An Initial Analysis of Nominalizations in Scientific Texts (1700-1900)

Introduction

The purpose of this paper is to provide an account of the current research on nominalizations I have been carrying out. Section 1 will contain an introduction serving as a justification for this research. Section 2 is concerned with the description of the corpus of data used in my research, the Coruña Corpus, A Collection of Samples for the Historical Study of English Scientific Writing (CC, henceforth). Section 3 is devoted to the study of the process of nominalization. In this section I will provide a brief description of the theories for nominalizations produced by linguistic schools and I will attempt to supply an original method for analyzing nominalizations. Finally, in section 4 I will offer hints for incoming research.

1. Justification

Some authors such as Halliday (2004) point out the importance of studying the language of science, given that science and the language of science are two indissoluble entities. The difficulty of the scientific language is not limited to the lexical level but it also applies to a range of specific grammar structures that characterize discourse. Among the difficulties that scientific English poses both to learners and professional researchers, Halliday mentions syntactic ambiguity (2004:162). Nominalization is one of the grammar structures causing a higher degree of ambiguity in scientific texts. When nominalizing, semantic information is usually left out, which increases the degree of ambiguity and the difficulty in correctly decoding the sentence. The diachronic study of scientific language is one of the main purposes of my research. This study may be framed within the scope of the history of English for Specific Purposes. More specifically, my aim is to attempt to outline the evolution in the use of the process of nominalization in scientific texts from the period 1700-1900. To pursue this objective, I will consider purely linguistic issues, as well as sociolinguistic, sociopragmatic and cognitive regards.

This study covers the period from 1700 to 1900. In this time-span, several social changes affected the methodologies and approaches employed in science. This evolution had its effects on language, too. The situation of the language at the turn of the 17th century was increasingly expanded to cover new aspects of life. At that time, English was struggling to become a respectable language for scientific use. The 17th and 18th centuries witnessed a constant dispute between empiricism and rationalism. The theory of empiricism highlighted the role of experience and sensation as opposed to the traditional process of reasoning inherited from medieval scholasticism, which relied on purely theoretical deductions derived from a set of previously established principles (Crespo, 2004). In the 18th century, the Age of reason aimed at establishing an authoritative system that would banish superstition and irrationality and propose a systemized organization for society. The coexistence of the two schools implied not only a different way of doing science but also a different language use, as Banks (2005) demonstrated in his study of two contemporaneous scientific texts written in late 17th century written by Newton, an empiricist, and Huygens, a rationalist.

Traditionally, vocabulary expansion, either through loanwords or neologisms, was considered the major linguistic change connected with the development of science. However, grammar structures are also permeable to language change. As Barber suggests, “The rise of scientific writing in English helped to establish a simple referential kind of prose as the central kind in Modern English” (1993: 214). Nominalizations serve as complex encodings of processes into
nouns and they contribute to the increase of complexity in texts. Consequently, Barber's assertion may lead to the hypothesis that the earlier the text is, the least chances we have to find either grammatical metaphors or nominalizations. Similarly, it is reasonable to expect an increase in the use of nominalizations as science advances and new scientific theories are enunciated.

2. Corpus material

The corpus material for this study has been taken from the Coruña Corpus, A Collection of Samples for the Historical Study of English Scientific Writing. The corpus has been designed to contribute to the diachronic study of English at several linguistic levels. Still under compilation, the CC will be made up of different subcorpora, according to different disciplines, such as Astronomy, Philosophy, Mathematics, Life Sciences, History, etc. For my study I have selected one discipline from the field of Exact Natural Sciences – Astronomy-, one from the field of Life Sciences and also one from the field of the Humanities – History-. My purpose is to study to what degree language use and language change are related to scientific disciplines.

The corpus contains two texts per decade and discipline. The general aim was to include one text from the beginning of each decade and one from the end. Samples contain around 10,000 words excluding figures, tables, formulae and graphs, which total 200,000 words per century. The corpus has been designed so as to fit adequate standards of representativeness and balance (Moskowich-Spiegel and Crespo, 2007: 349; Lareo and Estévez, 2008: 70; Montoya, 2008: 140). In the corpus, first editions of texts written in English by English-speaking authors have been preferred. Similarly no more than one text by the same author has been included, so as to avoid personal idiosyncrasies. The time-span of the CC, 1700-1900, was chosen according to extralinguistic considerations. Hence, the starting point of the corpus, 1700, coincides with a revolution of old epistemological patterns (Taavitsainen and Pahta, 1997). Several discoveries that took place at the end of the 19th century set a crisis in the basis of mechanical physics and serve as a good end-point for the corpus (Moskowich-Spiegel and Crespo, 2007: 348).

