

# Subject-auxiliary inversion in academic prose

**Carlos Prado-Alonso**

Universidad de Oviedo (Spain)

pradocarlos@uniovi.es

## Abstract

This paper describes a corpus-based analysis of subject-auxiliary inversion in academic prose texts. The focus of the analysis is Chen's (2013) *x Auxiliary Subject* construction (XASC), where *x* codes the fronting of a constituent which triggers the inversion of the auxiliary and the subject, as in *Little did she understand what was being dealt with* or *What does that mean?*

It is argued that the distribution of subject-auxiliary inversion in academic prose is related to the degree of an addressor's involvement in the texts. It will be shown that the more involvement in an academic prose text, the more inversions are to be expected. Furthermore, the data will show that subject-auxiliary inversion is far more frequently attested in learned exposition academic prose texts than in scientific exposition academic prose texts, and that the construction itself can be seen to serve as a discourse marker through which an addressor's involvement is coded in academic writing.

**Keywords:** Corpus, subject-auxiliary inversion, textual variation, academic prose, learned and scientific exposition.

## Resumen

### *La inversión de sujeto y auxiliar en textos académicos*

Este artículo examina la inversión de sujeto y auxiliar en textos de prosa académica inglesa mediante un análisis de corpus lingüísticos. El análisis toma como punto de partida la construcción *x Auxiliary Subject*, tratada en Chen (2013), donde la *x* representa un constituyente topicalizado que provoca la inversión del auxiliar y del sujeto como, por ejemplo, en las siguientes oraciones: *Little did she understand what was being dealt with* / *What does that mean?*

En el artículo se argumenta que, en la prosa académica del inglés, la distribución

de la inversión de sujeto y auxiliar está relacionada con el grado de implicación del emisor en los textos. El estudio demuestra que cuanto mayor sea el grado de presencia del autor en el texto, más uso de inversiones de sujeto y auxiliar cabe esperar. El análisis de corpus también demuestra que la inversión de sujeto y auxiliar es mucho más frecuente en la prosa académica de las ciencias humanas y sociales que en la de las ciencias exactas, y también pone de manifiesto que la construcción se puede utilizar como marcador del discurso que permite la presencia del emisor en este tipo de prosa.

**Palabras clave:** Corpus, inversión de sujeto y auxiliar, variación textual, prosa académica, ciencias humanas y sociales y ciencias exactas.

## 1. Introduction

Subject-auxiliary inversions are syntactic constructions in which the subject follows the first auxiliary in the verb phrase, as in (1)-(2).<sup>1</sup>

- (1) Nowhere has bankruptcy of the PDRY's system been more apparent than in the petroleum sector. (J37; political sciences)
- (2) How can these be reassembled as a reconstruction of the original plant? (J02; natural sciences)

Two main types of subject-auxiliary inverted constructions are distinguished in English: *verb-first inverted* and *verb-second inverted* constructions. These two types of inversion are labelled *Auxiliary Subject Constructions* (ASC) and *X Auxiliary Subject Constructions* (XASC) in Chen (2013). ASC (verb-first inversions) are subject-auxiliary inversions in which the auxiliary is the first syntactic constituent in the clause and is followed by the subject and the main verb, as in examples (3)-(4), below.

- (3) Should you find this memoir in its present state too tedious to pass on to our respectable Society, you will make only extracts from it. (J02; natural sciences)
- (4) Be it enacted by the Senate and House of Representatives of the United States of America in Congress assembled. (J55; humanities)

By contrast XASC, henceforth *XAS inversions*, are subject-auxiliary inversions in which the auxiliary is not the first constituent in the clause (verb-second inversions), and are triggered by the fronting of a constituent which is followed by the auxiliary and the subject, as illustrated in (5)-(6).

- (5) What does DNA do exactly? (J13; medical sciences)
- (6) Only now do they share a limit, namely the end-point itself. (J43; political sciences)

The two types of subject-auxiliary inversions mentioned above have been the subject of extensive research from a functional perspective (cf. Dorgeloh, 1997; Goldberg, 2006; Granath, 2007; Chen, 2013; Duffly, 2016; or Kim, 2018, among others). This article is a further contribution to this line of research and offers a comprehensive corpus-based analysis of XAS inverted constructions (verb-second inversions) in academic prose. There have been other corpus-based analyses of XAS inversions in present-day English, but these suffer from limitations of various kinds, as they focus only on particular types of subject-auxiliary inversions. For instance, Dorgeloh (1997), hailed as a major contribution to the field, does not deal with the analysis of subject-auxiliary inversions in interrogative clauses (cf. 5) and restricts her investigation to the declarative mood. Similarly, Granath (2007) is restricted to subject-auxiliary inversions triggered by deictics (cf. 6) or Kim (2018) to inversions in exclamatives. A more comprehensive corpus-based account, taking into account all XAS inversion types, is thus needed, in that only in such a way can we achieve a conclusive picture of the distribution and functions of these constructions in present-day English texts. The present study is a first step in this direction.

The current corpus-based analysis is based on Chen's (2013) classification of XAS inverted constructions (cf. Section 2). In what follows, I present an in-depth corpus-based analysis of the factors that drive the distribution and pragmatic use of XAS inversions in academic prose texts. The data are taken from four computerised corpora of British and American Present-day English, namely the FLOB, FROWN, BE06, and AmE06 corpora, comprising academic prose texts dealing with *medical*, *natural*, *political*, and *social sciences*, as well as with texts related to the study of *humanities* and *technology* and *engineering* from the 1990s and 2000s (for details see Hofland et al., 1999, and Baker, 2009). The distribution and frequency of the subject-auxiliary inversions in the corpus texts will be compared with Biber's (1988) well-known multidimensional textual analysis. On the basis of such a comparison, I will argue that XAS inversions do not only serve an interpersonal function in academic prose, but they are also inherently related to the degree of involvement of the academic prose texts in which they occur.

The paper is organised as follows. Section 2 offers a discussion of Chen's

(2013) constructional analysis of subject-auxiliary inversions, plus a classification of the XAS inverted constructions used in this research. Section 3 offers information on the academic prose texts analysed in this study. Section 4 sets out the results of a corpus analysis of the types and distribution of XAS inversions, which will be the basis for the pragmatic and textual analysis developed in Section 5. Finally, Section 6 offers some concluding remarks.

## 2. A classification of subject-auxiliary inversions: Chen (2013)

There have been numerous ways of classifying subject-auxiliary inversions in the literature. Some classifications are based on syntactic criteria alone. Newmeyer (1998) in fact argues that subject-auxiliary inversion includes a disparate group of structures, which do not share many functional properties, and the way to capture their commonality can only be formal.

By contrast, other classifications are based on semantic criteria alone. Jacobsson (1951), for instance, follows a basic semantic approach and distinguishes between subject-auxiliary inversions with *restrictive introductory members*—in which the clause-initial constituent triggers the inversion (cf. 7)—and *non-negative introductory members*, where the inversion does not include a clause-initial constituent with negative meaning, as in (8).

