ATLANTIS
Journal of the Spanish Association of Anglo-American Studies
42.2 (December 2020): 50-71
e-ISSN 1989-6840
DOI: http://doi.org/10.28914/Atlantis-2020-42.2.03

Linguistic Complexity across Two Early Modern English Scientific Text Types

Jesús Romero-Barranco
Universidad de Granada
jesusromero@ugr.es

In linguistics the concept of complexity has been analysed from various perspectives, among them language typology and the speech/writing distinction. Within intralinguistic studies, certain key linguistic features associated with reduced or increased complexity have been identified. These features occur in different patterns across various registers and their frequency is an indicator of the level of complexity of different kinds of texts. The concept of complexity has not, to date, been evaluated in early English medical writing, especially in terms of different text types. Thus, the present article analyses linguistic complexity in two Early Modern English medical texts, a surgical treatise (ff. 34r-73v) and a collection of medical recipes (ff. 74r-121v) housed as MS Hunter 135 in Glasgow University Library. Since they represent two different types of medical text, they can be productively compared in terms of linguistic complexity. The results obtained confirm that the surgical treatise is more complex than the collection of medical recipes owing to the higher presence of linguistic features denoting increased complexity in the former and of those indicating reduced linguistic complexity in the latter.

Keywords: linguistic complexity; Early Modern English; scientific writing; scientific text types; manuscript studies

. . .

La complejidad lingüística en dos tipos de texto científico de inglés moderno temprano

El concepto de complejidad ha sido estudiado en lingüística desde diferentes perspectivas, entre ellas la tipología del lenguaje y la distinción lenguaje oral/escrito. Dentro de los estudios intralingüísticos, se han identificado algunos rasgos asociados a menor o mayor complejidad lingüística. Estos rasgos característicos presentan diferentes niveles de ocurrencia según el registro y el análisis de su frecuencia permite evaluar la complejidad lingüística en diferentes tipos de texto. El concepto de complejidad no ha sido evaluado hasta ahora en el inglés médico en el período moderno temprano, especialmente teniendo en cuenta diferentes tipos de texto. Así, el presente artículo analiza la complejidad lingüística en dos textos médicos escritos en inglés moderno temprano, un tratado de cirugía (ff. 34r-73v) y una colección de recetas médicas (ff. 74r-121v) alojados como MS Hunter 135 en la biblioteca de la Universidad de Glasgow. Al representar dos tipos de texto médico, estos dos textos pueden ser comparados de forma productiva en cuanto a su complejidad lingüística. Los resultados obtenidos confirman que el tratado de cirugía es más complejo que la colección de recetas médicas dada la frecuencia superior en el primero de rasgos lingüísticos que indican mayor complejidad frente a la frecuencia superior en la segunda de rasgos lingüísticos relacionados con una menor complejidad.

Palabras clave: complejidad lingüística; inglés moderno temprano; escritura científica; tipos de textos científicos; estudios de manuscrito

1. Introduction

The concept of complexity was defined by Nicholas Rescher as "a matter of the number and variety of an item's constituent elements and of the elaborateness of their interrelational structure, be it organizational or operational" (1998, 1), whereas Östen Dahl argued that complexity depends on the length of the account needed to describe the system (2004, 21). In linguistics, complexity refers to the fact that language is composed of small units that, when combined, form hierarchies where different governing levels can be observed. These linguistic building blocks are morphemes, words, phrases and clauses/sentences, the latter combined through mechanisms such as coordination or subordination, which give rise to hierarchical relationships among the elements.

Linguistic complexity has been hotly debated over the last century, albeit with a general consensus on its invariance "since all languages have about equally complex jobs to do, and what is not done morphologically has to be done syntactically" (Hockett 1958, 180-81; see also Sampson 2009, 2). However, this assumption has been recently criticised, since numerous studies have demonstrated that there are conditioning factors that increase or decrease the level of complexity. These arguments suggest that, even though a crosslinguistic assessment of linguistic complexity may be problematic, a variety of methods can be applied to carry out intralinguistic studies, that is, studies focused on different aspects within the same language—historical periods, genres, text types, clausal constituents, etc. In this way, the linguist can not only measure the levels of linguistic complexity in different genres or text types, but also the diachronic evolution of one particular genre or text type within the history of a language, to name but one of this method's many applications.

There have been studies trying to identify specific linguistic features that would enable different levels of complexity in a given text to be assessed. Among the synchronic approaches to the topic, David Crystal and Derek Davy's work (1969) was groundbreaking inasmuch as they not only identified the distinctive features of some English text types—i.e., the language of conversation, unscripted commentary, religion, newspaper reporting and legal documents—but also outlined a series of text types for future research—i.e., television and press advertising, public speaking, written instructions, civil service documents, spoken legal language, broadcast talks and news and scientific literature. They analysed the different text types at various levels: phonetic/graphetic, phonological/graphological, grammatical, lexical and semantic (Crystal and Davy 1969, 15). This enabled, for instance, the identification of inexplicitness, randomness of subject matter and normal nonfluency as characteristic elements of the language of conversations, and of the use of vocatives and simple sentence structure as characteristics of religious language.

Crystal and Davy's work served as the basis for Douglas Biber's analysis (1992), grounded on his own earlier work (1985, 1988), where a macroscopic analysis of different registers was carried. For this purpose, Biber examined the distribution of

67 linguistic features across 481 texts belonging to 23 distinct spoken and written registers and, by means of exploratory factor analysis, identified five dimensions of variation across registers in English (1988, 170-98).[1] Subsequently, Biber labelled 33 out of those 67 linguistic features as potential markers of linguistic complexity, some of them indicating reduced complexity and others pointing to increased complexity (1992). What is being measured by way of the frequency of these linguistic features is the product rather than the process, that is, the representation of the surface structure of language. In other words, linguistic features such as *that*-deletions, contractions or the use of the anaphoric pronoun *it* imply a less complex surface structure but, at the same time, these utterances might require a greater effort on the part of the listener/reader to decipher the message.

