The linguistic positivity bias in English literature of different genres: A diachronic study across 700 years

Yaqian Shi[*]

1School of Foreign Languages, Huazhong University of Science and Technology, Wuhan, China

[*]**Corresponding author Email: shiyaqian0719@163.com**

Abstract

Previous studies have found that human beings have a positive tendency, which is labelled as the linguistic positivity bias. Such a bias has been examined and confirmed in literary works such as children and youth literature. However, such literature works are written for children and adolescence and thus they are naturally more positive. It is necessary to examine the bias based on literature of other genres that are written for adults. In addition, it remains unexplored how the positivity bias has evolved over a long period of time and how the social and cultural factors interact with literature works to shape the bias. To address the aforementioned issues, the present study aims to examine the linguistic positivity bias in English literature across three genres (i.e., novels, plays, and poetries) and their temporal trend across 700 years. The results show that most of the literature works have a positive tendency and no significant difference is found among the three genres regarding the sentiment values, which further confirms the linguistic positivity bias in human languages. However, it is worth noting that the linguistic positivity bias displays a decreasing trend. Possible explanations are provided for such a trend from social and cultural perspectives.

Keywords:linguistic positivity bias; English literature; genres; sentiment analysis

1 Introduction

The linguistic positivity bias, also known as the Pollyanna hypothesis (Augustine et al., 2011) or the Pollyanna principle (Armstrong & Hogg, 2001), refers to a universal human tendency to use a wide range of positive words more frequently (Dodds et al., 2015). One possible reason for the positivity bias is that human beings are generally optimistic and prefer to look at the bright or positive side of life (Boucher & Osgood, 1969). Another possible reason is that positively biased utterances can facilitate social relationships and help individuals achieve their communicative purposes (Garcia et al., 2012). As a result, they may unconsciously use more positive words in communications to convey their positive feelings such as happiness (Kloumann et al., 2012).

The linguistic positivity bias in human languages have been examined in various contexts (Augustine et al., 2011; Cao et al., 2021; Wen & Lei, 2021; Yuan et al., 2025). One area of research is to investigate the linguistic positivity bias across different languages (Augustine et al., 2011; Boucher & Osgood, 1969; Garcia et al., 2012). For example, Armstrong (2001) investigated the linguistic positivity bias in French. The result showed that young speakers of French tended to create more positive words while the use of new negative words remained stable, which validated the Pollyanna hypothesis. Similarly, the linguistic positivity bias has been examined and confirmed in other languages such as English, German, and Spanish (Augustine et al., 2011; Garcia et al., 2012), which further lend support to the universality of the linguistic positivity bias in human languages.

Another area of research is to investigate the linguistic positivity bias from the cognitive perspective (Jacobs, 2015; Lüdtke & Jacobs, 2015). For example, Lüdtke and Jacobs (2015) compared the processing speed of positive and negative words based on the word recognition task. Thirty-six native adult Germans were recruited to complete the task. They found that positive words have faster response times than neutral or negative ones, which indicated that positive words are cognitively easier to process. The same result is also obtained in 6–12-year-old children (Sylvester et al., 2016), which further confirmed the findings made in previous studies. The result can be explained by the higher subjective exposure frequency of positive words, which leads positive words to be processed much faster since they are more internally activated in memory (Jacobs et al., 2020).

Still another line of research focuses on the linguistic positivity bias across different genres such as academic writing and literature (Cao et al., 2021; Jacobs et al., 2020; Vinkers et al., 2015; Wen & Lei, 2021). For example, Vinkers et al. (2015) explored the use of positive and negative words in the abstracts of medical articles published between 1974 and 2014. Their analysis revealed an 880% rise in positive words and a 257% increase in negative words over the last 40 years. The result showed that researchers tended to use more positive words in their abstracts to promote their research, which demonstrated the linguistic positivity bias in academic writing. As a follow-up study, Wen and Lei (2021) examined the abstracts of 12 research disciplines in life sciences based on both lexicon analysis and sentiment analysis. The results showed a higher frequency of positive words in abstracts based on the lexicon analysis and a positive sentiment based on the sentiment analysis. In addition, the results also showed a significant upward trend of linguistic positivity bias over the past five decades, which further confirmed the linguistic positivity bias in academic writing.

More recently, researchers shifted their attention to linguistic positivity bias in literature (Dodds et al., 2015; Jacobs et al., 2020). For example, Dodds et al. (2015) examined the linguistic positivity bias in several languages such as English, French, Spanish, and Chinese based on 24 corpora, one of which is Google Books Project that includes many world-famous literature works. They investigated the use of positive and negative words in literature with a special word list. They found that positive words were significantly more frequent than negative ones in those literature under investigation, which demonstrated a positivity bias in literature. However, one possible issue of Dodds et al.'s (2015) study is its limited number of words in the word list, which includes only the most frequent 5–10,000 words. Therefore, their findings may not be con-

clusive. To address the issue, Jacobs et al. (2020) examined the linguistic positivity bias in literature based on the technique of sentiment analysis. To be more specific, they calculated the sentiment values of the English and German children and youth literature from the Gutenberg Literary English Corpus and the child-Lex corpus. Results of their study showed a positivity trend in those literature works, which added further evidence to the universality of the linguistic positivity bias.

