



At close range: prefaces and other text types in the *Coruña Corpus of English Scientific Writing*

Begoña Crespo¹

Grupo MuStE, University of A Coruña

Isabel Moskowich

Grupo MuStE, University of A Coruña

ABSTRACT

What we nowadays term “front matter” was conceived of in the past as a direct address to the reader. Over time, standard formulae were developed and certain rhetorical devices consolidated. Late modern authors were familiar with the highly conventionalised patterns of prefaces and dedications and employed their “discursive freedom” in their scientific works even though the style used for the transmission of scientific knowledge was also changing and being standardised. This paper revolves precisely around the either parallel or divergent development of prefaces to scientific works and the body of the texts themselves. In order to study such evolutions we have analysed samples written by women between 1700 and 1900 in the *Coruña Corpus of English Scientific Writing*. The scrutiny of some linguistic elements generally admitted to express involvement have rendered a decline in the use of involvement features but we assume that frequency of use of the same features should be different in both prefaces and actual works. Unexpectedly, the overall frequency of these features is higher in the texts than in their corresponding prefaces.

Keywords: Text-types, scientific register, Coruña Corpus, involvement, female writing, prefaces

1. Introduction

¹ **Corresponding author** – University of A Coruña, Grupo MuStE, A Coruña (Spain). Email: bcrespo@udc.es



Much of what we nowadays term “front matter” was, in the past, conceived of as a direct address to the reader. Hence, prefaces, forewords, dedications and various ways of addressing the reader were located at the beginning of both literary and non-literary works, with the clear intention of attracting the attention of the audience. Over time, standard formulae from the classical tradition were adopted and developed, and certain rhetorical devices consolidated. In a similar vein, there was an evolution in the way in which scientific knowledge was transmitted (Atkinson, 1996, 1999) as well as in the style considered most appropriate for such objectives (Bazerman, 1988). A close reading of the texts suggests that authors were familiar with the patterns of prefaces and dedications, which were highly conventionalised, in contrast to what we might call the “discursive freedom” which they enjoyed in composing their scientific works. Academic prose in present-day English is said to be highly informative, that is, detached from any relation of empathy between author and reader or from any kind of linguistically manifested emotion (Biber, 1988, 2001; Frank, 2009). Scientific writers, as academic writers, seem to be no exception becoming users of what Lakoff (1990) has termed “considerate styles” in contrast with the so-called “high-involvement styles” (p. 50). The objectivity and universalism that science demands is conveyed by means of low involvement markers representing, likewise, their particular cultural ideology as communicative practices in science are culturally-shaped. This call for neutrality originated with the emergence of Empiricism and was gradually becoming the concern of scientific style thereafter. Hence, the change from an author-centred to an object-centred prose as claimed by Atkinson (1999). Today, the extent to which authors are present in scientific reports has risen as a matter of certain debate among those linguists devoted to the study of the language of science and academic writing in general (Hyland, 2000, 2002, 2008). Some of their findings (Hyland, 1998) reveal that authors are not so detached in some cases but that they let the addressee feel their presence in a subtle way, demonstrating that this register is not as “objective, impersonal and informational” as was previously thought (p. viii). This paper is a preliminary approach to the analysis of this authorial involvement in post-empiricist texts and their corresponding prefaces in an attempt to find out whether adherence to rhetorical clichés or discursive freedom could have been important in this issue. To be more precise, our interest here is whether the development of front matter and the main body of texts run parallel or whether they follow different paths as far as their degree of involvement is concerned because they are tied to some stylistic conventions that were especially firm and serious in the case of female writers. To this end, we will analyse scientific texts and their corresponding front matter written by women between 1700 and 1900 as this is the time-

span covered by the *Coruña Corpus of English Scientific Writing* (henceforth CC). All samples will be extracted from different sub-corpora of CC, namely, *CETA* (*Corpus of English Texts on Astronomy*) (Moskowich et al, 2012), *CEPhiT* (*Corpus of English Philosophy Text*, forthcoming), *CELiST* (*Corpus of English Life Sciences Texts*) and *CHET* (*Corpus of History English Texts*). Given that women writers in this period mainly devoted themselves to the creation of fictional works, the number of samples written by female authors in these corpora is not very high but it reflects eighteenth- and nineteenth-century scientific reality accurately.

Since we will be working with samples from different corpora devoted each to a particular scientific discipline, we will use this as one of the variables in the analysis. Disciplinary conventions might exert an influence on how the author communicates science. The second variable will be time. To this respect we will present and compare our results per century, considering that, as could be expected, the findings will point to an evolution towards detachment, mainly in texts. Finally, the text-type or genre to which the text belongs, that is, generic conventions, could also determine the final writing.

As part of our working hypothesis, we expect to find more of these features in prefaces and other front matter than in the texts themselves, as the former may have been used as a means of making contact with members of the epistemic community more directly (Narrog, 2012). This might be so because one of the primary pragmatic functions of preliminary material is in general to influence readers positively (Bradbury-Jones, 2007) so that they are more receptive to the contents of the text itself. Analysing the variables mentioned above gives us the opportunity to show whether such a hypothesis is true for all genres/text-types and disciplines under survey, as well as how such a phenomenon evolved during the late Modern English period.

This paper is divided into 4 sections. Section 1 discusses the general assumption that scientific writing is characterised by detachment and objectivity, although some recent authors claim that authorial presence can also be traced in such writing. Section 2 presents the material under survey: eighteenth and nineteenth-century samples of female writing in the CC. This material will be analysed in Section 3 and findings set out. Finally, in the last section we will provide some conclusions regarding our initial research question.

2. Scientific writing: detached or involved?

In terms of linguistic analysis, involvement means authorial presence, reader-writer interaction and the expression of emotions which is manifested through language. On occasions, the written medium is allowed to make use of structures frequently found in oral exchanges which endow the writing with a touch of closeness and intimacy between the participants in the communicative process. Lack of involvement is what has been gradually characterising scientific prose in the last two centuries (Daston and Galison, 1992) in such a way that “the all-to-human scientists must, as a matter of duty, restrain themselves from imposing their hopes, expectations, generalizations, aesthetics, even ordinary language on the image of nature. Where human self-discipline flagged, the machine would take over” (p. 81).

Our focus here will be on the use of some of the linguistic elements generally taken as denoting involvement (Biber, 1988). These elements have already been used in studies of contemporary English but also to examine the Involved vs Informational dimension in texts written in earlier periods. This was the case of Atkinson’s study of the Philosophical Transactions (1999) where he addressed register dimensions basing on these features and many others. In addition, as Tieken-Boon van Ostade has stated referring mainly to syntax, “According to traditional accounts of eighteenth-century English, nothing much happened to the language during this period”. Both ideas have guided our selection of items (Tieken-Boon Van Ostade, 2006, p. 254).