A different approach is proposed by Javier Pérez-Guerra, who assessed the concept of linguistic complexity by measuring the size, syntactic density, syntactic depth and syntactic efficiency of various constituents within the sentence (2007). Focusing on unmarked subjects in academic texts, newspapers and spoken language drawn from the *British National Corpus* (2007), Pérez-Guerra noted a higher level of linguistic complexity in spoken language due to the larger proportion of pronominal subjects and the overall length of nonpronominal subjects in academic texts and newspapers. Taking a diachronic point of view, Pérez-Guerra and Ana E. Martínez-Insua studied pre- and postverbal nominal subjects and nominal adverbials in news and letters taken from *A Representative Corpus of Historical English Registers* compiled by Biber et al. (1993-2013), and concluded that letters are less complex than news inasmuch as they feature more pronominal subjects and the average length of all the syntactic units is shorter (2010). Moreover, they demonstrated that there is no significant variation in this phenomenon over the time span covered by the study (1750-present). In like manner, Anu Lehto carried out a diachronic analysis to ascertain whether the level of complexity in legal writing increased over the Early Modern English period. To this end, she analysed six linguistic features—text structure, sentence length and punctuation, clausal coordination, phrasal coordination, subordination and lexical bundles. Among her findings was the fact that the use of punctuation is quite consistent in the data—even when comparing the fifteenth and eighteenth centuries—and that documents show a considerable degree of structural organisation (2015, 135). Furthermore, the results showed that the connector *and*, the most widely used in the period, declined towards the end of the Early Modern English period, a fact that coincides with an increase in the use of punctuation (170).

The studies above demonstrate that the concept of linguistic complexity can be dealt with in different ways by focusing on a variety of linguistic aspects. However, early English scientific writing, key to the standardisation of English in early

[1] A linguistic dimension is determined on the basis of a consistent co-occurrence pattern among features (Biber 1988, 13).

Modern Britain (Taavitsainen and Pahta 2011), has been somewhat disregarded in the literature. In this article, complexity is understood as a relative concept, that is, a text type will be more or less complex on the basis of its comparison with another text type. This is in line with Fred Karlsson's view that the term *complexity* can be sporadically employed to refer to "the whole parameter 'complex-intermediate-simple,'" implying that complexity will not be fully understood if the other end of the parameter—*simplicity*—is not taken into account (2014, 145-46). This article employs an intralinguistic approach to assess the level of linguistic complexity in two texts contained in MS Hunter 135 (H135), a surgical treatise (ff. 34r-73v) and a collection of medical recipes (ff. 74r-121v) written in the sixteenth century. To this effect, section 2 explains the methodology used; sections 3 and 4 provide a microscopic analysis where the frequency of Biber's indicators of complexity (1992) is compared in the two texts; and section 5 provides the conclusions.

2. Source and Method

The source of evidence is H135, an Early Modern English scientific volume that contains a version of Guy de Chauliac's surgery treatise with interpolations by Henry de Mondeville and others (19,348 words) and a collection of medical recipes (19,482 words) (Voigts 1995, 261). In the former, the author explains how to provide the patient with medicines to heal their injury, apart from a description of the surgical operation itself, whereas in the latter a list of remedies for different maladies is supplied.[2] The selection of this material is appropriate for the objectives of the present article since it contains two texts belonging to different traditions within Early Modern English scientific writing, which will allow for the detection of textual variation. The analysis of these texts will ultimately contribute to a better understanding of Early Modern English scientific writing; further, the determination of the linguistic complexity of the surgical treatise—created and read by learned people—and the collection of medical recipes—created and read by lay people—may shed new light on the readership of the time.

The collection and analysis of the data was carried out in three consecutive stages. First, H135 was transcribed following semidiplomatic conventions.[3] High-resolution images of the manuscript were used and, in order to decipher the script in some damaged folios, the original witness was examined in situ at the University of Glasgow Library.

[2] For a given medical condition, a remedy is provided according to the following structure: title, ingredients, preparation, application (use and dosage) and statement of efficacy (Hunt 1990, 16-24; Alonso-Almeida and Cabrera-Abreu 2002, 138; Romero-Barranco 2017, 340-43).

[3] According to David C. Greetham, a diplomatic transcription "dispenses with any attempt at such scrupulous fidelity to appearance, and concentrates primarily on the textual content of the original, reproducing the exact spelling, punctuation and capitalization" (1992, 347). The transcription used in the present article is labelled *semi-diplomatic* because, even though original spelling, punctuation and capitalisation are, abbreviations are expanded and superior letters and insertions are included in the body of the text.

Second, the Early Modern English spelling of the transcribed text was normalised to Present-Day English allowing us to POS-tag the text so that automatic searches and linguistic information retrieval could be conducted. After the retrieval process, the results were normalised to tokens per 1,000 words for the sake of comparison.[4]

3. Linguistic Features Associated with Reduced Complexity

Six linguistic features belonging to three discourse functions were identified by Biber as markers of reduced complexity: structural reduction, less specified reference and fragmented structure (1992, 140). It must be noted that some of the linguistic features mentioned by him were not found in the material studied, either because they are not characteristic of these text types or because they were used rarely, or not at all, in the Early Modern English period. Consequently, only those features related with less specified reference and fragmented structure were retrieved. Chi-square was calculated and the distribution of these two linguistic dimensions was found to be not statistically significant (Cantos-Gómez 2013, 75-80), meaning that they are independent variables.

3.1. Less Specified Reference

The linguistic features related to less specified reference are pro-verb *do* (1), pronoun *it* (2) and demonstrative pronouns (3) (Biber 1992, 140):

(1) make a greate shorte tente wett with the same and put in the nose do so euery daye (surgical treatise, f. 43r)

(2) when it beginnithe to cole in the water (surgical treatise, f. 35r)

(3) menge thes alltogether and make therof a powder (surgical treatise, f. 35r)

Pro-forms (1) are expressions used for "recapitulating or anticipating the content of a neighbouring expression, often with the effect of reducing grammatical complexity," thus facilitating sentence connection (Quirk et al. 1985, 76). In the case of *do*, it originated in Middle English, when it could be used to replace a lexical verb used in the preceding clause (Fischer 1992, 268). Research on register variation has demonstrated that these structures are commonly used in face-to-face communication where speakers share the situational context and are able to clarify the message immediately (Biber et al. 1999, 432). As shown in the example, pro-verb *do* is used to refer to a whole previous explanation, allowing the reader to save time in processing the intended message.

