The acquisition of morphology in a corpus of secondary school EFL learners:

a focus on possessive '-s'

Elpiniki Atzolidaki

Supervisor: Cristóbal Lozano

Trabajo Fin de Máster (TFM) del Máster de profesorado de educación secundaria obligatoria y bachillerato, formación professional y enseñanza de idiomas

Universidad de Granada

Septiembre 2012

To my parents

Acknowledgements

The present paper is based on a research project conducted as part of the Master's Degree in Education and Foreign Language Teaching at the University of Granada.

This study owes a lot to the support and collaboration of my supervisor Dr. Cristóbal Lozano. Without his insightful comments, the seminars he organised and the material he facilitated me with, I would not have completed this three-fold project. In the same line, I would like to give special thanks to Dr. Ana Díaz Negrillo for her assistance in elaborating the learner corpus and the tagging scheme used in this study.

Additionally, I gratefully acknowledge the support of my family and friends. Their help, patience and encouragement were vital for the completion of this project.

I am also grateful to the students that participated in this study and to the schools IES Pedro Soto de Rojas, IES Padre Suarez and the Centro de Lenguas Modernas (UGR) for giving me their permission to conduct my research in their premises and for all the help and support of the relevant teachers.

Finally, I would like to thank all those people (professors, fellow students, family, friends) that have been a great source of inspiration in my life.

Table of contents

Abstract

This paper consists of three separate studies. The principal one is a morpheme order study (MOS) based on our own learner corpus which contains learner language from secondary school students. For this study we were partly based on previous relevant projects, but our research is distinct in two core points. The first relates to the data elicitation instrument (learner corpus) which is a novel and promising approach in the field of second language acquisition research. The second regards the scoring method used in our project in order to establish the accuracy rates for each of the grammatical morphemes at issue. Regarding this, our study's novelty relies on the combination of the most accurate relevant scoring models that have been previously suggested. Our second study focuses on a specific functor, namely the *possessive –s*. In this regard we have studied not only the accuracy rates of the aforementioned inflectional possession structure, but also its frequency of use in relation to the other possession forms in English. In keeping with the perceived importance of the data elicitation instrument, we have decided to use two different methods (a learner corpus and an experiment) with our L2 English learners and thus deliver more accurate results. Our final study accounts for the use of the various possession forms in L3 German. This study has two distinct focal points. The first refers to the learners' preference for one of the possessive forms available in German as seen by our subjects' corresponding choices. In this regard, we have also compared the frequency of use of each of the possession structures in both L2 English and L3 German. The other point of interest of our third study was the possible influence of our students' L2 (English) on the acquisition of their L3 (German) as seen by our subjects' use of the various possessive forms. In all our studies we have classified our subjects according to their proficiency level, which we determined by means of a corresponding test. This is a very distinctive feature of our study since most previous studies on this field have either not considered the subjects' level of proficiency or they have determined it based on criteria other than an actual proficiency test.

1 Introduction

The description of learners' language (i.e., their interlanguage) has been of particular interest to second language acquisition research. This interest in language learners' interlanguage stems from the idea that its analysis may provide researchers with enriching insights into the process of language acquisition in general and second language acquisition in particular.

Ellis (1994) claimed that, by collecting and analysing samples of learner language, SLA researchers can achieve two goals:
(a) a description of the subjects' linguistic systems, i.e. their interlanguages and
(b) an explanation of the processes and factors involved in acquiring a foreign language.

Hence interlanguage research is only the first step. If we manage to understand the process of second language acquisition, then we will be able to apply the findings to a variety of practical aspects of language teaching: syllabus design, materials development, task design, and language testing.

Researchers' focus on the acquisition of morphemes started in the 1970s when they were investigating the "independent grammars assumption" (Cook, 1993). It soon evolved into the so-called "natural order" studies due to the majority of the researchers' interest in confirming the innatist view of language acquisition. In 1973 Roger Brown observed that when children start learning their first language they omit mainly grammatical rather than lexical morphemes, which eventually appear in their language in subsequent developmental stages. That prompted language acquisition researchers to check whether there is a consistency in the order of L1 and L2 acquisition of grammatical morphemes. A more detailed account of the aforementioned types of morphemes can be found in section 2.1.

The underpinning idea of comparing L1 and L2 morpheme acquisition orders was that if a universal morpheme acquisition order could be established, then the process of acquisition would be proved to be internally driven. Accordingly, it could be argued that second language acquisition is independent of external factors such as the age, the teaching method, the type of exposure (naturalistic vs. instructed) or the L1.

In the same line a number of different theoretical frameworks have been put forth as explanatory patterns. Ellis (1994:44) identifies four major approaches:

1. the study of learners' errors;
2. the study of developmental patterns;
3. the study of variability;
4. the study of pragmatic features.

In our study we have been guided mainly by the approach that focuses on the study of developmental patterns by investigating the acquisition of morphology (both the correct suppliance of morphemes as well as the errors produced). We do consider that MO studies have to offer the SLA researcher with valuable information regarding the process of second language acquisition. First, because, as Ellis and Barkhuizen (2005) put it, "the descriptive information it provides serves as a basis for testing the validity of different explanations of the order of acquisition" (p.79). Second, because, although the so-called "natural order" approach has received a lot of criticism, many methodologically rigorous studies show sufficiently consistent general findings. Nevertheless, the reason why we consider morpheme order studies to be valuable for SLA research is the general degree of commonalities that previously conducted morpheme studies have shown. This, as Larsen-Freeman and Long (1991) noted, provides strong evidence that interlanguages exhibit common accuracy/acquisition orders.

2 Linguistic structures under investigation

In this chapter we will outline the theoretical background that relates to the various studies of our project. Our principal study, as illustrated in our project's title, is the acquisition of morphological structures in L2 English. Therefore, in the first section of the present chapter (2.1) we will provide a definition and a brief description of the **English morphemes** in general and the English grammatical morphemes in particular. However, our project also focuses on the **expression of possession** in L2 English and L3 German. Hence in section 2.2 we outline the various structures of possession expression in English and in German. Additionally, we describe and exemplify the restrictions that each of these possession structures has. As we will see, the differences in the choice of possessive structures that English and German present will guide us in our research on the influence of the L2 English on the acquisition of the L3 German.

2.1 Morphemes in English

Morphemes are usually defined as the "smallest meaningful constituents of words that can be identified" (Haspelmath, 2002:3). In the same line, every lexical item that "[…] cannot be split into meaningful smaller units" (Cook, 1993:25) is a morpheme. Consequently, morphology contains the rules that concern these **minimal meaningful units** of a language and the way in which morphemes are combined to make up words. Morphemes may be *free* or *bound*. The former can stand on their own (e.g. the word *boy*), whereas the latter are attached to other items (e.g. the *genitive –s* in *the boy's book*). Bound morphemes are further divided into inflectional and derivational. Regarding nouns, for example, we can say that they allow "[…] various suffix morphemes –called inflections- to be appended to indicate

plurality (*dogs*), possession (*dog's*), and both plurality and possession (*dogs')*" (Butters, 2001:325). Bound morphemes are derivational if their use results in the formation of a new word, that is, if "they derive one word from another" (Cook, 1993:25). Quite often appears in the relevant literature a distinction between lexical and grammatical morphemes. Lexical morphemes are also known as "content words" (Cook, 1993:25), e.g. *dog, boy, book*, whereas **grammatical morphemes** are also known as **functors**. Brown (1973) defines the latter as:

> […] forms that do not, in any simple way, make reference. They mark grammatical structures and carry subtle modulatory meanings. The word classes or parts of speech involved (inflections, auxiliary verbs, articles, prepositions, and conjunctions) all have few members and do not readily admit new members (p.75).

Natural order studies are mainly concerned with grammatical morphemes as we mentioned at the beginning of this section. The following table presents a list of the grammatical morphemes under examination in the first study of this paper. Note that each of our studies is presented in detail in chapter 4. In the list presented below each of the morphemes is illustrated by an example.

Functor	**Example**
Past regular –ed	She smil<u>ed</u>.
Past irregular	Then, she <u>stood up</u> and <u>left</u>.
Third person singular –s (3SG)	The baby cri<u>es</u>.
Progressive –ing	He is play<u>ing</u> the guitar.
Be copula	She <u>is</u> happy.
Be auxiliary	He <u>is</u> always eating chocolate.
Plural	The <u>children</u> were standing still. Then some of the boy<u>s</u> and girl<u>s</u> left.
Possessive –s	This is John<u>'s</u> car./ She was shown to the girl<u>s'</u> changing

	room.
Articles	The room was cold./ An apple is enough./ Men drive fast.

Table 1: List of the nine functors studied in this paper

2.2 The expression of possession in English and in German

As we stated at the beginning of this chapter, in our study we do not only investigate the acquisition of morphology in English, but we also focus on the accuracy rates of the inflectional forms of possession expression in both L2 English and L3 German. For the study of the acquisition of the **English morphology** by L2 language learners we used exclusively the data found in our **learner corpus**. On the contrary, the study of the **expression of possession** in English and in German required the use of additional data that we collected by means of an **experimental method**. Note that the methods used in our project are presented in detail in chapter 5 that regards our empirical studies and thus includes information about the methods and the materials used for the elicitation and the collection of the data. Naturally, in subsequent chapters we will also be giving more explanation as to the reasons that prompted our special interest on the expression of possession in these two linguistically related languages. But before we start studying the relative frequency with which our language learners used the various possession structures in both English and German, we should present the corresponding **systems of possession expression**.

Possession is the relation between two entities. Most specifically it is a relational concept that can express the conceptual relations between entities. McGregor (2009) suggests that possession should be defined on grounds of the linguistic factors implied rather than conceptually. In keeping with this proposal we can claim that possession is a **relational concept** that regards the relationship(s) between the ***possessum*** (henceforth PM) and the ***possessor*** (henceforth PR) as illustrated in the following examples:

1. John’s book. PR PM
2. The book of John. PM PR
3. His book. PR PM

The **PM** refers to what is possessed, whereas the **PR** makes reference to the entity (person, animal, etc.) that possesses the PM. In our examples (1-3), the PM is always “the book”, whereas the PR is always “John” although in (3) the reference is realised by means of the possessive pronoun.

2.2.1 Possession in English

Following McGregor (2009), we can affirm that three types of possessive construction are distinguished: attributive, predicative and external. In the case of **attributive possession**, the PM and the PR form a single NP as shown in the examples (1),(2) & (3) above. Sometimes these constructions are also referred to as adnominal possession. In the **predicative possession** the possessive relation is expressed in the predicate (e.g. John has a book). In the **external possession** constructions the possessive relation is expressed “[…] at the level of a clausal construction as in *The dog bit Cliff on the ankle*” (McGregor, 2009:2). In **this study** we will be looking at the **attributive possession**.

The attributive possession or possessive NPs have been claimed to (a) be definite (Quirk et al., 1985) and (b) introduce new PM referents into the discourse (Taylor, 1996). Willemse et al. (2009) based on a qualitative and quantitative analysis of a corpus of possessive NPs claim that “[…] many PM referents have a discourse status in between fully given and fully new” (p.13) and thence suggest a continuum-like classification.

Furthermore, within the attributive possession constructions structures like the so-called recursive possessives have been identified. In these multiple possessors can be found in the

same sequence (e.g.: John's father's book). Although the number of recursion is structurally unlimited, it seems that practically no more than two *genitives -s* are found in a chain since the opposite would be "[…] stylistically objectionable, comic and difficult to comprehend" (Quirk et al., 1985: section 17.118).

For the purposes of this study we should also mention that several studies of the English possession system have identified a number of factors that determine the choice of the possessive construction. Some studies have focused on phonological, pragmatic, syntactic and morphological factors (Hawkins, 1994; Quirk et al., 1985). Other researchers have focused on the cognitive and psychological factors (Taylor, 1989; Heine, 1997; Rosenbach, 2005).

In this regard we present the results of the study by **Rosenbach** (2005) in Figure 3. This chart shows the relative **frequency** of the ***s-genitive*** and the ***of-genitive*** according to four factors related to *animacy* and *weight*.

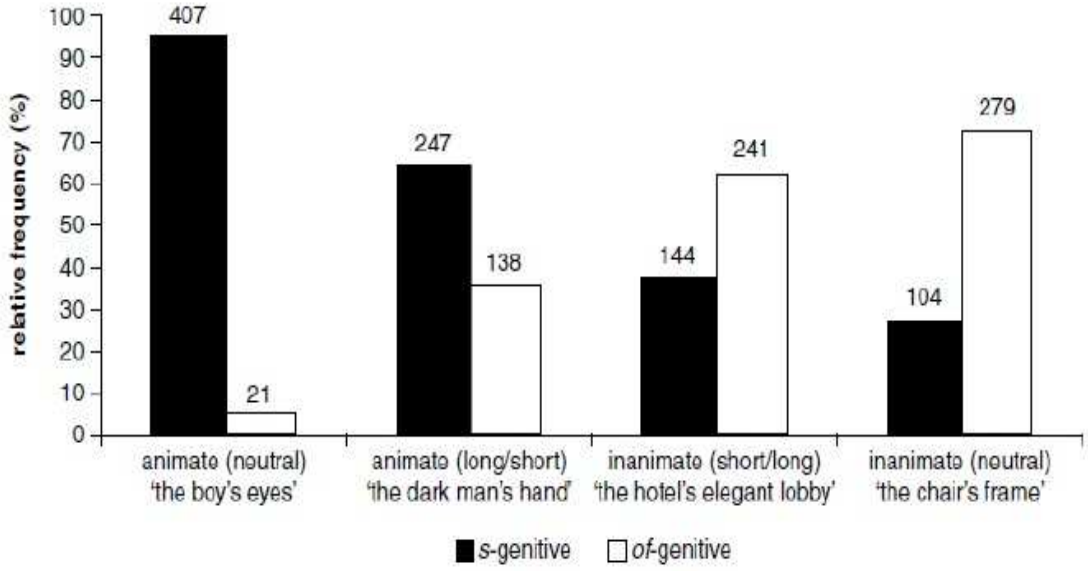


Figure 3: Relative frequency of the *s-genitive* and the *of*-genitive (from Rosenbach, 2005:620).

Rosenbach's (2005) study provides evidence from both a corpus analysis and an experimental study that **animacy** is a processing factor that influences the grammatical variation and it dominates the syntactic **weight**. She used a questionnaire based on a short text passage that provided context for both *s-genitive* and *of-genitive* constructions that included four conditional factors, two animate and two inanimate. The term "animacy" distinguishes between the animate referents and the inanimate referents. It has been defined as a dominant factor that causes variation between the *s-genitive* and the *of-genitive* constructions (Rosenbach, 2005). Her study includes 39 native speakers of American English, but the results are nevertheless interesting. As seen in Figure 3 **native speakers** of **English** prefer ***s-genitive*** constructions for **animate** possessors while ***of-genitive*** for **inanimate** possessors.

In our study we do not distinguish between animate and inanimate entities. In order to include this further distinction we would need additional data and time. Nevertheless, we understand that the results of Rosenbach's (2005) study are important for any further investigation related to the results of the present study. More details are presented in sections 7.4 and 7.5 where we regard the limitations of this paper and suggest possible avenues for future research. What is more, the outcomes of Rosenbach's (2005) study may not directly relate to our project but they do so inasmuch as they highlight yet another significant difference between the possession systems in English and in German (for the possession system in German see section 2.2.2). Indeed, it is this particular difference in the factors that influence the choice of the possessive form in English and in German that enables us to detect possible influences of the L2 English on the acquisition of L3 German but we will be looking into this in detail in section 6.3.

In the present study we are interested in investigating the use of the **synthetic/inflectional** vs. the **analytic/periphrastic** form of attributive possession. We decided to use the *possessive*

pronouns as a possession construction trigger in our sentence transformation task. That enabled us to check the L2 English learners preferences regarding both the other inflectional form (i.e. the *possessive –s*) and the periphrastic form (i.e. the prepositional phrase). Examples 4, 5 & 6 illustrate each of the aforementioned possessive forms.

4) This is his car. (inflectional/pronoun)[1]
5) This is my father's car. (inflectional/*possessive –s*)
6) This is the car of my father. (periphrastic/prepositional phrase)

2.2.2 Possession in German

It has already been mentioned that our study focuses also on another target language, namely, German. Therefore, we should now turn to the description of the expression of possession in German.

The first relevant point, that we would like to draw the reader's attention to, is that in **German** the expression of **possession** can be realised by **four** different **structures**. In English there are three forms of expressing possession, namely the *possessive pronoun* (ex.5), the *possessive –s* (ex.6), and the *genitive –of* (ex.7). The system of possession expression in German includes yet another inflectional form, that is, the *genitive case* as illustrated in the following example:

8) Das Auto meines Vaters.
   The.nom car.nom my.gen father.gen.

Another major difference is that in German there are certain **structural constraints** that determine the **choice of possessive form**. This condition is absent from the English possession

[1] Note that the students were told to rewrite the sentence using the NP *my father* instead of the possessive pronoun *his*. This example is actually taken from our sentence transformation task a full account of which can be found in Appendix 9.5 (A for English and B for German)

system. In English, as we mentioned in section 2.1, the choice of the PR>PM and PM>PR orders and the subsequent use of the equivalent possessive form has been explained on grounds of either phonological, pragmatic, syntactic and morphological factors (Hawkins, 1994; Quirk et al., 1985) or of cognitive and psychological factors (Taylor, 1989; Heine, 1997; Rosenbach, 2005). Rosenbach (2005) for instance claims that it is principally the animacy that prompts the use of a specific order (PR>PM or PM>PR) and thence the use of the equivalent possession form (see section 2.1 for more details on Rosenbach's study). In keeping with Rosenbach's (2005) findings we can claim that native speakers would prefer to use the PR>PM order and hence the *possessive –s* in those cases where the PR is an animate entity. In German, on the contrary, the PR>PM and PM>PR orders are not determined by conceptual factors. In that vein, both the PR>PM and the PM>PR orders are equally used. This difference is illustrated in the examples below.

English	9) My father's car.	PR=animate entity → PR>PM order
German	10) Vaters Auto. Father.gen car.nom	PR>PM order
	11) Das Auto meines Vaters. The.nom car.nom my.gen father.gen	PM>PR order
	12) Das Auto von meinem Vater. The.nom car.nom of my.dat father.dat	PM>PR order

Nevertheless, German, **unlike English**, does not allow the ***possessive –s*** to be **attached** to any noun. In German the *possessive –s* is affixed **only** to **proper names** and a **few kinship terms** (Eisenbeiß 2009) as examples 13.1, 14.1 & 14.2 below illustrate. Note that whenever the name ends in *"s"* an apostrophe is added instead of the *possessive -s*. This is the only difference between examples 14.1 and 14.2 below.

13.1)Vaters Auto. Father's car.nom	PR + possessing –s + PM PR: kinship
14.1)Marias Schwester. Maria's sister.nom 14.2)Tobias' Bruder. Tobias' brother.nom	PR + possessing –s + PM PR: proper name b) does not end in s → *possessing –s* c) ends in s → *apostrophe*

However if the **PR**, realised by either a **kinship term** or a **proper name**, is **modified** by an article or any other modifier, **then** the PM>PR order is required and hence the **use of** the ***possessive –s*** is **incorrect**. That is to say, if the PR in example number 13 was "my father" instead of "father" then we should use the PM>PR order and hence either the *genitive case* or the *prepositional phrase* would be required in order to express possession. Both cases are exemplified below.

13.2) Das Auto meines Vaters. The.nom car.nom my.gen father.gen.	PR: modified kinship term → PM>PR order → *genitive case*
13.3)Das Auto von meinem Vater. The.nom car.nom of my.dat father.dat	PR: modified kinship term → PM>PR order → *prepositional phrase*

At this point we would like to foreground that in German articles agree in number, gender and case with the nouns that they accompany. Nouns can also undergo some modification when the possession is expressed by means of the *genitive case.* Indeed, when the noun has a masculine or a neutral grammatical gender, then the suffix *–s/-es* is added in the formation of the genitive as seen in example 13.2 above.

Tables 2 and 3 illustrate all the cases of the expression of possession in German that we have described up to here.

1) PR>PM order	
Possessive -s	
a)Vaters Auto Father's car	PR + possessing –s + PM PR: kinship
b) Marias Schwester Maria's sister c) Tobias' Bruder Tobia's brother	PR + possessing –s + PM PR: name b) does not end in s → possessing –s c) ends in s → apostrophe

Table 2: Structural constraints: the use of the *possessor –s*

2) PM>PR order	
Genitive Case (GC) a) Das Auto meines Vaters. PM Pron:GC PR:GC The.nom car.nom my.gen father.gen b) Der Computer der Frau. PM Def.Art:GC PR:GC The.nom computer.nom the.gen woman.gen c) Das Fenster eines Autos. PM Indef.Art:GC PR:GC The.nom window.nom a.gen car.gen	(a): [PR= masculine → -s/-es suffix] + [PR = kinship noun + modifier → possessive –s ≠ possible] (b): PR = feminine → no modification applied on the noun. (c)]: [PR= neutral → -s/-es suffix] + [PR = common noun→ possessive –s ≠ possible]
P.P.: von + Dat. a) Das Auto von meinem Vater. PM + PP= von + NP: mod.=Pron(dat) The.nom car.nom of my.dat father.dat b) Der Computer von einer Frau PM + PP= von + NP: mod=	(a): PR = kinship noun + modifier → possessive –s ≠ possible

Indef. Art.(dat) The.nom computer.nom of a.dat woman.dat c) Das Fenster vom Auto PM + PP= von + NP: mod= Def. Art. (dat). The.nom window.nom of+the.dat car.dat	(c)]: PR = common noun→ possessive –s ≠ possible

Table 3: Structural constraints: The use of the *genitive case* or the *prepositional phrase*

To **sum up** chapter 2 regards the linguistic features studied in this paper. In the same line, **section 2.1** contains information about the morphemes in general and the **grammatical morphemes** that have been the pivot of our morpheme order study (i.e., the first of the three studies included in this paper). On the other hand, in **section 2.2** we outlined the basic aspects of the **possession expression systems in English and in German**. We understand this theoretical background to be important for two reasons. First, because the second study of this project focuses on the use of the *possessive –s* by L2 English language learners. As we will see in the following section the majority of the morpheme order studies have identified low accuracy rates for the *possessive –s.* This prompted our interest to study further the use of the inflectional form of possession especially because English is an analytic language. In the same line we wished to check the frequency of use of the synthetic (inflectional) possessive forms by learners whose target language is German. We chose German not only because it is a synthetic language, but also because it is also a West Germanic language. That induced the third study of our project in which we check the frequency of use of each of the possessive structures by our L3 German learners. The accurate knowledge of the structural restrictions outlined above is crucial in order to recognise the correct and the incorrect uses of the various possessive forms by our L3 German learners. Additionally, the

theoretical background that we present in section 2.2 enables us to recognise the commonalities, but also to distinguish the differences that the German and the English possession systems present. Indeed, these two systems are different not only in terms of the structures available for the expression of possession, but mainly regarding the way(s) that these forms are used. As we saw in sub-section 2.2.2 in German there are some structural restrictions regarding the choice of each possession form. The latter is a highly important point, as it is the one that defines the area where we actually see the influence of the previously learnt languages (L2 English) on the acquisition of a new foreign language (L3 German). Nevertheless, we will deal in depth with this issue in section 6.3 where we present the relevant results and comment upon their significance and possible explanation.

3 Literature review

This chapter is dedicated on the **brief review** of some major **previous studies relevant to** the research lines of **our project**. The **first section** focuses on the ones that relate to our principal study, that is, the **acquisition of some grammatical morphemes** by **L2 English** language learners. Since this is a historical overview we have decided to also include the two major works that regard the acquisition of morpheme in L1 English. Our decision is justified on the basis that these generated all the subsequent MOS. Furthermore, the L1 morpheme acquisition studies, and especially **Brown's** (1973) **study**, include elements that we have incorporated in our project, such as the use of **corpus** and the **list of functors** studied. In keeping with this decision we will start the first section of this chapter with a sub-section on the L1 MOS, which will be followed by a separate sub-section (3.1.2) in which we outline the major findings in relation to the acquisition of morphemes in foreign languages. Nevertheless, since the MO studies have received a lot of critique we have decided to include an additional sub-section (3.1.3) where we outline the major voices of criticism on the previously conducted MOS. The information found in this sub-section is also vital since it has guided our decisions regarding our data collection instruments and the scoring methods.

Decisions made on methodological issues are crucial and can determine the rigour of a study. In keeping with the perceived importance of the data elicitation method, we recognise the value of authentic learner language in SLA research. In the same vein we have dedicated a separate section within this chapter (**3.2**) to our principal elicitation instrument, that is, the **learner corpus**. Note that in section 3.2, except for occasional comments, we do not provide information regarding our own learner corpus. A full account of the latter along with

additional data regarding the instruments used and the procedures followed in the present study can be found in chapter 5.

The **final section** of chapter 3 regards the main studies on the influence of any previously learnt language on the acquisition of a new foreign language. Section 3.3 thus provides us with a **review** of the **studies** that have been conducted in relation to the other focal point of our study, that is, the **influence of** the **L2 on** the **acquisition of** an **L3**.

3.1 Morpheme order studies (MOS)

3.1.1 Research on L1 morpheme acquisition order

As we said in the introduction, morpheme acquisition research was partly prompted by the debate over behaviourist and nativist theories of language acquisition. The starting point was **Brown's (1973) longitudinal study** of three children who were native speakers of American English. Data for two of the children were collected over a five-year period, whereas for the third child the data were collected over a period of one year. At the beginning of Brown's study, the first two children were 27 months old, whereas the third was only 18 months old. For the analysis of the collected data, Brown studied the subjects' utterances in order to determine to which extent the grammatical morphemes in question were supplied in contexts where they were required or not. The underpinning idea was that certain contexts, also known as **obligatory occasions**, i.e. occasions when a native speaker is obliged to use particular morphemes, trigger the use of specific morphemes. The **accurate use** of the correspondent morpheme was thence seen as an **indication of** that morpheme's **acquisition**. As Brown (1973) put it:

> […] grammatical morphemes are obligatory in certain contexts, and so one can set an acquisition criterion not simply in terms of output, but in terms of output-where-required. Each obligatory context can be regarded as a kind of test item

> which the child passes by supplying the required morpheme or fails by supplying none or one that is not correct. This performance measure, the percentage of morphemes supplied in obligatory contexts, should not be dependent on the topic of conversation or the character of the interaction (p.255).

However, suppliance in obligatory contexts is not enough in order to decide whether a morpheme has been acquired or not. Another aspect that has to be taken into account is the **level of accuracy** of use that a learner must achieve in order to confirm the morpheme's acquisition. Brown (1973) set the level at 90 per cent on the grounds that it constitutes a level close to 100 per cent and corresponds to the level achieved by native speakers (Ellis & Barkhuizen, 2005).

Brown's study revealed a **similar order** of acquisition for grammatical morphemes in obligatory contexts for all three children although acquisition was not achieved at the same age, which indicates that the **route** (=order) of acquisition is similar to all children, but there may be variations in their **rate** (=speed) of acquisition. Table 4 presents the mean order of acquisition of L1 English morphemes according to Brown's findings. It is important to highlight that Brown (1973) was the first to use **corpus** data to study the L1 acquisition of morphology. The rest of the studies on L1 and L2 acquisition reviewed below, have used more controlled instruments. In this paper, we have also used L2 corpus data, in line with Brown's original study on L1.

ORDER	*MORPHEME*	*EXAMPLE*
1	Present Progressive '-ing'	He is eating
2 & 3	Prepositions 'in-on'	He is in the park
4	Plural '-s'	Two birds
5	Past Irregular	He went
6	Possessive '-s'	Daddy's book
7	Uncontractible Copula	She is good
8	Articles	The ice-cream
9	Past Regular	She looked
10	Third person singular '-s'	She looks
11	Third person singular irregular	She has
12	Auxiliary Be	She is walking
13	Contractible Copula	She's good
14	Contractible Auxiliar	She's walking

Table 4: Order of L1 Acquisition of English Morphemes (from Brown, 1973:275)

At around the same period **de Villers and de Villers (1973)** conducted a **cross-sectional study** in which they elicited spontaneous L1 speech data using **Brown's 14 functors** (see Table 4) and his coding rules to identify **obligatory contexts**. Speech samples were taken by 21 children aged 16-40 months. However, de Villers and de Villers (1973) did not set their cut-off point for acquisition at the 90% of accuracy. They simply ranked the functors according to the **relative accuracy** of use in obligatory occasions. Their findings were then compared to the acquisition order found by Brown (1973) for his three subjects. Their cross-sectional study revealed the **same order of acquisition** and came thus as a response to the critique that Brown (1973) based his conclusions on the observations of just three children.

3.1.2 Research in L2 morpheme acquisition order

As pointed out in the introduction the interest on morpheme acquisition orders was soon extended to L2 and second language acquisition research (henceforth SLA). As we will see most of the L2 studies in this field have used some variant of the obligatory occasion analysis.

Dulay and Burt (1973; 1974b) conducted two L2 morpheme acquisition studies. In their **first study** (1973) they studied three groups of **L1 Spanish child learners of English**. The first group, the 'East Harlem (New York)' group, received a bilingual (English and Spanish) education at school, but no formal instruction in English. The second group, the 'Sacramento (California)' group received only English education at school and formal instruction in English. The third group, the 'San Ysidro (Mexico)' group, crossed the border to attend an English school, but returned home daily where they spoke Spanish. Dulay and Burt **focused on** the acquisition of **eight** of the 14 **functors** proposed by Brown (1973) and found a **consistent order** across all the groups. Each obligatory context for a functor was scored according to the following schema (from Dulay & Burt, 1973:254):

> No functor supplied=0 (e.g. she's dance_)
> Misformed functor supplied=0.5 (e.g. she's dances)
> Correct functor supplied=1.0 (e.g. she's dancing)

The **accuracy score** for each functor was then calculated as a **ratio of** the **sum of** the **scores for each obligatory context** for that functor across the whole group. According to this model we could then calculate the accuracy rate for the three examples given above. In this case the total number of examples determined the OC which is hence equal to three. Applying the calculation process described above we end up with the following:

$$\frac{(1x0)+(1x0.5)+(1x1)}{3} \times 100 = \frac{1.5}{3} \times 100 = 50\%$$

Although each of Dulay and Burt's (1973) groups showed different rates of accuracy, they all followed the **same route**. Accordingly, they concluded that L2 morpheme acquisition order is not influenced by **exposure** to the target language (henceforth TL).

In 1974 Dulay and Burt conducted a **second study** (1974b) observing two groups of 6-8 year-old children learning English as a second language in the USA. The first group consisted of 60 children **native speakers of Spanish and** the second consisted of 55 children whose **L1** was **Cantonese**. In this study the authors found a **similar developmental pattern** for both groups **regardless** of their **L1**.

Like most L2 morpheme studies and unlike L1 morpheme studies, Dulay and Burt used controlled tasks. This decision was based on the fact that elicited L2 language is rarely spontaneous. The instrument they used was the **Bilingual Syntax Measure (BSM)** which is a syntax-based test of L2 proficiency designed for use with young children. The BSM was not designed to test order of acquisition but rather proficiency level. Nevertheless, according to Dulay and Burt (1974b), the **value of this method** is that the researcher "[…] can look to see how the child forms simple finite clauses (word order, gender, number and case for the pronoun, agreement for the verb, the form of the qualifier, etc)" (p.40). Additionally, the BSM was used as a measure by most researchers and that facilitated comparisons. The instrument consists of some **cartoon pictures and questions**. These prompts are used to elicit roughly predictable responses that include various obligatory contexts for grammatical morphemes. For instance the researcher may point to a very fat cartoon character and ask: "why is he so fat?" The expected answer is "Because he eats too much/a lot". Figure 1 exemplifies a case of the BSM. **Children's errors** were then placed into three categories: **developmental, interference, and unique**. Dulay and

Burt (1974b), based on the fact that the majority of errors fell in the developmental category, hypothesised that second language acquisition is similar to first language acquisition.

Figure 1: Example from the Bilingual Syntax Measure (from Dulay and Burt, 1974b:37-53)

Bailey et al. (1974) generalised the results of Dulay and Burt's studies to **adults** learning an L2 and found a similar order of acquisition for the same set of English morphemes. They studied 73 adults of **distinct L1 backgrounds**, 33 were native speakers of Spanish and 40 had different first languages (Greek, Persian, Turkish, etc.). They also used the **BSM** as elicitation method. The developmental patterns of both groups correlated significantly among them and with the patterns found in children learners of L2 English in Dulay and Burt's studies. **Krashen (1977)** suggested later that is was possible to group certain morphemes together on the basis of accuracy and create thus an **acquisition hierarchy** as shown in Figure 2. This organisational pattern states that the order of acquisition will progress from one group to the other, but makes no comment in relation to the acquisition order of the morphemes within each group.

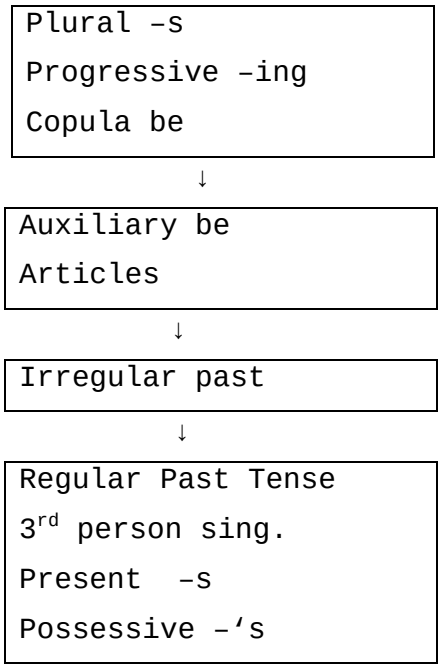


Figure 2: Order of Grammatical Morpheme Acquisition for L2 Learners of English, Krashen (1977)

In **1975 Fathman** carried out a **cross-sectional study** in order to investigate the relationship between **age** on one hand and rate as well as order of acquisition of English structures on the other. She studied 140 children aged 6-15 years, who had been in the USA for less than three years (70 had been there for one year, 40 for two years and 30 for three years) and had all been immersed in the verbal environment of the school. Fathman (1975) introduced the **Second Language Oral Production English (SLOPE)** test, which has twenty sections (each involving three items) designed to test 20 different grammatical phenomena (article, negation, wh-questions, etc.). The test usually consisted of two pictures and one question. For example, in order to elicit *plural –s* the researcher points to a picture of a boy and says "Here is a boy", then to a picture of two boys and asks "Here are two … ?". Both the **SLOPE** test and the **BSM** aim to the production of **spontaneous language** related to the L2 acquisition of grammatical knowledge. Fathman (1975) divided her subjects into two age groups (6-10 & 11-15 years) and then observed the changes between these two groups in terms of rate and order of acquisition of 20 morpheme categories or syntactic patterns.

Additionally, her subjects described orally a composite picture for which they were rated for correctness of grammar, pronunciation and general fluency. Fathman (1975), based on the results of her study, concluded that L2 acquisition process changes with the **age** in terms of **success in learning**, with the younger children showing higher pronunciation abilities and the older children higher morphological and syntactic abilities. In terms of order of acquisition the study revealed no changes and thence **Fathman** (1975) affirmed that **L2 acquisition order** remains **constant**.

Perkins and Larsen-Freeman (1975) performed a **cross-sectional** study of adults using multiple tasks in order to check the validity of the single task-based studies and found that the **order** of acquisition was similar to that of previous single-task studies. Indeed, they found significant correlations between the morpheme sequences produced by adult learners in the two **oral** production tasks and the morpheme sequences produced by child learners in Dulay and Burt's study. For the **written** task there were some differences such as a rise in the rank order of *plural –s* and *third person –s*. However, Ellis (1994) points out that these can be explained by the production conditions in the sense that speaking and writing are influenced by different sociolinguistic and psycholinguistic conditions, and they thence do not imply a different morpheme acquisition order. Additionally, Perkins and Larsen-Freeman (1975) investigated L2 acquisition in relation to the **type of instruction**. According to their study's results, when learners were exposed to naturalistic input, formal language instruction did not seem to affect accuracy orders. As Perkins and Larsen-Freeman (1975) put it: "[…] instruction does not radically alter order of acquisition" (p.241).

In **1978 Andersen** altered slightly the principal subject of **morpheme order**[2] research by **focusing on** the investigation of the

[2] Note that in this paper we will be using morpheme order and MO interchangeably. In the same line we use the acronym MOS when making reference to a Morpheme Order Study.

actual **explanation of the "natural order"** that previous studies had shown in both the L1 and the L2 morpheme acquisition. Unlike the explanations that had been proposed until then, mainly based on a nativist assumption, Andersen tried to account for the accuracy profiles of L2 learners in relation to the acquisition of **underlying syntactic knowledge**. He divided the data on the acquisition of English grammatical morphemes into two syntactic classes: **verb-related morphemes** and **noun-related morphemes** and analysed the accuracy rates accordingly. Although he based his data collection on the SOC model introduced by Brown (1973), he analysed the results using the **implicational scale** as scoring method. This technique is based on the idea that "[…] if a learner is accurate on a 'difficult' morpheme, she will also perform accurately on some 'easier' morpheme […] the reverse is not necessarily the case" (Hawkins, 2001:47). Andersen (1978) noted that the degree of difficulty could depend on the underlying syntactic properties of the morphemes, rather than on the actual morphemes. Said in other words, **Andersen (1978) argued that in order to use accurately a morpheme the L2 student must have previously acquired the underlying property of that the specific morpheme.** That is, the L2 learner must fully understand not only what a specific morpheme stands for, but also comprehend its underlying syntactic properties as such. In the same line, in order to accurately use the structure "*be + V-ing"* the L2 learner must have previously acquired what this morpheme represents, i.e. the progressive aspect and hence the difference between progressive and non-progressive (e.g. Mary is reading vs. Mary reads).

Makino (1979) attempted a morpheme order study with English learners in Japan in order to test the acquisition order hypothesis for **EFL learners**. He tested 777 adolescents using fill-in-the blank tests and stratified the data according to the subjects' grade level, English textbook used and location of the school (urban or rural). The order of acquisition correlated significantly with the orders obtained in most of the studies that had looked at ESL learners and thus supported the

hypothesis that all **L2 learners** follow the **same route** in terms of **morpheme acquisition despite** the **learning environment**.

Pica (1983) investigated the effects of **formal classroom instruction** on one hand **and** the influence of learning in a **naturalistic environment** on the other. The subjects of her study formed three groups. The first consisted of classroom-instructed adults, the second was formed by a group of adults who lived in an English-speaking environment in the absence of formal instruction and the third was a 'mixed' group that received both classroom instruction and input from native speakers in a naturalistic environment. All three groups had the same L1, namely Spanish. Ellis and Barkhuizen (2005) argue that Pica's study is "methodologically more sophisticated […] in that she examined learners' oversuppliance […] through a **target-like use** analysis as well as suppliance in obligatory contexts[3]"(p.87). Pica (1983) concluded that "different **conditions of exposure** to L2 English **do not** significantly **alter** the **accuracy order** in which grammatical morphemes are produced" (p.465). **Conditions of exposure**, however, do **influence** the **learners' interlanguages**. Pica found that learners with instructed exposure oversupplied morphemes in non-obligatory, and hence inappropriate, contexts. By contrast it was omission errors that were higher among those with naturalistic exposure. That indicates that different conditions of L2 exposure affect learners' hypotheses about the target language and their strategies for using it (Pica, 1983:495).

Lightbown (1983) in an attempt to determine what **factors** may affect the order of acquisition performed a study in which she investigated **frequency as putative determinant**. She looked at the relationship between the frequency of appearance of certain forms in the classroom and the frequency of their accurate use. Her subjects were instructed learners of English in Quebec aged between 11 and 17 years. They were asked to perform an **oral task**

[3] More details regarding the Suppliance in Obligatory Context (SOC) and the Target Like Use (TLU) are provided in section 3.1.3.

twice in two consecutive years. The results suggested that there is **no direct relationship** between input frequency and the accurate use of the morphemes.

After Andersen's (1978) study, most researchers were looking to establish not only a **descriptive** but also a solid **explanatory** model. That **tendency** is especially **prominent** in the studies conducted **after the 1980s**. As mentioned above, some explanation of the observed phenomena regarding order of acquisition had been previously attempted as well. For example Brown (1973), following a nativist perspective, argued that "some factor or some set of factors caused these grammatical morphemes to evolve in an approximately consistent order in these children" (p.272). In the same vein, during the 1970s a series of **putative determinants** have been suggested by L1 and L2 researchers including the following: (a) perceptual salience (Larsen-Freeman, 1976), (b) morphophonological regularity, (c) syntactic complexity, (d) frequency (Larsen-Freeman, 1976), (e) semantic complexity (Larsen-Freeman, 1976; Andersen, 1978), (f) native language transfer (Andersen, 1978), (g) individual variances (Andersen, 1978).

Ellis and Barkhuizen (2005) recognize two key studies in this account after the change of the researchers' focal point from description to explanation. The first is the **study** by **Zobl and Liceras (1994)** and the **second** is the one conducted by **Goldschneider and DeKeyser (2001)**.

Zobl and Liceras (1994) observe that **L1 acquisition order** is characterized by **nominal categories preceding verbal categories**, whereas **L2 acquisition order is cross-categorical**. They suggested that this is explained by abstract cognitive-linguistic principles. Indeed, they argue that the Nominal → Verbal pattern observed in L1 acquisition stems from the gradual maturation of **functional categories** in young children. Thence the difference in L1 and L2 acquisition patterns can be

explained by the fact that functional categories are already available from the beginning in L2 learning.

In keeping with this preference to find an explanation based on purely linguistic criteria, **Hawkins (2001)** has studied the patterns of acquisition order following the division between **verb-related** and **noun-related morphemes** put forth by Andersen (1978). **Hawkins (2001)** in his interpretation of Andersen's (1978) results **argues** that the scale of difficulty of the **verb-related morphemes** exhibit the addition of the following syntactic properties:

> **Copula → aspect (±progressive) →tense (±past) →subject-verb agreement (±3rd person singular)** (Hawkins, 2001:48).

Hawkins (2001) claims that this result is interesting mainly because it poses certain questions the answer to which may help our understanding of L2 morpheme acquisition patterns. As he put it "Why should a copular construction (John's hungry […]) become established in the mental grammar more accurately, apparently, than an aspectual one (John is cooking) […]" (Hawkins, 2001:48). He also suggested a working theory, called "**modulated structure building**" (2001:73), which includes insights from two theories known as "minimal trees" and "full transfer/full access". According to this model learners' initial L2 grammars consist of lexical projections like VP, NP, AP, PP and these have the structural properties of their L1 grammars (i.e. the position of the head, complement and specifier are initially determined by the L1). Initial L2 grammars consist of **lexical projections** only in principle since restructuring towards the L2 may be very rapid. **Functional projections** are **established later** than lexical projections (i.e. the establishment of English inflection and its projection IP appear after a VP-only stage). Finally, it is only once functional categories are established in the L2 grammar that the influence of L1 functional categories becomes evident. For example in initial stages L1 verb-subject agreement pattern is not relevant "because learners are not at the point

of development where they need to have a representation for specifier-head agreement" (Hawkins, 2001:74). But, as soon as they reach that point their L1 becomes influential and hence Spanish learners have an asset and indeed produce more accurately specifier-head agreement than Japanese students whose L1 does not have such agreement as the study by Stauble (1984) confirms. That is why Hawkins calls his structure building pattern **'modulated'**, because it "is **influenced** by **properties** of the **L1** at the **relevant point** in the **construction of** a **grammar** and not before" (Hawkins, 2001:74). The influence of the L1 on the acquisition of the TL has been criticised, as we will see later.

Goldschneider and DeKeyser (2001) carried out a **meta-analysis** of all the accuracy results obtained for oral production data in **twelve studies** involving 924 learners. They then suggested that the L2 acquisition order pattern can be explained by a combination of **five putative determinants** (i.e. functors features): perceptual salience, semantic complexity, morphophonological regularity, syntactic category and frequency. Other external factors such as L1 transfer could also be influential but their multiple-regression study did not allow them to take it into consideration. Goldschneider and DeKeyser **highlight** that **only** a **combination of variables** can **account for** the **"natural order"** and remark that the revision of these 12 studies concludes that the commonalities in their findings cannot be ignored. Additionally their study suggests that "L2 acquisition is the product of an interaction between the learner's internal mechanisms and the input" (Ellis & Barkhuizen, 2005:77). As **Hawkins and Lozano** (2006) put it: "The picture of **SLA** that emerges is one of a **complex interaction** among **innate knowledge**, previous knowledge from the **L1**, and **input** from the **L2**" (p.73).

In **2006 Muñoz** carried out a **cross-sectional study** in order to check the effects of the **age of onset**, i.e. the age in which the subjects started learning the L2, the **amount of exposure** to the

target language and the **proficiency** levels on both the order and rate of acquisition of English as a foreign language. She studied 6 groups of instructed learners of English who were asked to perform two oral tasks, a picture-elicited story and a semi-structured interview. Table 5 shows the composition in terms of onset age, time of exposure to the L2, test age and number of participants.

	Group AAO = 8		*Group BAO = 11*		*Group DAO = 18+*	
Time 1 200 h	A1	AT = 10;9 *N* = 30	B1	AT = 12;9 *N* = 30	D1	AT = 28;9 *N* = 20
Time 2 416 h	–	–	–	–	D2	AT = 30;4 *N* = 15
Time 3 726 h	A3	AT = 16;9 *N* = 30	B3	AT = 17;9 *N* = 30	–	–

Table 5: Groups' composition (Muñoz, 2006:116)

Muñoz (2006) performed an **Obligatory Occasion analysis** of the data and scored the responses following Pica's (1983) guidelines (2 points for the suppliance of the correct functor, 1 point for each incorrect functor supplied and 0 points when no morpheme was supplied). The results of her study confirmed the findings of earlier "natural sequence" studies in three points (Muñoz, 2006: 121-123): (a) that a **similar route** is to be found independently of learner's age, although the **age** seems to **influence** the **rate of acquisition** since older learners showed higher rates of accuracy in the use of morphological functors especially in initial stages of language acquisition, (b) that **proficiency level** plays a bigger role than first language and (c) learning context does not affect accuracy orders as much as proficiency level does, at least not until the learners have progressed beyond the very elementary levels of proficiency. As she put it "foreign language learners present accuracy orders that approach the average order once they have had a certain amount of exposure to the target language and have progressed beyond the very elementary levels of proficiency" (Muñoz, 2006:123). Proficiency has been proved influential in the acquisition of derivational morphology too. Muñoz's finding is

thus in keeping with other studies (Lardiere, 1995; Lardiere & Schwartz, 1997 cited in Muñoz 2006) which argue that L1 affects the acquisition of derivational morphology “in early L2 development, but this influence disappears with proficiency” (Hawkins & Lozano, 2006:71).

3.1.3 Identification of gaps in the literature

Morpheme order studies have been **questioned** in relation to some methodological presuppositions and techniques.

A fundamental point of the overall critique relates to the basic assumption of morpheme studies that **accuracy equals acquisition**. It has been argued that, from a conceptual point of view, accuracy cannot be equated with acquisition. That is, suppliance of a morpheme does not imply necessarily that the morpheme has been acquired. As this is a two-fold criticism we will present its points separately.

