revistadelta

https://dx.doi.org/10.1590/1678-460X202642472916

Articles

Dry writing? Applied linguistics texts are far from that

Escrita seca? Textos da linguística aplicada estão longe disso

Janailton Mick Vitor da Silva*
Deise Prina Dutra**

ABSTRACT

This article presents how stance bundles are used in applied linguistics texts, as well as illustrate how a renowned author of the same area makes use of them. By adopting corpus linguistics as an approach, we have extracted 4-word lexical bundles from two corpora, the Corpus of Articles of Applied Linguistics (CorAAL) and the Biber Corpus (BC). The lexical bundles were functionally analyzed following Biber, Conrad, and Cortes (2004) and Simpson-Vlach and Ellis (2010). After refining the bundle counts, we investigated if stance bundles were pervasive in CorAAL and the extent to which they may contribute to specific authorial characteristics in the case of BC. The results reveal that the use of attitudinal stance bundles is more frequent in both corpora. In BC, however, the evaluation and ability / possibility bundles are even more frequent, which shows the author's intent to add his interpretive voice in the stance he takes. Our results also highlight that academic writing, specifically in the area of applied linguistics, may not fully hold a dry, purely fact-oriented style as many have long believed.

Keywords: *stance lexical bundles; applied linguistics; academic writing; corpus linguistics; authorial voice*

* Universidade Federal da Paraíba. João Pessoa, Paraíba – Brasil. https://orcid.org/0000-0002-5137-5473. E-mail: janailtonm@gmail.com
** Universidade Federal de Minas Gerais. Belo Horizonte, Minas Gerais – Brasil. https://orcid.org/0000-0001-5799-5174. E-mail: deisepdutra@gmail.com

This content is licensed under a Creative Commons Attribution License, which permits unrestricted use and distribution, provided the original author and source are credited.

RESUMO

Este artigo apresenta como pacotes lexicais que revelam posicionamento são usados em textos na área da linguística aplicada, bem como ilustrar como um autor renomado da mesma área faz uso desses pacotes. Ao adotarmos a Linguística de Corpus como abordagem, extraímos pacotes de 4 palavras do *Corpus of Articles of Applied Linguistics (CorAAL)* e do *Biber Corpus (BC)*. Os pacotes foram analisados funcionalmente seguindo Biber, Conrad e Cortes (2004) e Simpson-Vlach e Ellis (2010). Após refinar as contagens desses pacotes, investigamos se os pacotes lexicais eram comuns no *CorAAL* e se poderiam contribuir para características autorais no *BC*. Os resultados revelam que o uso de pacotes lexicais de posicionamento atitudinal é mais frequente em ambos os *corpora*. No *BC*, contudo, os pacotes que expressam avaliação e habilidade / possibilidade são ainda mais frequentes, o que demonstra a intenção do autor de incorporar sua voz interpretativa no posicionamento que adota. Nossos resultados também evidenciam que a escrita acadêmica, especificamente na área de linguística aplicada, pode não conter um estilo seco ou puramente orientado à discussão de fatos como tem sido frequentemente defendido.

Palavras-chave: *pacotes lexicais de posicionamento; linguística aplicada; escrita acadêmica; linguística de corpus; voz autoral*

1. Introduction

Academic registers use language that not only conveys information, but also carries the writers' or speakers' opinions and attitudes (Biber, 2006; Hyland, 2002a, 2002b, 2008b; Ivanič, 1998; Tardy, 2012). In this regard, it seems arguable that linguistic choices may reveal a great deal about authors (i.e. how they convey (un)certainty or how they give opinion on what they are writing), serving to engage them with the audience. Despite the fact that several studies have been developed on academic writing over the last two decades, it has called our attention that academic writing is still oftentimes referred to as dry writing (Hyland, 2002a). In academic writing handbooks for international students (Bailey, 2018), for instance, there is an emphasis on exercises for students to practice the objective style of academic texts and not a single exercise on learning how to use stance markers[1]. Therefore, we

1. Stance can be expressed by "grammatical devices, word choice, and paralinguistic devices (…) such as loudness, pitch and duration, as well as on-linguistic devices such as body position and gestures." (Biber et al., 1999, p. 966). In this paper, we will not focus on stance grammatical devices such as adverbials or complement clauses, or the use of specific words such as *love* or

revistadelta
42.4
2026

build evidence against this common myth largely and classically advocated, which was questioned by Biber (2006, p. 87), that there is an "[…] idealized representation of university language, [in which] there would be no need for stance expressions". Stance choices contest this idealized image that university language has had over the years, because individuals cannot refrain themselves from leaving their fingerprints on what they touch and write. After all, "[…] academic writing is not absolutely objective and depersonalized" (Jalali, 2017, p. 38), and we want here to contribute to a clearer understanding of the academic register, more specifically of applied linguistics authors' stance expression.

In this article, we present how stance is characterized through the use of lexical bundles[2] (Biber, 2006; Biber et al., 2021) in written academic texts in applied linguistics to better understand how authors position themselves and express their authorial identity. While corpus-based studies have more frequently served to describe variations across registers in a language (Biber, 1988), or to characterize one specific register and how it deals with a topic, e.g. obesity in the news (Collins, Baker & Brookes, 2024), they have not often focused on investigating authors' particular writing identity. Studies on this topic, though, include Hyland's (2010) study on Deborah Cameron and John Swales, and Silva's (2025, 2026) investigation on Biber's metadiscourse style. Therefore, by analyzing two corpora[3], the *Corpus of Applied Linguistics Articles* (CorAAL) and the *Biber Corpus* (BC), we seek to demonstrate how applied linguists build personal, evaluative discourse in academic writing, considering the ways they employ epistemic and attitudinal stance-marking 4-word lexical bundles throughout their texts. It should also be noted that our choice of investigating Biber's writing out of other similarly important applied linguists comes from the fact that he has had a distinguished career, being widely cited and figuring among the world's top scientists, according to the 2022 Stanford list of the World's Top 2% Scientists (Ioannidis, 2023). Hence, the expression of stance in both corpora may show that academic writing does not hold a dry, purely fact-oriented style as many have long believed as our data reveals the extent to which specific bundles contribute to Biber's and other applied linguists'

happy. We have chosen to focus on how the combination of words (e.g. *it should be noted that*) that often co-occur in our corpora convey the authors' attitudes and evaluations.

2. A lexical bundle is a group "of words that show a statistical tendency to co-occur." (Biber et al., 1999, p. 989) and its characteristics will be fully discussed in the following sections of this article.

3. The corpora, which contain books, book chapters, and journal articles, will be described in section 3.

authorial characteristics, and may be used, in the future, in academic writing courses for novice writers in the area.

2. Theoretical Considerations on Stance and Lexical Bundles

Academic writers, like any other writer, are capable of construing authorial identity (Ivanič, 1998) in their texts, while also building personal and evaluative discourses: "academic writing, like all forms of communication, is an act of identity: it not only conveys disciplinary 'content' but also carries a representation of the writer." (Hyland, 2002b, p. 1092). This is to say that academic texts are not only suitable for presenting theoretical notions or facts, but also for expressing author's voice (Tardy, 2012).

