Scientometrics (2023) 128:4541–4558
https://doi.org/10.1007/s11192-023-04759-6

Promoting research by reducing uncertainty in academic writing: a large-scale diachronic case study on hedging in *Science* research articles across 25 years

Mingxin Yao[1] · **Ying Wei**[1] · **Huiyu Wang**[2]

Received: 26 October 2022 / Accepted: 26 May 2023 / Published online: 31 May 2023

Abstract
Hedges are important in academic writing since they indicate uncertainty and tentativeness about academic knowledge. However, few studies explore how hedges have changed in academic writing overtime. Among the existing studies, there is also divergent understandings. The current case study traced the diachronic development of hedges that express doubt and uncertainty in the full texts of *Science* research articles from 1997 to 2021. Our findings show that the use of such hedges has significantly decreased in the past 25 years in the research articles of the Journal *Science*. In addition, we propose that the drop of such hedges in *Science* research articles may be an implicit writing strategy for research promotion, and therefore may correlate with the rising linguistic positivity in academic writing. Our hypothesis was initially confirmed by the significant correlation between the evolution of hedges and Yuan and Yao's (Scientometrics 127:1–17, 2022) sentiment scores in academic writing. Our findings may reveal a bigger picture of promoting research by adopting not only explicit strategies such as more positive language (Yuan and Yao in Scientometrics 127:1–17, 2022) but also subtle and implicit writing strategies such as reducing uncertainty. Finally, we discussed the implications of this study for peer reviewers, editors, and researchers.

Keywords Hedges · Academic writing · *Science* · Linguistic positivity bias · Research promotion

Introduction

Hedges are abundant in science and play a critical role in academic writing (Hyland, 1998, p. 6), since they help researchers express tentativeness (Mur-Duenas, 2021) or subjective uncertainty (Hyland, 1998, p. 166) in their research articles to mitigate their scientific

✉ Ying Wei
ying_wei1213@163.com

[1] School of Foreign Studies, Nanjing University, Nanjing 210023, Jiangsu, People's Republic of China

[2] Faculty of Foreign Languages, Harbin Institute of Petroleum, Harbin 150028, Heilongjiang, People's Republic of China

claims (Martín, 2008) or avoid threatening peer readers and express modesty toward them (Myers, 1989). Given its significance in academic writing, hedges in research articles have received sufficient attention and thorough examination (Dontcheva-Navratilova, 2016; Fløttum et al., 2006; Hu & Cao, 2011; Hyland, 1996, 1998, 2005; Martín, 2005, 2008; Mur-Dueñas, 2016, 2021; Myers, 1989; Orta, 2010; Poole et al., 2019; Salager-Meyer, 1994; Salager-Meyer et al., 2003; Varttala, 1999, 2001; Vassileva, 2001; Yang, 2013). However, there is few diachronic explorations on hedges and its change in academic writing. Existing diachronic research also yields divergent insights. Furthermore, no study has concerned the possibility of using hedges as a writing strategy in relation to the current trend of promoting research with various strategies, such as a rising linguistic positivity (i.e., Cao et al., 2021; Liu & Zhu, 2023; Vinkers et al., 2015; Wen & Lei, 2022b; Yuan & Yao, 2022), an increasing proportion of hyping words (i.e., Hyland & Jiang, 2021a, 2021b; Millar et al., 2019), and a growing trend of confidence level (Wheeler et al, 2021), given the growing competitiveness in publication. Therefore, in the current study, we intend to trace the diachronic change of hedging in the academic writing of *Science* research article full texts, and also incorporate our analysis into the bigger picture of research promotion by studying the correlation between hedging and Yuan and Yao's (2022) linguistic positivity bias.

Research background

Hedging

In general, there are three lines of research concerning hedges or hedging in academic writing, namely a single-disciplinary approach, a cross-disciplinary approach, an intercultural approach. First, the studies on hedges with a single-disciplinary approach (e.g., Hyland, 1996; Salager-Meyer, 1994; Varttala, 1999) mainly address how hedges are used in particular disciplines, such as Medicine (Salager-Meyer, 1994; Varttala, 1999), Molecular Biology (Hyland, 1996), Clinical and Health Psychology (Martín, 2008), and Business Management (Mur-Dueñas, 2016; Orta, 2010). For instance, Salager-Meyer (1994) examines hedging in different sections in academic writing based on a corpus of 15 articles collected from five leading medical journals, and finds that shields, approximators, and compound hedges are the most commonly used hedges. The study also finds that hedges mostly appear in the Discussion and Comment sections, while the Methods and Case Report sections are the least-hedged sections in academic writing of Medicine.

Second, the cross-disciplinary approach of studying hedges in academic writing compares the differences in using hedges in research articles from a variety of disciplines (e.g., Hyland, 2005; Poole et al., 2019; Varttala, 2001), such as the comparison of hedging in research articles from economics, medicine, and technology (Varttala, 2001), or articles from biology, chemistry and physics (Poole et al., 2019). These studies contribute to revealing the cross-disciplinary difference of using hedges or metadiscourse in general, such as Hyland (2005)'s study, in which an analysis of metadiscourse is performed in 240 research articles from eight disciplines. His findings suggest that writers in humanities and social sciences adopt more hedges than hard sciences such as engineering.

Third, taking the intercultural approach, quite a number of studies compared the use of hedges in research articles in different languages (e.g., Dontcheva-Navratilova, 2016; Fløttum et al, 2006; Hu & Cao, 2011; Martín, 2005, 2008; Mur-Dueñas, 2016, 2021; Orta, 2010; Salager-Meyer et al., 2003; Vassileva, 2001; Yang, 2013). Such comparisons of

hedging in academic writing often involve two or more languages, such as the evaluation between English and other languages including Chinese (Hu & Cao, 2011; Yang, 2013), Spanish (Martín, 2005, 2008; Mur-Dueñas, 2021; Salager-Meyer et al., 2003), French (Fløttum et al, 2006; Salager-Meyer et al., 2003). In addition, comparisons of hedging are also made in academic writing in English as a first language and English as Lingua Franca (Mur-Dueñas, 2016), or between English as an L1 and as an L2 by speakers of a variety of language backgrounds such as Bulgarian (Vassileva, 2001), Spanish (Orta, 2010), Chinese (Hu & Cao, 2011; Yang, 2013), and Czech (Dontcheva-Navratilova, 2016). These studies contribute to our understanding about hedging in the academic writing across languages and cultures by revealing the linguistic or even cultural difference in the use of hedging in research articles. For instance, English academic writing is found to be more heavily hedged than other languages like Spanish (Martín, 2005; Mur-Dueñas, 2021; Orta, 2010; Salager-Meyer et al., 2003), Bulgarian (Vassileva, 2001), French (Salager-Meyer et al., 2003), and Chinese (Hu & Cao, 2011; Yang, 2013). Such differences, according to Mur-Dueñas (2021) could possibly be attributed to the factor that research articles written in English often deal with the international readership, and hence face more risks of criticism or even rejection.