The first inadequacy seems to stem from the incapability of the SOC to take into consideration a number of data that are also significant in terms of language acquisition. The fact that the learner **uses** a morpheme **correctly** in an obligatory context does **not imply** that the **learner** has **acquired** that particular morpheme **because** the subject may as well incorrectly **overuse** it in other occasions. That led scholars to establish the concept of **target-like use (henceforth TLU)** which implies a weighted scoring as well. **TLU takes into account correct suppliance in both obligatory contexts and non-obligatory context**. Researchers can thus account not only for the correct uses (TLU) but also for the incorrect uses, i.e. the **non target-like uses (henceforth NTLU)**. The NTLU includes the cases of **underuse**, **misuse**, and **overuse** and thus enables the researcher to tag and analyse a greater variety of morpheme uses. Therefore, scholars can draw a more **complete image** of the process of grammatical morpheme acquisition.

At this point we would like to give an **example** to illustrate the **difference** between the **SOC** and the **TLU models** in terms of the **results** obtained. Suppose that we wish to measure the accurate use of the *progressive –ing* morpheme in the following text:

> "I am now reading the second book of Harry Potter. Yesterday I finishing the first book. In general, I reading a lot."[4]

Following **Brown's SOC model** we should check for the **correct suppliance** of the *–ing* **in obligatory context**, which is only one in our example, and then **divide** the number of the correct suppliance **by** the **number of** the **obligatory contexts**, which is, again, just one in our example (i.e., 1:1). We would then have to conclude that the **accuracy rate for** the ***progressive –ing*** morpheme obtained by this (hypothetical) L2 learner **is (1:1) x 100 = 100%**. According to this accuracy rate we would be bound to affirm that **our hypothetical L2 learner has acquired the *progressive –ing***. However, the learner's actual production clearly indicates the contrary.

The **TLU model**, on the contrary, enables us to take into account the remaining two cases where our hypothetical student has used the *–ing.* These instances reveal an incorrect use of the morpheme in question since they represent non-obligatory contexts for the *progressive -ing*. So, following the TLU model we would **divide** the **number of correct suppliance** of the *–ing* in obligatory context **by** the **sum of** the total number of **OC and** the number of **NTLU** cases, i.e. the incorrect use in non-obligatory contexts. **On these grounds** we would conclude that the accurate **rate of the *progressive –ing* is [1:(1+2)] x 100 = 33%**.

Clearly the difference between the results produced by Brown's (1973) SOC and Pica's (1983) TLU models is highly significant. Considering the huge influence of the scoring method on the actual results, as reflected in our example, along with the

[4] This example is totally invented.

rationale behind the TLU we have decided to use this model in the analysis of our data.

Resuming our account of the **critiques of** the first **MO studies**, we should say that the central point of morpheme studies **fails to cover** phenomena proper of the acquisition development of certain morphemes. Indeed, the first morpheme studies disregarded the fact that some morphemes display a **U-shape pattern** of development. That is particularly observed with the acquisition of ***regular*** and ***irregular past*** in English. At the initial stages of *irregular past* acquisition, learners go through an early stage in which they use the forms of the *irregular past* correctly. Based on the *suppliance in obligatory context* condition, the researcher should conclude that by that moment the learner has acquired the *irregular past* forms. However, such assumption would be wrong as we now know that right after that initial stage of correct use learners experience a stage of over-generalized *–ed* forms, during which they replace the supposedly previously acquired *past irregular* forms with the *past regular* morpheme. That is, at the beginning of the acquisition of the *past irregular* morpheme our hypothetical L2 learner would produce a sentence like *"Yesterday I went to school"*. In the immediately following stage, nonetheless, the same hypothetical L2 learner would *regularise* the past tense of the verb *"to go"* and hence produce the incorrect sentence *"Yesterday I goed to school"*. Obviously **longitudinal studies** do not face the danger of not accounting for these developmental stages since they are based on data that refer to and reflect each subject's acquisition process for longer periods. Cases like over-generalisation are thus evident in this type of studies. **Cross-sectional studies,** on the other hand, can easily leave the aforementioned phenomenon unaccounted for. However, if their subjects are grouped according to their **level of proficiency** and the accuracy rates calculated separately for each group, then the developmental stages we described above become evident in cross-sectional studies as well. As Ellis and Barkhuizen (2005) put it, grouping on the

basis of the subjects' proficiency level "[...] will enable researchers to investigate proficiency as a covariate of accuracy order and to identify which morphemes display a lower level of accuracy at higher levels of proficiency"(p.78). In line with this suggestion we have decided to group our subjects on the basis of their proficiency level and give thus a more accurate and complete image of the accuracy orders displayed for each morpheme at each level. We give a full account of this procedure as well as of the way we measured our subjects' proficiency level in section 5.3.2.

Another criticism relates to the **small number of morphemes** that have been investigated in most morpheme order studies. Most morpheme studies have followed Brown's (1973) methodological guidelines and thence have investigated the whole or part of his list of functors. **Ellis and Barkhuizen (2005)** argue that the **set** of morphemes **could be expanded** but that would depend on whether the **instrument** used could ensure "a sufficient number of obligatory contexts for each morpheme" (p.78).

In the same line, both the SOC and the TLU studies have been **criticized** for **focusing** mainly on the **grammatical morphemes** and "thus ignoring the functional use of the language" (Muñoz, 2006:109). Along this line of thought another critique on natural order studies was developed. According to this, MOS should not extend the implications found for a set of morphemes to acquisition in general. We would like to point out that a study based on corpus data, presents certain advantages in this respect since it provides the researchers with additional information and hence enables them not only to regard the functional nature of language but also to extend their study to the other aspects of acquisition (e.g. regarding the stylistic aspects of language use).

Additionally, some categories include various **disparate features** such as case features, features of the verb phrase, features of the noun-phrase, etc. For example, all morpheme studies place

indefinite and *definite articles* in the same category when they do not constitute discrete features but rather cluster of forms. Nevertheless, this problem can be overcome though a methodological adaptation in order to group the morphemes on grounds of a specific principle. **J.D. Brown (1983)** suggested that morphemes should be classified into **open and bound classes**. **Andersen (1978)** distinguished into **verb-related and noun-related classes** and demonstrated thus an implicational ordering within each class as explained in the previous section.

Finally, morpheme order studies and their results have been questioned in terms of their **validity** since **most** natural order studies were **cross-sectional** and used the **BSM** (Bilingual Syntax Measure) as data elicitation instrument. Larsen-Freeman (1975), for example, believed that the sequence was an artefact of the BSM. Indeed, Rosansky's longitudinal study (1976) shows that the relative accuracy that has been obtained in cross-sectional studies does not predict the sequence of acquisition for groups of individuals. **Krashen (1978)**, on the contrary, **argued** that **spontaneous language** will always **reveal** the **order of acquisition**, whereas **language elicited** by means of **controlled activities** such as discrete-point grammar tests, in which learners' formal knowledge is tapped, **will show** the **order of learning**, which according to Krashen will be different to the order of acquisition. We should mention, however, that data obtained from children (Fathman, 1975; Kjarsgaard, 1979) and adults (Krashen, Sferlazza, Feldman, & Fathman, 1976) showed that for structures present in both the SLOPE and BSM a similar order was found despite the differences in the task and scoring method.

3.2 Learner Corpora and Learner Corpus Research

As we saw, Morpheme Order Studies were criticised on the grounds that their findings were based on very few samples of L2 learner language elicited in many cases through very controlled means. Recently the **use** of corpus known as **learner corpus** has been introduced in the **SLA research** providing thus a more **ample**

database of **language** produced by **L2 learners**. In that sense, we understand that learner corpora represent a very useful instrument of data collection. Additionally, learner corpora contain **authentic L2 production** which presents a twofold advantage. First, it reflects more accurately the acquisition process. Second it enables the researcher to account for other aspects of the language such as the functional or the pragmatic. In line with this thought we have decided to compile our own learner corpus, a detailed analysis of which can be found in chapter 5. In this section we will limit ourselves to the **review** of **language corpus** in general **and learner corpora** in particular.

Before proceeding with the actual description of the learner corpus, we should define what a language corpus is. According to **McEnery (2003)** a **language corpus** is:

> a **well-organized collection of data**, collected within the boundaries of a sampling frame designed to **allow** the **exploration of** a certain **linguistic** feature (or set of **features**) **via** the **collected data** (p.449).

The term learner corpus was first used for Longman's learner dictionaries, in which the information on EFL learners' common mistakes was provided, based upon the Longman Learners' Corpus. In 1990 a project called the International Corpus of Learner English (ICLE) was launched as part of the ICE (International Corpus of English) in order to collect L2 data (Granger, 1998 cited in Tono, 2000:124-125).

Learner corpora are defined by Granger (2008) as "[…] **electronic collections of texts produced by language learners**" (p. 259). Hence learner corpora differ from language corpora in that the data come from foreign language learners. It is important to highlight that the learner language included in the learner corpora reflects language learnt in **instructional environments**, not in naturalistic ones. That is, the learners do not live in a

country where the target language is spoken.[5] However, further classification in relation to the speakers involved and the type of data they produce is not only possible but also necessary.

Regarding the **status of the learners involved** two comments can be made. First, language learners whose linguistic performance is included in learner corpora are **foreign language learners**. That, apart from the implications outlined in the previous paragraph, also means that the data included in English learner corpora cannot come from speakers that belong to Kachru's (1985) outer circle (e.g. India or Nigeria where English has achieved the status of official language and/or language of education or administration). Secondly, the **distinction** made sometimes between **learner corpora** and **ELF** (English as a Lingua Franca) corpora is based on the "[…] researchers' orientation towards the data and the purposes they intend the corpora to serve" (Seidlhofer, 2004:224). That is, if the subjects are still in the process of learning then the researcher's focus will be the progress of the speaker's interlanguage and the gap that needs to be filled in order for the learner to become proficient speaker. On the contrary, if the subject is a proficient non-native speaker of English then the focus will be more on how they are able to communicate (Seidlhofer, 2004).

In terms of the learner data the most distinguishing feature that has been put forth by Sinclair (1996) is **authenticity** in the sense that they represent material gathered from the genuine communication of people. This feature raises a problem for learner corpus data as learners do not often use a foreign language in their genuine communication with other people. In order to overcome this obstacle, Granger (2008:261) suggests we should include in learner corpora only the learner production data that display a medium degree of naturalness such as that produced by picture description, summary or translation. In

[5] Since our corpus is a learner corpus and hence the data collected represent by definition language acquired in instructional environments, we will be using L2 and FL interchangeably in the chapters dedicated to the analysis of our data, the discussion and the conclusion. The same is true for the terms L2 learner and FL learner.

keeping with the importance of natural language for SLA research and following Granger's advice we have decided to compile our learner corpus based on language elicited by means of a picture description task. More information is provided in sections 5.1.1 and 5.3.3.

According to Granger (2008: 261-263), **learner corpora** can be **divided** in various types according to the underpinning criteria one uses for their typology. On these grounds learner corpora can be:

i. Commercial or academic

Commercial learner corpora are started by major publishing companies, whereas academic learner corpora are compiled in educational settings. In that sense, our learner corpus is academic.

ii. Big or small

Learner corpora that contain millions of words are naturally considered big and they do present an asset regarding representativeness of the data. However, small learner corpora that contain thousands or hundreds of words also are valuable. As pointed out by Granger (2008) "[a] detailed longitudinal study of one single learner is of great value if the focus is on individual language development" (p.262). The learner corpus used in our study is relatively small but it is our own learner corpus and this has other advantages such as the learner profile that we have included for each of our subjects and which gives us additional information (e.g. the learner's mother tongue, the hours of tuition received, additional FL spoken etc.; for more details please refer to sections 5.1.1 and 5.3.3.).

iii. English or non-English

English learner corpus is by far the largest collection of learner corpora nowadays. Only the International Corpus of Learner English (ICLE) contains 2.5 millions of words and covers learners from 11 different mother tongue backgrounds.

iv. Writing vs. speech

There are more written than spoken learner corpora. This is not surprising if we take into account that the major difficulties of "[…] collecting and transcribing speech is multiplied by a factor of 10 in the case of learner data […]" (Granger, 2008:263). In our study we have also opted for the written learner corpus, because we consider that it has certain advantages. First, it enables the researcher to collect a considerably larger amount of data in less time. Second, oral samples are not always clear for all the participants or in every environment. Finally, written corpora have an advantage bearing on the learner reaction to the means of data collection. As Ellis and Barkhuizen (2005) put it "[…] the presence of a cassette recorder may induce self-consciousness in learner's speech, thus making it less likely that the resulting samples will reflect their vernacular style" (p.27).

v. Longitudinal or cross-sectional

According to Granger the majority of learner corpora that focus on interlanguage are cross-sectional, that is, they contain data collected by various learners at a single point in time. Genuine longitudinal corpora, i.e. data from the same learners collected over time, are very difficult to compile. Therefore, researchers prefer the so-called **quasi-longitudinal corpora** that contain data gathered at a single point in time but from learners of **different proficiency levels**. In the same line we have opted for the compilation of a quasi-longitudinal corpus insofar as we have included language produced by learners of different proficiency levels (see section 5.2 for more details on the synthesis of our subjects).

vi. Immediate vs. delayed pedagogical use

Corpora compiled for delayed pedagogical use are not used directly as teaching/learning materials. They rather provide us with insights into learners' interlanguage and/or into the pedagogical tools used, which are thought to benefit similar-

type future learners. Learner corpora for immediate pedagogical use are, on the contrary, used directly by the learners who produced them.

Although the use of learner corpora in SLA and language teaching is quite recent, there are several reasons that explain researchers' growing interest in this field. Below we present the ones that we consider central and which justify our decision to compile and use our own learner corpus for the first part of our project that regards the morpheme accuracy rates obtained by our L2 English learners.

i. Learner corpora respond to the necessity of obtaining **quantitative data in SLA** (McEnery and Wilson, 1996:18). According to Granger (2009:16) one of the main assets of learner corpora is that it provides the SLA research with a much wider empirical basis. Additionally, by being systematic, and thus include data collected according to a number of criteria, learner corpora allow for diversified material (Díaz Negrillo, 2007:85). Therefore **learner corpora**, i.e. large systematically compiled databases of learners' language production, can be established as **representative** and thus **support generalisations**.

ii. Learner corpora contain basically **authentic language** and hence contextualised discourse. On one hand, this feature enables researchers to observe various aspects of learners' interlanguage (Granger, 2009:16). On the other hand, as Housen (2002) remarks, learner corpora make it possible to "empirically validate previous research findings obtained from smaller transcripts, as well as to test explanatory hypotheses about pace-setting factors in SLA" (p.108).

iii. Learner corpora are **computerised** and the amount of data available in machine-readable forms increases every year. That facilitates researchers' studies in many ways. First, it provides them with great amount of learner language data as seen by the fact that the CHILDES corpus is used in 3200 research

papers (Lozano & Mendikoetxea, forthcoming 2012). Secondly, computerization not only makes it possible for the researchers to manage large amounts of data, but it also provides them with tools of analysis. As Tono (2000) put it:

> The use of learner corpora opens up the possibility of filling the gap between small-scale, tightly controlled experimental research and large-scale, but impressionistic, survey-questionnaire type research (p. 132).

In line with this idea we have decided to found the principal study of our project on our own learner corpus. This has provided us with larger amounts of learner language, which, as Granger (2009:16) has claimed, provided us with a solid empirical base on which we observed the acquisition process of certain morpho-syntactic features. The fact that the elicitation method was systematically designed enabled us, as Díaz-Negrillo (2007:85) points out, to reach conclusions that we can claim valid for a relatively large amount of L2 learners. The importance of founding our research on authentic language has been emphasized many times in the previous section. We shall, therefore, simply repeat here that authentic language offers us insights into the learners actual interlanguage the description of which is crucial for SLA research in that they "[…] provide the evidence by which theories of L2 can be developed and tested"(Ellis and Barkhuizen, 2005:21). In this paper we have used our learner corpus to check the morpheme acquisition orders suggested by some previous MO studies. In the same line, we have tried to check the influence of the data collection method on the learner language by comparing the results of our corpus to the ones obtained by means of a controlled activity bearing on the *possessive –s*. In sections 5.1 and 5.4 we provide more information regarding the controlled data elicitation method used in our project.

3.3 L2 influence on L3 acquisition

At the beginning of this chapter we stated that we aim to review the major studies previously conducted in relation to every study of our project. Therefore, in section 3.1 we outlined the outcomes of some major MO studies and in section 3.2 we described the main characteristics of learner corpora emphasizing its importance as an instrument for SLA research. In the same line of thought, we have decided to devote the last section of chapter 3 to the review of the **major findings** regarding the **influence** of any **previously learnt language on** the **acquisition of a new foreign language**. As pointed out at the end of chapter 2, the last study of our project regards the influence of L2 English on the acquisition of the expression of possession in L3 German. Our initial idea was to check the accuracy rates of the *synthetic possessive forms* in German by learners of German as a foreign language and then compare it to the corresponding accuracy rates of the *possessive –s* by L2 learners of English. However, the results of our study indicated a possible influence of the English system of possession on the acquisition of the German expression of possession in those subjects that had learnt English as an L2 and were currently learning German as an L3. On these grounds we decided to investigate further on the effects of L2 on L3 acquisition. Therefore we would like to provide more information in relation to this area of research before proceeding with the presentation and analysis of our corresponding findings (section 6.3). In the present section we will present some previously conducted studies in this field as well as the three major explanatory models that have been put forth.

The influence that second or other background languages exert on the acquisition of a new target language is a very recent area of research. However, various studies have been conducted up to now, which have prompted the formulation of three major relevant models: the **Cumulative Enhancement Model** (**CEM**) (Flynn et al.,

2004), the **Typological Primacy Model** (**TPM**) (Rothman, 2010) and the **L2 Status Factor Hypothesis** (**LSFH**) (Bardel and Falk, 2007).

The **CEM** (Flynn et al., 2004) suggests that all previously learnt languages can influence L3 acquisition. It claims that in principle the L1 is the primary source of transfer. The L2 will have an influential role only if the feature in scope (i.e. the feature to be acquired) is not present in the L1. This model takes only syntactic overlap into consideration, that is, **transfer** is seen as a **facilitating factor** in TL acquisition. Hence the CEM suggests that **language acquisition** is a **cumulative process** according to which the learner can decide to use not only one, but all, previously acquired languages in L3 acquisition.

The **TPM** (Rothman, 2010) suggests that **(psycho)typology** determines whether the L1 or the L2 will be transferred in L3 acquisition. The term (psycho)typology can refer either to the proximity of the languages involved based on genetic relatedness, to a typological similarity of a particular structure, or to the learner's perception regarding languages' similarity. If a learner perceives the one or the other language to be "closer" to the L3, this particular language will be transferred. The difference between this hypothesis and the CEM is that the former recognises the possibility of non target-like use structures resulting from this transfer process, whereas, the CEM, as we said, assumes that the transfer is facilitating.

The **LSFH** (Bardel and Falk, 2007) claims that **in** an **L3 acquisition** process the primary source of **transfer** is **always the L2** regardless of the similarity, actual or assumed, between the L2 and the L3. The L2 status hypothesis is based on the study conducted by Williams and Hammarberg (1998) regarding L3 acquisition of the lexicon. They claimed that there is a general tendency to activate a previously learnt language, rather than to activate the L1 in the acquisition of a third, which they called L2 status factor. In 2007 Bardel and Falk conducted a

study in which they found that Williams and Hammarberg's observations in SLA coincided with their findings regarding the acquisition of L3 syntax and thence proposed the LSFH as an explanatory factor to the findings.

These hypotheses emerged as a response to the findings of various studies that revealed an L2 influence on the acquisition of an L3. The first of these studies focused mainly on lexical transfer (e.g. Hammarberg, 1998) but then the scholars' interest extended to the syntax as well. Some of the major studies of **syntactic transfer** are outlined below.

The first study that looked on the acquisition of L3 syntax was that conducted by **Zobl (1992)**. In this study Zobl found that **multilinguals** have **broader**, in the sense on non-restrictive, **IL grammars** than monolinguals. In the judgement tests he made he found that multilinguals accept more marked constructions and ungrammatical sentences, since they are being less restrictive in a new language.

Klein (1995) conducted a relevant study in which she tried to address the question of whether multilinguals are different to monolinguals in foreign language acquisition. She investigated the acquisition of English verbs and their prepositional complements as well as the acquisition of preposition stranding. In relation to this she compared a group of 17 L2 English learners with a group of multilingual L3/L4 English learners with various language backgrounds. She also used a group of 15 native speakers of English as a control group. This study confirmed Klein's initial hypothesis that multilinguals would acquire the preposition stranding structure faster, even though a similar structure did not exist in their L1 or their L2. This finding hence suggests that **multilinguals** are probably going through a **less conservative learning procedure** benefiting from their higher metalinguistic awareness. As Klein (1995) said, multilinguals have an "enhanced lexical knowledge which may help

to trigger parameter-setting and this propels Mls [multilinguals] further along the path of acquisition" (p.450).

Dentler (2000) also studied the **acquisition of L3 syntax**. She investigated the impact of L2 syntax in L3 German by focusing on the production of German main clauses by L1 Swedish and L2 English speakers. Her study's results show that the participants did not apply the V2 rule correctly in L3 German, although this is a feature that German shares with the participants' L1 but not with their L2. Dentler (2000) did not explain this phenomenon on the basis of L2 transfer but she rather sees it as an indication of a certain acquisition order followed by all subjects confirming thus Pienemann's (1998) Processability Theory. This, however, does not imply that her study's results do not indicate an L2 transfer.

Flynn et al. (2004) presented a study of the acquisition of relative clauses by **L2 English learners** with **different L1**, Spanish and Japanese chosen precisely due to the typological distance that characterises these two native languages. Spanish is a head-initial language and in that sense it is similar to English. Japanese, on the other hand, is a head-final language. Their findings showed that L1 Japanese learners of English performed better than L1 Spanish learners of English. The scholars thence claimed that "both 'determining' and experience with the consequences of the parametric value of this grammatical principle is necessary in acquisition in terms of the development of a language-specific grammar" (Flynn et al. 2004:8). That is, they argue that Japanese L1 speakers act just like monolingual English children. Since they have no previous experience with head-initial languages, they only have to set the correct parametric value for the English L1 or L2. They then compared these findings to others resulting from L3 acquisition studies. They found that "**when the L2 is still 'in progress', its influence on L3 acquisition is not the same as it is when L2 and L3 are sequential**" (Flynn et al., 2004:14).

Indeed, the importance of the TL and the L2 proficiency level has been suggested by other scholars as well. **Williams and Hammarberg (1998)** showed that **high proficiency in the L2** will enable this language to **influence** the acquisition of a new language. However, low proficiency in a background language has also been claimed to be a factor (De Angelis, 2005b). It is generally assumed that the lower the proficiency level in the TL the more the background languages will exert influences in order to solve communicative problems (Ringbom 1987).

In **2007, Bardel and Falk** conducted an empirical study in which they compared two groups of learners' acquisition of negation placement. They used both longitudinal Swedish data collected in the Netherlands and cross-sectional data collected in Sweden. The L3 Swedish as well as the L3 Dutch learners were at the very initial stage of L3 acquisition. They focused on the negation placement which depends on whether or not the finite verb raises up to the second place in the main clause (the V2 property). One group of learners (the English L2 group) had an L1 with V2 and an L2 without V2, and the other group (the Dutch/German L2 group) had an L1 without V2 and an L2 with V2. **They found that the group with the L2 that had the V2 property correctly placed the negation post-verbally from the very initial stages**.

4 Research Questions and Hypotheses

Our study is divided in **three** main **parts** according to the language at issue and the principal data elicitation method used. We should, therefore, formulate separately our hypotheses and questions that relate to each of these studies. First, we will present the research questions for the acquisition of morphology in a corpus of L2 English collected in an EFL secondary school (**study 1**). Then, we will focus on a particular morpheme: the expression of possession L2 English via the *genitive –s* morpheme vs. the analytic *PP* structure (**study 2**) and in L3 German via the *genitive –s* morpheme, the *genitive case* and the analytic *PP* structure (**study 3**).

4.1 Study I: MOS and Learner Corpus in L2 English

Q1. Our **first question** concerns the **proficiency** as an explanatory factor for the observed development of morphology in L2 English, that is, whether low and high proficient learners will display different accuracy orders as previous studies have shown (Muñoz, 2006).

> Our **first hypothesis** is that **accuracy** does not imply **acquisition** of the corresponding functor. Evidence from different studies proves that some functors' development does not display a linear evolution and hence accuracy at one stage does not imply that the same accuracy rates will be obtained in subsequent stages. That is, accuracy does not equate acquisition. If grouping our subjects according to their level of proficiency enables us to observe more phenomena of the acquisition process as Ellis and Barkhuizen (2005) have argued, then we will be able to study the different developmental stages of certain

morphemes (such as the *regular past* and the *irregular past*) and thus confirm that we should not assume **acquisition** has taken place on grounds of the **accuracy** rates achieved.

Our **second hypothesis** is that the accuracy orders will display a similar **route** for **all proficiency levels**. In relation to this we also argue that our findings will correlate with other EFL findings.

Q2. The **second question** we wish to answer through the analysis of our learner corpus from a MO perspective is whether **Hawkins'** (2001) **pattern** bearing on the difficulty of the morphemes can stand as an **explanatory model** of our study's findings. Therefore, as we will see in a subsequent section (5.4.4) we also organized our morphemes in **verb-related** and **noun-related** classes. Regarding this question we will be examining only the verb-related functors.

Our **third hypothesis**, therefore, is that the accuracy orders revealed in our study regarding the verb-related functors will have a **pattern similar** to the one suggested by **Hawkins** (2001) and Andersen (1978)[6]. In that vein, the development of the verb-related morphology in L2 English will depend on the difficulty of the underlying grammatical structures of each morpheme. However, we do **expect** to find **some variation** within the aforementioned sequence among our groups in accordance with our first hypothesis that regards the subjects' proficiency level. We furthermore argue that this deviation from Hawkins' developmental pattern will appear in relation to the ***past tense*** **morphemes**, that is, it will be **influenced by** the **non-linear development** that their acquisition exhibits. See section 3.1.3 for a discussion on this issue.

[6] A presentation and brief analysis of this pattern is presented in section 3.1.2.

4.2 Study II: Genitive –s in L2 English

Q1. The **first question** we wish to investigate in this section regards the influence of the **data elicitation instrument** on the research's results. In the same line, we will be checking Krashen's (1978) claim that the instrument influences the results and that the accuracy orders displayed in each case will be different according to whether the data come from **naturally** produced (corpus) or **experimentally** elicited language. Therefore we will compare the accuracy rates of the *possessive –s* obtained by the subjects that participated in our picture composition task (corpus) on the one hand and in the sentence transformation task (experiment) on the other. In this respect we will be using the SOC scores since it is acquisition we wish to check and hence we have to include both the TLU and the NTLU instances. Note that we do not calculate the SOC according to Brown's model, i.e. based only on the OC. More details regarding the model adopted in this study in order to calculate the accuracy rates is presented in the following chapter (section 5.4.5). When investigating the frequency, on the other hand, we shall compare the TLU and the NTLU scores in order to isolate and thus separate the use of the inflectional (TLU) and the periphrastic (NTLU) possession forms. We analyse in detail our tagging schemes for each of our three studies in chapter 5. Although our scoring model is structured on the basis of the TLU model, we call the total accuracy scores SOC instead of TLU in order to distinguish between the total accuracy rates, which contain both the TLU and the NTLU cases, and the actual TLU instances. Both the SOC and the TLU are used in comparisons in the analysis of the accuracy rates and the frequency of use respectively.

Our **first hypothesis** stems directly from our first question and regards the actual **influence** of the data collection **instrument on** the **accuracy rates** of the *possessive –s*. Following Krashen's idea that experimentally elicited data reflect the learnt but not necessarily the acquired knowledge, we expect to find that the *possessive –s* has been

more accurately used in our sentence transformation (experiment) than in the picture description (corpus) task.

Our **second hypothesis** refers to the relative frequency of the *possessive -s* in English. We believe that our L2 English learners will show a clear preference towards the use of the **periphrastic** form confirming thus the findings of previous MOS (Bailey et al., 1974; Larsen-Freeman, 1975; Rosansky, 1976) that rank the **inflectional** form among the last functors to be acquired.

4.3 Study III: Possessive structures in L3 German

Q1. The main **question** of this study contemplates the relative **frequency** of each of the **possessive forms** (i.e., periphrastic as well as inflectional) in our L3 German learners' language. In the same line we wish to compare the corresponding results with the ones obtained for our L2 English learners.

In relation to our previously posed **question** we cannot form a hypothesis. MO studies previously conducted bearing on the acquisition of L2 English morphology claim that the **route** is similar regardless of the subject's **L1** (see section 3.1.2). However, research on the acquisition of the possessive construction in German/Dutch[7] by L2 learners claim that in initial stages students tend to rely on their L1 expression of possession (Van de Craats et al., 2000; Matteini, 2007; Matteini, 2009).

Therefore, our **hypothesis** for this study regards only the **effect** that the subjects' **L2** exerts **on** their acquisition of the **L3**. In keeping with previous studies on this field (see section 3.3 for more details) our **hypothesis** predicts that **L3** German **acquisition** is **influenced by the learners' L2** English, especially in initial proficiency levels.

[7] Please note that the comparison is possible due to the similarities between the German and the Dutch systems of expressing possession.

5 The empirical studies

In this chapter we present the details of our empirical studies. In keeping with the perceived importance of the **methodology** we have decided to begin by introducing the two methods that we have adopted in this paper. Initially we provide information about the various types of method available. Then we discuss our own methodological options on which we dedicate two separate sub-sections (5.1.1 and 5.1.2).

In section 5.2 we analyse our **subjects' composition** first according to the educational institution in which they learn the TL (i.e., either English or German) and then in terms of their proficiency level.

The following section of this chapter regards the **materials** used in our empirical studies. In order to achieve clarity in their presentation we have decided to create one sub-section for each of the data elicitation instruments. In the same line, sub-section 5.3.1 refers to the format we used in order to create a learner's profile including thus useful information such as mother tongue, time of exposure to the TL, other FL spoken, etc. In sub-section 5.3.2 we present the placement tests used for the classification of our subjects' proficiency level in both English and German. In the final two sub-sections we describe and analyse the compilation of both our picture description task (corpus) and our sentence transformation task (experiment).

The final section of chapter 5 relates the actual **procedures** followed for the three studies of our project. Accordingly, we describe how we proceeded with the collection (5.4.1) and the transcription (5.4.2) of our data. After completing these stages we processed our data using an editorial tagging scheme which we present in sub-section 5.4.3. The actual tagging of the morphemes at issue is presented in the following sub-section

(5.4.4) Note that within this sub-section we distinguish three parts in keeping with the three studies of our project. Accordingly, each of these parts contains the detailed description of the tagging scheme that corresponds to every study. Finally, in sub-section 5.4.5 we introduce the calculation model used for the scoring of our data.

5.1 Method

First we should mention that our study is **cross-sectional** and that it focuses on the **accuracy order** of nine functors presented in sub-section 5.4.4 as well as on the **relative frequency** of the functors that relate to the expression of possession in both English and German.

In order to obtain samples of language from L2 learners, researchers can use different methods. Learner language production varies due to both internal and external factors. Research has shown that the order of acquisition of different grammatical structures can vary according to the kind of task used to elicit learners' language. Although some of the first MO studies reached similar conclusions regardless of the method used (see section 3.1), other studies support the hypothesis that the outcomes regarding the accuracy of use are influenced by the elicitation instrument. Indeed, Ellis (1987) found that the accuracy order of two past tense morphemes varied depending on how the data were collected.

In keeping with the perceived relationship between elicitation task and learner language we can distinguish three types of learner language samples corresponding to the three principal methods for collecting data.

The first method is the least controlled one and consists of obtaining samples of language produced in a real-life situation in order to satisfy some communicative need. The other two methods involve elicitation that, following Corder (1976), can

be distinguished in clinical and experimental. Figure 4 displays the three types of learner language.

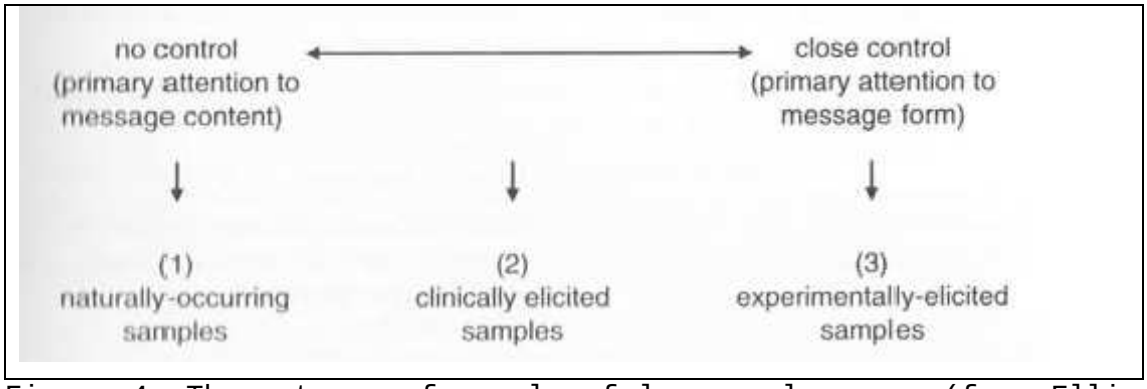


Figure 4: Three types of sample of learner language (from Ellis and Barkhuizen, 2005:23)

The difference between clinical and experimental elicitation "matches the distinction between task and exercise". (Ellis & Barkhuizen, 2005:23) Corder (1976) argued that the clinical elicitation is suitable for those studies that do not have a previously well-formed hypothesis and therefore intends to gather data of any sort from the participants. Experimental elicitation, on the other hand, is highly controlled because the researcher wishes to check the usage of a specific linguistic form. Experimentally elicited samples result in a constrained constructed response (Norris & Ortega, 2001), that is to say, in a short L2 segment.

In our case study we have opted for the use of **two methods**, ranging from less control (picture composition task: **corpus**) to more control (sentence transformation task: **experiment**). In the following sections (5.1.1 & 5.1.2) we provide a detailed description of the aforementioned methodologies.

5.1.1 Picture composition task (corpus)

The first method used was a **picture composition task** which represents clinical elicitation and enabled us to compile our own **learner corpus**. In fact, picture composition task is very close to naturally occurring samples of language. The only

difference between them is that the former is collected specifically for the purpose of the research. Additionally, picture composition tasks aim to the elicitation of a general, rather than a focused, sample of learner language. This further distinction is very important since the "learner's orientation to the elicitation task can have a profound effect on the language used" (Ellis and Barkhuizen, 2005:31).

The process we followed was to show the learners the picture composition (see Appendix 9.4) and ask them to write the story in their own words while they were in class. That is to say, our corpus did not result from a take-home composition task as we wanted the language used to be as authentic as possible. Students were allowed 30-45 minutes to complete the task and were provided with the entire series of pictures along with a reduced word-list that included mainly nouns they may not know. Our decision to give our subjects plenty of time is justified by the fact that we did not want them to feel stressed which would obstruct their concentration and probably distort their language. On the other hand, our choice to give learners a reduced list of relevant words was based on our previous decision in relation to the nature of the elicitation task. That is, the task should not be cognitively demanding in order to facilitate the production of **authentic learner language**. By giving our subjects a list of words we may have influenced the authenticity of the learner language but only in terms of lexical items, which obviously does not influence our research that focuses on functors.

5.1.2 Sentence transformation task

The second method was a controlled task, namely, a **sentence transformation task** intending to elicit the use of ***possessive -s*** in the case of L2 English learners and the equivalent synthetic forms for the expression of possession in the case of L3 German learners. Ellis & Barkhuizen (2005:37) include this type of experimental elicitation in the so-called discrete-points tests with the traditional language exercise format.

Our **decision to use two methods** one of which is very controlled is justified by the fact that the *possessive -s* is not easily elicited in naturally occurring samples. Additionally, it allowed us to check **Krashen's** argument that discrete-point tests trigger the learners' **formal linguistic knowledge** that results from the L2 learning process. If Krashen was correct, then the accuracy rates for the *possessive -s* will be different, and indeed significantly higher, from those found in our corpus analysis. Since the latter is authentic and thus to a high degree spontaneous language, it should show a morpheme accuracy order determined by unconscious acquisition as Krashen (1978) has argued.

5.2 Subjects

As we mentioned before, our study is cross-sectional and, as such, it includes a large number of subjects. For both parts of our study a total of **400 examples of learner language** were collected.

For the compilation of our learner corpus a total of **95 secondary education students** from the High School *IES Pedro Soto de Rojas* (Granada, Spain) participated in the study.

In the second part of our study consisting of the experimentally elicited data a total of **305 FL learners** took part. Out of the total of 305 students **256** were **secondary education students** from the High School *IES Pedro Soto de Rojas* (Granada, Spain) and had **English** as a foreign language whereas **7**, also **secondary education**, students were learning **German** as an L3 at the High School *IES Padre Suarez* (Granada, Spain). Due to the scarcity of L3 German learners, we also used samples collected from the language centre *Centro de Lenguas Modernas* (CLM) in Granada, Spain. Therefore, **42** subjects were **not secondary education** students. This, however, did not influence our study since we grouped our subjects not according to age but on the basis of their proficiency level which was determined by means of proficiency tests. The process is fully explained in section

5.3.2 and our L3 German learners' classification in terms of **proficiency level** is presented in table 6.

Additionally, a **learner's profile** was administered. It included information regarding age, gender, L1, proficiency level, etc. following the example that we provide in Appendix 9.1, has been compiled for all our subjects.

The participants were all volunteers who were informed about the objectives of the study from the very beginning.

Table 6 shows the composition of our subjects in terms of their proficiency level in L2 English. We then supply a chart of the aforementioned subjects' synthesis in Figure 5. As we can see the majority of the subjects in both parts of our research project (always in relation to L2 English) have a pre-intermediate or an intermediate level. Only few have reached the upper-intermediate level whereas just a minority is still at the elementary level. We should also mention that the age range of our L2 English learners was 11-18 years. In terms of educational level we included students from the 2ºESO to 1ºBachillerato.

	Learner Corpus	**Experimentally Elicited Data: *Possessive -s***	**TOTAL**
Elementary	5	18	**23**
Pre-Intermediate	43	106	**149**
Intermediate	39	101	**140**
Upper-Intermediate	8	31	**39**

Table 6: Distribution of **L2 English** subjects

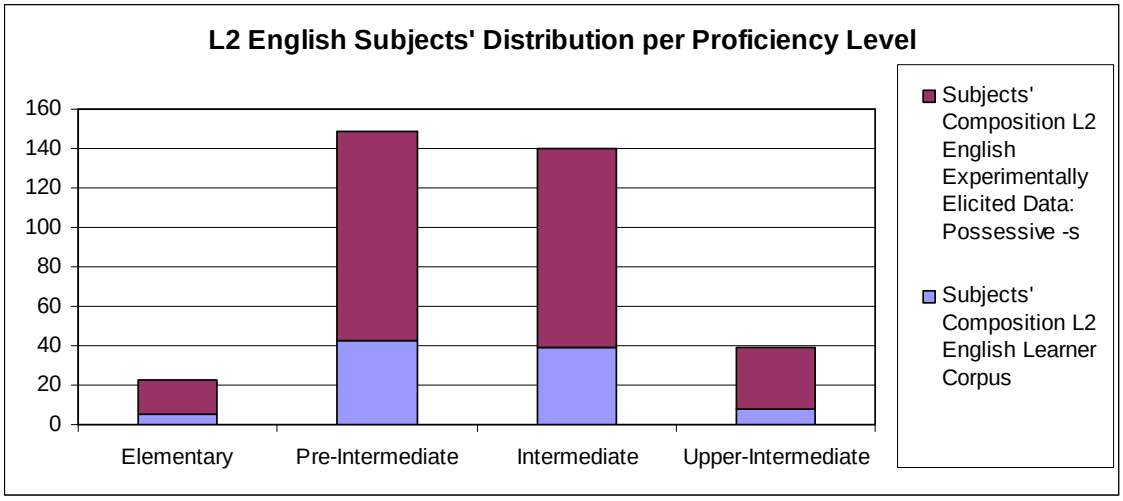


Figure 5: Distribution of the number of subjects per proficiency level in **L2 English**

Table 7 shows our subjects' distribution according to their proficiency level in L3 German. The chart of the same is presented in Figure 6, where we can observe that the majority of our subjects are at the elementary level. The pre-intermediate and the intermediate levels have almost the same number of students. Finally, we have managed to find only a few students at the upper-intermediate level. We would like to underline that, although some of our subjects come from a high school whereas the rest come from a language centre, a **comparison** of the final data is **possible** because in both cases the **learning** environment is **instructional**.

	IES Padre Suarez	**CLM**	**TOTAL**
Elementary	3	17	**20**
Pre-Intermediate	4	9	**13**
Intermediate	0	12	**12**
Upper-Intermediate	0	4	**4**

Table 7: Distribution of subjects in **L3 German**

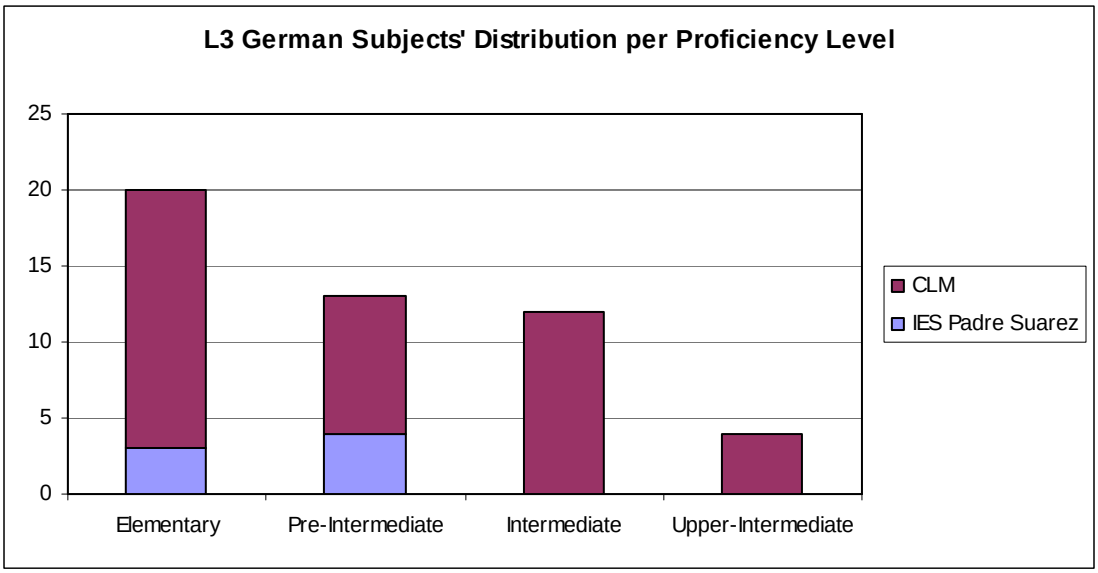


Figure 6: Distribution of the number of subjects per proficiency level in **L3 German**

Finally Table 8 and Figure 7 show the distribution of our subjects according to their proficiency level in both L2 English and L3 German.

Subjects' Composition L2 English & L3 German		
	German	**English**
Elementary	20	23
Pre-Intermediate	13	149
Intermediate	12	140
Upper-Intermediate	4	39

Table 8: Distribution of subjects in **L2 English and L3 German**

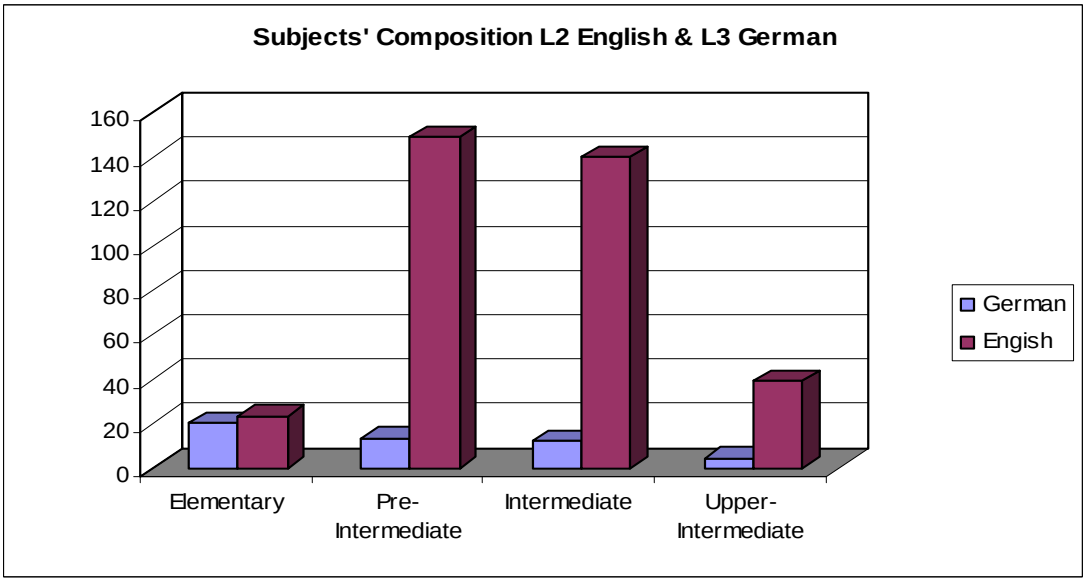


Figure 7: Distribution of our subjects' proficiency level in **L2 English and L3 German**.

5.3 Materials

In our case study we have used various types of data elicitation instruments.

5.3.1 Questionnaire (learner profile)

Learner language is influenced by a wide variety of linguistic, situational and psycholinguistic factors. Indeed, one of the critiques of the cross-sectional SLA studies is that "there is often no detailed information about the learners themselves and the linguistic environment in which production was elicited" (Gass, 2001:33). In that line of thought, we decided to include **additional information** regarding our **subjects**. Although Figure 8 shows the major variables that need to be controlled when compiling a learner corpus as traced by Granger (2008:264), we used it as a guideline for the collection of our experimentally elicited data as well. The underpinning idea of this two-part table is the distinction put forth by Ellis (1994:49) between learner variables and task variables. In our case study we have tried to incorporate as many of these variables as possible by creating a quite detailed **learner's profile** based on a **questionnaire**, a sample of which can be found in Appendix 9.1.