Even if there are specific linguistic features that scholars choose to use, which can characterize their own style (Biber & Conrad, 2009), we cannot disregard that they are or want to be accepted as part of a scientific community. Individual identity and agency, based on "class, ethnicity, and gender", in other words, life experiences, are intertwined with cultural and professional conventions influencing how scholars write (Hyland, 2010, p. 162). Therefore, understanding stance in academic writing involves considering that there is an individual voice as well as a set of disciplinary practices undertaken by members of that discipline. As Hyland (2008b, p. 144) highlights, "all writing contains 'voice' as it is this which situates us culturally and historically and allows us to locate ourselves in our communities through the linguistic resources our disciplines make available."

Differences in how scholars use specific linguistic features, such as reporting verbs (Hyland, 1999), or epistemic stance strategies, such as hedging and boosters (Hyland, 1998), confirm that disciplinary variation is not random, but reflect a community practice linked to what is central in the discipline. Gray (2015) found that situational characteristics of disciplines are linked to the linguistic choices writers make, and the cline of variation depends on the subject and what they study as well as how important it is for the area to be explicit about the research design, for instance. Hard sciences disciplines (e.g. biology and physics) "focus on describing in detail methodological procedures" and also concentrate on the research itself and its results; thus, their research articles carry high co-occurrence of agentless passive voice constructions, nouns as nominal premodifiers and non-finite relative clauses (Gray, 2015, p. 174). For instance, in the area of

revistadelta
42.4
2026

biochemistry, Poole, Gnann, and Hahn-Powel (2019, p. 1) discovered that authors have employed more epistemic stance markers indexing certainty than uncertainty.

Applied linguistics, in Gray's (2015) results, shared characteristics with political sciences, presenting previous research as well as focusing on how the research was developed and its implications. These disciplines, in a way, also share characteristics with philosophy and history as they portray academic involvement using stance markers. Earlier then, though, Hyland (1998, p. 357) had identified that stance expression through hedges and boosters were clearly more prominent in "[p]apers in philosophy, sociology, marketing, and linguistics [which] contained about two and a half times as many devices overall, with hedges particularly strongly represented" when compared to physics, mechanical and electrical engineering, and biology.

A more recent study shows that there has been a steady increase in authorial stance in academic writing over the past 50 years. There has been, however, a small, yet significant decrease in the expression of stance, with varying results according to the discipline: "[…] rises in hedges in engineering and biology, falls in the use of boosters by all disciplines but engineering, steady uses of attitude markers and, excepting applied linguistics, substantial rises in self-mention." (Hyland & Jiang, 2019, p. 134). The authors also note that such soft knowledge fields as applied linguistics have been gradually moving towards more self-effacing prose over the last 25 years, increasingly resembling hard science practices, partly due to an increase in more quantitative studies which limit opportunities for overt stance-taking. Despite these varying results, "writers cannot, in other words, *not* take a stance […]" (Hyland & Jiang, 2019, p. 133).

In academic discourse, many are the linguistic features used to express individual and community voices in writing (adjectives, e.g. *true* and *appropriate* and adverbial clauses, e.g. *as one might expect*). In our research, voice and stance are not understood as opposing phenomena, as "writers have to make important decisions on the voice they give to the stance and position they take." (Jiang, 2017, p. 86). Therefore, our focus is narrowed to formulaic language that carry stance markers. The understanding of the term *formulaic language* is rather ambiguous in academic literature since it has commonly referred to "chunks, collocations, fixed expressions, formulaic language, idioms, lexical bundles, multiword units, prefabricated units, set phrases" (Frankenberg-Garcia, 2018, p. 93). Although these terms

are interconnected, in this paper, we have chosen to use the term *lexical bundle* as part of the large superordinate *formulaic language* for the reasons introduced next.

Firstly, bundles are defined by Biber et al. (1999), Hyland (2008a), Biber (2006), and Biber et al. (2021) as frequently occurring uninterrupted word sequences, consisting of three or more elements. The scholars have usually focused on the 4-word pattern, mainly because they are more common than longer strings, such as 5-word or 6-word bundles, and their functions are more transparent than 3-word bundles. Because bundles are the "building blocks of discourse" (Biber et al., 2004, p. 371), our focus here underlies their noticeable and important presence in academic writing (Biber et al., 2021; Frankenberg-Garcia, 2018; Hyland, 2008a).

Lexical bundles play a substantial role in assisting writers establish a credible and reliable voice in their academic discipline (Pang, 2010); thus, analyzing how authors make use of them can help the understanding of disciplinary discourse as well as of authorial voice enacted therein. Lexical bundles can be categorized according to the communicative function(s) they perform in written and spoken discourse. Bearing this in mind, Biber, Conrad, and Cortes (2004) proposed a preliminary functional taxonomy of lexical bundles encompassing the following: *stance expressions, discourse organizers*, and *referential expressions*. Much of their work stemmed from Halliday's (1994) seminal work on linguistic macro functions: ideational (helps writers structure activities and experiences of the real world); textual (regards text organization and its elements as a message); and interpersonal (focuses on the writer or reader of the text and the relationships established between them). Hence, bundles can be classified in three:

> *Stance bundles* express attitudes or assessments of certainty that frame some other proposition. *Discourse organizers* reflect relationships between prior and coming discourse. *Referential bundles*[4] make direct reference to physical or abstract entities, or to the textual context itself, either to identify the entity or to single out some particular attribute of the entity as especially important. (Biber et al., 2004, p. 384, our highlights).

4. Referential expression lexical bundles have been found to be the most frequent bundles in corpus-based research on academic registers (Biber et al., 2004; Biber, 2009; Simpson-Vlach & Ellis, 2010). Even in learner corpus studies that concentrate on investigating one academic register (essays), these bundles have the highest frequency since "[r]eferential expressions play a major role in how argumentation is built up, being such a distinctive feature of academic texts, especially in argumentative essays." (Dutra & Berber Sardinha, 2013, p. 125).

revistadelta
42.4
2026

As mentioned in the introduction, our focus in this research lies on *stance bundles* only. To better analyze how they are used to mark authors' level of (un)certainty and attitudes, evaluations, etc. in their texts, the taxonomy adapted from Biber, Conrad, and Cortes (2004) and Simpson-Vlach and Ellis (2010) will be used in this paper (see Table 1). Stance bundles are separated into two broad categories: *epistemic* and *attitudinal / modality*. The *epistemic* bundles comment on the knowledge status of the information in the following proposition, expressing certainty or uncertainty. The *attitudinal / modality* bundles express the speaker's attitudes towards the actions or events described in the proposition, encompassing such meanings as desire, obligation / directive, intention / prediction, ability / possibility, and evaluation.