- (7) Aristotle did not necessarily intend these lectures to be treated as a single work, nor did he even necessarily regard them all as essays of logic. (J51; humanities)
- (8) The rate of divorce has increased, as has the proportion of men who are never or previously married. (J64; humanities)

Some classifications are based on a mixture of syntactic and semantic criteria (cf. Schmidt, 1980, or Dorgeloh, 1997). Schmidt (1980: 11-13), for instance, distinguishes between subject-auxiliary inversions after *adverbs with negative, restrictive or affective meaning*, after *other openers*, after *so, nor, neither* and *no more*, in *complex sentences with special heads*, in *correlative clauses* and in *comparative*, as in (9).

- (9) For high quality nesting cover, seed mixtures include rhizomatous grasses that provide more contiguous nesting cover than does the scattered cover associated with bunchgrasses (provide). (J70; technology and engineering)

Drawing on, among others, Lakoff & Brugman (1987), Lambrecht (1994), and Diessel (1997), Goldberg (2006) proposes a more detailed classification of subject-auxiliary inversion based on functional/cognitive grounds. She distinguishes eight types of subject-auxiliary inversions (cf. Goldberg, 2006: 166), namely inversions with *counterfactual conditions* (cf. 10), in *wishes/curses* (cf. 11), in *yes/no and wh-questions* (cf. 12), with *negative conjuncts* (cf. 13), with *initial negative or restrictive adverbs* (cf. 14), with *positive rejoinders* (cf. 15), in *exclamative clauses* (cf. 16) and in *comparative clauses* (cf. 9) –which deviate from a prototypical sentence– that is, an independent declarative sentence with a positive polarity and a predicate-focus information structure (cf. Goldberg, 2006: 166).

- (10) Should I leave this job to go to the bathroom, I risk being fired. (Dorgeloh, 1996: 21)
- (11) May God save the United States of America! [quoted from Prado-Alonso, 2011: 22]
- (12) Have you ever had a proper alcoholic drink, a whole drink, not just a sip? (J72; technology and engineering)
- (13) Nor do they appear under the headings “head” and “skull” for Salinan in Greenberg’s Hokan notebook. (J35; political sciences)
- (14) Only then did he prepare a final draft. (J56; humanities)
- (15) As percentage of boys in class increases so does the pace (or total number of discourse moves per hour) with boys. (J55; humanities)
- (16) God, have I seen attitudes change! (Green, 1982: 120)

Chen (2013) further groups Goldberg’s eight types of subject-auxiliary inversion into two main sets of constructions: *Auxiliary Subject Constructions* (ASC), and *X Auxiliary Subject Constructions* (XASC). The main difference between the two groups is that XAS constructions, which are the central concern of the present study, are constructions of focus, with X being the focal point in the entire sentence. As Chen (2013: 9) argues “they are results of fronting a particular element from a position in which it is not assured focus to the sentence initial position, whereby it *is* assured focus”.<sup>2</sup> The XAS group of constructions includes subject-auxiliary inversions in *wh-questions* (cf. 17), triggered by *negative conjuncts* (cf. 18), by *initial negative or restrictive adverbs* (cf. 19), by *positive rejoinders* (cf. 20), by *positive adverbs* (cf. 21), and by *deictics* (cf. 22).<sup>3</sup>

- (17) Why should the denial of it strike Protagoras. (J52; humanities)
- (18) Nor can it be understood solely or even mainly as the result of an English cultural or economic colonization of the so-called Celtic Fringe. (J57; humanities)
- (19) Only recently has attention shifted toward identifying genetic determinants of susceptibility and markers of the early stages of carcinogenesis. (J14, medical sciences)
- (20) Investor can do all right by doing good, so do the brokerage house and money managers who have developed this market niche. (J39; political sciences)
- (21) Particularly do they support universal coverage. (J28; social sciences)
- (22) Only now do they share a limit, namely the end-point itself. (J43; political sciences)

By contrast, ASC constructions are considered constructions of mood indicator, mapped with the notion of irreality, but they are not used to achieve focus. They include Goldberg's inversions in *counterfactual conditions* (cf. 10), in *wishes/curses* and/or *exclamatives* (cf. 11), and in *yes/no questions* (cf. 12).

Since the present study is functionally oriented, the classification of subject-auxiliary inversions in what follows is based on Chen's (2013) detailed functional/cognitive classification. The six XAS inversion types mentioned above will be the focus of this investigation. What follows offers information on the texts analysed in the study, as well as an account of the distribution of subject-auxiliary inversion in the textual categories of the corpus, which will serve as the basis for an in-depth textual and pragmatic analysis presented in Section 5.

### 3. The corpora

To analyse the behaviour and distribution of subject-auxiliary inverted structures in academic prose, four computerised corpora of British and American English texts from the 1990s and 2000s were selected. These are: the *Freiburg-Lancaster-Oslo/Bergen Corpus of British English* (FLOB; texts from 1991, released in 1999), 2) the *Freiburg-Brown Corpus of American English* (FROWN; texts from 1992, released in 1999), 3) the *British English 2006 Corpus*

(BrE06; texts from 2004-2008, released in 2008), and 4) *The American English 2006 Corpus* (AmE06; texts from 2004-2008, released in 2008); for details see Hofland et al. (1999) and Baker (2009). These four corpora were selected for two main reasons. Firstly, they match in their internal structure and, as will be shown presently, allow the compilation of a substantial amount and various types of academic prose texts. Secondly, FLOB, FROWN, BRE06 and AME06 were chosen because they match the structure of the *Lancaster-Oslo-Bergen Corpus of British English* (LOB; compilation date: 1961) corpus and the *Brown Corpus of American English* (BROWN; compilation date: 1962) corpus. LOB and BROWN are analysed in Biber's (1988) multidimensional analysis of linguistic variation which, as discussed in Section 5, will be used here for comparative purposes as tested criteria of linguistic variation.

The FLOB, FROWN, BrE06 and AmE06 corpora comprise 1,000,000 words each distributed into fifteen textual categories one of which is the *Academic Prose* textual category. The academic prose texts analysed comprise samples of approximately 2,000 words each, totalling 608,000 running words organised into six well-defined textual categories: *natural sciences, medical sciences, social behavioural science, political science, humanities, and technology and engineering*, as shown in Table 1.<sup>4</sup> Following Biber's (1988) multidimensional analysis of linguistic variation, these categories were further subdivided into two main types of academic prose texts: "scientific exposition" (medical sciences, natural sciences and technology and engineering) and "learned exposition" (social sciences, political sciences and humanities).

Textual categories		Words
SCIENTIFIC EXPOSITION	Technology and Engineering	96,000
	Medical Sciences	40,000
	Natural Sciences	96,000
LEARNED EXPOSITION	Social Behavioural Sciences	122,000
	Political Sciences, Law, Education	120,000
	Humanities	134,000
	<b>Total</b>	<b>608,000</b>

Table 1. Sources and distributions of the textual categories selected from the corpora.

