Interaction in written texts:
A bibliometric study of published research

Ken Hyland ✉

University of East Anglia, Norwich, United Kingdom
https://orcid.org/0000-0002-4727-8355
k.hyland@uea.ac.uk

Feng (Kevin) Jiang

Jilin University, Changchun, China
https://orcid.org/0000-0001-7369-9498
kevinjiang@jlu.edu.cn

Abstract

While writing involves interactions between writers and readers as each considers the other in creating and interpreting texts, research interest in written interaction is a fairly recent development. This paper uses a bibliometric analysis to trace the growing interest in written interaction over the past 30 years from its origins in philosophy, conversation analysis and sociocultural language pedagogy. To do so, we analyzed all 918 articles mentioning writing and interaction in the social science citation index since 1990, dividing the corpus into two periods following the massive increase in interest after 2005. We identify which topics have been most prevalent and which authors, publications, journals and countries most influential over time. The results indicate the growing importance of identity, genre, discipline, metadiscourse and stance, particularly drawing on corpus methods. We also note the participation of authors from more countries in publishing interaction research with the growth of authors from China becoming particularly visible. These findings may interest those working in written discourse analysis and scholarly publishing.

Keywords: bibliometrics; written interaction; research articles; scholarly publishing

1. Introduction

While most published work on English deals with the written language, this has only recently begun to address the concept of *interaction*. The ways we use language to cooperate with one another to build relationships, convey information and create our social worlds in written texts, however, is now a major area of study. It is with this body of research, or rather, with its changing preoccupations and predominant sources, that we are concerned with in this paper, both as a way of identifying major trends and of tracking the historical development of a key area of applied linguistics. To do this we adopt a bibliometric perspective to examine the literature on written interaction over the past 30 years (1990-2020), studying changes in its research topics (or most frequently occurring aspects of interaction), influential publications and authors, and its geographical sources. Specifically, we set out to answer the following questions:

1. What have been the most frequently explored topics of interaction and have these changed in frequency?
2. Which authors have been most influential and have these changed?
3. Which publications have been most influential and have these changed?
4. Which countries/regions have been most productive in contributing to this research over the period and have these changed?
5. Which journals have produced the most work in this area and have these changed?

2. Interaction in writing: The early days

It is a truism that writing involves interaction. Writers and readers consider each other, try to imagine each other's purposes and strategies, and write or interpret a text in terms of these imaginings. But this crucial aspect of communication largely evaded linguistic study until the 1980s with the emergence of a qualitative research agenda to describe and explain naturally occurring language use. The goal was to move beyond the long held view that written texts were just things, and to put the people back in. So, rather than regarding the linguistic features of written texts as representations of lexico-grammatical rules or regularities of style, analysts decided they could be examined as collective social practices: the attempts of writers to reach out to readers to achieve particular goals.

      The obvious place to look for inspiration at the beginning of this endeavor was the work in philosophy and sociology which had sought to model how people behaved in face-to-face conversations. Instead of examining texts themselves,

theorists built on the insights of speech act theory (e.g., Searle, 1975) and the notion of interaction rituals (Goffman, 1967) which suggested that interaction could be understood as sets of strategic practices and interpretive frames that facilitated joint actions among people. So conversation was taken as a starting point and theorists drew on simplified models of communication which assumed what people are like. Thus, for Grice (1975) successful interaction depended on participants presuming each other to be essentially cooperative, and both creating and interpreting utterances according to conformity or deviation from this principle. Similarly, Sperber and Wilson (1986) argued that successful interaction depended on participants finding relevance in the unfolding utterances of discourse.

Particularly influential in interactive writing research in the 1980s was Brown and Levinson's (1987) model of politeness. This draws on Goffman's (1967) notion of face to posit that interaction depends on the desire of individuals to have others see one in a favorable light, and to act unimpeded. Interaction for Brown and Levinson (1987) involves weighing up the interpersonal threats posed by various face threatening acts (FTAs), such as blaming, requesting, or apologizing. The model is based on face-to-face interaction as it involves calculating variables such as the imposition of the acts, the relationship between speaker and hearer and their relative power. It has, however, been adapted for the analysis of writing, either by focusing on texts where the writer and reader are in an identifiable relation to each other, as in letters (Cherry, 1988), or by seeing the audience as a composite of several possible types of reader, as in academic texts (Myers, 1989). This helps account for a writers use of questions to express a command or hedges to weaken statements (Hyland, 2004).

More directly related to interaction in writing was Nystrand's (1986) view that writers and readers approached a text with "mutual co-awareness" of the other following a "reciprocity principle" which is applicable to all social acts, and not just FTAs. For Nystrand (1986), an effective text is one which "balances the reciprocal needs of the writer for expression and the reader for comprehension" (p. 81) and where participants draw on certain 'elaborations' in texts to overcome interpretive difficulties. This view was influential in turning attention from the text to the context and the role of the social group's taken for granted conventions for effective interaction.

All this led to a more nuanced idea of *audience.* Ede and Lunsford (1984) introduced their notion of *audience invoked*, a creative construction of the writer, as opposed to *audience addressed* or real, concrete readers. This expanded the possibility of conceptualizing interaction with large unknown audiences which possessed community expectations for particular forms of engagement. With a more refined notion of audience, the door was open to recognizing the role of *community* in understanding written interaction. Thus, Faigley (1986) observed that writing "can be understood only from the perspective of a society rather than a single individual" (p. 535) and Geertz (1983) proposed that writing

depends on the actions of members of local communities. The notion of *discourse community* has therefore proved useful as it seeks to locate writers in particular contexts to identify how their rhetorical strategies are dependent on the purposes, setting, and audience of writing. Looking at writing in this way therefore evokes a social milieu which influences the writer and activates specific responses to recurring tasks.