Together with the texts, other extralinguistic information about the authors has been included in the corpus. This information proves specially useful to study the evolution of the process of nominalization according to place, sex and genre/text type parameters. Regarding place and sex, the CC includes information about the place of education and sex of the authors. The audience, either real or intended, delimits the parameter of genre/text type. Thus, in research articles, academic treatises and essays, originally aimed at the scientific community, I expect to find a higher degree of complexity and nominalizations than in texts intended for a less learned audience, such as lectures, textbooks, dialogues and letters.

3. Nominalizations: study and typology

Both the functionalist and the generative schools have tried to produce a consistent theory of nominalization. Within the generative tradition, Grefenstette & Teufel (98) define nominalization as a process that transforms a verbal phrase into a nominal form. The resulting nominal can take the form of a gerundive or a deverbal noun and it is usually accompanied by a highly-unpredictable, semantically-emptied verb that links the nominalization with the rest of the sentence. One of the main advantages of the generative approach is that it provides two different explanations for the process of nominalization. The transformational hypothesis (Lees, 1960; Newmeyer, 1971) focuses on the dependency between the nominal and the original Verb Phrase, whereas the lexicalist hypothesis (Chomsky, 1970) states the independence of the nominalized forms and acknowledges

---

1 J.J. Thompson's discovery of the electron and Planck's announcement of quantum mechanics in 1896, Einstein's first formulation of the Theory of Relativity in 1905
the differences between two types of nominalizations: gerundives –recognisable by their – -ing suffix- and derived nominals. The tendency has been to progressively follow the lexicalist approach (Grimshaw, 1990; Siloni, 1997; Zucchi, 1993). Semantic differences are better explained if we acknowledge that derived nominals, unlike gerundives, are nouns in deep structure, not deep-structure transformations.

Functionalists also consider nominalizations as processes of transformation, but they introduce the notion of “grammatical metaphor”, which Guillén (1998: 368) defines as “the transference of the linguistic representation of the semantic components of a situation between different lexicogrammatical categories”. For functionalists, semantic components have prototypical lexicogrammatical realizations. In nominalizations –as well in other kinds of grammatical metaphors- processes, which are normally encoded as verbs, are metaphorically or non-congruently encoded as nouns. (Banks, 2005a: 78) Consequently, nominalizations are mergers of two semantic components: that of “process”, from their verbal origin, and that of “entity”, from their actual codification as nouns. Functionalists have mostly tried to describe the functions and advantages involved in the use of nominalizations (Halliday: 1985, 2004, Ventola: 1996, Guillén: 1998, Ravelli:, Banks: 2005a, 2005b). A brief account of them would include: lexical cohesion (repetition and summarization); economy, conciseness and packing of information; backgrounding of information (related to theme, rhyme and information structures); advancement of discourse and reification. The packing of information and the dynamism nominalizations add to the thematic structure of a text are perhaps the most salient features. Nominalizations “made it possible on the one hand to construct hierarchies of technical terms, and on the other hand to develop an argument step by step, using complex passages ‘packaged’ in nominal forms as Themes” (Halliday, qtd. in Guillén, 1998: 371).

Whereas the generative approach takes only into consideration the form of the nominalized items so as to explain how their relationship with regard to the Verb Phrase, functionalists are not concerned with the form, but with the implications of the process itself. Both approaches are very useful but they minimize major issues. In spite of the functionalist claims, there is some evidence that shows that nominalizations are totally independent processes. In this sense, I will argue that it is necessary to differentiate between verbal and nominal encodings of processes and that nominalizations are a completely independent way of expressing a process. Formally, this claim is supported by the fact that there is no need in applying a word-formation process to derive a noun from a verb, since there are also nouns that encode a process even if they do not have a cognate verb.

Another point of conflict may be found in the definition of nominalization as a kind of metaphor. It has been argued (Banks, 2005b; Halliday, 1985) that grammatical metaphor could be compared to semantic metaphor because both metaphors are based in the shift of either the function or meaning to create certain effects in the text. There are reasons to think that this comparison is not completely accurate. Firstly, unlike grammatical metaphors, semantic metaphors require a referent in order to be interpreted. Grammatical metaphors, on the other hand, are easily interpreted and, indeed, sometimes it is rather difficult to recognize them in texts. Additionally, whereas nominalizations are a marker of adult speech, not every adult speaker is able to generate new semantic metaphors. Further evidence may be found in the existence of dead nominalizations and in the tendency to produce automatized expressions.

Another factor that should be borne in mind is the importance of the evolution of the process across time. Halliday (2004) has pointed out that the situation of contemporary scientific English is the result of an evolution that started 400 or 500 years ago. It is perhaps more productive to think of the independence of nominalizations as a continuum: from an original state in which the link between verbs and nominalizations was preponderant to a present-day situation in which
nominalizations do have some features that mark their independence.