[4] The text was transcribed with CLAWS (Rayson et al. 2007, n.p.). The automatic searches were made in *AntConc (Version 3.4.4)* (Anthony 2014).

The use of the pronoun *it* (2) is the second linguistic indicator associated with less specified reference. Third person pronouns are characteristically used in anaphoric expressions, while the other pronouns are generally used deictically (Lass 1999a, 147-48; Huddleston and Pullum 2002, 1468). Wallace L. Chafe and Jack Danielewicz associate this lack of referential explicitness to the spoken register, as speakers usually have limited time to produce utterances, resulting in the use of neuter pronouns such as *it*, which increase the vagueness of the text but at the same time boost production (1986, 90; see also Biber 1986).

The third linguistic indicator is the use of demonstrative pronouns (3), which, according to Biber, can "refer to an entity outside the text, an exophoric referent, or to a previous referent in the text" (1989, 226; see also Huddleston and Pullum 2002, 1504-09), that is, they may be used deictically or anaphorically. In the texts analysed in the present study, only anaphoric uses are witnessed since they are written texts.

Table 1 shows the distribution of these three linguistic indicators of less specified reference. As can be observed, they are more widely attested in the medical recipes than in the surgical treatise. This is understandable given the specific features of each text type. Even though surgical treatises were becoming more and more independent of the classical texts on which they were based—empiricism versus scholasticism[5]—they still preserved the structural conventions of that text type. Recipes, on the other hand, were read and produced by lay people, and thus featured language resembling the spoken register. This is supported by the wider distribution of the pro-verb *do*—even though the number of instances is very low—the pronoun *it* and demonstrative pronouns in the medical recipes—n.f. 0.6, 40.3 and 4.9, respectively— than in the treatise—n.f. 0.4, 29.8 and 2.8, respectively.[6]

Table 1. Less specified reference in H135

Linguistic features	Surgical treatise		Medical recipes	
	Raw	n.f.	Raw	n.f.
Pro-verb *do*	7	0.4	11	0.6
Pronoun *it*	572	29.8	784	40.3
Demonstrative pronouns	54	2.8	95	4.9
Total	633	33	890	45.8

[5] In early English scientific writing, the Early Modern period marks the transition "from scholasticism to empiricism, that is, from the medieval scholastic science that relied on classical authors such as Galen or Hippocrates, to new ways of constructing knowledge that were based on observation and cognition" (Romero-Barranco 2017, 2; see also Taavitsainen 2002, 204).

[6] *n.f.* stands for "normalised frequencies."

3.2. Fragmented Structure

The level of fragmented structure is quantified by means of the frequency of clausal coordination in a text (Biber 1992, 140), that is, syntactic arrangement by parataxis, where equal clauses with the same syntactic role are combined (Quirk et al. 1985, 918; Biber et al. 2007, 79). Chafe identifies fragmentation as the agglutination of idea units without connectives, although he also argues that new idea units are frequently introduced with coordinating conjunctions, *and* being the most frequent (1982, 38).

(4) Ther be two teese, one is callyd pia matter and that is the nather must , next the braines, the other is called dura matter and it is next the pan (surgical treatise, f. 57r).

Biber et al., on their part, link coordination only with orality and subordination to literacy (1999, 144-45), a fact that can also be supported from a diachronic point of view. In this sense, Thomas Kohnen carried out an analysis of coordinators and subordinators in Middle and Early Modern English sermons, and demonstrated that coordination decreases in the Early Modern English period coinciding with a spread of subordination (2007, 294). The reasons for this pattern change lie in the fact that literacy increased in the Early Modern English period, hence the wider distribution of mechanisms related to written registers such as subordination. In addition, there is corpus evidence supporting the notion that, between the seventeenth century and today, "medical, science and legal prose developed to become highly specialised registers [evolving] towards ever more 'literate' styles" (Biber and Finegan 1997, 269).

As shown in table 2, clausal coordination is more widely distributed in the recipes than in the surgical treatise. This could be explained from the perspective of text structure, since the recipes exclusively contain instructions for the preparation of remedies, where the coordinator *and* is frequently used to link the different steps. If clausal coordination is considered to measure linguistic complexity, medical recipes are less complex.

Table 2. Fragmented structure in H135

Linguistic features	Surgical treatise		Medical recipes	
	Raw	n.f.	Raw	n.f.
Clause coordination	107	5.6	166	8.5

4. Linguistic Features Associated with Increased Complexity

There are twenty-seven linguistic features associated with complexity, and these are classified into four different groups (Biber 1992): integrated structure, lexical specificity, passive constructions and dependent clauses. With the exception of lexical specificity,

where the scores were obtained from the calculation of mean word length and the type/token ratio, the distribution of integrated structure, use of passive constructions and use of dependent clauses was found to be statistically significant (X^2 62.6348, p < 0.00001) (Cantos-Gómez 2013, 75-80).

4.1. Integrated Structure

The level of integrated structure is calculated based on the frequency of nouns, prepositions, attributive adjectives, nominalisations and phrasal coordination. These constituents indicate a "high informational focus and a relatively dense integration of information in a text" (Biber 1992, 145).[7]

The frequency of nouns in different registers was examined by Biber et al. (1999, 65-66), who demonstrated that they are by far the most frequent lexical word class, being most common in news and, to a lesser extent, academic prose (see also Huddleston and Pullum 2002, 526). This distribution is understandable since the focus in these text types is on the transmission of information. In the case of the analysis that concerns us in the present section, the frequency of nouns was quite similar in both text types, although slightly higher in the recipes. The fact that, even though surgical treatises are considered academic material, recipes—a more informal register a priori—show a higher frequency of nouns—n.f. 31.2 over 29.5—is not that surprising, given the fact that they consist of long lists of plants and substances needed to prepare remedies.