Although the studies mentioned above have greatly enhanced our understanding of linguistic positivity bias in literature, their results remain incomplete and need further exploration. One issue is that most of the previous studies have not considered the factor of genres when examining the linguistic positivity bias in literature. Some studies such as Dodds et al. (2015) employed mixed genres of literature, which may confound the results. Other studies such as Jacobs et al. (2020) examined only one genre of literature, i.e., children and youth literature. Such literature works are written for children and adolescents who are in a critical period of value formation. It follows that they often convey positive information and conforms to the linguistic positivity bias. Therefore, the findings based on children and youth literature reveal only the features of children and youth literature and may not be generalized to literature of other genres. To paint a fuller picture of the linguistic positivity bias in literature, it is necessary to examine literature of other genres such as poetries and novels that are written for adults.

Another issue is that few of previous studies has examined the diachronic development of the linguistic positivity bias over a long period of time. Although previous studies have investigated the temporal trend of the positivity bias in academic writings, it remains underexplored how such a bias develops in other genres such as literature. Literature works are products of the age, which are significantly influenced by the social and cultural environments in which they were written. Therefore, it is interesting to diachronically investigate the trend of the linguistic positivity bias in literature based on a large dataset over a long period of time and explore how it interacts with the social and cultural factors (Iliev, et al., 2016; Iliev & Smirnova, 2016).

Still another issue is the method used to examine the linguistic positivity bias. Most of the previous studies examined the positivity bias based on a limited set of positive and negative words (except a few studies such as Jacobs et al., 2020 and Wen & Lei, 2021). That is, they examined the normalised frequency of positive and negative words used in texts based on a small predefined lexicon (Cao et al., 2021; Dodds et al., 2015). For example, Vinkers et al. (2015) investigated only 50 positive and negative words. The more frequent use of positive words indicates a positive sentiment and thus a linguistic positivity bias in the texts. However, this method is not a robust approach since it only includes a small number of words. In addition, they did not consider the intensity of positivity or negativity. For example, "bad" and "notorious" are negative words, but the latter word is obviously more negative than the former one. As a result, the findings may not be conclusive enough. Therefore, more sophisticated approaches are required to offer stronger and more reliable evidence for the linguistic positivity bias in human languages.

To address the aforementioned issues in previous studies, the present study aims to investigate the linguistic positivity bias in literature of different genres (i.e., novels, plays, and poetries) and their diachronic trend based on the new technique of sentiment analysis. More specifically, the purpose of the study is twofold. First, it intends to calculate the sentiments of the literature of three genres, i.e., novels, plays, and poetries, and examine whether the linguistic positivity bias still holds true for them. Second, it aims to explore the diachronic trend of the linguistic positivity bias in literature over a period time of approximately 700 years (from 1321 to 20th century).

2 Sentiment Analysis

Recently, a new emerging method for the examination of linguistic positivity bias is sentiment analysis (Jacobs et al., 2020; Wen & Lei, 2021). Sentiment analysis is a technique employed to evaluate the sentiment conveyed by a text, such as whether it is positive or negative (Mäntylä et al., 2018; Taboada, 2016). There are two primary methods for conducting sentiment analysis: machine learning-based approaches and lexicon-based approaches (Liu & Lei, 2018). The machine learning-based method is a classification approach that relies on a classifier trained with human-labelled sentiment datasets to identify the sentiment of a given sentence or text (D'Andrea et al., 2015). This approach has a better accuracy in determining the sentiment of a sentence or a text. However, it has a limited applicability across different domains since sentiment-labelled data are not readily available. The lexicon-based approach calculates the sentiment values or scores of a text based on a large lexicon of sentiment words that have been manually annotated with sentiment scores (van Houtan et al., 2020). When the sentiment value or score of a text is larger than zero, its sentiment is positive. In contrast, when the sentiment value or score of a text is less than zero, its sentiment is negative. Compared to the machine learning-based approach, this approach has a wider application in texts of various genres and domains since the lexicon of sentiment words is field or domain neutral (Taboada, 2016).