So far, these three lines of research mainly adopt a synchronic approach by studying or comparing the use of hedges in academic writing at one point in time (Dontcheva-Navratilova, 2016; Fløttum et al, 2006; Hu & Cao, 2011; Hyland, 1996, 1998, 2005; Martín, 2005, 2008; Mur-Dueñas, 2016, 2021; Myers, 1989; Orta, 2010; Salager-Meyer, 1994; Varttala, 1999, 2001; Vassileva, 2001; Yang, 2013). Few studies trace the diachronic development of hedges in academic writing (e.g., except for Hyland & Jiang, 2016; Poole et al., 2019; Salager-Meyer et al., 2003) to further explore how hedging has changed in academic writing over time. For instance, Poole et al. (2019) study the diachronic development of epistemic stance markers (i.e., hedges and boosters) in a corpus of 328 open-access biochemical research articles between 1972 and 2017 and find that epistemic stance markers expressing certainty had increased over time while those indexing doubt and uncertainty had decreased, which contradicts what Hyland and Jiang (2016) discover in their study of research articles from four academic disciplines from the years of 1965, 1985 and 2015. Most recently, Wheeler et al. (2021) trace the diachronic development of certainty in the academic psychology writing in a corpus of 790, 520 psychology journal article abstracts in a time-span of 46 years. Their study reveals a growing trend of certainty, further contradicting Hyland and Jiang (2016).

However, to date, the number of studies that trace the diachronic development of hedges is still limited. In addition, the findings generated by the previous studies also raise some interesting and controversial debts over whether scientific writers have adopted more (un) certainty in academic writing over time (Hyland & Jiang, 2016; Poole et al., 2019; Wheeler et al., 2021). Therefore, this study also intends to join the debate and offer some insights into the understudied and debated research of the diachronic change of hedges in academic writing, and finally shed light on whether academic writing is becoming more certain or less uncertain (Hyland & Jiang, 2016; Poole et al., 2019) by focusing on the use of a certain type of hedges (i.e., hedges that express uncertainty and doubt, or reliability hedges according to Hyland, 1998) in a corpus based on *Science* research articles across 25 years.

Finally, the size of corpora used in existing studies is still limited. For instance, Poole et al.'s (2019) study is based on a corpus of 328 open access research articles on a single topic in biology. With a larger scale, Hyland and Jiang (2016) studied the change of hedging as a part of changing attitude in a corpus of 2.2 million words collected from top five journals from four disciplines. Most recently, Wheeler et al. (2021) analyzed the change

of confidence in a database of 790,520 psychology journal article abstracts. Their corpora are either limited in size or to abstracts. Hence, existing findings and disputes should be testified and settled by incorporating corpora of a larger size. Therefore, the current study intends to study how hedging has changed by adopting a large-scale diachronic corpus (Yuan & Yao, 2022). Our corpus is significantly larger in size than those adopted in previous studies, containing a total of 11 million words (see "Data collection" for more description).

Research promotion strategies

Academic publishing has developed to be more competitive and challenging (i.e., Cao et al., 2021; Hyland & Jiang, 2021a, 2021b; Liu & Zhu, 2023; Vinkers et al., 2015; Wen & Lei, 2022b; Wheeler et al., 2021; Yuan & Yao, 2022). Meanwhile, writers are found to adopt various writing strategies to make their research articles more desirable or acceptable in the face of the rising competition in academic publishing. Currently, studies *observe and discover three potentially distinct writing strategies* including a rising trend of linguistic positivity (i.e., Cao et al., 2021; Liu & Zhu, 2023; Vinkers et al., 2015; Wen & Lei, 2022b; Yuan & Yao, 2022), a tendency to use more hyping words (i.e., Hyland & Jiang, 2021a, 2021b; Millar et al., 2019), and a growing level of confidence in academic writing (Wheeler et al., 2021). In this section, we introduce the three strategies of research promotion revealed by existing studies.

First, the linguistic positivity bias, which refers to the tendency to use more positive words than negative ones in human communication, is noticed and detected in the research abstracts and full texts of a variety of disciplines (i.e., Cao et al., 2021; Liu & Zhu, 2023; Vinkers et al., 2015; Wen & Lei, 2022b; Yuan & Yao, 2022). Such trend of favoring positive words over negative ones in academic writing is revealed by scholars who adopt different approaches. Some scholars discover linguistic positivity bias in academic writing based on a small word list of 25 positive words and 25 negative words (i.e., Cao et al., 2021; Vinkers et al., 2015). Others reveal a rising positive sentiment in academic writing by adopting sentiment analysis based on large sentiment lexicons (i.e., Liu & Zhu, 2023; Wen & Lei, 2022b; Yuan & Yao, 2022). For instance, Cao et al. (2021) examine the use of 25 positive and 25 negative words in 2, 200,710 open access articles in 9,106 journals on PubMed Centra from 1997 to 2019, and find increasingly more positive words than negative words in both the abstracts and full texts. Yuan and Yao (2022) conduct sentiment analysis based on larger sentiment lexicons on *Science* research articles between 1997 and 2021 and reveal an increasing more positive sentiment scores in their data. Such appetite for positive words in academic writing is attributed to the need to promote one's research as publishing grows to be increasingly more competitive (Cao et al., 2021; Liu & Zhu, 2023; Vinkers et al., 2015; Wen & Lei, 2022b; Yuan & Yao, 2022).

Second, researchers also find a tendency to "hype" research articles (i.e., Hyland & Jiang, 2021a, 2021b; Millar et al., 2019). "Hyping", as noted by Millar et al., (2019: 139), refers to "the tendency to use hyperbolic and/or subjective language to glamorize, promote and/or exaggerate aspects of research". With a close examination of hyping words in research articles, it is found that authors are more likely to adopt exaggerative and glamorous words in research articles (i.e., Hyland & Jiang, 2021a, 2021b; Millar et al., 2019). For instance, Hyland and Jiang (2021b) trace the use of 400 "hyping" words, which seek to promote, embellish or exaggerate aspects of research papers, in a corpus of 360 articles taken from leading journals of four disciplines over 50 years. They find that hard sciences

show a great increase of hyping that displays positive attitude and underlines contribution, and that authors are more willing to hype their research primacy, methods, as well as their prior research. Such phenomenon is recognized as a way of rhetorically "selling" one's study in a time when research publishing comes under intense pressure (Hyland & Jiang, 2021b, p. 182).

