

# There are differences between scientific and non-scientific English *indeed*: a case study



Begoña Crespo (University of A Coruña)

## ABSTRACT

This study considers the behaviour of one specific stance adverb, *indeed*. In a previous analysis of scientific texts, *indeed* was found to be one of the most frequently used adverbs in the expression of emphatic standpoint evincing authorial presence (Moskowich and Crespo 2014). Also noted was its differing use by male and female writers, as well as differences according to genre and the geographical provenance of authors. My aim in the present study is to see whether such behaviour of *indeed* is also found in non-scientific texts, and if so to what extent. The analysis will include both scientific and non-scientific texts from the nineteenth century, a period in which the general fixation of English in its contemporary form had already taken place. The initial hypothesis is that authors of scientific texts tended to express themselves with more caution, even tentativeness, in comparison to authors writing less “impersonal” texts. External factors might also lead to identifiable variations in use in scientific writing, these including the sex of the speaker, plus his or her self-confidence as a writer. Such factors will be used as variables in the analysis. Data for scientific writing will be drawn from the *Corpus of English Texts on Astronomy (CETA)* and the *Corpus of History English Texts (CHET)*; the *Penn Parsed Corpus of Modern British English (PPCMBE)* will be used for non-scientific texts.

## KEYWORDS

corpus linguistics, discourse analysis, Late Modern period, scientific English, stance

## DOI

<https://doi.org/10.14712/18059635.2020.2.5>

## 1 INTRODUCTION

In recent years stance-taking has been discussed widely in the literature, not only as a reflection of narrative voice in literary analyses but also from a linguistic point of view. It has been understood in terms of hedging, authorial presence, and propositional attitude, among others, even in supposedly neutral and objective text types such as scientific writing (Hyland 1998, 2005; Seoane 2016). Stance may be expressed by means of different grammatical realisations, and stance adverbs are just one of these. The current study will explore the behaviour of one stance adverb in particular, *indeed*, in both scientific and non-scientific English writing. In the following section a working hypothesis and the research questions will be given, as well as some socio-historical and theoretical background. Findings from previous works on scientific writing will also be noted. In Section 3 the corpus material and methodology are described, and the study itself, including results, is set out in Section 4. Finally, some conclusions will be offered in Section 5.



## 2 BACKGROUND AND RESEARCH QUESTIONS

The term *stance* entered the English language in the sixteenth century to denote “a standing-place, station, position” (OED). However, the linguistic study of authorial stance in texts is quite a recent innovation (Chafe 1986). In linguistic or pragmatic terms, stance can be interpreted in different ways, although it is always related in some form to interaction. The interpretation taken here involves stance as a component in a writer’s attitude, that is, how the communicative intention of an author has an effect on his or her language, or how a writer uses language, perhaps unconsciously, to produce a specific effect on the reader (Moskowich 2017).

When communicating in the register of science, “authors inevitably adopt a position, i.e. stance, both towards the information presented and the target audience” (Adams and Quintana-Toledo 2013: 13). Indeed, as Hyland (2005: 173) has noted, “over the past decade or so, academic writing has gradually lost its traditional tag as an objective, faceless and impersonal form of discourse and come to be seen as a persuasive endeavour involving interaction between writers and readers”. However, recent studies (Moskowich and Crespo 2014) suggest that stance had already become a productive device in the nineteenth century. Among the mechanisms used to express stance we find adverbs, and evidence indicates that *indeed* was used with some frequency in this sense in nineteenth-century scientific writing.

In terms of their semantics, stance adverbials have been classified as epistemic, attitudinal and style adverbials, depending on the communicative function they perform. The epistemic class includes meanings such as certainty, reliability, imprecision or the marking of perspective; attitudinal adverbials, as the name implies, express attitudes, feelings or value judgements; meanwhile, style adverbials describe how information is treated and presented (Conrad and Biber 1999). Thus, the word *indeed* as a stance adverb can serve as an epistemic stance adverb or an attitudinal one, as we will see in some of the examples below. As is the case with (11) for instance, there is a clear intention on the author’s part to manifest his/her opinion by reinforcing an utterance. The semantic classification of adverbs as well as the pragmatic intention with which it is used is not completely objectifiable but more amenable to subjective interpretation.

Markers of stance have also been classified according to the grammatical realisations they present, in this case that of the adverb. Placement in the clause is a further parameter used to characterise adverbial stance markers (Biber and Finegan 1988; Conrad and Biber 1999) and I will use clause position as an additional criterion to identify the communicative purpose of *indeed* in the analysis here.

