Scientometrics (2022) 127:6191–6207
https://doi.org/10.1007/s11192-022-04515-2

Is academic writing becoming more positive? A large-scale diachronic case study of Science research articles across 25 years

Zhou-min Yuan[1] · Mingxin Yao[2]

Received: 18 March 2022 / Accepted: 2 September 2022 / Published online: 1 October 2022

Abstract

Academic writing is developing to be more positive. This linguistic positivity bias is confirmed in academic writing across disciplines and genres. The current research adopted sentiment analysis and examined the diachronic change in linguistic positivity in the full texts of 2,556 research articles published in Science in 25 years. The results showed that academic writing in research articles in the journal Science has become significantly more positive in the past 25 years. The findings of this study confirm linguistic positivity bias in academic writing based on empirical data from Science. Reasons for the increasingly positive language use in science articles might include the popularization of science, the growing number of researchers, and the difficulty of publishing in high-impact journals. Finally, this study discussed the implications of our findings for researchers, editors, and peer reviewers.

Keywords Linguistic positivity bias · Academic writing · Research articles · Science · Sentiment analysis

Introduction

Linguistic positivity bias (Augustine et al., 2011; Rozin et al., 2010) has received wide attention from various disciplines, such as psychology (Augustine et al., 2011; Rozin et al., 2010) and data science (Dodds et al., 2015; Garcia et al., 2012). Existing research reveals a general tendency to use more positive words than negative words in human communication. (Rozin et al., 2010; Augustine, 2011; Garcia et al., 2012; Dodds et al., 2015). For example, Dodds et al. (2015) observe a universal positivity bias in 100,000 words spread across 24 corpora in 10 languages with diverse origins. Augustine et al. (2011) find evidence for the human tendency to use positive words more often than negative words in

✉ Mingxin Yao
yao_mingx@163.com

[1] School of Foreign Studies, Nanjing University of Posts and Telecommunications, Wenyuan Road #9, Nanjing, Jiangsu 210023, People's Republic of China

[2] School of Foreign Studies, Nanjing University, Xianlin Road #163, Nanjing, Jiangsu 210023, People's Republic of China

written and spoken English. Linguistic positivity bias is likely to be universal across languages and registers. Linguistic positivity bias has also been identified in recent studies on academic writing, which find an upward trend in linguistic positivity in academic writing (Cao et al., 2020; Holtz et al., 2017; Lerchenmueller et al., 2019; Vinkers et al., 2015; Weidmann et al., 2018; Wen & Lei, 2021). Evidence in support of this claim is unearthed in academic articles across disciplines, including medical and life science (Cao et al., 2020; Lerchenmueller et al., 2019; Vinkers et al., 2015; Wen & Lei, 2021), political science (Weidmann et al., 2018), and cross-cultural psychology (Holtz et al., 2017). Vinkers et al. (2015) found significant growth in the frequency of positive words used in all scientific abstracts in PubMed from 1974 to 2014. Following Vinkers et al. (2015), Weidmann et al. (2018) also observe a marked increase in the use of positive words in abstracts in political science over time. Lerchenmueller et al. (2019) unveil the role of gender differences in the positive presentation of academic writing by examining titles and abstracts from a large dataset of academic articles in clinical research and life science in PubMed from 2002 to 2007. To date, these studies are largely limited to the examination of a linguistic positivity bias in the abstracts based on a small list of predefined positive/negative words.

To overcome the limitations in previous studies, studies on linguistic positivity bias in academic writing also make various methodological modifications. First, full texts in academic journals are taken into account to validate the findings based on abstracts. For example, Holtz et al. (2017) find a general upward trend in the use of positive words based on the exploration of a linguistic positivity bias in the full texts of research articles from cross-cultural psychology. Cao et al. (2020) also find an increasingly positive trend in terms of linguistic positivity based on the examination of how the frequencies of positive and negative words change over time in both abstracts and full texts in journals from PubMed. Second, the latest studies resort to larger dictionaries and lexicons to tackle the limitation of the small list of positive and negative words (Vinkers et al., 2015), including studies by Holtz et al. (2017), Bordignon et al. (2021), and Wen and Lei (2021). In these studies, researchers adopt self-created dictionaries (Holtz et al., 2017), expanded list of positive/negative words (Bordignon et al., 2021), or sentiment analysis with large lexicons in R (Wen & Lei, 2021) to triangulate the results based on Vinkers et al.'s (2015) small list of positive and negative words. Third, regarding the limitation of findings generalised from one discipline, Bordignon et al. (2021) compare abstracts across disciplines between a pre-pandemic corpus and a corpus of preprints issued in response to the COVID-19 pandemic, and discover an increase of positive words and a slight decrease of negative words. The results are encouraging because in these studies, a growing trend of positive language is observed based on mixed methods, thus further confirming the linguistic positivity bias in academic writing.

While existing studies provide significant insights into linguistic positivity bias in academic writing, the findings are still limited in the following three aspects. First, all of the exiting studies, except for Bordignon et al. (2021), investigated linguistic positivity bias in academic writing in individual disciplines. Little study has researched linguistic positivity bias in academic writing across disciplines. Second, most findings in the existing research are based on abstracts, the generalisability of which is open to doubt since abstracts may fail to fully and accurately reflect the main body of research articles (Pitkin and Branagan, 1998, Pitkin et al., 1999). Third, among the small number of studies that examined linguistic positivity bias in full texts (Cao et al., 2020 and Holtz et al., 2017), findings were generated based on a small list of predefined positive/negative words (Vinkers et al., 2015), which may not be sufficiently robust to detect sentiment polarity in academic writing from various disciplines (Wen & Lei, 2021: 4).

To address these limitations in the literature, the current study intends to examine linguistic positivity bias in academic writing in the following ways. First, we attempt to investigate linguistic positivity bias in research articles in Science. As the leading outlet and the core journal of scientific discovery, Science is a multidisciplinary journal (Glänzel & Schubert, 2003; Glänzel et al., 1999) that welcomes research articles from all fields of science and any other source (https://www.science.org/content/page/mission-and-scope), and is thus an ideal sample to represent academic writing across disciplines. Second, we intend to explore linguistic positivity bias in the full texts of research articles published in Science for more in-depth findings. Finally, we adopt sentiment analysis with large lexicons of sentiment words (following Wen & Lei, 2021) in our data analysis for more robust and accurate results.