Currently, some sentiment analysis tools have been developed for sentiment analysis, such as the packages of syuzhet (Jockers, 2017) and sentimentr (Rinker, 2021) in programming language R. They are developed based on many sentiment lexicons that consider the intensity of positivity and negativity. For example, the lexicon of syuzhet includes more than 10,000 commonly used words and their sentiment values vary from -1 to 1. A positive value means a positive emotion while a negative value indicates a negative emotion. The larger the absolute sentiment value of a word is, the more positive/negative the word is. Due to its large lexicon and wide applicability, syuzhet has been employed in various lines of research regarding sentiment analysis (Liu & Lei, 2018; Vergeer, 2020). However, one possible concern in syuzhet is that it does not consider valence shifters such as negators, which thus may lead to errors in the sentiment analysis of a sentence (Rinker, 2021). For example, the sentence of "He likes playing football" is positive and with a negator, the sentence "He doesn't like playing football" becomes negative. Nevertheless, both sentences may be tagged as positive based on syuzhet since it fails to consider the negator, which makes the sentiment analysis less accurate. To address the concern, sentimentr that takes valence shifters into account has been developed as an alternative for sentiment analysis (Rinker, 2021), which can produce more accurate and valid results. Due to its satisfactory performance in sentiment analysis, sentimentr has been used in various domains (Idler et al., 2022; Ikoro et al., 2018; Roberts et al., 2022).

In this study, we will take the lexicon-based sentiment analysis for examining the linguistic positivity bias in literature due to its larger lexicon, domain neutral application, and more sophisticated algorithms. When conducting the sentiment analysis, we will adopt the package of sentimentr since it considers valence shifters and thus has a more satisfactory performance.

3 Methodology

In this section, we will first introduce the research questions that this study attempts to answer. Then, we will describe the data used in this study and the procedures of data processing and analyses.

3.1 Research questions

(1)   What is the sentiment in English literature of the three genres, i.e., novels, plays, and poetries? Do they conform to the linguistic positivity bias?

(2)   How has the linguistic positivity bias in English literature evolved in the past 700 years?

3.2 Corpus data

The data used in this study are English literature from The Corpus of the Canon of Western Literature (Green, 2017)[1]. This corpus consists of many canonical literary works produced by authors from the western countries such as Italy, France, Great Britain, Germany, and United States. These canonical literature works have great aesthetic beauty and play an important role in shaping other literature, as well as western thoughts and cultures (Adler & Weismann, 2000; Leavis, 1948). Based on the cultural and stylistic features in each era, the canonical literature works are categorized into four chronological ages: (1) The theocratic age (2000 BCE to 1321 CE); (2) The aristocratic age (1321 CE to 1832 CE); (3) The democratic age (1832 CE to 1900 CE); (4) The chaotic age (20th Century). In addition, the literature works are tagged with genres such as fictions, plays, poetries, and mixed genres. The aforementioned features of The Corpus of the Canon of Western Literature make them particularly appropriate for the present study since 1) it contains different genres of literature, and 2) more importantly, it covers a long span of time, which allows for a diachronic study of the linguistic positivity bias.

Table 1 The statistics of English literature used in this study

Ages	Time span	Country/region	No. of texts	No. of words
The aristocratic age	1321 to 1832	England and Scotland	143	14,416,044
The democratic age	1832 to 1900	Great Britain	191	19,287,528
		United States	80	7,734,357
The chaotic age	20th century	Great Britain and Ireland	84	5,937,856
		Australia and New Zealand	3	212,723
		United States	41	1,889,639
Total	/	/	542	49,478,147

In this study, we only used the literature written in English (see Table 1 for the statistics). It should be noted that this corpus has no English literature at the theocratic age and thus the evolution of the linguistic positivity bias in this period is not investigated. That is, we explore the evolution of the linguistic positivity bias in literature across three ages (approximately 700 years), i.e., the aristocratic age, the democratic age, and the chaotic age.

The factor of genre may also confound the results since literature of different genres convey different emotions (Kim & Klinger, 2018). Therefore, we investigate the evolution of the linguistic positivity bias in English literature of three different genres, i.e., fictions, plays, and poetries (see Table 2 for the statistics). We choose the three genres for two reasons. First, they are main and common genres in literature that have a wide range of readers. Second, these three genres occupy the largest proportion in the corpus, which can secure a wide number of texts for the follow-up study. In addition, the consideration of genres can also ensure the validity and reliability of our results.

---

1   https://www.dropbox.com/s/xtv2r37ytfc9pp7/Corpus%20of%20the%20Canon%20of%20Western%20Literature%20%281.0%29.rar?dl=0

Table 2 The statistics of English literature of four genres, i.e., fictions, plays, and poetries

Ages	Fictions		Plays		Poetries	
	No. of texts	No. of words	No. of texts	No. of words	No. of texts	No. of words
The aristocratic age (1321 to 1832)	27	3452368	31	1354638	24	1838910
The democratic age (1832 to 1900)	112	15387092	8	293861	71	4459031
The chaotic age (20th century)	64	6636013	11	286525	34	389027
Total	203	25880849	50	1935024	129	6686968

3.3 Data processing and analyses

The English literature in The Corpus of the Canon of Western Literature were processed and analysed as follows.

First, the texts were cleaned with information such as Gutenberg terms of use, license, and notes deleted.