These features include the use of first and second person pronouns, *wh*-words, hedges, amplifiers and private verbs. We believe that not all these features are equally direct when addressing the reader, but can be grouped according to different levels of subtlety. Personal pronouns and private verbs can be clearly regarded as direct devices, showing personal commitment in the use of I and the inclusion of the reader in the communicative process through the use of you or we. On the other hand, amplifiers and hedges can perhaps be seen as more subtle devices, in that they are less direct. In fact, when an author uses forms such as totally as in

- (1) ... which if superftition did not cloath with terrors muft prove totally unavailing. He therefore concerned himfelf no farther in the affair (Scott, 1762: 159)

she is not only indicating involvement but also expressing her own thoughts on the object of her writing. Finally, *wh*-words seem to occupy a medial point in this taxonomy of devices. They are not perceived by the addressee as either direct or subtle, and might perhaps be defined as somehow “neutral”. Example (2) below illustrates the use of *wh*-words as the initial element of indirect

questions. This mechanism evinces proximity with the addressee as it represents closeness to speech:

- (2) But to return to the question, "what causes Spring?" or to state it in another form, by what means does the Almighty produce the changes which this season presents? (Lincoln, 1832: 279)

Table 1 below shows the instances of each of these features we have searched for:

Hedges	At about, something like, more or less, almost, maybe, xxx sort of (where xxx is not determiner, possessive pronoun adjective or who), xxx kind of
Private verbs	Anticipate, assume, believe, conclude, decide, demonstrate, determine, discover, doubt, estimate, fear, feel, find, forget, guess, hear, hope, imagine, imply, indicate, infer, know, learn, mean, notice, prove, realise, recognise, remember, reveal, see, show, suppose, think, understand
Wh-questions	How, what, when, where, which, who, , whose, why, however,
Amplifiers	Absolutely, altogether, completely, enormously, entirely, extremely, fully, greatly, highly, intensely, perfectly, strongly, thoroughly, totally, utterly, very
1st and 2nd person pronouns	I, me mine, myself, ours, ourselves, us, we, you, yours, yourself, yourselves

Table 1. Linguistic features under study

The analysis of the above features in relation to the three variables mentioned (time, genre and discipline) will afford a glimpse of stylistic variation in scientific discourse. It is hoped that it will also provide a preliminary picture of the evolution towards a detached style in writers of specialised registers who are typically considered more involved: that is, female writers (Argamon et al., 2003). The focus of scientific writing is said to have shifted from authors to objects in order to provide a sense of objectivity (Atkinson, 1999), enhancing the credibility and verisimilitude of the information conveyed. Thus, the presence of the author, as in the following extracts from Newton's *Opticks* (1704):

I held the Prism (p. 22: line 3)

I looked through the Prism (p. 22: line 5)

I stopt the Prism (p. 23: lines 3-4)

I observed the length of its refracted Image (p. 23: lines 7-8)

I removed the Prism out of the Sun's Light and looked (p. 23: lines 11-12).

is gradually replaced by a more depersonalised style in which the author is hidden behind linguistic strategies such as the use of the passive voice and the avoidance of the first person singular pronoun. This can be seen in the following extract from a mid-twentieth century article published in the *New England Journal of Medicine*:

THE genus of fungi aspergillus includes approximately 100 species. Some of these species *are* frequently *encountered* as contaminants of specimens *submitted* to the laboratory for bacterial and fungal cultures. Thom and Church¹ called them the "weeds of the culture room." The frequent occurrence of aspergilli as contaminants has caused difficulty in establishing an etiologic relation in an illness in a given patient. *aspergillus fumigatus*, *A. niger*, *A. clavatus*, *A. flavus* and *A. versicolor* *have been* definitely *associated* with disease in man, but only the first two *can be said to be frequent causes* of mycotic disease. Because aspergilli may be . . . (Utz et al., 1959, p. 264)

The shifts just mentioned may be assumed to have a greater effect on the main body of texts than on their prefaces, since the latter may have not been subject to the same tendencies and varying trends in writing. In fact, a reasonably close relation between the writer and the reader may be observed in prefaces throughout history, whereas in scientific discourse this relation is disguised in different ways.

3. Corpus material

The material under survey has been extracted from different sub-corpora of the CC, namely, *CETA* (Corpus of English Texts on Astronomy), *CEPhiT* (Corpus of English Philosophy Texts forthcoming), *CELiST* (Corpus of English Life Sciences Texts) and *CHET* (Corpus of History English Texts)². Each of these sub-corpora contain two samples per decade of 10,000 words each which makes a total of ca. 200,000 words per century and discipline and ca. 400,000 words per discipline overall. The consequence of our attempt to reflect discourse realities in these corpora is twofold: on the one hand, there is a scarcity of samples written by female authors, and on the other the distribution of these samples in the different corpora is uneven, as indeed was the case in society at the time. Although other female authors figure in the CC, the samples that form our material here include only texts in which there is some form of front

² Beta versions of *CELiST* and *CHET* have been used for this survey.

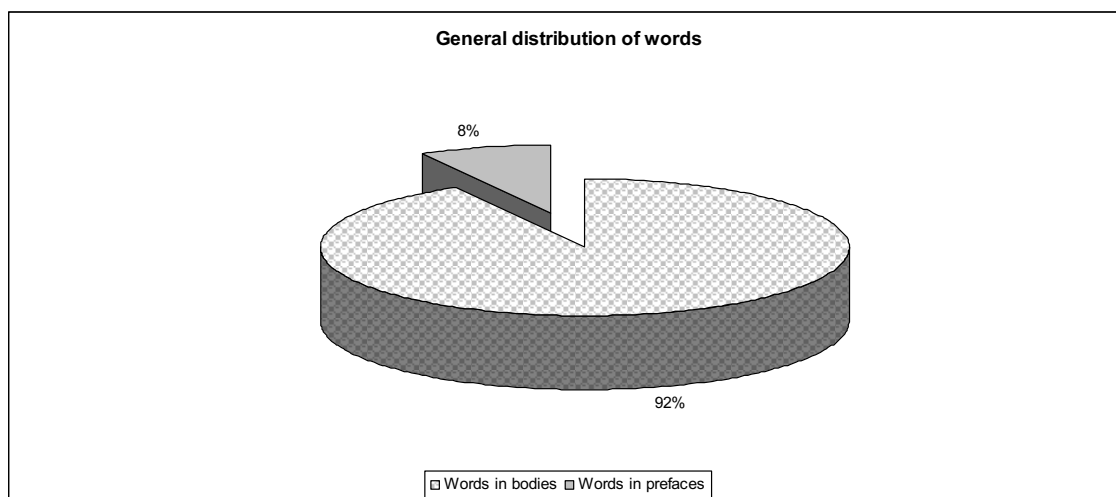
matter, so that the comparison of both extracts is possible. All the information regarding dates, samples and word counts is set out in Table 2:

Author	Year	Title	Discipline	Text type	Sample Words	Preface Words
Astell, Mary	1700	<i>Some reflections upon marriage</i>	Philosophy	Essay	10,079	248
Macaulay, Catherine	1783	<i>Treatise of the immutability of moral truth</i>	Philosophy	Treatise	10,060	1,650
Wollstonecraft, Mary	1792	<i>Vindication of the Rights of Women: : With Structures on Political and Moral Subjects</i>	Philosophy	Treatise	10,053	1,684
Wakefield, Priscilla	1816	<i>An introduction to the natural history and classification of insects, in a series of familiar letters</i>	Life Sciences	Letter	9,805	309
Lincoln, Almira Hart Phelps	1832	<i>Familiar lectures on botany, including practical and elementary botany</i>	Life Sciences	Lecture	10,028	792
Pratt, Anne	1840	<i>Flowers and their associations</i>	Life Sciences	Treatise	10,023	317
Agassiz, Elizabeth	1859	<i>A First Lesson in Natural History,</i>	Life Sciences	Letter	12,959	51
Lankester, Phebe	1879	<i>Wild flowers worth notice ... for their beauty uses and associations</i>	Life Sciences	Treatise	10,080	2,837
Bryan, Margaret	1797	<i>A compendious system of astronomy</i>	Astronomy	textbook	10,293	1,056
Clerke, Agnes Mary	1893	<i>A popular history of astronomy during the nineteenth century.</i>	Astronomy	Treatise	10,530	932
Justice, Elizabeth	1739	<i>Voyage to Russia: describing the Laws, Manners, and Cuftoms, of that great Empire, as govern'd, at this prefont, by that excellent Princefs, the Czarina. Shewing the Beauty of her Palace, the Grandeur of her Courtiers, the Forms of Building at Petersburgh, and other Places: with several entertaining Adventures, that happened in the Paffage by Sea, and Land</i>	History	Other	10,005	396
Scott, Sarah	1762	<i>The History of Mecklenburgh, from the Firt Settlement of the Vandals in that Country, to the Present Time; including a Period of about Three Thoufand Years</i>	History	Treatise	10,114	1,054
Warren, Mercy Otis	1805	<i>History of the rise, progress and termination of the American revolution. Interspersed with Biographical, Political and Moral Objervations. In three volumes. Vol. I.</i>	History	Treatise	10,032	1,609
Aikin, Lucy	1833	<i>Memoirs of the Court of King Charles the First. In two volumes. Vol. I</i>	History	Treatise	10,022	259