Based on these considerations, this study addresses how the sentiment in the full texts of Science has changed in the past 25 years. Finally, we intend to discuss the possible reasons for the development of sentiment of the full texts in Science in the past 25 years.

Methods

This section presents the data collection and data analysis that were employed in the current research.

Data Collection

According to the categories of manuscripts given by Science (https://www.science.org/content/page/science-information-authors), the journal accepts and publishes the following types of manuscripts: research articles, reports, reviews, and commentaries. In our study, we only focused on the category of research articles for the following reasons. First, analysis on this single category eliminates the influence of such factors as genre differences between research articles and other categories. Second, concerning that the categories may change over time, we chose the category of research articles since it has remained consistent over the past 25 years. Finally, the reason why we did not include reports in our study is that research articles and reports may vary in terms of academic significance and paper length. Specifically, the former presents research papers with major advances containing up to 4500 words while the latter only accepts important new research results of broad significance and is limited to 2500 words.

In the current research study, we developed a script in Python for data collection, based on which a diachronic corpus for academic writing was constructed. The data were retrieved from the journal Science through access provided by Nanjing University Library. During data collection, the crawl delay was set to 10, conforming to the Robots Exclusion Standard announced by the Journal website.[1] The collected data were stored in the format of Excel files to form a corpus with 25 years of research articles (full texts) published in Science.

[1] According to the Robots Exclusion Standard on https://www.science.org/robots.txt, the crawler programmes should set crawl-delay to at least 1.

Table 1 Descriptive statistics of the corpus used in the study

Year	Number of Articles	Number of Words in Full Texts	Mean Word Count	Number of Sentences Per Year	Mean Sentence Count Per Article
1997	35	148,055	4,230	8,712	248
1998	45	174,071	3,868	11,561	256
1999	44	169,413	3,850	10,800	245
2000	59	227,902	3,862	15,980	270
2001	67	252,738	3,772	15,894	237
2002	67	248,088	3,702	15,471	230
2003	60	226,624	3,777	14,663	244
2004	60	228,106	3,801	14,071	234
2005	72	266,953	3,707	16,793	233
2006	64	236,720	3,698	14,455	225
2007	58	226,335	3,902	14,653	252
2008	57	211,828	3,716	13,318	233
2009	69	275,520	3,993	19,008	275
2010	67	268,991	4,014	19,143	285
2011	75	292,817	3,904	20,723	276
2012	54	213,215	3,948	14,809	274
2013	90	386,548	4,294	26,018	289
2014	126	551,805	4,379	37,231	295
2015	114	507,989	4,456	33,636	295
2016	149	669,659	4,494	45,393	304
2017	172	780,369	4,537	52,982	308
2018	192	875,819	4,561	60,291	314
2019	258	1,179,668	4,572	86,932	336
2020	278	1,256,443	4,519	93,645	336
2021	224	1,039,839	4,642	79,431	354

The corpus consisted of 2,556 research articles dating from January 1997 to August 2021,[2] with a total of 10,915,515 words (see Table 1 for the descriptive statistics), representing academic writing samples for a wide variety of fields and disciplines. The full texts of these research articles were published across a span of 25 years, sufficient time to explore the diachronic change of linguistic positivity in academic writing. In addition, the investigation on the full texts allows us to gain a holistic understanding of academic writing, hence yielding more reliable and generalisable results than those from analysing only abstracts.

[2] By the time we finished data collection, the latest issue of the Science journal is released in August, 2021.

Data analysis

Sentiment analysis

To address the research question, we conducted sentiment analysis, a method that studies the positive or negative evaluations, attitudes, and views expressed in a text (Liu & Lei, 2018; Mäntylä et al., 2018; Serrano-Guerreroet al., 2015; Taboada, 2016; Wen & Lei, 2021). To date, two main approaches have been commonly adopted in sentiment analysis: machine learning and lexicon-based approaches (Taboada, 2016; Mukhtar, 2018; Van Houtan et al., 2020; Wen & Lei, 2021). The former approach runs on a classifier trained for determining the polarity of texts (Taboada, 2016). However, this approach is limited to the specific field or genre of the training dataset that the classifier is trained for (Wen & Lei, 2021: 7). The lexicon-based approach, on the other hand, is based on lexicons or dictionaries containing a large set of sentiment words and their polarities (Taboada, 2016; Wen & Lei, 2021). This approach, although less accurate, is not subject to a particular genre or domain of the trained texts and, therefore, can efficiently handle data from different domains (Mukhtar, 2018: 2182).

Before sentiment analysis, we preprocessed the raw texts in our corpus by removing all the non-English symbols such as β or Ä, which may also be used in equations, etc., to ensure that our analysis is not affected by such special symbols. To do so, we coded a regular expression in python, which deletes the non-English symbols in our data.

In the current study, we employ the lexicon-based approach in sentiment analysis for the following reasons. First, this method is proven to be robust across different domains without changing the dictionaries (Taboada et al., 2011: 9). Second, this approach allows us to compare our findings with previous studies using the same approach, such as Wen and Lei (2021). In detail, we coded an R script to run sentiment analyses on each research article (full text). Two packages embedded in R are used separately for performing sentiment analysis, namely, *Syuzhet* (Jockers, 2017) and *Sentimentr* (Rinker, 2018).

In the first sentiment analysis (SA1), we resort to *Syuzhet* (Jockers, 2017), a popular R package widely applied in studies with sentiment analysis (Bradley & James, 2019; Jensen & Bang, 2017; Liu & Lei, 2018; Vergeer, 2020; Wen & Lei, 2021). However, Syuzhet (Jockers, 2017) was found to be error-prone due to the lack of valence shifters (i.e., negators, intensifiers, or downtoners) in a sentence (See Rinker, 2018). Later, Rinker (2018) released a modified R package, *Sentimentr* (Rinker, 2018), based on the weakness of *Syuzhet* (Jockers, 2017). Therefore, the second sentiment analysis (SA2) is carried out based on *Sentimentr* (Rinker, 2018). In addition, with two packages, we are able to triangulate the results by comparing those of SA1 and SA2.