The language of the nineteenth century, as the focus of the present study, is certainly close in most respects to Present-day English (PDE) (Millward and Hayes 2012), and thus it is not unreasonable to take our findings here as a reflection of the beginning of the final shift away from the rigid objectivity in scientific discourse promoted by empiricism. Such a shift came after a period in which scientific writing, independent of other more popular communicative formats, had become well established, following the general acceptance of the empiricist scientific method and Boyle’s claims regarding scientific style (Allen, Qin and Lancaster 1994). Conse-



quently, scientific writing had come to be more and more standardised, both in style and intention. As one of the founders of the new science, in terms of both method and expression, Robert Boyle demanded a simple and clear style; others, such as Francis Bacon, claimed that only facts were to be presented. As a reaction against medieval scholasticism, the new empirical science was to be based not on the interpretation of authoritative statements by classical scholars, but on the objective and simple description of experimental works, observable phenomena or other facts, all expressed in clear and concise terms. Such objectivity, which has exerted a dominant influence on scientific writing from empiricism to the present day, has nevertheless shown itself to be flawed from our present-day perspective, with truthfulness and reliability being two of the key issues in modern scientific writing. In the final decade of the twentieth century the focus turned to the personal commitment of authors in writing science, to their presence, their attitudes towards the content of their work, and to the encoding of their implicit opinions (Hyland 2005). It is arguable that authors have never ceased to be perceptible through their writing, but some stylistic devices (presence of *we* instead of *I*, abundance of passive voice, etc.) certainly contributed to the sense of objectivity as it was originally intended.

In the present study, which is historical in nature, I will seek to demonstrate that, although less frequent than in non-scientific texts, authorial presence in scientific texts is not limited to the use of personal pronouns as an overt linguistic mechanism, but that devices such as adverbs can also successfully encode the author's view, seen as early as the nineteenth century, in that authorial intervention in science writing has been recognised as central in the construction of present-day scientific discourse (Hunston 1994; Hyland 1998, 2004; Hyland and Tse 2004, 2005).

The adverb *indeed* has been a part of the English lexicon since the fifteenth century and persists to the present day, although it has been affected by some notable linguistic processes, among them lexicalisation. From the point of view of its function, it may modify a particular word (noun, adjective, adverb) or the whole sentence, as the various examples in this paper will illustrate. In addition, from a pragmatic perspective, *indeed* has been associated with orality. In previous studies (see Busse 2012: 286), *indeed* has been found to play the role of a stance adverbial as a sentence modifier, with sentential scope that is pragmatically meaningful in that it can be both speaker- and hearer-oriented. Such use of *indeed* was found in Shakespeare's *Othello*, for instance. At this point, it should be noted that these examples from non-scientific texts have been taken from extracts of a dialogic nature, in other words, close to orality. The presence of stance adverbs across registers in the previous stages of the language has also been examined by Gray et al. (2011).

My working hypothesis is that, given the nature and meaning of *indeed* as an adverb indicating stance, it should be more abundant in non-scientific writing, where opinions may be expressed more openly, but has nevertheless been used by authors of scientific works as early as the nineteenth century. Hence, I will look first at its distribution across texts, and second, at any constraints on its distribution therein. The data and methodology will be described in the following section.



### 3 CORPUS MATERIAL AND METHOD

Since the main aim of this study is to compare the use of the stance adverb *indeed* in scientific and non-scientific writing from the nineteenth century, we have resorted to a specialised corpus, the *Coruña Corpus of English Scientific Writing (CC)*, and to a non-specialised one, the diachronic part of the *Penn-Helsinki Corpus*. From the former we have chosen samples belonging to the *Corpus of English Texts on Astronomy (CETA)* and the *Corpus of History English Texts (CHET)*, representing the natural and exact sciences and the humanities, respectively. From the latter we have used parts of the *Penn Parsed Corpus of Modern British English (PPCMBE)*, specifically those representing private letters, diaries, theatre, fiction, sermons and the law.

The structure and principles of compilation of these two corpora are different. Whereas all samples in the *CC* contain approximately 10,000 words (Moskowich 2012), those in the *PPCMBE* range from 3,302 to 28,271 words. However, where possible I have tried to avoid any skewing of the results by selecting samples of a similar size for each author. Hence, I have taken samples by 21 authors from *CETA*, 20 from *CHET*, and 25 from *PPCMBE*, yielding a total of some 652,235 words, as set out in Table 1 (for a more detailed account of the data, see the Appendix):

	Corpus	Words	TOTAL
Non-scientific	PPCMBE	247,919	247,919
Scientific	CETA	201,830	404,316
	CHET	202,486	
<b>TOTAL</b>			652,235

TABLE 1. Word-counts for the data

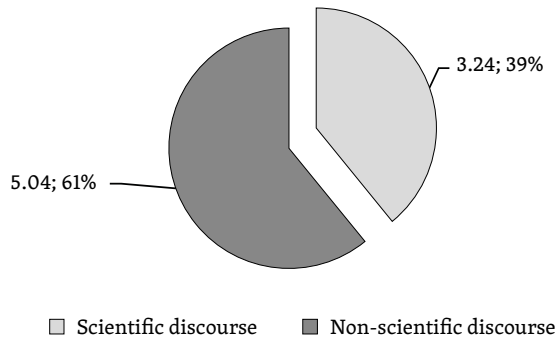
Given that the total number of words for scientific and non-scientific samples remains uneven, raw frequencies have been normalised to 10,000 words.

A further difference between the two corpora led to the use of different tools for the extraction of data: whereas texts in the *CC* are packed in an index and can be visualised in xml format using *Coruña Corpus Tool* (the software accompanying the corpus), texts from *PPCMBE* are in plain text files and have been handled using *AntConc* (version 3.2.2).

For the pragma-linguistic description of *indeed* the syntactic role played by this adverb and its position in the clause have also been considered.