Sewell, Elizabeth Missing	1857	<i>A first history of Greece</i>	History	Textbook	10,057	195
					154,140	13,389

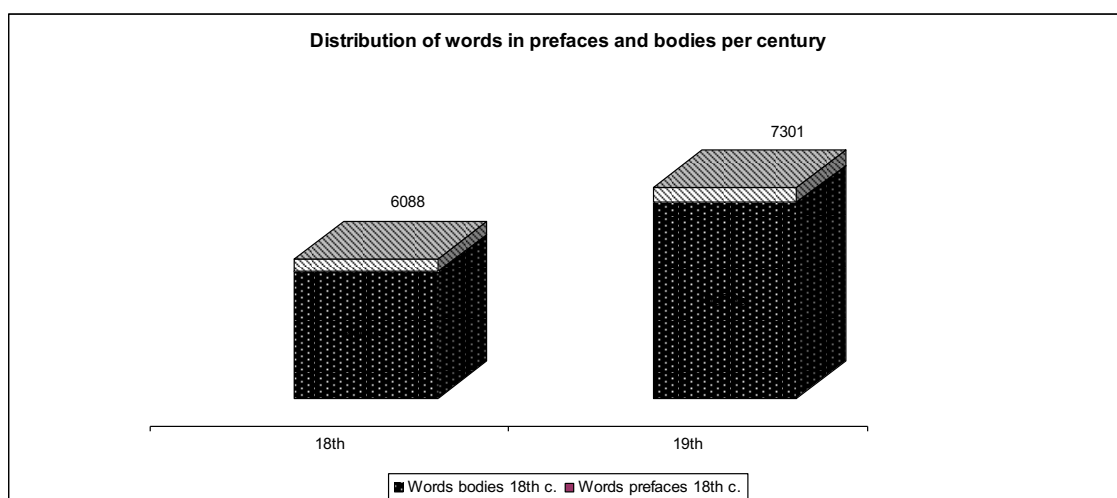
Table 2. Selected women scientists in the CC

Our material includes a total of 167,529 words, with a distribution into prefaces and main body texts as shown in Graph 1 below, with a split of 92% to 8% in favour of main texts:



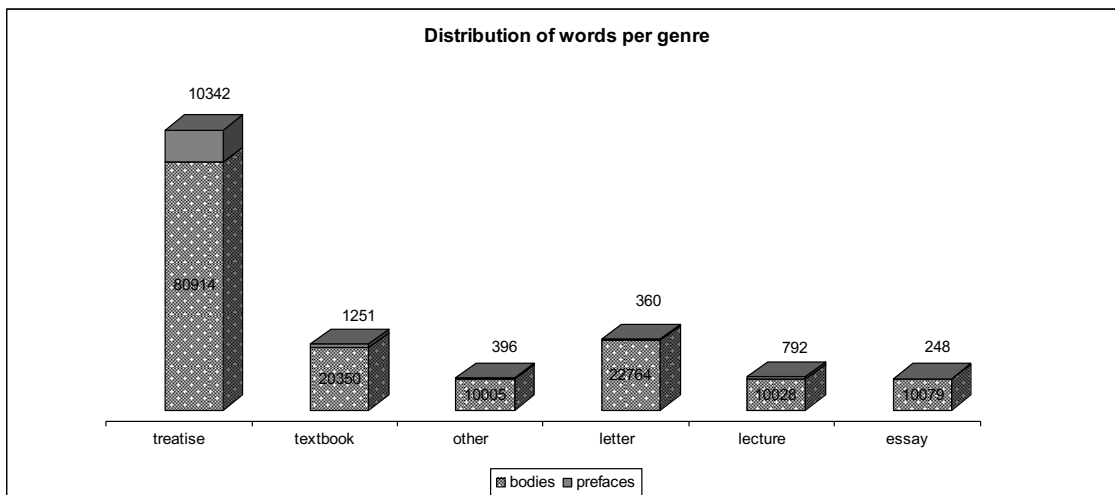
Graph 1. Distribution of words in front matter and main texts

Since we will consider three variables (time, text-type/genre and discipline) for the analysis of data, we will also show the distribution of our material according to these. As seen in Graph 2 below, the distribution of words in front matter and main text per century indicates that more words have been recorded in nineteenth-century texts in both cases (7,301 words in front matter and 93,536 in main texts, totalling 100,837). As already noted, these figures reflect not only the scarcity of women scientists at the time but also the lack of visibility they had. This might also explain why the eighteenth-century material contains fewer words (6,088 words of front matter and 60,604 of main texts, totalling 66,692).



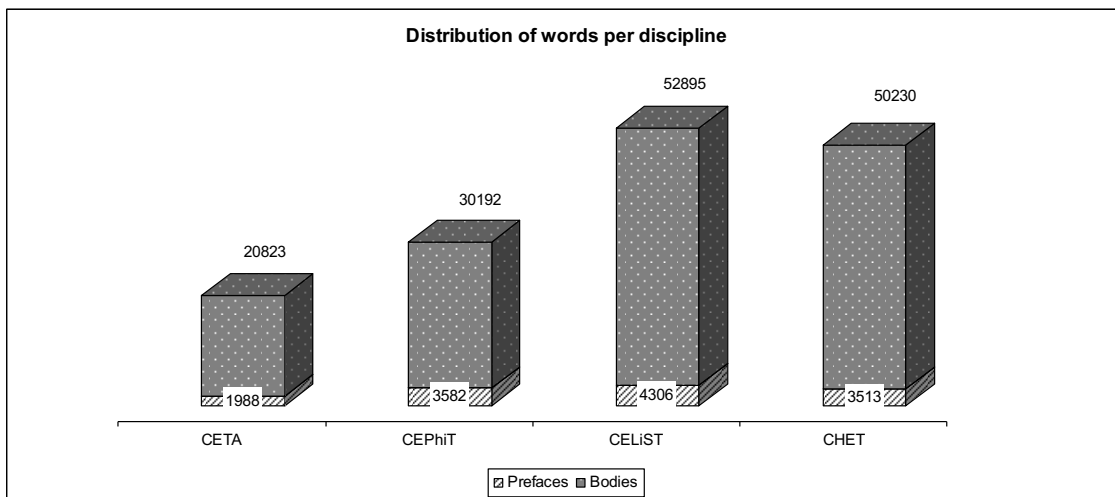
Graph 2. Distribution of words in front matter and main texts per century

The second variable used to present our material is that of genre. The present classification of genres in the CC includes 8 different categories: Treatise, Letter, Lecture, Essay, Textbook, Article, Dialogue and Other. However, our material contains no examples of Article or Dialogue. This is so, first, because the samples of these genres had no preface and were therefore disregarded, and second, because no Dialogues were among the samples by women in the CC. As Graph 3, below, makes clear, Treatise is the genre with the highest number of words, both in prefaces and main texts. Letter is second, although the proportion of words in prefaces and main texts here is unbalanced as compared with some of the other genres. The third genre is Textbook, with the split of words into prefaces and main texts reversed in comparison to that in Letter. Obviously, textbooks tend to require longer introductions, as a means of explaining students the purpose and possible use of the works in question. The three remaining genres are more or less evenly distributed as regards the number of words in front matter and main texts. The issue of note here is the fact that Lectures are the genre with the highest number of words in their prefaces. The reason for this could be that these lectures were first delivered in public and were subsequently published, and hence needed some kind of introductory material when offered to the readers.