These models of interaction, based as they are on how a reasoning writer or reader might smooth over communicative troubles, provide only a limited analysis, however. Another approach, starting with texts rather than general principles, emerged towards the end of the 1980s and focused on genres (Myers, 1990; Swales, 1990). These typically start with a detailed analysis of a range of empirically derived textual features and seek to explain form-function correspondences which create interaction.

Myers (1999) sees the analysis of written interaction as moving in two directions from the 1990s onwards, dematerializing the text to a set of features in relation to each other, or rematerializing texts to see them as things circulating in particular places and moments. He sees corpus analyses, such as Biber's (1988) factor analysis, as exemplifying the first and Canagarajah's (1996) account of how writing for publication in Sri Lanka, with its dependence on photocopiers, typewriter ribbons, and the postal service, as exemplifying the second.

> In either dematerializing texts and pursuing interaction as a bundle of linguistic features or rematerializing texts and pursuing interaction as a series of encounters situated in time and place, the powerful link between specific textual features and particular actions is made much more complex and less direct. Instead we have to make more tentative links, to say how this feature co-occurs with that one, or to say how this text works in this case. (Myers, 1999, p. 60)

This, then, has been the challenge of those researching the complex interactions represented in written texts. It is also the point at which we take up this story, using bibliometric techniques to analyze the most referenced topics, cited authors and sources relevant to the study of interaction in writing since 1990. We first describe our methods.

3. Methodology

3.1. Approach

Bibliometrics refers to "the application of mathematics and statistical methods to the analysis of academic publications" (Pritchard, 1969, p. 348). This is a

quantitative approach used in library and information sciences to describe patterns of publication within a given field or body of literature. It has helped to explore authorial networks (González-Alcaide et al., 2012), collaborations (Davarpanah & Aslekia, 2008), and gender inequalities in publication (Sebo et al., 2020). Most famously, the fact it facilitates the quantitative evaluation of publications, journals, and authors means that it informs the science citation index and is often used to study the research productivity of individual scholars and countries (Liu et al., 2015; Ma et al., 2014).

Generally, however, bibliometric studies have not sought to characterize particular areas of research. Pioneering exceptions are Lei and Liu (2019), Zhang (2020), and Hyland and Jiang (2020). 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, while Hyland and Jiang (2020) traced the emergence of English for academic purposes (EAP) over the past 40 years. This paper follows this general line of research while taking a narrower focus.

3.2. Corpus

The first step in answering our research questions was to create a corpus of journal articles. We chose to use the Web of Science (WoS) as our starting point as this allows searches of its 79 million papers from the SSCI journals. To do this, we had to ensure a robust search string to narrow down our corpus to relevant texts. We approached this task by searching for all occurrences of the terms "writing" and "interaction" anywhere in the same titles, abstracts and sets of keywords in all the journals in the core collection of SSCI journals. From this group we then limited the search to those journals in the WoS categories of linguistics, education and communication. This returned 918 journal articles from 1990 to 2020. Figure 1 shows the distribution of these publications across 30 years.
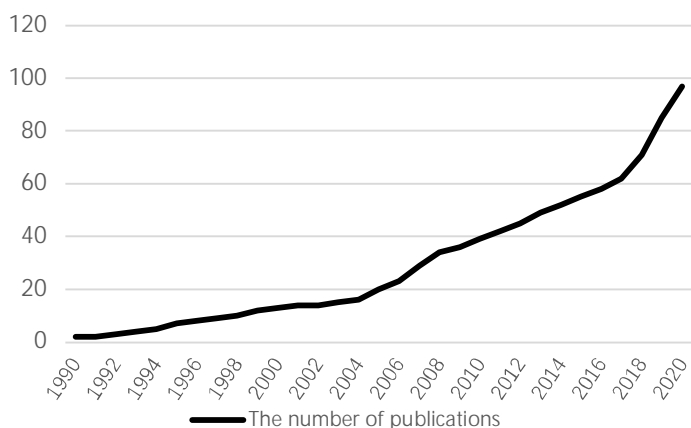
Figure 1 Distribution of the number of publications across time

As can be seen, there is a marked increase in the number of papers published on interaction in writing over the period, with a significant surge from 2005. This seems to be a watershed in academic interest in interaction in writing, a critical moment after which research in this area took off. We therefore decided to make this a break point and trace changes before and after this date. There are 126 articles in the first period 1990 to 2004 and 792 in the second from 2005 to 2020.

3.3. Data searches

We then interrogated these two corpora to answer the questions above, 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 books, chapters and articles in the two periods;
4) the most productive countries over the two periods;
5) the most active journals publishing and citing papers on interaction in writing.

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

1. Annotated all the identified abstracts with part-of-speech and lemma information[1] using Schmid's (1995) *TreeTagger* programme.
2. Searched the tagged corpora to identify all nouns and *n*-grams of 2 to 5 words using *AntConc* (Anthony, 2022) to identify candidate topics.

---

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

3. Filtered the n-grams using an automatic process (using stop words) to exclude function words (modals, pronouns, etc.) which do not occur in research topic phrases.
4. Manually checked all remaining cases to exclude phrases which did not constitute research topics.

We focused on only nouns in identifying potential topic areas as these are far more likely to constitute research subjects than other parts of speech, while 2-5 word *n*-grams seemed a realistic range to capture topics such as *peer feedback* (2-word gram) and *English as an additional language* (5-word gram). In cases where two or more terms referred to the same topic (e.g., "bundles" and "high frequency collocations"), we included only the more specific and most commonly occurring term. We then narrowed this down further by discarding three main categories: (a) words and clusters common in language in general (*aim of* and *in depth*), (b) concepts and issues not specific to interaction in writing (*analysis of, significant difference, the study*), and (c) concepts and issues common in writing but too general to be useful (e.g., *meaning, written language*) (see also Lei & Liu, 2019).

