Chapter 3

CETA as a tool for the Study of Modern Astronomy in English

Isabel Moskowich

1. Introduction

Modern culture can be said to depend on writing to such an extent that if scientific knowledge is not written it simply does not exist (Hyland, 1998). It is easy to observe, when glancing through any more or less old book, that conventions and practices have been subject to change, but such changes are not necessarily random. Hyland (1998: 18) claims that "The conventional linguistic practices for expounding and securing support for scientific knowledge are historical artefacts which date from the 1600s". And some scholars of the time saw it necessary to establish such discursive rules (as Boyle and his colleagues did when they proposed to separate the exposition of hypotheses and that of proven facts). As a consequence of the application of such discursive patterns, a particular reading public appeared. The subtle negotiation of knowledge that may be observed from the seventeenth century onwards as never before in history is therefore reflected in language and discourse as a vehicle for such negotiation. Compiling a corpus of scientific writing in Modern English seems a plausible idea as a means to study the development of the English language as well as the development of Science.

The Project *Coruña Corpus: A Collection of Samples for the Historical Study of English Scientific Writing* includes texts of a scientific character belonging to different fields of knowledge. The corpus has been conceived of as a collection of sub-corpora, one for each scientific discipline. The Coruña Corpus (CC) is a long- term project that will be coming out little by little, its first part being the *Corpus of English Texts on Astronomy (CETA)*. As many others, ours is a purpose-built electronic corpus conceived of as a means to have material for the study of scientific writing in English long before it became the *lingua franca* of science. Different pilot studies have demonstrated this part of the *Coruña Corpus* is a reliable tool for the study of the evolution of scientific writing in the field of Astronomy. *CETA* seems to be useful to observe and account for the linguistic habits of English-speaking astronomers as well as other factors relating to the way in which knowledge was conveyed depending on several extra-linguistic variables.

This first part of our greater *Coruña Corpus of English Scientific Writing* may,

therefore, provide researchers with important tools for a better understanding of all the issues just mentioned and some others to appear in the course of time. It has been built up by selecting samples of texts published between 1700 and 1900 (that is to say, the Modern English period[1]). Such text extracts can be of interest not only for linguists but for historians of Science as well. The Corpus offers the possibility of carrying out studies from a diachronic perspective but, since different genres have been included, it is suitable for other kinds of approaches, and comparative studies are obviously amongst them. As an electronic resource, *CETA* can be grouped together with other computerised corpora that have been designed to give access to specific insights into the so-called special languages, either on their own or by resorting to other available complementary tools.

This chapter is intended to give an account of some aspects relating to the compilation process and the corpus itself, and to the way in which the whole thing, as part of a major project, has been structured.

2. The principles governing *CETA*

To sketch the principles of *CETA* is to sketch those of the Coruña Corpus of English Scientific Writing. This is a collection of text samples that have been carefully selected put together (not "arbitrarily cut-out smaller text chunks" as the Lampeter Corpus manual says, Claridge *et al.*, 1999) in order to represent the particular manifestation of English in science writing, at a particular time and with the intention to facilitate research tasks. Both *CETA* and the *Coruña Corpus* as historical corpora are limited since we can only resort to written material, but we are pretty sure about their usefulness and this is, precisely, why we have devoted our efforts to establish the principles that are presented below.

We have included samples representing other categories besides the widely studies research article. Textbooks constitute an essential part in the transmission of

---

[1] Alternative dates such as 1660, 1725, 1776 or even 1800 (Görlach, 1994: 22) have been pointed as the borders between early and late Modern English. It is true that from the 18th century English scholars tend to use prescribed forms regardless of their dialectal origin. Regional and social dialects are considered inferior (Freeborn, 1992: 180). Besides, it is in the eighteenth century that we observe the outburst of all sorts of pamphlets, grammars and articles aiming at linguistic improvement.

scientific knowledge at least in the initial stages of academic life. However, they have seldom been the object of linguistic analysis except to compare them with other academic genres (Trauth, 1990; Littlefair, 1991).

Variation inside scientific and academic English can be also seen in the different text types —conceived of as having certain internal characteristics— and genres —as a way of socialising and, therefore, with certain external functions (García-Izquierdo & Montalt, 2002). In other words, inside one single discipline or domain discourse may undergo several modifications and changes depending on the genre to which that discourse sample belongs. Nwogu (1990) demonstrated that medical articles of an informative sort and their academic equivalents are composed differently; Myers (1990) found similar differences between informative and research biology articles. Bhatia (1993) also presented the linguistic differences existing between texts on legislation and those reporting cases and addressed to professional lawyers. Delimiting genres is complicated and it represents a problem that, according to Monzó (2002: 141) can be solved if we assume that genres may be sub-classified in variants attending to cultural criteria. She proposes the following:

a. Paragenre: genre belonging to one professional community

b. Diagenre: analogous genre in a culture that can be identified and recognised by the territory it occupies

c. Idiogenre: genre reflecting a particular author's idiosyncrasy in his/her texts in a constant way.

According to this, it is paragenres we are dealing with in *CETA*. Our classification of samples, therefore, has been based not only on linguistic characteristics but also on epistemological features and social functions. This way, we could say that, as compilers, we have searched for samples in different epistemological levels more or less (though not completely) equivalent to the three to be found today (Fortanet et al., 1998):

a. Highest epistemological level typical of research articles and *abstracts*;

b. High epistemological level (*abstracts* in *abstracting journals* and informative scientific articles);

c. A medium epistemic level for specialised non-academic articles.

3. Aim, scope and decisions on representativeness of *CETA*

As seems to be the general trend in Europe now, *CETA* is part of a bigger Project[2] aiming at offering a general view of the development of English Scientific Writing in several different disciplines from the beginning of the eighteenth century onwards. The time-span chosen is directly related with the foundation of the Royal Society of London and, of course, with the publication of the basic guidelines on how to present scientific works to the members of the Society with the ideas of clarity and simplicity behind it all. The political situation after the Restoration and the social changes it involved must be also borne in mind.

It is also worth mentioning that *CETA* in particular, as well as the whole *Coruña Corpus* in general (when completed), aims at covering the gap left by other historical corpora. That way, for instance, it will cover, in terms of chronology, what is left by the *Lampeter Corpus*, since the latter covers the 100-year period from 1640 to 1740 but does not go beyond that date. In terms of domain it is also more specific than the *Lampeter Corpus* that represents Science in general rather than particular disciplines or than the *Helsinki Corpus of English texts* that was not conceived of as a "specific" corpus in the same sense. As for disciplines, *CETA* represents a single discipline as is the case of *MEMT (Middle English Medical Texts)*.

Though both contain scientific writing, the *Coruña Corpus* and the ARCHER (*A Representative Corpus of Historical English Register*) do not overlap since the latter has material extracted from the Philological Transactions whereas the former (and that is the case of *CETA*) offers a representation of American English and longer formats.

When faced with the idea of compiling a second-generation corpus we first turned to the *UNESCO Classification of Science and Technology* (1988) in order to have a starting point of some kind. Of course, this classification cannot be applied to the Modern period without adapting it since the compartmentalisation and taxonomy of science in the 20th century is dramatically different from the one just beginning to develop in Modern Times in Europe. It is obvious that knowledge then, though far away from medieval scholasticism and seriously committed to the Scientific Method established by Empiricism, was still not organised as it is three centuries later.

---

[2] *CETA* is part of the Coruña Corpus of English Scientific Writing, a larger corpus under compilation. It will contain different subcorpora pertaining to different disciplines or domains, such as Philosophy, Mathematics, Life Sciences, History, etc.

UNESCO's taxonomy reflects the need for classification originated with Rationalism and is nothing but a result of it. This is why the field of Astronomy and Astrophysics that UNESCO divides in Cosmology and Cosmogony (code 2101), Planetary medium (code 2102), Optical Astronomy (code 2103), Planetology (code 2104), Radio-Astronomy (code 2105), Solar System (code 2106) and Other Astronomical Specialities (code 2199) do not exactly match Astronomy and Astrology, the disciplines more widely represented in *CETA*.

3.1 Categories represented in CETA

In order to obtain a representative corpus of writings on Astronomy during the Modern English period we tried to find samples that could be considered representative of all possible forms, genres and/or text-types (McEnery and Wilson 1996; Biber et al. 1998: 251-253).

In compiling this corpus, it was our intention to provide samples representative of the textual reality in the eighteenth and nineteenth centuries in English-speaking countries. In all cases we have resorted to prose texts only and all of them edited and printed. In an effort to obtain a really truthful representation of the language used in Astronomy texts, our first decision was to resort only to first editions taking two samples every ten years. We have resorted to first editions whenever possible and, when not, we have chosen those that were published within less than thirty years from the date the work first came out following Kytö, Rudanko and Smittenberg's (2000: 92) assumption that language change can be observed within 30-year periods.