### 4 FINDINGS AND DISCUSSION

From the 652,235 words in our data overall, there were 256 (3.92 normalised figures, hereafter *nf*) instances of *indeed*. However, the distribution is not homogeneous. Scientific texts account for 131 cases at a normalised frequency of 3.24 *nf*, with 68 instances in *CETA* and 63 in *CHET*; in the non-scientific samples, there are 125 cases, at 5.04 *nf* (see Figure 1):



**FIGURE 1.** Distribution of *indeed* in scientific and non-scientific texts

The frequency of use in scientific and non-scientific texts, then, does not differ wildly. We might take this as confirmation of the well-accepted use of *indeed* as a marker of authorial presence in scientific writing, at levels which, whilst lower than in non-scientific texts, are certainly significant, comparatively speaking, in this particular case. As stated by Simpson (2004: 38), academic writing is said to be rich in stance expressions, some of which overlap with other spoken and written registers.

The following examples illustrate some of the uses of *indeed* in both scientific (astronomy and history) and non-scientific texts. As can be seen, they can have different syntactic scope, whether on the sentence level (examples (1) and (3) below) or on the phrase level (example (2)):

- (1) not so wise in my trade; which, *indeed*, is more difficult to manage wisely (Carlyle 1835: 55)
- (2) a transparent fluid when viewed perpendicularly from above so faithfully *indeed* that it is hardly possible not to be impressed with (Herschel 1833: 208)
- (3) for their distrust promising confidence and obedience in future and *indeed* they were now ready to worship him as one inspired (Callcott 1828: 247)

Nevertheless, in all the three examples the pragmatic roles of the adverb consist of focusing on a linguistic segment (either a single word or the whole sentence) and highlighting its meaning, thus offering the reader an authorial viewpoint: “I know for sure that...” so “you can take it for certain”.

It is the sentence initial position that predominates in all hits of *indeed* found in the corpora. It is then mainly functioning as a disjunct, a peripheral adverbial manifesting the writer’s attitude towards a proposition.

The general findings for *indeed* in both blocks of texts, whether scientific or not, have been further examined from the point of view of the sex of the writer, exploring the frequency of use of the adverb in each case, as well as the meanings with which it has been used. Additionally, I have considered the communicative formats in which



authors are more likely to use this stance adverb, and whether this use is related in some way to the oral nature of the sample, the closeness between the author and his/her audience, and consequently to authorial involvement.

The analysis of these two variables will result in more information on the use of *indeed* in both science and non-science writing. As Bazerman (1988: 295–296) observes, “one cannot understand language without looking at the contexts in which it is used to convey meanings” and that “in making statements we bring together many elements — cultural, social, psychological and material — to accomplish our activities and create cognition”.

#### 4.1 SEX VARIABLE

If we turn to possible variables affecting text production here, the analysis is more enlightening. One such variable is the extra-textual issue of the author’s sex. As set out in Figure 2 below, women are in general more likely than men to use *indeed*. As already noted, all authors use *indeed* more often in non-scientific than in scientific texts. Texts written by women represent 15.26% of the total number of words in non-scientific texts (37,850 words), and the adverb *indeed* occurs more frequently here than in texts written by men (6.07 vs. 4.85 uses/10,000 words):

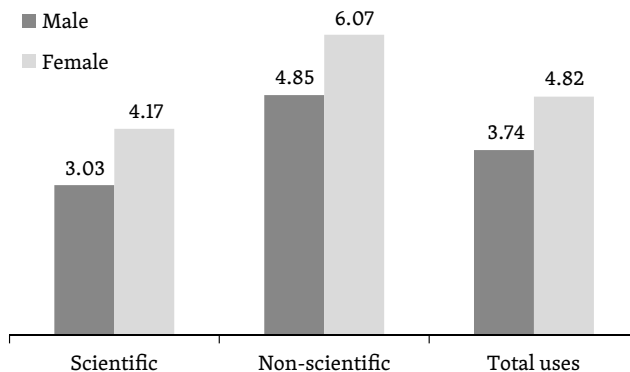


FIGURE 2. Male and female uses of *indeed*

Illustrations of these uses in non-scientific texts by women can be found in (4) and (5) below:

- (4) Pray tell Lizzy that if I had imagined her Teeth to be really out, I should have said before what I say now, that it was a very unlucky fall *indeed*, that I am afraid it must have given her a great deal of pain, & that I dare say her Mouth looks very comical. (Austen 180X, 183.522)
- (5) From the first avowal of Colonel Keith’s acquaintance with the Williamses, she had concluded him to be the nameless lover, and had been disappointed that



Alison, so far from completing the confidence, had become more reserved than ever, leaving her to wonder whether he were *indeed* the same, or whether his constancy had survived the change of circumstances. (Yonge 1865, 177.472)

In example 4, Austen resorts to *indeed* to indicate an intention to give a personal opinion on Lizzy's tooth. The adverb here is used in a highly interactive written format, one of the most effective ways of establishing verbal interaction between addresser and addressee. The intimate and personal tone of this communicative format is reinforced by the presence of first person singular pronouns and other expressions typical of the oral register ("am afraid", "dare say").

Example (5) shows a less oral-like use of *indeed*, but one which nevertheless has a pragmatic function, involving the confirmation of a previous statement through reinforcement and emphasis (*he = the same*) in a narrative work of fiction.

In scientific texts, female production represents 17.78% (71,908 words) of the total word count. Forms of *indeed* in texts written by women once more outnumber those used by men (4.17 vs. 3.03 nf), as is already shown in Figure 2 above.