Second, the linguistic positivity bias in English literature was examined across the three genres (i.e., fictions, plays, and poetries). That is, we calculated the sentiment values of texts of English literature across the three genres based on the sentiment analysis. Then, outliers were detected. A text was considered an outlier and excluded from further analysis if its sentiment value was deviated by three standard deviations from the mean (Shiffler, 1988). Two outliers within the poetry genre were identified and removed prior to further analysis. Subsequently, a one-way ANOVA was performed to determine whether there were significant differences among the three genres. In cases where significant results were found, Games–Howell post hoc tests were applied to assess differences between paired genres.

Finally, the linguistic positivity bias in English literature was analysed across three historical periods. Sentiment values for the literature from these periods were calculated with sentiment analysis techniques. After identifying and removing three outliers, the remaining texts were used for further analysis. A one-way ANOVA was then conducted to assess whether significant differences existed among the three periods. If significant results were found, Games–Howell post hoc tests were applied to explore differences between paired periods.

The work mentioned above was completed with homemade scripts in the programming language R. The sentiment analysis was conducted based on the R package of sentimentr (Rinker, 2021).

4 Results

In this section, we will first report the sentiment results of English literature across the three genres and then present the results across the three ages.

4.1 Sentiment analysis results of English literature across the three genres

The descriptive statistics of the sentiment values of English literature in the three genres are shown in Table 3 and Figure 1. The results show that most of the literature in the three genres have a positive sentiment since most of their sentiment values are larger than zero. In addition, the result of one-way ANOVA analysis

shows no significant difference among the three genres (df = 2, F = 0.907, p = 0.405). That is, the sentiment values of the three genres have no significant difference.

Table 3 The descriptive statistics of the sentiment values of English literature across the three genres

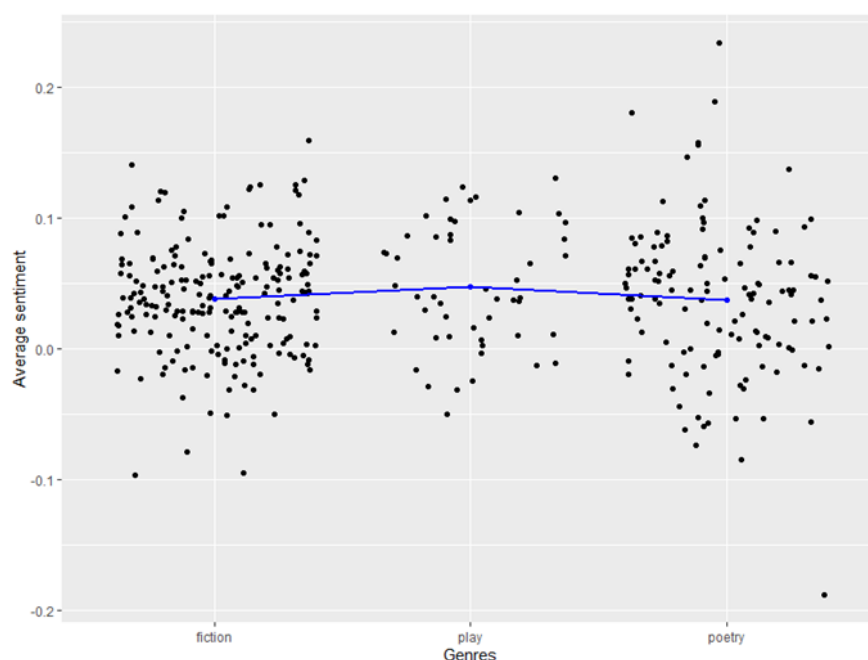| Genres	Max	Min	Mean	S.D.
Fiction	0.160	-0.097	0.038	0.043
Poetry	0.130	-0.050	0.047	0.047
Play	0.234	-0.188	0.037	0.058

Figure 1 Sentiment results across the three genres (i.e., fictions, plays, and poetries)

4.2 Sentiment analysis results of English literature across 700 years

The descriptive statistics of the sentiment values of English literature across the three ages are shown in Table 4 and Figure 2. Two points are worth noting regarding the results. First, most of the literature works have a positive sentiment in the three ages since most of their sentiment values are larger than zero. For example, there are 143 literature works at the aristocratic age, of which 134 have a positive sentiment (i.e., their sentiment values are larger than zero) and nine have a negative sentiment (i.e., their sentiment values are less than zero). In addition, as shown in Table 4, the mean sentiment values of English literature are all larger than zero in the three ages. Second, the sentiment values of the literature display a decreasing trend over the three ages. For example, the mean sentiment value of the literature at the aristocratic age is 0.061 while those at the democratic and chaotic ages are 0.052 and 0.015, respectively.