Since the textual categories in the corpora differ in size, the data were normalised following Biber's (1988: 14) proposal for a "normalised

frequency of a feature”. As Biber notes, “raw frequency counts cannot be used for comparison across texts when they are not at all of the same length”, since in this case longer texts would tend to have higher frequencies simply because there are more opportunities for a feature to occur within them. Using Biber’s procedure and comparing the frequency per 100; 1,000; 10,000, or 100,000 words –depending on the frequency of the feature under investigation– this possible bias is eliminated. In the present study, given that subject-auxiliary inverted structures are considered to be relatively rare syntactic constructions compared to unmarked SVX word order (cf. Biber et al., 1999: 926), raw frequencies are normalised by 100,000 words.

Given the size of the corpus, the use of software tools would have greatly facilitated the retrieval of inverted constructions. Unfortunately, the corpus-based search mainly had to be performed manually. Manual reading of the corpora was necessary since these corpora are neither parsed nor tagged. An automated analysis of the electronic database was only possible for the retrieval of subject-auxiliary inversion in *wh*-questions, as these subject-auxiliary inverted constructions have very specific *wh*-triggers, namely *what*, *which*, *who*, *whom*, *when*, *where*, *how* or *why*. To retrieve examples of XAS in *wh*-questions, the software tool Antconc 3.5.8 was used (cf. Anthony, 2019). The retrieval of the particular subject-auxiliary inversion in *wh*-questions by looking for the triggers also yielded a certain amount of junk, that is, results that did not belong to the kind of construction being sought, and a further manual reading of the examples retrieved from the automated search was required in the end.

#### **4. A corpus-based analysis of subject-auxiliary inversion in scientific and learned exposition**

The analysis of the corpora yielded 298 instances of XAS inverted structures, as shown in Table 2, below. As can be seen, XAS inversions are slightly more frequently attested in the declarative clauses (171 tokens) than in the interrogative clauses (127 tokens) but, overall, there are no dramatic discrepancies in the distribution of these constructions in the declarative and interrogative in terms of frequency. However, despite these similar frequencies, it should be noted that XAS inversions are still mainly associated with interrogative clauses in academic prose, since the overall number of



declarative clauses in this genre is much higher than the number of interrogative clauses.

Inversion in WH-questions is the only type of XAS inversion attested in the interrogative and the most frequent XAS inverted construction attested in the corpus (127 instances/normalised frequency 19.8). For reasons that will be discussed presently (cf. Section 5), this inversion type is more frequently attested in learned exposition (100 instances/normalised frequency 26.6) than in scientific exposition (27 instances/normalised frequency 11.6), where the construction is more frequently used in social sciences (38 instances/normalised frequency 31.1) and humanities (40 instances/normalised frequency 29.8). As Biber et al.'s (1999: 212) corpus-based analysis shows, there is a high frequency of wh-questions in spoken texts, and fictional dialogues which are typically modelled on conversation, because in these types of texts the situation often tends to be interactive, with a constant give-and-take among participants. It has been argued that the main function of wh-questions is to achieve such an interaction and that the construction is used to seek information from the addressee (cf. Kim, 2018: 481). By contrast, academic prose seldom includes interactions or direct speech clauses and would be expected to make an infrequent use of XAS inversions in wh-questions. The frequent occurrence of these constructions in the academic prose texts analysed here is therefore surprising. As will be discussed in Section 5, the motivation for this distribution is that inversion in wh-questions is used in academic writing to explain or describe something by posing a question and then providing an answer. It will also be shown that, in the textual categories analysed here, the construction serves an interpersonal function, which allows the addressor's presence in discourse, and its distribution is related to the degree of involvement in the text.

		Scientific exposition			Learned exposition					
		Technology/ Engineering	Medical sciences	Natural sciences	Social sciences	Political sciences	Huma- nities			
WORDS		96,000	40,000	96,000	122,000	120,000	134,000			
Mood	Trigger	TOTAL CORPORA	TOTAL			TOTAL				
Interrogative	Wh-word	127 (19.8)	15 (15.6)	6 (10.4)	6 (15)	27 (11.6)	38 (31.1)	22 (18.3)	40 (29.8)	100 (26.6)
Declarative	Positive rejoinder	101 (15.7)	16 (16.6)	3 (7.5)	8 (8.3)	27 (11.6)	29 (23.7)	23 (19.1)	22 (16.8)	74 (19.6)
	Negative conjunct	42 (6.5)	3 (3.1)	2 (5)	8 (8.3)	13 (5.6)	12 (9.8)	9 (7.5)	8 (5.9)	29 (7.7)
	Negative adverb	24 (3.7)	1 (0.4)	1 (2.5)	4 (4.1)	6 (2.5)	6 (4.9)	5 (4.1)	7 (5.2)	18 (4.8)
	Positive adverb	2 (0.3)	-	-	-	-	2 (1.6)	-	-	2 (0.5)
	Deictic element	2 (0.3)	-	-	-	-	-	2 (1.6)	-	2 (0.5)
	Total declarative	171 (26.7)	20 (20.8)	6 (15)	20 (20.8)	46 (19.8)	49 (40.1)	39 (32.5)	37 (27.6)	125 (33.2)

Table 2. Distribution of XAS inversion in the academic prose categories of FLOB, FROWN, BrE06, and AmE06.

As the data suggest, inversion triggered by a positive rejoinder is the most frequent type of XAS construction in declarative clauses (101 instances/normalised frequency 15.7). This XAS inversion involves syntactic structures which are triggered by a pro-form that stands for an entire predicate or a substantial part of it. This goes for XAS constructions triggered by *so*, *such*, *as*, etc. functioning as anaphoric adverbs, as shown in (23)-(24), below. In these types of XAS inverted structures, the pro-element is a grammaticalised device that stands for the preceding predicate. The construction itself is also fairly idiomatic, because, as noted by Biber et al. (1999: 916), there is often no completely equivalent SVO word-order. In (23a), for instance, the initial *so* stands for given information, and has a cohesive effect. Furthermore, its location in preverbal position emphasises the parallelism between the two clauses. The subject, which is the main communicative focus, is placed towards the end of the clause, in accordance with the principles of *Communicative Dynamism* and *End-focus* (cf. Hartvigson & Jakobsen, 1974: 62). This inverted pattern, however, can only be paraphrased with subject-verb order plus additive *too* (cf. 23b), i.e. *the kind of language selected fluctuates too*.