To ensure that we included only the most salient topics, we set a minimum threshold frequency of 10 occurrences per item (from different journal sources). While this involves some subjectivity, it was guided by the studies of Lei and Liu (2019), and Hyland and Jiang (2020), which set a threshold frequency of 30 occurrences. This was deemed high enough to ensure the significance of the selected items but not too high as to exclude important topics. Because our study addressed a more specialized area and involved a smaller corpus (918 compared to 10,000 articles), we reduced the target criterion to items which occurred at least 10 times. This produced an appropriate number of meaningful instances. We also added a range criterion, so that items had to occur in 10% of all the publications to ensure a reliably widespread appearance of a topic in the literature. We then normalized the frequencies by representing the number of occurrences of each topic per 100 papers. This was to allow comparisons between the huge disparities of papers in the two time periods. Finally, we ran a one-way *chi*-square test on the raw frequency for each of the topics across the two periods to determine statistical difference in the topic frequencies.

To address questions 2 and 3, concerning the most influential authors and publications, we first identified the most-cited authors in the references of the papers in each corpus using a frequency count. To find the most highly cited works, we part of speech tagged the reference lists then computed the frequencies of all the publications (books, chapters, and articles) from these lists in the corpus, using a regular expression search to identify the abbreviated titles.

Question 4, regarding the most productive countries in publishing papers on interaction in writing, we extracted the affiliation of every author of every paper in the

corpus. To answer question 5, concerning the journals publishing research on interaction in writing, we used a *Web of Science* facility to locate all the citations the identified papers had gained and generated a report of the journal sources of these citations.

4. Results and discussion

4.1. Most frequent interaction-relevant topics

The criteria discussed in section 3.3 (at least 10 occurrences in at least 13 papers in the first period and 80 or more in the second) produced 260 frequently discussed research topics related to interaction in writing over the 30 years. 179 new topics (68.8% of the total), were introduced between 2005 and 2020 as the field expanded, including *blog*, *workplace*, *voice* and *L3*. Table 1 shows the topics which statistically rose, fell and remained constant most significantly overall, with figures given for each period. The topics are organized by their percentage change in normed frequency.

Table 1 Most frequently discussed topics in interaction in writing (raw/per 100 papers)

	1990-2004		2005-2020		% normed	*chi-*	*p*
	raw	normed	raw	normed	change	value	
Significantly up							
identity	12	9.5	186	23.5	147.4	9.45	.00
genre	9	7.3	139	17.6	141.1	6.83	.00
discipline	10	8.0	150	18.9	136.3	7.11	.00
metadiscourse	11	8.7	160	20.2	132.2	7.35	.00
stance	10	8.1	149	18.8	132.1	6.87	.00
corpus	13	10.3	174	22.0	113.6	6.93	.00
engagement	13	10.3	172	21.7	110.7	6.66	.00
socialization	14	11.1	181	22.9	106.3	6.76	.00
assessment	23	18.3	298	37.6	105.5	11.08	.00
Significantly down							
intervention	49	38.9	42	5.3	-86.4	126.7	.00
classroom	11	8.7	11	1.4	-84.1	25.0	.00
peer feedback	18	14.3	25	3.2	-77.9	29.6	.00
test	10	7.9	18	2.3	-71.4	11.8	.00
ESL	22	17.5	42	5.3	-69.6	23.8	.00
revision	11	8.7	21	2.7	-69.6	11.9	.00
essay	18	14.3	40	5.1	-64.6	15.2	.00
writer	50	39.7	114	14.4	-63.7	40.4	.00
Constant							
discourse	29	23.0	179	22.6	-1.8	0.2	.73
audience	10	7.9	63	8.0	0.8	0.9	.33
literacy	20	15.9	126	15.9	0.2	0.6	.91
reader	13	10.3	82	10.4	0.8	0.9	.34
L2 learners	40	31.7	251	31.7	-0.1	0.3	1.45

It is clear that several of these topics have shown dramatic increases in popularity over the years, with *identity, genre, discipline, metadiscourse* and

stance rising over 130%. These are areas of interaction which have clearly captured researchers' imaginations, while issues concerning interaction in writing instruction have fallen heavily, with *intervention, classroom, peer feedback* and *testing* all showing substantial declines in research interest. Some of these popular topics have been rising more generally in applied linguistics and EAP (Hyland & Jiang, 2020) but *metadiscourse, stance, engagement* and *socialization* are specific to the growing interest in interaction. *Genre* and *corpus*, of course, are not studied as topics per se, but appear on the list as a result of their frequent association with studies of writing in interaction. This points to the popularity of text-oriented studies and the dominance of what Myers (1999) characterized as "dematerializing texts" discussed above.

Table 2, which shows topic changes over the period which are notable without being statistically significant, perhaps suggests that interest in pedagogy has not disappeared entirely from research on interaction in writing. *Graduate student, collaborative writing* and *writing activity* might indicate a shift of perspective in teaching to incorporate interaction among writers and between writers and readers. Social media and computer mediated writing are also in this list, indicating the impact of the web in providing teachers with qualitatively different types of instructional possibilities, enabling students to engage in collaborative multimedia authoring rather than only mediated interaction with others. From the mid-2000s, Web 2.0 technologies have redefined the Web as a social platform for collaboration, knowledge-sharing and networking, replacing the *informational web* with the *social web* (Pegrum, 2009, p.18). The decline in traditional concerns with accuracy, error and proficiency underscores this shift towards making-meanings interactively online.

