v. 34, n. 1, p. 393–421, 2026
ISSN: 2237-2083 · DOI: 10.17851/2237-2083.34.1.393-421
Submetido em: 09/12/2024 · Aprovado em: 06/11/2025

REVISTA DE ESTUDOS DA LINGUAGEM

"Strikingly", an Academic Writer's Style Can Be Interactional: Towards Characterizing Metadiscourse Style in Applied Linguistics Research Articles

"Surpreendentemente", o estilo de um autor acadêmico pode ser interacional: rumo à caracterização do estilo metadiscursivo em artigos de pesquisa da linguística aplicada

Janailton Mick Vitor da Silva
Universidade Federal da Paraíba (UFPB)
João Pessoa | PB | BR
janailtonm@gmail.com
https://orcid.org/0000-0002-5137-5473

Abstract: This paper aims at unveiling the interactional metadiscourse style of a prominent applied linguist, Douglas Biber. Comparisons are made between 25 of his research articles (RAs) to 126 RAs written by other applied linguists. In this work, a Corpus Linguistics (CL) methodological framework is proposed by using AntConc tools. The results show that the author's interactional stylistic choices convey levels of certainty and uncertainty about his propositions; make his views explicit, comment on the message, and express evaluations (importance and suitability) and emotions (surprise); and directly involve readers in the ongoing discourse by overtly pointing them to other parts in the RAs. Strikingly, Biber's interactional rhetorical practices are essentially evaluative and engaging, expressing solidarity, anticipating objections, and revealing the extent to which he works to jointly construct texts with his readers. As a result, the author's tendency to pursue an interactional metadiscourse style reveals how an expert academic writer crafts an authorial style that may empower novice and other expert writers to craft a style that involves the audience in meaningful reading and writing experiences through academic texts.

Keywords: Style; metadiscourse; interaction; corpus linguistics; linguistic features.

Resumo: Este artigo objetiva desvendar o estilo metadiscursivo interativo de um importante linguista aplicado, Douglas Biber. São feitas comparações entre 25 de seus artigos de pesquisa e 126 artigos escritos por outros linguistas aplicados. Neste trabalho, uma proposta metodológica baseada na Linguística de Corpus (LC) é seguida usando ferramentas do AntConc. Os resultados desta pesquisa mostram que as escolhas estilísticas interacionais do autor transmitem níveis de certeza e incerteza sobre suas proposições; tornam suas visões explícitas, comentam a mensagem e expressam avaliações (importância e adequação) e emoções (surpresa); e envolvem diretamente os/as leitores/as no discurso em andamento, direcionando-os explicitamente para outras partes dos artigos. Surpreendentemente, as práticas retóricas interacionais de Biber são essencialmente avaliativas e envolventes, expressando solidariedade, antecipando objeções e revelando até que ponto o autor trabalha para construir conjuntamente textos com seus/suas leitores/as. Como consequência, a tendência do pesquisador em buscar um estilo metadiscursivo interacional revela como um escritor acadêmico experiente cria um estilo autoral que pode empoderar novos e experientes autores/as a criar um estilo que envolva o público em experiências significativas de leitura e escrita por meio de textos acadêmicos.

Palavras-chave: Estilo; metadiscurso; interação; linguística de *corpus*; recursos linguísticos.

1 Introduction

In our everyday life, we frequently associate *style* with someone's preference for particular songs, dishes, clothes, etc., but we may not promptly associate it with someone's choice of particular words or phrases in a given language. For example, the way an individual uses an adjective as predicative (*That house is beautiful*) and not attributive (*That is a beautiful house*) may characterize a personal linguistic preference. This can also be seen when an academic writer shows predilection for *moreover* over *furthermore*, or when the prepositional phrase *for instance* is preferred over *for example*. In sum, these and other frequent patterns of language use may subsidize the identification of an individual's language style. As proposed in this paper, a writer leaves their fingerprint in texts, revealing a style that may be promptly associated with them.

Implicit in the word *way* mentioned in the previous paragraph lies the notion of *style*. In this research, *style* is analyzed within the academic writing realm (Bacon, 2013), owing to the fact that most style studies have mainly focused on literary texts and literary authors (Jeffries; McIntyre, 2010). *Style* is hereby understood as a writer's fingerprint or personal form of expression (Baker, 2000; Garcia, 2006), marked by a preference for some linguistic items and structures across different texts and compared to other writers' language choices (Malmkjaer; Carter, 2002). *Style* is also defined as an individual's personality in linguistic terms and an individual's voice in the collectiveness, whereby authorial identity(ies) may be similarly revealed (Hyland, 2002a; 2002b; 2008b; 2010; Ivanič, 1998; Tardy, 2012).

The relevance of carrying out this research can be verified from different perspectives. Drawing on Leech and Short (2007), style reveals how language is used in a given context, by a given person, for a given purpose. Since no one writes in a vacuum, detached from a socio-cultural context and without pre-determined purposes and audiences, academic language is hereby examined insofar as it can create interpersonal bonds with readers, precisely through interactional linguistic choices. According to Hyland (2005), writers can be interactional when they comment on their message, make their views, evaluations, and opinions explicit to readers, and jointly construct the texts with them.

These interpersonal bonds materialized in texts have been investigated from different angles in literature (Hyland; Tse, 2004). For the sake of specificity, however, metadiscourse has been selected for this investigation, because it can refer both to scholar's patterns of discourse organization and writer's ways of expressing stance and engagement (Hyland, 2004). In this regard, metadiscourse is understood as part of author's style in academic writing, since scholars may present other types of styles, even potentially non-metadiscoursal, when writing in and/or out of academia. As can be seen, this definition adds to long-standing discussions which have surrounded the concept of style: "the elusiveness of style lies in the age-old dispute as to whether style is extrinsic or intrinsic in nature; whether it is the icing on the cake – in effect, an optional extra – or whether it is an inherent part of the cake itself." (Burke, 2014, p. 24).

Furthermore, additional reasons for carrying out this investigation stem from Crystal's (2013) attempt to define authors' stylistic profile. Once one's style is defined, pedagogical materials highlighting important stylistic features to be used in training courses, for instance, might be developed. Moreover, by doing this research, Bacon's (2013) proposal to challenge a classical dichotomy in the field of style in academic writing, particularly in the context of teaching style to college students, is supported. For the author, this dichotomy establishes that a scholar should either write with style or write for the academy, as if writers could not write stylistically within academia. Finally, the myth of a dry, impersonal discourse in academic writing is similarly refuted in this research (Hyland, 2002a).

Considering the reasons stated in this introductory section, it is important to understand how renowned applied linguists craft their interactional metadiscourse style and, by extension, express stance and establish engagement with readers. Stemming from Silva's (2024) PhD dissertation, the main goal of this paper is to unveil the *interactional* metadiscourse style of an applied linguist, Douglas Biber, based on frequent and consistent use of metadiscourse language features across research articles (RAs) and in comparison to other RAs. Biber has been chosen as protagonist in this style investigation because of his distinguishing contributions to corpus linguistics (CL) research in a variety of topics, especially in register studies. In this regard, this investigation is guided by the following research question: how does Biber express stance and establish engagement with readers in his RAs through interactional metadiscourse elements?

To answer this question and accomplish the main goal of this paper, a CL methodological framework has been proposed by using *AntConc*, version 4.2.0, designed by Lawrence Anthony (2022), specifically through *Key-Word-In-Context (KWIC)*, *Cluster, Collocate*, and *Keyword* tools, both for compilation and contextual analysis of *Biber Corpus* (BC) and *Corpus of Journal Articles in Applied Linguistics (CorJAAL)*. These steps will be described and discussed in detail in the Methodology section. The linguistic data generated by the program has been submitted to corpus-based quantitative and qualitative linguistic analyses following the application of the *Stylistic Metadiscourse Model of Interpersonal Interaction in Academic Writing*, which will be introduced in the upcoming sections.

2 Theoretical background

The elusiveness inherent to the meaning of style has often been attributed to its polysemic condition (Burke, 2014). Nevertheless, many scholars have contributed to defining it. For example, Malmkjaer and Carter (2002, p. 510) see style as "[...] a consistent occurrence in the [spoken and written] text of certain items and structures, or types of items and structures, among those offered by the language as a whole." Furthermore, style is characterized as an individual's or group's "[...] personal form of expression [...]"[1] (Garcia, 2006, p. 123)[2], or "[...] the definition of a personality in linguistic terms"[3] (Câmara Junior, 1978, p. 13). In other words, writers select items and structures among many others similarly available in the language, resulting in the expression of their personality in linguistic terms.

Scholars' self-expression may also be imprinted in written texts. As a matter of fact, Hyland (2005, p. 53) asserts that "all writing carries information about the writer [...]", and that "writers cannot avoid projecting an impression of themselves and how they stand in relation to their arguments, their community and their readers." From these quotes, it can be inferred that writing is a means by which writers can express authorial identities through different language resources, including personal pronouns, adjectives, lexical bundles, adverbs, among others (Hyland, 2002a; 2002b; 2008b; 2010).

Because all writing carries information about the writer, it can thus be acknowledged that "[...] all writing contains 'voice' [...]" (or style, as is proposed here) (Hyland, 2008b, p. 144). Although individuals have a voice of their own, everything they say or write is constrained by a social context and its elements, such as genre, register, audience, etc. (Tardy, 2012). As a result, a writer's voice stems from *a mingling of voices* and discourses available in society, which resonates with Bakhtin's (1981) *heteroglossia* notion: "language is not a neutral medium that passes freely and easily into the private property of the speaker's intentions; it is populated – overpopulated – with the intentions of others" (Bakhtin, 1981, p. 294). In spite of this multivoicedness feature of language, Tardy (2012) asserts that the way a particular writer mingles other individuals' voices and discourses in a text is what characterizes their individual voice.

Style has traditionally been studied within the area of *stylistics*, which can be broadly "[...] defined as the (linguistic) study of style" (Leech; Short, 2007, p. 28), or the "[...] scientific

[1] "*[...] forma pessoal de expressão [...]*" (Garcia, 2006, p. 123).

[2] All translations provided here have been made by the author of this paper, unless otherwise stated.