Table 4 The descriptive statistics of the sentiment values of English literature across the three periods

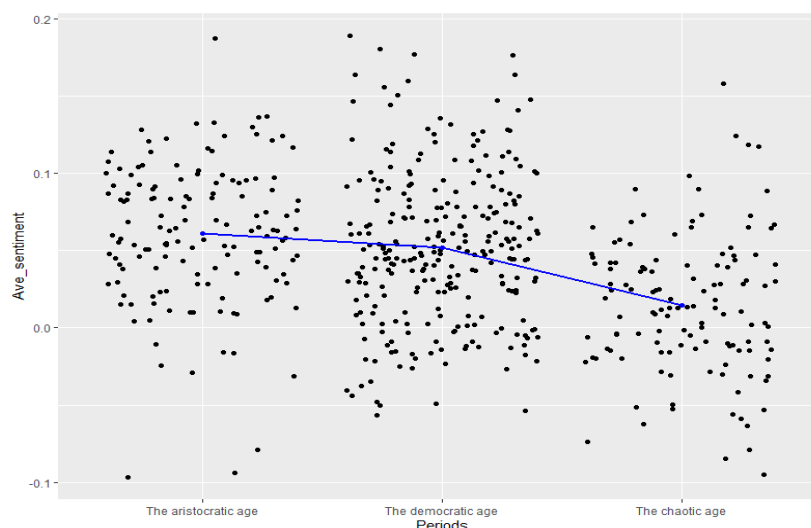| Periods	Max	Min	Mean	S.D.
The aristocratic age	0.187	-0.097	0.061	0.045
The democratic age	0.189	-0.057	0.052	0.049
The chaotic age	0.158	-0.095	0.015	0.044

Figure 2 Sentiment values of the English literature across the three periods

The results of one-way ANOVA analysis show significant differences among the three ages (df=2, F=37.37, p < 0.001). Results of the post hoc test are shown in Table 5. The results indicate no significant difference between the democratic age and the aristocratic age. However, significant differences are found between the chaotic age and the aristocratic age/the democratic age. That is, the sentiment values are significantly decreasing from the democratic age to the chaotic age.

Table 5 The results of the post hoc analysis of the English literature across the three ages

Paired-age comparisons	p value
The democratic age vs. The aristocratic age	0.140
The chaotic age-The aristocratic age	<0.001
The chaotic age-The democratic age	<0.001

The descriptive statistics of the sentiment values of English literature across the three genres in the three ages are shown in Table 6 and Figure 3. As shown in Table 6, the mean sentiment values of two literature genres, i.e., fictions and plays, are also larger than zero. Interestingly, the mean sentiment value of poetries is approximately zero at the chaotic age. The mean sentiment values of all the three genres also display a decreasing trend over the three ages.

The results of one-way ANOVA analysis show that significant differences exist among the three periods in the genres of fictions (df = 2, F = 12.76, p < 0.001) and poetries (df = 2, F = 18.33, p < 0.001) while no significant difference is found in the genre of plays (df = 2, F = 2.227, p = 0.119). The results of the post hoc test on the fictions and poetries between paired ages are shown in Table 7. The results indicate that the sentiment values of fictions and poetries are significantly decreasing from the democratic age to the chaotic age.

Table 6 The descriptive statistics of the English literature of the three genres across the three periods

|	fictions				plays				poetries			
	Max	Min	Mean	S.D.	Max	Min	Mean	S.D.	Max	Min	Mean	S.D.
The aristocratic age	0.124	-0.097	0.053	0.044	0.130	-0.031	0.057	0.050	0.137	0.011	0.071	0.030
The democratic age	0.160	-0.050	0.046	0.042	0.085	-0.011	0.041	0.035	0.189	-0.057	0.044	0.048
The chaotic age	0.088	-0.095	0.017	0.035	0.073	-0.050	0.024	0.039	0.158	-0.085	0.000	0.051
Total	0.160	-0.097	0.080	0.043	0.130	-0.050	0.047	0.047	0.234	-0.188	0.037	0.058
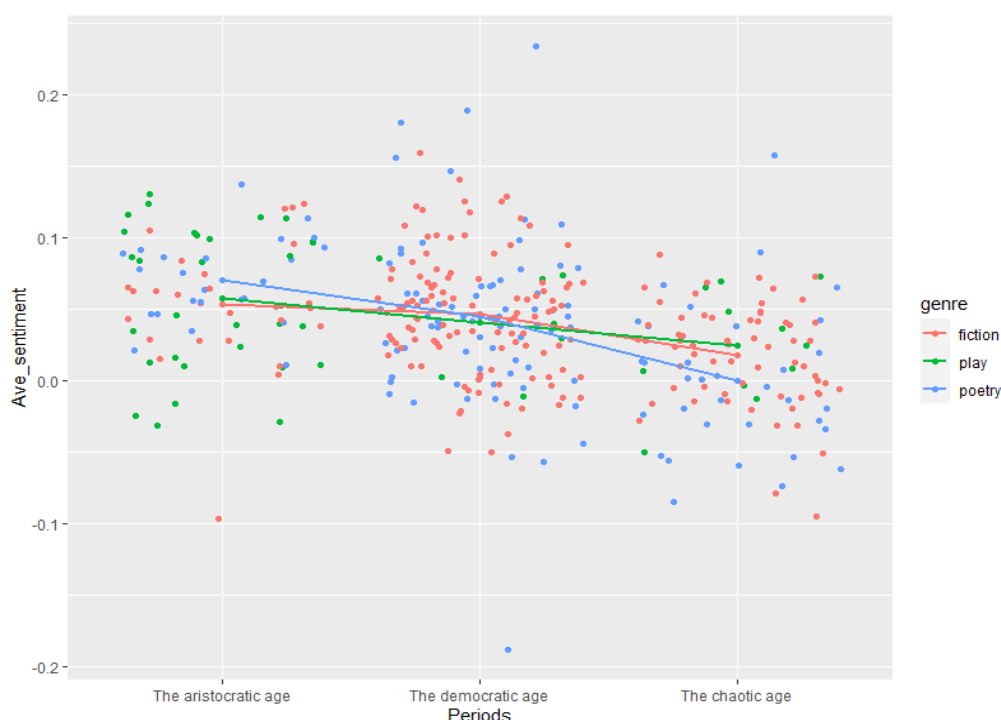
Figure 3 The diachronic trend of the sentiment values of the literature across the three genres in the three periods