Third, while linguistic positivity and hyping exhibit explicitness in research promotion since they mark an apparently positive linguistic presence in research articles, Wheeler et al. (2021) notice a more implicit strategy for research promotion, namely a growing level of confidence in academic writing. Specifically, Wheeler et al. (2021) calculate the confidence level in research abstracts of psychology and discovered a growing confidence level in the academic writing of psychology abstracts overtime. Such strategy is more implicit and subtle since authors do not explicitly make positive statements in academic writing but instead epistemically show more confidence about their research. It is also called for more caution and vigilance for over-confidence in academic writing (Wheeler et al., 2021).

Taken together, the above studies reveal a bigger picture of promoting research using various writing strategies, both explicitly and implicitly. However, no study considered the possibility of using hedges in academic writing to promote research. As previously discussed, hedges express uncertainty about the author's scientific claims. Given the heavy competition in the publishing industry and the current trend of research promotion by using different strategies, we hypothesize that hedges that express uncertainty would also have experienced a significant decline in research articles. Finally, we approach this hypothesis by studying how hedging has changed in *Science* research articles across 25 years.

If our hypothesis is confirmed, we intend to further propose a new strategy for research promotion, that is, reducing uncertainty in academic writing. To identify it as a new strategy, we also test whether there is a possibly significant correlation between the changing trends of hedging and the above strategies. To address the above knowledge gaps and test the hypotheses we proposed, we put forward two research questions as follows:

RQ1 How have the normalized frequencies of hedges in full texts of *Science* research articles developed in the last 25 years?
RQ2 Does the development of hedges in full texts of *Science* research articles correlate with the evolution of its linguistic positivity?

Methodology

This section introduces the methodology of the current paper by presenting the process of data collection and data analysis.

Data collection

In the current study, we collected research articles published in *Science* journal in the past 25 years (i.e., from January 1997 to December 2021) for our corpus investigation. Note that only research articles were collected as the data for this study out of the following considerations. To begin with, *Science* accepts and publishes a variety of manuscripts including research articles, reports, reviews, and commentaries. The reason why we focus on the category of research articles is that it may avoid the possible influence of such factors as genre differences caused by different types of manuscripts. Secondly, it ensures a good

Table 1 Descriptive specification of our full text corpus

Year	Number of articles	Number of words in articles	Mean word count per year
1997	35	148,055	4230
1998	45	174,071	3868
1999	44	169,413	3850
2000	59	227,902	3862
2001	67	252,738	3772
2002	67	248,088	3702
2003	60	226,624	3777
2004	60	228,106	3801
2005	72	266,953	3707
2006	64	236,720	3698
2007	58	226,335	3902
2008	57	211,828	3716
2009	69	275,520	3993
2010	67	268,991	4014
2011	75	292,817	3904
2012	54	213,215	3948
2013	90	386,548	4294
2014	126	551,805	4379
2015	114	507,989	4456
2016	149	669,659	4494
2017	172	780,369	4537
2018	192	875,819	4561
2019	258	1,179,668	4572
2020	278	1,256,443	4519
2021	320	1,463,817	4574

consistency of the diachronic research data since this category has not changed in the last 25 years. It is worth noting that reports were excluded from this study due to their differences in paper length and academic significance from those of research articles. To be specific, the former mainly introduces major scientific progress across disciplines with up to 4500 words while the latter, which simply reports important research results, is somehow limited to 2500 words.

Next, we describe how data is collected. In this research, a self-designed crawler script in Python was coded for data collection, resulting in a large-scale diachronic corpus of academic writing composed of *Science* research articles (full texts). The access to the *Science* website and hence these full texts were provided by Nanjing University Library. In case of bringing any potential pressure to the web server, the crawl delay in Python was adjusted to 10, which is in conformity to the Robots Exclusion Standard issued by the journal website. Finally, the retrieved data were stored in the format of Excel files to form our corpus. It should be noted that no full texts exceed the limit of an Excel cell, which stands at 32,767 characters.

The result of our data collection was a corpus containing 2652 full texts of *Science* research articles, with a total of 11,339,493 words (see Table 1), which is sufficient enough

for a diachronic investigation on the use of hedges to reflect how academic writing has evolved over time.

Data analysis

For data analysis, our study adopts a quantitative corpus analysis (Cao et al., 2021; Vinkers et al., 2015; Wen & Lei, 2022b; Yuan & Yao, 2022) by calculate the quantitative change of hedges in the *Science* corpus in a time-span of 25 years. Such analysis is conducted through running tests and calculating the normalized frequencies of hedges in the software Lancsbox and testing the statistical significance in R with homemade R scripts. Finally data visualization is made possible through Python scripts. Next, we describe how our methodology handles each research question.

RQ1 To answer the first research question, our data analysis contains two steps, namely a corpus linguistic analysis and a statistical analysis. The former provides this research with detailed linguistic variations of certain hedges in terms of normalized frequencies, while the latter reveals the statistical significance of such variations.

Since our study intends to explore the diachronic development of uncertainty in academic writing, it is important to select a list of hedges that expresses uncertainty, or what Hyland (1998) termed as "reliability hedges", for investigation before corpus analysis. Specifically, we referred to Mur-Duenas (2021)'s list of hedges in academic writing. Mur-Duenas (2021) compares the use of hedges in English and Spanish research articles by examining seven hedging categories, including modal verbs, semi-modal verbs, lexical verbs, adverbs, adjectives, nouns and phrases.

However, in our study, we only focus on the categories of hedging modal verbs, hedging semi-modal verbs, hedging adverbs and hedging adjectives, which are probably the most typical realizations of certainty (Biber & Finegan, 1989; Caffi, 2007, p. 142). First, modal verbs and modal adverbs are the typical linguistic devices that regulate the truth value, reliability, correctness, or validity of what is expressed, and therefore are the typical linguistic realizations of evidentiality (Caffi, 2007, p. 142). Second, Biber and Finegan (1989) note that doubt adjectives are likely to be employed in more indicative, structured, and integrated planned written texts. Therefore, it can be arguably said that modal verbs, adverbs and adjectives are possibly the most typical linguistic realizations of when expressing uncertainty. Table 2 displays the list of hedging modal verbs, hedging semi-modal verbs, hedging adverbs and hedging adjectives studied in this research.

