IN.AR^H^AP.TI.CL	0.07
DO-DF.AR^H^PP(of)^PP(about)	0.07	CP.IN(<AB.AT.RS)-DF.AR^AJ^H	0.07
DO-DF.AR^H^PP(on)	0.07	CP.IN(<AB.AT.RS)-DF.AR^N^H	0.07
DO-DF.AR^H^PP(to)	0.07	CP.IN(<AB.AT.RS)-DM.DT(THIS)^AJ^H	0.07
DO-DF.AR^NUM.CD^H	0.07	CP.IN(<AB.AT.RS)-DM.DT(THIS)^H	0.07
DO-DF.AR^QL.PV^AJ^H^RT.RV.CL	0.07	CP.IN(<AB.AT.RS)-IN.AR^AJ^H	0.07
DO-	0.07	CP.IN(<AB.AT.RS)-NG.DT^AJ^H	0.07

DF.AR^QL.PV^H^RT.RV.CL			
DO-DF.DV.GV.NP^AJ^H	0.07	CP.IN(<AB.AT.RS)-NUM.GO^H	0.07
DO-DF.DV.GV.NP^H^PP(for)	0.07	CP.IN(<AB.AT.RS)-PS.DT^H^AP.TI.CL	0.07
DO-DF.DV.GV.NP^H^PP(of)	0.07	CP.IN(<AB.AT.RS)-PS.DT^N^H	0.07
DO-DF.DV.GV.NP^H^PP(to)	0.07	CP.IN(<AB.AT.RS)-QT^H	0.07
DO-DF.DV.GV.NP^N^H	0.07	CP.IN(<AB.AT.RS)-UV.PDT^DM.DT(THAT)^H	0.07
DO-DM.DT(THAT)^H^AP.TI.CL	0.07	CP.IN(<AB.AT.SPA)-AJ^H	0.07
DO-DM.DT(THAT)^QL.PV^H	0.07	CP.IN(<AB.AT.SPA)-AJ^H^AP.TI.CL	0.07
DO-DM.DT(THIS)^N^H	0.07	CP.IN(<AB.AT.SPA)-DF.AR^AJ^H	0.07
DO-DM.DT(THIS)^AJ^AJ^H	0.07	CP.IN(<AB.AT.SPA)-DF.AR^AJ^H^PP(of)	0.07
DO-DM.DT(THOSE)^AJ^H	0.07	CP.IN(<AB.AT.SPA)-DM.DT(THESE)^QL.PV^AJ^H	0.07
DO-DM.DT(THOSE)^H	0.07	CP.IN(<AB.AT.SPA)-DM.DT(THIS)^H	0.07
DO-H^PP(about)	0.07	CP.IN(<AB.AT.SPA)-IN.AR^N^H	0.07
DO-H^PP(among)^AP.THAT.CL	0.07	CP.IN(<AB.AT.SPA)-IN.AR^AJ^H^PP(over)	0.07
DO-H^PP(other than)	0.07	CP.IN(<AB.AT.SPA)-IN.AR^H^PP(of)^NR.RV.CL	0.07
DO-IF.AV^IN.AR^AJ^H	0.07	CP.IN(<AB.AT.SPA)-NAS.DT^NUM.GO^H^PP(of)	0.07
DO-IF.AV^QT^H^PP(of)	0.07	CP.IN(<AB.AT.SPA)-NUM.CD^PV.PP(of PS.DT^H)	0.07
DO-IN.AR^N^H^AP.TI.CL	0.07	CP.IN(<AB.AT.SPA)-PS.DT^H^PP(for)	0.07
DO-IN.AR^AJ^H^PP(of)	0.07	CP.IN(<AB.AT.SPA)-UV.DT^H^PP(of)	0.07
DO-IN.AR^AJ^H^PP(of)^PP(in)	0.07	CP.IN(<AB.AT.SPA)-UV.PDT^DM.DT(THESE)^AJ^H^NR.RV.CL	0.07
DO-IN.AR^AJ^H^PP(to)	0.07	CP.IN(<AB.AT.TM)-DF.AR^NUM.GO^N^H^FL.ST.NR.IT.AP	0.07
DO-IN.AR^AJ^H^PP(with)	0.07	CP.IN(<AB.AT.TM)-H^PP(of)	0.07
DO-IN.AR^AJ^H^PT.WK.NR.IT.AP	0.07	CP.IN(<AB.AT.TM)-PS.DT^H^PP(for)	0.07
DO-IN.AR^AJ^N^H	0.07	CP.IN(<AB.DJ.CM.CL)-DF.AR^H^AP.THAT.CL	0.07
DO-IN.AR^AJ^N^H^PP(on)	0.07	CP.IN(<CP.AS<POCL)-AJ^H^PP(of)	0.07
DO-IN.AR^AJ^H^PP(in)	0.07	CP.IN(<CP.FOR<AB.AT.PU)-DM.DT(THIS)^N^H	0.07
DO-IN.AR^H^PL.ING.CL	0.07	CP.IN(<CP.OF<CP.FROM<AB.AT.MS)-DF.AR^H^PP(of)	0.07
DO-IN.AR^H^PP(of)^RT.RV.CL	0.07	CP.IN(<DO)-N^H	0.07
DO-IN.AR^H^PT.ST.NR.IT.AP	0.07	CP.IN(<DO)-DF.AR^AJ^AJ^H	0.07
DO-IN.AR^H^TI.CL	0.07	CP.IN(<DO)-DM.DT(THIS)^H	0.07
DO-IN.AR^N^H^PP(of)	0.07	CP.IN(<DO)-DM.DT(THIS)^H^PP(of)^PP(over)	0.07
DO-IN.AR^N^H^RT.RV.CL(whereby)	0.07	CP.IN(<DO)-H	0.07
DO-IN.AR^NUM.GO^H	0.07	CP.IN(<NA)-N^H	0.07
DO-IV.DT^QL.PV^H^PP(about)	0.07	CP.IN(<NA)-UV.PDT^H^PP(of)	0.07
DO-MR.GV^H^PP(with)	0.07	CP.IN(<SB(not))-DM.DT(THIS)^H	0.07
DO-N^H	0.07	CP.IN(<SB)-DF.AR^AJ^H	0.07
DO-N^H^FL.ST.NR.RF.AP	0.07	CP.IN(<SB)-DF.AR^N^H	0.07
DO-N^H^PT.ST.NR.EM.AP^RT.RV.CL	0.07	CP.IN(<SB)-DM.DT(THESE)^H	0.07
DO-NAS.DT^AJ^H	0.07	CP.IN(<SB)-DM.DT(THIS)^H	0.07
DO-NAS.DT^AJ^H^PP(of)	0.07	CP.IN(<SB)-PS.DT^AJ^AJ^AJ^H	0.07
DO-NAS.DT^H^PP(as to)	0.07	CP.IN(<SCL)-AS.DT^QL.PV^H	0.07
DO-NAS.DT^H^PP(between)	0.07	CP.IN(<SCL)-DF.AR^AJ^H^PP(of)	0.07

DO-NAS.DT^H^PP(in)	0.07	CP.IN(<SCL)-DM.DT(THAT)^AJ^H	0.07
DO-NAS.DT^H^PP(of)^PP(at)	0.07	CP.INCLUDING(<PT.ST.NR.EM.AP<SB)-AJ^H^PP(in)	0.07
DO-NAS.PN^PV.PP(of DM.DT(THESE)^H)	0.07	CP.INCLUDING(<PT.ST.NR.EM.AP<SB)-IN.AR^AJP^H^PP(by)^PP(of)	0.07
DO-NG.DT^H^AP.TI.CL	0.07	CP.INTO(<AB.AT.RU)-NUM.CD^H^FL.ST.NR.IT.AP	0.07
DO-NG.DT^H^PP(for)	0.07	CP.INTO(<AB.AT.SPA)-NUM.CD^AJ^H	0.07
DO-NG.DT^N^H	0.07	CP.INTO(<CP.FOR<AB.AT.CAU)-DF.AR^AJ^H^PP(of)	0.07
DO-NG.DT^NUM.GO^H	0.07	CP.INTO(<POCL)-H	0.07
DO-NG.DT^NUM.GO^H^TI.CL	0.07	CP.INTO(<SCL)-NUM.CD^H	0.07
DO-NUM.CD(QT)^H	0.07	CP.LIKE(<PSCL)-IN.AR^AJ^H	0.07
DO-NUM.CD(QT)^NUM.GO^H^PP(on)	0.07	CP.LIKE(<PSCL)-QT^IN.AR^H	0.07
DO-NUM.CD^H	0.07	CP.OF (<CP.UPON<AB.AT.TM)-IN.AR^N^H	0.07
DO-NUM.CD^H^PP(of)^RT.RV.CL	0.07	CP.OF(<AB.AT.CAU)-DF.DV.GV.NP^H	0.07
DO-NUM.CD^H^PT.WK.NR.IT.AP	0.07	CP.OF(<AB.AT.RS)-DF.AR^H	0.07
DO-NUM.CD^PV.PP(of DM.DT(THOSE)^AJ^H^RT.RV.CL)	0.07	CP.OF(<AB.AT.RS)-PS.DT^H^PL.ED.CL	0.07
DO-NUM.CD^PV.PP(of NUM.CD^H^AV^PT.WK.NR.IT.AP)	0.07	CP.OF(<AB.AT.TM)-DF.DV.GV.NP^H	0.07
DO-NUM.CD^PV.PP(of PS.DT^N^H)	0.07	CP.OF(<CF.FOR<DO)-AJ^N^H^PT.ST.NR.EM.AP^NR.RV.CL	0.07
DO-NUM.GO^H^PP(of)	0.07	CP.OF(<CP.ABOUT<DO)-DF.AR^H	0.07
DO-NUM.GO^H^PP(on)	0.07	CP.OF(<CP.AS<CP.FOLLOWING<AB.AT.TM)-DF.AR^H^AV	0.07
DO-PS.DT^AJ^H^AP.THAT.CL	0.07	CP.OF(<CP.AS<PSCL)-IN.AR^AJ^AJ^H^AP.TI.CL	0.07
DO-PS.DT^H^AP.THAT.CL	0.07	CP.OF(<CP.AT<DO)-DM.DT(THIS)^H	0.07
DO-PS.DT^H^AP.TI.CL	0.07	CP.OF(<CP.AT<SCL)-PS.DT^H	0.07
DO-PS.DT^H^PP(in)	0.07	CP.OF(<CP.BY<AB.AT.AG)-H	0.07
DO-PS.DT^N^H	0.07	CP.OF(<CP.FOR<DO)-H^PP(for)	0.07
DO-PS.DT^NUM.OR^H^PP(into)	0.07	CP.OF(<CP.FOR<SB)-DM.DT(THOSE)^H^PP(of)^PL.ING.CL	0.07
DO-PS.DT^NUM.OR^H^PP(over)	0.07	CP.OF(<CP.FROM<DO)-AJ^AJ^H	0.07
DO-PS.DT^QL.PV^AJ^H	0.07	CP.OF(<CP.FROM<SCL)-PS.DT^H^NR.RV.CL	0.07
DO-QT.IV^H	0.07	CP.OF(<CP.IN SPITE OF<AB.AT.CC)-DF.AR^H	0.07
DO-QT.PN^PV.PP(of DF.AR^H^PP of)	0.07	CP.OF(<CP.IN<AB.AT.RS)-AJ^H	0.07
DO-QT.PN^PV.PP(of DF.AR^H^RT.RV.CL)	0.07	CP.OF(<CP.IN<AB.AT.SPA)-N^H	0.07
DO-QT^H^FL.ST.NR.EM.AP	0.07	CP.OF(<CP.IN<SB)-N^H	0.07
DO-QT^H^PL.ING.CL	0.07	CP.OF(<CP.OF<DO)-H	0.07
DO-QT^H^PP(on)	0.07	CP.OF(<CP.TO<AB.AT.RS)-DF.AR^H^PP(of)	0.07
DO-QT^H^PP(on)^PT.ST.NR.EM.AP	0.07	CP.OF(<CP.TO<CP.OF<DO)-DM.DT(THESE)^H	0.07
DO-QT^H^TI.CL	0.07	CP.OF(<CP.TO<SB)-H	0.07

DO-QT^IN.AR^H	0.07	CP.OF(<CP.WITH<AB.AT.MAN)-H	0.07
DO-QV.PV^UV.PDT^PS.DT^H	0.07	CP.OF(<CP.WITH<AB.AT.RS)-IN.AR^N^H	0.07
DO-UV.DT^QL.PV^H	0.07	CP.OF(<CP.WITH<DO)-H	0.07
DO-UV.PDT^DF.AR^AJ^H^PP(about)	0.07	CP.OF(<CP.WITH<SCL)-H^PP(at)^PP(in)	0.07
DO-UV.PDT^DF.AR^H^PP(on)	0.07	CP.OF(<CP.WITH<SCL)-N^H^RT.RV.CL	0.07
DO-UV.PDT^H	0.07	CP.OF(<DO)-AJ^AJ^AJ^H	0.07
DO-UV.PDT^H^PP(for)	0.07	CP.OF(<DO)-DF.AR^H^PP(of)	0.07
DO-UV.PDT^QL.PV^NUM.GO^H	0.07	CP.OF(<DO)-DF.AR^H^FL.ST.RT.AP	0.07
DO-UV.PN^PV.PP(of DF.AR^AJ^NUM.CD^H)	0.07	CP.OF(<DO)-DM.DT(THAT)^H^PP(of)^FL.ST.NR.IT.AP	0.07
DO-UV.PN^PV.PP(of DM.DT(THESE)^H)	0.07	CP.OF(<DO)-H^PP(for)^PL.ED.CL	0.07
SB-AJ^AJ^H	0.07	CP.OF(<DO)-PDT^H	0.07
SB-AJ^H^AP.THAT.CL	0.07	CP.OF(<DO)-QT^AJ^H	0.07
SB-AJ^H^PP(from)^AP.THAT.CL	0.07	CP.OF(<NA)-N^N^H^PP(as to)	0.07
SB-AJ^H^PT.ST.NR.EM.AP	0.07	CP.OF(<NR.AP<DO)-DF.AR^H^PP(of)	0.07
SB-DF.AR^AJ^AJ^H	0.07	CP.OF(<PO)-DF.AR^AJ^H	0.07
SB-DF.AR^AJ^H^(for n TI.CL)	0.07	CP.OF(<PO)-IN.AR^AJ^N^H	0.07
SB-DF.AR^AJ^H^PP(between)	0.07	CP.OF(<PO)-N^N^H	0.07
SB-DF.AR^AJ^H^PP(for)	0.07	CP.OF(<PT.ST.NR.AP<DO)-PDT^H	0.07
SB-DF.AR^AJ^H^PP(of)^PP(for)	0.07	CP.OF(<SB(not))-AJ^AJ^H	0.07
SB-DF.AR^AJ^H^PP(of)^RT.RV.CL	0.07	CP.OF(<SB(not))-DF.DV.GV.NP^H^AP.TI.CL	0.07
SB-DF.AR^AJ^H^RT.RV.CL	0.07	CP.OF(<SB(not))-H	0.07
SB-DF.AR^AJ^H^TI.CL	0.07	CP.OF(<SB(not))-PS.DT^H^PP(at)	0.07
SB-DF.AR^AJ^N^H^RT.RV.CL	0.07	CP.OF(<SB)-N^H	0.07
SB-DF.AR^AJ^NUM.GO^H^PP(of)^PT.ST.NR.IT.AP	0.07	CP.OF(<SB)-DF.AR^H	0.07
SB-DF.AR^H^(for n AP.TI.CL)	0.07	CP.OF(<SB)-DF.AR^H^PL.ING.CL	0.07
SB-DF.AR^H^AV	0.07	CP.OF(<SB)-DF.DV.GV.NP^H^AP.TI.CL	0.07
SB-DF.AR^H^PP(of)^AP.TI.CL	0.07	CP.OF(<SB)-DM.DT(THIS)^H	0.07
SB-DF.AR^H^PP(on)	0.07	CP.OF(<SB)-H^PL.ED.CL	0.07
SB-DF.AR^H^RT.RV.CL^PP(of)	0.07	CP.OF(<SB)-H^PP(of)^PT.WK.NR.IT.AP	0.07
SB-DF.AR^N^H^PP(of)	0.07	CP.OF(<SB)-IN.AR^N^H	0.07
SB-DF.AR^NUM.CD^AJ^H^RT.RV.CL	0.07	CP.OF(<SB)-PDT^H	0.07
SB-DF.AR^NUM.CD^N^N^H^RT.RV.CL	0.07	CP.OF(<SB)-PS.DT^AJ^H	0.07
SB-DF.AR^NUM.CD^NUM.GO^H^RT.RV.CL	0.07	CP.OF(<SB)-PS.DT^H	0.07
SB-DF.AR^NUM.GO^AJ^H^TI.CL	0.07	CP.OF(<SB.AT)-IN.AR^AJ^H^NR.RV.CL	0.07
SB-DF.AR^NUM.GO^H	0.07	CP.OF(<SCL)-AJ^H	0.07
SB-DF.AR^NUM.GO^H^PP(in)	0.07	CP.OF(<SCL)-DF.AR^H^PP(with)	0.07
SB-DF.AR^NUM.OR^AJ^AJ^H^PP(of)	0.07	CP.OF(<SCL)-DF.DV.GV.NP^AJ^AJ^H	0.07
SB-DF.AR^NUM.OR^H	0.07	CP.OF(<SCL)-DM.DT(THAT)^H	0.07
SB-DF.AR^NUM.OR^H^TI.CL^PP(with)	0.07	CP.OF(<SCL)-DM.DT(THIS)^H	0.07
SB-DF.AR^NUM.OR^PV.PP(of	0.07	CP.OF(<SCL)-H	0.07

DM.DT(THESE)^H)			
SB-DF.AR^QL.PV^AJ^H^PP(of)	0.07	CP.OF(<SCL)-H^PP(against)	0.07
SB-DF.AR^QL.PV^H^			
RT.RV.CL(where)	0.07	CP.OF(<SCL)-H^RT.RV.CL	0.07
SB-DF.AR^UNC^H^RT.RV.CL	0.07	CP.OF(<SCL)-IN.AR^AJ^H	0.07
SB-DF.DV.GV.NP^H^AP.TI.CL	0.07	CP.OF(<SCL)-UV.PDT^PS.DT^AJ^H	0.07
SB-DF.DV.GV.NP^H^PP(for)	0.07	CP.OF(UNC(NA))-PS.DT^AJ^H	0.07
		CP.ON(<AB.AT.RS)-	
SB-DF.DV.GV.NP^NUM.CD^H	0.07	DF.DV.GV.NP^H^PP(of)	0.07
SB-DF.DV.GV.NP^NUM.GO^H	0.07	CP.ON(<AB.AT.RS)-N^H	0.07
SB-DM.DT(THAT)^H^PP(of)	0.07	CP.ON(<AB.AT.RS)-DF.AR^AJ^AJ^H^PP(of)	0.07
SB-DM.DT(THIS)^AJ^H^			
PP(against)	0.07	CP.ON(<AB.AT.RS)-DF.AR^AJ^H^PP(of)	0.07
SB-DM.DT(THIS)^H^PP(of)	0.07	CP.ON(<AB.AT.RS)-DF.AR^H^PL.ED.CL	0.07
SB-DM.DT(THIS)^NUM.GO^H	0.07	CP.ON(<AB.AT.RS)-DF.AR^H^PP(of)	0.07
SB-DM.DT(THIS)^QL.PV^H	0.07	CP.ON(<AB.AT.RS)-DM.DT(THESE)^AJ^H	0.07
SB-H^FL.ST.RT.EM.AP	0.07	CP.ON(<AB.AT.RS)-DM.DT(THIS)^H	0.07
SB-H^PL.ED.CL	0.07	CP.ON(<AB.AT.SPA)-AJ^IN.AR^H	0.07
SB-H^PP(of)^AP.TI.CL	0.07	CP.ON(<AB.AT.TM)-QT^H	0.07
SB-H^PP(with)	0.07	CP.ON(<CAJ<SB)-AJ^AJ^H	0.07
SB-IN.AR^N^H^PP(at)	0.07	CP.ON(<DO)-AJ^H^PP(of)	0.07
SB-IN.AR^AJ^H^PP(for)	0.07	CP.ON(<DO)-DF.AR^H^PP(of)	0.07
SB-IN.AR^AJ^N^H	0.07	CP.ON(<DO)-DM.DT(THIS)^H	0.07
SB-IN.AR^H^PP(like)	0.07	CP.ON(<DO)-H^PP(in)	0.07
SB-IN.AR^N^H	0.07	CP.ON(<PD.AT.MAT)-DF.DV.GV.NP^H	0.07
SB-IV.DT^H^RT.TV.CL	0.07	CP.ON(<SB)-N^H	0.07
SB-NAS.DT^AJ^N^H	0.07	CP.ON(<SCL)-DF.AR^AJ^H	0.07
		CP.ON(<SCL)-	
SB-NAS.DT^H^PP(of)	0.07	DF.AR^H^RT.RV.CL^PT.ST.NR.IT.AP	0.07
SB-NUM.CD^AJ^H^RT.RV.CL	0.07	CP.OVER(<AB.AT.CAU)-PS.DT^N^H	0.07
SB-			
NUM.CD^H^PP(of)^NR.RV.CL	0.07	CP.PLUS(<AB.AT.AC)-PS.DT^H^AP.TI.CL	0.07
SB-			
NUM.CD^NUM.GO^H^PP(of)	0.07	CP.THROUGH(<AB.AT.MS)-AJ^H	0.07
SB-NUM.CD^PV.PP(of			
DF.AR^AJ^AJ^H^TI.CL)	0.07	CP.THROUGH(<AB.AT.MS)-AJ^H^PP(with)	0.07
SB-NUM.CD^PV.PP(of			
DF.AR^AJ^H)	0.07	CP.TO(<AB.AT.RS)-DF.AR^H^RT.RV.CL	0.07
SB-NUM.CD^PV.PP(of			
DF.AR^AJ^H^NR.PL.ING.CL)	0.07	CP.TO(<AB.AT.RS)-DF.AR^H	0.07
SB-NUM.CD^PV.PP(of			
DF.AR^NUM.GO^H^RT.RV.CL)	0.07	CP.TO(<AB.AT.RS)-DF.AR^H^PL.ED.CL	0.07
SB-NUM.CD^PV.PP(of			
PS.DT^H^PP(behind))	0.07	CP.TO(<AB.AT.RS)-DF.AR^H^PP(of)	0.07
SB-NUM.CD^PV.PP(of			
PS.DT^H^PP(since))	0.07	CP.TO(<AB.AT.RS)-DF.AR^H^RT.RV.CL	0.07
SB-NUM.GO^AJ^AJ^H	0.07	CP.TO(<AB.AT.RS)-DF.DV.GV.NP^N^H	0.07
SB-NUM.GO^AJ^H	0.07	CP.TO(<AB.AT.RS)-IN.AR^AJ^H^PP(on)	0.07
SB-NUM.GO^H^PL.ED.CL	0.07	CP.TO(<AB.AT.TM)-PDT^H^RT.RV.CL(as)	0.07
SB-PDT^AJ^H	0.07	CP.TO(<AB.DJ.CO)-PS.DT^H	0.07
SB-PS.DT^AJ^H^PP(in)	0.07	CP.TO(<CP.FOR<SB)-H	0.07
SB-PS.DT^H^AP.TI.CL	0.07	CP.TO(<DO)-PS.DT^N^H	0.07
SB-PS.DT^H^PP(as)	0.07	CP.TO(<IO)-N^H	0.07
SB-PS.DT^N^H	0.07	CP.TO(<IO)-UV.PDT^DF.AR^H^PL.ED.CL	0.07
SB-PS.DT^NUM.GO^H	0.07	CP.TO(<NA)-N^H	0.07
SB-PS.DT^NUM.OR^H^PP(of)	0.07	CP.TO(<POCL)-H	0.07

SB-QT.PN^PV.PP(of DM.DT(THESE)^H)	0.07	CP.TO(<SB)-DF.AR^H^PL.ED.CL	0.07
SB-QT^AJ^H	0.07	CP.TO(<SB)-DF.AR^H^PP(between)	0.07
SB-QT^DF.AR^H	0.07	CP.TO(<SB)-H	0.07
SB-QT^H^RT.RV.CL	0.07	CP.TO(<SCL)-DF.AR^H^RT.RV.CL	0.07
SB-QV.PV^DM.DT(THAT)^H	0.07	CP.TO(<SCL)-DM.DT(THAT)^H	0.07
SB-RV.DT^AJ^H	0.07	CP.TO(<SCL)-IN.AR^AJ^H^FL.ST.NR.IT.AP^PL.ED.CL	0.07
SB-UV.PDT^DF.AR^H^RT.RV.CL	0.07	CP.TO(<SCL)-NAS.DT^QL.PV^H^RT.RV.CL	0.07
SB-UV.PDT^NUM.GO^H^PP(of)	0.07	CP.TO(CAJ<DO)-N^H	0.07
SCL-AJ^H^PL.ED.CL	0.07	CP.UNLIKE(<AB.AT.MAN)-DF.AR^H^PP(of)^NR.RV.CL	0.07
SCL-AJ^H^PP(of)	0.07	CP.UNTIL(<AB.AT.TM)-DF.AR^H^RT.RV.CL(where)	0.07
SCL-DF.AR^AJ^AJ^H^PP(in)	0.07	CP.UNTIL(<AB.AT.TM)-PDT^H^RT.RV.CL(as)	0.07
SCL-DF.AR^AJ^H^PP(to)	0.07	CP.UPON(<AB.AT.RS)-IN.AR^AJ^H^NR.RV.CL	0.07
SCL-DF.AR^AJ^H^RT.RV.CL	0.07	CP.UPON(<CAJ<SB)-PS.DT^H^PP(as)	0.07
SCL-DF.AR^H^PP(behind)	0.07	CP.WITH REGARD TO(<AB.AT.RS)-PS.DT^H^AP.TI.CL	0.07
SCL-DF.AR^H^PP(of)^AP.TI.CL	0.07	CP.WITH(<AB.AT.CAU)-DM.DT(THIS)^H^PP(in)	0.07
SCL-DF.AR^H^PP(of)^PP(in)	0.07	CP.WITH(<AB.AT.CAU)-DM.DT(THIS)^QL.PV^H	0.07
SCL-DF.AR^H^PP(of)^PP(on)	0.07	CP.WITH(<AB.AT.CAU)-H	0.07
SCL-DF.AR^H^PP(to)	0.07	CP.WITH(<AB.AT.CAU)-H^PP(at)	0.07
SCL-DF.AR^H^PP(with)	0.07	CP.WITH(<AB.AT.CG.CI)-AJ^H	0.07
SCL-DF.AR^H^RT.RV.CL(when)	0.07	CP.WITH(<AB.AT.CG.CI)-DF.AR^AJ^H^PT.WK.NR.AP	0.07
SCL-DF.AR^H^RT.RV.CL^PP(of)	0.07	CP.WITH(<AB.AT.CG.CI)-DF.AR^H^AP.THAT.CL	0.07
SCL-DF.AR^NUM.GO^H^RT.RV.CL	0.07	CP.WITH(<AB.AT.CG.CI)-NUM.CD^H	0.07
SCL-DF.AR^NUM.OR^H^PP(of)	0.07	CP.WITH(<AB.AT.MAN)-AJ^AJ^H	0.07
SCL-DF.DV.GV.NP^AJ^H	0.07	CP.WITH(<AB.AT.MAN)-AS.DT^H	0.07
SCL-DM.DT(THIS)^AJ^H	0.07	CP.WITH(<AB.AT.MAN)-H	0.07
SCL-DM.DT(THIS)^H^PP(of)	0.07	CP.WITH(<AB.AT.MAN)-PS.DT^AJ^AJ^H	0.07
SCL-H	0.07	CP.WITH(<AB.AT.MS)-AJ^H^NR.RV.CL(where)	0.07
SCL-H^AP.TI.CL	0.07	CP.WITH(<AB.AT.MS)-H^PP(in)	0.07
SCL-H^PP(of)^RT.RV.CL	0.07	CP.WITH(<AB.AT.MS)-H^PP(of)	0.07
SCL-H^RT.RV.CL	0.07	CP.WITH(<AB.AT.MS)-IN.AR^AJ^H	0.07
SCL-IF.AV^IN.AR^H	0.07	CP.WITH(<AB.AT.MS)-IN.AR^H^PP(for)	0.07
SCL-IN.AR^AJ^H^NR.AP.THAT.CL	0.07	CP.WITH(<AB.AT.MS)-NUM.GO^AJ^H^PL.ED.CL	0.07
SCL-IN.AR^AJ^H^NR.PL.ED.CL	0.07	CP.WITH(<AB.AT.RS)-DF.AR^H^PP(of)^PP(in)	0.07
SCL-IN.AR^AJ^H^NR.PL.ING.CL	0.07	CP.WITH(<AB.AT.RS)-DM.DT(THIS)^QL.PV^AJ^H^RT.RV.CL	0.07
SCL-IN.AR^AJ^H^PP(between)	0.07	CP.WITH(<AB.AT.RS)-IN.AR^AJ^H^PP(of)	0.07
SCL-IN.AR^AJ^H^PT.WK.NR.IT.AP	0.07	CP.WITH(<AB.AT.RS)-NAS.PN^PV.PP(of DF.AR^H^PP in)	0.07
SCL-IN.AR^AJ^H^RT.RV.CL	0.07	CP.WITH(<AB.AT.RS)-PS.DT^H	0.07
SCL-IN.AR^AJ^AJ^H	0.07	CP.WITH(<AB.AT.TM)-PS.DT^H^PP(on)	0.07

SCL-IN.AR^H^PL.ED.CL	0.07	CP.WITH(<AP.NR<SCL)-IN.AR^AJ^H^PP(for)	0.07
SCL-IN.AR^H^PP(of)^RT.RV.CL	0.07	CP.WITH(<CP.AROUND<AB.AT.SPA)-DF.AR^AJ^AJ^AJ^H	0.07
SCL-IN.AR^H^PP(to)	0.07	CP.WITH(<CP.OF<SB)-IN.AR^AJ^N^H	0.07
SCL-IN.AR^H^RT.RV.CL	0.07	CP.WITH(<CP.ON<CP.CONTRARY TO<AB.AT.MAN)-PS.DT^AJ^AJ^H^PT.WK.NR.PA.AP	0.07
SCL-IN.AR^N^H^PP(against)	0.07	CP.WITH(<DO)-NG.DT^H^TI.CL	0.07
SCL-IN.AR^N^H^PP(of)	0.07	CP.WITH(<DO)-N^H	0.07
SCL-NAS.DT^H	0.07	CP.WITH(<PT.WK.NR.AP<SB)-DF.AR^H^PP(of)	0.07
SCL-NUM.CD^H	0.07	CP.WITH(<PT.WK.NR.AP<SCL)-PS.DT^AJ^H	0.07
SCL-NUM.CD^NUM.GO^H^RT.RV.CL(where)	0.07	CP.WITH(<SB)-H^PP(of)	0.07
SCL-NUM.CD^PV.PP(of DF.AR^N^H)	0.07	CP.WITH(<SCL)-DF.AR^H^AP.TI.CL	0.07
SCL-NUM.CD^PV.PP(of DF.AR^AJ^H^PP(about))	0.07	CP.WITH(<SCL)-DF.AR^H	0.07
SCL-NUM.CD^PV.PP(of DF.AR^AJ^H^RT.RV.CL)	0.07	CP.WITH(<SCL)-H^PP(in)	0.07
SCL-NUM.CD^PV.PP(of DF.AR^H^RT.RV.CL)	0.07	CP.WITH(<SCL)-PS.DT^AJ^H	0.07
SCL-NUM.CD^QV.PV^H	0.07	CP.WITHIN(<AB.AT.SPA)-DM.DT(THESE)^H	0.07
SCL-NUM.GO^H^PP(of)	0.07	CP.WITHIN(<PO)-DF.DV.GV.NP^AJ^H	0.07
SCL-PDT(CV.RU)^IN.AR^AJ^H	0.07	CP.WITHOUT(<AB.AT.CC)-H	0.07
SCL-PS.DT^AJ^H	0.07	CP.WITHOUT(<AB.AT.CC)-DF.AR^AJ^H	0.07
SCL-PS.DT^H	0.07	FL.ST.NR.AP(<AB.AT.AG)-IN.AR^AJ^H^PP(of)	0.07
SCL-PS.DT^H^PP(behind)	0.07	FL.ST.NR.AP(<CP.AT<IO)-DF.AR^N^AJ^H	0.07
SCL-PS.DT^NUM.OR^H	0.07	FL.ST.NR.AP(<DO)-DF.AR^H	0.07
SCL-QV.PV^H^PT.WK.NR.IT.AP	0.07	FL.ST.NR.AP(<PO)-IN.AR^H^RT.RV.CL	0.07
SCL&DO-QT^H	0.07	FL.ST.NR.IT.AP(<CP.TO<AB.AT.SPA)-H^PP(of)	0.07
SB(not)-AJ^H^RT.RV.CL	0.07	FL.ST.NR.RF.AP(<CP.OF<DO&CP.ON<AB.AT.RS)-H^RT.RV.CL	0.07
SB(not)-AJ^N^H^RT.RV.CL	0.07	PT.ST.NR.AP(<CP.FOR<DO)-IN.AR^AJ^H	0.07
SB(not)-QT^AJ^H^PP(in)	0.07	PT.ST.NR.ATT.AP(<PO)-IN.AR^AJ^H	0.07
SB(not)-DM.DT(THAT)^H	0.07	PT.ST.NR.ATT.AP(<PO)-IN.AR^H^PP(of)	0.07
SB(not)-H	0.07	PT.ST.NR.DN.AP(<CP.FROM<SCL)-IN.AR^AJ^H^PL.ED.CL	0.07
SB(not)-H^(for n AP.TI.CL)	0.07	PT.ST.NR.EM.AP(<PO)-DF.AR^AJ^H^RT.RV.CL^PP(with)	0.07
		PT.ST.NR.IT.AP(<CP.ON<DO)-DF.AR^AJ^H	0.07

APPENDIX 12 ENCAPSULATING DIRECTIONS AND THEIR ANTECEDENTS (COMPLETE LIST)

Direction of encapsulation and antecedent (%)			
EXO^NA	8.09	CF.INTRA^AF.INTRA^AF.INTER& LC.CL(SI)^LC.SNP(*THAT*)^GB.ET.PR	0.07
AF.INTER&GB.ET.PR	5.74	CF.INTRA^AF.INTRA^AF.INTER& LC.CL(SI)^LC.SNP(*THESE*)^ GB.ET.PR	0.07
CF.INTRA&LC.CT	4.77	CF.INTRA^AF.INTRA^AF.INTER& LC.CL(SI)^LC.SNP(*THESE*)^LC.SNP	0.07
CF.INTRA&LC.CL	4.49	CF.INTRA^AF.INTRA^AF.INTER& LC.CL(SI)^LC.SNP(*THIS*)^LC.CL	0.07
CF.INTRA^EXO&LC.CT(SI)^NA	4.35	CF.INTRA^AF.INTRA^AF.INTER& LC.CL(SI)^LC.SNP(*THIS*)^LC.CL[QU]	0.07
CF.INTER&GB.ET.PR	3.73	CF.INTRA^AF.INTRA^AF.INTER& LC.CL(SI)^LC.SNP(*THOSE*)^ GB.ET.PR	0.07
CF.INTRA&LC.CL(AP)	3.73	CF.INTRA^AF.INTRA^AF.INTER& LC.CT(SI)^LC.SNP(*IT*)^LC.CNP	0.07
AF.INTRA^EXO& LC.PRM(SI)^NA	2.97	CF.INTRA^AF.INTRA^AF.INTER& LC.CT(SI)^LC.SNP(*IT*)^LC.SC	0.07
CF.INTRA^EXO&LC.CL(SI)^NA	2.35	CF.INTRA^AF.INTRA^AF.INTER& LC.CT(SI)^LC.SNP(*THAT*)^GB.ET.PR	0.07
CF.INTRA^CF.INTER& LC.CT(SI)^GB.ET.PR	2.28	CF.INTRA^AF.INTRA^AF.INTER& LC.CT(SI)^LC.SNP(*THAT*)^LC.CL	0.07
CF.INTRA^AF.INTER& LC.CT(SI)^GB.ET.PR	1.87	CF.INTRA^AF.INTRA^AF.INTER& LC.CT(SI)^LC.SNP(*THAT*)^LC.SNP	0.07
AF.INTER&LC.CL	1.52	CF.INTRA^AF.INTRA^AF.INTER& LC.CT(SI)^LC.SNP(*THEM*)^LC.CNP	0.07
CF.INTRA^EXO& LC.PM(SI)^NA	1.52	CF.INTRA^AF.INTRA^AF.INTER& LC.CT(SI)^LC.SNP(*THIS*)^LC.CNP	0.07
AF.INTRA^AF.INTER& LC.PRM(SI)^GB.ET.PR	1.31	CF.INTRA^AF.INTRA^AF.INTER^ AF.INTER(SN2)&LC.CT(SI)^ LC.SNP(*THEY*)^LC.SNP^ GB.ET.PR(SN2)	0.07
AF.INTRA&LC.SNP	1.17	CF.INTRA^AF.INTRA^AF.INTER(SN2) &LC.PM(SI)^LC.SNP^LC.SC(SN2)	0.07
AF.INTRA&LC.CNP	1.11	CF.INTRA^AF.INTRA^AF.INTRA& LC.CL(SI)^LC.SNP(*THIS*)^LC.SNP	0.07
AF.INTRA^CF.INTER& LC.PRM(SI)^GB.ET.PR	1.04	CF.INTRA^AF.INTRA^AF.INTRA& LC.CT(SI)^LC.SNP(*IT*)^LC.CNP	0.07
AF.INTER&LC.SC	0.97	CF.INTRA^AF.INTRA^AF.INTRA& LC.CT(SI)^LC.SNP(*THAT*)^LC.CNP	0.07
CF.INTRA^CF.INTER& LC.CL(SI)^GB.ET.PR	0.97	CF.INTRA^CF.INTRA^AF.INTER& LC.CT(SI)^LC.CL(SI)^GB.ET.PR	0.07
AF.INTRA&LC.PRM	0.90	CF.INTRA^CF.INTRA^AF.INTER& LC.CT(SI)^LC.CT(SI)^GB.ET.PR	0.07
CF.INTRA^AF.INTRA& LC.CT(SI)^LC.SNP	0.90	CF.INTRA^CF.INTRA^AF.INTRA& LC.CT(SI)^LC.CT(SI)^LC.CL	0.07
CF.INTRA^CF.INTRA& LC.CT(SI)^LC.CL	0.90	CF.INTRA^CF.INTRA^AF.INTRA& AF.INTER&LC.CT(SI)^ LC.CT(SI)^LC.SNP(*THESE*)^ LC.SC[QU]	0.07
AF.INTER&LC.SNP	0.76	CF.INTRA^CF.INTRA^CF.INTRA^	0.07

		AF.INTER&LC.CT(SI)^LC.CL(SI)^ LC.CNP(SI)^GB.ET.PR	
CF.INTRA^AF.INTER& LC.CL(SI)^GB.ET.PR	0.76	AF.INTER^AF.INTRA^CF.INTER& LC.ET.PR.CNP(SI)^ LC.PRM(SI)^GB.ET.PR	0.07
CF.INTER&LC.SC	0.76	AF.INTER^AF.INTRA^CF.INTRA& LC.ET.PR.SNP(SI)^ LC.PRM(SI)^LC.CNP	0.07
CF.INTRA^CF.INTER& LC.PM(SI)^GB.ET.PR	0.76	AF.INTER^AF.INTRA^CF.INTRA& LC.ET.PR.SNP(SI)^ LC.PRM(SI)^LC.SNP	0.07
CF.INTRA^CF.INTRA& LC.CT(SI)^LC.CL(AP)	0.69	AF.INTER^AF.INTRA^CF.INTRA& LC.SC(SI)^LC.PRM(SI)^ LC.RN.SG[QU]	0.07
AF.INTRA^CF.INTRA^EXO& LC.PRM(SI)^LC.CL(SI)^NA	0.62	AF.INTER^AF.INTRA^CF.INTRA& LC.SC(SI)^LC.SNP(SI)^LC.CL	0.07
AF.INTRA&LC.CL	0.55	AF.INTER^CF.INTER& GB.ET.PR(SI)^LC.TB	0.07
AF.INTRA^EXO& LC.SNP(SI)^NA	0.55	AF.INTER^CF.INTER& GB.ET.PR(SI)+LC.IM(SI)^GB.ET.PR	0.07
AF.INTER&LC.CNP	0.48	AF.INTER^CF.INTER& LC.CNP(SI)^GB.ET.PR	0.07
CF.INTRA^AF.INTRA& LC.CT(SI)^LC.CNP	0.48	AF.INTER^CF.INTER& LC.CNP(SI)^LC.FR	0.07
AF.INTRA^CF.INTER& LC.PRM(SI)^LC.SC	0.41	AF.INTER^CF.INTER& LC.ET.PR.SNP(SI)^GB.ET.PR	0.07
CF.INTRA&LC.CL(NR.AP)	0.41	AF.INTER^CF.INTER& LC.SC(SI)^GB.ET.PR	0.07
CF.INTRA&LC.CL[QU]	0.41	AF.INTER^CF.INTER& LC.SC(SI)^LC.FR	0.07
CF.INTRA^CF.INTRA& LC.CL(SI)^LC.CL	0.41	AF.INTER^CF.INTER& LC.SNP(SI)^LC.FG	0.07
CF.INTRA^CF.INTRA& LC.CL(SI)^LC.CNP	0.41	AF.INTER^CF.INTRA& GB.ET.PR(SI)^LC.CL	0.07
AF.INTRA^AF.INTRA^EXO& LC.PRM(SI)^LC.PRM(SI)^NA	0.41	AF.INTER^CF.INTRA& GB.ET.PR(SI)^LC.CL(NR.AP)	0.07
AF.INTRA^CF.INTRA^EXO& LC.PRM(SI)^LC.PM(SI)^NA	0.41	AF.INTER^CF.INTRA& GB.ET.PR(SI)^LC.CNP(NR.AP)	0.07
AF.INTER&LC.ET.PR.CL	0.35	AF.INTER^CF.INTRA& GB.ET.PR(SI)^LC.SNP+CNP(NR.AP)	0.07
CF.INTRA^AF.INTER& LC.PM(SI)^GB.ET.PR	0.35	AF.INTER^CF.INTRA& LC.CL[QU](SI)^GB.ET.PR[QU]	0.07
AF.INTRA^CF.INTRA& LC.PRM(SI)^LC.CL	0.35	AF.INTER^CF.INTRA& LC.CNP(SI)^LC.CNP	0.07
AF.INTRA^CF.INTRA& LC.PRM(SI)^LC.CL(AP)	0.35	AF.INTER^CF.INTRA& LC.ET.PR.CL(SI)^LC.CL	0.07
AF.INTRA^CF.INTRA& LC.PRM(SI)^LC.CT	0.35	AF.INTER^CF.INTRA& LC.ET.PR.SC(SI)^LC.RN.SG	0.07
CF.INTER&LC.CL	0.35	AF.INTER^CF.INTRA& LC.SC(SI)^LC.CNP	0.07
CF.INTRA&LC.CNP	0.35	AF.INTER^CF.INTRA& LC.SC(SI)^LC.SNP	0.07
CF.INTRA^CF.INTRA& LC.CT(SI)^LC.CNP	0.35	AF.INTER^CF.INTRA& LC.SC(SI)^LC.SNP+CNP	0.07
CF.INTRA^CF.INTRA& LC.PM(SI)^LC.CL	0.35	AF.INTER^CF.INTRA& LC.SNP(SI)^LC.CL	0.07

