English Language Teaching and Linguistics Studies
ISSN 2640-9836 (Print) ISSN 2640-9844 (Online)
Vol. 6, No. 5, 2024
www.scholink.org/ojs/index.php/eltls

Original Paper

Readability Trends in Four Genres: A Research Based on COCA

Deng Cheng[1] & Jingxiang Cao[1*]

1 School of Foreign Languages, Dalian University of Technology, Dalian, China

[*] Jingxiang Cao, Corresponding author

Received: July 29, 2024 Accepted: August 15, 2024 Online Published: September 06, 2024

doi:10.22158/eltls.v6n5p1 URL: http://dx.doi.org/10.22158/eltls.v6n5p1

Abstract

Readability measures the ease of reading a text, and high readability indicates easy texts. There has been a resurgence of interest in readability assessment due to the development of efficient natural language processing (NLP) systems. However, most studies have focused on one type of text, and few have compared the readability of different genres and described the diachronic changes in different genres. Based on the large-scale COCA corpus, this paper employs two readability indexes and three readability components to analyze the readability trend of different genres and uses SPSS to test the significance of the trend.

The results showed that the changing patterns of readability varied across the four studied genres. A general upward trend in readability was found in the fiction genre, a mixed trend in the academic and magazine genres, and a downward trend in the news genre. Based on the results, we concluded that there may be multiple factors at work behind the diachronic development of readability and that these factors may have different effects on different genres.

Keywords

genre, readability, COCA, diachronic study

1. Introduction

Readability refers to the ease of reading or understanding certain reading materials (Dale & Chall, 1948). Reading is an important way to gain knowledge and information, and thus high readability is vital for knowledge transmission and information communication (Hartley et al., 2002). It is also helpful for teachers and educators because it plays an important role in selecting appropriate reading materials for learners (Sheehan et al., 2014). In the last decade, the development of efficient natural language processing (NLP) systems has led to a resurgence of interest in readability assessment. Several studies have been conducted based on NLP-enabled feature extraction and state-of-the-art machine learning algorithms, with significant performance improvements over traditional readability

measures.

Genre is an important factor in describing the diachronic change of language although it is ignored in previous studies. Biber and Gray (2013) found that genres or registers are an important mediating factor for linguistic changes. Such a claim has also been supported by other previous investigations of changing patterns of grammatical features (Bao et al., 2018; Biber & Clark, 2002) and linguistic styles (Biber & Finegan, 2014) in different genres. More importantly, readability has been found to be affected by the genre of texts (Štajner et al., 2012). Therefore, further investigation is needed, in light of the possible genre effect on readability, to verify the trend of different genres.

Although genre is considered to be important in mediating linguistic change, the readability trends of various genres remain unknown. It is also unclear how genre differences affect the diachronic trend of readability. In addition, previous studies on the effect of genre on readability mostly studied from a synchronic perspective, hence it would be interesting to display the readability trend of various genres. Therefore, the present study sets out to examine the diachronic change of readability in the four genres (fiction, news, magazines, and academic articles) in the Corpus of Contemporary American English (COCA).

So this study tries to compare the readability trend of different genres and answer the following questions:

1) What are the trends of readability in the four genres of COCA?

2) What are the differences among the readability trends of different genres? And what are the possible factors behind different trends?

2. Literature Review

2.1 Readability Studies

Readability has been defined in various ways by different researchers, each tailoring the concept to their own study objectives. Dale and Chall (1948) formally defined readability as "the total (including all the interactions) of all those elements within a given piece of printed material that affect the success a group of readers has with it." This definition emphasizes both the text and the reader, focusing on how well the reader comprehends the text and the speed at which they read it. Klare (1963) defines readability as the ease with which a piece of text can be understood. This definition focuses on the writing style as a text element. Similarly, Ismail et al. (2016) proposed that readability determines whether a book is simple or complex to understand. It can be concluded that readability refers to the ease of reading or understanding certain reading materials (Dale & Chall, 1948; Klare, 1963).

Traditional readability formulas are primarily based on two factors: word difficulty and sentence complexity. For instance, word difficulty is often measured by the number of syllables, while sentence complexity is gauged by sentence length. Despite some recent criticism, traditional readability formulas remain widely used in many studies (Lei & Yan, 2016; Plavén Sigray et al., 2017; Wang et al., 2022). Klare (1980) argued that readability formulas, though imperfect, are more accurate than human

judgment. Bailin and Grafstein (2001) acknowledged the potential of readability formulas to assess text difficulty for certain readers.

The importance of readability and the availability of readability formulas have led to their application in various text types, such as speeches (Kayam, 2018), educational materials (Badarudeen & Sabharwal, 2010), academic articles, fiction, and contexts like healthcare (Ley & Florio, 1996). The decreasing readability of academic abstracts has garnered significant attention. Studies indicate that academic articles are often difficult to understand. Using the Flesch Reading Ease (FRE) formula, standard passages typically score between 60 and 70 (Zamanian & Heydari, 2012), whereas academic articles score lower, indicating higher difficulty. Lei and Yan (2016) assessed the readability of abstracts and full texts in information science, finding academic texts challenging to read. This trend is consistent across fields such as tourism (Dolnicar & Chapple, 2015) and psychology (Hartley et al., 2002), showing that academic publications are becoming increasingly difficult to comprehend. Plavén Sigray et al. (2017) analyzed 709,577 abstracts from 123 scientific journals (1881-2015) and found a notable decline in readability over time. Lei and Yan (2016) compared the readability of abstracts and full texts from 2003 to 2012, finding that abstracts are becoming more difficult to read while full texts are becoming slightly easier, though these differences were not statistically significant. This highlights discrepancies in readability trends within academic texts.

Previous studies also investigated the readability of other genres. News, written for the general public, is expected to use simpler language. However, news stories typically contain longer words and sentences than other writings (Fowler, 1978). Westin (2016) provided evidence for decreasing sentence length in English newspapers from 1900 to 2000, which could contribute to higher readability. However, Danielson (1992) found that the increased use of longer words has made news stories less readable. In contrast, fiction, which shares stylistic elements with news, shows an increasing readability trend due to shorter sentences and simpler words (Danielson, 1992). Tanaka (2012) confirmed this trend using Victorian Women's Writers Project Corpus which contains 200 original texts written by female writers from the 1820s to the 1920s. The readability of magazines has received less scholarly attention. Fowler and Smith (1979) found that magazines were more readable than newspapers during the same period. Early readability trend studies, however, often used unrepresentative corpora.

Most readability studies focus on one specific genre, particularly academic papers, highlighting the trend of increasing difficulty. There is limited research comparing the readability of different genres. Štajner et al. (2012) examined text complexity factors, including average sentence length and the Automated Readability Index, in British and American English genres from the 20th century, finding varied trends across genres. But the corpus used in this study is small, indicating the need for broader research. Given the limitations of previous studies, it is necessary to conduct comprehensive research to describe readability changes across various genres.