Table 2 Main topics showing notable but not significant change (normed to 100 papers)

	1990-2004		2005-2020		% normed change	Chi-value	*p*
	raw	normed	raw	normed			
Increase							
computer mediated	12	9.3	88	11.1	19.5	0.3	.66
gender	9	7.4	70	8.8	18.9	0.2	.58
graduate student	10	7.9	74	9.4	19.0	0.2	.67
social media	9	6.9	66	8.3	20.3	0.2	.70
collaborative writing	9	7.5	70	8.9	18.7	0.2	.58
writing activity	13	10.1	97	12.3	21.8	0.4	.60
syntactic complexity	11	8.9	85	10.7	20.2	0.3	.56
Decline							
accuracy	9	7.5	49	6.2	-17.5	0.4	.66
error	8	6.5	41	5.2	-20.4	0.4	.57
interview	12	9.8	65	8.2	-16.3	0.4	.60
language proficiency	13	10.2	67	8.5	-17.1	0.5	.48
writer	11	8.9	57	7.2	-19.1	0.5	.52

4.2. Most frequently cited authors

This growing interest in interactive writing can also be seen in the changes in influential authors. To identify these we conducted a frequency count of the reference lists in our corpus of 918 journal articles, divided them into the two distinct periods, and normalized the results to 100 papers. Table 3 lists the top 15 cited authors in these papers.

What may be most surprising about the two lists in Table 3 is there are so few names which appear in both periods. Hyland, van Lier, Swales and Vygotsky are the hardy perennials, casting their influence over the entire 30 years. Hyland and Swales have pioneered work which emphasizes the importance of interaction in disciplinary knowledge construction in EAP. Hyland has done this through numerous publications presenting models of metadiscourse and stance and engagement, while Swales' considerable influence has been through a focus on genre and human textographies. Van Lier's main contributions have been in second language learning and curriculum development, but he appears here as a result of his stress on the central role of interaction in language learning and teacher development. Vygotsky, of course, is well-cited for his sociocultural theory and the idea that human development is a socially mediated process where children acquire their cultural values and beliefs through interaction with others. This view has been the foundation of considerable research and theory in cognitive development, learning and written interaction.

Table 3 Most highly cited authors across the two periods

1990-2004			2005-2020		
Authors	raw	normed	Authors	raw	normed
Merrill Swain	73	57.9	Ken Hyland	228	28.8
Ken Hyland	52	41.3	Leo van Lier	226	28.5
Neomy Storch	40	31.7	John Heritage	189	23.9
Lev Vygotsky	39	31.0	Douglas Biber	126	15.9
Alison Mackey	31	24.6	Emanuel A. Schegloff	171	21.6
Amy Snyder Ohta	28	22.2	Lev Vygotsky	169	21.3
Mark Warschauer	26	20.6	M.K.A. Halliday	145	18.3
Gordon Wells	26	20.6	James Martin	132	16.7
Rod Ellis	23	18.3	Charles Goodwin	130	16.4
Susan Gass	23	18.3	Erving Goffman	102	12.9
Leo van Lier	22	17.5	William Labov	95	12.0
Penelope Brown	21	16.7	Carolyn R. Miller	93	11.7
James P. Lantolf	21	16.7	John Swales	92	11.6
John Swales	21	16.7	Gunther Kress	78	9.8
Richard Donato	20	15.9	Deborah Tannen	77	9.7

Note. Frequencies normed to 100 papers

In the earlier period there is a clear emphasis on citing writers who have brought interaction to writing through their work on pedagogy and learning. Some of the big names in the development of English language teaching figure prominently here including Merrill Swain, Rod Ellis, Susan Gass, James Lantolf, and Amy Snyder Ohta. Swain, for example, sees interaction as both a means of communication and a cognitive tool, promoting the importance of learners producing output so they can learn from gaps in their efforts to make meaningful contact. Gass advocates the interaction hypothesis, arguing that through input and interaction language learners can notice differences between their own formulations of the target language and the language of their conversational partners. Ellis similarly believes that social and intermental interaction are major forces in the acquisition of an L2. Lantoff and Snyder Ohta are known for their applications of Vygotsky's sociocultural theory to second language learning and promoting group interaction. Alison Mackey, on the other hand, is recognized for conversational interaction and Gordon Wells for his studies of literacy development in young children. All of these scholars, in different ways, stressed the importance of interaction in a learner's linguistic development.

While principally concerned with spoken interaction, the absence of any prior work on interaction in writing meant that this work was influential in informing later work on writing. Penelope Brown's work with Stephen Levinson on face was mentioned earlier for example. Research on writing appears here not only in the work of Swales and Hyland, but also that of Naomi Storch, who has championed work on collaborative writing, and by Mark Warschauer, the pioneer of interactive writer-reader engagement though Web-enabled platforms. It should be born in mind, however, that while citations take time to emerge, work on interaction in writing remained limited in this period and the raw citations are quite low. Serious interest in the topic did not accelerate until the second period of this analysis.

The variety of names, and specialized fields, in the latter period testify to both the diversity of interest in written interaction and the interdisciplinary influences on the research contributing to its understanding. The list shows a mixture of linguists (e.g., Labov, Halliday, Tannen), applied linguists (van Lier, Hyland, Biber, Martin, Swales), sociologists (Heritage, Goffman, Schegloff), semioticians (Kress, Goodwin), a psychologist (Vygotsky) and a rhetorician (Miller). Clearly theories and models founded on spoken registers continue to support research in written interaction after 2005, particularly those inspired by conversation analysts such as Heritage and Schegloff and the dramaturgical analyses of Erving Goffman. Conversation analysis (CA) has proved useful in describing task- and institution-centered interactions in a range of spoken encounters from medical consultations to classrooms. But while providing important insights into how participants orientate to each other, its exclusive focus on the behavior of interacting contributors limits its role in characterizing writing. The wider context remains a sealed box

and analysts ignore the constitutive elements of tasks and settings, such as disciplinary conventions, personal backgrounds, and task constraints.

