

# A cross-disciplinary study of verb boosters in research articles from Engineering, Medicine and Linguistics: Frequency and co-text variations

*Un estudio interdisciplinar de intensificadores verbales en artículos de investigación de ingeniería, medicina y lingüística: Frecuencias y variaciones en el co-texto*

Hanna Skorczynska

UNIVERSITAT POLITÈCNICA DE VALÈNCIA  
ESPAÑA  
hskorczy@upv.es

María Luisa Carrió-Pastor

UNIVERSITAT POLITÈCNICA DE VALÈNCIA  
ESPAÑA  
lcarrio@upv.es

Received: 02-X-2019 / Accepted: 30-XI-2020

DOI: 10.4067/S0718-09342021000200575

## Abstract

The present study looks into the variations in the frequencies and pragmatic functions of the metadiscourse markers known as boosters, and in particular, with regard to their verb forms. Three corpora have been compiled to this end, covering the fields of Engineering, Medicine and Linguistics. The corpora were manually annotated for metadiscourse markers, boosters included, by a group of annotators. A predetermined list was used for annotation, but throughout the annotation process, the list was modified to better reflect the use of metadiscourse in the corpora. The raw count of the occurrences of verb boosters shows clear differences between the corpora, which in turn confirms previous studies of this type. However, the range of verbs identified was very similar, pointing to a large overlap among the three. The three top frequency verb boosters also showed a clear overlap for Engineering and Medicine, but revealed considerable differences with Linguistics. This study has been conducted within a research project financed by the Spanish Ministry of Economy, Industry and Competitiveness (FFI2016-77941-P).

**Key Words:** Booster, metadiscourse, corpus, Medicine, Linguistics, Engineering.

## Resumen

El presente estudio analiza las variaciones de las frecuencias y las funciones pragmáticas de los marcadores metadiscursivos conocidos como intensificadores, y en particular de sus formas verbales. Se recopilieron tres corpus para este fin, que cubrían el campo de ingeniería, medicina y lingüística. Los corpus fueron anotados manualmente por un grupo de anotadores para identificar todos los marcadores discursivos, incluyendo los intensificadores. Un listado predeterminado de marcadores fue utilizado para la anotación. Dicho listado fue posteriormente completado durante el proceso de la anotación. La comparación de las frecuencias de los intensificadores verbales muestra claras diferencias entre los corpus, lo cual confirma estudios previos de este tipo. Sin embargo, el rango de verbos identificados fue muy similar indicando considerables coincidencias entre ellos. Asimismo, los tres verbos más frecuentes eran iguales en los corpus de ingeniería y medicina, pero diferentes en lingüística. Este estudio fue llevado a cabo dentro del proyecto de investigación financiado por el Ministerio de Economía, Industria y Competitividad del Gobierno de España (FFI2016-77941-P).

**Palabras Clave:** Intensificador, metadiscursos, corpus, medicina, lingüística, ingeniería.

## INTRODUCTION

The way writers express themselves in a genre tends to be similar but if we focus on a specific device or strategy some differences in the way writers communicate can be observed. This paper focuses on academic English as we notice that, although this is a genre that researchers have paid quite a lot of attention to, academic writers find themselves involved in a continuous negotiation with language, a continuous back and forth, in order to transmit meaning and ideas to their peers. We therefore consider that there is a need to study the different patterns used in specific contexts of academic English. Speakers of a language may be able to communicate an idea, but how this idea is communicated might differ depending on factors such as culture, social status, academic background, family education, political beliefs, genre, self-esteem, and so on.

One aspect that should be taken into account is the specific field of knowledge of academic discourse, as this may entail specific uses of strategies that convince readers about the researchers' ideas or the experiments they are conducting. As Hyland and Jiang (2018:19) explain, "Academic knowledge is the outcome of a process of getting people to believe things". In this paper, we focus on the use of interpersonal devices that engage readers in the rhetorical discourse narrated by researchers to interact with them.

Metadiscourse has been one standpoint used by numerous researchers to identify language interactions in academic English (Hyland, 1997, 1998a, 1998b, 2005; Hyland & Tse, 2004; Ädel, 2006; Gillaerts & Van de Velde, 2010; Carrió-Pastor, 2016, 2019b; Hyland & Jiang, 2018), but also in other genres and languages (Hu & Cao, 2011; Mur-Dueñas, 2011; Moya & Carrió-Pastor, 2018a, 2018b, 2018c; Carrió-Pastor, 2019a; Qin

& Uccelli, 2019). Quite recently, one of these researchers, who has dedicated most of his academic life to the analysis of metadiscourse, described it in the following terms:

“Metadiscourse is the commentary on a text made by its producer in the course of speaking or writing and it is a widely used term in current discourse analysis and language teaching” (Hyland, 2017:16).

He explores the strengths and shortcomings of metadiscourse, offering a general overview of the concept and pointing out that it is a powerful analytical tool; specifically, he states that “[...] metadiscourse has inspired a considerable amount of scholarship and continues to contribute enormously” (Hyland, 2017: 27). Metadiscourse is generally divided into two categories, i.e. textual (the devices that provide cohesion and coherence to a text and guide readers and listeners along the discourse) and interactional (the devices used to interact with listeners and readers, engaging them and showing readers the importance of the message). In this paper, we focus on one sub-category of the interactional metadiscourse devices, i.e. boosters, with the aim of studying the variation that can be caused by the nature of the specific field of knowledge. Our analysis is based on a quantitative analysis of academic corpora from three specific fields of knowledge, i.e. Engineering, Medicine and Linguistics.

Our starting point is that research articles should be seen as persuasive discourse even when they are commonly considered a factual, impersonal and objective type of discourse (Becher & Trowler, 2001; Rundbald, 2007). If we analyse academic writing, most writers aim at convincing readers about the importance of their findings and try to ‘sell’ their experiments, ideas or theories. In a sense, authors feel the need to persuade readers and editors of the veracity of their claims (Hunston & Thompson, 2000; Swales, 2004; Hyland, 2005). This may be considered a common rhetorical strategy that lends credibility to arguments by emphasising one’s own certainty about a proposition and part of a strong interpersonal view of metadiscourse comprising the ways speakers can organise a discourse and adopt a stance towards what is being discussed.

Specifically, boosting is a communicative strategy for expressing commitment to statements formulated by writers in order to persuade readers (and editors) of their veracity and validity (Holmes, 1984; Hyland, 1998a; Peacock, 2006). Several studies have approached cross-disciplinary variations in the use of boosters (Hyland, 1997, 1998a, 2005; Hyland & Tse, 2004; Peacock, 2006) and reported differences related to their frequencies and forms. However, to our knowledge, to date no study has undertaken a comparative study of specific booster forms, in this case, verbs, or has attempted to provide both quantitative and qualitative insights into cross-disciplinary variations.