AF.INTER^CF.INTER& GB.ET.PR	0.35	AF.INTER^CF.INTRA^CF.INTER& GB.ET.PR(SI)^LC.CT(SI)^GB.ET.PR	0.07
UNC&UNC(NA)	0.35	AF.INTER^CF.INTRA^CF.INTER& GB.ET.PR(SI)^LC.SNP(NR.AP)^ GB.ET.PR	0.07
AF.INTRA^AF.INTER& LC.SNP(*IT*)^GB.ET.PR	0.28	AF.INTER^CF.INTRA^CF.INTER& LC.ET.PR.SC(SI)^LC.PM(SI)^ GB.ET.PR	0.07
CF.INTRA^AF.INTRA& LC.CL(SI)^LC.CNP	0.28	AF.INTER^CF.INTRA^CF.INTRA& LC.CNP(SI)^LC.CT(SI)^LC.SNP+CNP	0.07
CF.INTRA^AF.INTRA& LC.CT(SI)^LC.CL	0.28	AF.INTER^CF.INTRA^CF.INTRA& LC.SNP(SI)^LC.CT(SI)^LC.SNP+CNP	0.07
AF.INTRA^CF.INTRA^ CF.INTER&LC.PRM(SI)^ LC.CT(SI)^GB.ET.PR	0.28	AF.INTER^CF.INTRA^CF.INTRA^ CF.INTRA&LC.CL(SI)^LC.CT(SI)^ LC.SNP(*THESE*)^LC.SNP	0.07
CF.INTRA^CF.INTER& LC.CT(SI)^LC.CL	0.28	AF.INTRA^AF.INTRA^AF.INTRA^ AF.INTRA^CF.INTER& LC.CNP(SI)^LC.PRM(SI)^ LC.PRM(SI)^LC.PRM(SI)^GB.ET.PR	0.07
CF.INTRA^CF.INTER& LC.CT(SI)^LC.SC	0.28	AF.INTRA^AF.INTRA^AF.INTRA^ CF.INTER&LC.PRM(SI)^LC.PRM(SI)^ LC.SNP(SI)^GB.ET.PR	0.07
AF.INTRA^CF.INTRA^EXO& LC.PRM(SI)^LC.CT(SI)^NA	0.28	AF.INTRA^AF.INTRA^CF.INTRA^ CF.INTRA&LC.PRM(SI)^LC.PRM(SI)^ LC.CL(SI)^LC.SNP	0.07
AF.INTRA^AF.INTER& LC.SNP(*THAT*)^LC.SC	0.21	AF.INTRA^CF.INTER& LC.PRM(SI)^LC.CL[QU]	0.07
AF.INTRA^AF.INTRA& LC.PRM(SI)^LC.SNP	0.21	AF.INTRA^CF.INTER& LC.PRM(SI)^LC.CNP	0.07
CF.INTRA^AF.INTER& LC.CT(SI)^LC.CL	0.21	AF.INTRA^CF.INTER& LC.PRM(SI)^LC.CT	0.07
CF.INTRA^AF.INTRA& LC.CL(SI)^LC.CL	0.21	AF.INTRA^CF.INTER& LC.PRM(SI)^LC.SC[QU]	0.07
CF.INTRA^AF.INTRA& LC.CL(SI)^LC.SNP	0.21	AF.INTRA^CF.INTER& LC.PRM(SI)^LC.SNP	0.07
CF.INTRA&LC.SNP	0.21	AF.INTRA^CF.INTRA& LC.CL(SI)^LC.CL	0.07
CF.INTRA&LC.SNP(RT.AP)	0.21	AF.INTRA^CF.INTRA& LC.CNP(SI)^LC.CL	0.07
CF.INTRA^CF.INTER& LC.CL^GB.ET.PR	0.21	AF.INTRA^CF.INTRA& LC.PRM(SI)^LC.CL(NR.AP)	0.07
CF.INTRA^CF.INTRA& LC.CL(SI)^LC.SNP	0.21	AF.INTRA^CF.INTRA& LC.PRM(SI)^LC.CNP(NR.AP)	0.07
CF.INTRA^CF.INTRA^EXO& LC.CT(SI)^LC.CL(SI)^NA	0.21	AF.INTRA^CF.INTRA& LC.PRM(SI)^LC.PP	0.07
AF.INTER^EXO& GB.ET.PR^NA	0.21	AF.INTRA^CF.INTRA& LC.SNP(SI)^LC.CNP	0.07
CF.INTRA^AF.INTER^ CF.INTER&LC.CT(SI)^ GB.ET.PR	0.21	AF.INTRA^CF.INTRA& LC.SNP(SI)^LC.SNP	0.07
CF.INTRA^AF.INTER^ CF.INTER&LC.PM(SI)^ GB.ET.PR	0.21	AF.INTRA^CF.INTRA^AF.INTRA^ AF.INTER^CF.INTER& LC.PRM(SI)^LC.CL(SI)^ LC.SNP(*THIS*)(SI)^LC.SNP(SI)^LC.CL	0.07
AF.INTER&GB.ET.PR[QU]	0.14	AF.INTRA^CF.INTRA^CF.INTER& LC.PRM(SI)^LC.CL(SI)^GB.ET.PR	0.07

AF.INTER&LC.CL[QU]	0.14	AF.INTRA^CF.INTRA^CF.INTER&LC.PRM(SI)^LC.CL(SI)^LC.SC	0.07
AF.INTER&LC.ET.PR.SNP	0.14	AF.INTRA^CF.INTRA^CF.INTER&LC.SNP(SI)^LC.CL(SI)^GB.ET.PR	0.07
AF.INTER&LC.SC[QU]	0.14	AF.INTRA^CF.INTRA^CF.INTRA&LC.PRM(SI)^LC.CL(SI)^LC.CL	0.07
AF.INTER^AF.INTER&LC.CL(SI)^GB.ET.PR	0.14	AF.INTRA^CF.INTRA^CF.INTRA&LC.PRM(SI)^LC.CT(SI)^LC.CNP	0.07
AF.INTER^AF.INTER&LC.ET.PR.CNP(SI)^GB.ET.PR	0.14	AF.INTRA^CF.INTRA^CF.INTRA&LC.PRM(SI)^LC.PM(SI)^LC.CL	0.07
AF.INTRA&LC.SNP+CNP	0.14	AF.INTRA^CF.INTRA^CF.INTRA&LC.PRM(SI)^LC.PM(SI)^LC.CNP	0.07
AF.INTRA^AF.INTER&LC.PRM(SI)^LC.CL	0.14	AF.INTRA^CF.INTRA^CF.INTRA&LC.PRM(SI)^LC.SNP(NR.AP)(SI)^LC.CL	0.07
AF.INTRA^AF.INTER&LC.PRM(SI)^LC.ET.PR.CL	0.14	AF.INTRA^CF.INTRA^CF.INTRA^CF.INTER&LC.SNP(SI)^LC.CT(SI)^LC.CL(SI)^GB.ET.PR	0.07
AF.INTRA^AF.INTER&LC.SNP(*IT*)^LC.CL	0.14	CF.INTER&LC.CL[QU]	0.07
AF.INTRA^AF.INTER&LC.SNP(*THAT*)^LC.SNP	0.14	CF.INTER&LC.ET.PR.CL	0.07
AF.INTRA^AF.INTER&LC.SNP(*THIS*)^GB.ET.PR	0.14	CF.INTER&LC.ET.PR.SC	0.07
AF.INTRA^AF.INTRA&LC.PRM(SI)^LC.CL	0.14	CF.INTER&LC.ET.PR.SC[QU]	0.07
AF.INTRA^AF.INTRA&LC.SNP(*IT*)^LC.CNP	0.14	CF.INTER&LC.SC[QU]	0.07
AF.INTRA^AF.INTRA&LC.SNP(*THAT*)^LC.CL	0.14	CF.INTER&LC.SNP+LC.AJ	0.07
AF.INTRA^AF.INTRA^AF.INTER&LC.PRM(SI)^LC.PRM(SI)^GB.ET.PR	0.14	CF.INTER&LC.SNP[NUM]	0.07
AF.INTRA^CF.INTRA^AF.INTRA&LC.PRM(SI)^LC.PM(SI)^LC.CL	0.14	CF.INTER&LC.SNP+CNP	0.07
CF.INTRA^AF.INTRA&LC.PM(SI)^LC.CL	0.14	CF.INTER&LC.SNP+CNP[QU]	0.07
CF.INTRA^AF.INTRA^AF.INTER&LC.CT(SI)^LC.SNP(*IT*)^GB.ET.PR	0.14	CF.INTRA&LC.CNP(RT.AP)	0.07
CF.INTRA^AF.INTRA^AF.INTER&LC.CT(SI)^LC.SNP(*IT*)^LC.SNP	0.14	CF.INTRA&LC.PM	0.07
CF.INTRA^AF.INTRA^AF.INTER&LC.CT(SI)^LC.SNP(*THAT*)^LC.SC	0.14	CF.INTRA&LC.PP(NR.AP)	0.07
CF.INTRA^AF.INTRA^AF.INTER&LC.CT(SI)^LC.SNP(*THESE*)^GB.ET.PR	0.14	CF.INTRA&UNC(NA)	0.07
CF.INTRA^AF.INTRA^AF.INTER(SN2)&LC.CT(SI)^LC.SNP^GB.ET.PR(SN2)	0.14	CF.INTRA^AF.INTRA^CF.INTRA&LC.CL(SI)^LC.SNP(SI)^LC.CL	0.07
CF.INTRA^CF.INTRA^AF.INTRA&LC.CT(SI)^LC.CL(SI)^LC.SNP	0.14	CF.INTRA^CF.INTER&LC.CL(SI)^LC.CL	0.07

AF.INTER^CF.INTER& GB.ET.PR(SI)^GB.ET.PR	0.14	CF.INTRA^CF.INTER& LC.CL(SI)^LC.SNP	0.07
AF.INTRA^CF.INTER& LC.PRM(SI)^LC.CL	0.14	CF.INTRA^CF.INTER&LC.CL(SI)^OV	0.07
AF.INTRA^CF.INTER& LC.SNP(SI)^GB.ET.PR	0.14	CF.INTRA^CF.INTER& LC.CNP^GB.ET.PR	0.07
AF.INTRA^CF.INTRA^ CF.INTER&LC.PRM(SI)^ LC.PM(SI)^GB.ET.PR	0.14	CF.INTRA^CF.INTER& LC.CT(SI)^GB.ET.PR[QU]	0.07
AF.INTRA^CF.INTRA^ CF.INTRA&LC.PRM(SI)^ LC.CL(SI)^LC.CNP	0.14	CF.INTRA^CF.INTER& LC.CT(SI)^LC.AJ	0.07
AF.INTRA^CF.INTRA^ CF.INTRA&LC.PRM(SI)^ LC.CL(SI)^LC.SNP(NR.AP)	0.14	CF.INTRA^CF.INTER& LC.CT(SI)^LC.CL[QU]	0.07
CF.INTER&LC.CNP	0.14	CF.INTRA^CF.INTER& LC.CT(SI)^LC.ET.PR.SC	0.07
CF.INTER&OV	0.14	CF.INTRA^CF.INTER& LC.CT(SI)^LC.ET.PR.TB	0.07
CF.INTRA&LC.SC	0.14	CF.INTRA^CF.INTER& LC.CT(SI)^LC.FG	0.07
CF.INTRA&LC.SNP(NR.AP)	0.14	CF.INTRA^CF.INTER& LC.CT(SI)^LC.IJ	0.07
CF.INTRA&LC.SNP+CNP	0.14	CF.INTRA^CF.INTER& LC.CT(SI)^LC.SC[QU]	0.07
CF.INTRA^CF.INTER& LC.CT(SI)^LC.SNP	0.14	CF.INTRA^CF.INTER& LC.PM(SI)^LC.ET.PR.CL	0.07
CF.INTRA^CF.INTER& LC.CT(SI)^LC.TB	0.14	CF.INTRA^CF.INTER& LC.SC^GB.ET.PR	0.07
CF.INTRA^CF.INTER& LC.PM(SI)^LC.CL	0.14	CF.INTRA^CF.INTER^ CF.INTER(SN2)&LC.CT(SI)^ LC.CNP^GB.ET.PR(SN2)	0.07
CF.INTRA^CF.INTER& LC.PM(SI)^LC.CNP	0.14	CF.INTRA^CF.INTER(SN2)& LC.SNP^GB.ET.PR(SN2)	0.07
CF.INTRA^CF.INTRA& LC.CL(SI)^LC.CT	0.14	CF.INTRA^CF.INTRA& LC.CL(SI)^LC.CL(AP)	0.07
CF.INTRA^CF.INTRA& LC.PM(SI)^LC.CNP	0.14	CF.INTRA^CF.INTRA& LC.CT(SI)^GB.ET.PR[QU]	0.07
CF.INTRA^CF.INTRA^ CF.INTER&LC.CL(SI)^ LC.CL^GB.ET.PR	0.14	CF.INTRA^CF.INTRA& LC.CT(SI)^LC.CL(NR.AP)	0.07
CF.INTRA^CF.INTRA^ CF.INTER&LC.CL(SI)^ LC.CT(SI)^GB.ET.PR	0.14	CF.INTRA^CF.INTRA& LC.CT(SI)^LC.CNP(NR.AP)	0.07
CF.INTRA^CF.INTRA^ CF.INTER&LC.CT(SI)^ LC.CL^GB.ET.PR	0.14	CF.INTRA^CF.INTRA& LC.CT(SI)^LC.PP	0.07
CF.INTRA^CF.INTRA^ CF.INTER&LC.PM(SI)^ LC.PM(SI)^GB.ET.PR	0.14	CF.INTRA^CF.INTRA& LC.CT(SI)^LC.SNP	0.07
AF.INTER^CF.INTRA^EXO& GB.ET.PR(SI)^LC.CT(SI)^NA	0.14	CF.INTRA^CF.INTRA& LC.CT(SI)^LC.SNP+CNP(NR.AP)	0.07
AF.INTER^EXO& GB.ET.PR(SI)^NA	0.14	CF.INTRA^CF.INTRA& LC.PM(SI)^LC.CL(AP)	0.07
AF.INTER^EXO&LC.CL(SI)^NA	0.14	CF.INTRA^CF.INTRA& LC.PM(SI)^LC.CL(NR.AP)	0.07

AF.INTER^EXO& LC.ET.PR.CNP(SI)^NA	0.14	CF.INTRA^CF.INTRA& LC.PM(SI)^LC.CNP(NR.AP)	0.07
AF.INTRA^CF.INTRA^EXO& LC.PRM(SI)^ LC.CNP(NR.AP)(SI)^NA	0.14	CF.INTRA^CF.INTRA& LC.PM(SI)^LC.SNP(NR.AP)	0.07
AF.INTRA^EXO& LC.CNP(SI)^NA	0.14	CF.INTRA^CF.INTRA^CF.INTER& LC.CL(SI)^LC.CNP(SI)^GB.ET.PR	0.07
AF.INTER^CF.INTER& LC.SNP^LC.SC	0.14	CF.INTRA^CF.INTRA^CF.INTER& LC.CL(SI)^LC.PM(SI)^LC.ET.PR.SC	0.07
AF.INTRA^EXO&LC.CL^NA	0.14	CF.INTRA^CF.INTRA^CF.INTER& LC.CT(SI)^LC.CT(SI)^LC.CL	0.07
AF.INTRA^EXO&LC.SNP^NA	0.14	CF.INTRA^CF.INTRA^CF.INTER& LC.PM(SI)^LC.SNP^LC.CL	0.07
AF.INTER&LC.CT	0.07	CF.INTRA^CF.INTRA^CF.INTRA& LC.CL(SI)^LC.PM(SI)^LC.SNP	0.07
AF.INTER&LC.ET.PR.CNP	0.07	CF.INTRA^CF.INTRA^CF.INTRA& LC.CT(SI)^LC.PM(SI)^LC.CNP	0.07
AF.INTER&LC.ET.PR.SC	0.07	AF.INTER^CF.INTRA^EXO& GB.ET.PR(SI)^LC.CL(SI)^NA	0.07
AF.INTER^AF.INTER& LC.CNP(SI)^LC.CL	0.07	AF.INTER^EXO&LC.CNP(SI)^NA	0.07
AF.INTER^AF.INTER& LC.ET.PR.SC(SI)^GB.ET.PR	0.07	AF.INTER^EXO&LC.CT(SI)^NA	0.07
AF.INTER^AF.INTER& LC.ET.PR.SC(SI)^ LC.ET.PR.SNP[NUM]	0.07	AF.INTER^EXO& LC.ET.PR.CL(SI)^NA	0.07
AF.INTER^AF.INTER& LC.SNP(SI)^LC.CNP	0.07	AF.INTER^EXO& LC.ET.PR.SNP(SI)^NA	0.07
AF.INTER^AF.INTER(SN2)& LC.CNP^GB.ET.PR(SN2)	0.07	AF.INTER^EXO&LC.SC(SI)^NA	0.07
AF.INTER^AF.INTRA& GB.ET.PR(SI)^LC.SNP	0.07	AF.INTER^EXO&LC.SNP(SI)^NA	0.07
AF.INTER^AF.INTRA& LC.CL(SI)^LC.PRM	0.07	AF.INTRA^AF.INTRA^AF.INTRA^EXO &LC.PRM(SI)^LC.PRM(SI)^ LC.PRM(SI)^NA	0.07
AF.INTER^AF.INTRA& LC.CL(SI)^LC.SNP[QT.PN]	0.07	AF.INTRA^AF.INTRA^AF.INTRA^EXO &LC.PRM(SI)^LC.PRM(SI)^ LC.SNP(SI)^NA	0.07
AF.INTER^AF.INTRA& LC.ET.PR.SNP(SI)^LC.CNP	0.07	AF.INTRA^AF.INTRA^CF.INTRA^EXO &LC.PRM(SI)^LC.PRM(SI)^ LC.PP(NR.AP)(SI)^NA	0.07
AF.INTER^AF.INTRA& LC.SC(SI)^LC.CNP	0.07	AF.INTRA^CF.INTRA^AF.INTRA^EXO &LC.PRM(SI)^LC.PM(SI)^ LC.SNP(SI)^NA	0.07
AF.INTER^AF.INTRA& LC.SC(SI)^LC.SNP	0.07	AF.INTRA^CF.INTRA^CF.INTRA^EXO &LC.PRM(SI)^LC.CL(SI)^ LC.PM(SI)^NA	0.07
AF.INTER^AF.INTRA^ AF.INTRA&GB.ET.PR(SI)^ LC.SNP(*IT*)^LC.SNP	0.07	AF.INTRA^EXO&LC.CL(SI)^NA	0.07
AF.INTRA&GB.ET.PR	0.07	CF.INTER^EXO&LC.CL(SI)^NA	0.07
AF.INTRA&LC.AJ	0.07	CF.INTRA^CF.INTER^EXO& LC.CT(SI)^GB.ET.PR(SI)^NA	0.07
AF.INTRA&LC.PP	0.07	CF.INTRA^CF.INTRA^EXO& LC.CL(SI)^LC.CNP(NR.AP)(SI)^NA	0.07
AF.INTRA^AF.INTER&	0.07	CF.INTRA^CF.INTRA^EXO&	0.07

LC.PRM(SI)^LC.CNP		LC.CT(SI)^LC.CT(SI)^NA	
AF.INTRA^AF.INTER&		CF.INTRA^CF.INTRA^EXO&	
LC.PRM(SI)^LC.ET.PR.SC[QU]	0.07	LC.CT(SI)^LC.PM(SI)^NA	0.07
AF.INTRA^AF.INTER&		CF.INTRA^CF.INTRA^EXO&	
LC.SNP^GB.ET.PR	0.07	LC.PM(SI)^LC.PM(SI)^NA	0.07
AF.INTRA^AF.INTER&			
LC.SNP(*IT*)^GB.ET.PR[QU]	0.07	CF.INTRA^EXO&LC.CNP(SI)^NA	0.07
AF.INTRA^AF.INTER&			
LC.SNP(*IT*)^LC.SC	0.07	CF.INTRA^EXO&LC.SNP(SI)^NA	0.07
AF.INTRA^AF.INTER&		AF.INTER^AF.INTER^EXO&	
LC.SNP(*IT*)^LC.SC[QU]	0.07	LC.CNP(SI)^GB.ET.PR^NA	0.07
AF.INTRA^AF.INTER&		AF.INTER^AF.INTER^CF.INTRA&	
LC.SNP(*IT*)^LC.SNP	0.07	LC.SNP(SI)&LC.CL&LC.CL	0.07
AF.INTRA^AF.INTER&		AF.INTER^AF.INTRA^CF.INTER&	
LC.SNP(SI)^GB.ET.PR	0.07	LC.SC[QU](SI)^LC.SC[QU]	0.07
AF.INTRA^AF.INTER&			
LC.SNP(SI)^LC.ET.PR.CL	0.07	AF.INTER^CF.INTER&LC.CL^LC.SC	0.07
AF.INTRA^AF.INTER&		AF.INTER^CF.INTER&	
LC.SNP(*THAT*)^LC.AJ	0.07	LC.CNP^GB.ET.PR	0.07
AF.INTRA^AF.INTER&			
LC.SNP(*THAT*)^LC.CL	0.07	AF.INTER^CF.INTER&LC.SC^LC.SC	0.07
AF.INTRA^AF.INTER&			
LC.SNP(*THAT*)^		AF.INTER^CF.INTRA&	
LC.ET.PR.SC[QU]	0.07	LC.ET.PR.SC[QU]^LC.AJ	0.07
AF.INTRA^AF.INTER&			
LC.SNP(*THIS*)^LC.CL	0.07	AF.INTER^CF.INTRA&LC.SC^LC.CL	0.07
		AF.INTER^CF.INTRA^AF.INTER^	
AF.INTRA^AF.INTER&		CF.INTRA^CF.INTER&LC.SC(SI)^	
LC.SNP(*THIS*)^LC.SC	0.07	LC.CL(SI)^LC.CL^LC.SC^LC.SC	0.07
		AF.INTER^CF.INTRA^AF.INTER^EXO	
AF.INTRA^AF.INTER(SN2)&		&LC.ET.PR.CNP(SI)^	
LC.SNP^GB.ET.PR(SN2)	0.07	LC.PM(SI)^LC.CNP^NA	0.07
		AF.INTER^CF.INTRA^CF.INTRA^	
AF.INTRA^AF.INTER(SN2)&		AF.INTER^CF.INTER&GB.ET.PR(SI)^	
LC.SNP^LC.ET.PR.SNP(SN2)	0.07	LC.CT(SI)^LC.CL(SI)^LC.SC^LC.SC	0.07
AF.INTRA^AF.INTRA&			
LC.CNP(SI)^LC.SNP	0.07	AF.INTER^EXO&LC.SC^NA	0.07
AF.INTRA^AF.INTRA&		AF.INTRA^AF.INTER^CF.INTER&	
LC.SNP(*IT*)^LC.CL	0.07	LC.PRM(SI)^GB.ET.PR	0.07
AF.INTRA^AF.INTRA&		AF.INTRA^AF.INTER^CF.INTRA&	
LC.SNP(*THAT*)^LC.CNP+CL	0.07	LC.PRM(SI)^GB.ET.PR^LC.CNP	0.07
AF.INTRA^AF.INTRA&		AF.INTRA^AF.INTER^CF.INTRA&	
LC.SNP(*THAT*)^LC.VR	0.07	LC.SNP(*THIS*)^LC.SC^GB.ET.PR	0.07
		AF.INTRA^AF.INTRA^AF.INTER^	
AF.INTRA^AF.INTRA&		CF.INTER&LC.PRM(SI)^	
LC.SNP(*THIS*)^LC.CL	0.07	LC.SNP(*IT*)^GB.ET.PR^LC.SC	0.07
AF.INTRA^AF.INTRA^		AF.INTRA^AF.INTRA^CF.INTRA&	
AF.INTER&LC.PRM(SI)^		LC.PRM(SI)^LC.SNP(*THAT*)^	
LC.SNP(*IT*)^LC.ET.PR.CNP	0.07	LC.CL(NR.AP)	0.07
AF.INTRA^AF.INTRA^			
AF.INTER&LC.PRM(SI)^		AF.INTRA^CF.INTER&	
LC.SNP(*THAT*)^GB.ET.PR	0.07	LC.SNP^GB.ET.PR	0.07
AF.INTRA^AF.INTRA^			
AF.INTER&LC.PRM(SI)^		AF.INTRA^CF.INTER&	
LC.SNP(*THIS*)^LC.SC	0.07	LC.SNP^LC.SNP	0.07
AF.INTRA^AF.INTRA^	0.07	AF.INTRA^CF.INTRA&	0.07

AF.INTRA&LC.PRM(SI)^ LC.SNP(*IT*)^LC.SNP		LC.CNP^LC.CNP	
AF.INTRA^AF.INTRA^ CF.INTRA^AF.INTRA& LC.SNP(SI)^LC.PRM(SI)^ LC.PM(SI)^LC.SNP	0.07	AF.INTRA^CF.INTRA& LC.SNP^LC.CNP	0.07
AF.INTRA^CF.INTRA^ AF.INTER&LC.PRM(SI)^ LC.CT(SI)^GB.ET.PR	0.07	AF.INTRA^CF.INTRA& LC.SNP(*THIS*)^LC.SNP	0.07
AF.INTRA^CF.INTRA^ AF.INTER&LC.PRM(SI)^ LC.PM(SI)^LC.ET.PR.CL	0.07	AF.INTRA^CF.INTRA^AF.INTER^ CF.INTER&LC.PRM(SI)^ LC.PM(SI)^GB.ET.PR	0.07
AF.INTRA^CF.INTRA^ AF.INTRA&LC.PRM(SI)^ LC.CL(SI)^LC.CL	0.07	AF.INTRA^CF.INTRA^AF.INTER(SN2) &LC.PRM(SI)^LC.PM^GB.ET.PR(SN)	0.07
AF.INTRA^CF.INTRA^ AF.INTRA&LC.PRM(SI)^ LC.CL(SI)^LC.SNP	0.07	AF.INTRA^CF.INTRA^AF.INTRA^ CF.INTRA&LC.SNP(SI)^ LC.CL(SI)^LC.CNP^LC.SNP	0.07
AF.INTRA^CF.INTRA^ AF.INTRA^AF.INTER& LC.PRM(SI)^LC.CL(SI)^ LC.SNP(SI)^GB.ET.PR	0.07	AF.INTRA^EXO&LC.SNP(*IT*)^NA	0.07
AF.INTRA^CF.INTRA^ AF.INTRA^AF.INTER& LC.PRM(SI)^LC.PM(SI)^ LC.SNP(*THIS*)^GB.ET.PR	0.07	CF.INTRA^AF.INTER^CF.INTER& LC.CL(SI)^GB.ET.PR	0.07
AF.INTRA^CF.INTRA^ AF.INTRA^AF.INTER& LC.SNP(SI)^LC.CL(SI)^ LC.SNP(*THAT*)^GB.ET.PR	0.07	CF.INTRA^AF.INTER^CF.INTER& LC.CT(SI)^GB.ET.PR^LC.SNP	0.07
CF.INTRA^AF.INTER& LC.CL(SI)^LC.ET.PR.SC	0.07	CF.INTRA^AF.INTER^CF.INTRA& LC.CL(SI)^GB.ET.PR^LC.CL	0.07
CF.INTRA^AF.INTER& LC.CL(SI)^LC.ET.PR.SNP	0.07	CF.INTRA^AF.INTER^CF.INTRA& LC.CL(SI)^GB.ET.PR^LC.SNP	0.07
CF.INTRA^AF.INTER& LC.CT(SI)^LC.CNP	0.07	CF.INTRA^AF.INTER^CF.INTRA& LC.CT(SI)^GB.ET.PR^LC.CL	0.07
CF.INTRA^AF.INTER& LC.CT(SI)^LC.ET.PR.CNP	0.07	CF.INTRA^AF.INTER^CF.INTRA& LC.CT(SI)^GB.ET.PR^LC.SNP	0.07
CF.INTRA^AF.INTER& LC.CT(SI)^LC.ET.PR.SC	0.07	CF.INTRA^CF.INTRA^AF.INTER^ CF.INTRA&LC.CT(SI)^ LC.PM(SI)^LC.CNP^LC.CL	0.07
CF.INTRA^AF.INTER& LC.CT(SI)^LC.ET.PR.TB	0.07	CF.INTRA^AF.INTER^EXO& LC.CL(SI)^GB.ET.PR^NA	0.07
CF.INTRA^AF.INTER& LC.CT(SI)^LC.SC	0.07	CF.INTRA^AF.INTER^EXO& LC.CL(SI)^LC.ET.PR.SC^NA	0.07
CF.INTRA^AF.INTER& LC.PM(SI)^LC.CNP	0.07	CF.INTRA^AF.INTER^EXO& LC.CT(SI)^GB.ET.PR^NA	0.07
CF.INTRA^AF.INTRA& LC.CL(SI)^LC.AJ	0.07	CF.INTRA^AF.INTER(SN2)& LC.CT^GB.ET.PR(SN2)	0.07
CF.INTRA^AF.INTRA& LC.CL(SI)^LC.SC	0.07	CF.INTRA^AF.INTRA^AF.INTER^ CF.INTER&LC.CT(SI)^ LC.SNP^GB.ET.PR	0.07
CF.INTRA^AF.INTRA& LC.CT(SI)^LC.AJ	0.07	CF.INTRA^AF.INTRA^AF.INTER^ CF.INTER&LC.CT(SI)^LC.SNP(*IT*)^ LC.CNP^GB.ET.PR	0.07
CF.INTRA^AF.INTRA&	0.07	CF.INTRA^AF.INTRA^CF.INTER&	0.07

LC.CT(SI)^LC.AJP		LC.CT(SI)^LC.SNP^LC.ET.PR.SC	
CF.INTRA^AF.INTRA&		CF.INTRA^AF.INTRA^EXO&	
LC.CT(SI)^LC.CL[QU]	0.07	LC.CT(SI)^LC.SNP^NA	0.07
CF.INTRA^AF.INTRA&		CF.INTRA^CF.INTRA^AF.INTER&	
LC.CT(SI)^LC.SNP+CNP	0.07	LC.CT(SI)^LC.CNP^GB.ET.PR	0.07
CF.INTRA^AF.INTRA&		CF.INTRA^CF.INTRA^AF.INTER&	
LC.PM(SI)^LC.CNP	0.07	LC.CT(SI)^LC.SNP(*THAT*)^GB.ET.PR	0.07
CF.INTRA^AF.INTRA&		CF.INTRA^CF.INTRA^EXO&	
LC.PM(SI)^LC.SNP	0.07	LC.PM(SI)^LC.SNP^NA	0.07
CF.INTRA^AF.INTRA&			
LC.PM(SI)^LC.CL	0.07	CF.INTRA^EXO&LC.CL^NA	0.07
CF.INTRA^AF.INTRA^			
AF.INTER&LC.CL(SI)^		CF.INTRA^UNC(NA)&	
LC.SNP^LC.SNP	0.07	LC.CL(SI)^UNC(NA)	0.07

APPENDIX 13 DISTRIBUTION OF ENCAPSULATING DIRECTIONS AND ANTECEDENTS (COMBINED) ACROSS THE FOUR MAIN ENCAPSULATING RELATIONS (COMPLETE LISTS)

Anaphora	%	Cataphora	%
AF.INTER&GB.ET.PR	18.65	CF.INTRA&LC.CT	11.84
CF.INTRA^AF.INTER& LC.CT(SI)^GB.ET.PR	6.07	CF.INTRA&LC.CL	11.15
AF.INTER&LC.CL	4.94	CF.INTER&GB.ET.PR	9.26
AF.INTRA^AF.INTER& LC.PRM(SI)^GB.ET.PR	4.27	CF.INTRA&LC.CL(AP)	9.26
AF.INTRA&LC.SNP	3.82	CF.INTRA^CF.INTER& LC.CT(SI)^GB.ET.PR	5.66
AF.INTRA&LC.CNP	3.60	AF.INTRA^CF.INTER& LC.PRM(SI)^GB.ET.PR	2.57
AF.INTER&LC.SC	3.15	CF.INTRA^CF.INTER& LC.CL(SI)^GB.ET.PR	2.40
AF.INTRA&LC.PRM	2.92	CF.INTRA^CF.INTRA& LC.CT(SI)^LC.CL	2.23
CF.INTRA^AF.INTRA& LC.CT(SI)^LC.SNP	2.92	CF.INTER&LC.SC	1.89
AF.INTER&LC.SNP	2.47	CF.INTRA^CF.INTER& LC.PM(SI)^GB.ET.PR	1.89
CF.INTRA^AF.INTER& LC.CL(SI)^GB.ET.PR	2.47	CF.INTRA^CF.INTRA& LC.CT(SI)^LC.CL(AP)	1.72
AF.INTRA&LC.CL	1.80	AF.INTRA^CF.INTER& LC.PRM(SI)^LC.SC	1.03
AF.INTER&LC.CNP	1.57	CF.INTRA&LC.CL(NR.AP)	1.03
CF.INTRA^AF.INTRA& LC.CT(SI)^LC.CNP	1.57	CF.INTRA&LC.CL[QU]	1.03
AF.INTER&LC.ET.PR.CL	1.12	CF.INTRA^CF.INTRA& LC.CL(SI)^LC.CL	1.03
CF.INTRA^AF.INTER& LC.PM(SI)^GB.ET.PR	1.12	CF.INTRA^CF.INTRA& LC.CL(SI)^LC.CNP	1.03
AF.INTRA^AF.INTER& LC.SNP(*IT*)^GB.ET.PR	0.90	AF.INTRA^CF.INTRA& LC.PRM(SI)^LC.CL	0.86
CF.INTRA^AF.INTRA& LC.CL(SI)^LC.CNP	0.90	AF.INTRA^CF.INTRA& LC.PRM(SI)^LC.CL(AP)	0.86
CF.INTRA^AF.INTRA& LC.CT(SI)^LC.CL	0.90	AF.INTRA^CF.INTRA& LC.PRM(SI)^LC.CT	0.86
AF.INTRA^AF.INTER& LC.SNP(*THAT*)^LC.SC	0.67	CF.INTER&LC.CL	0.86
AF.INTRA^AF.INTRA& LC.PRM(SI)^LC.SNP	0.67	CF.INTRA&LC.CNP	0.86
CF.INTRA^AF.INTER& LC.CT(SI)^LC.CL	0.67	CF.INTRA^CF.INTRA& LC.CT(SI)^LC.CNP	0.86
CF.INTRA^AF.INTRA& LC.CL(SI)^LC.CL	0.67	CF.INTRA^CF.INTRA& LC.PM(SI)^LC.CL	0.86
CF.INTRA^AF.INTRA& LC.CL(SI)^LC.SNP	0.67	AF.INTRA^CF.INTRA^CF.INTER& LC.PRM(SI)^LC.CT(SI)^GB.ET.PR	0.69
AF.INTER&GB.ET.PR[QU]	0.45	CF.INTRA^CF.INTER& LC.CT(SI)^LC.CL	0.69
AF.INTER&LC.CL[QU]	0.45	CF.INTRA^CF.INTER& LC.CT(SI)^LC.SC	0.69
AF.INTER&LC.ET.PR.SNP	0.45	CF.INTRA&LC.SNP	0.51

AF.INTER&LC.SC[QU]	0.45	CF.INTRA&LC.SNP(RT.AP)	0.51
AF.INTER^AF.INTER& LC.CL(SI)^GB.ET.PR	0.45	CF.INTRA^CF.INTER& LC.CL^GB.ET.PR	0.51
AF.INTER^AF.INTER& LC.ET.PR.CNP(SI)^GB.ET.PR	0.45	CF.INTRA^CF.INTRA& LC.CL(SI)^LC.SNP	0.51
AF.INTRA&LC.SNP+CNP	0.45	AF.INTER^CF.INTER& GB.ET.PR(SI)^GB.ET.PR	0.34
AF.INTRA^AF.INTER& LC.PRM(SI)^LC.CL	0.45	AF.INTRA^CF.INTER& LC.PRM(SI)^LC.CL	0.34
AF.INTRA^AF.INTER& LC.PRM(SI)^LC.ET.PR.CL	0.45	AF.INTRA^CF.INTER& LC.SNP(SI)^GB.ET.PR	0.34
AF.INTRA^AF.INTER& LC.SNP(*IT*)^LC.CL	0.45	AF.INTRA^CF.INTRA^CF.INTER& LC.PRM(SI)^LC.PM(SI)^GB.ET.PR	0.34
AF.INTRA^AF.INTER& LC.SNP(*THAT*)^LC.SNP	0.45	AF.INTRA^CF.INTRA^CF.INTRA& LC.PRM(SI)^LC.CL(SI)^LC.CNP	0.34
AF.INTRA^AF.INTER& LC.SNP(*THIS*)^GB.ET.PR	0.45	AF.INTRA^CF.INTRA^CF.INTRA& LC.PRM(SI)^LC.CL(SI)^ LC.SNP(NR.AP)	0.34
AF.INTRA^AF.INTRA& LC.PRM(SI)^LC.CL	0.45	CF.INTER&LC.CNP	0.34
AF.INTRA^AF.INTRA& LC.SNP(*IT*)^LC.CNP	0.45	CF.INTER&OV	0.34
AF.INTRA^AF.INTRA& LC.SNP(*THAT*)^LC.CL	0.45	CF.INTRA&LC.SC	0.34
AF.INTRA^AF.INTRA^ AF.INTER&LC.PRM(SI)^ LC.PRM(SI)^GB.ET.PR	0.45	CF.INTRA&LC.SNP(NR.AP)	0.34
AF.INTRA^CF.INTRA^ AF.INTRA&LC.PRM(SI)^ LC.PM(SI)^LC.CL	0.45	CF.INTRA&LC.SNP+CNP	0.34
CF.INTRA^AF.INTRA& LC.PM(SI)^LC.CL	0.45	CF.INTRA^CF.INTER& LC.CT(SI)^LC.SNP	0.34
CF.INTRA^AF.INTRA^ AF.INTER&LC.CT(SI)^ LC.SNP(*IT*)^GB.ET.PR	0.45	CF.INTRA^CF.INTER& LC.CT(SI)^LC.TB	0.34
CF.INTRA^AF.INTRA^ AF.INTER&LC.CT(SI)^ LC.SNP(*IT*)^LC.SNP	0.45	CF.INTRA^CF.INTER& LC.PM(SI)^LC.CL	0.34
CF.INTRA^AF.INTRA^ AF.INTER&LC.CT(SI)^ LC.SNP(*THAT*)^LC.SC	0.45	CF.INTRA^CF.INTER& LC.PM(SI)^LC.CNP	0.34
CF.INTRA^AF.INTRA^ AF.INTER&LC.CT(SI)^ LC.SNP(*THESE*)^GB.ET.PR	0.45	CF.INTRA^CF.INTRA& LC.CL(SI)^LC.CT	0.34
CF.INTRA^AF.INTRA^ AF.INTER(SN2)&LC.CT(SI)^ LC.SNP^GB.ET.PR(SN2)	0.45	CF.INTRA^CF.INTRA& LC.PM(SI)^LC.CNP	0.34
CF.INTRA^CF.INTRA^ AF.INTRA&LC.CT(SI)^ LC.CL(SI)^LC.SNP	0.45	CF.INTRA^CF.INTRA^CF.INTER& LC.CL(SI)^LC.CL^GB.ET.PR	0.34
AF.INTER&LC.CT	0.22	CF.INTRA^CF.INTRA^CF.INTER& LC.CL(SI)^LC.CT(SI)^GB.ET.PR	0.34
AF.INTER&LC.ET.PR.CNP	0.22	CF.INTRA^CF.INTRA^CF.INTER& LC.CT(SI)^LC.CL^GB.ET.PR	0.34
AF.INTER&LC.ET.PR.SC	0.22	CF.INTRA^CF.INTRA^CF.INTER& LC.PM(SI)^LC.PM(SI)^GB.ET.PR	0.34