Table 7 Results of the post hoc test on the sentiment values of the fictions and poetries between paired periods

Periods for comparison	p-value	
	fictions	poetries
The democratic age vs. The aristocratic age	0.688	0.05
The chaotic age vs. The aristocratic age	<0.001	<0.001
The chaotic age vs. The democratic age	<0.001	<0.001

5 Discussion

This study has examined the sentiments of English literature of three genres (i.e., fictions, plays, and poetries) and investigated their diachronic trend over a period of more than 700 years. The results reveal some points of interest.

First, the results reveal that the three genres show no significant difference in their sentiment values and most of them have a positive sentiment. The results confirm the linguistic positivity bias in these three genres, i.e., fictions, plays, and poetries, which are complementary to findings made in previous studies such as Jacobs et al. (2020). Jacobs et al. (2020) investigated the linguistic positivity bias in English and German children and youth literature works. Our study also found such a bias in other genres such as fictions, plays, and poetries. It indicates that the authors of different genres all tend to convey a positive feeling to readers. The result provides more evidence to the universality of the linguistic positivity bias in human languages.

Second, the results show that most of literature works have a positive sentiment in the three ages. The results also indicate a linguistic positivity bias in English literature, which is in line with the findings made in previous studies such as Jacobs et al. (2020) and Green (2017). For example, Green (2017, p. 289) analysed the most frequent lemmas used in the Corpus of the Canon of Western Literature and found that many of them were positive words such as good, great, and love. Based on the results, he concluded that the canon-

ical literature also shows a linguistic positivity bias despite authors from Homer to Hemmingway mainly address death, war, heartache, and tragedy. Our results further confirm his findings based on the new technique of sentiment analysis.

Third, the results reveal a downward trend of linguistic positivity in English literature over the three ages. In particular, the sentiment values of English literature at the chaotic age are significantly lower than those at the aristocratic and democratic ages. The results do not conform to those found in other genres such academic writings. For example, Wen and Lei (2021) examined the diachronic trend of the linguistic positivity bias in academic writings across 50 years and found a significant upward trend. The main reason for such an inconsistency is attributed to the data type used in these two studies. The data used in their study are academic writings. One of the purposes of academic writing is to "sell" the research to the journals. Therefore, writers need to use some linguistic devices such as the use of positive words to promote their research.

In contrast, the data used in our study are literature works. One of the purposes of literature works is to express the writers' thoughts, convey their emotions, and more importantly show the real world at their time (Fialho, 2019; Milner, 2005; Nurhamidah, 2019). That is, literary works are closely related with the social conditions in which those literary works are created. The aristocratic age (from 1321 to 1832) involves the Renaissance and Reformation period, the Enlightenment period, and the Romantic period, during which writers aimed to recover the philosophy in ancient Greece and Rome (Payne, 2014), emphasize reason and science (Hankins, 1985; Pinker, 2018), and focus on nature, imagination, and individuality (Izenberg, 1992). Therefore, during these periods, humanism, nature, and stories from classical mythology became the topics of literary creation. The democratic age (from 1832 to 1900) involves the Victorian period, during which writers longed for the morality of the medieval world (Timko, 1975). On the contrary, the chaotic age (the 20th century) involves the Modern Period in which the World Wars broke out. The society experienced enormous changes, such as the collapse of social order, the loss of faith in rational theory, and lack of authority and religion (Altaf, 2022). In this period, writers focused more on the inner feelings due to the disillusionment with the World Wars (Altaf, 2022; Yousef, 2017). As a result, the literary works in this period were more negative since individuals were more depressed and pessimistic due to the chaos and unrest in society (Van Hulle, 2018). For example, Eliot' plays focused on moral and religious issues and entertained the audience with shrewd social satire (Lehman, 2009), which resulted in a negative sentiment in his works. It follows that the average sentiment values at the chaotic age are decreasing, which are significantly lower than those at the aristocratic and democratic ages.