We also ran sentiment analysis on multiple lexicons in *Syuzhet* (Jockers, 2017) and Rinker (2018) to further triangulate the results. Specifically, in *Syuzhet* (Jockers, 2017), we opted for the Jockers sentiment lexicon (Jockers, 2017) and the NRC sentiment lexicon (Mohammad & Turney, 2010). In *Sentimentr* (Rinker, 2018), we used three lexicons, including the previous two lexicons and one additional lexicon, i.e., the SenticNet lexicon (Cambria et al., 2016). The reasons for using these lexicons are as follows. First, these lexicons are proven to be highly robust (Mohammad, 2010) and reliable (Wen & Lei, 2021) in terms of sentiment analysis. Second, embedded in the R packages, these lexicons are free and open for access.

The sentiment analysis procedure based on the above lexicons in *Syuzhet* (Jockers, 2017) and *Sentimentr* (Rinker, 2018) follows several steps. It should be noted that all of the algorithms used for sentiment analysis in this study are based on sentences. To calculate the sentiment scores for one article, the algorithms in the R packages first divide each research article (full text) into individual sentences. Next, they produce a raw sentiment score for each sentence in the article. Finally, a composite sentiment score for each article is calculated by adding up the raw scores of all the sentences in the article. However, the results produced by different lexicons are not comparable because the sentiment words included in these lexicons were tagged on different scales and intervals (Wen & Lei, 2021: 9). Therefore, standardization of the raw sentiment scores is necessary after the analysis. We followed Lennox et al. (2020) and Wen and Lei (2021) in standardizing the raw sentiment scores by calculating the mean, i.e., μ *(sentiments)* and the standard deviation, i.e., σ *(sentiments)* of the raw scores of all the research articles, and finally the standardized sentiment score for each article based on Lennox et al.'s method (2020), as displayed in Formula 1. Finally, to compare the sentiment scores across time on a yearly basis, we calculated the yearly means of the standardized sentiment scores.

Formula 1:

$$Standardizedsentiment = \frac{sentiment - \mu(sentiments)}{\sigma(sentiments)} + (sentiments)$$

In addition, it should be noted that sentiment analysis with these two packages is not affected by factors such as the prevalence of positive words and sentence length (Wen & Lei, 2021), due to a larger proportion of negative words in the lexicons and the design of the algorithms. However, due to the space limit and also the scope of the present paper, we do not specify the technical details for our instruments. For more technical details, please see Rinker (2019) as a consultation.

Statistical analysis

In terms of statistical analysis, we first conducted simple linear regression (Lei & Wen, 2020; Lei & Yan, 2016; Lei & Zhang, 2018; Wen & Lei, 2021) to examine the diachronic development of research articles in Science in terms of sentiment scores. Specifically, we performed simple linear regression on all five result samples from the five lexicons, two of which were based on *Syuzhet* (Jockers, 2017) and the other three based on *Sentimentr* (Rinker, 2018). In all the analyses of simple linear regression, we examined the developmental trajectory of the sentiment scores with the year as the independent variable and the standardized sentiment score of the full text of each research article as the dependent variable.

In addition, to further compare the results of sentiment analysis based on different packages and lexicons, we also performed Pearson's product-moment correlation analyses (Wen & Lei, 2021) to examine whether the five result samples are positively or negatively correlated with statistical significance.

Results

This section first reports the distribution of sentiment in the full texts across 25 years and then the results of statistical analysis, which may shed light on the trend of linguistic positivity bias.

Table 2 Distribution of sentiment in the full texts across 25 years

Year	*Syuzhet* (Jockers, 2017)		*Sentimentr* (Rinker, 2018)		
	Jockers sentiment lexicon	NRC sentiment lexicon	Jockers sentiment lexicon	NRC sentiment lexicon	SenticNet sentiment lexicon
1997	9.521331	9.527533	10.080014	10.661952	11.987791
1998	9.396531	9.406549	9.870166	10.315908	11.814906
1999	7.546121	7.578637	7.999714	8.21836	9.868191
2000	8.571431	8.608957	9.173225	9.526734	10.960807
2001	8.383277	8.433924	8.998861	9.42009	10.892345
2002	7.863707	7.898525	8.511171	8.917346	10.222025
2003	7.777968	7.813237	8.350494	8.488094	10.203304
2004	8.431991	8.462608	8.979032	9.344791	10.948697
2005	7.446901	7.439231	7.924671	8.110029	9.693471
2006	7.580008	7.58766	8.227586	8.489964	10.051596
2007	8.735593	8.775876	9.38116	9.682491	11.204487
2008	8.054496	8.105745	8.538112	8.94146	10.547337
2009	10.108575	10.153245	10.777659	11.069634	12.64925
2010	10.257882	10.277272	11.071934	11.166507	12.782353
2011	10.295619	10.318279	11.033113	11.376831	12.711169
2012	10.101425	10.115826	10.786342	11.124611	12.645313
2013	9.904227	9.947269	10.574727	10.815863	12.312222
2014	10.466433	10.50242	11.111915	11.368122	12.987398
2015	11.020506	11.050246	11.775518	11.977221	13.53401
2016	11.330756	11.350055	12.023898	12.192563	13.852181
2017	11.591671	11.617402	12.245307	12.457971	14.068735
2018	11.225818	11.246149	11.846535	12.016217	13.699407
2019	10.603499	10.620283	11.202951	11.355169	13.117856
2020	10.812	10.83708	11.454397	11.648572	13.344973
2021	10.413977	10.451182	10.988131	11.142709	12.929841

Distribution of sentiment across 25 years

We performed sentiment analyses based on five lexicons in two R packages, i.e., Jockers sentiment lexicon and NRC sentiment lexicon in *Syuzhet* (Jockers, 2017), as well as Jockers sentiment lexicon, NRC sentiment lexicon, and SenticNet sentiment lexicon in *Sentimentr* (Rinker, 2018). In this step, our data analysis generated five result samples for further analysis. To compare the sentiment scores across time on a yearly basis, we calculated the yearly means of the standardized sentiment scores from January 1997 to August 2021, as displayed in Table 2.