Table 1. *Functional taxonomy for stance bundles analysis*

Broad Categories	**Subcategories**	**4-word bundle examples**
Epistemic	expressions of uncertainty	*I don't know (what/if), I don't think so, I think it was, are more likely to, assume that the*
	expressions of certainty	*the fact that the, according to the, to show that*
Attitudinal / Modality	desire	*what I want to do, would like to, I don't want to*
	obligation / directive	*you need to know, you have to do, I want you to, it is important to, to ensure that (the), it is necessary to*
	intention / prediction	*what we're going to, is going to be, let me just, to do so, if you want/ed (to)*
	ability / possibility	*to be able to, to come up with, you can see, you can look at, and you're trying to*
	evaluation	*the importance of, is consistent with, it is obvious that, it doesn't matter*

Note. This table was elaborated following Biber, Conrad, and Cortes' (2004, p. 389-393) and Simpson-Vlach and Ellis' (2010, p. 19-20) bundle taxonomy

Upon laying out the theoretical considerations on stance and lexical bundles in this section, we introduce the step-by-step methodology carried out in this investigation next.

3. Methodological Considerations

In this paper, corpus linguistics (CL) is used as an approach (Tognini-Bonelli, 2001) to subsidize the corpus compilation steps, and both the data extraction and the linguistic analysis. Three specific tools from *AntConc*, version 4.2.0, designed by Anthony (2022), have been particularly useful, namely, *KWIC (Key-Word-In-Context), N-Gram,* and *Collocate*.

To investigate how stance bundles are used in applied linguistics texts, as well as illustrate how a renowned author of the same area makes use of them, we used two study corpora. The first corpus, the *Corpus of Articles in Applied Linguistics* (henceforth CorAAL) (Dutra et al., 2020; Prina Dutra; Berber Sardinha, 2021), contains 144 peer-reviewed, multi-authored journal research articles, published between 2013 and 2018. The articles were extracted from three major applied linguistics journals, namely *Journal of EAP* (EAP), including 50 texts; *Language Learning & Technology* (LLT), comprising 50 texts; and *TESOL Quarterly* (TQ), consisting of 44 texts (see Table 2). The journals were chosen because of the wide readership, of their availability to the research group and of the structure of the articles which included sections (e.g. introduction, methodology, etc.) (Dutra et al., 2020). The themes of the articles published in the journals are roughly similar to one another: EAP focused on English for Academic Purposes; LLT on foreign and second language teaching and learning, computer-assisted language learning; and TQ on English language teaching and learning.

Differently from the first corpus, the second one, *Biber Corpus* (henceforth BC), consists of texts written and published by a sole writer: Douglas Biber (see Table 2). The reason for this corpus layout was to illustrate how stance bundles could contribute to particular authorial characteristics. Hence, BC only included Biber's single-authored academic texts, specifically books, book chapters, and journal research articles. These texts were published across 35 years, starting with his first publications in 1984 and proceeding to 2019, year of his last single-authored text as revealed on his personal website[5], thus encompassing 43 texts overall. Lastly, it should be noted that the decision to include books and book chapters besides journal articles in BC stemmed from the fact that they all share similar situational contexts (Biber & Gray, 2016; Biber & Conrad, 2019). This is not to say, however, that there is no difference between each register

5. Available at: https://dougbiber.weebly.com/ Access on: 15 July 2026.

revistadelta
42.4
2026

and that such differences might not have influenced the author's choices. Furthermore, including other texts from the same academic register would be more representative and indicative of Biber's consistent use of stance throughout his writings.

Table 2. *CorAAL and BC text, type, and token numbers*

	CorAAL	BC*
Texts	144	43
Types	21,266	7,799
Tokens	1,022,936	391,166

* BC consists of 43 texts divided as follows: 15 book chapters, 3 books, and 25 journal articles.

In both corpora, files were originally in both .OCR and .PDF formats, and so we proceeded with converting them into .DOCX format. By using basic automated *Microsoft Word* tools, we edited the texts to correct typos and other conversion-resulting errors, such as in "distribu-tion → distribution". Later, we withheld in brackets (<>) pieces of text that were not to be included in the analysis. They included the abstracts and keywords, other writers' citations, titles of journal sections and subsections (<Introduction>), information related to the author (<Douglas Biber Northern Arizona University, Flagstaff>), years and pages of cited bibliographic works (<2000, p. 52>). We also added text-related identification tags at the beginning of each file (<text id = BC.JA.2011.01>). Lastly, we saved all these .DOCX files into .TXT (UTF-8) format. An example of a journal article from *Biber Corpus* can be checked in Figure 1 below:

Figure 1. *Example of a journal article from the Biber Corpus.*

<text id = BC.JA.2011.01>

<Corpus linguistics and the study of literature
Back to the future?>

<Douglas Biber
Northern Arizona University, Flagstaff>

<The present paper introduces corpus-based analytical techniques and surveys some of the specific ways in which corpus analysis has been applied to the study of literature. In recent years, those research efforts have been mostly carried out under the umbrella of 'corpus stylistics'. Most of these studies focus on the distribution of words (analyzing keywords, extended lexical phrases, or collocations) to identify textual features that are especially characteristic of an author or particular text. Corpus-based grammatical and pragmatic analyses of literary language are also briefly considered. Then, in the concluding part of the paper,
I briefly survey earlier computational and statistical research on authorship attribution and literary style. While that research tradition is in some ways the precursor to more recent work in corpus stylistics, it is also complementary to recent research in its application of sophisticated statistical and computational methods.>

<Keywords: corpus stylistics, keyword analysis, collocations, lexical phrases, key clusters, semantic prosody>

We used the *AntConc N-Gram tool* to extract 4-word lexical bundles. The bundles were retrieved from the program following Biber's (2006, p. 134) "frequency-driven approach", which means that they were obtained based on their occurrence and reoccurrence in both corpora and across texts. "[L]exical bundles are identified empirically, as the combinations of words that in fact recur most commonly in a given register." (Biber et al., 2021, p. 984). In this article, we report the investigation of 4-word bundles due to their transparent meaning and also because shorter bundles are often part of longer ones (Biber et al., 2004).

In this sense, for a group of words to be considered a bundle in this research, we set the following cut-off points: a bundle should have a minimum frequency of 10[6], appear in at least five texts, and with at least 10 times per 1,000 words as the normalized frequency. Our decision to use this cut-off point lies in the fact that we have only concentrated on stance bundles, and also because BC is not very large compared to CorAAL (391,166 tokens). This approach resolves methodological problems and aligns with other studies that have dealt with various types of bundles on larger corpora (e.g. Biber et al., 2004; Cortes, 2013; Dutra and Berber Sardinha, 2013). In the next section, the results are presented, showing the refinement criteria used to count the bundles. After that, the stance bundles were functionally classified according to Biber et al.'s (2004) and Simpson-Vlach & Ellis's (2010) taxonomies, always taking into consideration the context of use in each study corpus.