Connected to the notion of continuum, within the category of nominalizations, there are degrees concerning the expression of a process. Thus, between the two types of nominalizations described in the generative tradition, deverbal nouns focus on the entity while gerundives emphasize the nature of the grammatical metaphor: a noun denoting a process. In this sense, they are more process-like. This distinction is especially evident in the cases where the deverbal noun and the gerundive form are interchangeable and can occur in the same context, as in the following examples:

““The diminution of the moon's gravitation to the earth in the syzygies, and its augmentation in the quadratures, tends to flatten her orbit in the syzygies, and to lengthen it in the quadratures.””
(Ewing, 1809; emphasis added)

“The diminishing of the Appearance of a Body, removing from the Eye, lessens continually.”
(Lacy, 1779: 16; emphasis added)

The device of nominalization can, thus, be seen as part of a process. Now there may be more refined and more primitive nominalizations as far as the independence from the Verb group is concerned. There may be instances of nominalizations at different stages of the process VG/Process > NG/Process > NG/Entity, always considering that this is a continuum process rather than a clear-cut set of steps. The fact that, synchronically, some nominalizations are nearer the VG/Process stage and others approach to the NG/Entity do not interfere with the diachronic evolution of the process. Nominalizations should be considered as independent linguistic devices: the fact that a NG indicates the presence of a process does not show any deviance from regular use. This use is rather motivated by other series of parameters -sociolinguistics, cognition, thematic structure-. This consideration should not obstruct the study of nominalization as a changing process over time. Indeed, this linguistic device is key in the study of language change in scientific registers and, as Halliday asserts: “there has been a steady drift towards the nominalizing region” (2004: 175).

The main aim of this work is to study the evolution of nominalizations over a specific period of time - 17th to 19th centuries-. To pursue this goal, it is necessary to provide a theory of nominalizations. The flaws of most existing theories about nominalization do not relate to the veracity of their arguments but rather to the lack of variety in the range of parameters employed -Morphology, Syntax, information structures, sociolinguistics and Cognition-. According to this, four types of nominalizations can be distinguished. Each group has distinctive formal and functional features. Nevertheless, this classification aims at being flexible enough to acknowledge that there are central and peripheral elements.

3.1. Conditioned nominalization

In this type of nominalizations, the grammatical structure determines the use of a Noun Phrase. The author has to choose whether to keep the congruent Verb/Function realization and convolute the structure of the text or keep the structure simpler but resort to a metaphorical realization, as can be seen in the following example:

“(…) nor will the centrifugal power be sufficient to compensate the different gravitations of such an assemblage of bodies as constitute the solar system, which would come to ruin of itself, without some new regulation and adjustment of their original motions.” (1756 Ferguson; emphasis added)

The structure of this nominalization is highly unpredictable because it has to meet the requirements of the sentence in which it is included. Similarly, it is difficult to determine the reach
of semantic shift in this kind of nominalizations because any nuance in meaning the grammatical metaphor can be hindered by grammatical needs. This kind of nominalization summarizes information so as to make it fit in a context. It also serves as a simplifier of convoluted grammar structures, making the contents more easily accessible to the audience. As a result, it is expected that this kind of nominalizations will be found in lectures, textbooks, research articles and other texts in which any complexity is avoidable.

3.2. Lexical nominalization

This type of nominalization is usually found in contexts in which there is a high frequency of synonyms and words from the same lexical group in the same and preceding paragraphs, as can be seen in the following example:

“(...) then, the index remaining fixed, bring their limbs to the other wire, and if the same limbs be in contact, the axis is properly adjusted; but if they lap over, the object end of the telescope is inclined from the plane of the quadrant, and must be altered by the adjustment for that purpose; but if the limbs be separated, the object end is inclined to the quadrant, and must be adjusted accordingly, and repeat the operation till the limbs coincide at both wires, and the adjustment is made.” (1790 Vince; emphasis added)

This type of nominalization fulfills all the features described by functionalists (Banks 2001, 2005): a process that is codified into a Verb Group (congruent codification) and functions as rheme can be turned into a nominalization (grammatical metaphor) in the theme of the sentences at the end of the paragraph. This nominalization also serves as a cohesive device, repeating and summarizing information. Sometimes it is merely a matter of avoiding lexical repetition. If a verb appears multiple times, it is common that that verb may be replaced by a synonym verb or a grammatical metaphor. This last function is closely related to the following type of nominalization, the stylistic one.