- (23a) When these variables fluctuate, so does the kind of language selected.  
(J32; political sciences)
- (23b) When these variables fluctuate, the kind of language selected does too.
- (24) Discrimination is illegal in all countries but the definition of discrimination varies considerably, as does the quality of enforcement.  
(J22; social sciences)

XAS inversion triggered by positive rejoinders is attested far more frequently in learned exposition (74 instances/normalised frequency 19.6) than in scientific exposition (27 instances/ normalised frequency 11.6). Frequency of occurrence is also consistently higher in the textual categories of learned exposition, with social sciences (29 instances/normalised frequency 23.7) and political sciences (23 instances/normalised frequency 19.7) showing the highest scores. The same goes for XAS inversions triggered by negative conjuncts which, as Table 2 illustrates, occur more frequently in learned exposition (29 instances/normalised frequency 7.7) than in scientific exposition (13 instances/normalised frequency 5.6), with social sciences (12 instances instances/normalised frequency 9.8) and political sciences (9 instances/normalised frequency 7.5) also showing the highest scores. As Dorgeloh (1997: 91) notes, in XAS constructions triggered by negative conjuncts, such as *neither* or *nor*, there is both a connective (anaphoric) and a negative component in the meaning. In other words, the XAS construction performs a clause-linking function while at the same time it focuses on the negative semantics of the fronted conjunct. This is illustrated in (25)-(26), below, in which the fronted negative conjuncts, though equally anaphoric, are not just back-referring constituents, but also, due to their fronting positions, provide a stronger negative meaning for the sentence than their canonical SVO counterparts (cf. 25b)

On the basis of the corpus, *nor* is the most frequent negative conjunct used in this construction, especially in learned exposition, though examples of XAS inversion with *neither* are also attested. XAS inversions with *nor* are, as noted by Kjellmer (1979: 292), devices of textual cohesion “by means of which different parts of an argument can be held together, and which allow a writer insertions or deviations without jeopardising the stringency of his argument”.

- (25a) You cannot take pound of clean air away from the common stock and sell it on a market stall, nor can you exclude anybody from enjoying the benefits of clean air. (J41; political sciences)

- (25b) You cannot take pound of clean air away from the common stock and sell it on a market stall, and you cannot exclude anybody from enjoying the benefits of clean air either.
- (26) Just as stones or balls do not find entry through the compact surfaces of walls, neither do rays upon mirrors. (J12, natural sciences)

According to Green (1982: 130), inversion after negative conjuncts and pro-forms “seems to be much more frequent in speech”. However, the present corpus-based study suggests that they are also frequently attested in academic prose. Academic prose requires many cohesive constructions as it mostly takes place under conditions of *displacement*, that is, it deals with events, which are not part of the immediate environment of addressor or addressee. Formal cohesion must be tight because the situational context is not available to help fill in any possible gaps. Academic discourse exhibits a far more structured syntax than other types of genres. As an example, spoken texts tend to have a more fragmented character than academic prose texts and, as Chafe (1992: 68) notes, often string together connected clauses without connectives because the connections are ‘in the air’ and, if misunderstood, can be rapidly repaired. In speech, the relationship between ideas is encoded by means of paralinguistic features such as pitch, prominence, pauses, changes in tempo and voice quality, and gestures. This is not the case in academic prose, where there is only the linear flow of words, so this must be exploited to the full. The result is that –in order to integrate a succession of ideas into a more complex, coherent, and integrated whole– academic prose makes use of many cohesive devices, including XAS inversions triggered by a pro-form or a negative conjunct.

On the basis of the present data (cf. Table 2), XAS inversions triggered by a fronted restrictive or negative phrase, as illustrated in (27)-(28), represent the third most frequent type of XAS inversion in declarative clauses (cf. Table 2).

- (27) Only since the 1970’s, or so, have people in Indonesia called themselves gays or lesbians. (J43; political sciences)
- (28) Nowhere has this been more evident than in Burrell and Morgan’s (1979) collapsing of ontology and epistemology into a single subjective axis. (J11; natural sciences)

Inversions triggered by a negative or restrictive phrase were also more commonly attested in learned exposition (18 instances/normalised frequency 4.8) than in scientific exposition (6 instances/normalised

frequency 2.5). In learned exposition, they were more commonly attested in humanities (7 instances/normalised frequency 5.2) and social sciences (6 instances/normalised frequency 4.9). Inversions of this type are, as Biber et al. (1999: 915) note, more frequently found in texts where a strong rhetorical effect is required. As Huddleston & Pullum (2002: 820) note, “negators mark clausal negation more readily when positioned early in the clause”. This is certainly the case in XAS inversion triggered by a negative or restrictive phrase, in which, due to their prominent placement in preverbal position, there is a further intensification of the force of the fronted constituent. In other words, these constructions aim to connect clauses but also to produce a strong negative or restrictive emphatic effect (cf. Section 5). For instance, in (28) the placement of the negative adverb serves to emphasise the importance of the negative meaning of the proposition. As I have argued elsewhere (cf. Prado-Alonso, 2019), other types of texts make less consistent use of such inversions. This is the case of spoken texts, which rely not only on linear word-order. In speech, the speaker may also employ phonological features to structure and emphasise negative information. This is shown in (29b), where pitch prominence is given to the negative element and no inversion is required to achieve a negative emphasis on the clause.

(29a) Never before had women pursued college degrees.

(29b) Women had NEVER BEFORE pursued college degrees.

Finally, the other types of XAS inversions in the declarative mood –namely, inversions triggered by a fronted positive adverb or a deictic adverb (2 instances/normalised frequency 0.03)– are found very rarely (cf. Table 2) and their frequency is of very marginal significance in the data. Both subject-auxiliary inversions types would thus seem to be very infrequent in present-day English academic prose. As I have shown elsewhere, (cf. Prado-Alonso, 2016), Inversion with fronted deictics most frequently triggers the inversion of subject and main verb (cf. 30) –XVS inversion– rather than the inversion of subject and auxiliary (cf. 31), and they have been shown to be most commonly attested in speech than in writing.

(30) Here comes the first inequality. (J22; social sciences)

(31) And, there, will one find significant false-negative rates. (J36; political sciences)

The analysis presented thus far has provided a corpus-based account of instances of XAS inversion retrieved from academic prose texts, together with a preliminary discussion of the reasons behind the differences in distribution. The data clearly show that the different types of subject-auxiliary inverted constructions are more frequently attested in learned exposition than in scientific exposition. In order to examine the possible reasons for this marked difference in distribution, a more fine-grained analysis will now be given. Such an analysis will allow for a better understanding of the distribution and pragmatic function of this construction in academic prose. As will be seen in Section 5, this textual analysis is based on established criteria of linguistic variation (cf. Biber, 1988).

## 5. Subject-auxiliary inversion and the degree of addressor's involvement in academic prose

In *Variation across Speech and Writing* (1988), Biber analyses linguistic variation in the Lancaster-Oslo-Bergen Corpus of British English (LOB) and the Brown corpora. These two corpora, compiled in the 1960s, match the structure of FLOB, FROWN, BrE 2006 and AmE 2006 (for details see Hofland et al., 1999 and Baker, 2009). The textual categories represented in LOB and BROWN are analysed by Biber in terms of six parameters or dimensions:

- Dimension 1, which he calls *Involved versus Informational Production*, distinguishes discourse with interactional, affective or involved purposes and which is associated with strict real-time production and comprehension constraints, from discourse with highly informational purposes.
- Dimension 2, *Narrative versus Non-narrative Concerns*, distinguishes discourse with primary narrative purposes from discourse with non-narrative purposes, hence dealing with the difference between active, event-oriented discourse, and more static descriptive or expository types of discourse.
- Dimension 3, *Endophoric versus Situation-Dependent Reference*, distinguishes between discourse that identifies referents fully and explicitly through relativisation, and discourse that relies on non-specific deictics and reference to an external situation for the

purposes of identification. This dimension thus corresponds closely to the distinction between endophoric and exophoric reference (cf. Halliday & Hasan, 1976).

