Persuasion in English Philosophy Texts
(*CEPhiT*)[*]

Begoña Crespo[**], Isabel Moskowich[***]

Abstract:
The aim of this paper is to offer a description of the *Corpus of English Philosophy Texts (CEPhiT)* as well as to present a pilot study on persuasion strategies. Although this corpus contains samples from the eighteenth and nineteenth centuries, only eighteenth-century texts have been selected for this study. Methodologically speaking, some specific linguistic features indicating persuasion and argumentation (Biber, 1988) have been searched for: predictive modals, necessity modals, conditional subordinators and verbs with a suasive meaning. The interpretation of our findings will provide an overview of the author-reader relationship in late Modern English Philosophy writings, especially focusing on variables such as sex or genre.

Keywords: scientific English, corpus linguistics, persuasion strategies, late Modern period, philosophy writing, authorial presence

1. Introduction

Since every scientific field has its own writing traditions and restrictions, we have decided to compile different sub-corpora forming the *Coruña Corpus of English Scientific Writing* (*CC*). Each of them contains samples of texts published between 1700 and 1900 which correspond to a different scientific discipline. Overlapping of disciplines constitutes a basic difficulty in the selection of representative samples of scientific language, mainly when it is not present-day science we are dealing with. Instead of designing our own taxonomy of disciplines when compiling the *CC*, we resorted to the one published by UNESCO (1988) as a starting point. The first sub-corpus compiled was *CETA*, *Corpus of English Texts on Astronomy*. The second is the one we present here, *CEPhiT*, the *Corpus of English Philosophy Texts*.

[**] Professor PhD, Department of English Philology, Universidade da Coruña (Spain), bcrespo@udc.es
[***] Professor PhD, Department of English Philology, Universidade da Coruña (Spain), bcrespo@udc.es

2. Compilation Principles of *CEPhiT*

The compilation principles applied to *CEPhiT* are those applied to the whole *CC*. Therefore, we have tried to compile two 10,000 words text files per decade, so that each of the centuries represented contains approximately 200,000 words. Some pilot studies with our corpus have shown that 1,000-word samples are not really enough for the study of variation within the scientific register (Biber 1993) mainly because the scientific register was not as standardised at that time as it is nowadays. This corpus shares the structure and mark-up conventions used for the whole project which have proved to be extremely useful and valid for research since the sampling methods avoid idiosyncrasies and interference caused by translation.

We have been also careful to keep the principles of representativeness and balance (McEnery and Wilson 1996; Biber et al. 1998: 251–253). We have included only edited and printed prose texts. As with the other sub-corpora, first editions have been used whenever possible. When not, and assuming that language change can be observed within 30-year periods (Kytö et al. 2000: 92), those published within less than thirty years from the date the work was first published were chosen.

In order to have a complete representation of stylistic and pragmatic devices, we have collected extracts from different parts of the works sampled so that introductions, central chapters and conclusions are more or less equally represented. In *CEPhiT,* prefaces or dedications which are not scientific in their content have been excluded. With all this, we have obtained the word counts shown in Table 1.

Table 1. *Words in CEPhiT*	
Eighteenth century	200,022
Nineteenth century	201,107
Total	401,129

Some extra-linguistic factors such as age, sex, place of education and genre/text-type of each author and text are part of the information in the metadata files accompanying text files (Crespo and Moskowich 2010).

The general principles of compilation applied to the *CC* can be consulted in Moskowich and Crespo 2007; Moskowich and Parapar 2008; and Crespo and Moskowich 2010.

3. Time-Span Represented

Extra-linguistic considerations also determine the time-span of the *Coruña Corpus* and, therefore, of *CEPhiT. W*e have used landmarks in scientific thought rather than those in language change to set the time limits of our selection. Let us not forget that changes in scientific thought imply changes in scientific discourse (Moskowich 2011). *CEPhiT* has been compiled to cover the Modern English period.

The time-span chosen begins with the outburst of the scientific revolution, the foundation of the Royal Society and with the publication of the basic guidelines on how to present scientific works to its members with the ideas of clarity and simplicity behind it all. *CEPhiT* earliest texts date back to 1700 (Mary Astell) and 1705 (George Cheyne), a moment at which the old epistemological patterns of Scholasticism are suffering a radical transformation (Taavitsainen and Pahta 1997) and, therefore, a moment we considered ideal to start our compilation. This starting point coincides also with the new inductive method that John Stuart Mill (included in *CEPhiT*) systematised in 1845.