Next, we fed our corpus into the corpus analysis software Lancsbox developed by Lancaster University to examine the normalized frequency (frequency per 10 k words) of the above listed hedges in each year. The normalized frequencies of each category is achieved by summing up the normalized frequencies of each items in the category. Finally, we coded a local R script to perform simple linear regression (Cao et al., 2021; Wen & Lei, 2022a, 2022b; Yuan & Yao, 2022) to investigate the developmental trajectories of the above hedges in *Science* research articles (full texts) across a 25 year time-span, in terms of their normalized frequencies per year.

RQ2 To address the second research question, we ran Pearson's product-moment correlation analyses to triangulate our results on the use of hedges in the full texts of *Science* research articles with the findings of Yuan and Yao's (2022) study, which explores the

Table 2 List of hedges examined in this study

Category	Words
Modal verbs	May, would, can, might, could
Semi-modal verbs	Appear to, seem
Adverbs	Relatively, typically, likely, potentially, perhaps, not necessarily, somewhat, usually, commonly, nearly, partially, just, almost, theoretically, probably, marginally, roughly, approximately, basically, maybe, normally, ostensibly, partly, presumably, virtually, tentatively, sometimes, slightly, cautiously
Adjectives	Likely, potential, possible, hypothesized, common, unlikely, indicative, typical, apparent, feasible, presumed, probable, proposed, unclear, prone to

Note that we removed "quite" and "often" from Mur-Duenas (2021)'s list of adverbs and "indicative", "typical", "apparent", "common" from Mur-Duenas (2021)'s list of adjectives since these words do not express doubts or uncertainty. Other than these excluded words, all the words in Mur-Duenas (2021)'s list of hedging modal verbs, hedging semi-modal verbs, hedging adverbs and hedging adjectives are adopted for data analysis

diachronic development of linguistic positivity in *Science* research articles (full texts) between 1997 and 2021. In detail, the normalized frequencies of hedging modal verbs, hedging semi-modal verbs, hedging adverbs and hedging adjectives in 25 years were compared with the standardized sentiment scores across 25 years in Yuan and Yao's (2022) study.

Results

This section reports some major findings concerning the development of hedges in *Science* research articles across 25 years.

RQ1 The development of hedges in *Science* research articles across 25 years.

The development of all types of hedges

In general, the use of all types of hedges combined exhibited a significant downward trend over the past 25 years, as displayed in Fig. 1. The detailed statistics, which is presented in Table 3, shows that the result is statistically significant: ($F(1, 23) = 63$, $p = 4.905e{-}08$, Multiple $R^2 = 0.73$, Adjusted $R^2 = 0.72$). Specifically, the normalized of all types of hedges combined had dropped over 58% from 1997 to 2021, from its highest point standing at 115.80 in 1997 to the lowest 67.42 by 2021. In the following sections, we present the detailed development of each type of hedges.

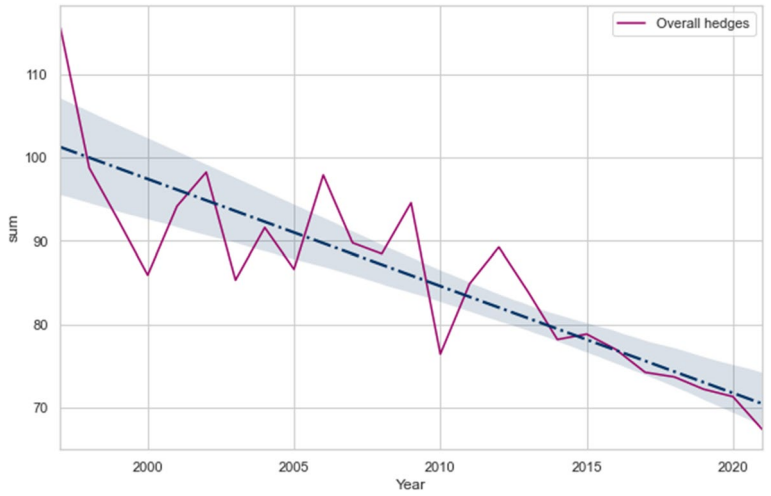

Fig. 1 Trajectory of all hedges in *Science* research articles (full texts) across 25 years

Table 3 Statistics of simple linear regression for overall hedges

Model	Variable	Estimate	Standard error	*t*-value	*p*-value
All types of hedges	(Intercept)	2659.9460	324.2930	8.202	2.79e−08***
	Year	−1.2813	0.1614	−7.938	4.91e−08***

*$p<0.05$; **$p<0.01$;***$p<0.001$

The development of hedging modal verbs and semi-modal verbs

First, our results show that the normalized frequency of both hedging modal verbs and hedging semi-modal verbs have declined significantly over the past 25 years, as shown in Figs. 2 and 3. Specifically, for modal verbs, the normalized frequency of hedging modal verbs went from its highest point ($n=56.57$) in 1997 to the lowest ($n=34.18$) in 2021. For semi-modal verbs, the normalized frequency went from 2.34 in 1997 to 0.63 in 2021.

Results of simple linear regression show that the trajectories of both modal verbs and hedging semi-modal verbs have witnessed significant decline within the past 25 years, with the statistics demonstrated in Table 4 (Modal verbs, $F(1, 23)=32.63$, $p=8.112e{-}06$, Multiple $R^2=0.59$, Adjusted $R^2=0.59$; Semi-modal verbs, $F(1, 23)=74.94$, $p=1.079e{-}08$, Multiple $R^2=0.77$, Adjusted $R^2=0.75$).

To be more specific, the normalized frequencies of "may", "would" and "can" have experienced major decrease, as displayed in Fig. 4, with that of "may" dropping from 15.31 in 1997 to 8.90 in 2021. Meanwhile, the normalized frequency of "would" dropped from 10.25 in 1997 to 4.10 in 2021, as demonstrated in Fig. 4. Finally, the normalized frequency of "can" plummeted from 19.60 in 1997 to 11.62. Note that the use of "might" and "could" have also declined to some degree, but such change is not as remarkable as that of the former three words.