Graph 3. Distribution of words per genre

Finally, the disciplines Life Sciences and History contain, in this order, the highest total number of words. The same order applies in the case of main texts but there is a slight variation in the case of front matter, since Life Sciences and Philosophy are the two disciplines with the highest number of words. The corresponding totals are displayed in Graph 4:



Graph 4. Distribution of words per discipline

Main texts, encoded in xml format, have been searched for using the *Coruña Corpus Tool*, a retrieval tool specifically designed for the *Coruña Corpus* Project in collaboration with the IRLab (Computing Department, University of A Coruña) and which accompanies every release of a sub-corpus. Prefaces, which are not xml-marked-up, have been searched for with Lawrence

Anthony's AntConc, freely available on his webpage (<http://www.laurenceanthony.net/software.html>).

Given that the raw numbers here are uneven, frequencies will be normalised to 10,000.

4. Analysis of data

As noted in Section 2, we will search for traces of author involvement through those linguistic features generally acknowledged to be representative of this phenomenon. They are set out in Table 3 below, together with their corresponding frequencies in all our material:

Involvement feature	Raw frequency	Normalised frequency
1st and 2nd person pronouns	1,737	103.68
Hedges	119	7.10
Amplifiers	548	32.71
Private verbs	1,950	116.39
<i>Wh</i> -words	3,198	190.89

Table 3. Frequencies of involvement features

Of the five features listed above, *wh*-words clearly predominate in our data (190.89). This may be due to the fact that, following Biber's (1988) initial proposal, we have included here forms such as *wherever*, *whatever*, *whence*, etc., rather than just *wh*- direct questions or reported speech, as some authors have proposed (Mischke, 2009). However, the use of mainly indirect questioning, not only in samples from prefatory material but also in those from the bodies of the scientific works under analysis, might work as a way of connecting with readers and in this way encouraging them to participate in authorial tenets, as can be seen in this example:

- (3) Modern Astronomers and Mathematicians deserve our highest esteem and commendation, who, by a communicable spirit, liberally extend the province of science, and thereby confer a lasting benefit on society; not like the ancients, who, contracted in their ideas, and wishing to be considered as superior beings, hoarded up their scientific principles for the purpose of ostentation; —and *what* was the consequence of their selfish and mysterious conduct? —Nothing less than in some cases the total subversion of all their knowledge, which, dying with them, the succeeding generations, uninstructed in their theories and investigations, and perhaps not possessing the same spirit of enquiry, the science became neglected or perverted; and thus all the advantages which had been produced were lost to mankind. (Bryan 1797: 101)

Such a linguistic mechanism provokes in readers the need to search for an answer, thus drawing them into the author's questions. Verbs of thinking and feeling, such as the one in example (4) below, come second in frequency (116.39 nf), perhaps as a result of the same tendency to use overt resources of author-reader interaction:

- (4) Notwithstanding the situation of the two insects was such that the spider could not see the fly, he perceived the victim the moment he entered the snare, repaired to the spot, and disabled him from escaping; which he attributes to the sense of hearing. But, as I *think* it possible that the approach of the fly was made known to the spider by the vibration of the thread, as well as by its tones of distress, I am more convinced of the fact by his inference, that, since it is well known that many insects have the power of uttering sounds, as the large beetle, the bee, the wasp, the gnat, and the fly, we cannot suppose that this power is given without a corresponding capacity in their fellows to understand their meaning. (Wakefield 1816: 9)

We should not overlook the fact that these forms represent an overt manifestation of the inner self, and it is the writer's own opinion that is transmitted through such verbs of thinking or feeling. First and second person pronouns come third (103.68 nf), emphasising the direct form of address between the two interlocutors in the communicative process (see examples 5 and 6):

- (5) I have seen several curious Representations, [viz]. A Garden so natural, that you would imagine you might gather Oranges from the Trees: The Walls of Peru, some of which appeared to be broke down: Their Alphabet which consists of Forty Letters; and their Academy, which is likewise beautifully illuminated. (Justice 1739: 23)
- (6) This was Rousseau's opinion respecting men: I extend it to women, and confidently assert that they have been drawn out of their Sphere by false refinement, and not by an endeavour to acquire masculine qualities. (Wollstonecraft 1792: 37)

Less obvious indications of the author's interest in establishing a relation with the reader through making herself visible are amplifiers (32.71 nf) and hedges (7.10 nf). Moreover, it is worth noting that these features are at least three times less frequent in our material than the other features analysed, which strongly suggests the preference for non-subtle devices in female writing. The use of amplifiers and hedges can be seen in examples 7 and 8 below:

- (7) Friendship or indifference inevitably succeeds love and this constitution seems perfectly to harmonize with the system of government ... (Wollstonecraft 1792: 58)

- (8) When a being, such as man, the mechanism of whose mind is framed of powers so various and so adverse, as often to act with *a kind of* hostility against each other, and whose imagination, instead of serving him, as a principle, to invigorate those sentiments which the cool dictates of his understanding produce, is often at war with that faculty, whose peculiar province it is to assist; (Macaulay 1783: 18)

Since this study aims to provide a comparison of the use of involvement features in prefaces and main texts in samples from the CC, Table 4 has been included, illustrating the corresponding overall distribution:

Features	rf	nf
Front matter	417	311.44
Main texts	7,179	465.74

Table 4. Frequencies of involvement features in the material under survey

Surprisingly, our data reveals that features denoting involvement are certainly more frequent in main texts than in front matter. As mentioned above, the dialogic nature of prefaces seems, in principle, to favour the presence of the features under discussion here, and on occasions we can even find overt references to the reader:

- (9) To those readers who have just commenced the study of botany, or who may intend to pursue it at some future time, the author hopes these pages may not be unacceptable. (Pratt 1840: iv)

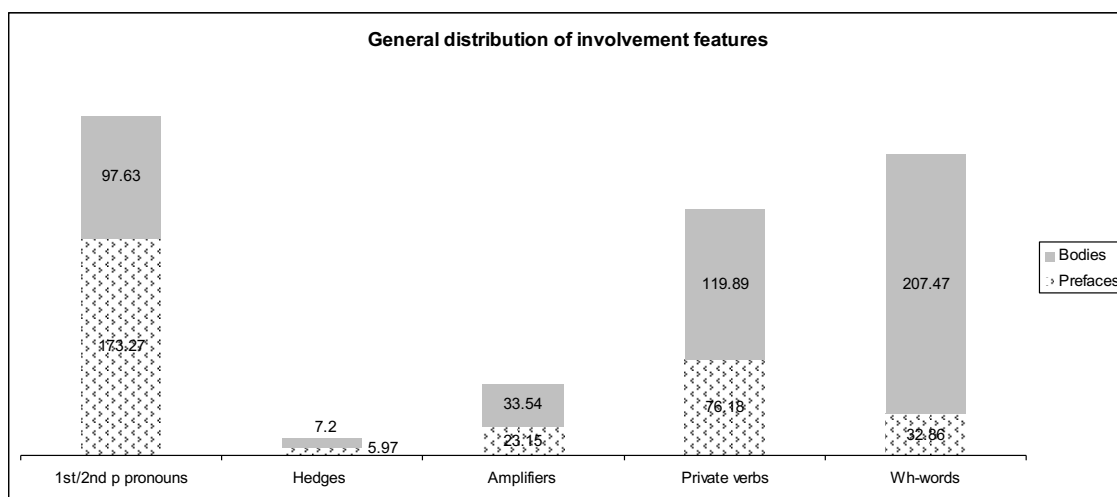
Such direct appeals to readers occur only very rarely in the main texts, constrained as these are by the dictates of the Royal Society, as advocated by Bacon and Boyle and summarised by Thomas Sprat (1635-1713), first historiographer of the Royal Society of London, in the following terms:

It will suffice my present purpose to point out what has been done by the Royal Society towards the correcting of its excesses in natural philosophy; to which it is, of all others, a most professed enemy.