Studies of spoken communication also inform written interaction research through Deborah Tannen's work on gender differences in communication style and how the language of everyday conversation affects relationships, as well as Labov's theory describing the influence of social variables on communication. The semioticists Charles Goodwin and Gunther Kress have also contributed to this literature, Goodwin by showing how the professionals demonstrate their expertise through discursive practices such as coding, highlighting, and graphical representation, to effectively interact with others and help them see phenomena of interest, and Kress through work on multimodality and an emphasis on the integration of forms of communication. Work by the linguists on the list is more directly relevant to writing. Citations to Halliday and Martin recognize the massively influential contribution of systemic functional linguistics which has helped researchers of interaction understand how lexicogrammar contributes to creating meaning and shaping relationships through texts. From a different direction, Biber's work on stance, the grammar of spoken and written English and of register variation using corpora has significantly informed both our understanding of interaction in writing and the methods we use to study it.

4.3. Most influential publications

We assume here that the most influential publications are those which receive the most citations, thus indicating the interest that writers see in the earlier work and its relevance to their own research. There are, however, difficulties with this as citation frequencies may be distorted by self-citation or negative citations, although their influence seems to be relatively minor (e.g., Hyland, 2003). We also have to be aware that not only are more works cited as research expands across the two periods, but that older publications have more time to accrue citations. But while our lists may fail to capture the impact of more recent publications, our results show how the field has changed and the influence of key publications in this. Tables 4 and 5 show the raw and normed frequencies of the top 15 most cited publications for each period.

Table 4 shows how researchers drew on a broad range of texts in the earlier period, with the importance of interaction in learning together with texts focusing on social interaction in academic and computer-mediated contexts. It is worth noting that the 30 most cited titles are dominated by books, with 9 in the first period and 12 in the second, illustrating the longevity of their influence. Articles comprise a small proportion, with two in each list, despite the immediate

visibility they provide to authors and the priority given them by institutions because of government assessment exercises.

Table 4 Most highly cited publications from 1990 to 2004

Publications	Cites	Normed
Swain, M. (2000). The output hypothesis and beyond: mediating acquisition through collaborative dialogue. In J. P. Lantolf (Ed.), *Sociocultural theory and second language learning* (pp. 97-114). Oxford University Press.	12	9.5
Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Harvard University Press.	12	9.5
Hyland, K. (2000). *Disciplinary discourses: Social interactions in academic writing*. Longman.	9	7.1
Swain, M. (1985). Communicative competence: Some roles of comprehensible input and comprehensible output in its development. In S. Gass & C. Madden (Eds.), *Input in second language acquisition* (pp. 235-253). Newbury House.	9	7.1
Stanley, J. (1992). Coaching student writers to be effective peer evaluators. *Journal of Second Language Writing, 1*(3), 217-233.	7	5.6
Grabe, W., & Kaplan, R. B. (1996). *Theory and practice of writing: An applied linguistic perspective*. Pearson.	6	4.8
Swain, M. (1995). Three functions of output in second language learning. In. G. Cook & G. Seidhofer (Eds.), *Principles and practices in applied linguistics: Studies in honor of H. G. Widdowson* (pp. 125-144). Oxford University Press.	6	4.8
Warschauer, M. (1996). Comparing face-to-face and electronic discussions in the foreign language classroom. *CALICO Journal, 13*(2), 7-26.	6	4.8
Pellettieri, J. (2000). Negotiation in cyberspace: The role of chatting in the development of grammatical competence. In M. Warschauer & R. Kern, (Eds.), *Network-based language teaching: Concepts and practice* (pp. 59-86). Cambridge University Press.	5	4.0
Swales, J. (1990). *Genre analysis: English in academic and research settings*. Cambridge University Press.	5	4.0
Bakhtin, M. (1981). *The dialogic imagination: Four essays*. University of Texas Press.	5	4.0
Hyland, K. (1998). *Hedging in scientific research articles*. John Benjamins.	5	4.0
Bereiter, C., & Scardamalia, M. (1987). *The psychology of written composition*. Lawrence Erlbaum.	4	3.2
Berkenkotter, C., & Huckin, T. N. (1995). *Genre knowledge in disciplinary communication: Cognition/culture/power*. Lawrence Erlbaum.	4	3.2
Block, D. (2003). *The social turn in second language acquisition*. Georgetown University Press.	4	3.2

Merrill Swain has three titles in the list, all book chapters, confirming the role that interaction has played in language pedagogy and its contribution to understanding writing. Vygotsky's *Mind and Society* and Hyland's *Disciplinary Discourses* top the book citations. The former a theoretical tour de force of cognitive development which shows the interactional relationships between humans and their surroundings, the latter is a series of studies describing the ways language is used to construct academic disciplines and scholarly identities. The enduring influence of Vygotsky's work can be seen by the fact that it also sits at the top of the list in the second period, head and shoulders above the second most cited work. The book is actually an edited collection of Vygotsky's key essays setting out his theory that social interaction within the family and with knowledgeable members of the community is the primary means by which children acquire relevant behaviors and cognitive processes.

The theoretical importance of social interaction in language learning proposed by Vygotsky is reinforced, in different ways, in the publications by Bereiter and Scardamalia and by Block, both of whom stress the significance of social involvement in the acquisition of language skills. More practical treatments regarding interaction in pedagogy are found in Stanley and in Swain. The Grabe and Kaplan book is an overview of text linguistic research, process writing and contrastive rhetoric, linking theory to concrete writing tasks and revealing the salience of interaction. The third major strand here is computer-mediated/internet interactions and, once again, interaction in writing informed by research in spoken discourse (Pellettieri) and by comparison of both modes (Warschauer).

Bakhtin's *Dialogic Imagination* introduced the key ideas of heteroglossia, polyglossia and intertextuality, thus transforming the study of writing (and language more generally) to emphasize context, hybridity and the connection between utterances. Finally there are studies of interaction in negotiating scholarly arguments found in Hyland's *Disciplinary Discourses* and *Hedging in Scientific Research Articles*, Swales' *Genre Analysis* and Berkenkotter and Huckin's *Genre Knowledge in Disciplinary Communication.* Each of these have had an enormous influence in showing how even the most apparently faceless genres only function successfully through writer-reader interactions.