Fig. 2 Trajectory of modal verbs in *Science* research articles (full texts) across 25 years

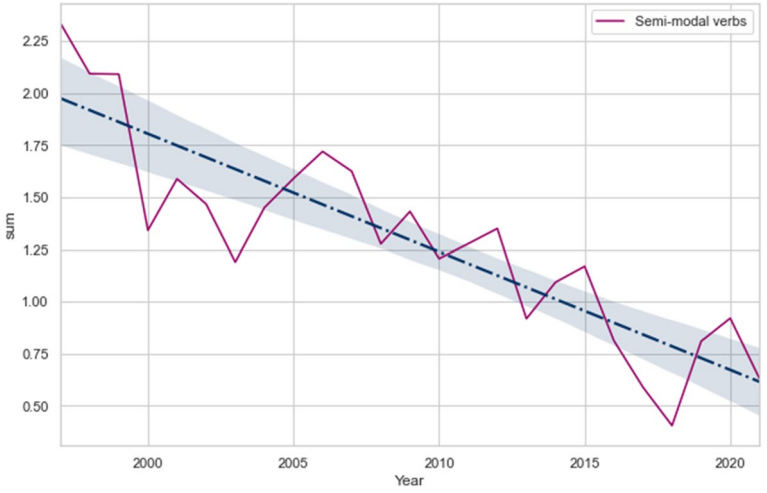

Fig. 3 Trajectory of semi-modal verbs in *Science* research articles (full texts) across 25 years

Table 4 Statistics of simple linear regression for modal verbs and semi-modal verbs

Model	Variable	Estimate	Standard error	*t*-value	*p*-value
Modal verbs	(Intercept)	1043.70985	175.48582	5.948	4.60e−06***
	Year	−0.49899	0.08735	−5.713	8.11e−06***
Semi-modal verbs	(Intercept)	115.123173	13.149164	8.755	8.82e−09***
	Year	−0.056659	0.006545	−8.657	1.08e−08***

*$p < 0.05$; **$p < 0.01$;***$p < 0.001$

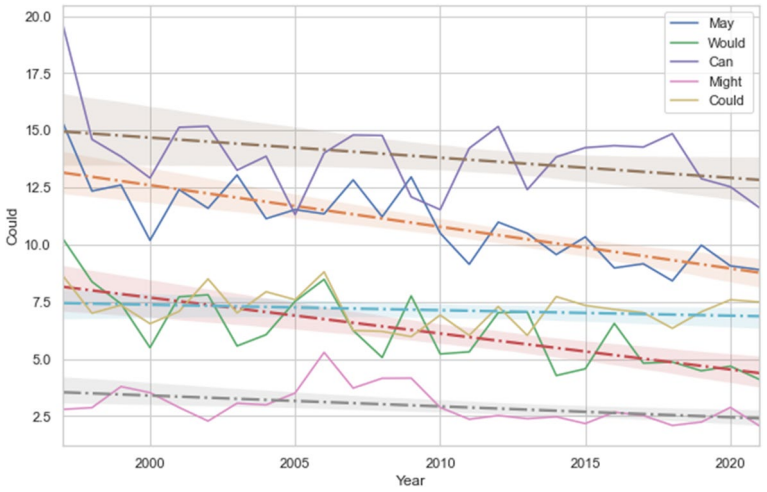

Fig. 4 Trajectory of all hedging modal verbs over time

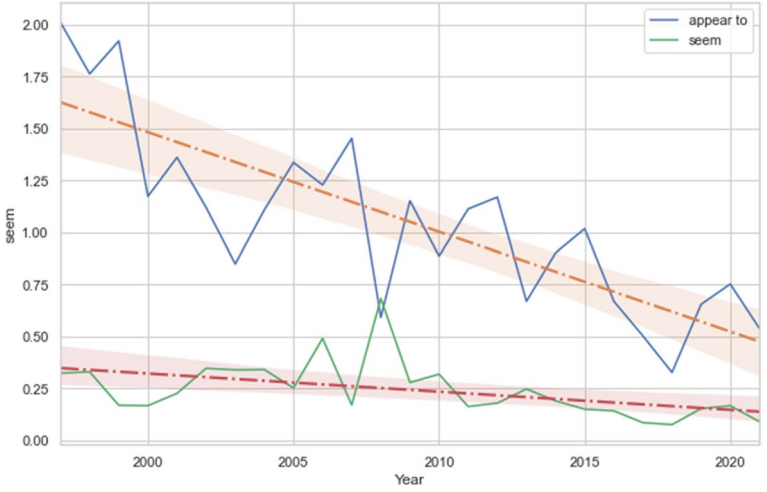

Fig. 5 Trajectory of "appear to" and "seem" over time

Likewise, the normalized frequencies of semi-modal verbs have also plunged over time. Figure 5 shows that the normalized frequency of "appear to" has significantly decreased from 2.01 in 1997 to 0.75 in 2021 and that of "seem" slightly dropped from 0.32 in 1997 to 0.09 in 2021, though with some fluctuation.

The development of hedging adverbs

Another simple linear regression was performed to trace the diachronic change of hedging adverbs. The trajectory of hedging verbs presented in Fig. 6 revealed a significant

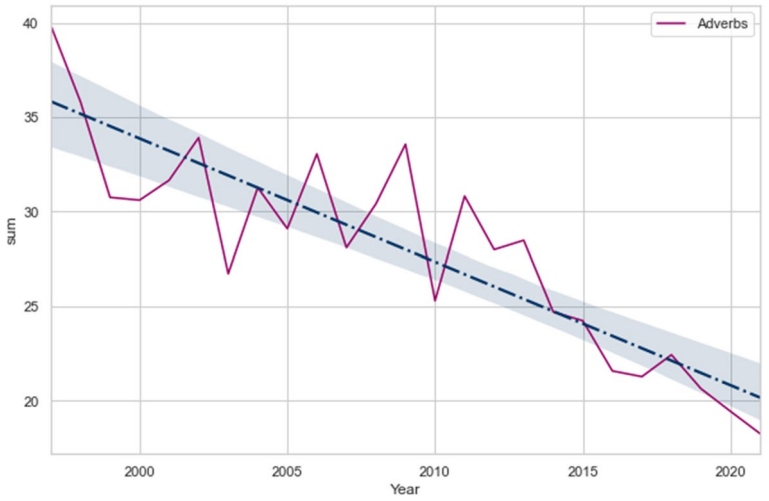

Fig. 6 Trajectory of hedging adverbs in *Science* research articles (full texts) across 25 years

Table 5 Statistics of simple linear regression for hedging adverbs and adjectives

Model	Variable	Estimate	Standard error	*t*-value	*p*-value
Hedging adverbs	(Intercept)	1338.74782	149.98560	8.926	6.23e−09***
	Year	−0.65244	0.07466	−8.739	9.11e−09***
Hedging adjectives	(Intercept)	224.06171	52.41095	4.275	0.000284***
	Year	−0.10411	0.02609	−3.991	0.000576***

*$p<0.05$; **$p<0.01$;***$p<0.001$

downward trend. The detail statistics were demonstrated in Table 5 (Hedging adverbs: ($F(1, 23)=76.37$, $p=9.113e{-}09$, Multiple $R^2=0.77$, Adjusted $R^2=0.76$)).