Attributive adjectives were found to have an influence on the integrated structure of the texts, as they are "highly integrative in their function," expanding and elaborating on the information presented in a text (Biber 1989, 237). Regarding their occurrence across different registers, they are found more frequently in academic prose, contrasting with their low frequency in conversation (Biber et al. 1999, 65). As shown in table 3, the occurrence of adjectives is lower than that of nouns. The reasons for this could lie in the fact that, even though both text types contain a high proportion of nouns, the information presented is rarely expanded on by means of adjectival qualification or description, so linguistic complexity is not increased in this sense.

Table 3. Integrated structure in H135

Linguistic features	Surgical treatise		Medical recipes	
	Raw	n.f.	Raw	n.f.
Nouns	4,126	214.9	4,344	223.5
Prepositions	3,182	165.7	2,006	103.2

[7] Discourse with informational purposes has been associated with carefully planned utterances (Biber 1988, 79-97).

Linguistic features	Surgical treatise		Medical recipes	
	Raw	n.f.	Raw	n.f.
Attributive adjectives	567	29.5	607	31.2
Nominalisations	121	6.3	86	4.4
Phrasal coordination	1,186	61.8	823	42.4
Total	9,182	478.2	7,866	404.7

Prepositions were also used as part of the measurement of the level of integrated structure of the texts, and they were found to occur more widely in the surgical treatise—n.f. 165.7. This indicates that the treatise presents a higher level of linguistic complexity, which is unsurprising since Early Modern English surgical treatises often combine elements belonging to the learned tradition with the experiences of surgeons. This leads to a text type full of descriptions and advice, together with instructions for the preparation of remedies, and the use of prepositions certainly helps organise this kind of information.

Nominalisations have traditionally been considered a distinctive feature of scientific writing, where they allow for the packaging of information and the expansion of idea units (Biber 1988, 227; Banks 2001, 2003, 2005; Bello 2016). As M. A. K. Halliday points out, "the device of nominalizing, far from being an arbitrary or ritualistic feature, is an essential resource for constructing scientific discourse" (1988, 169). Nominalisations serve three purposes, two of them grammatical and one semantic. On grammatical grounds, nominalisations allow for the addition of modifiers and quantifiers to the nominalised process as well as for the use of the nominalised process as subject, complement, etc. On semantic grounds, the process becomes more thing-like after the nominalisation (Banks 2005, 350). In H135, nominalisations occur more widely in the surgical treatise than in the collection of recipes—n.f. 6.3 and 4.4, respectively— demonstrating that the information is better integrated and packaged in the former and, consequently, its level of linguistic complexity is higher.

Phrasal coordination is the last feature associated with integrated structure, serving to expand idea units, as can be seen in (5) (Biber 1988, 245; see also Chafe 1982, 1985). In this respect, the surgical treatise is again more complex than the collection of recipes, as phrasal coordination is more widely attested—n.f. 61.8 over 42.4, respectively.

(5) the pen may thrust downe the fleshe vpon the neld and thow must thrust boldly and hard for thow shalt perceve it to require a greate thrust (surgical treatise, f. 38r)

To sum up, this section has evaluated five different linguistic features identified as markers of integrated structure and different tendencies have been identified (table 3). First, the distribution of nouns and attributive adjectives is similar in

both text types, even though the surgical text, a more academic type of writing, was expected to show a wider distribution of these items. However, the long list of ingredients contained in many of the medical recipes plausibly accounts for the similarity in the distribution of nouns and attributive adjectives in the two text types. Second, nominalisations, prepositions and phrasal coordination are witnessed overwhelmingly more frequently in the surgical treatise, evincing the more elaborate structure of this text type.

4.2. Lexical Specificity

Word length and type/token ratio are indicators of lexical specificity, signalling "potential and actual lexical variety" respectively (Finegan and Biber 2001, 258). From a register perspective, high levels of lexical variety are usually found in academic writing, while they are more restricted in the spoken domain due to the time requirements of online production (Biber 1988, 238; see also Biber 1986; Chafe and Denielewicz 1986). Table 4 shows the scores for mean word length and type/token ratio in the texts under study. As can be seen, potential lexical variety—word length—is operating roughly at the same level in the surgical text and the collection of recipes—4.06 and 3.96 respectively. However, when it comes to actual lexical variety—type/token ratio—the collection of recipes shows a higher frequency when compared to the surgical treatise—11.2 and 9.5 respectively. This could be explained by considering the specific characteristics of each text type; recipes feature more lexical variety as they include long lists of ingredients that are necessary for the preparation of remedies.

Table 4. Lexical specificity in H135[8]

Linguistic features	Surgical treatise	Medical recipes
Mean word length[9]	4.06	3.96
Type/token ratio	9.5	11.2

4.3. Passive Constructions

Passive constructions, where the subject is associated with a patient role and receives the action of the verb, are mechanisms that allow for information packaging. In this kind of construction, the subject is "demoted or dropped altogether, resulting in a static, more abstract presentation of information" (Biber 1988, 228; Huddleston and

[8] The figures in this table are not normalised, as they are individual counts for each text under study.

[9] The calculations of word length were carried out by means of WordSmith 7 (Scott 2017).

Pullum 2002, 1365; Toyota 2005, 319), thereby shifting the focus from the subject to the object. This has traditionally been identified as characteristic of the academic and scientific registers, together with the use of static verbs and impersonal constructions (Dorgeloh 2005, 85; see also Atkinson 1996, 340-46; Biber et al. 1999, 476). Passive structures, then, contribute to linguistic complexity inasmuch as they are the surface complex representation of a simpler counterpart with the same meaning. In H135, instances of agentless passives (6) and *by*-passives (7) were found:

(6) then ax the patient how he stode whenn he was hurte that thow may take a sercher (surgical treatise, f. 38v)

(7) SOMETYME The heedes of the cheke bones are out of their ionte which is knowne by thes tokens (surgical treatise, f. 44r)

As shown in table 5, agentless and *by*-passives are more widespread in the surgical treatise—n.f. 15.6 versus 5.9 and 0.6 versus 0.1 respectively. This demonstrates that the former presents a more complex information structure than the latter, since passives require more processing effort on the part of the reader.[10]

TABLE 5. Passive constructions in H135

Linguistic features	Surgical treatise		Medical recipes	
	Raw	n.f.	Raw	n.f.
Agentless passives	299	15.6	115	5.9
by-passives	11	0.6	2	0.1
Total	310	16.2	117	6

4.4. Dependent Clauses

Higher frequency of embedded or dependent clauses also contributes to increased linguistic complexity. Depending on their nature, they may accomplish various functions in discourse and, therefore, have been subdivided into: structural elaboration on reference, complement clauses, attitudinal clauses, adverbial clauses and participial clauses (Biber 1992). No instances of attitudinal clauses were found in my data.