Several events which were really important for the History of Science occurred around 1900, the last year of the period covered by *CEPhiT*. Some of these events were the discovery of the electron by J.J. Thompson in 1896, the crisis of the grounds of mechanical physics announced in this same year, Planck's proposal of quantum mechanics, or Einstein's publication of the Special Theory of Relativity in 1905 (Moskowich and Crespo 2010; Moskowich 2011). All these discoveries, as in the seventeenth century, were also accompanied by the need to change the discursive patterns of science announced by Thomas Huxley at the 1897 International Congress of Mathematics.

In the following paragraphs, all the extra-linguistic variables in the corpus will be presented.

4. Genres/Text Types

Contrary to what one may think, variation can be found within academic writing subject, among other things, to text type (the internal characteristics of texts) and to genre (as a way of socialising and, therefore, with certain external functions) (García-Izquierdo and Montalt 2002).

The classification we have used in the *CC* is based on Görlach (2004). All the categories proposed by this author were already used during the Modern Period. However, not all sub-corpora in the *CC* have the same genres or texts-types represented. In fact, in *CEPhiT* we have a lesser number of genres than in other disciplines. Table 2 below represents the number of samples compiled belonging to each genre:

Table 2. *Genres in Philosophy Texts*

Genres in *CEPhiT*	Samples
Treatise	22
Essay	10
Textbook	1
Lecture	5
Dialogue	1
Article	1

The ascription of texts to genres may be arguable (Fowler 1982), but we have examined very carefully both the whole texts from which samples had been extracted and their prefaces. This allowed us to conclude that *CEPhiT* contains samples of the six genres/text-types in the table above. In turn, this may be due to restrictions imposed by subject-matter: certain disciplines or domains seem to prefer just a few types of texts whereas others manifest themselves in a more varied way (Moskowich, 2011).

Modern authors writing about Philosophy seem to prefer treatise by large (we only have an example of text-book) as table 2 above shows. Essays come next, which points to a real liking for more formal genres. Other categories are also represented: the informative function is the most common one, but the instructive and entertaining functions are found here too in the shape of Lecture, Dialogue and Article.

Figure 1 displays the different genres gathered in all *CEPhiT* samples where 54% corresponds to treatise.
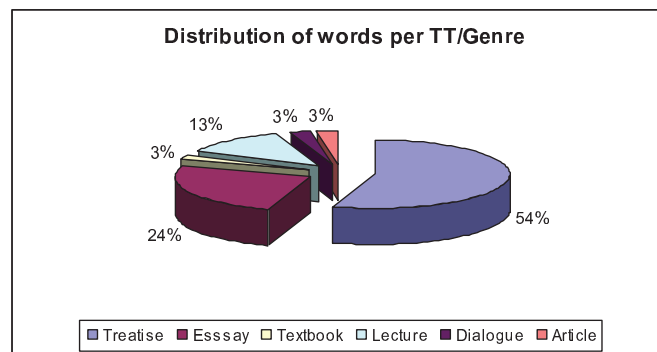

Figure 1. *Proportion of words per genre*

However, such distribution is not identical in the two centuries compiled. The graphs below show these differences reflecting the external reality which influenced text production in the field.
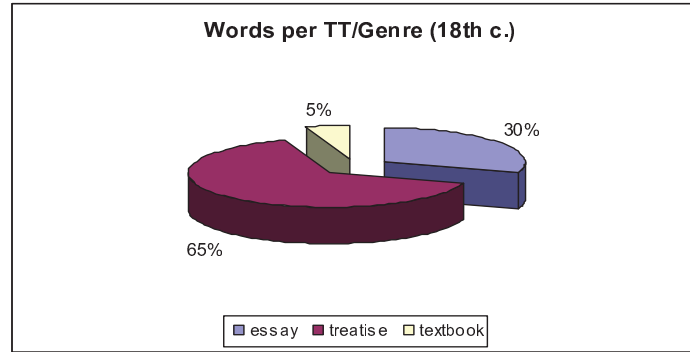

Words per TT/Genre (18th c.)