From all the samples of scientific texts written by women, some 23 instances are found. The distribution is unequal, with just one case in an astronomy text and the remaining 22 in history texts. Examples (6) and (7) illustrate such uses:

- (6) The inferred absence of an atmosphere is *indeed* scarcely reconcilable with some of the transit-phenomena just adverted to; but heights and hollows in abundance seem to exist. (Clerke 1893: 303)
- (7) And now, those who had murmured against the admiral, who had talked of putting him in chains, and forcing him to return to Spain, threw themselves at his feet, and entreated forgiveness for their distrust, promising confidence and obedience in future; and *indeed* they were now ready to worship him as one inspired by Heaven. (Callcott 1828: 247)

*Indeed* is placed in different positions within the sentence, to enhance the subject complement of a verbal group, as in (6), or as a sentential adverb modifying the whole clause (7).

The narrative character of the story and its similarity to fiction makes it possible for *indeed* to be used as a reinforcement of authorial presence and interaction with the reader. Thus, the truthfulness and reliability of what Clerke is stating in (6) remains beyond doubt, and *indeed* can be considered an epistemic stance adverb here. The interaction between author and reader consolidates the position of authority that the author occupies within the epistemic community. And this is precisely what is most needed by women who dare to take on a masculine role, that of scientist, in nineteenth-century society. The reliability conveyed by Callcott's *indeed* can be interpreted as an interaction with the readership in an even more intimate or personal tone, and might be classified as an attitudinal one.

An analysis of the variable communicative format (Moskowich and Crespo 2016), most generally known as genre, will be presented in what follows.



#### 4.2 COMMUNICATIVE FORMAT

Another variable, that of the communicative format, is in some way inherent to a text, in that the function or intention of a piece of writing itself influences language choice, that is, the form and function of the language used. Biber et al. (1999) have claimed that oral registers exhibit the highest number of stance adverbs. This may be due to a direct interaction between addresser and addressee, in that they often occupy the same time and/or space. The oral orientation of a communicative act determines the choice of format in the written medium. The use of *indeed* as a stance adverb is an overt manifestation of the oral nature of a written text in which interpersonal contact can be clearly appreciated.

In order to represent non-scientific texts, the more oral-like texts in the data, I have grouped all the categories from the PPCBME (see Appendix) into the following communicative formats: letters, sermons, drama and law. The data in each case is set out in Table 2 below:

	Words	Tokens	Nf
Letters	77,744	51	6.55
Sermons	27,414	22	8.02
Drama	76,813	50	6.50
Law	65,948	2	0.30
	247,919	125	5.04

TABLE 2. Data for non-scientific formats

In Figure 3, the normalised figures for uses of *indeed* in the various formats of non-scientific writing are shown:

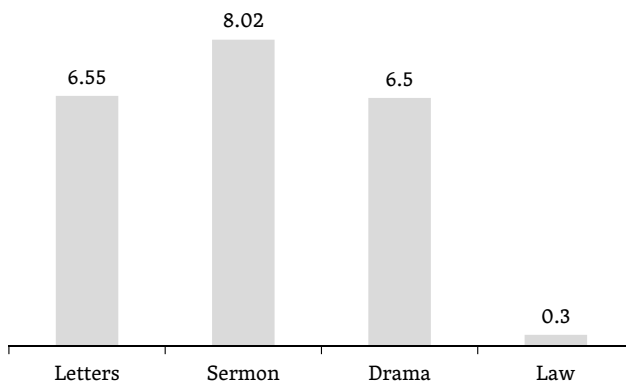


FIGURE 3. *Indeed* in non-scientific writing

As Figure 3 illustrates, sermons and legal texts occupy the two extremes in terms of the frequency of occurrence of *indeed* here, with the highest levels (8.02 nf) found



in sermons. Several reasons may account for this. Sermons pertain to the realm of speech as opposed to writing. They involve a direct address to the audience with the implicit pragmatic force of persuasion, and the use of *indeed* reinforces other pragmatic resources to this end. Such is the case in examples (8) and (9) below:

- (8) ... and rather than mock God by praying to Him in word, when *indeed* their hearts were fixed on other things, they made Him their only consolation, by voluntarily foregoing every other. (Froude 1830, 2,11.122)
- (9) It is *indeed* undoubtedly true that many persons are so constituted that such habits come much easier to them than to others. Many indeed there are, who, without much pains or restraint, may go on very comfortably and respectably, by means of those habits which they have already acquired, and in consequence of the discipline which they have already undergone. (Froude 1830, 2,26.258)

*Indeed*, as a sentence adverbial with an epistemic meaning, is supporting the speaker in his intention to emotionally engage the audience. The speaker is inviting the listener to feel in a certain way about what is being said. It is a way of seeking to shape the addressee's emotional and intellectual responses to an issue (Jensen 2015).