Table 5 Most highly cited publications from 2005 to 2020

Publications	cites	normed
Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Harvard University Press.	64	8.1
Hyland, K. (2005). *Metadiscourse: Exploring interaction in writing*. Continuum.	37	4.7
Heritage, J. (1984). A change-of-state token and aspects of its sequential placement. In J. M. Atkinson & J. Heritage (Eds.), *Structures of social action: Studies in conversation analysis* (pp. 299-345). Cambridge University Press.	27	3.4
Storch, N. (2013). *Collaborative writing in L2 classrooms*. Multilingual Matters.	27	3.4
Bakhtin, M. (1986). *Speech genres and other late essays*. University of Texas Press.	21	2.7
Labov, W. (1972). *Language in the inner city: Studies in the black English vernacular*. University of Pennsylvania Press.	21	2.7
Hyland K. (2002). *Teaching and research writing*. Pearson.	21	2.7
Hyland, K. (2005). Stance and engagement: A model of interaction in academic discourse. *Discourse Studies, 7*(2), 173-192.	18	2.3
Hyland, K. (2008). Genre and academic writing in the disciplines. *Language Teaching, 41*(4), 543-562.	18	2.3
Biber, D. (1988). *Variation across speech and writing*. Cambridge University Press.	17	2.1
Swales, J. (1990). *Genre analysis: English in academic and research settings*. Cambridge University Press.	17	2.1
Halliday, M. A. K., & Hasan, R. (1976). *Cohesion in English*. Longman.	16	2.0
Kress, G. (2003). *Literacy in the new media age*. Routledge.	15	1.9
Fairclough, N. (1992). *Discourse and social change*. Polity Press.	15	1.9
Berkenkotter, C., & Huckin, T. N. (1995). *Genre knowledge in disciplinary communication: Cognition/culture/power*. Lawrence Erlbaum.	14	1.8

In Table 5 we see the most highly cited publications in the last 15 years. About half of the authors mentioned in section 5 appear here, although the influence of other writers may have been dispersed over a greater number of publications. Nor were the high citations for some works sufficient to push their writers onto this list. While the raw citation numbers may seem relatively low, it should be remembered that the figures refer only to: (a) the relevance of these publications to work on interaction in writing, and (b) cited in SSCI listed journals. It is probable that they have also received considerable attention in other domains.

The titles in Table 5 show how interaction in writing has, by 2020, largely established itself as a distinct field with a much clearer focus on written language, on discourse analysis and on the role of interaction in sense making. Gone are the papers on language acquisition and interaction in speech, replaced by more discourse analytic and conceptual works. All but 3 of the 15 titles do not appear in the earlier list, with the work by Vygotsky, Swales and Berkenkotter and Huckin demonstrating their enduring importance. Bakhtin and Hyland also appear in both lists but with different titles than earlier. Bakhtin's *Speech Genres and Other Late Essays* deals more explicitly with the genres of everyday life and with language as a living dialogue, thus inspiring work in social construction and genre studies. Hyland has four titles in the list, all distinct from those in the earlier period. *Metadiscourse* and *Stance and Engagement* present the models for studying interaction in academic writing for which he is best known while *Genre and Academic Writing in the Disciplines* shows how various features of academic writing function to facilitate interaction, and *Teaching and Research Writing* presents a more general view of writing.

The crossover from spoken interaction is not completely missing in this table, however. The presence of Heritage's paper on the placement of the particle *oh* in conversation and Labov's study of Black English vernacular shows how perspectives from both CA and sociolinguistics endure in feeding into the study of writing. Storch's book on *Collaborative Writing in L2 Classrooms* continues the thread to interactions in writing pedagogy and reminds us of the spoken and written exchanges that surround writing. Biber's ground-breaking use of computational techniques both helps to characterize 23 spoken and written genres and reveals the various linguistic features which contribute to interaction in them. Kress' work contributes to the research on interaction by exploring the social and cultural effects of writing as it moves from the page to the screen and Fairclough's critical framework has widened the research on interaction to include the impact of social and political factors. Halliday and Hassan's classic treatise on *Cohesion in English*, by setting out a model for studying the resources for relating parts of texts, helps connects textual and interactional aspects of communication by showing how writers enhance a reader's comprehension.

4.4. Most productive countries

Scholarly publishing has been dominated for many years by institutions in the West, and particularly the United States, as a result of huge financial investments in education and research infrastructure, high salaries, and conditions conducive to scholarly networking. With increasing globalization, developing and non-metropolitan countries have sought to raise the visibility of their universities and gain a foothold in the global "knowledge economy" through scholarly publication. As a result, UNESCO (2017) reports that emerging economies have largely been behind the 4-5% annual growth in publishing output in recent years. There are, then, now more journals, more scholarly papers, more publishers, more co-authorship and more academics writing in a language which is not their native tongue (Hyland, 2015).

The affiliations of authors in our data only partially indicate this general expansion and do not reflect the considerable volume of submissions and papers now appearing from countries such as Iran, India, and Malaysia. No doubt the impact of this work will be seen over the next decade. Table 6 shows the 15 most productive publishing countries over the period based on the affiliation of every author in the corpus.