Thus, we have decided to focus on the analysis of verbs that act as boosters to identify the patterns followed in different disciplines. In this study, we show that verb boosters were the most frequent in Engineering and Linguistics, and the second most frequent type in Medicine, as shown in detail in the Results section. The identification of booster verb-noun patterns will be extremely helpful in designing academic English teaching materials.

The objectives of this paper are, on the one hand, to study interdisciplinary variation in the use of verb boosters in research articles from the fields of Engineering, Medicine and Linguistics and, on the other hand, to identify patterns that increase force, power, pressure or effectiveness in academic writing. The research questions we contemplate here are:

1. Is boosting context dependent? What are the main causes of different uses of boosting verbs in specific fields of knowledge?
2. What are the most common collocations of noun + boosting verbs? Is it possible to identify their patterns of use in correlation with their pragmatic functions?

This paper is structured as follows. First, the Introduction offers a general view of the use of metadiscourse and the framework of this study, and the objectives and research questions are set. Next, we focus on the different studies on boosters, the specific focus of this analysis. The corpora and method of study are then described and in section four, the results are shown and discussed. Finally, the conclusions are drawn.

## **1. Boosting: Pragmatic functions and cross-disciplinary variation**

Scholars interested in boosting have provided a number of definitions. Traditionally boosting has been defined as involving the expression of varying degrees of commitment or seriousness of intention (Holmes, 1984) and as “communicative strategies for increasing the force of statements” (Hyland, 1998a: 350), to which Peacock (2006: 65) added that it also serves to “emphasise certainty, strong commitment, conviction and accepted truth”. For Hyland (1998a: 353), boosters counterbalance and, in a way, contradict the use of hedges, which are so effective in the “conciliatory and defensive tactics” of the writers of research articles. Boosters, according to the same author, allow writers to express their conviction and at the same time, they limit the negotiating space available to the reader. For the purpose of this study, we will follow Mur-Dueñas (2011: 3070), who defines boosters as:

“Features which highlight the writers’ certainty and conviction about a proposition and which may be the result of certain pragmatic conventions in academic writing”.

We are interested in identifying these specific academic pragmatic conventions that make writers use some verb boosters more frequently than others. We believe there are differences in the way academic writers conceptualise and transmit certainty and conviction to readers.

Regarding the pragmatic functions of boosters, they have been referred to in previous studies in terms of expressing evidential or implicit truth, accepted truth, and solidarity (Peacock, 2006). When writers use boosters to express evidential or implicit truth, they do so to assert the truth resulting from the research in question, and not in order to make judgement claims. Examples of such boosting would be the use of the following verbs: 'show', 'demonstrate' and 'find' (Skelton, 1997). In the same line, Hunston (1993) argues that 'show', 'demonstrate' and 'establish', used as boosters, convey certainty arising from the convincing nature of the data reported and not from mere persuasion. Similarly, Swales (1990: 151) refers to these three verbs as "powerful rhetorical tools" used to signal that the claims made are to be taken as substantiated. Both Swales (1990) and Salager-Meyer (1994) argue that the verbs mentioned express commitment to a proposition. Peacock (2006) points out that the above verbs can clearly be considered as boosters because they emphasise a writer's certainty and commitment. However, he notes that this type of booster fulfils two functions: they express evidential truth but also play a rhetorical role.

According to Peacock (2006), expressing accepted truth is similar to conveying evidential or implicit truth. Modals (e.g. 'will' or 'must') or certain phrases (e.g. 'of course', 'clearly' or 'obviously') are often used to express accepted truth, that is, when a claim made is already widely accepted in a particular discipline. Even though the author's viewpoint is not meant to be involved in such cases, the boosters mentioned "mark involvement and solidarity with an audience, stressing shared information, group membership, and direct engagement with readers" (Hyland, 1998a: 350). The term "solidarity boosters", introduced by Vassileva (2001: 97), reflects more specifically the inclusion of the reader in a scientific or discourse community and his knowledge of the field (Myers, 1989; Harwood, 2005). The examples mentioned previously could also be considered as solidarity boosters depending on the context in which they have been used. Peacock (2006: 65) argues that boosters can perform different functions depending on the context and that the context itself is "vital in booster research". In this study, the identification of boosters has been carried out through a close analysis of their context of use. The co-text of each verb booster has also been examined in order to identify their patterns of use in correlation with their pragmatic functions.

Cross-disciplinary variation of boosting in research articles has been approached in a limited number of studies. Hyland (1998a) searched for 61 different boosters in a corpus of 56 research articles taken from seven journals representing eight disciplines,

and accounting for a total of 330,000 words. The boosters registered a normalised frequency of 5.88 per 1,000 words with 'will', 'the fact that', 'show', 'clear(ly)' and 'actually' as the most frequent items. Regarding the disciplinary variations, boosters were most frequent in Philosophy (9.7 per 1,000 words), followed by Marketing (7.1), Applied Linguistics (6.2), Physics (6.0), Mechanical Engineering (5.0), Sociology (4.6), Biology (3.9) and Electrical Engineering (3.2). Hyland (1998a) explains the variations in the use of boosters among the different scientific fields, and especially between humanities/social sciences on the one hand, and science and Engineering on the other, by pointing to distinct disciplinary writing conventions in these disciplines. According to the author (Hyland, 1998a: 370), writers in his corpus, and especially in 'soft disciplines', used boosters to emphasise the strength of their commitment to a proposition and, in this way, aimed to convince the reader of the "the logical force of the argument". Another way in which boosters were used was to comment impersonally on the validity of their propositions, and in particular to emphasise the relationship between the data and the claims made. Furthermore, writers in 'soft sciences' used more boosters as they rely more on "personal projection" (Hyland, 1998a: 372), while 'hard sciences' used fewer boosters because of a "preference for impersonal strategies" (Hyland, 1998a: 371). In another study, Hyland (1998b) analysed the use of boosters in 28 research articles from seven journals representing four disciplines. The average normalised frequency of boosters in that particular corpus was lower than in his previous study, with 3.9 boosters per 1,000 words. Boosters were most frequent in Applied Linguistics (4.6 per 1,000 words), followed by Marketing (4.2), Biology (3.5), and Astrophysics (3.0). Even though the results obtained mirrored cross-disciplinary variations reported in the previous study, the author did not provide details of the most frequent boosters, the total number of boosters searched for or their form.