They have therefore been most rigorous in putting in execution the only remedy that can be found for this extravagance, and that has been, a constant resolution to reject all the amplifications, digressions, and swellings of style; to return back to the primitive purity, and shortness, when men delivered so many things, almost in an equal number of words. They have exacted from all their members a close, naked, natural way of speaking; positive expressions; clear senses; a native easiness: bringing all things as near the mathematical plainness as they can; and preferring the

language of artisans, countrymen, and merchants, before that of wits or scholars (Graig, 1916).

The features expressing different degrees of authorial involvement are differently distributed in the two types of data, as can be seen in Graph 5:



Graph 5. Distribution of involvement features in front matter and main texts

As we have described, *wh*-words, private verbs and 1st and 2nd person pronouns were the most abundant features in the set of samples overall. However, this distribution is not the same in the two types of samples. Thus, this pattern can be observed in main texts, but not in the front matter. The difference is merely a question of order of frequency. In prefaces, 1st and 2nd person pronouns are most frequent (173.27 nf), followed by private verbs (76.18) and *wh*-words (32.86). Such findings might be accounted for by the fact that direct addresses to the reader require the use of devices such as personal pronouns and private verbs. Prefaces are inherently interactive, in that they could be considered as the written manifestation of authors trying to establish a dialogue with their addressees.

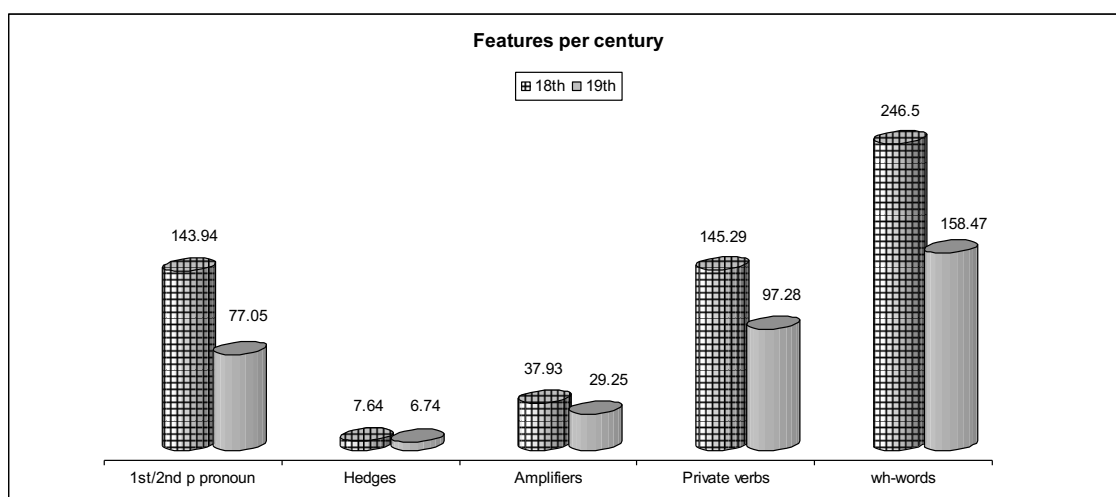
Returning to the overall picture, hedges and amplifiers are under-represented in comparison to the features already described. The distribution of both is quite balanced in front matter (5.97 nf hedges and 23.15 nf amplifiers) and in main texts (7.2 nf hedges and 33.54 nf amplifiers). The preference for non-subtle devices to express involvement, then, seems to extend to both kinds of extracts.

In what follows we will describe the data from the perspective of the variables time, genre and discipline.

4.1 Time

The use of the linguistic features under study here is clearly determined by time, as we will show below. In fact, an analysis of how they are distributed in both centuries shows a clear decline in the frequency of involvement devices in favour of a more detached style, as we claimed in our initial hypothesis. In contrast with the 581.32 (nf) involvement features present in the eighteenth-century samples, the extracts from the nineteenth century have only 368.81 nf.

A more detailed analysis of the behaviour of each feature over time is offered in Graph 6:

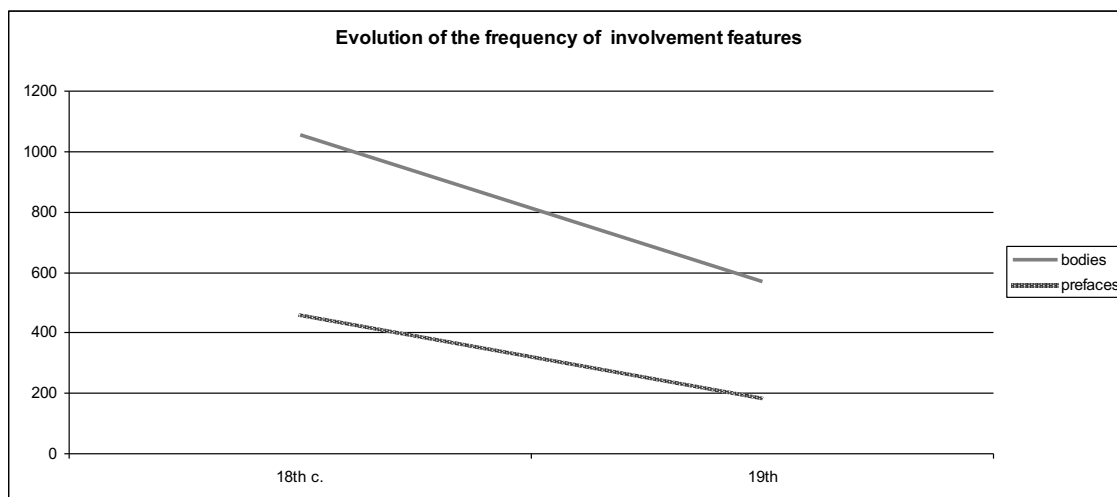


Graph 6. Features per century

Despite a general decrease in the use of involvement features, some important differences can be perceived in how this decrease occurs in each individual device. As can be inferred from the data in Graph 6, hedges (7.64 vs 6.74) and amplifiers (37.93 vs 29.25) are the only linguistic forms in which the difference in frequency of occurrence does not seem to change much over time. Perhaps due to the fact that personal pronouns (143.94 vs 77.05) and private verbs (145.29 vs 97.28) are more overt manifestations of authorial presence, they are the first to be avoided when scientific writers try to follow the dictates of the Royal Society. In this sense, women were no exception.