From eyeballing the data, we derived the hypothesis that the sentiment scores in the full texts went through a general increase from 1997 to 2021. However, since the statistical evaluation goes beyond what eyeballing the table can do (Hilpert & Gries, 2009: 390), we must rely on further statistics for precise insights into the diachronic change of sentiment scores in the data, which is presented in the next section.

Table 3 Descriptive statistics of standardized sentiment scores in SA1

Lexicon	Mean[a]	Standard deviation[a]	Maximum	Minimum
Jockers	10.743097	1.825873	38.412861	– 17.274664
NRC	9.49767	1.34048	38.418873	– 17.316931

[a]mean and standard deviation by year and full text

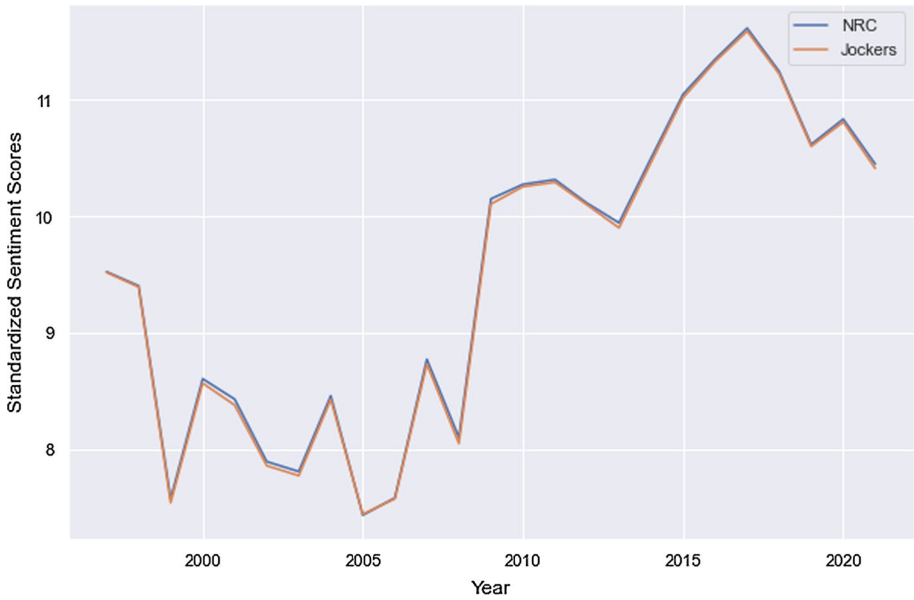

Fig. 1 Diachronic trajectory of linguistic positivity based on SA1

Trends of linguistic positivity based on sentiment analyses

In this section, we first report the statistical result of SA1 before that of SA2 because they are based on distinct packages.

Results of the SA1

In SA1, we ran sentiment analysis with two lexicons in *Syuzhet* (Jockers, 2017), namely, the Jockers sentiment lexicon and the NRC sentiment lexicon. Table 3 summarizes the descriptive statistics of the standardized sentiment scores based on each lexicon. Figure 1 demonstrates the yearly means of the standardized sentiment scores based on SA1 from January 1997 to August 2021.

The results of simple linear regression on SA1 suggested a significant increase in sentiment in the full texts, indicating an upward developmental trend of linguistic positivity in the last 25 years (Jockers: $F(1,23) = 34.15$, $p = 5.912\text{e-}06$, multiple $R^2 = 0.5975$, adjusted $R^2 = 0.58$; NRC: $F(1,23) = 34.33$, $p = 5.694\text{e-}06$, multiple $R^2 = 0.5988$, adjusted $R^2 = 0.5814$;). Table 4 presents the detailed statistics of the model.

Table 4 Detailed statistics of simple linear regression for SA1

Model	Variable	Estimate	Standard error	*t*-value	*p*-value
Jockers	(Intercept)	− 273.34805	48.40457	− 5.647	9.51e− 06 ***
	Year	0.14079	0.02409	5.843	5.91e− 06 ***
NRC	(Intercept)	− 273.42479	48.29356	− 5.662	9.18e− 06 ***
	Year	0.14084	0.02404	5.859	5.69e− 06 ***

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

Table 5 Descriptive statistics of the standardized sentiment scores in SA2

Lexicon	Mean[a]	Standard deviation[a]	Maximum	Minimum
Jockers	10.492273	1.628299	39.64096	− 18.735906
NRC	10.617342	1.689519	40.348968	− 19.887274
SenticNet	10.729428	1.835543	40.884626	− 17.719674

[a]mean and standard deviation by year and full text

Results of the SA2

In SA2, we analysed the sentiments in the full texts with *Sentimentr* (Rinker, 2018) and three lexicons in the package, i.e., the Jockers sentiment lexicon (Jockers, 2017), the NRC sentiment lexicon (Mohammad & Turney, 2010), and the SenticNet lexicon (Cambria

Fig. 2 Diachronic trajectory of linguistic positivity based on SA2

Table 6 Detailed statistics of simple linear regression for SA2

Model	Variable	Estimate	Standard error	*t*-value	*p*-value
Jockers	(Intercept)	− 283.97819	50.24481	− 5.652	9.40e− 06 ***
	Year	0.14639	0.02501	5.853	5.77e− 06 ***
NRC	(Intercept)	− 262.86763	52.22310	− 5.034	4.29e− 05 ***
	Year	0.14639	0.02501	5.853	5.77e− 06 ***
SenticNet	(Intercept)	− 281.32789	49.83593	− 5.645	9.56e− 06 ***
	Year	0.14599	0.02481	5.885	5.35e− 06 ***

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

Table 7 Pearson's correlation test for standardized sentiment scores in SA1 and SA2

Lexicon	1	2	3	4	5
1. Jockers (SA1)	–				
2. NRC (SA1)	0.9999409***	–			
3. Jockers (SA2)	0.9986053***	0.9986208***	–		
4. NRC (SA2)	0.9957922***	0.9958604***	0.9966371***	–	
5. SenticNet (SA2)	0.9990544***	0.9991685***	0.9981632***	0.9959328***	–

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

et al., 2016). Table 5 demonstrates the descriptive statistics of the standardized sentiment scores with SA2. Figure 2 displays the yearly means of the standardized sentiment scores of the full texts across 25 years.