- Dimension 4, *Overt Expression of Persuasion*, refers to those features associated with the addressor's expression of point of view or with argumentative styles intended to persuade the addressee.
- Dimension 5, labelled *Abstract versus Non-abstract Information*, distinguishes between texts with a highly abstract and technical informational focus, and those with non-abstract focus.
- Finally, Dimension 6, *On-line Informational Elaboration*, distinguishes between informational discourse produced under highly constrained conditions in which the information is presented in a relatively loose, fragmented manner, and other types of discourse, be it informational discourse that is highly integrated or discourse that is not informational in nature.

In addition to multidimensionality, variation is treated as continuously scalar in Biber's analysis. The six parameters, then, define continua of variation rather than discrete categories. For example, although it is possible to describe a text as simply abstract or non-abstract, it seems more accurate to describe it as more or less abstract.

Over the last thirty years, Biber's (1988) multidimensional analytical framework has come to be regarded a powerful tool for the analysis of register variation and genre. The application of this framework has been considered to yield well-tested findings in the study of linguistic variation, and has allowed linguists to investigate language in use and to formulate detailed descriptions, which in turn can encapsulate how language users make concrete language choices in particular linguistic contexts. What follows provides a comparison of the distribution of subject-auxiliary inverted constructions in the textual categories of the corpora analysed here with Biber's analysis of the same categories in terms of the Dimension 1, Involved versus Informational Production. The selection of this Dimension for the analysis of XAS constructions in academic prose is based on the fact that, as has been demonstrated (see Chen, 2003, or Prado-Alonso, 2019, among others), other inversion types, such as locative inversion, tend to occur more frequently in texts which exhibit a stronger interpersonal nature. If the distribution of XAS inversion –in academic prose– is sensitive to

involved or informational production, we can assume that it will be seen clearly in the present data.

Taking into account Dimension 1, Involved versus Informational Production, Biber’s analysis shows that, even though the academic prose texts of LOB and BROWN score low in involvement when compared to other text types, the categories of technology and engineering, medical sciences and natural sciences exhibit a very low score on the involved pole of Dimension 1, as illustrated in Figure 1, below. By contrast, social sciences, political sciences and humanities have moderately higher scores on the involved pole of this Dimension. Biber also shows that those categories with higher scores on Dimension 1 exhibit a higher degree of concern for interpersonal and affective meaning and are characterised by markedly infrequent occurrences of nouns, prepositions, and long words. They also exhibit a frequent occurrence of private verbs used for the overt expression

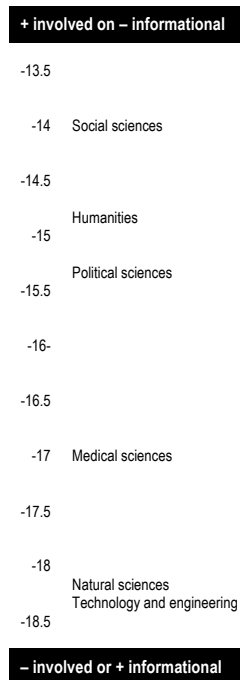


Figure 1. Mean scores of Dimension 1 –involved vs. informational production– in academic prose (adapted from Biber, 1988: 182-183).



of private attitudes, thought, and opinions (e.g. *think, feel*), that-deletions, present tenses, contractions, first and second person pronouns (referring directly to the addressor and addressee), emphatics and amplifiers used by the addressor to mark attitudinal comments, and pro-verb *do*, whose use may be considered to be related to XAS constructions, as some of these constructions require auxiliary *do* (Biber, 1988: 117). By contrast, the textual categories with low scores on this Dimension 1 –technology and engineering, medical sciences and natural sciences– have the opposite characteristics.

Looking first at the distribution of subject-auxiliary inversion in the interrogative mood, the comparison of Biber's mean scores on Dimension 1, illustrated in Figure 1, with the distribution of subject-auxiliary inverted constructions in the categories analysed here (cf. Table 2), shows that, in the interrogative, there is a tendency for those categories with a higher degree of addressor's involvement –namely social sciences, political sciences and humanities– to favour the use of the construction. This is also seen if we measure the correlation between the mean scores on Dimension 1 and the normalised frequencies of subject-auxiliary inversion in the interrogative by calculating a *Pearson Correlation Coefficient* test and a *Simple Linear Regression* test.<sup>5</sup> Correlation and regression are techniques for describing the relationships in data, and are used for answering such questions as whether high values of one variable go with high values of another, or whether one can predict the value of one variable when given the value of another. The Pearson Correlation Coefficient is a statistical test that allows establishing the strength of relationships in continuous variables. In other words, it is a precise measure of the way in which two variables correlate. The strength of the relationship between two variables, in this case the distribution of subject-auxiliary inversion in the academic prose categories and the degree of speaker involvement in those categories, can be expressed numerically using a Pearson Correlation Coefficient. Pearson's correlation coefficient is +1 if two variables vary together exactly. In general, a positive correlation coefficient shows that the two variables are positively correlated, where high values of the first variable are associated with high values of the second. A negative correlation where high values of the first variable are associated with low values of the second and vice versa is shown by a negative correlation coefficient. A value of -1 is obtained for a perfect negative correlation, and a value of 0 is obtained when the two variables are not correlated at all (cf. Butler, 1985; Baayen, 2008; and Johnson, 2008, among

others). As illustrated in Table 3, below, the result of the Pearson Correlation Coefficient for the distribution of subject-auxiliary inversion and Biber's mean scores on Dimension 1 in the categories of Academic Prose is 0.8171. This represents a strong positive correlation, which means that high X variable scores, i.e. frequency of subject-auxiliary inversion in the categories, goes with high Y variable scores, namely the degree of speaker's involvement (and vice versa). The statistical significance of the correlation can be further demonstrated by calculating its *P-Value* which represents the probability that the observed relationship between the two variables in the corpus occurred by chance. P-Values that are significant at the  $p \leq .0.1$  level are commonly considered statistically significant, and  $p \leq .0.05$  or  $p \leq .0.01$  levels are often called "highly" significant. In the present correlation the P-Value is 0.047119, which is significant at the  $p < 0.05$  level.

	Mean scores of the textual categories on Dimension 1	Normalised frequencies for XAS inversion in the interrogative mood
Technology and Engineering	-18.3	15.6
Natural Sciences	-18.2	15
Medical Sciences	-17	10.4
Political Sciences	-15.3	18.3
Humanities	-14.9	29.8
Social Sciences	-14	31.1
<b>Correlation coefficient</b>		<b>0.8171</b>
<b>P-Value (significant at <math>p &lt; 0.05</math>)</b>		<b>0.047119</b>
<b>Simple linear regression (<math>R^2</math>)</b>		<b>0.6677</b>

Table 3. Pearson correlation coefficient for the distribution of subject-operator inversion in the interrogative mood and Biber's (1988: 182-183) mean scores on Dimension 1.