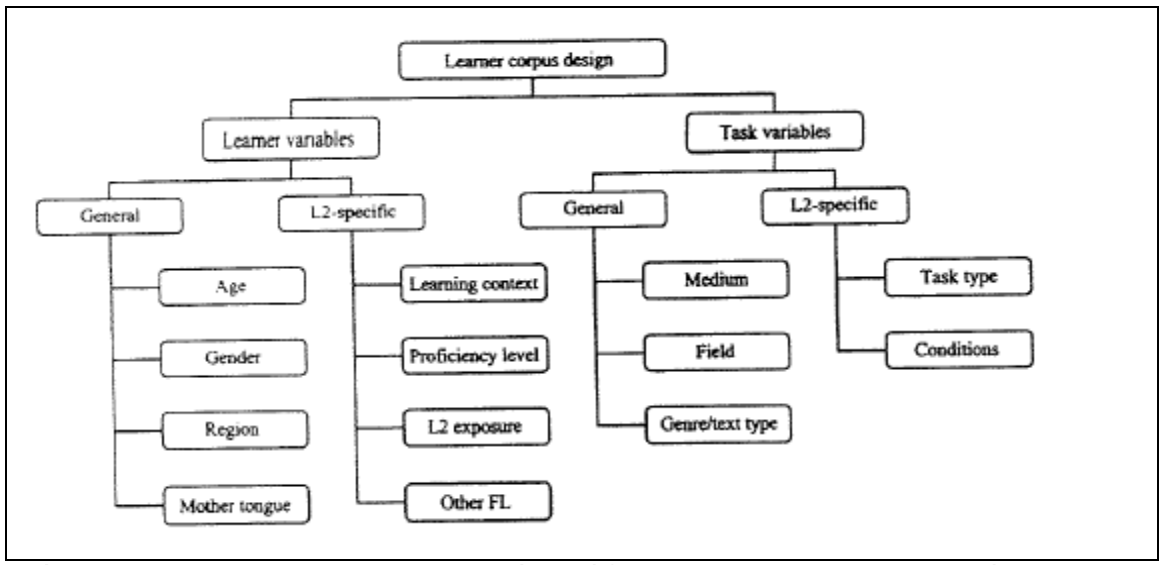


Figure 8: Learner Corpus Design (from Granger, 2008:264)

In the compilation of our **learner profile** we have been based on the **learner variables** as outlined in Figure 8. Indeed, we have included both the General and the TL-specific variables. Our subjects were asked to provide information about their *age* and *gender*. The variable *region* was not explicitly included in our learner profile since all our subjects were residents of Granada, Spain. Regarding the variable *mother tongue* the students had to explicitly state their mother tongue, their mother's native language, their father's mother tongue, as well as the language spoken at home. In relation to the *learning context* and the *exposure to the TL* we asked our subjects to provide all the relevant information in. That is, whether they have been in a country where the TL is spoken (and if yes, where and for how long), if they take part in a bilingual programme (and if yes, since when, for which subjects and how many hours per week), and finally whether they receive additional tuition on the TL or if they do other activities in the TL (e.g. watching films, reading books, etc.). Our subjects were also asked to state whether they speak or not other *foreign languages* and which. Finally, bearing on the subjects' *proficiency level* not only we conducted a placement test, but we also asked our students to give us their opinion regarding their proficiency level in each of the four areas (i.e. listening, reading, speaking and writing).

5.3.2 Placement tests

Granger (2008:264) mentions that of all the L2-specific learner variables, the **proficiency level** is the most important but also the most difficult to establish. Indeed, until now no L2 English learner corpus includes a proficiency test. Even the ICLE created by Granger et al. (see section 3.2) is a corpus that includes data from low advanced L2 English learners. In our learner corpus we have incorporated not only the learner profile described above but also a **proficiency test**. Accordingly, our learner corpus contains four sub-corpora according to the proficiency level (ranging from elementary to upper-intermediate).

We would like to emphasize that our decision to test our subjects' proficiency level helped us overcome the inconsistency provoked by the fact that "one researcher's advanced category may correspond to another's intermediate category" (Gass, 2001:37). Additionally, it allowed us to classify the subjects into different groups according to their level of proficiency and thus observe other phenomena within the SLA process such as the **U-shape developmental pattern** that certain morphemes present.

Our subjects' proficiency level was decided on grounds of each student's results on a placement test that we distributed before proceeding with the elicitation process.

In relation to **L2 English placement test**, we used the written placement test of the English Unlimited Test published by Cambridge University Press in 2010. This consists of 120 multiple-choice questions, 20 at each level from Starter to Advanced (covering CEF levels A1 to C1). We used the first 100 questions since it was a test given out to secondary education students, that is, we excluded CEF level C1. However, no student reached the upper limits of the upper-intermediate level. Students were asked to start at the beginning of the written test and stop when the questions became too difficult. We allowed 45 minutes for the test. A sample of the adapted placement test along with the teacher's guide can be found in Appendix 9.2.

Let us consider now the **L3 German proficiency test.** In order to check the proficiency level in German of the secondary education students of the High School "IES Padre Suarez" we adapted the placement test provided by the centre of modern languages (Escuela Oficial de Idiomas) in the Basque Country. Following the guidelines given out by the Escuela Oficial de Idiomas (henceforth EOI) we considered that a proficiency level was reached when the student had made at least 19 correct choices out of the 24 for each level. In keeping with our decision to conduct only a written test in English we excluded the reading

part that was originally found in the placement test of the aforementioned centre of modern languages. A sample of this can be found in Appendix 9.3. We did not apply that test to the students of the language centre *Centro de Lenguas Modernas* in Granada since the subjects are already grouped on the basis of a placement test compiled according to the Common European Framework of Reference for Languages (CEFR) by the *Centro de Lenguas Modernas* (henceforth CLM) in Granada. Although our L3 German subjects' proficiency level is defined on the basis of two different tests we consider that their classification is comparable and hence reliable for two reasons. First because the placement test of the EOI that we used with the students of the High School *IES Padre Suarez* and the placement test used by the CLM are compiled on the basis of the CEFR guidelines and thus share a common theoretical framework. Second because both the EOI and the CLM are certified FL examination centres.

5.3.3 Corpus (Picture Description Task)

According to Corder (1976), clinical elicitation suits best a generally formed hypothesis. Since our first aim is to check the accuracy order that previous L2 morpheme acquisition studies have put forth, we considered a clinically elicited language sample to be the most adequate option. As described in sub-section 5.1.1, our corpus was compiled on the basis of a **picture composition task** a sample of which can be found in Appendix 9.4. This sample is an adaptation of the elicitation task titled "Frog, where are you?" originally created by Mayer (1969). We decided to use a version of this picture composition task mainly because it has been used in other learner corpus based studies (e.g. CHILDES), as well as in one of the key studies on morpheme acquisition in L2 English (Muñoz 2006), which we reviewed above (sub-section 3.1.2).

Up to this point we have presented various reasons that support the usefulness of **learner corpora** even when they consist of raw data, that is to say, learner language with no added linguistic annotation. However, it is even more useful when it contains

extra information. **Annotation**, both grammatical and error, is a highly challenging and time-consuming process and hence researchers may choose to use the ready-made annotating tools. However, as Granger (2008) points out, these tools have been created on the basis of L1 corpora and "[…] there is no guarantee that they will perform as accurately when confronted with learner data" (p.265). In our attempt to deliver as reliable data as possible we have done all the grammatical annotation **manually**. The process is described in sub-section 5.4.4 where we outline the procedures of the present study.

5.3.4 Experimentally elicited language (Sentence Transformation Task)

The second question we wished to give answer to regards the so-called **Anglo-Saxon genitive**. Most previous studies have placed its accurate use, and thus assumed acquisition, at very low levels. We wanted to check the validity of this previously given position and also compare it to the expression of the genitive by means of the inflectional possession structures in another West Germanic language, namely in German.

Since the **elicitation** of the *possessive -s* is very **difficult** and we could not assure its use in our corpus, we decided to use an **experimental elicitation** as well. This instrument consists of 6 sentences that the students should transform in order to express possession. Note that in the given sentences the possession was expressed by means of the *possessive pronouns*. We used the **sentence transformation test** instead of the fill-in-the-blank questions that have been used in the study conducted by Wagner (2005) regarding the acquisition of the *possessive -s* because we considered that the latter actually leads the learner to use either the *possessive -s* or nothing. But it does not give the subject the option of selection between the synthetic form, i.e. the *possessive -s*, and the analytic form, that is, the *prepositional phrase* (introduced by the preposition *of*). We considered that the learner's choice between these two forms is significant in terms of acquisition of the inflectional

possessive form. We used the same sentences in both English and German. A sample of them can be found in Appendix 9.5.

Furthermore, the **combination** of both **clinically** and **experimentally elicited data** gave us some insight regarding the influence of the instrument on the results of morpheme studies and thus a way to check Krashen's hypothesis in relation to the elicitation of learned vs. acquired L2 knowledge. Note that this combination is justified since it has been recently argued that combining corpus and experimental data always provides **better insights** into the nature of **interlanguage** grammars (Gilquin & Gries 2009; Lozano & Mendikoetxea, forthcoming 2013; Mendikoetxea & Lozano, forthcoming 2013).

5.4 Procedure

5.4.1 Data collection

First we asked the subjects of our study to complete the **learner profile**. The following step was to carry out the **placement tests** which we then corrected and scored according to the guidelines of their creators as explained in section 5.3.2 and exemplified in Appendices 9.2 & 9.3. At the same time we asked the learners to complete the **sentence transformation task.** Then we did the **picture composition task** (corpus) for 95 secondary education students (from 2º ESO to 1º Bachillerato).

5.4.2 Transcription

After having collected the data, we **typed** each sentence transformation test and each composition in **plain text format** in order to avoid automatic corrections of the students' errors that "Word documents" usually perform. A sample of a transcribed file can be found in Appendix 9.6. In the same text we also included all the learner variables (see section 5.3.1), that is, the information found in the learner's profile regarding age, nationality, L1, etc. as well as the subject's proficiency level according to the placement test. Appendices 9.7.1, 9.7.2 and 9.7.3 contain examples of transcribed files that we then used to

compile our learner corpus and our sentence transformation task database for both English and German.

5.4.3 Tagging (editorial tagging scheme)

An **editorial tagging scheme** was used to code the learners' editions of their own writing. That is, we also coded rewritten material whether legible or illegible. In the first case we used the code $_RWR_ (the rewritten element). In the second case the code was $_RWU_¿. The tagging scheme for the learner's corrections is shown in Appendix 9.9. Table 9 shows the data that we included in each transcribed file name along with an example. From this the researcher can obtain information regarding the proficiency level that the specific student obtained in the placement test, the course in which the subject is at the moment of the data collection, the subject's age, the researcher's identification and the subject's initial. Accordingly, in our example below we can immediately see that our subject is at the upper-intermediate level, attends the 4ESO course and is 15 years old.

PlacementTest_Course_Age_Institution_Researcher's initials_student's initials.txt
B2_4ESO_15_PSR_EAT_IAC.txt

Table 9: Transcribed file name

5.4.4 Tagging: morphemes

After having transcribed all the texts, we used the **UAM Corpus Tool** software in order to code the data, i.e., to **tag the morphemes**. An example of the UAM Corpus Tool can be found in Appendix 9.8. In relation to the learner corpus we should mention that we tagged a set of morphemes based on the list found in previous morpheme studies. Table 10 presents the list of the morphemes investigated in the present study.

past reg: **past_reg**	V-related
past irreg: **past_irreg**	
3rd sing –s: **3sg**	
prog –ing: **ing**	
copula BE: **be_cop**	
aux BE: **be_aux**	
plural: **plu**	N-related
art (a/the): **art**	
possessive: **pos**	

Table 10: Tagged morphemes: Based on: Krashen (1977) The Monitor Model for adult second language performance. In: Burt, M., Dulay, H., and Finocchiaro, M. (eds.). *Viewpoints on English as a Second Language.* New York: Regents Publishing.

Based on one of the first MOS' critiques regarding the shortcomings of the SOC method, we decided to include in our tagging scheme not only the **SOC** (Suppliance in Obligatory Context) but also the **SNOC** (Suppliance in Non-Obligatory Context). Therefore both the Target-Like Use and the Non-Target-Like Use of each of the aforementioned morphemes have been examined. That enabled us to account also for the cases of incorrect use of the morphemes by the learners (that is, underuse, misuse and overuse, whose differences will be explained later) and thus trace a more complete picture of the actual accuracy rates. Our analysis of the subjects' performance in relation to the nine morphemes shown in Table 10 involved the **following steps**:

1. Identification of the obligatory contexts for the use of each of these morphemes and
2. identification of the cases of target-like-use and non-target-like-use of each morpheme for each learner.

❶Part I: Tagging scheme in the MOS Learner Corpus

In Table 11 we present the **tagging scheme** in UAM Corpus Tool for *progressive -ing* as used in the **MO study**. Following this pattern, the use of a morpheme can be classified as either **target-like**, which corresponds to Brown's (1973) idea of suppliance in

obligatory contexts, or **non-target-like**. Within the latter we further distinguish between **underuse**, that corresponds to the lack of suppliance, **misuse**, which may be a *misselection* (the use of an incorrect morpheme instead of the target morpheme in the obligatory context) or a *misrealisation* (the erroneous realisation of the morpheme required in a obligatory context), and **overuse**, which corresponds to what is also known as SNOC (suppliance in non-obligatory context). Through this scheme we could trace a more complete and hence more accurate image of each learner's interlanguage. The aforementioned categories are exemplified for *progressive -ing* in Table 12. Examples of the tagging scheme described above for each of the investigated morphemes can be found in Appendix 9.10.

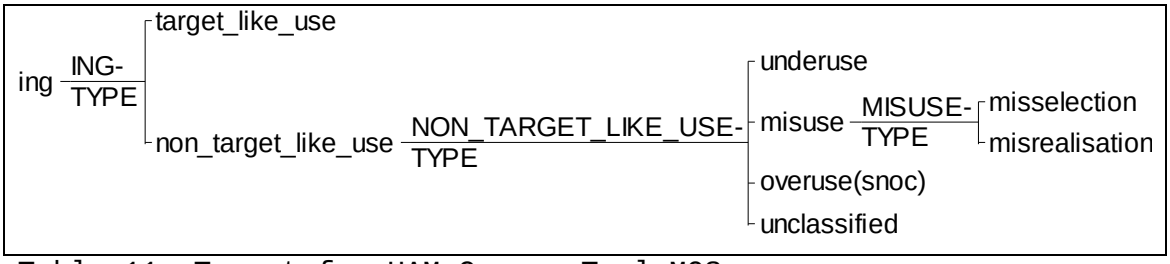


Table 11: Tagset for UAM Corpus Tool_MOS

OC:Progr. –ing (She is reading)			**S: Supplied form**
Target-like Use (TLU) (correct form supplied)			She is reading
Non-target-like Use	**Underuse** (no form supplied)		She is read_
	Misuse	**Misselection** (form exists)	She reads
		Misrealisation (form doesn't exist)	She is reads.
OC: 3rd Sing (She reads a lot)			**SNOC**
		Overuse (correct form supplied in NOC)	She reading a lot.

Table 12: Tagging scheme for *progressive -ing*

The data elicited through the sentence transformation task were also introduced in the UAM Corpus Tool in which we tagged it according to the rationale presented below.

❷Part II: Tagging scheme in the Possession Expression and the Experimentally Elicited Data in L2 English

All the cases that require a possession construction were considered as obligatory contexts (**OC**). This in English can be expressed by means of both the *possessive –s* and the *prepositional phrase* introduced by the preposition *'of'* (see discussion in section 2.2.1).

We decided to use the UAM tool for the tagging of the *possessive –s* for our experimentally elicited data as well. Although the latter do not present a corpus, we opted for the use of the corpus-software for various reasons. First and foremost, because we decided to maintain the formatting and thus facilitate the process of the tagging and the analysis for the researcher. Second, because it allowed us to introduce the additional data from the learner profile and create a separate file for each student but within the same project. Finally, because the UAM Corpus Tool gives the researcher the opportunity to add extra tagging schemes at any time and thus investigate other aspects and then contrast the elements studied.

For the purposes of our study we called **TLU** the use of the ***possessive -s*** and "**NTLU: misuse: misselection**" the use of the ***prepositional phrase***. This decision is justified by the fact that we investigate the accuracy rates for the *possessive -s* only and not for the possession construction in general. Hence if the use of the *possessive -s* is the TLU then anything else should be classified as NTLU. We specifically tagged the *PP* as "NTLU: misuse: misselection" so that we could also take into consideration the accuracy rates of the *PP*. We wish to clarify that our tagging scheme does not imply that the students have not acquired the possession construction in English. It rather indicates the preference of our subjects towards the use of

either the analytic (*PP*) or the synthetic (*pos. –s)* form of the attributive construction.

In Table 13 we present the tagging scheme used for the tagging of the *possessive –s* in our L2 English collected by means of the sentence transformation task (experiment). Note that each of the categories that appear in it is illustrated in the examples presented in Table 14.

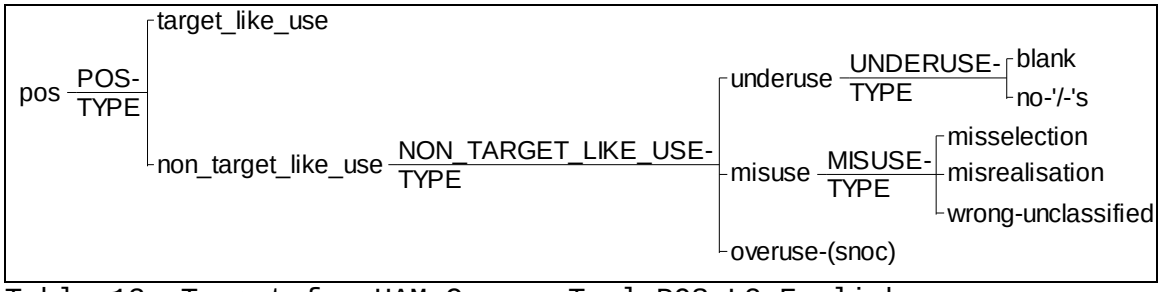


Table 13: Tagset for UAM Corpus Tool_POS_L2 English

OC: Possessive –s (My father's car)			**S: Supplied form**
Target-like Use (correct form supplied)			My father's car
Non-target-like Use	**Underuse** (no form supplied)	**Blank**	_____
		No '/'s	The boys_ shoes. My father_ car.
	Misuse (incorrect form supplied)	**Misselection** (form exists)	The car of my father.
		Misrealisation (form does not exist)	The boys's shoes.
		Wrong-unclassified	The shoes boys.
OC: Plural (Kids like toys)			**SNOC**
		Overuse (correct form supplied in NOC)	Kid's like toys.

Table 14: Morpheme tagging scheme for the expression of possession in L2 English

❸Part III: Tagging scheme in the Possession Expression and the Experimentally Elicited Data in L3 German

In keeping with our choice regarding the tagging scheme in English, we decided to call TLU the use of the synthetic possessive forms **in German**. That is, **both** the ***possessive -s*** and the ***genitive case*** have been tagged as **TLU**. Our tagging scheme is presented in table 15 below. Accordingly, each of the tagging categories included is exemplified in table 16. The **NTLU** is further divided into **three categories**. We called the first ***underuse*** to follow the model we used in our learner corpus study. This category marks the **lack of use** of any possession construction, either synthetic or analytic, in the identified OC. Under the label ***misuse*** we have distinguished **three subcategories**.

(i) The first is called ***misselection*** and includes, just like in the case of L2 English, the **use of the PP** instead of the genitive case. But in German, as we mentioned in section 2.2.2, the use of one of the three forms of attributive possession has specific **structural constraints**. Therefore, in German, unlike English, we recognise yet another case of *misselection*, that is, the **ungrammatical use of the *possessive -s*** in the cases where a *PP* or a *genitive case* is required. The same label (i.e., NTLU: misuse: misselection: *possessive –s*) has been attributed also to the cases where the learner has opted for the PR>PM instead of the PM>PR order (see section 2.2.2 for the significance of these orders in the possession system of German), even if the student has not added the suffix “-s”.

(ii) The second subcategory of misuse includes the two cases of ***misrealisation***, that is the wrong realisation of the *possessive –s* in German (i.e., -‘s) and the wrong genitive form.

(iii) The last subcategory of misuse was called ***wrong-unclassified*** and includes some exceptional cases that do not fit any other tag.

The final category of the NTLU is called ***overuse*** (also known as ***SNOC)***, that contains the cases of correct use of the possession construction in incorrect contexts.

Table 15 presents the tagging scheme in UAM Corpus Tool for the possession in L3 German. As mentioned above, each of the categories is exemplified in Table 16.

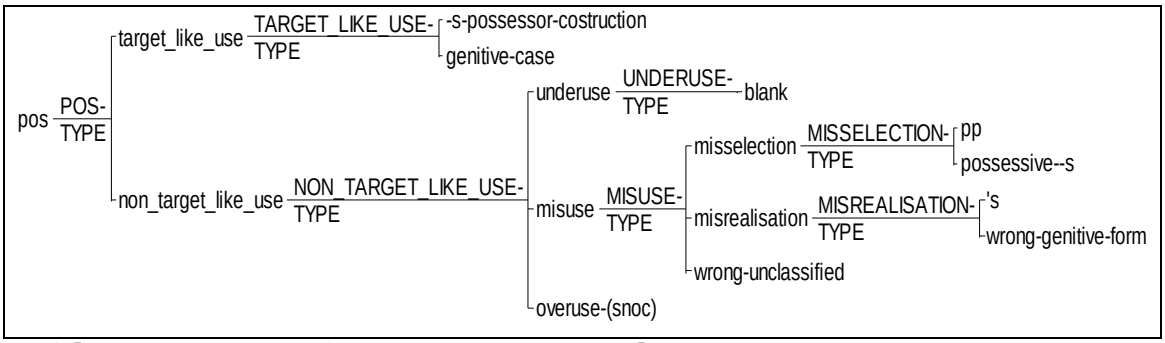


Table 15: Tagset for UAM Corpus Tool_POS_L3 German

OC: Possessive –s (My father's car)				**S: Supplied form**
Target-like Use (correct form supplied)		**-s-possessor construction**		Marias Schwester Lit: Maria's sister.nom
		Genitive case		Der Computer der Frau Lit: The.nom computer.nom the.gen woman.gen
Non-target-like Use	**Underuse** (no form supplied)	**Blank**		Maria Schwester. Lit: Maria sister.nom
	Misuse (incorrect form supplied)	**Misselection** (form exists)	**PP**	Der Computer von der Frau. Lit: The.nom computer.nom of the.dat woman.dat
			Pos –s	Kinders

				Spielzeuge. Lit: Children' toys.nom
		Misrealisation (form does not exist)	**'s**	Maria's Schwester. Lit: Maria's sister.nom.
			Wrong gen. form	Die Schuhe den Jungen. Lit: The.nom shoes.nom the.dat boys.dat
		Wrong-unclassified		Mein Auto's Vater. Lit: My car's father.
	Overuse (correct form supplied in NOC)			N.a

Table 16: Morpheme tagging scheme for the expression of possession in L3 German

5.4.5 Data scoring

In order to score our data we followed the guidelines suggested by both Dulay & Burt (1973) and Pica (1983), that is, we **combined** the **scoring** of each suppliance suggested by **Dulay and Burt(1973) and** the **TLU** model put forth by **Pica (1983)**. The former implied a weighted scoring according to the following schema:

- **0 points**: No suppliance
  (he walk__ yesterday) (two child__)
- **0.5 points**: Wrong morpheme supplied
  (he walk<u>s</u> yesterday) (two child<u>s</u>)
- **1 point**: Correct suppliance
  (he walk<u>ed</u> yesterday) (two <u>children</u>)

This was used in the calculation of the suppliance in obligatory context **(SOC)** and enabled us to include in the accuracy rates the cases of both *underuse* and *misuse*. The latter was included because its presence indicates that the student recognises the context as an obligatory one although he/she fails to make the correct choice. In our understanding this implies that the student has initiated the process of acquisition of that particular morpheme and in any case it indicates that his/her interlanguage is undergoing some change, which we consider important to account for.

More specifically, when working with the picture composition task we decided to tag the use of *past regular/past irregular* or *3rd person singular –s* according to the **student's initial choice** regarding the use of past or present for the narration of the story. So, if the student started narrating using the *present tense* and then changed to the *past tense*, we tagged the use of the *past* as "NTLU: misuse: misselection". To illustrate this we present a short extract from our corpus:

FILE_NAME: B1_2ESO_14_PSR_EAT_APL.txt
This story begins $_RWU_¿ at the night $_RWU_¿ in Ben's house. Ben is seven years old and he has got a frog and a dog. […] And he decided to find the frog, because he was sad and alone. […]

In this regard we also find what Muñoz (2006) observed in her study, that is, students "with higher levels of proficiency tended to narrate the story but often mixing tenses" (p.116). This can be seen in the following section where we present the features of the NTLU for past and present for all proficiency levels.

In our attempt to be as accurate as possible we decided to distinguish between the use of the **indefinite vs. definite article**, although both are included in one category. Therefore, we tag as "NTLU: misuse: misselection" the use of the definite article when the indefinite article should have been provided, and vice versa.

Pica's (1983:474) proposal refers to the general calculation pattern, also known as **TLU** (Target-Like Use), which "penalises" the cases of overuse by including them in the denominator.

$$TLU = \frac{SOC}{OC + SNOC}$$

Thence, after having calculated the suppliance of each morpheme in the corresponding obligatory context following Dulay and Burt's model, we proceeded with calculating the level of target-like use (TLU) for each morpheme. Note that for the purposes of this study we will be calling **SOC** the **final score** achieved after having applied the aforementioned calculation processes proposed by Dulay and Burt (1973) and Pica (1983). When we make reference to Brown's (1973) SOC model that will be always accompanied by Brown's name. In the same line, we will be using the term TLU when we study the relative frequency of the morpheme at issue and hence the cases of *misuse* are not included. As described in section 4.2, we have decided to use this distinction because we will be studying accuracy rates in relation to the acquisition process, as well as accuracy rates in relation to the frequency of use in which we cannot include neither the cases of *misuse*, which are included in the weighted scoring proposed by Dulay and Burt (1973), nor the cases of *overuse*, which represent the SNOC in Pica's model.

That is, we ended up with two accuracy indices for measuring the accuracy rates obtained for each morpheme: SOC and TLU. In our study the first is calculated on the basis of the following pattern:

$$SOC = \frac{\text{weighted scoring (according to the model of Dulay \& Burt)}}{OC + SNOC} \times 100$$

The TLU scores indicate the rates of target-like use instances as defined in our tagging schemes in section 5.4.4. In Appendix 9.11 we present in detail the descriptive statistics for each morpheme in all three studies of our project. The aforementioned accuracy indices, i.e. the SOC and the TLU, are also presented there.

Since we had already grouped our subjects according to their proficiency level, we did not calculate a target-like use score for each student but rather for each level. The same process was then followed for the experimentally elicited data in L2 English as well as in L3 German.

6 Results and Discussion

In this section we will present the **outcomes** of our study and comment on them. That is, we have decided to present the results together with the discussion so as to facilitate the reader's task, given that presenting the results for the three studies, plus a separate section with the implications of each study, would be too distracting.

Consequently, we have divided this chapter in **three parts** according to the instrument used for the data elicitation and the subjects' L2/L3. In the **first part** we make reference to the results of our **learner corpus study** and thence comment on our subjects' **accuracy rates and** the corresponding **morpheme orders**. In the **second section** we present the results of the ***possessive –s* accuracy rates** in **L2 English** for all proficiency levels. In the **final section** we present the results of our project that deals with the **accuracy rates** of each of the three investigated forms of **possession constructions** in **L3 German**.

6.1 Morpheme Order Study in L2 English

In order to **explore** the **progression** followed by subjects with different proficiency levels, **comparisons** are made between the **accuracy rates** obtained by each of the following groups: elementary, pre-intermediate, intermediate and upper-intermediate. Table 17 shows the accuracy rates and the corresponding rank for each morpheme obtained by each proficiency-level group. The accuracy rates are determined on the basis of our **SOC** formula as described in section 5.4.5 and are thus indicative of the acquisition level for each morpheme obtained by our subjects. Although accuracy does not equate

acquisition, we consider that our accuracy rates can be ***indicative*** of the latter because we do take into account the cases of *misuse, underuse and overuse*. Additionally, since we have calculated the accuracy rates for each proficiency level, we are able to check the acquisition process for those morphemes that do not present a linear developmental pattern. For more details regarding the aforementioned features see our discussion in section 3.1.3. Based on the accuracy rates we have determined the **rank** that each morpheme occupies in each proficiency level. The rank simply indicates that the accuracy rates obtained for one morpheme by one proficiency group are higher (or lower) than the accuracy rates obtained for another morpheme by the same proficiency group. It is in the same way that Brown (1973) defined the mean order of acquisition of L1 English morphemes presented in Table 4 (section 3.1.1). The same is true for the results of every previously conducted MOS bearing on L2/FL English (section 3.1.2). In our study we have found, for example, that our pre-intermediate group obtained higher accuracy rates for the *be_copula* morpheme than for the *be_auxiliary* morpheme. According to the percentages obtained the former is ranked in the first position whereas the second in the seventh (see Table 17 below). Note that whenever a "**tie**", i.e. two identical values, appeared in the data, a joined rank appears in Table 17. This represents the average of the ranks that they would otherwise occupy. Regarding the information displayed in Table 17, we would like to foreground that the **empty slots** do not imply absence of data, but rather reflect our decision to exclude the cases where our samples were less than 10 (**OC>10**) according to the criteria followed in previous studies (Muñoz, 2006; Goldschneider & DeKeyser, 2001), given that a group production of a morpheme lower than 10 is not sufficient data to reach a definite conclusion. Apart from table 17, for a detailed description of the accuracy rates for each morpheme by each of the four proficiency-level groups the reader is referred to Appendix 9.11.1. Note that these are also divided according to the proficiency level of the group studied. Additionally, each

of these descriptions contains a chart of the relation between TLU and NTLU (as defined in section 5.4.4) for every morpheme.

	El		PI		I		UI	
FUNCTOR	SOC	Rank	SOC	Rank	SOC	Rank	SOC	Rank
ART	63,63	**1**	84,05	**2**	91,74	**2**	96,53	**2**
PAST_IRREG	44,44	**2**	55,27	**5**	71,29	**6**	88,8	**4,5**
PAST_REG	38,46	**3**	57,18	**4**	70,79	**7**	88,8	**4,5**
3SG	5,5	**4**	21,86	**9**	17,9	**9**	0	**7**
BE_COP	-	**-**	92,46	**1**	95,67	**1**	100	**1**
ING	-	**-**	46,55	**8**	82,05	**4**	94,11	**3**
BE_AUX	-	**-**	47,25	**7**	79,23	**5**	100	**1**
POS	-	**-**	48,61	**6**	58,57	**8**	54,54	**6**
PLU	-	**-**	67,36	**3**	86,04	**3**	84,09	**5**

Table 17: SOC (accuracy rate) and rank orders for each proficiency level: El (elementary), PI (pre-intermediate), I (intermediate), UI (upper intermediate)

In the above table we can observe how the accuracy order for each morpheme (rank) changes according to the proficiency level. In that respect our study's results agree with Muñoz's (2006) findings that confirmed the **influence of the proficiency level** on the order of acquisition. In order to check Muñoz's (2006) claim that the accuracy orders of the foreign learners approach the average order "once they have progressed beyond the very elementary levels of proficiency" (p.123), we have compared the ranks achieved by each group to the ones achieved by the sum of all our subjects irrespective of their proficiency level. The results of this comparison are presented in Tables 18(a) and 18(b). The latter displays the SOC and rank orders for each group sorted by functor.

FUNCTOR	**All groups SOC**	**All groups Rank**	**Elem Rank**	**Pre-Int Rank**	**Int Rank**	**Upper-Int Rank**
ART	87,6	2	1	2	2	2
PAST_IRREG	73,7	4	2	5	6	4,5
PAST_REG	67,81	5	3	4	7	4,5
3SG	19,96	9	4	9	9	7

BE_COP	93,95	1	-	1	1	1
ING	67,56	6	-	8	4	3
BE_AUX	65,02	7	-	7	5	1
POS	53,52	8	-	6	8	6
PLU	75,8	3	-	3	3	5

Table 18(a): SOC and rank orders comparison

All groups	**Elementary Group**	**Pre-Intermediate Group**	**Intermediate Group**	**Upper-Intermediate Group**
BE_COP		BE_COP	BE_COP	BE_COP/AUX
ART	ART	ART	ART	ART
PLU		PLU	PLU	ING
PAST_IRREG	PAST_IRREG	PAST_REG	ING	PAST_IRREG
PAST_REG	PAST_REG	PAST_IRREG	BE AUX	PAST_REG
ING		POS	PAST_IRREG	PLU
BE_AUX		BE_AUX	PAST_REG	
POS		ING	POS	POS
3SG	3SG	3SG	3SG	3SG

Table 18(b): SOC and Rank orders comparison (sorted by functor)

In Table 18(b) we can see that the *be-copula* morpheme comes first for all groups except for the Elementary Level group for which we had not enough examples and hence excluded it from our study. The *article* and the *3SG* morphemes on the other hand seem to be stable in all groups, including the generic one (i.e., the group in which no distinction on grounds of the subjects' proficiency level was made). *Past regular* and *past irregular* appear simultaneously and occupy close positions in all groups. We will consider these two morphemes in detail below, but we wish to make a comment here bearing on the influence of the proficiency level. If we check the ranks occupied by *past irregular* and *past regular* in the generic group, we see that the former presents higher accuracy rates. This, however, can be misleading since the *regular* and the *irregular morphology* of the past tense in English interchange in a way their positions in

the different proficiency levels as our other groups rates reveal.

We can, therefore, claim that Tables 18(a) and 18(b) show the influence that the proficiency level exerts on the rank orders. Apart from our comment on the *regular past* and the *irregular past* ranks, we can also appreciate that there is a difference between the accuracy rates obtained by all our subjects (i.e. if we do not group them by proficiency level) and the corresponding rates obtained by the proficiency-level groups. Indeed, only the *be-copula*, the *article* and the *3SG* functors are ranked in similar positions by all proficiency-level groups and by the generic group. Note that although the *3SG* morpheme is ranked fourth (4) in our *Elementary Group*, it is still the morpheme with the least accuracy rates. To that observation we would like to add that there seems to be a high correlation between the average order, as presented by the results of our generic group, and the ones achieved by the more advanced students especially those that belong to the intermediate level. We can not consider the total number of subjects as an influential factor here, mainly because both the pre-intermediate and the upper-intermediate groups include an equally high number of participants as shown in Table 6(section 5.2). In that line we can argue that our study confirms Muñoz's claim that the **average order appears beyond the elementary level**.

Perhaps a more interesting conclusion can be reached through the observation of the data shown in Tables 17 & 18 regarding the **U-shape development** of the *past irregular* morpheme. Obviously, this observation can be made only when the proficiency level is taken into account, proving thus the relevant criticism of the first MO studies right. This, as we already mentioned, was related to the central assumption of Brown's (1973) SOC model that when a morpheme is being accurately used then it is also acquired (see sections 3.1.1 and 3.1.3). In relation to this developmental pattern Ellis and Barkhuizen (2005) say that

> […] when learners acquire English past irregular they frequently pass through an early stage of acquisition where they use some irregular forms correctly only to replace these later on with overgeneralised *-ed* forms […] (p.77)

The results of our study yield a similar pattern. Indeed, in the elementary level the SOC [8] percentage of the *past regular* morpheme is lower than that of the *past irregular* morpheme. In the following stage (pre-intermediate) the *regular morphology* surpasses the SOC percentage of the *past irregular* morpheme, only to fall back to a lower rank in the intermediate level. In the upper-intermediate level the differences are nil; indeed the percentage is exactly the same for both morphemes. The movement we just described can be appreciated in the Figure 9.

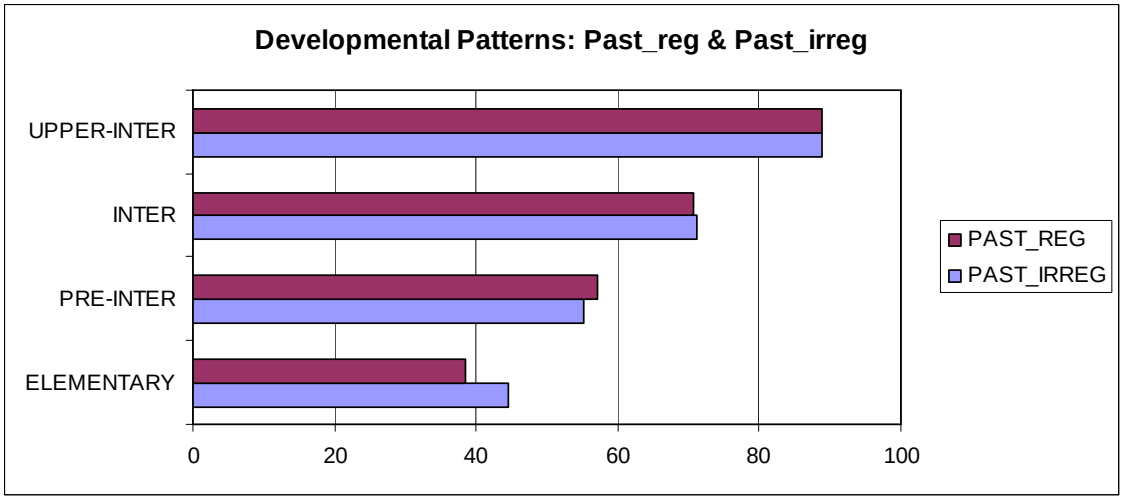


Figure 9: U-shape movement of the *past irregular* morpheme as seen by the comparison of the *past regular* and the *past irregular* accuracy rates.

Another interesting finding that results from our MOS project regards the morpheme accuracy orders for each proficiency level. We will not call it acquisition, as we just described that accuracy does not necessarily imply acquisition. We shall repeat, however, that the similarities in the route followed can be indicative of the acquisition order at each proficiency level

[8] Note that we refer to the SOC percentages calculated according to our scoring formula which includes Dulay and Burt's (1973) weighted scoring and uses Pica's (1983) TLU model. For more details we refer the reader to our section 5.4.5.

since we count with a large number of learner language data (corpus) and we have taken into account features that most of the first MO studies ignored (such as the cases of *underuse, misuse* and *overuse* or the subjects' *proficiency level)*. If we observe the accuracy rates for each of our four groups (i.e. from Elementary to Upper-Intermediate) in Table 17 we can see that there are **certain commonalities in the development** of the morpheme accuracy orders obtained by **each proficiency group**. This similar way of development is perhaps easier to appreciate in the chart displayed in Figure 10.

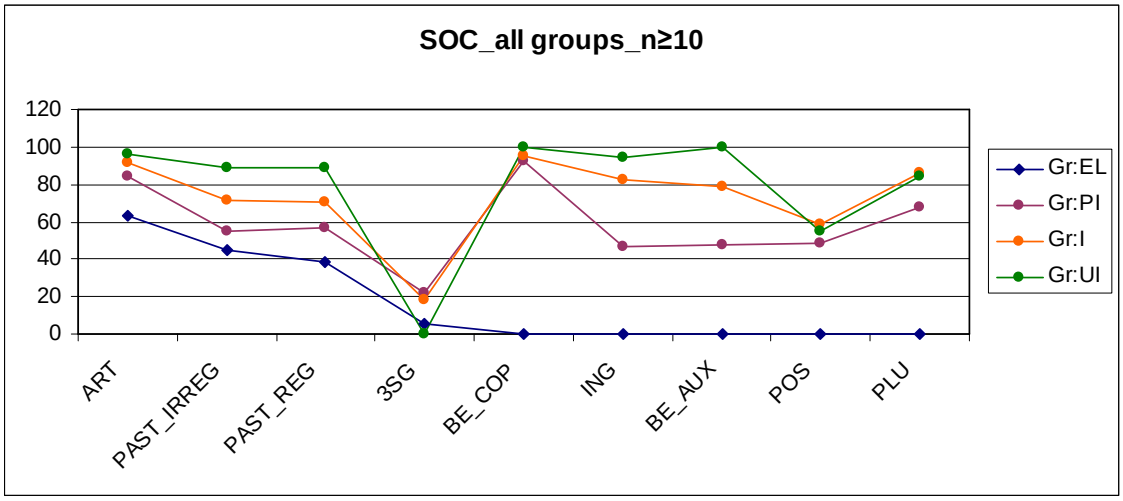


Figure 10: Accuracy order for each group (note that n ≥ 10)

In the above figure we can appreciate that the development of the morphemes' accuracy is similar for all four groups. There are **differences** in the **rates** obtained but in terms of the **route** followed the pattern displayed by each group is **similar**. We wish thus to repeat here the quote by Larsen-Freeman and Long (1991) where they affirmed that the observed similarities can not be ignored since they imply that "There is something moving in the bushes" (p.92).

The only significant differences are to be found in relation to the *past irregular* and the *past regular* morphemes on one hand and the *progressive -ing* and *be auxiliary* morphemes on the other. The former is explained by the U-shape developmental pattern as described above. The latter, however, requires our attention. In

relation to the *progressive -ing* and the *be_auxiliary* morphemes we only have data for the three more advanced groups, that is, the pre-intermediate (PI), the intermediate (I) and the upper-intermediate (UI). The last two groups exhibit the same accuracy order. In other words, at the **intermediate and upper-intermediate** levels students seem to use **more accurately** the ***progressive -ing*** **and** ***be auxiliary*** morphemes **than** the ***past irregular*** **and** the ***past regular*** **morphemes**. On the contrary, students at the **pre-intermediate** level seem to use **more accurately** the ***past irregular*** **and** the ***past regular*** morphemes **than** the ***progressive -ing*** **and** ***be auxiliary***. So, we could say that, in relation to these four morphemes, the pre-intermediate group follows the route: "past_reg" → "past_irreg" → "be_aux" → "-ing", the intermediate group follows the route "be_aux" → "-ing" →"past_irreg" →"past_reg" and the upper-intermediate group follow the route: "-ing" →"be_aux" →"past_reg"/"past_irreg". Table 19 displays the **routes** described for each group.

Level	**Morpheme**	**Morpheme**	**Morpheme**	**Morpheme**
PI	past_reg	past_irreg	be_aux	Progr_ing
I	be_aux	progr_ing	past_irreg	past_reg
UI	Progr_ing	Be_aux	past_irreg/reg	

Table 19: Route of accuracy obtained in four morphemes (past_reg, past_irreg, be_aux, progr_ing)

If we add to that route the accuracy rates of the *copula be* and the *3SG*, then we can create a table exhibiting the accuracy order for **all the verb-related morphemes** studied in our project.

Level	**Morpheme**	**Morpheme**	**Morpheme**	**Morpheme**	**Morpheme**	**Morpheme**
PI	Cop_be	Past_ Reg	Past_ Irreg	Be_aux	Progr_ Ing	3SG
I	Cop_be	Be_aux	Progr_ Ing	Past_ irreg	Past_ reg.	3SG
UI	Cop_be	Progr_ Ing	Be_aux	Past_ irreg/ reg		3SG

Table 20: Route of accuracy of the verb-related morphemes

If we transcribe the route shown in Table 20 according to the **syntactic properties** of each of the morphemes we will end up having the following schema:

(1) PI: Copula → tense (±past) → aspect (±progressive) →subject-verb agreement (±3rd person singular)
(2) I/UI: Copula → aspect (±progressive) → tense (±past) →subject-verb agreement (±3rd person singular)

The second pattern clearly fits to the schema proposed by Hawkins (2001) and Andersen (1978). The first, however, is slightly different, exhibiting higher accuracy levels in the use of the tense (±past) in comparison to the accurate use of the aspect (±progressive). **Two explanations** could be suggested for this phenomenon. The first would argue that the pattern displayed above by the students of the pre-intermediate level is due to the fact that our data elicitation task implied the telling of a story and thus directed the students towards the use of the past tense morphemes. Although we do admit that the instrument drastically influences the results, we do not consider this explanation to be very solid for two reasons. First, because the aforementioned accuracy rates are calculated as percentages after we assured that for each of the morphemes the minimum number of samples would be equal or higher than 10 and thence the assumed difference caused by the number of OC for each morpheme would immediately vanish. Second, because the same instrument was used with all our groups and yet the intermediate and the upper-intermediate level groups exhibit a different pattern. Therefore, the first option, although rational, cannot stand alone as an explanatory model of the described difference in route between the pre-intermediate group on the one hand and the intermediate and the upper-intermediate groups on the other.

We believe that our second explanation can more adequately account for this variance in route. It is related to the actual process of L2 acquisition. In that sense the aforementioned difference could be explained by the fact that in **initial stages**

L2 students produce output based on **rote memory** of individual words containing the past tense morpheme, that is, they **learn** the **past forms as chunks**. But when their **IL** goes through a **restructuring period** to accommodate additional features related to the expression of the past tense they incorrectly omit it. The same was argued by Lightbown (1983) in relation to the accurate use of the progressive morpheme. This idea proves Andersen's (1978) and Hawkins' (2001) claim right. Since the underlying syntactic properties related to the tense (±past) are more difficult than the ones found in the aspect (±progressive), as soon as the learner passes the stage of the memorization the accuracy order is reversed and fits the pattern proposed by Hawkins (2001) and Andersen (1978) and found in our study for the intermediate and the upper-intermediate levels.

This claim is also supported by the **dual-system model** suggested by Ullman (2001; 2005) in relation to the use of declarative and the procedural memory system in language acquisition. The former "[…] underlies the mental lexicon, whereas the procedural system subserves aspects of the mental grammar" (Ullman, 2001:718). His study shows that the forms of the ***irregular past*** are stored in the **declarative memory**, whereas the composition of ***regular forms*** is subserved by the **procedural memory**. Obviously, this does not imply that the declarative memory is not in use, but rather that the composition of the regular forms requires the activation of the memory system responsible for the mental grammar, i.e. the procedural memory. Ullman's model is supported by the U-shape pattern that the acquisition of the *past irregular* follows. The latter indicates that **in initial stages students memorize**, using their declarative memories, and are therefore more accurate in the production of the irregular forms, whereas in **higher proficiency levels** this is reversed because students proceed in the actual **manipulation of the language** by means of the procedural memory system. In that higher stage the aspect (±progressive) based on a more regular and thus easier set of rules, displays high accuracy rates, whereas the tense (±past) based on a more complex and hence difficult set of rules

exhibits low accuracy rates. On the contrary, in initial stages the use of the aspect and the tense morphemes rely on the use of the declarative memory and on which of the morphemes the student has best memorized.

If this rationale is right then the elementary group should display a similar pattern. As we mentioned before, we have decided to exclude the rates for those morphemes of which the overall number of samples was lower than 10 (n<10). At the elementary level these also included the use of the morphemes related to the aspect. However, we will present them in the following Table (No 21) and the corresponding chart (Figure 11) in order to see the elementary students' performance in relation to the verb-related morphemes.

