Applied Linguistics, **2025, 00, 1–21**
https://doi.org/10.1093/applin/amaf054
Advance access publication 16 September 2025
Article

Linguistic positivity in American English: A large-scale diachronic study over the past two centuries

Gui Wang, Zenan Chen, and Bin Shao[*]

School of International Studies, Zhejiang University, Hangzhou, Zhejiang 310058, China

[*]Corresponding author. 866 Yuhangtang Road, Hangzhou, Zhejiang 310058, China. E-mail: binshao@zju.edu.cn

This study examines the shifts in linguistic positivity in general American English from the 1820s to the 2000s using a 430-million-token structured historical corpus. Furthermore, we investigate the effects of external environmental and internal psychological factors on linguistic positivity. Results showed a significant decline in both positive and negative expressions, indicating a broader societal move from emotionality to rationality in the USA over the past two centuries. Our findings also demonstrated that major wars, unemployment rates, and subjective happiness significantly affect changes in positivity. Notably, while major conflicts generally exert downward pressure on the linguistic positivity, broader historical, socio-cultural, and technological contexts can shape and at times counterbalance this effect, as evidenced by the contrasting linguistic responses to the Korean and Vietnam Wars. This study offers valuable insights into how language mirrors and responds to societal values and historical events. By detailing the interaction between language use and various influencing factors, our research suggests that combining word-level and sentence-level linguistic positivity bias analysis might offer potential applications for monitoring collective well-being and societal trends, enabling more responsive interventions.

Introduction

Linguistic positivity bias (LPB) (Rozin et al. 2010; Augustine et al. 2011), or the Pollyanna hypothesis (Boucher and Osgood 1969), reveals that people tend to use positive words more frequently than negative words. This bias has been explored from various perspectives, including cognitive (Boucher and Osgood 1969; Matlin and Stang 1978) and socio-cultural (Augustine et al. 2011; Warriner and Kuperman 2015). From a cognitive angle, Matlin and Stang (1978) argued that positive words appear more frequently because positive information is cognitively privileged—it is easier to remember and more likely to be recalled. From a socio-cultural viewpoint, Augustine et al. (2011) suggested that the prevalence of positive words might result from their role in social facilitation. Using more positive words may lead to more positive social interactions, and those who experience more positive social interactions are likely to use positive words to a greater degree.

There are mainly two lines of investigation into linguistic positivity bias: synchronic and diachronic. Synchronic studies examine this hypothesis across multiple languages to determine if it represents a universal human tendency (Dodds et al. 2015; Diener et al. 2018). By using human evaluation of 100,000 words spread across twenty-four corpora in ten languages diverse in origin and culture, Dodds et al. (2015) observed that the words of natural human language possess a universal positivity bias. In addition, they demonstrated that the positivity bias is largely independent of the frequency of word use, thereby supporting the notion that the linguistic positivity bias is a fundamental aspect of human language.

In recent years, more attention has been paid to the diachronic aspect of linguistic positivity bias, considering that synchronic investigations alone cannot fully determine the stability or variation of this phenomenon over time and context within a given language (Iliev et al. 2016). Relevant studies vary in the genres of time-stamped text corpora they examined, with a predominant focus on academic writings (e.g. Vinkers et al. 2015; Holtz et al. 2017; Wen and Lei 2022; Liu and Zhu 2024). These studies consistently report an increase in the use of positive words or expressions in academic writing over past decades.

For instance, Holtz et al. (2017) examined the change of linguistic positivity bias in research articles from two cross-cultural journals (*Journal of Cross-Cultural Psychology* and *Cross-Cultural Research*). They found a consistent increase in positive framing and a rise in reports of marginally significant statistical findings. Holtz et al. (2017) proposed that these trends might stem from heightened publication pressures and the drive for visibility in a competitive academic environment, a phenomenon often referred to as academic capitalism. By combining a small lexicon analysis and sentiment analysis, Wen and Lei (2022) examined the linguistic positivity bias in a corpus of abstracts published between 1969 and 2019 in more than 100 scientific journals in life sciences. They found a marked increase in the linguistic positivity bias over the five decades examined. In addition to publication pressures, Wen and Lei (2022) identified political correctness as a potential factor. They argued that researchers may consciously use positive or at least non-negative language, especially when describing individuals from certain social groups, adhering to norms of political correctness. This analysis adds depth to our understanding of the factors that contribute to the evolution of linguistic positivity bias in academic writing, highlighting how external societal and cultural pressures can shape scholarly discourse over time.

Contrary to the numerous genre-specific investigations, few studies have explored linguistic positivity within general texts from a diachronic perspective, employing broad corpora. Thus, relatively little is known about how this phenomenon unfolds across a wide variety of texts consumed by the general public. A notable exception is the study by Morin and Acerbi (2017), which utilized the Google Books Ngram corpus (Michel et al. 2011) and two corpora of Anglophone fiction books. They observed a historical weakening in the linguistic positivity bias and tentatively suggested that this bias might not be an invariant feature of language, but rather subject to historical and cultural factors.

Iliev et al. (2016) expanded on the work of Morin and Acerbi (2017) by modeling both environmental factors (war and economy) and subjective factors (subjective happiness) in their investigation of linguistic positivity bias. Specifically, they examined linguistic positivity bias by leveraging two independent, time-stamped text corpora (Google Books Ngram and the New York Times), spanning approximately 200 years. Iliev et al. (2016) used LIWC (408 positive emotion words and 499 negative words) and calculated an LPB index as the ratio of the frequency of positive words to the frequency of negative words for each year to discern the temporary changes in the linguistic positivity bias. Their findings revealed a decline in the linguistic positivity bias in American English over the last two centuries. Furthermore, Iliev et al. (2016) observed that dynamic fluctuations in the linguistic positivity bias were predicted by changes in both the objective environment and national subjective happiness. Their findings provided robust empirical evidence that the linguistic positivity bias is dynamic and influenced by a confluence of subjective, objective, and societal factors.

In short, these studies have significantly advanced our understanding of the diachronic features of linguistic positivity bias in general texts and the potential factors contributing to the dynamics of this phenomenon. However, most of these studies relied on the Google Books Ngram corpus, which despite representing 4 per cent of published books, lacks detailed metadata and is unbalanced across different genres and time periods (Koplenig 2015; Pechenick et al. 2015).

Pechenick et al. (2015) observed that scientific texts have become an increasingly substantive portion of this corpus throughout the 1900s. Consequently, the changes in the linguistic positivity bias might not reflect any real literary, linguistic, or cultural trend, but simply changes in the composition of this corpus over the years (Morin and Acerbi 2017). This underscores the necessity for employing metadata-rich corpora for more accurate analyses. Moreover, while many studies have employed word-level analyses using tools like LIWC and WordNet Affect, few have adopted sentence-level analysis. However, word-level analysis might overlook critical contextual signals (e.g. negators and down-toners) that can reverse or reshape an otherwise positive or negative tone. These signals become clear only at the sentence level, where words interact syntactically and semantically. Thus, investigating changes from a sentence-level perspective could provide a deeper and more accurate insight into the nuances of linguistic positivity bias.

Building on these observations, the present study investigates the diachronic changes in the linguistic positivity bias in general American English by using the Clean Corpus of Historical American English (CCOHA; Alatrash et al. 2020), one of the largest balanced historical English corpora, spanning from the 1820s to the 2000s. We employed both word-level (LPB index) and sentence-level (polarity sentiment) analyses to explore the dynamics of linguistic positivity bias. By utilizing a balanced historical English corpus and investigating multiple levels of linguistic positivity bias, our study aims to further test the idea that the linguistic positivity bias is influenced by environmental and psychological factors. Additionally, we seek to explore the complex interplay between language use, cultural evolution, and shifts in sociality. As suggested by Snefjella et al. (2019), understanding the dynamics of change offers valuable insights into both the origins of present-day culture and society and the mechanisms driving evolution in human communities and their communicative tools.

The significance of the present study lies in the following points. By examining the CCOHA, which encompasses diverse genres of American English across a long time span, we can assess the linguistic positivity bias in contexts that reflect the lived experiences, social values, and public sentiments of the time. Such an investigation goes beyond purely theoretical curiosity; it offers insights into how language practices intersect with pivotal historical moments, such as economic upheavals, major conflicts, and cultural reforms, and how they may, in turn, shape public consciousness.

From a broader social and cultural perspective, understanding the diachronic trends in positivity bias is timely and consequential. The United States has long been a hub of global media influence, with its newspapers, magazines, novels, and digital platforms shaping, reflecting, and circulating dominant cultural narratives. Tracing how positivity in American English has evolved can thus shed light on broader rhetorical and discursive shifts that continue to resonate in the current public discourse. For instance, Wilson et al. (2020) and Levendusky and Malhotra (2016) have highlighted the heightened polarization in contemporary media landscapes, where language choice, tone, and framing can either amplify conflict or foster civility. Examining the historical ebbs and flows of positive language provides a key to contextualizing modern debates, potentially revealing whether the current patterns mark a departure from or a continuation of past trends. In parallel, shifts in social norms and cultural sensitivities, such as the rise of political correctness (Granath and Ullén 2019), underscore how positivity in language can serve as both a mirror and a driver of societal change.