AF.INTER^AF.INTER& LC.CNP(SI)^LC.CL	0.22	AF.INTER^AF.INTRA^CF.INTER& LC.ET.PR.CNP(SI)^ LC.PRM(SI)^GB.ET.PR	0.17
AF.INTER^AF.INTER& LC.ET.PR.SC(SI)^GB.ET.PR	0.22	AF.INTER^AF.INTRA^CF.INTRA& LC.ET.PR.SNP(SI)^ LC.PRM(SI)^LC.CNP	0.17
AF.INTER^AF.INTER& LC.ET.PR.SC(SI)^ LC.ET.PR.SNP[NUM]	0.22	AF.INTER^AF.INTRA^CF.INTRA& LC.ET.PR.SNP(SI)^ LC.PRM(SI)^LC.SNP	0.17
AF.INTER^AF.INTER& LC.SNP(SI)^LC.CNP	0.22	AF.INTER^AF.INTRA^CF.INTRA& LC.SC(SI)^LC.PRM(SI)^ LC.RN.SG[QU]	0.17
AF.INTER^AF.INTER(SN2)& LC.CNP^GB.ET.PR(SN2)	0.22	AF.INTER^AF.INTRA^CF.INTRA& LC.SC(SI)^LC.SNP(SI)^LC.CL	0.17
AF.INTER^AF.INTRA& GB.ET.PR(SI)^LC.SNP	0.22	AF.INTER^CF.INTER& GB.ET.PR(SI)^LC.TB	0.17
AF.INTER^AF.INTRA& LC.CL(SI)^LC.PRM	0.22	AF.INTER^CF.INTER& GB.ET.PR(SI)+LC.IM(SI)^GB.ET.PR	0.17
AF.INTER^AF.INTRA& LC.CL(SI)^LC.SNP[QT.PN]	0.22	AF.INTER^CF.INTER& LC.CNP(SI)^GB.ET.PR	0.17
AF.INTER^AF.INTRA& LC.ET.PR.SNP(SI)^LC.CNP	0.22	AF.INTER^CF.INTER& LC.CNP(SI)^LC.FR	0.17
AF.INTER^AF.INTRA& LC.SC(SI)^LC.CNP	0.22	AF.INTER^CF.INTER& LC.ET.PR.SNP(SI)^GB.ET.PR	0.17
AF.INTER^AF.INTRA& LC.SC(SI)^LC.SNP	0.22	AF.INTER^CF.INTER& LC.SC(SI)^GB.ET.PR	0.17
AF.INTER^AF.INTRA^ AF.INTRA&GB.ET.PR(SI)^ LC.SNP(*IT*)^LC.SNP	0.22	AF.INTER^CF.INTER& LC.SC(SI)^LC.FR	0.17
AF.INTRA&GB.ET.PR	0.22	AF.INTER^CF.INTER& LC.SNP(SI)^LC.FG	0.17
AF.INTRA&LC.AJ	0.22	AF.INTER^CF.INTRA& GB.ET.PR(SI)^LC.CL	0.17
AF.INTRA&LC.PP	0.22	AF.INTER^CF.INTRA& GB.ET.PR(SI)^LC.CL(NR.AP)	0.17
AF.INTRA^AF.INTER& LC.PRM(SI)^LC.CNP	0.22	AF.INTER^CF.INTRA& GB.ET.PR(SI)^LC.CNP(NR.AP)	0.17
AF.INTRA^AF.INTER& LC.PRM(SI)^LC.ET.PR.SC[QU]	0.22	AF.INTER^CF.INTRA& GB.ET.PR(SI)^LC.SNP+CNP(NR.AP)	0.17
AF.INTRA^AF.INTER& LC.SNP^GB.ET.PR	0.22	AF.INTER^CF.INTRA& LC.CL[QU](SI)^GB.ET.PR[QU]	0.17
AF.INTRA^AF.INTER& LC.SNP(*IT*)^GB.ET.PR[QU]	0.22	AF.INTER^CF.INTRA& LC.CNP(SI)^LC.CNP	0.17
AF.INTRA^AF.INTER& LC.SNP(*IT*)^LC.SC	0.22	AF.INTER^CF.INTRA& LC.ET.PR.CL(SI)^LC.CL	0.17
AF.INTRA^AF.INTER& LC.SNP(*IT*)^LC.SC[QU]	0.22	AF.INTER^CF.INTRA& LC.ET.PR.SC(SI)^LC.RN.SG	0.17
AF.INTRA^AF.INTER& LC.SNP(*IT*)^LC.SNP	0.22	AF.INTER^CF.INTRA& LC.SC(SI)^LC.CNP	0.17
AF.INTRA^AF.INTER& LC.SNP(SI)^GB.ET.PR	0.22	AF.INTER^CF.INTRA& LC.SC(SI)^LC.SNP	0.17
AF.INTRA^AF.INTER& LC.SNP(SI)^LC.ET.PR.CL	0.22	AF.INTER^CF.INTRA& LC.SC(SI)^LC.SNP+CNP	0.17
AF.INTRA^AF.INTER& LC.SNP(*THAT*)^LC.AJ	0.22	AF.INTER^CF.INTRA& LC.SNP(SI)^LC.CL	0.17
AF.INTRA^AF.INTER&	0.22	AF.INTER^CF.INTRA^CF.INTER&	0.17

LC.SNP(*THAT*)^LC.CL		GB.ET.PR(SI)^LC.CT(SI)^GB.ET.PR	
AF.INTRA^AF.INTER&		AF.INTER^CF.INTRA^CF.INTER&	
LC.SNP(*THAT*)^		GB.ET.PR(SI)^	
LC.ET.PR.SC[QU]	0.22	LC.SNP(NR.AP)^GB.ET.PR	0.17
		AF.INTER^CF.INTRA^CF.INTER&	
AF.INTRA^AF.INTER&		LC.ET.PR.SC(SI)^	
LC.SNP(*THIS*)^LC.CL	0.22	LC.PM(SI)^GB.ET.PR	0.17
AF.INTRA^AF.INTER&		AF.INTER^CF.INTRA^CF.INTRA&	
LC.SNP(*THIS*)^LC.SC	0.22	LC.CNP(SI)^LC.CT(SI)^LC.SNP+CNP	0.17
AF.INTRA^AF.INTER(SN2)&		AF.INTER^CF.INTRA^CF.INTRA&	
LC.SNP^GB.ET.PR(SN2)	0.22	LC.SNP(SI)^LC.CT(SI)^LC.SNP+CNP	0.17
		AF.INTER^CF.INTRA^CF.INTRA^	
AF.INTRA^AF.INTER(SN2)&		CF.INTRA&LC.CL(SI)^	
LC.SNP^LC.ET.PR.SNP(SN2)	0.22	LC.CT(SI)^LC.SNP(*THESE*)^LC.SNP	0.17
		AF.INTRA^AF.INTRA^AF.INTRA^	
		AF.INTRA^CF.INTER&LC.CNP(SI)^	
AF.INTRA^AF.INTRA&		LC.PRM(SI)^LC.PRM(SI)^	
LC.CNP(SI)^LC.SNP	0.22	LC.PRM(SI)^GB.ET.PR	0.17
		AF.INTRA^AF.INTRA^AF.INTRA^	
AF.INTRA^AF.INTRA&		CF.INTER&LC.PRM(SI)^	
LC.SNP(*IT*)^LC.CL	0.22	LC.PRM(SI)^LC.SNP(SI)^GB.ET.PR	0.17
		AF.INTRA^AF.INTRA^CF.INTRA^	
AF.INTRA^AF.INTRA&		CF.INTRA&LC.PRM(SI)^	
LC.SNP(*THAT*)^LC.CNP+CL	0.22	LC.PRM(SI)^LC.CL(SI)^LC.SNP	0.17
AF.INTRA^AF.INTRA&		AF.INTRA^CF.INTER&	
LC.SNP(*THAT*)^LC.VR	0.22	LC.PRM(SI)^LC.CL[QU]	0.17
AF.INTRA^AF.INTRA&		AF.INTRA^CF.INTER&	
LC.SNP(*THIS*)^LC.CL	0.22	LC.PRM(SI)^LC.CNP	0.17
AF.INTRA^AF.INTRA^			
AF.INTER&LC.PRM(SI)^		AF.INTRA^CF.INTER&	
LC.SNP(*IT*)^LC.ET.PR.CNP	0.22	LC.PRM(SI)^LC.CT	0.17
AF.INTRA^AF.INTRA^			
AF.INTER&LC.PRM(SI)^		AF.INTRA^CF.INTER&	
LC.SNP(*THAT*)^GB.ET.PR	0.22	LC.PRM(SI)^LC.SC[QU]	0.17
AF.INTRA^AF.INTRA^			
AF.INTER&LC.PRM(SI)^		AF.INTRA^CF.INTER&	
LC.SNP(*THIS*)^LC.SC	0.22	LC.PRM(SI)^LC.SNP	0.17
AF.INTRA^AF.INTRA^			
AF.INTRA&LC.PRM(SI)^		AF.INTRA^CF.INTRA&	
LC.SNP(*IT*)^LC.SNP	0.22	LC.CL(SI)^LC.CL	0.17
AF.INTRA^AF.INTRA^			
CF.INTRA^AF.INTRA&			
LC.SNP(SI)^LC.PRM(SI)^		AF.INTRA^CF.INTRA&	
LC.PM(SI)^LC.SNP	0.22	LC.CNP(SI)^LC.CL	0.17
AF.INTRA^CF.INTRA^			
AF.INTER&LC.PRM(SI)^		AF.INTRA^CF.INTRA&	
LC.CT(SI)^GB.ET.PR	0.22	LC.PRM(SI)^LC.CL(NR.AP)	0.17
AF.INTRA^CF.INTRA^			
AF.INTER&LC.PRM(SI)^		AF.INTRA^CF.INTRA&	
LC.PM(SI)^LC.ET.PR.CL	0.22	LC.PRM(SI)^LC.CNP(NR.AP)	0.17
AF.INTRA^CF.INTRA^			
AF.INTRA&LC.PRM(SI)^		AF.INTRA^CF.INTRA&	
LC.CL(SI)^LC.CL	0.22	LC.PRM(SI)^LC.PP	0.17
AF.INTRA^CF.INTRA^			
AF.INTRA&LC.PRM(SI)^		AF.INTRA^CF.INTRA&	
LC.CL(SI)^LC.SNP	0.22	LC.SNP(SI)^LC.CNP	0.17

AF.INTRA^CF.INTRA^ AF.INTRA^AF.INTER& LC.PRM(SI)^LC.CL(SI)^ LC.SNP(SI)^GB.ET.PR	0.22	AF.INTRA^CF.INTRA& LC.SNP(SI)^LC.SNP	0.17
AF.INTRA^CF.INTRA^ AF.INTRA^AF.INTER& LC.PRM(SI)^LC.PM(SI)^ LC.SNP(*THIS*)^GB.ET.PR	0.22	AF.INTRA^CF.INTRA^AF.INTRA^ AF.INTER^CF.INTER&LC.PRM(SI)^ LC.CL(SI)^LC.SNP(*THIS*)(SI)^ LC.SNP(SI)^LC.CL	0.17
AF.INTRA^CF.INTRA^ AF.INTRA^AF.INTER& LC.SNP(SI)^LC.CL(SI)^ LC.SNP(*THAT*)^GB.ET.PR	0.22	AF.INTRA^CF.INTRA^CF.INTER& LC.PRM(SI)^LC.CL(SI)^GB.ET.PR	0.17
CF.INTRA^AF.INTER& LC.CL(SI)^LC.ET.PR.SC	0.22	AF.INTRA^CF.INTRA^CF.INTER& LC.PRM(SI)^LC.CL(SI)^LC.SC	0.17
CF.INTRA^AF.INTER& LC.CL(SI)^LC.ET.PR.SNP	0.22	AF.INTRA^CF.INTRA^CF.INTER& LC.SNP(SI)^LC.CL(SI)^GB.ET.PR	0.17
CF.INTRA^AF.INTER& LC.CT(SI)^LC.CNP	0.22	AF.INTRA^CF.INTRA^CF.INTRA& LC.PRM(SI)^LC.CL(SI)^LC.CL	0.17
CF.INTRA^AF.INTER& LC.CT(SI)^LC.ET.PR.CNP	0.22	AF.INTRA^CF.INTRA^CF.INTRA& LC.PRM(SI)^LC.CT(SI)^LC.CNP	0.17
CF.INTRA^AF.INTER& LC.CT(SI)^LC.ET.PR.SC	0.22	AF.INTRA^CF.INTRA^CF.INTRA& LC.PRM(SI)^LC.PM(SI)^LC.CL	0.17
CF.INTRA^AF.INTER& LC.CT(SI)^LC.ET.PR.TB	0.22	AF.INTRA^CF.INTRA^CF.INTRA& LC.PRM(SI)^LC.PM(SI)^LC.CNP	0.17
CF.INTRA^AF.INTER& LC.CT(SI)^LC.SC	0.22	AF.INTRA^CF.INTRA^CF.INTRA& LC.PRM(SI)^LC.SNP(NR.AP)(SI)^ LC.CL	0.17
CF.INTRA^AF.INTER& LC.PM(SI)^LC.CNP	0.22	AF.INTRA^CF.INTRA^CF.INTRA^ CF.INTER&LC.SNP(SI)^ LC.CT(SI)^LC.CL(SI)^GB.ET.PR	0.17
CF.INTRA^AF.INTRA& LC.CL(SI)^LC.AJ	0.22	CF.INTER&LC.CL[QU]	0.17
CF.INTRA^AF.INTRA& LC.CL(SI)^LC.SC	0.22	CF.INTER&LC.ET.PR.CL	0.17
CF.INTRA^AF.INTRA& LC.CT(SI)^LC.AJ	0.22	CF.INTER&LC.ET.PR.SC	0.17
CF.INTRA^AF.INTRA& LC.CT(SI)^LC.AJP	0.22	CF.INTER&LC.ET.PR.SC[QU]	0.17
CF.INTRA^AF.INTRA& LC.CT(SI)^LC.CL[QU]	0.22	CF.INTER&LC.SC[QU]	0.17
CF.INTRA^AF.INTRA& LC.CT(SI)^LC.SNP+CNP	0.22	CF.INTER&LC.SNP+LC.AJ	0.17
CF.INTRA^AF.INTRA& LC.PM(SI)^LC.CNP	0.22	CF.INTER&LC.SNP[NUM]	0.17
CF.INTRA^AF.INTRA& LC.PM(SI)^LC.SNP	0.22	CF.INTER&LC.SNP+CNP	0.17
CF.INTRA^AF.INTRA& LC.PM(SI)^LC.CL	0.22	CF.INTER&LC.SNP+CNP[QU]	0.17
CF.INTRA^AF.INTRA^ AF.INTER&LC.CL(SI)^ LC.SNP^LC.SNP	0.22	CF.INTRA&LC.CNP(RT.AP)	0.17
CF.INTRA^AF.INTRA^ AF.INTER&LC.CL(SI)^ LC.SNP(*THAT*)^GB.ET.PR	0.22	CF.INTRA&LC.PM	0.17
CF.INTRA^AF.INTRA^ AF.INTER&LC.CL(SI)^	0.22	CF.INTRA&LC.PP(NR.AP)	0.17

LC.SNP(*THESE*)^GB.ET.PR			
CF.INTRA^AF.INTRA^			
AF.INTER&LC.CL(SI)^			
LC.SNP(*THESE*)^LC.SNP	0.22	CF.INTRA&UNC(NA)	0.17
CF.INTRA^AF.INTRA^			
AF.INTER&LC.CL(SI)^		CF.INTRA^AF.INTRA^CF.INTRA&	
LC.SNP(*THIS*)^LC.CL	0.22	LC.CL(SI)^LC.SNP(SI)^LC.CL	0.17
CF.INTRA^AF.INTRA^			
AF.INTER&LC.CL(SI)^		CF.INTRA^CF.INTER&	
LC.SNP(*THIS*)^LC.CL[QU]	0.22	LC.CL(SI)^LC.CL	0.17
CF.INTRA^AF.INTRA^			
AF.INTER&LC.CL(SI)^		CF.INTRA^CF.INTER&	
LC.SNP(*THOSE*)^GB.ET.PR	0.22	LC.CL(SI)^LC.SNP	0.17
CF.INTRA^AF.INTRA^			
AF.INTER&LC.CT(SI)^			
LC.SNP(*IT*)^LC.CNP	0.22	CF.INTRA^CF.INTER&LC.CL(SI)^OV	0.17
CF.INTRA^AF.INTRA^			
AF.INTER&LC.CT(SI)^		CF.INTRA^CF.INTER&	
LC.SNP(*IT*)^LC.SC	0.22	LC.CNP^GB.ET.PR	0.17
CF.INTRA^AF.INTRA^			
AF.INTER&LC.CT(SI)^		CF.INTRA^CF.INTER&	
LC.SNP(*THAT*)^GB.ET.PR	0.22	LC.CT(SI)^GB.ET.PR[QU]	0.17
CF.INTRA^AF.INTRA^			
AF.INTER&LC.CT(SI)^		CF.INTRA^CF.INTER&	
LC.SNP(*THAT*)^LC.CL	0.22	LC.CT(SI)^LC.AJ	0.17
CF.INTRA^AF.INTRA^			
AF.INTER&LC.CT(SI)^		CF.INTRA^CF.INTER&	
LC.SNP(*THAT*)^LC.SNP	0.22	LC.CT(SI)^LC.CL[QU]	0.17
CF.INTRA^AF.INTRA^			
AF.INTER&LC.CT(SI)^		CF.INTRA^CF.INTER&	
LC.SNP(*THEM*)^LC.CNP	0.22	LC.CT(SI)^LC.ET.PR.SC	0.17
CF.INTRA^AF.INTRA^			
AF.INTER&LC.CT(SI)^		CF.INTRA^CF.INTER&	
LC.SNP(*THIS*)^LC.CNP	0.22	LC.CT(SI)^LC.ET.PR.TB	0.17
CF.INTRA^AF.INTRA^			
AF.INTER^AF.INTER(SN2)&			
LC.CT(SI)^LC.SNP(*THEY*)^		CF.INTRA^CF.INTER&	
LC.SNP^GB.ET.PR(SN2)	0.22	LC.CT(SI)^LC.FG	0.17
CF.INTRA^AF.INTRA^			
AF.INTER(SN2)&LC.PM(SI)^		CF.INTRA^CF.INTER&	
LC.SNP^LC.SC(SN2)	0.22	LC.CT(SI)^LC.IJ	0.17
CF.INTRA^AF.INTRA^			
AF.INTRA&LC.CL(SI)^		CF.INTRA^CF.INTER&	
LC.SNP(*THIS*)^LC.SNP	0.22	LC.CT(SI)^LC.SC[QU]	0.17
CF.INTRA^AF.INTRA^			
AF.INTRA&LC.CT(SI)^		CF.INTRA^CF.INTER&	
LC.SNP(*IT*)^LC.CNP	0.22	LC.PM(SI)^LC.ET.PR.CL	0.17
CF.INTRA^AF.INTRA^			
AF.INTRA&LC.CT(SI)^		CF.INTRA^CF.INTER&	
LC.SNP(*THAT*)^LC.CNP	0.22	LC.SC^GB.ET.PR	0.17
CF.INTRA^CF.INTRA^		CF.INTRA^CF.INTER^	
AF.INTER&LC.CT(SI)^		CF.INTER(SN2)&LC.CT(SI)^	
LC.CL(SI)^GB.ET.PR	0.22	LC.CNP^GB.ET.PR(SN2)	0.17
CF.INTRA^CF.INTRA^			
AF.INTER&LC.CT(SI)^		CF.INTRA^CF.INTER(SN2)&	
LC.CT(SI)^GB.ET.PR	0.22	LC.SNP^GB.ET.PR(SN2)	0.17

CF.INTRA^CF.INTRA^ AF.INTRA&LC.CT(SI)^ LC.CT(SI)^LC.CL	0.22	CF.INTRA^CF.INTRA& LC.CL(SI)^LC.CL(AP)	0.17
CF.INTRA^CF.INTRA^ AF.INTRA&AF.INTER& LC.CT(SI)^LC.CT(SI)^ LC.SNP(*THESE*)^LC.SC[QU]	0.22	CF.INTRA^CF.INTRA& LC.CT(SI)^GB.ET.PR[QU]	0.17
CF.INTRA^CF.INTRA^ CF.INTRA^AF.INTER& LC.CT(SI)^LC.CL(SI)^ LC.CNP(SI)^GB.ET.PR	0.22	CF.INTRA^CF.INTRA& LC.CT(SI)^LC.CL(NR.AP)	0.17
		CF.INTRA^CF.INTRA& LC.CT(SI)^LC.CNP(NR.AP)	0.17
		CF.INTRA^CF.INTRA& LC.CT(SI)^LC.PP	0.17
		CF.INTRA^CF.INTRA& LC.CT(SI)^LC.SNP	0.17
		CF.INTRA^CF.INTRA& LC.CT(SI)^LC.SNP+CNP(NR.AP)	0.17
		CF.INTRA^CF.INTRA& LC.PM(SI)^LC.CL(AP)	0.17
		CF.INTRA^CF.INTRA& LC.PM(SI)^LC.CL(NR.AP)	0.17
		CF.INTRA^CF.INTRA& LC.PM(SI)^LC.CNP(NR.AP)	0.17
		CF.INTRA^CF.INTRA& LC.PM(SI)^LC.SNP(NR.AP)	0.17
		CF.INTRA^CF.INTRA^CF.INTER& LC.CL(SI)^LC.CNP(SI)^GB.ET.PR	0.17
		CF.INTRA^CF.INTRA^CF.INTER& LC.CL(SI)^LC.PM(SI)^LC.ET.PR.SC	0.17
		CF.INTRA^CF.INTRA^CF.INTER& LC.CT(SI)^LC.CT(SI)^LC.CL	0.17
		CF.INTRA^CF.INTRA^CF.INTER& LC.PM(SI)^LC.SNP^LC.CL	0.17
		CF.INTRA^CF.INTRA^CF.INTRA& LC.CL(SI)^LC.PM(SI)^LC.SNP	0.17
		CF.INTRA^CF.INTRA^CF.INTRA& LC.CT(SI)^LC.PM(SI)^LC.CNP	0.17

Exophora	%	Combined	%
EXO^NA	33.62	AF.INTER^CF.INTER&GB.ET.PR	7.69
CF.INTRA^EXO&LC.CT(SI)^NA	18.10	AF.INTER^EXO&GB.ET.PR^NA	4.62
AF.INTRA^EXO& LC.PRM(SI)^NA	12.36	CF.INTRA^AF.INTER^CF.INTER& LC.CT(SI)^GB.ET.PR	4.62
CF.INTRA^EXO&LC.CL(SI)^NA	9.77	CF.INTRA^AF.INTER^CF.INTER& LC.PM(SI)^GB.ET.PR	4.62
CF.INTRA^EXO&LC.PM(SI)^NA	6.32	AF.INTER^CF.INTER& LC.SNP^LC.SC	3.08
AF.INTRA^CF.INTRA^EXO& LC.PRM(SI)^LC.CL(SI)^NA	2.59	AF.INTRA^EXO&LC.CL^NA	3.08
AF.INTRA^EXO& LC.SNP(SI)^NA	2.30	AF.INTRA^EXO&LC.SNP^NA	3.08
AF.INTRA^AF.INTRA^EXO& LC.PRM(SI)^LC.PRM(SI)^NA	1.72	AF.INTER^AF.INTER^EXO& LC.CNP(SI)^GB.ET.PR^NA	1.54
AF.INTRA^CF.INTRA^EXO&	1.72	AF.INTER^AF.INTER^CF.INTRA&	1.54

LC.PRM(SI)^LC.PM(SI)^NA		LC.SNP(SI)&LC.CL&LC.CL	
AF.INTRA^CF.INTRA^EXO&		AF.INTER^AF.INTRA^CF.INTER&	
LC.PRM(SI)^LC.CT(SI)^NA	1.15	LC.SC[QU](SI)^LC.SC[QU]	1.54
CF.INTRA^CF.INTRA^EXO&			
LC.CT(SI)^LC.CL(SI)^NA	0.86	AF.INTER^CF.INTER&LC.CL^LC.SC	1.54
AF.INTER^CF.INTRA^EXO&		AF.INTER^CF.INTER&	
GB.ET.PR(SI)^LC.CT(SI)^NA	0.57	LC.CNP^GB.ET.PR	1.54
AF.INTER^EXO&			
GB.ET.PR(SI)^NA	0.57	AF.INTER^CF.INTER&LC.SC^LC.SC	1.54
		AF.INTER^CF.INTRA&	
AF.INTER^EXO&LC.CL(SI)^NA	0.57	LC.ET.PR.SC[QU]^LC.AJ	1.54
AF.INTER^EXO&			
LC.ET.PR.CNP(SI)^NA	0.57	AF.INTER^CF.INTRA&LC.SC^LC.CL	1.54
		AF.INTER^CF.INTRA^AF.INTER^	
AF.INTRA^CF.INTRA^EXO&		CF.INTRA^CF.INTER&	
LC.PRM(SI)^		LC.SC(SI)^LC.CL(SI)^LC.CL^	
LC.CNP(NR.AP)(SI)^NA	0.57	LC.SC^LC.SC	1.54
		AF.INTER^CF.INTRA^AF.INTER^	
AF.INTRA^EXO&		EXO&LC.ET.PR.CNP(SI)^	
LC.CNP(SI)^NA	0.57	LC.PM(SI)^LC.CNP^NA	1.54
		AF.INTER^CF.INTRA^CF.INTRA^	
AF.INTER^CF.INTRA^EXO&		AF.INTER^CF.INTER&GB.ET.PR(SI)^	
GB.ET.PR(SI)^LC.CL(SI)^NA	0.29	LC.CT(SI)^LC.CL(SI)^LC.SC^LC.SC	1.54
AF.INTER^EXO&			
LC.CNP(SI)^NA	0.29	AF.INTER^EXO&LC.SC^NA	1.54
		AF.INTRA^AF.INTER^CF.INTER&	
AF.INTER^EXO&LC.CT(SI)^NA	0.29	LC.PRM(SI)^GB.ET.PR	1.54
AF.INTER^EXO&		AF.INTRA^AF.INTER^CF.INTRA&	
LC.ET.PR.CL(SI)^NA	0.29	LC.PRM(SI)^GB.ET.PR^LC.CNP	1.54
AF.INTER^EXO&		AF.INTRA^AF.INTER^CF.INTRA&	
LC.ET.PR.SNP(SI)^NA	0.29	LC.SNP(*THIS*)^LC.SC^GB.ET.PR	1.54
		AF.INTRA^AF.INTRA^AF.INTER^	
		CF.INTER&LC.PRM(SI)^	
AF.INTER^EXO&LC.SC(SI)^NA	0.29	LC.SNP(*IT*)^GB.ET.PR^LC.SC	1.54
		AF.INTRA^AF.INTRA^CF.INTRA&	
AF.INTER^EXO&		LC.PRM(SI)^LC.SNP(*THAT*)^	
LC.SNP(SI)^NA	0.29	LC.CL(NR.AP)	1.54
AF.INTRA^AF.INTRA^			
AF.INTRA^EXO&			
LC.PRM(SI)^LC.PRM(SI)^		AF.INTRA^CF.INTER&	
LC.PRM(SI)^NA	0.29	LC.SNP^GB.ET.PR	1.54
AF.INTRA^AF.INTRA^			
AF.INTRA^EXO&			
LC.PRM(SI)^LC.PRM(SI)^		AF.INTRA^CF.INTER&	
LC.SNP(SI)^NA	0.29	LC.SNP^LC.SNP	1.54
AF.INTRA^AF.INTRA^			
CF.INTRA^EXO&			
LC.PRM(SI)^LC.PRM(SI)^		AF.INTRA^CF.INTRA&	
LC.PP(NR.AP)(SI)^NA	0.29	LC.CNP^LC.CNP	1.54
AF.INTRA^CF.INTRA^			
AF.INTRA^EXO&			
LC.PRM(SI)^LC.PM(SI)^		AF.INTRA^CF.INTRA&	
LC.SNP(SI)^NA	0.29	LC.SNP^LC.CNP	1.54
AF.INTRA^CF.INTRA^			
CF.INTRA^EXO&		AF.INTRA^CF.INTRA&	
LC.PRM(SI)^LC.CL(SI)^	0.29	LC.SNP(*THIS*)^LC.SNP	1.54

LC.PM(SI)^NA			
		AF.INTRA^CF.INTRA^AF.INTER^ CF.INTER&LC.PRM(SI)^	
AF.INTRA^EXO&LC.CL(SI)^NA	0.29	LC.PM(SI)^GB.ET.PR	1.54
		AF.INTRA^CF.INTRA^ AF.INTER(SN2)&LC.PRM(SI)^	
CF.INTER^EXO&LC.CL(SI)^NA	0.29	LC.PM^GB.ET.PR(SN2)	1.54
		AF.INTRA^CF.INTRA^AF.INTRA^ CF.INTRA&LC.SNP(SI)^	
CF.INTRA^CF.INTER^EXO& LC.CT(SI)^GB.ET.PR(SI)^NA	0.29	LC.CL(SI)^LC.CNP^LC.SNP	1.54
CF.INTRA^CF.INTRA^ EXO&LC.CL(SI)^ LC.CNP(NR.AP)(SI)^NA	0.29	AF.INTRA^EXO&LC.SNP(*IT*)^NA	1.54
CF.INTRA^CF.INTRA^EXO& LC.CT(SI)^LC.CT(SI)^NA	0.29	CF.INTRA^AF.INTER^CF.INTER& LC.CL(SI)^GB.ET.PR	1.54
CF.INTRA^CF.INTRA^EXO& LC.CT(SI)^LC.PM(SI)^NA	0.29	CF.INTRA^AF.INTER^CF.INTER& LC.CT(SI)^GB.ET.PR^LC.SNP	1.54
CF.INTRA^CF.INTRA^EXO& LC.PM(SI)^LC.PM(SI)^NA	0.29	CF.INTRA^AF.INTER^CF.INTRA& LC.CL(SI)^GB.ET.PR^LC.CL	1.54
CF.INTRA^EXO& LC.CNP(SI)^NA	0.29	CF.INTRA^AF.INTER^CF.INTRA& LC.CL(SI)^GB.ET.PR^LC.SNP	1.54
CF.INTRA^EXO& LC.SNP(SI)^NA	0.29	CF.INTRA^AF.INTER^CF.INTRA& LC.CT(SI)^GB.ET.PR^LC.CL	1.54
		CF.INTRA^AF.INTER^CF.INTRA& LC.CT(SI)^GB.ET.PR^LC.SNP	1.54
		CF.INTRA^CF.INTRA^AF.INTER^ CF.INTRA&LC.CT(SI)^ LC.PM(SI)^LC.CNP^LC.CL	1.54
		CF.INTRA^AF.INTER^EXO& LC.CL(SI)^GB.ET.PR^NA	1.54
		CF.INTRA^AF.INTER^EXO& LC.CL(SI)^LC.ET.PR.SC^NA	1.54
		CF.INTRA^AF.INTER^EXO& LC.CT(SI)^GB.ET.PR^NA	1.54
		CF.INTRA^AF.INTER(SN2)& LC.CT^GB.ET.PR(SN2)	1.54
		CF.INTRA^AF.INTRA^AF.INTER^ CF.INTER&LC.CT(SI)^ LC.SNP^GB.ET.PR	1.54
		CF.INTRA^AF.INTRA^AF.INTER^ CF.INTER&LC.CT(SI)^LC.SNP(*IT*)^ LC.CNP^GB.ET.PR	1.54
		CF.INTRA^AF.INTRA^CF.INTER& LC.CT(SI)^LC.SNP^LC.ET.PR.SC	1.54
		CF.INTRA^AF.INTRA^EXO& LC.CT(SI)^LC.SNP^NA	1.54
		CF.INTRA^CF.INTRA^AF.INTER& LC.CT(SI)^LC.CNP^GB.ET.PR	1.54
		CF.INTRA^CF.INTRA^AF.INTER& LC.CT(SI)^LC.SNP(*THAT*)^GB.ET.PR	1.54
		CF.INTRA^CF.INTRA^EXO& LC.PM(SI)^LC.SNP^NA	1.54
		CF.INTRA^EXO&LC.CL^NA	1.54

APPENDIX 14 LEMMA DISTRIBUTION OF ENCAPSULATING RELATIONS AND ANTECEDENTS (COMPLETE LISTS)

14.1 *Main encapsulating relations*

Anaphora	%	Cataphora	%	Exophora	%	Combined	%
Endorsement	100.00	Quest	100.00	Correction	100.00	Opposite	30.00
Leave	75.00	Proviso	85.71	Prejudice	63.64	Misfortune	28.57
Testimony	71.43	Chance	77.50	Detail	50.00	Terror	25.00
Phenomenon	69.44	Capacity	73.91	Evidence	50.00	Thing	15.00
Impetus	66.67	Sense	72.73	Recommendation	50.00	Dimension	12.50
Contradiction	62.50	Failure	69.57	Crime	45.00	Motivation	12.50
Anger	60.00	Surprise	64.00	Venture	41.38	Assessment	11.54
Terror	50.00	Suspicion	64.00	System	41.18	Point	10.00
Triumph	50.00	Motivation	62.50	Work	40.74	Part	9.38
Finding	46.43	Recollection	60.00	Answer	40.00	Area	7.50
Area	45.00	Warning	58.62	Scandal	38.46	Detail	7.50
Part	43.75	Irony	57.14	Contradiction	37.50	Way	7.50
Myth	43.33	Endeavour	55.56	Project	37.50	Challenge	5.88
Foreboding	42.86	Example	55.00	Experience	35.00	System	5.88
Venture	41.38	Objective	55.00	Problem	35.00	Phenomenon	5.56
Application	40.00	Way	55.00	Word	35.00	Crime	5.00
Opposite	40.00	Application	50.00	Dimension	33.33	Example	5.00
Characteristic	37.93	Myth	46.67	Facet	28.57	Experience	5.00
Practice	37.50	Point	45.00	Foreboding	28.57	Practice	5.00
Project	37.50	Problem	45.00	Irony	28.57	Word	5.00
Word	37.50	Time	45.00	Practice	27.50	Surprise	4.00
Work	37.04	Vision	44.83	Philosophy	27.27	Suspicion	4.00
Philosophy	36.36	Facet	42.86	Assessment	26.92	Finding	3.57
Challenge	35.29	Assessment	42.31	Finding	25.00	Characteristic	3.45
Warning	34.48	Anger	40.00	Joke	25.00	Vision	3.45
Example	32.50	Joke	40.00	Time	25.00	Myth	3.33
Joke	32.50	Thing	40.00	Vision	24.14	Answer	2.50
Scandal	30.77	Challenge	38.24	Sense	22.73	Chance	2.50
Failure	30.43	Characteristic	37.93	Characteristic	20.69	Joke	2.50
Thing	30.00	Experience	37.50	Challenge	20.59	Objective	2.50
System	29.41	Triumph	37.50	Recollection	20.00	Problem	2.50
Facet	28.57	Philosophy	36.36	Part	18.75	Project	2.50
Misfortune	28.57	Answer	35.00	Area	17.50	Time	2.50
Vision	27.59	Impetus	33.33	Objective	17.50	Anger	0.00
Point	27.50	Detail	32.50	Point	17.50	Application	0.00
Time	27.50	Scandal	30.77	Way	15.00	Capacity	0.00
Dimension	25.00	Area	30.00	Misfortune	14.29	Contradiction	0.00
Motivation	25.00	Evidence	30.00	Thing	12.50	Correction	0.00
Objective	25.00	Opposite	30.00	Triumph	12.50	Endeavour	0.00
Answer	22.50	Practice	30.00	Suspicion	12.00	Endorsement	0.00
Experience	22.50	Dimension	29.17	Endeavour	11.11	Evidence	0.00
Way	22.50	Foreboding	28.57	Phenomenon	11.11	Facet	0.00
Endeavour	22.22	Misfortune	28.57	Application	10.00	Failure	0.00
Recommendation	21.88	Testimony	28.57	Capacity	8.70	Foreboding	0.00
Crime	20.00	Part	28.13	Surprise	8.00	Impetus	0.00
Evidence	20.00	Recommendation	28.13	Chance	7.50	Irony	0.00
Recollection	20.00	Crime	27.50	Example	7.50	Leave	0.00
Surprise	20.00	Finding	25.00	Warning	6.90	Philosophy	0.00

Suspicion	20.00	Leave	25.00	Myth	6.67	Prejudice	0.00
Assessment	19.23	Terror	25.00	Anger	0.00	Proviso	0.00
Prejudice	18.18	System	23.53	Endorsement	0.00	Quest	0.00
Problem	17.50	Project	22.50	Failure	0.00	Recollection	0.00
Capacity	17.39	Word	22.50	Impetus	0.00	Recommendation	0.00
Irony	14.29	Work	22.22	Leave	0.00	Scandal	0.00
Proviso	14.29	Prejudice	18.18	Motivation	0.00	Sense	0.00
Chance	10.00	Venture	17.24	Opposite	0.00	Testimony	0.00
Detail	10.00	Phenomenon	11.11	Proviso	0.00	Triumph	0.00
Sense	4.55	Contradiction	0.00	Quest	0.00	Venture	0.00
Correction	0.00	Correction	0.00	Terror	0.00	Warning	0.00
Quest	0.00	Endorsement	0.00	Testimony	0.00	Work	0.00

14.2 *Specific encapsulating relations (top ten)*

CF.INTRA	%	AF.INTER	%	CF.INTER	%	AF.INTRA	%
Quest	80.00	Contradiction	50.00	Triumph	37.50	Leave	75.00
Chance	77.50	Endorsement	50.00	Assessment	34.62	Impetus	66.67
Capacity	69.57	Phenomenon	44.44	Myth	33.33	Testimony	42.86
Failure	69.57	Foreboding	42.86	Joke	30.00	Opposite	30.00
Recollection	60.00	Anger	40.00	Opposite	30.00	Facet	28.57
Proviso	57.14	Terror	37.50	Way	30.00	Failure	26.09
Suspicion	56.00	Project	32.50	Facet	28.57	Endorsement	25.00
Surprise	52.00	Finding	32.14	Foreboding	28.57	Triumph	25.00
Endeavour	44.44	Myth	30.00	Irony	28.57	Phenomenon	22.22
Time	42.50	Application	27.50	Proviso	28.57	Anger	20.00
Warning	41.38	Philosophy	27.27	Testimony	28.57	Area	20.00
Sense	40.91	Work	25.93	Sense	27.27	Example	17.50
Anger	40.00	Motivation	25.00	Detail	25.00	Practice	17.50
Application	40.00	Triumph	25.00	Dimension	25.00	Misfortune	14.29
Motivation	37.50	Warning	24.14	Example	25.00	Characteristic	13.79
Objective	35.00	Point	22.50	Terror	25.00	Capacity	13.04
Irony	28.57	Time	22.50	Practice	22.50	Answer	12.50
Vision	27.59	Characteristic	20.69	System	20.59	Application	12.50
Problem	27.50	Venture	20.69	Evidence	20.00	Challenge	11.76
Thing	27.50	System	20.59	Experience	20.00	Endeavour	11.11
Challenge	26.47	Experience	20.00	Point	20.00	Work	11.11
Example	25.00	Objective	20.00	Quest	20.00	Venture	10.34
Leave	25.00	Suspicion	20.00	Prejudice	18.18	Crime	10.00
Way	25.00	Thing	20.00	Finding	17.86	Myth	10.00
Philosophy	22.73	Part	18.75	Answer	17.50	Part	9.38
Area	22.50	Challenge	17.65	Objective	17.50	Surprise	8.00
Characteristic	20.69	Vision	17.24	Characteristic	17.24	Scandal	7.69
Venture	20.69	Dimension	16.67	Warning	17.24	Joke	7.50
Point	20.00	Recommendation	15.63	Impetus	16.67	Word	7.50
Answer	17.50	Scandal	15.38	Recommendation	15.63	Finding	7.14
Crime	17.50	Area	15.00	Scandal	15.38	Vision	6.90
Experience	17.50	Evidence	15.00	Project	15.00	Recommendation	6.25
Impetus	16.67	Practice	15.00	Word	15.00	System	5.88
Scandal	15.38	Word	15.00	Misfortune	14.29	Evidence	5.00
Work	14.81	Irony	14.29	Vision	13.79	Project	5.00
Facet	14.29	Misfortune	14.29	Philosophy	13.64	Way	5.00
Misfortune	14.29	Proviso	14.29	Motivation	12.50	Dimension	4.17
Part	12.50	Testimony	14.29	Part	12.50	Assessment	3.85
Recommendation	12.50	Way	12.50	Thing	12.50	Warning	3.45