Functor	Gr:El	Gr:PI	Gr:I	Gr:UI
BE_COP	75	92,46	95,67	100
PAST_IRREG	44,44	55,27	71,29	88,8
PAST_REG	38,46	57,18	70,79	88,8
ING	33,3	46,55	82,05	94,11
BE_AUX	33,3	47,25	79,23	100
3SG	5,5	21,86	17,9	0

Table 21: Accuracy rates (SOC) of the verb-related morphemes obtained by all four groups

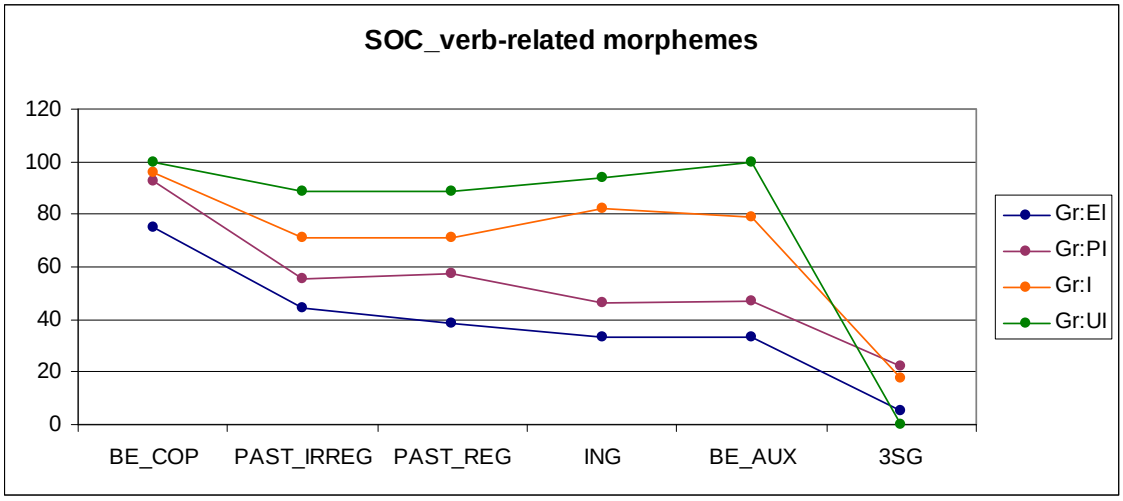


Figure 11: Accuracy rates (SOC) of the verb-related morphemes obtained by all four groups

Indeed, as we can see in Table 21 and in its chart (fig. 11), the very initial stages, namely the **elementary and** the **pre-intermediate** display the **same route** regarding the use of the **verb-related morphemes** as shown in the SOC scores. The same is true for the two higher levels, that is, the **intermediate and** the **upper-intermediate**. Hence we can affirm that Hawkins' (2001) schema works in higher proficiency levels since it is there that students actually proceed in the composition of linguistic forms. In former stages, Hawkins' (2001) pattern is observed only partly due to the extended use of the declarative memory system, which disregards the actual difficulty of each morpheme based on the underlying structural properties.

Salience could account for the perceived "preference" of the students to memorize *irregular past* first, then *regular past* and finally the *aspect*. Note that this was one of the variables predicting accuracy in MO studies that Goldschneider and DeKeyser (2001) included in their meta-analysis as mentioned in section 3.1.2. In their study salience is defined as the **ease with which a morpheme is perceived**. They state that the perceptual salience of the morphemes reviewed in their meta-analysis was calculated on the basis of "[…] the number of phones, the presence/absence of a vowel in the surface form, and the total relative sonority of the functor" (Goldschneider and DeKeyser, 2005:23). They also mention "stressed/unstressed and serial position" (Goldschneider and DeKeyser, 2005:23) as possible subfactors composing the score of perceptual salience, but they do not include them in their study. In our study we do not determine the score of perceptual salience for each of the aforementioned morphemes but we understand that they all share these factors. It may be true that *progressive –ing* and *past –ed* are placed at the end and that gives them an asset in comparison to the *irregular past* morpheme as argued by Slobin (1971, cited by Goldschneider and DeKeyser, 2005:22). Nevertheless, *irregular past* is *stressed* and that could counterbalance the effect of *serial position* exhibited by the *progressive –ing* and the *past regular*. In order to conclude regarding the role of perceptual

salience on the accuracy rates obtained for each of the verb-related morphemes at issue, we need to determine the score of the perceptual salience according each of the aforementioned subfactors. We consider that this surpasses the limits of the present paper, and hence would have to deal with it in a separate study. Another possible explanation is **classroom input,** in that in initial stages most of the classroom talk relates to what students did, saw, etc., and thus provides the learners with more opportunities not only to receive input in relation to the past tense morphemes but also to use them and hence memorize them. That would partly correspond to what Goldschneider and DeKeyser call **frequency in the input** which is defined as "[…] the number of times a given structure occurs in speech addressed to the learner" (Goldschneider and DeKeyser, 2001:29). In the same line, the *irregular past* morpheme is more likely to receive frequent **teacher feedback.** This would probably correspond to the "other factors that cause some parts of the input to *become* salient" as Dulay and Burt suggested (1978, cited by Goldschneider and DeKeyser, 2001:22). We do agree with Goldschneider and DeKeyser (2001) in that a combination of factors is responsible for the accuracy rates in each morpheme obtained by L2/FL students (see section 3.1.2 for more details), but, as we mentioned above, the determination and study of each of the proposed factors exceeds the scope of this paper.

Additional morpheme order analysis: our study compared to previous research

In the first part of the present sub-section we have presented and commented only on our study's results. We would like to compare our learner corpus-based study to the findings of previous morpheme order studies. For this purpose we will use the studies by **Muñoz (2006)**, **Lightbown (1983)**, **Dulay and Burt (1974)** and Krashen **et al. (1977)**. The last two are included as representative of the first MO studies in **ESL** contexts. The other two are included because they are both MO studies of **EFL** and in that aspect they share a fundamental element with our

study. Additionally, our study and Muñoz's study also share the distinction between proficiency levels and the subjects' L1 (Spanish).

The following table includes the results of each of these studies as well as of our study. We should remind here that Muñoz's groups A1, B1, D1, D2, A3 & B3 were constructed on the basis of the onset age and their proficiency level according to the hours they had been exposed to English.

FUNCTOR	Krashen's rank	Dulay& Burt's rank	Lightbown's rank	Muñoz's A1 rank	Muñoz's B1 rank	Elem rank
ING	1	3	3,5	3	5	-
PLU	2	8	3,5	1	1,5	-
BE_COP	3	2	2	2	1,5	-
BE_AUX	4	5	5	-	4	-
ART	5	1	1	4	3	1
PAST_IRREG	6	7	-	-	-	2
PAST_REG	7	6	-	-	-	3
3SG	8	9	6	5	6	4
POS	-	8	-	-	-	-

FUNCTOR	Muñoz's A3 rank	Muñoz's D1 rank	Muñoz's B3 rank	Muñoz's D2 rank	Pre-inter rank	Inter rank	Upper-inter rank
ING	5	5	5	4	8	4	3
PLU	2	2	1	2	3	3	5
BE_COP	1	1	2	1	1	1	1
BE_AUX	3	4	3,5	5	7	5	1
ART	4	3	3,5	3	2	2	2
PAST_IRREG	6	6	6	6	5	6	4,5
PAST_REG	7	7	7	7	4	7	4,5
3SG	8	8	8	8	9	9	7
POS	-	-	-	-	6	8	6

Table 22: Comparison of rank orders for all our proficiency level groups, Muñoz's groups, Lightbown's group and Krashen's average order.

❶Comparison between Muñoz's (2006) data and our data:

At first it seems that our accuracy orders are very different from the findings of the other studies. Nonetheless, a more careful examination of the data reveals a correlation especially between the accuracy rates of our pre-intermediate to upper-intermediate groups and those found in Muñoz's A3, B3, and D2 groups. There are only a few differences, like the order for the *be_auxiliary* in which our groups and Muñoz's groups do not correlate, or the rank of the *progressive -ing* functor in our pre-intermediate group which is different, but the rest of the groups do correlate.

The upper-intermediate group in our case also shows orders different to those of Muñoz's high proficiency level group. Two explanations could be offered for such a divergence. The first is the small number of upper-intermediate items that we have obtained in our study. The second explanation is that Muñoz (2006) determines the proficiency level of her subjects on the basis of the hours of exposure to the target language according to which the most proficient group has been exposed to English during 726 hours. This could correspond to our intermediate group. Indeed, Muñoz (2006) presents three proficiency levels starting from the elementary one. Therefore, we can assume that there is a correspondence, in terms of proficiency level, between Muñoz's (2006) groups and our groups as seen in Table 23. The chart of the accuracy rates for each proficiency correspondence displayed in Table 23 is seen in Figure 12.

Muñoz's (2006) groups	Our groups
A1, B1, D1: 200h of exposure to the target language	Elementary Group (El) Elementary level scored in the placement test
D2: 416h of exposure to the target language	Pre-intermediate Group (PI) Pre-intermediate level achieved in the placement test
A3, B3: 726h of exposure to the target language	Intermediate Group (I) Intermediate level scored in

	the placement test
-	Upper-intermediate Group Upper-intermediate level scored in the placement test

Table 23: A comparison of our subjects' composition and the one found in Muñoz's (2006) study

Figure 12: Comparison of the accuracy rates obtained by the pre-intermediate and the intermediate groups in Muñoz's study and in our project

In Figure 12 we can better appreciate the **commonalities** in the orders that each group exhibits. There are differences in the actual **rates** obtained but in general the **routes** are similar. Note that our upper-intermediate group had no equivalent in Muñoz's study, since we brought Muñoz's groups together on the basis of the hours of exposure to the target language. Additionally, we do not present a comparison between our

elementary group and Muñoz's A1, B1 and D1 groups because in every case there are many rates that have been excluded (due to the "n≥10" condition followed in both studies) so an actual chart is impossible. Furthermore, we have already proved, as Muñoz did before, that proficiency level is highly influential and the average rates appear beyond the very initial stages.

We could therefore say that our study does correlate with the results of the other EFL study that grouped the subjects according to proficiency level. Any differences found could be explained by the fact that we have not grouped our subjects according to the onset age, which could have influenced the final data. Additionally, the difference displayed in relation to certain verb-related morphemes among our pre-intermediate group and Muñoz's D2 group could probably have been caused by the difference on the two groups' age of onset as well as age of testing. For more details we refer the reader to section 5.2 where we present the subjects of our learner corpus as well as to the relevant discussion previously displayed (section 6.1). Another explanatory factor could be the data collection method: oral in Muñoz's study and written in our case. Differences between previous studies (see for example Perkins & Larsen-Freeman vs. Dulay & Burt, section 3.1.2) have been explained in a similar way based on the fact that speaking and writing are influenced by different sociolinguistic and psycholinguistic conditions (Ellis: 1994).

❷Comparison between Lightbown's (1983) data and our data:

A first comparison of our findings with Lightbown's rank would yield a sole correlation between her group's orders and our elementary group's orders for just two functors. On the other hand, these are the only two functors with data that are common in both studies. That is, the only two functors common in the data for our elementary group and in Lightbown's study are the *article* and the *3SG*. For both groups (i.e., our elementary group and Lightbown's group) the *article* morpheme comes first in rank, whereas the *3SG* morpheme comes last. Indeed, further examination

of Lightbown's findings in relation to our other groups' orders leads to the conclusion that there is no correlation. This could be explained by the fact that Lightbown's group consists only of Grade 6 learners but no further distinction is made with regard to the subjects' proficiency level.

❸Comparison between Dulay & Burt's (1974) and Krashen's (1977a) data and our data:

In relation to the correspondence between our orders and those suggested by the first MO studies as seen by Dulay & Burt's and Krashen's example, we should say that our findings do not generally support the order suggested by these scholars. It is true that regarding some functors we do find similarities in the accuracy rates, but that cannot be regarded as significant of the overall order suggested. In that sense, we disagree with Wagner's (2005) claim that her study provides evidence that "the acquisition order of the grammatical morphemes by English as a Foreign Language learners is similar to the order of English as a Second Language learners" (p.34). Wagner compares her findings with those put forth in the study of Dulay and Burt (1974) and she indeed finds a significant correlation between the two studies' findings. But that correlation regards only 3 out of the 9 morphemes investigated by Dulay and Burt. Hence, even though there is a significant correlation regarding these 3 morphemes, we disagree with her decision to extend said relationship to the accuracy order of the grammatical morphemes in EFL as the aforementioned quote suggests. Indeed, we also find commonalities between our findings and the ones suggested by Dulay and Burt (1974b) regarding three morphemes (3SG, Art, Pos). Furthermore, we could say that there is a relative correlation between our findings and the order suggested by Krashen (1977a) as displayed in Figure 2. We say that there is a relative correlation because again there are similarities in relation to some morphemes only. We shall, however, say that the difference between the average orders proposed by Krashen and Dulay and Burt, on the one hand, and our orders, on the other hand, can probably be explained by the fact that their study was

an ESL study whereas ours was an EFL study. The proficiency level does not seem to be influential since, if we compare Krashen's average order (Table 22) and the orders of our general group for which we have not taken into account the subjects' proficiency level (Table 18a), we see that there is still no correlation except maybe for the case of the 3SG functor. Another possible explanation could be the influence of the data elicitation instrument or even the scoring method. Indeed, Pica's (1983) calculation model (included in our study's scoring) gives lower scores than those achieved by Brown's (1973) model (used in the other two studies). For more details regarding Brown's (1973), Pica's (1983) and our calculation model see section 5.4.5. In relation to the influence of the scoring method on the accuracy rates and thus the morpheme orders obtained we refer the reader to our discussion in section 3.1.3.

Let us focus now in more detail on the results regarding the *genitive –s* morpheme in the L2 English sentence transformation task.

6.2 Genitive –s in L2 English

Before presenting the results of our **second study**, we would like to remind that the majority of our findings regarding the accuracy rates of the *possessive –s* have resulted from our sentence transformation task and hence represent **experimentally elicited data** (see sections 5.1, 5.1.1 & 5.1.2 for a discussion of corpus vs. experimental data). Recall that the idea of using an elicitation task like this was justified by the low rates of *genitive –s* produced in the corpus, due to the nature of the corpus task. In this section, we shall present the results of our experimentally elicited data and then compare them to the corresponding findings from our learner corpus.

In relation to the use of the ***possessive –s*** morpheme we wish to investigate **two issues**:

(i) The first is Krashen's (1978) claim that **discrete-point grammar tests**, in which learners' formal

knowledge is appointed, will show the order of learning, whereas the **naturally produced language** will be displaying different order of acquisition (that is, different according to Krashen). If his claim is right then the SOC rates for the *possessive -s* in our learner corpus will be significantly different from those obtained in our sentence transformation task.

(ii) The second question relates to the **L2 learner preference for** the analytic form of expressing the possession, that is to say, **the use of the *PP*** construction. This preference is assumed by the fact that in the first MO studies the synthetic form of possession, i.e. the *possessive –s*, occupies one of the last position in their morpheme acquisition orders.

Note that regarding the first question we will be comparing the **SOC scores** since we wish to check the **acquisition order**, whereas the **TLU scores** will be contrasted to the NTLU rates in relation to our second question, where we will be studying the **frequency rates** (for more details regarding this distinction please check section 5.4.5).

Question 1 (Will different methods yield different SOC rates?). Table 24 displays the **accuracy rates** of the ***possessive –s*** obtained by all our groups in the **sentence transformation task**. A detailed description of the TLU and NTLU rates along with their chart and a full account of the NTLU types are displayed in Appendix 9.11.2.

Group	SOC	OC
Gr: El	44,44	108
Gr: PI	67,92	636
Gr: I	90,09	606
Gr: UI	90,59	186

Table 24: SOC scores for *possessive -s* in the sentence transformation task

Based on the data presented in the table above we can affirm that the accuracy orders for the *possessive –s* are very high in the three groups that correspond to higher proficiency levels. In that sense, Muñoz's (2006) claim that the **proficiency level** is "a stronger determinant factor" (p.122) is confirmed also by the results of our experimentally elicited data. However, if we wish to check the influence of the instrument on the actual results and hence study Krashen's hypothesis on the learning/acquisition distinction we should compare the results displayed in Table 24 to the relevant accuracy rates obtained by our subjects that participated in the learner corpus study. Table 25 shows the corresponding contrast.

Group	SOC in sentence transformation task	SOC in story narration (corpus)	Difference
Gr: El	44,44	0	44,44
Gr:PI	67,92	48,61	19,31
Gr:I	90,09	58,57	31,52
Gr:UI	90,59	54,54	36,05

Table 25: SOC of the *possessive –s* in the two tasks

As we can observe in the above table the **accuracy rates** obtained by the participants of the **sentence transformation task surpass** the **results** obtained by the **learner corpus** groups. Since the SOC rates and the corresponding difference are calculated as a percentage we consider that the difference is significant in all cases. The least significant difference is to be found among the learners of the **pre-intermediate level**, but mainly due to the

low accuracy rates found in the experimentally elicited data. This could be explained by the **plateau effect** (Richards, 2008) in SLA. Indeed, that seems to be the right explanation in this case since it is only in the transition from the elementary to the pre-intermediate level that the progress rate seems to be significantly low. The difference in accuracy rates between the intermediate and the upper-intermediate group is also low, but this is due to high accuracy scores in both cases (>90%).

Going back to the comparison of the data elicited by means of the two different tasks and the relevant question we posed in the beginning regarding the **influence of the research instrument** on the results we can affirm that there is a clear correlation and that Krashen's claim is right. That is to say, it seems that discrete-point grammar tasks indeed trigger the subjects' learnt rules, i.e. formal knowledge, and hence the accuracy rates obtained are significantly higher than those found in naturally produced language.

Question 2 (Will rates for the analytic *genitive PP* be higher than for the *synthetic –s* morpheme?).
Before dealing with the second issue we wished to investigate, that is, the **choice** of the L2 English learners **between** the **analytic** and the **synthetic forms** in order to express **possession** in **English**, we should repeat our **initial hypothesis**. According to this **L2 English learners** will exhibit a clear **preference** for the **analytic form** especially in the initial stages.

In order to (dis)confirm this hypothesis, we should compare not only the TLU and NTLU rates but also the types of NTLU found in each proficiency level. We remind here that we have tagged the use of the analytic possession construction (i.e. the *PP*) as "NTLU: misuse: misselection" in order to be able to identify it and thus make the comparison. Recall that this does not imply low rates of accuracy in the use of the possession structure in general. These rates are represented by the SOC, as shown in Table 24. The following table (no. 26) displays the TLU and the

NTLU rates obtained by the students that participated in our sentence transformation task. In Figure 13 we can see the chart of these results.

Group	**TLU**	**NTLU**
Gr: El	25	75
Gr: PI	58,6	41,4
Gr: I	83,3	16,7
Gr: UI	89,2	10,8

Table 26: TLU and NTLU rates in the use of the *possessive –s* by L2 learners of English (sentence transformation task)

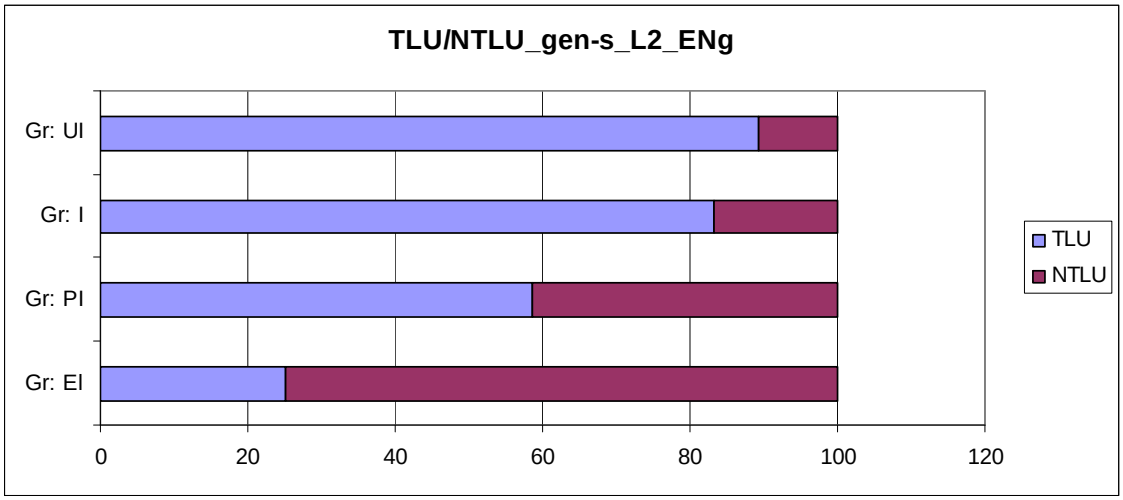


Figure 13: TLU/NTLU of the *possessive –s* by L2 learners of English (sentence transformation task)

As we can see the accurate use of the ***possessive –s*** (i.e., TLU) **correlates** with the learner's **proficiency level**. That is, the higher the student's proficiency level the more accurate the use of the *possessive –s*. We should remind here that we constructed our sentence transformation task in such a way that both the synthetic (*-s possessor*) and the analytic (*PP*) forms could be used. That enabled us to secure equal possibilities of use for each of the forms and thus diminish the possible effects of the instrument used. For more details we refer the reader to the corresponding section (5.3.4).

However, in order to determine whether this correlation also displays a preference of the high proficiency students for the

synthetic *genitive –s* form (as seen in the TLU rates) and the opposite tendency of the low proficiency learners, we should further analyse the data. Indeed, if we observe the types of NTLU we can see that the majority of NTLU cases exhibited by all four groups do not belong to what we classified as *PP*. The results are shown in Table 27.

	Gr:El	**Gr:PI**	**Gr:I**	**Gr:UI**
NTLU_Type	75% (n=81)	41,4% (n=263)	16,7% (n=101)	10,8% (n=20)
Underuse	77,8%	35,9%	18,8%	75%
Misuse	22,2%	64,1%	81,2%	25%
Overuse	0%	0%	0%	0%
Misuse_Type	N=18	N=118	N=82	N=5
Misselection	22,2%	28%	34,1%	0%
Misrealisation	27,8%	34,7%	41,5%	60%
Wrong-unclassified	50%	37,3%	24,4%	40%

Table 27: NTLU types of the *genitive –s*

Group	**TLU**	**NTLU: PP**
Gr:El	25%	2,7%
Gr:PI	58,6%	12,54%
Gr:I	83,3%	27,72%
Gr:UI	89,2%	0%

Table 28: Relative frequency of the *possessive –s* and the *PP* as possession expressions

The data in Tables 27 and 28 clearly show that the misselection type, that is, the use of the *periphrastic possession* structure *(PP)* is significantly low.

According to these data we should say that our hypothesis is disconfirmed, that is, learners at all levels show a clear **preference for** the synthetic ***(genitive –s)*** form of the possession structure in English. In the same line we can argue that the low accuracy rates observed in the initial proficiency

level groups are not due to the students' preference for the use of the analytic form for the expression of possession.

The last comment we would like to make here regards the relation between our findings and Rosenbach's (2005) study, i.e., we want to contrast **L2 English learners vs. English natives** on the use of the synthetic *genitive –s* morpheme vs. the *possessive PP* analytic structure. In Figure 3 above we showed Rosenbach's (2005) findings on the frequency of the *genitive –s* among native speakers of English. According to that, native speakers prefer the use of the *genitive –s* when the possessor is an animated entity. In our study's task all possessors were animated entities and the students showed a clear preference for the use of the *possessive –s*. That tendency cannot be explained by an L1 transfer hypothesis. The subjects of our study were native speakers of Spanish which displays a different possession structure. In Spanish one can express possession either by means of the *possessive pronouns* (e.g. *He leído todos sus libros*), which would be the synthetic form, or through a *prepositional phrase* (e.g. *Este es el coche de María*), which would obviously reflect the analytic form. English, on the other hand, has two synthetic forms. One is the *possessive –s* which has no equivalent in the Spanish possessive system. The other, that is, the *possessive pronouns*, is found in both the English and the Spanish system of possession. However, as we already mentioned in sections 2.2.1 and 5.3.4, *possessive pronouns* were included in the given sentences. Hence the only common possession form that was still available for the students to use in the task was the periphrastic form (i.e., the *prepositional phrase*). But we saw that the students preferred the use of the synthetic *possessive –s* instead. Therefore the **L1 transfer explanation is rejected** in this case. Nonetheless, we cannot suggest that the possessor's characteristic as an animated entity is what triggered the use of the *genitive –s* by our learners of L2 English. We can say that our findings imply such a correlation, but in order to make a positive statement in this regard we should check the whole spectrum of possibilities and thus

include also inanimate entities and distinguish between neutral and long/short. As we mentioned in section 2.2.1, we did not include a similar distinction in our study, but we consider that our findings could be used in a future research that would also contain data in relation to inanimate and neutral or long/short possessors. Regarding this issue we refer the reader to the relevant discussion presented in sections 7.4 and 7.5.

After having analysed the expression of possession in L2 English, we shall move to our 3rd study: the expression of possession in L3 German via the *genitive -s* morpheme and via other structures. Recall that a full account of the German system of possession is given in section 2.2.2.

6.3 Possessive structures in L3 German

In this section we wish to explore the use of each of the possession structures in German by students who learn German as a third language. As we mentioned in section 2.2.2 in German there are **structural restrictions** as to which type of possession structure one can use. Furthermore, in German there are four ways of expressing possession. The first is, as in English and Spanish, the *possessive pronouns* (e.g. Das ist *sein* Auto; lit.: This is *his* car). In keeping with our choice in the case of English, we decided to include the *possessive pronouns* in the given sentences of the transformation task in German. This enabled us to follow a similar tagging scheme in English and in German and thus allowed us to make comparisons. Additionally, it helped us check whether there is or not an L2 English influence on the acquisition of the L3 German possession system, but to this we will return later.

Before proceeding to the presentation of our findings, we should mention our principal aims:

(i) First we wish to check the **general preferences** of the **L3 German learners** in relation to the synthetic and the analytic forms for **expressing possession**.

(ii) Then we would like to **compare** these results to the **relative frequency** of the *possessive –s* and the *PP* **in L2 English** (see study no. 2 in the previous section).

(iii) Finally, we intend to determine whether there is or not an **influence** of the **L2** English on the acquisition of the **L3** German.

The majority of our subjects (77,6%), were learning English as an L2. The remaining 22,4% either did not mention any other L2 or mentioned that they were learning a different L2. In most of these cases, that is, at least regarding the young subjects of the remaining 22,4%, we believe that they have gone through the process of learning English as a foreign language, at school for example. Nevertheless, since we cannot know whether they meant that they are not currently learning or that they have never learnt English as foreign language, we decided to exclude these samples from the last subsection of our third study.

In order to study the tendencies in the use of the **synthetic** and the **periphrastic possession** structures by L3 German learners, we should compare the **TLU** and the "**NTLU: misuse: misselection**" accuracy rates for each proficiency level. Table 29 displays the relevant information as resulted from the analysis of our raw data. These represent the accuracy rates of the two synthetic forms of possession in German, i.e. the *possessive –s* and the *genitive case*, as well as the accuracy rates of the periphrastic form, that is, the *PP*. Note that the percentage has been calculated on the basis the total number of OC that appeared in each group. A detailed description of the TLU and NTLU rates along with their chart and a full account of the NTLU types are displayed in Appendix 9.11.3.1.

Group	**TLU**	**NTLU:**	**misuse:**	**misselection:**	**PP**
Elementary (Gr:El)	5,8%	94,2%:	85,8%:	89,7%:	40,2%
Pre-intermediate (Gr:PI)	14,1%	85,9%:	83,6%:	83,9:	59,6%
Intermediate (Gr:I)	12,7%	87,3%:	85,5%:	96,2%:	70,6%
Upper-intermediate (Gr:UI)	41,7%	58,3%:	100%:	50%:	71,4%

Table 29: Relative frequency of the synthetic and the analytic possessive forms in L3 German learners' language

The third column in Table 29 presents the accuracy rates of the NTLU, the "NTLU:misuse", the "NTLU:misuse:misselection", and finally the "NTLU:misuse:misselection:PP". So, for example, in the case of the elementary level group, these data mean that in relation to all the obligatory contexts (OC) for the expression of possession, the students of the elementary group presented a 94,2% of **NTLU**, out of which the 85,8% were cases of **misuse**. That is, the remaining 14,2% represents other types of NTLU such as underuse or the so-called SNOC. Then, the 89,7% out of this 85,8% of "NTLU:misuse" reflects the **misselection** type which includes both the use of the *PP* and the use of the *possessive –s* in contexts where a different possession form was required (for more details on our tagging scheme for L3 German see section 5.4.4 part 3). We decided not to tag this as a case of SNOC because these reflect OC for a possession structure. The mistake lies in the choice of the particular form of possession expression (*possessive –s*) in those occasions where the corresponding structural restrictions allow the use of either the *genitive case* or the *PP*. If we had included it in the SNOC we would have erroneously distorted the calculation of the SOC, which we will be using for the L2 English (study no. 2 above) and L3 German (this study, no. 3) comparison later. Finally, the 40,2% of the misselection cases represents the use of the PP.

This is the full analysis of the percentages presented in Appendix 9.11.3.1. Note that displaying just the 40,2% of the PP use would be misleading since it is not estimated on the basis of the total number of possession expression OC. However, since it is difficult to reach a conclusion from these percentages, we have transformed the PP use percentage into a rate based on the overall OC. The rates were calculated following the aforementioned process for every group and therefore represent the percentage of *PP* occurrences in the corresponding obligatory contexts. We present the results in the following table.

Group	**TLU: genitive 's & genitive case (in %)**	**NTLU: PP (in %)**
Elementary Gr:El	5,8	29,16
Pre-intermediate Gr:PI	14,1	35,89
Intermediate Gr:I	12,7	50,70
Upper-intermediate Gr:UI	41,7	20,83

Table 30: Relative frequency of the *genitive –s*, the *genitive case* and the *PP* in the L3 German learners' language

In Table 30 we can see that in general the **accurate use** of the ***PP*** has not obtained very high rates, as they are 50% or below. Nevertheless, these rates are still higher than the equivalent rates of the TLU. Only the upper-intermediate group obtained higher accuracy rates in the use of the synthetic forms of possession expression *(genitive -s and genitive case)*. Yet, the difference between the use of the synthetic and the analytic forms at this level is significantly lower than the corresponding difference in the other three levels. This can be clearly appreciated in Figure 14.

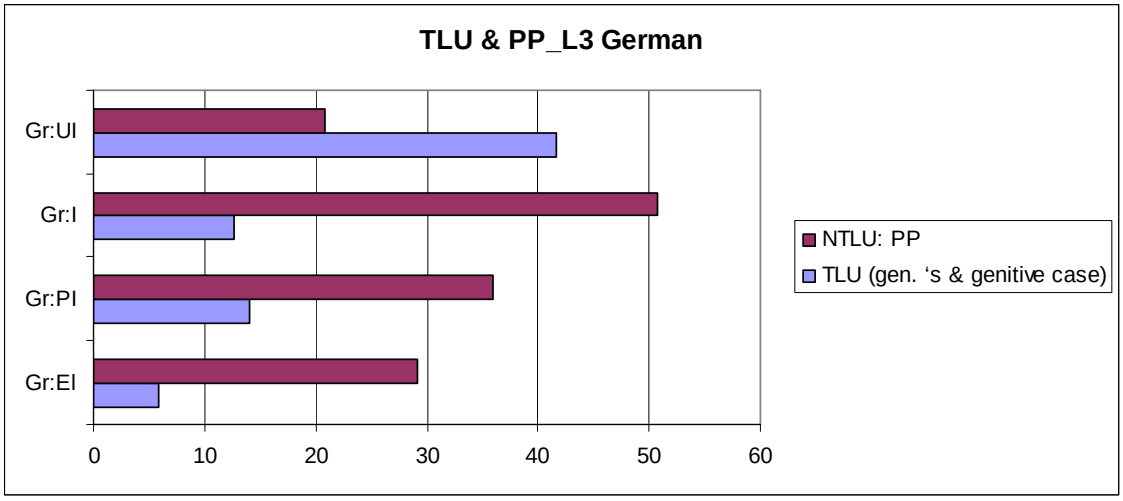


Figure 14: Relative frequency of the *s-genitive* (TLU), the *genitive case* (TLU) and the *of-genitive* (NTLU:PP)

Based on the accuracy rates obtained for the *s-genitive* and the genitive case on the one hand, and those obtained for the *of-genitive* on the other, we can claim that in general the **L3 learners of German** show a **preference for the use of the analytic form**. However, explaining this preference on grounds of an L1 transfer would be a mistake for two reasons. First, because the *of-genitive* structure may be part of the possession expressions in our subjects' L1 Spanish, but it also forms part of their L2 English. This implies that **L1 transfer cannot account for the L3 German data** here. Second, because the overall percentage of the periphrastic form of possessive expression is low in our data. Additionally, the preference shown in Figure 14 is determined only in relation to the TLU cases, but that does not mean that the *PP* is the learners' first choice when it comes to possession expression in German. The remaining percentages in each level represent the cases where the *genitive –s* has been used incorrectly. We shall deal with these rates later when we will be examining the influence of the L2 English on the L3 German.

We should now pass to our second question that refers to the comparison of relative frequency of the **synthetic** and the **analytic** forms of possession **in both L2 English and L3 German**. We have decided to use the experimentally elicited L2 English data (see study no. 2 in the previous section) in order to avoid

possible interferences of the instrument. In the following table we present the data for both the L2 English and the L3 German.

	TLU		**NTLU: PP**	
Group	**L2_Eng**	**L3_Ger**	**L2_Eng**	**L3_Ger**
Gr:El	25	5,80	3,7	29,16
Gr:PI	58,6	14,10	5,1	35,89
Gr:I	83,3	12,70	4,6	50,7
Gr:UI	89,2	41,70	0	20,83

Table 31: Accuracy rates for the use of the synthetic and the analytic forms of possession expression in L2 English and L3 German

Based on the data displayed in Table 31 we can argue that the use of the **synthetic form** (*genitive* morphology) is **more extended** among the learners of **L2 English** than among the learners of **L3 German**. In the same line, the use of the periphrastic form (*PP* structure) in L2 English displays very low rates. In comparison to this, the use of the *PP* made by L3 learners of German is significantly higher. That could probably be explained by the fact that in English there are no structural restrictions and thus the achievement of the TLU of the *genitive -s* is much easier for the learners. On the contrary, the various structural restrictions posed in the use of each of the possession forms in German probably has a negative interference in students' accuracy rates.

Before we set about checking whether there is or not an **L2 influence on L3 acquisition**, we should clarify where this influence, if it exists, should appear. As we mentioned in section 2.2.1, English has two ways of expressing possession, namely the synthetic and the analytic. The former is constructed by means of the *possessive pronouns* or the *genitive -s*. The analytic form is constructed by means of the *prepositional phrase*. In German, on the other hand, there are also two ways of expressing possession, that is, the inflectional and the

periphrastic. The inflectional possession expression in German includes the *possessive pronouns, the genitive –s and the genitive case*. The latter is the only formal feature that English lacks. Furthermore, in German the use of the *genitive –s* is restricted in those cases where the *possessor* is expressed by means of either a proper name or a kinship term (see section 2.2.2 for more details). English, on the contrary, does not pose such restrictions to the use of the *possessive –s*. As we saw briefly in section 2.2.1 various factors have been suggested as a possible explanation of the native speakers' tendency to use one possessive form over the other. But there are no rules that allow or prohibit the use of each of the possessive forms except for the relative order of the *possessor* and the *possessum* (i.e., the PR>PM or the PM>PR). For more details in this respect we refer the reader to our sections 2.2, 2.2.1 and 2.2.2. So, the only differences in the possession expression systems in English and in German are the two extra features described above that we find in the German possession. These are the ***genitive case*** and the ***structural restrictions*** regarding the choice of the possession form. In relation to the first, we could say that the use of other possession structures in those occasions where the *genitive case* is required would imply an influence of the L2 English. Nonetheless, this is not a sound assumption, since all the OC for the *genitive case* in German are also the OC for the *PP* (see section 2.2.2). Therefore, the only area where we can look for a possible influence is the use of the *genitive –s* in those structures where either a *genitive case* or a *PP* is allowed. As we explained when we described our tagging process, we called these cases "NTLU: misuse: misselection: possessive –s" in order to distinguish them from the cases where a *genitive –s* was the correct choice.

Before presenting our results we should highlight that we have included in our task the OC for all the forms of the possession expression. Additionally, we have excluded, as we said in the beginning of this subsection, those samples where the subjects did not expressively said that they were or had been at some

point in their lives learning English. The relevant results of the remaining samples are displayed in the following table. The whole description of the statistic features for this sub-group of L3 German learners can be found in Appendix 9.11.3.2.

	TLU		**NTLU**	
Group	**gen_s**	**gen-case**	**PP**	**Pos. –s**
Gr:El	5,3	0	28,07	43,85
Gr:PI	7,5	7,5	28,78	28,78
Gr:I	6,25	8,3	35,41	27,08
Gr:UI	-	-	-	-

Table 32: Relative frequency of the *genitive –s*, the *genitive case*, the *of-genitive* and the misused *possessive –s* in the language of L3 German learners with L2 English

In the above table we can see that the **accurate use** of the three forms of possession expression (i.e., the *genitive –s*, the *genitive case*, and the *PP*) correlates with the **proficiency level**. In the same vein, we observe that the use of the *PP* is lower than the wrong use of the *possessive -s* at the elementary level. The same rates are equal in the pre-intermediate level, and in the intermediate level the use of the *PP* is higher than the inaccurate use of the *possessive –s*.

However, what we find even more interesting, is the fact that the rates of the inaccurate use (i.e. the use in the wrong contexts) of the *possessive –s* is significantly higher not only compared to the accuracy rates for each of the synthetic forms, i.e. the *genitive -s* and the *genitive case*, but also in relation to the overall TLU rates for both synthetic forms. That is, **the *genitive –s* morpheme is most frequently used by our L3 German learners in the wrong contexts** (i.e., wrong according to the structural restrictions imposed in German). That can only be explained by the **influence of their L2 English** where these structural restrictions do not exist and hence the use of the *possessive –s* would not be wrong in the corresponding contexts. Indeed, this tendency of the wrong use of the *genitive –s*

correlates negatively with the proficiency level. That is, the **less proficient the student the higher the L2 influence**.

Our results correspond to the findings of previous studies on the role of the L2 in the acquisition of the L3 syntax that we presented in section 3.3. Additionally, our findings confirm our initial hypothesis that there will be an **L2 influence especially in initial stages**. In that aspect, we agree with the hypothesis put forth by Ringbom (1987) that claimed that background languages exert a high influence when the L3 proficiency level is low (see section 3.3).

The **L1 transfer cannot explain these data** since in our subjects' L1 the *genitive-s* feature does not exist. The fact that their L2 does not restrict the use of the *genitive –s* in those cases where the PR=proper name/kinship term explains the learners' tendency to incorrectly extend the use of the *genitive –s* without considering the *possessor*. Accordingly, we argue that the **Cumulative Enhancement Model (CEM**, Flynn et al. 2004) is not supported by our study. The CEM suggests that all previously learnt languages can be transferred in the L3 acquisition, when there is a structural overlap in two of the languages involved. In our case there is a structural overlap between the possession expression systems in L1 Spanish and L3 German on the one hand and L2 English and L3 German on the other hand. However, as we commented above based on the data displayed in Table 31, the use of the structural overlap between L2 English and L3 German, i.e. the *genitive –s*, is more extended than the structural overlap found in L1 Spanish and L3 German, i.e. the *PP*. To that we should add that the *PP* as a possession expression is actually common in all three languages. Even so, our subjects did not seem to prefer that structure, but rather the one that was shared only by the L2 and the L3.

The **Typological Primacy Model (TPM**, Rothman, 2010) states that according to economy of acquisition the most similar, typologically, language becomes the source for transfer,

independently of the language being an L1 or an L2. Based on our data we could argue in favour of this model since English is a Germanic language and seems to be the source of influence. However, we believe that such an assumption would be erroneous for two reasons. The first reason why a confirmation of the TPM would be inaccurate if based on our study's results is the special case that English presents as a language. English is indeed a Germanic language, but it is also one that has been extensively influenced by Romance languages especially in terms of its lexicon. Now, if we consider that in initial stages learners do not apply rules but rather assimilate chunks of language, as the declarative/procedural model suggests and our findings confirm (for more details see section 6.1), then we can argue that in these initial stages learners do not conceive the morphosyntactic similarities between two languages simply because they do not analyse the target language's morphological and syntactic properties. Additionally, in our case, similarities in language can be found between our subjects' L1 and L3 as well as between our subjects' L2 and L3. Yet only the L2 influence seems to be particularly high in initial stages. Second, and most importantly, we cannot accept the TPM as an explanatory model of our study's results because our data do not contrast various L1s and L2s which are, typologically speaking, similar or dissimilar to our L3 German. In that sense, our study lacks data in order to reach a clear conclusion in relation to the TPM model. Therefore, **we should not argue in favour or against the TPM model** on the grounds of this study.

The final hypothesis that has been proposed in relation to the L2 transfer in the acquisition of an L3 is known as the **L2 status hypothesis** (Williams and Hammarberg, 1998). The underpinning idea is that in L3 acquisition there is an L2 transfer due to the subject's "[…] desire to suppress L1 as being 'non-foreign' and to rely rather on an orientation towards a prior L2 as a strategy to approach the L3" (Hammarberg, 2001:36-37). We believe that this model can be proposed as the explanatory factor of our L3 German learners' extended use of

the genitive -s in non-obligatory for these possession structure contexts. Indeed, this model is the only that manages to account for our L3 German subjects' preference for the *genitive –s* even though the *PP* was common in all three languages (L1_Spanish, L2_English, L3_German).

7 Conclusions

After having presented our data and commented on the results of each of our three studies, we shall now introduce the conclusions that can be reached regarding the studies' questions and hypotheses presented in section 4. In this chapter we will also proceed with the distinction of three sections each of which corresponds to one of our project's studies. In this line of thought, section 7.1 deals with the questions and hypotheses of our first study that regards the morphemes' accuracy rates obtained by our L2 English subjects in our picture description task (corpus). Accordingly, section 7.2 presents our conclusions in relation to the use of the *possessive –s* as revealed by our data from both the learner corpus and the experimentally elicited learner language (sentence transformation task). Finally, in section 7.3 we regard the importance and the implications of the use of the various possessive structures in German by our L3 German learners.

7.1 Study I: MOS & Learner Corpus in L2 English

In relation to our first question, which referred to the possibility that the proficiency level can stand as an explanatory factor for the development of morphology in L2 English, our study provides evidence that the **proficiency level** is in general a covariate of **accuracy** order. In the same line, we can argue that the average order of accuracy appears after the very initial stages of proficiency. In this respect we agree with Muñoz's (2006) findings.

Our first hypothesis was that **accuracy** in use would not imply **acquisition** of the corresponding functor. In order to confirm that we needed to show that there are functors for which the accuracy rates are lower in higher proficiency levels. This

hypothesis was confirmed, since the grouping of our subjects on the basis of their proficiency level enabled us to successfully observe the U-shape pattern of development of the *past irregular* functor.

Our second hypothesis was that we would find a **similar pattern of morphemes accuracy order** for each proficiency level. This hypothesis was also confirmed as shown in Figure 10 (section 6.1). Additionally, our findings are similar to those suggested by Muñoz (2006) for those subjects that had a proficiency level beyond the elementary one. Our findings do not conform to Dulay and Burt's (1974) and Krashen's (1977) average orders, although they do display some similarities.

Our second question contemplated the possible **explanatory value** that Andersen's (1978) and Hawkins' (2001) pattern could have for the results of our study in relation to the **verb-related morphemes**. In that respect, our study confirms that Andersen's (1978) and Hawkins'(2001) pattern can explain the development of the verb-related morphemes studied in our project. In the same line, we can argue that our third hypothesis was confirmed. The slight difference in order found in our pre-intermediate group in comparison to Hawkins' (2001) pattern is explained by the SLA process and the function of the subjects' interlanguage (IL).

7.2 Study II: Genitive –s in L2 English

In relation to this project we formulated one question and two hypotheses. The question we posed in this respect was the **influence of the data elicitation** method on the results. Accordingly, we wished to check the validity of Krashen's claim that naturally occurred language will display the order of acquisition which will be different to the order of learning. The latter, following Krashen, should appear in tasks that trigger the subject's formal knowledge, such as the discrete-point grammar tests. Indeed, our findings yield a clear influence of the research instrument on the final results. They also confirm the claim that the accuracy orders found in

learners' language elicited naturally will be different to the corresponding orders found in experimentally elicited learner language. In that vein, our first hypothesis that the accuracy rates for the *possessive –s* obtained through our corpus task would be lower than the corresponding rates found in the data collected through our experiment (sentence transformation task) is confirmed. This, we believe, could imply that Krashen's distinction between acquisition and learning is supported.

Our second hypothesis regarding the relative **frequency** of the ***genitive –s*** by L2 English learners was that they would show a clear preference towards the use of the **analytic** form (PP) especially in initial stages, over the **synthetic form** (*genitive –s* morpheme). Our data disconfirmed this hypothesis, since they display high rates of use of the synthetic possession expression, that is, of the *possessive –s*. In the same vein, we find a correlation between our L2 English learners' relevant choice and the native speakers' preference as suggested by Rosenbach (2005). However, we would need more data in order to confirm such correspondence. On the other hand, our study's data provide some evidence that **L1 transfer cannot stand alone** as an explanatory model of the SLA/FLA processes.

7.3 Study III: Possessive structures in L3 German

In this part we formulated one hypothesis and one question. The latter relates to the **relative frequency** of the **inflectional and** the **periphrastic forms** of possession expression in both L2 English and L3 German. Based on the comparison of our data we were prompted to say that the **synthetic** form of possession expression is used more often by **L2 English** learners, whereas the **analytic** form of possession expression is used more often by **L3 German** learners. Indeed, what we found was that the analytic form was used more often than the synthetic forms in the corresponding OC.

Bearing on our hypothesis for this study, we should repeat our initial claim that the use of the possession expressions by L3 German learners would display an **L2 English influence** especially in initial stages. Indeed, this hypothesis was confirmed by our data. The highest frequency rates were achieved in relation to the use of the *genitive –s* but in wrong contexts, according to the structural restrictions of the German possessive system, which reveals an L2 English transfer, since in the English system of possession allows these restrictions do not exist. The lack of structural restrictions that is found in English was reproduced in the samples of our subjects' L3 German. The model that we consider most likely to account for this phenomenon is the so-called **L2 status hypothesis**. The Typological Primacy Model (TPM), although we believe that is not a valid explanatory model for our findings due to the reasons we exposed in section 6.3, cannot actually be rejected since in our study we have not included various and typologically distant background languages.

7.4 Limitations of the study

We hold that this study could be improved in various aspects. First, the participation of **more students**, especially from the elementary and upper-intermediate proficiency levels, would have allowed for a more precise description of the morpheme accuracy orders in each proficiency level. In the same vein it would have enabled us to compare our MOS results with previous studies that include data from younger learners.

Time limitations prevented us from considering additional information such as the subjects' **onset age** and the actual exposure to the target language. These are data that we have in our possession since relevant questions were included in our learner's profiles. However, we did not manage to take this information into consideration when we grouped our participants. Therefore we had to consider only the proficiency level when we separated our subjects into different groups.

Additionally, we would like to have tested our students' **proficiency level based on an oral test as well**. When we were correcting the written tests we came across some cases in which two students were put in the same level although one was just two points above the minimum required for that level, whereas the other was just two points below the minimum required for the next level. We therefore believe that an oral test would have clarified these edge cases.

Furthermore, our study would have been improved if we had included **extra means of data elicitation**. We believe that a recording of **spontaneous conversations** would have enabled us not only to make more accurate comparisons between our data and the findings of previous studies, but also to investigate the differences between completely naturally produced and clinically elicited language.