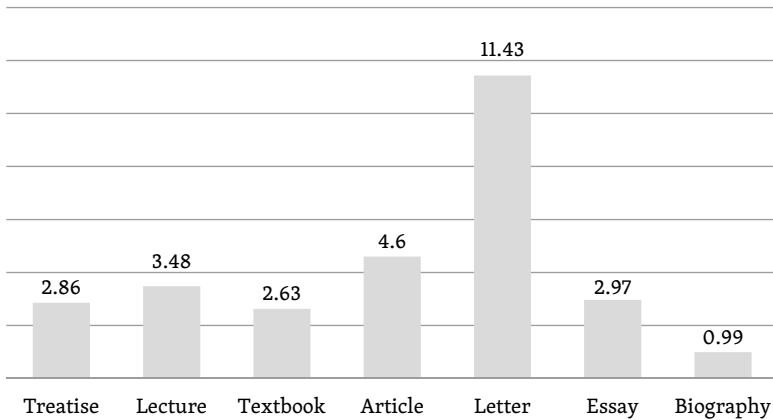
On the other hand, legal texts, although they might reproduce oral interactions such as those taken directly from trials, contain the fewest instances of *indeed*. There are in fact just two uses of this adverb in the current data, both found within dialogues, and both given here, in examples (10) and (11):

- (10) *indeed* I was satisfied, from their declarations (Watson 1817, 1,150.1796)
- (11) "Did you give your note to Mr. Hone's publication?"  
"No, *indeed*, I did not." (Watson 1817, 1,163.2112)

Again, it is the feature of orality that is playing a part here in the legal format. However, when comparing it with sermons, a different communicative intention on the speaker's part should also be noted. In this case, the speaker puts the emphasis on personal reaffirmation as a means of achieving the persuasive effect on the listener that he seeks. In this sense, the adverb could be considered as pertaining to the attitudinal stance adverbs group.

The other two formats analysed, letters and drama, display similar frequencies of use (6.55 and 6.5 nf), and are situated between the higher use in sermons and the markedly low use in legal texts.

As for scientific discourse, Figure 4 shows the occurrences of *indeed* in the seven formats represented. Letters show a heavy reliance on these forms (11.43 nf), followed at some distance by another very dynamic form, articles (4.6 nf). Looking at the realm of scientific writing, we note that letters are perhaps the least formal of formats used. In the scientific domain these are not letters in the normal sense, but rather are often compositions of a highly technical nature in epistolary form, including forms of ad-



**FIGURE 4.** Use of *indeed* in scientific writing per format

dress such as “Dear XX” and ending with some kind of leave-taking formula. As such they were an established means of conveying and exchanging knowledge within the scientific community. This goes a long way towards explaining the high number of occurrences of *indeed* in the data here.

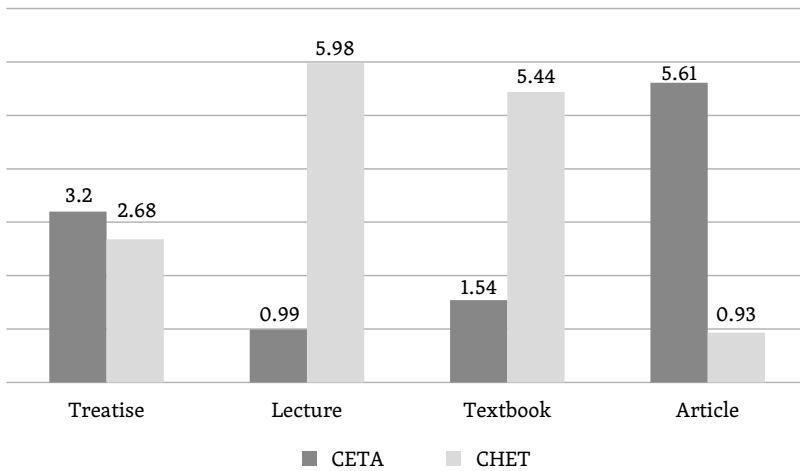
A more detailed analysis of the scientific texts shows that *indeed* is not uniform across the data. 68 instances (3.35 nf) were found in nineteenth-century samples of astronomy texts (*CETA*), and some 63 cases (3.11 nf) in history texts (*CHET*), which in principle does not imply a great difference between the two disciplines. In fact, the real difference between the two can be seen not in the overall use of *indeed* here, then, but in how the adverb is distributed across the text formats in each field of knowledge.

From the seven formats represented in scientific writing, four are found in both disciplines (astronomy and history). As shown in Figure 5 below, a similar number of instances were found in treatise (3.2 in *CETA* vs. 2.68 in *CHET*) whereas dramatic differences were observed in lecture (0.99 in *CETA* vs. 5.98 in *CHET*), textbook (1.54 in *CETA* vs. 5.44 in *CHET*) and article (5.61 in *CETA* vs. 0.93 in *CHET*).

Although it may be surprising that the format article contains the highest frequency of the form *indeed*, we can perceive in the following extract from an article by Young, published in 1880, how the language has been formulated with the intention of persuading the audience of the virtues of solar astronomy as a science:

NO branch of astronomical inquiry is of more general interest than that which concerns itself with the sun. The investigations as to the distance and mass of this body, its constitution, its radiation and temperature, its atmosphere and appendages of chromosphere and corona, these and all allied researches have attracted the attention and interest not only of professional men of science, but of the whole intelligent public. (Young 1880: 89)

An example of this use of *indeed* in the same article, as a sentential marker with persuasive force, can be found in (12):