In the present quantitative investigation, there is one variable, i.e. the dependent variable (the reason for the distribution of subject-auxiliary inversion in the textual categories), that we want to explain and another variable, i.e. the independent variable (degree of speaker's involvement), that is believed to affect the dependent variable. Calculating a Simple Linear Regression Test is a way of predicting the behaviour of the dependent variable according to the values of the independent variable. The key output of a regression analysis is called the *Coefficient of Determination* –denoted by  $R^2$ – which is interpreted as the proportion of the variance in the dependent variable that is predictable from the independent variable. In other words, the Coefficient of Determination indicates the percentage of the variation in the dependent variable that the independent variable explains collectively; its result ranges from 0 to 1. An  $R^2$  of 0 means that the dependent variable cannot be predicted from the independent variable, an  $R^2$  of 1 shows that

the dependent variable can be predicted without error from the independent variable and an  $R^2$  between 0 and 1 indicates the extent to which the dependent variable is predictable from the independent variable. In the present data, the Coefficient of Determination ( $R^2$ ) is 0.6677, which suggests that the distribution of the textual categories in terms of involvement explains at least 66% of the examples of the distribution of subject-auxiliary inversion in the interrogative mood.

The statistical analysis therefore shows that, in the interrogative, the preponderance of inversion in learned exposition correlates with the degree of the addressor's involvement in these text types. The XAS inversions attested in the interrogative are well suited to the interpersonal features of these textual categories, and the data further show that the more involved a text, the more XAS inversions in the interrogative are to be expected.

In the interrogative, the subject-auxiliary inversions analysed here are used to seek information about non-subject parts of the sentence, as illustrated in (32)-(35), below. As Chen (2013: 4) notes, they differ from inversions in Yes/No questions –i.e. verb first Auxiliary-Subject Constructions (cf. 36)– in that the latter seek information about the proposition expressed by the entire sentence. To seek information about a particular part of the sentence, the addressor fronts the *wh*-lexeme whose referent is automatically placed at the centre of attention and obligatorily triggers the inversion. It is in this sense that this type of inversion has been lexicalised in present-day English, and such a process of lexicalisation fixes the construction as a distinct unit (cf. Brinton & Traugott, 2005; Traugott & Trousdale, 2013).

- (32) How can Protogoras be encouraged that man is not the measure? (J52; humanities)
- (33) What does this mean in terms of railway operations? (J74; technology and engineering)
- (34) WHERE did ALL THESE FRENQUENCY DATA come from? (J32; social sciences)
- (35) Why should the synthetic chemist seek out novel oxidation? (J07; natural sciences)
- (36) Did you know that our genes are located on your chromosomes? (J13; medical sciences)

Since the questioned unit in XAS *wh*-inversions is picked out and placed clause-initially for focus, this results in a stressed-unstressed-stressed

phonological sequence (cf. Chen, 2013: 10), as shown in (34), whereby the fronted *wh*-element is phonologically emphasised and the subject, which typically does not receive stress in an SVO word-order, is also phonologically emphasised. In interrogative clauses, the need to focus is therefore the motivation for this subject-auxiliary inversion. It is in this sense that, in the interrogative, this construction is a discourse marker which has a focus management function. In other words, it is a construction that changes the addressee's current focus of interest or attention and thus belongs to "the speaker's organisation of discourse" (cf. Halliday & Hassan, 1976: 239). The corpus-based data show that this interpersonal function of subject-auxiliary inversion is most frequently attested in texts with a higher degree of interpersonal features. In examples (37)-(38), below, for instance, the writer poses questions to focus the attention on particular aspects of content in the text. In (37), the three XAS inversions triggered by *how* and *what* signal a strong emotional involvement on behalf of the writer, who makes use of a frequent strategy in academic prose: posing a question to announce a research theme or to focus the attention on a particular aspect or aspects of the phenomenon under discussion. In (37), the two XAS inversions are used to express a forceful statement. Here the addressor follows a natural didactic strategy used in academic writing: that of explaining by posing a *wh*-question and then providing an answer. In both instances, the subject-auxiliary inversions allow the addressor to be present in the text and to shift the reader's attention.

- (37) How then can teacher educators respond –and in essence positively counter– prospective teachers' resistance to teach for diversity and for understanding? What are some examples of promising pedagogical strategies that teacher educators could use in their courses to help prospective teachers meet the expectations of their teacher education programs? How can teacher educators better prepare prospective teachers to meet the challenges of helping increase the achievement and participation of all students in mathematics and science? We tackle these questions head-on by providing rich narratives of our experiences [...]. (J22; social sciences)
- (38) Why does Feagin find meta-response to tragedy praiseworthy? First, she claims that the sympathy we enjoy when we respond to tragedy also underlies our capacity for moral action. But what does this establish in terms of moral status? Here, we must inquire further into the grounds of pleasure. (J61; humanities)

In the declarative mood, the preponderance of subject-auxiliary inversion in political sciences, social sciences and humanities, i.e. learned exposition, is also best explained by the degree of addressor's involvement. This can be noticed if we assess the correlation between Biber's mean scores on Dimension 1, illustrated in Figure 1, and the normalised frequencies of subject-auxiliary inversion in declarative clauses by calculating a (Pearson) Correlation Coefficient, as illustrated in Table 4, below. The result is 0.8357, with a P-Value 0.038591, which is significant at the  $p \leq .05$  level. This represents a strong positive correlation, which indicates that, in declarative clauses, there is a tendency for texts with a higher concern for interpersonal or affective content, namely the texts in learned exposition to favour the use of subject-auxiliary inversion. Similarly, the Coefficient of Determination ( $R^2$ ) in the Simple Linear Regression test is 0.6984 which suggests that the distribution of the textual categories in terms of involvement explains, at least, 69% of the examples of the distribution of subject-auxiliary inversion in the declarative mood.

	Mean scores of the textual categories on Dimension 1	Normalised frequencies for XAS inversion in the interrogative mood
Technology and Engineering	-18.3	20.8
Natural Sciences	-18.2	20.8
Medical Sciences	-17	15
Political Sciences	-15.3	32.5
Humanities	-14.9	27.6
Social Sciences	-14	40.1
<b>Correlation coefficient</b>		<b>0.8357</b>
<b>P-Value (significant at <math>p &lt; 0,05</math>)</b>		<b>0.038591</b>
<b>Simple linear regression (<math>R^2</math>)</b>		<b>0.6984</b>

Table 4. Pearson correlation coefficient for the distribution of subject-operator inversion in the declarative mood and Biber's (1988: 182-183) mean scores on Dimension 1.