Evidence	10.00	Assessment	11.54	Surprise	12.00	Detail	2.50
Joke	10.00	Endeavour	11.11	Challenge	11.76	Experience	2.50
Myth	10.00	Chance	10.00	Endeavour	11.11	Objective	2.50
Assessment	7.69	Crime	10.00	Application	10.00	Problem	2.50
Detail	7.50	Joke	10.00	Crime	10.00	Time	2.50
Practice	7.50	Opposite	10.00	Problem	10.00	Chance	0.00
Project	7.50	Prejudice	9.09	Suspicion	8.00	Contradiction	0.00
Word	7.50	Answer	7.50	Area	7.50	Correction	0.00
Finding	7.14	Detail	7.50	Work	7.41	Foreboding	0.00
Phenomenon	5.56	Example	7.50	Venture	6.90	Irony	0.00
Dimension	4.17	Sense	4.55	Phenomenon	5.56	Motivation	0.00
Contradiction	0.00	Capacity	4.35	Capacity	4.35	Philosophy	0.00
Correction	0.00	Problem	2.50	Time	2.50	Point	0.00
Endorsement	0.00	Correction	0.00	Anger	0.00	Prejudice	0.00
Foreboding	0.00	Facet	0.00	Chance	0.00	Proviso	0.00
Opposite	0.00	Failure	0.00	Contradiction	0.00	Quest	0.00
Prejudice	0.00	Impetus	0.00	Correction	0.00	Recollection	0.00
System	0.00	Leave	0.00	Endorsement	0.00	Sense	0.00
Terror	0.00	Quest	0.00	Failure	0.00	Suspicion	0.00
Testimony	0.00	Recollection	0.00	Leave	0.00	Terror	0.00
Triumph	0.00	Surprise	0.00	Recollection	0.00	Thing	0.00

AF.INTRA& AF.INTER	%	AF.INTER& CF.INTER	%	AF.INTRA& AF.INTRA	%	AF.INTER& CF.INTRA	%
Recollection	20.00	Terror	12.50	Endorsement	25.00	Motivation	12.50
Testimony	14.29	Opposite	10.00	Failure	4.35	Area	5.00
Contradiction	12.50	Dimension	8.33	Surprise	4.00	Thing	5.00
Part	12.50	Part	6.25	Part	3.13	Assessment	3.85
Word	12.50	System	5.88	Area	2.50	Vision	3.45
Joke	10.00	Point	5.00	Example	2.50	Phenomenon	2.78
Problem	10.00	Practice	5.00	Joke	2.50	Objective	2.50
Prejudice	9.09	Assessment	3.85	Point	2.50	Point	2.50
Surprise	8.00	Finding	3.57	Practice	2.50	Way	2.50
Scandal	7.69	Myth	3.33	Problem	2.50	Anger	0.00
Area	7.50	Challenge	2.94	Thing	2.50	Answer	0.00
Thing	7.50	Phenomenon	2.78	Word	2.50	Application	0.00
Finding	7.14	Chance	2.50	Anger	0.00	Capacity	0.00
Warning	6.90	Experience	2.50	Answer	0.00	Challenge	0.00
Example	5.00	Problem	2.50	Application	0.00	Chance	0.00
Way	5.00	Way	2.50	Assessment	0.00	Characteristic	0.00
Philosophy	4.55	Anger	0.00	Capacity	0.00	Contradiction	0.00
Assessment	3.85	Answer	0.00	Challenge	0.00	Correction	0.00
Characteristic	3.45	Application	0.00	Chance	0.00	Crime	0.00
Vision	3.45	Area	0.00	Characteristic	0.00	Detail	0.00
Challenge	2.94	Capacity	0.00	Contradiction	0.00	Dimension	0.00
Phenomenon	2.78	Characteristic	0.00	Correction	0.00	Endeavour	0.00
Answer	2.50	Contradiction	0.00	Crime	0.00	Endorsement	0.00
Objective	2.50	Correction	0.00	Detail	0.00	Evidence	0.00
Point	2.50	Crime	0.00	Dimension	0.00	Example	0.00
Practice	2.50	Detail	0.00	Endeavour	0.00	Experience	0.00
Time	2.50	Endeavour	0.00	Evidence	0.00	Facet	0.00
Anger	0.00	Endorsement	0.00	Experience	0.00	Failure	0.00
Application	0.00	Evidence	0.00	Facet	0.00	Finding	0.00
Capacity	0.00	Example	0.00	Finding	0.00	Foreboding	0.00
Chance	0.00	Facet	0.00	Foreboding	0.00	Impetus	0.00

Correction	0.00	Failure	0.00	Impetus	0.00	Irony	0.00
Crime	0.00	Foreboding	0.00	Irony	0.00	Joke	0.00
Detail	0.00	Impetus	0.00	Leave	0.00	Leave	0.00
Dimension	0.00	Irony	0.00	Misfortune	0.00	Misfortune	0.00
Endeavour	0.00	Joke	0.00	Motivation	0.00	Myth	0.00
Endorsement	0.00	Leave	0.00	Myth	0.00	Opposite	0.00
Evidence	0.00	Misfortune	0.00	Objective	0.00	Part	0.00
Experience	0.00	Motivation	0.00	Opposite	0.00	Philosophy	0.00
Facet	0.00	Objective	0.00	Phenomenon	0.00	Practice	0.00
Failure	0.00	Philosophy	0.00	Philosophy	0.00	Prejudice	0.00
Foreboding	0.00	Prejudice	0.00	Prejudice	0.00	Problem	0.00
Impetus	0.00	Project	0.00	Project	0.00	Project	0.00
Irony	0.00	Proviso	0.00	Proviso	0.00	Proviso	0.00
Leave	0.00	Quest	0.00	Quest	0.00	Quest	0.00
Misfortune	0.00	Recollection	0.00	Recollection	0.00	Recollection	0.00
Motivation	0.00	Recommendation	0.00	Recommendation	0.00	Recommendation	0.00
Myth	0.00	Scandal	0.00	Scandal	0.00	Scandal	0.00
Opposite	0.00	Sense	0.00	Sense	0.00	Sense	0.00
Project	0.00	Surprise	0.00	Suspicion	0.00	Surprise	0.00
Proviso	0.00	Suspicion	0.00	System	0.00	Suspicion	0.00
Quest	0.00	Testimony	0.00	Terror	0.00	System	0.00
Recommendation	0.00	Thing	0.00	Testimony	0.00	Terror	0.00
Sense	0.00	Time	0.00	Time	0.00	Testimony	0.00
Suspicion	0.00	Triumph	0.00	Triumph	0.00	Time	0.00
System	0.00	Venture	0.00	Venture	0.00	Triumph	0.00
Terror	0.00	Vision	0.00	Vision	0.00	Venture	0.00
Triumph	0.00	Warning	0.00	Warning	0.00	Warning	0.00
Venture	0.00	Word	0.00	Way	0.00	Word	0.00
Work	0.00	Work	0.00	Work	0.00	Work	0.00

CF.INTRA&CF.INTER (%)			
Problem	7.50	Impetus	0.00
Point	5.00	Irony	0.00
Vision	3.45	Joke	0.00
Myth	3.33	Leave	0.00
Part	3.13	Misfortune	0.00
System	2.94	Motivation	0.00
Example	2.50	Opposite	0.00
Objective	2.50	Phenomenon	0.00
Anger	0.00	Philosophy	0.00
Answer	0.00	Practice	0.00
Application	0.00	Prejudice	0.00
Area	0.00	Project	0.00
Assessment	0.00	Proviso	0.00
Capacity	0.00	Quest	0.00
Challenge	0.00	Recollection	0.00
Chance	0.00	Recommendation	0.00
Characteristic	0.00	Scandal	0.00
Contradiction	0.00	Sense	0.00
Correction	0.00	Surprise	0.00
Crime	0.00	Suspicion	0.00
Detail	0.00	Terror	0.00
Dimension	0.00	Testimony	0.00
Endeavour	0.00	Thing	0.00
Endorsement	0.00	Time	0.00

Evidence	0.00	Triumph	0.00
Experience	0.00	Venture	0.00
Facet	0.00	Warning	0.00
Failure	0.00	Way	0.00
Finding	0.00	Word	0.00
Foreboding	0.00	Work	0.00

14.3 *Antecedents (top ten)*

GB.ET.PR	%	LC.CL	%	LC.SNP	%	LC.CNP	%
System	70.83	Leave	100.00	Opposite	50.00	Testimony	50.00
Detail	70.00	Proviso	85.71	Facet	40.00	Work	43.75
Terror	70.00	Capacity	66.67	Area	38.46	Anger	40.00
Project	69.23	Endeavour	62.50	Word	28.13	Endorsement	40.00
Myth	67.74	Failure	62.50	Failure	25.00	Impetus	33.33
Experience	65.38	Recollection	60.00	Philosophy	22.22	Phenomenon	31.43
Irony	60.00	Objective	58.33	Phenomenon	20.00	Motivation	30.00
Point	52.38	Recommendation	56.25	Prejudice	20.00	Scandal	22.22
Evidence	50.00	Suspicion	52.17	Challenge	18.75	Characteristic	20.00
Practice	48.48	Surprise	51.85	Example	18.18	Facet	20.00
Characteristic	48.00	Time	45.16	Crime	17.39	Contradiction	16.67
Assessment	45.83	Anger	40.00	Impetus	16.67	Example	15.91
Philosophy	44.44	Irony	40.00	Characteristic	16.00	Misfortune	14.29
Part	44.12	Venture	40.00	Part	14.71	Triumph	14.29
Work	43.75	Chance	39.47	Misfortune	14.29	Crime	13.04
Triumph	42.86	Warning	37.93	System	12.50	Assessment	12.50
Way	42.11	Finding	33.33	Practice	12.12	Application	11.11
Facet	40.00	Impetus	33.33	Way	10.53	Area	10.26
Foreboding	40.00	Thing	31.82	Evidence	10.00	Evidence	10.00
Prejudice	40.00	Crime	30.43	Motivation	10.00	Terror	10.00
Joke	37.84	Evidence	30.00	Joke	8.11	Venture	10.00
Testimony	37.50	Motivation	30.00	Answer	8.00	Problem	8.82
Dimension	36.84	Problem	29.41	Project	7.69	Suspicion	8.70
Challenge	34.38	Misfortune	28.57	Warning	6.90	Failure	8.33
Finding	33.33	Point	28.57	Myth	6.45	Answer	8.00
Scandal	33.33	Triumph	28.57	Time	6.45	Vision	8.00
Answer	32.00	Challenge	25.00	Recommendation	6.25	Experience	7.69
Vision	32.00	Philosophy	22.22	Work	6.25	Project	7.69
Recommendation	31.25	Scandal	22.22	Objective	5.56	Surprise	7.41
Venture	30.00	Joke	21.62	Sense	5.56	Thing	6.82
Thing	29.55	Dimension	21.05	Dimension	5.26	Myth	6.45
Problem	29.41	Way	21.05	Detail	5.00	Part	5.88
Misfortune	28.57	Endorsement	20.00	Venture	5.00	Objective	5.56
Phenomenon	28.57	Quest	20.00	Capacity	4.76	Philosophy	5.56
Application	25.00	Application	19.44	Thing	4.55	Dimension	5.26
Endeavour	25.00	Experience	19.23	Assessment	4.17	Detail	5.00
Word	21.88	Contradiction	16.67	Vision	4.00	Capacity	4.76
Crime	21.74	Sense	16.67	Problem	2.94	Finding	4.17
Opposite	21.43	Part	14.71	Application	2.78	System	4.17
Warning	20.69	Example	13.64	Chance	2.63	Warning	3.45
Area	20.51	Assessment	12.50	Point	2.38	Challenge	3.13
Example	20.45	Practice	12.12	Anger	0.00	Practice	3.03
Anger	20.00	Vision	12.00	Contradiction	0.00	Chance	0.00
Endorsement	20.00	Area	10.26	Correction	0.00	Correction	0.00
Objective	19.44	Detail	10.00	Endeavour	0.00	Endeavour	0.00

Time	19.35	Phenomenon	8.57	Endorsement	0.00	Foreboding	0.00
Suspicion	17.39	Answer	8.00	Experience	0.00	Irony	0.00
Contradiction	16.67	Project	7.69	Finding	0.00	Joke	0.00
Impetus	16.67	Myth	6.45	Foreboding	0.00	Leave	0.00
Proviso	14.29	Word	6.25	Irony	0.00	Opposite	0.00
Sense	11.11	Characteristic	0.00	Leave	0.00	Point	0.00
Motivation	10.00	Correction	0.00	Proviso	0.00	Prejudice	0.00
Surprise	3.70	Facet	0.00	Quest	0.00	Proviso	0.00
Chance	2.63	Foreboding	0.00	Recollection	0.00	Quest	0.00
Capacity	0.00	Opposite	0.00	Scandal	0.00	Recollection	0.00
Correction	0.00	Prejudice	0.00	Surprise	0.00	Recommendation	0.00
Failure	0.00	System	0.00	Suspicion	0.00	Sense	0.00
Leave	0.00	Terror	0.00	Terror	0.00	Time	0.00
Quest	0.00	Testimony	0.00	Testimony	0.00	Way	0.00
Recollection	0.00	Work	0.00	Triumph	0.00	Word	0.00

LC.SC	%	LC.CT	%	LC.SNP(DM.PN)	%	LC.SNP(PN.PN)	%
Foreboding	60.00	Quest	80.00	Recollection	20.00	Endorsement	20.00
Prejudice	40.00	Chance	50.00	Word	18.75	Contradiction	16.67
Contradiction	33.33	Application	27.78	Misfortune	14.29	Testimony	12.50
Sense	33.33	Sense	27.78	Thing	11.36	Scandal	11.11
Dimension	26.32	Vision	24.00	Joke	10.81	Part	8.82
Warning	17.24	Recollection	20.00	Venture	10.00	Problem	8.82
Finding	16.67	Time	16.13	Finding	8.33	Assessment	8.33
Opposite	14.29	Capacity	14.29	Area	7.69	Surprise	7.41
Triumph	14.29	Crime	13.04	Surprise	7.41	Warning	6.90
Joke	13.51	Suspicion	13.04	Point	7.14	Practice	6.06
System	12.50	Challenge	12.50	Example	6.82	Philosophy	5.56
Thing	11.36	Scandal	11.11	Part	5.88	Venture	5.00
Surprise	11.11	Terror	10.00	Problem	5.88	Failure	4.17
Motivation	10.00	Area	7.69	Answer	4.00	Vision	4.00
Word	9.38	Surprise	7.41	Characteristic	4.00	Challenge	3.13
Example	9.09	Problem	5.88	Time	3.23	Joke	2.70
Suspicion	8.70	Objective	5.56	Challenge	3.13	Way	2.63
Assessment	8.33	Finding	4.17	Phenomenon	2.86	Point	2.38
Answer	8.00	Characteristic	4.00	Objective	2.78	Example	2.27
Way	7.89	Experience	3.85	Way	2.63	Thing	2.27
Point	7.14	Project	3.85	Anger	0.00	Anger	0.00
Myth	6.45	Warning	3.45	Application	0.00	Answer	0.00
Time	6.45	Myth	3.23	Assessment	0.00	Application	0.00
Recommendation	6.25	Practice	3.03	Capacity	0.00	Area	0.00
Work	6.25	Anger	0.00	Chance	0.00	Capacity	0.00
Part	5.88	Answer	0.00	Contradiction	0.00	Chance	0.00
Problem	5.88	Assessment	0.00	Correction	0.00	Characteristic	0.00
Phenomenon	5.71	Contradiction	0.00	Crime	0.00	Correction	0.00
Capacity	4.76	Correction	0.00	Detail	0.00	Crime	0.00
Characteristic	4.00	Detail	0.00	Dimension	0.00	Detail	0.00
Experience	3.85	Dimension	0.00	Endeavour	0.00	Dimension	0.00
Practice	3.03	Endeavour	0.00	Endorsement	0.00	Endeavour	0.00
Objective	2.78	Endorsement	0.00	Evidence	0.00	Evidence	0.00
Chance	2.63	Evidence	0.00	Experience	0.00	Experience	0.00
Anger	0.00	Example	0.00	Facet	0.00	Facet	0.00
Application	0.00	Facet	0.00	Failure	0.00	Finding	0.00
Area	0.00	Failure	0.00	Foreboding	0.00	Foreboding	0.00
Challenge	0.00	Foreboding	0.00	Impetus	0.00	Impetus	0.00

Correction	0.00	Impetus	0.00	Irony	0.00	Irony	0.00
Crime	0.00	Irony	0.00	Leave	0.00	Leave	0.00
Detail	0.00	Joke	0.00	Motivation	0.00	Misfortune	0.00
Endeavour	0.00	Leave	0.00	Myth	0.00	Motivation	0.00
Endorsement	0.00	Misfortune	0.00	Opposite	0.00	Myth	0.00
Evidence	0.00	Motivation	0.00	Philosophy	0.00	Objective	0.00
Facet	0.00	Opposite	0.00	Practice	0.00	Opposite	0.00
Failure	0.00	Part	0.00	Prejudice	0.00	Phenomenon	0.00
Impetus	0.00	Phenomenon	0.00	Project	0.00	Prejudice	0.00
Irony	0.00	Philosophy	0.00	Proviso	0.00	Project	0.00
Leave	0.00	Point	0.00	Quest	0.00	Proviso	0.00
Misfortune	0.00	Prejudice	0.00	Recommendation	0.00	Quest	0.00
Philosophy	0.00	Proviso	0.00	Scandal	0.00	Recollection	0.00
Project	0.00	Recommendation	0.00	Sense	0.00	Recommendation	0.00
Proviso	0.00	System	0.00	Suspicion	0.00	Sense	0.00
Quest	0.00	Testimony	0.00	System	0.00	Suspicion	0.00
Recollection	0.00	Thing	0.00	Terror	0.00	System	0.00
Scandal	0.00	Triumph	0.00	Testimony	0.00	Terror	0.00
Terror	0.00	Venture	0.00	Triumph	0.00	Time	0.00
Testimony	0.00	Way	0.00	Vision	0.00	Triumph	0.00
Venture	0.00	Word	0.00	Warning	0.00	Word	0.00
Vision	0.00	Work	0.00	Work	0.00	Work	0.00

LC.PRM	%	LC.SNP^CNP	%
Practice	12.12	Vision	16.00
Application	11.11	Example	6.82
Dimension	5.26	Answer	4.00
Capacity	4.76	Problem	2.94
Answer	4.00	Application	2.78
Project	3.85	Way	2.63
Warning	3.45	Area	2.56
Time	3.23	Anger	0.00
Anger	0.00	Assessment	0.00
Area	0.00	Capacity	0.00
Assessment	0.00	Challenge	0.00
Challenge	0.00	Chance	0.00
Chance	0.00	Characteristic	0.00
Characteristic	0.00	Contradiction	0.00
Contradiction	0.00	Correction	0.00
Correction	0.00	Crime	0.00
Crime	0.00	Detail	0.00
Detail	0.00	Dimension	0.00
Endeavour	0.00	Endeavour	0.00
Endorsement	0.00	Endorsement	0.00
Evidence	0.00	Evidence	0.00
Example	0.00	Experience	0.00
Experience	0.00	Facet	0.00
Facet	0.00	Failure	0.00
Failure	0.00	Finding	0.00
Finding	0.00	Foreboding	0.00
Foreboding	0.00	Impetus	0.00
Impetus	0.00	Irony	0.00
Irony	0.00	Joke	0.00
Joke	0.00	Leave	0.00
Leave	0.00	Misfortune	0.00

Misfortune	0.00	Motivation	0.00
Motivation	0.00	Myth	0.00
Myth	0.00	Objective	0.00
Objective	0.00	Opposite	0.00
Opposite	0.00	Part	0.00
Part	0.00	Phenomenon	0.00
Phenomenon	0.00	Philosophy	0.00
Philosophy	0.00	Point	0.00
Point	0.00	Practice	0.00
Prejudice	0.00	Prejudice	0.00
Problem	0.00	Project	0.00
Proviso	0.00	Proviso	0.00
Quest	0.00	Quest	0.00
Recollection	0.00	Recollection	0.00
Recommendation	0.00	Recommendation	0.00
Scandal	0.00	Scandal	0.00
Sense	0.00	Sense	0.00
Surprise	0.00	Surprise	0.00
Suspicion	0.00	Suspicion	0.00
System	0.00	System	0.00
Terror	0.00	Terror	0.00
Testimony	0.00	Testimony	0.00
Thing	0.00	Thing	0.00
Triumph	0.00	Time	0.00
Venture	0.00	Triumph	0.00
Vision	0.00	Venture	0.00
Way	0.00	Warning	0.00
Word	0.00	Word	0.00
Work	0.00	Work	0.00

APPENDIX 15 FORMAL PATTERNS AND THEIR ENCAPSULATING RELATIONS (COMPLETE LISTS)

15.1 *Specific encapsulating relations*

Formal structure & Direction of encapsulation (Detailed)			
DF.AR^H-AF.INTER	2.35	DM.DT(THIS)^AJ^H^PP(against)-AF.INTER	0.07
H-EXO	2.28	DM.DT(THIS)^H^PP(in)-AF.INTER	0.07
DF.AR^H^PP(of)-CF.INTRA	2.00	DM.DT(THIS)^H^PP(of)-AF.INTER	0.07
AJ^H-EXO	1.66	DM.DT(THIS)^H^PP(of)-CF.INTRA	0.07
DM.DT(THIS)^H-AF.INTER	1.31	DM.DT(THIS)^H^PP(of)^PP(over)-AF.INTRA	0.07
DF.AR^H-CF.INTRA	1.24	DM.DT(THOSE)^H^PP(of)^PL.ING.CL-AF.INTER	0.07
DF.AR^H-EXO	1.17	H^PP(against)-AF.INTER	0.07
DF.AR^H^PP(of)-CF.INTER	0.97	H^PP(among n)^AP.THAT.CL-CF.INTRA	0.07
H-CF.INTER	0.97	H^PP(at)-CF.INTRA&AF.INTER(SN2)	0.07
AJ^H-CF.INTER	0.90	H^PP(at)^PP(in)-EXO	0.07
DF.AR^H-CF.INTER	0.90	H^PP(by n)-EXO	0.07
H-AF.INTER	0.90	H^PP(for)-CF.INTER	0.07
IN.AR^AJ^H-AF.INTRA	0.83	H^PP(for)^PL.ED.CL-CF.INTRA	0.07
DM.DT(THESE)^H-AF.INTER	0.83	H^PP(from n)-AF.INTER	0.07
PS.DT^H-EXO	0.83	H^PP(from n)-CF.INTER	0.07
DF.AR^H^PP(of)-EXO	0.76	H^PP(in)-CF.INTER	0.07
IN.AR^H-AF.INTRA&AF.INTER	0.76	H^PP(in)-AF.INTRA	0.07
DF.AR^AJ^H-AF.INTER	0.69	H^PP(of)-CF.INTRA&CF.INTRA	0.07
H^PP(of)-EXO	0.69	H^PP(of)-AF.INTRA&AF.INTER(SN2)	0.07
IN.AR^H-EXO	0.69	H^PP(of)^AP.TI.CL-CF.INTRA	0.07
DM.DT(THAT)^H-AF.INTER	0.69	H^PP(of)^PP(of)-EXO	0.07
NUM.CD^H-CF.INTER	0.69	H^PP(of)^RV.CL-AF.INTRA	0.07
DF.AR^H^PP(of)-AF.INTER	0.62	H^PP(on)-CF.INTER	0.07
DF.AR^AJ^H-CF.INTRA	0.55	H^PP(other than n)-CF.INTRA&CF.INTER	0.07
H^PP(of)-CF.INTER	0.55	H^PP(to)-AF.INTRA	0.07
H^PP(of)-CF.INTRA	0.55	H^PP(to)-AF.INTRA&AF.INTER	0.07
IN.AR^H-AF.INTRA	0.55	H^PP(with)-CF.INTER	0.07
PS.DT^H-AF.INTER	0.55	H^PP(with)-EXO	0.07
QT^H-EXO	0.48	IF.AV^QT^H^PP(of)-CF.INTRA	0.07
IN.AR^AJ^H-CF.INTER	0.48	IN.AR^N^H^PP(at)-AF.INTER	0.07
DF.AR^AJ^H^PP(of)-EXO	0.48	IN.AR^AJ^H^PP(between)-AF.INTRA&AF.INTER	0.07
IN.AR^H^PP(of)-CF.INTRA	0.48	IN.AR^AJ^H^PP(between)-EXO	0.07
DF.AR^AJ^H^PP(of)-CF.INTER	0.48	IN.AR^AJ^H^PP(for)-CF.INTRA	0.07
H^AP.TI.CL-CF.INTRA	0.48	IN.AR^AJ^H^PP(for)-AF.INTER	0.07
AJ^H-AF.INTER	0.48	IN.AR^AJ^H^PP(of)-EXO	0.07
AJ^H-CF.INTRA	0.48	IN.AR^AJ^H^PP(over)-AF.INTRA	0.07
DF.AR^AJ^H-EXO	0.48	IN.AR^AJ^H^PP(to)-AF.INTRA&AF.INTER&AF.INTER(SN2)	0.07
DF.AR^H^RV.CL-EXO	0.48	IN.AR^AJ^H^PP(with)-CF.INTER	0.07
H-AF.INTRA	0.48	IN.AR^AJ^N^H^PP(on)-CF.INTRA&AF.INTER(SN2)	0.07

IN.AR^H^AP.TI.CL-CF.INTRA	0.48	IN.AR^H^PP(about)-AF.INTER	0.07
QT^H-CF.INTER	0.48	IN.AR^H^PP(about)-CF.INTRA	0.07
N^H-EXO	0.41	IN.AR^H^PP(for)-CF.INTRA	0.07
DF.AR^AJ^H-CF.INTER	0.41	IN.AR^H^PP(for)-AF.INTER	0.07
DF.AR^AJ^H^PP(of)-CF.INTRA	0.41	IN.AR^H^PP(for)-EXO	0.07
DF.AR^N^H-AF.INTER	0.41	IN.AR^H^PP(like n)-CF.INTER	0.07
DF.AR^H^AP.TI.CL-CF.INTRA	0.41	IN.AR^H^PP(of)-AF.INTER	0.07
IN.AR^AJ^H-EXO	0.41	IN.AR^H^PP(of)^RV.CL-AF.INTRA	0.07
AS.DT^H-CF.INTER	0.41	IN.AR^H^PP(to)-AF.INTRA	0.07
N^H-AF.INTRA	0.35	IN.AR^N^H^PP(against)-AF.INTRA&AF.INTER(SN2)	0.07
N^H-CF.INTER	0.35	MR.GV^H^PP(with)-EXO	0.07
DF.DV.GV.NP^H^PP(of)-CF.INTRA	0.35	NAS.DT^H^PP(as to n)-EXO	0.07
DF.DV.GV.NP^H-EXO	0.35	NAS.DT^H^PP(between)-EXO	0.07
DF.DV.GV.NP^H^AP.TI.CL-CF.INTRA	0.35	NAS.DT^H^PP(of)-CF.INTRA	0.07
DF.DV.GV.NP^H-AF.INTER	0.35	NAS.DT^H^PP(of)^PP(at)-CF.INTER	0.07
IN.AR^AJ^H^AP.THAT.CL-CF.INTRA	0.35	NAS.DT^NUM.GO^H^PP(of)-AF.INTRA&EXO	0.07
IN.AR^N^H-AF.INTRA	0.35	NAS.PN^PV.PP(of DF.AR^H^PL.ED.CL)-EXO	0.07
DF.AR^H^PP(of)^AP.TI.CL-CF.INTRA	0.35	NAS.PN^PV.PP(of DF.AR^H^PP in n)-CF.INTER	0.07
H^PP(of)-AF.INTRA&AF.INTER	0.35	NAS.PN^PV.PP(of DM.DT(THESE)^H)-CF.INTER	0.07
IN.AR^H^PP(of)-AF.INTRA	0.35	NG.DT^H^PP(on)-EXO	0.07
DF.AR^H^RV.CL-CF.INTRA	0.35	NG.DT^H^PP(with)-EXO	0.07
H-CF.INTRA	0.35	NUM.CD(QT)^NUM.GO^H^PP(on)-AF.INTER&CF.INTER	0.07
IN.AR^AJ^H-CF.INTRA	0.35	NUM.CD^H^PP(of)^RV.CL-EXO	0.07
PS.DT^H^AP.TI.CL-CF.INTRA	0.35	NUM.CD^NUM.GO^H^PP(of)-CF.INTRA	0.07
N^H-AF.INTER	0.28	NUM.CD^PV.PP(of DF.AR^N^H)-AF.INTRA&AF.INTER&CF.INTER	0.07
DF.DV.GV.NP^AJ^H-CF.INTER	0.28	NUM.CD^PV.PP(of DF.AR^AJ^AJ^H^TI.CL)-CF.INTRA	0.07
IN.AR^AJ^N^H-EXO	0.28	NUM.CD^PV.PP(of DF.AR^AJ^H)-CF.INTRA	0.07
AJ^H^PP(in)-EXO	0.28	NUM.CD^PV.PP(of DF.AR^AJ^H^PP(about))-AF.INTRA&AF.INTER	0.07
H^PP(for)-CF.INTRA	0.28	NUM.CD^PV.PP(of DF.AR^AJ^H^RV.CL)-AF.INTRA	0.07
H^PP(of)-AF.INTRA	0.28	NUM.CD^PV.PP(of DF.AR^AJ^H^RV.CL)-CF.INTRA	0.07
IN.AR^AJ^H^PP(to)-AF.INTRA	0.28	NUM.CD^PV.PP(of DF.AR^AJ^H^RV.CL)-CF.INTRA&CF.INTER	0.07
PS.DT^H^PP(for)-CF.INTRA	0.28	NUM.CD^PV.PP(of DF.AR^H)-AF.INTRA&CF.INTRA	0.07
DF.AR^H^PP(to)-EXO	0.28	NUM.CD^PV.PP(of DF.AR^H)-CF.INTRA	0.07
H^PP(of)-AF.INTER	0.28	NUM.CD^PV.PP(of DF.AR^H^RV.CL)-AF.INTRA	0.07
QT^H^PP(of)-CF.INTER	0.28	NUM.CD^PV.PP(of	0.07

		DF.AR^NUM.GO^H^RV.CL)-AF.INTER&CF.INTRA	
		NUM.CD^PV.PP(of DM.DT(THOSE)^AJ^H^RV.CL)-	
DF.AR^H^RV.CL-CF.INTER	0.28	AF.INTER	0.07
PS.DT^H-CF.INTRA	0.28	NUM.CD^PV.PP(of NUM.CD^AJ^H)-CF.INTRA	0.07
AJ^AJ^H-EXO	0.28	NUM.CD^PV.PP(of NUM.CD^AJ^H)-CF.INTER	0.07
DF.AR^H-AF.INTRA	0.28	NUM.CD^PV.PP(of PS.DT^H^PP(since n))-CF.INTRA	0.07
DF.AR^H^AP.THAT.CL-CF.INTRA	0.28	NUM.CD^PV.PP(of PS.DT^N^H)-EXO	0.07
DM.DT(THIS)^AJ^H-AF.INTER	0.28	NUM.GO^H^PP(of)-AF.INTER&CF.INTRA	0.07
H^RV.CL-EXO	0.28	NUM.GO^H^PP(on)-AF.INTER&EXO	0.07
IN.AR^H-AF.INTER	0.28	NUM.GO^H^PP(to)^RV.CL-AF.INTER&CF.INTER	0.07
IN.AR^H-AF.INTRA&AF.INTRA	0.28	PS.DT^AJ^H^PP(in)-AF.INTRA	0.07
NAS.DT^H-EXO	0.28	PS.DT^H^PP(as)-CF.INTRA	0.07
NG.DT^H-CF.INTRA	0.28	PS.DT^H^PP(as)-EXO	0.07
PDT^H-AF.INTER	0.28	PS.DT^H^PP(during)-EXO	0.07
PS.DT^AJ^H-CF.INTER	0.28	PS.DT^H^PP(for)-AF.INTER	0.07
PS.DT^N^H-EXO	0.21	PS.DT^H^PP(for)-CF.INTRA&CF.INTER	0.07
IN.AR^AJ^H^PP(of)-CF.INTRA	0.21	PS.DT^H^PP(in)-AF.INTER	0.07
IN.AR^AJ^H-AF.INTER	0.21	PS.DT^H^PP(of)-AF.INTRA	0.07
H^PP(to)-EXO	0.21	PS.DT^H^PP(on)-AF.INTER	0.07
IN.AR^AJ^H^PP(of)-AF.INTRA	0.21	PS.DT^NUM.OR^H^PP(into)-AF.INTER	0.07
H^PP(for)-EXO	0.21	PS.DT^NUM.OR^H^PP(of)-AF.INTER	0.07
IN.AR^AJ^H^PP(for)-AF.INTRA	0.21	PS.DT^NUM.OR^H^PP(over)-EXO	0.07
NUM.CD^H^PP(of)-CF.INTRA	0.21	QT.PN^PV.PP(of DF.AR^H^PP of n)-EXO	0.07
IN.AR^H^PP(of)-EXO	0.21	QT.PN^PV.PP(of DF.AR^H^RV.CL)-AF.INTER	0.07
IN.AR^AJ^H^PP(of)-CF.INTER	0.21	QT.PN^PV.PP(of DM.DT(THESE)^H)-AF.INTRA	0.07
PS.DT^AJ^H^AP.TI.CL-CF.INTRA	0.21	QT^AJ^H^PP(on)-CF.INTER	0.07
AJ^H-AF.INTRA	0.21	QT^H^PP(concerning)-CF.INTER	0.07
AS.DT^H-EXO	0.21	QT^H^PP(from n)-CF.INTER	0.07
DF.AR^AJ^H-AF.INTRA	0.21	QT^H^PP(on)-EXO	0.07
DF.AR^NUM.OR^H^RV.CL-CF.INTRA	0.21	UV.DT^H^PP(of)-EXO	0.07
H^RV.CL-CF.INTER	0.21	UV.PDT^DF.AR^AJ^H^PP(about)-EXO	0.07
H^TI.CL-CF.INTER	0.21	UV.PDT^DF.AR^H^PP(on)-EXO	0.07
H^TI.CL-EXO	0.21	UV.PDT^H^PP(for)-CF.INTRA	0.07
IN.AR^AJ^H-AF.INTRA&AF.INTER	0.21	UV.PDT^H^PP(of)-CF.INTER	0.07
IN.AR^AJ^H-AF.INTRA&AF.INTRA	0.21	UV.PDT^NUM.GO^H^PP(of)-AF.INTER&CF.INTER	0.07
IN.AR^AJ^H^AP.TI.CL-CF.INTRA	0.21	UV.PN^PV.PP(of DF.AR^AJ^NUM.CD^H)-AF.INTER	0.07
IN.AR^H-CF.INTRA	0.21	UV.PN^PV.PP(of DM.DT(THESE)^H)-	0.07

		EXO	
IN.AR^H^AP.THAT.CL-CF.INTRA	0.21	AJ^H^PP(in)-AF.INTER	0.07
IN.AR^H^RV.CL-AF.INTRA	0.21	PS.DT^AJ^H^PP(in)-AF.INTER	0.07
IN.AR^N^H-EXO	0.21	AS.PN^PV.PP(of DF.AR^H^PP(in))-CF.INTER	0.07
PS.DT^H-CF.INTER	0.21	DF.AR^H^PP(with)-CF.INTER	0.07
PS.DT^AJ^H-CF.INTRA	0.21	DF.AR^H^RV.CL(when)-AF.INTRA&AF.INTER	0.07
UV.DT^H-EXO	0.21	DF.AR^H^RV.CL(where)-CF.INTRA	0.07
N^H-CF.INTRA	0.14	DF.AR^NUM.GO^H^PP(in)-AF.INTER&CF.INTRA	0.07
DF.DV.GV.NP^H^PP(for)-CF.INTRA	0.14	DF.AR^QL.PV^H^RV.CL(where)-CF.INTRA	0.07
DF.DV.GV.NP^H^PP(of)-AF.INTER	0.14	H^RV.CL(when)-CF.INTRA	0.07
DF.DV.GV.NP^H^PP(of)-CF.INTER	0.14	H^RV.CL(when)-EXO	0.07
DF.DV.GV.NP^AJ^H-AF.INTER	0.14	IN.AR^AJ^H^PP(of)^PP(in)-AF.INTER&CF.INTRA	0.07
DF.DV.GV.NP^H-CF.INTRA	0.14	IN.AR^AJ^H^PP(in)-EXO	0.07
DF.DV.GV.NP^N^H-CF.INTER	0.14	IN.AR^H^RV.CL(where)-CF.INTRA	0.07
DF.AR^H^FL.ST.RT.AP-CF.INTRA	0.14	IN.AR^H^RV.CL(where)-AF.INTRA&AF.INTER	0.07
IN.AR^H^PT.WK.NR.AP-CF.INTRA	0.14	IN.AR^H^RV.CL(where)-AF.INTRA&AF.INTRA	0.07
QT^H^FL.ST.NR.AP-CF.INTRA	0.14	IN.AR^N^H^PP(of)-AF.INTRA	0.07
DF.AR^N^H-AF.INTRA	0.14	NAS.DT^H^RV.CL(when)-CF.INTER	0.07
IN.AR^AJ^AJ^H-AF.INTRA	0.14	NUM.CD^H^PP(of)-EXO	0.07
DM.DT(THIS)^AJ^AJ^H-AF.INTER	0.14	NUM.CD^NUM.GO^H^RV.CL(where)-AF.INTRA&AF.INTER	0.07
DF.AR^N^H-CF.INTER	0.14	QT^H^PL.ING.CL^PP(as to)-CF.INTER	0.07
DF.AR^N^H-EXO	0.14	QT^NUM.GO^H^PP(in)-AF.INTER&EXO	0.07
IN.AR^N^H-AF.INTER	0.14	PS.DT^H^PP(for)-EXO	0.07
DF.AR^AJ^H^PP(of)-AF.INTRA	0.14	NUM.CD^PV.PP(of PS.DT^H^PP(behind))-CF.INTRA	0.07
IN.AR^H^PP(of)-CF.INTER	0.14	IN.AR^H^PP(of)^RV.CL-EXO	0.07
H^PP(in)-EXO	0.14	IN.AR^AJ^H^PP(of)-AF.INTER	0.07
AJ^H^PP(about)-CF.INTER	0.14	IN.AR^AJ^H^PP(for)-EXO	0.07
AJ^H^PP(of)-AF.INTRA&AF.INTER	0.14	DF.AR^H^PP(of)-CF.INTRA	0.07
DF.AR^AJ^AJ^H^PP(of)-EXO	0.14	AJ^H^PP(of)-CF.INTRA	0.07
DF.AR^AJ^H^PP(about)-CF.INTRA	0.14	AJ^H^PP(with)-CF.INTER	0.07
DF.AR^AJ^H^PP(of)-AF.INTER	0.14	DF.AR^AJ^AJ^H^PP(in)-AF.INTER&CF.INTER	0.07
DF.AR^H^PP(for)-CF.INTRA	0.14	DF.AR^H^PP(behind)-CF.INTER	0.07
DF.AR^H^PP(of)-AF.INTRA	0.14	DF.AR^H^PP(for)-EXO	0.07
DF.AR^H^PP(of)^AP.THAT.CL-CF.INTRA	0.14	H^PP(of)-CF.INTRA&CF.INTER	0.07
DF.AR^H^PP(on)-CF.INTER	0.14	NAS.DT^AJ^H^PP(of)-CF.INTRA	0.07
DF.AR^H^RV.CL^PP(of)-CF.INTER	0.14	NAS.DT^H^PP(in)-EXO	0.07
DF.AR^NUM.CD^AJ^H^PP(of)-	0.14	NUM.GO^H^PP(of)-	0.07