5%
30%
65%

☐ essay ■ treatise ☐ textbook

Figure 2. *Words per genre in 18ᵗʰ c. Philosophy texts*

Both Figures 2 and 3 illustrate a wider variety of genres used in the nineteenth century as compared with those used by authors in the preceding century. The fact that Philosophy was considered to deserve dissemination at different social and cultural levels may have caused this.
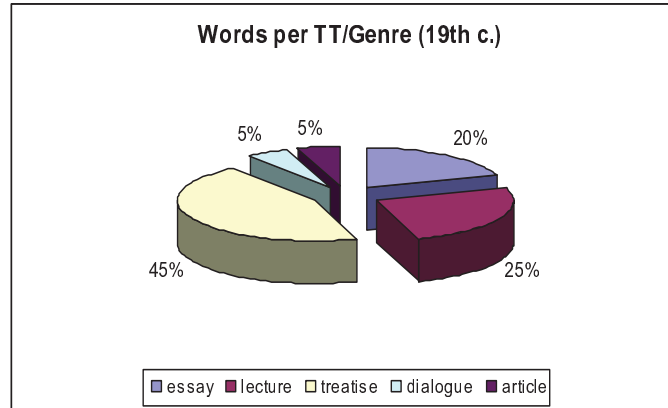

Words per TT/Genre (19th c.)

5% 5% 20%
45%
25%

☐ essay ■ lecture ☐ treatise ☐ dialogue ■ article

Figure 3. *Proportion of words per genre in 19ᵗʰ c. Philosophy texts*.

The information contained in *CEPhiT* metadata files suggests that Philosophy opens itself to a larger readership from 1800 and does so by resorting to a wider range of genres.

5. Sex

Not many works written by women can be regarded as philosophical texts in the period under survey. *CEPhiT* contains only three samples of female writing representing 8% of all the words in it (see Figure 4). These women are Mary Astell (1700), Catharine Macaulay (1783) and Mary Wollstonecraft (1792).
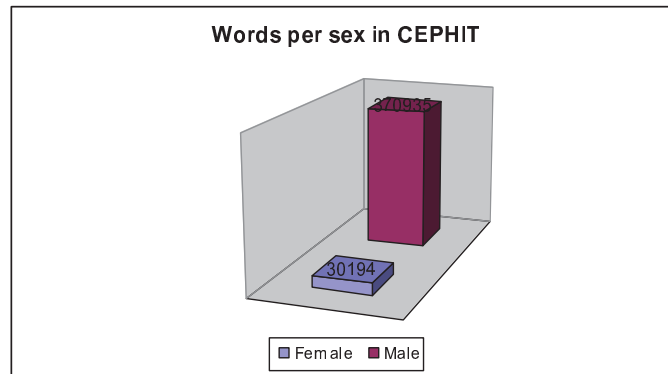

Words per sex in CEPHIT

370935

30194

☐ Female ☐ Male

Figure 4. *Words written by male and female authors*

CEPhiT reflects this scarcity of overt female activity. In fact, no women writing philosophy in the nineteenth century have been included in it; therefore the 30,194 words of female writing were produced in the period prior to the beginning of the suffragist movement.

6. Mapping *CEPhiT*

It has been already mentioned that the corpus is valid not only for the diachronic study of English scientific writing but also for that of variation depending on other variables such as geographical origin. This is why we have resorted to texts by authors whose linguistic habits could be traced.

In compliance with the *CC* principles, we have selected English-speaking authors writing in English, avoiding any sort of translation. When referring to "geographical distribution of authors" we are not considering the places where they were born but, instead, those where they received formal education, and where they acquired the linguistic habits to be found in their writings.

No American authors have been included in this sub-corpus though they abound in other parts of *CC*. It was Europe that was producing most works on philosophy, whereas North America had lived a convulsive eighteenth century and was, in the nineteenth, more worried about the practical application of scientific advances than about

metaphysical ones. In this sense, *CEPhiT* is a small-scale mirror of reality.

An overview of the different places where the authors contained in *CEPhiT* learned to write is offered in Figure 5.