2.2 Genre Studies

Language is a continuously evolving and changing entity (Milroy, 2001). Genres, as conventional

instances of organized text (Couture, 1986), also evolve in this process. The history of English is composed of the histories of its registers and genres (Diller, 2001), each with its own dynamics and diachronic developments. Genre variation focuses on the conventional ways in which complete texts of different types are structured and it is a fundamental aspect of human language(Biber, 2009). Given the importance of genre variation, it is of central importance for both the description of particular languages and the development of language.

However historical linguistics is neglected or understudied from this angle in early genre studies. Biber (1991) verified the co-occurrences of 67 linguistic features in data that included over 481 texts extracted from 23 different genres. Although it studied genres from a synchronic perspective, it provided an innovative and multidimensional methodology that could be applied to diachronic study. Then a trend towards a more colloquial direction was found in fiction, essays, and letters, but the rates of change were different (Biber & Finegan 1989). Biber and Finegan (2014) conducted a long-term diachronic study and found that drama and medical academic writing evolved in different directions. And recently registering can be a predictor of language change has been proved by a number of studies (Biber, 2012; Biber & Gray, 2013, 2016). For example, Biber and Gray (2013) argue that registers are a mediating factor for diachronic change in language showing that "minor differences in register can correspond to meaningful and systematic differences in the patterns of linguistic change". With the development of corpus linguistics and NLP tools, corpus linguistic methods now can provide solid empirical evidence for how genres develop and how they evolve from a diachronic perspective.

Empirical evidence demonstrates the impact of genre on the historical development of language. Consequently, the question arises as to whether and to what extent readability assessment is genre-independent. The existing literature on readability indicates that the degree of readability is connected to genre. The study conducted by Kate et al. (2010) improves the accuracy of readability predictions by using genre-specific features. Similarly, Štajner et al. (2012) compared the readability of various genres from different genres in American English and British English and demonstrated that readability and linguistic features correlated with readability exhibit different trends. Furthermore, research into the readability of Japanese textbooks has demonstrated that text genres are related to a textbook's readability and ultimately influence readers' comprehension (Li, 2022). Consequently, we could conclude that readability is influenced by the genre of texts (Kate et al., 2010; Štajner et al., 2012).

Despite the genre effect on readability and its importance to linguistic change, no research has been conducted to compare and measure the readability trend of different genres using large-scale data. Having precise information about how a given change occurred is a necessary prerequisite for discussions of why the change occurred in the way it did.

3. Methods

3.1 Readability Measurement

Two readability formulas were used in this research to examine texts: the Flesch Reading Ease (FRE) (Flesch, 1948) and the New Dale-Chall Readability Formula (NDC) (Chall & Dale, 1995). These two measures use different language metrics: syllable count, sentence count, word count, and percentage of difficult words. A high FRE score means a high readability, while a high New Dale-Chall grade reflects a low readability. FRE uses statistics like sentence length and average number of syllables per word, while New Dale-Chall also takes difficult words into consideration. Two different readability measurements were chosen to ensure that the results were not induced by a single measurement. FRE was chosen due to its popularity and consistency with other readability metrics (Didegah & Thelwall, 2013), and because it has previously been applied to trends over time (Lim, 2008; Danielson et al., 1992; Jatowt & Tanaka, 2012; Stevenson, 1964). NDC was chosen since it is both well-established and compares well with more recent methods for analyzing readability (Benjamin, 2012). Their formulas and linguistic features are summarized as follows:

Flesch Reading Ease Readability Formula rates texts on a 100-point scale; the higher the score, the easier it is to understand the text. Most standard passages have approximately a readability score of 60 to 70.

The Flesch Reading Ease Readability Formula is:

$206.835 – (1.015 \times ASL) – (84.6 \times ASW)$

Where, ASL is the Average Sentence Length (the number of words divided by the number of sentences), and ASW is the Average of Syllables per Word (the number of syllables divided by the number of words).

Another popular readability formula is the Dale-Chall Formula. The original Dale-Chall Formula was developed for adults and children above the 4th grade level. They designed it to correct certain shortcomings in the Flesch Reading Ease Formula. It was a sentence-length variable plus a percentage of hard words—words not found on the Dale-Chall long list of 3000 easy words, 80 percent of which are known to fourth-grade readers.

The Dale-Chall Raw Score is given by,

$\text{Raw Score} = 0.1579\ PDW + 0.496\ ASL + 3.6365$

Raw Score refers to the reading grade of a reader who can answer one half of the test questions on a passage, PDW is Percentage of Difficult Words (words not on the Dale-Chall word list), and ASL is Average Sentence Length in Words.

The two readability measures consist of three components, i.e., the number of syllables per word (FRE), the number of words per sentence (FRE and NDC), and the share of difficult words (NDC). A natural question is thus which of the components changed to drive the overall decrease in readability? Thus we also count the syllables, lexicons, sentences, and difficult words in these texts to calculate the three components used in these two measures. These readability statistics were calculated with the

readability of the Python package, developed by the NLTK team (available at https://github.com/shivam5992/textstat).

3.2 Material

The corpus employed in this study is the Corpus of Contemporary American English (COCA). As one of the most widely used corpora of Contemporary English, COCA consists of 1,002,889,754 tokens of English texts from 1990 to 2019.

The corpus contains more than one billion words, including 20 million words each year from 1990-2019 (with the same genre balance year by year). It is the only corpus of English that is large, recent, and has a wide range of genres. It means that in addition to seeing variation by genre, we can also map out recent changes in English in ways that are not possible with any other corpus. With a well-designed and balanced composition, COCA allows for research in a wide range of language changes and variations, including syntactic changes. Table 1 shows the genres and texts in the corpus:

Table 1. Genres and texts in COCA

Genre	Texts	Words
Spoken	44,803	127,396,932
Fiction	25,992	119,505,305
Magazines	86,292	127,352,030
Newspapers	90,243	122,958,016
Academic	26,137	120,988,361
Web(Genl)	88,989	129,899,427
Web(Blog)	98,748	125,496,216
TV/Movies	23,975	129,293,467

There are 8 genres in COCA but only 4 genres in this corpus are employed in this study: fiction, magazines, newspapers, and academic articles. Sub-corpus of web genres and spoken are ignored in this study because the readability formulas used in this study are designed based on some traditional written texts. There are a number of previous studies that have demonstrated the unreliability of traditional readability measures for Web pages and other types of non-traditional documents (Collins Thompson & Callan, 2004; Feng et al., 2009).

Before we calculate the readability of each passage, texts are cleaned to remove extra marks that might influence the readability. Firstly, we removed sentences with special characters from the corpus. As described on the official website of the corpus, ten randomly selected words in every two hundred words in the texts were replaced with special characters such as "@" for the purpose of copyright protection. Therefore, sentences containing such special characters were removed in the present study to eliminate possible disturbing effects. Also the common tags such as "<P>" need to be removed.