[3] "*[...] a definição de uma personalidade em termos lingüísticos*" (Câmara Junior, 1978, p. 13).

study of patterns of variation in written language. The object of study is the language of a single individual (idiolect), resulting in a description of his or her identifying linguistic characteristics." (McMenamin, 2002, p. 176). Particularly in this paper, the focus is on the application of CL tools to help compile and analyze large portions of texts, including patterns of linguistic phenomena, which had earlier (in traditional stylistics) been unfeasible to undertake given the techniques available (McIntyre; Walker, 2019).

Taking these concepts presented so far into consideration, for the purposes of this investigation, style is defined within the context of academic writing as "[…] the particular way a scholar varies the use of language in academic genres and registers for specific purposes, resulting in the expression of voice(s), authorial identity(ies), stance(s), and potential effects in texts and audiences." (Silva, 2024, p. 23). This definition of style encompasses some underlying elements which will be discussed next.

Firstly, when style is defined as the particular way an individual makes variation in the use of language in academia, it can be argued that scholars (synonymously regarded as authors or writers in this paper), show consistent preference for particular language items and/or structures over others, representing distinctive individual variation in the register used within the academic context in academic genres.

Moreover, this individual variation may contribute to the expression of academic writers' voice(s), authorial identity(ies), and stance(s) in academic registers and genres, as they inevitably leave marks and traces on what they write. This expression, nevertheless, does not represent sole and individual voice(s), identity(ies), and stance(s) per se, because, as socially engaged subjects, scholars take part in discourse practices enacted in and out of academia, which is a social milieu consisting of specific principles and traditions.

Furthermore, authors' personal form of expression in academia is not arbitrary, especially because it needs to acquiesce to academic and linguistic norms and conventions that regulate academic discourse and represent linguistic disciplinary variation (Hyland, 2009). As a result, language used by academics is imprinted in a myriad of text varieties, such as RAs, which reveals how style is also construed in and derived from the social discourse communities to which scholars, readers, and other agents belong in academia. For example, RAs undergo several publishing steps, and their content is affected by different literacy brokers, such as editors, translators, peer reviewers, academic peers, and friends (Lillis; Curry, 2010).

Finally, authors' style is adopted with the intention of reaching particular purposes and causing potential effects in texts and audiences. Hence, in addition to expressing research-related facts, information, data, viewpoints, among other meanings, they may also end up creating interpersonal bonds with audiences by means of *interactive* and *interactional* metadiscourse features. For instance, previous research has shed light on the use of these features to shape authorial style in RAs (Silva, 2024, 2025).

In view of these concepts, a theoretical model can be applied to the analysis of academic writer's texts and their language style. This model, entitled *Stylistic Metadiscourse Model of Interpersonal Interaction in Academic Writing* (Silva, 2024, 2025), is a critical, blended up-do-date version of Hyland's (2005) *interpersonal model of metadiscourse*, Hyland's (2008a) *taxonomy for bundles functional analysis*, Hyland's (2009) *key resources of academic interaction*, and Biber's (2006) and Biber *et al*.'s (2021) *lexico-grammatical and semantic features used for stance analyses*. But prior to introducing the model, it is paramount to understand some key concepts: metadiscourse, stance, and formulaic language.

Similar to *style*, *metadiscourse* is also a fuzzy term. To start with, Crismore (1983, p. 2) defined *metadiscourse* as "the author's intrusion into the discourse, either explicitly or non--explicitly, to direct rather than inform, showing readers how to understand what is said and meant in the primary discourse and how to 'take' the author". Similarly, *metadiscourse* is deemed as some linguistic material which adds no propositional information to the text, but rather points to the author's presence therein (Vande Kopple,1985). In more recent studies, *metadiscourse* is characterized as aspects of a particular text, such as language features "[...] which explicitly refer to the organization of the discourse or the writer's stance towards either its content or the reader" (Hyland, 2004, p. 109). However, Hyland also warns that metadiscourse has mistakenly been comprehended as either "discourse about discourse" or "talk about talk" (Hyland, 2005, p. 16), limiting its definition to textual over interpersonal aspects.

It is particularly on this interpersonal perspective that this investigation relies. Therefore, it is argued that the nature of metadiscourse is essentially dialogic, because "[...] all metadiscourse is interpersonal in that it takes account of the reader's knowledge, textual experiences and processing needs and that it provides writers with an armoury of rhetorical appeals to achieve this" (Hyland; Tse, 2004, p. 161). It involves two main types of interaction, *interactive* and *interactional*, whereby Hyland (2005) pays homage to the works produced by Thompson and Thetela (1995) and Thompson (2001):

> *The interactive dimension*. This concerns the writer's awareness of a participating audience and the ways he or she seeks to accommodate its probable knowledge, interests, rhetorical expectations and processing abilities. The writer's purpose here is to shape and constrain a text to meet the needs of particular readers, setting out arguments so that they [readers] will recover the writer's preferred interpretations and goals. The use of resources in this category therefore addresses ways of organizing discourse, rather than experience, and reveals the extent to which the text is constructed with the readers' needs in mind.
> *The interactional dimension*. This concerns the ways writers conduct interaction by intruding and commenting on their message. The writer's goal here is to make his or her views explicit and to involve readers by allowing them to respond to the unfolding text. This is the writer's expression of a textual 'voice', or community-recognized personality, and includes the ways he or she conveys judgements and overtly aligns him- or herself with readers. Metadiscourse here is essentially evaluative and engaging, expressing solidarity, anticipating objections and responding to an imagined dialogue with others. It reveals the extent to which the writer works to jointly construct the text with readers. (Hyland, 2005, p. 50-51, emphasis in original).

In other words, the *interactive* dimension regards the extent to which writers have empathy with their readers by organizing the discourse so that they can fulfill the demands of the audience and make sure readers retrieve their objectives. The *interactional* dimension aims at writers' direct, explicit intrusions along the discourse, making their epistemic and attitudinal stances clear to the audience. This is a more dialogic connection between reader and writer, engaging the audience in the discussion and inviting it to co-construct the texts.

In Hyland's (2005) work, both of these dimensions present particular corresponding functional categories (interactive = *transitions, frame markers, endophoric markers, evidentials, code glosses*; interactional = *hedges, boosters, attitude markers, self-mentions, engagement markers*). In a more recent book, Hyland (2009) revisits his earlier model and splits the interactional

dimension into two, by making the differences between stance from engagement more explicit. Stance now encompasses all the four previous elements (*hedges, boosters markers, attitude mention,* and *self*) from the interactional dimension, whereas engagement consists of the following subcategories: *reader, directives, questions reference, shared knowledge,* and *asides*.

In addition to *metadiscourse,* another term has been used to refer to writer's presence in texts: *stance*. In literature, *stance* figures as one among many other language features used to express writers' personal feelings, attitudes, assessments, and value judgements, having been labeled differently over time: "'evaluation'; 'intensity'; 'affect'; 'evidentiality'; 'hedging'; 'modality' and 'stance'" (Biber, 2006, p. 87). Despite this varied terminology, the expression of stance is important in academia: "in claiming a right to be heard, and to have their work taken seriously, writers must simultaneously take a stance towards what they discuss and display competence as disciplinary insiders." (Hyland; Jiang, 2019, p. 131).

Furthermore, another key term in this paper is *formulaic language*, which includes "[...] chunks, collocations, fixed expressions, formulaic language, idioms, lexical bundles, multiword units, prefabricated units, set phrases" (Frankenberg-Garcia, 2018, p. 93). Nevertheless, although these terms are interconnected, the term *lexical bundles* is hereby adopted as part of the large superordinate *formulaic language*, because bundles are considered "[...] sequences of word forms that commonly go together in natural discourse" (Biber *et al*., 2021, p. 982).

Considering the theoretical notions laid out in this section, Table 01 summarizes the *Stylistic Metadiscourse Model of Interpersonal Interaction in Academic Writing* (Silva, 2024, 2025), which was elaborated based on a critical analysis of Hyland (2005; 2008a; 2009), Biber (2006), and Biber *et al*. (2021):

Table 01–Stylistic metadiscourse model of interpersonal interaction in academic writing (Silva, 2024, 2025)

General features of interaction	Categories		Subcategories
Interactive (helps guide readers through the text)	Discourse organizers (research-oriented and text-oriented)		Transitions
			Frame markers
			Evidentials
			Code glosses
Interactional (involves readers in the text)	Participant-oriented systems of interaction	Stance markers	Hedges (epistemic)
			Boosters (epistemic)
			Attitude markers (attitudinal)
			Self-mentions (epistemic)
		Engagement markers	Reader pronouns
			Personal asides
			Endophoric markers and directives
			Questions
			Appeals to shared knowledge

Source: elaborated by the author.

Based on the model introduced above, in the interactive dimension, *transitions* express additive, causative, and contrastive relations between main clauses (*moreover; furthermore; and*); *frame markers* refer to research/discourse acts, sequences, stages, or topic shifts (*first; then; a wide range of*); *evidentials* point to information from other texts, reporting on but not taking a stance towards other scholars' research and findings (*argue; assert; according to; claim*); and *code glosses* supply additional information by rephrasing, explaining, or elaborating on what has been stated (*for instance; for example*) (Hyland, 2005; 2008a; 2009; Biber, 2006a; Biber *et al.*, 2021).

In addition to these interactive elements, the model also presents interactional (sub) categories. To begin with, in the stance category, *hedges* withhold writer's commitment to a proposition, allowing information to be presented as an opinion rather than a fact and opening that position to negotiation (*believe; doubt; may; might*); *boosters* emphasize writer's certainty and confident voice, strengthening an argument but closing down possible alternatives to a proposition (*determine; discover; find; must; should*); a*ttitude markers* express writer's affective attitudes, emotions, evaluations to/on a proposition, conveying surprise, agreement, importance, obligation, frustration, etc. (*amazed; amazing; astonishingly*); and *self--mentions* make explicit reference to the writer (*me; we; I*). As for engagement markers, *reader pronouns* overtly build relationship with reader (*your; you;* inclusive *we*); *personal asides* interrupt the argument to offer a comment on what has been stated (through appositive phrases or clauses in between dashes or parentheses); *endophoric and directive markers* direct readers to another part of the text or to another text, inform them how to carry out some action in the real world, and instruct them how to interpret an argument (*as shown in figure; as noted above; as can be seen*); *questions* invite engagement, encourage curiosity, and bring interlocutors into the discourse arena (through interrogative and/or rhetorical questions); and *appeals to shared knowledge* ask readers to recognize something as familiar or accepted, identify themselves with particular views, and assign to them a role in creating the argument (*of course; by the way; obviously*) (Hyland, 2005; 2008a; 2009; Biber, 2006a; Biber *et al.*, 2021).