Geographical distribution per
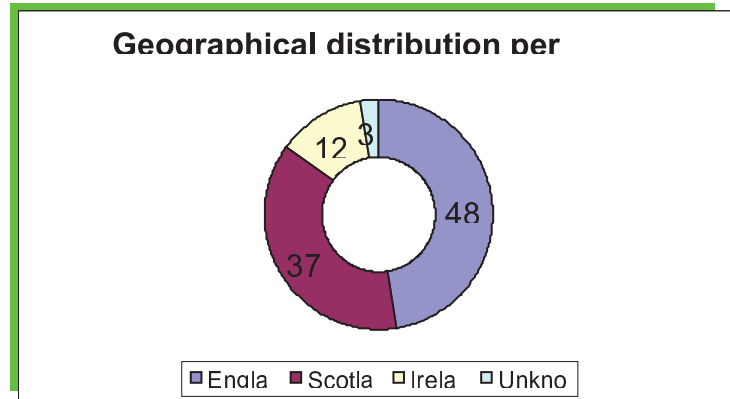
12 3

48

37

☐ Engla ☐ Scotla ☐ Irela ☐ Unkno

Figure 5. *The provenance of authors in CEPhiT*

Figures 6 and 7 below reflect the way in which the distribution per centuries is not exactly the same as the overall one:
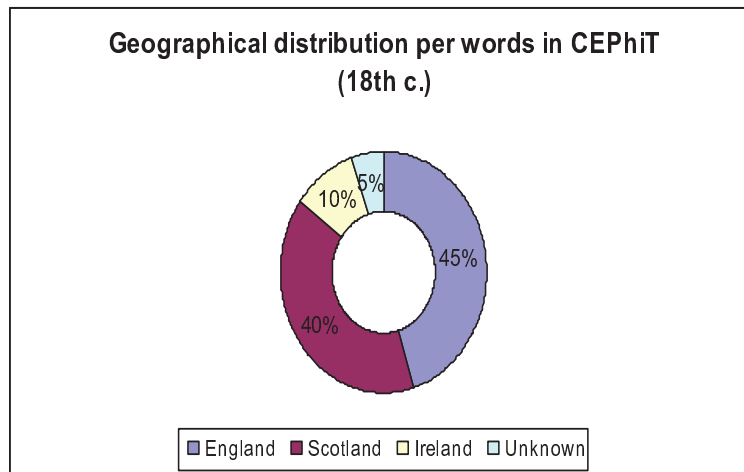
Geographical distribution per words in CEPhiT (18th c.)

5%

10%

45%

40%

☐ England ☐ Scotland ☐ Ireland ☐ Unknown

Figure 6. *Geographical distribution in the eighteenth century*

Geographical distribution per words in CEPhiT (19th c.)

England 50%
Scotland 35%
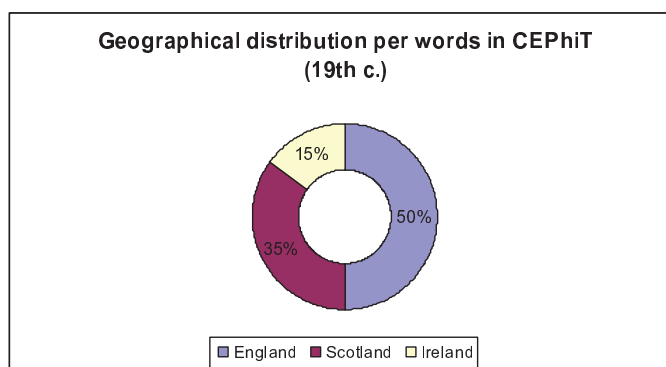Ireland 15%

England ▪ Scotland □ Ireland

Figure 7. *Geographical distribution in the 19th century*

The outstanding presence of Scottish authors in both centuries should be noted in contrast to the low presence of Irish authors. Social and political changes have a deep impact on the development of language. The way in which *CEPhiT* has been sampled reflects such social and political shifts. For instance, the fact that during the eighteenth century, Ireland lived the Protestant Ascendancy implied that the native Irish population was excluded from power and public life (Claydon and McBride 1999).

In the following pages we will explore to what an extent some of these authors use persuasive strategies in their writings.