[10] According to Knut Lambrecht, "information-structure analysis is centered on the comparison of semantically equivalent but formally and pragmatically divergent sentence-pairs, such as active vs. passive" (1994, 6; see also Seoane 2012).

4.4.1. Structural Elaboration on Reference

Structural elaboration on reference is measured by means of the frequency of different instances of relatives (table 6). Relatives are finite postmodifying clauses that allow for the addition of new information about the antecedent (Biber et al. 1999, 195; Hundt et al. 2012, 210). According to Biber and Susan Conrad, these structures function to provide information and they are "complex syntactic constructions, difficult to produce in real-time situations, but well-suited to the focused informational purposes of textbooks" (2009, 67).[11]

Table 6. Structural elaboration on reference in H135

Linguistic features	Surgical treatise		Medical recipes	
	Raw	n.f.	Raw	n.f
Pied-piping relative clauses	11	0.6	4	0.2
that relative clauses in subject position	23	1.2	33	1.7
that relative clauses in object position	40	2.1	26	1.3
Total	74	3.9	63	3.2

Three different constructions were found in our data: pied-piping relative clauses, *that* relative clauses in subject position and *that* relative clauses in object position. Pied-piping relative clauses (8) are constructions where the linking relative appears together with a preposition. In my data, the frequency of these constructions in the surgical treatise is three times higher than in the collection of recipes—n.f. 0.6 and 0.2 respectively—meaning that the former features a more information-oriented structure and, therefore, a higher level of syntactic complexity.

(8) put of this oyle euery daye warme ones into the sonne eer, on whiche syde the waxe kyrnells bredithe (surgical treatise, f. 51r)

The occurrence of relative clauses in subject (9) and object (10) position also contributes to increasing the level of syntactic complexity. In my data, *that* relative clauses in subject position show a slightly higher occurrence in the medical recipes as opposed to the surgical treatise—n.f. 1.7 and 1.2 respectively. The distribution of *that* relative clauses in object position, in contrast, is more widely attested in the surgical treatise than in the medical recipes—n.f. 2.1 and 1.3 respectively.

(9) when pia matter is hurte thowe shalte se all the foresaide tokens with those that folow (surgical treatise, f. 34r)

[11] Chafe (1982, 1985) also identified these structures as devices for integration and idea unit expansion.

(10) then shalt thou first cut the vttermost skyne on crose wise and the hole gobbet that thow fyndest therin (surgical treatise, f. f. 41r)

In spite of the low frequency of these three linguistic features across the two texts, the quantitative analysis evidences a wider distribution of *that* relative clauses in subject position in the medical recipes. However, the sum of the three linguistic features demonstrates that the surgical treatise is linguistically more complex with regard to structural elaboration on reference.

4.4.2. Complement Clauses

Complement clauses are "a type of dependent clause used to complete the meaning relationship of an associated verb or adjective in a higher clause" (Biber et al. 1999, 658). There are various kinds of complement clauses, each with a different purpose: *wh*-clauses can express an indirect question (e.g., *Jill was asking what happened*) or a relative clause (e.g., *Burbidge road is where Carlos used to live*); *that*-clauses are employed to report the speech, thoughts, attitudes or emotions of humans (e.g., *I think that Stuart's gone a bit mad*); and infinitive clauses can report speech and cognitive states (e.g., *I'm just trying to get away early*) (Biber et al. 1999, 657-97).[12]

(11) then shalt thow Vnderstand that the wound is appostemid (surgical treatise, ff. 58r-58v)

(12) And thowe may take a goose pen beyng open at the end to thrust again the neld poynt vpon the skyne syde (surgical treatise, f. 37v)

As for the functions of these clauses at the discourse level and their influence on linguistic complexity, they have been identified as mechanisms that allow for integration and idea unit expansion (Biber 1988, 231; Chafe 1982, 1985). These devices are more prone to be witnessed in planned written registers. As shown in table 7, the instances of these complexity features in the surgical treatise outnumber those in the collection of recipes.

(13) And if it be so depe that the larde cannot reche the Bottome Then take a tent of lynnen clothe and anoynt it (surgical treatise, f. 38v)

[12] No instances of interrogative or relative clauses were found in the corpus studied.

Table 7. Complement clauses in H135

Linguistic features	Surgical treatise		Medical recipes	
	Raw	n.f.	Raw	n.f
that complement clauses to verbs	12	0.6	3	0.2
that complement clauses to adjectives	9	0.5	5	0.3
Infinitives	384	20	332	17.1[2]
Total	405	21.1	340	17.6

4.4.3. Adverbial Clauses

Adverbial clauses also add information to main clauses, either as adjuncts or disjuncts. While the former denote the circumstances of the situation in the main clause, the latter give information about the style or form of what is said in the main clause. There are various subclasses of adverbial clauses, including condition, reason/cause and concession (Quirk et al. 1985, 1070-118). These three kinds of adverbial clause, together with a general category encompassing all the other types, were identified by Biber (1988, 235-36) as potential indicators of linguistic complexity.