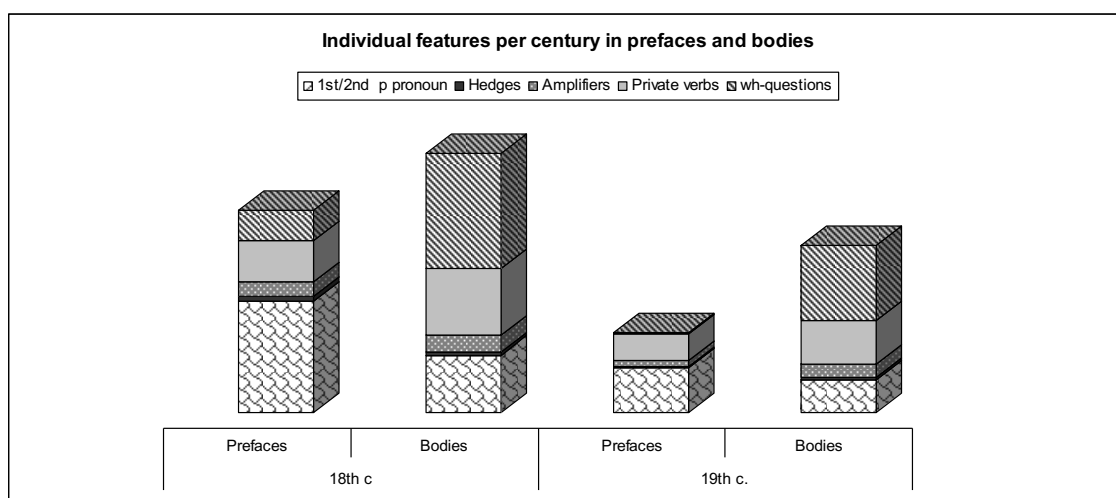
When considering the two kinds of samples selected for this study it seems necessary to present a diachronic outline of involvement features in scientific texts written by women. These figures can be further interpreted in terms of their distribution in the different parts of the samples, as Graph 7 below shows. Here we can see that main texts show a more marked decline than

front matter, which is in line with the evolution of discursive patterns in science towards what we find nowadays, a more object-centred style, in Atkinson's words (1999).



Graph 7. Evolution of the use of involvement features over time

Perhaps this tendency for objectivity in women authors can also be accounted for in terms of certain social influences, such as these women's own positions as writers: not many women participated in the formal academic world, and thus it is probable that they had to make an extraordinary use of objectivity markers in order to achieve the same visibility and respect as male scientists. Interestingly enough, the features in eighteenth century front matter are more abundant than in any of the nineteenth century samples. A comparison of involvement features between front matter and main texts per century shows that the gap in the use of personal pronouns in eighteenth-century prefaces and main texts was greater than in the nineteenth century. This finding could, in principle, suggest that the same tendency to objectivity is found in prefaces.



Graph 8. Individual features per century in prefaces and main texts

As can be seen in Graph 8 above, hedges and amplifiers are scarce in both sets of samples when compared to the rest of the features under survey. Our data also reveal that private verbs and *wh*-words occur more frequently in eighteenth-century main text samples, and the same pattern, after the overall decline in the number of features is taken into account, is repeated in the nineteenth-century main texts.

- (10) such an accumulation never actually takes place in this supposition. I do not presumptuously mean to confute his assertion as I conceive that great man by no means intended to convey... (Bryan 1797: 93)
- (11) I AM well aware, that to write on subjects which have been so extensively considered, and fully delineated, by the ablest Mathematicians, and by Philosophers of the most penetrating genius, will not procure K any honor on the score of originality; yet I trust I shall not incur censure by publishing a Compendium which, according to my ideas, will render subjects, generally thought obscure, clear to the understanding of young people: —If I have failed in the attempt, through the imbecility of my judgment, I hope the motive may be my apology. (Bryan 1797: vii)
- (12) ... the conditions of the seeds in germination are worthy of attention. I need scarcely describe the plant so familiar to us... (Lankester 1879: 69)
- (13) The success which has attended my little book on “Wild Flowers worth Notice,” in the various forms in which for several years it has appeared before the public, determined the publisher to produce it again in a revised and improved condition, yet retaining its old character, only adding to it, rather than altering it. I am therefore glad to have the opportunity of correcting any error which may now exist in work done some years ago, and have endeavoured to bring the information given, up to the present time. (Lankester 1879: v)

Thus, the variable of time seems to confirm the decay of the author-centred style in favour of a more object-centred one in women's writings. Authorial detachment is progressively more obvious in the use of particular linguistic forms not only in main texts (see examples (10) and (12)) but also in front material, as examples (11) and (13) illustrate.

4.2 Genre

The second variable under study here is genre. As mentioned in Section 2 above, the materials under analysis belong to six different genres: Treatise, Textbook, Other, Lecture, Letter and Essay.

Genre	total nf
Treatise	366.81
Textbook	597.65
Other	712.43
Letter	544.45
Lecture	385.39
Essay	550.01

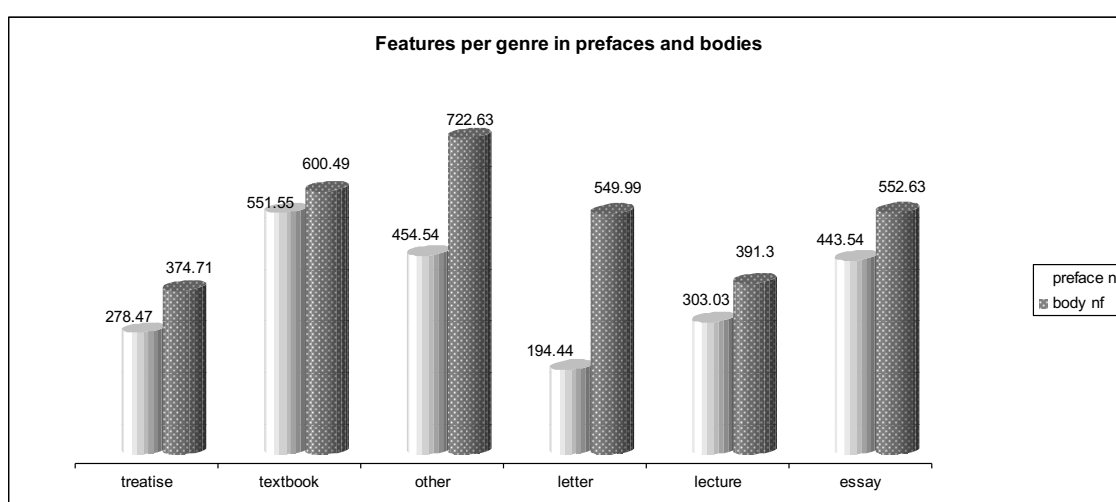
Table 5. Overall distribution of features per genre (normalised figures)

Features denoting involvement occur most frequently in the genre Travelogue (712.43 nf), the sample here taken from *CHET* (see Table 5). It includes place descriptions and "experiences encountered in the course of travel" (*OED*) and seems to favour the use of a free narrative style, one which is more personal than other genres as illustrated in example (14). The difference between the presence of these features in the text itself and the corresponding preface (722.63 vs 454.54), the second highest difference, could reinforce such an explanation (see Graph 9 below).