In relation to the frequency of use of the *genitive –s* in English we suggest that further study should be conducted including **all other types of possessors**. That would make possible a comparison between L2 English learners' use of the various possession forms and native speakers' relevant preferences. Furthermore, although we tried to limit the instrument's influence on our results, a different elicitation task that would result in more authentic language would be desirable, provided that it would ensure a sufficient number of OC for the *genitive –s*.

Finally, we would like to have been able to include more information on the **influence of various background languages** on the acquisition of a new foreign language. For this we should have collected data of various and typologically distant first and second/foreign languages and then compare the learners' results in the production of L3 language.

7.5 Avenues for future research

In this final section we would like to suggest some possible lines for future research according to the findings but also the limitations of the present studies.

In the same line, we would argue that future MO studies should be conducted taking into consideration additional information with regard to the **subjects' background**. In our study we have regarded the subjects' proficiency level, but we consider that it is possible and desirable to further analyse our data as to include the subjects' onset age, exposure to TL and other relevant information found in our learner profiles.

Additionally, we believe that, for a better understanding of the morpheme acquisition process, future research should include various **data elicitation instruments**, which we did in our study, but only in relation to one of the functors at issue, namely the *genitive –s*.

Regarding the study of the developmental pattern of individual morphemes, the *genitive –s* in our case, we understand that further research should be carried out including additional elicitation tasks designed for the production of the functor under examination. This, as we mentioned in the previous section, is a highly demanding activity since it should aim to the designing of an instrument that would ensure not only that the specific morpheme is used in sufficient occasions, but also that the learner language produced will be as natural as possible.

Perhaps a more realistic future project regarding the use of the various possessive forms by **L2 English** learners can be proposed, in which various **types of *possessors*** (e.g., animate vs. inanimate, etc.) would be included. This would allow the researcher to study L2 English learners' preferences and compare them to the corresponding ones shown by L1 English speakers. In our understanding this type of comparison would provide us with some enriching insights bearing on the relation between the L1 and the foreign language (FL) acquisition processes.

Finally, we consider that **further research** should be conducted in the direction of the **influence** that **background FL** may exert on the **acquisition of a new target language**. This field of research is relatively new and, although a number of relevant studies have already been conducted, we understand that there is still much to be discovered. Nevertheless, it is **of particular interest** for the SLA research since it can indeed provide us with useful information and thus shed light on the actual process of language acquisition. In line with this thought, we trust that the **introduction of learner corpora** in this type of research would provide the SLA researcher with important information regarding the various aspects of language (e.g. the strategies adopted bearing on the functional use of language, the acquisition of FL pragmatics, etc.), and hence enable us to obtain a more holistic view of foreign language acquisition.

8 References

Andersen, R.W. 1978. ‘An implicational model for second language research.’ *Language Learning*, 28: 221-282.

Archer, D., P. Rayson, A. Wilson and **A. McEnery** (eds.) 2003. *Proceedings of the Corpus Linguistics*. Lancaster: UCREL.

Bailey, N., C. Madden and **S. D. Krashen** 1974. 'Is there a natural sequence in adult second language learning?' *Language Learning,* 24: 235-243.

Bardel, C. and Falk, Y. 2007. ‘The role of the second language in the third language acquisition: the case of Germanic syntax.’ *Second Language Research,* 23:459-484.

Brown, R. 1973. *A First Language: The Early Stages*. Cambridge, Mass: Harvard University Press.

Brown, H. D., C. Yorio and R. Crymes (eds.) 1977. On TESOL '77: *Teaching and Learning English as a Second Language. Trends in Research and Practice*. Washington, DC: TESOL.

Brown, J.D. 1983. ‘An exploration of morpheme group interactions’ in Bailey, K., Long, M., and Peck, S.(eds.) *Second Language Acquisition Studies*. Rowley, MA: Newbury House.

Butters, R.R. 2001. ‘Grammatical Structure’ in Algeo J. (ed.) *English in North America*. Cambridge: Cambridge University Press. 325-339

Cook, V. 1993. *Linguistics and Second Language Acquisition*. New York: St. Martin's Press.

Corder, S.P. 1976. ‘The study of interlanguage’ in *Proceedings of the Fourth International Conference of Applied Linguistics Volume 3: Techniques in Applied Linguistics.* Oxford: Oxford University Press.

de Villiers, J. and P. de Villiers. 1973. ‘A cross-sectional study of the acquisition of grammatical morphemes in child speech.’ *Journal of Psycholinguistic Research,* 2: 267-278.

De Angelis, G. 2005b. ‘Interlanguage transfer of function words.’ *Language Learning*, 55: 379-414.

Dentler, S. 2000. Deutsch und English - das gibt immer Krieg [German and English – there is always a war]. In Dentler, S., Hufeisen, B. and Lindemann, B.(eds.) *Tertiär und Drittsprachen. Projekte und empirische Untersuchungen. [Tertiary Languages. Projects and empirical research.]* Tübingen: Stauffenburg Verlag: 77-97.

Díaz Negrillo, A. 2007. *A Fine-Grained Error Tagger for Learner Corpora*. Unpublished PhD Thesis. University of Jaen.

Dulay, H. C. and **M. K. Burt.** 1973. 'Should we teach children syntax?' *Language Learning,* 23: 245-258.

Dulay, H. C. and **M. K. Burt.** 1974b. 'Natural sequences in child second language acquisition.' *Language Learning,* 24: 37-53.

Eisenbeiß, S. et al. 2009. 'Learning to encode possession.' in McGregor, W.(ed.) *The expression of possession.* Berlin: Mouton de Gruyter. 143-212.

Ellis, R. 1987. 'Interlanguage variability in narrative discourse: Style shifting in the use of the past tense.' *Studies in Second Language Acquisition* 11: 305- 328.

Ellis, R. 1994. *The Study of Second Language Acquisition.* Oxford: Oxford University Press.

Ellis, R. and **G. Barkhuizen** 2005. *Analysing Learner Language*. Oxford: Oxford University Press.

Fathman, A. 1975. 'The relationship between age and second language productive ability.' *Language Learning*, 25: 245-253.

Flyn, S. 1987. A parameter setting model of L2 Acquisition: Experimental Studies in Anaphora. Dordrecht: Reidel.

Flyn, S. et al. 2004. 'The cumulative-enhancement model of language acquisition: comparing adults' and children's patterns of development first, second and third language acquisition of relative clauses.' *International Journal of Multilingualism* 1(1): 3-16.

Gass, S. 2001. 'Sentence matching: a re-examination.' *Second Language Research*, 17:421-441.

Goldschneider, J. and R. DeKeyser 2001. 'Explaining the "Natural Order of L2 Morpheme Acquisition" in English: A meta-

analysis of multiple determinants.' *Language Learning*, 51(1): 1-50.

Granger, S. 2004 'Computer learner corpus research: current status and future prospects.' in Connor and Upton (eds.), 123-145.

Granger, S. 2008. 'Learner corpora.' in Lüdeling and Kytö (eds.), 259-275.

Granger, S. 2009 'The contribution of learner corpora to second language acquisition and foreign language teaching.' in Aijmer, K. (ed.) *Corpora and Language Teaching.* Amsterdam: John Benjamins. 13-33.

Gilquin, G., & Gries, S. T. 2009 'Corpora and experimental methods: a state-of-the-art review'. *Corpus Linguistics and Linguistic Theory*, 5(1), 1–26.

Hammarberg, B. 2001. 'Roles of L1 and L2 in L3 production and acquisition.' In Cenoz, J., Hufeisen, B. and Jessner, U. (eds.) *Cross-linguistic Influence in Third Language Acquisition: Psycholinguistic Perspectives.* Clevedon: Multilingual Matters. 21-41.

Haspelmath, M., 2002. *Understanding Morphology*. Oxford: Oxford University Press.

Hawkins, J.A. 1994. *A performance theory of order and constituency.* Cambridge: Cambridge University Press.

Hawkins, R. 2001. *Second Language Syntax. A Generative Introduction.* Oxford: Blackwell.

Hawkins, R. and Lozano, C. 2006. 'Second Language Acquisition of Phonology, Morphology, and Syntax' in Brown, K.(ed.) *The Encyclopaedia of English Language and Linguistics (2nd edition).* London: Elsevier. 67-74.

Heine, B. 1997. *Possession: Cognitive Sources, Forces and Grammaticalization.* New York: Cambridge University Press.

Housen, A. 2002. 'A corpus-based study of the L2 acquisition of the English verb system' in Granger, S., Hung, J. and Petch-Tyson, S.(eds.) *Computer Learner Corpora, Second Language Acquisition and Foreign Language Teaching.* Amsterdam: John Benjamins. 77-116.

Kachru, Y. 2003. ‘On definite reference in world Englishes.’ *World Englishes*, 22(4): 497-510.

Klein, E. 1995. ‘Second versus third language acquisition: Is there a difference?’ *Language Learning*, 45: 419-465.

Krashen, S. 1977. 'Some issues relating to the monitor model.' in Brown, Yorio and Crymes (eds.), *TESOL’ 77*: 144-158. TESOL, Washington, DC.

Krashen, S. 1977. ‘The Monitor Model for adult second language performance.’ In: Burt, M., Dulay, H., and Finocchiaro, M. (eds). *Viewpoints on English as a Second Language.* New York: Regents Publishing.

Krashen, S. 1978. 'Is the “Natural Order” an artefact of the bilingual syntax measure?' *Language Learning,* 28(1): 187-191.

Krashen, S., Sferlazza, V., Feldman, L. and Fathman, A. 1976. ‘Adult performance on the SLOPE test: More evidence for a natural sequence in adult second language acquisition.’ *Language Learning*, 26: 145-151.

Kjarsgaard, M. 1979. *The order of English morpheme category acquisition by Vietnamese children.* Unpublished PhD Thesis. Arizona State University, Phoenix.

Kwon, E. Y. 2005. 'The natural order of morpheme acquisition: A historical survey and discussion of three putative determinants.' *Working Papers in TESOL Applied Linguistics,* 5/1: 1-20.

Larsen –Freeman, D.E. 1975. ‘The acquisition of grammatical morphemes by adult ESL students.’ *TESOL Quarterly,* 9:409-419

Larsen –Freeman, D.E. 1976. ‘An explanation for the morpheme acquisition order of second language learners.’ *Language Learning,* 26:125-134.

Larsen –Freeman, D.E., & Long, M.H. 1975. *An Introduction to second language acquisition research.* New York: Longman

Lightbown, P. 1983. ‘Exploring relationships between developmental and instructional sequences in L2 acquisition’ in Selinger H.B. and Long M.H.(eds.)*Classroom*

Oriented Research in Second Language Acquisition. Rowley, MA: Newbury House. 217-245.

Lozano, C. and Mendikoetxea, A. (forthcoming 2012). 'Learner corpora and second language acquisition: the design and collection of CEDEL2.' In Ballier, N., Díaz-Negrillo, A. and Thompson, P.(eds.) *Automatic Treatment and Analysis of Learner Corpus Data.* Amsterdam: John Benjamins.

Lozano, C., & Mendikoetxea, A. (forthcoming 2013). 'Corpus and experimental data: subjects in second language research'. In L. Degand & S. Granger (Eds.), *Corpora and Language in Use*. Louvain: Presses universitaires de Louvain.

Lüdeling, A. and **M. Kytö** (eds.) 2008. *Corpus Linguistics. An International Handbook. Vol. 1.* Berlin: Mouton de Gruyter.

Mackey, A. and **S. Gass** 2012. *Research Methods in Second Language Acquisition: A Practical Guide.* London: Wiley-Blackwell.

Makino, T. 1979. *English morpheme acquisition order of Japanese secondary school students.* Unpublished PhD Thesis. University of New Mexico, Albuquerque.

Matteini, S. 2007. *On the morphosyntactic acquisition of the German DP by Italian adult L2 learners: Possessive constructions, Noun placement and inflectional morphology on nominal modifiers.* Unpublished PhD Thesis. University of Firenze.

Matteini, S. 2009. 'L2 strategies of morphological variability in the morphosyntactic acquisition of German DPs' in *Actes du colloque AcquisiLyon 09*. Available at (http://www.ciscl.unisi.it/doc/doc_pub/proceedings_acquisilyon09.pdf).

Mayer, M. 1969. *Frog, Where are You?* New York: Dial Press.

Mendikoetxea, A., & Lozano, C. (forthcoming 2013). 'Conceptual and methodological interfaces in SLA research: triangulating corpus and experimental data in L2 Subject-Verb and Verb-Subject alternations.' (S. De Knop & F. Meunier, Eds.) *Corpus Linguistics and Linguistic Theory* [Special issue on Learner corpus research, cognitive linguistics and second language acquisition].

McEnery, A. M. 2003. 'Corpus linguistics' in Mitkov (ed.), 448-463.

McEnery, T. and **A. Hardie** 2012. *Corpus Linguistics: Method, Theory and Practice*.Cambridge: Cambridge University Press.

McEnery, A. M. and **A. Wilson** 2001. *Corpus Linguistics.* Edinburgh: Edinburgh University Press.

McEnery, A. M., R. Xiao and **Y. Tono** 2006. *Corpus-based Language Studies.* London:Routledge.

McGregor, W. 2009. 'Introduction' in McGregor, W.(ed.) *The expression of possession.* Berlin: Mouton de Gruyter. 1-13.

Muñoz, C. 2006. *Age and the Rate of Foreign Language Learning.* Clevedon, Buffalo: Multilingual Matters.

Norris, J. and Ortega, L. 2001. 'Does type of instruction make a difference? Substantive findings from a meta-analytic review' in Ellis, R.(ed.) *Form-Focused Instruction and Second Language Learning*. Malden, MA: Blackwell.

O'Donnell, M. 2009. 'The UAM Corpus Tool: Software for corpus annotation and exploration.' in Callejas et al. (eds.), 1433-1447.

Perkins, K. and Larsen-Freeman, D. 1975. 'The effect of formal language instruction on the order of morpheme acquisition.' *Language Learning*, 25(2): 237-243.

Pica, T. 1983. 'Adult acquisition of English as a second language under different conditions of exposure.' *Language Learning*, 33:465-497.

Pienemann, M. 1998. *Language Processing and Second Language Development: Processability Theory.* Amsterdam: John Benjamins.

Quirk, R. et al. 1985. *A Comprehensive Grammar of the English Language.* London:Longman.

Ringbom, H. 1987. *The Role of the First Language in Foreign Language Learning.* Clevedon: Multilingual Matters.

Rosenbach, A. 2007. 'Animacy versus weight as determinants of grammatical variation in English.' *Language* 81(3): 613-644.

Rothman, J. 2010. 'On the typological economy of suntactic transfer: word order and high/low attachment preference in relative clause interpretations in L3 Brazilian

Portuguese.' *International Review of Applied Linguistics in Language Teaching*, 48: 245-275.

Rothman, J and Abrelli-Amaro, J. 2010. 'What variables condition syntactic transfer? A look at the L3 initial state.' *Second Language Research,* 26:189-218.

Rosansky, E. 1976. 'Methods and morphemes in second language acquisition research.' *Language Learning,* 26(2): 409-425.

Seidlhofer, B. 2004. 'Research Perspectives on Teaching English as a lingua franca.' *Annual Review of Applied Linguistics,* 24: 209-239.

Sinclair, J. 1996. *EAGLES. Preliminary Recommendations on Corpus Typology.* Available from (http://www.ilc.cnr.it/EAGLES/corpustyp/corpustyp.html).

Stauble, A.M. 1984. 'A comparison of a Spanish-English and a Japanese-English second language continuum: Negation and verb morphology.' In R. Andersen (ed.), *Second languages: A cross-linguistic perspective.* Rowley, MA: Newbury House.

Taylor, J.R. 1996. *Possessives in English.* Oxford: Oxford University Press.

Tono, Y. 2000. 'A computer learner corpus based analysis of acquisition' in Burnard and McEnery (eds.), 123-133.

Tono, Y. 2002. *The Role of Learner Corpora in SLA Research and Foreign Language Teaching: The Multiple Comparison Approach.* Unpublished PhD Thesis.Lancaster University, UK.

Tono, Y. 2003. 'Learner corpora: design, development and applications' in Archer, Rayson, Wilson and McEnery (eds.), 800-809.

Tono, Y. 2005. *Corpus-based SLA Research: State-of-the-Art of Learner Corpus Studies*. Tokio:Kurosio Publishers.

Ullman, M. 2001. 'The neural basis of lexicon and grammar in first and second language: the declarative/procedural model.' *Bilingualism: Language and Cognition*, 4: 105-122.

Ullman, M. 2005. 'A cognitive neuroscience perspective on second language acquisition: declarative/procedural model.' In Sanz, C.(ed.) *Mind and Context in Adult Second Language Acquisition. Methods, Theory and Practice*. Washington, DC: Georgetown University Press. 141-178.

Van de Gratts, I. Corver, N. and R. Van Hout. 2000. ‘Conservation of Grammatical Knowledge: on the acquisition of possessive noun phrases by Turkish and Moroccan learners of Dutch’. *Linguistics,* 38(2): 221-314.

Wagner, J. 2005. ‘Acquisition of Grammatical Morphemes in English as a Foreign Language Learners’ Available at (http://www.ielanguages.com/documents/papers/Masters%20Seminar.pdf)

Willemse, P., Davidse, K. and Heyvaert, L. 2009. ‘English possessives as reference-point constructions and their function in the discourse.’ in McGregor, W.(ed.) *The expression of possession.* Berlin: Mouton de Gruyter. 13-50

Williams, S. and Hammarberg, B. 1998. ‘Language switches I L3 production: Implications of a polyglot speaking model.’ *Applied Linguistics*, 19:295-333.

Zobl, H. and J. Liceras 1994. ‘Functional categories and the acquisition orders.’ *Language Learning*, 44(1): 159-180.

9 Appendices

9.1 Learner Profile for L2 English and for L3 German

FILE I.D.

Universidad de Granada

[Last saved: 3-may-12]

Fecha:_____

INFORMACIÓN PERSONAL

■TUS INICIALES____ ■TU NICK_____ ■EDAD____ ■SEXO:☐Hombre ☐Mujer

■CURSO:☐1ºESO ☐2ºESO ☐3ºESO ☐4ºESO ☐1ºBachi. ☐2ºBachi.
☐PCPI ☐Grado administrativo ☐Otro_____

■INSTITUTO DONDE ESTÁS ESTUDIANDO:_____

INFORMACIÓN LINGÜÍSTICA

■Lengua materna: ☐español ☐otra (indicar):_____

■Lengua materna de tu padre: ☐español ☐otra (indicar):_____

■Lengua materna de tu madre: ☐español ☐otra (indicar):_____

■Lengua(s) que hablas en casa:☐español ☐otras (indicar):_____

■Edad a la que empezaste a aprender inglés _____

■¿Cuál crees tú que es tu nivel de inglés?

SPEAKING:	LISTENING:	READING:	WRITING:
☐Principiante bajo (A1)	☐Principiante bajo (A1)	☐Principiante bajo (A1)	☐Principiante bajo (A1)
☐Principiante alto (A2)	☐Principiante alto (A2)	☐Principiante alto (A2)	☐Principiante alto (A2)
☐Intermedio bajo (B1)	☐Intermedio bajo (B1)	☐Intermedio bajo (B1)	☐Intermedio bajo (B1)
☐Intermedio alto (B2)	☐Intermedio alto (B2)	☐Intermedio alto (B2)	☐Intermedio alto (B2)
☐Avanzadobajo (C1)	☐Avanzadobajo (C1)	☐Avanzadobajo (C1)	☐Avanzadobajo (C1)
☐Avanzado alto (C2)	☐Avanzado alto (C2)	☐Avanzado alto (C2)	☐Avanzado alto (C2)

■¿Estás aprendiendo otro idioma además del inglés? ☐Sí ☐No
Si tu respuesta es SÍ, ¿Cuál? _____

■Nota del curso pasado (a rellenar por el profesor de inglés): _____

EXPOSICIÓN LINGÜÍSTICA

■¿Has hecho alguna estancia en un país de habla inglesa? ☐Sí ☐No
Si tu respuesta es SÍ, ¿dónde? _____
¿cuándo? _____
¿cuántas semanas o meses estuviste allí? _____

■¿Has estudiado o estudias inglés fuera del instituto? ☐Sí ☐No
Si has contestado SÍ, ¿en qué año y cuánto tiempo (semanas/meses)?
_____

■¿Haces algo fuera del colegio relacionado con el inglés? (ej: ver películas en inglés, leer internet en inglés, etc.) ☐Sí ☐No

Especifica: _____
_____

■¿Estás en algún programa de bilingüismo en el Instituto? ☐Sí ☐No
Si tu respuesta es SÍ, ¿en qué curso empezaste el bilingüismo?_____
¿Qué asignaturas bilingües tienes? _____
¿Cuántas horas semanales de inglés tienes en esas asignaturas?_____

■**Consentimiento**: ☐ marca aquí para dar el consentimiento de que tus datos sean usados con fines de investigación sobre el aprendizaje del inglés. Esto NO es un examen. Todos tus datos serán anónimos y tratados confidencialmente. **Gracias por tu colaboración.**

FILE I.D.

Fecha:_____

[Last saved: 19-jul-12]

INFORMACIÓN PERSONAL

■TUS INICIALES____ ■TU NICK_____ ■EDAD____ ■SEXO:☐Hombre ☐Mujer

■CURSO:☐1ºESO ☐2ºESO ☐3ºESO ☐4ºESO ☐1ºBachi. ☐2ºBachi.
☐PCPI ☐Grado administrativo ☐Otro_____

■Centro DONDE ESTÁS ESTUDIANDO:_____

INFORMACIÓN LINGÜÍSTICA

■Lengua materna: ☐español ☐otra (indicar):_____

■Lengua materna de tu padre: ☐español ☐otra (indicar):_____

■Lengua materna de tu madre: ☐español ☐otra (indicar):_____

■Lengua(s) que hablas en casa:☐español ☐otras (indicar):_____

■Edad a la que empezaste a aprender alemán _____

■¿Cuál crees tú que es tu nivel de alemán?

Sprechen:	Hören:	Lesen:	Schreiben:
☐Principiante bajo (A1)	☐Principiante bajo (A1)	☐Principiante bajo (A1)	☐Principiante bajo (A1)
☐Principiante alto (A2)	☐Principiante alto (A2)	☐Principiante alto (A2)	☐Principiante alto (A2)
☐Intermedio bajo (B1)	☐Intermedio bajo (B1)	☐Intermedio bajo (B1)	☐Intermedio bajo (B1)
☐Intermedio alto (B2)	☐Intermedio alto (B2)	☐Intermedio alto (B2)	☐Intermedio alto (B2)
☐Avanzadobajo (C1)	☐Avanzadobajo (C1)	☐Avanzadobajo (C1)	☐Avanzadobajo (C1)
☐Avanzado alto (C2)	☐Avanzado alto (C2)	☐Avanzado alto (C2)	☐Avanzado alto (C2)

■¿Estás aprendiendo otro idioma además del alemán? ☐Sí ☐No
Si tu respuesta es SÍ, ¿Cuál? _____

■Nota del curso pasado (a rellenar por el profesor de alemán): ____

EXPOSICIÓN LINGÜÍSTICA

■¿Has hecho alguna estancia en un país de habla alemana? ☐Sí ☐No
Si tu respuesta es SÍ, ¿dónde? _____
¿cuándo? _____
¿cuántas semanas o meses estuviste allí? _____

■¿Has estudiado o estudias alemán fuera del instituto? ☐Sí ☐No
Si has contestado SÍ, ¿en qué año y cuánto tiempo (semanas/meses)?
Note that this question was included only in the learner profile given to the secondary education students of IES Padre Suarez
_____

■¿Haces algo fuera del colegio relacionado con el alemán? (ej: ver películas en alemán, leer internet en alemán, etc.) ☐Sí ☐No

Especifica: _____
_____

■**Consentimiento**: ☐ marca aquí para dar el consentimiento de que tus datos sean usados con fines de investigación sobre el aprendizaje del inglés. Esto NO es un examen. Todos tus datos serán anónimos y tratados confidencialmente. **Gracias por tu colaboración.**

9.2 Proficiency test: L2 English

English Unlimited Placement test

Written test

A1 to C1

- Choose the best answer for each question.
- Stop when the questions become too difficult.
- Spend no more than 40 minutes on the test.

1 Where _____ from?
I'm from Russia.
A you are B you C are you

2 We have _____ house in Moscow.
A any B a C an

3 I have two _____ , a boy and a girl.
A sons B daughters C children

4 I work in a _____ . I'm a doctor.
A hospital B hotel C supermarket

5 This is my brother. _____ name's Paul.
A Her B His C He's

6 _____ five people in my family.
A They are B There is C There are

7 I get up _____ 7 o'clock in the morning.
A for B at C in

8 I like apples, but I _____ bananas.
A don't like B like C do like

9 Excuse me, _____ speak French?
A do you B you do C you

10 How much are _____ shoes?
A this B these C that

11 Where are my glasses?
They're _____ the table.
A at B on C in

12 My sister _____ tennis very well.
A plays B play C playing

13 I usually go to work _____ train.
A on B with C by

14 I don't see my parents very often _____ they live in South Africa.
A so B but C because

15 Rosie stayed _____ home yesterday afternoon.
A in B at C to

16 Last night I _____ to the cinema.
A went B did go C was

17 The _____ is quite expensive but the food there is excellent.
A film B restaurant C book

18 Do you want to listen to music or _____ TV?
A see B look C watch

19 _____ were you at the weekend?
I was in Scotland.
A When B Where C What

20 _____ you have a good time at the party?
Yes, it was fun.
A Did B Were C Had

21 Are you _____ English teacher?
A Maria B Marias' C Maria's

22 Bob will meet _____ at the airport.
A us B we C our

23 I'm going to a concert tonight. _____ you like to come?
A Do B Are C Would

24 _____ use your dictionary?
Sure. Here you are.
A Could I B Could you C Do I

25 I like this apartment but the _____ is too expensive for me.
A money B rent C cost

26 Excuse me, how do I _____ to the bus station?
A come B get C arrive

27 Do you sell stamps?
Yes, we do. How _____ do you want?
A any B many C much

28 Sorry I'm so late.
That's _____ .
A OK B great C right

29 I'd like _____ milk in my coffee, please.
A some B any C a

30 _____ a bus stop near my flat.
A It's B Here's C There's

31 Is this a good time to talk?
Sorry, no. I _____ dinner.
A cook B am cooking C cooking

32 I think cycling is more dangerous _____ driving.
A as B like C than

33 We _____ going to the theatre next Saturday.
A will B do C are

34 _____ meet for coffee some time soon.
A Let's B Do you C Shall they

35 Kamal has got a holiday home near _____ sea.
A a B the C some

36 If you've got a headache, you _____ go home.
A should B did C had

37 _____ ever been to New York?
A Have you B Are you C Did you

38 I only get about five hours' sleep a night.
That's not _____ .
A enough B lot C too much

39 Did Amina finish the report?
No. She _____ it tomorrow.
A finishes B is going to finish C finished

40 Paula _____ loves working with children.
A very B really C much

41 Is Ottawa the capital of Canada?
I think _____ .
A is B yes C so D right

42 We never _____ a television when I was a child.
A have had B hadn't C had D didn't have

43 We paid the restaurant bill _____ credit card.
A to B with C on D by

44 The last time I _____ Joanna was in Paris.
A have seen B saw C see D was seeing

45 If you _____ money from a friend, you should always pay it back promptly.
A borrow B earn C spend D lend

46 Can I make myself a cup of coffee?
Of course. You _____ to ask.
A haven't B mustn't C needn't D don't have

47 I _____ a lot of sport in my free time.
A do B practise C make D exercise

48 _____ anywhere interesting recently?
A Do you go B Have you been C Are you going D Will you go

49 It's Walter's birthday on Friday. He _____ be 30, I think.
A should B can C will D shall

50 Learning the piano isn't as difficult _____ learning the violin.
A like B so C than D as

51 If the weather _____ bad tomorrow, we can go to a museum.
A will be B was C is D would be

52 About a billion cans of Coca-Cola _____ drunk around the world every day.
A is B are C was D were

53 My mum's not very well.
Oh, _____ .
A it doesn't matter B I do apologise C sorry to hear that D not bad, thanks.

54 Hans isn't here. He _____ to see his grandmother. He'll be back tomorrow.
A has gone B had been C has been D had gone

55 Would you mind changing my appointment? _____ time on Friday is fine.
A Next B All the C Every D Any

56 When I was a child, I _____ climb the wall and jump into our neighbours' garden.
A would B did C have D used

57 Have you finished _____ the wall yet?
A paint B to paint C painting D painted

58 Can you help me? I've tried _____ hotel in the city and can't find a room.
A many B any C every D all

59 Lena used to find work boring _____ she became a nurse.
A unless B until C if D since

60 If I _____ closer to my office, I could walk to work.
A lived B would live C had lived D live

61 I _____ outside the cinema when suddenly a police car arrived.
A stood B was standing C have stood D am standing

62 Shall we go to *The Riceboat* for dinner?
It _____ be fully booked. They're sometimes busy on a Monday.
A will B may C can D must

63 We've _____ come back from a trip to India. It was amazing.
A already B yet C just D only

64 I've got to be at work in five minutes.
Don't worry, I _____ you a lift if you want.
A give B am giving C 'll give D 'm going to give

65 My doctor advised me _____ more exercise.
A take B taking C having taken D to take

66 I couldn't _____ up with the noise in the city, so we moved to the countryside.
A put B live C set D take

67 There's no name on this dictionary.
It _____ be mine then. Mine's got my name on the front.
A might not B mustn't C won't D can't

68 Julia _____ married since she was 20.
A is B was C has been D is being

69 Don't worry if I _____ late tonight. I'm going to the gym after work.
A am B will be C would be D was

70 I've got a terrible headache, and it won't go away.
Have you tried _____ some aspirin?
A to take B take C took D taking

71 Boxing is a sport _____ requires a lot of speed and fitness.
A it B that C what D where

72 Jon _____ working on this project for a couple of months so he hasn't made much progress yet.
A is only B has only been C was only D had only been

73 I was wondering _____ I could ask you some questions.
Sure, go ahead.
A what B if C that D how

74 What clothes should I pack for a trip to Boston?
Well, it depends _____ the time of year that you go.
A on B with C up D to

75 I've finished this salad and I'm still hungry. I _____ ordered something more filling.
A must have B would have C should have D may have

76 Do you ever ask your neighbours to do favours _____ you?
A for B to C with D about

77 Some married couples seem to get more _____ over time.
A alike B same C like D equal

78 I don't know how much this card costs. The price label's _____ off.
A gone B taken C done D come

79 Ben got the job because he _____ a very good impression at his interview.
A made B did C put D took

80 Salsa music always _____ me of my trip to Cuba.
A remembers B realises C recognises D reminds

81 I _____ to be picking Tom up at the station but I've lost my keys.
A am supposed B am requested C am intended D am obliged

82 How about going to *Colours* nightclub?
There's no _____ I'm going there. It's awful!
A hope B way C time D opportunity

83 By the age of 18, I _____ not to go to university.
A had decided B decided C have decided D was deciding

84 I'm afraid your car _____ repaired before next week.
A hasn't been B wasn't C wouldn't be D can't be

85 The amount of organically grown food on sale has _____ enormously in recent years.
A raised B lifted C increased D built

86 Can you believe it? A woman has been _____ for hacking into the computer of her online virtual husband.
A accused B suspended C arrested D suspected

87 You may borrow my laptop _____ you promise to look after it.
A unless B in case C as long as D although

88 It's a huge painting. It _____ taken ages to complete.
A must have B can't have C should have D won't have

89 Pierre tends to put _____ dealing with problems, rather than dealing with them immediately.
A down B off C over D away

90 If the taxi hadn't stopped for us, we _____ standing in the rain.
A were still B would still be C are still D will still be

91 My mother's Italian, so _____ the language has been quite easy for me.
A to learn B learn C having learned D learning

92 _____ I had the talent, I still wouldn't want to be a movie star.
A In case B Even if C Provided that D However much

93 The factory workers threatened _____ on strike if they didn't get a pay rise.
A going B to go C that they go D to have gone

94 I was about to go to sleep when it _____ to me where the missing keys might be.
A remembered B happened C appeared D occurred

95 There's going to be a new department at work. They've asked me to _____ it up.
A take B set C put D bring

96 If the film is a _____ success, the director will get most of the credit.
A big B high C large D good

97 By the end of today's seminar I will _____ to each of you individually.
A speak B have spoken C be speaking D have been speaking

98 This is a photo of my little sister _____ ice cream on the beach.
A eat B eating C was eating D having eaten

99 Our students take their responsibilities very _____ .
A considerably B thoroughly C seriously D strongly

100 Pia was _____ delighted with the birthday present.
A very B completely C fairly D absolutely

English Unlimited Placement test

Teacher's guide

This placement test is designed to help teachers decide at which level of *English Unlimited* new students should be placed.

The test contains:

- 120 written multiple-choice questions, 20 at each level from Starter to Advanced (covering CEF levels A1 to C1). The questions focus on the language taught across the six levels of *English Unlimited*.
- an oral placement test, designed to be used in conjunction with the written test. The questions in this test are linked to the goals in the *English Unlimited* coursebooks, which in turn are based on the CEF 'can-do' statements. The oral test is designed to be flexible. This should allow teachers to get an overview of students' speaking ability in order to 'fine-tune' their assessment of the correct level at which to place students.
- keys and tables for interpreting scores.

Written test procedure

1 Ask the student to start at the beginning of the written test and stop when the questions become too difficult. Allow 30–40 minutes for the test.

2 Mark the test and use the table below to place the student provisionally at one of the six levels of *English Unlimited*.

3 Use the oral placement test to confirm or adjust this placement (see Oral test procedure below).

	Starter	Elementary	Pre-intermediate	Intermediate	Upper Intermediate	Advanced
Written test score	0–15	16–35	36–55	56–75	76–95	96+

9.3 Proficiency Level Test: L3 German

Prueba de Nivel de Alemán **Nivel Básico 1 A1**
Grammatik- und Wortschatztest

1.- Tobias kommt _____ Berlin.
a) von b) aus c) in
2.- Am Morgen fahre ich mit _____ Bus.
a) der b) den c) dem
3.- Das Kind _____ Tobias.
a) heiβt b) bin c)heiβe
4.- Mein Mann _____ in Berlin.
a) arbeite b) wohnst c) arbeitet
5.- Wir reisen oft _____ Bahn.
a) bei b) mit c) mit der
6. Morgens trinke ich _____ Kaffee.
a) keinen b) nicht c) kein
7.- Tobias _____ immer sehr schnell.
a) sprecht b) spreche c) spricht
8.- Sonntags _____ ich um 9.30 _____.
a) stehe/- b) aufstehe/- c) stehe/auf
9.- Jeden Tag haben wir vier _____Unterricht.
a) Uhren b) Stunden c) Stunde
10.- Auf dem Tisch gibt es drei _____.
a) Buch b) Bücher c) Heft
11.- Am Wochenende _____.
a) ich nicht lerne b) lerne ich nicht c) nicht ich lerne
12.- Heute habe ich _____.
a) nicht Zeit b) Zeit nicht c) keine Zeit
13.- _____ du nach Hause?
a) Kommt b) Kommst c) Kommen
14.- _____ Sie bitte!
a) Wartet b) Warten c) Warte
15. Ich verstehe dich nicht. _____bitte lauter!
a) Sprichst du b) Sprechen Sie c) Sprich
16.- Lekeitio liegt _____ Meer.
a) an b) am c) ans
17.- _____ August fahren wir nach Deutschland.
a) Am b) Im c) –
18.- Ich bin _____1983 geboren.
a) in b) am c) –
19.- Das Bild hängt an _____ Wand.
a) die b) das c) der
20.- Ich habe _____ Brief bekommen.
a) dein b) deinen c) deiner
21.- Maria hat schon _____ Prüfung gemacht.
a) seine b) ihre c) deine
22.- Ich _____ in die Stadt gelaufen.
a) habe b) bin c) ist
23.- Tobias hat Pasta _____.
a) kochen b) gekocht c) kocht
24.- Wir haben unsere Bücher _____.
a) vergisst b) vergesst c) vergessen
25.- Ich muss zum Arzt _____.
a) gehen b) gegangen c) geht

26.- Hier _____ man nicht rauchen. Es ist verboten.
a) muss b) kann c) darf
27.- Komm zu mir, _____.
a) wenn du willst! b) wenn willst du! c) wenn du will!

***Prueba de Nivel de Alemán* Nivel Básico 2 A2**
Grammatik- und Wortschatztest

1.- Hans _____ eigentlich Ingenieur werden. Aber er ist Mechaniker geworden.
a) möchte b) wollte c) will
2.- Die Chefin hat mit _____ Sekretärin gesprochen.
a) ihrer b) ihrem c) seiner
3.-Ich möchte in einem Land leben, _____ schöne Landschaften hat.
a) der b) den c) das
4.- Deine Mutter hat angerufen. Du _____ heute Abend früh nach Hause kommen.
a) magst b) sollst c) willst
5.- Heute Morgen haben wir _____ von unseren Kollegen in der Firma verabschiedet.
a) uns b) einander c) es
6.- Beate ist ein bisschen größer _____ ihre Zwillingsschwester.
a) ob b) wie c) als
7.- Peter hat die Prüfung bestanden. _____ feiern wir heute eine Party bei ihm.
a) Denn b) Deshalb c) Obwohl
8.- Die Frau mit dem _____ Kleid ist die Schwester von Frank.
a) roten b) rotem c) rotes
9.- _____ ich mein Studium abgeschlossen habe, war ich schon 30.
a) Wenn b) Wann c) Als
10.- Du, ich brauche deinen Wagen. Würdest du _____ leihen?.
a) er mir b) mir es c) ihn mir
11.- In _____ Monat fahre ich nach Berlin.
a) ein b) einen c) einem
12.- _____ ich keine Zigaretten mehr rauche, atme ich viel besser.
a) Als b) Nach c) Seit
13.- Sie haben _____ beim Tanzen kennen gelernt.
a) sich b) ihr c) ihnen
14.- Ich habe meinen Laptop mitgebracht, _____ du den Text lesen kannst.
a) deshalb b) denn c) damit
15.- Mein Vater interessiert sich sehr _____ Politik.
a) an b) für c) von
16.- Ich muss mit dem Arzt einen Termin _____ .
a) vereinbaren b) verbinden c) verbleiben
17.- Ich habe mich sehr _____ deine Einladung gefreut.
a) vor b) mit c) über
18.- Ich weiß nicht, _____ Peter meine E-mail bekommen hat.
a) wenn b) dass c) ob

19.- Ich habe meine Brille auf den Tisch _____, und jetzt ist sie nicht mehr da.
a) gelegen b) gelegt c) gelogen
20.- _____ Wochenende haben wir Zeit uns zu erholen.
a) Am b) An c) An die
21.- Der Freund, _____ ich heute nach Bochum reisen wollte, hat einen Unfall gehabt.
a) den b) von dem c) mit dem
22.- _____ ich mit der Arbeit fertig bin, gehe ich zu meinen Freunden.
a) Als b) Wenn c) Wann
23.- _____ weiβ, warum er uns so belogen hat.
a) Alle b) Keiner c) Jeden
24.- Gestern _____ ich keine billige Eintrittskarte für die Oper besorgen.
a) konnte b) mochte c) sollte
25.- Die Lehrerin hat das Kind gelobt, _____ es hat eine sehr schöne Geschichte geschrieben.
a) damit b) deshalb c)denn
26.- Das Wochenende haben wir _____ dem Land verbracht.
a) auf b) in c) an
27.- Ist _____ in Ordnung? Du siehst müde aus.
a) alle b) alles c)etwas
28.- Das Auto _____ vom Mechaniker repariert.
a) hat b) bin c) wird
29.- _____ Sie mir bitte sagen, wo ich meinen Laptop anschließen kann?
a) Konnten b) Könnten c) Hätten

Prueba de Nivel de Alemán **Nivel Intermedio 1 B1.1**
Grammatik- und Wortschatztest

1. - Karl, ……. Freundin im siebten Monat schwanger ist, hat gestern seine Arbeit verloren.
   a) deren b) seine c) dessen
2. - „Ich bin sehr ……. deiner Meinung interessiert.“
   a) an b) für c) von
3. - Eine Woche ……. er seinen Führerschein gemacht hatte, hatte er bereits seinen ersten Unfall.
   a) seitdem b) nachdem c) seit
4. - Sie versucht es immer wieder auf ……. Weise, aber dieses Mal wird es nicht funktionieren.
   a) ihrer b) derselben c) dieselbe
5. - „Lass die Tasche ruhig liegen. Du ……. sie nicht wegzuräumen.“
   a) brauchst b) musst c) verstehst
6. - „Kannst du mir mal sagen, ……. du Angst hast?“
   a) vor wen b) wovor c) was
7. - „Ich bin wütend ……. Klaus, weil er nicht mit mir ausgehen will.“
   a) wegen b) für c) auf
8. - „……. ich enttäuscht von ihm bin, bin ich doch seine Freundin.“
   a) Trotzdem b) Deswegen c) Obwohl
9. - ……. einer Arbeit im väterlichen Betrieb wählte er einen Lehrberuf bei der Stadtverwaltung.

a) An b) Statt c) In

10. - „Wenn sie noch nicht hier sind, dann sind sie bestimmt wieder den ……. Weg gefahren.“
a) weitesten b) größten c) stundenlangen

11. - „Er ist wieder einmal zu spät gekommen, ……. wir ohne ihn anfangen mussten.“
a) warum b) zumal c) so dass

12. - „Das ist der Kollege, ……. ich dir erzählt habe.“
a) von dem b) über den c) wovon

13. - Letzten Sommer in Hamburg hat María ……. Deutsch gesprochen.“
a) viel b) viele c) mehrere

14. - ……. länger man übt, ……. besser klappt es dann.
a) Je / desto b) Zwar / aber c) Sowohl / als auch

15. - Es ist Herbst, da verlieren die Bäume ihre ..... .
a) Äste b) Büsche c) Blätter

16. - ..... bezahlen 3 Euro Eintritt, Kinder unter 16 Jahren die Hälfte.
a) Eltern b) Große c) Erwachsene

17. - Die Sekretärin klebte den Brief zu und warf ihn in ..... .
a) die Mailbox b) den Briefkasten c) die Post

18. - Mein Computer hat eine Maus, einen Monitor und einen ..... .
a) Schreiber b) Laser c) Drucker

19. - Zum Kaffee essen die Müllers gerne ein Stück ..... .
a) Brötchen b) Teigware c) Kuchen

20. - “Ich war schon in Lateinamerika und Asien.” - “Dann bist du aber schon viel..... .“
a) gefahren b) gegangen c) gereist

21. - Die Brücke trägt nur 5 Tonnen. ..... dürfen sie nicht benutzen.
a) Fahrräder b) Fußgänger c) LKWs

22. - Diese Informationen möchte ich auf einer CD ..... .
a) lagern b) speichern c) liegen

23. - Den Sommer möchten wir bei unseren ..... in London verbringen.
a) Bekanntem b) Bekannten c) Bekannte

24. - Der Postbote konnte den Brief nicht zustellen, weil Herr Braun die ..... falsch notiert hatte.
a) Bankleitzahl b) Postleitzahl c) Postnummer

25. - In der Eile hatte Peter ganz vergessen, ... einzupacken. Jetzt musste er so ins Bett gehen.
a) den Schlafanzug b) das Schlafhemd c) die Nachthose

26. - Seit ich Sport treibe, funktioniert mein ..... besser.
a) Blutlauf b) Blutsystem c) Kreislauf

27. - Um endlich eine nette Partnerin zu finden, hat Klaus in der Sektion “Lonely Hearts” seiner Lokalzeitung ..... aufgegeben.
a) eine Werbung b) eine Anzeige c) einen Hinweis

28. - Petra möchte ihrer Mutter beim Tischdecken helfen und hat schon einmal die ..... aus der Schublade geholt.
a) Dosenöffner b) Bestecke c) Kochlöffel

Prueba de Nivel de Alemán **Nivel Intermedio 2 B1.2**

Grammatik- und Wortschatztest

1- Sich _____ das Wetter zu ärgern, hat überhaupt keinen Sinn.
a) auf b) an c) über
2- Er _____ das Essen seiner Frau mit dem seiner Mutter.
a) vergeht b) vergibt c) vergleicht
3- Ich habe _____ in dieser Sache getäuscht.
a) mich b) mir c) es
4- _____ eines Maschinenschadens kam die U-Bahn heute Morgen verspätet an.
a) Trotz b) Während c) Wegen
5- Es ist verboten einen Wagen _____ .
a) zu überholen b) überholen c) überzuholen
6- _____ sie den Chef sprechen konnte, war schon die Kündigung eingetroffen.
a) Bevor b) Wenn c) Nachdem
7- Die Arbeitslosigkeit müsste _____ .
a) bekämpfen b) bekämpft werden c) bekämpft worden
8- Nur wenige Menschen haben einen so guten Geruchssinn, _____ sie vermuten.
a) wie b) als c) da
9- . _____ sie den ganzen Tag arbeitet, hat sie immer noch Zeit für ihre alte Tante.
a) Dennoch b) Deshalb c) Obwohl
10- Sie dürfen nicht alles durch eine _____ Brille sehen.
a) rosa b) rote c) schwarze
11-Die _____ Zeitung berichtet von einer ganz aktuellen Entwicklung.
a) heutliche b) heutige c) heuzutage
12- Wenn Sie einen _____ suchen, rufen Sie uns umgehend an.
a) Beschäftigung b) Nebenjob c) Stelle
13- Auch mit deiner Hilfe hätte ich nichts _____ .
a) gemacht können b) machen gekonnt c) machen können.
14- Der Opersänger lernt Deutsch, _____ er interessiert sich für deutsche Komponisten.
a) denn b) da c) darum
15- Wir müssen uns _____ Klima anpassen.
a) an das b) an den c) am
16- Im verlassenen Dorf sind Häuser _____ .
a) vertragen b) verfallen c) verzogen
17- Die Studentin hat ein Stipendium _____, um in Deutschland ihre Doktorarbeit schreiben zu können.
a) beantragt b) beworben c) bestellt
18- Die Studenten, _____ die Prüfung gut gelungen ist, sind zufrieden.
a) denen b) die c) dessen
19- _____ diesem Wetter bleiben wir lieber zu Hause.

a) Mit b) Während c) Bei
20- Ein Angestellter, der nicht pünktlich ist, _____.
a) werde entlassen b) wäre entlassen c) wird entlassen
21- Die _____ des Fußballspiels beginnt um 19:00 Uhr.
a) Niederlassung b) Übertragung c) Veröffentlichung
22- Hätte ich _____ eine größere Wohnung!
a) ja b) mal c) bloß
23- Das Buch soll ein _____ Publikum erreichen können.
a) weites b) breites c) braves
24- _____ der Computer kaputt ging, hatte ich nur die Hälfte meiner Arbeit gespeichert.
a) Wenn b) Wann c) Als
25- Er konnte eine vom Lehrer _____ Frage nicht beantworten.
a) gestellte b) gestellten c) gestellter
26- _____ ihrer Mühe hat sie keine guten Ergebnisse bekommen.
a) Infolge b) Trotz c) Dank
27-Die Teilnahme _____ der Demo war gefährlich.
a) auf b) in c) an
28- Er ist eine sehr _____ Person, die immer bereit zu helfen ist .
a)hilfsbedürftige b) hilflose c) hilfsbereite

9.4 Elicitation of Learner Language: Picture Composition Task

FROG WHERE ARE YOU?