In the statistical analysis for SA2, simple linear regression revealed an upward trend in the development of linguistic positivity because the sentiments have significantly increased in the full texts in the past 25 years (Jockers: $F(1,23)=34.26$, $p=5.773\text{e}-06$, multiple $R^2=0.5983$, adjusted $R^2=0.5809$; NRC: $F(1,23)=27.38$, $p=2.625\text{e}-05$, multiple $R^2=0.5435$, adjusted $R^2=0.5236$; SenticNet: $F(1,23)=34.63$, $p=5.346\text{e}-06$, multiple $R^2=0.6009$, adjusted $R^2=0.5836$). Table 6 demonstrates the results of the simple linear regression model.

Generally, the statistical results of SA2 are in line with those of SA1, thus offering further triangulated evidence for the increasing linguistic positivity in academic writing during the 25 years evaluated.

Results of the correlation test

Finally, we performed Pearson's correlation test (Wen & Lei, 2021) on the five samples of standardized sentiment scores in SA1 and SA2 to test the reliability of sentiment analysis. Table 7 shows the interrelation of the standardized sentiment scores in SA1 and SA2. According to the results, all the sentiment measures are highly significantly correlated ($r=0.9999409>0.9$, $p=2.2\text{e}-16<0.001$), providing further triangulated evidence for the reliability of our sentiment analysis.

Table 8 List of predefined positive and negative words (Vinkers et al., 2015)

Category	Words
Positive	amazing, assuring, astonishing, bright, creative, encouraging, enormous, excellent, favourable, groundbreaking, hopeful, innovative, inspiring, inventive, novel, phenomenal, prominent, promising, reassuring, remarkable, robust, spectacular, supportive, unique, unprecedented
Negative	detrimental, disappointing, disconcerting, discouraging, disheartening, disturbing, frustrating, futile, hopeless, impossible, inadequate, ineffective, insignificant, insufficient, irrelevant, mediocre, pessimistic, substandard, unacceptable, unpromising, unsatisfactory, unsatisfying, useless, weak, worrisome

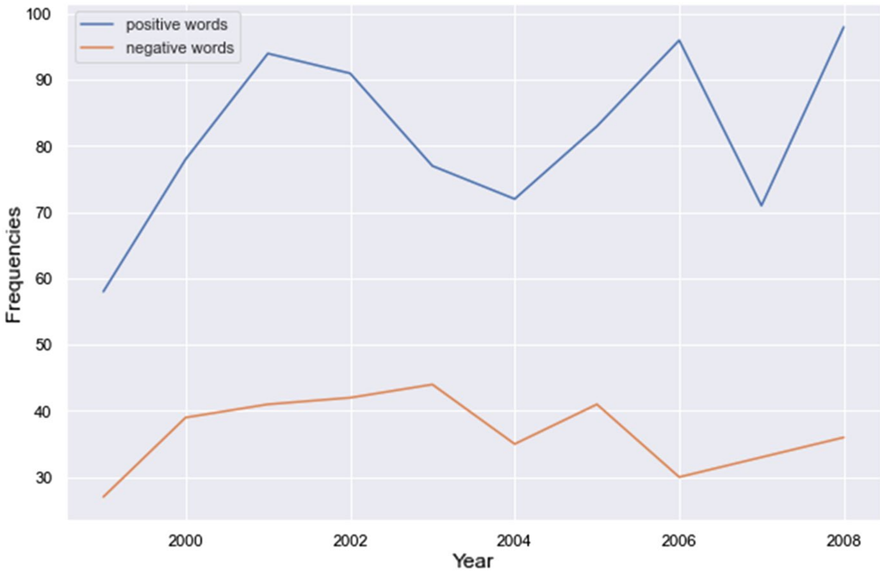

Fig. 3 Frequency of Vinkers et al's (2015) positive and negative words between 1999 and 2008

Fluctuations in the sentiment scores

The current study investigated linguistic positivity bias in academic writing based on a diachronic corpus of academic research articles published between January 1997 and August 2021 in Science. To date, this research is most likely the first study that employs sentiment analysis with large lexicons to examine linguistic positivity bias in the full texts, rather than abstracts, of academic writing from a diachronic perspective.

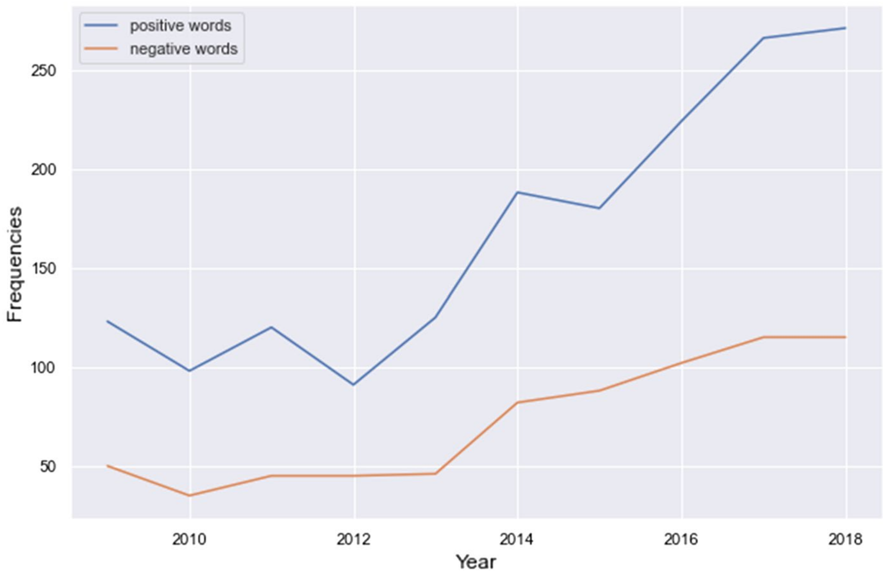

Fig. 4 Frequency of Vinkers et al's (2015) positive and negative words between 2009 and 2018