4. Results And Discussion: Stance Bundles In CorAAL And BC

To begin with, the functional classification of all bundles retrieved from CorAAL and BC revealed that most of them are referential and discourse organizer bundles. Despite the fact that they account for 90% of the bundles in CorAAL and 95% in BC, it is relevant to concentrate on stance bundles as our interest is to better understand authors' voices in both corpora. An investigation focused on the extent to which applied linguistics texts reveal authors' stands and evaluations will enlighten our understanding of the discourse construction in this discipline, specifically by one of its

6. In Biber (2006), considering that his sample sizes of university registers were small (e.g. 50,400 words for office hours), different criteria were used. For instance, a bundle that occurred three times was extracted and analyzed as long as it occurred in at least three different texts.

revistadelta
42.4
2026

more prominent authors, Biber. Therefore, we concentrate our thorough investigation on stance over referential and discourse bundles.

In order to have results that reflect stance bundles' real frequency in CorAAL and BC, we refined the bundle counts so as to avoid overlaps. For instance, the 4-word bundle *it is important to* overlapped with the 6-word bundle *it is important to note that*, since the former bundle is part of the latter. Hence, the refinement of the bundle count followed Grondona's methodology (2015)[7] and only the 4-word stance bundles were maintained with an updated number of frequencies (see Table 3).

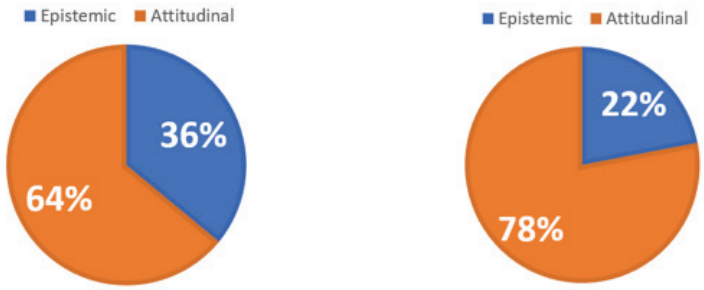
Table 3. *Number of stance bundles before and after the removal of overlapping bundles*

Corpus	Before refinement		After refinement	
	Types	Tokens	Types*	Tokens
CorAAL	63	1,192	39	707
BC	40	722	27	511

*Along the analysis, we will refer to the bundle types that were common in both corpora. Only eight (08) stance bundles out of the 66 extracted from the corpora are exactly the same in CorAAL and BC.

First, our analysis reveals a higher preference for attitudinal / modality bundles (see Figure 2) in both corpora.

Figure 2. *Percentages of stance bundles in CorAAL (left) and in BC (right)*



7. Grondona (2015) proposed an automated methodology to eliminate topic-related and overlapping bundles extracted from corpora to guarantee that the bundles submitted to analysis reflect authors' choices and are not, for instance, a copy of the task instruction, or have their frequency number inflated with bundles that overlap.

As can be seen, the stance bundle distribution shows that, in applied linguistics, not only is expressing certainty and uncertainty about ideas and claims important (see example 1), but that there is also room for writers, like Biber himself, to express their attitudes towards what is being presented, for instance, in relation to the importance of distinguishing between sets of multi-dimensional-related approaches and studies (see example 2). Similar to Hyland and Jiang's (2019) study, the attitudinal stance expression of *importance* has been favored in applied linguistics articles over time.

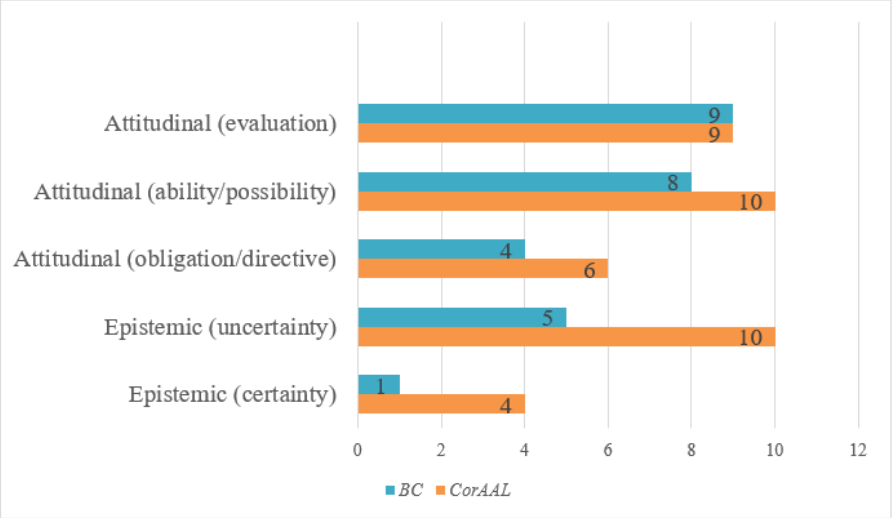
1. Some choices that students make put them at risk: enrolling part-time, delaying entering PSE, and completing only a few credits in their first year of PSE. *It is likely that* financial hardship motivates some of these choices. (CorAAL.TQ.Art.2013.003).

2. As noted in the introduction, *it is important to* distinguish between the multi-dimensional approach to register variation and multi-dimensional studies of particular discourse domains in particular languages. (BC. JA.1995.01).

The higher use of attitudinal over epistemic markers across our corpora seems to resonate with Hyland and Jiang (2019) as well, in that while there has been a decrease in the use of these features in applied linguistics articles across time, attitude markers usage continues steadier than epistemic bundles. More specifically, a careful look into the attitudinal / modality category uncovered that the expression of evaluation and possibility is more favored in both corpora (see Figure 3), which, added to the number of obligation / directive bundle types, show a greater variety in the attitudinal expression.

revistadelta
42.4
2026

Figure 3. *Classification of 4-word stance bundle types across both corpora*



A few points must be highlighted about the types of stance bundles categorized according to Biber, Conrad, and Cortes (2004) and Simpson-Vlach and Ellis (2010) across our corpora (see Figure 3). Firstly, these results promptly reveal that CorAAL authors and Biber have built rather evaluative and personal texts by resorting to a variety of attitudinal stance bundles that clearly made attitudes and evaluations towards their propositions evident to the audience. This finding illustrates how attitudinal stance markers bluntly highlight writers' alignment with the interests of the applied linguistics community in which they participate (Hyland & Jiang, 2019). Second, the corpora texts also contained considerable levels of confidence by means of such epistemic stance bundles as *the fact that the* (see example 3):

3. Despite *the fact that the* Foundation Certificate develops group work skills and includes assessed team projects, this project, they said, was longer and much more complex. (CorAAL.EAP.Art.2015.007).

As can be seen, authors' positioning has also been attested in other studies, such as in Biber (2006), Hyland (2002a, 2002b, 2008b), Ivanič (1998), and Tardy (2012). Biber (2006), for instance, reveals that textbooks from different disciplines, such as business, engineering, natural sciences, social sciences, and humanities use more attitudinal bundles, mainly ability and importance, than the epistemic ones. Our analysis does not disregard the fact that, although articles and textbooks convey knowledge, they are

quite different academic registers. Textbooks from "engineering and natural sciences are based on 'laws' of nature, which are often treated as if they do not have exceptions, resulting in a lesser need for epistemic lexical bundles", while the same types of bundles are more important in humanities and social sciences textbooks as the latter disciplines value exposing their students to different perspectives of a topic and their status (Biber, 2006, p. 168). Journal articles, the academic register that encompasses the majority of the texts in both of our study corpora, also convey knowledge. However, they need to strongly convince readers of the relevance and importance of their findings, even presenting a new perspective, methodology, or contradicting well-accepted knowledge of an area. Hence, stance bundles play an important role in journal articles and so, from now on, we will concentrate on analyses that not only present the bundle functions, but also pinpoint the specific authorial bundle choices and their frequency. In this regard, Tables 4 to 7 display the 4-word stance bundles ordered by frequency in CorAAL and BC and classified functionally according to Biber, Conrad, and Cortes (2004) and Simpson-Vlach and Ellis (2010).