3.3. Stylistic nominalization

In this type of nominalization there is no functional or formal reason in the text that helps completely explain the use of grammatical metaphor. A verbal codification would be perfectly possible without adding complexity. However, the difference with lexical nominalizations lies in the fact that the frequency of words from the same semantic family is very low. The use of a nominalization is, consequently, a free choice of the writer and it belongs to the domain of text style. The use of stylistic nominalization is clear in the following example:

“If we were to pause upon the interesting and sublime facts which even this superficial view discloses, we should be filled with an amazement at their grandeur and beauty, to which no words can do justice, because we cannot either convey or entertain any adequate conception of their magnificence.” (1817, Philips; emphasis added)

The author could have written down “we should be amazed at their grandeur and beauty” instead of resorting to the nominalization. In that case, the nuance of “being filled” would be lost unless another expression was introduced. But, for stylistic reasons he favored the expression of the nuance in a nominalization. Possible functions of stylistic nominalizations are associated to the status of the text as a product written by specific individuals -scholars- with a specific receiver in mind -the scientific community-. This type of nominalizations add complexity to the text and are more likely to be found in research articles, academic treatises and formal essays. Stylistic complexity may also be
linked to the receiver of the text, the scientific community. Ventola (1996) deals with the topic of academic writing and links it directly to the scientific community, the sociolinguistic group that receives the text. Grammar complexity can be considered a way of “guild codification”, a code that only members of the community master and that distinguishes outsiders and novices from well-established members.

3.4. Term nominalization

There are contexts in which a low amount of words from the same lexical family together with an appropriate grammatical structure could facilitate a verbal realization of a process. However, in some cases a nominalized codification is preferred. In those cases, it is the semantic component that makes the difference. If we consider the process “congruent Verb/process codification > metaphorical Noun/process codification > congruent Noun/Entity” codification as a continuum, this kind of nominalization would be somewhere between the last two stages, while the other types would be in the central stage.

This kind of nominalizations are nearer the semantic codification of entities as nouns. The nominalizations in this group can be labeled as “terms”. Terms are cognitive devices we create and use to study reality by establishing a set of differences and frontiers (Calvin 1996, Eckardt 1993, Lakoff, 1990, Thagard, 1996). They are specially useful in scientific disciplines because they provide semantic traces of entity to both processes and entities. The extensive use of terminology is, in fact, one of the defining features of the scientific register in any language. Banks refers to this phenomenon as reification (Banks 2005b: 349). Another piece of evidence that helps distinguish this kind of nominalizations is the use of semantically-emptied verbs that form a collocation together with the nominalization, as in the following example:

“THE Earth which we inhabit, is one of thoſe Planets which moves round the Sun, performing its Revolution in the Space of a Year, or 365 Days, 5 Hours, and 49 Minutes, (...)” (1726 Gordon; emphasis added)

The motivation for this nominalization may be that the turning of a verb into a noun facilitates readers the process of identifying the processes and events that are being subject to study – it is easier to identify that we are studying the REVOLUTION of the Earth than that we are studying the fact that the Earth revolves around the Sun. Cognitively, this process can be similar to the one employed when providing indexes at the end of a book to facilitate quick searches or when giving a title to a book or a chapter2; It is related to a cognitive process linked not only to the reification of science but also to the organization of information in our minds.

4. Further research

Currently, I am in the process of counting and analyzing the first discipline of my study - Astronomy. Then, I will also include the other two disciplines - Life sciences and History. My intention is to further develop my theory of nominalizations so as to be able to apply it to the results of corpus exploitations. For this purpose, I am now working on the development of a typology of nominalizations based on formal and functional issues. I am specially interested in the analysis of nominalizations from the point of view of cognitive science. I believe nominalizations are very productive linguistic devices intervening in the construction and advancement of discourse and, thus, they are closely related to the organization of information in our minds and other cognitive considerations. Apart from cognition, I am including sociolinguistic issues - sex and origin of writers, intended audience - in order to study how these parameters influence language use and

2 In fact, there are also multiple examples of nominalizations within titles, as in “[SECT]. IV. Containing a Description of the CELESTIAL GLOBE.” (Gordon 1726; emphasis added)
language change.

**Corpus references**


Ferguson, James (1756). *Astronomy explained upon Sir Isaac Newton's principles and made easie to those who have not studied mathematics* (146-167). London: J. Ferguson.

Gordon, George (1726). *An introduction to geography, astronomy, and dialling. Containing the most useful elements of the said sciences, adapted to the meanest capacity, by the description and uses of the terrestrial and celestial globes. With an introduction to chronology* (pp. 63-99, 101-123). London: J. Senex.

Lacy, John (1779). *The universal system: or mechanical cause of all the appearances and movements of the visible heavens* (1-33). London: J. Buckland.


**References**