A further and more thorough study of boosters in different disciplines was conducted by Peacock (2006). The author looked into six disciplines, some of them not examined by Hyland (1998a, 1998b): Business (Marketing and Management), Language and Linguistics, Public and Social Administration, Law, Physics and Materials Science, and Environmental Science. The corpus used was also larger than those in Hyland's studies, as it included 216 articles and contained about 1,250,000 words. Peacock (2006) used a pre-established list of 118 booster candidates compiled from previous research on the topic (e.g. Hyland, 1998a; Hunston, 1993; Skelton, 1997; Vassileva, 2001). The corpus was searched for all booster candidates and the cases where they were actually used as boosters were filtered out by evaluating their context of use. The average frequency of boosters in Peacock's corpus was 9.15 per 1,000 words, nearly twice as much as in Hyland's (1998a) study, most probably due to the fact that a more comprehensive list of boosters was used (116 as compared to 61), and also because of the different disciplines examined. The highest proportion of boosters was found in Language and Linguistics (10.98 per 1,000 words), followed by

Law (10.05), Public and Social Administration (9.61), Physics and Materials Science (8.53), Business (Marketing & Management) (7.84), and Environmental Science (7.57). The frequencies in similar disciplines, that is, Language and Linguistics (10.98) and Business (7.84) differ from Hyland (1998a): Applied Linguistics (4.6) and Marketing (4.2). The differences in the data obtained for similar disciplines reveal the importance of the corpus design and the method of corpus analysis (search items) used in this type of research.

One of the most interesting of Peacock's findings was the identification of booster clusters, the use of which resulted in a notable increase in the persuasive force of boosting. Regarding the cross-disciplinary variation in the frequency of booster use, writers in the two sciences (Physics and Materials Science, as well as Environmental Science) used a much higher proportion of boosters of the "evidential or implicit truth" type (Peacock, 2006: 73), such as 'show', 'demonstrate', 'find' or 'establish' than authors in other disciplines. Peacock suggests that writers in sciences seek to minimise their personal involvement in their findings and appear more objective, which is in accordance with Hyland's (1998a) line of argumentation. In addition, however, they present their claims as evidential, that is, fully based on convincing data rather than being attributable to the writer's persuasive skills. The author argues that the choice of the verbs mentioned may reflect a distinct stance towards findings in sciences, and so a different way of validating findings and expressing commitment (Peacock, 2006). Finally, writers in sciences also avoid using first-person references in combination with the verbs mentioned and show preference for 'figure', 'data', or 'model' as nouns in the subject position for those verbs (e.g. 'the data show...'). In comparison to humanities/social sciences, writers in these disciplines resort to more frequent boosting and a broader and distinct range of boosters than in sciences. Following Hyland (1998a), Peacock (2006) claims that this is due to the predominance of more personal presentation and persuasion in 'soft sciences'.

In view of the previous studies, this research will focus not only on the frequency variations of one booster form, the verb, but it will also examine their co-text in the corpora of Engineering, Medicine and Linguistics. The Engineering corpus covers a broad range of Engineering fields, such as industrial, construction or reliability and safety systems, which had not been included by Hyland (1998a). The Linguistics corpus includes Applied Linguistics, Translation Studies, Cognitive Linguistics and Language, so it is more varied than Hyland's and more in line with Peacock's. Finally, the Medicine corpus is a novel research article corpus, which has not been analysed for boosting in previous studies. A close examination not only of booster verb frequencies in the three corpora but also of their context and co-text of use will allow for a more insightful account of boosting usage, which, in turn, will provide useful information for teaching academic writing.

## 2. Corpora and method of study

Three corpora of research articles were compiled for this study and, as has been previously stated, they represented the disciplines of Engineering, Medicine and Linguistics. The disciplines chosen are notably different in terms of research objectives, data and methodologies used and these differences are likely to be reflected in the writing style employed (Hyland, 1998a; Peacock, 2006). For instance, medical research articles are considerably shorter than articles from Engineering and Linguistics, and this difference seems to point to a disciplinary convention followed in that particular field.

The texts for the corpora were extracted from scientific journals included in the Social Science Citation Index and Science Citation Index Expanded (Web of Science). These sources of texts were chosen bearing in mind the need to ensure the highest quality of research reported and of the English language use. In this sense, only texts written by native-proficiency speakers of English were selected, and in case of multiple authorship, by at least one such speaker. Native proficiency was established by the personal details included (name, surname, affiliation) and in case of doubt, the authors were contacted to confirm it. The list of journals used to compile the corpora is included in Appendix I. As can be seen, the number of journals used per field is different in each case. This is due to the availability of JCR journals in each field, for instance, a limited number of such journals are from the Engineering field. Another reason for these differences is the average length of the articles. As medical articles are considerably shorter, many more texts were needed and a broader variety of journals was used. All of the texts were published in the period from 2014 to 2017. It should also be pointed out here that the corpus is still being built and, thus, Table 1 shows the corpora statistics as of the end of 2018.

**Table 1.** Corpora statistics.

Corpus	Tokens	Number of texts	Average number of tokens per text
Engineering	684,743	65	10,535
Medicine	308,200	67	4,600
Linguistics	636,620	80	7,958

As can be seen in Table 1, the Engineering texts are the longest with more than 10,000 tokens per text, followed by the Linguistics texts with nearly 8,000 tokens, and Medicine with the lowest average extension of 4,600 tokens.

The corpora were uploaded to the METOOL online tool (<http://metool.idm.upv.es/>), designed at the Research Institute for Information and Language Processing at the University of Wolverhampton (UK) with the aim of identifying and tagging rhetorical devices in discourse<sup>1</sup>. The three corpora were tagged with METOOL for metadiscourse devices by a group of annotators using a pre-



established list of such devices compiled for this purpose and based on Hyland (2005) and Mur-Dueñas (2011). The working definition for boosters used by the annotators was the one by Mur-Dueñas (2011), emphasising the fact that they are devices that focus on transmitting certainty and conventions and could be the indication of pragmatic conventions in specific academic writing.

In METOOL, the entire texts were read by annotators and the metadiscourse devices (boosters among them) were manually tagged in the tool with the aim of identifying the rhetorical devices associated to academic English in context. This project followed a methodology based on inter-rater reliability, i.e. the documents have been coded by more than one user and the degree of agreement was 100%. Any additional metadiscourse devices found throughout the annotation process were discussed following this process and, finally, added to the list of the search items or rejected, depending on the decision taken. Table 2 includes the list of 57 lexical items tagged as boosters.

**Table 2.** Lexical items tagged as boosters in the corpora.

Nouns	Verbs	Adjectives	Adverbs	Phrases
evidence	show	significant	constantly	fully
fact	determine	extraordinary	clearly	in fact
majority	demonstrate	extensively	significantly	for the most part
assertion conclusion	reveal	considerable	generally	of course
	highlight	clear	largely	to a large extent
	confirm	vast	particularly	in effect
	emphasise	evident	indeed	
	assert	thrilling	widely	
	hold	paramount	highly	
	underscore	utmost	primarily	
	stress		consistently	
	establish		strongly	
	prove		actually	
	know		mostly	
			especially	
			entirely	
			essentially	
			dramatically	
			substantially	
			always	
			exceptionally	
			well within	

Once the corpora had been tagged, they were queried for verb booster tags using the same online tool. In the last stage of this study, the concordances obtained were examined manually in order to analyse the co-text of boosters and evaluate their pragmatic functions.