7. A Study on Persuasion Strategies

In scientific register authors had to argue in defense of their opinions, findings and positions, and although no overt manifestation of authorial presence was permitted by the dominant canon of Bacon and Boyle's style, they could resort to other subtle means to persuade their readership. Texts that seek to encourage a certain attitude in the reader are not only audience-focused but also well-structured writings, using rhetorical mechanisms of persuasion and argumentation in a more or less explicit way. This coincides with the concept and characteristics of authorial self mentioned by Ivanić (1998) and Hyland (2002) and which is analyzed for the description of present-day academic writing.

The discourse of scientific authors was one of tentative power with regard to their target audience: other authors and learners. Logic and deduction constituted the grounds of the scientific method as well as determined the manifestation of persuasion. According to Crespo (2011: 202):

> Logic implies to appeal to a person's sense of reason because the speaker or writer coaches an opinion in such a way that it, rather than someone else's, seems to be

the most rational solution. Deduction implies the use of a logical, well-justified method which clearly demonstrates how elements in nature, or claims about nature, are built up. It runs from the most general to the most specific, highlighting principles, shared assumptions, values and beliefs [...] What the author wants to accomplish in his/her use of language, then, is a change in the perspectives of the audience.

The interpretation of a bundle of linguistic features as markers of persuasion and argumentation was first considered by Biber in his 1988 work. These features were later analyzed by other authors in their studies on genres and registers (Atkinson 1999; Mischke 2006; Nesi 2009; Włodarczyk 2010, etc).

Atkinson (1999: 123–125), for instance, considers some of the features Biber (1988) claims to index the expression of persuasion: suasive verbs such as *suggest* or *promise*, prediction modals (*will, shall, would*), necessity modals (*ought to, should, must*), conditional subordination *(if, unless)* and split auxiliaries (They *are* objectively known). He reveals that in the articles of the Philosophical Transactions of the Royal Society between 1675 and 1975 there is an increasing tendency towards the use of non-persuasive markers. This trend may be viewed in connection with the change from author-centered to object-centered kind of prose and to the increasingly more abstract nature of texts as they approach the last quarter of the twentieth century.

7.1. Describing Methodology and Corpus of Data

To examine persuasion strategies in 18th-century Philosophy texts we have resorted to some of the linguistic features included in Biber's Dimension 4 "Overt expression of persuasion" (1988): conditional subordinators, persuasive verbs, predictive and necessity modals. In Biber (1995: 161), he argues that in present-day English professional letters and editorials show many more persuasive elements than press reviews or broadcasts. However, official documents and academic prose occupy an intermediate position between the two aforementioned registers. This "neutrality" in the presence of persuasive strategies may be explained in opposition to the degree of abstractness and the high objectivity of late twentieth century academic writing (Biber 1995: 165). Biber's conclusions raise another issue: the type of genre employed by the author might be reckoned as an interesting variable to examine the persuasive style of eighteenth century men of science.

We have carried out a quantitative study of these features for which both absolute as well as normalized figures to 10,000, when necessary, will be offered. After the general findings, we will consider the sex

variable: female persuasive strategies will be compared with overall figures and figures corresponding to male writings.

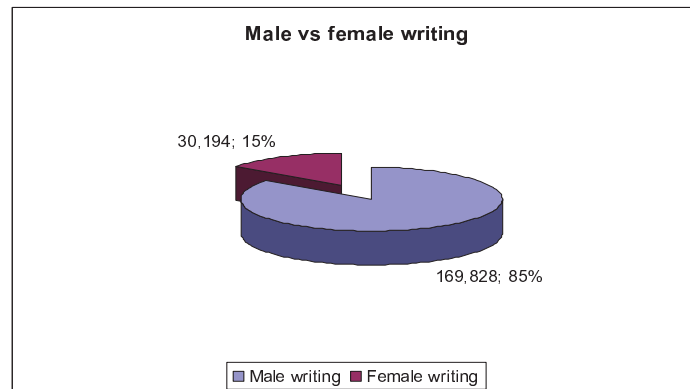Figure 8 below shows the number of words corresponding to the sex variable.

Male vs female writing

30,194; 15%

169,828; 85%

☐ Male writing ■ Female writing

Figure 8. *Male vs. female writing*

In what follows all the data retrieved will be studied in detail.