CF.INTRA		AF.INTRA&AF.INTER	
DM.DT(THAT)^H^PP(of)-AF.INTER	0.14	PS.DT^H^PP(at)-CF.INTRA	0.07
IN.AR^H^PP(of)-CF.INTRA&AF.INTER	0.14	PS.DT^H^PP(behind)-CF.INTER&CF.INTER(SN2)	0.07
IN.AR^H^PP(with)-EXO	0.14	PS.DT^H^PP(in)-CF.INTRA	0.07
NUM.CD^PV.PP(of DF.AR^H^RV.CL)-CF.INTRA	0.14	PS.DT^H^PP(in)-CF.INTER	0.07
NG.DT^H^PP(for)-EXO	0.14	QT^H^PP(of)-CF.INTRA	0.07
PS.DT^H^PP(of)-EXO	0.14	QT^H^PP(of)-EXO	0.07
DF.AR^H^PP(of)-AF.INTER&CF.INTER	0.14	AJ^H^AP.THAT.CL-CF.INTRA	0.07
H^RV.CL(when)-AF.INTER	0.14	AJ^N^H^RV.CL-CF.INTRA	0.07
RV.DT^H^PP(of)-CF.INTRA	0.14	QT^H^RV.CL-AF.INTER&EXO	0.07
PS.DT^H^PP(of)-AF.INTER	0.14	DF.AR^AJ^AJ^AJ^H-EXO	0.07
PS.DT^H^PP(of)-CF.INTRA	0.14	DF.AR^N^AJ^H^RV.CL-CF.INTER	0.07
NUM.GO^H^PP(of)-AF.INTER&CF.INTER	0.14	DF.AR^UNC^H^RV.CL-EXO	0.07
NUM.CD^H^PP(of)-CF.INTER	0.14	IN.AR^H^PL.ING.CL-CF.INTRA	0.07
NG.DT^H^PP(of)-EXO	0.14	IN.AR^N^AJ^H-AF.INTRA&AF.INTER	0.07
NG.DT^H^PP(of)-CF.INTRA	0.14	IV.DT^H-EXO	0.07
IN.AR^H^PP(of)-AF.INTRA&AF.INTER	0.14	NAS.DT^H^RV.CL-CF.INTER	0.07
H^PP(about)-EXO	0.14	PS.DT^AJ^AJ^H-EXO	0.07
AJ^H-AF.INTER&CF.INTER	0.14	QT^H^RV.CL-CF.INTER	0.07
AJ^H^PL.ED.CL-EXO	0.14	H-CF.INTRA	0.07
AJ^H^RV.CL-EXO	0.14	AJ^AJ^AJ^H-CF.INTER	0.07
AS.DT^AJ^H-CF.INTER	0.14	AJ^AJ^H-AF.INTER	0.07
AS.DT^AJ^H-EXO	0.14	AJ^AJ^H-CF.INTER	0.07
AS.DT^QL.PV^H-EXO	0.14	AJ^AJ^H-CF.INTRA	0.07
DF.AR^AJ^AJ^H-EXO	0.14	AJ^H-AF.INTER&EXO	0.07
DF.AR^AJ^H-AF.INTRA&AF.INTER	0.14	AJ^H^AP.TI.CL-CF.INTRA	0.07
DF.AR^AJ^H^RV.CL-EXO	0.14	AJ^N^AJ^H-EXO	0.07
DF.AR^AJ^H^RV.CL-AF.INTER	0.14	AJ^N^H-CF.INTER	0.07
DF.AR^H-AF.INTRA&AF.INTER	0.14	AJ^N^H^RV.CL-EXO	0.07
DF.AR^H^PL.ED.CL-AF.INTER	0.14	AS.DT^H-CF.INTRA	0.07
DF.AR^H^PL.ED.CL-EXO	0.14	AS.DT^N^H-CF.INTER	0.07
DF.AR^H^RV.CL-AF.INTRA&AF.INTER	0.14	DF.AR^AJ^(NG)H-CF.INTRA	0.07
DF.AR^H^RV.CL-AF.INTER	0.14	DF.AR^AJ^AJ^H-CF.INTRA	0.07
DF.AR^H^RV.CL-AF.INTRA	0.14	DF.AR^AJ^AJ^H^PL.ING.CL-EXO	0.07
DF.AR^QL.PV^H^RV.CL-AF.INTRA&AF.INTER	0.14	DF.AR^AJ^H-AF.INTRA&CF.INTER	0.07
DM.DT(THESE)^H-CF.INTER	0.14	DF.AR^AJ^H-AF.INTRA&AF.INTRA	0.07
DM.DT(THIS)^H-AF.INTRA	0.14	DF.AR^AJ^H^(for n TI.CL)-CF.INTRA	0.07
H-AF.INTRA&CF.INTRA	0.14	DF.AR^AJ^H^AP.THAT.CL-CF.INTRA	0.07
IN.AR^AJ^H-CF.INTRA&CF.INTER	0.14	DF.AR^AJ^H^AP.TI.CL-CF.INTRA	0.07
IN.AR^AJ^H^PL.ED.CL-AF.INTRA	0.14	DF.AR^AJ^H^RV.CL-UNC	0.07
IN.AR^H-AF.INTRA&AF.INTER(SN2)	0.14	DF.AR^AJ^H^TI.CL-CF.INTRA	0.07
IN.AR^H^PL.ED.CL-AF.INTRA	0.14	DF.AR^AJ^N^H-AF.INTER	0.07
IV.DT^H-CF.INTER	0.14	DF.AR^AJ^N^H^RV.CL-CF.INTRA	0.07
NAS.DT^AJ^H-AF.INTER	0.14	DF.AR^H-AF.INTER&CF.INTER	0.07

NG.DT^AJ^H-CF.INTRA	0.14	DF.AR^H-AF.INTRA&AF.INTER&CF.INTRA	0.07
PS.DT^H-AF.INTER&CF.INTER	0.14	DF.AR^H-CF.INTRA&CF.INTER	0.07
PS.DT^H-UNC	0.14	DF.AR^H-UNC	0.07
NUM.CD^AJ^H-CF.INTRA	0.14	DF.AR^H^(for n AP.TI.CL)-CF.INTRA	0.07
NUM.GO^H-AF.INTER&CF.INTER	0.14	DF.AR^H^AV-AF.INTRA	0.07
NUM.GO^H-AF.INTRA&EXO	0.14	DF.AR^H^AV-EXO	0.07
PDT^H^RV.CL(as)-CF.INTRA	0.14	DF.AR^H^PL.ED.CL-CF.INTER	0.07
PS.DT^AJ^H-AF.INTRA	0.14	DF.AR^H^PL.ED.CL-CF.INTRA	0.07
PS.DT^AJ^H-EXO	0.14	DF.AR^H^PL.ING.CL-CF.INTRA	0.07
PS.DT^N^H-AF.INTER	0.14	DF.AR^H^TI.CL-EXO	0.07
QT^AJ^H-CF.INTER	0.14	DF.AR^N^H^AP.THAT.CL-CF.INTRA	0.07
QT^AJ^H-EXO	0.14	DF.AR^NUM.CD^AJ^H^RV.CL-CF.INTRA&CF.INTER	0.07
QT^H-AF.INTER	0.14	DF.AR^NUM.CD^H-EXO	0.07
QT^H^RV.CL-EXO	0.14	DF.AR^NUM.CD^H-AF.INTER	0.07
UV.PDT^DF.AR^H-EXO	0.14	DF.AR^NUM.CD^N^N^H^RV.CL-CF.INTRA	0.07
DF.AR^AJ^PP(of DF.DV.GV.NP^H)-CF.INTRA	0.07	DF.AR^NUM.CD^NUM.GO^H^RV.CL-AF.INTER&CF.INTRA	0.07
AS.PN^PV.PP(of DF.DV.GV.NP^H^PP of n)-EXO	0.07	DF.AR^NUM.GO^AJ^H^TI.CL-AF.INTRA&CF.INTRA	0.07
DF.DV.GV.NP^H^PP(to)-CF.INTER	0.07	DF.AR^NUM.GO^H-CF.INTRA	0.07
DF.DV.GV.NP^AJ^AJ^H-AF.INTER	0.07	DF.AR^NUM.GO^H-AF.INTER&CF.INTRA	0.07
DF.DV.GV.NP^AJ^H-AF.INTRA	0.07	DF.AR^NUM.GO^H^RV.CL-CF.INTRA&EXO	0.07
DF.DV.GV.NP^H-CF.INTER	0.07	DF.AR^NUM.OR^H-CF.INTER	0.07
DF.DV.GV.NP^N^H-AF.INTER	0.07	DF.AR^QL.PV^AJ^H^RV.CL-AF.INTRA	0.07
DF.DV.GV.NP^NUM.CD^H-AF.INTER	0.07	DF.AR^QL.PV^H^RV.CL-CF.INTER	0.07
DF.DV.GV.NP^NUM.CD^H-CF.INTER	0.07	DF.AR^QL.PV^H^RV.CL-EXO	0.07
DF.DV.GV.NP^NUM.GO^H-AF.INTER	0.07	DM.DT(THAT)^AJ^H-AF.INTER	0.07
N^H^RV.CL-EXO	0.07	DM.DT(THAT)^H-AF.INTRA	0.07
AJ^H^FL.ST.RT.AP-CF.INTRA	0.07	DM.DT(THAT)^H-CF.INTER	0.07
H^FL.ST.RT.AP-CF.INTRA	0.07	DM.DT(THAT)^H-EXO	0.07
AJ^H^PP(in)^PT.WK.NR.AP-CF.INTRA	0.07	DM.DT(THAT)^H^AP.TI.CL-CF.INTRA	0.07
AJ^H^PT.WK.NR.AP-CF.INTRA	0.07	DM.DT(THAT)^QL.PV^H-AF.INTER	0.07
DF.AR^AJ^H^PT.WK.NR.AP-CF.INTRA	0.07	DM.DT(THESE)^AJ^H-AF.INTER	0.07
H^PP(of)^PT.WK.NR.AP-CF.INTRA	0.07	DM.DT(THESE)^H-AF.INTER&CF.INTER	0.07
IN.AR^AJ^H^PT.WK.NR.AP-AF.INTRA&CF.INTRA	0.07	DM.DT(THESE)^H-EXO	0.07
IN.AR^AJ^H^PT.WK.NR.AP-CF.INTRA	0.07	DM.DT(THESE)^H-AF.INTRA	0.07
NUM.CD^H^PT.WK.NR.AP-CF.INTRA	0.07	DM.DT(THESE)^QL.PV^AJ^H-AF.INTER	0.07
NUM.CD^PV.PP(of NUM.CD^H^AV^PT.WK.NR.AP)	0.07	DM.DT(THIS)^AJ^H-AF.INTRA	0.07

-CF.INTRA			
PS.DT^AJ^AJ^H^PT.WK.NR.AP -EXO	0.07	DM.DT(THIS)^H-EXO	0.07
QT^AJ^H^PP(for)^ PT.WK.NR.AP-CF.INTRA	0.07	DM.DT(THIS)^H- CF.INTRA&CF.INTER	0.07
QV.PV^H^PT.WK.NR.AP- CF.INTRA	0.07	DM.DT(THIS)^N^H-AF.INTRA	0.07
DF.AR^AJ^H^FL.WK.NR.AP- CF.INTRA	0.07	DM.DT(THIS)^NUM.GO^H-AF.INTER	0.07
AJ^N^H^PT.ST.NR.AP^RV.CL- CF.INTRA	0.07	DM.DT(THIS)^QL.PV^AJ^H^RV.CL- CF.INTER	0.07
DF.AR^AJ^NUM.GO^H^PP(of)^ PT.ST.NR.AP-CF.INTRA	0.07	DM.DT(THIS)^QL.PV^H-AF.INTER	0.07
QT^H^PP(on)^PT.ST.NR.AP- CF.INTRA	0.07	DM.DT(THIS)^QL.PV^H-CF.INTER	0.07
AJ^H^PT.ST.NR.AP-CF.INTRA	0.07	DM.DT(THOSE)^AJ^H-AF.INTER	0.07
DF.AR^H^RV.CL^PT.ST.NR.AP -EXO	0.07	DM.DT(THOSE)^H-AF.INTER	0.07
DF.AR^N^H^PT.ST.NR.AP-EXO	0.07	H-AF.INTER&CF.INTER	0.07
IN.AR^H^PT.ST.NR.AP- CF.INTRA&CF.INTER	0.07	H^(for n AP.TI.CL)-CF.INTRA	0.07
N^H^PT.ST.NR.AP^RV.CL- CF.INTRA	0.07	H^AP.THAT.CL-CF.INTRA	0.07
NUM.CD^AJ^H^PT.ST.NR.AP- CF.INTRA	0.07	H^PL.ED.CL-CF.INTRA	0.07
DM.DT(THAT)^H^PP(of)^ FL.ST.NR.AP-CF.INTRA	0.07	H^PL.ED.CL-EXO	0.07
DF.AR^NUM.GO^N^H^ FL.ST.NR.AP-EXO	0.07	H^RV.CL-AF.INTRA&EXO	0.07
IN.AR^AJ^H^FL.ST.NR.AP^ PL.ED.CL-CF.INTRA	0.07	IF.AV^IN.AR^AJ^H-EXO	0.07
N^H^FL.ST.NR.AP-CF.INTRA	0.07	IF.AV^IN.AR^H-CF.INTRA	0.07
NUM.CD^H^FL.ST.NR.IT.AP- CF.INTRA	0.07	IN.AR^AJ^AJ^H-EXO	0.07
IN.AR^AJ^H^PL.ED.CL^ AP.THAT.CL-CF.INTRA	0.07	IN.AR^AJ^H-AF.INTRA&EXO	0.07
NUM.CD^PV.PP(of DF.AR^AJ^H^PL.ING.CL)- CF.INTRA	0.07	IN.AR^AJ^H^PL.ED.CL-CF.INTRA	0.07
IN.AR^AJ^H^PL.ED.CL- AF.INTER	0.07	IN.AR^AJ^H^RV.CL-AF.INTRA	0.07
IN.AR^AJ^H^PL.ING.CL- CF.INTER	0.07	IN.AR^H-AF.INTRA&CF.INTER	0.07
IN.AR^AJ^H^RV.CL-CF.INTRA	0.07	IN.AR^H-CF.INTER	0.07
DF.AR^H^PP(of)^RV.CL-EXO	0.07	IN.AR^H-UNC	0.07
IN.AR^H^PP(of)^RV.CL- CF.INTRA	0.07	IN.AR^H^PL.ED.CL-CF.INTER	0.07
NUM.CD^H^PP(of)^RV.CL- CF.INTER	0.07	IN.AR^H^RV.CL- AF.INTRA&AF.INTER	0.07
AJ^H^RV.CL(where)-EXO	0.07	IN.AR^H^TI.CL-CF.INTER	0.07
DF.AR^AJ^N^H^RV.CL-EXO	0.07	IN.AR^N^AJ^H-CF.INTER	0.07
IN.AR^AJ^H^RV.CL-EXO	0.07	IN.AR^N^H-AF.INTRA&AF.INTER	0.07
PS.DT^H^RV.CL-AF.INTER	0.07	IN.AR^N^H-CF.INTER	0.07
UV.PDT^DM.DT(THESE)^AJ^H ^RV.CL-AF.INTER	0.07	IN.AR^NUM.GO^H-EXO	0.07
N^N^H-AF.INTRA	0.07	IV.DT^H^RT.TV.CL-AF.INTRA	0.07

DM.DT(THIS)^N^H-CF.INTER	0.07	N^N^H-AF.INTER	0.07
IN.AR^N^H^AP.TI.CL-CF.INTRA	0.07	NAS.DT^AJ^H-EXO	0.07
DF.AR^AJ^H^PP(of)-AF.INTER&CF.INTER	0.07	NAS.DT^AJ^N^H-EXO	0.07
AS.DT^QT^H^TI.CL-AF.INTER&EXO	0.07	NAS.DT^H-AF.INTER	0.07
QT^H-AF.INTER&EXO	0.07	NAS.DT^H-AF.INTRA&AF.INTRA	0.07
IN.AR^AJ^AJ^H^AP.TI.CL-CF.INTRA	0.07	NAS.DT^H-CF.INTRA	0.07
NAS.DT^QT^H-AF.INTER&CF.INTER	0.07	NAS.DT^H^AP.THAT.CL-EXO	0.07
PDT(CV.RU)^IN.AR^AJ^H-AF.INTRA	0.07	NAS.DT^QL.PV^H^RV.CL-AF.INTER	0.07
AS.DT^AJ^AJ^H^PP(in)-EXO	0.07	NG.DT^AJ^AJ^H-EXO	0.07
IN.AR^AJ^H^PP(by)^PP(of)-CF.INTER	0.07	NG.DT^AJ^H-AF.INTER	0.07
IN.AR^AJ^H^PP(of)-AF.INTRA&AF.INTER(SN2)	0.07	NG.DT^AJ^H-AF.INTRA	0.07
NUM.CD^AJ^H^PP(in terms of)-CF.INTER	0.07	NG.DT^AJ^H-AF.INTRA&AF.INTER	0.07
PS.DT^AJ^H^PP(of)-AF.INTER	0.07	NG.DT^AJ^H-AF.INTRA&AF.INTRA	0.07
AJ^AJ^H-AF.INTRA	0.07	NG.DT^AJ^H-CF.INTER	0.07
AJ^H-AF.INTRA&AF.INTER	0.07	NG.DT^AJ^H^RV.CL-EXO	0.07
AJ^IN.AR^H-AF.INTER	0.07	NG.DT^H-AF.INTER	0.07
DF.AR^AJ^H^AJ-CF.INTRA	0.07	NG.DT^H-AF.INTRA	0.07
H^AJ-CF.INTRA	0.07	PS.DT^H-AF.INTER&AF.INTER(SN2)	0.07
IN.AR^AJ^AJ^H^AJ-CF.INTRA	0.07	PS.DT^H-AF.INTRA	0.07
N^H^AJ-CF.INTER	0.07	PS.DT^H-AF.INTRA&AF.INTER	0.07
NUM.CD^AJ^H-AF.INTER	0.07	NG.DT^H^AP.THAT.CL-EXO	0.07
NUM.CD^H^AJ-CF.INTRA&EXO	0.07	NG.DT^H^AP.TI.CL-CF.INTRA	0.07
DF.AR^N^AJ^H-EXO	0.07	NG.DT^H^TI.CL-EXO	0.07
DF.AR^N^H-AF.INTER&CF.INTER	0.07	NG.DT^N^H-EXO	0.07
N^N^H^PP(as to n)-CF.INTER	0.07	NG.DT^NUM.GO^H-AF.INTER&EXO	0.07
IN.AR^N^H^RV.CL(whereby)-CF.INTER	0.07	NG.DT^NUM.GO^H^TI.CL-AF.INTER&EXO	0.07
N^H^PP(for)-CF.INTRA	0.07	NUM.CD(QT)^H-EXO	0.07
AJ^H^PP(of)-AF.INTRA	0.07	NUM.CD^AJ^H-AF.INTRA&AF.INTER	0.07
AS.PN^PV.PP(of DF.AR^AJ^N^H)-EXO	0.07	NUM.CD^AJ^H-CF.INTER	0.07
DF.AR^QL.PV^AJ^H^PP(of)-AF.INTER	0.07	NUM.CD^AJ^H^RV.CL-CF.INTRA	0.07
H^PP(of)-AF.INTRA&AF.INTRA	0.07	NUM.CD^H-AF.INTRA&AF.INTER	0.07
IN.AR^AJ^H^PP(on)-EXO	0.07	NUM.CD^H-CF.INTRA	0.07
IN.AR^H^PP(about)-CF.INTER	0.07	NUM.CD^H-EXO	0.07
IN.AR^N^H^PP(of)-CF.INTER	0.07	NUM.CD^N^H-AF.INTRA	0.07
IV.DT^QL.PV^H^PP(about)-CF.INTER	0.07	NUM.CD^NUM.GO^H^PL.ED.CL-AF.INTER&CF.INTRA&CF.INTER	0.07
NUM.CD(QT)^H^PP(about)-CF.INTER	0.07	NUM.CD^NUM.GO^QV.PV^H-CF.INTRA	0.07
NUM.CD^PV.PP(of PS.DT^H)-CF.INTRA	0.07	NUM.CD^QV.PV^H-CF.INTRA	0.07
PS.DT^H^PP(on)-EXO	0.07	NUM.GO^AJ^AJ^H-AF.INTER&CF.INTRA	0.07
AJ^H^PP(about)-EXO	0.07	NUM.GO^AJ^H-AF.INTRA&CF.INTER	0.07
AJ^H^PP(against)-CF.INTER	0.07	NUM.GO^AJ^H^PL.ED.CL-EXO	0.07

AJ^H^PP(for)-EXO	0.07	NUM.GO^H-AF.INTER&CF.INTRA	0.07
AJ^H^PP(from n)^AP.THAT.CL-CF.INTRA	0.07	NUM.GO^H-AF.INTER&EXO	0.07
AJ^H^PP(in)-CF.INTER	0.07	NUM.GO^H-CF.INTRA	0.07
AJ^H^PP(of)-CF.INTER	0.07	NUM.GO^H-CF.INTRA&CF.INTER(SN2)	0.07
AJ^H^PP(of)-EXO	0.07	NUM.GO^H-EXO	0.07
AJ^H^PP(on)-EXO	0.07	NUM.GO^H^PL.ED.CL-AF.INTER&CF.INTER	0.07
AJ^H^PP(to)-CF.INTER	0.07	PDT^AJ^H-AF.INTER	0.07
AJ^H^PP(with)-EXO	0.07	PDT^H-CF.INTER	0.07
AS.DT^H^PP(for)-EXO	0.07	PDT^IN.AR^H-AF.INTER	0.07
AS.DT^H^PP(in)-EXO	0.07	PS.DT^AJ^AJ^AJ^H-EXO	0.07
AS.DT^H^PP(of)-EXO	0.07	PS.DT^AJ^H-AF.INTER	0.07
AS.DT^H^PP(to)-EXO	0.07	PS.DT^AJ^H-UNC	0.07
AS.PN^PV.PP(of DM.DT(THESE)^H)-AF.INTER	0.07	PS.DT^AJ^H^AP.THAT.CL-CF.INTRA	0.07
QT.PN^PV.PP(of DF.AR^H^PP of)-AF.INTER	0.07	PS.DT^H^AP.THAT.CL-CF.INTRA	0.07
QT^AJ^H^PP(in)-CF.INTRA	0.07	PS.DT^H^PL.ED.CL-CF.INTER	0.07
QT^H^PP(on)-CF.INTER	0.07	PS.DT^NUM.GO^H-CF.INTRA	0.07
DF.AR^AJ^AJ^H^PP(of)-AF.INTRA	0.07	PS.DT^NUM.OR^AJ^H-AF.INTER	0.07
DF.AR^AJ^H^PP(between)-CF.INTRA	0.07	PS.DT^NUM.OR^H-AF.INTER	0.07
DF.AR^AJ^H^PP(for)-AF.INTRA	0.07	PS.DT^QL.PV^AJ^H-AF.INTER	0.07
DF.AR^AJ^H^PP(for)-AF.INTRA&AF.INTER	0.07	PS.DT^QT^H-EXO	0.07
DF.AR^AJ^H^PP(for)-CF.INTRA	0.07	QT.IV^H-CF.INTER	0.07
DF.AR^AJ^H^PP(from within n)-EXO	0.07	QT^DF.AR^H-CF.INTER	0.07
DF.AR^AJ^H^PP(in)-CF.INTER	0.07	QT^DS.GV^H-EXO	0.07
DF.AR^AJ^H^PP(of)^PP(for)-CF.INTRA	0.07	QT^H-AF.INTRA	0.07
DF.AR^AJ^H^PP(of)^RV.CL-AF.INTER	0.07	QT^H-AF.INTRA&AF.INTER	0.07
DF.AR^AJ^H^PP(over)-CF.INTRA	0.07	QT^H-CF.INTRA	0.07
DF.AR^AJ^H^PP(to)-AF.INTRA	0.07	QT^H^AP.THAT.CL-EXO	0.07
DF.AR^AJ^H^RV.CL^PP(with)-EXO	0.07	QT^H^PL.ING.CL-CF.INTRA	0.07
DF.AR^H^PP(between)-AF.INTER	0.07	QT^H^TI.CL-CF.INTER	0.07
DF.AR^H^PP(for)-AF.INTER&CF.INTRA	0.07	QT^H^TI.CL-CF.INTRA	0.07
DF.AR^H^PP(from n)-CF.INTRA	0.07	QT^IN.AR^H-AF.INTRA	0.07
DF.AR^H^PP(in)^AP.THAT.CL-CF.INTRA	0.07	QT^IN.AR^H-CF.INTRA	0.07
DF.AR^H^PP(of)-AF.INTRA&AF.INTER	0.07	QV.PV^DM.DT(THAT)^H-AF.INTRA	0.07
DF.AR^H^PP(of)-AF.INTRA&AF.INTER&CF.INTER	0.07	QV.PV^H-CF.INTRA&CF.INTER	0.07
DF.AR^H^PP(of)-AF.INTRA&AF.INTRA	0.07	QV.PV^UV.PDT^PS.DT^H-AF.INTER	0.07
DF.AR^H^PP(of)-	0.07	RV.DT^AJ^H-CF.INTRA	0.07

AF.INTRA&CF.INTER			
DF.AR^H^PP(of)^PP(about)-AF.INTER	0.07	RV.DT^H-AF.INTER	0.07
DF.AR^H^PP(of)^PP(in)-CF.INTER	0.07	RV.DT^H-EXO	0.07
DF.AR^H^PP(of)^PP(in)-EXO	0.07	UV.DT^H-AF.INTER	0.07
DF.AR^H^PP(of)^PP(on)-AF.INTRA&AF.INTER	0.07	UV.DT^H^RV.CL-CF.INTRA	0.07
DF.AR^H^PP(to)-AF.INTRA&AF.INTER	0.07	UV.DT^QL.PV^H-EXO	0.07
DF.AR^H^PP(with)-AF.INTRA	0.07	UV.PDT^DF.AR^H^PL.ED.CL-EXO	0.07
DF.AR^N^H^PP(of)-CF.INTER	0.07	UV.PDT^DF.AR^H^RV.CL-EXO	0.07
DF.AR^NUM.OR^AJ^AJ^H^PP(of)-CF.INTRA	0.07	UV.PDT^DM.DT(THAT)^H-AF.INTER	0.07
DF.AR^NUM.OR^H^PP(of)-AF.INTRA&AF.INTER&CF.INTER	0.07	UV.PDT^H-EXO	0.07
DF.AR^NUM.OR^H^TI.CL^PP(with)-CF.INTRA	0.07	UV.PDT^PS.DT^AJ^H-EXO	0.07
DF.AR^NUM.OR^PV.PP(of DM.DT(THESE)^H)-AF.INTER	0.07	UV.PDT^QL.PV^H-AF.INTER	0.07
DM.DT(THAT)^H^PP(of)-CF.INTRA	0.07	UV.PDT^QL.PV^NUM.GO^H-AF.INTRA&EXO	0.07

15.2 *Main encapsulating relations*

Formal structure & Direction of encapsulation (AF, CF, EXO)			
DF.AR^H^PP(of)-CF	3.04	DF.AR^NUM.OR^H^TI.CL^PP(with)-CF	0.07
DF.AR^H-AF	2.76	DF.AR^NUM.OR^PV.PP(of DM.DT(THESE)^H)-AF	0.07
H-EXO	2.28	DM.DT(THAT)^H^PP(of)-CF	0.07
DF.AR^H-CF	2.21	DM.DT(THIS)^AJ^H^PP(against)-AF	0.07
IN.AR^H-AF	2.00	DM.DT(THIS)^H^PP(in)-AF	0.07
AJ^H-EXO	1.66	DM.DT(THIS)^H^PP(of)-CF	0.07
IN.AR^AJ^H-AF	1.45	DM.DT(THOSE)^H^PP(of)^PL.ING.CL-AF	0.07
DM.DT(THIS)^H-AF	1.45	H^PP(against)-AF	0.07
H-AF	1.38	H^PP(among)^AP.THAT.CL-CF	0.07
AJ^H-CF	1.38	H^PP(at)-CF&AF(SN2)	0.07
H-CF	1.38	H^PP(for)^PL.ED.CL-CF	0.07
H^PP(of)-CF	1.24	H^PP(from)-AF	0.07
DF.AR^H-EXO	1.17	H^PP(from)-CF	0.07
DF.AR^AJ^H-AF	1.11	H^PP(in)-CF	0.07
DF.AR^AJ^H-CF	1.04	H^PP(in)-AF	0.07
IN.AR^AJ^H-CF	0.97	H^PP(of)^AP.TI.CL-CF	0.07
DF.AR^H^PP(of)-AF	0.90	H^PP(of)^RV.CL-AF	0.07
DF.AR^AJ^H^PP(of)-CF	0.90	H^PP(on)-CF	0.07
DM.DT(THESE)^H-AF	0.90	H^PP(other than)-CF	0.07
PS.DT^H-EXO	0.83	H^PP(with)-CF	0.07
DF.AR^H^PP(of)-EXO	0.76	IF.AV^QT^H^PP(of)-CF	0.07
AJ^H-AF	0.76	IN.AR^N^H^PP(at)-AF	0.07
PS.DT^H-AF	0.76	IN.AR^AJ^H^PP(between)-AF	0.07
NUM.CD^H-CF	0.76	IN.AR^AJ^H^PP(for)-CF	0.07
DM.DT(THAT)^H-AF	0.76	IN.AR^AJ^H^PP(over)-AF	0.07

H^PP(of n)-EXO	0.69	IN.AR^AJ^H^PP(with)-CF	0.07
IN.AR^H-EXO	0.69	IN.AR^AJ^N^H^PP(on)-CF&AF(SN2)	0.07
IN.AR^H^PP(of)-CF	0.62	IN.AR^H^PP(about)-AF	0.07
H^PP(of)-AF	0.62	IN.AR^H^PP(for)-CF	0.07
N^H-AF	0.62	IN.AR^H^PP(for)-AF	0.07
DF.AR^H^RV.CL-CF	0.62	IN.AR^H^PP(like)-CF	0.07
IN.AR^H^PP(of)-AF	0.55	IN.AR^H^PP(of)^RV.CL-AF	0.07
IN.AR^N^H-AF	0.55	IN.AR^H^PP(to)-AF	0.07
DF.AR^N^H-AF	0.55	NAS.DT^H^PP(of)-CF	0.07
QT^H-CF	0.55	NAS.DT^H^PP(of)^PP(at)-CF	0.07
DF.DV.GV.NP^H^PP(of)-CF	0.48	NAS.DT^NUM.GO^H^PP(of)-AF&EXO	0.07
DF.AR^AJ^H^PP(of)-EXO	0.48	NAS.PN^PV.PP(of DF.AR^H^PP in)-CF	0.07
N^H-CF	0.48	NAS.PN^PV.PP(of DM.DT(THESE)^H)-CF	0.07
PS.DT^AJ^H-CF	0.48	NUM.CD(QT)^NUM.GO^H^PP(on)-AF&CF	0.07
PS.DT^H-CF	0.48	NUM.CD^NUM.GO^H^PP(of)-CF	0.07
QT^H-EXO	0.48	NUM.CD^PV.PP(of DF.AR^N^H)-AF&CF	0.07
DF.AR^AJ^H-EXO	0.48	NUM.CD^PV.PP(of DF.AR^AJ^AJ^H^TI.CL)-CF	0.07
DF.AR^H^RV.CL-EXO	0.48	NUM.CD^PV.PP(of DF.AR^AJ^H)-CF	0.07
AS.DT^H-CF	0.48	NUM.CD^PV.PP(of DF.AR^AJ^H^PP(about))-AF	0.07
H^AP.TI.CL-CF	0.48	NUM.CD^PV.PP(of DF.AR^AJ^H^RV.CL)-AF	0.07
IN.AR^H^AP.TI.CL-CF	0.48	NUM.CD^PV.PP(of DF.AR^H)-AF&CF	0.07
IN.AR^AJ^H^PP(of)-CF	0.41	NUM.CD^PV.PP(of DF.AR^H)-CF	0.07
H^PP(of n)-AF	0.41	NUM.CD^PV.PP(of DF.AR^H^RV.CL)-AF	0.07
DF.AR^H^RV.CL-AF	0.41	NUM.CD^PV.PP(of DF.AR^NUM.GO^H^RV.CL)-AF&CF	0.07
IN.AR^AJ^H-EXO	0.41	NUM.CD^PV.PP(of DM.DT(THOSE)^AJ^H^RV.CL)-AF	0.07
DF.AR^H^AP.TI.CL-CF	0.41	NUM.CD^PV.PP(of PS.DT^H^PP(since))-CF	0.07
H^PP(for)-CF	0.35	NUM.GO^H^PP(on)-AF&EXO	0.07
NUM.CD^H^PP(of)-CF	0.35	NUM.GO^H^PP(to)^RV.CL-AF&CF	0.07
IN.AR^AJ^H^PP(of)-AF	0.35	PS.DT^H^PP(as)-CF	0.07
QT^H^PP(of)-CF	0.35	PS.DT^H^PP(for)-AF	0.07
PS.DT^H^PP(for)-CF	0.35	PS.DT^H^PP(in)-AF	0.07
IN.AR^AJ^H^PP(to)-AF	0.35	PS.DT^H^PP(on)-AF	0.07
DF.AR^H^PP(of)^AP.TI.CL-CF	0.35	PS.DT^NUM.OR^H^PP(into)-AF	0.07
DM.DT(THIS)^AJ^H-AF	0.35	PS.DT^NUM.OR^H^PP(of)-AF	0.07
DF.DV.GV.NP^H-EXO	0.35	QT.PN^PV.PP(of DF.AR^H^RV.CL)-AF	0.07
IN.AR^AJ^H^AP.THAT.CL-CF	0.35	QT.PN^PV.PP(of DM.DT(THESE)^H)-AF	0.07
DF.DV.GV.NP^H^AP.TI.CL-CF	0.35	QT^AJ^H^PP(on)-CF	0.07
DF.DV.GV.NP^H-AF	0.35	QT^H^PP(concerning)-CF	0.07
PS.DT^H^AP.TI.CL-CF	0.35	QT^H^PP(from)-CF	0.07
DF.AR^H^PP(of)-AF&CF	0.28	UV.PDT^H^PP(for)-CF	0.07
DF.AR^AJ^H^PP(of)-AF	0.28	UV.PDT^H^PP(of)-CF	0.07
IN.AR^AJ^H^PP(for)-AF	0.28	UV.PDT^NUM.GO^H^PP(of)-AF&CF	0.07
AJ^H^PP(in)-EXO	0.28	UV.PN^PV.PP(of	0.07

		DF.AR^AJ^NUM.CD^H)-AF	
DF.AR^H^PP(to)-EXO	0.28	AJ^H^PP(in)-AF	0.07
QT^H-AF	0.28	AS.PN^PV.PP(of DF.AR^H^PP(in))-CF	0.07
IN.AR^H^RV.CL-AF	0.28	DF.AR^H^PP(with)-CF	0.07
IN.AR^H-CF	0.28	DF.AR^NUM.GO^H^PP(in)-AF&CF	0.07
N^H-EXO	0.28	IN.AR^N^H^PP(of)-AF	0.07
IN.AR^AJ^N^H-EXO	0.28	QT^H^PL.ING.CL^PP(as to)-CF	0.07
AJ^AJ^H-EXO	0.28	QT^NUM.GO^H^PP(in)-AF&EXO	0.07
		NUM.CD^PV.PP(of	
H^RV.CL-EXO	0.28	PS.DT^H^PP(behind))-CF	0.07
NAS.DT^H-EXO	0.28	AJ^H^PP(with)-CF	0.07
DF.DV.GV.NP^AJ^H-CF	0.28	DF.AR^AJ^AJ^H^PP(in)-AF&CF	0.07
DF.AR^H^AP.THAT.CL-CF	0.28	DF.AR^H^PP(behind)-CF	0.07
NG.DT^H-CF	0.28	NAS.DT^AJ^H^PP(of)-CF	0.07
PDT^H-AF	0.28	PS.DT^H^PP(at)-CF	0.07
NG.DT^AJ^H-AF	0.28	PS.DT^H^PP(behind)-CF	0.07
NUM.GO^H^PP(of)-AF&CF	0.21	QT^H-AF&EXO	0.07
PS.DT^H^PP(of)-AF	0.21	PS.DT^AJ^AJ^H^PT.WK.NR.AP-EXO	0.07
AJ^H^PP(of)-AF	0.21	DF.AR^H^RV.CL^PT.ST.NR.AP-EXO	0.07
H^PP(to)-EXO	0.21	DF.AR^N^H^PT.ST.NR.AP-EXO	0.07
		DF.AR^NUM.GO^N^H^FL.ST.NR.AP-	
H^PP(for)-EXO	0.21	EXO	0.07
IN.AR^H^PP(of)-EXO	0.21	AJ^H^RV.CL(where)-EXO	0.07
PS.DT^AJ^H-AF	0.21	DF.AR^AJ^N^H^RV.CL-EXO	0.07
NUM.GO^H-AF&CF	0.21	IN.AR^AJ^H^RV.CL-EXO	0.07
NUM.GO^H-AF&EXO	0.21	DF.AR^N^AJ^H-EXO	0.07
NUM.CD^AJ^H-CF	0.21	NP^H^RV.CL-EXO	0.07
NG.DT^AJ^H-CF	0.21	H^RV.CL(when)-EXO	0.07
IN.AR^AJ^H^PL.ED.CL-AF	0.21	DF.AR^AJ^AJ^AJ^H-EXO	0.07
DF.DV.GV.NP^H-CF	0.21	DF.AR^UNC^H^RV.CL-EXO	0.07
DF.DV.GV.NP^AJ^H-AF	0.21	IV.DT^H-EXO	0.07
PS.DT^N^H-EXO	0.21	PS.DT^AJ^AJ^H-EXO	0.07
AS.DT^H-EXO	0.21	AJ^N^AJ^H-EXO	0.07
H^TI.CL-EXO	0.21	AJ^N^H^RV.CL-EXO	0.07
IN.AR^N^H-EXO	0.21	DF.AR^AJ^AJ^H^PL.ING.CL-EXO	0.07
UV.DT^H-EXO	0.21	DF.AR^H^AV-EXO	0.07
H-AF&CF	0.21	DF.AR^H^TI.CL-EXO	0.07
PS.DT^AJ^H^AP.TI.CL-CF	0.21	DF.AR^NUM.CD^H-EXO	0.07
DF.AR^NUM.OR^H^RV.CL-CF	0.21	DF.AR^QL.PV^H^RV.CL-EXO	0.07
H^RV.CL-CF	0.21	DM.DT(THAT)^H-EXO	0.07
H^TI.CL-CF	0.21	DM.DT(THESE)^H-EXO	0.07
IN.AR^AJ^H^AP.TI.CL-CF	0.21	DM.DT(THIS)^H-EXO	0.07
IN.AR^H^AP.THAT.CL-CF	0.21	H^PL.ED.CL-EXO	0.07
PS.DT^AJ^H^PP(in)-AF	0.14	IF.AV^IN.AR^AJ^H-EXO	0.07
DF.AR^AJ^H^PP(for)-AF	0.14	IN.AR^AJ^AJ^H-EXO	0.07
AJ^H^PP(of)-CF	0.14	IN.AR^NUM.GO^H-EXO	0.07
IN.AR^H^PP(about)-CF	0.14	NAS.DT^AJ^H-EXO	0.07
IN.AR^H^PP(of)-CF&AF	0.14	NAS.DT^AJ^N^H-EXO	0.07
H^PP(in)-EXO	0.14	NAS.DT^H^AP.THAT.CL-EXO	0.07
DF.AR^AJ^AJ^H^PP(of n)-EXO	0.14	NG.DT^AJ^AJ^H-EXO	0.07
IN.AR^H^PP(with n)-EXO	0.14	NG.DT^AJ^H^RV.CL-EXO	0.07
NG.DT^H^PP(for)-EXO	0.14	NG.DT^H^AP.THAT.CL-EXO	0.07
PS.DT^H^PP(of)-EXO	0.14	NG.DT^H^TI.CL-EXO	0.07
NG.DT^H^PP(of)-EXO	0.14	NG.DT^N^H-EXO	0.07
H^PP(about)-EXO	0.14	NUM.CD(QT)^H-EXO	0.07
DF.DV.GV.NP^H^PP(for)-CF	0.14	NUM.CD^H-EXO	0.07