### 3. Results and discussion

#### 3.1. *Booster frequency patterns*

This section will first report on the average frequency of all the boosters identified in the corpora. After that, verb boosters will be discussed in more detail. Table 3 shows the average frequencies of all boosters found in the corpora.

**Table 3.** Overall frequencies of all the boosters in the three corpora.

	Engineering	Medicine	Linguistics
Number of occurrences	2,349	951	2,777
Normalised frequency per 1,000 words	3.428	3.083	4.362

Fisher's exact test for count data with simulated p-value (based on 2,000 replicates)<sup>2</sup>, applied to check the statistical significance of the frequencies, gave a p-value of 0.000499. For  $p < 0.001$ , the results are significant at the 0.1% level, that is, the relative frequencies reported here show a high level of significance.

As can be seen in Table 4, the highest frequency (4.363) was registered in the Linguistics corpus, while the lowest (3.083) in Medicine. Engineering registered the medium normalised frequency of 3.428 occurrences per 1,000 words. In comparison to the frequency data in the previous studies, the average frequency obtained here for Linguistics is similar to the figure reported by Hyland (1998): 4.6 per 1,000 words, but considerably lower than that given in Peacock (2006): 10.58 occurrences per 1,000 words. As has already been pointed out, many of the corpus analysis results depend on the corpus design and the method of analysis. This study has used the smallest number of boosters (57), compared to Hyland (1998a) with 61 and Peacock (2006) with 116 items, and this is clearly reflected in the range of the frequencies obtained, especially for Linguistics (10.98 in Peacock, 2006). The differences regarding Engineering (3.428) are not so notable, as the frequency obtained in this study is close to Hyland's: 3.2 for Electrical Engineering and 5.0 for Mechanical Engineering. Regarding Medicine (3.083), the frequency is again close to Hyland's data for Biology (3.9), but not to other sciences in his study (Physics with 6.0). In comparison to Peacock (2006), the frequency obtained in this study is considerably lower (Physics and Material Sciences with 8.53 and Environmental Science with 7.57).

**Table 4.** The occurrences and normalised frequencies (per 1,000 words) of booster word forms per corpus.

Booster word form	Engineering		Medicine		Linguistics	
Noun	263	0.384	168	0.545	483	0.759
Verb	983	<b>1.434</b>	264	0.855	964	<b>1.516</b>
Adjective	203	0.296	95	0.308	290	0.456
Adverb	860	1.256	414	<b>1.343</b>	916	1.439
Phrase	40	0.058	10	0.032	124	0.195
Total	2,349	3.428	951	3.083	2,777	4.363

Table 4 shows the usage frequencies of different booster word forms in the three corpora. Verb boosters were the most frequent in Engineering and Linguistics, and the second most frequent in Medicine. As has been mentioned earlier, the high frequency of these boosters in the three corpora had motivated this study. Table 4 also indicates that adverbs were the second most frequent boosters in Engineering and Linguistics, and the most frequent boosters in Medicine. The frequency data point to notable differences in the patterns of use of the different booster word forms in the corpora studied.

### ***3.2. Verb booster frequency patterns***

A similar number of verb boosters was used in each corpus: 16 in Engineering, 14 in Medicine and 15 in Linguistics (see Table 5). Fourteen of the verbs identified were overlapping items in the three corpora: ‘show’, ‘determine’, ‘demonstrate’, ‘prove’, ‘hold’, ‘stress’, ‘establish’, ‘highlight’, ‘conclude’, ‘reveal’, ‘confirm’, ‘know’, ‘emphasise’ and ‘underscore’. It can be argued, therefore, that these verbs tend to perform boosting functions regardless of the field, at least as far as Engineering, Medicine and Linguistics are concerned. Only one verb, ‘indicate’ was found in Engineering and Linguistics, but not in Medicine, and ‘assert’ was used exclusively as a booster in Engineering. In this sense, the variations found for the three fields are notably small.

**Table 5.** Normalised frequency per 1,000 words of verb boosters identified per corpus.

Engineering		Medicine		Linguistics	
show	0.486	show	0.282	show	0.638
determine	0.229	determine	0.188	determine	0.156
demonstrate	0.150	demonstrate	0.110	demonstrate	0.140
prove	0.111	establish	0.065	establish	0.124
hold	0.098	highlight	0.045	hold	0.090
stress	0.082	reveal	0.045	know	0.090
establish	0.074	know	0.026	reveal	0.074
highlight	0.064	prove	0.019	confirm	0.038
conclude	0.038	confirm	0.019	prove	0.030
reveal	0.035	hold	0.013	stress	0.030
confirm	0.035	emphasise	0.013	conclude	0.028
know	0.015	stress	0.010	emphasise	0.030
emphasise	0.010	conclude	0.010	highlight	0.025
underscore	0.003	underscore	0.010	assert	0.017
indicate	0.003			underscore	0.006
assert	0.001				
Total	1.434		0.855		1.516

Another significant finding is the three most frequent verb boosters, which were found to be the same in the three corpora: ‘show’, ‘determine’ and ‘demonstrate’. This pattern is in line with the previous finding of the high number of overlapping verb boosters in the three corpora. In addition, if a mean frequency of verb boosters is calculated for each corpus<sup>3</sup>, the following values are given: 0.090 for Engineering, 0.061 for Medicine and 0.101 for Linguistics. This was used to identify the verb boosters with a significant normalised frequency, that is, a frequency above the mean value (highlighted in bold in Table 4). There are five verbs like that in the Engineering corpus (‘show’, ‘determine’, ‘demonstrate’, ‘prove’ and ‘hold’) and four verbs in the Medicine and Linguistics corpora (‘show’, ‘determine’, ‘demonstrate’ and ‘establish’). As can be seen, a few different verbs appear on these two lists: ‘prove’ and ‘hold’ in Engineering and ‘establish’ in Medicine and Linguistics. These partly overlapping verb boosters registering normalised frequencies above the mean value point to a consistent pattern in verb booster usage in the three corpora. The verbs with a frequency below the mean value are more numerous in each corpus, accounting for between 10 and 11 depending on the corpus. This is an interesting finding, as it reveals that most of the boosting performed through verbs is carried out by means of four to five verbs that are continuously repeated, while the remaining items are used more occasionally or on a one-off basis. This is especially true for ‘show’, which registered a significantly higher frequency than the next verb on the list in each corpus. For instance, in the Engineering corpus, ‘show’ registered a normalised frequency of 0.486 as compared to 0.229 for ‘determine’. Similarly, ‘show’ registered 0.282 in Medicine, while the following verb, ‘determine’ had a normalised frequency of 0.188. Finally, in Linguistics, the difference is even more remarkable: 0.638 for ‘show’ and 0.156 for ‘determine’. As can be seen, ‘show’ performs a prominent

boosting role in the discourse of the three scientific fields. The fact that the same verbs ('show', 'determine', 'demonstrate', 'establish') used to convey evidential or implicit truth are consistently used as boosters in the three corpora questions Peacock's (2006) claim that writers in sciences use a higher proportion of them than in other disciplines, such as humanities/social sciences. In this study, for instance, 'show' and 'demonstrate' registered a higher frequency than in Medicine (Table 4).