7.2. Analysis of data

In a total of 200,022 words we have found 3,553 tokens (1.77%) with persuasive implications, which are distributed as shown in Figure 9:

Suasive
16

Predic
modals:

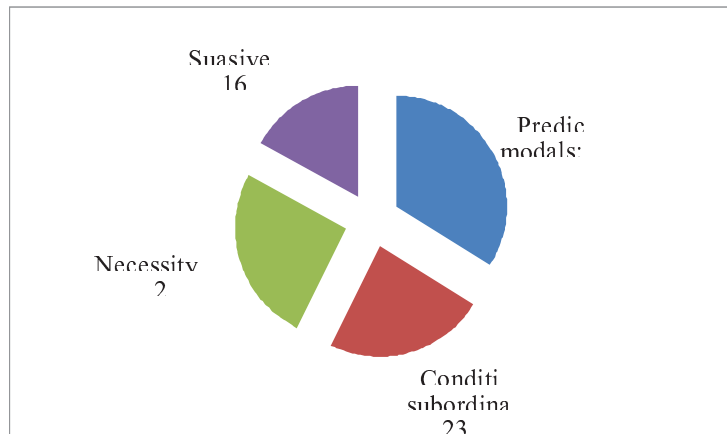Necessity
2

Conditi
subordina
23

Figure 9. *Linguistic parameters: general figures*

Predictive modals exhibit the highest proportion of tokens indicating persuasion or argumentation. This is a content-based strategy to transmit

information from the writer's point of view and to make the reading public process that information in the same way. It is the author's deliberate evaluation of the topics that is perceived by the reader, and this is done to reassure the truthfulness of the author's opinions. It can be seen clearly in example (1):

> (1)   one judgment on evident propoſitions it will follow that men **will** be ſo far irrational and by conſequence imperfeſt agents as… (Collins, 1717: 63)

The same strategy of trying to influence the reader's perception of the message applies also to the remaining set of modals, those indicating necessity, which occupy the second position as in (2):

> (2)   depend on circumſtances peculiar to any age or nation but **muſt** be the reſult of human nature or the ſuggeſtion of… (Ferguson, 1769: 123)

Conditional subordination, with the emphasis on argumentation, illustrates how some structural elements of a complex syntax play a role in the expository prose in which principles and ideas are rightly and overtly expressed. It is the third linguistic feature in order of frequency:

> (3)   that the agent himſelf is the cauſe is saying nothing **unleſs** it can be proved that he chooſes one action rather… (Crombie, 1793: 18)

The last feature refers to suasive verbs, which might be seen as the most obvious linguistic manifestation of persuasion. Nevertheless, they represent only 16.85% of all the features which have been taken into consideration, as in (4):

> (4)   … his deſires and appetites were well rul'd and did not **move** themſelves but by the comandment of reaſon and this reaſon… (Dunton, 1710: 331).

It appears, therefore, that persuasion was not as overtly or frequently present in 18th century philosophy texts as might have been thought. Rather, a more covert persuasion, one which favours a rhetorical form of argumentation, seems to emerge.

In the following sections we will present the analysis of the variable 'sex'.

7.2.1. The sex variable. From a total of 200,022 words under survey, only 30,194 belong to female authors. Although we are conscious of the fact that the findings obtained from these data cannot be generalized due to the low number of words, this scarcity of female production as

contained in our corpus reflects the social reality of the period. Only 15.09% of the material included in the eighteenth century section of *CEPhiT*, then, was written by women (see Section 5). The female authors included are the following:

Table 3. *Number of words per female writer*

Author	Date	Words
Astell	1700	10,077
Macaulay	1783	10,059
Wollstonecraft	1792	10,058
	Total	30,194