DF.DV.GV.NP^H^PP(of)-AF	0.14	NUM.GO^AJ^H^PL.ED.CL-EXO	0.07
AJ^H^PP(about)-CF	0.14	NUM.GO^H-EXO	0.07
DF.AR^AJ^H^PP(about)-CF	0.14	PS.DT^AJ^AJ^AJ^H-EXO	0.07
DF.AR^H^PP(for)-CF	0.14	PS.DT^QT^H-EXO	0.07
DF.AR^H^PP(of)^AP.THAT.CL-CF	0.14	QT^DS.GV^H-EXO	0.07
DF.AR^H^PP(on)-CF	0.14	QT^H^AP.THAT.CL-EXO	0.07
DF.AR^H^RV.CL^PP(of)-CF	0.14	RV.DT^H-EXO	0.07
DF.AR^NUM.CD^AJ^H^PP(of)-CF	0.14	UV.DT^QL.PV^H-EXO	0.07
DM.DT(THAT)^H^PP(of)-AF	0.14	UV.PDT^DF.AR^H^PL.ED.CL-EXO	0.07
NUM.CD^PV.PP(of DF.AR^H^RV.CL)-CF	0.14	UV.PDT^DF.AR^H^RV.CL-EXO	0.07
RV.DT^H^PP(of)-CF	0.14	UV.PDT^H-EXO	0.07
PS.DT^H^PP(of)-CF	0.14	UV.PDT^PS.DT^AJ^H-EXO	0.07
NG.DT^H^PP(of)-CF	0.14	DF.AR^N^H-AF&CF	0.07
H^PP(to)-AF	0.14	AJ^H-AF&EXO	0.07
NUM.CD^PV.PP(of DF.AR^AJ^H^RV.CL)-CF	0.14	DF.AR^AJ^H-AF&CF	0.07
NUM.CD^PV.PP(of NUM.CD^AJ^H)-CF	0.14	IN.AR^H-AF&CF	0.07
PS.DT^H^PP(in)-CF	0.14	DM.DT(THESE)^H-AF&CF	0.07
NUM.CD^AJ^H-AF	0.14	IN.AR^AJ^H-AF&EXO	0.07
AJ^AJ^H-AF	0.14	QT^DF.AR^H-CF	0.07
N^N^H-AF	0.14	AJ^H^FL.ST.RT.AP-CF	0.07
DF.AR^N^H-EXO	0.14	H^FL.ST.RT.AP-CF	0.07
NP^H-EXO	0.14	AJ^H^PT.WK.NR.AP-CF	0.07
AJ^H^PL.ED.CL-EXO	0.14	DF.AR^AJ^H^PT.WK.NR.AP-CF	0.07
AJ^H^RV.CL-EXO	0.14	IN.AR^AJ^H^PT.WK.NR.AP-AF&CF	0.07
AS.DT^AJ^H-EXO	0.14	IN.AR^AJ^H^PT.WK.NR.AP-CF	0.07
AS.DT^QL.PV^H-EXO	0.14	NUM.CD^H^PT.WK.NR.AP-CF	0.07
DF.AR^AJ^AJ^H-EXO	0.14	QV.PV^H^PT.WK.NR.AP-CF	0.07
DF.AR^AJ^H^RV.CL-EXO	0.14	DF.AR^AJ^H^FL.WK.NR.AP-CF	0.07
DF.AR^H^PL.ED.CL-EXO	0.14	AJ^N^H^PT.ST.NR.AP^RV.CL-CF	0.07
PS.DT^AJ^H-EXO	0.14	AJ^H^PT.ST.NR.AP-CF	0.07
QT^AJ^H-EXO	0.14	IN.AR^H^PT.ST.NR.AP-CF	0.07
QT^H^RV.CL-EXO	0.14	N^H^PT.ST.NR.AP^RV.CL-CF	0.07
UV.PDT^DF.AR^H-EXO	0.14	NUM.CD^AJ^H^PT.ST.NR.AP-CF	0.07
AJ^H-AF&CF	0.14	IN.AR^AJ^H^FL.ST.NR.AP^PL.ED.CL-CF	0.07
PS.DT^H-AF&CF	0.14	N^H^FL.ST.NR.AP-CF	0.07
DF.AR^H-AF&CF	0.14	NUM.CD^H^FL.ST.NR.IT.AP-CF	0.07
DF.AR^H^FL.ST.RT.AP-CF	0.14	IN.AR^AJ^H^PL.ED.CL^AP.THAT.CL-CF	0.07
IN.AR^H^PT.WK.NR.AP-CF	0.14	IN.AR^AJ^H^PL.ING.CL-CF	0.07
QT^H^FL.ST.NR.AP-CF	0.14	IN.AR^AJ^H^RV.CL-CF	0.07
IN.AR^AJ^AJ^H-AF	0.14	PS.DT^H^RV.CL-AF	0.07
DM.DT(THIS)^AJ^AJ^H-AF	0.14	UV.PDT^DM.DT(THESE)^AJ^H^RV.CL-AF	0.07
DF.AR^N^H-CF	0.14	DM.DT(THIS)^N^H-CF	0.07
DF.DV.GV.NP^N^H-CF	0.14	IN.AR^N^H^AP.TI.CL-CF	0.07
H^RV.CL(when)-AF	0.14	AS.DT^QT^H^TI.CL-AF&EXO	0.07
AS.DT^AJ^H-CF	0.14	IN.AR^AJ^AJ^H^AP.TI.CL-CF	0.07
DF.AR^AJ^H^RV.CL-AF	0.14	NAS.DT^QT^H-AF&CF	0.07
DF.AR^H^PL.ED.CL-AF	0.14	PDT(CV.RU)^IN.AR^AJ^H-AF	0.07
DF.AR^QL.PV^H^RV.CL-AF	0.14	AJ^IN.AR^H-AF	0.07

DM.DT(THESE)^H-CF	0.14	DF.AR^AJ^H^AJ-CF	0.07
IN.AR^H^PL.ED.CL-AF	0.14	H^AJ-CF	0.07
IV.DT^H-CF	0.14	IN.AR^AJ^AJ^H^AJ-CF	0.07
NAS.DT^AJ^H-AF	0.14	N^H^AJ-CF	0.07
PS.DT^H-UNC	0.14	NUM.CD^H^AJ-CF&EXO	0.07
PDT^H^RV.CL(as)-CF	0.14	IN.AR^N^H^RV.CL(whereby)-CF	0.07
PS.DT^N^H-AF	0.14	DF.DV.GV.NP^AJ^AJ^H-AF	0.07
QT^AJ^H-CF	0.14	DF.DV.GV.NP^N^H-AF	0.07
IN.AR^H^RV.CL(where)-AF	0.14	DF.DV.GV.NP^NUM.CD^H-AF	0.07
AJ^AJ^H-CF	0.14	DF.DV.GV.NP^NUM.CD^H-CF	0.07
DF.AR^H^PL.ED.CL-CF	0.14	DF.DV.GV.NP^NUM.GO^H-AF	0.07
NAS.DT^H-AF	0.14	DF.AR^H^RV.CL(when)-AF	0.07
NG.DT^H-AF	0.14	DF.AR^H^RV.CL(where)-CF	0.07
NUM.GO^H-CF	0.14	DF.AR^QL.PV^H^RV.CL(where)-CF	0.07
QT^H^TI.CL-CF	0.14	H^RV.CL(when)-CF	0.07
IN.AR^N^H^PP(against)-AF	0.07	IN.AR^H^RV.CL(where)-CF	0.07
DM.DT(THIS)^H^PP(of)-AF	0.07	NAS.DT^H^RV.CL(when)-CF	0.07
DM.DT(THIS)^H^PP(of)^ PP(over)-AF	0.07	NUM.CD^NUM.GO^H^RV.CL(where)-AF	0.07
NUM.GO^H^PP(of)-AF	0.07	AJ^H^AP.THAT.CL-CF	0.07
DF.AR^H^PP(of)^PP(about)-AF	0.07	AJ^N^H^RV.CL-CF	0.07
DF.AR^H^PP(of)^PP(on)-AF	0.07	QT^H^RV.CL-AF&EXO	0.07
DF.AR^AJ^H^PP(of)-AF&CF	0.07	DF.AR^N^AJ^H^RV.CL-CF	0.07
IN.AR^AJ^H^PP(of)^PP(in)-AF&CF	0.07	IN.AR^H^PL.ING.CL-CF	0.07
DF.AR^H^PP(of)^RV.CL-EXO	0.07	IN.AR^N^AJ^H-AF	0.07
AS.DT^AJ^AJ^H^PP(in)-EXO	0.07	NAS.DT^H^RV.CL-CF	0.07
AS.PN^PV.PP(of DF.DV.GV.NP^H^PP of n)-EXO	0.07	QT^H^RV.CL-CF	0.07
AS.PN^PV.PP(of DF.AR^AJ^N^H)-EXO	0.07	AJ^AJ^AJ^H-CF	0.07
IN.AR^AJ^H^PP(on)-EXO	0.07	AJ^H^AP.TI.CL-CF	0.07
PS.DT^H^PP(on)-EXO	0.07	AJ^N^H-CF	0.07
AJ^H^PP(about n)-EXO	0.07	AS.DT^N^H-CF	0.07
AJ^H^PP(for n)-EXO	0.07	DF.AR^AJ^AJ^H-CF	0.07
AJ^H^PP(of n)-EXO	0.07	DF.AR^AJ^H^(for n TI.CL)-CF	0.07
AJ^H^PP(on n)-EXO	0.07	DF.AR^AJ^H^AP.THAT.CL-CF	0.07
AJ^H^PP(with n)-EXO	0.07	DF.AR^AJ^H^AP.TI.CL-CF	0.07
AS.DT^H^PP(for n)-EXO	0.07	DF.AR^AJ^H^RV.CL-UNC	0.07
AS.DT^H^PP(in n)-EXO	0.07	DF.AR^AJ^H^TI.CL-CF	0.07
AS.DT^H^PP(of n)-EXO	0.07	DF.AR^AJ^N^H-AF	0.07
AS.DT^H^PP(to n)-EXO	0.07	DF.AR^AJ^N^H^RV.CL-CF	0.07
DF.AR^AJ^H^PP(from within n)-EXO	0.07	DF.AR^H-UNC	0.07
DF.AR^AJ^H^RV.CL^ PP(with n)-EXO	0.07	DF.AR^H^(for n AP.TI.CL)-CF	0.07
DF.AR^H^PP(of n)^PP(in n)-EXO	0.07	DF.AR^H^AV-AF	0.07
H^PP(at n)^PP(in n)-EXO	0.07	DF.AR^H^PL.ING.CL-CF	0.07
H^PP(by n)-EXO	0.07	DF.AR^N^H^AP.THAT.CL-CF	0.07
H^PP(of n)^PP(of n)-EXO	0.07	DF.AR^NUM.CD^AJ^H^RV.CL-CF	0.07
H^PP(with n)-EXO	0.07	DF.AR^NUM.CD^H-AF	0.07
IN.AR^AJ^H^PP(between n)-EXO	0.07	DF.AR^NUM.CD^N^N^H^RV.CL-CF	0.07
IN.AR^AJ^H^PP(of n)-EXO	0.07	DF.AR^NUM.CD^NUM.GO^H^RV.CL-AF&CF	0.07

IN.AR^H^PP(for n)-EXO	0.07	DF.AR^NUM.GO^AJ^H^TI.CL-AF&CF	0.07
MR.GV^H^PP(with n)-EXO	0.07	DF.AR^NUM.GO^H-CF	0.07
NAS.DT^H^PP(as to n)-EXO	0.07	DF.AR^NUM.GO^H-AF&CF	0.07
NAS.DT^H^PP(between n)-EXO	0.07	DF.AR^NUM.GO^H^RV.CL-CF&EXO	0.07
NAS.PN^PV.PP(of DF.AR^H^PL.ED.CL)-EXO	0.07	DF.AR^NUM.OR^H-CF	0.07
NG.DT^H^PP(on n)-EXO	0.07	DF.AR^QL.PV^AJ^H^RV.CL-AF	0.07
NG.DT^H^PP(with n)-EXO	0.07	DF.AR^QL.PV^H^RV.CL-CF	0.07
NUM.CD^H^PP(of n)^RV.CL-EXO	0.07	DM.DT(THAT)^AJ^H-AF	0.07
NUM.CD^PV.PP(of PS.DT^N^H)-EXO	0.07	DM.DT(THAT)^H-CF	0.07
PS.DT^H^PP(as n)-EXO	0.07	DM.DT(THAT)^H^AP.TI.CL-CF	0.07
PS.DT^H^PP(during n)-EXO	0.07	DM.DT(THAT)^QL.PV^H-AF	0.07
PS.DT^NUM.OR^H^PP(over n)-EXO	0.07	DM.DT(THESE)^AJ^H-AF	0.07
QT.PN^PV.PP(of DF.AR^H^PP of n)-EXO	0.07	DM.DT(THESE)^QL.PV^AJ^H-AF	0.07
QT^H^PP(on n)-EXO	0.07	DM.DT(THIS)^H-CF	0.07
UV.DT^H^PP(of n)-EXO	0.07	DM.DT(THIS)^N^H-AF	0.07
UV.PDT^DF.AR^AJ^H^PP(about n)-EXO	0.07	DM.DT(THIS)^NUM.GO^H-AF	0.07
UV.PDT^DF.AR^H^PP(on n)-EXO	0.07	DM.DT(THIS)^QL.PV^AJ^H^RV.CL-CF	0.07
UV.PN^PV.PP(of DM.DT(THESE)^H)-EXO	0.07	DM.DT(THIS)^QL.PV^H-AF	0.07
IN.AR^AJ^H^PP(in)-EXO	0.07	DM.DT(THIS)^QL.PV^H-CF	0.07
NUM.CD^H^PP(of)-EXO	0.07	DM.DT(THOSE)^AJ^H-AF	0.07
PS.DT^H^PP(for)-EXO	0.07	DM.DT(THOSE)^H-AF	0.07
IN.AR^H^PP(of)^RV.CL-EXO	0.07	H^(for n AP.TI.CL)-CF	0.07
IN.AR^AJ^H^PP(for)-EXO	0.07	H^AP.THAT.CL-CF	0.07
DF.AR^H^PP(for)-EXO	0.07	H^PL.ED.CL-CF	0.07
NAS.DT^H^PP(in)-EXO	0.07	H^RV.CL-AF&EXO	0.07
QT^H^PP(of)-EXO	0.07	IF.AV^IN.AR^H-CF	0.07
AJ^H^PP(in)^PT.WK.NR.AP-CF	0.07	IN.AR^AJ^H^PL.ED.CL-CF	0.07
H^PP(of)^PT.WK.NR.AP-CF	0.07	IN.AR^AJ^H^RV.CL-AF	0.07
NUM.CD^PV.PP(of NUM.CD^H^AV^PT.WK.NR.AP)-CF	0.07	IN.AR^H-UNC	0.07
QT^AJ^H^PP(for n)^PT.WK.NR.AP-CF	0.07	IN.AR^H^PL.ED.CL-CF	0.07
DF.AR^AJ^NUM.GO^H^PP(of)^PT.ST.NR.AP-CF	0.07	IN.AR^H^TI.CL-CF	0.07
QT^H^PP(on)^PT.ST.NR.AP-CF	0.07	IN.AR^N^AJ^H-CF	0.07
DM.DT(THAT)^H^PP(of)^FL.ST.NR.AP-CF	0.07	IN.AR^N^H-CF	0.07
NUM.CD^PV.PP(of DF.AR^AJ^H^PL.ING.CL)-CF	0.07	IV.DT^H^RT.TV.CL-AF	0.07
IN.AR^H^PP(of)^RV.CL-CF	0.07	NAS.DT^H-CF	0.07
NUM.CD^H^PP(of)^RV.CL-CF	0.07	NAS.DT^QL.PV^H^RV.CL-AF	0.07
IN.AR^AJ^H^PP(by)^PP(of)-CF	0.07	NG.DT^H^AP.TI.CL-CF	0.07
NUM.CD^AJ^H^PP(in terms of)-CF	0.07	NG.DT^NUM.GO^H-AF&EXO	0.07
PS.DT^AJ^H^PP(of)-AF	0.07	NG.DT^NUM.GO^H^TI.CL-AF&EXO	0.07
DF.AR^AJ^	0.07	NUM.CD^AJ^H^RV.CL-CF	0.07

PP(of DF.DV.GV.NP^H)-CF			
DF.DV.GV.NP^H^PP(to)-CF	0.07	NUM.CD^H-AF	0.07
N^N^H^PP(as to)-CF	0.07	NUM.CD^N^H-AF	0.07
		NUM.CD^NUM.GO^H^PL.ED.CL-	
N^H^PP(for)-CF	0.07	AF&CF	0.07
DF.AR^QL.PV^AJ^H^PP(of)-AF	0.07	NUM.CD^NUM.GO^QV.PV^H-CF	0.07
IN.AR^N^H^PP(of)-CF	0.07	NUM.CD^QV.PV^H-CF	0.07
IV.DT^QL.PV^H^PP(about)-CF	0.07	NUM.GO^AJ^AJ^H-AF&CF	0.07
NUM.CD(QT)^H^PP(about)-CF	0.07	NUM.GO^AJ^H-AF&CF	0.07
NUM.CD^PV.PP(of PS.DT^H)-			
CF	0.07	NUM.GO^H^PL.ED.CL-AF&CF	0.07
AJ^H^PP(against)-CF	0.07	PDT^AJ^H-AF	0.07
AJ^H^PP(from)^AP.THAT.CL-			
CF	0.07	PDT^H-CF	0.07
AJ^H^PP(in)-CF	0.07	PDT^IN.AR^H-AF	0.07
AJ^H^PP(to)-CF	0.07	PS.DT^AJ^H-UNC	0.07
AS.PN^PV.PP(of			
DM.DT(THESE)^H)-AF	0.07	PS.DT^AJ^H^AP.THAT.CL-CF	0.07
QT.PN^PV.PP(of DF.AR^H^			
PP of)-AF	0.07	PS.DT^H^AP.THAT.CL-CF	0.07
QT^AJ^H^PP(in)-CF	0.07	PS.DT^H^PL.ED.CL-CF	0.07
QT^H^PP(on)-CF	0.07	PS.DT^NUM.GO^H-CF	0.07
DF.AR^AJ^AJ^H^PP(of)-AF	0.07	PS.DT^NUM.OR^AJ^H-AF	0.07
DF.AR^AJ^H^PP(between)-CF	0.07	PS.DT^NUM.OR^H-AF	0.07
DF.AR^AJ^H^PP(for)-CF	0.07	PS.DT^QL.PV^AJ^H-AF	0.07
DF.AR^AJ^H^PP(in)-CF	0.07	QT.IV^H-CF	0.07
DF.AR^AJ^H^PP(of)^PP(for)-			
CF	0.07	QT^H^PL.ING.CL-CF	0.07
DF.AR^AJ^H^PP(of)^RV.CL-AF	0.07	QT^IN.AR^H-AF	0.07
DF.AR^AJ^H^PP(over)-CF	0.07	QT^IN.AR^H-CF	0.07
DF.AR^AJ^H^PP(to)-AF	0.07	QV.PV^DM.DT(THAT)^H-AF	0.07
DF.AR^H^PP(between)-AF	0.07	QV.PV^H-CF	0.07
DF.AR^H^PP(for)-AF&CF	0.07	QV.PV^UV.PDT^PS.DT^H-AF	0.07
DF.AR^H^PP(from)-CF	0.07	RV.DT^AJ^H-CF	0.07
DF.AR^H^PP(in)^AP.THAT.CL-			
CF	0.07	RV.DT^H-AF	0.07
DF.AR^H^PP(of)^PP(in)-CF	0.07	UV.DT^H-AF	0.07
DF.AR^H^PP(to)-AF	0.07	UV.DT^H^RV.CL-CF	0.07
DF.AR^H^PP(with)-AF	0.07	UV.PDT^DM.DT(THAT)^H-AF	0.07
DF.AR^N^H^PP(of)-CF	0.07	UV.PDT^QL.PV^H-AF	0.07
DF.AR^NUM.OR^AJ^AJ^H^			
PP(of)-CF	0.07	UV.PDT^QL.PV^NUM.GO^H-AF&EXO	0.07
DF.AR^NUM.OR^H^PP(of)-			
AF&CF	0.07		

REFERENCES

Aarts, Bas. 2001. *English syntax and argumentation*, 2nd edn. Basingstoke: Palgrave.

Aarts, F.G.A.M. 2004. On the distribution of noun-phrase types in English clause-structure. In Geoffrey Sampson & Diana McCarthy (eds.), *Corpus linguistics: Readings in a widening discipline*, 34–48. London: Continuum.

Aijmer, Karin. 2007. The interface between discourse and grammar. In Agnès Celle & Ruth Huart (eds.), *Connectives as discourse landmarks*, 31–46. Amsterdam, Philadelphia, PA: John Benjamins Publishing Company.

Aktas, Rahime Nur & Viviana Cortes. 2008. Shell nouns as cohesive devices in published and ESL student writing. *Journal of English for Academic Purposes* 7(1), 3–14.

Anthony, Laurence. 2011. *AntConc* (Version 3.2.4). Tokyo, Japan: Waseda University. http://www.antlab.sci.waseda.ac.jp/antconc_index.html (accessed 22 June 2013)

Ariel, Mira. 1990. *Accessing noun-phrase antecedents*. London: Routledge.

Ashby, Michael (ed.). 2005. *Oxford advanced learner's dictionary of current English*, 7th edn (CD-ROM). Oxford: Oxford University Press.

Asher, Nicholas. 1993. *Reference to abstract objects in discourse*. Dordrecht: Kluwer Academic Publishers.

Baker, Mona, Gill Francis & Elena Tognini-Bonelli (eds.). 1993. *Text and technology*: *In honour of John Sinclair*. Amsterdam, Philadelphia, PA: John Benjamins Publishing Company.

Bednarek, Monika. 2009. Corpora and discourse: A three-pronged approach to analyzing linguistic data. In Michael Haugh, Kate Burridge, Jean Mulder & Pam Peters (eds.), *Selected proceedings of the 2008 HCSNet Workshop on Designing the Australian National Corpus: Mustering languages*, 19–24. Somerville, MA: Cascadilla Proceedings Project.

Bhatia, Vijay K. 1993. *Analysing genre: Language use in professional settings*. London: Longman.

Biber, Douglas. 2006. Stance in spoken and written university registers. *Journal of English for Academic Purposes* 5(2), 97–116.

Biber, Douglas & Susan Conrad. 2009. *Register, genre, and style*. Cambridge: Cambridge University Press.

Biber, Douglas & Bethany Gray. 2010. Challenging stereotypes about academic writing: Complexity, elaboration, explicitness. *Journal of English for Academic Purposes* 9(1), 2–20.

Biber, Douglas & Bethany Gray. 2011. Grammatical change in the noun phrase: The influence of written language use. *English Language and Linguistics* 15(2), 223–50.

Biber, Douglas, Stig Johansson, Geoffrey Leech, Susan Conrad & Edward Finegan. 1999. *Longman grammar of spoken and written English*. London: Longman.

BNC Sampler, XML version. 2005. Distributed by Oxford University Computing Services on behalf of the BNC Consortium.

BNC Sampler: powered by CQPweb. 2013. http://cqpweb.lancs.ac.uk/bncsampler/ (accessed 22 June 2013)

BNCweb, version 4.2. 2008. http://bncweb.lancs.ac.uk/ (accessed 22 June 2013)

Bolinger, Dwight. 1977. *Pronouns and repeated nouns*. Bloomington, IN: Indiana University Linguistics Club.

Botley, Simon Philip. 2006. Indirect anaphora: Testing the limits of corpus-based linguistics. *International Journal of Corpus Linguistics* 11(1), 73–112.

Brown, Gillian & George Yule. 1983. *Discourse analysis*. Cambridge: Cambridge University Press.

Burnard, Lou. 2007. *Reference guide for the British National Corpus (XML edition)*. http://www.natcorp.ox.ac.uk/docs/URG/ (accessed 22 June 2013)

Burnard, Lou. 2008a. *Reference guide for the BNC Sampler (XML edition)*. http://www.natcorp.ox.ac.uk/corpus/sampler/sampler.pdf (accessed 22 June 2013)

Burnard, Lou. 2008b. *Reference guide for the BNC Baby*, 2nd edn. http://www.natcorp.ox.ac.uk/corpus/baby/manual.pdf (accessed 22 June 2013)

Caldwell, Candice. 2009. *Lexical vagueness in student writing*: *Are shell nouns the problem?.* Saarbrücken: VDM, Verlag Dr. Müller.

Carter, Ronald. 1998. *Vocabulary*: *Applied linguistic perspectives*, 2nd edn. London: Routledge.

Carter, Ronald & Michael McCarthy. 2006. *Cambridge grammar of English*: *A comprehensive guide. Spoken and written English grammar and usage*. Cambridge: Cambridge University Press.

Charles, Maggie. 2003. 'This mystery...': A corpus-based study of the use of nouns to construct stance in theses from two contrasting disciplines. *Journal of English for Academic Purposes* 2(4), 313–26.

Charles, Maggie. 2007. Argument or evidence? Disciplinary variation in the use of the noun *that* pattern in stance construction. *English for Specific Purposes* 26(2), 203–18.

Chomsky, Noam. 1970. Remarks on nominalization. In Roderick A. Jacobs & Peter S. Rosenbaum (eds.), *Readings in transformational grammar*, 184–221. Washington, WA: Georgetown University Press.

Chomsky, Noam. 1981. *Lectures on government and binding*. Dordrecht: Foris.

Consten, Manfred, Mareile Knees & Monika Schwarz-Friesel. 2007. The function of complex anaphors in texts: Evidence from corpus studies and ontological considerations. In Monika Schwarz-Friesel, Manfred Consten & Mareile Knees (eds.), *Anaphors in text*: *Cognitive, formal and applied approaches to anaphoric reference*, 81–102. Amsterdam, Philadelphia, PA: John Benjamins Publishing Company.

Conte, Maria-Elisabeth. 1996. Anaphoric encapsulation. In Walter de Mulder & Liliane Tasmowski (eds.), *Coherence and Anaphora*. *Belgian Journal of Linguistics* 10, 1–10.

Coulthard, Malcolm (ed.). 1994. *Advances in written text analysis*. London: Routledge.

Curme, George O. 1947. *English grammar. The principles and practice of English grammar applied to present-day usage*. New York, NY: Barnes & Noble, Inc.

de Beaugrande, Robert A. & Wolfgang U. Dressler. 1981. *Introduction to text linguistics*. London, New York, NY: Longman.

de Haan, Pieter. 1991. On the exploration of corpus data by means of problem-oriented tagging: Postmodifying clauses in the English noun phrase. In Stig Johansson & Anna-Brita Stenström (eds.), *English computer corpora*: *Selected papers and research guide*, 51–65. Berlin: Mouton de Gruyter.

Delahunty, Gerald P. 2012. An analysis of *the thing is that* sentences. *Pragmatics* 22(1): 41–78.

Downing, Angela & Philip Locke. 2006. *English grammar*: *A university course*, 2nd edn. London: Routledge.

Fairclough, Norman. 1995. *Critical discourse analysis: The critical study of language*. London: Longman.

Fairclough, Norman. 2000. *New labour, new language?* London, New York, NY: Routledge.

Flowerdew, John. 2002. A pedagogic grammar of signalling nouns in discourse. *Revista Canaria de Estudios Ingleses* 44, 141–55.

Flowerdew, John. 2003a. Signalling nouns in discourse. *English for Specific Purposes* 22(4), 329–46.

Flowerdew, John. 2003b. Register-specificity of signalling nouns in discourse. In Pepi Leistyna & Charles F. Meyer (eds.), *Corpus analysis: Language structure and language use*, 35–46. Amsterdam: Rodopi.

Flowerdew, John. 2006. Use of signalling nouns in a learner corpus. *International Journal of Corpus Linguistics* 11(3), 345–62.

Flowerdew, John. 2010. Use of signalling nouns across L1 and L2 writer corpora. *International Journal of Corpus Linguistics* 15(1), 36–55.

Flowerdew, John. 2011. Small corpora, larger corpora and discourse analysis in the study of lexical cohesion. Presented at Corpus Linguistics Conference, University of Birmingham.

Flowerdew, John & Lynne Flowerdew. 2013. Discourse properties of signalling nouns in academic genres. Presented at the Conference on Genre- and Register-related Text and Discourse Features in Multilingual Corpora, Institut Libre Marie Haps, Brussels.

Flowerdew, Lynne. 2008. *Corpus-based analyses of the problem-solution pattern*: *A phraseological approach*. Amsterdam, Philadelphia, PA: John Benjamins Publishing Company.

Fox, Barbara A. 1987. *Discourse structure and anaphora: Written and conversational English*. Cambridge: Cambridge University Press.

Fox, Chris & Rosalind Combley (eds.). 2009. *Longman dictionary of contemporary English*, 5th edn (CD-ROM). Harlow: Pearson Education Limited.

Francis, Gill. 1986. *Anaphoric nouns* (Discourse Analysis Monograph 11). Birmingham: English Language Research, University of Birmingham.

Francis, Gill. 1988. The teaching of techniques of lexical cohesion in an ESL setting. In Verner Bickley (ed.), *Language in a bilingual or multilingual Setting*, 325–38. Hong Kong: Institute of Language in Education.

Francis, Gill. 1991. Nominal group heads and clause structure. *Word* 42(2), 145–56.

Francis, Gill. 1993. A corpus-driven approach to grammar: Principles, methods and examples. In Baker, Francis & Tognini-Bonelli (eds.), 137–56.

Francis, Gill. 1994. Labelling discourse: An aspect of nominal-group lexical cohesion. In Coulthard (ed.), 83–101.

Francis, Gill, Susan Hunston & Elizabeth Manning. 1998. *Collins Cobuild grammar patterns 2*: *Nouns and adjectives*. London: HarperCollins Publishers.

Fraurud, Kari. 1988. Hierarchical relations between discourse referents and the interpretation of antecedentless definite noun phrases. In Victoria Rosén (ed.), *Papers from the Tenth Scandinavian Conference of Linguistics, Bergen, June 11-13, 1987*, 196–210. Bergen: Department of Linguistics and Phonetics, University of Bergen.

Fraurud, Kari. 1990. Definiteness and the processing of NP's in natural discourse. *Journal of Semantics* 7, 395–433.

Fraurud, Kari. 1992. *Situation reference (What does 'it' refer to?)* (GAP Working Paper 24). Hamburg: Fachbereich Informatik, Universität Hamburg.

Garnham, Alan & Jane Oakhill. 1990. Mental models as contexts for interpreting texts: Implications from studies of anaphora. *Journal of Semantics* 7: 379–93.

Goldberg, Adele E. 2006. *Constructions at work*. *The nature of generalization in language*. Oxford: Oxford University Press.

Gray, Bethany. 2010. On the use of demonstrative pronouns and determiners as cohesive devices: A focus on sentence-initial *this/these* in academic prose. *Journal of English for Academic Purposes* 9(3), 167–83.

Greenbaum, Sidney. 1996. *The Oxford English grammar*. Oxford: Oxford University Press.

Halliday, Michael A.K. 1978. *Language as social semiotic*: *The social interpretation of language and meaning*. London, New York, NY: Edward Arnold.

Halliday, Michael A.K. 1993. On the language of physical science. In Halliday & Martin (eds.), 54–68.

Halliday, Michael A.K. 1994. *An introduction to functional grammar*, 2nd edn. London: Edward Arnold.

Halliday, Michael A.K. 1998. Things and relations: Regrammaticising experience as technical knowledge. In James R. Martin & Robert Veel (eds.), *Reading science: Critical and functional perspectives on discourses of science*, 185–235. London: Routledge.

Halliday, Michael A.K. & Ruqaiya Hasan. 1976. *Cohesion in English*. London, New York: Longman.

Halliday, Michael A.K. & James R. Martin. 1993. The model. In Halliday & Martin (eds.), 22–50.

Halliday, Michael A.K. & James R. Martin (eds.). 1993. *Writing science: Literacy and discursive power*. London, Washington, WA: The Falmer Press.

Halliday, Michael A.K. & Christian M.I.M. Matthiessen. 2004. *An introduction to functional grammar*, 3rd edn. London: Hodder Education.

Hardie, Andrew. 2012. CQPweb - combining power, flexibility and usability in a corpus analysis tool. *International Journal of Corpus Linguistics* 17(3), 380–409.

Hasselgård, Hilde. 2012. Facts, ideas, questions, problems, and issues in advanced learners' English. *Nordic Journal of English Studies* 11(1), 22–54.

Hawkins, John A. 1978. *Definiteness and indefiniteness: A study in reference and grammaticality prediction*. London: Croom Helm.

Hinkel, Eli. 2004. *Teaching academic ESL writing*: *Practical techniques in vocabulary and grammar*. Mahwah, NJ, London: Lawrence Erlbaum Associates Publishers.

Hoey, Michael. 1979. *Signalling in discourse* (Discourse Analysis Monograph 6). Birmingham: English Language Research, University of Birmingham.

Hoey, Michael. 1983. *On the surface of discourse*. London: George Allen & Unwin.

Hoey, Michael. 1993. A common signal in discourse: How the word *reason* is used in texts. In Sinclair, Hoey & Fox (eds.), 67–82.

Hoey, Michael. 1994. Signalling in discourse: A functional analysis of a common discourse pattern in written and spoken English. In Coulthard (ed.), 26–45.

Hoey, Michael. 2005. *Lexical priming*: *A new theory of words and language*. London: Routledge.

Hoey, Michael & Matthew B. O'Donnell. 2008. Lexicography, grammar, and textual position. *International Journal of Lexicography* 21(3), 293–309.

Hoffmann, Sebastian, Stefan Evert, Nicholas Smith, David Lee & Ylva Berglund Prytz. 2008. *Corpus linguistics with BNCweb*: *A practical guide*. Frankfurt am Main: Peter Lang.

Huddleston, Rodney. 2002. Content clauses and reported speech. In Huddleston & Pullum (eds.), 947–1030.

Huddleston, Rodney & Geoffrey Pullum (eds.). 2002. *The Cambridge grammar of the English language*. Cambridge: Cambridge University Press.

Hunston, Susan. 2001. Colligation, lexis, pattern, and text. In Michael Scott & Geoff Thompson (eds.), *Patterns of text*: *In honour of Michael Hoey*, 13–33. Amsterdam, Philadelphia, PA: John Benjamins Publishing Company.

Hunston, Susan. 2002. *Corpora in applied linguistics*. Cambridge: Cambridge University Press.

Hunston, Susan. 2008. Starting with the small words: Patterns, lexis and semantic sequences. *International Journal of Corpus Linguistics* 13(3), 271–95.

Hunston, Susan & Gill Francis. 2000. *Pattern grammar*: *A corpus-driven approach to the lexical grammar of English*. Amsterdam, Philadelphia, PA: John Benjamins Publishing Company.

Ivanič, Roz. 1991. Nouns in search of a context: A study of nouns with both open- and closed-system characteristics. *International Review of Applied Linguistics in Language Teaching* 29, 93–114.

Jespersen, Otto. 1927. *A modern English grammar on historical principles. Part III: Syntax. Second volume.* London: George Allen & Unwin.

Jespersen, Otto. 1949. *A modern English grammar on historical principles*. *Part VII: Syntax*. London: George Allen & Unwin.

Kanté, Issa. 2010a. Head nouns as modal stance markers – Academic texts vs. legal texts. *Revue de Linguistique et de Didactique des Langues* 41, 121–35.

Kanté, Issa. 2010b. Mood and modality in finite noun complement clauses: A French-English contrastive study. *International Journal of Corpus Linguistics* 15(2), 267–90.

Keizer, Evelien. 2007. *The English noun phrase: The nature of linguistic categorization*. Cambridge: Cambridge University Press.

Kolhatkar, Varada, Heike Zinsmeister & Graeme Hirst. 2013. Annotating anaphoric shell nouns with their antecedents. In *Proceedings of the 7th Linguistic Annotation Workshop & Interoperability with Discourse, August 8-9, 2013, Sofia, Bulgaria*, 112–21. Stroudsburg, PA: The Association for Computational Linguistics.

Kruisinga, Etsko. 1932. *A handbook of present-day English*. *Part II*: *English Accidence and Syntax 2*, 5th edn. Groningen: P. Noordhoff.

Kruisinga, Etsko. 1932. *A handbook of present-day English*. *Part II*: *English Accidence and Syntax 3*, 5th edn. Groningen: P. Noordhoff.

Langacker, Ronald W. 1987. Nouns and verbs. *Language* 63(1), 53–94.

Lee, David Y.W. 2001. Genres, registers, text types, domains, and styles: Clarifying the concepts and navigating a path through the BNC jungle. *Language Learning & Technology* 5(3), 37–72.

Levi, Judith N. 1978. *The syntax and semantics of complex nominals*. New York, NY: Academic Press.

Levinson, Stephen C. 1983. *Pragmatics*. Cambridge: Cambridge University Press.

Lin, Ming-chia. 2012. Shell nouns on the move: Expert and L2 student abstracts in applied linguistics. *Arab World English Journal* 3, 223–45.

Lorés, Rosa. 2006. The referential function of metadiscourse: Thing(s) and idea(s) in academic lectures. In Ana María Hornero, María José Luzón

& Silvia Murillo (eds.), *Corpus linguistics*: *Applications for the study of English*, 315–33. Bern: Peter Lang.

Lyons, John. 1977. *Semantics* (2 vols.). Cambridge: Cambridge University Press.

Mahlberg, Michaela. 2003. The textlinguistic dimension of corpus linguistics: The support function of English general nouns and its theoretical implications. *International Journal of Corpus Linguistics* 8(1), 97–108.

Mahlberg, Michaela. 2005. *English general nouns*: *A corpus theoretical approach*. Amsterdam, Philadelphia, PA: John Benjamins Publishing Company.

Martin, James R. 1992. *English text: System and structure*. Amsterdam, Philadelphia, PA: John Benjamins Publishing Company.

Martin, James R. 1997. Analysing genres: Functional parameters. In Frances Christie & James R. Martin (eds.), *Genre and institutions*: *Social processes in the workplace and school*, 3–39. London: Cassell.

Martin, James R. 2000. Beyond exchange: APPRAISAL systems in English. In Susan Hunston & Geoff Thompson (eds.), *Evaluation in text*: *Authorial stance and the construction of discourse*, 142–75. Oxford: Oxford University Press.

Martin, James R., Christian M.I.M. Matthiessen & Clare Painter. 1997. *Working with functional grammar*. London, New York, NY: Arnold.

Matthiessen, Christian M.I.M. 1995. *Lexicogrammatical cartography*. Tokyo: International Language Sciences Publishers.

Matthiessen, Christian M.I.M. 1999. The system of TRANSITIVITY: An exploratory study of text-based profiles. *Functions of Language* 6(1), 1–51.

Matthiessen, Christian M.I.M. 2006. Frequency profiles of some basic grammatical systems: An interim report. In Geoff Thompson & Susan Hunston (eds.), *System and corpus: Exploring connections*, 103–42. London, Oakville: Equinox.

Mauranen, Anna. 1993. Theme and prospection in written discourse. In Baker, Francis & Tognini-Bonelli (eds.), 95–114.

McCarthy, Michael. 1984. *Some pragmatic features of written text*. Birmingham: English Language Research, University of Birmingham.

McEnery, Tony & Andrew Hardie. 2012. *Corpus linguistics*: *Methods, theory and practice*. Cambridge: Cambridge University Press.

McIntosh, Colin, Ben Francis & Richard Poole (eds.). 2009. *Oxford collocations dictionary for students of English*, 2nd edn (CD-ROM). Oxford: Oxford University Press.

Melrose, Robin. 2003. 'Having things both ways': Grammatical metaphor in a systemic-functional model of language. In Simon-Vandenbergen, Taverniers & Ravelli (eds.), 417–42.

Mihatsch, Wiltrud. 2009. Nouns are THINGS: Evidence for a grammatical metaphor?. In Klaus-Uwe Panther, Linda L. Thornburg & Antonio Barcelona (eds.), *Metonymy and metaphor in grammar*, 75–97. Amsterdam, Philadelphia, PA: John Benjamins Publishing Company.

Moreno, Ana I. 2004. Retrospective labelling in premise-conclusion metatext: An English-Spanish contrastive study of research articles on business and economics. *Journal of English for Academic Purposes* 3(4), 321–39.

Nelson, Gerald, Sean Wallis & Bas Aarts. 2002. *Exploring natural language*: *Working with the British component of the International Corpus of English*. Amsterdam, Philadelphia, PA: John Benjamins Publishing Company.

Nesi, Hilary & Emma Moreton. 2012. EFL/ESL writers and the use of shell nouns. In Ramona Tang (ed.), *Academic writing in a second or foreign language*: *Issues and challenges facing ESL/EFL academic writers in higher education contexts*, 126–45. London: Continuum.

O'Donnell, Michael, Michele Zappavigna-Lee & Casey Whitelaw. 2008. A survey of process type classification over difficult cases. In Carys Jones & Eija Ventola (eds.), *From language to multimodality*: *New developments in the study of ideational meaning*, 47–64. London, Oakville: Equinox.

Oxford English dictionary, 2nd edn, 1989; online version. http://www.oed.com/ (accessed 22 June 2013)

Paradis, Carita. 2004. Towards a theory of lexical meaning as ontologies and construals. *The Department of English in Lund University: Working Papers in English Linguistics* 4, 51–79.

Partington, Alan. 1998. *Patterns and meanings*: *Using corpora for English language research and teaching*. Amsterdam, Philadelphia, PA: John Benjamins Publishing Company.

Payne, John & Rodney Huddleston. 2002. Nouns and noun phrases. In Huddleston & Pullum (eds.), 323–523.

Petch-Tyson, Stephanie. 2000. Demonstrative expressions in argumentative discourse: A computer corpus-based comparison of non-native and native English. In Simon Botley & Tony McEnery (eds.), *Corpus-based and computational approaches to discourse anaphora*, 43–64. Amsterdam, Philadelphia, PA: John Benjamins Publishing Company.

Poutsma, Hendrik. 1905. *A grammar of late modern English*: *For the use of continental, especially Dutch students*. *Part I*: *The sentence. Section II*: *The composite sentence*. Groningen: P. Noordhoff.

Prince, Ellen F. 1981. Toxard a taxonomy of given-new information. In Peter Cole (ed.), *Radical pragmatics*, 223–55. London, New York, NY: Academic Press.

Pullum, Geoffrey K. & Rodney Huddleston. 2002. Adjectives and adverbs. In Huddleston & Pullum (eds.), 525–95.

Quirk, Randolph, Sidney Greenbaum, Geoffrey Leech & Jan Svartvik. 1985. *A comprehensive grammar of the English language*. London, New York, NY: Longman.

Ravelli, Louise J. 2003. Renewal of connection: Integrating theory and practice in an understanding of grammatical metaphor. In Simon-Vandenbergen, Taverniers & Ravelli (eds.), 37–64.

Renouf, Antoinette & John, M. Sinclair. 1991. Collocational frameworks in English. In Karin Aijmer & Bengt Altenberg (eds.), *English corpus linguistics: Studies in honor of Jan Svartvik*, 128–43. London, New York, NY: Longman.

Schleppegrell, Mary J. 2004. Technical writing in a second language: The role of grammatical metaphor. In Louise J. Ravelli & Robert A. Ellis (eds.), *Analysing academic writing: Contextualized frameworks*, 172–89. London, New York, NY: Continuum.

Schmid, Hans-Jörg. 1997. The historical development and present-day use of the noun *idea* as documented in the OED and other corpora. *Poetica* 47, 87–128.