In declarative clauses, the preponderance of subject-auxiliary inversion in political sciences, social sciences and humanities is therefore best explained by the degree of involvement in these text-styles. The different types of inversions in the declarative clauses analysed here are well suited to these interpersonal features. As argued by Chen (2013), the function of these constructions is the fronting of the preverbal constituent. In other words, he argues that they serve a focalising function and allow the addressor to move the addressee's attention to the constituent placed in front position in the clause, which is given prominence. For instance, inversions triggered by a pro-form or a negative conjunct are not only textual devices which simply

point back to a referent but, as Dorgeloh (1997: 116) notes, produce a further “emotionally expressive” effect, which arises from the status of inversion as a marked construction and the process of reordering. Since English is an SVX word order language, and a “paraphrased” SVX word order is generally available in subject-auxiliary inversion triggered by pro-forms or additive adverbs (cf. 39), the use of this construction involves the breaking of expectations about the use of the unmarked sentence pattern. In these inversions, the addressor thus breaks the unmarked discourse conditions of an SXV word order to emphasise the linking parallelism between the inverted clause and the preceding clause.

- (39a) When these variables fluctuate, so does the kind of language selected and used in conversation. (J32; social sciences)
- (39b) When these variables fluctuate, the kind of language selected and used in conversation does too.

In (39a), for instance, the fronting of the constituent triggers the inversion of the auxiliary and the subject, which attracts informational focus. In this example, both the trigger and the inverted subject receive prominence (which would be represented phonologically in speech). The result is a bi-focal construction in which focus is given to two constituents (the fronted element “so” and the subject “the kind of language selected and used in conversation”), which would not be given prominence in the canonical SVX counterpart (cf. 39b), while at the same time the new information is introduced in discourse and emphasised in the postverbal subject. The inversion reflects the concern of the addressor with the relative weight or prominence of a single element from the clause, namely the pro-form which is emphasised in preverbal position, thus achieving a stronger comparative effect. This is in line with the importance in science of comparison as a means of understanding reality; it is, quite simply, difficult to explain the nature of something without describing how it resembles or contrasts with some other comparable thing. As argued in Stein (1995: 137), the breaking of normal expectations and the emphasis achieved by the fronting of the pro-form has an interpersonal meaning, namely an “affectual or emotive component”.

Similarly, inversion triggered by a negative or restrictive element allows the presence of the addressor in discourse in that, through the use of this construction, she/he can change the addressee’s focus of interest or attention in the text towards an individual constituent, namely the preverbal restrictive

or negative trigger. The negative or restrictive trigger gains emphatic meaning since it is fronted, and the inversion serves a direct function in the explanation or evaluation provided in scientific writing: namely to emphasise the negative meaning of the clause. With the use of the inversion, the addressor singles out a stronger negative meaning and thereby adds to it an interpersonal meaning. This can be seen in (40a) where the initial placement of “only” emphasises the restrictive interpretation of the utterance. The constituent followed by the restrictive trigger, namely the subject, is more strongly affected by the restrictive scope of the trigger. By contrast, in (40b) the restrictive element is part of what is being predicated, it is not moved out of its position, and lacks such an emphatic effect. Subject-auxiliary inversion triggered by a restrictive or a negative element therefore contains an interpersonal meaning: “a judgement relative to the speaker’s beliefs” (cf. Stein, 1990: 279) of the negative or restrictive interpretation of the utterance.

- (40a) Only recently has attention shifted toward identifying genetic determinants of susceptibility and markers of the early stages of carcinogenesis. (J14, medical sciences)
- (40b) Attention has shifted toward identifying genetic determinants of susceptibility and markers of the early stages of carcinogenesis attention only recently.
- (41) Here comes the bus.
- (42) Now do they share a limit, namely the end-point itself. (J43; political sciences)

Finally, inverted constructions triggered by deictic adverbs designate the spatial or temporal location of the addressor while at the same time designating the location of the postposed subject as being close to the addressor. In xvs inversions (cf. 41), Bolinger (1977: 93), for instance, regards the constructions as “presenting something on the immediate stage” by bringing something literally before the addressee’s presence. This claim is further elaborated by Dubrig (1988: 91) who claims that, as illustrated in (41), the construction encompasses a pragmatic presentative function, which consists of “directing the addressee’s conscious attention to an object in his environment by making him focus on a region in his perceptual field”. In subject-auxiliary inversions triggered by deictics (cf. 42), the addressor makes a procedural use of the construction and instructs the addressee to mentally reconstruct a certain lay-out of a temporal or spatial location. In other

words, the preverbal deictic points to a spatial or temporal location and once the addressee's attention is directed towards this location, she/he can focus 'more easily' on the postposed clause constituents, which introduce the new information and receive prominence in discourse, as illustrated in (42), above. The predicates are concrete locative or temporal relations, but, due to a displacement from the situation itself, a conceptual reconstruction in the mind of the addressee is required. The viewpoint established by the inversion becomes a device of discourse focus management in academic prose, that is, it allows the addressor to change the addressee's current focus of interest or attention, and therefore the construction serves an interpersonal function.

## 6. Summary and conclusions

Drawing on Chen's (2013) classification of *X Auxiliary Subject* inversions, the present paper has offered a corpus-based analysis of the distribution and pragmatic use of these types of constructions in academic writing.

The study has shown that, in both the interrogative and declarative moods, these inverted constructions serve an interpersonal function and may be considered discourse markers through which addressor's involvement is reflected in academic prose texts. It is in this sense that the constructions are used by the addressor to change the addressee's focus of interest towards a particular constituent, namely the fronted X. The front shifting of the X constituent triggers the inversion of the auxiliary and the subject, which is also given prominence and attracts the informational focus. Such an interpersonal focus management function is also revealed in the distribution of XAS inversions in the academic prose textual categories analysed here. The inversions occur in linguistic contexts in which the X is fronted in an attempt to accommodate the post-verbal new information into the addressee's knowledge base. The corpus-based analysis has further shown that there is a tendency for XAS inversions to be more frequent in academic prose text-types with a larger number of interpersonal discourse features, namely texts related to learned exposition: humanities, political sciences, and social sciences. The statistical analysis has in fact demonstrated that the more interpersonal in nature an academic prose text is, the more XAS inversions are to be expected.