From these 30,194 words we have found 396 tokens (1.31%) which are indicative of persuasion, as seen in Figure 10:
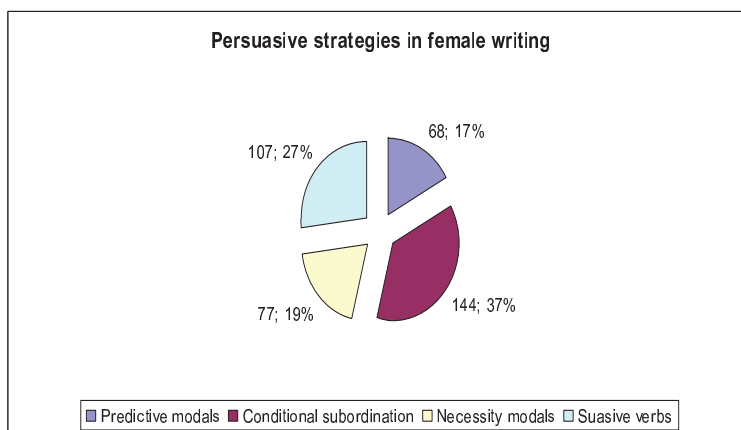

Figure 10. *Strategies of persuasion in female works*

If we compare these results with the findings from works written by men, amounting to 169,828 words, we find 3,157 tokens, that is, 18.58 instances per 10,000 words. Normalized figures (see Table 4) also demonstrate that the occurrences of persuasive strategies in female writing amount to 13.11 per 10,000 words.

Table 4. *Male vs. female persuasive strategies*

	Words	Tokens	nf/10,000
Male	169,828	3,157	18.58
Female	30,194	396	13.11

This overall presentation of the data extracted from *CEPhiT* confirms the tendency for the use of more persuasive strategies on the part of male authors than on the part of female ones in relation to raw numbers. Normalized figures confirm this male tendency to persuasion in scientific writing which surpasses the female actual usage of these strategies. Still, we can go into further detail and compare each individual feature so as to ascertain whether any more specific tendencies in the use of these strategies are present:

Table 5. *Comparison of each linguistic strategy*

Features	Male strategies	nf/10,000	Female strategies	nf/10,000
Predictive modals	1,140	67.12	68	22.52
Conditional subordination	678	39.92	144	47.69
Necessity modals	847	49.87	77	25.50
Suasive verbs	492	28.97	107	35.43

As can be seen from Table 5, female writers are more likely to include complex syntax of the subordinating kind, that is, conditional subordination with contingent and hypothetical clauses as well as a higher number of suasive verbs, which brings about the overt manifestation of authorial presence. In contrast, male preferences revolve around modality devices, mainly strong modality, which seems to modulate the writer's presence in their works.

8. Conclusions

From a general standpoint, the predominant persuasive or argumentative strategies are predictive and necessity modals. This suggests that modality dominates scientific discourse in eighteenth century Philosophy texts. Less frequently used are conditional subordination devices and, even less, suasive forms.

In this respect, and in contrast to what might generally be expected, suasive verbs are the least represented linguistic feature, which may corroborate the assumption that persuasion or argumentation, if any, is not as overtly represented as we might anticipate. This could lead us to think that the tendency is to present any kind of persuasive mechanism in a more subtle and covert manner.

It is also noticeable that the study of the sex variable sheds some light on the different linguistic uses of men and women in the eighteenth century. More strategies of persuasion have been recorded in male (18.58/10,000) than female (13.11/10,000) writing in relation to the overall number of words analyzed. Moreover, female writing has been attested to employ more conditional subordination and suasive verbs that male writing where modality of both predictive and necessity meanings are the prominent linguistic devices.

In sum, we can say that eighteenth century scientific discourse, as attested in the discipline of Philosophy, makes a moderate use of persuasive strategies, especially of visible strategies. Nevertheless, this does not mean that it is altogether uniform. There are gender differences which seem to point to opposite views on how to treat and convey scientific information.

REFERENCES:

Atkinson, Dwight, *Scientific Discourse in Sociohistorical Context: The Philosophical Transactions of the Royal Society of London, 1675–1975*, Mahwah, NJ, Lawrence Erlbaum, 1999.

Bhatia, Vijay, *Analysing genre: Language use in professional settings*, London, Longman, 1993.

Biber, Douglas, *Variation across speech and writing*, Cambridge, UK, Cambridge University Press, 1988.

Biber, Douglas, *Representativeness in corpus design*, in "Literary and Linguistic Computing", no. 8(4), 1993, p. 243–257.

Biber, Douglas, *Dimensions of register variation: A cross-linguistic comparison*, Cambridge, UK, Cambridge University Press, 1995.