Glossary: Dog (perro), frog (rana), boy (niño), bed (cama), vase (vaso), floor (suelo), look at (mirar a), smell (oler), day (día), night (noche), sleep (dormir), escape (escapar), worried (preocupado), look for (buscar), shout (gritar), forest (bosque), bee (abeja), rock (roca), hold (sostener), branch (rama) deer (ciervo), drop (caer), push (empujar), fall (caer), river (río), water (agua), trunk (tronco), find (encontrar), family (familia), leave (dejar), hand (mano) and wave goodbye (decir adios).

1

2

Pictures reproduced by Mayer (1969), with permission of the author/artist and publisher.

9.5 Experimental Elicitation of Learner Language: Sentence Transformation Task: English & German

Rewrite the sentences below replacing the underlined words with the words in brackets.

A. English:

1. This is his car. (my father)
2. Her sister fell off the swing. (Maria)
3. Their shoes are in the locker. (boys)
4. Their toys were left out in the rain. (children)
5. Her computer was broken. (woman)
6. His truck was dirty. (man)

B. German:

1. Das ist sein Auto. (mein Vater)
2. Ihre Schwester fiel von der Schaukel. (Maria)
3. Ihre Schuhe sind im Schrank. (Jungen)
4. Ihre Spielzeuge blieben im Regen draußen liegen. (Kinder)
5. Ihr Computer/Rechner war kaputt. (Frau)
6. Sein LKW war schmutzig/dreckig. (Mann)

9.6 Transcription Template

```
FILE_NAME:
LEVEL:
INITIALS:
NICK:
AGE:
SEX:
COURSE:
SCHOOL:
L1:
FATHERS_L1:
MOTHERS_L2:
LANGUAGE_AT_HOME:
AGE_EXPOSURE:
SPEAKING_SELF:
LISTENING_SELF:
READING_SELF:
WRITING_SELF:
OTHER_LANGUAGE:
WHICH_LANG:
ENGLISH_MARK_LAST_YEAR:
STAY_ABROAD:
WHERE_STAY?:
WHEN_STAY?:
MONTHS_STAY:
EXTRAMURAL_INSTRUCTION:
WHEN_EXTRAMURAL:
MONTHS_EXTRAMURAL:
ADDITIONAL_EXTRAMURAL:
WHICH_ADDITIONAL_EXTRAMURAL:
BILINGUAL_PROGRAM:
WHEN_BILINGUAL_PROGRAM:
BILINGUAL_SUBJECTS:
BILINGUAL_EXTRA_HOURS_SUBJECTS:
COMPOSITION:
```

9.7 Examples

9.7.1 Learner Corpus Transcribed File

```
FILE_NAME: B2_4ESO_15_PSR_EAT_IAC.txt
LEVEL:     Upper-Intermediate
INITIALS:  IAC
NICK: Iván
AGE:  15
SEX:  Hombre
COURSE:    4ºESO
SCHOOL:    IES Pedro Soto de Rojas
L1:  Español
FATHERS_L1:     Español
MOTHERS_L2:     Español
LANGUAGE_AT_HOME:     Español
AGE_EXPOSURE:   6
SPEAKING_SELF:  B2
LISTENING_SELF: B1
READING_SELF:   B1
WRITING_SELF:   B1
```

OTHER_LANGUAGE: yes
WHICH_LANG: Francés
ENGLISH_MARK_LAST_YEAR:
STAY_ABROAD: yes
WHERE_STAY?: Londres
WHEN_STAY?: 2012
MONTHS_STAY: 5 días
EXTRAMURAL_INSTRUCTION: no
WHEN_EXTRAMURAL:
MONTHS_EXTRAMURAL:
ADDITIONAL_EXTRAMURAL: sí
WHICH_ADDITIONAL_EXTRAMURAL: series en inglés, partidos de basket en inglés, hablar en inglés con amigos de otros países por internet
BILINGUAL_PROGRAM: no
WHEN_BILINGUAL_PROGRAM:
BILINGUAL_SUBJECTS:
BILINGUAL_EXTRA_HOURS_SUBJECTS:
COMPOSITION: One night a boy was $_RWU_¿ having some fun with his pets in his bedroom. These pets were a frog and a dog. The frog was in a bottle and the dog was looking at the frog while standing next to the boy. When the boy feel asleep the frog got out of the bottle and went out of the bedroom. The frog $_RWR_e scaped during the night so the boy didn't feel anything and he continued sleeping even the dog which was sleeping with him knew that the frog was scaping. Early in the morning, the boy woke up and saw an empty bottle in front of his bed, then he understood that the frog have scaped and he needed to look for it if he didn't want to lose it. Next, the boy and the dog ran to the forest $_RWU_¿ and tried to find the frog, both of them love it so much so they were ready to do everything to find it. They were looking for the frog in all the places of the forest, the boy started searching at $_RWR_the some trees $_RWU_¿ while the dog was helping him looking at other different trees. The boy $_RWR_went climbed a little rock and called his frog as loud as he could. After that, he saw that he was not $_RWR_keeping $_RWR_a $_RWR_tree

holding a branch to maintain his equilibrium, he was holding a deer! Then, the deer pushed him down the mountain and the boy dropped with his dog. The boy and the dog fell $_RWU_¿ in the water, but they didn't suffer any injury or hurt. They returned to loonk for the frog pulling a trunk that they found next to them. When they got on the trunk, they could see a family of little frogs where the dog were looking for was. The boy caught his frog and then, $_RWU_¿ all of them $_RWU_¿ went back home. The boy and his dog felt very happy and they wave goodbye to the frog family.

9.7.2 Sentence Transformation Transcribed File English

FILE_NAME: B1_1BACH_17_PSR_EAT_MRV.txt
LEVEL: INTERMEDIATE(70)
INITIALS: MRV
NICK:
AGE: 17
SEX: MUJER
COURSE: 1ºBACHILLERATO
SCHOOL: IES Pedro Soto de Rojas
L1: Español
FATHERS_L1: Español
MOTHERS_L2: Español
LANGUAGE_AT_HOME: Español
AGE_EXPOSURE: 7
SPEAKING_SELF: A2
LISTENING_SELF: B1
READING_SELF: A2
WRITING_SELF: A2
OTHER_LANGUAGE: Yes
WHICH_LANG: Francés
ENGLISH_MARK_LAST_YEAR:
STAY_ABROAD: INGLATERRA
WHERE_STAY?: 2011
WHEN_STAY?: DOS SEMANAS
MONTHS_STAY:

EXTRAMURAL_INSTRUCTION: SÍ
WHEN_EXTRAMURAL:DURANTE LOS ÚLTIMOS AÑOS DE PRIMARIA
MONTHS_EXTRAMURAL:
ADDITIONAL_EXTRAMURAL: SÍ
WHICH_ADDITIONAL_EXTRAMURAL: ACTIVIDADES EXTRAESCOLARES EN INGLÉS
BILINGUAL_PROGRAM: no
WHEN_BILINGUAL_PROGRAM:
BILINGUAL_SUBJECTS:
BILINGUAL_EXTRA_HOURS_SUBJECTS:
COMPOSITION: 1. This is his car (my father): This is my father's car.
2. Her sister fell off the swing (Maria): Maria's sister fell off the swing
3. Their shoes are in the locker (boys): Boys' shoes are in the locker
4. Their toys were left out in the rain (children): Children's toys were left out in the rain.
5. Her computer was broken (woman): Woman's computer was broken.
6. His truck was dirty (man): Man's truck was dirty.

9.7.3 Sentence Transformation Transcribed File German

FILE_NAME: A2.2_2BACH_17_PS_EAT_NVCG.txt
LEVEL: A2.2
INITIALS: NVCG
NICK:
AGE: 17
SEX: MUJER
COURSE: 2ºBACH
SCHOOL: IES PADRE SUAREZ
L1: ESPAÑOL
FATHERS_L1: ESPAÑOL
MOTHERS_L2: ESPAÑOL
LANGUAGE_AT_HOME: ESPAÑOL

AGE_EXPOSURE: 12

SPEAKING_SELF: B2

LISTENING_SELF: B2

READING_SELF: C1

WRITING_SELF: C1

OTHER_LANGUAGE: SÍ

WHICH_LANG: INGLÉS

STAY_ABROAD: NO

WHERE_STAY?:

WHEN_STAY?:

MONTHS_STAY:

EXTRAMURAL_INSTRUCTION: NO

WHEN_EXTRAMURAL:

MONTHS_EXTRAMURAL:

ADDITIONAL_EXTRAMURAL: NO

WHICH_ADDITIONAL_EXTRAMURAL:

BILINGUAL_PROGRAM: SÍ

WHEN_BILINGUAL_PROGRAM: 1ºESO

BILINGUAL_SUBJECTS: HISTORIA, PROYECTOS

BILINGUAL_EXTRA_HOURS_SUBJECTS: 2HORAS POR SEMANA

COMPOSITION: 1. Das ist sein auto (mein Vater): Das ist das Auto meinem Vater

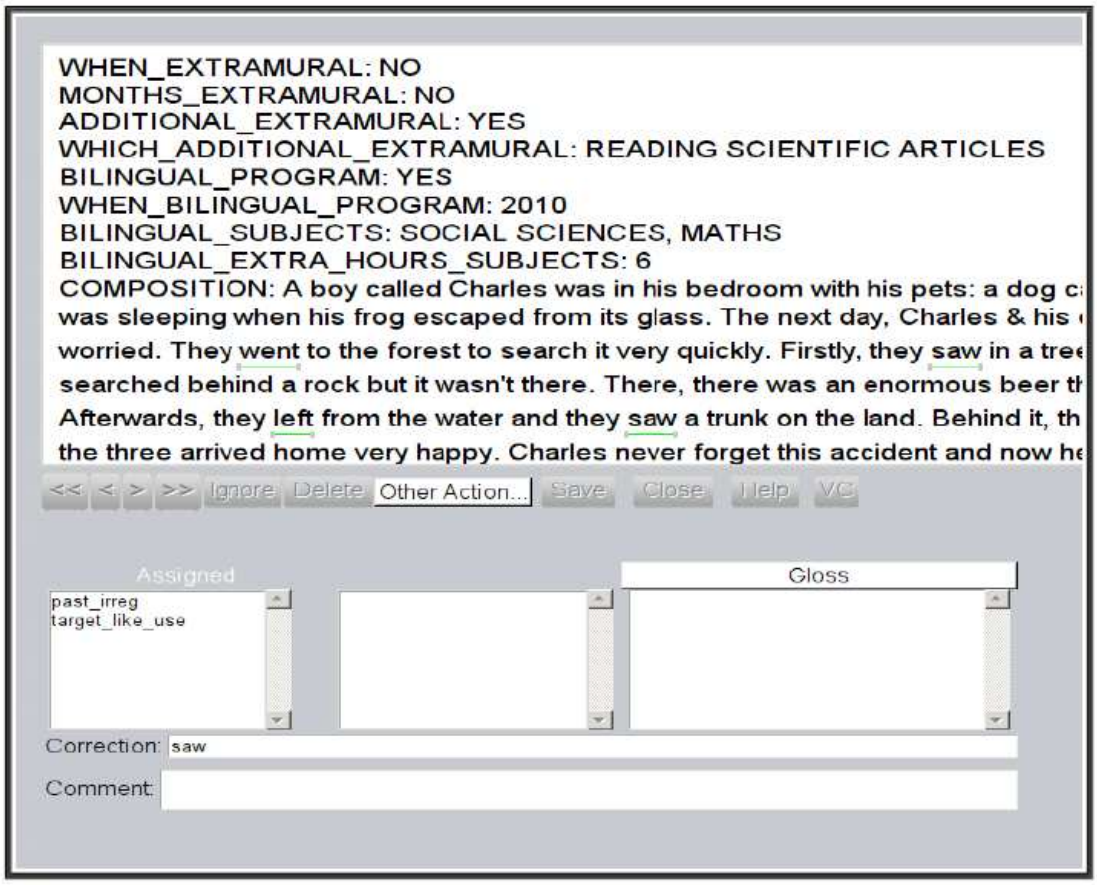
2. Ihre Schwester fiel von der Schaukel (Maria): Maria's Schwester fiel von der Schaukel

3. Ihre Schuhe sind im Schrank(Jungen): Die Schuhe den Jungen sind im Schrank

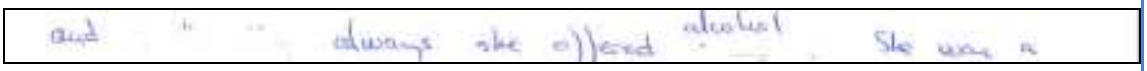
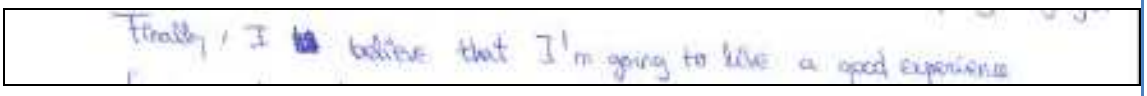
4. Ihre Spielzeuge blieben im Regen draussen liegen (Kinder): Die Spielzuege von den Kinder blieben im Regen draussen liegen.

5. Ihr Computer war kaputt. (Frau): Der Computer der Frau war kaputt.

9.8 UAM Corpus Tool software

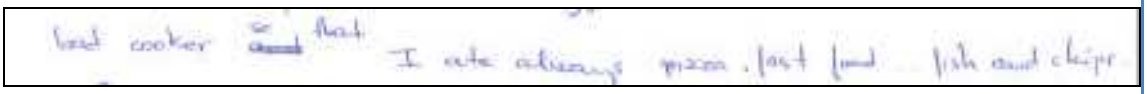


9.9 Tagging Manual: Learners' correction tagging scheme

Rewriting unreadable: [the student edits by rewriting; the original formulation is illegible]
Examples:


[…] and $_RWU_¿ always she offered $_RWU_¿ alcohol. Finally, I $_RWU_¿ believe that I'm going to live a good experience […] $_RWU_¿ rewriting unreadable

Rewriting readable: [the student edits by rewriting; the original formulation is legible]

Example:

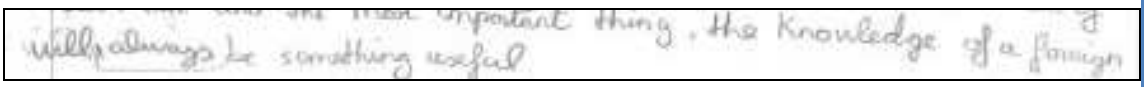


[…] bad cooker $_RWR_and so that I ate always pizza, fastfood... […]

$_RWR_*word* rewriting readable

Reallocation: [the student edits by reallocating units]

Example:



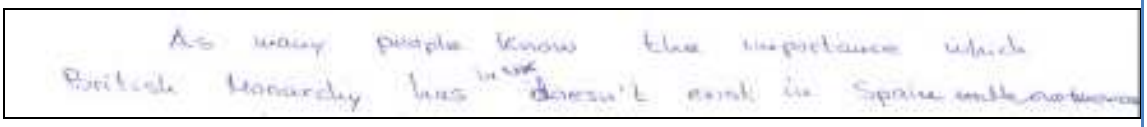
[…] the knowledge of a foreign will $_NLoc _be always $_OLoc__be something useful.

Two tags are needed here: one for the old location and the other the new location. The word that is reallocated occurs after BOTH tags.

$_OLoc_*word* old location
$_NLoc_*word* new location

Late insertion: [the student edits by inserting new units]

Example:



[…] the importance which British Monarchy has $_Ins_in $_Ins_UK doesn't exist in Spain […]

$_Ins_*word* insertion

Unreadable: [the student’s writing is unintelligible]
Example:
$_UNR_¿ unreadable

9.10 Tagging Manual: MOS Project

Morpheme tagging scheme for Past regular

OC: Past reg (Peter walked yesterday)				**S: Supplied form**
Target-like Use (correct form supplied)				Peter walked yesterday
Non-target-like Use	**Underuse**			Peter walk__ yesterday
	Misuse	**Misselection** (form exists)		Peter walking yesterday
		Misrealisation (form does not exist)		n/a
OC: ing (Peter is walked [=walking])				**SNOC**
		Overuse (correct form supplied but in NOC)		Peter is walked

Morpheme tagging scheme for Past Irregular

OC: Past irreg (Peter stole yesterday)				**S: Supplied form**
Target-like Use (correct form supplied)				Peter stole yesterday
Non-target-like Use	**Underuse** (no form supplied)			Peter steal__ yesterday
	Misuse (incorrect form supplied)	**Misselection** (form exists)		Peter stealing yesterday
		Misrealisation (form does not exist)		Peter stealed yesterday
				Peter stoled yesterday
OC: 3rd sing (Peter never stole [=steals])				**SNOC**
		Overuse (correct form supplied but in NOC)		Peter never stole

Morpheme tagging scheme for Third person singular *-s*

OC: 3rd sing (He has a friend) / (He never eats pasta)				**S: Supplied form**
Target-like Use (correct form supplied)				He has a friend / He never eats
Non-target-like Use	**Underuse** (no form supplied)			He have_ a friend / He never eat_
	Misuse (incorrect form supplied)	**Misselection** (form exists)		He had a friend / He never eating
		Misrealisation (form does not exist)		
OC: Past irreg (He has a problem [=had] yesterday)				**SNOC**
		Overuse (correct form supplied but in NOC)		He has a problem yesterday

Morpheme tagging scheme for Present progressive *-ing*

OC: Progressive (Peter is swimming)				**S: Supplied form**
Target-like Use (correct form supplied)				Peter is swimming
Non-target-like Use	**Underuse** (no form supplied)			Peter is swim___
	Misuse (incorrect form supplied)	**Misselection** (form exists)		Peter is swims
		Misrealisation (form does not exist)		n.a.
OC: 3rd sing (Peter swimming [=swims] every day)				**SNOC**
		Overuse (correct form supplied but in NOC)		swimming

Morpheme tagging scheme for Copula *Be*

OC: Copula (Peter is happy)			**S: Supplied form**
Target-like Use (correct form supplied)			Peter is happy
Non-target-like Use	**Underuse** (no form supplied)		Peter __ happy
	Misuse (incorrect form supplied)	**Misselection** (form exists)	Peter be happy
		Misrealisation (form does not exist)	n.a.
OC: 3rd sing HAVE (Peter is a house [=has])			**SNOC**
Overuse (correct form supplied but in NOC)			
		Overuse (correct form supplied but in NOC)	Peter is

Morpheme tagging scheme for Auxiliar *Be*

OC: Aux BE (Peter is swimming)			**S: Supplied form**
Target-like Use (correct form supplied)			Peter is swimming
Non-target-like Use	**Underuse** (no form supplied)		Peter __ swimming
	Misuse (incorrect form supplied)	**Misselection** (form exists)	Peter be swimming
		Misrealisation (form does not exist)	n.a.
OC: aux HAVE (Peter is left [=has])			**SNOC**
		Overuse (correct form supplied but in NOC)	Peter is left

Morpheme tagging scheme for Indefinite article

OC: Indefinite article (A friend of mine came)				**S: Supplied form**
Target-like Use (correct form supplied)				A friend of mine
Non-target-like Use	**Underuse** (no form supplied)			__ friend of mine
	Misuse (incorrect form supplied)	**Misselection** (form exists)		The friend of mine
		Misrealisation (form does not exist)		
OC: generic (I like a pasta [Ø])				**SNOC**
		Overuse (correct form supplied but in NOC)		I like a pasta

Morpheme tagging scheme for Possessive *-s*

OC: Possessive (My mum's car)				**S: Supplied form**
Target-like Use (correct form supplied)				mum's car
Non-target-like Use	**Underuse** (no form supplied)			mum_ car
	Misuse (incorrect form supplied)	**Misselection** (form exists)		mums car
		Misrealisation (form does not exist)		n.a.
OC: Plural (Kid's like toys [=Kids like toys])				**SNOC**
		Overuse (correct form supplied but in NOC)		Kid's like toys

9.11 Descriptive Statistics

9.11.1 Part I: MOS & Learner Corpus

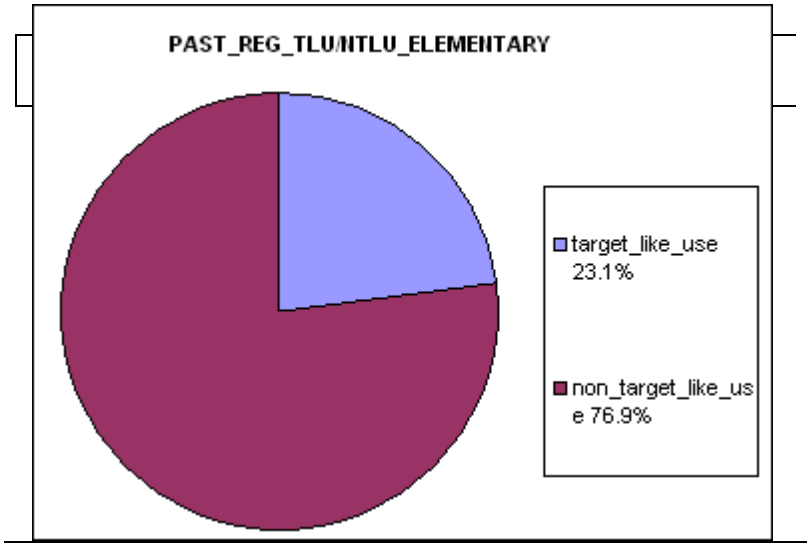
9.11.1.1 Specific Descriptive Statistics (grouping criterion: proficiency level)

Descriptive Statistics: Features

Project:	MOS: Elementary Level

Feature	Percent	N
PAST_REG-TYPE		***N=13***

target_like_use	23.1%	3
non_target_like_use	76.9%	10



NON_TARGET_LIKE_USE-TYPE		N=10
underuse	60.0%	6
misuse	40.0%	4
overuse(snoc)	0.0%	0
unclassified	0.0%	0

MISUSE-TYPE		N=4
misselection	25.0%	1
misrealisation	75.0%	3

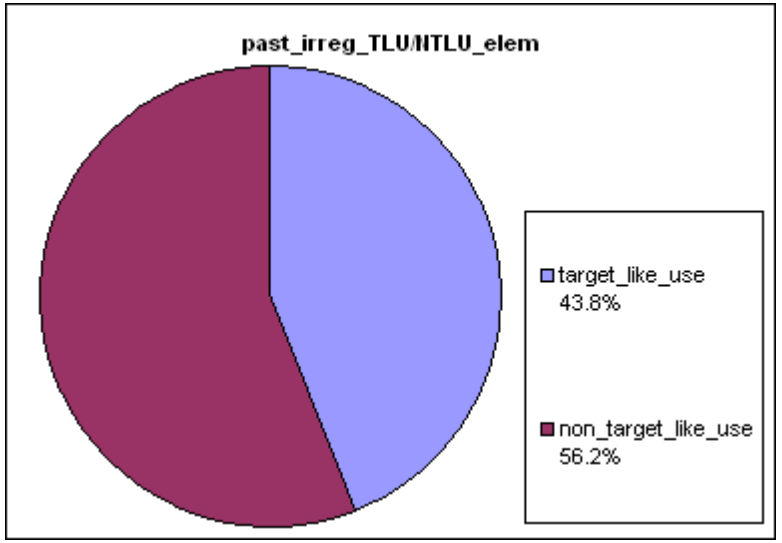
SCORE: [SOC ÷ (OC+SNOC)] × 100

SOC = underuse (0 points)
misuse (0,5 points)
target_like_use (1 point)

SOC_past_reg_elem: (6×0) + (4 × 0,5) + (3 × 1) = 5
OC_past_reg_elem: 13
SNOC_past_reg_elem: 0
Score_past_reg_elem: 5 ÷ (13+0) × 100= 38,46%

Feature	Percent	N
PAST_IRREG-TYPE		***N=16***

target_like_use	43.8%	7
non_target_like_use	56.2%	9



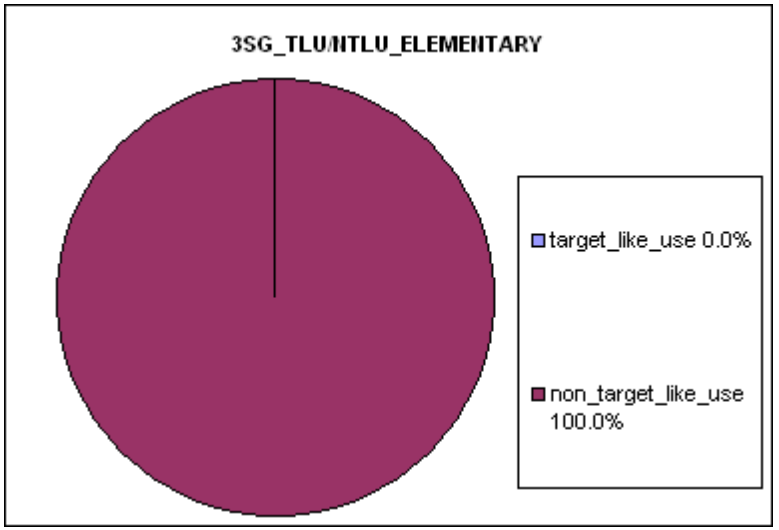
NON_TARGET_LIKE_USE-TYPE		N=9
underuse	55.6%	5
misuse	22.2%	2
overuse(snoc)	22.2%	2
unclassified	0.0%	0

MISUSE-TYPE		N=2
misselection	0.0%	0
misrealisation	100.0%	2

SCORE: [SOC ÷ (OC+SNOC)] × 100
SOC = underuse (0 points)
misuse (0,5 points)

	target_like_use (1 point)
SOC_past_irreg_elem:	(5×0) + (2 × 0,5) + (7 × 1) =8
OC_past_irreg_elem:	16
SNOC_past_irreg_elem:	2
Score_past_irreg_elem:	8 ÷ (16+2) × 100= 44,44 %

Feature	Percent	N
3SG_TYPE		***N=15***
target_like_use	0.0%	0
non_target_like_use	100.0%	15
NON_TARGET_LIKE_USE-TYPE		N=15
underuse	86.7%	13
misuse	13.3%	2



overuse(snoc)	0.0%	0
unclassified	0.0%	0
MISUSE-TYPE		N=2
misselection	0.0%	0
misrealisation	100.0%	2

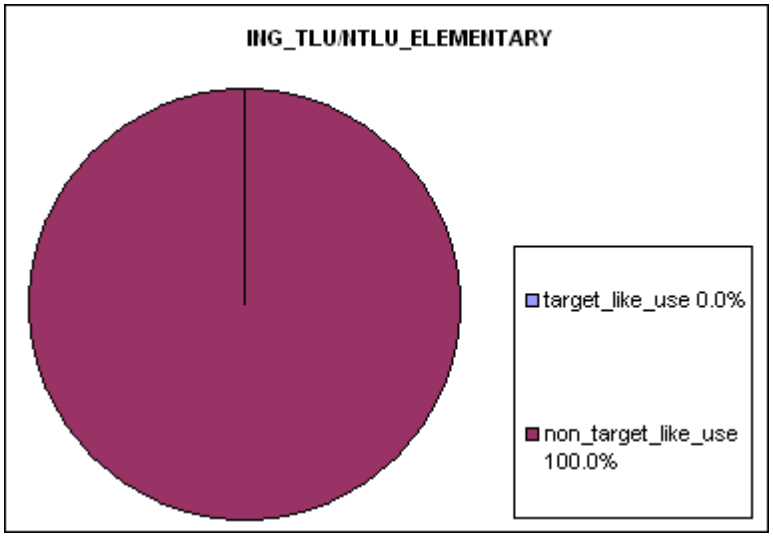
SCORE: [SOC ÷ (OC+SNOC)] × 100

SOC = underuse (0 points)
misuse (0,5 points)
target_like_use (1 point)

SOC_3SG_elementary: (13×0) + (2 × 0,5) + (0× 1) = 1
OC_3SG_elementary: 15
SNOC_3SG_elementary: 0
Score_3SG_elem: 1 ÷ (15+3) × 100= 5,5%

Feature	Percent	N
ING_TYPE		***N=5***

target_like_use	0.0%	0
non_target_like_use	100.0%	5

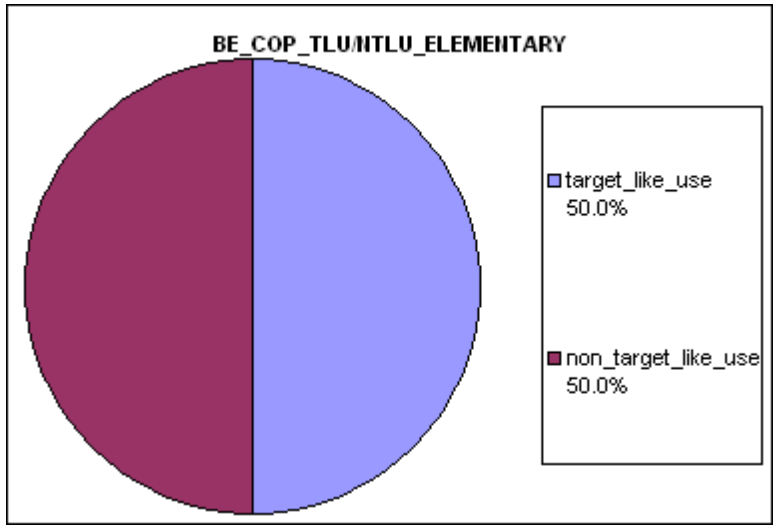


NON_TARGET_LIKE_USE-TYPE		N=5
underuse	0.0%	0
misuse	80.0%	4
overuse(snoc)	20.0%	1
unclassified	0.0%	0
MISUSE-TYPE		N=4
misselection	0.0%	0
misrealisation	100.0%	4

SCORE: [SOC ÷ (OC+SNOC)] × 100
SOC = underuse (0 points)
misuse (0,5 points)
target_like_use (1 point)

SOC_ING_elementary: (0×0) + (4 × 0,5) + (0× 1) = 2
OC_ING_elementary: 5
SNOC_ING_elementary: 1
Score_ING_elem: 2 ÷ (5+1) × 100= 33,3%

Feature	Percent	N
BE_COP_TYPE		***N=8***
target_like_use	50.0%	4
non_target_like_use	50.0%	4
NON_TARGET_LIKE_USE-TYPE		N=4
underuse	0.0%	0
misuse	100.0%	4
overuse(snoc)	0.0%	0
unclassified	0.0%	0



MISUSE-TYPE		N=4
misselection	0.0%	0
misrealisation	100.0%	4

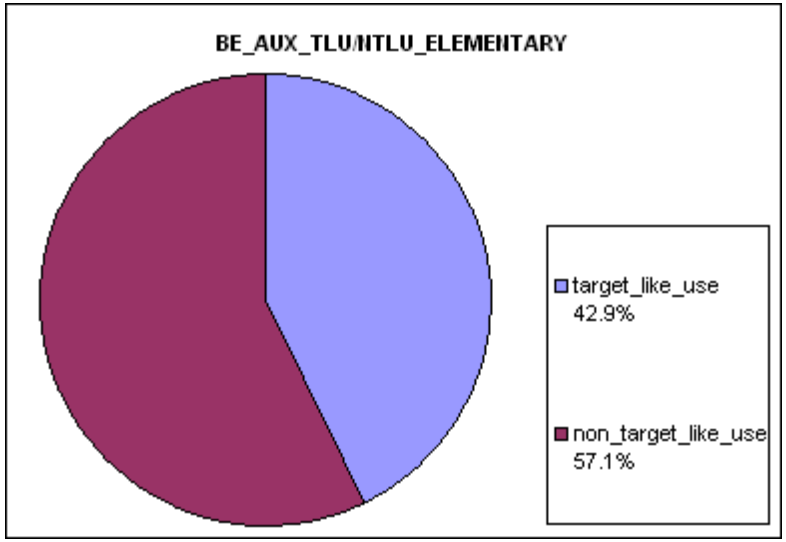
SCORE: [SOC ÷ (OC+SNOC)] × 100

SOC = underuse (0 points)
misuse (0,5 points)
target_like_use (1 point)

SOC_BE_COP_elementary: (0×0) + (4 × 0,5) + (4× 1) = 6
OC_BE_COP_elementary: 8
SNOC_BE_COP_elementary: 0
Score_BE_COP_elem: 6 ÷ (8+0) × 100= 75%

Feature	Percent	N
BE_AUX_TYPE		***N=7***

target_like_use	42.9%	3
non_target_like_use	57.1%	4

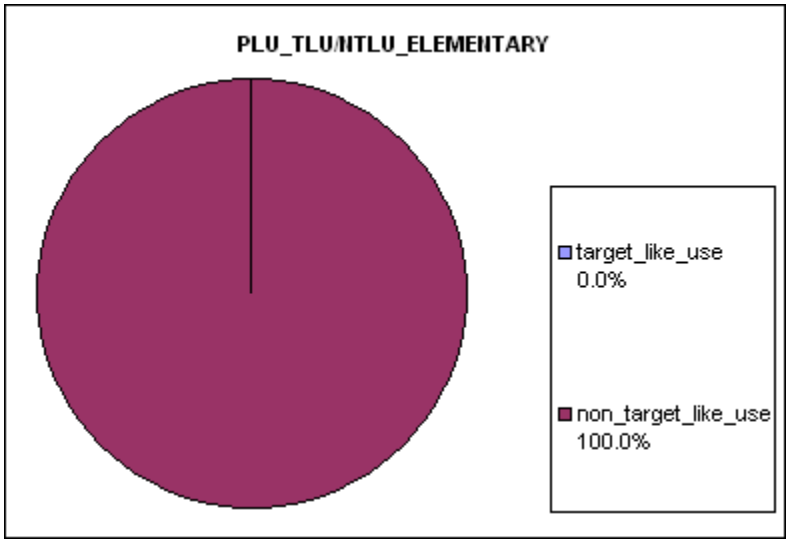


NON_TARGET_LIKE_USE-TYPE		N=4
underuse	50.0%	2
misuse	0.0%	0
overuse(snoc)	50.0%	2
unclassified	0.0%	0
MISUSE-TYPE		N=0
misselection	0.0%	0
misrealisation	0.0%	0

SCORE: [SOC ÷ (OC+SNOC)] × 100
SOC = underuse (0 points)
misuse (0,5 points)
target_like_use (1 point)

SOC_BE_AUX_elementary: (2×0) + (0 × 0,5) + (3× 1) =3
OC_BE_AUX_elementary: 7
SNOC_BE_AUX_elementary: 2
Score_BE_AUX_elem: 3 ÷ (7+2) × 100= 33,3%

Feature	Percent	N
PLU_TYPE		***N=5***
target_like_use	0.0%	0
non_target_like_use	100.0%	5
NON_TARGET_LIKE_USE-TYPE		N=5
underuse	80.0%	4
misuse	0.0%	0
overuse(snoc)	20.0%	1



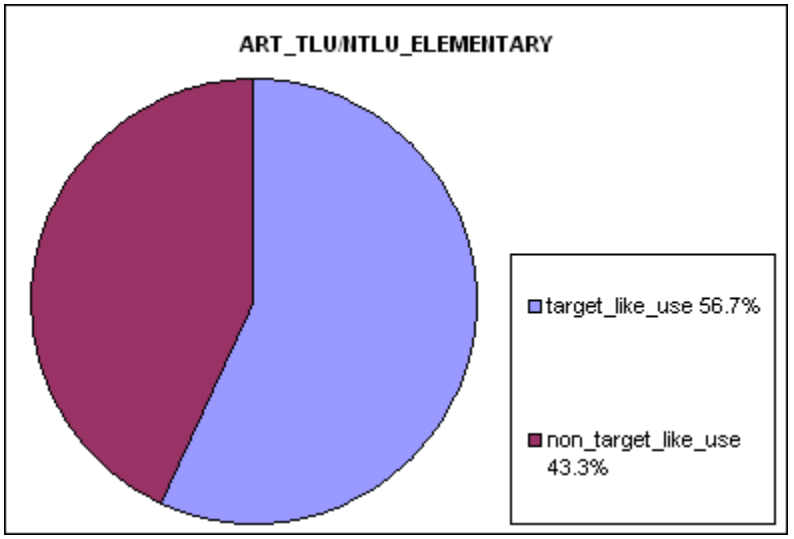
unclassified	0.0%	0
MISUSE-TYPE		N=0
misselection	0.0%	0
misrealisation	0.0%	0

SCORE: [SOC ÷ (OC+SNOC)] × 100

SOC =	underuse (0 points)
	misuse (0,5 points)
	target_like_use (1 point)
SOC_PLU_elementary:	(4×0) + (0× 0,5) + (0× 1) =0
OC_PLU_elementary:	5
SNOC_PLU_elementary:	1
Score_PLU_elem:	0 ÷ (5+1) × 100= 0%

Feature	Percent	N
ART_TYPE		***N=60***

target_like_use	56.7%	34
non_target_like_use	43.3%	26

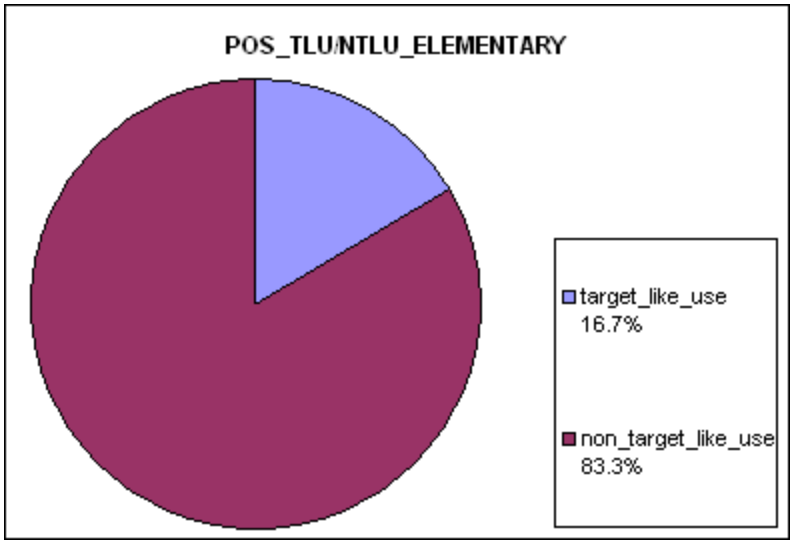


NON_TARGET_LIKE_USE-TYPE		N=26
underuse	15.4%	4
misuse	61.5%	16
overuse(snoc)	23.1%	6
unclassified	0.0%	0
MISUSE-TYPE		N=16
misselection	93.8%	15
misrealisation	6.2%	1

SCORE: [SOC ÷ (OC+SNOC)] × 100
SOC = underuse (0 points)
misuse (0,5 points)
target_like_use (1 point)

SOC_ART_elementary: (4×0) + (16× 0,5) + (34× 1) =42
OC_ART_elementary: 60
SNOC_ART_elementary: 6
Score_ART_elem: 42 ÷ (60+6) × 100=63,63%

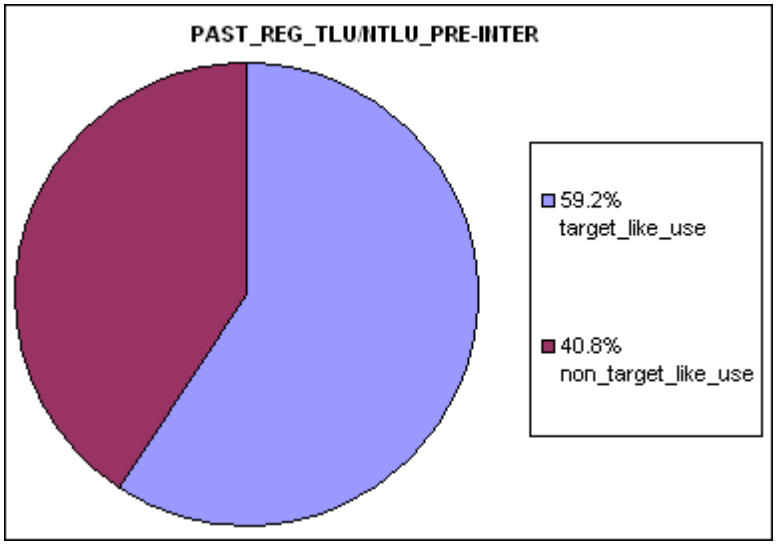
Feature	Percent	N
POS_TYPE		***N=6***
target_like_use	16.7%	1
non_target_like_use	83.3%	5
NON_TARGET_LIKE_USE-TYPE		N=5
underuse	40.0%	2
misuse	40.0%	2
overuse(snoc)	20.0%	1



```
unclassified                      0.0%           0
MISUSE-TYPE                                    N=2
misselection                     50.0%           1
misrealisation                   50.0%           1

SCORE: [SOC ÷ (OC+SNOC)] × 100
                  SOC = underuse (0 points)
                        misuse (0,5 points)
                        target_like_use (1 point)
SOC_POS_elementary:     (2×0) + (2× 0,5) + (1× 1) =2
OC_POS_elementary:                  6
SNOC_POS_elementary:                1
Score_POS_elem:           2 ÷ (6+1) × 100=28,57%
```

Project: MOS: Pre-Intermediate Level

Feature	Percent	N
PAST_REG-TYPE		***N=147***
target_like_use	59.2%	87
non_target_like_use	40.8%	60
NON_TARGET_LIKE_USE-TYPE		N=60
underuse	61.7%	37
misuse	15.0%	9
overuse(snoc)	21.7%	13
unclassified	1.7%	1



MISUSE-TYPE		N=9
misselection	33.3%	3
misrealisation	66.7%	6

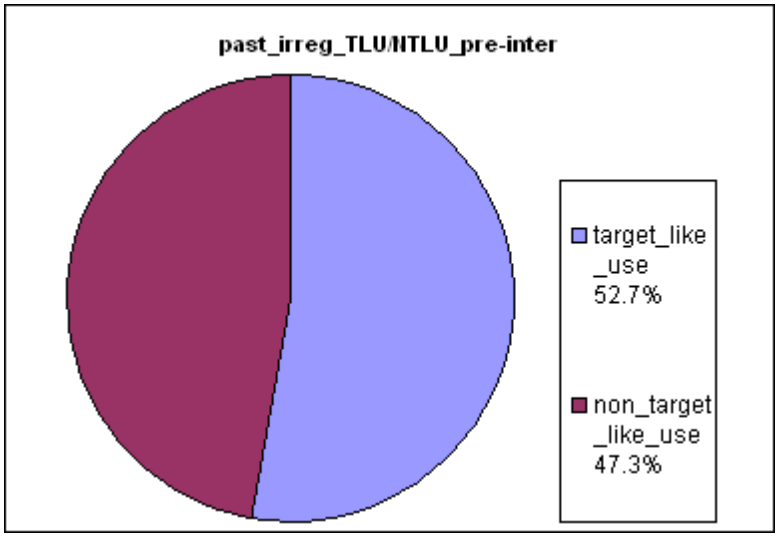
SCORE: [SOC ÷ (OC+SNOC)] × 100

SOC = underuse (0 points)
misuse (0,5 points)
target_like_use (1 point)

SOC_past_reg_pre-intermediate: (37×0) + (9 × 0,5) + (87× 1) = 91,5
OC_past_reg_pre-intermediate: 147
SNOC_past_reg_pre-intermediate: 13
Score_past_reg_pre-inter: 91,5 ÷ (147+13) × 100= 57,18%

Feature	Percent	N
PAST_IRREG-TYPE		***N=148***

target_like_use	52.7%	78
non_target_like_use	47.3%	70



NON_TARGET_LIKE_USE-TYPE		N=70
underuse	50.0%	35
misuse	31.4%	22
overuse(snoc)	18.6%	13
unclassified	0.0%	0
MISUSE-TYPE		N=22
misselection	4.5%	1
misrealisation	95.5%	21

SCORE: [SOC ÷ (OC+SNOC)] × 100
SOC = underuse (0 points)
misuse (0,5 points)
target_like_use (1 point)

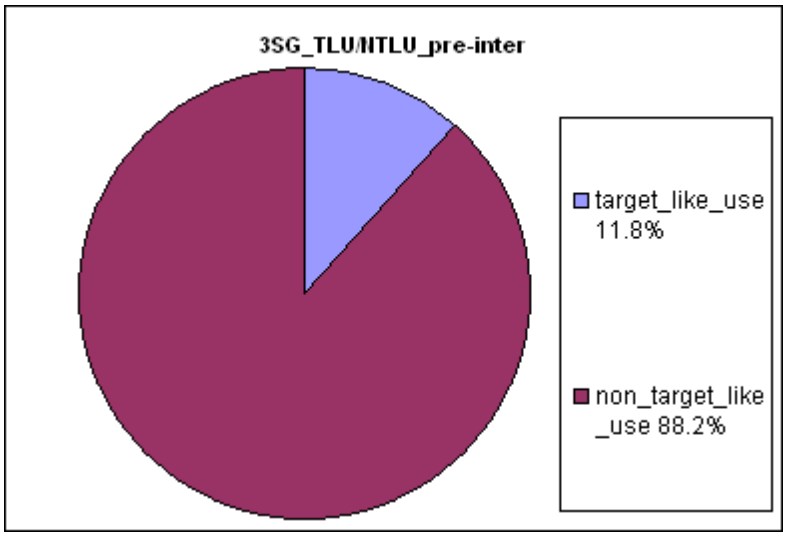
SOC_past_irreg_pre-intermediate: (35×0) + (22 × 0,5) + (78 × 1) = 89

OC_past_irreg_pre-intermediate: 148

SNOC_past_irreg_pre-intermediate: 13

Score_past_irreg_pre-inter: 89 ÷ (148+13) × 100= 55,27 %

Feature	Percent	N
3SG_TYPE		***N=212***
target_like_use	11.8%	25
non_target_like_use	88.2%	187
NON_TARGET_LIKE_USE-TYPE		N=187
underuse	74.9%	140
misuse	23.5%	44
overuse(snoc)	1.6%	3

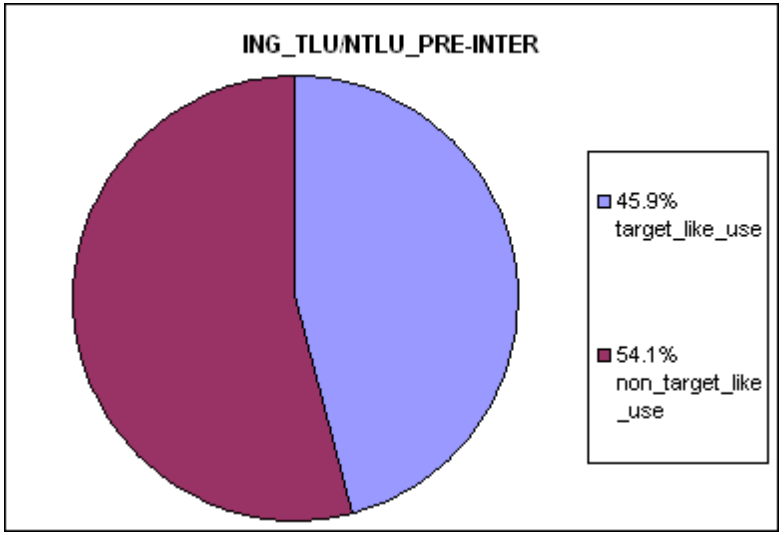


unclassified	0.0%	0
MISUSE-TYPE		N=44
misselection	68.2%	30
misrealisation	31.8%	14