Epistemic Stance Bundles

As we look at Tables 4 to 7, we can notice important differences in the epistemic and attitudinal stance bundles. First of all, most of these frequent lexical bundles represent lexical variation across CorAAL and BC, particularly in the former corpus. This lexical variation in the stance bundles, though, does not reveal an explicit inclusion of a research theme. They help the author put together ideas (e.g. *similar to each* other). Only the bundle *the importance of register*, which is part of longer bundles like *the importance of register variation,* is related to Biber's main research topic: *register variation*. This is, unsurprisingly, an expected result as both corpora differ in their themes, with CorAAL, for example, covering papers on different scholars' investigation agendas, which may concentrate on language proficiency tests, teacher education, research writing, technology and language education, to mention a few.

revistadelta
42.4
2026

Table 4. *Epistemic stance bundles across corpora*

Epistemic	Bundle types	Raw freq. CorAAL (tokens)	Norm freq. CorAAL	Raw freq. BC (tokens)	Norm freq. BC
Certainty	*the fact that the*	46	65	20	39
	to the fact that	22	31	-*	-
	it is clear that	21	30	-	-
	by the fact that	12	17	-	-
Uncertainty	*are more likely to*	29	41	-	-
	little is known about	21	30	-	-
	are likely to be	20	28	10	20
	is likely to be	19	27	-	-
	it is likely that	15	21	10	20
	were more likely to	15	21	-	-
	seems to be a	14	20	-	-
	is more likely to	11	16	-	-
	it seems that the	11	16	-	-
	does not seem to	10	14	-	-
	on the assumption that	-	-	33	65
	based on the assumption	-	-	26	51
	parameters seem to be	-	-	14	27
TOTAL	17	266	377	113	222

*This symbol (-) means that the exact same bundle was not produced in this corpus above the cut-off point chosen in our study.

Second, several stance bundles in our corpora were attested as frequent in academic prose in other studies (Biber et al., 2004) (e.g. *the fact that the*, *are more likely to*, *it is necessary to*, *it is possible to,* see Tables 4, 6, and 7). Interestingly, out of the 66 bundles extracted from both corpora, only eight were exactly the same in structure (e.g. *the fact that the*, *it is important to*, *of the most important*, *it is difficult to,* etc.), showing that stance is expressed in both corpora, but with a considerable preference for the use of different bundle structures. This leads us to observe that, to express specific meanings, like *certainty*, bundles carry basically the same words, thus varying very little: *the fact that the*, *to the fact that*, *by the fact that*. This result may be due to a methodological decision in the extraction of 4-word bundles only. If

we had considered 5-word bundles, we would have detected bundles such as *to the fact that the* and fewer occurrences of some of the bundles mentioned above. Despite the elimination of overlapping bundles, our results continue to be relevant and reveal that applied linguistics in general, including Biber, show certainty by using 4-word stance bundles which contain the word *fact*. The use of the head noun *fact* as part of the noun+that+complement structure is a recurring pattern that seems to reflect disciplinary conventions as well, since this stance expression is similarly found in texts written in the humanities, social sciences, engineering, among others (Jiang, 2017). In our research, for instance, Biber refers to examples that distinguish two modes of communication (speech and writing) by highlighting the differences to the reader in a clear way (see example 04):

4. These simple examples illustrate *the fact that the* two modes of communication have quite different strengths and weaknesses, and they therefore tend to be used in complementary situations (BC.BS.1988.01).

In example 05, a CorAAL author justifies his/her viewpoint by arguing that language varieties are available in some listening input, which stresses, once again, a transparent and confident position in applied linguistics texts:

5. This claim is based on *the fact that the* listening input in the test includes a range of native-speaker accents (North American, Australian, New Zealand, and British), and that all Standard varieties of English are accepted in evaluating candidates' responses to the test (CorAAL.EAP.Art.2015.003).

Furthermore, it can be seen that the epistemic expression of uncertainty is more favored than that of certainty in both corpora, resonating with Hyland and Jiang's (2019) finding as this has been a disciplinary convention shared by applied linguists over time. In CorAAL and BC, most of the bundles that express this function are formed with the adverb *likely* or with the verb *seem* (see Table 4). In example 06, Biber hypothesizes that spoken registers are similar to written registers:

6. However, to the extent that they have changed, *it is likely that* spoken registers have become more similar to written registers, resulting in an underassessment of the changes caused by the addition of written registers. (BC.BS.1995.01).

In example 07, the stance expression of uncertainty is observed in a statement in which a CorAAL author assumes that such factors as ages and contexts are considered unimportant by test-takers in a particular study:

revistadelta
42.4
2026

7. *It is likely that* the different ages and contexts of the different tests made some of these important items irrelevant for specific groups of test-takers. (CorAAL.TQ.Art.2014.006).

Overall, the variety of uncertainty bundles (11 types) in CorAAL is higher than in BC (5 types) (see Table 4). Out of these five types, three of them have a normalized frequency above 20 and carry a noun within (e.g. *on the assumption that*, *based on the assumption*, *parameters seem to be*). The noun *assumption*, even if conveying some uncertainty, does show that the author is laying out a claim that he accepts as true, as seen in a study in which Biber assumes that linguistic co-occurrence reflects underlying communicative functions (see example 08). As for the use of the noun *parameters*, it lends the bundle a tone of certainty as well, since *parameters* are interpreted as guidelines or limits of a certain proposition as, in the case of Biber's text, the author arranges two separate communicative *parameters* based on his research results (see example 09):

8. After the statistical analysis is completed, dimensions are interpreted functionally, *based on the assumption that* linguistic co-occurrence reflects underlying communicative functions. (BC.JA.2014.01).

9. Two separate communicative *parameters seem to be* represented here: the primary purpose of the writer/speaker (involved versus informational), and the production circumstances (those restricted by real-time constraints versus those enabling careful editing possibilities). (BC.BOC.2010.01).

As can be seen, the use of epistemic bundles is higher in CorAAL than in BC both in terms of expressing certainty and uncertainty. However, the most frequent uncertainty bundles in BC are formed with nouns (*assumptions* and *parameters*), expressing a certain level of confidence from Biber's part. Interestingly, the head noun *assumption* as part of noun+that+complement structure is more common in natural sciences and technical engineering texts (Jiang, 2017). This characteristic may be interpreted as the authors' effort to create a context of reliance which we will see more clearly orchestrated as we analyze the choice of attitudinal stance bundles (ability / possibility) next.