Article history:

Received 19 September 2019

Received in revised form 2 November 2020

Accepted 15 November 2021



## References

- Anthony, L. (2019). *AntConc* (Version 3.5.8) [Computer Software]. Tokyo, Japan: Waseda University.
- Baker, P. (ed.). (2009). *Contemporary Corpus Linguistics*. Continuum.
- Bolinger, D. (1977). *Meaning and Form*. Longman.
- Baayen, R. H. (2008). *Analyzing Linguistic Data: A Practical Introduction to Statistics Using R*. Cambridge University Press. DOI: <<https://doi.org/10.1017/CBO9780511801686>>.
- Barlow, M. (2000). *Corpus of Spoken Professional American English* (CD-ROM version). Houston, TX.
- Biber, D. (1988). *Variation across Speech and Writing*. Cambridge University Press. DOI: <<https://doi.org/10.1017/CBO9780511621024>>.
- Biber, D., S. Johansson, G. Leech, S. Conrad & E. Finegan (1999). *The Longman Grammar of Spoken and Written English*. Longman.
- Brinton, L. J. & E. C. Traugott (2005). *Lexicalization and Language Change Research Surveys in Linguistics*. Cambridge University Press. DOI: <<https://doi.org/10.1017/CBO9780511615962>>.
- Butler, C. S. (1985). *Statistics in Linguistics*. Oxford: Blackwell.
- Chafe, W. L. (1992). "Information flow in speaking and writing" in P. Downing, S. D. Lima & M. Noonan (eds.), *The Linguistics of Literacy*, 17-29. John Benjamins. DOI: <<https://doi.org/10.1075/tsl.21.05cha>>.
- Chafe, W. L. (1994). *Discourse, Consciousness, and Time: The Flow and Displacement of Conscious Experience in Speaking and Writing*. University of Chicago Press.
- Chen, R. (2013). "Subject-auxiliary inversion and linguistic generalization: Evidence for functional/cognitive motivation in language". *Cognitive Linguistics* 24(1): 1-32. DOI: <<https://doi.org/10.1515/cog-2013-0001>>.
- Diessel, H. (1997). "Verb-first constructions in German" in M. Verspoor, K. Lee & E. Sweetser (eds.), *Lexical and Syntactical Constructions and the Construction of Meaning*, 51-68. John Benjamins. DOI: <<https://doi.org/10.1075/cilt.150.07die>>.
- Dorgeloh, H. (1997). *Inversion in Modern English: Form and Function*. John Benjamins. DOI: <<https://doi.org/10.1075/sidag.6>>.
- Duffy, P. (2016). "The role of do-auxiliary in subject-auxiliary inversion: Developing Langacker's notion of existential negotiation". *Cognitive Linguistics* 27(2): 269-287. DOI: <<https://doi.org/10.1515/cog-2015-0125>>.
- Goldberg, A. E. (2006). *Constructions at Work: The Nature of Generalization in Language*. Oxford University Press. DOI: <https://doi.org/10.1093/acprof:oso/9780199268511.001.0001>
- Granath, S. (2007). "Size matters – or thus can meaningful structures be revealed in large corpora" in R. Facchinetti (ed.), *Corpus Linguistics 25 Years on*, 169-188. Rodopi. DOI: <[https://doi.org/10.1163/9789401204347\\_011](https://doi.org/10.1163/9789401204347_011)>.
- Green, G. M. (1982). "Colloquial and literary uses of inversion" in D. Tannen (ed.), *Spoken and Written Language: Exploring Orality and Literacy*, 119-154. Ablex.
- Jacobsson, B. (1951). *Inversion in English with Special Reference to the Early Modern English Period*. Almqvist Wiksell.
- Kjellmer, G. (1979). "On clause-introductory *nor* and *neither*". *English Studies* 60(3): 280-295. DOI: <<https://doi.org/10.1080/00138387908597971>>.
- Halliday, M. A. K. & R. Hasan (1976). *Cohesion in English*. Longman. DOI: <<https://doi.org/10.4324/9781315836010>>.
- Hartvigson, H & L. K. Jakobsen (1974). *Inversion in Present-day English*. Odense University Press.
- Hofland, K., A. Lindebjerg & J. Thunestvedt (1999). *ICAME Collection of English Language Corpora* (2<sup>nd</sup> ed., CD-ROM version). The HIT Centre.
- Huddleston, R & G. K. Pullum (2002). *The Cambridge Grammar of the English Language*. Cambridge University Press. DOI: <<https://doi.org/10.1017/9781316423530>>.
- Johnson, K. (2008). *Quantitative Methods in Linguistics*. Blackwell.
- Kim, J. (2018). "Inverted wh-exclamative constructions: A construction-based perspective". *Korean Journal of Linguistics* 43(3): 465-484. DOI: <<https://doi.org/10.18855/lisoko.2018.43.3.004>>.
- Lakoff, G & C. Brugman (1987). *The Semantics of Aux-inversion and Anaphora Constraints*. Paper presented at the Annual Meeting of the Linguistic Society of America, San Francisco.
- Lambrecht, K. (1994). *Information Structure and Sentence Form*. Cambridge University Press. DOI: <<https://doi.org/10.1017/CBO9780511620607>>.
- Newmeyer, F. D. (1998). *Language Form and*

*Language Function*. MIT Press.

Prado-Alonso, C. (2011). *Full-verb Inversion in Written and Spoken English*. Peter Lang.

Prado-Alonso, C. (2016). "A constructional analysis of obligatory xvs syntactic structures". *Studia Anglica Posnaniensia* 51(1): 51-82. DOI: <<https://doi.org/10.1515/stap-2016-0002>>.

Prado-Alonso, C. (2019). "Obligatory inversion in scientific texts". *Studia Neophilologica* 91(3): 314-335. DOI: <<https://doi.org/10.1080/00393274.2019.1652112>>.

Quirk, R, S. Greenbaum, G. Leech & J. Svartvik

(1985). *A Comprehensive Grammar of the English Language*. Longman.

Schmidt, D. A. (1980). *A History of Inversions in English*. Unpublished Ph.D. Dissertation: The Ohio State University.

Talmy, L. (2000). *Towards a Cognitive Semantics*. MIT Press.

Traugott, E. C. & G. Trousdale (2013). *Constructionalization and Constructional Changes*. Oxford University Press. DOI: <<https://doi.org/10.1093/acprof:oso/9780199679898.001.0001>>.

**Carlos Prado-Alonso** works as a full-time lecturer in *English Language and Linguistics* at the University of Oviedo and is a member of the Research Group *Variation Linguistic Change and Grammaticalization* based at the University of Santiago de Compostela.

Most of his published work has been on the analysis –from a functional corpus-based perspective– of several English syntactic constructions, including inverted constructions, verbal anaphoric structures, and parenthetical clauses.

## NOTES

<sup>1</sup> Unless otherwise stated, examples of inversions have been retrieved from academic prose texts in four computerised corpora, namely FLOB, FROWN, BrE06, and AmE06. For details of these corpora and the textual categories analysed here, see Section 3.

<sup>2</sup> *Focus* is here understood as a cognitive notion and is defined as the special attention that the addressor draws to a particular part of a sentence (cf. Talmy, 2000).

<sup>3</sup> It should be borne in mind that Goldberg's comparative subject-auxiliary inversion, as seen in (17), is not included in Chen's analysis, in that he considers this construction to be the result of subject postponement rather than of the inversion of subject and auxiliary (cf. Chen, 2013: 10-12). Following Chen's classification, comparative subject-auxiliary inversions are not included in the corpus-based analysis here.

<sup>4</sup> LOB, BROWN, FLOB, FROWN, BrE2006 and AmE2006 also contain a textual category called *mathematics* as part of the Academic Prose category. However, this was excluded from the analysis because it has a limited number of texts, and these contain fewer than 20,000 words.

<sup>5</sup> All statistical analyses were conducted using R (version 3.6.0). See Baayen (2008).