Moreover, a diachronic view of linguistic positivity bias can inform discussions about the emotional and psychological tenor of public life. Recent attention to mental health and collective well-being underscores the importance of recognizing how language both reflects and reinforces prevailing attitudes. Notably, depression ranks among the most common mental disorders in the United States, according to the National Institutes of Health, and recent data indicate that an estimated 21 million adults, i.e. 8.3 per cent of all US adults, experience at least one major depressive episode each year (Center for Behavioral Health Statistics and Quality 2022). Against this backdrop, researchers have increasingly turned to text analysis as a potential tool for proactive screening and early detection of depression (e.g. Neuman et al. 2012; Seabrook et al. 2018). Since our use of language embodies and communicates our attitudes (Lakoff 1975: 45), shifts in positivity bias may not only reflect broader cultural emphases on optimism, self-help, and resilience but also illuminate how evolving discourse practices intersect with mental health trends. Conversely, detecting a decline in positivity during certain historical periods might reveal widespread anxieties or disillusionment resulting from war, economic recessions, or other societal disruptions. By leveraging both word-level and sentence-level sentiment analyses, our study might uncover the nuanced ways in which positivity bias manifests, evolves, and intersects with historical and cultural milieus, thereby providing a richer context for understanding how language use can signal, shape, or even help mitigate mental health concerns. In so doing, we not only build upon existing research by moving beyond genre-specific examinations but also respond to calls for more richly contextualized perspectives on language change.

Specifically, our study addresses the following three research questions:

RQ1: How has the linguistic positivity of American English evolved from the 1820s to the 2000s at word level?
RQ2: How has the linguistic positivity of American English evolved from the 1820s to the 2000s at sentence level?
RQ3: How do environmental (i.e. wars and economic conditions) and psychological factors (i.e. subjective happiness) contribute to the changes in linguistic positivity of American English from the 1820s to the 2000s?

Methods

Corpus data

To investigate the diachronic changes in linguistic positivity in American English, we utilized the CCOHA (Alatrash et al. 2020). The CCOHA offers several advantages over the original Corpus of Historical American English (COHA; Davies 2012), including an increased number of word tokens (431 million compared to 406 million), fewer non-words (64 million compared to 66 million), and a reduced number of invalid tokens. More specifically, there are two main reasons for adopting this corpus.

First, spanning from 1810 to 2009 and encompassing over 431 million words, the CCOHA stands as the largest structured corpus of historical English. It serves as an ideal resource for portraying a comprehensive picture of changes in the linguistic positivity of American English across a long time span. As Biber and Gray (2013: 108) suggested, such a large corpus permits analysis of historical change on a scale previously unachievable with earlier corpus designs, enabling nearly continuous tracking of changes over time. Table 1 presents the tokens for each decade in the CCOHA from the 1820s to the 2000s.

Second, the CCOHA is a structured collection of carefully selected historical English texts sourced from newspapers, popular magazines, fiction, and nonfiction books across decades. For example, fiction accounts for 48–55 per cent of the total in each decade, and the corpus maintains balance across sub-genres and domains (e.g. by Library of Congress classification for nonfiction and by sub-genre for fiction—prose, poetry, drama, etc.) (Davies 2012: 123). This balance across genres and sub-genres allows researchers to examine changes with reasonable certainty that the corpus data can well represent changes in the authentic language use, rather than just being artifacts of differences in genre balance. Therefore, the CCOHA may offer useful features not available in larger, unstructured corpora like the Google Books Ngram corpus.

Data processing

Corpus preprocessing

The present study utilized texts from the years 1820 to 2009 in the CCOHA, as summarized in Table 1. Texts from 1810 to 1819 were excluded due to their limited token count. In addition,

Table 1. Composition of CCOHA by decade.

Decades	Tokens
1820s	6,563,804
1830s	13,068,602
1840s	15,186,641
1850s	15,594,938
1860s	15,987,728
1870s	17,580,971
1880s	19,035,366
1890s	19,243,417
1900s	20,782,902
1910s	21,457,302
1920s	24,169,546
1930s	23,222,619
1940s	22,992,886
1950s	23,180,560
1960s	22,631,770
1970s	22,385,946
1980s	23,708,172
1990s	26,233,779
2000s	27,700,738

we removed sentences containing special characters (i.e. '@') from the corpus. As outlined on the COHA website (https://www.corpusdata.org/), in the offline version of the corpus, ten consecutive tokens of every 200 tokens were replaced with '@' characters for copyright protection. Furthermore, given the large size of the CCOHA, which includes over 100,000 text files, we aggregated all files from the same year into a single, larger file. This consolidation allowed for more efficient processing and handling of the data, ensuring a streamlined approach to our study.

Emotion lexicon selection

To examine changes in the linguistic positivity in American English from the 1820s to the 2000s, we employed the NRC emotion lexicon (Mohammad and Turney 2013) for all three research questions. This lexicon comprises a list of English words and their associations with eight basic emotions (anger, fear, anticipation, trust, surprise, sadness, joy, and disgust), along with two sentiments (negative and positive), annotated through crowdsourcing.

We selected the NRC emotion lexicon for three reasons. First, it is manually sentiment-tagged with great care, which guarantees high accuracy in determining the association between terms and their sentiments. Second, the NRC includes 14,182 emotion words, providing substantial coverage necessary for our comprehensive analysis. Third, this lexicon is designed for general use, unlike other lexicons tailored to a particular genre, such as financial news (the Loughran and McDonald lexicon; Loughran and Mcdonald 2011), customer reviews (the Opinion lexicon; Hu and Liu 2004), or literary texts (the Jockers sentiment lexicon; Jockers 2017). Mohammad and Turney (2013) created this lexicon using words from general language resources (e.g. Macquarie Thesaurus and General Inquirer) rather than domain-specific terminology. The words were selected to represent common English usage across various parts of speech, making the lexicon particularly well-suited for examining the broader spectrum of linguistic positivity across general American English texts.

Linguistic positivity calculation

To discern the temporal changes of linguistic positivity across multiple facets, we adopted two positivity indices: the LPB index (Iliev et al. 2016) and polarity sentiment, which target the word and sentence levels, respectively. For both indices, we used the NRC emotion lexicon.

The LPB index, calculated as the ratio of positive to negative words (Iliev et al. 2016), is a measure used to detect shifts in text positivity over time. A higher LPB index at one time point compared to another indicates an increase in text positivity. To mitigate the potential effect of varied token sizes across each year in the CCOHA, we segmented each yearly combined file into 10,000-word segments and calculated the LPB index for each segment. The final LPB index for each year was then determined as the mean LPB index of all segments for that year.

Additionally, we calculated the normalized frequency of positive and negative words per year to further clarify the underlying drivers behind observed shifts in the LPB index. For instance, an increase in the LPB index alone does not reveal whether this increase results from more frequent positive words, fewer negative words, or a combination of both. Specifically, the normalized frequency is defined as the expected count of positive or negative words per 10,000 words. Consistent with the LPB index calculation, we divided each annual text into uniform 10,000-word segments (discarding any remainder), counted positive and negative words in each segment, and averaged these counts across segments.

In contrast to the LPB index, polarity sentiment assesses positivity at the sentence level. This index is crucial because word-level sentiment analysis may omit contextual nuances, like valence shifters—negators, intensifiers, or down-toners—that significantly alter sentence meaning. For example, the emotion word 'like' in the sentence 'He does not like me' would misleadingly indicate positivity if analyzed without considering the negator 'not.' We utilized the R package *sentimentr* (Rinker 2019) to calculate polarity sentiment. This package is designed to rapidly compute text polarity sentiment in English at the sentence level, taking into account valence shifters. According to Mahmoudi et al. (2024), *sentimentr* consistently performs well across diverse datasets among all thirteen lexicon and rule-based packages in R and Python for sentiment analysis examined in their study.

To manage the variability in sentence counts across years in the CCOHA, particularly in the earlier years, we employed a random sampling approach for the calculation of polarity sentiment. This approach is also more computationally efficient, as processing the entire corpus at the sentence level would require substantially greater computational resources. For each yearly combined file, we randomly sampled 500 sentences and calculated their mean polarity sentiment. This process was repeated 1,000 times (i.e. 1,000 iterations of 500 randomly sampled sentences) for each file to obtain a robust average sentiment value for each year. Notably, we did not control for sentence length in these calculations, as the *sentimentr* algorithm accounts for various factors within sentences, making sentence length a non-critical factor in the sentiment assessment (Rinker 2019: 44–48).

Data analysis

For the first research question, we employed simple linear regression analyses to trace the diachronic trajectories of the normalized frequency of positive/negative words and the LPB index in American English. Specifically, we sought to determine whether there was a significant upward or downward trend in linguistic positivity at the word level in the CCOHA from the 1820s to the 2000s. In these models, the year functioned as the independent variable, and the dependent variable was the normalized frequency of positive/negative words and the LPB index, respectively.