Though it has been said that 1,000-word samples are more than enough for the study of variation within the scientific register (Biber, 1993), we were conscious that some of the available "types" of texts were not that technical or scientific[3]. Besides, the scientific register was not as standardised as it is nowadays so that the possibilities to find variation are greater. In many aspects, texts seem less repetitive both in their structure and in their lexical choice. In that case, larger samples would provide a better idea of the type of language used as well as serve the purpose of comparison with other non-scientific samples (Lareo, 2006; Lareo & Moskowich, 2007). Our samples contain

---

[3] That is the case, for instance of John Harris's *Astronomical dialogues between a gentleman and a lady: wherein the doctrine of the sphere, uses of the globes, and the elements of Astronomy.*

also non-analysable items that have been represented too though not counted. That is the case of zodiacal signs, references to points or coordinates in the sky or even quotations whose deletion would have caused the non-understanding of texts but that are, by no means, representative of the language of the author.

English Scientific writing developed from very early in the Modern Age. Though Latin was still the language of learning and science until well into the eighteenth century the vernacularisation process had long begun (Taavitsainen and Pahta, 2004) and was by now well established. Scientific writing was being also produced in English as a first option and this is why *CETA* contains only samples directly written in English rather than translated from Latin or any other language even if the translator and the author were the same person. We thought that this was safer to avoid any linguistic interference. Already in the Middle Ages the conventions of Latin scientific writing were translated into the vernaculars of Europe (see Crossgrove 1998, Pahta and Taavitsainen 2004), but during the Enlightenment and soon after it, scientific and technological matters were dealt with as a matter of economic benefit (Jardine, 1999) and maybe this too in some way accelerated the publication of scientific works in English (which was, no doubt, the language of sponsors). As a matter of fact, knowledge is socially constructed (Hyland, 1998: 13) and as a construct it serves a social and economic goal.

We aimed at collecting samples representing the different predominant written manifestations of this knowledge (excluding poetic and fiction pieces) existing in the time-span covered by *CETA*. One of our main concerns (it still is) was to set the limits between genre, text-type, register and style. In the definition of genre the elements directly depending on the texts (form, style, purpose) are not enough. Moessner (2001: 132) claims that it is basic to consider also the reader's perspective, that is to say, "which features make a reader interpret a text as a prototypical novel, short story, parody, etc?". Genre division must count on extralinguistic factors such as subject matter, purpose and discourse situation (Rissanen, 1996).

As Görlach (2004: 1) states in his preface "Proper definitions, and investigations including diachronic developments and diatopic contrasts seem to be indispensable before, for instance, corpus linguistics can claim to make reliable statements based on a representative text selection". Therefore, we tried to outline the picture of which were the existing functional text categories at the time (that is to say, which were the different

genres[4] to be produced in eighteenth- and nineteenth-century Astronomy) and we also checked if that production was more or less homogeneous in all English-speaking countries at the time, or if some types existing in, say, England, were not to be found in the United States. Our hypothesis is that no significant difference is to be observed in the genres considered for the discipline Astronomy neither as period or place of writing are concerned.

As in the ones compiled in the *Lampeter Corpus* various functions can be found in our texts: the informative function is the commonest, but the instructive, and even the entertaining, functions are not uncommon either. Therefore, we resorted to the following categories: Essay, Treatise, Textbook, Lecture, Letter, Dialogue, Article, Encyclopaedia or Others when they present miscellaneous features.

Görlach (2004: 88), though referring to them as text-types, defines these types as follows:

a. Article: non-fictional composition or dissertation in a newspaper, journal or read at conferences.

b. Dialogue: Literary work in conversational form

c. Encyclopaedia: book containing information in all branches of knowledge, arranged alphabetically.

d. Essay: short prose composition, first draft

e. Lecture: formal discourse delivered to students. Piece of writing intended to be read aloud.

f. Letter: written communication (not necessarily sent by post)

g. Treatise: discussion of some topic including some methodological issues

h. Textbook: book used as a standard book

Of all these, only two are not considered as major or more generic text-types by Görlach, namely, encyclopaedia and dialogue. When our samples seemed not to fit in any of the above, they have been classed as "others".

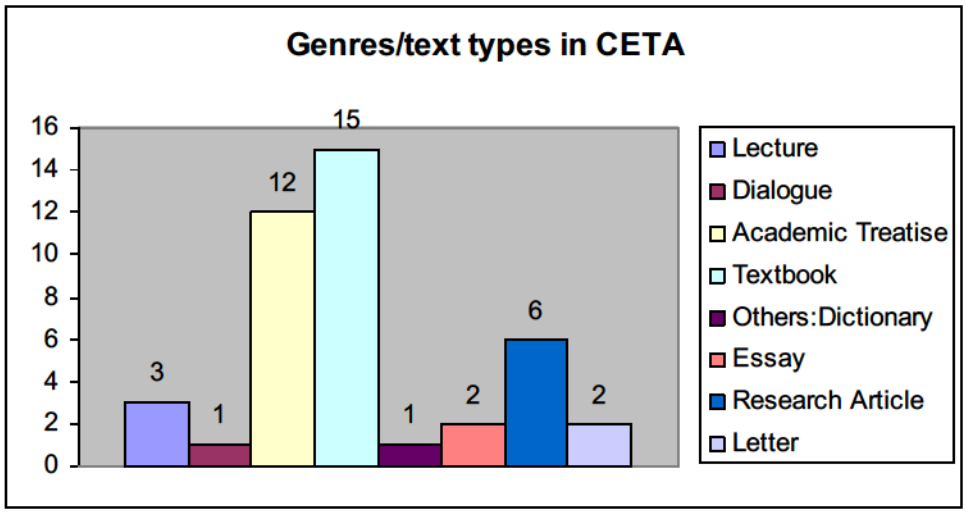
All these categories were already in use when our authors were producing their texts and many of them have been recorded as early as the end of the fourteenth century.

---

[4] The ascription of a sample to one or another genre is arguable. As Fowler (1982: 41) puts it genres may be considered as family members who "are related in various ways without necessarily having any single feature in common by all".

Such is the case of treatise, first recorded with its meaning of "a book or writing which treats of some particular subject" (*OED*). In modern times, however, the meaning includes also the idea of a book "containing a formal or methodological discussion or exposition of the principles of the subject". In this respect, many of the samples contained in *CETA* belong to treatises (as is the case of Curson, 1702, *The Theory of Sciences illustrated, or the grounds and principles of the seven arts; grammar, logick, rhetorick, musick, arithmetick, geometry, Astronomy. Accurately demonstrated and reduced to practice,* whose title is most illustrative). One of the more painstaking tasks of the team, no doubt, was the classification of samples, though we are pretty sure they are all there...

Graph 1. illustrates the different genres gathered in *CETA* samples.



Graph 1. Genres in CETA

All the different categories we have gathered seem to reflect the social reality of a world in which knowledge was not exclusive of Universities or other institutions (where the taxonomies for lecture, treatise and textbook/handbook would perfectly fit), but was also wanted outside such institutions as was mentioned earlier in sections 1 and 2. The vernacularisation of science and technology brought about its popularisation too and new ways of communication had to be used. Letters, dialogues and other forms were also found though, obviously, not all disciplines were so prone to be spread just because they were not equally popular. Let us add to this the already mentioned idea of economic benefit in a country about to commence the Industrial Revolution with all its socio-economic and political implications and the panorama is complete.

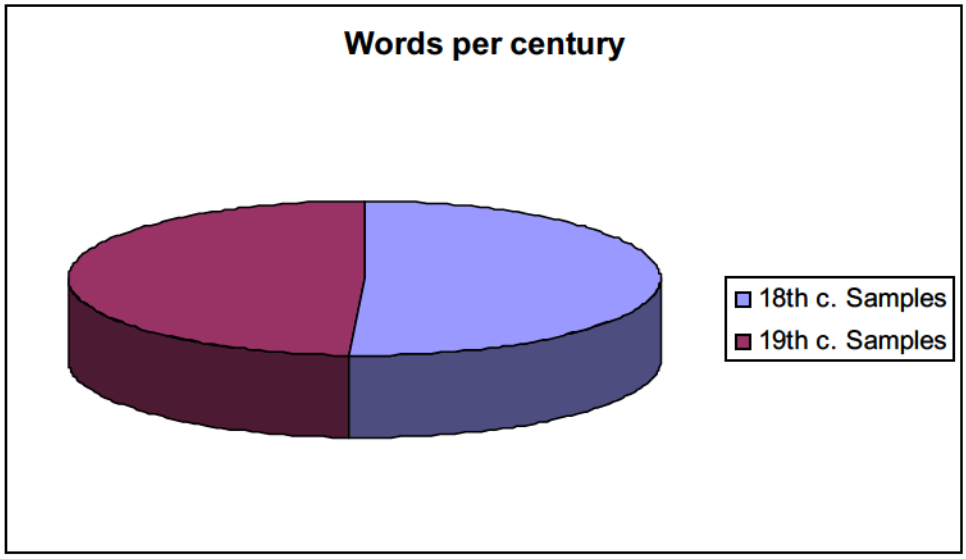
In general we can say that this distribution displayed in graph 1 reflects the production at the time (Görlach, 2004: 1). If there are more samples from treatises and lectures it is simply because Astronomy, as part of the Quadrivium, was not one of the most popular fields and it was mainly communicated in an academic setting[5].