Schmid, Hans-Jörg. 1999. Towards a functional-cognitive lexicology of nouns. In Wolfgang Falkner & Hans-Jörg Schmid (eds.), *Words, lexemes, concepts – Approaches to the lexicon. Studies in honour of Leonhard Lipka*, 213–26. Tübingen: Gunter Narr.

Schmid, Hans-Jörg. 2000. *English abstract nouns as conceptual shells*: *From corpus to cognition*. Berlin, New York, NY: Mouton de Gruyter.

Schmid, Hans-Jörg. 2001. Presupposition can be a bluff: How abstract nouns can be used as presupposition triggers. *Journal of Pragmatics* 33, 1529–52.

Schmid, Hans-Jörg. 2007. Non-compositionality and emergent meaning of lexico-grammatical chunks: A corpus study of noun phrases with sentential complements as constructions. *ZAA* 55(3), 313–40.

Schmid, Hans-Jörg. 2010. Does frequency in text instantiate entrenchment in the cognitive system? In Dylan Glynn & Kerstin Fischer (eds.), *Quantitative methods in cognitive semantics: Corpus-driven approaches*, 101–33. Berlin, New York, NY: De Gruyter Mouton.

Simon-Vandenbergen, Anne-Marie, Miriam Taverniers & Louise J. Ravelli (eds.). 2003. *Grammatical metaphor: Views from systemic functional linguistics*. Amsterdam, Philadelphia, PA: John Benjamins Publishing Company.

Sinclair, John M. 1991. *Corpus, concordance, collocation*. Oxford: Oxford University Press.

Sinclair, John M. 1993. Written discourse structure. In Sinclair, Hoey & Fox (eds.), 6–31.

Sinclair, John M. 2004. *Trust the text: Language, corpus and discourse* (Edited with Ronald Carter). London: Routledge.

Sinclair, John M. (ed.). 2006. *Collins Cobuild advanced learners' English dictionary*, 5th edn (CD-ROM). Glasgow: HarperCollins Publishers.

Sinclair, John M., Gwyneth Fox, Stephen Bullon, Ramesh Krishnamurthy, Elizabeth Manning & John Todd. 1990. *Collins Cobuild English grammar*. London: HaperCollins Publishers.

Sinclair, John M., Michael Hoey & Gwyneth Fox (eds.). 1993. *Techniques of description: Spoken and written discourse, a festschrift for Malcolm Coulthard*. London, New York, NY: Routledge.

Stirling, Lesley & Rodney Huddleston. 2002. Deixis and anaphora. In Huddleston & Pullum (eds.), 1449–1564.

Swales, John M. 1990. *Genre analysis: English in academic and research settings*. Cambridge: Cambridge University Press.

Sweet, Henry. 1891. *A new English grammar logical and historical*. *Part I: Introduction, phonology and accidence*. Oxford: Oxford University Press.

Tadros, Angele. 1985. *Prediction in text* (Discourse Analysis Monograph 10). Birmingham: English Language Research, University of Birmingham.

Tadros, Angele. 1994. Predictive categories in expository text. In Coulthard (ed.), 69–82.

Thompson, Paul. Forthcoming. Exploring Hoey's notion of *textual colligation* in a corpus of student writing. In Alejandro Alcaraz Sintes & Salvador Valera Hernández (eds.), *Diachrony and synchrony in English corpus studies*, 349–73. Frankfurt am Main: Peter Lang.

Tognini-Bonelli, Elena. 2001. *Corpus linguistics at work*. Amsterdam, Philadelphia, PA: John Benjamins Publishing Company.

van Dijk, Teun A. 1977. *Text and context: Explorations in the semantics and pragmatics of discourse*. London, New York, NY: Longman.

van Dijk, Teun A. 1997. Discourse as interaction in society. In Teun A. van Dijk (ed.), *Discourse as social interaction, Volume 2*, 1–37. London: Sage.

van Leeuwen, Teun. 2006. Critical discourse analysis. In Keith Brown (ed.), *Encyclopedia of language and linguistics, Volume 3*, 290–4. Oxford: Elsevier.

van Leeuwen, Teun. 2008. *Discourse and practice. New tools for critical discourse analysis*. Oxford: Oxford University Press.

Vendler, Zeno. 1968. *Adjectives and nominalizations*. The Hague: Mouton.
Winter, Eugene O. 1977. A clause-relational approach to English texts: A study of some predictive lexical items in written discourse. *Instructional Science* 6, 1–92.
Winter, Eugene O. 1982. *Towards a contextual grammar of English*. London: George Allen & Unwin.
Winter, Eugene O. 1992. The notion of unspecific versus specific as one way of analysing the information of a fund-raising letter. In William C. Mann & Sandra A. Thompson (eds.), *Discourse description*: *Diverse linguistic analyses of a fund-raising text*, 131–70. Amsterdam, Philadelphia, PA: John Benjamins Publishing Company.
Wodak, Ruth & Michael Meyer. 2009. Critical discourse analysis: History, agenda, theory and methodology. In Ruth Wodak & Michael Meyer (eds.), *Methods of critical discourse analysis*, 1–33. London: Sage.
Yamasaki, Nozomi. 2008. Collocations and colligations associated with discourse functions of unspecific anaphoric nouns. *International Journal of Corpus Linguistics* 13(1), 75–98.
Zandvoort, Reinard W. 1972. *A handbook of English grammar*, 6th edn. London: Longman.

SPANISH SUMMARY OF THE THESIS/ RESÚMEN EXTENSO DE LA TESIS EN ESPAÑOL (MENCIÓN INTERNACIONAL)

1 INTRODUCCIÓN

1.1 *Investigación previa*[42]

La investigación de la forma y uso de nombres como *fact*, *idea* o *warning* ha sido objeto de especial atención durante las últimas cuatro décadas. Esto se refleja en la variedad de términos que se han propuesto con el fin d enfatizar diferentes aspectos de su comportamiento: 'general nouns' (Halliday & Hasan 1976, Mahlberg 2005), 'Vocabulary 3 items' (Winter 1977), 'lexical signalling' (Hoey 1979), 'enumerables' y 'advance labelling' (Tadros 1985, 1994), 'anaphoric nouns' (Francis 1986), 'carrier nouns' (Ivanič 1991), 'advance' y 'retrospective labels' (Francis 1994), 'shell nouns' (Hunston & Francis 2000, Schmid 2000) y 'signalling nouns' (Flowerdew 2003a). Todos estos conceptos comparten una serie de características, que pueden resumirse en cinco ideas principales:

i) Abstracción: La descripción de los nombres cápsula se relaciona estrechamente con la noción de abstracción, cuya identificación depende de criterios formales y semánticos. Vendler (1968) es un claro ejemplo de la aplicación de criterios formales, dado que su clasificación ontológica de unidades nominales (eventos o hechos) se basa en su aparición en patrones estructurales pre-establecidos, los llamados 'containers' (Vendler 1968: 33). Uno de esos patrones es 'N *is* N', en el cuál el nombre cápsula aparece como complemento del verbo copulativo *be*, como en el ejemplo (1). La descripción que Winter (1977: 9, 85) hace de las unidades 'Vocabulary 3' es también formal, como se refleja en su afirmación de que sólo nombres concretos prototípicos (e.g. *animal*) serían adecuados en un ejemplo como (2).

[42] En este resumen, al igual que en el proyecto de tesis, se utiliza el término español 'nombre cápsula' en lugar del término inglés 'shell noun'.

(1) 'That he died/His death' is a **fact** (Vendler 1968: 73)
(2) 'It is like **an effect** [...]' (Winter 1977: 85)

Lyons (1977, II: 442–7) utiliza un enfoque más semántico, ya que no establece ninguna conexión entre la naturaleza ontológica de las unidades nominales y patrones oracionales típicos. La clasificación de nombres que plantea este autor consta de entidades de primer orden, de segundo orden y de tercer orden. Las entidades 'de primer orden' se corresponden con nombres concretos prototípicos (e.g. *table*, *spoon*), las 'de segundo orden', con nominalizaciones de procesos y eventos (e.g. *activity*, *destruction*), y las 'de tercer orden', con hechos y proposiciones (e.g. *issue*, *argument*). Según Schmid (2000: 68), en una escala de prototipicalidad, las entidades de tercer orden constituyen el núcleo de la clase de nombres cápsula.

ii) Unidades léxicas y gramaticales, significado inherente y contextual: Otro rasgo distintivo de estas unidades es el de la clasificación a medio camino entre unidades léxicas o abiertas y unidades gramaticales o cerradas (e.g. Halliday & Hasan 1976: 275; Winter 1977: 2; Francis 1986: 3; Ivanič 1991: 103). Su estatus abierto proviene de su significado inherente, mientras que su estatus cerrado se relaciona con su significado contextual (Ivanič 1991: 95). Los nombres cápsula, por lo tanto, se asocian con un significado que permanece constante, y uno que, al igual que con los pronombres, varía dependiendo del contexto en el que el nombre aparece. Este significado variable puede explicarse en función de la 'deixis textual impura' de Lyons (1977: 668) y la noción de 'referencia situacional' de Fraurud (1992: 4). Ambos términos se usan para referirse a algunos casos de referencia anafórica en los cuales el pronombre no es co-referencial con una entidad de primer orden, sino con segmentos discursivos previos. El ejemplo (3) ilustra el tipo más concreto de anáfora, mientras que en (4a) y (4b), a pesar de expresar la misma información, (4b) utiliza un pronombre demostrativo y (4a) un sintagma nominal cápsula. El uso de un nombre cápsula en (4a) permite al escritor evaluar el segmento discursivo subrayado. Es precisamente el potencial descriptivo y caracterizador de los nombres cápsula lo que distingue a estas unidades de los pronombres. Por consiguiente, aunque similares a los pronombres en su significado contextual, el potencial evaluador de los nombres cápsula es claramente léxico o abierto.

(3) 'The boys went home. **They** were tired' (Fraurud 1992: 3)
(4) (a)'[...] foul-smelling algae, the product of exceptionally high temperatures and high levels of sea pollution, which led to a huge drop in bookings. Fortunately **this problem** does not seem to have recurred this summer' (Schmid 2000: 124)
(b)'[...] foul-smelling algae, the product of exceptionally high temperatures and high levels of sea pollution, which led to a huge drop in bookings. Fortunately **this** does not seem to have recurred this summer'

iii) Antecedente largo: En la literatura, se suele asociar a estos nombres con la encapsulación de cláusulas, oraciones o segmentos discursivos de mayor longitud (como en el ejemplo (5)). Por lo tanto, en principio, la encapsulación de palabras no garantiza la propiedad de '[...] shell-nounhood' (Schmid 2000: 13). Algunas referencias, sin embargo, contemplan la posibilidad de antecedentes nominales para nombres cápsula, como se refleja en el ejemplo (6) (Ivanič 1991: 109; Flowerdew 2003a: 336; Gray 2010: 179):

(5) '[...] In reply to that question a golfing colleague of mine offered **two reasons**. The first was that beginners usually start with handed-down clubs, which are usually right-handed. The second was that, for technical reasons, left-handed individuals make good right-handed golfers.' (Francis 1994: 84)
(6) 'It is interesting to read about **the items electors mentioned as having, in their view, specially affected the election**. […] rash Labour promises – cost of new pension scheme – bribery of electorate [...] strikes' (Ivanič 1991: 109)

iv) Anáfora y realización inter-oracional: Los usos anafóricos, como en (4) arriba, ocupan una posición importante en la literatura (e.g. Halliday & Hasan 1976; Francis 1986; Conte 1996; Charles 2003; Moreno 2004; Gray 2010), lo cuál se desprende de las muchas referencias sobre la resolución de la anáfora pronominal (e.g. Chomsky 1981, Fox 1987, Fraurud 1988, Asher 1993). Otros tipos de encapsulación, sin embargo, son menos frecuentes en la investigación sobre nombres cápsula (e.g. Winter 1977, Hunston & Francis 2000 sobre los usos catafóricos de los nombres cápsula; Francis 1994, Schmid 2000 sobre anáfora y catáfora; Ivanič 1991, Flowerdew 2003a sobre usos endofóricos y exofóricos).

Dada la importancia que se ha venido dando a la encapsulación retrospectiva, no es de extrañar que la realización inter-oracional de los nombres cápsula sea tan frecuente en la literatura, como en

(4) arriba o (7) debajo. Son pocos los estudios que hacen referencia a la encapsulación intra-oracional (e.g. Hunston & Francis 2000; Biber 2006; Charles 2007, como en (8)), o a la encapsulación intra- e inter-oracional (e.g. Ivanič 1991; Winter 1992; Schmid 2000; Flowerdew 2003a; Caldwell 2009).

(7) '[…] the Soviet Union has 'shot its bolt', and that only the unreconstructed Cold Warriors are losing any sleep about the Russian menace. James Reston has readily and complacently echoed this assessment in his criticism of the Reagan équipe' (Francis 1986: 27)
(8) '[...] the best the White House has been able to conjure up is the tired accusation that they are liberals and lackeys of special interest groups.' (Schmid 2000: 135)

v) Sintagmas nominales cápsula específicos y los patrones N-cl y N-*be*-cl: En cuanto a estructura formal, los nombres cápsula suelen relacionarse con sintagmas definidos y demostrativos, como en (4), (6), (7) y (8) arriba. Son pocos los estudios en los que se hace referencia a sintagmas indefinidos (e.g. Ivanič 1991: 111; Partington 1998: 92–3; Aktas & Cortes 2008: 10), como en (9):

(9) 'In a move to tighten control of a far-reaching empire and to improve the group's own image, Maurice and Charles Saatchi, credited with building up the company, have stepped down from the day-to-day running of the group' (Partington 1998: 94)

Se presta atención considerable también a dos estructuras de post-modificación, i.e. las cláusulas nominales de complemento de *that* y *to* infinitivo (Biber et al. 1999: 645), como en (10) y (11) debajo. Los sintagmas preposicionales, sin embargo, tienden a omitirse en la gran mayoría de las descripciones disponibles. Winter (1992: 157), Flowerdew (2003a: 337) y Caldwell (2009: 176) son algunas excepciones (ver (12)).

(10) 'The Association will give a warning that poll tax bills in some Conservative districts will exceed government guidelines […]' (Schmid 2000: 135)
(11) 'The first action was to place the vessel under cover and remove the deckhouse' (Schmid 2000: 263)
(12) 'its function of providing mechanical strength' (Flowerdew 2003a: 337)

La gran mayoría de las generalizaciones sobre el comportamiento de los nombres cápsula se basa en resultados obtenidos de corpus pequeños y

sobre un género específico. El discurso académico es primordial en este respecto, ya que se ha hecho mucha investigación sobre el análisis de sub-géneros académicos como libros de texto, artículos de investigación o ensayos (e.g. Tadros 1985; Flowerdew 2003a; Moreno 2004; Charles 2007, etc.). La prosa periodística ocupa también una posición destacada en las descripciones sobre nombres cápsula, especialmente en aquellos casos en los que la recuperación de datos se basa en corpus de inglés extensos como el *Bank of English* (de aquí en adelante, *BoE*; e.g. Hunston & Francis 2000; Schmid 2000). A pesar de ofrecer extensión (225 millones de palabras en Schmid 2000), el *BoE* no proporciona una representación equilibrada de géneros, puesto que el 70% del corpus se corresponde con prosa periodística. El uso de otros corpus generales del inglés como el *Lancaster-Oslo Bergen Corpus* (de aquí en adelante, *LOB*) o el *British National Corpus* (de aquí en adelante, *BNC*) ofrece más equilibrio, pero, en estos casos, la investigación se centra sólo en resultados cualitativos (e.g. Ivanič 1991) o en datos cuantitativos de modo (i.e. tan sólo oral frente a escrito; e.g. Aijmer 2007, Yamasaki 2008).

El análisis de datos en la literatura suele basarse en la recuperación automática de patrones pre-definidos, en particular, N-cl y N-*be*-cl, los cuáles, como se menciona arriba, se asocian prototípicamente a los nombres cápsula (e.g. Francis 1993; Hunston & Francis 2000; Schmid 2000 and 2007, etc.). Los enfoques más manuales al análisis de datos de corpus identifican otros patrones (e.g. N-*of*, N-*which*, etc.). Sin embargo, el énfasis recae en ciertos sub-géneros y nombres en particular (e.g. Francis 1986 sobre el periódico mensual *Encounter*, Flowerdew 2003a sobre libros de textos y clases de biología, Hoey 1993 sobre el nombre *reason* y Lorés 2006 sobre *thing* y *idea*).

1.2 *La tesis*

1.2.1 *Justificación*

Como puede intuirse de la revisión en 1.1, los nombres cápsula constituyen un área ampliamente estudiada. A pesar de esto, la literatura revela ciertas lagunas, dado que:

i) Se ha prestado mucha atención a la prosa académica y periodística, y muy poca a otros géneros.
ii) Son muy pocos los estudios que proporcionan una descripción sistemática y detallada del uso de los nombres cápsula en diferentes áreas de investigación lingüística.

iii) El análisis de datos suele limitarse a patrones y relaciones encapsuladoras específicas (i.e. N-cl y N-*be*-cl; anáfora).
iv) Los enfoques analíticos automáticos prevalecen sobre los manuales.
v) Los análisis de corpus pequeños suelen hacerse sólo con géneros específicos.

De esto se desprende la necesidad de un estudio en el cuál:

i) El uso de los nombres cápsula se describa en función de una muestra pequeña pero representativa de la lengua inglesa en general.
ii) Se investiguen niveles formales, sintácticos, semánticos y textuales de análisis lingüístico.
iii) Se incluyan todos los patrones y usos.

1.2.2 *Método*

Esta tesis adopta una metodología 'corpus-driven' (a diferencia de 'corpus-based'; véase Tognini-Bonelli 2001: 67–71) para el análisis de datos sobre nombres cápsula. Aunque un análisis completamente automático puede aparecer beneficioso en su rápido procesamiento de grandes cantidades de datos, el ámbito de investigación puede quedarse limitado a causa de la naturaleza restrictiva de las búsquedas automatizadas, las cuáles tienden a forzar al investigador a ignorar ciertas características lingüísticas. Siguiendo a Sinclair (2004: 23), este estudio propone un enfoque abierto al análisis de datos lingüísticos, en el que no se impongan ideas preconcebidas (e.g. nombres cápsula y cláusulas nominales de complemento), en el que, en pocas palabras, se confíe en el texto. Como en el enfoque teórico de corpus que Mahlberg (2005: 31–8) utiliza para la descripción de los 'general nouns' o nombres generales (e.g. *thing*, *fact*, *people*, *world*), en este estudio tan sólo dos suposiciones mínimas guían el análisis. Una de esas suposiciones relaciona a los nombres cápsula con entidades de segundo y tercer orden con una limitada especificidad semántica (e.g. *action*, *event*, *idea*, *point*). La segunda suposición enfatiza la interpretación contextual de estas unidades, independientemente de la forma del sintagma nominal (definido o indefinido). Por lo tanto, la interpretación de cualquier nombre cápsula viene siempre influida por el contexto en el que aparece. Además de Sinclair (2004) y Mahlberg (2005), esta tesis también se inspira en la teoría del 'lexical priming' o priming léxico de Hoey (2005), ya que el análisis contempla todos los niveles de análisis lingüístico, i.e. formal/estructural, sintáctico, semántico y textual.

El estudio utiliza un pequeño corpus de inglés británico contemporáneo oral y escrito, el *BNC Sampler*. El corpus consta de 2 millones de palabras, y constituye una muestra de un 2% del *BNC* completo (con 100 millones de palabras). A pesar de su tamaño, el *Sampler* es más equilibrado que el *BNC*, dado que, a diferencia del *BNC*, dónde el 90% es inglés escrito y el 10% inglés hablado, los géneros orales y escritos en el *Sampler* se distribuyen de un modo más uniforme (50%-50%).

La muestra consta de 1447 concordancias correspondientes a 60 lemas. Las unidades provienen de los rangos de frecuencia altos, medios y bajos de una lista de 922 nombres cápsula. La lista reúne a 15 listas de unidades cápsula encontradas en la literatura (siendo la más extensa la de Schmid 2000, con 670 unidades). Para cada uno de los 60 lemas, se analiza un conjunto aleatorio de 40 concordancias (20 para el singular y 20 para el plural).

Cada línea se anota en base a nueve variables, lo cual refleja la perspectiva multidimensional que se aplica en esta tesis:

i) Género del texto (e.g. prosa académica, conversación, etc.).
ii) Estructura semántica del sintagma nominal (e.g. Deíctico, Epíteto, etc.)
iii) Estructura formal (e.g. artículo definido, sintagma preposicional, etc.).
iv) Función sintáctica (e.g. objeto directo, sujeto).
v) Tipo de participante (e.g. Atributo, Identificador).
vi) Tema-Rema.
vii) Dirección de la encapsulación (e.g. anáfora inter-oracional, catáfora intra-oracional, etc.).
viii) Antecedente (e.g. discurso extenso, oración, etc.).
ix) Tipo semántico de nombre cápsula (e.g. factual, mental, etc.).

1.2.3 *Objetivos*

Esta tesis tiene tres objetivos generales y cuatro específicos.

Los tres objetivos generales son:

i) Identificar los criterios léxico-gramaticales, sintáctico-semánticos y textuales relacionados con el uso de los nombres cápsula desde una perspectiva sincrónica.
ii) Examinar como los criterios se interrelacionan unos con otros.
iii) Investigar la conexión entre cada criterio y grupos de unidades cápsula.

Los objetivos específicos son:

i) Analizar manualmente una muestra de 1447 concordancias para 60 lemas extraídos del *BNC Sampler*.
ii) Identificar criterios que distingan a nombres cápsula de aquellos que no lo son.
iii) Codificar ejemplos de nombres cápsula en función de nueve variables léxico-gramaticales, sintáctico-semánticas y textuales.
iv) Establecer el grado en el que estas unidades contribuyen a la organización de diferentes tipos de discurso hablado y escrito.

1.2.4 *Contenidos de la tesis*

Esta tesis se divide en siete capítulos, incluyendo la introducción (capítulo 1). El capítulo 2 profundiza en el marco teórico de la descripción de los nombres cápsula, con especial atención a definiciones, criterios de identificación y variación con respecto a género. El capítulo 3 sienta las bases teóricas para el enfoque analítico de esta tesis. El capítulo arroja luz sobre las lagunas que inspiraron la elección del corpus, de las variables y de los métodos. El capítulo 4 explica las decisiones sobre el corpus y la muestra de estudio, y describe el enfoque teórico de corpus manual aplicado en esta tesis. El capítulo 5 detalla los principios que subyacen cada una de las nueve variables analíticas, combinando procedimientos metodológicos con observaciones cualitativas sobre algunas de las variables. El capítulo 6 presenta y discute los resultados cuantitativos para cada variable por separado, y en relación unas con otras. El capítulo 7 resume los contenidos de esta tesis y propone futuras vías de investigación.

2 METODOLOGÍA DE CORPUS Y PROCEDIMIENTOS ANALÍTICOS EN LA INVESTIGACIÓN SOBRE NOMBRES CÁPSULA

Esta sección trata sobre la metodología de corpus que se ha seguido en la literatura sobre nombres cápsula. Se hace referencia tanto a los corpus como a los procedimientos analíticos utilizados.

La mayor parte de las generalizaciones sobre el uso de nombres cápsula resulta del análisis de corpus pequeños y restringidos a géneros específicos. La prosa académica, prominente en este tipo de investigación, se ha estudiado en corpus de textos profesionales disciplinares (como en libros de texto, artículos científicos y tesis: e.g. Tadros 1985; Moreno 2004; Charles 2007; Gray 2010), ensayos argumentativos producidos por estudiantes universitarios no-nativos (e.g. Francis 1988; Flowerdew 2006; Hasselgård 2012), textos nativos (normalmente, profesionales) frente a

textos no-nativos (e.g. Aktas & Cortes 2008; Caldwell 2009), y clases de universidad (e.g. Flowerdew 2003a; Lorés 2006). Otro género que destaca en la literatura es la prosa periodística, como se refleja en el uso que Francis (1986 y 1994) hace de artículos de los periódicos *The Times* y *Encounter*. Cabe mencionar también que, aunque la investigación hasta el momento se ha centrado en corpus sincrónicos, hay casos como Schmid (1997) y Kanté (2010a), donde se hace uso de evidencia de corpus histórica o diacrónica.

Comparado con la cantidad de estudios sobre géneros específicos, hasta el momento, la investigación usando corpus generales del inglés ha sido escasa. Los tres corpus que se han empleado en estos estudios son el *LOB Corpus* (Ivanič 1991), el *BoE* (a lo largo de sus diferentes etapas de compilación: Hoey 1993; Hunston & Francis 2000; Schmid 2000; Mahlberg 2005) y el *BNC* (Aijmer 2007; Yamasaki 2008). *LOB*, con alrededor de 1 millón de palabras de inglés escrito, es el más pequeño de los tres corpus. Ivanič (1991) utiliza este corpus como una fuente de ejemplos, pero no ofrece ningún resultado cuantitativo acerca de la distribución de unidades y patrones con respecto a género. El *BoE*, cuya composición asciende a 450 millones de palabras en Mahlberg (2005), es el más extenso de los corpus generales del inglés disponibles. Sin embargo, como reconoce Schmid (2000: 43), este corpus ofrece tamaño, pero no equilibrio, dado que la prosa periodística abarca el 70% del corpus. Finalmente, el *BNC*, con alrededor de 100 millones de palabras, no es tan extenso como el *BoE*, pero en cuanto a distribución, es más equilibrado que el *BoE* (a pesar de la proporción de 90% frente a 10% de lengua escrita y oral). Los dos estudios que usan el *BNC*, no obstante, no incluyen resultados de género, ya que se presta atención sólo a modo (i.e. lenguaje oral: Aijmer 2007, y lenguaje oral frente a escrito: Yamasaki 2008).

De lo anterior se desprende que la elección de un corpus equilibrado puede jugar un papel fundamental en los resultados obtenidos. El uso de un corpus pequeño pero restringido a un género específico puede ayudar a identificar el comportamiento de los nombres cápsula en un tipo específico de texto oral o escrito, pero, al mismo tiempo, a menos que se comparen los resultados con un corpus más equilibrado, no deberían hacerse generalizaciones acerca de ese comportamiento en la lengua inglesa en general.

Además del corpus, otro aspecto a tener en cuenta en la investigación sobre nombres cápsula es el procedimiento analítico utilizado en el análisis del corpus. La literatura consultada revela tres tipos de análisis, i.e. análisis automatizados de corpus extensos, análisis semi-automatizados de corpus extensos o pequeños, y análisis manuales de corpus pequeños. Los análisis de corpus extensos y pequeños forman parte del enfoque

tridimensional de análisis del discurso a través de corpus que propone Bednarek (2009: 20–2). El tercer método, correspondiente al análisis de textos individuales, suele utilizarse en el análisis crítico del discurso.

El análisis automatizado de corpus extensos (e.g. Francis 1993; Biber et al. 1999; Hunston & Francis 2000; Schmid 2000 y 2007) se realiza mediante la recuperación automática de ejemplos a través de búsquedas pre-definidas de patrones, siendo los más comunes N-cl y N-*be*-cl. A pesar de la abundancia de datos sobre estos patrones, la rigidez de las búsquedas de corpus conlleva la exclusión de muchos ejemplos potenciales de nombres cápsula.

Un menor grado de automatización se refleja en los análisis semi-automatizados de corpus extensos o pequeños, en los cuales las búsquedas automáticas van seguidas de análisis manuales. Yamasaki (2008) y Caldwell (2009) ilustran este enfoque. Yamasaki (2008) utiliza el *BNC* con el fin de explorar el potencial evaluativo de los patrones *th*-N (e.g. *this problem*) y *th-be*-N (e.g. *this is a problem*). La decisión de centrarse en ambos patrones sigue a la recuperación y posterior análisis manual de un grupo aleatorio de concordancias correspondientes a 73 nombres cápsula. Tras la búsqueda automática de estos patrones con cinco nombres (i.e. *change*, *shift*, *failure*, *mistake* y *problem*), el análisis manual posterior identifica el tipo de pre-modificación y distribución de modo (lenguaje oral frente a escrito) de los nombres seleccionados. Caldwell (2009: 78) lleva a cabo búsquedas sintácticas de patrones cápsula típicos, así como de otros que se suelen excluir de análisis automatizados (e.g. nombre + sintagma preposicional), con el fin de comparar su frecuencia en prosa nativa y no-nativa no-profesional (a nivel de grado) y prosa nativa profesional (artículos científicos). Un análisis manual posterior establece el tipo de referencia que los nombres cápsula definidos muestran en las primeras 200 palabras de cada uno de los textos de los que constan los tres sub-corpus.

El grado de automatización es mínimo en los análisis manuales de corpus pequeños. En estos estudios, se hace uso de muestras pequeñas de géneros específicos, y se prefiere el análisis manual e intensivo frente a la búsqueda de patrones predefinidos. El foco de investigación de las referencias consultadas (e.g. Francis 1986; Hoey 1993; Flowerdew 2003a; Mahlberg 2005; Lorés 2006), sin embargo, es demasiado general o demasiado específico. Por ejemplo, Francis (1986) y Flowerdew (2003a) incluyen una gran variedad de nombres en su análisis, pero la atención recae principalmente en los aspectos discursivos y semánticos de estas unidades. Los datos (que se presentan en forma de generalizaciones) tienden a no estar acompañados de resultados cuantitativos. Además, mientras que Francis (1986) investiga prosa periodística en un periódico en particular, Flowerdew (2003a) examina un corpus de libros de texto y

clases de biología. Mahlberg (2005: 59) usa el *BoE* para describir las funciones textuales locales de 20 nombres generales. Estas funciones representan los significados que estos nombres adquieren cuando se utilizan en contexto (e.g. *time* con el significado de 'medida', 'historia', 'vida', etc.). Aunque hay algunos nombres generales que también son nombres cápsula (e.g. *thing*, *fact*, *system*, *problem*), hay muchos otros que se comportan como nombres cápsula solo con respecto a su falta de especificidad semántica, pero no con relación a su referencia, puesto que tan sólo los nombres cápsula pueden referirse a segmentos discursivos. Finalmente, Hoey (1993) y Lorés (2006), a diferencia de otros estudios manuales, ofrecen resultados cuantitativos, pero estos se restringen al nombre *reason* en Hoey (1993) y a *thing(s)* y *idea(s)* en Lorés (2006).

Un análisis de datos de corpus enteramente manual es importante en ciertas áreas de análisis del discurso, especialmente en aquellas en las que se pretende profundizar en nuestro conocimiento sobre el comportamiento lingüístico y discursivo general de algunas unidades. Este enfoque tiene limitaciones, ya que la lectura intensiva de segmentos de discurso extensos fuerza al investigador a reducir considerablemente el número de unidades del análisis. Un análisis enteramente automatizado tiene la ventaja de que permite el procesamiento de grandes cantidades de datos y es menos difícil para el investigador, pero, al mismo tiempo, es menos revelador con respecto al uso de patrones textuales (Mahlberg 2005: 58). En base a la perspectiva tridimensional de Bednarek (2009) para la investigación de corpus, el estudio en esta tesis se correspondería con la categoría de análisis manuales de corpus pequeños, ya que la contextualización del análisis prevalece sobre los análisis automatizados de cohesión léxica tan frecuentes en la literatura sobre nombres cápsula (Bednarek 2009: 21–2). Con casi 2 millones de palabras, el corpus que se emplea en esta tesis (i.e. el *BNC Sampler*) no es tan pequeño como los que se suelen encontrar en análisis manuales de corpus pequeños (aproximadamente 70.000-80.000 palabras), sino un corpus, que, desde la perspectiva de Flowerdew (2011), podría considerarse pequeño y grande: lo suficientemente extenso para ser representativo de la lengua inglesa, pero lo suficientemente pequeño para permitir un enfoque manual al análisis de la cohesión léxica.

3 LA TESIS: MÉTODO

3.1 *La elección del corpus*

A la hora de elegir el corpus, se plantearon cuatro posibilidades*: BNC* (100 millones de palabras), *BNC Baby* (4 millones de palabras), *BNC Sampler* (2 millones de palabras) y *ICE-GB* (1 millón de palabras). Ya que se pretendía

estudiar los nombres cápsula en una muestra representativa de la lengua inglesa, era necesario elegir el corpus más representativo y equilibrado en cuanto a géneros.

La Tabla 4.1 en 4.2.2 (tesis) muestra que, pese a diferencias en las proporciones, el *Sampler* ofrece la cobertura más completa de los géneros del *BNC*. Mientras que el *Sampler* excluye 32 de los 70 géneros en el *BNC*, dicha cifra asciende a 37 y 48 en *ICE-GB* y *Baby* respectivamente. Esto se traduciría en 12.45% de géneros del *BNC* que no están presentes en el *Sampler*, 44.64% en *ICE-GB* y 54.51% en el *BNC Baby*. De los tres corpus pequeños, *BNC Sampler,* por tanto, podría ofrecer el mayor grado de representatividad de los géneros del *BNC*, y, asimismo, de la lengua inglesa en general. Las Figuras 4.1 a la 4.4 en 4.2.2 (tesis) muestran de nuevo al *Sampler* como el corpus pequeño más representativo del *BNC*. Tan sólo 10 de los 34 macro-géneros en el *BNC* se excluyen de este corpus (cf. 16 y 30 en *ICE-GB* y *Baby* respectivamente).

En cuanto a distribución, la línea de tendencia ascendente en la Figura 4.1 (4.2.2, tesis) ilustra la mayor proporción de lenguaje escrito en el *BNC* (90%). Hay cuatro macro-géneros que predominan en este corpus: prosa académica y no-académica, ficción y textos misceláneos. La Figura 4.2 permite observar una distribución mucho más uniforme en el *Sampler* que la existente en el *BNC*, dado que la mayor parte de los géneros no sobrepasan el 5%, y aquellos en los que se sobrepasa (reuniones formales, ficción, textos misceláneos y periódicos) permanecen dentro del rango de 5%-10%. La Figura 4.3 muestra el predominio de cuatro macro-géneros en el *BNC Baby*: conversación, prosa académica, ficción y periódicos. Finalmente, la Figura 4.4 sugiere una cobertura más irregular de los macro-géneros del *BNC* en *ICE-GB* que en el *Sampler*. Mientras que el área desde S:courtroom hasta S:parliament es más uniforme que en el *Sampler*, el lenguaje escrito en *ICE-GB* contiene más áreas para las que no existe evidencia textual. La línea de tendencia descendente en *ICE-GB* indica una mayor proporción de textos orales en este corpus (60%). Aunque en el *Sampler* se observa también una línea descendente, la inclinación es más evidente en *ICE-GB*.

Teniendo en cuenta las cuatro figuras, el *Sampler* representa el corpus con la cobertura de géneros más amplia y equilibrada. Se decidió por tanto, utilizar este corpus para el estudio de la tesis. La combinación en el *Sampler* de tamaño reducido y de cobertura de una amplia variedad de géneros se ajusta perfectamente al análisis multidimensional de los nombres cápsula que se propone en esta tesis. Una posible crítica a esta decisión es que el tamaño del corpus (2 millones de palabras) no ofrece suficientes datos acerca del uso de los nombres cápsula. Teniendo también en cuenta las reservas de Lee (2001: 53–4) acerca de la representatividad

del *Sampler*, podría decirse que la escasa representatividad textual de ciertos géneros no puede conducir a conclusiones infalsificables sobre la distribución genérica de los nombres cápsula. Aunque no se niega la verdad de estas afirmaciones, es importante enfatizar que, de los corpus analizados, el *Sampler* es el que ofrece la distribución de géneros y modos más uniforme (50% escrito-50% oral). En este respecto, el *Sampler* es más adecuado que el más extenso pero meno equilibrado *BNC* y *BoE*.

3.2 *La muestra*

La muestra proviene inicialmente de una lista de frecuencia de 922 lemas extraídos de 15 listas de nombres cápsula en la literatura. La lista está dividida en tres rangos de frecuencia principales: alto (1), medio (2) y bajo (3). Cada rango de frecuencia está, a su vez, dividido en tres sub-rangos (i.e. alto: A, medio: B y bajo: C). El análisis consta de 1447 concordancias correspondientes a 60 lemas extraídos de los tres rangos principales (1, 2, 3), y luego de los sub-rangos altos y medios dentro de cada rango (i.e. A y B). Se decidió excluir los sub-rangos bajos (i.e. C) debido a la escasa evidencia de corpus para 3C. Esta tesis por tanto analiza los 10 primeros lemas de los rangos 1A, 1B, 2A, 2B, 3A y 3B (60 en total) (véase Tabla 4.4 en la Tesis para las 60 unidades y su distribución). Para cada uno de los lemas, se decidió, como regla general, analizar 20 concordancias para el singular y 20 para el plural (40 en total; véase la Tabla 4.5 para excepciones a esta regla).

3.3 *El análisis de datos*

No todas las concordancias extraídas del corpus representan usos encapsuladores. Con el fin de identificar dichos usos, cada concordancia se analiza manualmente y con respecto al contexto en el que aparece el nombre. Este tipo de análisis textual intensivo hace necesario ampliar los límites de la concordancia al contexto más extenso (Mahlberg 2005: 58).

Un simple clic sobre cualquier palabra nodo en *CQPweb* permite acceder al contexto extendido para cualquier concordancia. Aunque la realización léxica suele aparecer en el texto inmediato, el contexto inicial que da el corpus ha de ampliarse en muchas ocasiones a segmentos discursivos extensos y no contiguos para identificar la realización léxica del nombre. La lectura intensiva y detallada de textos individuales de corpus que se requiere para este tipo de análisis, según Botley (2006: 102), se contradice con los métodos de corpus típicos basados en concordancias. El ejemplo (13) ilustra la cantidad de evidencia textual que tuvo que considerarse para la interpretación de algunos ejemplos de nombres cápsula. En el ejemplo,

his vision no se refiere a información adyacente, sino a los segmentos subrayados, los cuáles, como se evidencia en las líneas numeradas, aparecen lejos de la palabra nodo.

(13) 106 After all Koresh's apocalyptic vision, his talk of the seven seals that only the lamb of God can unlock, sounded like refinements of what they already believed. [...] david koresh:
174 Someone is going to rule whether the big world likes it or not.
175 Thou shalt break them with the rod of iron, thou shalt dash them in pieces like a potter's vessel. [...] Well David said that we were at the end, everyone believed that they were that the time has come that we was at the end of the world. [...]
388 In order to make his own sort of erm prophetic vision come true. 389 Erm he decided to stage a fire in which it would make it appear as though er this was a result of some sort of erm armed confrontation between law enforcement and his group. (*BNC Sampler*: HE3, S:brdcast: documentary)

El análisis exhaustivo de datos de corpus en este estudio se relaciona con un enfoque teórico de corpus. Según Mahlberg (2005: 31–8), este enfoque critica el análisis basado en patrones y búsquedas pre-definidas (e.g. N-cl y N-*be*-cl) a favor de un tratamiento de datos más 'corpus-driven'. A diferencia de la versión más extrema de la investigación 'corpus-driven' (e.g. Tognini-Bonelli 2001: 67–71; Sinclair 2004: 23), el análisis de datos desde una perspectiva teórica de corpus se lleva a cabo con ciertas preconcepciones teóricas. Por tanto, como afirma McEnery & Hardie (2012: 161), '[...] it is arguably impossible to approach corpus evidence with no preconceptions about language'. Esas preconcepciones vienen a menudo impuestas por nociones teóricas y categorías útiles para un análisis 'corpus-driven' específico. Además de la dependencia de teorías relevantes, un enfoque teórico de corpus comienza el análisis con una serie de suposiciones mínimas acerca del fenómeno lingüístico que se pretende estudiar (Mahlberg 2005: 35–8). Hay tres suposiciones mínimas que guían el análisis que Mahlberg hace sobre los nombres generales: son muy frecuentes, desempeñan funciones textuales locales y son nombres.