- (14) for, I muft *confefs*, the Sea-Terms were fuch, that I could not *believe* they talk'd Englifh. I went upon Deck; and, in a fmall Time, had the Happinefs to fee every Sail hoifted, and hear the Captain fay to the Boy at the Helm, Steady, my Lad! Steady! Which is all can be wifh'd for at Sea. After being thus entertained, I went into the Cabbin to Breakfast; which was exceeding neat, and very genteely furnifhed with a rich Crimfon Damask; and very handfome China, Glafs, and Sconces, with gilt Frames: The Mouldings of the Wainfcot were gilt with Gold; and the Tea Things were fet in the fame Order, as if *we* had been at Land: And I cannot fay but every Thing *we* had was very agreeable, and the Reverse of what I *expected*. (Justice 1739: 3)

The genre Textbook comes next (597.65). It is surprising that the third position is not occupied by Lecture, since together with the first two genres, lectures are intended to engage the readership, mainly in that they are written to be read aloud and might thus manifest some traces of orality. In this vein, it is also surprising that in the third position here is Essay, since this should in principle involve a more detached way of communication. On the contrary, the findings for Treatise are expected, in that it occupies the last position on our frequency scale.

A detailed picture of the distribution of features in front matter and main texts across the different genres is presented in Graph 9.



Graph 9. Distribution of features per genre in front matter and main texts

One of the predictable findings is the abundance of these features in prefaces to textbooks. All the devices are used in these samples, presumably as a means of being in contact with readers and preparing them for what they should expect from the textbook itself. This may be interpreted in terms of the type of intended addressee of textbooks. At that time, women wrote textbooks as well as other instructional material for children or for other women (Moskovich, 2013; Crespo, 2015) as illustrated in the examples below:

- (15) There are but few elementary works on these subjects, adapted to young readers, either from their high price, or their scientific manner, which is more likely to alarm than attract the inexperienced pupil. (Wakefield 1816: iv-v)
- (16) Even to read it with a map does not materially lessen the difficulty; for children, especially girls, rarely know much of ancient geography; and whilst they are laboriously searching for the unknown places, the

thread of the history is interrupted, and their attention perhaps irrecoverably distracted. (Sewell 1857: 5)

The preface to the travelogue by Elizabeth Justice comes next (454.54 nf), with a slightly higher incidence of involvement devices than the essay by Mary Astell (443.54 nf). Although we only have one sample of each of these genres, and some idiosyncrasies of these particular authors could be one of the reasons for these results, it is also worth noting that both extracts belong to humanist-related disciplines. At the other end of the scale we find that the genre Letter, illustrated by example (17) below, contains only 194.44 nf devices in the prefatory material.

- (17) THIS little book, which it is hoped may be interesting for children, and perhaps of some use to parents whose children share the general juvenile delight in Aquariums, has been prepared under the direction of Professor Agassiz, and owes any little merit it may possess to his advice and assistance. (Agassiz 1859: 3)

The samples of letters we have are preceded by short pieces of writing. This may also account for the notable difference in findings for the letters themselves, a genre which is *per se* closer to the reader (Earle, 1999). Personal letters lie at the basis of learned letters by means of which various phenomena of nature were reported. Following letter-writing conventions which include a direct address to a particular reader may be interpreted as a sign of intimacy and an attempt to establish rapport with the addressee. The rest of the main texts, corresponding to the other genres, seem to follow the general distribution of features per genre as discussed above.

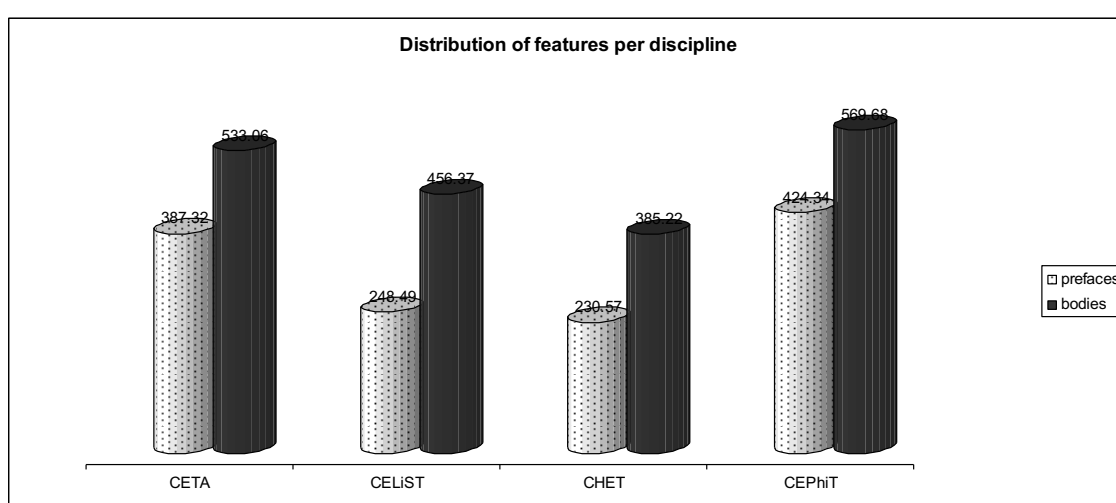
4.3 Discipline

The distribution of features in front matter and main texts according to the variable discipline (Graph 10) shows that Philosophy contains the highest number of features in normalised frequencies, both in main texts (569.68) and front matter (424.34). An example of this can be found in the dedication by Mary Wollstonecraft:

- (18) SIR, HAVING read with great pleasure a pamphlet, which *you* have lately published, on National Education, I dedicate this volume to *you* —the first dedication that *I* have ever written, to induce *you* to read it with attention; and, because *I think* that *you* will *understand me*> which *I* do not *suppose* many pert wittlings will, who may ridicule the arguments they are unable to answer. But, Sir, *I* carry my respect for your

understanding ftill farther; fo far, that / am confident <you> will not throw my work aside, and haftily *conclude* that / am in the wrong, becaufe *you* did not view the subject in the fame light *yourself*. (Wollstonecraft 1792: iii).

Philosophy is followed by Astronomy with 533.06 nf in the main texts and 387.32 in front matter, and by Life Sciences (456.27 nf and 248.49 nf, respectively). History is the discipline with the lowest number of such devices found (385.22 nf for main texts and 230.57 for front matter) although, as can be seen, prefaces to Life Sciences and History texts display quite similar proportions.



Graph 10. Distribution of features per discipline

As explained elsewhere the four sub-corpora used for this study contain samples of disciplines arranged in such a way that two of them belong to the Humanities (Philosophy in *CEPhiT* and History in *CHET*) and the other two to the Natural Sciences (Life Sciences in *CELIST* and Astronomy in *CETA*) (Crespo and Moskowich, 2010). Such a classification makes it possible to perceive that involvement features are not necessarily more frequent in the Humanities, as might be expected. On the contrary, Natural Sciences texts offer a greater density of these features, perhaps due to the genre used in each case (Textbook and Letter).

5. Conclusions

The linguistic features we have analysed here are present in our material in different ways. Overall in the data, *wh*-words, private verbs (that is, verbs of

thinking and feeling) and 1st and 2nd person pronouns appear in this order. However, this order is reversed when we look only at front matter, where 1st and 2nd person pronouns are the most abundant elements, followed by private verbs and *wh*-words. This seems to reflect the communicative purpose of prefaces and other preliminary material.

As regards the behaviour of these markers of involvement over time we note a decline, and this corroborates the tendency towards detachment (Lareo and Montoya, 2007) which has been observed in other historical studies of scientific texts (Moskowich and Crespo, 2015). This decline is not so clearly marked in front matter, perhaps due to the more flexible rhetorical patterns applied here and the need of authors to interact with readers in some way.

As for genre, we have observed that Treatise contains fewer instances of these markers than any other genre in our material, which is perhaps indicative of the level of formality which treatises typically have. Travelogue is in fact the genre with the highest proportion of involvement devices in our data. In terms of our analysis of how subject-matter may have influenced the expression of involvement, we might recall that although our single sample of the genre Travelogue belongs to History, the discipline itself is the one containing fewest involvement devices. On the other hand, it is another discipline from the Humanities, Philosophy, which displays the largest number of linguistic elements conveying authorial involvement.