**FIGURE 5.** Hits of *indeed* in CETA and CHET

- (12) If *indeed*, as is now almost universally admitted, the sun be almost entirely a mass of permanent gases and uncondensed vapors, the heat of the central portions of the globe must be unimaginable. Strictly, then, we cannot speak of the sun as having a temperature. (Young 1880: 93)

In the history texts from CHET, instances of *indeed* predominate in the formats lecture and textbook, as illustrated in examples (13) and (14):

- (13) It will be seen hereafter that the Mur Ollamhan is not mentioned among the vestiges described by the ancient topographers; and *indeed* there is every reason to believe, that it had no existence save in the etymological inference of O’Flaherty... (Petrie 1839: 31)
- (14) After the death of Alfred, however, England, and *indeed* the whole of Britain, became a prey to the Scandinavian freebooters from Denmark and Norway. (Masson 1855: 82)

Since one of the functions of these formats is that of teaching and persuading the readership, it is not surprising that the two formats exemplified above are the ones containing the highest occurrences of the adverb in history texts.

For the high use of *indeed* in article (CETA) and lecture (CHET) there may be an additional external reason which helps to explain its notable presence, especially when compared with the frequency of these forms in non-scientific texts. In the nineteenth century, when science was in the process of becoming a professional activity, “editors and publishers sought to find the most appropriate and appealing forms in which to package science for both general and professional audiences” (Shuttleworth and Chandley 2016: 298). In addition, as these same authors have explained, “...there



was no clear division through most of the nineteenth century between professional and popular scientific journals, or between amateur and professional communities. Journals provided a space both for interaction and for self-definition, as established norms of scientific communication were gradually set in place” (Shuttleworth and Chandley 2016: 299).

Both articles and lectures were common formats in this realm of pre-professional science, when the dissemination of scientific knowledge was not limited to the epistemic community but also open to an increasingly literate public.

## 5 CONCLUSIONS

In this paper I have considered the behaviour and distribution of the adverb *indeed* in scientific and non-scientific writing from the nineteenth century. *Indeed* is a stance adverb, and manifests the author’s opinion or point of view on a specific issue. The initial research questions here have been answered quite clearly. First, how is *indeed* distributed across texts? The analysis has revealed that *indeed* was more frequently used in non-scientific texts, thus confirming my initial hypothesis, and underlining the fact that this adverb is of particular use in more personal and intimate styles of writing. Second, looking at the constraints on such distribution, I have noted that both extra- and intra-linguistic factors play a role. As for the former, findings have shown that the sex of the author is determinant, with higher frequencies of use in women’s writing in both scientific and non-scientific writing; women might thus be said to be more personal in their communicative acts than men, using a more involved style. As for intra-linguistic factors, the communicative format in which *indeed* is used most often is that of letter. I argued that this may be due to the relative proximity of this format to orality, and to the nature of epistolary communication itself, which evidently conveys the human relationships on which it depends. In fact, sermon, the predominant non-scientific format in using *indeed*, plays a similar function, even with a more prominent persuasive tone. Finally, the use of stance adverbs such as *indeed* reveals an author’s value system and his or her relationship with the addressee, these obviously within the social constraints of the period and the specific need of each individual author to reinforce either factuality (epistemic stance) in a search for reliability, or attitude (attitudinal stance) as a means of expressing opinions and value judgements.

### Acknowledgements

The research reported here has been funded by the Spanish Ministerio de Economía y Competitividad (MINECO), grant number FFI-2016-75599-P. This grant is hereby gratefully acknowledged.

## REFERENCES

- Adams, H. and E. Quintana-Toledo (2013) Adverbial Stance Marking in the Introduction and Conclusion Sections of Legal Research Articles. *Revista de Lingüística y Lenguas Aplicadas* 8, 13–22.
- Allen, B., J. Qin 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.
- Bazerman, C. (1988) *Shaping Written Knowledge. The genre and activity of the experimental article in science*. Wisconsin: University of Wisconsin Press.
- Biber, D. and E. Finegan (1988) Adverbial Stance Types in English. *Discourse Processes* 1, 1–34.
- Biber, D., S. Johansson, G. Leech, S. Conrad and E. Finegan (1999) *Longman Grammar of Spoken and Written English*. Essex: Longman.
- Busse, B. (2012) Historical Text Analysis: Underlying Parameters and Methodological Procedures. In: Ender, A., A. Leemann and B. Wälchli (eds) *Methods in Contemporary Linguistics*, 285–308. Berlin: Mouton de Gruyter.
- Conrad, S. and D. Biber (1999) Adverbial Marking of Stance in Speech and Writing. In: Hunston, S. and G. Thompson (eds) *Evaluation in text. Authorial stance and the construction of discourse*, 56–73. New York: Oxford University Press.
- Chafe, W. (1986) Evidentiality in English Conversation and Academic Writing. In: Chafe, W. and J. Nichols (eds) *The Linguistic Coding of Epistemology*, 261–272. Norwood, NJ: Ablex.
- Gray, B., D. Biber and T. Hiltunen (2011) The Expression of Stance in Early (1665–1712) Publications of the *Philosophical Transactions* and Other Contemporary Medical Prose: Innovations in a Pioneering Discourse. In: Taavitsainen, I. and P. Pahta (eds) *Medical Writing in Early Modern English*, 221–247. Cambridge: Cambridge University Press.
- Hunston, S. (1994) Evaluation and Organization in a Sample of Written Academic Discourse. In: Coulthard, M. (ed) *Advances in Written Text Analysis*, 191–218. London: Routledge.
- Hyland, K. (1998) *Hedging in Scientific Research Articles*. Amsterdam: John Benjamins.
- Hyland, K. (2004) Disciplinary Interactions: Metadiscourse in L2 Postgraduate Writing. *Journal of Second Language Writing* 13/2, 133–151.
- Hyland, K. (2005) Stance and Engagement: A Model of Interaction in Academic Discourse. *Discourse Studies* 7/2, 173–192.
- Hyland, K. and P. Tse (2004) Metadiscourse in Academic Writing: A reappraisal. *Applied Linguistics* 25/2, 156–177.
- Hyland, K. and P. Tse (2005) Hooking the reader: a corpus study of evaluative *that* in abstracts. *English for Specific Purposes* 24/2, 123–139.
- Jensen, T. (2015) Investigating the Passions of Persuasion. Newsletter Autzen House. Oregon State University: The Center for the Humanities. Available at: <https://humanities.oregonstate.edu/investigating-passions-persuasion> [last accessed 22 November 2016].
- Millward, C. M. and M. Hayes (2012) *A Biography of the English Language*. Wadsworth: Cengage Learning.
- Moskowich, I. (2012) CETA as a Tool for the Study of Modern Astronomy in English. In: Moskowich, I. and B. Crespo (eds) *Astronomy 'playne and simple': The Writing of Science between 1700 and 1900*, 35–57. Amsterdam: John Benjamins.
- Moskowich, I. (2017) Pronouns as Stance Markers in the Coruña Corpus: An Analysis of the CETA, CEPHiT and CHET. In: Alonso-Almeida, F. (ed.) *Stancetaking in Late Modern English scientific writing: Evidence from the Coruña Corpus*, 73–92. Valencia: Universitat Politècnica de Valencia.
- Moskowich, I. and B. Crespo (2014) Stance is Present in Scientific Writing, *Indeed: Evidence from the Coruña Corpus of English Scientific Writing*. *Token: A Journal of English Linguistics* 3, 91–114.
- Moskowich, I. and B. Crespo (2016) Classifying Communicative Formats in CHET, CECHeT



and Others. *EPiC Series in Language and Linguistics* 1, 308–320.

*Oxford English Dictionary Online* (2009) Oxford: Oxford UP. Available at: <https://www.oed.com> [last accessed 28 November 2016].

Seoane, E. (2016) Authorial Presence in Late Modern English Philosophical Writing: Evidence from *CEPhiT*. In: Moskowich, I., G. Camiña-Rioboo, I. Lareo and B. Crespo (eds) *The Conditioned and the Unconditioned. A Corpus of English Philosophy Texts*, 123–143. Amsterdam: John Benjamins.

Shuttleworth, S. and B. Charnley (2016) From Science Periodicals in the Nineteenth and Twenty-First Centuries. *Notes and Records* 70, 297–304.

Simpson, R. (2004) Stylistic Features of Academic Speech: The Role of Formulaic Expressions. In: Connor, U. and T. A. Upton (eds) *Discourse in the Professions. Perspectives from Corpus Linguistics*, 37–64. Amsterdam: John Benjamins.

### Begoña Crespo

English Studies

Dpto. de Letras — Universidade da Coruña

Campus da Zapateira s/n

15071 A Coruña — SPAIN

ORCID ID: 0000-0001-5125-9416

bcrespo@udc.es

## APPENDIX

Year	Author	Title of Work	Words per sample
1804	Small, Robert	<i>An account of the astronomical discoveries of Kepler: including an historical review of the systems which had successively prevailed before his time</i>	10,435
1809	Ewing, John	<i>A plain, elementary and practical system of natural experimental philosophy: including astronomy and chronology</i>	9,985
1811	Brewster, Sir David	<i>Ferguson's astronomy explained upon Sir Isaac Newton's Principles: with notes and supplementary chapters</i>	9,824
1818	Phillips, William	<i>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</i>	10,130
1822	Gummere, John	<i>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</i>	10,507

Year	Author	Title of Work	Words per sample
1828	Luby, Thomas	<i>An Introductory Treatise on Physical Astronomy</i>	10,704
1833	Herschel, John F. W.	<i>The Cabinet Encyclopedia. Conducted by the Rev. Dionysius Lardner... Assisted by eminent literary and scientific men. Natural Philosophy. Astronomy. A treatise on Astronomy</i>	10,224
1838	Garland, Landon C.	<i>Address on the Utility of Astronomy</i>	9,608
1841	Olmsted, Denison	<i>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</i>	8,742
1845	Bradford, Duncan	<i>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</i>	10,268
1855	Bartlett, W. H. C., (William Holms Chambers)	<i>Elements of natural philosophy (Spherical Astronomy)</i>	10,858
1858	Whewell, William	<i>The plurality of worlds</i>	10,079
1860	Mitchel, Ormsby McKnight	<i>Popular astronomy. A concise elementary treatise on the sun, planets, satellites and comets</i>	10,183
1868	Loomis, Elias	<i>A treatise on Astronomy</i>	10,323
1871	Chauvenet, William	<i>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</i>	9,895
1874	Steele, Joel Dorman	<i>Fourteen weeks in descriptive astronomy</i>	9,979
1880	Young, LL. D.	<i>Recent Progress in Solar Astronomy</i>	6,454
1880	Darwin, George Howard	<i>On the Secular Changes in the Elements of the Orbit of a Satellite, revolving about a Tidally Distorted Planet</i>	5,181
1889	Croll, James	<i>Stellar Evolution and its relation to Geological Time</i>	9,390
1893	Clerke, Agnes Mary	<i>A popular history of astronomy during the nineteenth century</i>	10,530
1895	Lowell, Percival	<i>Mars: Canals. The Atlantic Monthly: Mars III. Canals</i>	8,531

TABLE 3. Samples from CETA





Year	Author	Title of Work	Words per sample
1802	Adolphus, John	<i>The history of England from the accession of King George the Third, to the conclusion of peace in the year one thousand seven hundred and eighty-three. Vol. III</i>	10,158
1805	Warren, Mercy Otis	<i>History of the rise, progress and termination of the American revolution. Interspersed with Biographical, Political and Moral Observations. In three volumes. Vol. I</i>	10,214
1810	Bigland, John	<i>The history of Spain, from the earliest period to the year 1809. In two volumes. Vol. I</i>	10,065
1814	Britton, John	<i>The history and antiquities of the cathedral church of Salisbury; illustrated with a series of engravings, of views, elevations, plans, and details of that edifice: also etchings of the ancient monuments and sculpture: including biographical anecdotes of the bishops, and other eminent persons connected with the church</i>	10,017
1820	Hardiman, James	<i>The history of the town and county of the town of Galway, from the earliest period to the present time. Embellished with several engravings. To which is added a copious appendix, containing the principal charters and other original documents</i>	10,255
1828	Callcott, Maria	<i>A Short history of Spain. In two volumes. Vol. II</i>	10,332
1833	Aikin, Lucy	<i>Memoirs of the Court of King Charles the First. In two volumes. Vol. I</i>	10,022
1839	Petrie, George	<i>On the History and Antiquities of Tara Hill</i>	10,117
1840	Smyth, William	<i>Lectures on Modern History, from the Irruption of the Northern Nations to the Close of the American Revolution. In two volumes. Vol. II</i>	9,933
1844	D'Alton, John	<i>The history of Drogheda, with its environs; and an introductory memoir of the Dublin and Drogheda railway. In two volumes. Vol. I</i>	10,008
1855	Masson, David	<i>Medieval history</i>	10,166
1857	Sewell, Elizabeth Missing	<i>A first history of Greece</i>	10,037
1860	Freer, Martha Walker	<i>History of the reign of Henry IV. King of France and Navarre. In two volumes.</i>	10,043
1862	Bennett, George	<i>The History of Bandon</i>	10,005
1872	Gray, John Hamilton	<i>Confederation; or, The Political and Parliamentary History of Canada, from the Conference at Quebec, in October, 1864, to the Admission of British Columbia, in July, 1871. In two volumes. First volume.</i>	10,045
1875	Killen, William Dool	<i>The ecclesiastical history of Ireland. From the earliest period to the present times. Vol. II</i>	10,083





Year	Author	Title of Work	Words per sample
1884	Breese, Sidney	<i>The Early History of Illinois, from its Discovery by the French, in 1673, until its Cession to Great Britain in 1763. Including the Narrative of Marquette's Discovery of the Mississippi</i>	10,057
1887	Kingsford, William	<i>The history of Canada. Vol. I. [1608-1682]</i>	10,041
1893	Cooke, Alice M.	<i>The Settlement of the Cistercians in England. The English Historical Review, Vol. 8, No. 32.</i>	10,730
1895	Burrows, Montagu	<i>The History of the Foreign Policy of Great Britain</i>	10,158

**TABLE 4.** Samples from CHET

Wordcount summary by individual text in PPCMBE			
Text	Date	Genre	Wordcount
austen-180x	1805-1808	LETTERS_PRIV	9,650
benson-190x	1905-1906	DIARY	9,986
brougham-1861	1861	DRAMA_COMEDY	10,049
carlyle-1835	1835	LETTERS_PRIV	9,343
collier-1835	1835	DRAMA_COMEDY	9,459
colman-1805	1805	DRAMA_COMEDY	10,161
dickens-1837	1837	FICTION	9,437
froude-1830	1830	SERMON	9,254
godwin-1805	1805	FICTION	9,343
haydon-1808	1808	DIARY	10,015
meredith-1895	1895	FICTION	9,322
nightingale-188x	1888-1889	LETTERS_PRIV	3,302
nightingale-189x	1890	LETTERS_PRIV	6,201
pusey-186x	1865-1866	SERMON	9,022
ruskin-1835	1835	DIARY	9,882
statutes-1805	1805	LAW	9,440
statutes-1835	1835	LAW	9,370
statutes-1865	1865	LAW	9,456
statutes-1895	1895	LAW	9,411
talbot-1901	1901	SERMON	9,138
thring-187x	1870-1872	DIARY	9,997
victoria-186x	1863-1865	LETTERS_PRIV	9,368
watson-1817	1817	PROCEEDINGS_TRIAL	28,271
wilde-1895	1895	DRAMA_COMEDY	9,713
yonge-1865	1865	FICTION	9,329

**TABLE 5.** Samples from PPCMBE