La perspectiva teórica de corpus que se adopta en esta tesis se basa en un marco teórico ecléctico (cf. Schmid 2000: 20 y Mahlberg 2005: 33, dónde se defiende también una aproximación ecléctica). Este eclecticismo se desprende del análisis multidimensional (formal, textual y semántico-pragmático) que se propone en la tesis. El análisis consta de nueve variables extraídas de gramáticas descriptivas contemporáneas del inglés (Quirk et al. 1985) y de la Gramática Sistémico-Funcional (de aquí en adelante, GSF; Halliday & Matthiessen 2004). En lo que sigue, se

presentan las nueve variables de estudio (para más detalles, ver el capítulo 5 de la tesis):

i) Género del texto: Las categorías de género que se usan en este estudio (e.g. W:ac, i.e. prosa académica; S:meeting, i.e. reuniones formales) son las que aparecen en *BNCweb* (véase 4.2.2 en la tesis para más información).
ii) Estructura semántica: Esta variable se relaciona con la estructura semántica del sintagma nominal, como se describe en la GSF (Halliday & Matthiessen 2004), y consta de Deíctico, post-Deíctico, Epíteto, Clasificador y Cualificador. La gramática de Halliday ofrece una clasificación más sistemática de los pre-modificadores semánticos que la que se utiliza en la investigación sobre nombres cápsula (véase 3.2.3 y la Tabla 3.1 en la tesis).
iii) Estructura formal: Las categorías incluidas en esta variable se adoptan de la gramática de Quirk et al. (1985) (e.g. artículo definido, cláusula de relativo, cláusula nominal de complemento, etc.). La gramática de Halliday no se utilizó para esta variable debido a que la gramática de Quirk et al. (1985) da una descripción más detallada de la forma del sintagma nominal que la propuesta por Halliday & Matthiessen (2004).
iv) Función sintáctica: El marco que se sigue aquí es de nuevo el de la gramática de Quirk et al. (1985) (e.g. objeto directo, sujeto, adverbial adjunto, complemento del objeto, etc.). Frente a la clasificación detallada en Quirk et al. (1985), Halliday & Matthiessen (2004) distinguen solo cuatro tipos de función sintáctica (sujeto, predicado, complemento y adjunto). En la GSF, la función de complemento incluye lo que en Quirk et al. (1985) se corresponde con el objeto directo e indirecto, el complemento del sujeto y el complemento del objeto.
v) Tipo de participante: El análisis de los roles semánticos que desempeñan los nombres cápsula sigue el sistema de la Transitividad de Halliday & Matthiessen (2004) (e.g. Actor, Identificador, Circunstancia, etc.).
vi) Tema/Rema: Para esta variable, se usa el sistema de Tema de Halliday & Matthiessen (2004).
vii) Dirección de la encapsulación: Esta variable consta de referencia endofórica (anáfora y catáfora), de sus realizaciones (inter-oracional, intra-oracional) y de referencia exofórica. El análisis de esta variable se basa en la literatura sobre encapsulación que se describe en 3.2.2 (ver tesis).

viii) Antecedente: Siguiendo a Stirling & Huddleston (2002: 1455), el término 'antecedente' se emplea aquí con referencia a segmentos discursivos anafóricos y catafóricos (ver 3.2.2 en la tesis). Las categorías que abarca esta variable se adoptan de Gray (2010), dónde se distingue entre 'Discurso Extenso Global' (donde el antecedente ocupa varias oraciones o es difícil de delimitar) y 'Discurso Local', que se divide en Sintagma Nominal (Simple: sin post-modificación y Complejo: con post-modificación) y Oración/Cláusula (ver 2.3.1.2.1 en la tesis).

ix) Tipo semántico: Esta variable se aplica al significado que el nombre cápsula muestra en contexto. La clasificación semántica que se sigue aquí es la que propone Schmid (2000): nombres cápsula factuales, mentales, lingüísticos, circunstanciales, modales y eventivos (ver 3.2.5 en la tesis).

La aproximación teórica de corpus de este estudio se inspira, por tanto, en la GSF para las variables semánticas y textuales (i.e. estructura semántica, tipo de participante y Tema/Rema), y en la gramática de Quirk et al. (1985) para las variables formales o estructurales (i.e. estructura formal y función sintáctica). Las cuatro variables restantes, i.e. género, dirección de la encapsulación, antecedente y tipo semántico, son más heterogéneas en sus bases teóricas. Una posible crítica a la naturaleza ecléctica de este estudio es que no se utiliza un único marco. Sin embargo, si se hubiera utilizado un único enfoque teórico, el foco de investigación se habría ceñido o a aspectos más formales (i.e. Quirk et al. 1985) o a aspectos más semántico-pragmáticos y textuales (i.e. GSF). La utilización de un enfoque ecléctico permite, por tanto, explorar una mayor cantidad de variables, y cada variable puede, al mismo tiempo, codificarse con el mayor detalle posible.

Para concluir esta sección, cabe mencionar que, aparte del marco teórico de las variables, el enfoque teórico de corpus también se evidencia en dos suposiciones mínimas. Una de esas suposiciones es que los nombres cápsula son entidades abstractas de segundo y tercer orden como *action*, *event*, *idea* o *point*. Hay otros nombres abstractos, como *love, arrival* o *democracy* que no funcionan como unidades encapsuladoras, ya que son semánticamente más genéricos, y por tanto, menos dependientes de realización léxica que ejemplos típicos de nombres cápsula (ver Schmid 1999: 223 en 3.2.1, tesis). Por ejemplo, mientras que podría preguntarse lo que *action* o *idea* implican en un contexto específico, no es muy frecuente hacer la misma pregunta acerca de nombres como *love*, *arrival* o *democracy*, ya que representan conceptos generales para los cuales no se aprecia la necesidad de información contextual para su entendimiento.

La segunda suposición se relaciona, de hecho, con el contexto. Siguiendo a Ivanič (1991: 111), se asume que, con independencia del determinante que acompaña al nombre cápsula (específico, no específico o cero), el contexto inmediato influirá siempre en su interpretación. A medida que el discurso se desarrolla, los lectores se esfuerzan por entender cualquier entidad discursiva, apoyándose en el modelo mental creado por el texto hasta ese momento, además de en modelos extra-textuales de conocimiento (ver Brown & Yule 1983: 201 y Garnham & Oakhill 1990: 380 en 3.2.3, en la tesis). En casos en los que la lectura intensiva de textos de corpus no proporciona información contextual para entender el significado de un nombre cápsula, se asume que el lector recurrirá a conocimiento extra-textual, y, por consiguiente, el ejemplo se trata como exofórico. Por ejemplo, en (14), el sintagma nominal *an unambiguous policy recommendation* tiene referencia genérica, ya que implica *any potential recommendation on policy that might be considered unambiguous* (i.e. *one such recommendation*). Puesto que no se especifica ninguna de esas recomendaciones ni en el discurso precedente ni posterior, *recommendation* se trata aquí como exofórico. El ejemplo (15), pese a aparecer con el artículo indefinido, es diferente, ya que *a clear vision* puede interpretarse en función del segmento discursivo subrayado (i.e. *the BBC's vision would be to introduce a wide range of high quality programmes, greater efficiency*, etc.). En este caso, el sintagma nominal cápsula es endofórico.

(14) Clearly this might not be a task that can be undertaken with any great degree of precision. Yet if we are to pronounce on the desirability or otherwise of individual monopoly situations, or proposed mergers, this is clearly what is required to provide an unambiguous policy recommendation. (*BNC Sampler*: HXN, W:commerce)

(15) 'The BBC must therefore have a clear vision if it is to retain its role as the cornerstone of British broadcasting and continue to command respect and admiration in Britain and throughout the world. 'In January John Birt laid out that vision of a wide range of high quality programmes, greater efficiency and accountability with value for money for licence payers. 'We, the ten members of the board of management, believe that John Birt is the best person to lead the BBC and he has our unanimous support. (*BNC Sampler*: CF6, W:newsp:other:report)

3.4 *Nombres cápsula frente a aquellos que no lo son*

En la sección 4.4.2 (tesis) se describen 12 categorías de casos que no representan nombres cápsula, y que, por tanto, no aparecen en la base de datos (ver 4.4.2 para más detalles):

i) Nombres pre-modificadores: Los nombres con función de pre-modificación se excluyen del análisis debido a su estatus no referencial. El ejemplo se incluye en la base de datos si el nombre desempeña una función a nivel de cláusula o como complemento en un sintagma preposicional.
ii) Repetición completa: Tan sólo se incluyen nombres repetidos parcialmente siempre y cuando el nombre más simple provenga de un sintagma nominal o segmento discursivo más complejo e informativo.
iii) Clase de palabra errónea: La palabra nodo, con forma verbal o adjetival, aparece etiquetada erróneamente como nombre.
iv) Unidad incorrecta: El hablante utiliza un nombre cápsula por error.
v) Títulos de libros, periódicos, etc.: El nombre forma parte de expresiones no referenciales como títulos, nombres propios, etc.
vi) Entidad de primer orden: conjuntos cerrados y rangos de valores: El nombre denota o medidas (e.g. *dimension*) o conjuntos cerrados de unidades (e.g. *human senses*). Los nombres circunstanciales como *time* o *point* se incluyen en la base de datos sólo si encapsulan información eventiva.
vii) Entidad de primer orden: unidad visible y tangible: En términos generales, aquellos nombres que denotan entidades físicas reales se excluyen del análisis, ya que su significado no es ni abstracto ni meta-discursivo. Por lo tanto, unidades como *area*, *application* y *part* se tratan como nombres cápsula sólo en sus sentidos abstractos.
viii) Usos genéricos: Se excluyen de la base de datos aquellos casos en los que el significado del nombre hace referencia a cualidades generales o conceptos sin necesidad de especificación contextual. Hay, sin embargo, casos aparentemente genéricos, que dependen del contexto inmediato. Estos sí se incluyen en el análisis.
ix) Nominalizaciones: Tan solo las nominalizaciones de producto suelen asociarse a la necesidad de especificación contextual. En general, las nominalizaciones procesuales se excluyen de la base de datos.
x) Adverbiales conectores y enfatizadores: Dado que la atención en estos casos se centra no en el nombre que complementa a la

preposición, sino en el sintagma preposicional completo, a no ser que el significado sea contextual y referencial (lo cual se asocia típicamente a *this*), el ejemplo se excluye de la base de datos.

xi) Frases hechas: Sólo aquellos casos en los que la frase hecha necesita especificación contextual se incluyen en el análisis.

xii) Discurso incompleto: La naturaleza a veces difusa del discurso oral espontáneo hace que algunas variables (frecuentemente, dirección de encapsulación y antecedente) se etiqueten como 'unclear'.

4 RESULTADOS

En función de los resultados obtenidos en el capítulo 6, el comportamiento de los nombres cápsula analizados puede resumirse en 8 puntos principales:

i) Su utilización se concentra principalmente en macro-géneros escritos, especialmente en prosa periodística y académica. Al comparar el uso de nombres en la parte oral del *BNCweb* con el uso de nombres cápsula en la parte oral de la muestra de estudio, hay tres macro-géneros que destacan en el uso de nombres cápsula: conversaciones, reuniones y discursos. Las unidades de mayor frecuencia en lenguaje escrito son semánticamente más específicas y con una distribución más uniforme que lo que se aprecia con las unidades que aparecen en lenguaje oral. Desde el punto de vista semántico, la conexión más evidente es la que existe entre prosa académica y nombres factuales.

ii) Los tres determinantes más frecuentes son el artículo definido, el artículo cero y el artículo indefinido. Los nombres cápsula definidos son, en general, más comunes en el lenguaje oral, especialmente en programas de televisión/radio, discursos y reuniones. Estos muestran además una fuerte preferencia por posiciones de sujeto y Tema, lo cual se asocia a su vez a prosa académica y periodística. En cuanto a dirección de la encapsulación, los usos catafóricos intra-oracionales y los exofóricos predominan. Frente al predominio del artículo definido en macro-géneros orales, el artículo cero destaca en macro-géneros escritos, especialmente en prosa académica divulgativa, periodística y en tratados sobre economía. Sintácticamente, estos casos favorecen posiciones de complemento y de objeto preposicional, mientras que, textualmente, son a menudo exofóricos. Los nombres cápsula indefinidos aparecen comúnmente en conversaciones y en ficción, además de en posiciones de complemento del sujeto. Finamente, en lo que se

refiere a asociaciones semánticas, el artículo definido suele aparecer con nombres factuales y lingüísticos, el artículo cero, con nombres lingüísticos y eventivos, y el artículo indefinido, con nombres modales y eventivos (seguidos de unidades factuales).

iii) La pre-modificación semántica es relativamente infrecuente en la muestra de estudio. Pese a esto, cabe destacar la estrecha conexión entre Epíteto y posiciones atributivas, lo que, a su vez, explica su frecuente asociación con sintagmas nominales indefinidos. De un modo más específico, los Epítetos interpersonales predominan con nombres cápsula mentales en macro-géneros orales (i.e. programas de televisión/radio, discursos y conversaciones) y en ficción. Los Epítetos experienciales, por el contrario, favorecen claramente la prosa académica, además de significados lingüísticos y modales. Los Clasificadores, dada su objetividad, no es de extrañar que aparezcan frecuentemente en prosa expositiva (i.e. tratados de economía, textos misceláneos y prosa periodística). Los nombres eventivos y modales predominan con este pre-modificador. Los post-Deícticos de identidad son comunes en reuniones, y como podría esperarse, muestran una fuerte conexión con posiciones de sujeto en cláusulas relacionales intensivas (e.g. *the only thing is*, *the main problem is*).

iv) En cuanto a estructuras de post-modificación, los sintagmas preposicionales de *of* predominan en la muestra de estudio. La distribución genérica en este caso no tiene nada de especial, ya que la post-modificación preposicional tiene una frecuencia similar en todos los macro-géneros. Las cláusulas de relativo, no obstante, son muy comunes en macro-géneros orales, especialmente en conversaciones, discursos y reuniones. Mientras que las clausulas de relativo tienden a unirse a nombres circunstanciales, los sintagmas preposicionales prevalecen con nombres factuales, mentales y lingüísticos.

v) Desde el punto de vista sintáctico, la mayor concentración de nombres cápsula aparece en posiciones de objeto directo, sujeto y complemento del sujeto. Los nombres cápsula sujeto predominan en prosa académica y periodística, mientras que los casos de objeto directo y complemento del sujeto destacan en conversaciones y ficción. Los nombres sujeto tienden a mostrar significados modales y factuales, mientras que los casos de objeto prefieren significados lingüísticos, modales y mentales. Cabe mencionar que los nombres modales y lingüísticos son los más frecuentes en posiciones de Rema, mientras que Tema suele estar

ocupado por nombres factuales (como es el caso también con los casos de sujeto).

vi) Con respecto a Transitividad, los participantes relacionales destacan claramente en todos los macro-géneros, seguidos de circunstancias. Los contextos relacionales muestran una marcada preferencia por nombres factuales, mientras que las circunstancias suelen asociarse a nombres circunstanciales. Los resultados indican una correspondencia entre tipo semántico de nombre cápsula y tipo de proceso (i.e. relacional y factual, material y eventivo, verbal y lingüístico, circunstancia y circunstancial).

vii) En general, los nombres cápsula en la muestra de estudio parecen preferir usos catafóricos, seguidos de casos de anáfora y exófora. Los casos de catáfora prevalecen en prosa periodística, los de anáfora, predominan en programas de televisión/radio y en prosa académica, y los de exófora, en tratados de economía y conversaciones. En cuanto a preferencias semánticas, los usos catafóricos son especialmente prominentes con nombres modales, los anafóricos, con nombres factuales, y los exofóricos, con nombres lingüísticos y eventivos.

viii) En cuanto a realización léxica, la catáfora intra-oracional se muestra como la relación de encapsulación más frecuente, seguida de la anáfora y catáfora inter-oracionales, y de la anáfora intra-oracional. Los antecedentes de cláusula y de complemento preposicional destacan con los casos de catáfora intra-oracional, mientras que los antecedentes nominales predominan con ejemplos de anáfora intra-oracional. Los dos tipos de relaciones inter-oracionales (anáfora y catáfora) muestran una fuerte preferencia por antecedentes globales extensos. Los resultados más evidentes en la relación entre antecedente, género y significado se desprenden de cláusulas, sintagmas nominales y pronombres. Las cláusulas suelen aparecer en prosa periodística, conversaciones y tratados de economía, y con nombres eventivos y mentales. Los antecedentes nominales (especialmente, sintagmas complejos) tienden a asociarse con prosa expositiva (i.e. académica profesional y divulgativa) y nombres factuales. Los antecedentes pronominales también tienden a aparecer con nombres factuales, pero, en este caso, el uso en macro-géneros orales (especialmente, en conversaciones) predomina.

5 CONCLUSIÓN

5.1 *El estudio y su contribución a la investigación sobre los nombres cápsula ingleses*

Hace más de una década, Schmid (2000: 379) concluía su estudio sobre nombres cápsula sugiriendo dos posibles vías de investigación. Desde su punto de vista, la investigación sobre nombres cápsula podría beneficiarse de métodos de corpus abarcando el mayor número de patrones. Esto podría conseguirse sólo a través de un enfoque menos mecánico, y más basado en el análisis detallado de datos de corpus. La segunda vía enfatiza la necesidad de profundizar en la distribución de estas unidades con respecto a género y modo.

Tras Schmid (2000), varios estudios ampliaron el ámbito de investigación a otros tipos de patrones (no sólo N-cl y N-*be*-cl; e.g. Flowerdew 2003a; Aktas & Cortes 2008; Caldwell 2009). A pesar de esto, una gran parte de la investigación hasta el momento se ha centrado en discurso académico. Incluso Flowerdew (2003a: 331), dónde se plantea una descripción exhaustiva de estas unidades, extrae resultados de una muestra de inglés académico. Como demuestra la Tabla 6.1 en 6.2.1.1 (ver tesis), aunque la prosa académica ocupa una posición importante en el uso de nombres cápsula, hay otros macro-géneros que también destacan (e.g. periódicos, reuniones formales, ficción, discursos y conversaciones).

Esta tesis explora las dos vías de investigación que sugiere Schmid (2000: 379), ya que, por una parte, se utiliza un corpus pequeño pero equilibrado, y, por otra, el enfoque analítico es manual y multidimensional. Por lo tanto, siguiendo a Mahlberg (2005: 188), se cree aquí que solamente un análisis altamente detallado e interpretativo de datos de corpus puede actualmente arrojar luz sobre fenómenos textuales. En el caso de los nombres cápsula, dónde el contexto tiene tal influencia en la forma, función y significado de estas unidades, la lectura intensiva de segmentos discursivos desde oraciones hasta páginas enteras es el único modo de llegar a resultados más detallados sobre el comportamiento de estas unidades. Esto se refleja en la cantidad de detalle que se da en muchos de los ejemplos en esta tesis, dónde se expanden los límites de la concordancia a extensos segmentos con el fin de ofrecer una representación más o menos acertada de los modelos discursivos de lectores y oyentes en situaciones discursivas concretas (Prince 1981: 235). Este análisis, por tanto, cuestiona la validez de enfoques puramente cuantitativos y ceñidos a la concordancia en casos en los que el foco de atención se encuentra no sólo en forma y significado dentro de la concordancia, sino también en como significado y forma emergen y se relacionan con la función de una entidad discursiva en un

punto específico en el desarrollo del discurso oral o escrito. Esta es precisamente una de las mayores contribuciones de esta tesis a la investigación sobre nombres cápsula, dado que, hasta dónde yo sé, este es el primer estudio en el cuál se investigan sistemáticamente los rasgos formales, sintácticos, semánticos y textuales de estas unidades.

5.2 *¿Qué son los nombres cápsula?*

El análisis cuantitativo y cualitativo en esta tesis muestra que los nombres cápsula son primordialmente nombres. Aunque esto pueda parecer una obviedad, la literatura no ha discutido suficientemente el grado de afinidad existente entre nombres cápsula y nombres en general. La evidencia cuantitativa en el capítulo 6 recoge una serie de similitudes entre ambos:

i) El artículo definido prevalece en ambos, seguido por el artículo indefinido, los determinantes posesivos y demostrativos.
ii) Los determinantes posesivos predominan en ficción, y los demostrativos, en prosa académica.
iii) La pre-modificación es más común en los géneros escritos, en particular, en prosa periodística y académica.
iv) Los adjetivos evaluativos son frecuentes en ficción y conversación, mientras que los adjetivos descriptivos y clasificadores predominan en prosa periodística y académica.
v) La post-modificación preposicional (especialmente *of*) aparece como la estructura de post-modificación más frecuente en nombres cápsula y nombres en general.
vi) El uso de nombres cápsula y de cualquier otro nombre prevalece en posiciones de objeto y de complemento preposicional.
vii) Los nombres con función de sujeto suelen utilizarse en prosa académica y periodística. La función de objeto, sin embargo, predomina en ficción.
viii) Los sintagmas simples (sin post-modificación) son comunes en posiciones de sujeto, mientras que los sintagmas complejos (con post-modificación) tienden a preferir posiciones post-verbales.
ix) El artículo definido suele aparecer en posición de sujeto, y el indefinido, en posición de atributo.
x) Los usos anafóricos son muy frecuentes en prosa académica, y los exofóricos, en conversación.

Teniendo en cuenta estas similitudes, la idea que Hunston & Francis (2000: 185) plantean sobre la pertenencia de estas unidades a una clase separada de nombres no parece ser completamente acertada, al menos, en cuanto a

forma y función sintáctica. Los nombres cápsula se comportan como nombres en su asociación con los mismos tipos de determinantes, pre-modificadores, estructuras de post-modificación y funciones sintácticas. Esto, por tanto, plantea la pregunta sobre qué hace a un nombre un nombre cápsula. Los siguientes tres puntos intentan dar respuesta a dicha pregunta en función de los resultados de la tesis:

i) Limitada especificidad semántica y realización léxica: En la literatura, el uso de nombres cápsula se asocia a la falta inherente de especificidad semántica en entidades de segundo y tercer orden. Se cree aquí que esta falta de especificidad no es un concepto monolítico, sino que varía en función de significado y contexto. El significado de los nombres cápsula se refleja en unidades altamente frecuentes como *thing*, *way*, *area* o *point*, y en aquellas de rango de frecuencia bajo como *contradiction*, *endeavour*, *foreboding* o *recollection*. La tesis muestra una conexión entre frecuencia y la función encapsuladora, puesto que una frecuencia elevada implica una mayor asociación con información contextual, mientras que una frecuencia baja conlleva usos más genéricos (y por tanto, usos no encapsuladores).

El contexto juega un papel fundamental en la desambiguación e identificación de nombres cápsula en esta tesis. El análisis adopta la perspectiva del lector o interlocutor en una situación discursiva específica para determinar el significado contextual del nombre cápsula. En este punto, es importante señalar que la información contextual no es siempre ni claramente delimitable ni se encuentra en el contexto inmediato, como afirman Francis (1994: 83) y Hunston & Francis (2000: 185). Este énfasis en antecedentes inmediatos y precisos proviene de la frecuente asociación entre la propiedad encapsuladora, el uso de sintagmas definidos y demostrativos, y el predominio de los patrones N-cl y N-*be*-cl. Por tanto, ejemplos como (16) y (17) representarían, según la literatura, usos encapsuladores prototípicos, ya que el sintagma nominal en ambos es definido y la información contextual es claramente delimitable. Sin embargo, el ejemplo (18) es menos claro en su significado contextual, ya que el antecedente no es fácilmente delimitable. A pesar de esto, *terror* depende del contexto, dado que no es ni terror ni acciones violentas en general lo que inspiró esta declaración, sino la violencia usada por el gobierno rumano durante el fin de semana. Por esta razón, se puede decir que la falta de especificidad semántica no se relaciona única y exclusivamente con entidades abstractas definidas o demostrativas. Puede

aparecer con cualquier realización formal del sintagma nominal con la condición de que el nombre sea significativo en contexto. Por consiguiente, la observación que hace Ivanič (1991: 112) de que en el estudio de estos nombres, '[...] the context seems to play a role in their interpretation even when they are accompanied by indefinite reference, or when they occur as uncountables' debería expresarse de nuevo en estos términos: 'the context always plays a role in the interpretation of shell-noun instances, regardless of the determiner used'. La importancia del contexto es tal que incluso entidades concretas de primer orden pueden adquirir usos encapsuladores en casos en los que el foco de atención está en el contenido metadiscursivo del nombre (siendo éste proposicional o eventivo). Esto es evidente en el ejemplo (19). Consten et al. (2007: 82) analizan *this grafitto* como un nombre encapsulador en base a su denotación de una entidad visible de primer orden (i.e. pintura en una pared). El contexto, sin embargo, parece sugerir una interpretación encapsuladora, ya que el significado incluye el producto y el contenido proposicional o la información asociada a ese producto (i.e. el segmento subrayado).

(16) Lethal Force tested the suspicion that men who could have been captured or wanted to surrender had been shot dead [...] (*BNC Sampler*: J1L, W:misc)

(17) Mr Baker said the main area of concern for over the next few years was 'serious downturn' in the economies of export markets in Europe and Japan. (*BNC Sampler*: CF8, W:newsp:other:report)

(18) By Foreign Staff OUTRAGE and concern at the violence in Romania during the weekend came from all parts of the world yesterday, with the Soviet Union and the US in the lead. [...]European foreign ministers, meeting in Brussels, condemned 'in the strongest possible terms' the measures taken by the Romanian security forces. [...]he said 'if indeed some loss of life has occurred, I can only express my very profound regret'. [...] It reported that Romanian state institutions and factories were under intensified guard, and that the frontiers were closed to tourists. [...]The Polish Parliament stood for a minute's silence after approving unanimously a resolution which accused the Romanian authorities of 'exceptional brutality' leading to children being shot at and expressed solidarity with the 'victims of terror.' The World Council of Churches, the World Alliance of Reformed Churches, the Lutheran World Federation and the Conference of European Churches said in a joint telegram they were 'deeply disturbed by the disquieting intervention of the army against civilian people in Timisoara.' (*BNC Sampler*: AAB, W:newsp:brdsht_nat:report)

(19) 'At the end of the eighties a rogue characterized the decadent gluttony: 'Buy nicer and nicer, fly further and further, shag faster and faster'. During several months this graffito ornamented the walls of an old comfort station at Hamburg's Yuppie district […]' (Consten et al. 2007: 82)

ii) Estatus abierto y cerrado, y el criterio de la proforma: La literatura suele inspirarse en la definición que Halliday & Hasan (1976: 274) hacen de los nombres generales (e.g. *thing*, *fact*, *person*, *creature*) como unidades a medio camino entre el significado gramatical y el léxico para describir el uso de los nombres cápsula. Siguiendo a Halliday & Hasan (1976), los nombres cápsula suelen equipararse a pronombres y pro-formas (e.g. *this*, *they*, *it*, *so*) en función de su frecuente estatus definido y anafórico y su falta de especificidad semántica (cf. e.g. Francis 1986: 4; Francis 1994: 85). Ivanič (1991: 107–8) dice que, aunque sus nombres 'carrier' se parecen a los pronombres en su inespecificidad semántica, son nombres en todo lo demás (i.e. más informativos, potencial evaluativo, pre- y post-modificación). Mahlberg (2005: 177) resuelve la cuestión de la indeterminación categorial de los nombres generales afirmando que una perspectiva lingüística de corpus al estudio de estas unidades no reconoce distinciones gramaticales precisas. La indeterminación es la base fundamental de los resultados de corpus, ya que, los nombres y cualquier otra clase de palabra muestran limites difusos cuando se analizan a través de un corpus. Por consiguiente, el foco de investigación debería estar en el significado como uso (Mahlberg 2005: 177) o en como el significado se plasma formalmente en tipos de textos específicos. Esto explica la identificación que Mahlberg (2005) hace de las categorías semánticas específicas asociadas al uso de los nombres generales en contexto (i.e. las funciones locales textuales; e.g. orientación temporal, medida, gente vs. gobierno, humanos, etc.).

En esta tesis, no se considera que los nombres cápsula tengan una indeterminación categorial inherente. Como se especifica arriba, los nombres cápsula se comportan como nombres en su forma, función sintáctica y significado léxico, y como tal, se tratan como nombres. Lo que les hace especiales es su significado abstracto y su necesidad de información contextual para clarificar su contribución al discurso. En vista de las asociaciones formales, sintácticas y semánticas de los nombres cápsula con cualquier otro nombre, lo verdaderamente interesante no es su categorización nominal o pronominal, sino como los significados contextuales de

entidades de segundo y tercer orden en tipos de textos específicos influyen en su comportamiento formal, sintáctico y textual.

iii) Los nombres cápsula, y su relación con género y tipo semántico: De lo anterior, se desprende que una definición acertada y exhaustiva de los nombres cápsula no puede basarse en criterios formales y sintácticos, como suele hacerse en la literatura. Tan solo las variables semánticas y textuales en esta tesis parecen arrojar luz sobre la propiedad encapsuladora. De especial relevancia es el uso extendido de estas unidades con procesos relacionales (véase la Figura 6.22b en 6.2.3.1, tesis) y antecedentes largos (cláusulas, oraciones, segmentos extensos: 64.96%; véase la Tabla 6.34 en 6.2.4.1, tesis), lo cuál no refleja la conexión esperada entre nombres concretos de primer orden, procesos materiales y antecedentes cortos. Por lo tanto, los nombres cápsula son cápsula en función de su frecuente aparición con procesos relacionales y su encapsulación de segmentos discursivos. Aunque este último hallazgo refleja la afirmación frecuentemente repetida sobre la referencia de los nombres cápsula a segmentos discursivos, a diferencia de otros estudios, los resultados aquí vienen apoyados por datos cuantitativos detallados que incluyen una gran variedad de tipos de antecedente. La taxonomía de antecedentes que se propone en esta tesis (véase 5.3.4.2, tesis) muestra que, a pesar de la conexión con antecedentes extensos, 35.04% de los nombres cápsula se corresponde con antecedentes sintagmáticos (principalmente sintagmas nominales y complementos preposicionales).

Una definición de nombres cápsula en función solo de tipo de participante y antecedente no es lo suficientemente exhaustiva, ya que no captura la influencia que tanto significado como tipo de texto pueden tener en el comportamiento formal, sintáctico y semántico de estas unidades. Para ilustrar esta influencia, la Tabla 7.1 en 7.3 (tesis) recoge las cinco categorías más frecuentes para las variables de género, estructura formal y semántica, función sintáctica, participante, Tema/Rema, dirección de encapsulación y antecedente (las dos columnas del extremo izquierdo), junto a los tres macro-géneros y tipos semánticos de nombre cápsula más comunes para cada categoría (columnas restantes). Las seis columnas del extremo derecho exploran la conexión entre estructura (formal y semántica), dirección de encapsulación y otras variables (ver el capítulo 6 para los resultados completos). Las categorías en rojo son aquellas para las que las líneas en las

figuras en el capítulo 6 destacan. Las celdas en amarillo indican casos en los cuales el comportamiento de los nombres cápsula en la muestra coincide con el comportamiento de los nombres en general. Los seis puntos siguientes se basan en las categorías en rojo para describir algunas de las asociaciones más marcadas en la muestra (ver el capítulo 6 para todos los detalles):

- La prosa académica y las reuniones formales suelen preferir nombres factuales y lingüísticos respectivamente. Las preferencias semánticas se relacionan estrechamente con las funciones comunicativas típicas de cada macro-género (e.g. informar, ofrecer sugerencias, etc.; véase 6.2.1.3 en la tesis).
- Los nombres cápsula definidos suelen aparecer en posiciones de sujeto y con realizaciones catafóricas intra-oracionales. Los nombres indeterminados son frecuentes con significados lingüísticos, posiciones de complemento preposicional y usos exofóricos. Los nombres indefinidos prevalecen en conversación y en posiciones atributivas. Los determinantes demostrativos son muy comunes en prosa académica, y aparecen casi siempre como casos de anáfora inter-oracional.
- La post-modificación preposicional predomina con nombres factuales, lingüísticos y mentales. Las clausulas de relativo, por su parte, muestran una fuerte preferencia por nombres circunstanciales y por macro-géneros hablados (especialmente, conversaciones y discursos). Las cláusulas nominales de complemento de *to*-infinitivo prevalecen con significados modales, y destacan en prosa periodística y conversaciones. Las cláusulas de *that* son muy frecuentes con nombres mentales, y suelen aparecer en posiciones de sujeto nocional o existencial y de objeto preposicional.
- En cuanto a pre-modificación semántica, las asociaciones más relevantes son aquellas entre Clasificadores, significados eventivos y modales, entre Epítetos interpersonales, ficción, programas de radio/televisión, nombres mentales y posiciones atributivas, entre Epítetos experienciales, prosa académica y posiciones atributivas, y entre post-Deícticos de identidad, reuniones formales y posiciones intensivas relacionales de sujeto.

- Los nombres cápsula objeto son frecuentes en conversación, ficción y con significados lingüísticos y modales, mientras que los nombres sujeto y Tema predominan en prosa académica y periodística, y con significados modales y factuales.
- Los resultados revelan una estrecha relación entre los componentes del sistema de la Transitividad y los tipos semánticos de nombre cápsula. Los participantes relacionales muestran una fuerte preferencia por nombres factuales, los participantes materiales, por nombres eventivos, y los participantes verbales, por nombres lingüísticos (ver 6.2.3.5 para más detalles).
- En cuanto a encapsulación, los usos exofóricos son muy comunes en tratados de economía y en conversaciones, con significados lingüísticos y eventivos, y en posiciones de Rema y de Tema marcado. La catáfora intra-oracional aparece frecuentemente en periódicos, y se relaciona con significados modales y posiciones de Tema. La anáfora inter-oracional predomina en prosa académica y en programas de radio/televisión, además de en posiciones de Tema. La catáfora inter-oracional, sin embargo, destaca en folletos, manuales de instrucciones y otros textos. Su distribución es casi igual en posiciones de Rema y Tema. Finalmente, la anáfora intra-oracional se relaciona claramente con nombres eventivos y factuales en posiciones de Rema.
- Las asociaciones más destacadas en lo que se refiere a tipo de antecedente son entre cláusulas y nombres modales, entre sintagmas nominales simples y programas de radio/televisión, entre sintagmas nominales complejos y discurso académico (profesional y popularizado), y entre oraciones y ficción.

5.3 *Posibles vías de investigación*

A pesar de lo positivo del enfoque multidimensional de esta tesis, las conclusiones sobre el comportamiento de los nombres cápsula en función de significado y género son solamente tentativas, dado que se basan en un grupo limitado de lemas (solo 60) y de datos (1447 concordancias). Se necesita más evidencia para determinar si las conexiones observadas aquí se relacionan únicamente con los datos considerados, o si, en verdad, indican tendencias generales. Esta sección sugiere una serie de posibles

vías de investigación para hacer frente a las limitaciones de este estudio, y como resultado, para profundizar en áreas de investigación inexploradas hasta el momento. Las seis vías propuestas se dividen en dos sub-secciones, 5.3.1 y 5.3.2, la primera dedicada a áreas de investigación relacionadas con el uso de los nombres cápsula con respecto a género, estructura discursiva y propósito, y la segunda, relacionada con los límites y aplicación de la función encapsuladora.

5.3.1 *Géneros y sub-géneros, estructura discursiva e implicaciones sociales*

i) Se necesitan más datos para confirmar o refutar las asociaciones entre nombre cápsula, género y significado que se presentan en esta tesis. Esto implicaría expandir el foco de análisis a una mayor cantidad de lemas y a más evidencia para cada lema (e.g. 100 concordancias, como en Mahlberg 2005, en lugar de solo 40).
ii) Aunque se ha investigado mucho sobre el uso de los nombres cápsula en una serie de disciplinas y sub-géneros académicos, se presta todavía demasiada atención a patrones formales específicos y a sus funciones (e.g. *th*-N, N-cl). Existe, por tanto, la necesidad de análisis multidimensionales sobre el uso de nombres cápsula en sub-géneros y disciplinas académicas específicas, análisis en los cuales se investiguen de un modo exhaustivo los aspectos formales, sintácticos, semánticos y textuales de los nombres cápsula. Esta necesidad se aplicaría no sólo a prosa académica, sino también a cualquier otro macro-género.
iii) Otra posible vía de investigación es la influencia de la estructura discursiva en el comportamiento de los nombres cápsula. Hasta dónde yo sé, la escasa investigación disponible sobre esta influencia se ciñe a la identificación de los nombres cápsula que aparecen en movimientos y actos retóricos específicos. Sin embargo, se dice muy poco sobre como el comportamiento formal, sintáctico-semántico y textual de los nombres cápsula varía a través de diferentes movimientos y actos retóricos.
iv) En la literatura, se suele afirmar que la función caracterizadora o evaluativa de los nombres cápsula puede también conllevar intenciones manipuladoras (e.g. Conte 1996: 6; Schmid 2000: 8; Schmid 2001). A la hora de utilizar un nombre cápsula en particular, el escritor o hablante puede estar persuadiendo al lector o interlocutor para que acepte una idea o afirmación sin cuestionarla. Según Schmid (2000: 8), esto es algo en lo que políticos y otra gente con experiencia de debate son expertos. En

un artículo sobre las presuposiciones inherentes en el uso de nombres cápsula, Schmid (2001) encuentra que el patrón *the* N-*be*-cl es especialmente adecuado cuando existen intenciones de manipulación (e.g. *the fear is that*, *the problem is that*), ya que, a través del artículo definido, el escritor o interlocutor presenta antecedentes como verdades generales (cf. *the fear is that* vs. *my fear is that*). Aunque el potencial manipulador de las unidades cápsula se reconoce en la literatura, en los pocos casos en los que se menciona, se presta atención sólo a ciertos patrones (e.g. N-*be*-cl). Por esta razón, se cree que la descripción de estas unidades se beneficiaría de una atención más explícita a las connotaciones ideológicas que se desprenden de significados y patrones específicos. Este objetivo de investigación es consistente con los principios que inspiran al enfoque del Análisis Crítico del Discurso, cuyos mayores exponentes son Fairclough (1995, 2000), van Dijk (1997), van Leeuwen (2006, 2008) y Wodak & Meyer (2009). La combinación de este enfoque con la perspectiva multidimensional de esta tesis podría arrojar luz sobre como géneros altamente ideológicos (e.g. discursos políticos, debates y prosa periodística) hacen uso de nombres cápsula para transmitir sus mensajes.

5.3.2 *Los límites formales de la función encapsuladora*

i) Schmid (1999: 222–3) hace referencia al estatus encapsulador de entidades nominalizadas de segundo y tercer orden. Su explicación parece sugerir la existencia de un gradiente que incluye unidades que prototípicamente dependen del contexto (e.g. *warning*, *claim* o *assumption*) y unidades abstractas genéricas (y por lo tanto, no encapsuladoras) como *inflation*, *love* o *derivation*. En la sección 4.4.2.9 (tesis), el análisis de la muestra lleva a una conclusión parecida. Los nombres como *warning* o *recommendation* son típicos ejemplos de la propiedad encapsuladora, puesto que la falta de especificidad de su estructura semántica viene impuesta por los complementos de sus equivalentes verbales (i.e. *warning that*<*warn somebody that*, *recommendation that*<*recommend that*). En casos como *assessment* y *endorsement*, cuyos verbos no permiten proyección de cláusulas de *that* o de *to*-infinitivo, la propiedad encapsuladora es una cuestión de grado. Sólo si la nominalización es de producto, más bien que de acto, puede apreciarse una interpretación encapsuladora. Sin embargo, hay casos en los que incluso nominalizaciones de acto pueden considerarse como nombres cápsula (véase 4.4.2.9 en la tesis para

más información). El contexto es, por tanto, esencial en la distinción entre nombres cápsula y usos procesuales (y, por tanto, no encapsuladores). En el futuro, se podría profundizar en la relación entre morfología derivativa y nombres cápsula, con el fin de cuantificar y arrojar luz sobre el grado en el que nominalizaciones de verbos sin proyección (e.g. *assessment*, *endorsement*, *correction*) permiten interpretaciones encapsuladoras. Este objetivo podría alcanzarse examinando la conexión entre procesos de formación de palabras (e.g. afijación, conversión, etc.) y usos encapsuladores dependientes del contexto.

ii) Al concluir su estudio sobre nombres anafóricos, Francis (1986: 104) sostiene que cualquier unidad puede usarse para conseguir cohesión anafórica o catafórica si el contexto permite una u otra interpretación. Esto plantea la pregunta de si la propiedad de encapsulación se relaciona solo con nombres abstractos de segundo y tercer orden, o si se aplica también a cualquier nombre dependiente del contexto. Los ejemplos (19) arriba y (20) debajo ilustran esto último. Caldwell (2009: 46) mantiene que *this cat* en (20) no implica un uso meramente deíctico de *cat* (i.e. *this is a cat*), sino un significado que se ve influido aquí por la percepción que el escritor tiene del animal. Por tanto, lo que se destaca aquí no es tanto la entidad de primer orden, como la representación de esa entidad en un punto concreto en el universo de discurso (i.e. *not any cat, but this cat with two small piercing eyes, the attitude of a landmine,* etc.).

(20) 'Two small piercing eyes. The attitude of a temperamental landmine. Ten years old and with timing that put my alarm clock to shame. This cat is what stood between me and my first cup of coffee every morning […]' (Caldwell 2009: 46)

Francis (1986: 104) llega incluso a sugerir la existencia no sólo de nombres anafóricos, sino también de verbos anafóricos (e.g. *aware of what was happening*), adjetivos anafóricos (e.g. *even more extraordinary is the notion that*) y adverbios anafóricos (e.g. *more disastrously*). Posteriormente, Partington (1998: 101–4) contempla la existencia de verbos generales (e.g. *happen*, *occur*), los cuáles se usan para sustituir a acciones o eventos más específicos. Este énfasis en los usos anafóricos o catafóricos de cualquier palabra se relaciona con la categoría de unidades 'Vocabulary 3' que propone Winter (1977: 20), la cual consta de

nombres como *attribute*, *justification* y *method*, verbos como *affirm*, *confirm* y *repeat*, y adjetivos como *hypothetical*, *similar* y *converse*. Aunque se ha prestado atención a estas unidades en términos de su señalización de relaciones entre cláusulas y patrones discursivos (Hoey 1979, 1983, 1994; Winter 1977, 1982, 1992; e.g. Causa-Consecuencia, Problema-Solución, etc.), el foco de investigación ha tendido a centrarse en el análisis detallado de textos individuales. Citando a Hoey en 1993, '[...] Winter's (1977) Vocaublary 3 items are more complex in their functioning than either he or I bargained for [...] we have only just begun the proper description of signalling in discourse', existe una clara necesidad de un estudio en el que la señalización metadiscursiva se investigue de un modo exhaustivo a través de análisis multidimensionales de evidencia extraida de corpus, incluyendo variables formales, semánticas y textuales.