Specifically, except the normalized frequency of "potentially", which had a major increase from 0.13 in 1997 to 1.26 in 2021,most of the hedging adverbs in our list exhibited significant decline in normalized frequency. Due to limited space, we will not go into detail about the normalized frequencies of all the hedging adverbs in the current paper.

The development of hedging adjectives

Finally, the normalized frequencies of hedging adjectives also showed remarkable decrease, though with a fluctuant curve, as indicated in Fig. 7, with the statistics demonstrated in Table 5 ($F(1, 23)=15.93$, $p=0.0005$, Multiple $R^2=0.41$, Adjusted $R^2=0.38$). It should be noted that we paid special attention to the adjective "hypothesized" and "unclear", since the former may appear to be a past tense form of the verb "hypothesize" and therefore conveys predication, as in "we hypothesized that…" while the latter may serve as a criticism of others' claims. To eliminate the possible influences of such poor examples, we extracted all the examples that contain "hypothesized" and "unclear". In total, we found 759 examples using "hypothesized" and 546 examples using "unclear". Next, we manually examined

Fig. 7 Trajectory of hedging adjectives in *Science* research articles (full texts) across 25 years

all the examples and removed poor cases. In this step, we found and removed 694 poor cases for "hypothesized" and 65 valid cases. For "unclear", we did not find any poor cases.

Overall, the normalized frequency of hedging adjectives dropped significantly over time. The use of many hedging adjectives, such as "possible" and "unlikely", dropped significantly overtime. However, it should be acknowledged that the curve is fluctuant, as indicated by Fig. 7 and the R^2 of the simple linear regression (Multiple $R^2 = 0.41$, Adjusted $R^2 = 0.38$). Such fluctuation may be attributed to a significant rise of "unclear" from 0.45 in 1997 to 0.60 in 2021. Other than this, the normalized frequency of a number of adjectives, such as "potential", "likely", "unclear", showed no significant change in the past 25 years.

RQ2 The correlation between the use of hedges and sentiment scores.

The correlation between the use of hedges and sentiment scores

In this section, we report the results of a series of Pearson's correlation analyses, which compares the normalized frequencies of hedges in this study and the standardized sentiment scores in Yuan and Yao's (2022) study.

As shown in Fig. 8, all hedges somehow negatively correlate with Yuan and Yao's (2022) standardized sentiment scores with statistical significance. Specifically, the development of modal verbs had a slightly negative correlation with the sentiment scores. Meanwhile, the trajectory of semi-modal verbs had a more significant and negative correlation with the sentiment scores. Finally, the diachronic development of hedging adverbs and adjectives also had a negative correlation with sentiment scores with a considerable significance. The results show that the diachronic trend of hedges and that of linguistic positivity are negatively correlated, which indicates that there are gradually fewer hedges over time as the linguistic positivity increases.

This result indicates that the rising linguistic positivity (Yuan & Yao, 2022) has a significant association with the decreasing trend of uncertainty found in this study. Although our analysis does not allow us to interpret this association causally, we tentatively propose that

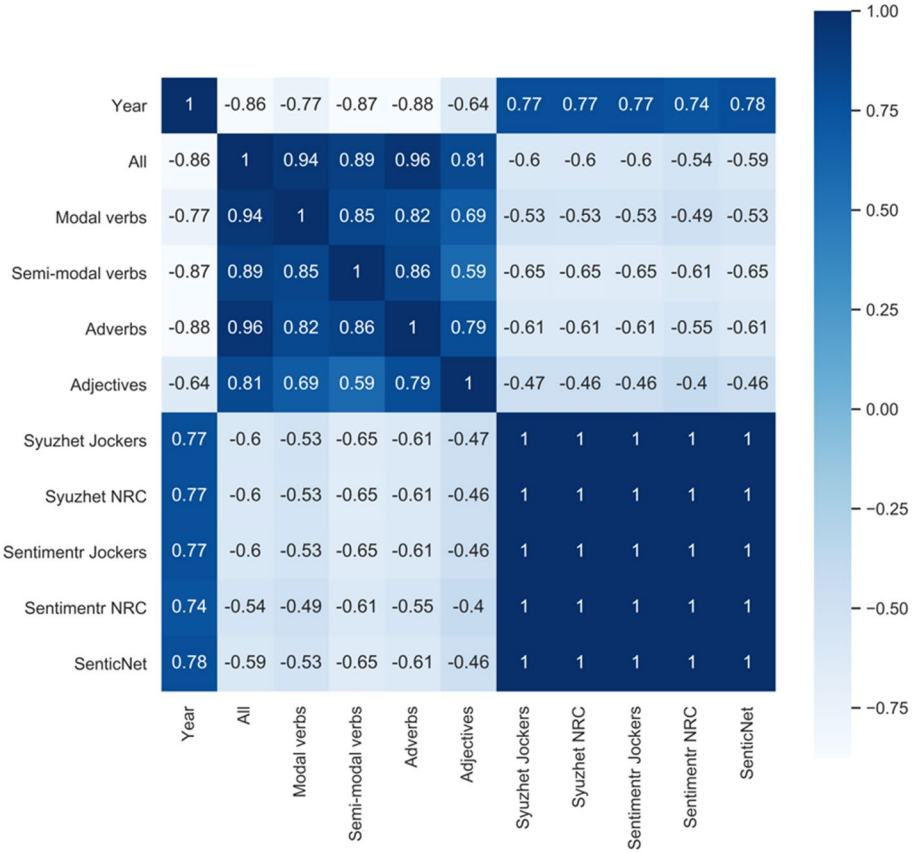

Fig. 8 The correlation between hedges and Yuan and Yao's (2022) sentiment scores of *Science* research articles (full texts) over 25 years

in the context of today's competitive publishing industry, Yuan and Yao's (2022) findings and our results may represent different strategies of research promotion. While the former is explicit and direct, the latter is more implicit and subtle. Specifically, other than obvious and explicit ways of using positive or exaggerative language to promote research in academic writing (Cao et al., 2021; Hyland & Jiang, 2021a, 2021b; Millar et al., 2019; Vinkers et al., 2015; Wen & Lei, 2022b; Yuan & Yao, 2022), there may be more implicit and subtle ways to promote one's research, that is, by showing confidence (Wheeler et al., 2021). Hedges are the linguistic realizations of (un)certainty for one's scientific claims. Therefore, declining uncertainty may serve as implicit positive writing strategies for research promotion. In other words, researchers make their research look desirable by showing less uncertainty through the use of less hedges in their academic writing.