### 3.3 Noun-verb combinations with 'show', 'determine' and 'demonstrate'

'Show', 'determine' and 'demonstrate' were the most frequent verb boosters in the three corpora, but with distinct frequencies registered in each corpus (Table 5). A further examination of the co-text of these verbs was conducted to identify possible similarities and differences in the patterns of use in the three corpora.

'Show', as a booster, was most often used in the predicate position followed by a noun clause ('...show that...') or by a non-finite clause ('...was shown to be...'). Therefore, the nouns used in the subject position with 'show' have been examined. Table 6 includes the most frequent comparable nouns for each corpus, that is, with similar percentage ranges. These percentages refer to the proportion of use of each noun in comparison to the total number of nouns combining in the subject position with the booster 'show' in each corpus. The full list of nouns with the corresponding percentages has been included in Appendix II.

**Table 6.** Nouns used in the subject position with the booster 'show'.

Engineering		Medicine		Linguistics	
noun	%	noun	%	noun	%
figure	17.0	study	22.22	we	10.03
results	15.61	data	9.72	example	9.76
we	11.81	results	5.56	authors	8.71
study	5.91	finding	4.17	I	6.86
authors	3.38	research	4.17	data	5.28
		table	4.17	study	5.28
		trials	4.17	results	5.01

The full list of nouns (Appendix II) shows notable differences among their numbers in each corpus: 61 in Engineering, 35 in Medicine and 85 in Linguistics. These figures may be related to the size of the corpora, with Medicine being the smallest. However, the similar size of the Engineering and Linguistics corpora and a distinct number of nouns combining with 'show' reveal certain differences. These differences can be explained by a higher frequency of this verb in Linguistics, but also the specificity of the field. As has been mentioned previously (Hyland, 1998a), writers in humanities/social sciences may tend to use a more personal writing style, but may

also use more synonymous words to avoid repetition (e.g. participants-persons-speakers, and so forth).

Table 6 includes the most frequent and comparable in terms of the range of percentages of nouns combining with the booster ‘show’. The grey boxes highlight the nouns that appear in the three corpora, while the blue ones show the nouns that are overlaps in any combination of two corpora. For instance, ‘results’ and ‘study’ were overlaps in all three corpora within the highest range of percentages, ‘we’ and ‘authors’ were used both in Engineering and Linguistics, and ‘data’ overlapped in Medicine and Linguistics. As can be seen, Linguistics shares the most overlaps. The nouns exclusively used in each corpus with that percentage range were fewer in Engineering and Linguistics, but not in Medicine, with four non-overlapping nouns. One thing that caught our attention is the use of the personal pronouns ‘I’ and ‘we’ in Linguistics, which confirms the view of a more personal stance taken in this type of academic discourse (Hyland, 1998a). Example (3) shows the use of ‘I’ in this corpus. However, ‘we’ was also frequently used as the subject of ‘show’ in Engineering (see example 1) with a similar percentage to that found in Linguistics (11.81 and 10.03). This finding clearly contradicts the claims made by Hyland (1998a) about a more impersonal style of writing in this field. Finally, example 3 shows the use of ‘finding’ in the Medicine corpus: a combination with ‘show’ that is characteristic of this corpus.

- (1) In our results *we show* averages of six simulations, each with a different starting year to reduce the effect of weather variability of individual years. (Engineering)
- (2) In addition to the intrinsic clinical significance of these exposures (Nanni et al., 2012), our *findings show* that they are also associated with the co-occurrence of psychotic experiences and depression, which further indicates elevated clinical significance (Wigman et al., 2012). (Medicine)
- (3) In Section 5, *I show* that using this conceptualisation of a POV makes better predictions for English spatial expressions as well. (Linguistics)

Regarding the second most frequent verb booster in the three corpora, ‘determine’, Table 7 includes the nouns combining with this verb in the object position when it was used in the active voice and in the subject position when it was used in the passive voice.

**Table 7.** Nouns used in the object position with the booster ‘determine’ in the active voice and in the subject position in the passive voice.

Engineering		Medicine		Linguistics	
noun	%	noun	%	noun	%
indices	23.76	trends	9.76	case	6.0
value	5.94	classifications	4.88	value	5.0
distributions	3.96	factors	4.88	extent	4.0
relationship	3.96	profile	4.88	assignment	3.0
		variability	4.88	direction	3.0
		association	4.88	order	3.0
				relationship	3.0
				structure	3.0

In comparison to noun-verb combinations with ‘show’, the booster ‘determine’ was used with many more nouns specific to each corpus within the highest percentage range. In other words, fewer overlaps were identified in this group of nouns. Only two of them, ‘value’ and ‘relationship’, are overlapping items in the Engineering and Linguistics corpora. No overlaps were identified in this group for Medicine and another corpus, thereby showing that the pattern of use of the booster ‘determine’ is entirely distinct from the other two corpora and bears notable differences with regard to Engineering and Linguistics. The examples that follow (4-6) are instances of the noun-verb combinations specific to each corpus.

- (4) Onsite process engineers have years of collective experience and knowledge about how to *determine* the unit process operational *indices* from plant production indices. (Engineering)
- (5) However, our study suggests that use of this outcome as a sole performance measure has significant limitations. First, overall in-hospital mortality is low in this population, and thus small numbers of deaths could *determine* RSMR *classifications*. (Medicine)
- (6) *Case* in the (c) examples may be *determined* either via case overwriting, on the raising analysis, or via normal mechanics of case assignment to NPs, on the non-raising alternative. (Linguistics)

Appendix III includes the full list of the nouns combining with ‘determine’. As with ‘show’, the number of nouns varies in each corpus, the highest being in Engineering with 72 items, followed by Linguistics with 66 items, and Medicine with 33. However, the broadest range of nouns combining with the booster ‘determine’ was found in Engineering and not in Linguistics, as was the case of the booster ‘show’. The patterns of use of these two verbs vary slightly in this sense.

Finally, and with regard to the booster ‘demonstrate’, Table 8 shows the most frequent nouns combining with this verb in the subject position and within the highest percentage ranges. Like ‘show’, ‘demonstrate’ was most often found to be

used in the predicate position followed by noun clauses ('...demonstrate that...') or non-finite clauses ('...demonstrated to be...'). Appendix IV shows the list of all the nouns combining with 'demonstrate'. The data reveal similarities regarding the number of nouns identified between Engineering (35 items) and Linguistics (36 items), with Medicine registering the lowest number of nouns, namely 13. Overall, 'demonstrate' combined with fewer nouns as compared to 'show' and 'determine', which is most probably due to a lower frequency of this verb in the three corpora.

**Table 8.** Nouns used in the subject position with the booster 'demonstrate'.

Engineering		Medicine		Linguistics	
noun	%	noun	%	noun	%
authors	12.90	study	28.57	examples	23.29
we	9.68	data	9.52	authors	12.33
study	8.06	analysis	9.52	I	4.11
figure	4.84	we	9.52	section	4.11
research	4.84	findings	4.76	study	4.11
		recommendation	4.76	work	4.11
		discussion	4.76		
		research	4.76		
		results	4.76		
		benefit	4.76		
		need	4.76		
		work	4.76		

As can be seen in Table 8, overlapping nouns combining with the booster 'demonstrate' in three or two corpora mirror the data reported for 'show' in Table 5. That is, there are few overlapping verbs, such as 'study' in the three corpora, and 'authors', 'research' and 'work' in any combination of two corpora. Here again, personal pronouns such as 'I' and 'we' were found in the top positions as subjects of 'demonstrate', in a similar way to 'show'. In other words, 'we' was the second most frequent subject in Engineering and 'I' was the third most frequent in Linguistics, reflecting the personalisation features in the two corpora. Finally, many more nouns were found to be exclusively used in Medicine, and this pattern also resembles the usage of both 'show' and 'determine'. The examples that follow (7-9) show the use of the exclusive nouns in each corpus.

- (7) *Figs. 3 and 4 demonstrate* clear differences between the relative importance of the attributes for welfare compared to their financial importance to the farm business. (Engineering)
- (8) However, these *data*, also modest in size, *demonstrate* an accumulating radial experience in treating CTO disease with higher success rates later in the learning. (Medicine)
- (9) *Example (38) demonstrates* the interaction of NP loci with directional verbs. (Linguistics)



Finally, it should be noted that nearly all the nouns identified as combining with the three verb boosters are either semi-technical or general meaning words. This is interesting, as one would expect to find many more technical terms reflecting the specificity of the three fields. As this is not the case, and because the range of the non-overlapping nouns is broad, the data provided in this study are of special interest for teaching academic writing in English. The possibility of having access to a list of nouns combining with the three most frequent verb boosters in each discipline is highly valuable as a source of information and materials for teaching activities in an academic English classroom. As has been previously discussed, boosting and hedging (Hyland, 1998a, 1998b; Peacock, 2006) are the two main discourse pragmatic functions identified in research articles that should be successfully used by scientific writers. This study provides useful corpus-based data to help develop this particular persuasive skill not only with regard to the verbs themselves, but also to their context and co-text of use.

## **CONCLUSIONS**

This study reports variations in the use of verb boosters and their co-text in Engineering, Medicine and Linguistics corpora. With this analysis we have answered the objectives and research questions posed at the beginning of this study. Regarding the first research question, we have proved that boosting is context dependent and we found different frequencies of it in the specific fields of knowledge analysed. The causes of these differences may be the extent to which the studies described by academic writers are of an empirical nature and if the authors felt it was important to convince readers with data. In the second research question, we have shown the most common noun-verb combinations with ‘show’, ‘determine’ and ‘demonstrate’ in section 4.3 of this paper, Tables 6, 7 and 8, and also, in Appendices II, III and IV. In the data shown, several patterns have been identified and discussed to establish their pragmatic functions.

As a summary of the research carried out here, verb boosters have been examined since they turned out to be the most frequent booster forms in Engineering and linguistics, and the second most frequent in Medicine. Their importance for boosting in research articles in these three fields is therefore especially relevant. Even though verb boosters have been found to be most frequent in Linguistics, followed by Engineering and Medicine, which confirms previous research (Peacock, 2006), significant similarities in their usage have also been identified. Boosting in the three disciplines is performed by means of the same verbs, ‘show’, ‘determine’ and ‘demonstrate’, used in a repetitive manner. As has been discussed previously (Peacock, 2006), they have been used to convey implicit truth, that is, to reinforce the evidentiality of the claims raised, regardless of the discipline. The analysis of the co-text of these verbs, and more specifically of the nouns combining either in the subject

or object position, unveiled further patterns not described to date in the literature. Quite to the contrary of what Hyland (1998a) and Peacock (2006) argued, the personal noun 'we' was combined with 'show' and 'demonstrate' in the Engineering corpus together with other nouns such as 'figure', 'results' or 'study'. In the Linguistics corpus, two personal pronouns, 'we' and 'I' were used with these two verbs, thus confirming the claims raised about a more personal style of scientific reporting in this field.

This study does not only provide quantitative data on the use of verb boosters in corpora that had not previously been examined to that end (e.g. Medicine and some Engineering fields), but can also usefully inform the teaching of English for academic purposes. The data concerning the typical verb-noun combinations reported here have clear pedagogical applications in cases where students struggle to formulate claims about their research and reinforce the use of boosting.

## REFERENCES

- Ädel, A. (2006). *Metadiscourse in L1 and L2 English*. Amsterdam/ Philadelphia: John Benjamins.
- Becher, T. & Trowler, P. (2001). *Academic tribes and territories*. Buckingham: The Society for Research into Higher Education and Open University Press.
- Carrió-Pastor, M. L. (2016). A contrastive study of interactive metadiscourse in academic papers written in English and in Spanish. In F. Almeida, L. Cruz García & V. González Ruiz (Eds.), *Corpus-based studies on language varieties*. Bern: Peter Lang.
- Carrió-Pastor, M. L. (2019a). Authorial engagement in business emails. In C. Sancho Guinda (Ed.), *Engagement in Professional Genres* (pp. 47-65). Amsterdam: John Benjamins.
- Carrió-Pastor, M. L. (2019b). Do writers express the same attitude in historical genres? A contrastive analysis of attitude devices in the Corpus of History English Texts. In I. Moskovic, E. Sánchez Barreiro, I. Lareo & P. Lojo Sandino (Eds.), *Writing history in Late Modern English. Explorations of the Coruña Corpus* (pp. 237-259). Amsterdam: John Benjamins.
- Gillaerts, P. & Van de Velde, F. (2010). Interactional metadiscourse in research article abstracts. *Journal of English for Academic Purposes*, 9, 128-139.
- Harwood, N. (2005). 'We Do Not Seem to Have a Theory ... The Theory I Present Here Attempts to Fill This Gap': Inclusive and Exclusive Pronouns in Academic Writing. *Applied Linguistics*, 26(3), 343-375.
- Holmes, J. (1984). Modifying illocutionary force. *Journal of Pragmatics*, 8, 345-65.

- Hu, G. & Cao, F. (2011). Hedging and boosting in abstracts of applied linguistics articles: A comparative study of English-and-Chinese-medium journals. *Journal of Pragmatics*, 43, 2795-2809.
- Hunston, S. (1993). *Evaluation and ideology in scientific writing*. In M. Ghadessy (Ed.), *Register analysis: Theory and practice* (pp. 57-73). London and New York: Pinter Publishers.
- Hunston, S. & Thompson, G. (2000). *Evaluation in text. Authorial stance and the construction of discourse*. Oxford: Oxford University Press.
- Hyland, K. (1997). Scientific claims and community values: Articulating an academic culture, *Language and Communication*, 17(1), 19-31.
- Hyland, K. (1998a). Boosting, hedging and the negotiation of academic knowledge, *Text*, 18(3), 349-82.
- Hyland, K. (1998b). Persuasion and context: The pragmatics of academic metadiscourse. *Journal of Pragmatics*, 30, 437-55.
- Hyland, K. (2005). *Metadiscourse: Exploring interaction in writing*. London: Continuum International Publishing Group.
- Hyland, K. (2017). Metadiscourse: What is it and where is it going? *Journal of Pragmatics*, 113, 16-29.
- Hyland, K. & Tse, P. (2004). Metadiscourse in academic writing: A reappraisal. *Applied linguistics*, 25(2), 156-177.
- Hyland, K. & Jiang, E. K. (2018). In this paper we suggest: Changing patterns of disciplinary metadiscourse. *English for Specific Purposes*, 51, 18-30.
- Moya, P. & Carrió-Pastor, M. L. (2018a). Estrategias de intensificación en los comentarios digitales sobre noticias. *Spanish in Context*, 15(3), 369-391.
- Moya, P. & Carrió-Pastor, M. L. (2018b). Análisis comparativo de los marcadores de compromiso en los comentarios sobre noticias digitales en España y Chile. *Onomázein. Revista de lingüística, filología y traducción. Special issue*, 4, 26-48.
- Moya, P. & Carrió-Pastor, M. L. (2018c). La atenuación en los comentarios sobre las noticias digitales en periódicos de España y Chile. *Onomázein. Revista de lingüística, filología y traducción*, 40, 56-76.
- Mur-Dueñas, P. (2011). An intercultural analysis of metadiscourse features in research articles written in English and in Spanish. *Journal of Pragmatics*, 43, 3068-3079.

- Myers, F. (1989). The language politicians use, (review of *The Language of Politics* by M. L. Geis, 1987). *World Englishes*, 8(2), 243-246.
- Peacock, M. (2006). A cross-disciplinary comparison of boosting in research articles. *Corpora*, 1(1), 61-84.
- Qin, W. & Uccelli, P. (2019). Metadiscourse: Variation across communicative contexts. *Journal of Pragmatics*, 139, 22-39.
- Rundbald, G. (2007). Impersonal, general, and social: The use of metonymy versus passive voice in medical discourse. *Written Communication*, 24(3), 250-277.
- Salager-Meyer, F. (1994). Hedges and textual communicative function in medical English written discourse. *English for Specific Purposes*, 13, 149-170.
- Skelton, A. (1997). Studying hidden curricula: Developing a perspective in the light of postmodern insights. *Curriculum Studies*, 5(2), 177-193.
- Swales, J. (1990). *Genre analysis: English in Academic and Research Settings*. Cambridge: Cambridge University Press.
- Swales, J. (2004). *Research genres: Explorations and applications*. Cambridge: Cambridge University Press.
- Vassileva, I. (2001). Commitment and detachment in English and Bulgarian academic writing. *English for Specific Purposes*, 20(1), 83-102.

## APPENDIX I

### Corpora journal list

Engineering corpus	Medicine corpus	Linguistics corpus
Landslides	American Academy of Dermatology	The Translator
IEEE Transactions on Industrial Engineering	American College of Cardiology	Metaphor and Symbol
Journal of Construction Engineering & Management	American College of Cardiology	Language, Cognition and Neuroscience
Reliability Engineering & System Safety	American Journal of Infection Control	Mind & Language
	American Journal of Preventive Medicine	Journal of Linguistics
	British Journal of Oral and Maxillofacial Surgery	Applied Linguistics
	British Medical Journal	Language and Literature
	Cancer	English for Specific Purposes
	Cancer Epidemiology	Journal of English for Academic Purposes
	Cardiovascular interventions	Cognitive Linguistics
	Cardiovascular Radiation Medicine	

	Clinical Pediatric Emergency medicine"	
	Immunity (Cambridge)	
	JMIR PUBLIC HEALTH AND SURVEILLANCE	
	Journal of Psychiatric Research	
	Journal of Surgical Research	
	Lancet Public Health	
	Patient Education and Counselling	
	Perspectives on Medical Education	
	The Lancet	
	The Lancet Oncology	

## APPENDIX II

Nouns combining with the booster 'show' in the subject position.

Engineering	%	Medicine	%	Linguistics	%
figure	17.30	study	22.22	we	10.03
results	15.61	data	9.72	example	9.76
we	11.81	results	5.56	authors	8.71
study	5.91	finding	4.17	I	6.86
authors	3.38	research	4.17	data	5.28
model	2.95	table	4.17	study	5.28
table	2.95	trials	4.17	results	5.01
analysis	2.53	analysis	2.78	analysis	3.17
data	2.53	authors	2.78	languages	2.37
tests	2.53	BMMNCs	2.78	figure	2.11
orientation	2.11	burns	2.78	model	2.11
mass	1.69	drugs	2.78	research	2.11
example	1.27	figure	2.78	table	1.85
research	1.27	network	2.78	experiments	1.58
section	1.27	tracts	2.78	interaction	1.58
values	1.27	we	2.78	speakers	1.58
comparison	0.84	amitriptyline	1.39	evidence	1.32
evidence	0.84	approach	1.39	findings	1.32
estimates	0.84	care	1.39	participants	1.32
experiment	0.84	cells	1.39	comparisons	1.06
literature	0.84	diagram	1.39	excerpt	1.06
measurements	0.84	doctors	1.39	section	1.06
method	0.84	information	1.39	system	1.06
proof	0.84	injection	1.39	work	1.06
sensors	0.84	interventions	1.39	form	0.79
simulations	0.84	investigation	1.39	line	0.79
validation	0.84	irregularities	1.39	pronouns	0.79
application	0.42	movements	1.39	relativization	0.79
approach	0.42	papers	1.39	situation	0.79
calculation	0.42	patients	1.39	contrast	0.53
cluster	0.42	ratios	1.39	effects	0.53
computation	0.42	sucrose	1.39	English	0.53

contrast	0.42	test	1.39	nouns	0.53
correlation	0.42	therapy	1.39	questions	0.53
curve	0.42	transplantation	1.39	recording	0.53
ellipticity	0.42			responses	0.53
failures	0.42			tokens	0.53
findings	0.42			variables	0.53
indicator	0.42			variation	0.53
inspection	0.42			accents	0.26
investigations	0.42			account	0.26
mass	0.42			adjectives	0.26
matrix	0.42			agreements	0.26
mechanisms	0.42			argument	0.26
module	0.42			assumption	0.26
paper	0.42			clauses	0.26
plot	0.42			conversation	0.26
probabilities	0.42			co-occurrence	0.26
problem	0.42			diagnostics	0.26
projects	0.42			differences	0.26
proof	0.42			element	0.26
records	0.42			explanation	0.26
researchers	0.42			expressions	0.26
samples	0.42			features	0.26
scenarios	0.42			females	0.26
solution	0.42			inspection	0.26
stability	0.42			interface	0.26
transformations	0.42			judgments	0.26
treatment	0.42			manipulation	0.26
trend	0.42			males	0.26
variables	0.42			marker	0.26
				measures	0.26
				men	0.26
				nasals	0.26
				observations	0.26
				patterns	0.26
				plot	0.26
				processes	0.26
				production	0.26
				proficiency	0.26
				projects	0.26
				prompts	0.26
				ratio	0.26
				researchers	0.26
				script	0.26
				sibilant	0.26
				simulation	0.26
				status	0.26
				stops	0.26
				structure	0.26
				support	0.26
				tests	0.26
				transcript	0.26
				understanding	0.26
				words	0.26

## APPENDIX III

Nouns combining in the object position with the booster ‘determine’ in the active voice and in the subject position in the passive voice.

Engineering	%	Medicine	%	Linguistics	%
indices	23.76	trends	9.76	case	6.0
value	5.94	classifications	4.88	values	5.0
distributions	3.96	factors	4.88	extent	4.0
relationship	3.96	profile	4.88	assignment	3.0
hazards	2.97	variability	4.88	direction	3.0
parameters	2.97	association	4.88	order	3.0
clusters	1.98	content	2.44	relationship	3.0
effect	1.98	course	2.44	structure	3.0
emissions	1.98	cycles	2.44	agreement	2.0
flux	1.98	degree	2.44	axes	2.0
magnitude	1.98	differences	2.44	behaviour	2.0
mass	1.98	duration	2.44	decision	2.0
performance	1.98	growth	2.44	effect	2.0
strength	1.98	instances	2.44	expressions	2.0
threshold	1.98	level	2.44	form	2.0
volume	1.98	progress	2.44	impact	2.0
weights	1.98	proliferation	2.44	marking	2.0
analysis	0.99	rate	2.44	position	2.0
area	0.99	reduction	2.44	position	2.0
attribute	0.99	relationships	2.44	properties	2.0
behaviour	0.99	results	2.44	allomorphy	1.0
characteristics	0.99	significance	2.44	architecture	1.0
choice	0.99	source	2.44	association	1.0
coefficient	0.99	stability	2.44	clusters	1.0
cohesion	0.99	standards	2.44	conditions	1.0
combination	0.99	stenosis	2.44	curve	1.0
competition	0.99	studies	2.44	degree	1.0
conditions	0.99	types	2.44	differences	1.0
conductivity	0.99	variables	2.44	domain	1.0
constants	0.99	evidence	2.44	elements	1.0
correction	0.99	calculations	2.44	exponence	1.0
costs	0.99	receipt	2.44	factors	1.0
datasets	0.99	cities	2.44	features	1.0
deviation	0.99			frame	1.0
dimensions	0.99			function	1.0
direction	0.99			groups	1.0
discrimination	0.99			head	1.0
duration	0.99			inputs	1.0
efficacy	0.99			involvement	1.0
equipment	0.99			mean	1.0
error	0.99			meaning	1.0
function	0.99			messages	1.0
growth	0.99			mismatches	1.0
impact	0.99			models	1.0
increase	0.99			movement	1.0
influence	0.99			nature	1.0
interfaces	0.99			novelty	1.0
lengths	0.99			outcome	1.0

level	0.99			outputs	1.0
likelihood	0.99			pattern	1.0
location	0.99			phaschood	1.0
mechanisms	0.99			phrasings	1.0
nature	0.99			placement	1.0
patterns	0.99			predicate	1.0
possibility	0.99			processes	1.0
potential	0.99			productions	1.0
ratio	0.99			pronoun	1.0
reproduction	0.99			realisation	1.0
resources	0.99			reduplication	1.0
scale	0.99			relation	1.0
sequence	0.99			rule	1.0
significance	0.99			scope	1.0
sizes	0.99			selection	1.0
states	0.99			sets	1.0
subsystems	0.99			way	1.0
surface	0.99			word	1.0
term	0.99				
time	0.99				
timing	0.99				
trade-offs	0.99				
transition	0.99				
variation	0.99				

## APPENDIX IV

Nouns combining with the booster ‘determine’ in the subject position.

Engineering		Medicine		Linguistics	
noun	%	noun	%	noun	%
authors	12.90	study	28.57	examples	23.29
we	9.68	data	9.52	authors	12.33
study	8.06	analysis	9.52	I	4.11
figure	4.84	we	9.52	section	4.11
research	4.84	findings	4.76	study	4.11
paper	3.23	recommendation	4.76	work	4.11
illustration	3.23	discussion	4.76	data	2.74
data	3.23	research	4.76	facts	2.74
results	3.23	results	4.76	patterns	2.74
examples	3.23	benefit	4.76	speakers	2.74
analysis	3.23	need	4.76	we	2.74
sample	3.23	work	4.76	agreement	1.37
literature	1.61	authors	4.76	aim	1.37
section	1.61			analysis	1.37
abnormality	1.61			attitudes	1.37
investigation	1.61			cases	1.37
technology	1.61			cluster	1.37
report	1.61			constructions	1.37
topography	1.61			derivations	1.37
method	1.61			figure	1.37
test	1.61			metrists	1.37
theorem	1.61			outputs	1.37
ratio	1.61			paper	1.37



initiation	1.61			polling	1.37
development	1.61			possibility	1.37
table	1.61			processes	1.37
practices	1.61			report	1.37
procedure	1.61			research	1.37
efficacy	1.61			results	1.37
detection	1.61			selection	1.37
group	1.61			speaker	1.37
methodology	1.61			structures	1.37
omission	1.61			subjects	1.37
program	1.61			teenager	1.37
probabilty	1.61			tests	1.37
				utility	1.37

## NOTES

<sup>1</sup> The tool was funded by the Ministerio de Economía y Competitividad (Spain), research project number FFI2016-77941-P.

<sup>2</sup> R software for statistical computing was used in this case.

<sup>3</sup> This was done by adding the frequencies included in Table 4 and, then, dividing the figure obtained by the number of verb boosters.