The most surprising of our findings is that the overall proportion of features expressing involvement in main texts surpasses that of cases found in front matter. That is to say, contrary to our initial hypothesis, and perhaps also to any common sense supposition, prefaces and other preliminary material are more detached than the main body of scientific texts written by women during the late Modern English period, at least as far as the elements analysed here are concerned. This may be the result of the constant effort of women to sympathise with their readers, even in the main body of their works. Direct references to their “juvenile friends” (Wakefield, 1816), “girls” (Sewell, 1857) and “pupils” (Bryan, 1797) in front matter seem to confirm this, and the predominance of involvement features in the main texts could be said to play that same role, creating a sense of closeness between author and reader.

This study has involved an analysis of the inevitable inter-relationship between genres, disciplines and the diachronic development of the features under study. The application of these three variables to the data illustrates, once more, that texts by women authors are subject to different influences and forces, none of these being independent of each other.

About the authors

Dr Begoña Crespo is a Senior Lecturer at the University of A Coruña (Spain). She has published on lexicon enlargement in the History of English, genre/text-type taxonomy and the diachrony of English scientific writing. She is currently supervising some PhD dissertations on the language of science. She has published in journals such as *Studia Neophilologica*, *Nordic Journal of English Studies*, *Neophilologische Mitteilungen*, *English Studies*, *SAP and Literary and Linguistic Computing* among others as well as book chapters in D.S Brewer, Rodopi, Peter Lang, Lincom Europa or John Benjamins.

Dr Moskowich is a senior lecturer in the University of A Coruña. Her research covers different aspects of the History of English, mainly the Scandinavian lexical elements and sociolinguistic implications during the Middle Ages. In the last few years, she has been involved in the compilation of the *Coruña Corpus of English Scientific Writing*. She has published extensively in international journals and has edited different volumes as well as being responsible for funded several research projects.

Acknowledgments

The research here reported on has been funded by the Spanish Ministerio de Economía y Competitividad (MINECO), grant number FFI2013-42215-P. This grant is hereby gratefully acknowledged.

Article history

Paper received: 30th November 2015

Paper received in revised form and accepted for publication: 22nd February 2016

References

Argamon, S; Koppel, M; Fine, J. and Shimoni, A. (2003). Gender, Genre, and Writing Style in Formal Written Texts. *Text and Talk* 23 (3), 321-346.

- Atkinson, D. (1996). The Philosophical Transactions of the Royal Society of London, 1675-1975: A sociohistorical discourse analysis. *Language in Society*, 25, 333-371.
- Atkinson, D. (1999). *Scientific discourse in sociohistorical context: The Philosophical Transactions of the Royal Society of London 1675-1975*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Bazerman, Ch. (1988). *Shaping Written Knowledge: The Genre and Activity of the Experimental Article in Science*. Madison: University of Wisconsin Press.
- Besnier, N. (1994). Involvement in linguistic practice: An Ethnographic Appraisal. *Journal of Pragmatics*, 22, 279-299.
- Biber, D. (1988). *Variation across Speech and Writing*. Cambridge: Cambridge University Press.
- Biber, Douglas. (2001). Dimensions of variation among eighteenth-century speech-based and written registers. In D. Biber and S. Conrad (Eds.), *Variation in English: Multi-Dimensional Studies*, (pp. 200–214). Essex: Pearson Education.
- Bradbury-Jones, C; Irvine, F. and Sambrook, S. (2007). Unity and Detachment: A Discourse Analysis of Doctoral Supervision. *International Journal of Qualitative Methods*, 6(4), 81-96.
- Craig, H. (Ed.) (1916). *History of the Royal Society*. <http://www.bartleby.com/209/587.html> [03/05/2015]
- Crespo, B. and Moskowich, I. (2010). CETA in the Context of the *Coruña Corpus*. *Literary and Linguistic Computing*, 25(2). 153.164.
- Crespo, B. (2015). Women writing science in the eighteenth century: a preliminary approach to their language in use. *Anglica. An International Journal of English Studies*. 24(2), 103-114.
- Daston, L. and Galison, P. (1992). The Image of Objectivity. *Representations*, 40, 81-128.
- Dickerson, F.C. (2009). *The Way We Write is All Wrong. A Profile of and Prescription for Fixing The Broken Discourse of Fund Raising*. Unpublished Dissertation. http://hightouchgreetings.com/uploads/The_Way_We_Write_is_All_Wrong.pdf [12/09/2015]
- Earle, R. ed (1999). Epistolary Selves: Letters and Letter-Writers, 1600–1945. *Warwick Studies in the European Humanities*. Aldershot: Ashgate.
- Hyland, K. (1998). *Hedging in Scientific Research Articles*. Amsterdam: John Benjamins.
- Hyland, K. (2000). *Disciplinary Discourses: Social interactions in academic writing*. Harlow: Pearson Education.
- Hyland, K. (2002). Authority and invisibility: Authorial identity in academic writing. *Journal of Pragmatics*, 34, 1091-1112.
- Hyland, K. (2008). Disciplinary voices: Interactions in research writing. *English Text Construction*, 1(1), 5-22.
- Lakoff, R.T. (1990). *Talking power: The politics of language in our lives*. New York: Basic Books.
- Lareo, I. and Montoya, A. (2007). Scientific Writing: Following Robert Boyle's Principles in Experimental Essays –1704 and 1998. *Revista Alicantina de Estudios Ingleses* 20, 119-137.

- Mischke, G.E. (2009). *Analysing involvement in distance-education study-guides: an appraisal-based approach*. UNISA. <http://uir.unisa.ac.za/bitstream/handle/10500/1864/00front.pdf?sequence=3> [25/09/2015]
- Moskowich, I. (2013). Eighteenth-century female authors: women and science in the *Coruña Corpus of English Scientific Writing*. *Australian Journal of Linguistics*, 33(4), 467-487.
- Moskowich, I; Lareo, I; Camiña, G. and Crespo, B. (2012). *A Corpus of English Texts on Astronomy (CETA)*. Amsterdam: John Benjamins.
- Moskowich, I. and Crespo, B. (2015). Involved in writing science: nineteenth-century women in the *Coruña Corpus*. *International Journal of Language and Linguistics*.
- Narrog, H. (2012). *Modality, Subjectivity, and Semantic Change: A Cross-Linguistic Perspective*. Oxford: Oxford University Press.
- Newton, I. (1704). *Opticks*. London : Printed for Sam Smith, and Benj. Walford, printers to the Royal Society, at the prince's Arms in St. paul's Church-yard. <http://archive.org/details/opticksortreatis00newt> [09/03/2013].
- Oxford English Dictionary Online*. 2nd edition. Oxford University Press. <http://www.oed.com/>.
- Prelli, L.J. (1989). The rhetorical construction of scientific ethos. In H.W. Simon (Ed.), *Rhetoric in the human science*, (pp. 87-104). London: Sage.
- Tieken-Boon van Ostade, I. (2006). English at the onset of the normative tradition. In L. Mugglestone (Ed.), *The Oxford History of English*, (pp. 240-273). Oxford: Oxford University Press.
- Utz, J.P.; German, J.L; Louria, D.B; Emmons, Ch.W. and Bartter, F.C. (1959). Pulmonary Aspergillosis with Cavitation — Iodide Therapy Associated with an Unusual Electrolyte Disturbance. *New England Journal of Medicine*, 260, 264-268. doi: 10.1056/NEJM195902052600603.