SCORE: [SOC ÷ (OC+SNOC)] × 100
SOC = underuse (0 points)
misuse (0,5 points)
target_like_use (1 point)
SOC_3SG_pre-intermediate: (140×0) + (44 × 0,5) + (25× 1) = 47
OC_3SG_pre-intermediate: 212
SNOC_3SG_pre-intermediate: 3
Score_3SG_pre-inter: 47 ÷ (212+3) × 100=21,86 %

Feature	Percent	N
ING_TYPE		***N=74***

target_like_use	45.9%	34
non_target_like_use	54.1%	40

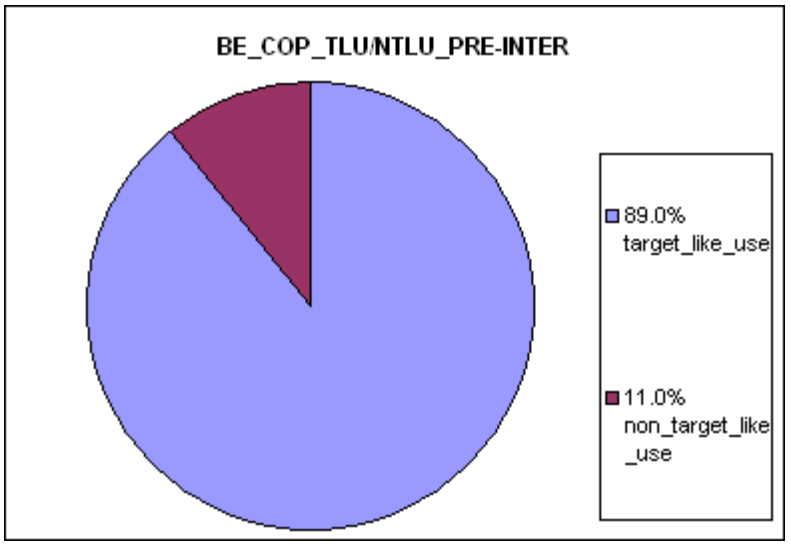


NON_TARGET_LIKE_USE-TYPE		N=40
underuse	35.0%	14
misuse	32.5%	13
overuse(snoc)	32.5%	13
unclassified	0.0%	0
MISUSE-TYPE		N=13
misselection	7.7%	1
misrealisation	92.3%	12

SCORE: [SOC ÷ (OC+SNOC)] × 100
SOC = underuse (0 points)
misuse (0,5 points)
target_like_use (1 point)

SOC_ING_pre-intermediate: (14×0) + (13 × 0,5) + (34× 1) = 40,5
OC_ING_pre-intermediate: 74
SNOC_ING_pre-intermediate: 13
Score_ING_pre-inter: 40,5 ÷ (74+13) × 100= 46,55%

Feature	Percent	N
BE_COP_TYPE		***N=73***
target_like_use	89.0%	65
non_target_like_use	11.0%	8
NON_TARGET_LIKE_USE-TYPE		N=8
underuse	37.5%	3
misuse	62.5%	5
overuse(snoc)	0.0%	0

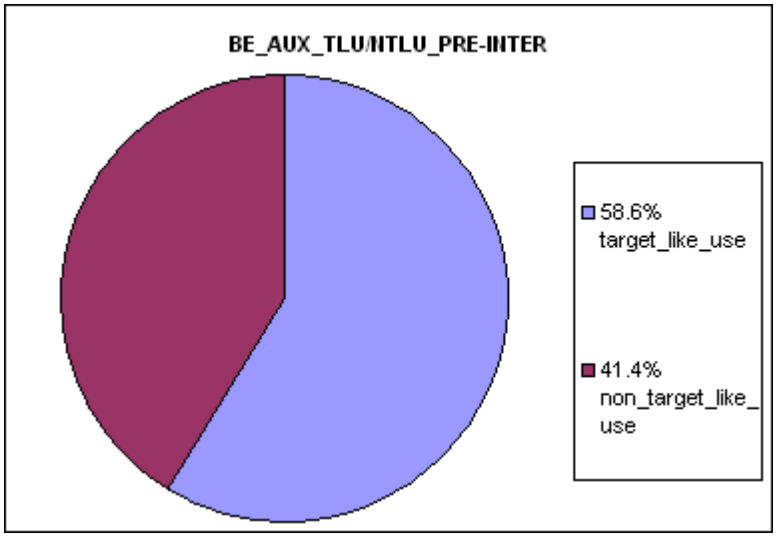


unclassified	0.0%	0
MISUSE-TYPE		N=5
misselection	0.0%	0
misrealisation	100.0%	5

SCORE: [SOC ÷ (OC+SNOC)] × 100
SOC = underuse (0 points)
misuse (0,5 points)
target_like_use (1 point)
SOC_BE_COP_pre-intermediate: (3×0) + (5 × 0,5) + (65× 1) = 67,5
OC_BE_COP_pre-intermediate: 73
SNOC_BE_COP_pre-intermediate: 0
Score_BE_COP_pre-inter: 67,5 ÷ (73+0) × 100= 92,46%

Feature	Percent	N
BE_AUX_TYPE		***N=70***

target_like_use	58.6%	41
non_target_like_use	41.4%	29

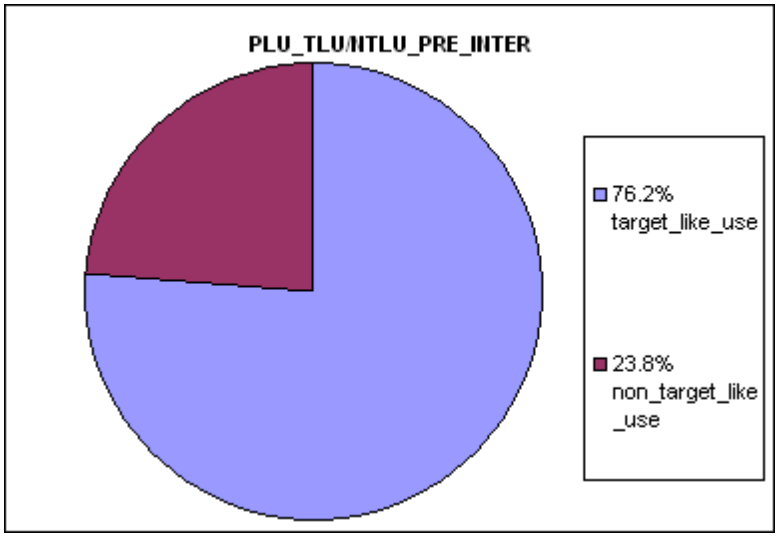


NON_TARGET_LIKE_USE-TYPE		N=29
underuse	13.8%	4
misuse	13.8%	4
overuse(snoc)	72.4%	21
unclassified	0.0%	0
MISUSE-TYPE		N=4
misselection	0.0%	0
misrealisation	100.0%	4

SCORE: [SOC ÷ (OC+SNOC)] × 100
SOC = underuse (0 points)
misuse (0,5 points)
target_like_use (1 point)

SOC_BE_AUX_pre-intermediate: (4×0) + (4 × 0,5) + (41× 1) = 43
OC_BE_AUX_pre-intermediate: 70
SNOC_BE_AUX_pre-intermediate: 21
Score_BE_AUX_pre-inter: 43 ÷ (70+21) × 100= 47,25%

Feature	Percent	N
PLU_TYPE		***N=63***
target_like_use	76.2%	48
non_target_like_use	23.8%	15
NON_TARGET_LIKE_USE-TYPE		N=15
underuse	33.3%	5
misuse	6.7%	1
overuse(snoc)	60.0%	9
unclassified	0.0%	0



MISUSE-TYPE		N=1
misselection	0.0%	0
misrealisation	100.0%	1

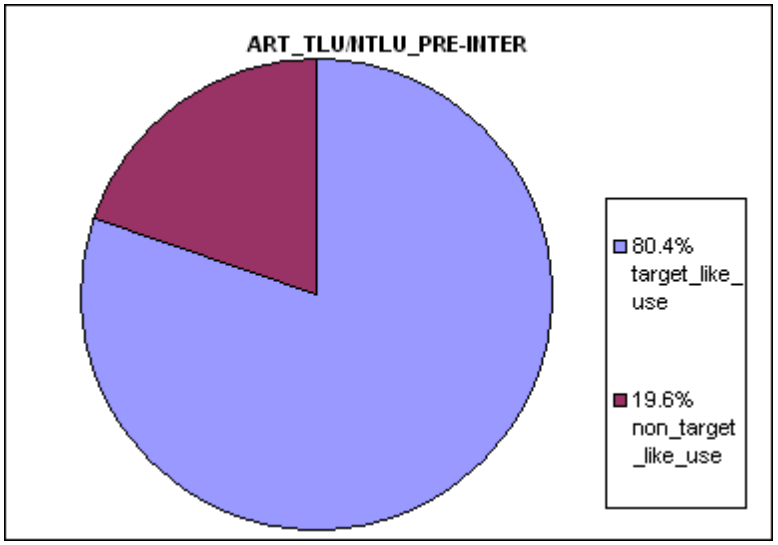
SCORE: [SOC ÷ (OC+SNOC)] × 100

SOC = underuse (0 points)
misuse (0,5 points)
target_like_use (1 point)

SOC_PLU_pre-intermediate: (5×0) + (1× 0,5) + (48× 1) =48,5
OC_PLU_pre-intermediate: 63
SNOC_PLU_pre-intermediate: 9
Score_PLU_pre-inter: 48,5 ÷ (63+9) × 100=67,36%

Feature	Percent	N
ART_TYPE		***N=628***

target_like_use	80.4%	505
non_target_like_use	19.6%	123

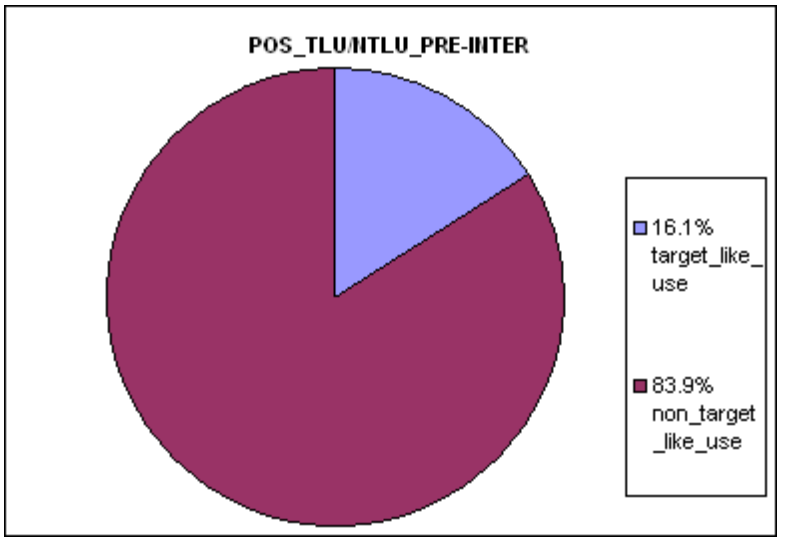


NON_TARGET_LIKE_USE-TYPE		N=123
underuse	23.6%	29
misuse	61.8%	76
overuse(snoc)	14.6%	18
unclassified	0.0%	0
MISUSE-TYPE		N=76
misselection	89.5%	68
misrealisation	10.5%	8

SCORE: [SOC ÷ (OC+SNOC)] × 100
SOC = underuse (0 points)
misuse (0,5 points)
target_like_use (1 point)

SOC_ART_pre-intermediate: (29×0) + (76×0,5) + (505×1) =543
OC_ART_pre-intermediate: 628
SNOC_ART_pre-intermediate: 18
Score_ART_pre-inter: 543 ÷ (628+18) × 100=84,05%

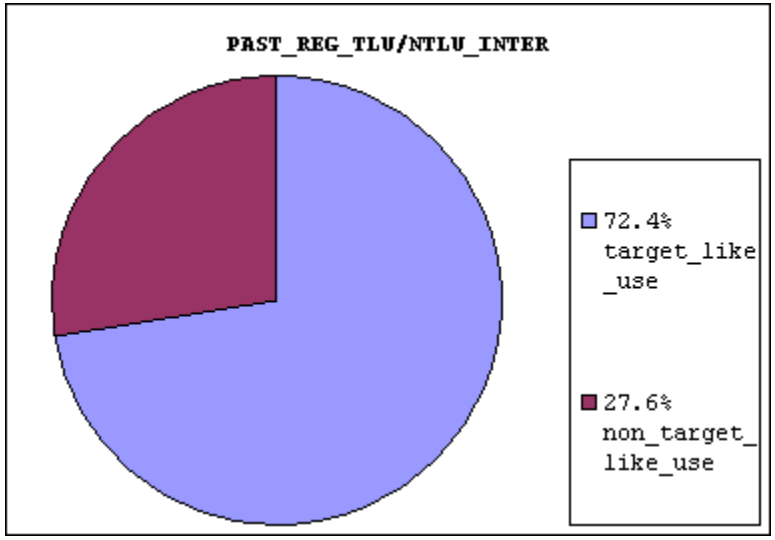
Feature	Percent	N
POS_TYPE		***N=31***
target_like_use	16.1%	5
non_target_like_use	83.9%	26
NON_TARGET_LIKE_USE-TYPE		N=26
underuse	0.0%	0
misuse	96.2%	25
overuse(snoc)	3.8%	1



```
unclassified                                    0.0%     0
MISUSE-TYPE                                             N=25
misselection                                   76.0%    19
misrealisation                                 24.0%     6

SCORE: [SOC ÷ (OC+SNOC)] × 100
                          SOC = underuse (0 points)
                                misuse (0,5 points)
                                target_like_use (1 point)
SOC_POS_pre-intermediate:       (0×0) + (25×0,5) + (5×1) =17,5
OC_POS_pre-intermediate:                          31
SNOC_POS_pre-intermediate:                         5
Score_POS_pre-inter:              17,5 ÷ (31+5) × 100=48,61%
```

Project: MOS: Intermediate Level

Feature	Percent	N
PAST_REG-TYPE		***N=217***
target_like_use	72.4%	157
non_target_like_use	27.6%	60
NON_TARGET_LIKE_USE-TYPE		N=60
underuse	73.3%	44
misuse	10.0%	6
overuse(snoc)	15.0%	9
unclassified	1.7%	1
MISUSE-TYPE		N=6



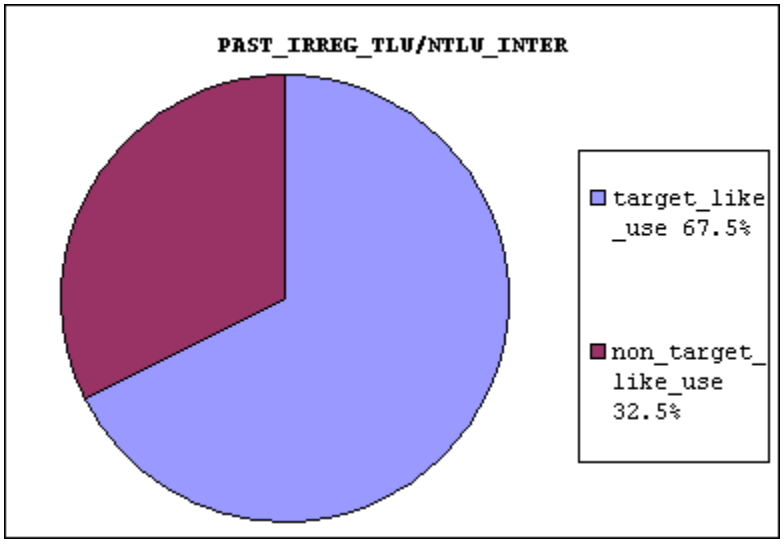
misselection 16.7% 1
misrealisation 83.3% 5

SCORE: [SOC ÷ (OC+SNOC)] × 100
SOC = underuse (0 points)
misuse (0,5 points)
target_like_use (1 point)
SOC_past_reg_intermediate: (44×0) + (6 × 0,5) + (157× 1) = 160
OC_past_reg_intermediate: 217
SNOC_past_reg_intermediate: 9
Score_past_reg_inter: 160 ÷ (217+9) × 100= 70,79%

Feature Percent N

PAST_IRREG-TYPE ***N=255***

target_like_use	67.5%	172
non_target_like_use	32.5%	83



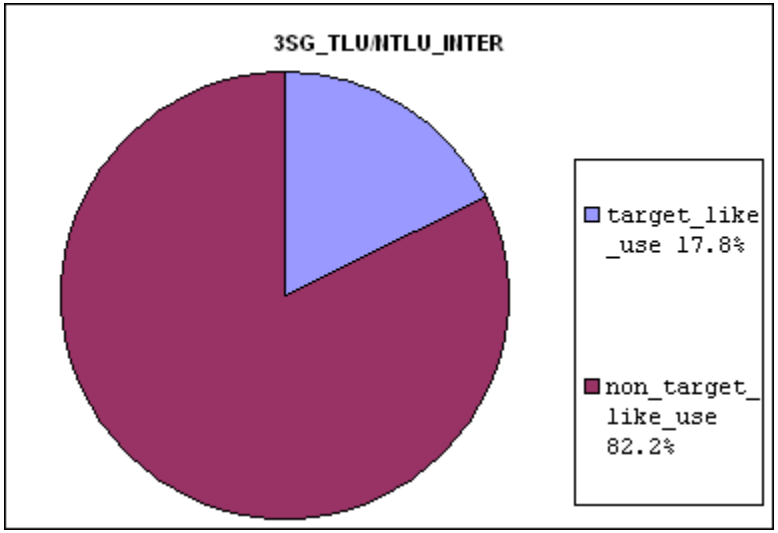
NON_TARGET_LIKE_USE-TYPE		N=83
underuse	51.8%	43
misuse	37.3%	31
overuse(snoc)	9.6%	8
unclassified	1.2%	1

MISUSE-TYPE
misselection
misrealisation

SCORE: [SOC ÷ (OC+SNOC)] × 100
SOC = underuse (0 points)
misuse (0,5 points)
target_like_use (1 point)
SOC_past_irreg_intermediate: (43×0) + (31×0,5) + (172 × 1) = 187,5

OC_past_irreg_intermediate: 255
SNOC_past_irreg_intermediate: 8
Score_past_irreg_inter: 187,5 ÷ (255+8) × 100= 71,29 %

Feature	Percent	N
3SG_TYPE		***N=73***
target_like_use	17.8%	13
non_target_like_use	82.2%	60
NON_TARGET_LIKE_USE-TYPE		N=60
underuse	81.7%	49
misuse	5.0%	3
overuse(snoc)	13.3%	8
unclassified	0.0%	0
MISUSE-TYPE		N=3

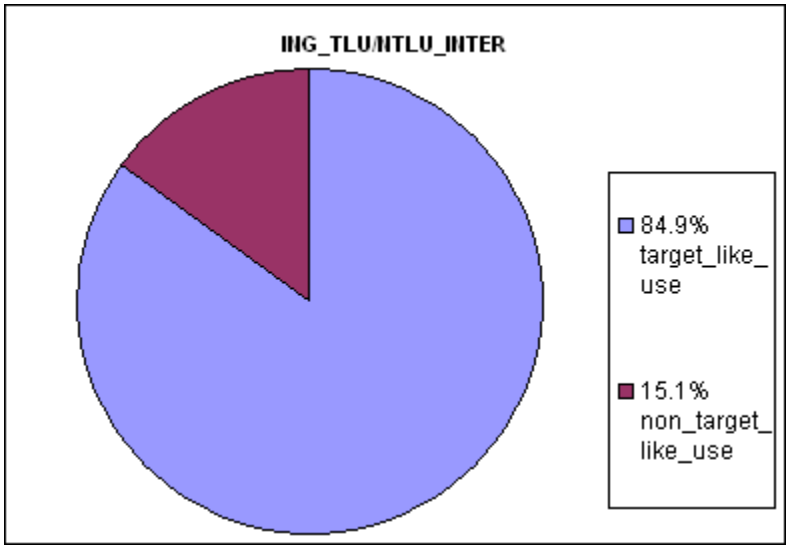


misselection 66.7% 2
misrealisation 33.3% 1

SCORE: [SOC ÷ (OC+SNOC)] × 100
SOC = underuse (0 points)
misuse (0,5 points)
target_like_use (1 point)
SOC_3SG_intermediate: (49×0) + (3 × 0,5) + (13× 1) = 14,5
OC_3SG_intermediate: 73
SNOC_3SG_intermediate: 8
Score_3SG_inter: 14,5 ÷ (73+8) × 100=17,9 %

Feature Percent N

ING_TYPE ***N=73***



target_like_use	84.9%	62
non_target_like_use	15.1%	11
NON_TARGET_LIKE_USE-TYPE		N=11
underuse	18.2%	2
misuse	36.4%	4
overuse(snoc)	45.5%	5
unclassified	0.0%	0
MISUSE-TYPE		N=4
misselection	50.0%	2
misrealisation	50.0%	2

SCORE: [SOC ÷ (OC+SNOC)] × 100
SOC = underuse (0 points)

misuse (0,5 points)
target_like_use (1 point)

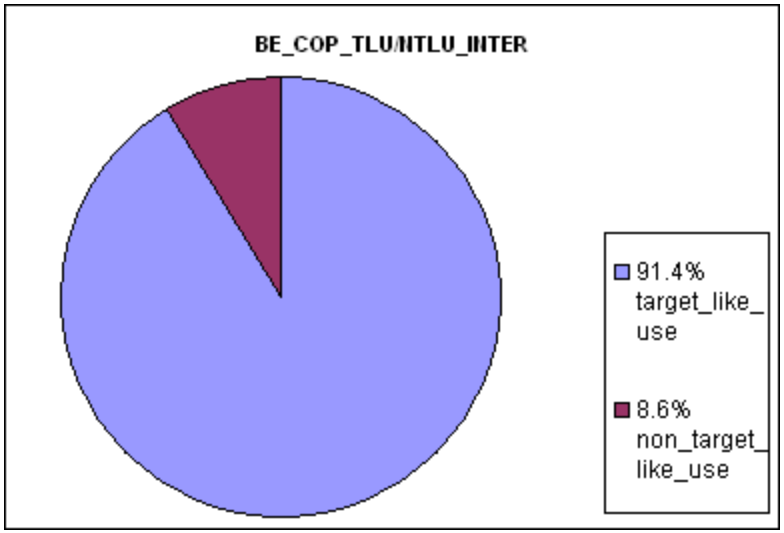
SOC_ING_intermediate: (2×0) + (4×0,5) + (62× 1) = 64

OC_ING_intermediate: 73

SNOC_ING_intermediate: 5

Score_ING_inter: 64 ÷ (73+5) × 100= 82,05%

Feature	Percent	N
BE_COP_TYPE		***N=81***
target_like_use	91.4%	74
non_target_like_use	8.6%	7
NON_TARGET_LIKE_USE-TYPE		N=7
underuse	0.0%	0
misuse	100.0%	7



overuse(snoc)	0.0%	0
unclassified	0.0%	0
MISUSE-TYPE		N=7
misselection	0.0%	0
misrealisation	100.0%	7

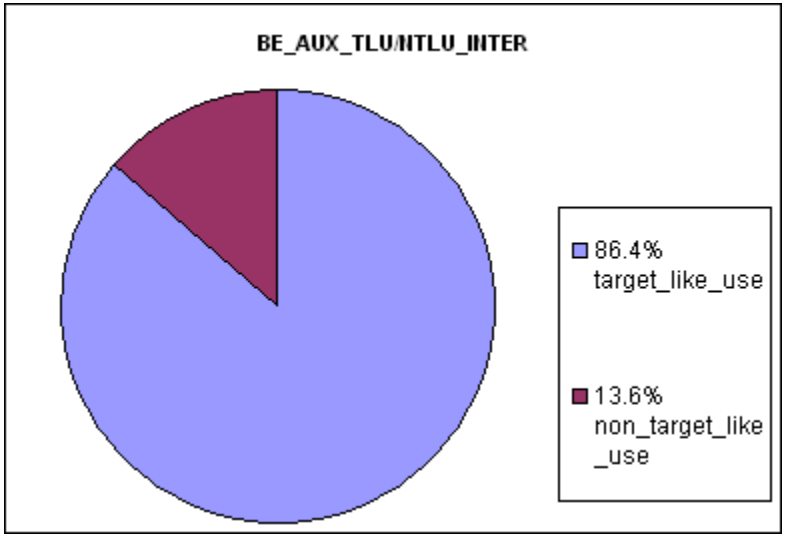
SCORE: [SOC ÷ (OC+SNOC)] × 100

SOC = underuse (0 points)
misuse (0,5 points)
target_like_use (1 point)

SOC_BE_COP_intermediate: (0×0) + (7× 0,5) + (74× 1) =77,5
OC_BE_COP_intermediate: 81
SNOC_BE_COP_intermediate: 0
Score_BE_COP_inter: 77,5 ÷ (81+0) × 100= 95,67%

Feature	Percent	N
BE_AUX_TYPE		***N=59***

target_like_use	86.4%	51
non_target_like_use	13.6%	8

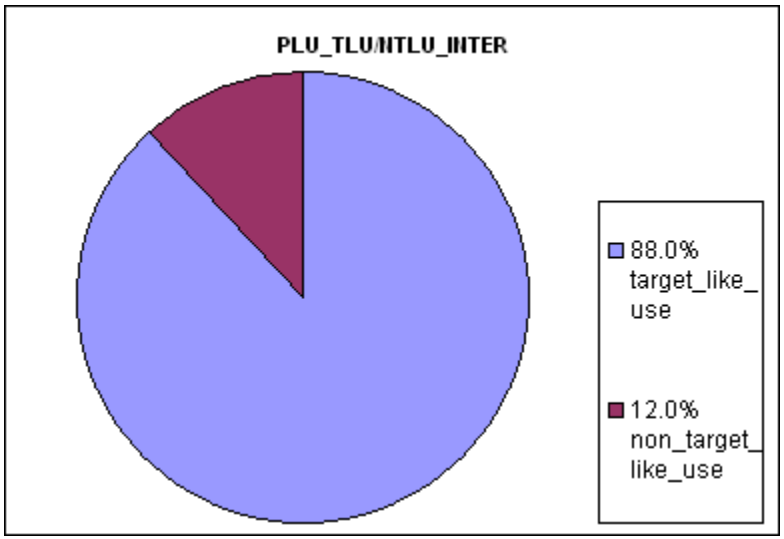


NON_TARGET_LIKE_USE-TYPE		N=8
underuse	12.5%	1
misuse	12.5%	1
overuse(snoc)	75.0%	6
unclassified	0.0%	0
MISUSE-TYPE		N=1
misselection	0.0%	0
misrealisation	100.0%	1

SCORE: [SOC ÷ (OC+SNOC)] × 100
SOC = underuse (0 points)
misuse (0,5 points)
target_like_use (1 point)
SOC_BE_AUX_intermediate: (1×0) + (1 × 0,5) + (51 × 1) = 51,5

OC_BE_AUX_intermediate: 59
SNOC_BE_AUX_intermediate: 6
Score_BE_AUX_inter: 51,5 ÷ (59+6) × 100=79,23 %

Feature	Percent	N
PLU_TYPE		***N=83***
target_like_use	88.0%	73
non_target_like_use	12.0%	10
NON_TARGET_LIKE_USE-TYPE		N=10
underuse	50.0%	5
misuse	20.0%	2
overuse(snoc)	30.0%	3
unclassified	0.0%	0
MISUSE-TYPE		N=2

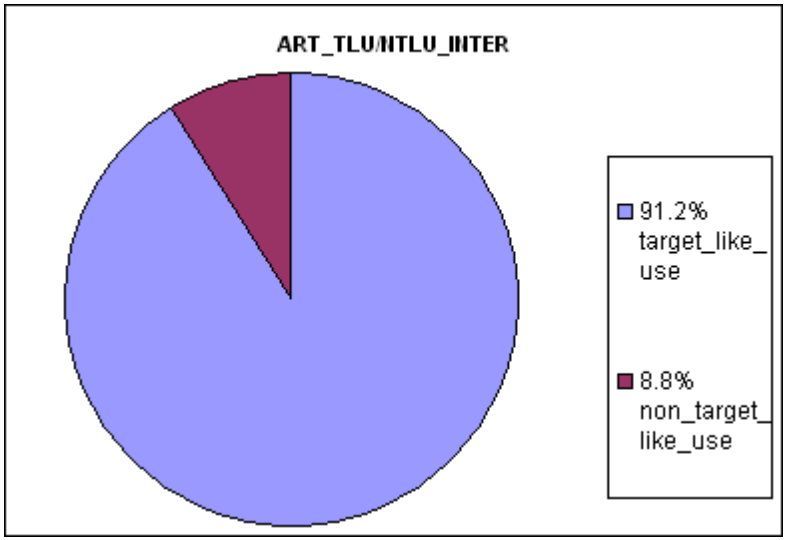


misselection 0.0% 0
misrealisation 100.0% 2

SCORE: [SOC ÷ (OC+SNOC)] × 100
SOC = underuse (0 points)
misuse (0,5 points)
target_like_use (1 point)
SOC_PLU_intermediate: (5×0) + (2× 0,5) + (73× 1) =74
OC_PLU_intermediate: 83
SNOC_PLU_intermediate: 3
74 ÷ (83+3) ×
Score_PLU_inter: 100=86,04%

Feature Percent N

ART_TYPE ***N=638***



target_like_use	91.2%	582
non_target_like_use	8.8%	56
NON_TARGET_LIKE_USE-TYPE		N=56
underuse	7.1%	4
misuse	64.3%	36
overuse(snoc)	28.6%	16
unclassified	0.0%	0
MISUSE-TYPE		N=36
misselection	94.4%	34
misrealisation	5.6%	2

SCORE: [SOC ÷ (OC+SNOC)] × 100
SOC = underuse (0 points)

misuse (0,5 points)
target_like_use (1 point)

SOC_ART_intermediate: (4×0) + (36×0,5) + (582×1) = 600

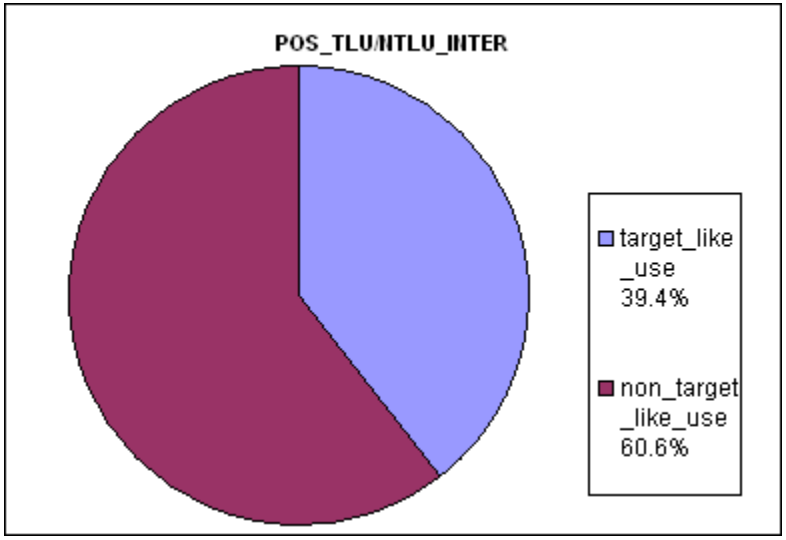
OC_ART_intermediate: 638

SNOC_ART_intermediate: 16

Score_ART_inter: 600 ÷ (638+16) × 100=91,74%

Unit: POS:pos

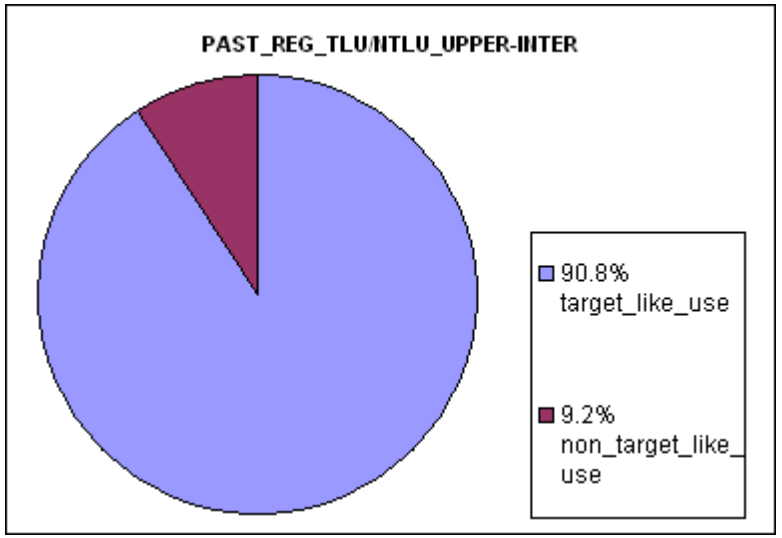
Feature	Percent	N
POS-TYPE	***N=33***	
target_like_use	39.4%	13
non_target_like_use	60.6%	20



```
NON_TARGET_LIKE_USE-TYPE                  N=20

                      underuse                15.0%     3

                        misuse                75.0%    15

                 overuse(snoc)                10.0%     2

SCORE: [SOC ÷ (OC+SNOC)] × 100
                        SOC = underuse (0 points)
                              misuse (0,5 points)
                              target_like_use (1 point)
SOC_POS_intermediate:         (3×0) + (15×0,5) + (13×1) =20,5
OC_POS_intermediate:                            33
SNOC_POS_intermediate:                           2
Score_POS_inter:                 20,5 ÷ (33+2) × 100=58,57%
```

Project:	MOS:Upper Intermediate Level	

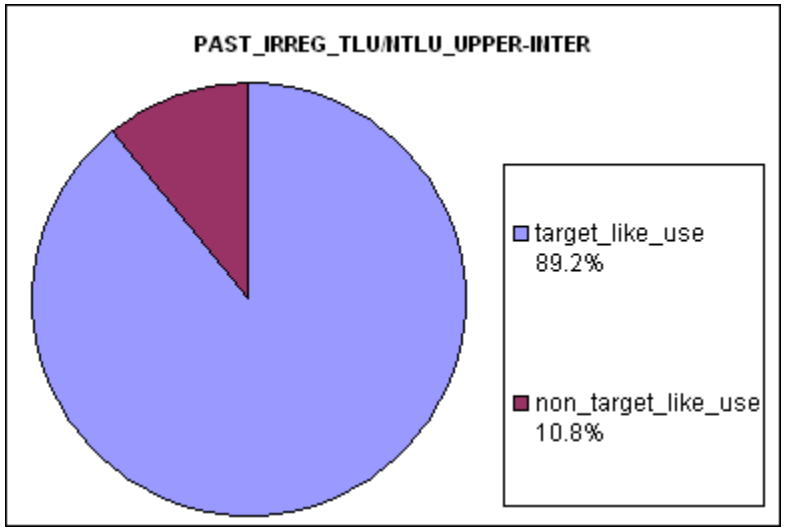
Feature	Percent	N
PAST_REG-TYPE		***N=65***
target_like_use	90.8%	59
non_target_like_use	9.2%	6
NON_TARGET_LIKE_USE-TYPE		N=6
underuse	50.0%	3
misuse	16.7%	1
overuse(snoc)	33.3%	2
unclassified	0.0%	0
MISUSE-TYPE		N=1



misselection 100.0% 1
misrealisation 0.0% 0

SCORE: [SOC ÷ (OC+SNOC)] × 100
SOC = underuse (0 points)
misuse (0,5 points)
target_like_use (1 point)
SOC_past_reg_upper-intermediate: (3×0) + (1 × 0,5) + (59× 1) =59,5
OC_past_reg_upper-intermediate: 65
SNOC_past_reg_upper-intermediate: 2
Score_past_reg_upper-inter: 59,5 ÷ (65+2) × 100= 88,8%

Feature	Percent	N
PAST_IRREG-TYPE		***N=65***



target_like_use	89.2%	58
non_target_like_use	10.8%	7
NON_TARGET_LIKE_USE-TYPE		N=7
underuse	28.6%	2
misuse	42.9%	3
overuse(snoc)	28.6%	2
unclassified	0.0%	0
MISUSE-TYPE		N=3
misselection	0.0%	0
misrealisation	100.0%	3

SCORE: [SOC ÷ (OC+SNOC)] × 100

SOC = underuse (0 points)
misuse (0,5 points)
target_like_use (1 point)

SOC_past_irreg_upper- (2×0) + (3×0,5) + (58 × 1) =59,5

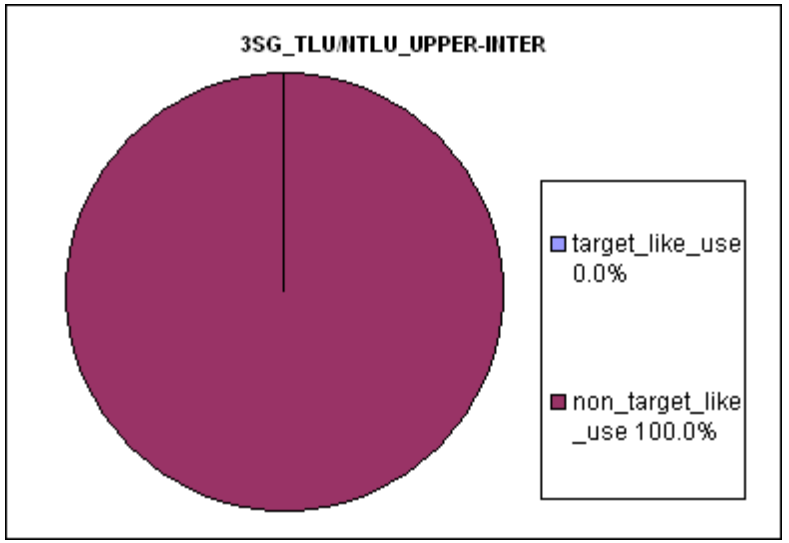
intermediate:

OC_past_irreg_upper-intermediate: 65

SNOC_past_irreg_upper-intermediate: 2

Score_past_irreg_upper-inter: 59,5 ÷ (65+2) × 100= 88,80 %

Feature	Percent	N
3SG_TYPE		***N=1***
target_like_use	0.0%	0
non_target_like_use	100.0%	1
NON_TARGET_LIKE_USE-TYPE		N=1
underuse	0.0%	0
misuse	0.0%	0
overuse(snoc)	100.0%	1
unclassified	0.0%	0



MISUSE-TYPE		N=0
misselection	0.0%	0
misrealisation	0.0%	0

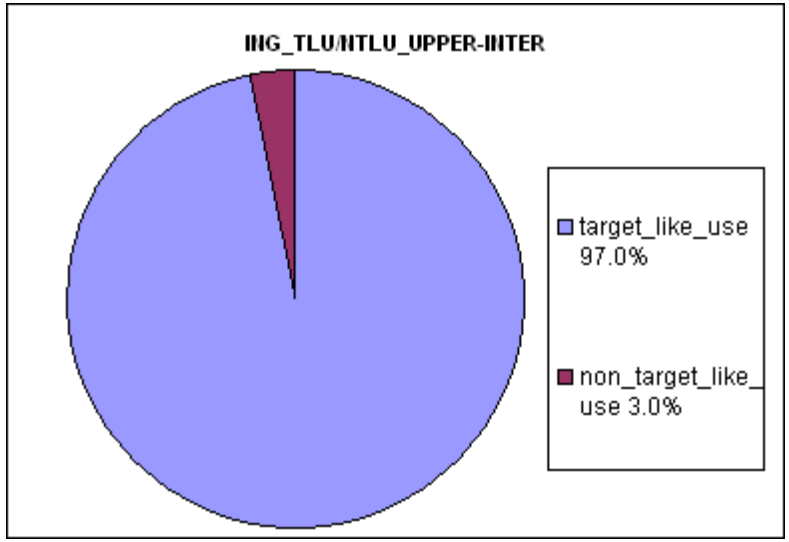
SCORE: [SOC ÷ (OC+SNOC)] × 100

SOC = underuse (0 points)
misuse (0,5 points)
target_like_use (1 point)

SOC_3SG_upper-intermediate: (0×0) + (0 × 0,5) + (0× 1) = 0
OC_3SG_upper-intermediate: 1
SNOC_3SG_upper-intermediate: 1
Score_3SG_upper-inter: 0 ÷ (1+1) × 100=0 %

Feature	Percent	N
ING_TYPE		***N=33***

target_like_use	97.0%	32
non_target_like_use	3.0%	1
NON_TARGET_LIKE_USE-TYPE		N=1
underuse	0.0%	0
misuse	0.0%	0
overuse(snoc)	100.0%	1
unclassified	0.0%	0
MISUSE-TYPE		N=0
misselection	0.0%	0
misrealisation	0.0%	0



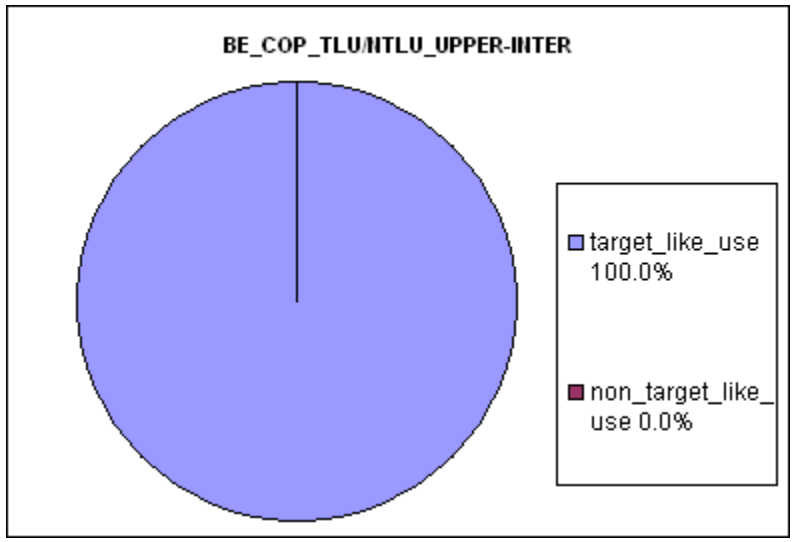
SCORE: [SOC ÷ (OC+SNOC)] × 100

SOC = underuse (0 points)
misuse (0,5 points)
target_like_use (1 point)

SOC_ING_upper-intermediate: (0×0) + (0×0,5) + (32× 1) = 32

OC_ING_upper-intermediate: 33
SNOC_ING_upper-intermediate: 1
Score_ING_upper-inter: 32 ÷ (33+1) × 100=94,11 %

Feature	Percent	N
BE_COP_TYPE		***N=20***
target_like_use	100.0%	20
non_target_like_use	0.0%	0
NON_TARGET_LIKE_USE-TYPE		N=0
underuse	0.0%	0
misuse	0.0%	0
overuse(snoc)	0.0%	0
unclassified	0.0%	0
MISUSE-TYPE		N=0

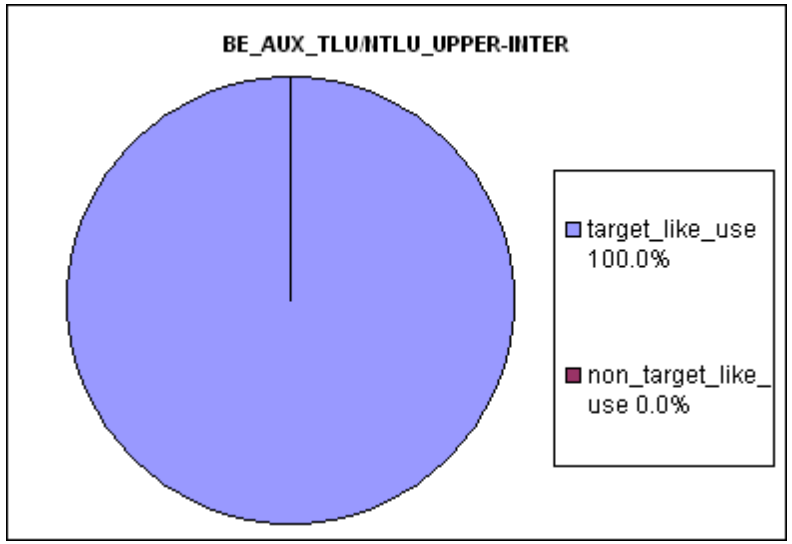


misselection 0.0% 0
misrealisation 0.0% 0

SCORE: [SOC ÷ (OC+SNOC)] × 100
SOC = underuse (0 points)
misuse (0,5 points)
target_like_use (1 point)
SOC_BE_COP_upper-intermediate: (0×0) + (0× 0,5) + (20× 1) =20
OC_BE_COP_upper-intermediate: 20
SNOC_BE_COP_upper-intermediate: 0
Score_BE_COP_upper-inter: 20 ÷ (20+0) × 100= 100%

Feature	Percent	N
BE_AUX_TYPE		***N=28***

target_like_use	100.0%	28
non_target_like_use	0.0%	0
NON_TARGET_LIKE_USE-TYPE		N=0
underuse	0.0%	0
misuse	0.0%	0
overuse(snoc)	0.0%	0
unclassified	0.0%	0
MISUSE-TYPE		N=0
misselection	0.0%	0
misrealisation	0.0%	0



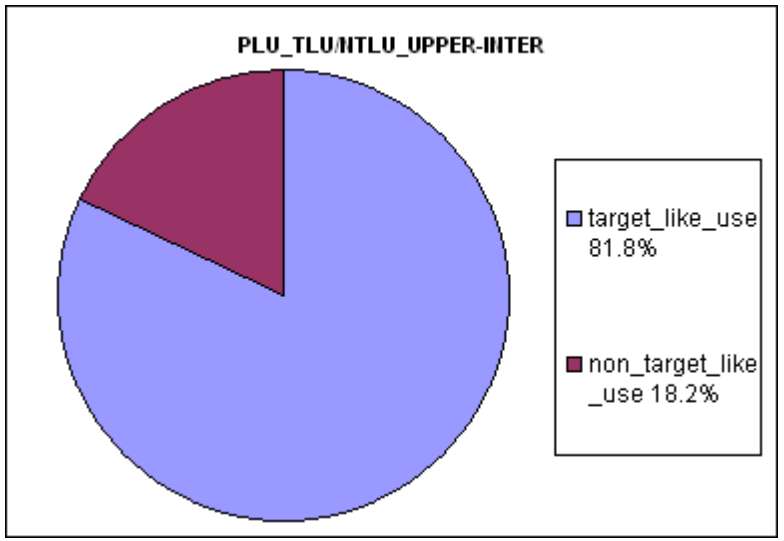
SCORE: [SOC ÷ (OC+SNOC)] × 100

SOC = underuse (0 points)
misuse (0,5 points)
target_like_use (1 point)