It should be noted that this model is quite comprehensive. However, considering the goal and research question laid out for this paper, only the features found in the research corpus and classified as *interactional* will be investigated, including *boosters, hedges, attitude markers*, *endophoric* and *directive markers*. These features, however, will only be introduced and analyzed in section 4. Before that, methodological procedures adopted for the current investigation are introduced next.

3 Methodology

In this paper, CL has been used as an approach (Tognini-Bonelli, 2001) to carry out stylistic analyses by comparing study and reference corpora. The methodological steps followed to carry out this investigation are presented and discussed in this section.

3.1 The study corpus

The first corpus compiled in this research was *BC*, the study corpus. *BC* consists of 25 single-authored RAs, mostly quantitative empirical (Gray, 2015), written by Douglas Biber and published in a myriad of peer-reviewed journals. These RAs span a time period of 35 years and were published from 1984 to 2019. Table 02 displays an overview of the texts, their time range, and the number of types and tokens obtained from *AntConc* (Anthony, 2022):

Table 02 –Overview of the study corpus

Biber Corpus - BC	
Sub-register/Sub-corpus	Journals RAs (JA)
Time range	1984-2019
Number of texts	25
Types	5,263
Tokens	144,719

Source: elaborated by the author

A note should be made regarding the time range informed on the table, for Biber published more than 25 texts throughout his career, including books, book chapters, reviews, monographs, and proceeding papers, especially in collaboration with other scholars, as listed on his personal website CV[4]. Nevertheless, in view of the style focus of this investigation, only 25 single-authored journal RAs were selected for the study corpus compilation. Because this collaborative practice is part of academic practice, Biber (2013, p. 375) has stated the following on his collaborative work and authorship attribution:

> I like collaborating with a coauthor and working as part of a collaborative team. […] But I somehow have ended up doing as much collaborative research as individual research. […] The first author listed on my books and articles is the person who takes the lead in the analysis, and who actually does most of the writing. But all authors are usually fully involved in the analysis process and help with commenting and revising the paper.

Four are the major reasons for compiling only journal RAs in the present research. First, Biber's greatest number of single-authored texts are, in fact, RAs. Hence, similar to Hyland's (2008b; 2010) and McIntyre and Walker's (2019) sample corpora, BC was considered an appropriate sample to investigate Biber's style in comparison to his contemporaries'. Second, including only journal RAs in the current sample and excluding other academic genres (i.e. books or book chapters) would not pose a problem, especially because such genres share particular features in comparison to other genres (Biber; Gray, 2016) and are often produced in similar situations (Biber; Conrad, 2019). Third, a great amount of metadiscourse research focuses on RAs (Pearson; Abdollahzadeh, 2023), so it seemed feasible to carry on this research trend.

[4] BIBER, D. Home. Arizona: Northern Arizona University, [201-?]. Available at: https://dougbiber.weebly.com/. Accessed on: 11 Nov. 2025.

Last, it is much easier to track Biber's prominence as a scholar (and potentially as an academic writer) on the basis of journal RAs citation metrics.

3.2 The reference corpus

A reference corpus was compiled, *CorJAAL*, consisting of texts published by three peer-reviewed journals: *Applied Linguistics*, *System*, and *TESOL Quarterly*. In style research, it is important to have a comparison parameter to illustrate what is peculiar to a writer's language repertoire and is probably not so to other academic writers'. For that, a methodological procedure suggested by Baker (2000) concerns the comparison of more than one author's text to another writer's set of works. A reference corpus works as a norm to which the study corpus will be compared, but as a relative norm, as "[...] an absolute norm for English cannot be relied on [...]" (Leech; Short, 2007, p. 41). Therefore, since Biber writes RAs in this area, the purpose is to investigate how he varies language within this expected relative norm as far as the use of interactional metadiscourse elements is concerned.

Despite the importance of having a reference corpus for style investigation, "the issue of reference corpus selection is far from decided" (Scott; Tribble, 2006, p. 64). Many scholars have struggled to establish the best layout for a reference corpus, without reaching much consensus. However, some researchers agree that size, genre, register, content, and research goals are some of the issues that should be borne in mind. In this regard, the reference corpus of this investigation considers various elements, including size (five times larger than the study corpus), register (academic writing), sub-register and genre (empirical and theoretical RAs), authors, journals, themes, topics, and fields within applied linguistics.

Regarding size, *CorJAAL* is five times larger than *BC*, following Berber Sardinha's (2000) study. Table 03 presents an overview of the reference corpus and its time range, as well as the number of types and tokens obtained from *AntConc:*

Table 03 –Overview of the reference corpus

Corpus of Journal Articles in Applied Linguistics - CorJAAL	
Sub-register/Sub-corpus	Journal RAs (JA)
Time range	1984-2019
Number of texts	126
Types	21,311
Tokens	819,612

Source: elaborated by the author

Concerning genre and register, selecting only RAs published in journals represents a way to control variables that could otherwise impact the results. In addition, as to Geluso and Hirch (2019), when a reference corpus contains a similar register[5] to the study corpus,

[5] The authors are referring to the following academic sub-register: academic writing in applied linguistics. In their research, they made different comparisons. To reach the conclusion that keywords characterize the study corpus and not the register, Geluso and Hirch (2019) compared their study corpora (made of papers published by faculty members from two universities in the United States, including book chapters and journal publications) to a reference corpus of the same sub-register: research articles in the applied linguistics discipline.

the results from a keyword search are very particular to the study corpus and not to the register. Put differently, the words identified characterize the study corpus (and potentially the author's style, as will be shown) and not the register. However, this is not to say that register may not influence someone's individual style, since research like Grieve's (2023) has shown evidence on this issue.

Moreover, *CorJAAL* contains empirical (qualitative and quantitative) and theoretical RAs (Gray, 2015). As to empirical RAs, there is a balanced number of both qualitative and quantitative texts published in *Applied Linguistics*, but more qualitative RAs in *System* and *TESOL Quarterly*. The journals and RAs which compose *CorJAAL* were selected based on the following criteria:

- Criteria for selecting *CorJAAL* journals:
  - Be among top 20 with the highest impact factor (IF)[6] throughout 2022 and across a 5-year period as published in Clarivate's (2023) Journal Citation Reports;
  - Publish RAs on a variety of topics, themes, and fields within applied linguistics.
- Criteria for selecting RAs from these journals:
  - Have been published over a 35-year period, the same time span of the RAs in *BC* (from 1984 to 2019);
  - Show the highest number of citations in the volumes and issues published in the same time span as the *BC* articles.

Furthermore, a plethora of authors and their texts compose *CorJAAL*, which serve a consistent diversity of authorship against which Biber's texts have been compared.

3.3 The compilation of the study and reference corpora

Regarding corpus compilation, all study and reference corpora texts have been downloaded from the web. The *.pdf* files were converted into *.txt* documents (UTF-8) using free online pro-

[6] IF has figured as the dominant bibliometric for evaluating academic output for decades, and this was the major reason why it was chosen to select journals for the ongoing research. However, it raises many problems, especially when it is wrongly used by institutions to assess articles' quality and prevent scholars from getting jobs, tenure, or grants. In general, a common critique of IF is that "[…] citations are a shallow measure of research quality or impact" (Lillis; Curry, 2010, p. 9). In addition, Lillis and Curry (2010) notice that journals with higher IF receive more submissions; authors make immoderate self-citations and deliberate citations of other scholars publishing in the same journal; journals may pressure academics to cite papers from the same journal and limit scholars' citations to rival journals; university library subscriptions favor journals with higher IF. Furthermore, Waltman and Traag (2020) uphold that, under certain conditions, the number of citations an article receives is a more accurate indicator of the value of the article than IF. Hyland (2022, p. 18) also notes that "[…] a few highly cited papers can dramatically increase the impact of average or less significant papers."

grams, such as *I love Pdf*[7], *Ocr 2 Edit*[8], and *Online 2 PDF*[9]. By using basic automated *Microsoft Word* tools, the texts were edited to correct typos and other conversion-resulting errors, such as *distribu-tion → distribution*, so as not to compromise files processing on *AntConc*. Later, pieces of text from the original papers were withheld in brackets (< >), including the following cases: in-text short and long citations (only in the case of *BC*); titles of paper sections and subsections (e.g., <Conclusion>); information related to Biber and other writers (e.g., <Douglas Biber, Northern Arizona University, Flagstaff>). Because only the IMRD sections from *BC* and *CorJAAL* were used for the corpora compilation, such post-textual elements as authors' bios, notes, and acknowledgements sections, as well as images, tables, graphs, and lists of bibliographical works cited in the references section were deleted. For all these cases, the tag <DELETED EXCERPT> was inserted in place. Moreover, text-related identification tags were added at the beginning of each file, as displayed in Table 04:

Table 04– Example of text identification tags used in the corpora files

BC	CorJAAL
<text id = BC.JA.1993.04> <title = > <author = >	<text id = CorJAAL.AL.1993.04> <title = > <author = >
BC = Biber Corpus	CorJAAL = Corpus of Journal Articles in Applied Linguistics
JA = Journal RAs (subcorpus)	SYS = abbreviation for the journal System (AL = Applied linguistics; TQ = TESOL Quarterly)
Year of publication = 1993	Year of publication = 1993
Order of texts per year[10] = 04	Order of texts per year = 04
Title = the title of each research article (RA) is written	Title = the title of each RA is written
Author = the name of each author is given	Author = the name of each author is given

Source: elaborated by the author

At last, all the edited *.docx* files were saved into *.txt* (UTF-8) format. All these procedures were similarly followed in the reference corpus compilation.

3.4 Selecting keywords for analysis

The main *AntConc* tool used to identify interactional stylistic words was the *Keyword tool*. According to Scott (2022), the definition of keyness in CL is given below:

> A word is said to be 'key' if
> a) it occurs in the text at least as many times as the user has specified as a minimum frequency

[7] I LOVE PDF. [S. l.: s. n], 2025. Available at: https://www.ilovepdf.com/ Accessed on: 11 Nov. 2025.