This study has some theoretical implications for the hypothesis of the linguistic positivity bias. Our results indicate a more complex mechanism behind the linguistic positivity bias. It has been hypothesized that human beings tend to use more positive words since they often look at the bright side of the life (Augustine et al., 2011). That is, the linguistic positivity bias is driven by our cognition. Based on this hypothesis, the linguistic positivity bias should present an increasing trend or remain the similar level. However, our results show a decreasing trend of the bias over the past 700 years. This finding may reveal that other factors such as cultural and social ones also contribute to the linguistic positivity bias (Iliev, et al., 2016). That is, the linguistic positivity bias may be shaped by both the internal factor such as our cognition and the external factors such as social and cultural ones. Therefore, it is necessary to consider mixed factors when researchers investigate the linguistic positivity bias and explore their complex relationships behind such a bias.

6 Conclusion

This study examined the linguistic positivity bias in literature of three genres (i.e., novels, plays, and poetries) and explored their diachronic trend across 700 years. The results show that most of the literature works have a positive sentiment and no significant difference is found among the three genres, which further confirm the linguistic positivity bias in the field of literature. However, it is worth noting that the linguistic positivity bias displays a decreasing trend. This may be attributed to the social and cultural factors. The finding indicates a more complex mechanism of linguistic positivity bias. That is, the linguistic positive

bias may be not only driven by our cognition but also shaped by other external factors such as social and cultural ones.

This study may be limited in the following aspects. First, it examined only three genres. Future research may examine the linguistic positivity bias based on other genres of literature such as prose and drama since different genres may convey different emotions (Kim & Klinger, 2018). Second, this study investigated the temporal trend of the linguistic positivity bias based on literature. The findings may not be generalized to other genres or registers. Future studies may diachronically examine the bias based on other written corpora such as the Corpus of Historical American English since it includes other genres such as newspaper and magazines. In addition, researchers may also explore its temporal trend based on the spoken corpora since they may better reflect the feelings or emotions of the general public at that time. Finally, the literature used in this study only cover the first half of the 20th century. Future research may investigate the temporal trend of the linguistic positivity bias in recent decades since the society has changed significantly due to the rapid development of science and technology. It is of interest to investigate how the cognitive factor and social factors interact with each other to shape the trend of the linguistic positivity bias in recent decades.

Funding

Conflict of interest disclosure

The author(s) declared that there are no conflicts of interest with respect to the authorship or the publication of this article

References

Adler, M. J., & Weismann, M. (2000). How to think about the great ideas: From the great books of western civilization. Chicago: Open Court.

Altaf, H. (2022). Modernism and Disillusionment: A Study of the Prominent Works of the British Modern Period. The Journal for English Language and Literary Studies, 12(3), 10-10.

Armstrong, N., & Hogg, C. (2001). The Pollyanna Principle in French: A study of variable lexis. Journal of Pragmatics, 33(11), 1757–1785. https://doi.org/10.1016/S0378-2166(00)00081-3

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.

Bloom, H. (1994). The western canon: The books and school of the ages. Boston: Houghton Mifflin Harcourt.

Boucher, J., & Osgood, C. E. (1969). The pollyanna hypothesis. Journal of Verbal Learning and Verb Behavior, 8, 1–8.

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.

D'Andrea, A., Ferri, F., Grifoni, P., & Guzzo, T. (2015). Approaches, tools and applications for sentiment analysis implementation. International Journal of Computer Applications, 125(3), 26–33.

Dodds, P. S., Clark, E. M., Desu, S., Frank, M. R., Reagan, A. J., Williams, J. R., Mitchell, L., Harris, K. D., Kloumann, I. M., Bagrow, J. P., Megerdoomian, K., McMahon, M. T., Tivnan, B. F., & Danforth, C. M. (2015). Human language reveals a universal positivity bias. In Proceedings of the National Academy of Sciences of the United States of America, 112(8), 2389–2394.

Fialho, O. (2019). What is literature for? The role of transformative reading. Cogent Arts & Humanities, 6(1), 1692532.

Garcia, D., Garas, A., & Schweitzer, F. (2012). Positive words carry less information than negative words. EPJ Data Science, 1(3), 1–12.

Green, C. (2017). Introducing the Corpus of the Canon of Western Literature: A corpus for culturomics and stylistics. Language and Literature, 26(4), 282–299. https://doi.org/10.1177/0963947017718996

Hankins, T. L. (1985). Science and the Enlightenment. Cambridge: Cambridge University Press.

Idler, E., Bernau, J. A., & Zaras, D. (2022). Narratives and counter-narratives in religious responses to COVID-19: A computational text analysis. PloS one, 17(2), e0262905.