For the second research question, we performed a simple linear regression analysis with the year as the independent variable and polarity sentiment as the dependent variable. This allowed us to detect diachronic changes in positivity at the sentence level in general American English across the examined years.

For the third research question, we examined how environmental and psychological factors influence changes in linguistic positivity. Our investigation was guided by two established hypotheses regarding the mechanisms behind LPB. The first hypothesis suggests that LPB may be driven by objective circumstances. According to this framework, a higher relative frequency of positive words reflects a greater frequency of positive events in speakers' environments (Rozin et al. 2010). The second hypothesis focuses on psychological factors, proposing that the evaluative content of language can mirror people's affective states (e.g. Rude et al. 2004; Tausczik and Pennebaker 2010; Mehl et al. 2017).

We operationalized environmental factors as wars and economic conditions, while psychological factors were measured through subjective happiness. In general, these indices were selected based on the availability of extensive historical records and their relatively clear valence. Wars, for instance, are well-documented in the USA since the 1770s (beginning with the Revolutionary War) and generally hold strong negative connotations that might lower LPB. Acerbi et al. (2013b) observed a 'sad' peak in linguistic content during World War II (WWII) in the Google Books Ngram corpus (covering both British and American English). In a similar vein, economic conditions, particularly inflation and unemployment, are known to profoundly affect collective well-being (Di Tella et al. 2001; Gandelman and Hernández-Murillo 2009; Blanchflower et al. 2014) and may thus influence the overall positivity in written language. Lastly, national subjective happiness offers an individual-level psychological perspective that complements these broader environmental indicators.

For war-related information, we extracted data from the US Department of Veterans Affairs (Department of Veterans Affairs 2023), which details the US military casualties for each war that the USA was engaged in from 1775 to 2000. It is worth noting that this casualty data is presented by war rather than annually. Following Iliev et al. (2016), we assumed a uniform distribution of casualties over the duration of each war for simplicity. For the data from 2001 to 2009, we referred to American War and Military Operations Casualties (Blum 2020), which provided yearly figures for military deaths caused by hostile action. To further determine the impact of major wars on positivity, we focused on those conflicts in the US history with total military casualties exceeding 10,000.

Economic conditions covered three indices: inflation rates, unemployment rates, and the Misery Index. The time-series data for these indices from 1948 to 2009 can be freely accessed from www.miseryindex.us. We chose the Misery Index because it serves as a widely used indicator of short-term economic distress and a reverse measure of economic well-being (Nessen 2008). It is calculated as the sum of the current unemployment rate and the current inflation rate. This index has proven valuable in approximating how macroeconomic conditions influence population well-being, as evidenced by correlations with consumer sentiment (Lovell and Tien 2000), crime rate (Tang and Lean 2009), and suicide rate (Yang and Lester 1992). We included separate analyses of unemployment and inflation rates alongside the Misery Index, which allows us to examine their distinct effects on linguistic positivity.

The psychological factor covered in our study is the national subjective happiness in the USA. To operationalize this construct, we selected the World Database of Happiness (Veenhoven 2024). It is a peer-reviewed repository that aggregates standardized subjective well-being data from over 40,000 research findings (e.g. distributional and correlational studies) as of 2021, with ongoing updates expanding its coverage to become the largest and most comprehensive archive in the field (Veenhoven 2024). For our analysis, we specifically used the US national happiness time series from this database, which provides annual average happiness scores from 1946 to 2009 on a standardized scale of 0–10 (higher scores indicating greater happiness).

The World Database of Happiness was chosen for three key reasons. First, it provides longitudinal coverage (1946–2009) that aligns with our corpus timeline, enabling temporal alignment of linguistic and psychological trends. Second, it standardizes heterogeneous survey items (e.g. life

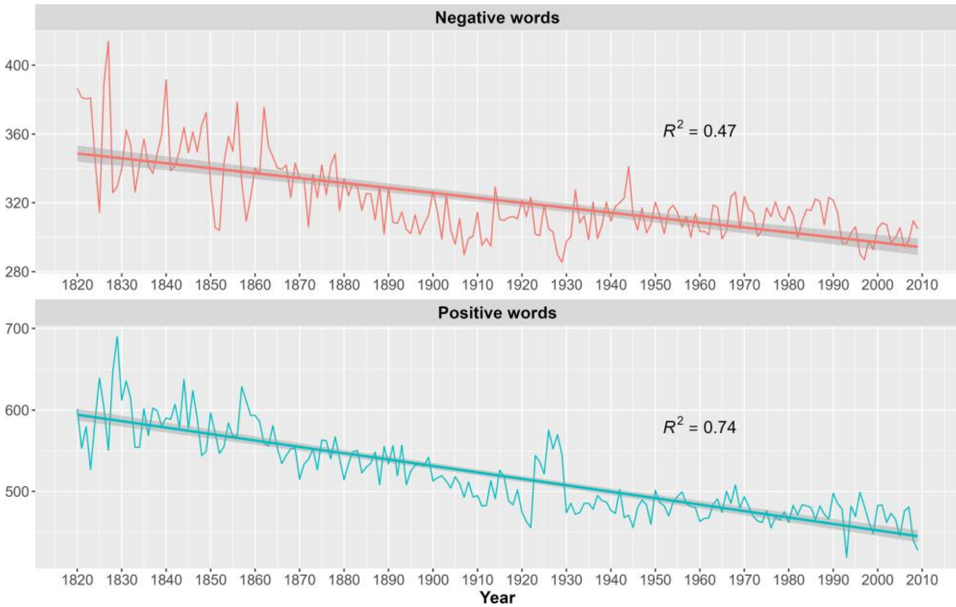

Figure 1. Diachronic changes in the normalized frequency of positive/negative words in the CCOHA.

satisfaction, happiness scales) into a validated 0–10 metric, ensuring cross-temporal comparability and mitigating measurement inconsistency risks inherent in historical happiness research. Third, the World Database of Happiness is widely recognized in interdisciplinary studies of societal well-being, such as environmental psychology (Samavati and Veenhoven 2024) and economics (Fischer 2008), which enhances the interpretability of our findings within broader academic discourse. By including this measure, we bridge psychological states (affective well-being) with linguistic patterns, offering a multidimensional perspective on how collective happiness may co-vary with linguistic positivity over time.

Specifically, we performed three groups of linear models to examine the effects of wars, economic conditions, and subjective happiness on linguistic positivity. It is worth noting that we controlled for the effect of the year when examining the influence of wars, economic conditions, and subjective happiness on positivity. Incorporating time into the models allowed us to test whether there is an observable relationship between positivity indices and the potential factors, independent of the effect of time. However, we did not construct a combined model that includes all three groups of variables (war, economic conditions, and subjective happiness) into a single model, as the data points for each variable cover significantly different time ranges. In particular, we have nearly 200 years of data on wars (1820–2009), while the earliest available data for unemployment rates begin in 1948. Therefore, a combined model would necessitate excluding a substantial portion of the war data, thus diminishing the historical scope of our analysis.

Results

Diachronic changes in linguistic positivity at word level

To explore shifts in positivity in American English at the word level, we examined the diachronic changes in the normalized frequency of positive/negative words, alongside the LPB index, in the CCOHA from the 1820s to the 2000s.

Figure 1 illustrates that both negative and positive words experienced a general downward trend over the past two centuries. In addition, both categories indicate significant decreases with large

Table 2. Statistics of simple linear regressions of year and negative/positive words.

	Variable	Estimate	Standard error	t-value	P-value
Negative words	(Intercept)	870.30	42.47	20.49	8.61×10^{-50}*
	Year	–0.29	0.02	–12.93	9.22×10^{-28}*
Positive words	(Intercept)	2,031.23	64.61	31.44	8.84×10^{-77}*
	Year	–0.79	0.03	–23.40	1.35×10^{-57}*

*<0.001.

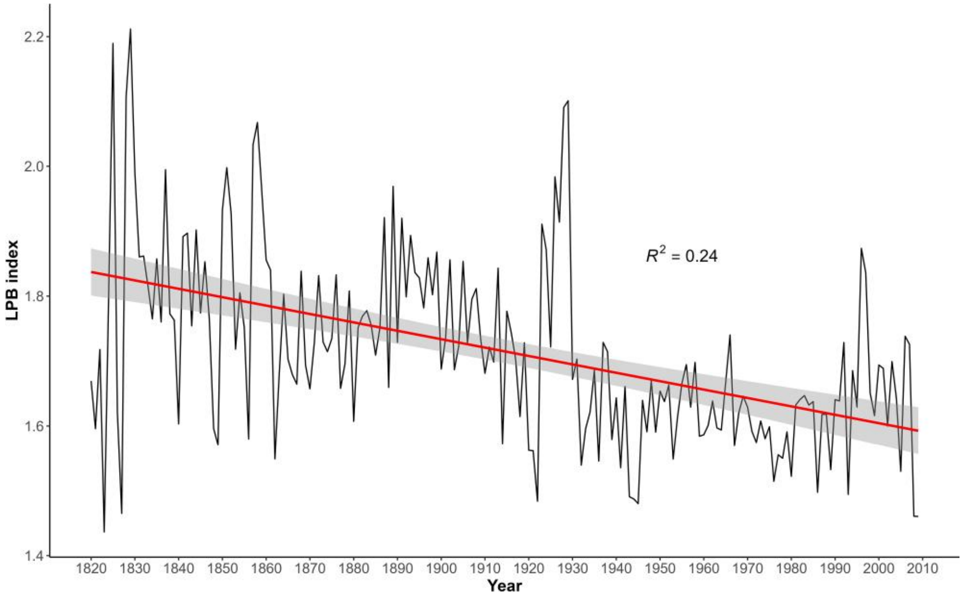

Figure 2. Diachronic changes in the LPB index in the CCOHA.