Attitudinal Stance Bundles

As previously mentioned, attitudinal stance expressions were the most frequent in our data compared to the epistemic ones, similarly resonating

with Hyland and Jiang's (2019) results. In CorAAL, evaluation bundles were often used by the authors to express effectiveness (*the effectiveness of the*), importance (*the importance of the*), reliability (*the reliability of the*), consistency (*is consistent with the*), and degrees of difficulty (*it is difficult to*) and understanding (*to better understand the*) (see Table 5). In BC, evaluation bundles mostly revolved around Biber's expression of importance (*the importance of register*), but others expressed similarity (*are similar in their*) and peculiarity (*that are peculiar to*) (see Table 5). Despite the fact that some bundles were similar in meaning in CorAAL and BC (e.g. *the importance of the* and *the importance of register,* respectively), only one evaluation bundle appeared in both corpora with the exact same structure (*it is difficult to*) (see examples 10 and 11). Clearly, as previously mentioned, Biber's preference for the stance bundle *the importance of register* is related to the author's research topic: *register variation.*

Table 5. *Evaluation stance bundles across corpora*

Evaluation bundles	Raw freq. CorAAL (tokens)	Norm freq. CorAAL	Raw freq. BC (tokens)	Norm freq. BC
the effectiveness of the	24	34	-	-
the importance of the	24	34	-	-
it is difficult to	22	31	18	35
an important role in	18	25	-	-
to better understand the	18	25	-	-
the reliability of the	14	20	-	-
is consistent with the	13	18	-	-
a better understanding of	13	18	-	-
of the importance of	12	17	-	-
are similar in their	-	-	19	37
the importance of register	-	-	19	37
similar to one another	-	-	18	35
similar to each other	-	-	13	25
sample is typical of	-	-	11	22
that are peculiar to	-	-	11	22
and a generally fragmented	-	-	10	20
TOTAL	158	222	119	233

revistadelta
42.4
2026

The expression of difficulty is common among applied linguists. For instance, in example 10, Biber recognizes the difficulty of determining what the object of study of register and genre is:

10. However, *it is difficult to* pin down specific differences in the object of study referred to by the two terms; both are used in reference to linguistic varieties associated with particular situations of use and particular communicative purposes. (BC.BS.2006.01).

Moreover, in example 11, a CorAAL author attests the difficulty of making assertions about the proficiency of some types of students:

11. *It is difficult to* draw strong conclusions about the English proficiency of redesignated students. (CorAAL.TQ.Art.2017.010).

Particularly in the case of BC, Biber favors the expression of evaluation more often than CorAAL because he sets comparisons between different registers (science fiction vs. general fiction; official documents vs. academic prose), reflecting stance choices aligned with the topic of his texts (see example 12):

12. In fact, some registers *are similar in their* central tendencies but differ markedly in their ranges of variation (science fiction versus general fiction, and official documents versus academic prose, where the first register of each pair has a more restricted range of variation). (BC.JA.1993.03).

In addition to evaluation bundles, the expression of ability / possibility is also an authorial characteristic across the applied linguists' writings, particularly Biber's (see Table 6):

Table 6. *Ability / possibility stance bundles across corpora*

Ability / possibility bundles	Raw freq. CorAAL	Norm freq. CorAAL	Raw freq. BC	Norm freq. BC
can be found in	22	31	-	-
it is possible that	30	42	-	-
it is possible to	18	25	49	96
can be attributed to	15	21	11	22
would be able to	15	21	-	-
as can be seen	15	21	-	-
can be seen in	11	16	-	-

can be seen as	11	16	-	-
could be attributed to	10	14	-	-
can be used as	10	14	-	-
can be used to	-	-	53	104
can be interpreted as	-	-	26	51
can be regarded as	-	-	17	33
can be considered as	-	-	15	29
can be used for	-	-	14	27
can also be used	-	-	13	25
can be interpreted in	-	-	10	20
TOTAL	157	221	208	407

There is also a considerable amount of stance bundles expressing ability / possibility in CorAAL and BC, used to frame or introduce some possible or actual action or proposition, particularly in passive constructions utilizing the modal verbs *can* and *could.* Overall, Biber is the applied linguist who mainly expresses ability / possibility, and this may be justified by the registers composing BC, as all the ability / possibility bundles, with no exception, recurred more often in books and book chapters than in journal articles. The remaining stance bundles expressed possibility through phrases containing *possible* and *be able to*, and only two bundles recur in both corpora: *can be attributed to* and *it is possible to.*

In order to express possibility in passive constructions, the authors introduce some justification or explanation about something stated in the following proposition. For instance, in example 13, Biber lays out two reasons for the historical developments involving the distributions of first and second-person pronouns across English and Somali and their structural and functional similarity. In example 14, a CorAAL author states that the lack of knowledge of field-related collocations and formulaic sequences justifies second language (L2) students' difficulty in acquiring discipline specific academic style:

13. These historical developments *can be attributed to* two influences: 1) an increasing need for written prose with dense informational content, associated with the 'informational explosion' of recent centuries, and 2) an increasing awareness among writers of the production possibilities of the written mode, permitting extreme manipulation of the text (BC. BOC.2010.02)

revistadelta
42.4
2026

14. The difficulty in attaining nativelike academic style *can be attributed to* L2 writers' unawareness of the collocations and formulaic sequences that are representative of their fields. (CorAAL.EAP.Art.2015.002)

Moreover, in examples 15 and 16, it is possible to recognize the authors' attitudes as they describe and argue about elements available in the discourse (*figures* in example 15 and *affordances* in example 16):

15. On the basis of Figures 3 and 4, *it is possible to* describe the distinguishing linguistic characteristics of each of the eight text types. (BC.JA.1989.01)

16. *It is possible to* argue that the available affordances for gamers to act and interact online provides a potential social communal pace for them to develop autonomy in L2 learning […] (CorAAL.LLT.Art.2014.003).

In the examples 13 to 16, the authors' attitudes in the expression of possibility reflect their effort to create a context of reliance built towards their propositions, similarly done in their expression of epistemic stance as previously analyzed in subsection *Epistemic Stance Bundles*.

Lastly, the remaining attitudinal / modality bundles were obligation / directive stance expressions (see Table 7).

Table 7. *Obligation / directive stance bundles across corpora*

Obligation / directive bundles (08)	Raw freq. CorAAL	Norm freq. CorAAL	Raw freq. BC	Norm freq. BC
it is important to	60	85	17	33
to ensure that the	18	25	-	-
to be aware of	15	21	-	-
of the most important	12	17	17	33
it is also important	11	16	-	-
is a need for	10	14	-	-
it is necessary to	-	-	19	37
must be based on	-	-	18	35
TOTAL	126	178	71	138

Obligation / directive stance bundles direct readers to carry out actions that the writers want them to complete and guide the audience to achieve

specific conclusions, like the bundle *it is important to,* a similar bundle shared both by CorAAL and BC, and which highlights that the attitudinal stance expression of *importance* has been favored in applied linguistics articles over time (Hyland & Jiang, 2019). As a matter of fact, "bundles consisting of *it* clausal fragments […] are usually used to act as expressions of stance" (Jalali, 2017, p. 32). Particularly in Biber's case, as Silva (2026) has shown, this use reflects the author's metadiscourse style. For instance, Biber draws the reader's attention to and outlines the importance of the quantitative way of identifying some language patterns (see example 17). In addition, a CorAAL author states the importance of comparing two concepts, learner autonomy and self-regulation, but stresses that they differ in a specific way (see example 18):

17. *It is important to* note that the co-occurrence patterns underlying dimensions are identified quantitatively (by a statistical factor analysis) and not on any a priori basis. (BC.JA.1995.01)

18. *It is important to* note, however, that autonomy encompasses control over a wider range of phenomena than self-regulation (CorAAL. TQ.Art.2014.005)

Among the obligation / directive stance bundles (see Table 7), the necessity bundles also called our attention. While in CorAAL only the noun *need* forms a bundle with a normalized frequency of 14, in BC the adjective *necessary* and the modal verb *must* form bundles which have a much higher normalized frequency, 37 and 35, respectively. These bundles interestingly occur in BC only, highlighting the author's attitudinal stance as part of his larger metadiscourse style (Silva, 2026). In example 19, a CorAAL author urges for the necessity of carrying out additional research, using a different structure bundle in a much less frequent way (normalized frequency of 14):

19. Among these four perspectives, there *is a need for* more research on game-enhanced L2 learning that occurs in the wild [...] (CorAAL.LLT. Art.2014.003)

In example 20, Biber acknowledges the need to have a complete description of the distributional patterns within each language prior to carrying out a cross-linguistic comparison of register variation. Later, in example 21, the author endorses the necessity of having a robust methodology before scholars can reach comprehensive conclusions:

revistadelta
42.4
2026

20. Before undertaking a cross-linguistic comparison of register variation, *it is necessary to* have a complete description of the distributional patterns within each language. (BC.BS.1995.01)

21. Textual variation across any sociolinguistic domain is complex and multi-faceted; hence comprehensive conclusions concerning variation within such domains *must be based on* a comprehensive methodology. (BC.JA.1986.02)

As has been presented in this section, the use of attitudinal over epistemic bundles is higher in both corpora. Next, we proceed with laying out the final remarks of this research article.

5. Final Remarks

In this paper, we have demonstrated how Biber and other applied linguists have crafted personal and evaluative discourses in their texts. Consequently, we have built evidence against the mistaken belief that academic writing is devoid of authors' personal forms of (self-)expression and is home to a dry, fact-oriented discourse that opts to set theoretical notions over express writers' views, opinions, beliefs, assessments, etc.

Several were the evaluative meanings conveyed by different stance bundles used by the applied linguists, including the expression of similarity, reasonability, importance, surprise, adequacy, peculiarity, effectiveness, importance, excitement, reliability, consistency, and degrees of difficulty and understanding.

As we have demonstrated through contextualized corpus-based analyses of stance-marking 4-word lexical bundles, we have recognized how academic writing also contains authors' voices as they take particular stances. Such registers as journal articles, books, and book chapters similarly allow for scholars' voice expression, helping shed light on their authorial personas and identities as they simultaneously establish relationships and bonds with their reading audience and their scientific communities.

Concerning the functional classification of our corpora stance bundles, most of them fit the attitudinal / modality category, especially in CorAAL authors' texts. Several stance bundles in our corpora were attested as frequent in academic prose in other studies. Only eight were exactly the same in

structure, showing that stance is expressed with a considerable preference for the use of different bundle structures.

Although there has been a decrease in the use of stance expression over the last 25 years, as Hyland and Jiang (2019) discovered, our corpora results demonstrate that the expression of stance continues to be widespread in our area, illustrating not only 'individual' options (in the case of Biber), but also larger disciplinary conventions shared by the applied linguistics community in which CorAAL authors (and Biber himself) participate.

Our data shows that the use of epistemic bundles is higher in CorAAL than in BC, both in terms of lexical variability and of expressing certainty and uncertainty, resonating with Hyland and Jiang's (2019) finding as this is a disciplinary convention shared among applied linguists. Nevertheless, the more frequent uncertainty bundles in BC are formed with nouns, and, when expressing certainty, both corpora shared stance bundles with the use of the head noun *fact* as part of the noun+that+complement structure, representing disciplinary conventions.

Furthermore, the higher use of attitudinal / modality bundles over epistemic across both corpora promptly reveals that CorAAL authors and Biber have built rather evaluative and personal texts by resorting to a variety of attitudinal stance bundles that clearly made attitudes and evaluations towards their propositions evident to the audience. This finding also demonstrates how these stance markers highlight the writers' alignment with the interests of the applied linguistics community. Not only that, the higher use of attitudinal over epistemic markers appears to reverberate with Hyland and Jiang (2019), in that attitude markers usage continues steadier than epistemic bundles.

Particularly in the case of BC, Biber favors the expression of evaluation more often than CorAAL because of the way he chooses to address the topics of his texts, thus endorsing his authorial style as was previously revealed in Silva (2026). Moreover, Biber is the applied linguist who mainly expresses ability / possibility, and this may be justified by the registers of the texts composing BC. Obligation / directive stance bundles were preferred by CorAAL authors, and they used them with the intent to direct readers to carry out actions that they wanted them to complete, serving as more engaging patterns to establish rapport with the audience. Also, the choice of *importance*-expressing bundles by both corpora highlights that this

revistadelta
42.4
2026

attitudinal stance expression has been commonplace in applied linguistics texts.

Finally, in addition to contributing to the understanding of authorial characteristics, this study could similarly address some pedagogical implications. First of all, it is important to raise students' awareness of authentic data obtained from corpus-based lists and concordances. Students may be presented with a functional classification of bundles so as to better understand their differences. Particularly in the case of stance bundles, we recognize the relevance of expressing authorial voice and presence throughout the texts, especially for novice writers who sometimes may feel coerced to behave otherwise. This may be done through repeated exposure and activities such as matching or item identification, derived from the academic registers students will need to write and read. In order to cater for disciplinary variation, it is similarly relevant to help students understand the different features and discourse patterns they will encounter in their particular courses. Comparative corpus analysis may also favor the understanding of sociocultural factors which affect writers' linguistic choices. Over time, constructing a small corpus of students' own writings may boost their self-esteem and sense of achievement and accountability. We thus expect that these suggestions equip novice (and expert) writers with contextually appropriate rhetorical strategies in the use of stance bundles and other stance expressions as they make their voice heard in the academic setting. At the same time, these practices will take us to see that academic writing, specifically in the area of applied linguistics, may not fully hold a dry, purely fact-oriented style as many have long believed.

Conflict of Interests

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Authors Contributions

Both authors were responsible for conceptualizing the study, designing methodology, data curation, formal analysis, and project administration, as well as writing the original manuscript. They both supervised the research process, contributed to the validation of the study, and were involved in the review and editing of the final manuscript.