In medical writing, adverbial clauses of condition may describe a physical state and a specific treatment (14) or the quantity of medicine produced if a series of steps are followed. These clauses are more frequently found in the collection of recipes—n.f. 16.2 over 14.9 (table 8). The reason for this distribution is that *if*-clauses are one of the devices used for the introduction of recipes, in which a condition is described and then the treatment is explained:

(14) IF THE Throte be wounded and the wesand or throte bole partid (surgical treatise, f. 47r)

Causative clauses express "how one event or state is contingent upon another" (Biber et al. 1999, 779). In H135, such clauses are introduced by causative *for* (15) or *because*, and are used to add relevant information related to the surgical operation or the recipe being described. Table 8 shows that their occurrence in the surgical treatise is almost double that in the recipe collection—n.f. 2.7 and 1.6, respectively. These clauses constitute one more reason to think that empiricism was being consolidated in medical practice and that each practitioner would include all information they considered relevant for the correct accomplishment of the instructions.

[13] This figure does not include infinitive clauses appearing in the titles.

(15) Then I give the counsell not to meddle therwithe For it lyethe not in mans cure to heale yt (surgical treatise, f. 67r)

Table 8. Adverbial clauses in H135

Linguistic features	Surgical treatise		Medical recipes	
	Raw	n.f.	Raw	n.f
Conditional adverbial subordination	286	14.9	315	16.2
Causative adverbial subordination	52	2.7	32	1.6
Concessive adverbial subordination	-	-	10	0.5
Other adverbial subordination	71	3.7	23	1.2
Total	409	21.3	380	19.5

Concessive clauses "indicate that the situation in the matrix clause is contrary to expectation in the light of what is said in the concessive clause" (Quirk et al. 1985, 1098; Biber et al. 1999, 779). These clauses occur infrequently in H135, and are in fact only found in the collection of recipes. In (16), the concessive clause is used to assure that the treatment will be efficient irrespective of the duration of the condition:

(16) for him that hathe lost his sight <u>althoughe it be by the space of. x. yeres</u> (medical recipes, f. 76r)

Finally, all other adverbial clauses, comprising those introduced by other adverbial subordinators such as *while*, *whilst*, *whereby*, *so that* and *as long as* (17), were combined in the category "other adverbials" (Biber 1988, 236). As Table 8 shows, the frequency of this group of adverbials is three times higher in the surgical treatise—n.f. 3.7 over 1.2—thereby demonstrating that it is more complex than the collection of recipes.

(17) IF THE breakinge of the pane be greate and the wounde aboue is straite <u>so that</u> thow cannot be certeyne of the Quantitie of breking (surgical treatise, f. 35r)

The occurrence of adverbial clauses in H135 allows four conclusions to be drawn. First, adverbial subordination as a whole is more common in the surgical treatise—n.f. 21.3 compared to 19.5. Second, conditional clauses are used to a similar extent in both text types—n.f. 14.9 and 16.2 in the surgical treatise and the collection of recipes respectively. Third, causative adverbial subordination is more widely attested in the surgical treatise—n.f. 2.7 vs 1.6. Fourth, the occurrence of concessive adverbial subordination is negligible—n.f. 0 and 0.5 in the surgical treatise and the collection of recipes respectively.

4.4.4. Participial Clauses

According to Biber, participial clauses these clauses are found more frequently in writing, where they are used for integration and structural elaboration (1988, 233). In H135, only present participial postnominal clauses and present and past participial adverbial clauses were found.

Table 9. Participial clauses in H135

Linguistic features	Surgical treatise		Medical recipes	
	Raw	n.f.	Raw	n.f.
Present participial postnominal clauses	48	2.5	16	0.8
Present participial adverbial clauses	17	0.9	41	2.1
Past participial adverbial clauses	28	1.5	66	3.4
Total	93	4.9	123	6.3

The distribution of present participial postnominal clauses in H135 (18) is three times higher in the surgical treatise than in the recipes—n.f. 2.5 vs 0.8. According to Sandra A. Thompson, these clauses are used for depictive functions, that is, for the elaboration of descriptions by means of the creation of mental images, hence indicating higher linguistic complexity (1983, 51).

(18) make yt to boyle strewing in all the powder of litarge ouer (surgical treatise, f. 49r)

Present and past participial adverbial clauses—(19) and (20)—function as reduced relative clauses and have been identified as devices used for producing "highly informational discourse under severe time constraints" (Biber 1988, 233; see also Janda 1985, 447). In H135, the occurrence of these constructions is twice as frequent in the recipes than in the surgical treatise—n.f. 2.1 and 3.4 and 0.9 and 1.5 respectively. This distribution is perfectly understandable, since the collection of recipes is less elaborate than the surgical treatise in terms of structure, indeed sharing many of the linguistic features of spoken discourse.

(19) strew theron a powder which is good to fret awaie cankers Being in smow placs and maid in thus manner (surgical treatise, f. 50v)

(20) haue awaie all the corruption therof with paring of some instrument maid therfore (surgical treatise, f. 71v)

5. Conclusions

This article focuses on the level of linguistic complexity of two text types within the field of Early Modern English scientific writing—a surgical treatise and a collection of medical recipes. Linguistic indicators of reduced and increased complexity were retrieved and their relative frequency in the two text types allowed us to reach the following conclusions.

The linguistic features related to reduced linguistic complexity were more widely witnessed in the collection of recipes, which was expected since this text type is more informal than the surgical treatise. In line with this, linguistic features such as the pro-verb *do*, the pronoun *it*, demonstrative pronouns and clausal coordination occur almost twice as often in the collection of recipes than in the surgical treatise. The instances of linguistic features associated with increased linguistic complexity in the surgical treatise outnumber those in the collection of recipes in some contexts, while in others the opposite occurs. This shows that even though the surgical treatise is more complex in terms of the overall count, the nature of the recipes entails the occurrence of certain features related with higher linguistic complexity. Specifically, the surgical treatise presents a higher incidence of integration of structure, passive constructions, structural elaboration on reference and complement and adverbial clauses. The collection of recipes, in turn, shows a higher occurrence of linguistic features associated with lexical specificity and the use of participial clauses. For these reasons, I conclude that the surgical treatise clearly features a higher level of linguistic complexity than the recipes. These results are consistent with the nature of the text types under study. Surgical treatises were produced and consumed by learned professionals, accustomed to this kind of literature, while collections of medical recipes were transmitted from generation to generation from the Old English period onward and were created and consumed by lay people for domestic use, hence their more colloquial nature.