Ikoro, V., Sharmina, M., Malik, K., & Batista-Navarro, R. (2018, October). Analyzing sentiments expressed on Twitter by UK energy company consumers. In 2018 Fifth international conference on social networks analysis, management and security (SNAMS) (pp. 95-98). IEEE.

Iliev, R., Hoover, J., Dehghani, M., & Axelrod, R. (2016). Linguistic positivity in historical texts reflects dynamic environmental and psychological factors. Proceedings of the National Academy of Sciences, 113(49), E7871-E7879.

Iliev, R., & Smirnova, A. (2016). Revealing word order: Using serial position in binomials to predict properties of the speaker. Journal of Psycholinguistic Research, 45, 205-235.

Izenberg, G. N. (1992). Impossible Individuality: Romanticism, Revolution, and the Origins of Modern Selfhood, 1787-1802. Princeton: Princeton University Press.

Jacobs, A. M. (2015). Neurocognitive poetics: Methods and models for investigating the neuronal and cognitive-affective bases of literature reception. Frontiers in Human Neuroscience, 9, 186.

Jacobs, A. M., Herrmann, B., Lauer, G., Lüdtke, J., & Schroeder, S. (2020). Sentiment Analysis of Children and Youth Literature: Is There a Pollyanna Effect? Frontiers in Psychology, 11, 574746.

Jockers, M. L. M. (2017). Syuzhet: Extracts sentiment and sentiment-derived plot arcs from text. https://CRAN.R-project.org/packhttps://cran.r-project.org/package=syuzhet

Kim, E., & Klinger, R. (2018). A survey on sentiment and emotion analysis for computational literary studies. arXiv preprint arXiv:1808.03137.

Kloumann, I. M., Danforth, C. M., Harris, K. D., Bliss, C. A., & Dodds, P. S. (2012). Positivity of the English language. PloS One, 7(1), e29484.

Leavis, F. R. (1948). The great tradition: George Eliot, Henry James, Joseph Conrad. London: Chatto & Windus.

Lehman, R. S. (2009). Eliot's Last Laugh: The Dissolution of Satire in" The Waste Land". Journal of Modern Literature, 65-79.

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.

Lüdtke, J., & Jacobs, A. M. (2015). The emotion potential of simple sentences: Additive or interactive effects of nouns and adjectives? Frontiers in Psychology, 6, 1137.

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.

Milner, A. (2005). Literature, culture and society. New York: Psychology Press.

Nurhamidah, I., Purwanto, S., & Ekaningsih, N. (2019). A literary work as self-reflection of the author: Why and how it is manifested. EduLite: Journal of English Education, Literature and Culture, 4(2), 194-203.

Payne, M. (2014). Renaissance or Reformation?. A Companion to British Literature, 70-81.

Pinker, S. (2018). Enlightenment now: The case for reason, science, humanism, and progress. London: Penguin UK.

Rinker, T. (2021). Calculate text polarity sentiment. https://cran.r-project.org/web/packages/sentimentr/sentimentr.pdf

Roberts, A., Jellicoe, M., & Fox, K. (2022). How does a move towards a coaching approach impact the delivery of written feedback in undergraduate clinical education?. Advances in Health Sciences Education, 1-15.

Shiffler, R. E. (1988). Maximum Z scores and outliers. The American Statistician, 42, 79–80.

Sylvester, T., Braun, M., Schmidtke, D., & Jacobs, A. M. (2016). The Berlin Affective Word List for Children (kidBAWL): Exploring Processing of Affective Lexical Semantics in the Visual and Auditory Modalities. Frontiers in Psychology, 7, 969. https://doi.org/10.3389/fpsyg.2016.00969

Taboada, M. (2016). Sentiment Analysis: An Overview from Linguistics. Annual Review of Linguistics, 2(1), 325–347.

Timko, M. (1975). The Victorianism of Victorian Literature. New Literary History, 6(3), 607-627.

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. https://doi.org/10.1016/j.patter.2020.100005

Van Hulle, D. (2018). Introduction Negative Modernism: Beckett's Poetics of Pejorism and Literary Enactment. Beckett and Modernism, 1-18.

Vergeer, M. (2020). Artificial intelligence in the Dutch press: An analysis of topics and trends. Communication Studies, 71(3), 373–392. https://doi.org/10.1080/10510974.2020.1733038

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. Clinical Research, 351, h6467. https://doi.org/10.1136/bmj.h6467

Wen, J., & Lei, L. (2021). Linguistic positivity bias in academic writing: A large-scale diachronic study in life sciences across 50 years. Applied Linguistics. Advance online publication. https://doi.org/10.1093/applin/amab037

Yousef, T. (2017). Modernism, postmodernism, and metamodernism: A critique. International Journal of Language and Literature, 5(1), 33-43.

Yuan, Z. M., Jiang, C., & Yao, M. (2025). The art of positive reporting in media discourse: Time magazine's linguistic positivity bias over the 20th century. Digital Scholarship in the Humanities, fqaf039.