Table 6 Most productive countries across the 30 years (by author affiliation)

1990-2004			2005-2020		
Country	No.	%	Country	No.	%
USA	68	54.0	USA	275	34.7
UK	11	8.7	UK	87	11.0
Canada	10	7.9	Mainland China	65	8.2
Australia	7	5.6	Australia	46	5.8
Germany	7	5.6	Canada	42	5.3
Israel	6	4.8	Spain	42	5.3
Hong Kong, China	6	4.8	Taiwan	34	4.3
Italy	5	4.0	Germany	25	3.2
Mainland China	5	4.0	Hong Kong, China	25	3.2
Brazil	4	3.2	France	24	3.0
Denmark	4	3.2	Sweden	23	2.9
Norway	4	3.2	Belgium	21	2.7
Scotland	4	3.2	Finland	21	2.7
Spain	4	3.2	New Zealand	21	2.7
Finland	3	2.4	Italy	20	2.5

As can be seen, authors working in the USA, England, Australia and Canada continue to dominate work in this field over the entire period, although the proportion of papers originating in the USA has declined markedly. The smaller percentage of the papers provided by the top 15 to the total, moreover, indicates a wider range of participating countries further down the most recent list.

Canada, Germany, Hong Kong and Italy have all fallen, both in position and the proportion they have added to the whole, while Israel, Brazil, Denmark, Norway and Scotland have dropped out of the top 15 entirely. Taiwan, France, Sweden, Belgium and New Zealand have all entered the more recent list and England, Australia, Spain and Finland have increased the proportion of work to the sum.

Most strikingly, Mainland China has emerged as a major participant in this area, as in its contribution to scholarly publishing more generally. The main reason for this is huge government investment in research in recent years (e.g., Zhang et al., 2013), with expenditure rising almost 33 times between 1995 and 2013 (National Bureau of Statistics of China, 2015). One result has been an increase in articles in SCI journals written by authors from Chinese institutions from 120,000 in 2009 to 450,000 in 2019 (Mallapaty, 2020). Xie and Freeman (2019) even suggest that Chinese authors, from anywhere in the world, co-authored 34.5% of all papers published in 2016 – a 22% increase since 2000. In the case of contributions to written interaction, Table 6 shows that China jumped from 9th to 3rd place over the period and doubled the proportion of work it adds to the topic.

Our results, of course, are based on author affiliations rather than nationality and we cannot confirm a writer's country of origin, first language, or ethnic background from this. In today's scholarly diaspora, academics are scattered around the world in institutions outside their home countries. The findings suggest, however, something of the interest that the topic has generated among academics around the world.

4.5. Most productive journals

Finally, we analyzed our corpus to determine the journals which had published the most articles relating to interaction in writing. It should be remembered that articles form only a small percentage of the most cited sources discussing writing and interaction (Tables 4 and 5). However, given the importance of journals in disseminating research, it is likely that articles are likely to be widely read, if not always cited. It is also worth noting that the ranking in Table 7 is likely to be skewed by the volume of papers each journal publishes rather than the high proportion of work they might include on written interaction. So, the *Journal of Pragmatics* and *Computers in Human Behavior*, for example, are published monthly, and so publish a larger number of papers overall. Many of these papers relate to written interaction, so placing the former at the top of the list and pushing the latter into it. Other journals, such as *TESOL Quarterly* and *English for Specific Purposes*, publish only four issues annually, although they may have a higher proportion of papers on interaction.

Table 7 Journals publishing most papers on interaction in writing (1990-2020)

Journals	Number	% of all
Journal of Pragmatics	262	2.5
System	200	1.9
Computer Assisted Language Learning	191	1.8
Modern Language Journal	171	1.6
Journal of Second Language Writing	164	1.5
Journal of English for Academic Purposes	151	1.4
Language Learning & Technology	140	1.3
English for Specific Purposes	128	1.2
TESOL Quarterly	119	1.1
Foreign Language Annals	118	1.1
Language Teaching Research	106	1.0
Applied Linguistics	92	0.9
Text & Talk	81	0.8
Discourse Studies	81	0.8
Computers in Human Behavior	79	0.7

Interestingly, the journals listed in Table 7 cover a broad swathe of the applied linguistic spectrum. Some, like *System* and *Computer Assisted Language Learning*, are largely aimed at practitioner readership, while others provide more theoretical and challenging treatments (*Applied Linguistics, Modern Language Journal*). There are journals which address computer and technology in communication (*Language Learning & Technology* and *Computers in Human Behavior*) and those that focus on particular areas of the field, such as pragmatics, discourse, second language writing, and specialist varieties of English.

The range of journals also suggests the breadth of interest in the topic of interaction in writing and the different ways it is being studied. Academics working in pragmatics, computer communication, writing instruction, second language acquisition, EAP and discourse studies have come to find this an interesting and productive area of research.

5. Conclusions

In this study we have used bibliometric methods to track the emergence of an area of applied linguistics and language pedagogy which has grown considerably in recent years. In a corpus of SSCI articles from 1990 to the present, we have found a marked rise in research output after 2005 which continues to this day. This surge may be due to the publication of a number of influential books around that time. Those by Hyland (2005) and Ädel (2006) on metadiscourse, Martin and White's (2005) book on evaluation using the appraisal model, and the re-issue of Hyland's (2004) book on disciplinary discourses with a new publisher seem to have captured

the zeitgeist. These works coincided with greater interest in written discourse analysis more generally and a growing body of work in English for academic purposes, seeking to describe the characteristics of various scholarly written genres.

Among our main findings we note the growing importance of identity, genre, discipline, metadiscourse and stance in the terms referred to and in the work of authors and texts cited. Clearly, research focusing on interaction in academic written genres continues apace, particularly that drawing on corpus methods, while more classroom oriented studies on the interactions among collaborating writers and the acquisition of writing competence has declined. The earlier period is characterized by work which either focuses on interaction as a cognitive and social tool of language pedagogy or which explores how the interactive features of spoken texts can be shaped to explain written texts. The more recent 15 years have seen an explosion of work on interactive written discourse while work on interaction in social media and computer-mediated writing has also become more prominent and might be expected to influence research in the future. We also note the participation of authors from more countries in publishing interaction research, with the growth of authors from China becoming particularly visible.