The occurrence of the linguistic features indicating reduced or increased linguistic complexity not only sheds light on the levels of complexity of the text types under analysis, but it also reveals their most characteristic features as text types belonging to the same genre. Three different tendencies have been identified. First, there is a high frequency of nouns and adjectives, infinitive clauses and conditional adverbial subordination in both text types, explained by the fact that these are characteristic features of the scientific genre to which both texts belong. Second, the surgical treatise is characterised by the high frequency of prepositions, nominalisations, phrasal coordination, passives, pied-piping relative clauses, relative clauses in object position, *that* complement clauses to adjectives and verbs, causative adverbial subordination and present participial postnominal clauses. Third, medical recipes are characterised by the high frequency of *that* relative clauses in subject position and present and past participial clauses.

The results obtained and the conclusions reached in the present study pave the way for future research in the field of early English scientific writing and, more specifically,

the linguistic complexity of the different text types within this genre. In this regard, other text types—e.g., theoretical treatises, philosophical transactions, etc.—could be analysed and their diachronic evolution studied in order to better understand the genre of scientific writing in early English.

Works Cited

Alonso-Almeida, Francisco and Mercedes Cabrera Abreu. 2002. "The Formulation of Promise in Medieval English Medical Recipes: A Relevance-Theoretic Approach." *Neophilologus* 86 (1): 137-54.

Anthony, Laurence. 2014. *AntConc (Version 3.4.4)*. Tokyo: Waseda U. [Accessed online on October 10, 2020].

Atkinson, Dwight. 1996. "The Philosophical Transactions of the Royal Society of London, 1675-1975: A Sociohistorical Discourse Analysis." *Language in Society* 25 (3): 333-71.

Banks, David. 2001. "The Reification of Scientific Process: The Development of Grammatical Metaphor in Scientific Discourse." In Mayer 2001, 555-63.

—. 2003. "The Evolution of Grammatical Metaphor in Scientific Writing." In Simon-Vandenburgen, Taverniers and Ravellieds 2003, 127-47.

—. 2005. "On the Historical Origins of Nominalized Process in Scientific Text." *English for Specific Purposes* 24: 347-57.

Beadle, Richard and A. J. Piper, eds. 1995. *New Science out of Old Books: Studies in Manuscripts and Early Printed Books in Honour of A. I. Doyle*. Aldershot: Scolar Press.

Bello, Iria. 2016. "Cognitive Implications of Nominalizations in the Advancement of Scientific Discourse." *International Journal of English Studies* 16 (2): 1-23.

Biber, Douglas. 1985. "Investigating Macroscopic Textual Variation through Multifeature/Multidimensional Analyses." *Linguistics* 23 (2): 337-60.

—. 1986. "Spoken and Written Textual Dimensions in English: Resolving the Contradictory Findings." *Language* 62 (2): 384-414.

—. 1988. *Variation across Speech and Writing*. Cambridge: Cambridge UP.

—. 1989. "A Typology of English Texts." *Linguistics* 27 (1): 3-43.

—. 1992. "On the Complexity of Discourse Complexity: A Multidimensional Analysis." *Discourse Processes* 15 (2): 133-63.

Biber, Douglas and Edward Finegan. 1997. "Diachronic Relations among Speech-based and Written Registers in English." In Nevalainen 1997, 253-75.

Biber, Douglas et al. 1999. *Grammar of Spoken and Written English*. London: Longman.

Biber, Douglas, Ulla Connor and Thomas A. Upton. 2007. *Discourse on the Move: Using Corpus Analysis to Describe Discourse Structure*. Amsterdam and Philadelphia: John Benjamins.

Biber, Douglas and Susan Conrad. 2009. *Register, Genre and Style*. Cambridge: Cambridge UP.

Blake, Norman, ed. 1992. *The Cambridge History of the English Language*. Vol 2, *1066-1476*. Cambridge: Cambridge UP.

Campoy, Mari Carmen and María José Luzón, eds. 2007. *Spoken Corpora in Applied Linguistics*. Bern: Peter Lang.

Cantos-Gómez, Pascual. 2013. *Statistical Methods in Language and Linguistic Research*. Sheffield: Equinox.

Chafe, Wallace L. 1982. "Integration and Involvement in Speaking, Writing and Oral Literature." In Tannen 1982, 35-54.

—. 1985. "Linguistic Differences Produced by Differences between Speaking and Writing." In Olson, Torrance and Hildyard 1985, 105-23.

Chafe, Wallace L. and Jack Danielewicz. 1986. "Properties of Spoken and Written Language." In Horowitz and Samuels 1986, 83-116.

Crystal, David and Derek Davy. 1969. *Investigating English Style*. London: Longman.

Dahl, Östen. 2004. *The Growth and Maintenance of Linguistic Complexity*. Amsterdam and Philadelphia: John Benjamins.

Davies, Matthew et al., eds. 2007. *Proceedings of the Corpus Linguistics 2007 Conference* (CL2007). Birmingham: U of Birmingham.

Dorgeloh, Heidrun. 2005. "Patterns of Agentivity and Narrativity in Early Science Discourse." In Skaffari et al. 2005, 83-94.

Dorgeloh, Heidrun and Anja Wanner, eds. 2010. *Syntactic Variation and Genre*. Berlin and New York: De Gruyter.

Eckert, Penelope and John R. Rickford, eds. 2001. *Style and Sociolinguistic Variation*. Cambridge: Cambridge UP.

Fanego, Teresa, Belén Méndez-Naya and Elena Seoane, eds. 2002. *Sounds, Words, Texts and Change: Selected Papers from 11 ICEHL*. Vol 2. Amsterdam and Philadelphia: John Benjamins.

Finegan, Edward and Douglas Biber. 2001. "Register Variation and Social Dialect Variation." In Eckert and Rickford 2001, 235-67.

Fischer, Olga. 1992. "Syntax." In Blake 1992, 207-408.

Ghadessy, Mohsen, ed. 1988. *Registers of Written English*, *Situational Factors and Linguistic Features*. London: Pinter.

Greetham, David C. 1992. *Textual Scholarship: An Introduction*. New York and London: Garland.