[8] OCR2EDIT. *OCR2Edit - Extract text from scans, images and more*. [S. l.: s. n], 2025. Available at: https://www.ocr2edit.com/ Accessed on: 11 Nov. 2025.

[9] ONLINE 2 PDF. *Conversor PDF online.* [S. l.: s. n], 2007. Available at: https://online2pdf.com/ Accessed on: 11 Nov. 2025.

[10] It was necessary to add this information in the tag as more than one paper was published in the same year.

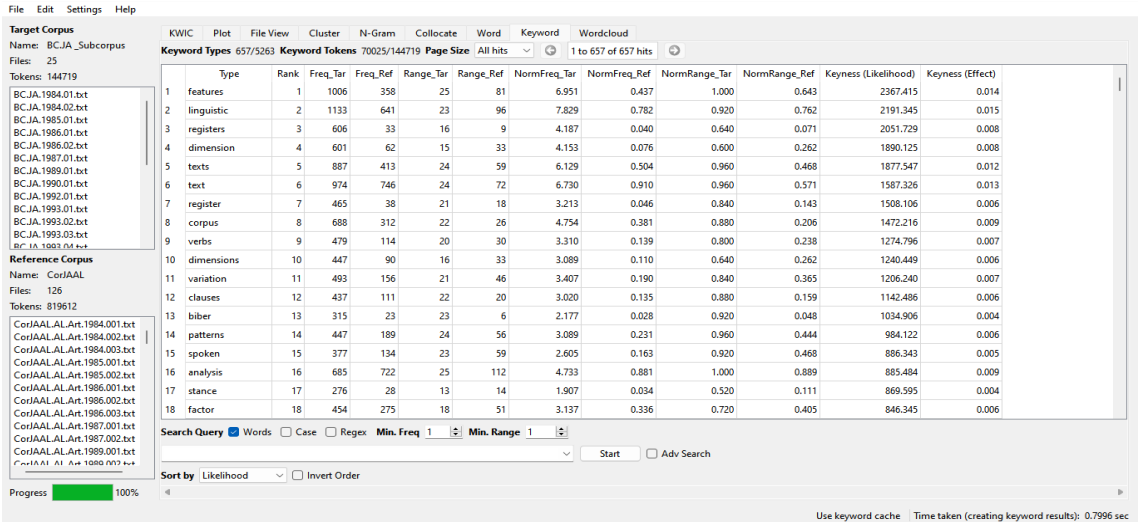
> b) its frequency in the text when compared with its frequency in a reference corpus is such that the statistical probability as computed by an appropriate procedure is smaller than or equal to a p value specified by the user
> c) in addition, the strength of keyness must be at least as great as the minimum log ratio set by the user. (Scott, 2022).

To obtain such keywords, the program compares two wordlists automatically generated by *AntConc Word* tool (in this investigation: each wordlist from a respective subcorpus of *BC*, which is the study corpus – or target corpus, as named by Anthony – and the second one from *CorJAAL*, or reference corpus). The results displayed by *Keyword Tool* are classified based on their statistical significance and labeled as either positive or negative:

> a word which is *positively* key occurs *more often* than would be expected by chance in comparison with the reference corpus.
> A word which is *negatively* key occurs *less often* than would be expected by chance in comparison with the reference corpus. (Scott, 2022, emphasis added).

For the purposes of the keyword analysis, *BC* was compared to *CorJAAL* following the default Keyness Values available in the program. In terms of data normalization, the value 1,000 words was set in the *Statistics global settings*, although this option may be altered at any time to suit the user's preference. This calculus is exemplified in Biber (2006a): the number of types is divided by the number of tokens and the result is multiplied by one thousand (1000/normed rate). For instance, take the case of the keyword *registers:* it occurs 606 times across *BC* (which has 144,719 tokens): (606 / 144,719) * 1000 = 4.187. Thus, the normed rate of the keyword *registers* in *BC* is 4.187 per one thousand words. After setting these values, both *BC* and *CorJAAL* were loaded in the program by using the *Corpus Manager* function. *BC* was used as the target corpus and *CorJAAL* as the reference. Upon selecting both corpora, the button *Start* in the *Keyword tool* was pressed, and a list of 657 keywords was shown on screen, as seen in Figure 01:

Figure 01– Screenshot of *AntConc* keyword tool



	Type	Rank	Freq_Tar	Freq_Ref	Range_Tar	Range_Ref	NormFreq_Tar	NormFreq_Ref	NormRange_Tar	NormRange_Ref	Keyness (Likelihood)	Keyness (Effect)
1	features	1	1006	358	25	81	6.951	0.437	1.000	0.643	2367.415	0.014
2	linguistic	2	1133	641	23	96	7.829	0.782	0.920	0.762	2191.345	0.015
3	registers	3	606	33	16	9	4.187	0.040	0.640	0.071	2051.729	0.008
4	dimension	4	601	62	15	33	4.153	0.076	0.600	0.262	1890.125	0.008
5	texts	5	887	413	24	59	6.129	0.504	0.960	0.468	1877.547	0.012
6	text	6	974	746	24	72	6.730	0.910	0.960	0.571	1587.326	0.013
7	register	7	465	38	21	18	3.213	0.046	0.840	0.143	1508.106	0.006
8	corpus	8	688	312	22	26	4.754	0.381	0.880	0.206	1472.216	0.009
9	verbs	9	479	114	20	30	3.310	0.139	0.800	0.238	1274.796	0.007
10	dimensions	10	447	90	16	33	3.089	0.110	0.640	0.262	1240.449	0.006
11	variation	11	493	156	21	46	3.407	0.190	0.840	0.365	1206.240	0.007
12	clauses	12	437	111	22	20	3.020	0.135	0.880	0.159	1142.486	0.006
13	biber	13	315	23	23	6	2.177	0.028	0.920	0.048	1034.906	0.004
14	patterns	14	447	189	24	56	3.089	0.231	0.960	0.444	984.122	0.006
15	spoken	15	377	134	23	59	2.605	0.163	0.920	0.468	886.343	0.005
16	analysis	16	685	722	25	112	4.733	0.881	1.000	0.889	885.484	0.009
17	stance	17	276	28	13	14	1.907	0.034	0.520	0.111	869.595	0.004
18	factor	18	454	275	18	51	3.137	0.336	0.720	0.405	846.345	0.006

Source: elaborated by the author

Traditionally, researchers have limited themselves to analyzing only "top 10, 50 or 100 keywords" (Brezina, 2018, p. 85). Although dealing with long lists of keywords is time-consuming and may lead to errors, selecting some keywords over others may overlook important features of the corpus. In this regard, "a crucial next step when dealing with lists of key items [...] is working out which ones, if any, help most to answer your research questions." (McIntyre; Walker, 2019, p. 283). Therefore, the keywords were grouped into three major categories: text-aboutness (keywords that regarded the topic/theme of the RAs); non-text-aboutness metadiscourse (interactional and interactive metadiscourse keywords); non-metadiscourse and non-text-aboutness (keywords that are neither text-aboutness nor metadiscourse). Despite these three major categories, only the non-text-aboutness metadiscourse keywords underwent contextual analysis in this paper. It should be mentioned that this entire selection included keywords that were not only in the top 100, but were also above this classical threshold (Brezina, 2018). For this reason, the issue of sorting by likelihood or effect size keyness values did not bear much importance, as all the keywords were eventually categorized, including those which had shown varying likelihood and effect size values (this is also the reason why normalized and percentage values are used for discussion in the next section). Since the main goal here was to focus on interactional keywords of the study corpus, *BC* keywords that fit this metadiscourse dimension were identified and used for the subsequent contextual analysis, as will be discussed in the next section.

During this selection process, some issue was raised and needed to be attended to. It was observed that some keywords collocated with other words (some even key), forming lexical bundles, illustrating Sinclair's (2004) co-occurrence principle and that isolated tokens may be little informative (Berber Sardinha, 2004). To better locate these bundles, the *Cluster tool* was used and Biber *et al.*'s (2021, p. 985) cut-off points adopted. To be considered a bundle in this investigation, the expression should appear in at least five (05) texts (minimum range) and recur at least 10 times (minimum frequency). Another issue was raised even when adopting this cut-off point as the program returned many bundles: they needed to be classified as metadiscoursal according to the stylistic model presented in section 02, otherwise they would be deleted from the analysis. In sum, single and compound (bundles) interactional keywords were selected for analysis, as will be discussed next.

4 Results

The number of types and tokens of key interactional language items categorized according to the *Stylistic Metadiscourse Model of Interpersonal Interaction in Academic Writing* is presented in Table 05:

Table 05– Number of stylistic metadiscourse interactional types and tokens across *BC*

Interactional features	types	raw tokens	normed tokens
boosters (stance)	43	2,267	15.66
attitude markers (stance)	8	730	5.04
endophoric and directive markers (engagement)	12	226	1.56
hedges (stance)	1	45	0.31
total	**64**	**3,268**	**22.58**

Source: elaborated by the author

To express stance, Biber uses *boosters*, *attitude markers*, and *hedges*. To establish engagement with readers, he opts for *endophoric markers and directives*. Overall, these participant-oriented subcategories were realized by key language items through adverbs, adjectives, verbs, and bundles. In order to answer the research question posed in the introduction, this section will show how Biber expresses stance and establishes engagement with readers in his RAs and, as a consequence, favors an interactional metadiscourse style. More specifically, it will be explored how Biber's individual stylistic choices convey levels of certainty (through *boosters*) and uncertainty (through *hedges*) about his propositions; make his views explicit, comment on the message, and express evaluations and emotions through *attitude markers*; and directly involve readers in the ongoing discourse by overtly pointing them to other parts in the texts through *endophoric markers and directives*. As will be seen, these rhetorical practices are "[...] essentially evaluative and engaging, expressing solidarity, anticipating objections [...], reveal[ing] the extent to which [Biber] works to jointly construct the text with readers". (Hyland, 2005, p. 51).