Table 3. Statistics of simple linear regressions of year and LPB index.

Variable	Estimate	Standard error	t-value	P-value
(Intercept)	4.19080	0.32183	13.02	4.78×10^{-28}*
Year	–0.00129	0.00017	–7.70	7.70×10^{-13}*

*<0.001.

accounting of variation (negative words: $F(1, 189) = 167.09$, $P = 9.22 \times 10^{-28}$, multiple $R^2 = 0.4706$, adjusted $R^2 = 0.4677$; positive words: $F(1, 189) = 547.67$, $P = 1.35 \times 10^{57}$, multiple $R^2 = 0.7445$, adjusted $R^2 = 0.7431$) (see Table 2 for more details). Furthermore, the rate of decline was steeper for positive words (estimate: –0.79 vs. –0.29). These results suggest that the US general audience publications have shown a marked decline in the use of emotion-related terms over the last nearly 200 years, with positive terms decreasing at a faster rate per year than the negative ones.

As shown in Fig. 2, the LPB index follows a similar downward trajectory. Specifically, there is a significant decrease in the LPB index with moderate accounting of variation ($F(1, 189) = 59.23$, $P = 7.70 \times 10^{-13}$, multiple $R^2 = 0.2396$, adjusted $R^2 = 0.2355$) (see Table 3 for more details). In short, these results suggest a historical weakening in linguistic positivity at the word level in

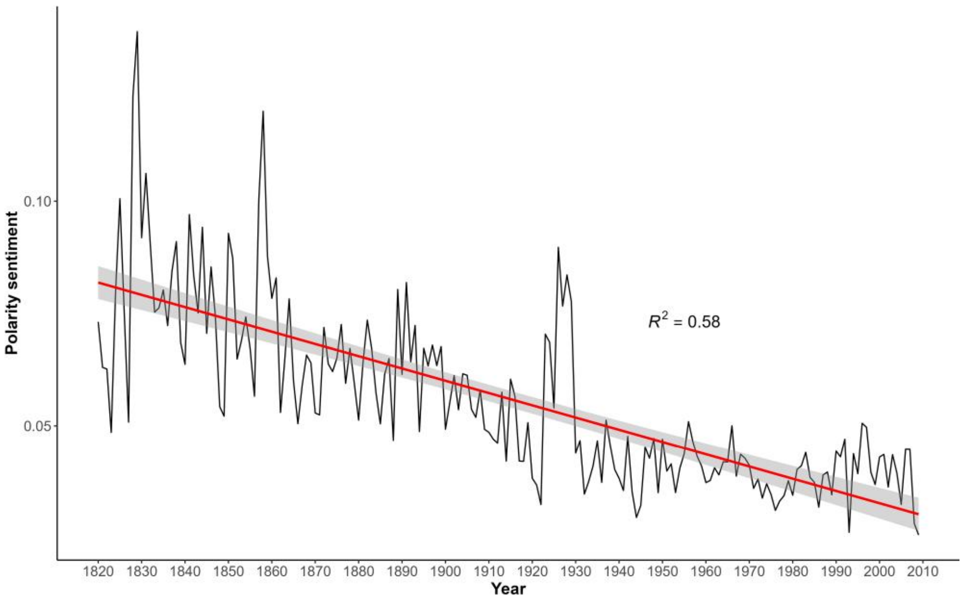

Figure 3. Diachronic changes in the polarity sentiment in the CCOHA.

Table 4. Statistics of simple linear regressions of year and polarity sentiment.

Variable	Estimate	Standard error	t-value	P-value
(Intercept)	0.57904	0.03256	17.79	3.79×10^{-42}*
Year	−0.00027	0.00002	−16.07	3.91×10^{-37}*

*<0.001.

American English. Importantly, this decline is not due to an increase in negative words but rather a decrease in positive words, aligning with findings by Morin and Acerbi (2017). However, our study also noted a significant decrease in negative words, a trend not observed by Morin and Acerbi (2017). This discrepancy may stem from differences in the corpora used between the two studies.

Diachronic changes in linguistic positivity at the sentence level

To investigate the changes in linguistic positivity in American English at the sentence level, we analyzed the diachronic changes in polarity sentiment in the CCOHA. Figure 3 shows that polarity sentiment over the past two centuries follows a general downward trend. This trend parallels the decrease observed at the word level. The results from a simple linear regression model, detailed in Table 4, reveal significant declines in polarity sentiment with large accounting of variation ($F(1, 189) = 258.25$, $P = 3.91 \times 10^{-37}$, multiple $R^2 = 0.5787$, adjusted $R^2 = 0.5765$). These findings suggest that American English has increasingly adopted a more negative tone at the sentence level.

Factors contributing to linguistic positivity

To examine the effect of wars on positivity in American English, we conducted two simple linear regression analyses. Specifically, we modeled the US military casualties as predictors of the LPB index and polarity sentiment.

Table 5. Statistics of simple linear regressions of the US military casualties and LPB index/polarity sentiment.

	Variable	Estimate	Standard error	t-value	P-value
LPB index	(Intercept)	4.17867	0.31453	13.29	8.5×10^{-29}***
	Casualties	-6.3×10^{-7}	2.0×10^{-7}	−3.14	0.0020**
	Year	−0.00128	0.00016	−7.81	4.0×10^{-13}***
Polarity sentiment	(Intercept)	0.57801	0.03207	18.02	9.5×10^{-43}***
	Casualties	-5.3×10^{-8}	2.1×10^{-8}	−2.60	0.0102*
	Year	−0.00027	0.00002	−16.26	1.3×10^{-37}***

*<0.05; **<0.01; ***<0.001.

Table 6. Wars with more than 10,000 US military casualties.

War	Period	Casualties
Mexican War	1846–1848	17,435
Civil War	1861–1865	780,213
World War I	1917–1918	320,518
World War II	1941–1945	1,076,245
Korean War	1950–1953	153,530
Vietnam War	1964–1975	243,523

As shown in Table 5, the US military casualties (1820–2009) significantly predicted the LPB index and polarity sentiment with large accounting of variation (LPB index: $F(1, 188) = 35.94$, $P = 6.18\times10^{-14}$, multiple $R^2 = 0.2777$, adjusted $R^2 = 0.2700$; polarity sentiment: $F(1, 188) = 136.44$, $P = 2.89\times10^{-37}$, multiple $R^2 = 0.5934$, adjusted $R^2 = 0.5890$). These results indicate that higher casualties are associated with lower values of both the LPB index and polarity sentiment, suggesting a more negative tone in texts. This finding aligns with Iliev et al. (2016), who demonstrated that the LPB index decreases with higher casualty counts. Our study extends their findings by examining linguistic positivity at both the word level (LPB index) and the sentence level (polarity sentiment).

To further investigate the effect of wars on positivity in American English, we created two corresponding plots depicting the relationship between major wars and LPB index/polarity sentiment across decades. Table 6 provides details of these major wars, including their periods and casualties, compiled from Blum (2020) and the Department of Veterans Affairs (2023). Figure 3 presents these plots, displaying the z-scores of the mean LPB index and polarity sentiment for each decade on the y-axis. It is worth noting that we marked the *Vietnam War* under the 1970s for convenience, even though it spanned both the 1960s and 1970s.

As shown in Fig. 4, compared to the LPB index of the decades preceding major wars, the LPB index of the decades during which major wars occurred consistently experienced a marked decrease, except for the 1950s during the Korean War. The same pattern was observed in the changes of polarity sentiment related to major wars. These results further confirm that wars may be associated with more negative language use.

To investigate the effect of economic conditions on positivity in American English, we examined the effects of inflation rates, unemployment rates, and the Misery Index on predicting the fluctuation of both the LPB index and polarity sentiment. Table 7 presents the results of simple linear regression models with the three economic conditions as independent variables and the two positivity indices as dependent variables.