Discussion

The current research explored the diachronic development of uncertainty in *Science* research articles (full texts) by examining the use of hedges that expressed uncertainty or doubts. Specifically, we analyzed the use of hedging modal verbs, semi-modal verbs, adverbs and adjectives in our corpus of *Science* research articles across 25 years. Our findings show that there have been significantly fewer hedges from 1997 to 2021. In detail, the normalized frequencies of modal verbs, semi-modal verbs, adverbs, and adjectives all decreased significantly in the past 25 years. Considering these results, it can be generally argued that academic writing, taking *Science* research articles (full texts) as a case, has evolved to be less uncertain, since we found significantly fewer hedges that expressed uncertainty in the corpus of academic writing in a 25-year-timespan. Our findings also agree with Poole et al. (2019) and Wheeler et al. (2021), who also find that academic writing has evolved to be less uncertain and more confident over time.

Also, we found significant correlations between the decreasing uncertainty and the increasing linguistic positivity in academic writing (Cao et al., 2021; Vinkers et al., 2015; Wen & Lei, 2022b; Yuan & Yao, 2022), both of which, we hypothesize, could be influenced by an increasingly more competitive environment for publication and the high visibility of the *Science* journal.

First, similar to the increasingly positive language in academic writing, less uncertainty in academic writing may also be a result of the pressure and competition in academic publishing (Cao et al., 2021; Vinkers et al., 2015; Wen & Lei, 2022b; Yuan & Yao, 2022). Since using hedges, especially reliability hedges (Hyland, 1998), in writing research articles contribute to expressing tentativeness (Mur-Duenas, 2021), or the writers' subjective uncertainty (Hyland, 1998, p. 166) about their scientific claims, using fewer hedges in research articles may indicate less uncertainty in reporting their scientific findings.

As academic publishing develops to be more competitive and challenging (Hyland & Jiang, 2021a, 2021b; Millar et al., 2019; Vinkers et al., 2015), as a result of a rising number of researchers (Lillis & Curry, 2013), higher demand for publication (Nosek et al., 2012), researchers may adopt a more positive strategy in writing research articles to make the paper more appealing (Cao et al., 2021; Hyland & Jiang, 2021a, 2021b; Vinkers et al., 2015; Wen & Lei, 2022b; Yuan & Yao, 2022). Likewise, reducing uncertainty in one's research may an implicit writing strategy in research promotion, compared to employing explicitly positive language, that contributes to a higher likelihood of acceptance in high-impact journals.

Second, the high visibility and reputation of the journal *Science* may also be a potential reason why hedging decreased significantly in *Science* research articles. As a leading journal for advanced research and findings, the journal has a worldwide readership of estimated more than one million[1] and ranks among the world's most cited journal. It is also one of the most widely recognized journals in the global academic community. Therefore, scientific discoveries published in *Science* would also benefit from a worldwide readership. With such a high visibility, research findings would often receive more discussion, debate, criticism, or even doubt within and outside the academic community. As a strategy, reduced uncertainty in research articles may display the authors' affirmative stance towards

[1] See https://www.science.org/content/page/mission-and-scope.

their methods and findings, and thus may help authors avoid the risk of doubt or criticism from editors, peer reviewers, and future readers.

The current study may have the following implications. Empirically, since not many studies have traced the use of hedges in academic writing, our study poses a salient question as to whether academic has evolved to be less uncertain by using fewer hedges, or specifically doubt hedges or reliability hedges. Our findings based on *Science* research articles found a significant decline in the use of such hedges, which may contribute to less uncertainty in academic writing. Practically, we confirmed our hypothesis that less use of hedges may correlate with the ongoing linguistic positivity bias in academic writing (Cao et al., 2021; Vinkers et al., 2015; Wen & Lei, 2022b; Yuan & Yao, 2022). We also argue that such variation in the use of hedges may be an implicit writing strategy to make research papers more desirable. Therefore, for peer reviewers and editors, our findings may have implications in terms of whether too little uncertainty is presented in research papers, since even the most assured scientific propositions have an inherently limited period of acceptance (Hyland, 1998, p. 6). For researchers and writers, our findings may serve as a reminder of the importance of using hedges in academic writing since the nature of academic knowledge is indeterminate (Mur-Duenas, 2021, p. 2), and categorical assertions of truth in scientific writing can be decidedly hazardous (Hyland, 1998, p. 6).

Finally, it should be acknowledged that our research is limited in many ways. First, we only examined a limited number of hedges in our study, excluding more hedges such as nouns (i.e., hypothesis), verbs (i.e., hypothesize) and phrasal hedges (i.e., to our knowledge). Therefore, our findings may be limited to the use of modal verbs, semi-modal verbs, adverbs, and adjectives. Future studies may fill this knowledge gap by incorporating a more inclusive list of hedges. Second, we only examined whether academic writing is becoming less uncertain by sampling only one journal, the findings based on which may be confined to this journal and thus lack certain generalizability. Future studies may extend our research to other journals. Third, in our study, we only explored the diachronic development of a certain type of hedges (ones that express uncertainty and doubt) in research articles as a whole, thus ignoring the internal variations of hedging in different sections of academic writing. Future studies may also trace and compare the diachronic development of different types hedges in different sections of research articles. Finally, we proposed a new research promotion strategy and confirmed that the use of fewer hedges in academic writing may correlate with linguistic positivity bias. However, there is still room for discussion concerning whether it may also correlate with other strategies, such as "confidence" (Wheeler et al., 2021). More in-depth studies on the correlation between hedging and other strategies of research promotion are yet to be conducted to validate or refute our findings.

Conclusion

In the current study, we traced the diachronic change of how researchers use hedges to express uncertainty, in a corpus of research articles published in the journal *Science* from 1997 to 2021. Our results show that over time there is fewer hedges expressing uncertainty used in *Science* full texts, indicating a declining of uncertainty in academic writing. We also found significant correlation between such trend of hedging and a positive trend for linguistic positivity found in Yuan and Yao (2022). Finally, we argue that such declining uncertainty can be an implicit writing strategy that writers adopt to make their research more desirable to editors, reviewers and readers, concerning the current competitive

publishing environment. Finally, this study offers implication for academic writers, peer reviewers, and editors.