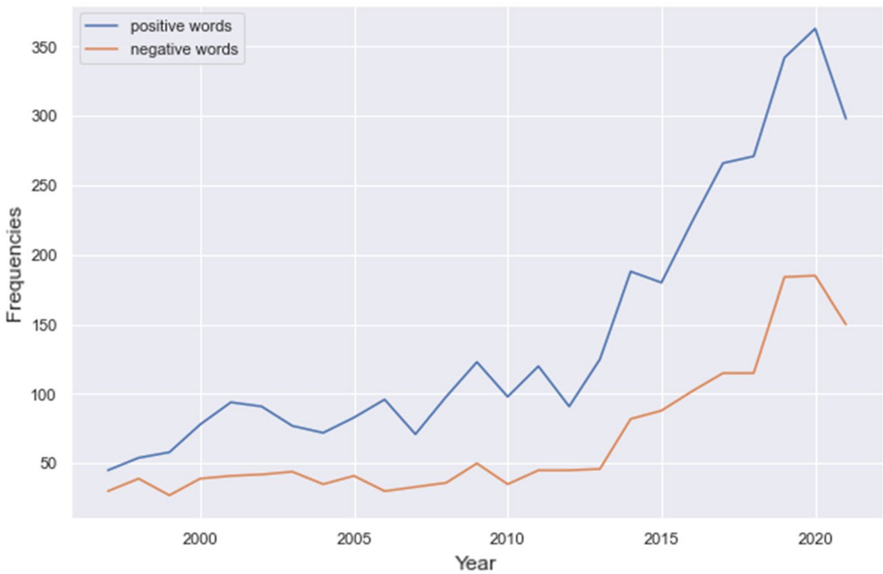

Fig. 5 Overall distribution of Vinkers et al's (2015) positive and negative words between 1997 and 2021

The findings of this study revealed a generally increasingly positive trend in academic writing in Science over the past 25 years. Our findings were in line with those in Cao et al. (2020) and Wen and Lei (2021) with evidence from Science. Our results also fill the gap that no previous study has examined linguistic positivity bias by employing sentiment analysis with large lexicons on the full texts of research articles from a diachronic perspective.

However, some fluctuations were observed in the data. Roughly between 1999 and 2008 there were some fluctuations in the sentiment scores, which is followed by a general increase in the scores approximately from 2009 to 2018. We find it necessary to report and discuss why there were such fluctuations in our data. To do so, we investigated the linguistic nature of our corpus with the help of a list of predefined positive and negative words proposed by Vinkers et al. (2015), as displayed in Table 8. Specifically, we coded a python script to calculate the frequency of these positive and negative words in the corpus. The above-mentioned fluctuations between 1999 and 2008 were also observed in the frequency of Vinkers et al's (2015) positive and negative words in our corpus, as displayed in Fig. 3. From 2009 to 2018 there was a general increase in the sentiment scores, which was also in line with the frequency of Vinkers et al's (2015) positive and negative words between 2009 and 2018, as shown in Fig. 4. To be specific, there appeared to be more increasingly more positive words than negative words from 2009 to 2018.

The overall distribution of Vinkers et al's (2015) positive and negative words between 1997 and 2021 is illustrated in Fig. 5. From Fig. 5, we are able to see a general increase in the frequency of positive words, with less fluctuations observed. This result might further reinforce our arguments based on the results of the previous sentiment analysis that academic writing in Science research articles has become increasingly positive.

However, we have to acknowledge the limitations of Vinkers et al's (2015) list of positive and negative words since they are small in number, and therefore could not entirely represent the fluctuations in our sentiment scores. Therefore, such an analysis merely provide an aspect of the linguistic variations in the corpus. In other words, the fluctuations in our sentiment scores could possibly be a result of the inner linguistic variations in the corpus, with Vinkers et al's (2015) list of positive and negative words serving as a case in point.

Discussion

In this section, we discuss the potential reasons for the use of increasingly positive language in academic writings. First, the popularization of science (Bell & Turney, 2014; Pilkington, 2016) might be one reason why academic writing has developed to be more positive. The popularization of science is a long-standing tradition that includes a variety of practices in making scientific information more accessible to general and nonexpert audiences (Bucchi & Trench, 2014). As a result, both publishers and researchers are keen to promote scientific advancements and research brands with the public (Bell & Turney, 2014; Pilkington, 2016). By doing so, researchers tend to adopt narratives that shape a positive image of themselves as creative thinkers when describing their discoveries (Pilkington, 2016). In relation to Science, the journal seeks to not only advance scientific understanding but also publish papers that merit recognition by the wider scientific community and the general public (https://www.science.org/content/page/mission-and-scope). Therefore, it is possible that the popularization of science has played a vital role in the increasing linguistic positivity trend in research articles in Science over the past 25 years.

Second, we agree with Wen and Lei (2021), Cao et al. (2020), and Vinkers et al. (2015) that the linguistic positivity bias is possibly influenced by competition in publications in the academic community. Specifically, positive language as a technique or strategy has been increasingly adopted in recent decades (Cao et al., 2020: 4), during which high-quality publications have gained such importance that they can influence various aspects of a researcher's career (Nicolini & Nozza, 2008; Nosek et al., 2012; Wen & Lei, 2021), such as hiring, salary, promotion, tenure, and grant awards (Nosek et al., 2012: 616). At the same time, however, it has become increasingly more difficult to have one's research published (Wen & Lei, 2021: 17) due to the high demand for publication (Nosek et al., 2012), the growing number of researchers (Lillis & Curry, 2013), the competitive process of publication selection (Millar et al., 2019), and the more thorough and critical editorial and peer review process in high-impact journals (Vinkers et al., 2015: 3). Consequently, researchers may adopt a more positive writing approach in research articles (Vinkers et al., 2015) to promote their research for publication purposes (Cao et al., 2020; Wen & Lei, 2021).