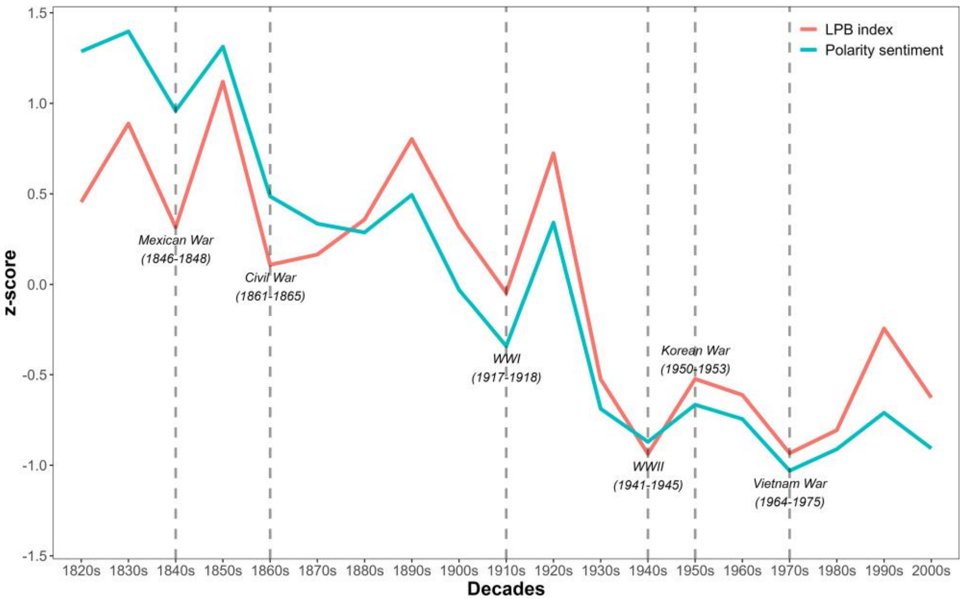

Figure 4. Corresponding plot between wars with more than 10,000 US military casualties and LPB index/polarity sentiment in the CCOHA.

Table 7. Statistics of simple linear regressions of economy conditions and LPB index/polarity sentiment.

	Variable	Estimate	Standard error	t-value	P-value
LPB index	(Intercept)	1.22027	1.07678	1.1333	0.2617
	Inflation rates	−0.00388	0.00335	−1.1596	0.2509
	Year	0.00021	0.00054	0.3887	0.6989
LPB index	(Intercept)	0.97753	1.06344	0.5562	0.5802
	UR	−0.00543	0.00274	−2.4031	0.0194*
	Year	0.00035	0.00054	1.0285	0.3079
LPB index	(Intercept)	0.59753	1.07433	0.9192	0.3617
	Misery Index	−0.01560	0.00649	−1.9786	0.0525
	Year	0.00056	0.00055	0.6541	0.5156
Polarity sentiment	(Intercept)	0.15355	0.07719	1.9893	0.0513
	Inflation rates	−0.00025	0.00024	−1.0229	0.3105
	Year	−0.00006	0.00004	−1.4615	0.1492
Polarity sentiment	(Intercept)	0.13608	0.07609	1.3689	0.1762
	UR	−0.00038	0.00020	−2.6671	0.0099**
	Year	−0.00005	0.00004	−0.7478	0.4576
Polarity sentiment	(Intercept)	0.10410	0.07604	1.7884	0.0788
	Misery Index	−0.00123	0.00046	−1.9599	0.0547
	Year	−0.00003	0.00004	−1.2139	0.2296

UR indicates unemployment rates; *<0.05; **<0.01.

Table 7 shows that neither the Misery Index nor inflation rates could significantly predict changes in the LPB index or polarity sentiment. In contrast, unemployment rates demonstrated a significant effect on predicting both LPB index ($P = .0194$) and polarity sentiment ($P = .0099$).

Table 8. Statistics of simple linear regressions of subjective happiness and LPB index/polarity sentiment.

	Variable	Estimate	Standard error	t-value	P-value
LPB index	(Intercept)	0.36287	1.19779	0.3029	0.7633
	Happiness	0.10914	0.03352	3.2562	0.0021**
	Year	0.00024	0.00061	0.3891	0.6989
Polarity sentiment	(Intercept)	0.15121	0.08160	1.8532	0.0701
	Happiness	0.00859	0.00228	3.7627	0.0005***
	Year	−0.00009	0.00004	−2.1075	0.0404*

*<0.05; **<0.01; ***<0.001.

This finding suggests that unemployment rates, compared to other economic indicators, may more directly reflect the economic hardships faced by individuals, thus having a stronger influence on the emotional content expressed in their language. This observation leads to a tentative hypothesis that unemployment rates are a more pertinent indicator of the economic conditions affecting linguistic positivity in written texts.

To examine the effect of psychological factors on diachronic changes in positivity in American English, we analyzed the relationship between subjective happiness (1946–2008) and both the LPB index and polarity sentiment in the CCOHA.

Table 8 shows that subjective happiness significantly predicted both the LPB index and polarity sentiment with moderate accounting of variation (LPB index: $F(2, 48) = 5.70$, $P = .0061$, multiple $R^2 = 0.1951$, adjusted $R^2 = 0.1608$; polarity sentiment: $F(2, 48) = 8.2958$, $P = .0008$, multiple $R^2 = 0.2609$, adjusted $R^2 = 0.2295$). These results indicate that the higher level of subjective happiness is associated with the increased use of positive words and more positive sentiment in sentences in general American English.

Discussion

Our results indicate a consistent decline in emotionality (both positive and negative) in the US general audience publications over the past two centuries. This pattern corroborates earlier studies (e.g. Acerbi et al. 2013a; Morin and Acerbi 2017; Scheffer et al. 2021). For instance, Acerbi et al. (2013a) identified a general decrease in emotion words from 1900 to 2000 by analyzing the Google Books Ngram corpus. Our study extends these findings by examining CCOHA, which encompasses not only books (both fiction and nonfiction) but also newspapers and popular magazines. This broader dataset allows us to confirm and expand upon the observed decline in emotionality in American English, suggesting that this trend is pervasive across multiple types of written genres.

One possible explanation for the decline in emotionality is a broader societal shift in the USA from a focus on emotional expression to a preference for rationality (Iliev and Axelrod 2016; Scheffer et al. 2021). Iliev and Axelrod (2016) reported a marked rise in causal language, which they attributed to the increasing influence of science, education, and technology in Western society. Similarly, Scheffer et al. (2021) found an increase in fact-based argumentation in the Google Books Ngram corpus beginning in the mid-nineteenth century, coupled with a decline in sentiment-laden words. They proposed that the scientific approach gained cultural prestige and shaped various institutions, spurring significant socioeconomic benefits. In turn, these scientific norms permeated everyday language and diminished explicit emotional expressions. Such findings underscore the interwoven nature of linguistic shifts and broader cultural transformations: as scientific thinking and rational discourse gained prominence, language adapted accordingly, reflecting and reinforcing new intellectual and cultural priorities.

In addition to a broader decline in emotionality, our findings reveal a general trend toward less positivity in American English. Multiple, potentially overlapping factors may account for this shift, including changing social norms around emotional expression, evolving literary styles, and a media environment that increasingly emphasizes conflict over consensus. For instance, as news coverage sharpened its focus on contentious or polarizing topics, its language usage may naturally reflect higher levels of negativity. Another possibility is that diminishing prosocial attitudes in American society have played a role, given prior work linking positive language use and prosocial behaviors (Williams 2006; Rand et al. 2015). Several longitudinal studies (e.g. Kesebir and Kesebir 2012; Greenfield 2013) also hint at weakening social cohesion and individualistic shifts, which might be consistent with reduced prosocial sentiment. However, caution is warranted in interpreting these patterns as evidence of an outright erosion of prosocial values; instead, our findings underscore how linguistic positivity may be shaped by a complex interplay of cultural, economic, and historical forces—an interplay that will be examined in greater depth in the following sections.

Furthermore, our analysis provides evidence that wars significantly affect LPB, which is consistent with Iliev et al. (2016) and Acerbi et al. (2013a). Together, these studies suggest that large-scale conflicts can lower overall levels of linguistic positivity. However, we found that the effects of wars on LPB can vary considerably across specific historical contexts. The Korean War and the Vietnam War exemplify this complexity.

While most major US conflicts prompted a drop in the LPB, the Korean War coincided with an increase. One possible explanation lies in the unique Cold War climate of the early 1950s: many Americans perceived the Korean War as a justified and essential stand against communism, at least at the beginning (Cumings 1981; Stueck 1997; Armstrong 2004). Accordingly, relevant reports and writings may have conveyed a relatively positive or at least less negative tone, reflecting a sense of legitimacy. Another potential explanation could be a lack of public awareness of the Korean War compared to other wars studied. Frequently overshadowed by the post-WWII boom in the USA, the Korean War is often referred to as 'The Forgotten War' (Blair 1987; Stein 1994). Specifically, the economic prosperity of the 1950s, driven by technological advances and a booming automobile industry, likely sustained a sense of optimism that, in turn, tempered the otherwise negative linguistic impact of warfare.