Acknowledgements The authors would like to extent our greatest gratitude to the editor and the reviewers for their professional and constructive comments, especially about the conceptual connections between linguistic positivity bias, hyping, confidence, and research promotion, which would not have been made clearer without their help.

Declarations

Conflict of interest The authors declare that they have no conflict of interest.

Ethical approval Ethical approval corpus data used in the current research is collected through institutional access, and hence no ethical concerns are reported.

References

Biber, D., & Finegan, E. (1989). Styles of stance in English: Lexical and grammatical marking of evidentiality and affect. *Text-Interdisciplinary Journal for the Study of Discourse, 9*(1), 93–124.
Caffi, C. (2007). *Mitigation*. Elsevier.
Cao, X., Lei, L., & Wen, J. (2021). Promoting science with linguistic devices: A large-scale study of positive and negative words in academic writing. *Learned Publishing, 34*(2), 82–88. https://doi.org/10.1002/leap.1322
Dontcheva-Navratilova, O. (2016). Cross-cultural variation in the use of hedges and boosters in academic discourse. *Prague Journal of English Studies, 5*(1), 163–184.
Fløttum, K., Dahl, T., & Kinn, T. (2006). *Academic voices: Across languages and disciplines* (Vol. 148). John Benjamins Publishing.
Hu, G., & Cao, F. (2011). Hedging and boosting in abstracts of applied linguistics articles: A comparative study of English-and Chinese-medium journals. *Journal of Pragmatics, 43*(11), 2795–2809.
Hyland, K. (1996). Writing without conviction? Hedging in science research articles. *Applied Linguistics, 17*(4), 433–454.
Hyland, K. (1998). *Hedging in scientific research articles*. John Benjamins.
Hyland, K. (2005). *Mitigation*. Continuum.
Hyland, K., & Jiang, F. (2016). Change of attitude? A diachronic study of stance. *Written Communication, 33*(3), 251–274.
Hyland, K., & Jiang, F. K. (2021a). The Covid infodemic: Competition and the hyping of virus research. *International Journal of Corpus Linguistics, 26*(4), 444–468.
Hyland, K., & Jiang, F. K. (2021b). 'Our striking results demonstrate…': Persuasion and the growth of academic hype. *Journal of Pragmatics, 182*, 189–202.
Lillis, T., & Curry, M. J. (2013). 10 English, scientific publishing and participation in the global knowledge economy. In E. J. Erling & P. Seargeant (Eds.), *English and Development* (pp. 220–242). Multilingual Matters.
Liu, X., & Zhu, H. (2023). Linguistic positivity in soft and hard disciplines: temporal dynamics, disciplinary variation, and the relationship with research impact. *Scientometrics, 128*, 1–21.
Martín, P. M. (2005). *The rhetoric of the abstract in English and Spanish scientific discourse: A cross-cultural genre-analytic approach* (Vol. 279). Peter Lang.
Martín, P. M. (2008). The mitigation of scientific claims in research papers: A comparative study. *International Journal of English Studies, 8*(2), 133–152.
Millar, N., Salager-Meyer, F., & Budgell, B. (2019). "It is important to reinforce the importance of…": 'Hype' in reports of randomized controlled trials. *English for Specific Purposes, 54*, 139–151.
Mur-Dueñas, P. (2016). Modal hedging verbs in English as a Lingua Franca (ELF) Business Management research articles. *Kalbotyra, 69*, 153–178.
Mur-Dueñas, P. (2021). There may be differences: Analysing the use of hedges in English and Spanish research articles. *Lingua, 260*, 103131.
Myers, G. (1989). The pragmatics of politeness in scientific articles. *Applied Linguistics, 10*(1), 1–35.
Nosek, B. A., Spies, J. R., & Motyl, M. (2012). Scientific utopia: II. Restructuring incentives and practices to promote truth over publishability. *Perspectives on Psychological Science, 7*(6), 615–631.

Orta, I. V. (2010). A contrastive analysis of the use of modal verbs in the expression of epistemic stance in Business Management research articles in English and Spanish. *Ibérica, Revista de la Asociación Europea de Lenguas para Fines Específicos, 19*, 77–95.
Poole, R., Gnann, A., & Hahn-Powell, G. (2019). Epistemic stance and the construction of knowledge in science writing: A diachronic corpus study. *Journal of English for Academic Purposes, 42*, 100784.
Salager-Meyer, F. (1994). Hedges and textual communicative function in medical English written discourse. *English for Specific Purposes, 13*(2), 149–170.
Salager-Meyer, F., Ariza, M. Á. A., & Zambrano, N. (2003). The scimitar, the dagger and the glove: Intercultural differences in the rhetoric of criticism in Spanish, French and English medical discourse (1930–1995). *English for Specific Purposes, 22*(3), 223–247.
Varttala, T. (1999). Remarks on the communicative functions of hedging in popular scientific and specialist research articles on medicine. *English for Specific Purposes, 18*(2), 177–200.
Varttala, T. (2001). *Hedging in scientifically oriented discourse: Exploring variation according to discipline and intended audience*. Tampere University Press.
Vassileva, I. (2001). Commitment and detachment in English and Bulgarian academic writing. *English for Specific Purposes, 20*(1), 83–102.
Vinkers, C. H., Tijdink, J. K., & Otte, W. M. (2015). Use of positive and negative words in scientific PubMed abstracts between 1974 and 2014: retrospective analysis. *BMJ, 351*, 6467.
Wen, J., & Lei, L. (2022a). Adjectives and adverbs in life sciences across 50 years: Implications for emotions and readability in academic texts. *Scientometrics, 127*(8), 4731–4749.
Wen, J., & Lei, L. (2022b). Linguistic positivity bias in academic writing: A large-scale diachronic study in life sciences across 50 years. *Applied Linguistics, 43*(2), 340–364.
Wheeler, M. A., Vylomova, E., McGrath, M. J., & Haslam, N. (2021). More confident, less formal: Stylistic changes in academic psychology writing from 1970 to 2016. *Scientometrics, 126*, 9603–9612.
Yang, Y. (2013). Exploring linguistic and cultural variations in the use of hedges in English and Chinese scientific discourse. *Journal of Pragmatics, 50*(1), 23–36.
Yuan, Z. M., & Yao, M. (2022). Is academic writing becoming more positive? A large-scale diachronic case study of Science research articles across 25 years. *Scientometrics, 127*, 1–17.