SOC_BE_AUX_upper-intermediate: (0×0) + (0 × 0,5) + (28 × 1) = 28

OC_BE_AUX_upper-intermediate:		28
SNOC_BE_AUX_upper-intermediate:		6
Score_BE_AUX_upper-inter:	28 ÷ (28+0) × 100=100 %	

Feature	Percent	N
PLU_TYPE		***N=22***
target_like_use	81.8%	18
non_target_like_use	18.2%	4
NON_TARGET_LIKE_USE-TYPE		N=4
underuse	75.0%	3
misuse	25.0%	1
overuse(snoc)	0.0%	0
unclassified	0.0%	0
MISUSE-TYPE		N=1



misselection	0.0%	0
misrealisation	100.0%	1

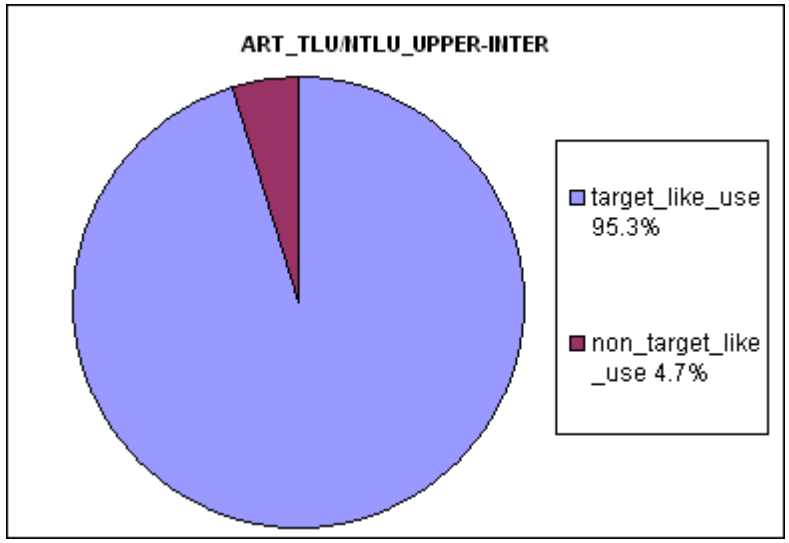
SCORE: [SOC ÷ (OC+SNOC)] × 100

SOC = underuse (0 points)
misuse (0,5 points)
target_like_use (1 point)

SOC_PLU_upper-intermediate: (3×0) + (1× 0,5) + (18× 1) =18,5
OC_PLU_upper-intermediate: 22
SNOC_PLU_upper-intermediate: 0
Score_PLU_upper-inter: 18,5 ÷ (22+0) × 100=84,09%

Feature	Percent	N
ART_TYPE		***N=129***

target_like_use	95.3%	123
non_target_like_use	4.7%	6
NON_TARGET_LIKE_USE-TYPE		N=6
underuse	0.0%	0
misuse	83.3%	5
overuse(snoc)	16.7%	1
unclassified	0.0%	0
MISUSE-TYPE		N=5
misselection	100.0%	5
misrealisation	0.0%	0



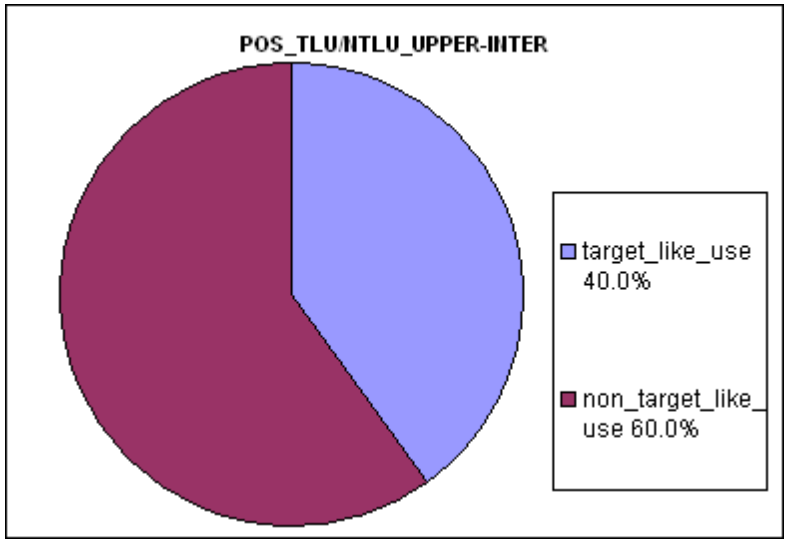
SCORE: [SOC ÷ (OC+SNOC)] × 100

SOC = underuse (0 points)
misuse (0,5 points)
target_like_use (1 point)

SOC_ART_upper-intermediate: (0×0) + (5×0,5) + (123×1) =125,5

OC_ART_upper-intermediate: 129
SNOC_ART_upper-intermediate: 1
Score_ART_upper-inter: 125,5 ÷ (129+1) × 100=96,53%

Feature	Percent	N
POS_TYPE		***N=10***
target_like_use	40.0%	4
non_target_like_use	60.0%	6
NON_TARGET_LIKE_USE-TYPE		N=6
underuse	16.7%	1
misuse	66.7%	4
overuse(snoc)	16.7%	1
unclassified	0.0%	0
MISUSE-TYPE		N=4



misselection 25.0% 1
misrealisation 75.0% 3

SCORE: [SOC ÷ (OC+SNOC)] × 100

SOC = underuse (0 points)
misuse (0,5 points)
target_like_use (1 point)

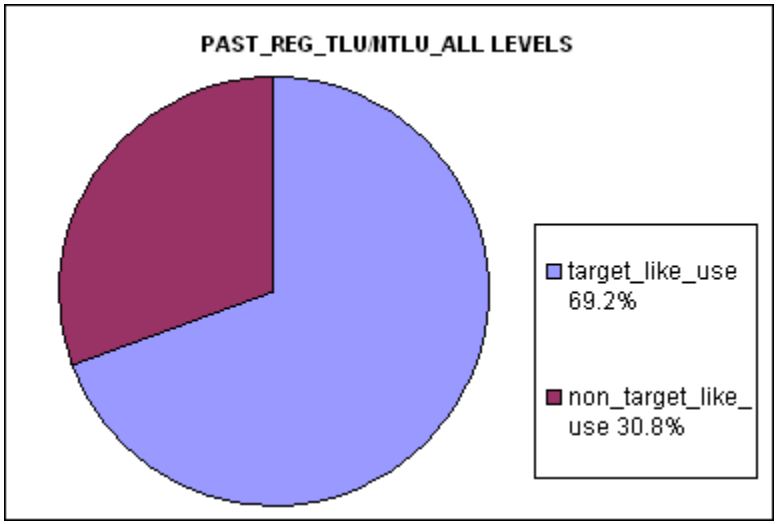
SOC_POS_upper-intermediate: (1×0) + (4×0,5) + (4×1) =6
OC_POS_upper-intermediate: 10
SNOC_POS_upper-intermediate: 1
Score_POS_upper-inter: 6 ÷ (10+1) × 100=54,54%

9.11.1.2 General Descriptive Statistics (grouping criterion: none)

Descriptive Statistics: Features

MOS: ALL
PROFICIENCY
Project: LEVELS

Feature	Percent	N
PAST_REG-TYPE		***N=442***
target_like_use	69.2%	306
non_target_like_use	30.8%	136
NON_TARGET_LIKE_USE-TYPE		N=136



underuse	66.2%	90
misuse	14.7%	20
overuse(snoc)	17.6%	24
unclassified	1.5%	2
MISUSE-TYPE		N=20
misselection	30.0%	6
misrealisation	70.0%	14

SCORE: [SOC ÷ (OC+SNOC)] × 100

SOC = underuse (0 points)
misuse (0,5 points)
target_like_use (1 point)

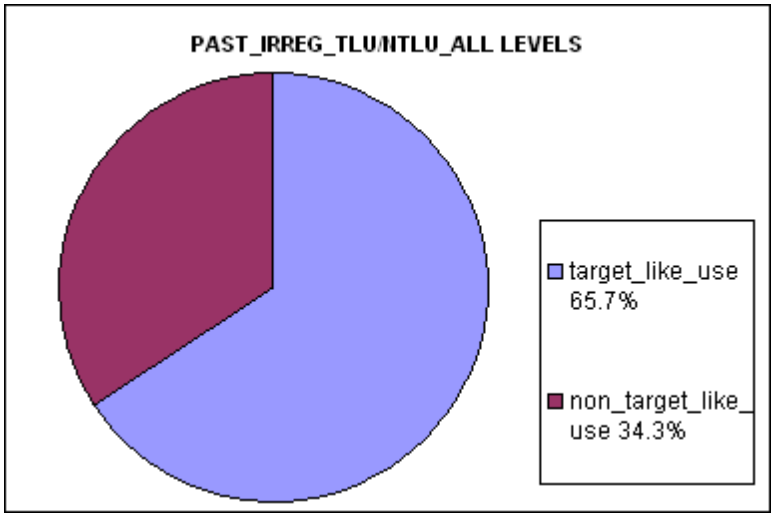
SOC_past_reg: (90×0) + (20 × 0,5) + (306 × 1) =316

OC_past_reg: 442

SNOC_past_reg: 24

Score_past_reg: 316 ÷ (442+24) × 100= 67,81%

Feature	Percent	N
PAST_IRREG-TYPE		***N=478***



target_like_use	65.7%	314
non_target_like_use	34.3%	164
NON_TARGET_LIKE_USE-TYPE		N=164
underuse	51.2%	84
misuse	34.1%	56
overuse(snoc)	14.6%	24
unclassified	0.0%	0
MISUSE-TYPE		N=56
misselection	10.7%	6
misrealisation	89.3%	50

SCORE: [SOC ÷ (OC+SNOC)] × 100

SOC = underuse (0 points)
misuse (0,5 points)
target_like_use (1 point)

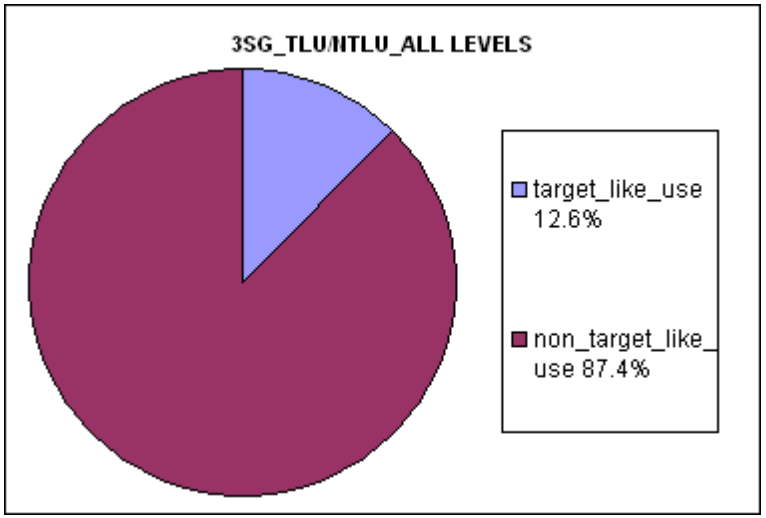
SOC_past_irreg: (84×0) + (56 × 0,5) + (314 × 1) =370

OC_past_irreg: 478

SNOC_past_irreg: 24

Score_past_irreg: 370 ÷ (478+24) × 100= 73,70%

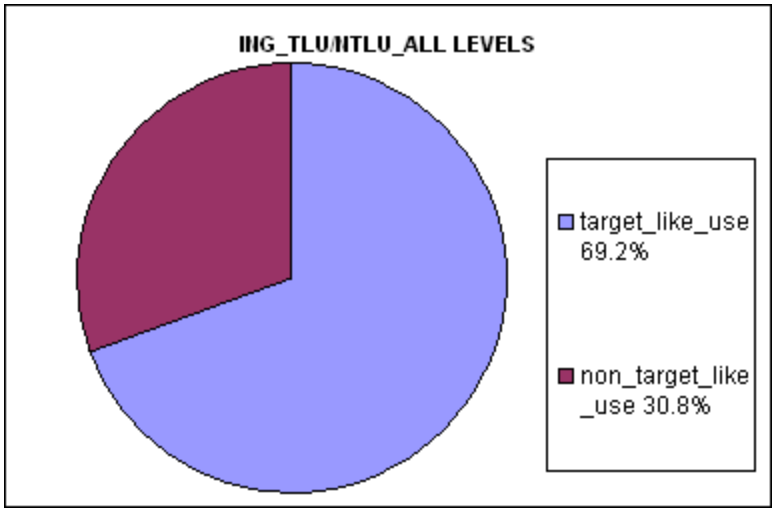
Feature	Percent	N
3SG-TYPE		***N=301***
target_like_use	12.6%	38
non_target_like_use	87.4%	263
NON_TARGET_LIKE_USE-TYPE		N=263
underuse	76.8%	202
misuse	18.6%	49
overuse(snoc)	4.6%	12



```
         unclassified                  0.0%     0
MISUSE-TYPE                                    N=49
         misselection                 65.3%    32
       misrealisation                 34.7%    17
SCORE: [SOC ÷ (OC+SNOC)] × 100
                SOC = underuse (0 points)
                      misuse (0,5 points)
                      target_like_use (1 point)
SOC_3sg:              (202×0) + (49 × 0,5) + (38 × 1) =62,5
OC_3sg:                                301
SNOC_3sg:                               12
Score_3sg:             62,5 ÷ (301+12) × 100= 19,96%
```

Feature	Percent	N
ING-TYPE		***N=185***

target_like_use	69.2%	128
non_target_like_use	30.8%	57



NON_TARGET_LIKE_USE-TYPE		N=57
underuse	28.1%	16
misuse	36.8%	21
overuse(snoc)	35.1%	20
unclassified	0.0%	0
MISUSE-TYPE		N=21
misselection	14.3%	3
misrealisation	85.7%	18

SCORE: [SOC ÷ (OC+SNOC)] × 100

SOC = underuse (0 points)
misuse (0,5 points)
target_like_use (1 point)

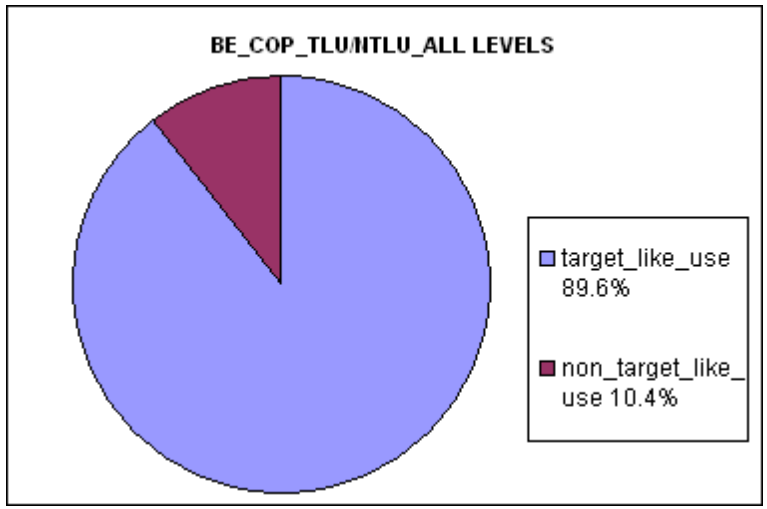
SOC_ing: (16×0) + (21 × 0,5) + (128 × 1) = 138,5

OC_ing: 185

SNOC_ing: 20

Score_ing: 138,5 ÷ (185+20) × 100= 67,56%

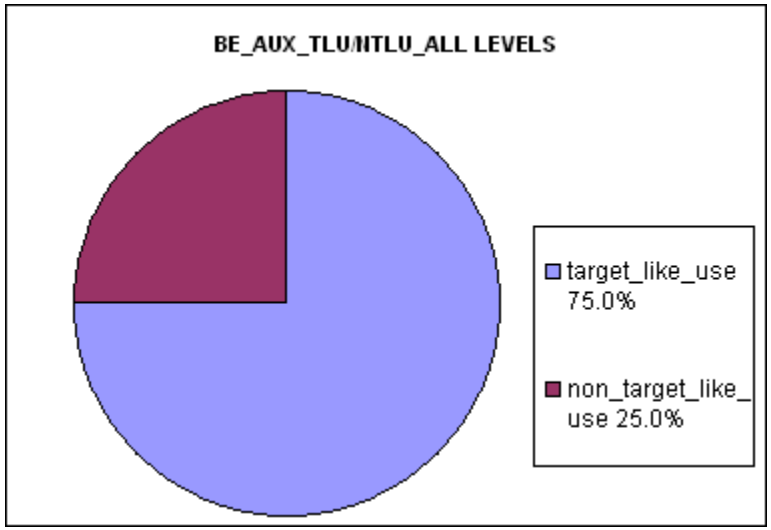
Feature	Percent	N
BE_COP-TYPE		***N=182***
target_like_use	89.6%	163
non_target_like_use	10.4%	19
NON_TARGET_LIKE_USE-TYPE		N=19
underuse	15.8%	3
misuse	84.2%	16
overuse(snoc)	0.0%	0
unclassified	0.0%	0
MISUSE-TYPE		N=16



misselection 0.0% 0
misrealisation 100.0% 16
SCORE: [SOC ÷ (OC+SNOC)] × 100
SOC = underuse (0 points)
misuse (0,5 points)
target_like_use (1 point)
SOC_be_cop: (3×0) + (16 × 0,5) + (163 × 1) = 171
OC_be_cop: 182
SNOC_be_cop: 0
Score_be_cop: 171 ÷ (182+0) × 100= 93,95%

Feature	Percent	N
BE_AUX-TYPE		***N=164***

target_like_use	75.0%	123
non_target_like_use	25.0%	41



NON_TARGET_LIKE_USE-TYPE		N=41
underuse	17.1%	7
misuse	12.2%	5
overuse(snoc)	70.7%	29
unclassified	0.0%	0
MISUSE-TYPE		N=5
misselection	0.0%	0
misrealisation	100.0%	5

SCORE: [SOC ÷ (OC+SNOC)] × 100

SOC = underuse (0 points)
misuse (0,5 points)
target_like_use (1 point)

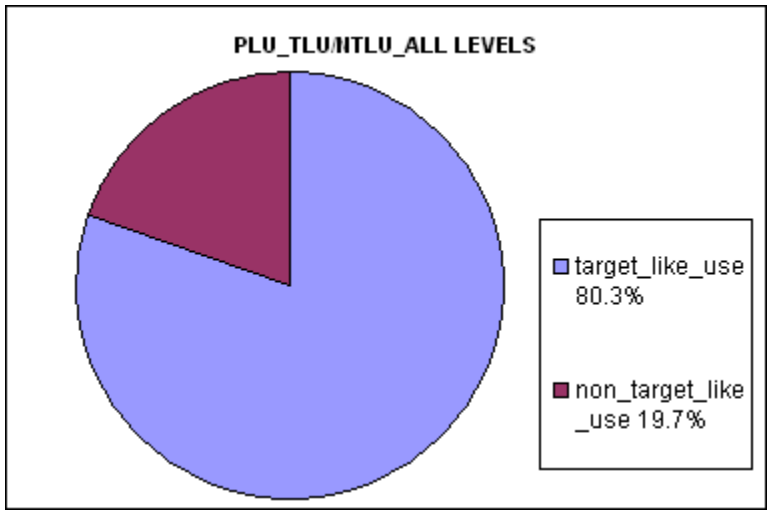
SOC_be_aux: (7×0) + (5 × 0,5) + (123 × 1) = 125,5

OC_be_aux: 164

SNOC_be_aux: 29

Score_be_aux: 125,5 ÷ (164+29) × 100= 65,02%

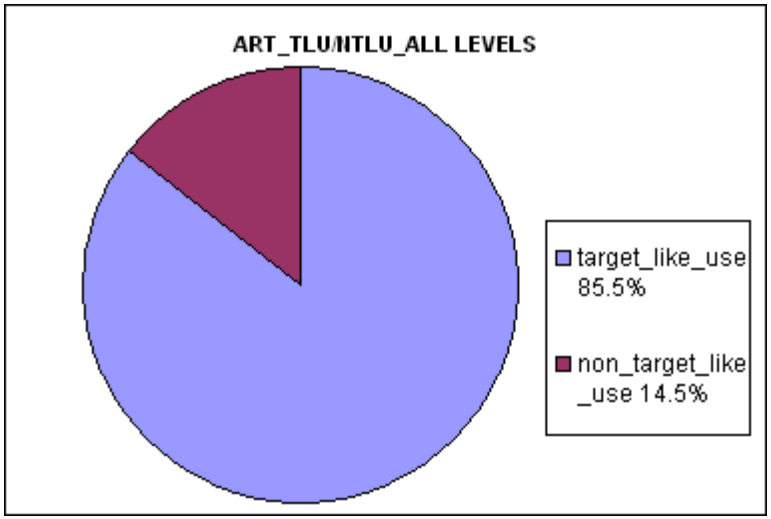
Feature	Percent	N
PLU-TYPE		***N=173***
target_like_use	80.3%	139
non_target_like_use	19.7%	34
NON_TARGET_LIKE_USE-TYPE		N=34
underuse	50.0%	17
misuse	11.8%	4
overuse(snoc)	38.2%	13
unclassified	0.0%	0
MISUSE-TYPE		N=4



```
         misselection                  0.0%      0
       misrealisation                100.0%      4
SCORE: [SOC ÷ (OC+SNOC)] × 100
                SOC = underuse (0 points)
                      misuse (0,5 points)
                      target_like_use (1 point)
SOC_plu:              (17×0) + (4 × 0,5) + (139 × 1) = 141
OC_plu:                                 173
SNOC_plu:                                13
Score_plu:             141 ÷ (173+13) × 100= 75,80%
```

Feature	Percent	N
ART-TYPE		***N=1455***

target_like_use	85.5%	1244
non_target_like_use	14.5%	211



NON_TARGET_LIKE_USE-TYPE		N=211
underuse	17.5%	37
misuse	63.0%	133
overuse(snoc)	19.4%	41
unclassified	0.0%	0
MISUSE-TYPE		N=133
misselection	91.7%	122
misrealisation	8.3%	11

SCORE: [SOC ÷ (OC+SNOC)] × 100

SOC = underuse (0 points)
misuse (0,5 points)
target_like_use (1 point)

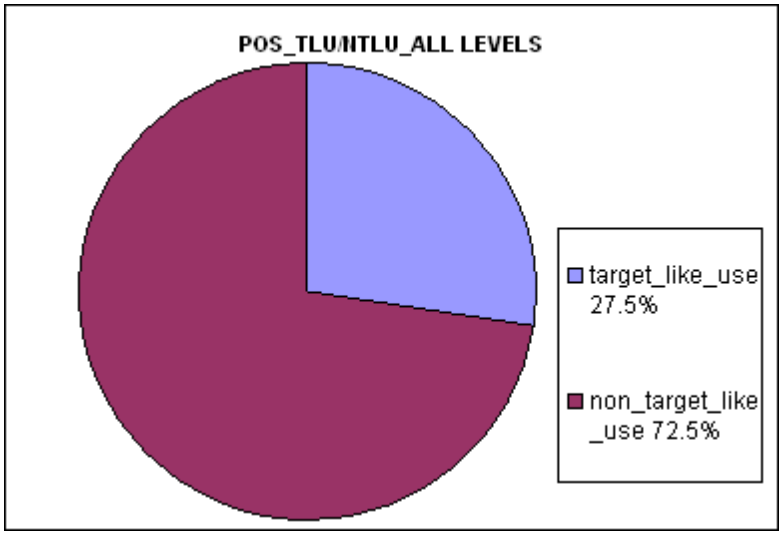
SOC_art: (37×0) + (133 × 0,5) + (1244 × 1) = 1310,5

OC_art: 1455

SNOC_art: 41

Score_art: 1310,5 ÷ (1455+41) × 100= 87,60%

Feature	Percent	N
POS-TYPE		***N=80***
target_like_use	27.5%	22
non_target_like_use	72.5%	58
NON_TARGET_LIKE_USE-TYPE		N=58
underuse	10.3%	6
misuse	81.0%	47
overuse(snoc)	8.6%	5
unclassified	0.0%	0
MISUSE-TYPE		N=47



```
         misselection                70.2%    33
       misrealisation                29.8%    14
SCORE: [SOC ÷ (OC+SNOC)] × 100
                SOC = underuse (0 points)
                      misuse (0,5 points)
                      target_like_use (1 point)
SOC_pos:              (6×0) + (47 × 0,5) + (22 × 1) = 45,5
OC_pos:                                  80
SNOC_pos:                                 5
Score_pos:             45,5 ÷ (80+5) × 100= 53,52%
```

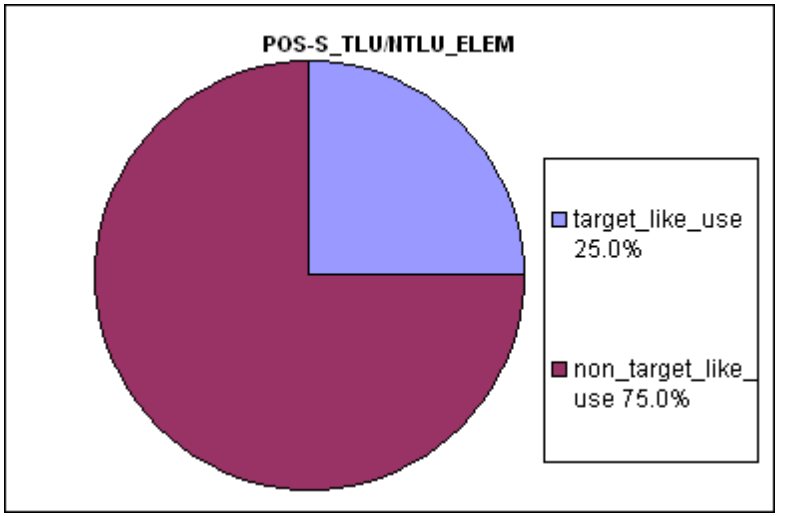
9.11.2 Part II: The use of the possessive –s by L2 English learners: an experimentally elicited data study

Descriptive Statistics: Features

Project	possessive -s

Elementary Level

Feature	Percent	N
POS-TYPE	***N=108***	



target_like_use	25.0%	27
non_target_like_use	75.0%	81

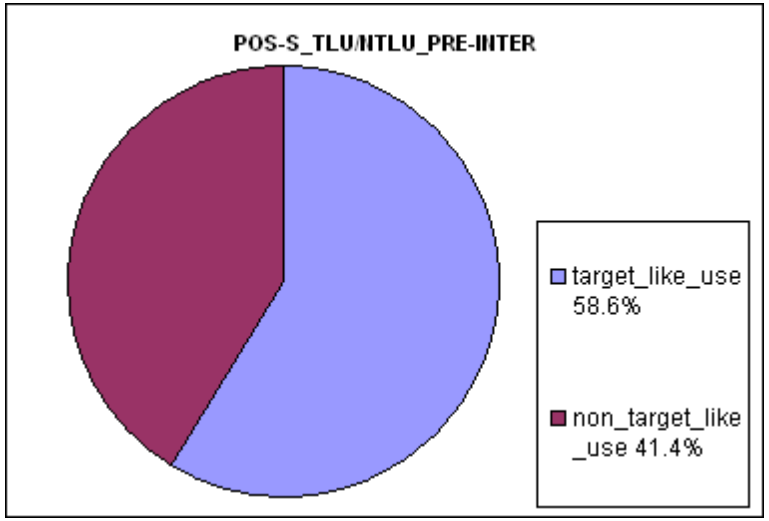
NON_TARGET_LIKE_USE-TYPE	N=81	
underuse	77.8%	63
misuse	22.2%	18
overuse-(snoc)	0.0%	0
UNDERUSE-TYPE	**N=63**	
blank	41.3%	26
no-'/-'s	58.7%	37
MISUSE-TYPE	**N=18**	
misselection	22.2%	4
misrealisation	27.8%	5
wrong-unclassified	50.0%	9

SCORE: [SOC ÷ (OC+SNOC)] × 100
SOC = underuse (0 points)

	misuse (0,5 points) target_like_use (1 point)
SOC_POS_elementary:	(63×0) + (18× 0,5) + (27× 1) =36
OC_POS_elementary:	81
SNOC_POS_elementary:	0
Score_POS_elem:	36 ÷ (81+0) × 100=44,44%

Pre-intermediate Level

Feature	Percent	N
POS-TYPE	***N=636***	
target_like_use	58.6%	373
non_target_like_use	41.4%	263



NON_TARGET_LIKE_USE-TYPE	N=263	
underuse	55.1%	145
misuse	44.9%	118
overuse-(snoc)	0.0%	0
UNDERUSE-TYPE	N=145	
blank	35.9%	52
no-'/-'s	64.1%	93
MISUSE-TYPE	N=118	
misselection	28.0%	33
misrealisation	34.7%	41
wrong-unclassified	37.3%	44

SCORE: [SOC ÷ (OC+SNOC)] × 100

SOC = underuse (0 points)
misuse (0,5 points)
target_like_use (1 point)

SOC_POS_pre-intermediate: (145×0) + (118× 0,5) + (373× 1) =432

OC_POS_pre-intermediate: 636

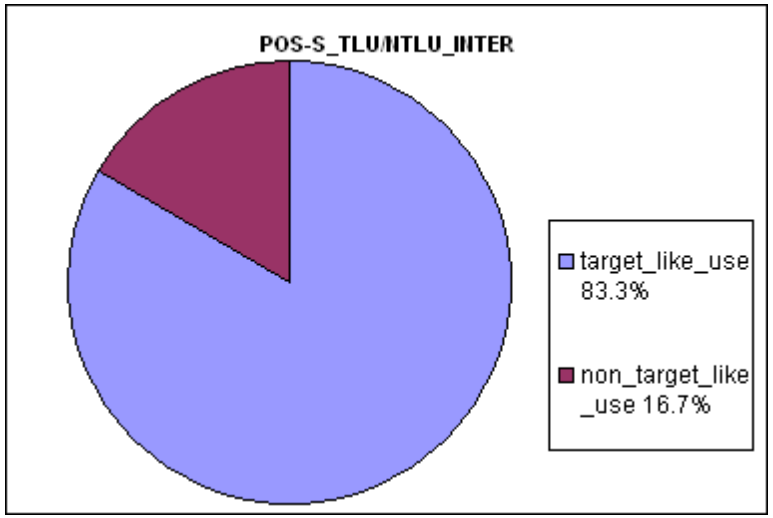
SNOC_POS_pre-intermediate: 0

Score_POS_pre-inter: 432 ÷ (636+0) × 100=67,92%

Intermediate Level

Feature	Percent	N

POS-TYPE ***N=606***



target_like_use	83.3%	505
non_target_like_use	16.7%	101

NON_TARGET_LIKE_USE-TYPE	N=101	
underuse	18.8%	19
misuse	81.2%	82
overuse-(snoc)	0.0%	0
UNDERUSE-TYPE	**N=19**	
blank	21.1%	4
no-'/-'s	78.9%	15
MISUSE-TYPE	**N=82**	
misselection	34.1%	28
misrealisation	41.5%	34
wrong-unclassified	24.4%	20

SCORE: [SOC ÷ (OC+SNOC)] × 100

SOC = underuse (0 points)
misuse (0,5 points)
target_like_use (1 point)

SOC_POS_intermediate: (19×0) + (82× 0,5) + (505× 1) = 546

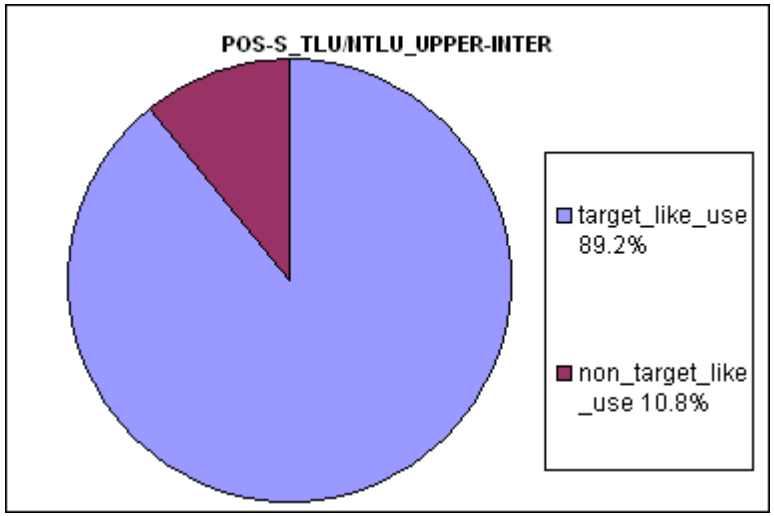
OC_POS_intermediate: 606

SNOC_POS_intermediate: 0

Score_POS_inter: 546 ÷ (606+0) × 100=90,09%

Upper-intermediate Level

Feature	Percent	N
POS-TYPE	***N=186***	



target_like_use	89.2%	166
non_target_like_use	10.8%	20

NON_TARGET_LIKE_USE-TYPE	N=20	
underuse	75.0%	15
misuse	25.0%	5
overuse-(snoc)	0.0%	0
UNDERUSE-TYPE	**N=15**	
blank	40.0%	6
no-'/-'s	60.0%	9
MISUSE-TYPE	**N=5**	
misselection	0.0%	0
misrealisation	60.0%	3
wrong-unclassified	40.0%	2

SCORE: [SOC ÷ (OC+SNOC)] × 100
SOC = underuse (0 points)

	misuse (0,5 points) target_like_use (1 point)
SOC_POS_upper- intermediate:	(15×0) + (5× 0,5) + (166× 1) =168,5
OC_POS_upper-intermediate:	186
SNOC_POS_upper- intermediate:	0
Score_POS_upper-inter:	168,5 ÷ (186+0) × 100=90,59%

9.11.3 Part III: The use of the possessive structures by L3 German learners: an experimentally elicited data study

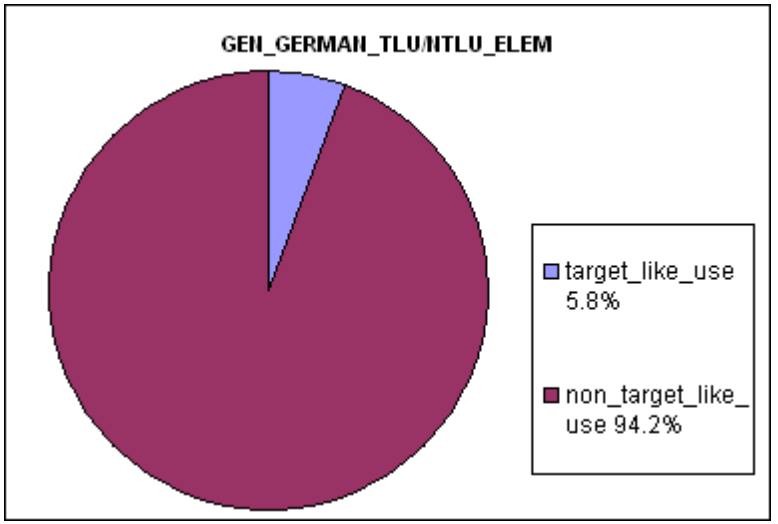
9.11.3.1 The relative frequency of the various possession constructions in German by L1 Spanish learners

Descriptive Statistics: Features

Project	genitive_ German

Elementary Level

Feature	Percent	N
POS-TYPE	***N=120***	
target_like_use	5.8%	7
non_target_like_use	94.2%	113
TARGET_LIKE_USE-TYPE	N=7	
-s-possessor-construction	100.0%	7
genitive-case	0.0%	0
NON_TARGET_LIKE_USE-TYPE	N=113	
underuse	14.2%	16
misuse	85.8%	97
overuse-(snoc)	0.0%	0
UNDERUSE-TYPE	N=16	
blank	100.0%	16
MISUSE-TYPE	N=97	



misselection	89.7%	87
misrealisation	2.1%	2
wrong-unclassified	8.2%	8
MISSELECTION-TYPE	N=87	
pp	40.2%	35
possessive--s	59.8%	52
MISREALISATION-TYPE	N=2	
's	100.0%	2
wrong-genitive-form	0.0%	0

SCORE: [SOC ÷ (OC+SNOC)] × 100

SOC = underuse (0 points)
misuse (0,5 points)
target_like_use (1 point)

SOC_Pos_elem_German: (16×0) + (97× 0,5) + (7× 1) =55,5

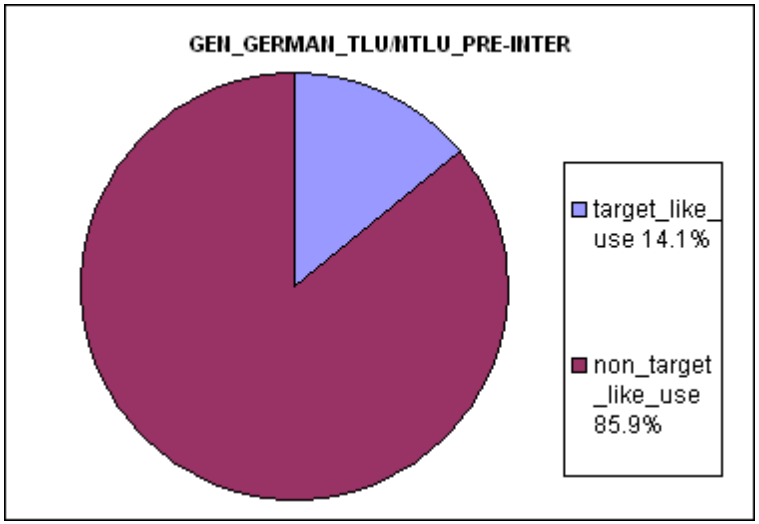
OC_POS_elem_German: 120

SNOC_POS_elem_German 0

Score_POS_elem_G: 55,5 ÷ (120+0) × 100=46,25%

Pre-intermediate Level

Feature	Percent	N
POS-TYPE	***N=78***	



target_like_use	14.1%	11
non_target_like_use	85.9%	67
TARGET_LIKE_USE-TYPE	N=11	
-s-possessor-costruction	54.5%	6
genitive-case	45.5%	5
NON_TARGET_LIKE_USE-TYPE	N=67	
underuse	16.4%	11
misuse	83.6%	56
overuse-(snoc)	0.0%	0
UNDERUSE-TYPE	N=11	
blank	100.0%	11
MISUSE-TYPE	N=56	
misselection	83.9%	47
misrealisation	16.1%	9
wrong-unclassified	0.0%	0
MISSELECTION-TYPE	N=47	
pp	59.6%	28

	possessive--s	40.4%	19
MISREALISATION-TYPE	N=9		
	's	44.4%	4
	wrong-genitive-form	55.6%	5

SCORE: [SOC ÷ (OC+SNOC)] × 100

SOC = underuse (0 points)
misuse (0,5 points)
target_like_use (1 point)

SOC_Pos_pre-inter_German: (11×0) + (56× 0,5) + (11× 1) =39

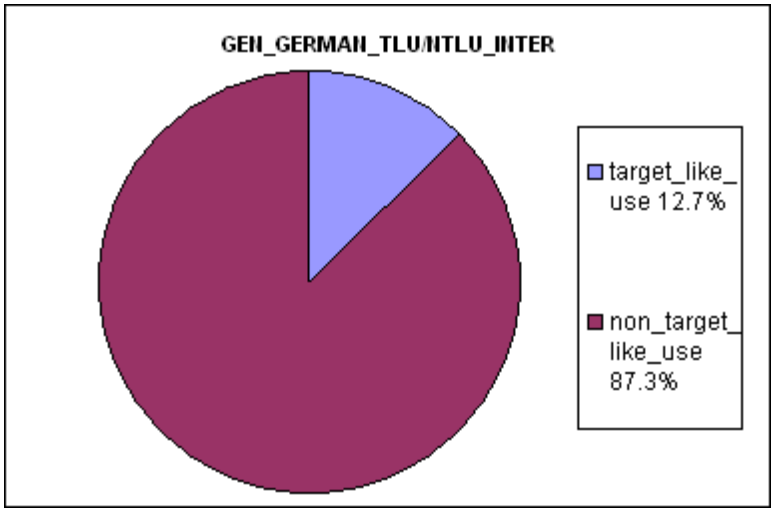
OC_POS_pre-inter_German: 78

SNOC_POS_pre-inter_German 0

Score_POS_pre-inter_G: 39 ÷ (78+0) × 100=50%

Intermediate Level

Feature	Percent	N
POS-TYPE	***N=71***	



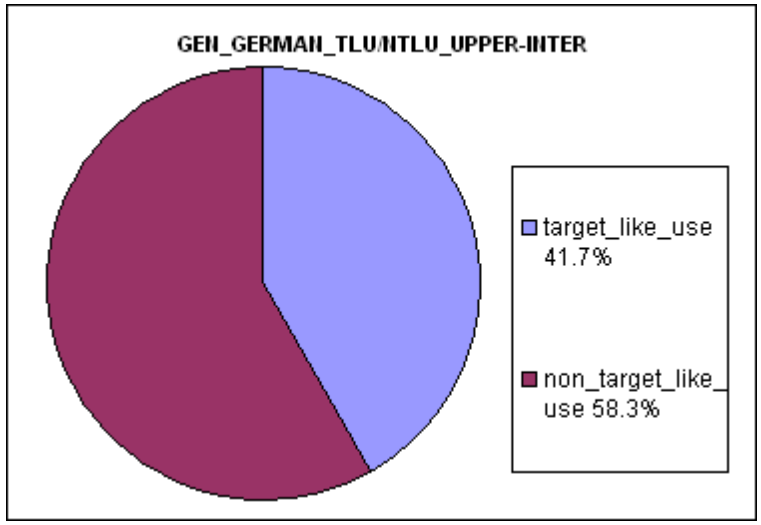
target_like_use	12.7%	9
non_target_like_use	87.3%	62
TARGET_LIKE_USE-TYPE	N=9	
-s-possessor-costruction	44.4%	4
genitive-case	55.6%	5
NON_TARGET_LIKE_USE-TYPE	N=62	
underuse	14.5%	9
misuse	85.5%	53
overuse-(snoc)	0.0%	0
UNDERUSE-TYPE	N=9	
blank	100.0%	9
MISUSE-TYPE	N=53	
misselection	96.2%	51
misrealisation	1.9%	1
wrong-unclassified	1.9%	1
MISSELECTION-TYPE	N=51	
pp	70.6%	36

```
             possessive--s                29.4%   15
MISREALISATION-TYPE                  N=1
                         's               0.0%    0
       wrong-genitive-form              100.0%    1

SCORE: [SOC ÷ (OC+SNOC)] × 100
                     SOC = underuse (0 points)
                           misuse (0,5 points)
                           target_like_use (1
                           point)
SOC_Pos_inter_German:      (9×0) + (53× 0,5) + (9× 1) =35,5
OC_POS_inter_German:                          71
SNOC_POS_inter_German                          0
                            35,5 ÷ (71+0) ×
Score_POS_inter_G:         100=50%
```

Upper-intermediate Level

Feature	Percent	N
POS-TYPE	***N=24***	

target_like_use	41.7%	10
non_target_like_use	58.3%	14
TARGET_LIKE_USE-TYPE	N=10	
-s-possessor-costruction	10.0%	1
genitive-case	90.0%	9
NON_TARGET_LIKE_USE-TYPE	N=14	
underuse	0.0%	0
misuse	100.0%	14
overuse-(snoc)	0.0%	0
UNDERUSE-TYPE	N=0	
blank	0.0%	0
MISUSE-TYPE	N=14	
misselection	50.0%	7
misrealisation	28.6%	4
wrong-unclassified	21.4%	3
MISSELECTION-TYPE	N=7	
pp	71.4%	5



```
             possessive--s                28.6%    2
MISREALISATION-TYPE                  N=4
                        's                25.0%    1
       wrong-genitive-form                75.0%    3

SCORE: [SOC ÷ (OC+SNOC)] × 100
                     SOC = underuse (0 points)
                           misuse (0,5 points)
                           target_like_use (1
                           point)
SOC_Pos_upper-
inter_German:              (0×0) + (14× 0,5) + (10× 1) =17
OC_POS_upper-
inter_German:                                24
SNOC_POS_upper-
inter_German                                  0
                            17 ÷ (24+0) ×
Score_POS_upper-inter_G:    100=70,83%
```

9.11.3.2 The relative frequency of the various possession constructions in German by L1 Spanish and L2 English learners

Descriptive Statistics: Features

Project: Genitive German_L1 Spanish_L2 English

Elementary Level

Feature	Percent	N
POS-TYPE	***N=114***	
target_like_use	5.3%	6
non_target_like_use	94.7%	108
TARGET_LIKE_USE-TYPE	N=6	
-s-possessor-construction	100.0%	6
genitive-case	0.0%	0
NON_TARGET_LIKE_USE-TYPE	N=108	
underuse	14.8%	16
misuse	85.2%	92
overuse-(snoc)	0.0%	0
UNDERUSE-TYPE	N=16	
blank	100.0%	16
MISUSE-TYPE	N=92	

misselection	89.1%	82
misrealisation	2.2%	2
wrong-unclassified	8.7%	8
MISSELECTION-TYPE	N=82	
pp	39.0%	32
possessive--s	61.0%	50
MISREALISATION-TYPE	N=2	
's	100.0%	2
wrong-genitive-form	0.0%	0

Pre-intermediate Level

Feature	Percent	N
POS-TYPE	***N=66***	
target_like_use	15.2%	10
non_target_like_use	84.8%	56
TARGET_LIKE_USE-TYPE	N=10	
-s-possessor-costruction	50.0%	5
genitive-case	50.0%	5
NON_TARGET_LIKE_USE-TYPE	N=56	
underuse	16.1%	9
misuse	83.9%	47
overuse-(snoc)	0.0%	0

UNDERUSE-TYPE	N=9	
blank	100.0%	9
MISUSE-TYPE	N=47	
misselection	80.9%	38
misrealisation	19.1%	9
wrong-unclassified	0.0%	0
MISSELECTION-TYPE	N=38	
pp	50.0%	19
possessive--s	50.0%	19
MISREALISATION-TYPE	N=9	
's	44.4%	4
wrong-genitive-form	55.6%	5

Intermediate Level

Feature	Percent	N
POS-TYPE	***N=48***	
target_like_use	14.6%	7
non_target_like_use	85.4%	41
TARGET_LIKE_USE-TYPE	N=7	
-s-possessor-costruction	42.9%	3
genitive-case	57.1%	4
NON_TARGET_LIKE_USE-TYPE	N=41	
underuse	22.0%	9

misuse	78.0%	32
overuse-(snoc)	0.0%	0
UNDERUSE-TYPE	N=9	
blank	100.0%	9
MISUSE-TYPE	N=32	
misselection	93.8%	30
misrealisation	3.1%	1
wrong-unclassified	3.1%	1
MISSELECTION-TYPE	N=30	
pp	56.7%	17
possessive--s	43.3%	13
MISREALISATION-TYPE	N=1	
's	0.0%	0
wrong-genitive-form	100.0%	1