By contrast, the Vietnam War corresponds to a substantial drop in the LPB, marking the lowest levels among all decades from the 1820s to the 2000s despite ranking only fourth in casualties among the six major wars examined. We attribute this marked decline in part to the extensive anti-war sentiment in the USA (Small 2004). Influential public figures and large segments of society, notably university students, spoke out fervently against American involvement (Miller 1994). By 1971, polls indicated that 61 per cent of Americans believed participation in Vietnam to be a mistake (The New York Times 1971). This wave of public opposition was amplified by the rapid development and pervasive reach of television. Dubbed 'the Living Room War,' the Vietnam conflict was the first to be widely broadcast into people's homes, offering immediate and often graphic images of battle. Unlike earlier wars, where government agencies largely controlled the flow of information, television coverage of the Vietnam War brought the conflict's harsh realities into everyday American life. This unfiltered exposure elicited strong anti-war sentiment, fueling the use of more negative language in media reports and personal writings and thus contributing to Vietnam's record-low LPB in our dataset.

Taken together, these findings illustrate that the effects of wars on the LPB extend beyond casualty numbers. While major conflicts generally exert downward pressure on linguistic positivity, broader historical, socio-cultural, and technological contexts can shape and at times counterbalance this effect. The distinct linguistic outcomes associated with the Korean and Vietnam Wars highlight this complexity. The Korean War, 'The Forgotten War,' unfolded during a period characterized by economic optimism, ideological legitimacy, and lower public engagement, factors that collectively mitigated the typical negative linguistic consequences of warfare. In contrast, the Vietnam War, 'the Living Room War,' sharply differed as intensive television coverage dramatically heightened public awareness and anti-war sentiment through vivid depictions of combat. This unprecedented immediacy significantly amplified negative linguistic portrayals.

Ultimately, these comparisons underscore how wartime linguistic positivity emerges from an intricate interplay among media exposure, public perception, socioeconomic conditions, and collective memory.

As for the effect of economic conditions on the LPB, one striking aspect of our findings is that unemployment rates show a significant effect on linguistic positivity, whereas inflation rates do not. Prior research on the LPB often merges these two measures into a single Misery Index, potentially overlooking their divergent impacts. We argue that losing one's job entails profound psychological and material consequences beyond the gradual squeeze of rising prices. Previous studies have shown that unemployment disrupts an individual's sense of identity and daily structure, exacerbates financial insecurity, and fosters social stigma—factors that directly undermine overall well-being (e.g. Jahoda 1982; Warr 1987; Feather 1990). Additionally, unemployment has been linked to increased mortality rates related to alcohol and smoking (Sullivan and von Wachter 2009), creating a ripple effect of economic and social strain not only for individuals but also for their families and communities (Kivimaki 2003; Lenthe et al. 2005). In contrast, although inflation can erode purchasing power, people find ways to adapt through wage adjustments or selective spending, and thus the day-to-day emotional toll is often less acute. Di Tella et al. (2001) found that people would tradeoff a 1-percentage-point increase in unemployment for a 1.7-percentage-point increase in the inflation rate. Blanchflower et al. (2014) even suggested that unemployment may be up to five times as detrimental as inflation, helping explain why it emerges as a stronger predictor of linguistic positivity shifts. Against this backdrop, it makes sense that changes in unemployment, rather than price increases alone, would more strongly surface in language use, manifesting as shifts in linguistic positivity that reflect heightened anxiety or distress when jobs become scarce.

Despite these distinct effects, we found no significance for the combined Misery Index on LPB changes contrasting with the highly significant effect reported by Iliev et al. (2016). This discrepancy may stem from differences in corpora and research scope. However, our study draws on a broad spectrum of text types (fiction, nonfiction, popular magazines, and newspapers), Iliev et al. (2016) relied on the Google Books Ngrams corpus and New York Times articles, both of which may be more responsive to real-time economic events. By encompassing literature and news content tied closely to current affairs, their corpora may more directly capture signals tied to overall economic conditions. In contrast, the more varied genres in the CCOHA incorporate editorial norms, literary traditions, and broader authorial choices, potentially diluting any direct effect of inflation on language use. Our findings thus suggest that future research should disaggregate measures of economic conditions, particularly unemployment, rather than relying solely on aggregated indicators like the Misery Index. Such an approach could better illuminate the nuanced relationships between specific economic stressors and shifts in linguistic positivity. Furthermore, we propose that economic indices which more closely track individuals' personal financial situations may offer more accurate predictions of changes in positivity observed within textual analyses.

Finally, our analysis reveals that subjective happiness exerts a significant effect on the LPB at both the word and sentence levels. This dual-level influence aligns with psychological frameworks suggesting collective well-being shapes not only what is expressed, such as the accessibility of positive lexemes like 'unity' or 'prosperity,' but also how ideas are framed contextually, as seen in shifts in sentence-level polarity (Fredrickson 2001; Pennebaker et al. 2015). Critically, these findings complement our earlier focus on objective factors like wars and economic indicators, demonstrating that linguistic shifts in American English arise from an interplay of external societal pressures and internal psychological states. By integrating word-level and sentence-level measures of positivity, we disentangle environmental influences (e.g. wartime stress) from affective evaluations (e.g. declining optimism), offering a more nuanced understanding of linguistic evolution. Ultimately, these results position language not merely as a reflection of societal conditions but as a dynamic system shaped by the confluence of material realities, cultural narratives, and collective emotionality.

Conclusion

The present study investigated the changes in linguistic positivity in the US discourses in popular printed material over the past two centuries using the CCOHA. In addition, we tested the effects of environmental and psychological factors on changes in positivity. We analyzed the positivity of general American English at both word and sentence levels using the LPB index and polarity sentiment with a comprehensive emotion dictionary (NRC emotion lexicon). To our knowledge, this is the first study to examine these two levels of positivity by employing a large-scale structured corpus of historical English.

Our findings reveal a general downward trend in both positive and negative words in American English over the past two centuries. We proposed that this decrease may result from a societal shift in public interest from emotionality to rationality in American society. Furthermore, we identified a significant decline in positivity at both word and sentence levels, which may be associated with a decrease in prosociality in the USA. Regarding the drivers of changes in positivity, we found that both environmental (e.g. wars and economic conditions) and psychological factors (i.e. subjective happiness) play a role in predicting these changes. Notably, we highlighted that unemployment rates emerged as a particularly strong predictor of changes in positivity, given their more direct impact on individual economic hardships and subsequent language use. We also observed that major wars generally lower linguistic positivity, although this effect varies with historical contexts, technological advancements, and media developments.

The implications of the present study are manifold. Methodologically, by integrating both word-level and sentence-level measures of linguistic positivity within a large, balanced corpus, we offer a more comprehensive and robust framework for tracking long-term trends in general American English. The sentence-level measure helps mitigate biases inherent in strictly word-based metrics, such as disregarding negations or intensifiers. Theoretically, our study advances the understanding of war-LPB dynamics by revealing that conflict's influence on linguistic positivity is not uniform, but rather contingent upon a complex interplay of societal factors, such as public consensus, media coverage, and economic context. Moreover, we refined the effect of economic conditions on the LPB by specifically identifying unemployment rate as a more reliable economic predictor, thus underscoring the importance of factors that directly affect individual livelihoods when examining diachronic changes of LPB. These findings further our understanding of how societal upheavals and economic stressors shape collective linguistic expression. Practically, our approach may serve as a diagnostic tool for gauging social health. Given the correlations among LPB, economic conditions, and subjective happiness, the method we propose may help detect changes in collective well-being. Previous works (e.g. Rude et al. 2004; Bollen et al. 2021) have demonstrated the validity of text analysis for psychological conditions, such as, depression. However, most existing approaches rely primarily on word-level analyses using relatively small emotion lexicons. For example, Rude et al. (2004) employed a fixed list of approximately 2,000 positively or negatively valenced words and primarily examined their relationship to pronouns, whereas larger lexicons such as the NRC contain over 15,000 emotion-related lexical items. By adopting a broader lexicon like NRC alongside polarity-based sentiment analysis, researchers and healthcare professionals may more accurately capture nuanced expressions of negative and positive effects in Pennebaker-style writing tasks (Tausczik and Pennebaker 2010). Consequently, our combined use of word-level and sentence-level measures holds promise for improving clinical assessments of psychological well-being, enabling schools, hospitals, and mental health professionals to develop more precise and responsive interventions.

Some limitations should be acknowledged. First, this study relies solely on the CCOHA, which is specific to general American English, limiting its applicability to other languages or cultural contexts. Future studies could expand this scope by utilizing large-scale corpora across various languages, further exploring general patterns in changes in linguistic positivity. Furthermore, our study analyzed environmental factors and psychological factors separately rather than in an integrated model due to temporal coverage differences and multicollinearity concerns.

Future research could benefit from an integrated approach examining how economic indicators might correlate differently with well-being vs. linguistic expression, potentially revealing mediating relationships between objective conditions, subjective experiences, and linguistic patterns that our separate analyses could not capture.