It should be mentioned, however, that Biber's style could have been more interactional if other metadiscourse features had been identified as positively key in *BC*. In other words, the corpus was bereft of statistically significant results of *self-mentions*, *reader pronouns*, *personal asides*, *questions*, and *appeals to shared knowledge*. These are more self-expressing and engaging features, particularly two of them, which will be commented on in the next paragraph. This finding equally illustrates the recent trend of texts in applied linguistics concerning the decrease of engagement features, as well as "[...] an increase in more quantitative studies which restrict opportunities for overt stance-taking." (Hyland; Jiang, 2019, p. 134). Biber's texts fit this quantitative paradigm and thus prove Hyland and Jiang's (2019) assumption.

Self-mentions are by far the most powerful means of self-representation in academic writing, revealing "[...] the degree of explicit author presence in the text measured by the frequency of first-person pronouns and possessive adjectives [...] (Hyland, 2005, p. 53). This decision is not merely by chance, because "the presence or absence of explicit author reference is generally a conscious choice by writers to adopt a particular stance and a contextually situated authorial identity." (Hyland, 2005, p. 53). Another feature absent in *BC*, *reader pronouns*, "[...] explicitly address readers, either to focus their attention or include them as discourse participants", serving to "[...] highlight or downplay the presence of readers in the text." (Hyland, 2005, p. 53). The style guidelines of *BC* journals publishing companies posed no restriction for the use of *self-mentions* or *reader pronouns*, suggesting Biber's decision to avoid using them altogether or using them less than expected.

Despite the absence of these features, Biber's texts show significant use of other interactional elements, which will be explained as of the next subsection. For that, this section presents and discusses a few examples (at least one) from each participant-oriented subcategory as used in their respective context of occurrence across *BC*.

4.1 Boosters

Biber opts for verbs and a single adverb to stress his confidence in attesting claims in texts. Biber's *boosters* stylistic choices recur in over 50% of all 25 texts and in almost all three time slots, revealing how he consistently opts for these markers across his academic production and time. *Boosters*, like *hedges*, convey epistemic modality, but there is a plausible difference between them, especially in terms of either opening (*hedges*) or closing (*boosters*) space for readers to reach specific conclusions while interpreting texts. Despite the fact that *boosters* tend to limit the variability of viewpoints on the part of readers, they still end up joining together writers and readers, as both agents are part of the same discourse community and, to a certain extent, agree with one another:

> Hedges imply, then, that a statement is based on plausible reasoning rather than certain knowledge, and allow readers the freedom to dispute it. At first glance, boosters seem to contradict such conciliatory and defensive tactics. They emphasize the force of propositions and display commitment to statements, thereby asserting the writer's conviction and restricting the negotiating space available to the reader. But while an apparently risky tactic, boosters nevertheless allow writers to strategically engage with colleagues, effecting interpersonal solidarity and membership of a disciplinary in-group. (Hyland, 1998b, p. 353).

According to Silva (2024), booster verbs state the writers' stance towards their views and claims; booster verbs are endophoric in nature, pointing readers to some internal aspects of the authors' research and building solidarity with the audience. These verbs were identified in BC , and their semantic classification and normed frequencies (all occurring 15,55 times per 1,000 words) following Biber *et al*. (2021) are displayed in Table 06:

Table 06 –Semantic classification of booster verbs and their normed frequencies in *BC*

Activity	Communication	Mental	Existence	Occurrence	Causative
3.18	0.49	2.96	6.67	1.71	0.55
analyze; analyzed; analyzing; carried; combines; investigate; show; shows; shown	describe; describing	compares; compute; considered; considering; identify; identified; interpreted;	characterize; characterized; comprises; differ; mark; illustrate; illustrates; illustrated; include; included; presents; marks; reflect; reflects; reflecting; represents; representing; represent; represented; uncover	occur; occurring; resulting	required

Source: elaborated by the author

As can be seen in Table 06, Biber favors existence verbs, followed by activity, mental, occurrence, causative, and communication verbs. Biber's stance expression for existence verbs serves to report states that exist between entities, representing the state of existence or a particular relationship between entities. In example 1 in *BC*, Biber explains what a comparison corpus is used for within stylistic studies. By opting for the verb *represent*, he expresses certainty about the state of *representing* that this type of corpus is submitted to because of the way it functions within a stylistic investigation. Moreover, the author inhibits the negotiating space available to readers by claiming that this corpus purpose is to *represent*, but nothing else, illustrating Biber's choice of one form out of the array of all available forms:

(1) The comparison corpus is used to *represent* 'typical' patterns of use, making it possible to empirically identify distinctive linguistic patterns in the target corpus that depart from those typical patterns. (Biber, 2011, p. 16)

In the conclusion section of another text, Biber revisits and reiterates what has been his purpose throughout the RA: to show that corpora of different sizes can be useful (example 2). For that, he chooses the activity verb *show* to denote the action of *showing* while ascertaining the usefulness of such corpora. Furthermore, he prevents readers from identifying other research goals other than the one already laid out. At the same time, however, this limitation offers readers interpersonal solidarity, preventing them from going astray:

(2) Rather, my purpose has been to *show* that existing corpora are adequate in many respects [...] (Biber, 1990, p. 269)

Elsewhere in *BC* (example 3), while Biber admits the systematicity and importance of lexico-grammatical association patterns, he nonetheless recognizes some limitation in interpreting these findings only this way, making explicit to the audience that additional reflections must be made. For that, he employs the mental verb *interpreted* in a passive voice structure to denote the interpreting action that must be executed by someone, either himself as researcher/writer or another scholar investigating the same subject. Consequently, he makes an apparent effort to engage with the disciplinary community and so displays a supportive authorial identity. At the same time, Biber restricts his readership to dispute other possibilities for this case besides the urgent need to interpret it:

(3) Although patterns such as these must be *interpreted* much more fully, the present section has illustrated the systematicity and importance of these association patterns in describing the use of related grammatical features. (Biber, 1996, p. 185)

Biber's certainty and tendency to restrict disputing alternatives to his propositions can also be recovered through the occurrence verb *occur*. In example 4, the author reports on his research findings concerning fillers that complete the lexical frame *the * of the*. These results are even shown in an illustration followed by the statement, proving that Biber's claim is accurate:

(4) In most cases, the high frequency patterns in academic writing are extremely productive. For example, there are numerous fillers that *occur* in the highest frequency frame the * of the, and as Table 4 shows, the most common of these fillers accounts for only 2% of the total occurrences of the pattern. (Biber, 2009, p. 298)

At the end of another publication (example 5), Biber urges for the need for additional research on use factors and relations between functions and association patterns. To make this demand, he opts for the causative verb *required*, especially in passive voice structures, also indicating that someone else (or perhaps himself) should carry on the research he started. This choice signals the author's certainty for his claim, displaying a confident authorial identity, as can also be seen by the attitude marker *obviously* at the beginning of the sentence. Although it is not obvious who is *required* to undertake additional research, this absence may also be seen as an attempt to engage with the disciplinary community by inviting it to follow on Biber's investigation:

(5) Obviously, future research is *required* to investigate the relative importance of use factors and the ways in which particular functional considerations relate to particular kinds of association patterns. (Biber, 1996, p. 193)

Another way of ascertaining Biber's claims is through the communication verb *describe*. In example 6, the author lays out three features of a text-linguistic approach to linguistic variation, and limits how his readership may see such an approach. To address issues about the first feature, he admits that the research goal of this approach entails *describing* the characteristics of texts, and not *identifying* or *illustrating* them, actions which belong to other verb semantic categories:

(6) The text-linguistic approach to linguistic variation differs in all three respects: the research goal is to *describe* the characteristics of texts, rather than the characteristics of a linguistic feature; each text constitutes an observation; and the quantitative findings represent the rates of occurrence of linguistic features in texts rather than the proportional preference for a linguistic variant in comparison to other variants. (Biber, 2012, p. 33)

Biber also uses a single adverb to stress his confident, assertive authorial identity in *BC*. In example 7, in order to highlight two important issues not considered by his analysis, he opts for the booster adverb *reliably* to refer to particular texts that could truthfully represent mean scores calculated on the basis of six linguistic features compared across 10-text samples:

(7) Secondly, this analysis addressed how many texts were needed to *reliably* represent mean scores, but did not address the representation of linguistic diversity in registers. (Biber, 1993a, p. 253)

Given this overall analysis on the use of *boosters* in *BC*, Biber's individual boosting choices disclose an interactional metadiscourse style whereby the author stresses the strength of his propositions, displays commitment and declares his conviction to statements, even if this entails, on the one hand, restricting the negotiating space available to the participating audience and, on the other hand, offering it interpersonal solidarity occasionally. These choices are also responsive to the rhetorical conventions of the genre and illustrate how the author expresses stance in *BC* to reveal confident, assertive, and supportive authorial identities. Next in subsection 4.2, the focus will be given to the feature the author uses to convey uncertainty about his propositions, *hedges*, which is the counterpart of *boosters*.

4.2 Hedges

Biber opts for only one adverb to show "a lack of complete commitment to the truth value of an accompanying proposition [...] or a desire to express that commitment categorically" (Hyland, 1998a, p. 1). This stylistic choice recurs in almost all *BC* texts and in all three time slots, proving the author's consistent decision for this marker across his academic production and time. As a mechanism of author's self-effacing, epistemic modality, hedging is realized by only one key adverb in *BC: nearly.* This result resonates with Biber *et al*.'s (2021) claim that *hedges* are considered a sub-class of adverbs and that, like conversation, academic prose shows a heavy reliance on single adverbs as stance devices, especially those indicating epistemic stance. Although only one single hedging adverb was found key, this result still represents Biber *et al*.'s (2021) finding that authors in academic prose tend to favor this grammatical device for the expression of epistemic modality, like Biber himself across his RAs.

Hyland (1998a) advocates for the scientific investigation of *hedges* in academia because they are essential to building the rhetoric of academic discourse, especially when it comes to writer's providing interpretative statements on paper's theory, method, or data:

> Hedging enables writers to express a perspective on their statements, to present unproven claims with caution and to enter into a dialogue with their audiences. It is therefore a substantial means by which scientists confirm their membership of the scientific community [...]. (Hyland, 1998a, p. 6).