Third, the increasing linguistic positivity bias in academic writing may also be attributed to positive outcome reporting bias (Dwan et al., 2013) or positive publication bias (Mlinarić et al., 2017). Recent studies found that studies with positive or statistically significant results have greater odds of being published (Dwan et al., 2013; Mlinarić et al., 2017), which contributes to a scientific culture that favors positive outcomes (Wen & Lei, 2021). As a result, researchers are more likely to report statistically positive or significant outcomes (Dwan et al., 2013) as a strategy to impress the audience (Wen & Lei, 2021), editors, and peer reviewers (Chiu et al., 2017). In this process of promoting their research (Cao et al., 2020; Millar et al., 2019), researchers may be inclined to use more positive writing, such as hyperbolic and/or subjective language, to glamorize and promote and/or exaggerate aspects of their research (Millar et al., 2019: 139). Researchers who have observed this phenomenon are now warning the scientific community about the risk that this practice may undermine the objectivity and interpretation of newly discovered scientific knowledge (Millar et al., 2019; Wen & Lei, 2021) as well as the trustworthiness of published research findings (Ioannidis, 2005).

This study has implications for researchers, editors, and peer reviewers. On the one hand, positive language is useful in terms of selling the paper and promoting science to the general public; however, extensive use of positively subjective language could erode the accuracy of the information conveyed and, hence, result in doing a disservice to science (Millar et al., 2019). Therefore, researchers should adopt the right judgment of intention in academic writing (Millar et al., 2019) and a prudent use of promotional language (Cao et al., 2020) to preserve the integrity of the scientific findings. Wen and Lei (2021) also argue that researchers take responsibility for the language used in academic writing. The findings of this research and previous studies may also have some implications for editors and peer reviewers regarding the need for heightened vigilance for promotional language use (Millar et al., 2019) and more tolerance and even rewards for negative research results for the purpose of keeping science on track. (Nature Editorial, 2017).

Empirically, our study features originality in the following fronts. First, we investigate the full texts of research articles on Science. Therefore, our findings based on the full texts may be more generalisable than those generated from abstracts. It may further validate and reinforce previous findings based on abstracts, whose generalisability might be questionable. Second, our study may better represent academic writing. Specifically, since Science publishes research articles from across disciplines, our findings could reveal the diachronic development of linguistic positivity in academic writing across disciplines instead of one or two disciplines.

However, our study is also limited as follows. First, as a case study on Science, the findings on the full texts are limited to this journal. Future research may extend the size of the corpus by incorporating full texts from more scientific journals such as Nature and other Nature indexed journals. Second, our study revealed only a positive trend in the diachronic development of academic writing in science. However, it may fail to account for exactly how the use of positive language evolved over time. Future studies may approach the same issue with in-depth qualitative methods. Third, our study is limited in that we did not take into account that language use and hence sentiments may vary in different sections of a research article, which is a potentially interesting and relevant topic for future research. Finally, although our instruments exhibit robustness and high reliability, they are still limited in terms of accuracy. Future research may adopt tools with better accuracy in detecting sentiments in academic writing, such as more accurate algorithms or machine learning models.

Acknowledgements The authors would like to extend their kind regards to the editorial office and the reviewers for their insightful comments and suggestions.

Funding This research was funded by the National Social Science Fund of China (project number 20AYY009).

Declarations

Conflict of interest The authors have no financial or non-financial interests to disclose.

Ethical approval The data used in this research are texts of journal articles with institutional access. The authors have no ethical issues to report.

References

Augustine, A. A., Mehl, M. R., & Larsen, R. J. (2011). A positivity bias in written and spoken English and its moderation by personality and gender. *Social Psychological and Personality Science, 2*(5), 508–515.

Bell, A., & Turney, J. (2014). Popular science books: from public education to science bestsellers. In *Routledge Handbook of Public Communication of Science and Technology* (pp. 31–42). Routledge.

Bordignon, F., Ermakova, L., & Noel, M. (2021). Over-promotion and caution in abstracts of preprints during the COVID-19 crisis. *Learned Publishing, 34*(4), 622–636.

Bradley, A., & James, R. J. (2019). How are major gambling brands using Twitter? *International Gambling Studies, 19*(3), 451–470.

Bucchi, M., & Trench, B. (2014). Science communication research: themes and challenges. In *Routledge Handbook of Public Communication of Science and Technology* (pp. 17–30). Routledge.

Cambria, E., Poria, S., Bajpai, R., & Schuller, B. (2016, December). SenticNet 4: A semantic resource for sentiment analysis based on conceptual primitives. In Proceedings of COLING 2016, the 26th international conference on computational linguistics: Technical papers (pp. 2666–2677).

Cao, X., Lei, L., & Wen, J. (2020). Promoting science with linguistic devices: A large-scale study of positive and negative words in academic writing. *Learned Publishing, 34*(2), 82–88.

Chiu, K., Grundy, Q., & Bero, L. (2017). 'Spin'in published biomedical literature: A methodological systematic review. *PLoS Biology, 15*(9), e2002173.

Dodds, P. S., Clark, E. M., Desu, S., Frank, M. R., Reagan, A. J., Williams, J. R., & Danforth, C. M. (2015). Human language reveals a universal positivity bias. *Proceedings of the National Academy of Sciences, 112*(8), 2389–2394.

Dwan, K., Gamble, C., Williamson, P. R., & Kirkham, J. J. (2013). Systematic review of the empirical evidence of study publication bias and outcome reporting bias—an updated review. *PLoS ONE, 8*(7), e66844.

Editorial, N. (2017). Rewarding negative results keeps science on track. *Nature, 551*, 414. https://doi.org/10.1038/d41586-017-07325-2

Garcia, D., Garas, A., & Schweitzer, F. (2012). Positive words carry less information than negative words. *EPJ Data Science, 1*(1), 1–12.