Acknowledgements

We would like to extend our sincere gratitude to the anonymous reviewers for their valuable comments and suggestions, which greatly improved the quality of this paper. We would also like to thank Prof. Ruut Veenhoven, who generously shared the World Database of Happiness.

Funding

This paper is supported by the Fundamental Research Funds for the Central Universities of China.

Ethical approval

Ethical approval is not applicable for this article.

Statement of human and animal rights

This article does not contain any studies with human or animal subjects.

Statement of informed consent

There are no human subjects in this article, and informed consent is not applicable.

Notes on Contributors

Bin Shao is a professor of linguistics in the School of International Studies at Zhejiang University. His research interests mainly include language variation and change, English lexicology, corpus linguistics, and contrastive linguistics. He recently published research articles in the *Journal of English Linguistics*, *Corpus Linguistics and Linguistic Theory*, *Lingua*, etc. Address for correspondence: School of International Studies, Zhejiang University, No. 866, Yuhangtang Road, Hangzhou, Zhejiang, China. binshao@zju.edu.cn.

Gui Wang is a PhD candidate majoring in Foreign Linguistics & Applied Linguistics in the School of International Studies at Zhejiang University. His research focuses on corpus linguistics, culturomics, and the intersections of language, culture, and history, utilizing a quantitative analysis of large-scale corpus data. He recently published articles in *TESOL Quarterly*, *Language Sciences,* and *International Journal of Applied Linguistics*. Address for correspondence: School of International Studies, Zhejiang University, No. 866, Yuhangtang Road, Hangzhou, Zhejiang, China. gui.wang@zju.edu.cn

Zenan Chen is a PhD candidate majoring in Foreign Linguistics & Applied Linguistics in the School of International Studies at Zhejiang University. His research focuses on the variation and change of language structures, through the lens of corpus-based analysis and multivariate statistical modeling. He recently published an article in *Lingua*. Address for correspondence: School of International Studies, Zhejiang University, No. 866, Yuhangtang Road, Hangzhou, Zhejiang, China. chenzenan@zju.edu.cn

References

Acerbi, A. et al. (2013b) 'The Expression of Emotions in 20th Century Books', *PLoS One*, 8: e59030. https://doi.org/10.1371/journal.pone.0059030

Acerbi A., Lampos V., Bentley R. A. (2013a) 'Robustness of Emotion Extraction From 20th Century English Books', in *2013 IEEE International Conference on Big Data*, Silicon Valley, pp. 1–8. California, USA: IEEE. https://doi.org/10.1109/BigData.2013.6691663.

Alatrash, R., Schlechtweg, D., and Kuhn, J. (2020) 'CCOHA: Clean Corpus of Historical American English', in *Proceedings of the 12th Conference on Language Resources and Evaluation* (LREC 2020), pp. 6958–66. Marseille, France: European Language Resources Association.

Armstrong, C. K. (2004). *The North Korean Revolution, 1945–1950*, Studies of the Weatherhead East Asian Institute, Columbia University. Ithaca, NY: Cornell University Press. https://doi.org/10.7591/9780801468803

Augustine, A. A., Mehl, M. R., and 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: 508–15. https://doi.org/10.1177/1948550611399154

Biber, D., and Gray, B. (2013) 'Being Specific About Historical Change: The Influence of Sub-Register', *Journal of English Linguistics*, 41: 104–34. https://doi.org/10.1177/0075424212472509

Blair, C. (1987) *The Forgotten War: America in Korea, 1950–1953*, 1st edn. New York: Times Books.

Blanchflower, D. G. et al. (2014) 'The Happiness Trade-Off Between Unemployment and Inflation', *Journal of Money, Credit, and Banking*, 46: 117–41. https://doi.org/10.1111/jmcb.12154

Blum, D. A. (2020) 'American War and Military Operations Casualties: Lists and Statistics', https://sgp.fas.org/crs/natsec/RL32492.pdf, accessed 15 Jan. 2024.

Bollen, J. et al. (2021) 'Historical Language Records Reveal a Surge of Cognitive Distortions in Recent Decades', *Proceedings of the National Academy of Sciences of the United States of America*, 118: e2102061118. https://doi.org/10.1073/pnas.2102061118

Boucher, J., and Osgood, C. E. (1969) 'The Pollyanna Hypothesis', *Journal of Verbal Learning and Verbal Behavior*, 8: 1–8. https://doi.org/10.1016/S0022-5371(69)80002-2

Center for Behavioral Health Statistics and Quality. (2022) '2021 National Survey on Drug Use and Health (NSDUH): Methodological Summary and Definitions', https://catalog.data.gov/dataset/2021-methodological-summary-and-definitions, accessed 20 Jan. 2024.

Cumings, B. (1981) *The Origins of the Korean War*, 3. Princeton, NJ: Princeton University Press.

Davies, M. (2012) 'Expanding Horizons in Historical Linguistics With the 400-Million Word Corpus of Historical American English', *Corpora*, 7: 121–57. https://doi.org/10.3366/cor.2012.0024

Department of Veterans Affairs (2023) 'America's Wars', https://department.va.gov/americas-wars/, accessed 15 May 2024.

Diener, E. et al. (2018) 'Revisiting "Most People Are Happy"—and Discovering When They Are Not', *Perspectives on Psychological Science: A Journal of the Association for Psychological Science*, 13: 166–70. https://doi.org/10.1177/1745691618765111

Di Tella, R., MacCulloch, R. J., and Oswald, A. J. (2001) 'Preferences Over Inflation and Unemployment: Evidence From Surveys of Happiness', *American Economic Review*, 91: 335–41. https://doi.org/10.1257/aer.91.1.335

Dodds, P. S. et al. (2015) 'Human Language Reveals a Universal Positivity Bias', *Proceedings of the National Academy of Sciences of the United States of America*, 112: 2389–94. https://doi.org/10.1073/pnas.1411678112

Feather, N. T. (1990) *The Psychological Impact of Unemployment*, Springer Series in Social Psychology. New York, NY: Springer. https://doi.org/10.1007/978-1-4612-3250-6

Fischer, C. S. (2008) 'What Wealth-Happiness Paradox? A Short Note on the American Case', *Journal of Happiness Studies*, 9: 219–26. https://doi.org/10.1007/s10902-007-9047-4

Fredrickson, B. L. (2001) 'The Role of Positive Emotions in Positive Psychology. The Broaden-and-Build Theory of Positive Emotions', *The American Psychologist*, 56: 218–26. https://doi.org/10.1037//0003-066x.56.3.218

Gandelman, N., and Hernández-Murillo, R. (2009) 'The Impact of Inflation and Unemployment on Subjective Personal and Country Evaluations', *Federal Reserve Bank of St. Louis Review*, 91: 107–26. https://doi.org/10.2139/ssrn.2296104

Granath, S., and Ullén, M. (2019) '"The Elevation of Sensitivity Over Truth": *Political Correctness* and Related Phrases in the *Time* Magazine Corpus', *Applied Linguistics*, 40: 265–87. https://doi.org/10.1093/applin/amx019

Greenfield, P. M. (2013) 'The Changing Psychology of Culture From 1800 Through 2000', *Psychological Science*, 24: 1722–31. https://doi.org/10.1177/0956797613479387

Holtz, P., Deutschmann, E., and 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: 1410–31. https://doi.org/10.1177/0022022117724902

Hu, M., and Liu, B. (2004) 'Mining and Summarizing Customer Reviews'. in *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery & Data Mining (KDD-04)*, pp. 168–77, New York, NY, USA. Association for Computing Machinery.

Iliev, R. et al. (2016) 'Linguistic Positivity in Historical Texts Reflects Dynamic Environmental and Psychological Factors', *Proceedings of the National Academy of Sciences of the United States of America*, 113: E7871–9. https://doi.org/10.1073/pnas.1612058113

Iliev, R., and Axelrod, R. (2016) 'Does Causality Matter More Now? Increase in the Proportion of Causal Language in English Texts', *Psychological Science*, 27: 635–43. https://doi.org/10.1177/0956797616630540

Jahoda, M. (1982) *Employment and Unemployment: A Social-Psychological Analysis*. The Psychology of Social Issues. New York: Cambridge University Press.

Jockers, M. L. (2017) 'Syuzhet Package v1.0.4.' https://www.rdocumentation.org/packages/syuzhet/versions/1.0.4, accessed 9 Jun. 2024.

Kesebir, P., and Kesebir, S. (2012) 'The Cultural Salience of Moral Character and Virtue Declined in Twentieth Century America', *The Journal of Positive Psychology*, 7: 471–80. https://doi.org/10.1080/17439760.2012.715182

Kivimaki, M. (2003) 'Temporary Employment and Risk of Overall and Cause-Specific Mortality', *American Journal of Epidemiology*, 158: 663–8. https://doi.org/10.1093/aje/kwg185

Koplenig, A. (2015) 'The Impact of Lacking Metadata for the Measurement of Cultural and Linguistic Change Using the Google Ngram Data Sets—Reconstructing the Composition of the German Corpus in Times of WWII', *Digital Scholarship in the Humanities*, 31: 169–88. https://doi.org/10.1093/llc/fqv037

Lakoff, R. (1975) ''Language and Woman's Place'', *Language In Society*, 2: 45–79. https://doi.org/10.1017/S0047404500000051

Lenthe, F. J. v. et al. (2005) 'Neighbourhood Unemployment and All Cause Mortality: A Comparison of Six Countries', *Journal of Epidemiology and Community Health*, 59: 231–7. https://doi.org/10.1136/jech.2004.022574

Levendusky, M., and Malhotra, N. (2016) 'Does Media Coverage of Partisan Polarization Affect Political Attitudes?', *Political Communication*, 33: 283–301. https://doi.org/10.1080/10584609.2015.1038455

Liu X. and Zhu H. (2024) 'The Diachronic Change in Linguistic Positivity in the Academic Book Reviewing of Language Studies: A Text-Mining Analysis', *Scientometrics*, 130:133–57. https://doi.org/10.1007/s11192-024-05208-8

Loughran, T., and Mcdonald, B. (2011) 'When Is a Liability Not a Liability? Textual Analysis, Dictionaries, and 10Ks', *The Journal of Finance*, 66: 35–65. https://doi.org/10.1111/j.1540-6261.2010.01625.x

Lovell, M. C., and Tien, P.-L. (2000) 'Economic Discomfort and Consumer Sentiment', *Eastern Economic Journal*, 26: 1–8. https://doi.org/10.2139/ssrn.222510

Mahmoudi, A., Jemielniak, D., and Ciechanowski, L. (2024) 'Assessing Accuracy: A Study of Lexicon and Rule-Based Packages in R and Python for Sentiment Analysis', *IEEE Access*, 12: 20169–80. https://doi.org/10.1109/ACCESS.2024.3353692

Matlin, M. W., and Stang, D. J. (1978) *The Pollyanna Principle: Selectivity in Language, Memory, and Thought*. Cambridge, MA: Schenkman.

Mehl, M. R. et al. (2017) 'Natural Language Indicators of Differential Gene Regulation in the Human Immune System', *Proceedings of the National Academy of Sciences of the United States of America*, 114: 12554–9. https://doi.org/10.1073/pnas.1707373114

Michel, J.-B. et al. (2011) 'Quantitative Analysis of Culture Using Millions of Digitized Books', *Science*, 331: 176–82. https://doi.org/10.1126/science.1199644

Miller, J. (1994) *Democracy is in the Streets*. Cambridge, MA: Harvard University Press.

Mohammad, S. M., and Turney, P. D. (2013) 'Crowdsourcing a WordEmotion Association Lexicon', *Computational Intelligence*, 29: 436–65. https://doi.org/10.1111/j.1467-8640.2012.00460.x

Morin, O., and Acerbi, A. (2017) 'Birth of the Cool: A Two-Centuries Decline in Emotional Expression in Anglophone Fiction', *Cognition & Emotion*, 31: 1663–75. https://doi.org/10.1080/02699931.2016.1260528

Nessen, R. (2008) 'Arthur Okun Father of the Misery Index', *The Brookings Institution*, December 17. https://www.brookings.edu/articles/the-brookings-institutions-arthur-okun-father-of-the-miseryindex/, accessed 9 May 2024.

Neuman, Y. et al. (2012) 'Proactive Screening for Depression Through Metaphorical and Automatic Text Analysis', *Artificial Intelligence in Medicine*, 56: 19–25. https://doi.org/10.1016/j.artmed.2012.06.001

Pechenick, E. A., Danforth, C. M., and Dodds, P. S. (2015) 'Characterizing the Google Books Corpus: Strong Limits to Inferences of Socio-Cultural and Linguistic Evolution', *PLoS One*, 10: e0137041. https://doi.org/10.1371/journal.pone.0137041

Pennebaker, J. W. et al. (2015) *The Development and Psychometric Properties of LIWC2015*. Austin, TX: University of Texas at Austin.

Rand, D. G., Kraft-Todd, G., and Gruber, J. (2015) 'The Collective Benefits of Feeling Good and Letting Go: Positive Emotion and (dis)Inhibition Interact to Predict Cooperative Behavior', *PLoS One*, 10: e0117426. https://doi.org/10.1371/journal.pone.0117426

Rinker, T. (2019) 'Calculate Text Polarity Sentiment', https://cran.r-project.org/web/packages/sentimentr/sentimentr.pdf, accessed 10 Jan. 2024.

Rozin, P., Berman, L., and Royzman, E. (2010) 'Biases in Use of Positive and Negative Words Across Twenty Natural Languages', *Cognition & Emotion*, 24: 536–48. https://doi.org/10.1080/02699930902793462

Rude, S., Gortner, E.-M., and Pennebaker, J. (2004) 'Language Use of Depressed and Depression-Vulnerable College Students', *Cognition & Emotion*, 18: 1121–33. https://doi.org/10.1080/02699930441000030

Samavati, S., and Veenhoven, R. (2024) ''Happiness in Urban Environments: What We Know and Don't Know Yet'', *Journal of Housing and the Built Environment: HBE*, 39: 1649–707. https://doi.org/10.1007/s10901-024-10119-4

Scheffer, M. et al. (2021) 'The Rise and Fall of Rationality in Language', *Proceedings of the National Academy of Sciences of the United States of America*, 118: e2107848118. https://doi.org/10.1073/pnas.2107848118

Seabrook, E. M. et al. (2018) 'Predicting Depression From Language-Based Emotion Dynamics: Longitudinal Analysis of Facebook and Twitter Status Updates', *Journal of Medical Internet Research*, 20: e9267. https://doi.org/10.2196/jmir.9267

Small, M. (2004) *Antiwarriors: The Vietnam War and the Battle for America's Hearts and Minds, Vietnam, America in the War Years*, 1st edn. Lanham, MD: SR Books.

Snefjella, B., Généreux, M., and Kuperman, V. (2019) 'Historical Evolution of Concrete and Abstract Language Revisited', *Behavior Research Methods*, 51: 1693–705. https://doi.org/10.3758/s13428-018-1071-2

Stein, R. C. (1994) *The KoreanWar: 'The Forgotten War'*. Hillside, NJ: Enslow Publishers.

Stueck, W. W. (1997) *The Korean War: An International History*, Princeton Studies in International History and Politics, Paperback Print, With Corr. Princeton, NJ: Princeton University Press.

Sullivan, D., and von Wachter, T. (2009) 'Job Displacement and Mortality: An Analysis Using Administrative Data*', *The Quarterly Journal of Economics*, 124: 1265–306. https://doi.org/10.1162/qjec.2009.124.3.1265

Tang, C. F., and Lean, H. H. (2009) 'New Evidence From the Misery Index in the Crime Function', *Economics Letters*, 102: 112–5. https://doi.org/10.1016/j.econlet.2008.11.026

Tausczik, Y. R., and Pennebaker, J. W. (2010) 'The Psychological Meaning of Words: LIWC and Computerized Text Analysis Methods', *Journal of Language and Social Psychology*, 29: 24–54. https://doi.org/10.1177/0261927X09351676

The New York Times. (1971) '61% in Poll Assert Entry Into the War Was U.S. "Mistake"', *The New York Times*, 52.

Veenhoven, R. (2024) Trends in Nations; Data File to be Used for the Longitudinal Analysis of Societal Conditions for Happiness. World Database of Happiness', Rotterdam, The Netherlands: Erasmus University.

Vinkers, C. H., Tijdink, J. K., and Otte, W. M. (2015) 'Use of Positive and Negative Words in Scientific PubMed Abstracts Between 1974 and 2014: Retrospective Analysis', *BMJ*, 351: h6467. https://doi.org/10.1136/bmj.h6467

Warr, P. B. (1987) *Work, Unemployment, and Mental Health*. Oxford; New York: Clarendon Press; Oxford University Press.

Warriner, A. B., and Kuperman, V. (2015) 'Affective Biases in English Are Bi-Dimensional', *Cognition & Emotion*, 29: 1147–67. https://doi.org/10.1080/02699931.2014.968098

Wen, J., and Lei, L. (2022) 'Linguistic Positivity Bias in Academic Writing: A Large-Scale Diachronic Study in Life Sciences Across 50 Years', *Applied Linguistics*, 43: 340–64. https://doi.org/10.1093/applin/amab037

Williams, F. (2006). 'The Relationship Among Linguistic Patterns, Thwarted Belongingness, Perceived Burdensomeness, and Suicidal Behavior: A Test of Joiner's Theory of Suicide', PhD dissertation, Florida State University, Tallahassee, FL.

Wilson, A. E., Parker, V. A., and Feinberg, M. (2020) 'Polarization in the Contemporary Political and Media Landscape', *Current Opinion in Behavioral Sciences*, 34: 223–8. https://doi.org/10.1016/j.cobeha.2020.07.005

Yang, B., and Lester, D. (1992) 'The Misery Index and an Index of Misery', *Atlantic Economic Journal*, 20: 98. https://doi.org/10.1007/BF02300179