In *BC*, the only key hedge is *nearly*, often collocating with such words as *all*, *identical*, and *every*, as well as with numbers and percentage. These patterns demonstrate that "hedges are also very common with numbers, measurements, and quantities. These forms are also called *approximators*, and typically function as modifiers of numerical or other quantifying expressions." (Biber *et al*., 2021, p. 554, emphasis in original). Biber's choice of *nearly* over other available forms in the language also depicts an individual stylistic preference.

For instance, in example 8, Biber is unsure as to whether specialized dimensions arise from all MD analyses, and in doing so enters into a dialogue with his audience to dispute this claim. In the meantime, he reveals hesitant and dialogic authorial identities:

(8) And to some extent, this expectation is met, with specialized dimensions emerging in *nearly* all MD analyses. (Biber, 2019, p. 52)

In example 9, Biber is cautious to determine that every discourse domain of English and other languages and cultures favors the presence of a narrative dimension, allowing his readership to refute whether this claim holds true:

(9) Similar to the oral/literate dimensions, a narrative dimension emerges in the MD analysis of *nearly* every discourse domain of English, and in *nearly* every language/culture. (Biber, 2014, p. 25)

Moreover, when reporting on numerical results from two studies, Biber hedges his commitment to, respectively, providing absolute values for model improvement (example 10) and

two syntactic functions of prepositional phrases (example 11). It can be seen that the author opens possibilities for his readers to notice how approximate (and not absolute) the results are:

(10) It represents a significant improvement over all other models and a relatively strong improvement in Delta over other models – *nearly* a 5% improvement over the 4-D model, and *nearly* a 10% improvement over the 1-D model. (Biber, 1992, p. 147)

(11) For editorials, this table shows that there is *nearly* a 50/50 split for prepositional phrases attached as nominal versus verbal modifiers [...] (Biber, 1993b, p. 226)

Given this overall analysis, Biber's individual hedging choice of *nearly* uncovers an interactional metadiscourse style by which he refrains from committing himself to the truth values of his propositions. At the same time, however, he allows readers to respond to the unfolding text, opening up possibilities for (re)assessing the claims and propositions. This stylistic choice also reveals the author's hesitant and dialogic authorial identities. Next in subsection 4.3, the focus will be allocated to the last feature the author uses with a stance-marking purpose: *attitude markers*.

4.3 Attitude markers

Biber opts for adjectives and adverbs to overtly comment on the message and express evaluations and emotions in the meantime. This stylistic choice recurs in over 60% of *BC* texts and in almost all three time slots, proving the author's consistent decision for these markers across his academic production and time. Furthermore, there is one keyword that works as metadiscoursal both on its own and as part of a lexical bundle: *particular*.

Differently from *hedges* and *boosters*, which are epistemic in nature and convey degrees of (un)certainty and (un)likelihood, *attitude markers* are attitudinal. This is to say that they are used to express writers' affective attitudes, emotions, evaluations, and feelings to/on a proposition, conveying surprise, agreement, importance, obligation, frustration, etc. (Biber *et al.*, 2021; Biber, 2006; Hyland, 2005). As a result, this self-expression helps construct an interactional metadiscourse style in texts. Compared to frequently used epistemic stylistic choices, Biber's underuse of *attitude markers* appears to reflect academic prose register: "overall, attitudinal stance markers are less common than epistemic markers [...]." (Biber *et al.*, 2021, p. 966).

It is perceivable that Biber favors the expression of evaluations and emotions while overtly commenting on the messages, echoing his voice plainly to the readers. In the former case, he opts for the adjectives *important, particular, primary*, the adverb *especially*, and the bundle *in particular* to convey importance; and the adjective *adequate* to transmit a sense of suitability. In the latter, he chooses the adjective *surprising* and the adverb *strikingly* to express his surprise towards claims.

To begin with, Biber's expression of importance unveils his evaluative authorial identity. As seen in example 12, the author opts for the adjective *important* to highlight that textual co-occurrence patterns between frequent nouns, adjectives, and prepositional phrases are relevant for the identification of linguistic characteristics of registers and styles. It is thus possible to detect the way Biber comments on his message and leaves his fingerprint for readers to spot, also representing an individual stylistic choice:

(12) For example, frequent nouns, adjectives, and prepositional phrases commonly co-occur in academic prose texts, working together to provide a dense integration of information. Textual co-occurrence patterns such as these are *important* in identifying the salient linguistic characteristics of registers and styles. (Biber, 1996, p. 175)

The expression of importance is also recognized by the adverb *especially*. In some cases, this adverb is even followed by the stance adjective *important*. In other cases, though, it collocates with other stance adjectives (*noteworthy, robust, conservative*) that did not return key by *AntConc*, but which can disclose Biber's evaluative authorial identity. In example 13, the author recognizes that, out of all dimensions of MD analysis often identified in many different studies, two bear especial importance because they help build basic human communication:

(13) Two such dimensions are *especially important*, regardless of the discourse domain: a dimension associated with oral vs. literate discourse, and a dimension associated with narrative discourse. The robustness of narrative dimensions across languages and discourse domains indicates that this rhetorical mode is basic to human communication, whether in speech or in writing. (Biber, 2014, p. 31)

Another evaluative stance adjective used to transmit a sense of suitability in *BC* is *adequate*. In example 14, Biber refutes some criticisms and leads readers to grasp that existing computer-based corpora are suitable for many studies of linguistic variation:

(14) However, the findings reported here argue against the claim that existing computer--based corpora are unsuited for the study of linguistic variation; instead, they indicate that, with respect to the above questions, the corpora are *adequate* for many studies of linguistic variation across texts. (Biber, 1990, p. 258)

The expression of emotion in *BC* is observed through the stance *attitude markers surprising* and *strikingly*, disclosing Biber's emotional authorial identity. In example 15, Biber is fully surprised with some research results because they may have been influenced by the layout of the corpus used. Again, it is possible to see how Biber comments on his message and emits his voice for readers to hear:

(15) It is perhaps not *surprising* that Dimension 1 in the original 1988 MD analysis was strongly associated with the oral/literate opposition, given that the corpus in that study ranged from spoken conversational texts to written expository texts. (Biber, 2014, p. 20-21)

In another use, the adjective *surprising* is modified by the adverb *especially* presented earlier. Below, it can be perceived how Biber's evaluative and emotional authorial identities are revealed in the sentence, with particular stress on the former. As can be seen in example 16, a common assumption about syntactic complexity held by many linguists is weakened when research revealed unexpected findings. Once again, Biber's interactional choice evidently discloses how he feels about the results found:

(16) The oral style relies on clauses to construct discourse – including a dense use of dependent clauses. In contrast, the complexity of the literate style is phrasal. This finding, replicated across languages and across discourse domains, is *especially surpri-*

sing, because it runs counter to assumptions about syntactic complexity held by many linguists. (Biber, 2019, p. 53)

Furthermore, emotion is also conveyed through the adverb *strikingly*. As seen in example 17, Biber's stance choice of this word reveals his surprise towards particular lexico-grammatical patterns detected in some analyses:

(17) These analyses show that *strikingly* different lexico-grammatical patterns are associated with each complement clause type and with each register, and that those associations can be explained in terms of the typical topics and communicative purposes of each register. (Biber, 1999, p. 136)

Given this overall analysis on the use of stance-marking *attitude markers* in *BC*, Biber's individual stylistic attitudinal choices unveil an interactional metadiscourse style whereby the author overtly comments on the messages, taking the participating audience to recognize his evaluative and emotional authorial identities. Lastly, in subsection 4.4, the focus will be allocated to the only feature the author uses with a more overtly engaging purpose: *endophoric and directive markers*.

4.4 Endophoric and directive markers

Biber opts for bundles to explicitly direct readers to particular points in his RAs. This stylistic choice recurs in over 50% of *BC* texts and in almost all three time slots, proving the author's consistent decision for these markers across his academic production and time. In the first assessment of the keyword list, however, *figure, shows, table, sample, illustrates, presents, above, section, shown*, and *summarizes* were identified as key, but when their collocates were searched for, it was observed that they formed bundles and could also function as metadiscoursal in BC.

Prior to illustrating some examples of *endophoric and directive markers* from *BC*, it is important to remember that, *endophoric markers* were originally classified as an interactive feature by Hyland (2005), but in a revisited version of the model, the author excluded it altogether and created *directives* (Hyland, 2009). However, in this paper, it is hereby understood that there is a connection between *endophoric markers* (originally interactive) and *directives* (originally interactional), for they have the same purpose: to show perceptible assessments of the material and the audience, direct readers to particular points in the text, and influence them to reach specific understandings and conclusions (Hyland, 2005). For this reason, in this paper, *endophoric markers* are a type of *interactional engagement markers*, together with *directives*, for both involve "[...] rhetorically positioning the audience, pulling readers into the discourse at critical points, predicting possible objections and guiding them to particular interpretations." (Hyland, 2005, p. 54). These features are also concerned with genre conventions (Swales, 1990). In sum:

> *Directives* instruct the reader to perform an action or to see things in a way determined by the writer. [...] they are signalled mainly by imperatives (like *consider, note,* and *imagine*), modals of obligation addressed to the reader (e.g. *must, should, ought*) [...] Directives are the only engagement feature which occur more frequently in papers in the sciences and engineering. (Hyland, 2005, p. 154).

> *Endophorics* […] refer the reader to explanatory or related material elsewhere in the text. They represent the writer's assessment of both the material and the audience, relating the propositions to the reader's assumed ability to process, and accept, the ongoing argument. In the soft fields this largely means reinforcing an argument by ensuring the reader has immediate access to relevant data or arguments located elsewhere in the text […] Endophorics, however, are overwhelmingly a feature of writing in the hard disciplines, where the referent is usually a nearby table or graph. (Hyland, 2005, p. 156-157).