Glänzel, W., & Schubert, A. (2003). A new classification scheme of science fields and subfields designed for scientometric evaluation purposes. *Scientometrics, 56*(3), 357–367.
Glänzel, W., Schubert, A., & Czerwon, H. J. (1999). An item-by-item subject classification of papers published in multidisciplinary and general journals using reference analysis. *Scientometrics, 44*(3), 427–439.
Hilpert, M., & Gries, S. T. (2009). Assessing frequency changes in multistage diachronic corpora: Applications for historical corpus linguistics and the study of language acquisition. *Literary and Linguistic Computing, 24*(4), 385–401.
Holtz, P., Deutschmann, E., & Dobewall, H. (2017). Cross-cultural psychology and the rise of academic capitalism: Linguistic changes in CCR and JCCP articles, 1970–2014. *Journal of Cross-Cultural Psychology, 48*(9), 1410–1431.
Ioannidis, J. P. (2005). Why most published research findings are false. *PLoS Medicine, 2*(8), e124.
Jensen, M. J., & Bang, H. P. (2017). Populism and connectivism: An analysis of the Sanders and Trump nomination campaigns. *Journal of Political Marketing, 16*(3–4), 343–364.
Jockers, M. L. 2017. 'Syuzhet: Extracts sentiment and sentiment-derived plot arcs from text,' available at https://cran.r-project.org/web/packages/syuzhet. Accessed 20 September 2021.
Lei, L., & Wen, J. (2020). Is dependency distance experiencing a process of minimization? A diachronic study based on the State of the Union addresses. *Lingua, 239*, 102762.
Lei, L., & Yan, S. (2016). Readability and citations in information science: Evidence from abstracts and articles of four journals (2003–2012). *Scientometrics, 108*(3), 1155–1169.
Lei, L., & Zhang, Y. (2018). Lack of improvement in scientific integrity: An analysis of WoS retractions by Chinese researchers (1997–2016). *Science and Engineering Ethics, 24*(5), 1409–1420.
Lennox, R. J., Veríssimo, D., Twardek, W. M., Davis, C. R., & Jarić, I. (2020). Sentiment analysis as a measure of conservation culture in scientific literature. *Conservation Biology, 34*(2), 462–471.
Lerchenmueller, M. J., Sorenson, O., & Jena, A. B. (2019). Gender differences in how scientists present the importance of their research: observational study. *BMJ*. https://doi.org/10.1136/bmj.l6573
Lillis, T., & Curry, M. J. (2013). 10. English, Scientific Publishing and Participation in the Global Knowledge Economy. In *English and Development* (pp. 220–242). Multilingual Matters.
Liu, D., & Lei, L. (2018). The appeal to political sentiment: An analysis of Donald Trump's and Hillary Clinton's speech themes and discourse strategies in the 2016 US presidential election. *Discourse, Context & Media, 25*, 143–152.
Mäntylä, M. V., Graziotin, D., & Kuutila, M. (2018). The evolution of sentiment analysis—A review of research topics, venues, and top cited papers. *Computer Science Review, 27*, 16–32.
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.
Mlinarić, A., Horvat, M., & Šupak Smolčić, V. (2017). Dealing with the positive publication bias: Why you should really publish your negative results. *Biochemia Medica, 27*(3), 447–452.
Mohammad, S. M. 2010. 'Sentiment and emotion lexicons,' available at http://saifmohammad.com/WebPages/lexicons.html. Accessed 8 November 2021.
Mohammad, S., & Turney, P. (2010, June). Emotions evoked by common words and phrases: Using mechanical turk to create an emotion lexicon. In Proceedings of the NAACL HLT 2010 workshop on computational approaches to analysis and generation of emotion in text (pp. 26–34).
Mukhtar, N., Khan, M. A., & Chiragh, N. (2018). Lexicon-based approach outperforms supervised machine learning approach for Urdu sentiment analysis in multiple domains. *Telematics and Informatics, 35*(8), 2173–2183.
Nicolini, C., & Nozza, F. (2008). Objective assessment of scientific performances world-wide. *Scientometrics, 76*(3), 527–541.
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.
Pilkington, O. A. (2016). *Presented Discourse Analysis in Popular Science Narratives of Discovery* (Doctoral dissertation, University of Birmingham).
Pitkin, R. M., & Branagan, M. A. (1998). Can the accuracy of abstracts be improved by providing specific instructions?: A randomized controlled trial. *JAMA, 280*(3), 267–269.
Pitkin, R. M., Branagan, M. A., & Burmeister, L. F. (1999). Accuracy of data in abstracts of published research articles. *JAMA, 281*(12), 1110–1111.
Rinker, T. (2018). Trinker/sentimentr: Dictionary Based Sentiment Analysis that Considers Valence Shifters (version 2.6.1), available at http://github.com/trinker/sentimentr. Accessed 20 September 2021.
Rinker, T. (2019). Calculate Text Polarity Sentiment, available at https://cran.r-project.org/web/packages/sentimentr/sentimentr.pdf. Accessed 20 September 2021.

Rozin, P., Berman, L., & Royzman, E. (2010). Biases in use of positive and negative words across twenty natural languages. *Cognition and Emotion, 24*(3), 536–548.

Serrano-Guerrero, J., Olivas, J. A., Romero, F. P., & Herrera-Viedma, E. (2015). Sentiment analysis: A review and comparative analysis of web services. *Information Sciences, 311*, 18–38.

Taboada, M. (2016). Sentiment analysis: An overview from linguistics. *Annual Review of Linguistics, 2*, 325–347.

Taboada, M., Brooke, J., Tofiloski, M., Voll, K., & Stede, M. (2011). Lexicon-based methods for sentiment analysis. *Computational LinguistIcs, 37*(2), 267–307.

Van Houtan, K. S., Gagne, T., Jenkins, C. N., & Joppa, L. (2020). Sentiment analysis of conservation studies captures successes of species reintroductions. *Patterns, 1*(1), 100005.

Vergeer, M. (2020). Artificial intelligence in the Dutch press: An analysis of topics and trends. *Communication Studies, 71*(3), 373–392.

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*. https://doi.org/10.1136/bmj.h6467

Weidmann, N. B., Otto, S., & Kawerau, L. (2018). The use of positive words in political science language. *Political Science & Politics, 51*(3), 625–628.

Wen, J. U., & Lei, L. (2021). Linguistic positivity bias in academic writing: A large-scale diachronic study in life sciences across 50 years. *Applied Linguistics*. https://doi.org/10.1093/applin/amab049