3.3 Data Analysis

To obtain a comprehensive understanding of the readability trends, we performed a series of statistical analyses.

For each of the two readability measures and three readability components, the mean values at 30 different time points were plotted on a line chart using Excel 2022. This approach allows for a direct visual representation of changes in averages over time. Previous diachronic studies have frequently employed line graphs for similar purposes (Biber & Gray, 2013; 2016). However, these studies often utilized corpora with longer time spans and smaller samples of texts. In contrast, our study calculates the average readability for each year within a shorter 1-year interval, providing a more granular depiction of fluctuations over time. To further elucidate the overall trend or direction of change in readability, a trend line was added to each line graph. This trend line was derived via linear regression, a statistical method that identifies the straight line minimizing the sum of the distances between each data point and the line. This trend line thus offers the best estimate of the overarching direction of the data points, whether upward, downward, or stable.

We examined the diachronic change of readability and its components per year by performing a series of Mann-Kendall trend tests. To evaluate the strength of its diachronic change, the change rate across the examined period was assessed with Theil-Sen's slope estimator (Hussain & Mahmud, 2019). The Mann-Kendall trend test (Kendall, 1975; Mann, 1945) is a widely used nonparametric test for detecting significant trends in time series. Most previous studies measured the strength of diachronic change by comparing the average frequency of use of relevant linguistics features at the beginning and the end of the target time period (Hyland & Jiang, 2019; 2020). Direct comparisons of frequencies at the first and last time points may be "less than optimal" (Gries, 2006) to capture the process of diachronic change.

By employing these methods, our study aims to provide a more detailed and accurate depiction of the diachronic trends in readability to capture its changes.

4. Results and Discussion

4.1 The Readability Trends of the Four Genres

From Table 2 we could see the trend of readability of the four genres in COCA. For each readability measures, we list the results of Mann-Kendall trend tests for the readability time series, including the trend of readability, P-value, and the change rate. As Table 2 shows, each different genres displays different trend.

Table 2. Results of Mann-Kendall Tests for Readability Across Four Genres

Genres	Readability index	Trend	Change rate	P
News	FRE	decreasing	-0.087	0.012

Magazines	NDC	increasing	0.012	0.002
	FRE	no trend	0.048	0.568
Fictions	NDC	increasing	0.029	0.000
	FRE	increasing	0.042	0.002
	NDC	decreasing	-0.004	0.009
Academic articles	FRE	increasing	0.112	0.004
	NDC	increasing	0.011	0.000

4.1.1 The readability Trend of Fiction

The analysis of the diachronic trend in readability for fiction in COCA reveals significant increase. Figure 1 shows an upward trend in FRE scores, supported by a positive slope in the linear regression trend line. Figure 2, displaying NDC scores, reveals a downward trend with a negative slope in the trend line. Both trends are statistically significant, with p-values less than 0.05, underscoring that fiction has become more readable over the examined period. The Mann-Kendall test further corroborates these findings, identifying a statistically significant upward trend in FRE scores and a significant downward trend in NDC scores, with both trends indicating increased readability. Collectively, these results robustly demonstrate a significant increase in the readability of fiction texts in COCA over the past 30 years.

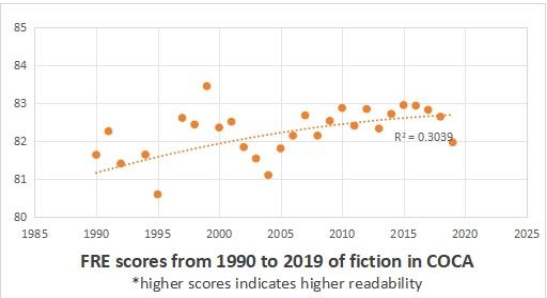


Figure 1. Trends in FRE Scores for Fiction

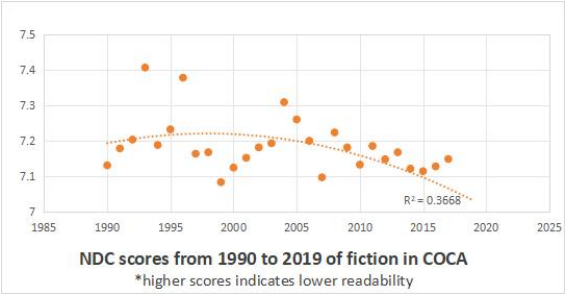
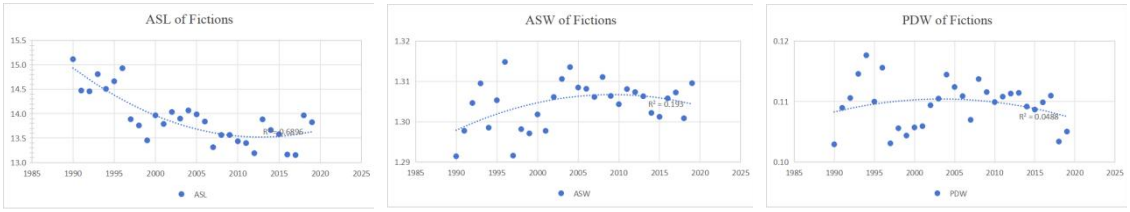


Figure 2. Trends in NDC Scores for Fiction

From following Figure 3 and Table 3, we could tell that the decreasing average sentence length could account for the increasing readability. From 1990 to 2019, the average sentence length of fiction presents a significant decreasing trend and the change in words seems to be insignificant. Thus shorter sentence could be main reason for the increasing readability.



A **B** **C**

Figure 3. Trends in Readability Components for Fiction

Table 3. Statistical Analysis of Readability Components in Fiction

Genres	Readability components	Trend	Change rate	P
Fictions	ASL	decreasing	-0.046	0.000
	ASW	no trend	0.000	0.162
	PDW	no trend	0.000	0.871

The most obvious finding of this part is that the readability of fiction in COCA increased over the 30 years, which means fiction is becoming easier to read. The observed increase in readability could be attributed to the decreasing average sentence length. These results suggest that shorter sentences are effective in enhancing text readability. Our findings are consistent with Tanaka (2012), who also reported a significant improvement in fiction readability. In accordance with the present results, previous studies have demonstrated that the sentence length of fiction has decreased (Rudnicka, 2018). These results are in line with those of previous studies.

The decrease in average sentence length in COCA appears to be influenced by a multifaceted evolution in writing style and reader engagement strategies. Notably, there is a trend towards shorter sentences and a reduction in sentence complexity, which contribute to easier text. This shift may be partly attributed to the integration of more colloquial language into fiction, reflecting a desire to mirror natural speech patterns and make narratives feel more conversational and immediate to readers. Previous studies found that fiction contains a large number of conversations (Biber & Gray, 2016), and these oral features of fiction were found to have become more prominent during the past two centuries (Biber & Finegan, 1989). Recent studies also found a significant increase in dialogue over time in fiction (Muzny et al., 2017). Therefore, the increased proportion of colloquial dialogues in the Fiction genre would result in a decrease in average sentence length and an increase in readability.