In *BC*, Biber explicitly directs readers to other parts of the texts through textual acts by means of different bundles, such as *see table* and *see section*. It should be noted that the journal publishing companies' style guidelines encouraged authors to explicitly make cross-references along the texts, with two publishers suggesting the use of mental verbs like *see.* In this regard, Biber's choices seem to have undergone style guidelines influences, suggesting that he may have ended up abiding by academic conventions administered by literacy brokers and gatekeepers, including peer reviewers. Despite that, an interesting pattern across *BC* is that all uses of *see table* and *see section* come in parentheses and mostly at the end of sentences, implying that these directives behave similarly to non-integral citations, receiving minor focus and attention from the author. For instance, through textual acts, Biber discloses a supportive authorial identity by directing readers to *see:* a) a list of texts used to compare British and American writing, classified according to genre (example 18); and b) a section where it has been discussed how some dimensions characterized English and some discourse domains (example 19):

(18) Each dimension will be discussed briefly with respect to both written and spoken genres, but the focus here is on the 796 British and American written texts taken from the LOB and Brown corpora, representing the same nine written genres in each dialect (*see table* 3). (Biber, 1987, p. 109)

(19) Subsequent MD studies have shown that some of these dimensions turn out to be peculiar to English or to particular discourse domains (*see Section* 5). (Biber, 2014, p. 15)

Moreover, the keyword *shows* often collocates with *figure* and *table*. In example 20, it is possible to notice that Biber motivates readers to look at an illustration in the text, in which a marked decline with respect to the use of modals in different registers over a 5-decade period is depicted. The author's choice regards a textual act to be undertaken by readers:

(20) The study here provides additional evidence that modals have undergone a recent decline in use. As *Figure* 4 *shows*, modals have undergone a marked decline in all registers, especially over the last 50 years. (Biber, 2004, p. 117)

As can be seen, the use of *endophoric and directive markers* demonstrates an involvement between different semiotic resources available in the writer's texts to which he makes explicit reference. This connection can also be identified in example 21, proving that "[…] writing is a constant switching between written and diagrammatic representation because scientific concepts are typically semiotic hybrids which combine verbal, mathematical and visual elements." (Hyland, 2005, p. 157). The way Biber refers to data in tables illustrates this hybrid combination of elements. Again, readers are encouraged to achieve a textual act by

looking at a specific table in the article which presents numerical data on mean score and standard deviation:

(21) *Table 4 presents* the mean score and standard deviation of seven linguistic features in the pilot corpus, together with the computed tolerable error for each feature. It can be seen that the tolerable error ranges from 9.03 for nouns (which have a mean of 180.5) to 0.13 for conditional clauses (which have a mean of only 2.5) (Biber, 1993a, p. 253)

A further textual act is fostered in another excerpt in *BC*. In example 22, Biber leads readers to another part of the text, although not certainly where, encouraging them to note that text-linguistic register studies have now unveiled previously undiscovered language patterns. This choice also signals an individual stylistic decision:

(22) Second, *as noted above*, text-linguistic register studies have also been important because they have uncovered patterns of use that were not even anticipated in previous research. (Biber, 2019, p. 49)

As seen in this subsection, Biber's preference for some *endophoric and directive markers* demonstrates how he engages with readers by leading them to particular text extracts, some even made of hybrid semiotic resources (*figures, tables)*. The author's stylistic engaging choices disclose a supportive authorial identity and unveil an interactional metadiscourse style by which he explicitly bonds with the participating audience so that it can promptly recognize relevant information referred to in the RAs. In some cases, though, these choices may have been influenced by literacy brokers. Final Remarks are presented next as a way to summarize the findings in this research and illustrate how Biber has constructed a metadiscourse style through interactional features.

5 Conclusion

This research has attempted to unveil the interactional metadiscourse style of Biber by pursuing the following research question: how does Biber express stance and establish engagement with readers in his RAs through interactional metadiscourse elements? In view of the analyses carried out in this paper, it has been detected that Biber has used *boosters, attitude markers,* and *hedges* to express stance, and *endophoric markers and directives* to establish engagement with readers. Consequently, these patterns result in the creation of an interactional metadiscourse style through an individual preference for particular adverbs, adjectives, verbs, and bundles.

As seen across *BC*, the author's individual stylistic choices across texts and time convey levels of certainty (through *boosters*) and uncertainty (through *hedges*) about his propositions; make his views explicit, comment on the message, and express evaluations (importance and suitability) and emotions (surprise) through *attitude markers*; and directly involve readers in the ongoing discourse by overtly pointing them to other parts in the RAs through *endophoric markers and directives*.

More specifically, Biber's use of *boosters* has stressed the strength of his propositions and displayed his commitment and conviction to statements, revealing confident, assertive, and supportive authorial identities. Moreover, the author's *hedge* choice has demonstrated bivalent decisions on his part: he refrains from committing himself to some claims, but also allows the audience to respond to the unfolding text, opening up possibilities for (re)assessing the propositions. Together, these choices have disclosed the author's hesitant and dialogic authorial identities. Furthermore, Biber's adoption of *attitude markers* has allowed him to overtly comment on the discourse through the expression of evaluations and emotions, leading the participating audience to perceive his evaluative and emotional authorial identities. At last, the author's preference for some *endophoric and directive markers* have demonstrated how he engages with readers by leading them to promptly recognize relevant information located in figures, tables, sections, and other parts of the RAs, disclosing a supportive authorial identity.

Notwithstanding the accomplishment of this research, broader claims about academic writing may be substantiated through further comparative analyses of multiple authors or contexts. Therefore, some limitations have been noticed and may be resolved in future studies. *CorJAAL* may not have been the best reference corpus, but has nonetheless proven so on the basis of the research goals. Unarguably, the best layout of a reference corpus is an unsolved issue and has continued to foster additional discussions. Nevertheless, as previous research (Geluso; Hirch, 2019) has shown, a different layout might affect the findings. Furthermore, because a style analysis is comprehensive, it can be argued that gatekeepers/literacy brokers may have influenced Biber's stylistic choices. However, the attention has hereby been directed to the texts resulting from such editing and publishing processes, since tracking other people's editions to Biber's texts might have been a daunting, if not an impossible task to carry out here. In terms of keywords chosen for analysis, like the research carried out by Pojanapunya and Watson Todd (2016), future studies might compare keyword lists generated separately by likelihood and effect size measures and see how they illustrate potentially different stylistic authorial choices. At last, the *Stylistic Metadiscourse Model of Interpersonal Interaction in Academic Writing* (Silva, 2024,2025) was devised for the purposes of this research, but if applied in additional investigations, it could be verified whether it continues to disclose academic writers' style.

In view of the data analyzed in this paper, "strikingly" (an interactional keyword used in *BC*), an academic writer's style can be interactional in applied linguistics research articles. Biber's interactional rhetorical practices are essentially evaluative and engaging, expressing solidarity, anticipating objections, and revealing the extent to which he works to jointly construct texts with his readers. By debunking the myth of a dry, impersonal discourse in academic writing and taking on Crystal's (2013) early proposal, it is possible to see that Biber's tendency to an interactional metadiscourse style shows potential implications for the understanding of how an expert academic writer crafts an authorial style that may empower novice and other expert writers to craft a style that involves the audience in meaningful reading and writing experiences through academic texts.

Acknowledgements

The author of this paper is thankful to Professor Deise Prina Dutra, PhD, from Federal University of Minas Gerais (UFMG), for advising the PhD dissertation from which this paper has originated.

References

ANTHONY, L. *AntConc* (Version 4.2.0) [Computer Software]. Tokyo: Waseda University, 2022. Available at: https://www.laurenceanthony.net/software/antconc/ . Accessed on: 11 Nov. 2025.

BACON, N. Style in academic writing. *In:* DUNCAN, M.; VANGURI, S. M. (ed.). *The centrality of style*. Anderson: Parlor Press, 2013. p. 173-190.

BAKER, M. Towards a methodology for investigating the style of a literary translator. *Target*, Amsterdam/Philadelphia, v. 12, n. 2, p. 241-266, 2000. DOI : https://doi.org/10.1075/target.12.2.04bak

BAKHTIN, M. *The dialogic imagination*: four essays by M. M. Bakhtin. Austin: University of Texas Press, 1981.

BERBER SARDINHA, A. P. Comparing Corpora with Wordsmith Tools: How Large Must the Reference Corpus Be?. *In*: Annual Meeting of the Association for Computational Linguistics, 38, 2000, Hong Kong. *Proceedings of the Workshop on Comparing Corpora*... Hong Kong: Hong Kong University of Science and Technology, 2000. p. 7-13.

BERBER SARDINHA, A. P. *Lingüística de corpus*. Barueri: Manole, 2004.

BIBER, D. *et al*. *Grammar of spoken and written English*. Amsterdam/Philadelphia: John Benjamins Publishing Company, 2021.

BIBER, D. A textual comparison of British and American writing. *American Speech*, *[S. l.]*, v. 62, n. 2, p. 99-119, 1987. DOI: https://doi.org/10.2307/455273

BIBER, D. Methodological issues regarding corpus-based analyses of linguistic variation. *Literary and Linguistic Computing*, *[S. l.]*, v. 5, n. 4, p. 257–269, 1990. DOI: https://doi.org/10.1093/llc/5.4.257

BIBER, D. On the complexity of discourse complexity: A multidimensional analysis. *Discourse Processes*, *[S. l.]*, v. 15, n. 2, p. 133–163, 1992. DOI: https://doi.org/10.1080/01638539209544806

BIBER, D. Representativeness in Corpus Design. *Literary and Linguistic Computing*, *[S. l.]*, v. 8, n. 4, p. 243-257, 1993a. DOI: https://doi.org/10.1007/978-0-585-35958-8_20

BIBER, D. Using register-diversified corpora for general language studies. *Computational Linguistics*, *[S. l.]*, v. 19, n. 2, p. 219–241, 1993b.

BIBER, D. Investigating language use through corpus-based analyses of association patterns. *International Journal of Corpus Linguistics*, *[S. l.]*, v. 1, n. 2, p. 171–197, 1996. DOI: https://doi.org/10.1075/ijcl.1.2.02bib

BIBER, D. A register perspective on grammar and discourse: variability in the form and use of English complement clauses. *Discourse Studies*, *[S. l.]*, v. 1, n. 2, p. 131–150, 1999. DOI: https://doi.org/10.1177/1461445699001002001

BIBER, D. Historical patterns for the grammatical marking of stance: A cross-register comparison. *Journal of Historical Pragmatics*, *[S. l.]*, v. 5, n. 1, p. 107–136, 2004. DOI: https://doi.org/10.1075/jhp.5.1.06bib:

BIBER, D. *University language*: A corpus-based study of spoken and written registers. Amsterdam/Philadelphia: John Benjamins Publishing Company, 2006.

BIBER, D. A corpus-driven approach to formulaic language in English: Multi-word patterns in speech and writing. *International Journal of Corpus Linguistics*, *[S. l.]*, v. 14, n. 3, p. 275–311, 2009. DOI: https://doi.org/10.1075/ijcl.14.3.08bib

BIBER, D. Corpus linguistics and the study of literature: Back to the future?. *Scientific Study of Literature, [S. l.]*, v. 1, n. 1, p. 15–23, 2011. DOI: https://doi.org/10.1075/ssol.1.1.02bib

BIBER, D. Register as a predictor of linguistic variation. *Corpus Linguistics and Linguistic Theory*, *[S. l.]*, v. 8, n. 1, p. 9–37, 2012. DOI: https://doi.org/10.1515/cllt-2012-0002

BIBER, D. Interview with Douglas Biber. [Interview given to] Bethany Gray. *Journal of English Linguistics*, Thousand Oaks, v. 41, n. 4, p. 359-379, 2013. DOI: https://doi.org/10.1177/0075424213502237

BIBER, D. Using multi-dimensional analysis to explore cross-linguistic universals of register variation. *Languages in Contrast*, *[S. l.]*, v. 14, n. 1, p. 7–34, 2014. DOI: https://doi.org/10.1075/lic.14.1.02bib

BIBER, D. Text-linguistic approaches to register variation. *Register Studies*, *[S. l.]*, v. 1, n. 1, p. 42–75, 2019. DOI: https://doi.org/10.1075/rs.18007.bib

BIBER, D. *Home*. Arizona: Northern Arizona University, [201-?]. Available at: https://dougbiber.weebly.com/. Accessed on: 11 Nov. 2025.

BIBER, D.; CONRAD, S. *Register, genre, and style*. 2. ed. Cambridge: Cambridge University Press, 2019.

BIBER, D.; GRAY, B. *Grammatical complexity in academic English*: linguistic change in writing. Cambridge: Cambridge University Press, 2016.

BREZINA, V. *Statistics in corpus linguistics*: a practical guide. Cambridge/New York: Cambridge University Press, 2018.

BURKE, M. Rhetoric and poetics: The classical heritage of stylistics. *In:* BURKE, M. (ed.). *The Routledge Handbook of Stylistics*. London/New York: Routledge, 2014. p. 11-30.

CÂMARA JUNIOR, J. M. *Contribuição à estilística portuguesa*. 3. ed. Rio de Janeiro: Ao livro técnico, 1978.

CLARIVATE. *Journal Citation Reports*. London: Clarivate, 2023.

CRISMORE, A. *Metadiscourse*: what is it and how is it used in school and non-school social science texts. Urbana-Champaign: University of Illinois, 1983.

GARCIA, O. M. *Comunicação em prosa moderna*: aprenda a escrever, aprendendo a pensar. 26. ed. Rio de Janeiro: Editora FGV, 2006.

GELUSO, J.; HIRCH, R. The reference corpus matters: comparing the effect of different reference corpora on keyword analysis. *Register Studies*, Amsterdam/Birmingham, v. 1, n. 2, p. 209–242, 2019.

GRAY, B. *Linguistic variation in research articles*: when discipline tells only part of the story. Amsterdam/Philadelphia: John Benjamins Publishing Company, 2015.

GRIEVE, J. Register variation explains stylometric authorship analysis. *Corpus Linguistics and Linguistic Theory*, Berlin, v. 19, n. 1, 1-31, 2023.

HYLAND, K. The scholarly publishing landscape. In: HANGANU-BRESCH, C.; ZERBE, M. J.; CUTRUFELLO, G.; MACI, S. M. (ed.). *The Routledge handbook of scientific communication*. London/New York: Routledge, 2022. p. 15-25.

HYLAND, K. Community and individuality: performing identity in applied linguistics. *Written Communication*, Thousand Oaks, v. 27, n. 2, p. 159-188, 2010.

HYLAND, K. *Academic Discourse*: English in a global context. London/New York: Continuum International Publishing Group, 2009.

HYLAND, K. As can be seen: Lexical bundles and disciplinary variation. *English for Specific Purposes*, Amsterdam, v. 27, n. 1, p. 4-21, 2008a.

HYLAND, K. Small bits of textual material: A discourse analysis of Swales' writing. *English for Specific Purposes*, Amsterdam, v. 27, n. 2, p. 143-160, 2008b.

HYLAND, K. *Metadiscourse*: Exploring Interaction in writing. London/New York: Continuum International Publishing Group, 2005.

HYLAND, K. *Disciplinary discourses*: social interactions in academic writing. Ann Arbor: The University of Michigan Press, 2004.

HYLAND, K. Authority and invisibility: authorial identity in academic writing. *Journal of Pragmatics*, Amsterdam, v. 34, p. 1091-1112, 2002b. DOI: https://doi.org/10.1016/S0378-2166(02)00035-8

HYLAND, K. Options of identity in academic writing. *ELT Journal Volume*, Oxford, v. 56, n. 4, p. 351-358, 2002a. DOI: https://doi.org/10.1093/elt/56.4.351

HYLAND, K. *Hedging in scientific research articles.* Amsterdam/Philadelphia: John Benjamins Publishing Company, 1998a.

HYLAND, K. Boosting, hedging and the negotiation of academic knowledge. *Text*, Berlin, v. 18, n. 3, p. 349-382, 1998b.

HYLAND, K; TSE, P. Metadiscourse in academic writing: a reappraisal. *Applied Linguistics*, Oxford, v. 25, n. 2, p. 156-77, 2004. DOI: https://doi.org/10.1093/applin/25.2.156

HYLAND, K.; JIANG, F. (K.). *Academic discourse and global publishing*: disciplinary persuasion in changing times. London/New York: Routledge, 2019.

I LOVE PDF. [S. l.: s. n], 2025. Available at: https://www.ilovepdf.com/ Accessed on: 11 Nov. 2025.

IVANIČ, R. *Writing and identity*: the discoursal construction of identity in academic writing. Amsterdam/Philadelphia: John Benjamins Publishing, 1998.

JEFFRIES, L.; MCINTYRE, D. *Stylistics*. Cambridge: Cambridge University Press, 2010.

LEECH, G.; SHORT, M. *Style in fiction*: a linguistic introduction to English fictional prose. Harlow: Pearson/Longman, 2007.

LILLIS, T.; CURRY, M. *Academic writing in a global context*: the politics and practices of publishing in English. London/New York: Routledge, 2010.

MALMKJAER, K.; CARTER, R. Stylistics. *In:* MALMKJAER, K. (ed.) *The linguistics encyclopedia*. 2. ed. London/New York: Routledge, 2002. p. 510-520.

MCINTYRE, D.; WALKER, B. *Corpus stylistics*: theory and practice. Edinburgh: Edinburgh University Press Ltd, 2019.

MCMENAMIN, G. R. (ed.) *Forensic Linguistics*: advances in forensic stylistics. Boca Raton: CRC Press, 2002.

ONLINE 2 PDF. *Conversor PDF online*. [S. l.: s. n], 2007. Available at: https://online2pdf.com/ Accessed on: 11 Nov. 2025.

OCR2EDIT. *OCR2Edit - Extract text from scans, images and more.* [S. l.: s. n], 2025. Available at: https://www.ocr2edit.com/ Accessed on: 11 Nov. 2025.

PEARSON, W. S.; ABDOLLAHZADEH, E. Metadiscourse in academic writing: a systematic review. *Lingua*, Amsterdam, v. 263, p. 1-32, 2023. DOI: https://doi.org/10.1016/j.lingua.2023.103561

POJANAPUNYA, P.; WATSON TODD, R. Log-likelihood and odds ratio: Keyness statistics for different purposes of keyword analysis. *Corpus Linguistics and Linguistic Theory*, Berlin, v. 14, n. 1, 133-167, 2016. DOI: 10.1515/cllt-2015-0030

SCOTT, M. *WordSmith Tools manual*. Stroud: Lexical Analysis Software Ltd., 2022. Available at: https://lexically.net/downloads/version_64_8/HTML/index.html . Accessed on: 11 Nov. 2025.

SCOTT, M.; TRIBBLE, C. *Textual patterns*: key words and corpus analysis in language education. Amsterdam/Philadelphia: John Benjamins Publishing Company, 2006.

SILVA, J. M. V. *(Un)surprisingly, style can be found in academic writing*: towards characterizing Douglas Biber's metadiscourse style. 2024. 176f. Dissertation (Ph.D. in Linguistic Studies) – Faculdade de Letras, Universidade Federal de Minas Gerais, Belo Horizonte, 2024.

SILVA, J. M. V. Being Interactive in Applied Linguistics Research Articles: Towards Characterizing an Academic Writer's Metadiscourse Style. *Revista Brasileira de Linguística Aplicada*, Belo Horizonte, v. 25, n. 1, p. 1-29, 2025. DOI: https://doi.org/10.1590/1984-6398202549625

SINCLAIR, J. *Trust the text*: language corpus and discourse. London: Routledge, 2004.

SWALES, J. M. *Genre analysis*: English in academic and research settings. Cambridge: Cambridge University Press, 1990.

TARDY, C. Current conceptions of voice. *In:* HYLAND, K.; GUINDA, C. (ed.). *Stance and voice in written academic genres*. London: Palgrave Macmillan, 2012. p. 34-47.

THOMPSON, G. Interaction in academic writing: learning to argue with the reader. *Applied Linguistics*, Oxford, v. 22, n. 1, p. 58-78, 2001.

THOMPSON, G.; THETELA, P. The sound of one hand clapping: the management of interaction in written discourse. *TEXT*, Berlin, v. 15, n. 1, 103-27, 1995. DOI: https://doi.org/10.1515/text.1.1995.15.1.103

TOGNINI-BONELLI, E. *Corpus linguistics at work*. Amsterdam/Birmingham: John Benjamins Publishing Company, 2001.

VANDE KOPPLE, W. Some exploratory discourse on metadiscourse. *College Composition and Communication*, New York/ Ann Arbor, *v.* 36, n. 1, p. 82-93, 1985. DOI: https://doi.org/10.2307/357609

WALTMAN, L.; TRAAG, V. A. Use of the journal impact factor for assessing individual articles: Statistically flawed or not? *F1000Research*, London, v. 9, p. 1-29, 2000.