Halliday, M. A. K. 1988. "On the Language of Physical Science." In Ghadessy 1988, 162-78.

Hocket, Charles F. 1958. *A Course in Modern Linguistics*. Basingstoke: Palgrave Macmillan.

Horowitz, Rosalind and S. Jay Samuels, eds. 1986. *Comprehending Oral and Written Language*. New York: Academic Press.

Huddleston, Rodney and Geoffrey K. Pullum. 2002. *The Cambridge Grammar of the English Language*. Cambridge: Cambridge UP.

Hundt, Marianne, David Denison and Gerold Schneider. 2012. "Relative Complexity in Scientific Discourse." *English Language and Linguistics* 12 (2): 209-40.

Hunt, Tony. 1990. *Popular Medicine in Thirteenth-Century England*. Cambridge: D. S. Brewer.

Janda, Richard D. 1985. "Note-Taking as a Simplified Register." *Discourse Processes* 8 (4): 437-54.

Karlsson, Fred. 2014. "Complexity in Linguistic Theorizing." *The Mental Lexicon* 9 (2): 144-69.

Klein-Andreu, Flora, ed. 1983. *Discourse Perspectives on Syntax*. New York: Academic Press.

Kohnen, Thomas. 2007. "'Connective Profiles' in the History of English Texts." In Lenker and Meurman-Solin 2007, 289-308.

Lambrecht, Knud. 1994. *Information Structure: Topic*, *Focus and the Mental Representations of Discourse Referents*. Cambridge: Cambridge UP.

Lass, Roger. 1999a. "Introduction." In Lass 1999, 1-12.

—, ed. 1999b. *The Cambridge History of the English Language*. Vol 3, *1476-1776*. Cambridge: Cambridge UP.

Lehto, Anu. 2015. "The Genre of Early Modern English Statutes: Complexity in Historical Legal Language." PhD diss., University of Helsinki.

Lenker, Ursula and Anneli Meurman-Solin, eds. 2007. *Connectives in the History of English*. Amsterdam and Philadephia: John Benjamins.

Mayer, Felix, ed. 2001. *Language for Special Purposes: Perspectives for a New Millennium*. Tübingen: Gunter Narr.

Meurman-Solin, Anneli, María José López-Couso and Bettelou Los, eds. 2012. *Information Structure and Syntactic Change in the History of English*. Oxford: Oxford UP.

Nevalainen, Terttu, ed. 1997. *To Explain the Present: Studies in the Changing English Language in Honour of Matti Rissanen*. Helsinki: Société Neophilologique.

Olson, David R., Nancy Torrance and Angela Hildyard, eds. 1985. *Literacy*, *Language and Learning: The Nature and Consequences of Reading and Writing*. Cambridge: Cambridge UP.

Pérez-Guerra, Javier. 2007. "'Am I more Complex When I Speak or When I Write?' A Corpus-Based Study on Linguistic Complexity in Spoken and Written Present-day English." In Campoy and Luzón 2007, 127-46.

Pérez-Guerra, Javier and Ana E. Martínez Insua. 2010. "Do Some Genres or Text Types Become more Complex than Others?" In Dorgeloh and Wanner 2010, 111-40.

Quirk, Randolph et al. 1985. *A Comprehensive Grammar of the English Language*. London and New York: Longman.

Rayson, Paul et al. 2007. "Tagging the Bard: Evaluating the Accuracy of a Modern POS Tagger on Early Modern English Corpora." In Davies et al. 2007, n.p.

Rescher, Nicholas. 1998. *Complexity: A Philosophical Overview*. New Brunswick, NJ and London: Transaction.

Romero-Barranco, Jesús. 2017. "Early Modern English Scientific Text Types: Edition and Assessment of Linguistic Complexity of the Texts in MS Hunter 135 (ff. 34r-121v)." PhD diss., University of Málaga.

Sampson, Geoffrey. 2009. "A Linguistic Axiom Challenged." In Sampson, Gil and Trudgill 2009, 1-18.

Sampson, Geoffrey, David Gil and Peter Trudgill, eds. 2009. *Language Complexity as an Evolving Variable*. Oxford: Oxford UP.

Scott, Michael. 2017. *WordSmith Tools (Version 7)*. Stroud: Lexical Analysis Software.

Seoane, Elena. 2012. "Givenness and Word Order: A Study of Long Passives from Early Modern English to Present-Day English." In Meurman-Solin, López-Couso and Los 2012, 139-63.

Simon-Vandenburgen, Anne-Marie, Miriam Taverniers and Louise Ravelli, eds. 2003. *Grammatical Metaphor: Views from Systemic Functional Linguistics*. Amsterdam and Philadelphia: John Benjamins.

Skaffari, Janne et al., eds. 2005. *Opening Windows on Texts and Discourses of the Past*. Amsterdam and Philadelphia: John Benjamins.

Taavitsainen, Irma. 2002. "Historical Discourse Analysis: Scientific Language and Changing Thought-styles." In Fanego, Méndez-Naya and Seoane 2002, 201-26.

Taavitsainen, Irma and Päivi Pahta, eds. 2011. *Medical Writing in Early Modern English*. Cambridge: Cambridge UP.

Tannen, Deborah, ed. 1982. *Spoken and Written Language: Exploring Orality and Literacy*. Norwood, NJ: ABLEX.

Thompson, Sandra A. 1983. "Grammar and Discourse: The English Detached Participial Clause." In Klein-Andreu 1983, 43-65.

Toyota, Junichi. 2005. "Politeness as a Distancing Device in the Passive and in Indefinite Pronouns." In Skaffari et al. 2005, 319-42.

Voigts, Linda E. 1995. "A Doctor and his Books: The Manuscripts of Roger Marchall (d. 1477)." In Beadle and Piper 1995, 249-314.

Received 15 October 2019 Revised version accepted 23 March 2020

Jesús Romero-Barranco is a member of the Department of English and German at the University of Granada. His research interests are historical linguistics, manuscript studies and corpus linguistics. He has published in journals such as *Studia Neophilologica*, *Nordic Journal of English Studies* and *English Worldwide*, among others.