Another factor is the influence of social changes with respect to the readers' background and the impact of this on the diachronic changes in the language of English fiction. According to Biber and Conrad (2019), the increase in literacy brought about a rise in demand for fiction that was simpler and easier to read. This could provide a reason why the language used in fiction became easier to read, namely such

language encourages a larger readership, and this, in turn, has commercial benefits. Fiction seems to have become an essential component of mass (popular) culture, although this thesis is controversial (Fluck, 1988). The evidence for this is that children's fiction, science fiction, and fantasy were all created in the 19th century (or earlier) but became popular in the 20th century (Leavis, 2011; McCracken, 1998). These diversified fictional genres were written in plain language in order to attract more readers.

4.1.2 The readability Trend of News

Different from the downward trend of readability in the Fiction genre, the trend of News displays a curve as shown in Figure 4 and Figure 5.

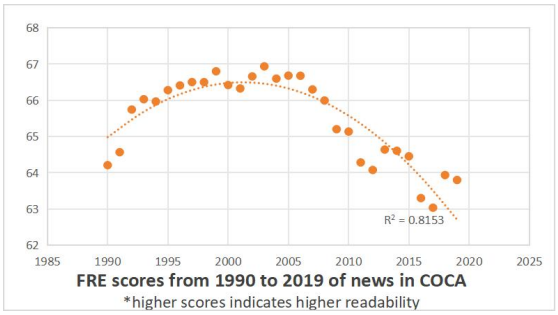


Figure 4. Trends in FRE Scores for News

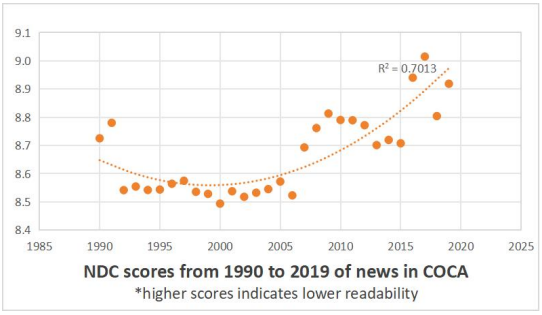
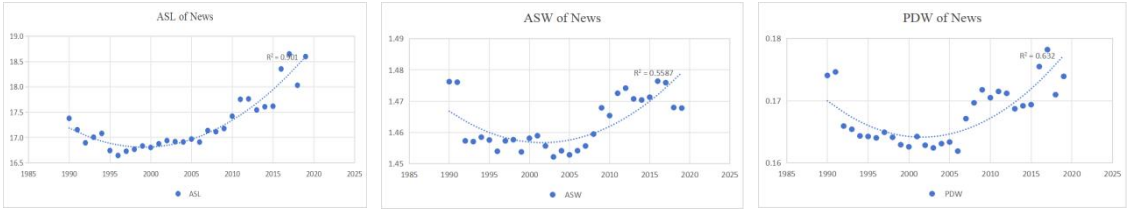


Figure 5. Trends in NDC Scores for News

In terms of FRE, their readability decreases at first and then increases. It shows a decreasing trend in terms of FRE (P=0.012) And NDC shows a significant upward trend (P=0.002) which indicates that news is becoming more difficult to read. All these statistics demonstrate that the readability of news has had a significant decreasing trend over the last 30 years. The result reveals a significant increasing trend in the average sentence length and the average syllable per word in the News. This trend matches with the trend of its readability, which could serve as a reason behind its readability change. The ASL trend of the news is fitted to the polynomial regression model. (R^2=0.901) Although difficult words in the news do not exhibit a significant trend over the 30 years, they changed in these years and also display a curve.



A **B** **C**

Figure 6. Trends in Readability Components for News

The results of this study show that the readability of news has decreased over the past 30 years, indicating that news articles are becoming harder to read. Further investigation suggests that this trend is due to increased sentence length and the use of more polysyllabic words. These findings align with previous research, which also observed a decline in news readability (Dalecki & Lewis, 2009).

Table 4. Statistical Analysis of Readability Components in News

Genres	Readability components	Trend	Change rate	P
News	ASL	increasing	0.045	0.000
	ASW	increasing	0.001	0.030
	PDW	no trend	0.000	0.064

Rudnicka (2018) reported a decrease in the average sentence length of news in the Corpus of Historical American English (COHA). Similarly, Schneider (2002) and Westin (2002) found that the sentence length in newspapers has been decreasing since the 1700s. However, our study reveals a different trend: the average sentence length of news articles has increased since 1990. This discrepancy might be due to differences in the periods studied. While Rudnicka (2018) examined data from 1860 to 2000, our study focuses on the years 1990 to 2019. These findings suggest that news articles in the 21st century exhibit a different trend in sentence length compared to the past. Nonetheless, our study confirms that the decreasing readability of newspapers is related to an increased vocabulary burden (Danielson et al., 1992).

From a sociolinguistic perspective, the decline in readability may be linked to changes in the linguistic features of media texts, driven by evolving demands for information conveyance in recent decades. Society has become more complex, with increases in governmental activities and advancements in science and technology. Consequently, newspapers now cover topics less directly connected to the everyday lives of their readers. Linguistic devices in media texts have become more sophisticated in response to the information explosion in modern society (Biber & Gray, 2011). To convey more diversified, sophisticated, and fine-grained information, media reportage has adopted longer and more complex terms and phrases, particularly noun phrases (Biber & Gray, 2011).

4.1.3 The Readability Trend of Magazine

The FRE scores of magazines increase in the first 15 years and then decrease. No significant increasing or decreasing trend is shown based on the result of the Mann-Kendall tests. In terms of NDC, the readability of magazines shows an upward trend. It shows a significant increasing trend which means that magazines are becoming more difficult to read over the 30 years in terms of NDC.

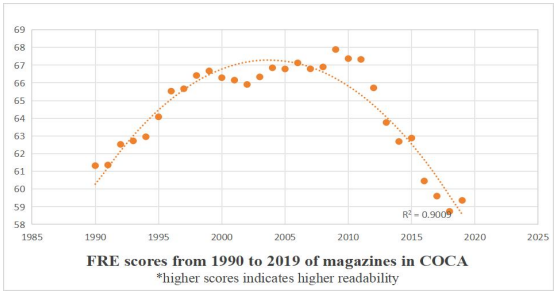


Figure 7. Trends in FRE Scores for Magazine

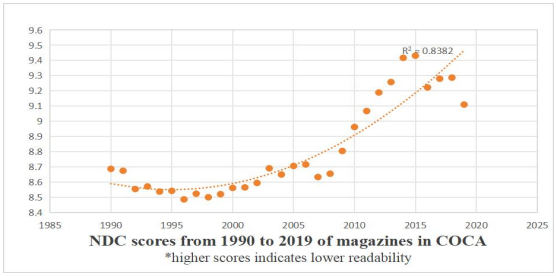


Figure 8. Trends in NDC Scores for Magazine

In particular, Magazine is significantly different from all the other three genres in that the ASL shows no significant trend. PDW of magazines displays a significant increasing trend which could be the main reason for the increasing NDC scores. The ASL and ASW exhibit no significant trend but we could see that it did change over the 30 years. Due to the mixed effect, the FRE scores show no significant trend.

A **B** **C**

Figure 9. Trends in Readability Components for Magazines

Table 5. Statistical Analysis of Readability Components in Magazines

Genres	Readability components	Trend	Change rate	P
Magazines	ASL	no trend	-0.013	0.454
	ASW	no trend	0.000	0.353
	PDW	increasing	0.001	0.009

The FRE scores of magazines showed an increase from 1990 to 2005 but then experienced a decline. In contrast, the NDC scores of magazines have consistently increased over the same period. Understanding these trends requires a closer look at the components and methodologies of these readability formulas, as well as potential shifts in magazine content and language use over time. The FRE score measures readability based on sentence length and syllable count per word. Higher scores indicate easier readability. The increase in FRE scores from 1990 to 2005 suggests that magazine content during this period became more accessible, potentially due to shorter sentences and simpler word choices. This could have been driven by a shift towards a more casual, reader-friendly writing

style aimed at a broader audience. The introduction of the Internet during the 1990s has a role to play here (David, 2010). Many people assume the advent of the new era would bring out the death of print magazines. Therefore magazines choose to use shorter sentences and simpler words to compete with other media. After 2005, the decline in FRE scores suggests a reversal in this trend, with magazines potentially adopting longer sentences and more complex words. This change could be attributed to a shift in editorial policies, targeting a more specialized audience, or a response to changing reader expectations for more in-depth, nuanced content.

The NDC score, on the other hand, assesses readability based on sentence length and the percentage of words not found on a list of familiar words. An increasing NDC score indicates that magazines have been using more complex vocabulary over time. The consistent rise in NDC scores suggests a steady trend towards more sophisticated language use, reflecting perhaps a growing emphasis on specialized content or an attempt to address more educated readers. According to the theory of magazines, magazines always target a precisely defined group of readers (Holmes, 2020). The expanding number of mass media inventions, for example, radio, movies, and television served to whet the appetite for new sources of information that magazines provided. This expanding range of interests created a demand for larger numbers of special interest magazines and a declining market for general interest magazines (David, 2010). It is reasonable that more difficult words will be used in magazines when they become more specialized to attract their target audiences.

In summary, the initial increase in FRE scores from 1990 to 2005 may reflect efforts to simplify language and make content more accessible. The subsequent decline, coupled with the continuous rise in NDC scores, suggests a trend towards more complex language use and specialized content in magazines over time. The differing trends between FRE and NDC scores highlight how readability can be influenced by various factors. While sentence structure and word complexity impact both measures, the specific criteria and focus of each formula can yield different insights.

4.1.4 The Readability Trend of Academic Articles

The change in academic articles is more complicated compared to other genres. Academic articles are becoming more and more difficult to read over time based on the NDC scores. But this finding is consistent with the results of Plavén-Sigray et al. (2017) in PubMed papers, Lei and Yan (2016) in information science journals, and Dolnicar and Chapple (2015) in tourism articles. But the trend is less significant compared with their studies and their overall scores of academic scores are lower than theirs. It could be explained by the materials used in this study. The corpus used in this study is COCA in which academic articles in this corpus are taken from more than 200 different peer-reviewed journals and include full texts, sections, and abstracts. However previous studies focus on the readability of academic abstracts which is less readable than the full texts.

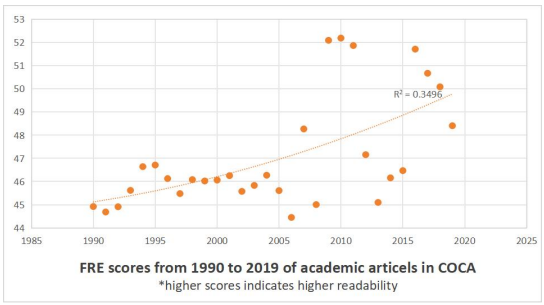


Figure 10. Trends in FRE Scores for Academic articles

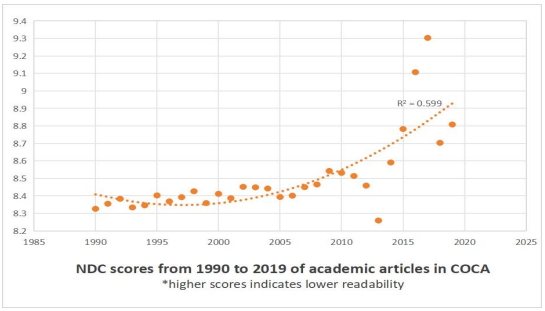
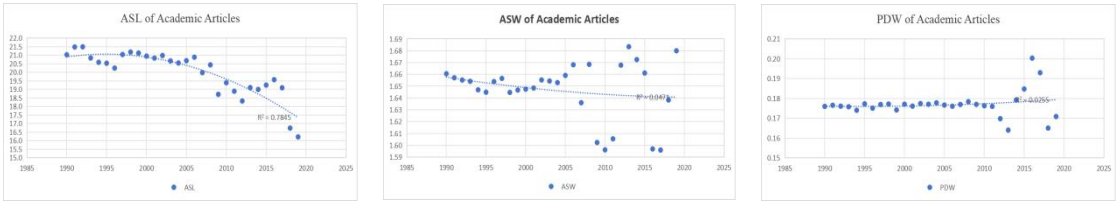


Figure 11. Trends in NDC Scores for Academic articles

Considering the FRE scores, the readability of academic articles has increased over the 30 years. The trend of FRE scores is similar to a study measuring the readability of full texts. Lei and Yan (2016) explored the readability of academic texts and compared the readability of abstracts and full texts from 2003 to 2012. They found that the difficulty levels of abstracts are increasing while those of full texts are dropping and the differences weren't statistically significant.



A **B** **C**

Figure 12. Trends in Readability Components for Academic Articles

Table 6. Statistical Analysis of Readability Components in Academic Articles

Genres	Readability components	Trend	Change rate	P
Academic articles	ASL	decreasing	-0.104	0.000
	ASW	no trend	0.000	0.858
	PDW	no trend	0.000	0.399

The observed increase in both FRE scores and NDC scores for academic articles may initially appear contradictory, but these trends can be explained by changes in writing style and vocabulary usage within the academic community. The rising FRE scores indicate that academic articles are increasingly using shorter sentences and simpler words. This trend reflects a growing emphasis on clarity in academic writing. Reporting scientific findings clearly and accurately is essential for the dissemination

of knowledge and the reproducibility of results (Plavén-Sigray et al., 2017). Modern academic writing prioritizes clear and concise communication to make research accessible to a broader audience, including non-specialists. Consequently, journals may impose stricter guidelines on readability to enhance the impact and citation of articles. Conversely, the increasing NDC scores suggest a rise in the use of more complex and less familiar vocabulary. This can be attributed to the increasing specialization of academic fields, which often necessitates the use of technical jargon and domain-specific terminology. Plavén-Sigray et al. (2017) demonstrated an increase in general scientific jargon over the years. These general science jargon words are frequently used by scientists in scientific texts, indicating a progressively increasing in-group scientific language.

The mixed findings of increasing FRE and NDC scores are consistent with the research of Gross et al. (2002), who identified two opposing trends in the readability of contemporary scientific prose. On one hand, scientific writing is becoming more difficult to read due to the increasing use of complex and compact noun phrases. On the other hand, it is becoming easier to read due to declining sentence length and a reduced number of clauses per sentence. These trends reflect an effort to balance the clarity of presentation with the precision of terminology required in advanced academic discourse. The goal is to make the overall structure of the text easier to follow while still conveying the complex and specialized information necessary for academic work.

In summary, the increasing FRE scores highlight a shift towards clearer and more concise writing in academic articles, while the rising NDC scores indicate the continued use of specialized vocabulary. These trends collectively aim to improve the readability and accessibility of academic texts without compromising the precision and depth required in scholarly communication.

4.2 Genre effects on Measuring Readability and Diachronic Linguistic Changes

This study found that genre exerts a significant effect on the measure of readability and the diachronic trend of readability. It could be concluded from the statistics that different genres may display significantly different readability scores at a given time and the four genres display different diachronic trends across time. In terms of this, this study could contribute to the existing literature from both synchronic and diachronic perspectives.

As is shown in Figure 13-14 and Table 7, significant differences are found to exist in these cross-genre comparisons. Firstly, different genres have different readability levels. Fiction is the most easy-to-read with its average FRE scoring 82.11 while academic articles are always hard to read with an FRE score of 47.04. The average scores of news and magazines are quite close, both of these two types of text could be easily read by the average middle school student. But magazines are a bit harder to read compared with new ones.

Table 7. The Interpretation of Readability Scores of the Four Genres

Genres	Readability index	Mean	School level	Verbal description
News	FRE	65.47	8th grade–9th grade	Standard, plain English. Easily understood by 13–15-year-old students
	NDC	8.67	10th grade–12th grade	Fairly difficult to read
Magazines	FRE	64.47	8th grade–9th grade	Standard, plain English. Easily understood by 13–15-year-old students
	NDC	8.84	10th grade–12th grade	Fairly difficult to read
Fictions	FRE	82.11	6th grade	Easy to read. Conversational English for consumers
	NDC	7.17	9th grade–10th grade	Standard, plain English. Easily understood by 14–16-year-old students
Academic articles	FRE	47.04	College	Difficult to read
	NDC	8.50	10th grade–12th grade	Fairly difficult to read

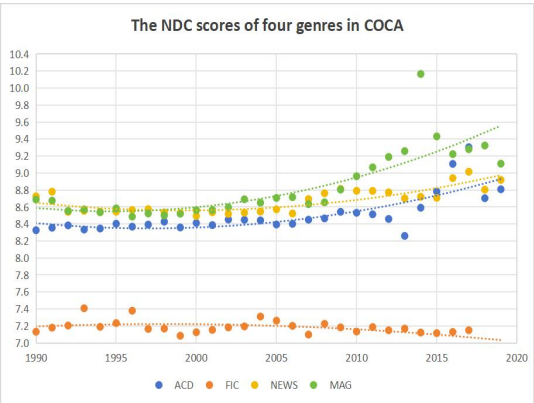


Figure 13. Trends in FRE Scores of Four Genres

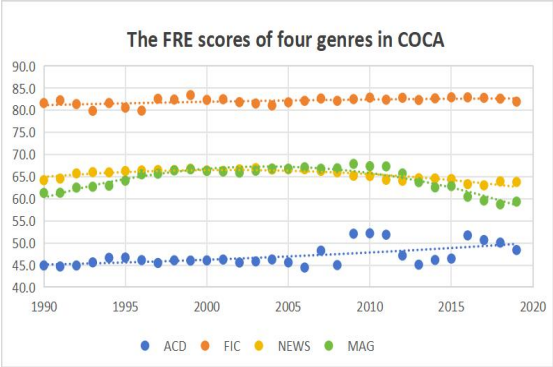


Figure 14. Trends in NDC Scores of Four Genres

Not only they are different in their readability level, but they also show different trends. The difference between these four genres is most obvious in their FRE trend. Although magazines and news seem to have similar trends, the change magazines have undergone is more significant and shows no significant decreasing or increasing trend. In particular, Fiction is significantly different from all the other three genres in FRE trend in that only fiction displays a downward trend.

However, the genre differences in NDC scores are less obvious. As is shown in Figure 14, the NDC

scores of Fiction are significantly lower than those of the other three genres, and only fiction displays a downward trend. While the comparisons between the FRE scores of Magazine, News, and Academic articles show significant differences from each other, the NDC trends of the three genres are relatively close to each other. To summarize, the major distinction of the NDC trend lies between Fiction and the other three genres.

Based on these statistics, this study has affirmed the findings of previous studies that genre exerts a significant effect on readability. The figure with the same readability measurement conducted on time-aligned texts showed that significant differences in readability exist in many of the cross-genre comparisons. Such a distinction may be explained by two reasons. The reason is that different genres differ in the extent to which they have informational purposes and specialized audiences. For example, fictions especially those popular ones aiming at providing entertainment are easy to read since writers would not assume that their readers have any background knowledge (Biber & Gray, 2016). Thus it contains a larger proportion of narration and dialogues compared with other genres, which would result in higher readability. However, the other three genres of News, Magazine, and academic articles share an informational purpose (Biber & Gray, 2016), which emphasizes their usefulness in providing information (e.g., facts and events) (Hollis, 2023). These three are also different because they are written for different audiences with assumed background knowledge. The purpose of conveying information would require heavy use of complicated modifiers such as attributive adjectives and nouns as nominal premodifiers, and hence decrease the readability.

Second, we complement the research of previous studies concerning genre effect on readability from a diachronic perspective. The importance of genre or register on historical linguistic change has been proved. However, there is no such research that compares the readability of different genres from a diachronic perspective. Rudnicka Karolina (2018) investigated the average sentence length of four genres in COHA and found a decrease in sentence length for all the investigated genres but the degree of the decrease is slightly different for each genre. Štajner et al. (2012) compared the distribution and various linguistic features of readability in different genres from a synchronic perspective and found that readability and linguistic features correlated with readability are also genre-dependent. Our diachronic exploration has taken this conclusion a step further and demonstrated notable differences between genres in the development of readability.

From a more macroscopic perspective of historical linguistics, the present study provides further evidence for the claim that historical linguistic change is a process mediated by specific genres and registers (Leech & Smith, 2009; Biber & Gray, 2013). It has been shown that genre differences exist in linguistic changes at various levels such as collocations (Baker & McEnery, 2005), lexical bundles (Culpeper & Kytö, 2010), and lexico-grammatical features (Biber, 2012; Biber & Gray, 2013; Bao et al., 2018). The results in our study further support this view and demonstrate that genre differences also exist in linguistic changes on the vocabulary and sentence level, since the readability measurement we used takes into consideration word and sentence.

This study therefore highlights the importance of considering genre differences in diachronic analysis of readability. This study agrees with Biber and Gray's (2013) idea that language changes should be captured in terms of specific genres or registers, rather than attempting to describe the overall change of the language as a whole. Therefore, an important methodological implication of this study is that diachronic changes in readability can be better studied from a genre-specific perspective. It should be acknowledged that previous studies on a specific genre are important because their findings reveal important underlying trends. However, one should therefore be cautious when making generalizations about language change based on case studies from one or a few genres. Furthermore, there may also be risks in trying to describe "overall" change with a "balanced" corpus composed of mixed genres. Because linguistic phenomena may exhibit different historical patterns between genres, lumping these genres into one overarching trend may ignore some of the underlying patterns at work in certain genres. That is, although multiple sets of data individually support a hypothesis, the combined data set may support the opposite hypothesis (Wang & Rousseau, 2021). Therefore, corpus linguists should also pay attention to potential factors such as genre and register effects when conducting diachronic studies to avoid confounding the results.

5. Conclusion

Based on the large-scale COCA corpus, the study examined the diachronic change in readability in the four genres of contemporary English over 30 years. The results showed that the changing patterns of readability varied across the four genres studied. A general upward trend in readability was found in the fiction genre, while a mixed trend was found in the academic and magazine genres, and a downward trend was found in the news genre. Based on the results, we conclude that there may be multiple factors at work behind the diachronic development of readability in Contemporary English and that these factors may exert varying effects on different genres. The Fiction genre experienced an increase in readability, perhaps driven by the increased proportion of conversational dialogues in fiction (Biber & Finegan, 1989; Muzny et al., 2017). However, in the genres of magazines, news, and academic articles, the mainstream driving force might have been the need to convey more detailed, complex, and sophisticated information, which has resulted in decreasing readability. Methodologically, it is also argued that researchers should be aware of genre distinctions in readability studies and be cautious in drawing generalized conclusions from a comprehensive corpus.

While this study provides valuable insights into the diachronic changes in readability across different genres, several limitations need to be addressed in future research. First, the study relies exclusively on the COCA corpus, which represents American English. This limitation means the findings may not be generalizable to other varieties of English. Future studies should aim to validate these results using corpora from different English-speaking regions, despite potential challenges in data availability. Second, the study covers a relatively short period of 30 years. This time span may not be sufficient to capture long-term trends and shifts in readability. Extending the period of analysis could provide a

more comprehensive understanding of diachronic changes. Third, the research focuses primarily on traditional readability measures. Future investigations should consider examining other linguistic properties that contribute to readability, such as grammar, cohesion, and sentence structure, with a genre-specific approach. This broader perspective could offer deeper insights into how readability evolves over time. Lastly, traditional readability indexes were not designed for scientific articles, which are typically read by highly educated individuals. These indexes were originally intended for evaluating school textbooks, political speeches, newspapers, and local government brochures. Therefore, the use of these measures for academic and complex texts might not fully capture their readability. Critics might argue that this approach is not entirely valid. However, our main goal is to challenge the assumption that historical linguistic changes should be assessed uniformly across all genres. We contend that genre plays a crucial role as a mediating factor in these developments, and changes should be studied relative to specific genres rather than attempting to average the language as a whole. By addressing these limitations, future research can build on our findings and contribute to a more nuanced understanding of the diachronic changes in readability across different genres of English.

Acknowledgments

The work was supported by the MOE (Ministry of Education) Foundation for Humanities and Social Sciences, China (Grant No. 18YJA740002).

References

Badarudeen, S., & Sabharwal, S. (2010). Assessing readability of patient education materials: current role in orthopaedics. *Clinical Orthopaedics and Related Research*, *468*(10), 2572-2580.

Bailin, A., & Grafstein, A. (2001). The linguistic assumptions underlying readability formulae: A critique. *Language & Communication*, *21*, 285-301.

Baker, P., & McEnery, T. (2005). A corpus-based approach to discourses of refugees and asylum seekers in UN and newspaper texts. *Journal of language and politics*, *4*(2), 197-226.

Bao, C., Zhang, X., Qu, Y., & Feng, Z. (2018). American English perfect construction across registers. *Journal of Quantitative Linguistics*, *25*(4), 314-341.

Benjamin, R. (2012). Reconstructing readability: Recent developments and recommendations in the analysis of text difficulty. *Educational Psychology Review*, *24*(1), 63-88.

Biber, D., & Jones, J. (2009). 61. Quantitative methods in corpus linguistics. In A. Lüdeling & M. Kytö (Ed.), *An International Handbook* (pp. 1286-1304, Vol. 2). Berlin, New York: De Gruyter Mouton. https://doi.org/10.1515/9783110213881.2.1286

Biber, D. (1991). *Variation across speech and writing*. Cambridge University Press.

Biber, D. (2012). Register as a predictor of linguistic variation. *Corpus linguistics and linguistic theory*, *8*(1), 9-37.

Biber, D., & Clark, V. (2002). Historical shifts in modification patterns with complex noun phrase

structures. *English historical syntax and morphology: Selected papers from*, *11*, 43-66.

Biber, D., & Conrad, S. (2019). *Register, genre, and style*. Cambridge University Press.

Biber, D., & Finegan, E. (1989). Drift and the evolution of English style: A history of three genres. *Language*, 487-517.

Biber, D., & Finegan, E. (2014). Diachronic relations among speech-based and written registers in English 1. In *Variation in English* (pp. 66-83). Routledge.

Biber, D., & Gray, B. (2011). Grammatical change in the noun phrase: The influence of written language use. *English Language & Linguistics*, *15*(2), 223-250.

Biber, D., & Gray, B. (2013). Being specific about historical change: The influence of sub-register. *Journal of English Linguistics*, *41*(2), 104-134.

Biber, D., & Gray, B. (2016). *Grammatical complexity in academic English: Linguistic change in writing*. Cambridge University Press.

Chall JS, Dale E. (1995). *Readability revisited: The New Dale-Chall readability formula*. Brookline Books.

Collins-Thompson, K, & Callan, J. (2004). A Language Modeling Approach to Predicting Reading Difficulty. In *Proceedings of HLT-NAACL 2004* (pp. 193-200).

Couture, B. (1986). Effective ideation in written text: a functional approach to clarity and exigence.

Culpeper, J., & Kytö, M. (2010). *Early Modern English dialogues: Spoken interaction as writing*. Cambridge University Press.

Dale, E., & Chall, J. S. (1948). A formula for predicting readability: Instructions. *Educational Research Bulletin*, *27*(2), 37-54.

Dalecki, L., Lasorsa, D. L., & Lewis, S. C. (2009). The news readability problem. *Journalism Practice*, *3*(1), 1-12.

Danielson, W. A., Lasorsa, D. L., & Im, D. S. (1992). Journalists and novelists: A study of diverging styles. *Journalism Quarterly*, *69*(2), 436-446.

Didegah, F., & Thelwall, M. (2013). Which factors help authors produce the highest impact research? Collaboration, journal and document properties. *Journal of informetrics*, *7*(4), 861-873.

Diller, H., & Görlach., M. (2001). *Towards a history of English as a history of genres*.

Dolnicar, S., & Chapple, A. (2015). The readability of articles in tourism journals. *Annals of Tourism Research*, *52*, 161-166. https://doi.org/10.1016/j.annals.2015.03.007

Feng, L., Elhadad, N., & Huenerfauth, M. (2009). Cognitively motivated features for readability assessment. In *The 12th Conference of the European Chapter of the Association for Computational Linguistics (EACL 2009)*.

Flesch, R. (1948). A new readability yardstick. *Journal of Applied Psychology*, *32*, 221-233.

Fluck, W. (1988). Fiction and fictionality in popular culture: Some observations on the aesthetics of popular culture. *Journal of Popular Culture*, *21*(4), 49-62.

Fowler Jr, G. L. (1978). The comparative readability of newspapers and novels. *Journalism Quarterly*,

55(3), 589-591.

Fowler Jr, G. L., & Smith, E. J. (1979). Readability of newspapers and magazines over time. *Newspaper research journal*, *1*(1), 3-8.

Gries, T. (2006). Exploring variability within and between corpora: Some methodological considerations. *Corpora*, *1*(2), 109-151.

Hartley, J., Sotto, E., & Pennebaker, J. (2002). Style and substance in psychology: Are infuential articles more readable than less infuential ones? *Social Studies of Science*, *32*(2), 321-334. https://doi.org/10.1177/030631270203200 2005

Hollis, H. (2023). Readers' experiences of fiction and nonfiction influencing critical thinking. *Journal of Librarianship and Information Science*, *55*(1), 18-32.

Holmes, T. (2020). Magazines, Megazines, and Metazines. *The Handbook of Magazine Studies*, 1-19.

Hussain, M., & Mahmud, I. (2019). pyMannKendall: A python package for non parametric Mann Kendall family of trend tests. *Journal of open source software*, *4*(39), 1556.

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

Hyland, K., & Jiang, F. (2020). Changing patterns of self–citation: Cumulative inquiry or self-promotion? *Text & Talk*, *38*(3), 365-387.

Ismail, A., Yusof, N., & Yunus, K. (2016). The readability of Malaysian English children books: A multilevel analysis. *International Journal of Applied Linguistics and English Literature*, *5*(6), 214-220.

Jatowt, A., & Tanaka, K. (2012). Longitudinal analysis of historical texts' readability. *Proceedings of the 12th ACM/IEEE-CS Joint Conference on Digital Libraries* (pp. 353-354).

Kate, R. J., Luo, X., Patwardhan, S., Franz, M., Florian, R., Mooney, R. J., Roukos, S., & Welty, C. (2010). Learning to predict readability using diverse linguistic features. In *Proceedings of the 23rd International Conference on Computational Linguistics (COLING 2010)* (pp. 546-554).

Kayam, O. (2018). The readability and simplicity of Donald Trump's language. *Political Studies Review*, *16*(1), 73-88.

Kendall, M. G. (1975). *Rank Correlation Methods* (4th Ed.). Charles Grifin, London.

Klare, G. R. (1963). *Measurement of readability*. Iowa State University Press.

Klare, G. R. (1980). *A manual for readable writing*. Glen Burnie, MD: REM Company

Leavis, L. (2011). *Fiction and the Reading public*. Random House

Leech, G., & Smith, N. (2009). Change and constancy in linguistic change: How grammatical usage in written English evolved in the period 1931-1991. In *Corpus Linguistics* (pp. 173-200). Brill.

Lei, L., & Yan, S. (2016). Readability and citations in information science: Evidence from abstracts and articles of four journals (2003-2012). *Scientometrics*, *108*(3), 1155-1169.

Ley, P., & Florio, T. (1996). The use of readability formulas in health care. *Psychology, Health & Medicine*, *1*(1), 7-28.

Li, W. (2022). Text genres, readability and readers' comprehensibility. *European Journal of Computer Science and Information Technology*, *10*(4), 52-64.

Lim, E. T. (2008). *The anti-intellectual presidency: The decline of presidential rhetoric from George Washington to George W. Bus*h. Oxford University Press.

Mann, H. B. (1945). Nonparametric tests against trend. *Econometrica: Journal of the econometric society*, 245-259.

McCracken, S. (1998). *Pulp: Reading popular fiction*. Manchester University Press.

Milroy, J. (2001). Language ideologies and the consequences of standardization. *Journal of sociolinguistics*, *5*(4), 530-555.

Muzny, G., Algee-Hewitt, M., & Jurafsky, D. (2017). Dialogism in the novel: A computational model of the dialogic nature of narration and quotations. *Digital Scholarship in the Humanities, 32(suppl_2)*, ii31-ii52.

Plavén-Sigray, P., Matheson, G. J., Schiffler, B. C., & Thompson, W. H. (2017). The readability of scientific texts is decreasing over time. *eLife*, *6*, e27725.

Rudnicka, K. (2018). Variation of sentence length across time and genre. *Diachronic corpora, genre, and language change*, 220-240.

Schneider, K. (2002). *"The" Development of Popular Journalism in England from 1700 to the Present: Corpus Compilation and Selective Stylistic Analysis* (Doctoral dissertation).

Sheehan, K. M., Kostin, I., Napolitano, D., & Flor, M. (2014). The TextEvaluator tool: Helping teachers and test developers select texts for use in instruction and assessment. *The Elementary School Journal*, *115*(2), 184-209.

Štajner, S., Evans, R., Orasan, C., & Mitkov, R. (2012). What can readability measures really tell us about text complexity. In *Proceedings of workshop on natural language processing for improving textual accessibility* (pp. 14-22).

Stevenson, R. L. (1964). Readability of conservative and sensational papers since 1872. *Journalism Quarterly*, *41*(2), 201-206.

Sumner, D. E. (2010). The magazine century: American magazines since 1900. Peter Lang.

Wang, S., Liu, X., & Zhou, J. (2022). Readability is decreasing in language and linguistics. *Scientometrics*, *127*(8), 4697-4729.

Wang, