Metadiscourse: The evolution of an approach to texts

Abstract

Metadiscourse is the commentary on a text made by its producer in the course of speaking or writing, revealing something of how communication involves the personalities, attitudes and assumptions of those who are communicating. It offers a framework for understanding communication as social engagement and helps reveal how writers and speakers consider their audience in creating texts. This paper uses a bibliometric analysis to trace the growing interest in metadiscourse since its early incarnations in the 1980s. To do so we analysed all 431 papers relating to metadiscourse in the core collection of the Web of Science between 1983 and 2020, dividing the corpus into two periods following the massive increase in interest after 2006. We identify which topics have been most prevalent, which authors and publications most influential and which disciplines and journals most active in citing the metadiscourse literature. The findings show the importance of academic and business writing, cross-disciplinary, language and genre studies, and the increasing predominance of an interpersonal model. These findings may be of interest to those working in discourse analysis and the study of social interaction.

Keywords

1. Introduction

Metadiscourse is the commentary on a text made by its producer in the course of speaking or writing. This commentary allows writers to guide, direct and interact with their readers and demonstrate a concern for their ability to understand a text as it is intended, recognising the writer's stance, seeing connections between ideas and feeling involved in what is being communicated. Metadiscourse, then, is a recipient design filter, which shapes a text out of consideration for its readers or hearers based on an estimation of how best they can be helped to process, comprehend and agree with what is being discussed. By explicitly stepping into a text in this way, we see the writer's assessment of his or her audience and therefore something of the context in which language is being used. As a result, metadiscourse has become one of the leading approaches to studying texts, so that a *Google Scholar* inquiry returns 29,500 articles containing the term, 11,000 of these being in the last 5 years.

In this paper we seek to track the emergence of metadiscourse since it was first brought to prominence in the early 1980s (Crismore 1983) to its current popularity. Adopting a bibliometric approach, we explore the literature on the topic over the past 37 years (1983-2020), seeking to identify changes in published metadiscourse papers related to research topics, influential publications and authors and their geographical sources. Specifically, we address these four questions:

(1) What have been the most frequently explored topics and have these changed?

(2) Which authors have been most influential and have these changed?

(3) Which publications have been most influential and have these changed?

(4) Which disciplines and journals have been most active in citing this literature?

We first offer a brief overview of metadiscourse, go on to describe our methods, then present our results in terms of the four research questions above. We conclude by showing the significance and implications of the research.

2. Metadiscourse: a brief overview

Briefly, metadiscourse captures the ways writers organize their texts to help readers understand and assess propositional information (Ädel and Mauranen 2010; Hyland 2005). Its roots lie in the idea that language not only refers to the world, concerned with exchanging information of various kinds, but also to itself: with material which helps readers to organise, interpret and evaluate what is being said. This view connects metadiscourse to earlier ideas in linguistics such as Jacobson's (1980) 'metalinguistic function' of language, which refers to language focusing on the text itself, and Halliday's (1985: 271) 'metaphenomena', which are "categories of the language, not of the real world". The idea was also present in pragmatics, where Watzlawick, Beavin and Jackson (1968) commented:

> The ability to metacommunicate appropriately is not only the *condition sine qua non* of successful communication, but is intimately linked with the enormous problem of awareness of self and others. (p.53)

The term 'metadiscourse' itself, however, was originally introduced by the structural linguist Zelig Harris (1959) but it was not taken up in discourse studies until Williams' (1981) and Crismore's (1983) attempts to refine and operationalise it by creating functional sub-categories. The work by Vande Kopple (1985) and Crismore (1989) was instrumental in helping the concept gain traction by making a firm distinction between propositional material, what is talked *about*, and metadiscourse, which signals the presence of a text-organising and content-evaluating author rather than the subject matter. Their work was foundational in identifying two broad categories of metadiscourse, borrowing Halliday's (1994) terms for language metafunctions: textual and interpersonal. Unlike Halliday, however, they saw textual resources not as intrinsic to all language use, but as constituting a neatly separable set clearly distinguished from either propositional or interpersonal aspects.

Later models of metadiscourse have sought to address this limitation, and we can see the different perspectives as spread along a theoretical continuum ranging from a

narrow text-centred view to a broad interpersonal one (Hyland 2017). At one end sits a view championed by Mauranen (1993), which she refers to as *metatext* or *text reflexivity*. This attempts to sharpen the concept by excluding all 'non-reflexive' features and restricting metadiscourse to expressions such as *'the presentation is in four parts'* or *'we now turn to look at the disadvantages',* which refer to the direction, purpose and internal structure of the text itself. Further along the continuum, we find studies which extend this 'reflexive' view of metadiscourse to include how writers refer to themselves, their readers and their texts. Ädel (2006), for example, adds features which refer to the writer and the imagined reader of the text to those which comment on the text (e.g. Zhang 2016).

Also positioned along the cline are alternative conceptions, such as Beauvais' (1989) attempts to limit metadiscourse to explicit illocutionary predicates or Ifantidou's (2005) reformulation based on a relevance framework. We also find studies which subscribe to a broader definition but focus on a limited range of features, such as code glosses (Hyland 2007), hedges (Hu and Cao 2011), stance (McGrath and Kuteeva 2012) or interactional features (Gillaerts and Van de Velde 2010).

At the far end of the cline, analysts see a writer or speaker's commentary on his or her unfolding text as representing a set of interpersonal options. Here metadiscourse comprises the resources writers employ to both organise textual material and make it comprehensible and persuasive by projecting themselves into their discourse. Hyland (Hyland 2005; Hyland and Tse 2004), for example, seeks to capture the interpersonal character of communication in the concept, with a distinction between *interactive* and *interactional* resources (Thompson 2001). *Interactive* refers to the writer's management of the information flow to guide readers through a text, setting out the structure, referring to sources and linking parts of the discourse; and *interactional* refers to the writer/speaker's interventions to express a stance and build a relationship with readers.

This appears to be a natural and logical extension of a concept which seeks to consolidate the linguistic devices speakers and writers use to shape their messages for particular listeners or readers. Here metadiscourse is understood as a coherent set of interpersonal resources used to organise a discourse or the writer's stance towards either its content or the reader. It is an umbrella term for a heterogeneous array of features which assist readers not only to connect and organise material but also to interpret it in a way preferred by the writer and with regard to the understandings and values of a particular discourse community. As such it offers an appropriate frame for the interpretation of discourse. In this view both organisational and evaluative features are necessary to create an interpersonally effective text, although some researchers feel that this broad interpretation weakens the term by trying to include too much (Ädel and Mauranen 2010). The proponents, however, argue that the interpersonal model offers a more inclusive view of how we monitor our production by making decisions about the effects we are having on our audience and that a finished text is an outcome of this awareness of the reader.

Critics of metadiscourse, however, argue that these diverse perspectives point to fundamental conceptual difficulties with the term and an inability to pin down what it is. Hyland (2005), for example, points out that the traditional distinction between propositional and metadiscoursal elements of a text cannot be sustained in practice. Items often identified as metadiscoursal, such as *therefore, in contrast* and *as a result of*, can act as metadiscourse by connecting steps in an argument or work 'propositionally' to connect events in the world outside the text. There are other analytical difficulties, such as deciding what constitutes an instance of metadiscourse as longer units might encompass smaller units, while the use of predefined sets of lexical items for corpus searches can imply that the function of each form is unvarying (Ädel and Mauranen 2010). A third, related, criticism is that the formal heterogeneity of metadiscourse means that functions may be performed in different ways or linguistic forms may perform more than one function simultaneously. The same forms, for example, can convey different categories of metadiscourse, so that *quite* can be a hedge (*quite good*) or a

booster (*quite extraordinary*), while particular functions can be expressed in numerous ways (*even if, of course, admittedly, although*).

This kind of category overlap is well known in discourse analysis, however, and is a consequence of the multi-functionality of language itself. It points to the need to see lists of items as *potentially* performing metadiscourse functions; as a starting point of high frequency items that commonly work as metadiscourse in a particular register. Researchers stress the need for items to be examined in their sentential contexts to ensure they are performing metadiscourse functions as well as recommending sweeps of the corpus to discover unexpected realisations, such as 'metadiscourse nouns' (Jiang & Hyland, 2016). Metadiscourse studies therefore underline, rather than resolve, the problem of polypragmatic meanings.

Despite these issues, however, metadiscourse is seen as a useful way of seeing how writers understand the context in which they are working by using language to take readers' needs, understandings and existing knowledge into account. Analysts regard writers' successful management of these local rhetorical resources as a key means of achieving their immediate communicative objectives.

We now turn to discuss the methods and findings of our study.

3. Bibliometric analysis: data and procedures
Bibliometrics is the 'statistical analysis of books, articles, or other publications' (Oxford English Dictionary 2021). Used in library and information sciences, it seeks to describe patterns of publication within a given field or body of literature, most famously focusing on citations for a scholarly publication as evidence of its value. It has also helped to explore authorial networks (Fahimnia, Sarkis and Davarzani 2015), compare publishing countries (Fan et al. 2020), and document citation patterns (Ahmad et al. 2020). It is also often used to study the research productivity of individual scholars and countries (Ellegaard and Wallin 2015).

Generally, bibliometric studies have not sought to identify the most popular research themes, cited authors, or publications in a field. There have been recent exceptions, however, with studies by Lei and Liu (2019), Zhang (2020) and Hyland & Jiang (2021a, 2021b) in linguistics. Lei and Liu (2019) explored the most cited topics, authors and publications in 40 Social Science Citation Index (SSCI) applied linguistics journals between 2005–2016. Zhang (2020) used a bibliometric method to track changes in second language acquisition research between 1997 and 2018. Hyland and Jiang traced the emergence of English for Academic Purposes (EAP) over the past 40 years (2021a) and English for Specific Purposes (ESP) since 1990 (2021b). The value of the approach is that it uses mathematical and statistical methods to characterise the development of an area. It can therefore complement more qualitative State of the Art reviews, which tend to summarize current and emerging trends from a particular angle (e.g. Hyland, 2017; D'Angelo & Consonni 2020), by bringing a quantitative dimension to understand the historical publication trends of a field more broadly.

This paper follows this general line of research while adopting a more specific focus. It also adds a diachronic dimension to Hyland's (2017) overview of metadiscourse. We conducted a search for *metadiscourse* in the core collection of *Web of Science* (WoS), scanning titles, abstracts and keywords between 1983 and 2020. Unlike Google Scholar, which trawls the internet and finds papers of variable quality and provenance, the *Web of Science* is a curated database containing articles in peer reviewed journals in the Social Science Citation Index. We chose 1983 as our starting point as this was the date of Avon Crismore's (1983) report of metadiscourse in social science textbooks, the earliest explicit empirical exposition on the topic we are aware of. The search produced 431 papers, as shown in Figure 1.

Figure 1 Number of metadiscourse publications across time
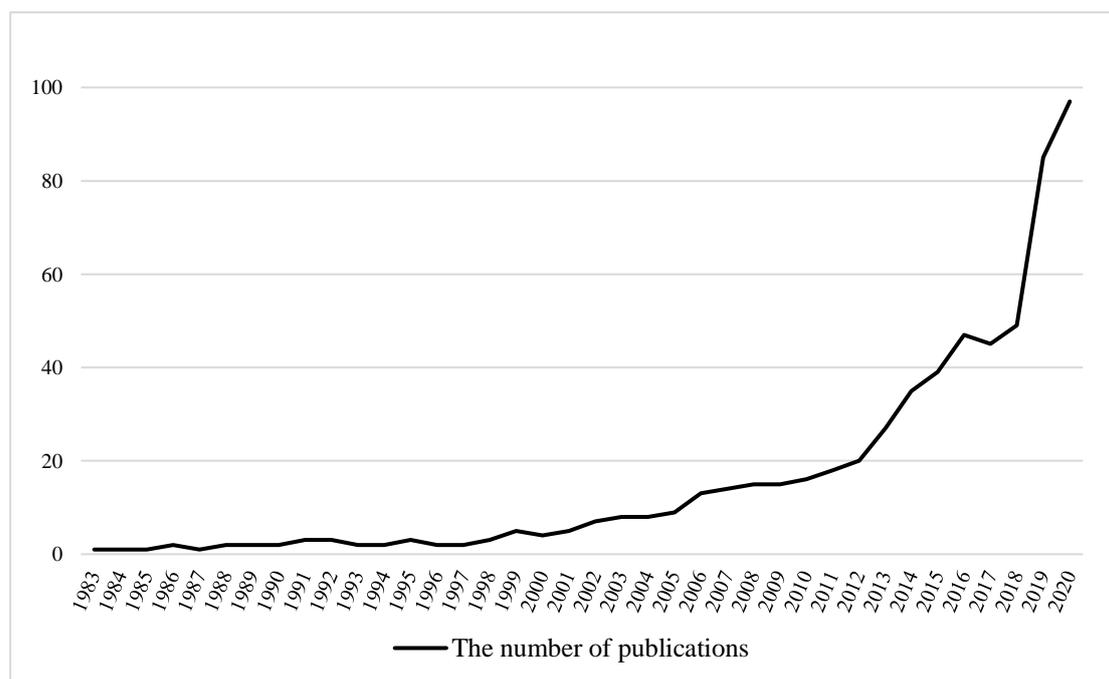
We can see from the table that there is a marked increase in the number of papers published on metadiscourse over the years with a surge from around 2005 onwards. Hyland and Jiang (2016) report a doubling of papers in applied linguistics between 1965 and 2015, although this does not account for this rise in metadiscourse research. Perhaps more relevant is that 2005 saw the publication of two major books on the topic, Adel's (2006) *Metadiscourse in L1 and L2 English* and Hyland's (2005) *Metadiscourse*, which have had a significant impact on metadiscourse studies. We therefore decided to make 2006 a break point and trace changes before and after this date. There are 61 articles in the WoS in the first period from 1983 to 2005, and 370 in the second from 2006 to 2020. We then interrogated these two corpora to discover:

    (1) The most frequently explored topics overall and in each period

    (2) The most cited authors overall and in each period

    (3) The most cited publications in the two periods

    (4) The most active journals citing metadiscourse papers

To answer question 1, about frequent topics, we did the following:

1) Annotated all the target abstracts with part-of-speech and lemma information[1] using Schmid's (1995) *TreeTagger* programme.

2) Extracted n-grams of 2 to 5 words from the tagged corpora to identify candidate topics using *AntConc* (Anthony 2019).

3) Manually checked all the results to exclude nouns and lexical strings which did not constitute research topics.

We included only nouns when identifying potential metadiscourse subject areas as these are far more likely to constitute research topics than other parts of speech, while 2-5 word *n*-grams seemed a realistic range to capture topics such as *research articles* (2-word gram) and *personal and impersonal authorial references* (5-word gram). We then discarded three main categories: a) words and clusters common in language in general (e.g. *frequency* and *in depth*); b) concepts and issues not specific to metadiscourse (*basis of*, *under investigation*, *the study*); and c) concepts and issues common in metadiscourse but too general to be useful (e.g. *function, language*) (see also Lei and Liu 2019).

Lei and Liu (2019) and Hyland and Jiang (2021a) set a threshold frequency of 30 occurrences for inclusion and showed this as sufficient to ensure the significance of the selected items but not too high as to exclude important topics. In the current study our topic is more specialised, so the corpus size is much smaller, i.e., 431 articles compared to 10,000. We therefore reduced this constraint to include topics occurring at least 5 times. We also added a range criterion, so that items must occur in no less than 5 publications to ensure a reliably widespread appearance of a topic in the literature. We then normalised the frequencies by representing the number of occurrences of each topic per 100 papers to compare differences between the two time periods. Finally, we ran a one-way chi-square test for each of the topics across the two periods to determine the statistical significance of differences in the topic frequencies.

---

[1] A lemma is the dictionary or citation form of a word, e.g. run, runs, ran and running are indexed by the lemma 'run'. This ensured that we captured all examples of a form.

To address questions 2 and 3, concerning the most influential authors and publications, we first identified the most-cited authors in the reference lists of the papers in each corpus using a frequency count. To find the most highly cited works, we computed the frequencies of all the publications (books, chapters, and articles) from the reference lists in the corpus, using a regular expression search, and then normed the citations per 100 papers.

To address question 4, regarding the most active sources citing metadiscourse papers, we explored the disciplinary field and publication language of the journals citing metadiscourse. The information was extracted using a tool provided in the *Web of Science*.

4. Metadiscourse research in historical perspective

In this section we discuss our findings taking each research question in turn.

4.1 The most frequently studied metadiscourse topics

The criteria discussed above (at least 5 occurrences in at least 5 papers) produced 196 frequently discussed research topics related to metadiscourse over the 37 years. Table 1 shows those topics with consistently high frequencies throughout. It underlines the overwhelming concern of metadiscourse researchers in describing the character of the academic register. This interest in unpacking the mechanisms of knowledge production and communication in the academy is shown in the frequencies of *academic discourse, academic writing* and *research articles*. These have been perennial themes since the beginning of metadiscourse research and prominent in the work of the early pioneers discussed above. Hyland and Jiang (2021) also found *writing in higher education* to be among the most frequent themes in articles published in the journal *English for Specific Purposes* (ESP) over the past 40 years as metadiscourse research reflects, or perhaps drives, the huge interest in interaction in academic texts.

The key role of *research articles* in the promotion of careers, knowledge and disciplines means that interest in metadiscourse as a central element of the *rhetorical strategies* authors need to establish a persuasive argument has grown. It is also interesting to find *pedagogical implication* in the list, indicating the influence of the concept in English for Academic Purposes (EAP) and academic writing classes.

Table 1 Topics with consistently high frequency over the periods (by normed change)

	1983-2005		2006-2020		% normed change	Chi-value	*p*
	raw	normed	raw	normed			
Constant							
academic discourse	25	41.0	166	44.9	9.5	0.15	.70
research articles	13	21.3	86	23.2	9.0	0.07	.79
pedagogical implications	31	50.8	205	55.4	9.0	0.17	.68
rhetorical strategies	21	34.4	137	37.0	7.6	0.08	.78
academic writing	32	52.5	201	54.3	3.6	0.02	.88

While the topics listed in Table 1 were consistently high over the period, some topics rose and fell considerably as the field grew. Table 2 shows those topics recording the most statistically significant changes. Once again, the topics are listed by their percentage change in normed frequency.

Table 2 The metadiscourse topics with the most dramatic changes over the period

	1983-2005		2006-2020		% normed change	Chi-value	*p*
	Raw	normed	raw	normed			
Significantly up							
business communication	5	8.2	20	32.4	295.7	10.51	0.00
student writing	5	8.2	18	31.9	289.1	10.21	0.00
expert writers	6	9.8	39	37.7	283.0	11.92	0.00
across languages	7	11.5	39	37.6	227.4	10.42	0.00
disciplinary variation	8	13.1	51	40.8	211.2	10.78	0.00
different genres	8	13.1	50	40.5	209.1	10.63	0.00
corpus-based study	18	29.5	98	80.5	172.6	18.41	0.00

Significantly down						
communication theory	7	11.5	10	2.7	76.5	10.33	0.00
letters	9	14.8	14	3.8	74.4	11.95	0.00
Italian and English	8	13.1	13	3.5	73.2	10.03	0.00
coherence	14	23.0	24	6.5	71.7	16.31	0.00
modality	12	19.7	25	6.8	65.7	10.32	0.00

It is interesting to find *business communication* achieving the greatest gains as ESP research has grown and become more interested in the texts of the commercial world. Studies of academic genres on the other hand, as we have seen above, have remained consistently high. The focus on particular types of writers, whether *student, expert* or *disciplinary*, have increased massively, often as a way of unpacking their texts for pedagogic purposes or as variables in comparative studies. Variations in the use of metadiscourse across *writers, languages, disciplines and genres* have tended to be mainstays of published research since the beginning and this interest has increased massively over time.

The study of metadiscourse in these different rhetorical and social contexts has been hugely influential in understanding how interaction works in written and (occasionally) spoken discourse. In addition, it has also been instrumental in revealing something of the different communicative purposes of genres and the epistemological norms and social practices of different linguistic and professional communities. *Corpus* studies have played a significant role in this research, revealing the regular patterns of metadiscourse use in particular genres and registers and by given groups of writers. The results suggest the routine, almost automatic, practices of disciplinary insiders in selecting the most effective argument forms in a particular language, discipline or genre.

The items with the most significant falls have very low raw frequencies and perhaps little can be gained from their analysis. The decrease in *communication theory* as a topic in metadiscourse research, for example, may not indicate a falling interest in how metadiscourse fits into wider theories of interaction but simply be a fringe issue.

Similarly, studies of the scientific *letters* genre and comparisons of the use of metadiscourse by *Italian and English* writers have also fallen, together with an explicit focus on *modality* and *coherence*. These two latter features are central to metadiscourse, with modality being a key part of stance-taking and coherence pointing to the ways metadiscourse is used to fit texts together in convincing ways for readers. The decline in normed references is therefore in the *terms* rather than in the *concepts* themselves.

The continuing interest in what metadiscourse is, and not just what it can tell us about texts, is shown in the topics which have emerged after 2006, shown in Table 3. These do not appear in the earlier list and so indicate a new interest in metadiscourse models, its pedagogical potential and emergent types.

Table 3   Most frequent new topics after 2006

	2006-2020		
	raw	normed	No of papers
Emerging			
interactional metadiscourse	112	30.27	101
interactive metadiscourse	56	15.14	49
reflexive metadiscourse	40	10.81	35
pragmatic competence	33	8.92	33
graduate students	26	7.03	21
metadiscursive nouns	10	2.70	8
visual metadiscourse	10	2.70	7

The table shows that, in this data, Hyland's (2005) interpersonal model has attracted the greatest interest with *interactional metadiscourse* receiving the most research attention. This is partly a consequence of the 'interactive turn' in discourse analysis (e.g. Hyland 2004), which has sought to identify the ways that writers (and particularly successful academic writers) have their projected readers in mind as they craft their texts. Despite Swales' (2019) sense that this does not coincide with his own practice, the concepts of stance, engagement and authorial presence have continued to appeal to researchers and inform studies of writing. *Interactive metadiscourse* is the other side of the interpersonal coin, concerned with the organisation of discourse for particular readers rather than the

construction of a viable writer and rhetorical relationship. The appearance of *reflexive metadiscourse* (Mauranen 1993; Ädel 2006) indicates a model restricted to the ways that writer/speakers talk about the evolving text itself.

Pragmatic competence appears on the list as researchers realise that metadiscourse is a key pragmatic construct concerning how writers project themselves to effect communication, signalling their intentions and attracting reader/hearers. G*raduate students* points to the increasing attention given to advanced learners' texts and the growing role of metadiscourse as a resource for teaching academic writing. More recently, *metadiscursive nouns* and *visual metadiscourse* have been proposed as important extensions of the term. Metadiscursive nouns (such as *fact*, *analysis*, *belief*) offer writers a way of organizing discourse into a cohesive flow of information and constructing a stance towards it (Jiang & Hyland, 2016). Visual metadiscourse, in contrast, refers to a set of devices that designers use to influence audience's interpretation of the text (De Groot et al 2015).

4.2 Most cited authors

Metadiscourse research is not only characterised by a growing number of topics and themes, but also in attracting an ever-increasing range of authors. To determine the most influential academics in this field we identified the most highly cited by conducting a frequency count of all items in the reference lists in our corpus of 431 articles. We then divided these into pre- and post-2005 periods and normalised the results to 100 papers. Table 4 lists the top 15 cited authors in these papers.

Table 4 Most highly cited authors across the two periods (raw and per 100 papers)

1983-2005			2006-2020		
Authors	**Raw**	**normed**	**Authors**	**raw**	**normed**
Avon Crismore	91	149.9	Ken Hyland	1675	452.7
Ken Hyland	70	114.8	Avon Crismore	346	93.5
William Vande Kopple	39	63.9	Anna Mauranen	250	67.6
M. A. K. Halliday	36	59.0	John M. Swales	240	64.9
John M. Swales	30	49.2	M. A. K. Halliday	177	47.8

Margaret S. Steffensen	29	47.5	Douglas Biber	162	43.8
Anna Mauranen	28	45.9	Geoff Thompson	149	40.3
Douglas Biber	19	31.1	William Vande Kopple	122	33.0
Raija Markkanen	17	27.9	Vijay K. Bhatia	122	33.0
Deborah Schiffrin	17	27.9	Ana I. Moreno	121	32.7
Robert T. Craig	16	26.2	Nigel Harwood	82	22.2
Rodney Farnsworth	16	26.2	Geoffrey Leech	73	19.7
Susan Hunston	15	24.6	Ulla Connor	72	19.5
Geoffrey Leech	14	23.0	David Bunton	71	19.2
Greg Myers	14	23.0	Susan Conrad	67	18.1

Table 4 shows that almost half the names appear in both lists, demonstrating the continuing influence of early writers. The pioneering work of Avon Crismore, William Vande Kopple and Anna Mauranen remain central to later developments and Hyland's contribution is evident in the fact that his publications continue to attract citations. The other names in both periods – Swales, Halliday and Biber – are not directly acknowledged as advocates for metadiscourse, or known for their use of it, but they have been inspirational through their theoretical and empirical work on language use and text analysis.

This, then, is a diverse pool of academics. In the earlier period we find the idea fermenting in the work of composition scholars such as Crismore, Vande Kopple and Steffensen, the rhetorician Rodney Farnsworth, and communication theorist Robert Craig. Steffensen, Markkanen and Farnsworth have all worked with Crismore (see Table 5). The remaining names on the first list are all linguists of various hues: Halliday a theoretical linguist; Leech, Hunston and Biber corpus grammarians; Shiffrin, Myers, Hyland and Swales discourse analysts. Each has addressed a different conceptual or empirical facet of metadiscourse: Hyland largely focusing on disciplinary variations and Mauranen on cross-cultural and linguistic differences, Swales on genre, Biber on register variation, and Myers on metadiscourse features as strategic politeness.

In the last 15 years, many of the writers in the 1983-2005 list have consolidated their appeal to metadiscourse analysts with seven names from the earlier period in the top eight. As might be expected from the topics in Table 3, Hyland is the most cited author in the metadiscourse sources, with the early works of Crismore, Mauranen and Vande Kopple continuing to influence the field. The addition of Geoff Thompson is a result of his paper, for the first time distinguishing interactional and interactive aspects of language, which was picked up by Hyland in his model. David Bunton's detailed paper examining doctoral dissertations helped researchers understand the role of interactive metadiscourse in structuring long pieces of writing.

Bhatia's work on professional texts has encouraged the study of metadiscourse in workplace discourses and Harwood's numerous studies of doctoral supervision have also interested metadiscourse writers. Susan Conrad, like Douglas Biber and Geoffrey Leech, has advanced the understanding of natural language use through the use of corpora while Ana Moreno and Ulla Connor have explored various rhetorical features of English texts written by speakers of other languages. This eclectic group of researchers have all, whether directly through their studies of metadiscourse or of the contexts in which language is used, provided inspiration or insights for the many authors researching metadiscourse.

4.3 Most influential publications

We assume here that the most influential publications are those which receive the most citations. Citation is the standard means by which authors demonstrate their indebtedness to earlier work and indicate its relevance to their own studies, thus demonstrating the impact of that work. As a result they are widely used by funding bodies, promotion boards and appointment committees. They can, however, be distorted by self-citation or negative citations, although their influence seems to be relatively small (e.g. Hyland & Jiang 2018a). When considering the massive rise in citations after 2005 it is also important to recognise that not only are there more works to be cited in the second period, but also that older publications have more time to gather citations.

Because of this, bibliometricians therefore prefer to count citations for papers of similar age when they want to measure a paper's value (e.g. Cooper 2015). But while our lists may fail to capture the impact of more recent publications, our results show changes in the field and the influence of key publications. Tables 5 and 6 present the raw and normed frequencies of the top 15 most cited publications for each period.

Table 5 Most highly cited publications from 1983 to 2005 (raw and per 100 papers)

Publications	Cites	normed
Hyland, K. (1998). Persuasion and context: the pragmatics of academic metadiscourse. *Journal of Pragmatics*, *30*(4), 437–455.	8	13.11
Vande Kopple, W. J. (1985). Some exploratory discourse on metadiscourse. *College Composition and Communication*, 82–93.	7	11.48
Crismore, A., & Farnsworth, R. (1989). Mr. Darwin and his readers: exploring interpersonal metadiscourse as a dimension of ethos. *Rhetoric Review*, *8*(1), 91–112.	7	11.48
Halliday, M. A. K. (1978). *Language as social semiotic: The social interpretation of language and meaning*. Baltimore: University Park Press.	7	11.48
Kitagawa, C., & Lehrer, A. (1990). Impersonal uses of personal pronouns. *Journal of Pragmatics*, *14*(5), 739–759.	6	9.84
Mauranen, A. (1993). Contrastive ESP rhetoric: metatext in Finnish-English economics texts. *English for Specific Purposes*, *12*(1), 3–22.	6	9.84
Austin, J. L. (1962). *How to do things with words.* Oxford: Oxford University Press.	6	9.84
Barton, E. L. (1995). Contrastive and non-contrastive connectives. *Written Communication*, *12*(2), 219–239.	6	9.84
Bazerman, C., & Paradis, J. G. (Eds.). (1991). *Textual dynamics of the professions: Historical and contemporary studies of writing in professional communities*. Madison: University of Wisconsin Press.	6	9.84
Beauvais, P. J. (1989). A speech act theory of metadiscourse. *Written Communication*, *6*(1), 11–30.	5	8.20
Bhatia, V. K. (1993). *Analysing genre: Language use in professional settings*. London: Longman.	5	8.20
Biber, D., & Finegan, E. (1989). Styles of stance in English: lexical and grammatical marking of evidentiality and affect. *TEXT*, *9*(1), 93–124.	5	8.20
Biber, D., Johansson S., Leech G., Conrad, S., & Finegan E. (1999). *Longman grammar of spoken and written English*. Harlow: Longman.	5	8.20

Publications		
Hyland, K. (2002). Authority and invisibility: authorial identity in academic writing. *Journal of Pragmatics*, *34*(8), 1091–1112.	5	8.20
Crismore, A., & Farnsworth, R. (1990). Metadiscourse in popular and professional science discourse. In W. Nash (Ed.), *The writing scholar: Studies in academic discourse* (pp. 118–136). Newbury Park: SAGE.	4	6.56
Hyland, K. (1999). Talking to students: metadiscourse in introductory coursebooks. *English for Specific Purposes*, *18*(1), 3–26.	4	6.56

Table 6 Most highly cited publications from 2006 to 2020 (raw and per 100 papers)

Publications	Cites	normed
Hyland, K., & Tse, P. (2004). Metadiscourse in academic writing: a reappraisal. *Applied Linguistics*, *25*(2), 156-177+288.	62	16.76
Hyland, K. (2005). Stance and engagement: a model of interaction in academic discourse. *Discourse Studies*, *7*(2), 173–192.	61	16.49
Hyland, K. (2005). *Metadiscourse: Exploring interaction in writing*. London: Continuum.	55	14.86
Vande Kopple, W. J. (1985). Some exploratory discourse on metadiscourse. *College Composition and Communication*, 82–93.	49	13.24
Hyland, K. (1998). Persuasion and context: the pragmatics of academic metadiscourse. *Journal of Pragmatics*, *30*(4), 437–455.	49	13.24
Thompson, G. (2001). Interaction in academic writing: learning to argue with the reader. *Applied Linguistics*, *22*(1), 58–78.	47	12.70
Swales, J. M. (1990). *Genre analysis: English in academic and research settings*. Cambridge: Cambridge University Press.	42	11.35
Crismore, A., Markkanen, R., & Steffensen, M. S. (1993). Metadiscourse in persuasive writing: a study of texts written by American and Finish university students. *Written Communication*, *10*(1), 39–71.	41	11.08
Hyland, K. (2007). Applying a gloss: exemplifying and reformulating in academic discourse. *Applied Linguistics*, *28*(2), 266–285.	38	10.27
Bhatia, V. K. (2004). *Worlds of written discourse*. London: Continuum.	37	10.00
Ädel, A., & Mauranen, A. (2010). Metadiscourse: diverse and divided perspectives. *Nordic Journal of English Studies*, *9*(2), 1–11.	35	9.46
Biber, D., Johansson S., Leech G., Conrad, S., & Finegan E. (1999). *Longman grammar of spoken and written English*. Harlow: Longman.	33	8.92
Hyland, K. (2004). Disciplinary interactions: metadiscourse in L2 postgraduate writing *Journal of Second Language Writing*, *13*(2), 133–151.	32	8.65

Abdi, R. (2002). Interpersonal metadiscourse: an indicator of interaction and identity. *Discourse Studies*, *4*(2), 139–145.	27	7.30
Biber, D. (2006). Stance in spoken and written university registers. *Journal of English for Academic Purposes*, *5*(2), 97–116.	27	7.30
Myers, G. (1989). The pragmatics of politeness in scientific articles. *Applied Linguistics*, *10*(1), 1–35.	25	6.76

It is interesting to note that the 30 most cited titles are dominated by journal articles, with 10 in the first period and 12 in the second. This is in marked contrast to Hyland and Jiang's (2021a, 2021b) bibliometric analyses of EAP and ESP research which found that books predominate. Certainly institutions have pushed academics towards research papers for their quicker turnaround and more rapid accumulation of citations, which benefit them in assessment exercises. While books appear to have greater longevity, it seems that articles are favoured when examining a specific area, such as metadiscourse, as they not only give authors swift publication and immediate visibility, but also scholarly currency over an extended period.

Most of the authors discussed in the previous section appear somewhere on these two lists of most cited publications. Some key figures, such as Hyland (9 appearances), Crismore (3), Biber (3), Mauranen (2), Bhatia (2) and Vande Kopple (2) comprise 2/3 of the authors. Perhaps surprisingly, three items, Hyland's 1998 pragmatics paper on academic metadiscourse, Vande Kopple's 1985 work setting out the precepts of the model and Biber et al.'s 1999 Longman Grammar, appear on both lists, demonstrating their enduring influence.

In the first period, five titles explore the emerging concept of metadiscourse through an analysis of a particular genre or register. Hyland's *persuasion and context* discusses research writing and *Talking to students*, textbook features; Crismore and Farnsworth's *Mr. Darwin and his readers* describes a famous scientific treatise, and a book chapter addresses metadiscourse in professional and popularised science. Mauranen's *Contrastive ESP rhetoric*, based on her PhD dissertation, compares student writing in

two languages. The articles by Beauvais and Vande Kopple set out very different models of metadiscourse. Beauvais sees metadiscoursal features as indicators of expositive illocutionary acts and Vande Kopple provides a taxonomy of metadiscursive functions. A number of papers address particular features of metadiscourse: Kitagawa and Lehrer focus on personal pronouns, Barton on connectives, Biber and Finnegan on stance and Hyland on authorial self-mention.

The remaining works provide grammatical or pragmatic discussions of language use, which metadiscourse researchers have drawn on to support their work. Halliday (1994), for example, never mentions the term metadiscourse in his work, but the essays in *Language as Social Semiotic* expound his view of language as a signalling system, embedded within a cultural matrix, which is a rich source of ideas for metadiscourse writers. Similarly, Austin's (1975) speech act theory, set out in *How to do things with words*, takes the same focus on function, discourse and context as Halliday but situates it in philosophical pragmatics, again providing insights on how understanding is created in interaction and informing metadiscourse work. The books by Bazerman and Paradis and Vijay Bhatia examine the way writers manipulate rhetorical resources to communicate successfully in professional settings, providing a broad theoretical reference which has been used by metadiscourse analysts to understand occupational texts. Finally, Biber et al.'s *Longman Grammar of Spoken and Written English* provides a comprehensive, corpus-based grammar for the analysis of texts and often a starting point for analysts' exploration of metadiscoursal features.

In Table 6 we see a higher proportion of the cited work dealing with metadiscourse more specifically as the field becomes more established and with less need to explore its antecedents or theoretical supports. Eight items have metadiscourse in their titles and another three focus on specific functions: Hyland's *Stance and engagement* and *applying a gloss*, and Biber's treatise on *stance in spoken and written university registers.* The remaining five texts all deal with written discourse in various ways and

all, excepting Bhatia's *Worlds of Written Discourse*, with academic texts. All have been mentioned in the previous section.

One striking aspect of Table 6 is that Hyland has 6 titles in the list, indicating his considerable influence on metadiscourse research. We noted his interpersonal model in section 2 and the texts setting this out in detail appear in Table 6. The papers in *Applied Linguistics* and *Discourse Studies* and the monograph with Continuum are the main sources of the model and have attracted the most citations since 2006. The model draws on Thompson's (2001) interactive and interactional distinction from *Interactions in Academic Writing* and Swales' (1990) pioneering discussion relating language to context in *Genre Analysis,* which appear high in the list, and is indebted to Vande Kopple's early speculative *exploratory discourse on metadiscourse* as well as Crismore, Markkanen and Steffensen's more empirical study, comparing *metadiscourse in persuasive writing* by Finnish and American students.

The reflexive model is represented on the list by Ädel and Mauranen's discussion in *Metadiscourse: Diverse and Divided Perspectives*, which introduces a journal special issue on metadiscourse. Here they refer to the two main models as integrative (including interpersonal elements) and non-integrative (text-reflective) approaches, favouring the latter. Hyland's interpersonal model is elaborated and applied to professional research writing in *Persuasion and context* and to postgraduate thesis writing in *Disciplinary interactions.* Two papers focus on a particular aspect of metadiscourse: Abdi's *interpersonal metadiscourse* compares the use of hedges, boosters and attitude markers across disciplines and Hyland looks at the use of code glosses in research articles in *Applying a gloss,* both studies, once again, using corpora of research articles.

Biber's paper on *Stance in spoken and written university registers* builds on his paper with Finegan in the earlier period detailing the lexico-grammatical marking of evidentiality and effect and is one of only three of the 15 most cited works which was actually published in the second period. Bhatia's book extends genre analysis beyond a

traditional focus on academic contexts to advertising, business, law and media texts and, like Myers' *Pragmatics of politeness* and Biber et al.'s *Grammar of Spoken and Written English*, has provided metadiscourse analysts with fruitful insights into the forms and functions of language use.

4.4 Most active journals citing metadiscourse papers

Our final question explores the journals which cite the 431 metadiscourse papers in our corpus. Using the Web of Science 'Cited Reference Search' function, we extracted the names of the citing journals, their disciplinary field and publication language and frequency of citations. Table 7 shows the source disciplines of the citing journals.

Table 7 Disciplinary categories of journals citing metadiscourse papers

Disciplinary categories	Journals	% of all
Linguistics/Applied linguistics	246	27.32
Communication	153	17.02
Education	133	14.77
Computer Science	124	13.75
Psychology	33	3.73
Literature	32	3.51
Sociology	23	2.56
Information Science / Library Science	21	2.32
Anthropology	19	2.08
Behavioral Sciences	16	1.81

As can be seen, while the term originated with Harris (1959), a structural linguist, it has been referred to and used in a wide range of other disciplines, not all of which are obviously language-focused.

Table 8 confirms that the majority of journals citing metadiscourse papers are in linguistics and indicates the main citing journals in this discipline. These are overwhelmingly the *Journal of English for Academic Purposes, English for Specific Purposes* and *Journal of Pragmatics*. The first two reflect the predominant metadiscourse focus on academic and professional texts and the third encourages papers which contribute to data-informed understandings of the language-context relationship. Several journals seek to span related disciplines such as linguistics and education

(*System*, *Assessing Writing*) and linguistics and communication (*Discourse Studies, Discourse, Context & Media*), while the list also contains a communication journal (*Communication Theory*).

Table 8   The journals most actively citing metadiscourse papers

Journals	citations	%
Journal of English for Academic Purposes	143	15.91
English for Specific Purposes	121	13.46
Journal of Pragmatics	107	11.90
Journal of Second Language Writing	48	5.34
Ibérica	42	4.67
Text & Talk	40	4.45
Communication Theory	37	4.12
System	30	3.34
Written Communication	27	3.00
Applied Linguistics	25	2.78
Discourse Studies	25	2.78
International Journal of Corpus Linguistics	22	2.45
Assessing Writing	21	2.34
Journal of Research in Applied Linguistics	20	2.22
Discourse Context & Media	19	2.11

Finally, it is worth observing that metadiscourse has attracted the interest of research in many languages other than English. Table 9 presents the languages of journals citing metadiscourse papers.

Table 9 Languages of journals citing metadiscourse papers

Languages	frequency	%
English	814	90.51
Spanish	55	6.12
Russian	9	0.99
Portuguese	5	0.58
French	4	0.44
Catalan	3	0.33
Chinese	3	0.33
Afrikaans	2	0.22
Dutch	2	0.22
German	2	0.22

English dominates the list in large part because the vast majority of journals listed in the web of Science are in English. Western countries, and particularly the United States, have been at the centre of scholarly publishing, supported by massive investments in research and the necessary scientific, technical and educational infrastructure. More recently, however, advances in technology, increased funding for research, more efforts by developing countries to raise the status of their universities, and more pressures on individual authors has meant greater participation of previously peripheral countries.

The enthusiasm with which Spanish authors have taken up metadiscourse is reflected in the fact that Spanish accounts for 6.1% of the citing journals. There are, in fact, a large number of comparative studies of metadiscourse between English and Spanish texts (e.g. Pérez-Llantada 2010) and between English texts written by English and Spanish authors (e.g. Mur Dueñas 2011). While *Google Scholar*, which trawls the web for dissertations and unpublished material as well as books and articles, shows 9% of the material written in Farsi and 5% in Chinese (Hyland 2017), the WoS data show Chinese to be the only Asian language to make to the list of citing journals. Metadiscourse has made a huge impact on researchers in China, indicating the growing importance of publishing in English and the need to understand the features of effective texts (Hu and Cao 2011).

5. Conclusions

Now one of the most commonly employed methods for analysing specialist written texts, metadiscourse is a concept which seems to have found its time. In this bibliometric study we have sought to track the emergence and rise of the concept from its early exploratory period to the substantial presence in discourse studies it occupies today.

Our data indicate a clear rise in the numbers of papers, citations, journals, disciplines and languages embracing the concept, with a steep increase since 2006. The WoS

database of published material shows that research heavily privileges written academic texts, particularly research articles, with a strong focus on comparing features of metadiscourse across disciplines, genres and languages. It is also the case that a great deal of the most cited work actually focuses on a specific feature of metadiscourse, particularly stance. There is a risk that the approach might remain too closely associated with the description of a limited range of text types and fail to realise its potential as a systematic means of gaining insights into interaction more generally. We have noted, however, an emerging literature exploring business texts and, to a lesser extent, visual metadiscourse, and an expanding interest in the pedagogical applications of metadiscourse. Methodologically, studies are understandably dominated by discourse-analytic procedures, particularly using corpora, although these are occasionally supplemented with the views of text users.

The list of the most cited authors shows a strong core of influential writers who have driven the concept forward and remain influential as the research expands beyond linguistics into new disciplines and other languages. It is clear that metadiscourse research is embedded in the weave of applied linguistics and discourse analysis more generally, with Swales, Bhatia and Myers all heavily cited. There is also a solid reliance on a foundation in functional linguistics (such as the work of Halliday), corpus linguistics (Biber and Hunston) and pragmatics (Austin).

One clear outcome of this study is that it appears the term 'metadiscourse' has largely, although not entirely, come to mean work conducted using a broad, interpersonal understanding of the term. While restricting metadiscourse to 'text reflexivity' makes it easier to identify cohesive and text organisational features, it removes much of what makes metadiscourse a powerful analytic tool. Overall, the term *metadiscourse*, then, seems to be cohering around a core conception of audience-awareness and the rhetorical features of texts which support a writer/speaker's communicative goals.

The study, of course, has the limitation of privileging the apex of published research by focusing on the authors, publications and citations of work indexed in the SSCI databases. These are the most available and celebrated sources, and we have neglected the mass of metadiscourse research found in regional, local and university publications, as well as studies within conference proceedings, dissertations, theses and academic blogs. This huge body of work, however, is difficult to systematically access and analyse and includes material of variable quality. It is, though, almost certainly influenced by the research we have discussed in this study and may be an interesting topic for further research, perhaps through *Google Scholar* sources.

We have, however, shown how the current interest in written interaction has emerged and grown over the past 30 years. Today this is a thriving area of research, grounded in the study of authentic texts and promising further insights into language use, the connections between communities and communicative behaviours, and how teaching might be best informed. Metadiscourse is, then, a concept which has inspired a considerable amount of scholarship and continues to contribute enormously to how language works as (meta)communication.

References //follow T&T house style strictly// //update all references//

Ädel, Annelie. 2006. *Metadiscourse in L1 and L2 English.* Amsterdam: John Benjamins.

Ädel, Annelie & Anna Mauranen. 2010. Metadiscourse: Diverse and divided perspectives. *Nordic Journal of English Studies* 9(2). 1-11.

Ahmad, Nisar, Amjad Naveed, Shabbir Ahmad & Irfan Butt. 2020. Banking sector performance, profitability and efficiency: A citation-based systematic literature review. *Journal of Economic Surveys* 34(1). 185-218.

Anthony, Laurence. 2019. *AntConc* (Version 3.5.8). Tokyo: Waseda University. http://www.antlab.sci.waseda.ac.jp/

Austin, John L. 1975. *How to do things with words 2nd ed*. Oxford: Oxford University Press.

Beauvais, Paul J. 1989. A speech-act theory of metadiscourse. *Written Communication* 6 (1). 11-30.

Cooper, Diane I. 2015. Bibliometrics basics. *Journal of the Medical Library Association* 103(4). 217–218.

Crismore, Avon. 1983. *Metadiscourse: What is it and how is it used in school and non-school social science texts.* Urbana-Champaign: University of Illinois.

Crismore, Avon. 1989. *Talking with readers: Metadiscourse as rhetorical act*. New York: Peter Lang.

D'Angelo, L., & Consonni, S. (2020). A tale of three waves: Or, concerning the history and theory of metadiscourse. *Iberica*, (40). 13-34.

De Groot, Elizabeth, Catherine Nickerson, Hubert Korzilius & Marinel Gerritsen. 2015. Picture this: Developing a model for the analysis of visual metadiscourse. *Journal of Business and Technical Communication* 30(2). 165-201

Ellegaard, Ole & Johan A. Wallin. 2015. The bibliometric analysis of scholarly production: How great is the impact? *Scientometrics* 105(3).1809-1831.

Fahimnia, Behnam, Joseph Sarkisb & Hoda Davarzani. 2015. Green supply chain management: A review and bibliometric analysis. *International Journal of Production Economics* 162. 101-114.

Fan, Jingchun, Ya Gao, Na Zhao, Runjing Dai, Hailiang Zhang, Xiaoyan Feng, Guoxiu Shi, Jinhui Tian, Che Chen, Brett D. Hambly & Shisan Bao. 2020. Bibliometric analysis on COVID-19: A comparison of research between English and Chinese studies. *Public Health* 8. 1-10.

Gillaerts, Paul & Freek Van de Velde. 2010. Interactional metadiscourse in research article abstracts. *Journal of English for Academic Purposes* 9(2). 128-139.

Halliday, Michael Alexander Kirkwood. 1994. *An introduction to functional grammar.* London: Edward Arnold.

Halliday, Michael Alexander Kirkwood. 1985. *Spoken and written language*. Oxford: Oxford University Press.

Harris, Zellig S. 1959. *Computable syntactic analysis: Transformations and discourse analysis papers*. Philadelphia: University of Pennsylvania.

Hu, Guangwei & Feng Cao. 2011. Hedging and boosting in abstracts of applied linguistics articles: A comparative study of English- and Chinese-medium journals. *Journal of Pragmatics* 43. 2795-2809.

Hyland, Ken 2004 *Disciplinary Discourses: social interactions in academic writing*. Ann Arbor: University of Michigan Press

Hyland, Ken. 2005. *Metadiscourse*. London: Continuum.

Hyland, Ken 2007. Applying a gloss: Exemplifying and reformulating in academic discourse. *Applied Linguistics* 28 (2), 266-285

Hyland, Ken 2017. Metadiscourse: what is it and where is it going? *Journal of Pragmatics*, *113*, 16–29.

Hyland, Ken & Feng K. Jiang 2018. "In this paper we suggest": changing patterns of disciplinary metadiscourse. *English for Specific Purposes*, *51*, 18–30.

Hyland, Ken & Feng K. Jiang. 2021a. A bibliometric study of EAP research: Who is doing what, where and when? *Journal of English for Academic Purposes*.

Hyland, Ken & Feng K. Jiang. 2021b. Delivering relevance: the emergence of ESP as a discipline. *English for Specific Purposes*. 64: 13-25

Hyland, Ken and Polly Tse 2004. Metadiscourse in scholastic writing: a reappraisal. *Applied Linguistics, 25*(2), 156–177.

Ifantidou, Elly. 2005. The semantics and pragmatics of metadiscourse. *Journal of Pragmatics* 37(9). 1325-1353.

Jiang, Feng K. & Ken Hyland 2016. Nouns and academic interactions: a neglected feature of metadiscourse. *Applied Linguistics*. 39 (4) 508–531

Lei, Lei & Dilin Liu. 2019. Research trends in applied linguistics from 2005 to 2016: a bibliometric analysis and its implications. *Applied Linguistics*, 40(3), 540–561.

Mauranen, Anna. 1993. Cultural differences in academic rhetoric: A textlinguistic study. Frankfurt am Main: Peter Lang.

McGrath, Lisa & Maria Kuteeva. 2012. Stance and engagement in pure mathematics research articles: Linking discourse features to disciplinary practices. *English for Specific Purposes* 31(3). 161–173.

Mur Dueñas, Pilar. 2011. An intercultural analysis of metadiscourse features in research articles written in English and in Spanish. *Journal of Pragmatics* 43(12). 3068-3079.

Oxford English Dictionary. 2021. https://www.lexico.com/definition/bibliometrics (accessed 10 September 2021)

Pérez-Llantada, Carmen. 2010. The discourse functions of metadiscourse in published academic writing: Issues of culture and language. *Nordic Journal of English Studies* 9(2). 41-68.

Schmid, Helmut. 1995. *TreeTagger*. http://www.cis.unimuenchen.de/~schmid/tools/TreeTagger/

Swales, John. 2019. The futures of EAP genre studies: A personal viewpoint. *Journal of English for Academic Purposes* 38. 75-82.

Thompson, Geoff. 2001. Interaction in academic writing: Learning to argue with the reader. *Applied Linguistics* 22(1). 58-78.

Vande Kopple, William J. 1985. Some exploratory discourse on metadiscourse. *College Composition and Communication* 26. 82-93.

Watzlawick, Paul, Janet Beavin and Don Jackson. 1968. *Pragmatics of Human Communication*. New York: Norton.

Williams, Joseph M. 1981. *Style: Ten lessons in clarity and grace*. Boston: Scott Foresman.

Zhang, Man. 2016. A multidimensional analysis of metadiscourse markers across written registers. *Discourse Studies* 18(2). 204-222.

Zhang, Xian. 2020. A bibliometric analysis of second language acquisition between 1997 and 2018. *Studies in Second Language Acquisition* 42(1). 199-222.