Data availability

The data used in the research are not publicly available.

References

Anthony, L. (2022). *AntConc (Version 4.2.0) [Computer Software].* Waseda University.

Bailey, S. (2018). *Academic writing: A handbook for international students*. Routledge.

Biber D. (1988). *Variation across speech and writing*. Cambridge University Press.

Biber, D. (2006). *University Language: A corpus-based study of spoken and written registers.* John Benjamins Publishing Company.

Biber, D. (2009). A corpus-driven approach to formulaic language in English: Multiword patterns in speech and writing. *International Journal of Corpus Linguistics 14*(3), 275-311.

Biber, D., Conrad, S., & Cortes, V. (2004). If you look at... Lexical bundles in university teaching and textbooks. *Applied Linguistics 25(*3), 371-405.

Biber, D., & Gray, B. (2016). *Grammatical complexity in academic English*: *Linguistic change in writing*. Cambridge University Press.

Biber, D., & Conrad, S. (2019). *Register, genre, and style*. Cambridge University Press.

Biber, D., Johansson, S., Leech, G., Conrad, S., & Finegan, E. (2021). *Grammar of spoken and written English.* John Benjamins Publishing Company.

Chen, Y., & Baker, P. (2010). Lexical bundles in L1 and L2 academic writing. *Language Learning & Technology, 14*(2), 30-49.

Collins, L., Baker, P., & Brookes, G. (2024) Representations of obesity in Australian and UK news coverage: A diachronic comparison. *Applied Corpus Linguistics*, *4*(2), p. 1-10.

Cortes, V. (2013) The purpose of this study is to: Connecting lexical bundles and moves in research article introductions. *Journal of English for Academic Purposes, 12*(1), 33–43.

Dutra, D. P., & Berber Sardinha, T. (2013). Referential expressions in English learner argumentative writing. In S. Granger, G. Gilquin, & F. Meunier (Eds.), *Twenty years of learner corpus research: Looking back, moving ahead* (pp. 117-127). Presses universitaires de Louvain.

Dutra, D. P., Queiroz, J. M. S., Macedo, L. D. de, Costa, D. D., & Mattos, E. (2020). Adjectives as nominal premodifiers in chemistry and applied linguistics corpora. In U. Römer, V. Cortes, & E. Friginal (Eds.), *Advances in corpus-based research on academic writing effects of discipline, register, and writer expertise* (pp. 205-226). John Benjamins Publishing Company.

Frankenberg-Garcia, A. (2018). Investigating the collocations available to EAP writers. *Journal of English for Academic Purposes*, *35*, 93-10.

Gray, B. (2015). *Linguistic variation in research articles: When discipline tells only part of the story*. John Benjamins.

Grondona, C. B. (2015). *Eliminação de pacotes lexicais relacionados ao tópico e de pacotes lexicais em contexto de sobreposição: Uma proposta metodológica para os estudos da Linguística de Corpus* [Unpublished master's thesis, Universidade Federal de Minas Gerais].

Halliday, M. A. K. (1994). *Functions of language.* Edward Arnold.

Hyland, K. (1998). Boosting, hedging and the negotiation of academic knowledge. *Text 18*(3), 349-382.

Hyland, K. (1999) Academic attribution: citation and the construction of disciplinary knowledge. *Applied Linguistics, 20*(3), 341–367.

Hyland, K. (2000). *Disciplinary Discourses*: *Social interactions in academic writing.* Longman.

Hyland, K. (2002a). Options of identity in academic writing. *ELT Journal, 56*(4), 351-358.

Hyland, K. (2002b). Authority and invisibility: Authorial identity in academic writing. *Journal of Pragmatics*, *34*(8), 1091-1112.

Hyland, K. (2005). *Metadiscourse: Exploring interaction in writing*. Continuum International Publishing Group.

Hyland, K. (2008a). As can be seen: Lexical bundles and disciplinary variation. *English for Specific Purposes*, *27*(1), 4-21.

Hyland, K. (2008b). Small bits of textual material: A discourse analysis of Swales' writing. *English for Specific Purposes*, *27*(2), 143-160.

Hyland, K. (2010). Community and individuality: Performing identity in applied linguistics. *Written Communication, 27*(2), 159-188.

Hyland, K. (2012). Bundles in academic discourse. *Annual Review of Applied Linguistics*, *32*, 150-169.

Hyland, K., & Jiang, F. K. (2019). *Academic discourse and global publishing*: *Disciplinary persuasion in changing times.* Routledge.

Ioannidis, J. P. A. (2023). *October 2023 data-update for "Updated science-wide author databases of standardized citation indicators"* (Version 6) [Data set]. Elsevier Data Repository. https://elsevier.digitalcommonsdata.com/datasets/btchxktzyw/6

Ivanič, R. (1998). *Writing and identity*: *The discoursal construction of identity in academic writing.* John Benjamins Publishing.

Jalali, H. (2017). Reflection of stance through *it* bundles in applied linguistics. *Ampersand*, *4*, 30-39.

Jiang, F. K. (2017). Stance and voice in academic writing: the "noun + that" construction and disciplinary variation. *International Journal of Corpus Linguistics, 22*(1), 85-106.

Pang, W. (2010). Lexical bundles and the construction of an academic voice: A pedagogical perspective. *Asian EFL Journal*, *47*, 1-13.

Poole, R., Gnann, A., & Hahn-Powell, G. (2019). Epistemic stance and the construction of knowledge in science writing: A diachronic corpus study. *Journal of English for Academic Purposes*, *42*, 1-11.

Prina Dutra, D., & Berber Sardinha, T. (2021, March 20–23). *A multi-dimensional typology of English research article sections* [Paper presentation]. American Association for Applied Linguistics Conference (AAAL), Online.

Scott, M., & Tribble, C. (2006). Concordances: The immediate context. In M. Scott, & C. Tribble. *Textual patterns*: *Key words and corpus analysis in language education* (pp. 33-54). John Benjamins Publishing Company.

Silva, J. M. V. (2025). Being interactive in applied linguistics research articles: Towards characterizing an academic writer's metadiscourse style. *Revista Brasileira De Linguística Aplicada*, *25*(1), 1-29.

Silva, J. M. V. (2026). "Strikingly", an academic writer's style can be interactional: Towards characterizing metadiscourse style in applied linguistics research articles. (2026). *Revista De Estudos Da Linguagem*, *34*(1), 393–421.

Simpson-Vlach, R., & Ellis, N. C. (2010). An academic formulas list: New methods in phraseology research. *Applied Linguistics, 31*(4), 1-26.

Sinclair, J. (2004). *Trust the text*: *Language corpus and discourse*. Routledge.

Tardy, C. (2012). Current conceptions of voice. In K. Hyland, & C. Guinda (Eds.), *Stance and voice in written academic genres* (pp. 34-47). Palgrave Macmillan.

Tognini-Bonelli, E. (2001). *Corpus linguistics at work.* John Benjamins Publishing Company.

Recebido em: 11.08.2025
Aprovado em: 11.08.2026

EDITORA GERAL: Marcia Veirano Pinto