There are, of course, limitations to our study. For one thing, it privileges the apex of published work by focusing on the authors, publications and citations of work indexed in the SSCI databases. These are the most available and celebrated works, and we have neglected the majority of research in this area found in regional, local and university publications, or which never sees the light of publication at all. This no doubt comprises a huge body of work, but it is difficult to systematically access and analyze. This wider body of research, however, is also almost certainly influenced by the research we have discussed in this study and may be an interesting topic for further research. We have, however, shown how the current interest in written interaction has emerged and grown over the past 30 years, building on the work of natural philosophers, speech act theorists, conversation analysts, composition researchers and others. Today this is a thriving area of research, grounded in the study of authentic texts and promising to provide further insights into language use, the connections between communities and communicative behaviors, and the teaching practices by which these behaviors can be nurtured.

References

Ädel, A. (2006). *Metadiscourse in L1 and L2 English.* John Benjamins.

Anthony, L. (2022). *AntConc* (Version 4.2.0) [Computer Software]. Waseda University. https://www.laurenceanthony.net/software

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

Brown, P., & Levinson, S. (1987). *Politeness: Some universals in language usage.* Cambridge University Press.

Canagarajah, A. S. (1996). "Nondiscursive" requirements in academic publishing: Material resources of periphery scholars, and the politics of knowledge production. *Written Communication, 14*(4), 435-472.

Cherry, R. (1988). Politeness in written persuasion. *Journal of Pragmatics, 12*, 63-81.

Davarpanah, M., & Aslekia, S. (2008). A scientometric analysis of international LIS journals: Productivity and characteristics. *Scientometrics, 77*(1), 21-39.

Ede, L., & Lunsford, A. (1984). Audience addressed/audience invoked: The role of audience in composition theory and pedagogy. *College Composition and Communication, 35*(2), 155-171.

Faigley, L. (1986). Competing theories of process: A critique and a proposal. *College Composition and Communication, 48*, 527-542.

Geertz, C. (1983). *Local knowledge: Further essays in interpretive anthropology.* Basic Books.

Goffman, E. (1967). *Interaction ritual.* Anchor Books.

González-Alcaide, G., Jinseo, P., Huamaní, C., Gascón, J., & Rincón, J. (2012). Scientific authorships and collaboration network analysis on Chagas disease: Papers indexed in PubMed (1940-2009). *Revista do Instituto de Medicina Tropical de São Paulo, 54*, 219-228.

Grice, P. (1975). Logic and conversation. In P. Cole & J. Morgan (Eds.), *Syntax and semantics* (pp. 41-58). Academic Press.

Hyland, K. (2003). Self-citation and self-reference: Credibility and promotion in academic publication. *Journal of American Society for Information Science and Technology, 54*(3), 251-259.

Hyland, K. (2004). *Disciplinary discourses: Social interactions in academic writing.* University of Michigan Press.

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

Hyland, K. (2015) *Academic publishing: Issues and challenges in the production of knowledge.* Oxford University Press.

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

Hyland, K., & Jiang, K. (2020). "This work is antithetical to the spirit of research": An anatomy of harsh peer reviews. *Journal of English for Academic Purposes, 46*, 100867. https://doi.org/10.1016/j.jeap.2020.100867

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

Liu, W., Hu, G., Tang, L., & Wang, Y. (2015). China's global growth in social science research: Uncovering evidence from bibliometric analyses of SSCI publications (1978-2013). *Journal of Informetrics*, *9*, 555-569.

Ma, F., Li, Y., & Chen, B. (2014). Study of the collaboration in the field of the Chinese humanities and social sciences. *Scientometrics*, *100*, 439-458.

Mallapaty, S. (2020). China bans cash rewards for publishing papers. *Nature, 579*(7797), 18. https://doi.org/10.1038/d41586-020-00574-8

Martin, J. R., & White, P. R. (2005). *The language of evaluation: Appraisal in English*. Palgrave.

Myers, G. (1989). The pragmatics of politeness in scientific articles. *Applied Linguistics, 10*, 1-35.

Myers, G. (1990). *Writing biology: Texts in the social construction of scientific knowledge.* University of Wisconsin Press.

Myers, G. (1999). Interaction in writing: Principles and problems. In C. Candlin & K. Hyland (Eds.), *Writing: Texts, processes and practices* (pp. 40-61). Longman.

National Bureau of Statistics of China. (2015). *China statistical yearbook on science and technology* (1996-2014). China Statistics Press.

Nystrand, M. (1986). *The structure of written communication: Studies in reciprocity between writers and readers.* Academic Press.

Pegrum, M. (2009). *From blogs to bombs.* UWA Publishing.

Pritchard, A. (1969). Statistical bibliography or bibliometrics? *Journal of Documentation*, *25*, 348-349.

Schmid, H. (1995). *TreeTagger software* [Computer software]. http://www.cis.unimuenchen.de/~schmid/tools/TreeTagger/

Searle, J. (1975). Indirect speech acts. In P. Cole & J. L. Morgan (Eds.), *Speech acts* (pp. 59-82). Brill.

Sebo, P., Maisonneuve, H., & Fournier, J. P. (2020). Gender gap in research: A bibliometric study of published articles in primary health care and general internal medicine. *Family Practice*, *37*(3), 325-331.

Sperber, D., & Wilson, D. (1986). *Relevance: Communication and cognition.* Blackwell.

Swales, J. (1990). *Genre analysis: English in academic and research settings.* Cambridge University Press.

UNESCO. (2017). *Science report: Towards 2030.* https://en.unesco.org/unesco_science_report/

Xie, Q., & Freeman, R. B. (2019). Bigger than you thought: China's contribution to scientific publications and its impact on the global economy. *China & World Economy*, *27*(1), 1-27.

Zhang, H., Patton, D., & Kenney, M. (2013). Building global-class universities: Assessing the impact of the 985 Project. *Research Policy, 42*, 765-775.

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