Biber, Douglas; Susan Conrad; Randi Reppen, *Corpus linguistics: Investigating language structure and use*, Cambridge, Cambridge University Press, 1998.

Claydon, Tony; Ian McBride (eds.), *Protestantism and national identity: Britain and Ireland, c. 1650-c. 1850*, Cambridge, Cambridge University Press, 1999.

Crespo, Begoña, *Persuasion markers and ideology in eighteenth-century philosophy texts (CEPhiT)*, in "Lenguas para Fines Específicos", no. 17, 2011, p. 191–228.

Crespo, Begoña; Isabel Moskowich, *CETA in the context of the Coruña Corpus*, in "Literary and Linguistic Computing", no. 25(2), 2010, p. 153–164.

Fowler, Alastair, *Kinds of literature. An introduction to the theory of genres and modes*, Oxford, Clarendon Press, 1982.

Görlach, Manfred, *Text types and the history of English*, Berlin/New York, Mouton de Gruyter, 2004.

Hyland, Ken. *Authority and Invisibility: Authorial Identity in Academic Writing*, in "Journal of Pragmatics", no. 34, 2002, p. 1091–1112.

Ivanić, Roth, *Writing and Identity: The Discoursal Construction of Identity in Academic Writing*, Amstrdam, ohn Benjamins, 1998.

Kytö, Merja; Junahi Rudanko; Erik Smitterberg, *Building a bridge between the present and the past: A corpus of 19-century English*, in "ICAME", no. 24, 2000, p. 85–97.

McEnery, Tony; Andrew Wilson, *Corpus linguistics*, Edinburgh, Edinburgh University Press, 1996.

Mischke, Gertruida Elizabeth, *Analysing 'involvement' in Distance Education Study Guides: An Appraisal-based Approach*, University of South Africa, 2005.

Monzó Nebot, Esther, *La profesió del traductor juridic i jurat: Descripció sociòlogica del professional i anàlisi discursiva del trasgenere*, Unpublished PhD Dissertation, Universitat Jaume I, 2002.

Moskowich, Isabel; Begoña Crespo, Presenting the *Coruña Corpus*: A collection of samples for the historical study of English scientific writing, in Javier Pérez Guerra, Dolores González-Álvarez, Jorge L. Bueno Alonso & Esperanza Rama Martínez (eds.), *'Of varying language and opposing creed': New insights into Late Modern English*, Bern, Peter Lang, 2007, p. 341–357.

Moskowich, Isabel; Javier Parapar López. Writing science, compiling science. *The Coruña Corpus of English Scientific Writing*, in María Jesús Lorenzo Modia (ed.), *Proceedings from the 31st AEDEAN Conference*, A Coruña, Universidade da Coruña, 2008, p. 531–544.

Moskowich, Isabel, "A smooth homogeneous globe" in *CETA*: Compiling late modern astronomy texts in English, in Nila Vázquez (ed.), *Creation and use of historical English corpora in Spain*, Cambridge, Cambridge Scholar Publishin, 2013.

Myers, Greg, *Writing biology: Texts in the social construction of scientific knowledge*, Madison, University of Wisconsin Press, 1990.

Nesi, Hilary, *A multidimensional analysis of student writing across levels and disciplines*, in M. Edwardes (ed.), *Taking the Measure of Applied Linguistics: Proceedings of the BAAL Annual Conference*, London, BAAL/Scitsiugnil Press, 2009.

OED Online. Oxford, Oxford University Press. http://dictionary.oed.com (accessed 28 April 2009). 2009.

Sokól, Malgorzata, *Discoursal construction of academic identity in cyberspace (On the example of an E-seminar)*, Szczecin, Szczecin University PhD thesis, 2006.

Taavitsainen, Irma; Päivi Pahta, *The Corpus of Early English Medical Writing*, in "ICAME Journal", no. 21, 1997, p. 71–78.

UNESCO, *Proposed international standard nomenclature for fields of science and technology*. UNESCO/ROU257 rev. 1, Paris, 1988.

Włodarczyk, Matylda, *Infinitives in the 1820 Settler Letters of Denunciation: What can a Contextualised Application of Corpus-Based Results Tell us About the Expression of Persuasion?*, in "Poznań Studies in Contemporary Linguistics", no. 46(4), 2011, p. 533–564.