3.2 Representativeness and size

We cannot agree with Claridge (1999) when in her introduction to the *Lampeter Corpus* she says that they have taken complete texts because any other option would have been "arbitrarily cut-out smaller text chunks" put together. Our samples have been selected so that all parts of texts (introductions, central chapters and conclusions) are more or less equally represented. Arbitrariness is to be found only in the choice of the paragraph that would be the last one keyed in and that depended on word counts.

We have tried to compile a similar number of words and samples for each century. Therefore, we have obtained a total of 208,083 words for the eighteenth century part and 202,403 for the nineteenth-century one. The same number of samples, twenty one, appears for the two periods. However, not all genres/text types are equally represented in the sense that articles are more common in the later part of the Corpus and the only sample we have in the category "others" corresponds to the eighteenth century. Of course, this is just reflecting the surrounding reality.
The graph below shows the overall distribution in terms of word counts:

[5] Other disciplines were more easily found outside specialists circles, as happened with Natural History where we find texts where the description of the anatomy of fish is mixed with recipes explaining how to cook them properly (Dodd, 1752).



Graph 2. Words in CETA

Of course, it goes without saying that this selection has often been determined by the availability of texts.

3.3 Authors represented in CETA

Some works from the nineteenth century which were notably shorter (articles) have been included "in toto" but that does not unbalance our corpus in any way. The list below (table 1.) presents authors in alphabetical order accompanied by the title of the text sampled and the corresponding file name contained in the Corpus.

Adams, George. 1777. *A Treatise describing the construction and explaining the use of celestial and terrestrial globes*. London.	astr 1777 Adams 1-57.xml
Bartlett, William Holms Chambers. 1855. *Elements of natural philosophy (Spherical Astronomy*. New York: Barnes & Co.	astr 1855 Bartlett 1-33.xml
Bonnycastle, John. 1786. *An introduction to Astronomy in a series of letters from a preceptor to his pupil, in which the most useful and interesting parts of the science are clearly and familiarly explained.* London.	astr 1786 Bonnycastle 19-68.xml
Bradford, Duncan. 1845. *The wonders of the heavens, being a popular view of Astronomy, including a full illustration of the mechanism of the heavens; embracing the Sun, Moon, and stars.* Boston: American Stationers Co.	astr 1845 Bradford 82-95.xml
Brewster, Sir David. 1811. *Ferguson's Astronomy explained upon Sir Isaac Newton's Principles: with notes and supplementary chapters.* Philadelphia.	astr 1811 Brewster 241-277.xml
Bryan, Margaret. 1797. *A compendious system of Astronomy in a course of familiar lectures: in which the principles of the science*	astr 1797 Bryan 91-

are clearly elucidated so as to be intelligible to those who have not studied mathematics: also trigonometrical and celestial problems, with a key to the ephemeris, and a vocabulary of the terms of science used in the lectures which later are explained agreeably to their application in them. London.	122.xml
Charlton, Jasper. 1735. *The Ladies Astronomy and Chronology in four parts*. London.	astr 1735 Charlton 13-53.xml
Chauvenet, William. 1871. *A manual of spherical and practical Astronomy, embracing the general problems of spherical Astronomy, the special applications to nautical Astronomy, and the theory and use of fixed and portable astronomical instruments, with an appendix on the method of least squares*. Vol I. Philadelphia	astr 1871 Chauvenet 9-37. xml
Clerke, Agnes Mary. 1893. *A Popular History of Astronomy in the Nineteenh Century*. London.	astr 1893 Clerke 300-329.xml
Costard, George. 1767. *The history of Astronomy, with its application to geography, history, and chronology; occasionally exemplified by the globes*. London: James Lister.	astr 1767 Costard 270-298.xml
Croll, James.1889. *Stellar Evolution and its relation to Geological Time*. New York: Appleton.	astr 1889 Croll 12-52.xml
Curson, Henry. 1702. *The Theory of Sciences illustrated, or the grounds and principles of the seven arts; grammar, logick, rhetorick, musick, arithmetick, geometry, Astronomy. Accurately demonstrated and reduced to practice*. With a variety of questions, problems and propositions both delightful and profitable. London: Richard Smith.	astr 1702 Curson 337-400.xml
Darwin, George Howard. 1880. On the Secular Changes in the Elements of the Orbit of a Satellite, revolving about a Tidally Distorted Planet. *Philosophical transactions*. London.	astr 1880 darwin 864-873.xml
Ewing, John. 1809. *A plain, elementary and practical system of natural experimental philosophy: including Astronomy and chronology*. Philadelphia: Hopkins and Earle.	astr 1809 Ewing 492-523.xml
Ferguson, James. 1756. *Astronomy explained upon Sir Isaac Newton's Principles and made easy to those who have not studied Mathematics*. London.	astr 1756 Ferguson 146-167.xml
Fuller, Samuel. 1732. *Practical Astronomy, in the description and use of both globes, orrery and telescopes wherein the most useful elements, and most valuable modern discoveries of the true Astronomy are exhibited, after a very easy and expeditious manner, in an exact account of our solar system, with ten curious copper plates*. Dublin.	astr 1732 Fuller 1-27.xml
Garland, Landon C.1838. "An Address on the Utility of Astronomy". *Southern Literary Messenger; devoted to every department of literature and the fine arts*. Richmond, VA: T.W. White (etc.), Volume 4, Issue: 2, Feb 1838.	astr 1838 Garland 123-130.xml
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*. London: J. Senex; G. Strahan; W. and J. Innys; J. Osborn and T. Longman.	astr 1726 Gordon 63-99, 101-123.xml

Gummere, John. 1822. *An elementary treatise on Astronomy. In two parts. The first, containing a clear and compendious view of the theory. The second, a number of practical problems. To which are added, Solar, Lunar and some other Astronomical Tables.* Philadelphia.	astr 1822 Gummere 200-237.xml
Harris, John. 1719. *Astronomical dialogues between a gentleman and a lady: wherein the doctrine of the sphere, uses of the globes, and the elements of Astronomy.* London: T. Wood.	astr 1719 Harris 1-52.xml
Herschel, John F. W. 1833. "A treatise on Astronomy". *The Cabinet Encyclopedia. Conducted by the Rev. Dionysius Lardner ... Assisted by eminent literary and scientific men. Natural Philosophy.* London: Longman, Rees, Orme, Brown, Green and Longman, and John Taylor.	astr 1833 Herschel 205-251.xml
Hill, John. 1754. *Urania: or, a compleat view of the heavens; containing the antient and modern Astronomy, in form of a dictionary: illustrated with a great number of figures comprising all the constellations, with the stars laid down according to their exact situations and magnitudes: from repeated and accurate observations (Vol.I. Being the first of A compleat system of natural and philosophical knowledge*). London: T. Gardner.	astr 1754 Hill 1-17-xml
Hodgson, James. 1749. *The Theory of Jupiter's Satellites: with the construction and use of the tables for computing their eclipses.* London: W. and J. Mount and T. Page on Tower Hill, and H. Whitridge.	astr 1749 Hodgson 83-111
Lacy, John. 1779. *The universal system: or mechanical cause of all the appearances and movements of the visible heavens: shewing the true powers which move the Earth and Planets in their Central and annual Rotations with A Dissertation on Comets, the Nature, Cause, Matter, and Use of their Tails, and the Reasons of their Long Trajectories; likewise and attempt to prove what it is that moves the Sun around its Axis.* London: J. Auckland.	astr 1779 lacy 1-35.xml
Long, Roger. 1742. *Astronomy, in five books*. Cambridge.	astr 1742 Long 61-82.xml
Loomis, Elias. 1868. *A treatise on Astronomy.* New York: Harper and Brothers (???	astr 1868 Loomis 9-37.xml
Lowell, Percival. 1895. "Mars: Canals". *The Atlantic Monthly: Mars III. Canals*. Vol 76 July 1895: 106-119	astr 1895 Lowell 106-119
Luby, Thomas. 1828. *Introductory Treatise on Physical Astronomy.* London and Dublin: Baldwin and Cradock.	astr 1828 Luby 1-34.xml
Mitchel, Ormsby McKnight. 1860. *Popular Astronomy. A concise elementary treatise on the Sun, planets, satellites and comets.* New York: Phinney, Blakeman & Mason.	astr 1860 Mitchel 864-873.xml
Morden, Robert. 1702. *An Introduction to Astronomy, geography, navigation, and other mathematical sciences made easier by the description and uses of the celestial and terrestrial Globes. In seven parts.* London.	astr 1702 Morden 1-42.xml
Nicholson, William. 1782. *An introduction to natural philosophy. Illustrated with copper plates*. London: Printed for J. Johnson.	astr 1782 Nicholson 100-151, 154-156.xml

Olmsted, Denison. 1841. *Letters on Astronomy, addressed to a lady in which the elements of the science are familiarly explained in connexion with its literary history. With numerous engravings.* Boston: Marsh, Capen, Lyon and Webb.	astr 1841 Olmstead 312-339.xml
Phillips, William. 1817. *Eight familiar lectures on Astronomy: intended as an introduction to the science: for the use of young persons and others not conversant with the mathematics.* New York: James Eastburn and Co.	astr 1817 Phillips 46-80.xml
Small, Robert. 1804. *An account of the astronomical discoveries of Kepler: including an historical review of the systems which had successively prevailed before his time.* London: J. Mawman.	astr 1804 Small 70-105.xml
Steele, Joel Dorman. 1874. *Fourteen weeks in descriptive Astronomy.* New York: A.S. Barnes.	astr 1874 Steele 13-42, 45-64.xml
Stewart, Matthew. 1761. *Tracts, Physical and Mathematical. Containing an Explanation of several important points in Physical Astronomy; and a new Method for ascertaining the Sun's distance from the Earth, by the Theory of Gravity.* Edinburgh: printed for W. Sands and A. & J. Bell/ London: A. Millar and J. Nourse.	astr 1761 Stewart 340-398.xml
Vince, Samuel. 1790. *A treatise on practical Astronomy.* Cambridge, Mass: J. Archdeacon, J. and J. Merrill, J. Nicholson and W. Lunn.	astr 1790 Vince 6-31.xml
Watts, Isaac. 1726. *The knowledge of the heavens and the Earth made easy: or, the first principles of Astronomy and geography explain'd by the use of globes and maps: with a Solution of the Common Problems by a plain Scale and Compasses as well as by the Globe*. London.	astr 1726 Watts 1-50.xml
Whewell, William. 1858. *The plurality of worlds. With an introduction by Edward Hitchcock. I Astronomical discoveries p17/ II Astronomical objection to religion.* New York; Boston	astr 1858 Whewell 17-51.xml
Whiston, William. 1715. *Astronomical lectures, read in the publick schools at Cambridge.* London: R. Senex and W. Taylor.	astr 1715 Whiston 1-37.xml
Wilson, Alexander. 1774. "Observations on the solar spots". *Philosophical Transactions,* Vol. 64: 1-19.	astr 1774 Wilson 1-19.xml
Young, Prof., LL. D., Ph. D. 1880. "Recent Progress in Solar Astronomy" *The Princeton Review*. Volume 1: 88-104.	astr 1880 Young 88-104.xml

Table 1. Authors in CETA

Women are seldom mentioned in History of Science books or in Biographical Dictionaries. It was not common to make public female activity in certain social fields and science (Astronomy in particular) was one of those traditionally defined as masculine[6]. This means many outstanding female scientists were never publicly recognised. It is difficult to trace their lives because many lost their own name when

[6] Vid. Chapter 2

married and some used a masculine pseudonym to make sure their work was taken seriously (Herrero, 2007: 75). Excluded from official science, women learnt by absorbing and listening to other women, from mothers to daughters. Academies first admitted women as late as the 20th century, which means these scientific societies evolved for 150 years without a female point of view. Therefore, female authorship is very difficult to establish. On some occasions women did not sign their own works, as is the case of the Catalogue of Stars by German female astronomers in the seventeenth century. Women should not observe the sky at night as this was considered an indecorous behaviour (Herrero, 2007: 82). Although women participated intensively in the field of Astronomy once the Copernican system was accepted, their access to study and scientific work was limited to the role of mere assistants.

As could be expected for scientific writing, then, women are not abundant as authors of scientific works in general and this is more so in Astronomy. Only two have been included in *CETA* though others are known to have existed. That is the case of Adelaide Ames (18??-1932), who worked as a researcher for Harvard Observatory under Harlow Shapley from 1923-1932. She published a comprehensive catalogue of 2,778 nebulae which could not be used for the type of linguistic research *CETA* wants to serve. Others are Mary Ibertson (d. 1914) and Mary Ashlay (c. 1880) or Elizabeth Bardwell (d. 1899). The two ladies we have chosen, Margaret Bryan for the eighteenth century and Agnes Mary Clerke for the nineteenth, signed the work they authored and both made significant research that resulted in important advances and discoveries[7].

3.4 Time-span represented

The time-span covered by *CETA* is based on extra-linguistic considerations, since we have resorted to changes in scientific thought rather than to changes in the language itself to set the limit dates of our text selection. Of course, changes in scientific thought imply changes in the way in which knowledge is conveyed (Moskowich & Parapar 2007; Crespo 2008b).

CETA earliest texts date back to 1700 when the old epistemological patterns based on authorities undergo a dramatic transformation (Taavitsainen and Pahta, 1997).

---

[7] The metadata files in *CETA* contain detailed information on their lives and work but it is worth mentioning here the fact that one of the biggest craters on the Moon is named Clerke's Crater after its discoverer.

This change in the transmission of knowledge marks the beginning of a new scientific era and, consequently, it was considered a good starting point for our compilation.

Modern science was thought of as a brand-new resource founded on what would to be labelled induction (John Stuart Mill would systematise his methods of induction some time later) as well as on an experimental methodology and the use of a mathematical language to convey new results. In this sense it was diametrically opposed to the scholastic trend, that is, the understanding of science as deduction from well-established principles. Empiricism promoted the development of Science outside Universities for the first time, probably favoured by a better economic situation and population demanding practical applications for scientific thought in the market. At the same time as the importance of religion decreases and that of quantification of data as a means to reach valid conclusions grows.

These social and methodological changes resulted in the conscious creation of a new language to transmit science on the part of authors (Swales, 1990), a representative sample of which we have tried to compile in *CETA*.

Our collection of texts stretches as far as 1900, when several facts really relevant for the History of Science took place. Such is the case of the discovery of the electron by J.J. Thompson in 1896, the crisis of the grounds of mechanical physics announced by Mach, Kirchhoff or Bolzmann in this same year, Planck's announcement of quantum mechanics, or Einstein's publication (be it his idea or Mileva Maric's original idea) of a paper proposing what is today called the Special Theory of Relativity in 1905. it is evident that all these discoveries can be compared to the ones establishing the turning-point in the seventeenth century we referred to above and they were also accompanied by a new conception in the way Science should be conveyed. In fact, Thomas Huxley chanpioned for a new scientific style at the 1897 International Congress of Mathematics. No doubt from that moment scientific discourse changed dramatically again.

3.5 Geographical distribution of samples

If we assume that the *Corpus of English Texts on Astronomy* is not only conceived of as a tool for the study the evolution of English scientific writing in time but also for that of variation depending on different sociolinguistic variables, it is obvious that the more we know about the texts compiled and their authors, the better. For this reason we have resorted, when possible, to texts by authors about whom we could find basic biographical information and, therefore, whose linguistic habits we could infer. Such

biographical information has been compiled in a set of metadata files accompanying the samples themselves. The structure of such files is such that information inside them is also searchable by our search engine *Coruña Corpus Tool (CCT)*. No biographical information has been provided in the metadata files only in those cases in which there was no text available (complying with the *Coruña Corpus* principles) for a particular tiime-span of whose author we had some information.

Our decision was also determined by another idea. The whole *Coruña Corpus of English Scientific Writing* in general and *CETA* in particular have been devised from a social-constructionist position (Hyland, 1998: 82).. In other words, we believe that knowledge and its creation depend on context, and this implies that science is the interpretation of the world of a particular individual (the scientist) and not an independent entity or an absolute truth. This being so language is a central element both for interpretation of facts and for its transmission.

We have selected English-speaking authors writing in English, avoiding even translations made by the authors themselves, thus making it more difficult to find samples. This was especially do for the eighteenth century, since many of them also used Latin and this may have caused some interference on their use of their native language. When speaking about the geographical distribution of authors we are referring mainly to the places where they were educated, since it is there we think they acquired the linguistic habits to be found in their writings as sampled in *CETA*.

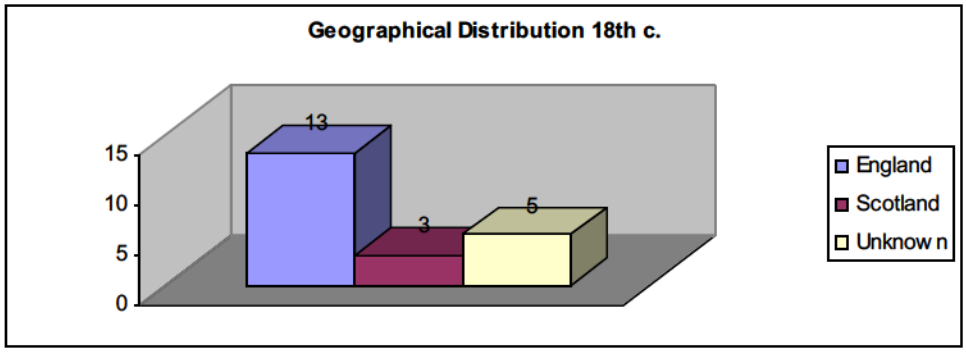
Table 2. below shows the distribution of authors according to the geographical variable (also contained in the metadata files in the *Corpus*) and, as can be seen, 33% are of American and 45% are of European origin. We have not been able to find information about the places where five authors were educated, all of them from the eighteenth century.

AUTHOR	**YEAR**	**PLACE OF EDUCATION**
Henry Curson	1702	?
Robert Morden	1702	England
William Whiston	1715	Cambridge (England)
John Harris	1719	Oxford (England)
George Gordon	1726	?
Isaac Watts	1726	Southampton and Newington (England)
Samuel Fuller	1732	?
Jasper Charlton	1735	?
Roger Long	1742	Cambridge (England)

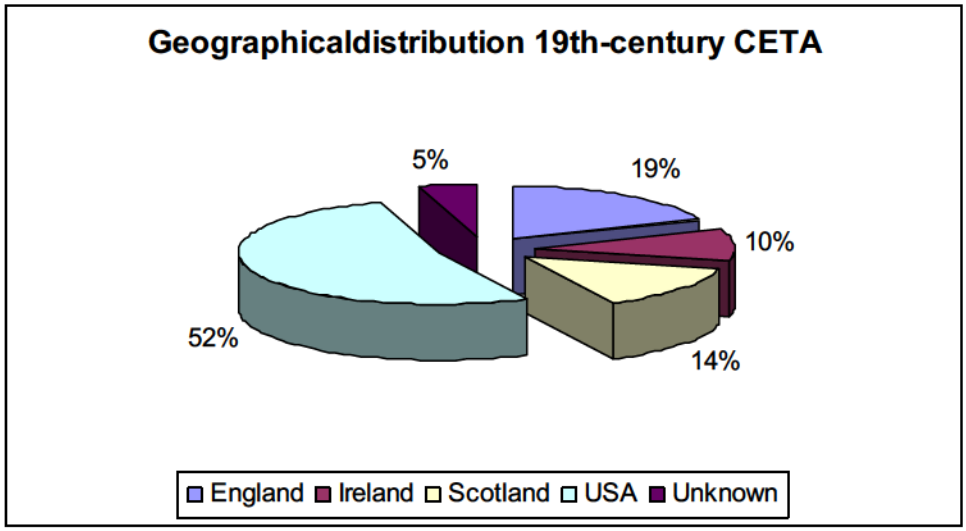
James Hodgson	1749	England
John Hill	1754	Peterborough/Westminster, UK
James Ferguson	1756	Banffshire (Scotland)
Matthew Stewart	1761	Rothesay (Scotland)
George Costard	1767	Oxford (England)
Alexander Wilson	1774	St. Andrews, Scotland
George Adams	1777	Southampton (England)
John Lacy	1779	?
William Nicholson	1782	North Yorkshire (England)
John Bonnycastle	1789	Buckinghamshire (England)
Samuel Vince	1790	Cambridge (England)
Margaret Bryan	1797	London (England)
Robert Small	1804	Dundee, (Scotland)
John Ewing	1809	Princeton, New Jersey (USA)
David Brewster	1811	Jedburgh and Edingburgh (Scotland)
William Phillips	1818	London (England)
John Gummere	1822	Moreland (USA)
Thomas Luby	1828	Dublin (Ireland)
John Frederick William Herschel	1833	Eton College and Cambridge (England)
Landon Cabell Garland	1838	Virginia, Alabama (USA)
Denison Olmsted	1841	East Hartford, Connecticut (USA)
Duncan Bradford	1845	?
William Holms Chambers Bartlett	1855	West Point, New York (USA)
William Whewell	1858	Lancaster, Cambridge (England)
Ormsby McKnight Mitchel	1860	Ohio (USA)
Elias Loomis	1868	Connecticut (USA)
William Chauvenet	1871	Philadelphia, Pennsylvania (USA)
Dorman Steele	1874	Syracuse, New York (USA)
George Howard Darwin	1880	Kent and Cambridge (England)
Charles Augustus Young	1880	New Hampshire (USA)
James Croll	1889	Perthshire (Scotland)
Agnes Mary Cerke	1893	Cork (Ireland)
Percival Lowell	1895	Boston (USA)

Table 2. Geographical origin of authors in CETA.

Graphs 3 and 4 below illustrate the fact that there are no American authors in the eighteenth century, though the situation is totally different for the nineteenth century: 52% of all authors come from the USA. English authors represent 19% and, very closely, Scottish authors represent 14%. It is evident that, once more, some external conditions have played an important role for this change: the spread of education and the social demand for practical results to improve the economy in the States.

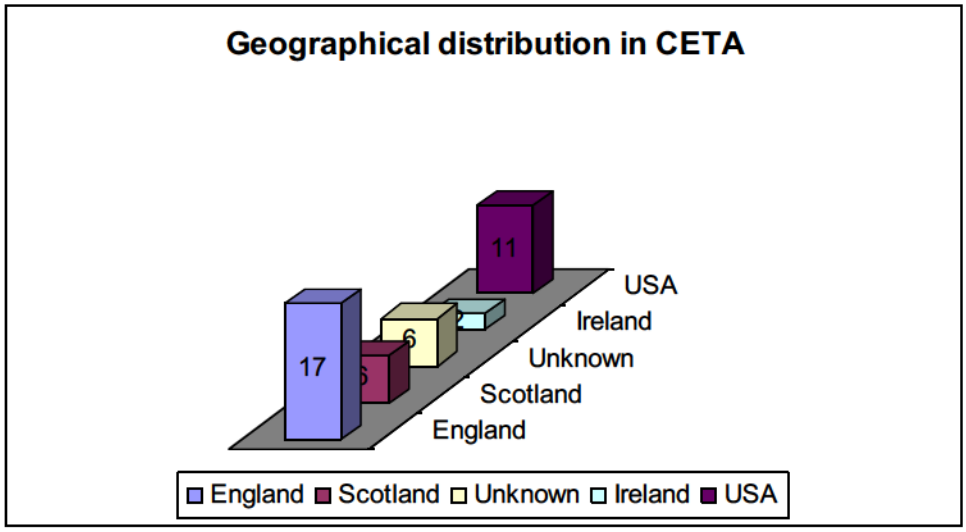


Graph 3. Geographical distribution in the 18th century



Graph 4. Geographical distribution in the 19th century

An overview of the different places where the authors contained in *CETA* learned to write is the one offered in graph 5.



Graph 6. The provenance of authors in CETA

4. Editorial policy and related software

Corpora constitute, in themselves, an editorial task. From the mere selection of one particular extract rather than another to the application of the different representation conventions to be used, many decisions have to be made. The texts in the *Coruña Corpus of English Scientific Writing* of which *CETA* is a part have been edited to represent even special graphemes in their XML format (visible in one of the windows of the *Coruña Corpus Tool, CCT*). In our effort to be as close to the original as possible we have also preserved certain symbols used to refer to constellations and other specific matters but have avoided the representation of all those elements that did not represent the language of the author, thus eliminating quotations

The difference in spelling across the two hundred years covered by the Corpus and which we wanted to represent as faithfully as possible implied that old-fashioned characters such as <ſ> (long <s>), <ſ> (italicised long <s>) or the ligatured digraph <ct>, made OCR under modern standards completely unfeasible. Therefore, manual typing was always needed at some stage.

We have tried to keep the output computerised text as truthful to the original as possible. However, we had to find a balance between two possibilities: a) showing the text the way it originally was as an image, and b) offering researchers the possibility of working with the information stored in the texts in an open, flexible and productive way. Balance between these two options implied taking some editorial decisions that will be accounted for at this point.

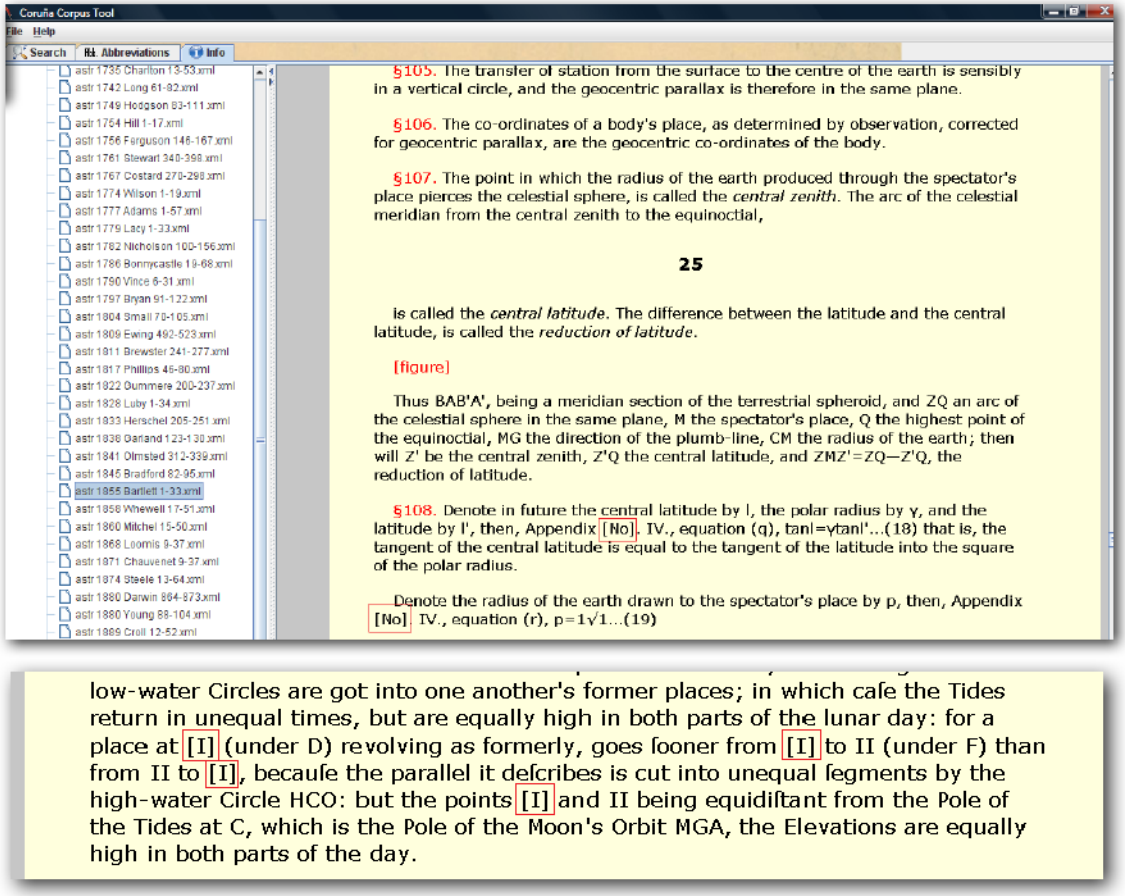
For each sample we have included TEI-compliant headers with information about the file, full name of the research group behind this corpus, sponsors and director, name of this Astronomy sub-corpus (*CETA*) and the amount of analysable material in the file. The header box concludes with a reduced version of the full title of the text, pages selected the name of the author and the year of publication. We have kept the original numbering of the text, our only alteration being the centring of all page numbers on the screen in a bold font type between blank lines. We have also resorted to a bigger bold blue font for titles and chapters to make the visual revision of texts more appealing to the researcher.

Our commitment to respect the original versions of the texts met some difficulties: we have nonetheless decided to omit editorial material -such as page headers, footers and margin notes- for the indexing though they can be read in the XML window. We have also got rid of extra blank spaces before some punctuation signs, such as colons and semi-colons. Finally, and whenever possible, a few spelling errors have been corrected, because they are likely to have been made by the printer rather than by the author. We have considered the different spellings across time and checked all the items in the *Oxford English Dictionary*. Those items impossible to identify or missing elements have been marked as [unclear].

Apart from the TEI tags we have included a set of editorial marks between square brackets in order to make analysis straightforward. They are discussed in the following paragraphs.

The square brackets contain information such as the location of quotations, figures, formulae, etc. in the original text. But, at the same time, they are used to disambiguate homographic forms that the *CCT* could consider a word. For instance, the Roman number *I* has been enclosed in brackets to avoid the miscounting of the personal pronoun *I*. Thus, in the wordlist generated by the *CCT* the first personal pronoun will appear, i.e. as i-325 (implying that the author has used 325 times the personal pronoun) and the Roman number *I* will appear as [i]-20 (showing that the Roman number *I* has been found only 20 times in that sample). Squeare brackets have been used for other strings of characters that could be ambiguous. For instance, the phrase *the points BE* will appear as *the points [BE]*; the abbreviation for number *No* will appear as [no] to distinguish it from the negative particle *no*, etc. (see Figure 1).

Figure 1: use of square brackets



Another editorial change regards a habit typical of eighteenth-century editors: the last word on one page was repeated as the first on the following. The texts samples in CETA, as well as in other sub-corpora of the CC have avoided such repetitions that would have altered word counts and frequencies.

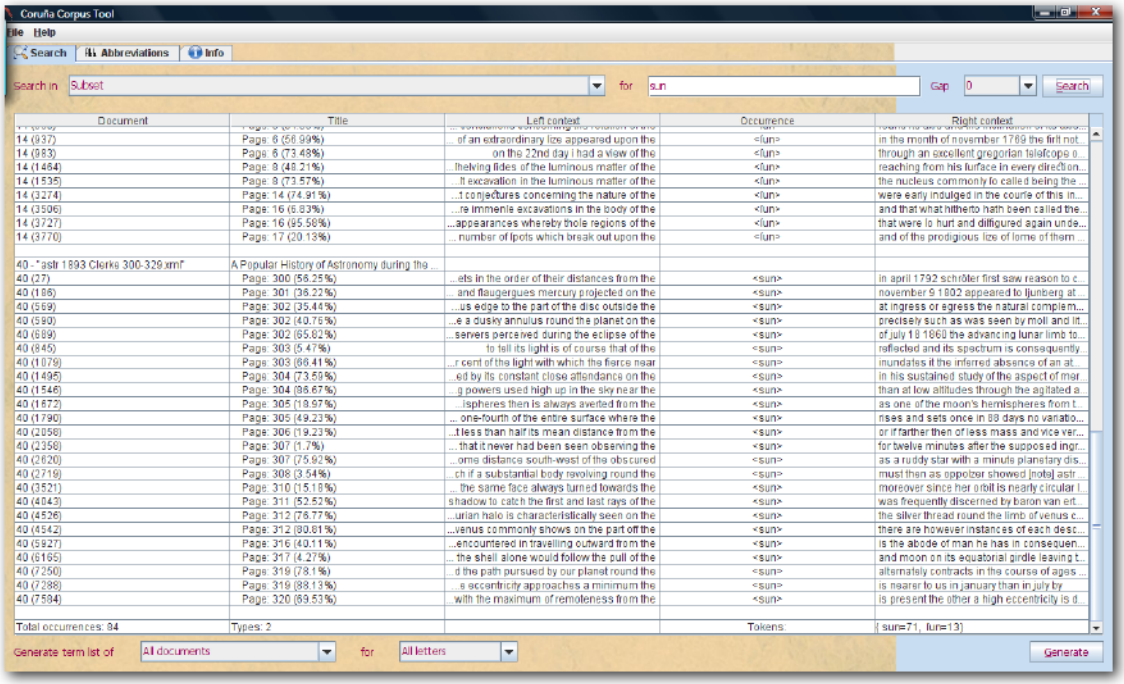
Original paragraphs, but not lines, are kept in *CETA* files. That implies the exclusion of truncated words at the end of a line[8]. There is only a case in which the original form of paragraphs cannot be respected: footnotes, since TEI restrictions prevent the division of a footnote into different paragraphs. Therefore, the information included in footnotes is written in one single paragraph in the *CETA* files. TEI restrictions affect also the place where a footnote appears in the Info Display window of the *CCT*. They are placed below the word they refer to, in a separate paragraph. We have also placed all note references after the word they refer to. This makes the electronic text easier to understand.

[8] Hyphens have been limited to compound words when they were hyphenated in the original. Therefore, when a hyphen has been used as a layout mark by the author or printer, an EM-dash has been placed instead

Other decisions concerning notes are the exclusion of editorial notes for indexing, though they can be viewed as we mentioned above, since they do not represent the author's own language[9].

One of the aims of the compilers was to keep original spelling variants. The *CCT* has been developed taking these variants into consideration. Therefore, when a searched word could have been written with two or more different spellings, the *CCT* shows all the possible spellings and distinguishes them as different types (see Figure 2, bottom row).

Figure 2: Search window. Spelling variants



Though the *Coruña Corpus of English Scientific Writing* in general, and *CETA* in particular, could be more easily searched if the files were left as raw XML files, the implementation of the CCT as its own search-engine has rendered better results, mainly in cases of disambiguation and in its treatment of editorial marks and punctuation. We hope it will be felt as frienfly and useful by researchers as it is by compilers.

[9] For a detailed list of the editorial marks used in *CETA*, see the Introduction to the Corpus.

Bibliography

A&A Language Editors. 2007. "Astronomy and Astrophysics English guide". *Astronomy and Astrophysics. An International Weekly Journal*, 466: 3.

Aaronson, Steve. 1977. "Style in Scientific Writing", *Essays on an Information Scientist.* Vol. 3, 1977-78: 4-13

Adams, George. 1777. *A Treatise Describing the Construction and Explaining the Use of Celestial and Terrestrial lobes*. Printed for and sold by the autor, at Tycho Brahe's Head, N. 60, in Fleet-Street.

Allen, Bryce, Qin, Jian. and F.W. Lancaster. 1994. "Persuasive Communities: A Longitudinal Analysis of References in the Philosophical Transactions of the Royal Society", 1665-1990. *Social Studies of Science* 24/2: 279-310.

Atkinson, Dwight. 1999. *Scientific Discourse in Sociohistorical Context. The Philosophical Transactions of the Royal Society of London, 1675-1975.* Mahwah, New Jersey: Lawrence Erlbaum Associates, Publishers.

Banks, David. "Emerging scientific discourse in the late seventeenth century. A comparison of Newton's *Opticks*, and Huygens' *Traité de la lumière"*. In *Functions of Language*, 2005: 65-86

Beal, Joan. 2004. *English in Modern Times, 1700-1945*. London: Arnold

Bello Piñón, Nuria. 2007. "Aiming to feature lectures: diachronic study of code-switching and borrowing in scientific English", http://ifa.amu.edu.pl/plm/files/Presentations/PLM2007_Presentation_Pinon.pdf. Accesed september 2008.

Bello Piñón, Nuria. 2008. "Switching the Code in Science: (extra)Linguistic Motivations for Code-switching in Late Modern English" paper delivered at the *Congreso Internacional sobre el Lenguaje de la Ciencia y la Tecnología,* Universitat Jaume I. Castellón (Spain) 16-18 October 2008.

Bello Piñón, Nuria. (forthcoming) "Code-switching in 18th century Scientific English: some considerations based on nonlinguistic features" *Proceedings of HEL-LEX 2*. Somerville, MA: Cascadilla Press.

Bhatia, V. 1993. *Analysing Genre: Language Use in Professional Settings*. London: London.

Biber, Douglas. 1993. *Representativeness in Corpus Design*. *Literary and Linguisitic Computing* 8 (4), 2243–257.

Biber, Douglas, Finnegan, E. & Atkinson, D. 1994. "ARCHER and its challenges:

compiling and exploring a representative corpus of historical English registers" In *Creating and using English language corpora. Papers from the 14th International Conference on English language research on computerized corpora,* Fries, U. Tottie, G and Schneider, P. (eds.), 1-13. Amsterdam: Rodopi

Biber, Douglas, Susan Conrad & Randi Reppen. 1998. *Corpus Linguistics*. Cambridge: Cambridge University Press.

Bonnycastle, John. 1789. *An introduction to Astronomy in a series of letters from a preceptor to his pupil.* London: printed for J. Johnson, No. 72, St. Paul's Church-yard.

Bradford, Duncan. 1845. *The wonders of the heavens, being a popular view of Astronomy, including a full illustration of the mechanism of the heavens; embracing the Sun, Moon, and stars.* Boston: American Stationers Co.

Camiña Rioboó, Gonzalo. 2005. "Tmesis and In(ter)fixation: the unity of words and/or morphemes in question". In *Reinterpretations of English. Essays on Language, Linguistics and Philology (II),* Moskowich, I. and Crespo, Begoña (eds.), 71-81. A Coruña: Universidade da Coruña.

Catholic Encyclopedia. http://www.newadvent.org/. Accessed 11th September 2008.

Chomsky, Noam. 1965. *Aspects of a Theory of Syntax*. Cambridge. Mass: MIT Press.

Claridge, Claudia, Josef Schmied & Rainer, Siemund. 1999. "The Lampeter Corpus of Early Modern English Tracts". In *ICAME Collection of English Language Corpora* (CD-Rom), Knut Hofland, Anne Lindebjerg, Jørn Thunestvedt, 2nd ed. The HIT Centre, University of Bergen, Norway.

Clerke, Agnes. 1893. *A Popular History of Astronomy in the Nineteenth Century.* London.

Costard, George. 1767 *The history of Astronomy, with its application to geography, history, and chronology; occasionally exemplified by the globes.* London: printed by James Lister; and sold by J. Newbery.

Crespo García, Begoña. 2004. "The scientific register in the History of English: a corpus-based study". *Studia Neophilologica*, 76/2: 125-139.

Crespo García, Begoña. 2008a. "Specific and non-specific nouns in lME: when Robert grows from man to herb". *English Studies,* 89/5:587-605.

Crespo García, Begoña. 2008b. "Astronomy, Language, Culture". Participation in the workshop entitled *The compilation of Astronomy texts in late Modern English: CETA in the context of the Coruña Corpus* at IULMA Conference, Castelló, 16-18 October, 2008.

Crespo García, Begoña. 2008c. "The transmission of 18th century astronomical knowledge: initiating new discursive paths?" *First ISLE Conference*. 8-11 October 2008. Freiburg (Germany).

Crespo García, Begoña. Forthcoming. "Rosewater, wheel of fortune: Compounding and lexicalisation in seventeenth century scientific terminology".

Crespo García, Begoña & Moskowich, Isabel. 2004. "Enlarging the lexicon: The field of technology and administration from 1150 to 1500". *Studia Anglica Posnaniensia,* 40: 163-180.

Crespo García, Begoña & Moskowich, Isabel. 2005. "Latin Forms in Vernacular Scientific Writing: Code-Switching or Borrowing?". In *Selected Proceedings of the 2005 Symposium on New Approaches in English Historical Lexis (Hel-Lex)*, McConchie, R. et al. (eds), 51-59. Somerville, MA: Cascadilla Press.

Crespo García, Begoña & Moskowich, Isabel 2007. "Medicine, astronomy, affixes and others: An account of verb formation". In *Some Early Scientific Works*. *SELIM,* 13: 179-197.

Crespo García, Begoña & Moskowich, Isabel. (forthcoming). "The limits of my language are the limits of my world: The scientific lexicon from 1350 to 1640".

Crespo García, Begoña & Moskowich, Isabel. 2008. *Revisiting the Principles of CETA. Advances in the Coruña Corpus,* Icame 29th International Conference, Ascona, 14-18 May. University of Zurich.

Crombie, Alistair. 1993. *Historia de la ciencia de San Agustín a Galileo*. Madrid: Alianza.

Crossgrove, William. 1998. "Introduction". In Crossgrove et al. (eds), 81-87.

Crossgrove, William, Margaret Schleissner & Linda Ehrsam Voigts (eds). 1998. *Early Science and Medicine: A Journal for the Study of Science, Technology and Medicine in the Pre-modern Period* 3/2, Special issue: *The Vernacularization of Science, Medicine, and Technology in Late Medieval Europe*. Leiden: Brill.

Curson, Henry 1702. *The Theory of Science Illustrated, or the Grounds and Principles of the Seven Arts; Grammar, Logick, Rhetorick, Musick, Arithmetick, Geometry, Astronomy. Accurately Demonstrated and Reduced to Practice.* London: Richard Smith.

Dodd, James S. 1752. *An Essay Towards a Natural History of the Herring*. London: printed for T. Vincent. Near Chancery-Lane, Fleet-Street.

Encyclopedia Britannica. 1943-1973. Macropedia. Vol. 2. Knowledge in Depth. 15th Edition. William Benton, Publisher, 245-250.

Fortanet, I., S. Posteguillo, J. C. Palmer & J. Coll. 1998. "Disciplinary variations in the writing of research articles in English". In *Genre Studies in English for Academic Purposes*, vol. 9, I. Fortanet, S. Posteguillo, J. C. Palmer and J. F. Coll (eds.), 59-78. Valencia: Universitat Jaume I.

Foster Jones, Richard. 1930. "Science and English prose style in the third quarter of the seventeenth century". *PMLA*, 45/4: 977-1009.

Fowler, Alastair. 1982. *Kinds of Literature. An Introduction to the Theory of Genres and Modes*. Oxford: Clarendon Press.

Freeborn, D. 1992. *From Old English to Standard English*. London: Macmillan.

García Izquierdo, I. & V. Montalt. 2002. "Translating into Textual Genres". *Linguistica Antverpiensia*, 1: 135-143.

Gillespie, Ch.C. (ed). 1973. *Dictionary of Scientific Biography*. New York: Cjarles Scribner's Sons.

Görlach, Manfred. 1994. *The Linguistic History of English.* London: Macmillan.

Görlach, Manfred. 2004. *Text Types and the History of English*. Berlin/New York: Mouton de Gruyter.

Gotti, Maurizio. 1996. *Robert Boyle and the Language of Science*. Milan: Guerini Cientifica.

Gross, Alan G., Joseph E. Harmon & Michael Reidy. 2002. *Communicating Science: The Scientific Article From the 17th Century to the Present*. Oxford: Oxford University Press.

Gummere, John. 1822. *An Elementary Treatise on Astronomy. In Two Parts. The First, Containing a Clear and Compendious View of the Theory. The Second, a Number of Practical Problems. To Which are Added, Solar, Lunar and Some other Astronomical Tables.* Philadelphia.

Harris, John. 1719. *Astronomical dialogues between a gentleman and a lady: wherein the doctrine of the sphere, uses of the globes, and the elements of Astronomy.* London: T. Wood.

Herrero López, C. 2007. "Las mujeres en la investigación científica". *Criterios*, 8: 73-96.

Hoskin, Michael (ed.). 1999. *The Cambridge Concise History of Astronomy*. Cambridge: Cambridge University Press.

Hough, Samuel J. 1968. Henry Curson. *Notes and Queries*, 15/2: 64.

Hyland, Ken. 1998. *Hedging in Scientific Research Articles*. John Benjamins.

Inkster, Ian. 1982. “Advocates and audience- Aspects of popular astronomy in England, 1750-1850”. In *Popular Astronomy in England, 1750-1850*. British Astronomical Association, 119-123.

Jardine, Lisa. 1999. *Ingenious Pursuits: Building the Scientific Revolution*. London: Little, Brown and Company.

Jones, Peter Murray. 1999. "Medicine and science". In *The Cambridge History of the Book in Britain 1400-1558*. Vol. III, Hellinga, Lotte and J. B. Trapp (eds), 433-448. Cambridge: Cambridge University Press.

Kaufer, David & Carley, Kathleen M. 1993. *Communication at a Distance: The Influence of Print on Socio-Cultural Organization and Change*. Hillsdale, NJ: Lawrence Erlbaum.

Kytö, M., Rudanko, J. & Smitterberg, E. 2000. “Building a Bridge between the Present and the Past: A Corpus of 19-century English”. *ICAME* 24: 85-97.

Lareo, Inés. 2008. “Make-collocations in nineteenth-century scientific English”. *Studia Neophilologica*, 80: 3-18.

Lareo, Inés. (forthcoming). “Analysing a type of collocations. Make complex predicates in nineteenth-century science and fiction”. *Revista Canaria de Estudios Ingleses.*

Lareo, Inés & Esteve Ramos, María José. 2008. “18th century scientific writing: a study of make complex predicates in the Coruña Corpus”. *ICAME*, 32: 69-96.

Lareo, Inés & Moskowich, Isabel. (forthcoming). “Make plus adjective in eighteenth century science and fiction: Some issues made public”. *English Studies*.

Lareo Martín, I., Crespo García, B. & Moskowich, I. 2008. “The compilation of astronomy texts in late modern English: CETA in the context of the Coruña Corpus”. *International Conference on the Language of Science and Technology,* IULMA, Castelló (Spain) 16-18 October 2008.

Lavoisier, A.L. 1789. *Traité elementaire de chimie.* In *Oeuvres de Lavoisier, 1864.* Paris: Imprimerie Imperiale.

Littlefair, Alison B. 1991. *Reading all Types of Writing: the Importance of Genre and Register for Reading Development.* Philadelphia: Open University Press

Long, Roger. 1742. *Astronomy, in Five Books.* Cambridge, printed for the author.

Luby, Thomas. 1828. An *Introductory Treatise on Physical Astronomy.* London and Dublin: Baldwin and Cradock.

Manzalaoui, M. A. 1977. *Secretum Secretorum: Nine English Versions.* Vol. 1: *Text*. (Early English Text Society O.S. 276.) Oxford: Oxford University Press.

Matthew, H. Colin G. & Harrison, Brian (eds), 2004. *Oxford Dictionary of National*

Biography: From the Earliest Times to the Year 2000). Oxford: Oxford University Press.

McEnery, Tony & Wilson, Andrew. 1996. *Corpus Linguistics*. Edinburgh: Edinburgh University Press.

McIntosh, Carey. 2000. "Elementary rhetorical ideas and eighteenth-century English". In *Language Sciences* 22: 231-249.

Mellan, Charles, Davis, William and Finucane, Brendan (eds). 2002. *Irish Innovators in Science and Technology*. Dublin: Royal Irish Academy.

Moessner, Lilo. 2001. "Genre, text type, style, register: A terminological maze?". *European Journal of English Studies*, 5/2, 131-138.

Montoya Reyes, Ana. 2008. "The presence of cognitive verbs in mathematical texts (1800-1900) of the Coruña Corpus. *ICAME*, 32: 139-155.

Monzó Nebot, E. 2002. *La Profesió del Traductor Juridic i Jurat: Descripció Sociòlogica del Professional i Anàlisi Discursiva del Trasgenere.* Unpublished PhD Dissertation. Castelló: Universitat Jaume I.

Morden, Robert. 1702. *An Introduction to Astronomy, Geography, Navigation, and other Mathematical Sciences Made Easier by the Description and Uses of the Celestial and Terrestrial Globes. In Seven Parts.* London.

Moskowich, Isabel. 2001. "Morfología flexiva del inglés moderno". In *Lingüística Histórica Inglesa,* De la Cruz Cabanillas, Isabel & Martín Arista, Javier (eds.), 624-654. Barcelona: Ariel

Moskowich, Isabel. 2008. "'To lerne *sciences* touching nombres and proporciouns': The proportion of affixation in early scientific writing". *English Studies* 89/1: 39-52.

Moskowich, Isabel. 2008. "Charting the development of English Scientific Writing: The case of adjectives in 18th century Astronomy". *First ISLE International Conference* (8-11 October 2008). Freiburg (Germany)

Moskowich, Isabel. (forthcoming). Morphologically complex nouns in English Scientific Texts after Empiricism.

Moskowich, Isabel & Crespo García, Begoña. 2006. "*Lop-webbe* and *henne cresse*: Morphological aspects of the scientific register in Late Middle English". *Studia Anglica Posnaniensia.* 42: 133-145.

Moskowich, Isabel & Crespo García, Begoña. 2007. "Presenting the Coruña Corpus: A collection of samples for the historical study of English Scientific Writing". In *'Of Varying Language and Opposing Creed': New Insights into Late Modern English. Bern: Peter Lang (Linguistic Insights Series),* Pérez Guerra et al. (eds).

Berlin/New York: Peter Lang.

Moskowich, Isabel & Parapar, Javier. 2007. "Writing Science, Compiling Science: The Coruña Corpus of English Scientific Writing". *31st AEDEAN Conference*, A Coruña 14-17 november 2007.

Myers, G. 1990. *Writing Biology: Texts in the Social Construction of Scientific Knowledge*. Madison: University of Wisconsin Press.

Nicholson, William. 1782. *An Introduction to Natural Philosophy.* London: Printed for J. Johnson.

Nwogu, K.N. 1990. *Discourse Variation in Medical Texts: Scheme, Theme and Cohesion in Professional and Journalistic Accounts*. Nottingham: University of Nottingham.

O'Céirín, Kit & O'Céirín, Cyril. 1996. *Women of Ireland. Biographic Dictionary*. Tír Eolas.

OED CD-ROM version 3.0. 2002. Oxford: Oxford University Press.

Olmsted, Denison. 1841. *Letters on Astronomy, Addressed to a Lady in which the Elements of the Science are Familiarly Explained in Connexion with its Literary History. With Numerous Engravings.* Boston: Marsh, Capen, Lyon and Webb.

Parapar, Javier and Isabel Moskowich. 2007. The Coruña Corpus Tool. *Revista del Procesamiento de Lenguaje Natural* 39: 289-290.

Rissanen, Matti. 1996. "Genres, texts and corpora in the study of medieval English". In *Anglistentag 1995 Greifswald. Proceedings,* Jürgen Klein and Dirk Vanderbeke (eds), 229–242. Tübingen: Max Niemeyer Verlag.

Shapin, Steven. 1984. "Pump and circumstance: Robert Boyle's literary technology". In *Social Studies of Science*, Vol. 14, No.4. Sage Publications, Ltd. (481-520).

Sprat, Thomas. 1667. *The History of the Royal Society of London, for the Improving of Natural Knowledge,* London: Printed by T.R. for J. Martyn, and J. Allestry.

Taavitsainen, Irma. 2000a. "Science". In *The Chaucer Companion,* Brown, Peter (ed), 378-396. Oxford: Blackwell..

Taavitsainen, Irma. 2000b. "Metadiscursive comments and the evolution of early English medical writing 1375-1550". In *Corpora Galore: Analyses and Techniques in Describing English,* Kirk, John M. (ed), 191-207. Amsterdam: Rodopi. .

Taavitsainen, Irma. 2004. "Transferring classical discourse conventions into the vernacular". In *Medical and Scientific Writing in Late Medieval English,*

Taavitsainen, Irma & Pahta, P ivi (eds) 2004. Cambridge: Cambridge University Press.

Taavitsainen, Irma & Pahta, P ivi. 1997. "The corpus of early English medical writing: Linguistic variation and prescriptive collocations in scholastic style". In *To Explain the Present: Studies in Changing English Language in Honour of Matti Rissanen*, Nevalainen, Terttu & Leena Kahlas-Tarkka (eds), 209-225. Mémoires de la Société Néophilologique de Helsinki, 52. Helsinki: Société Néophilologique .

Taavitsainen, Irma & Pahta, P ivi. (eds.) 2004. *Medical and Scientific Writing in Late Medieval English*. Cambridge: Cambridge University Press.

Taavitsainen,I., Pahta, P. & M kinnen. 2005. *Middle English Medical Texts.* Amsterdam/Philadelphia: John Benjamins.

Trauth, G.P. 1990. *German and English Academic Genres: A Contrastive Study of Astronomical Sublanguage and Discourse*. PhD Dissertation. University of California, Berkeley.

UNESCO. 1988. *Proposed International Standard Nomenclature for Fields of Science and Technology*. UNESCO/ROU257 rev. 1. Paris.

Valle, Ellen.1999. *A Collective Intelligence. The Life Sciences in the Royal Society as a Scientific Discourse Community, 1665-1965*. University of Turku: Anglicana Turkuensia.

Van Doren, Charles & McHenry, Robert (eds) 1974. *Webster's American Biographies*. Springfield, MA: C. & G. Merriam Company Publishers.

Voigts, Linda Ehrsam. 1989. "Scientific and medical books". In *Book Production and Publishing in Britain, 1375-1475,* Griffiths, Jeremy and Derek Pearsall (eds), 345-402. Cambridge: Cambridge University Press.

Watts, Isaac. 1726. *The Knowledge of the Heavens and the Earth Made Easy: Or, the First Principles of Astronomy and Geography Explain'd by the Use of Globes and Maps.* London: Printed for J. Clark and R. Hett, at the Bible and Crown in the Poultry near Cheapside; E. Mathews at the Bible in Pater-noster Row, and R. Ford at the Angel in the Poultry near Stocks-Market.

Webster's Biographical Dictionary. 1980. Springfield, MA: G. & C. Merriam Company Publishers.

Yeomans, Donald K. 1977. *The Origin of North American Astronomy—Seventeenth Century*. *ISIS* 68/3: 414-425.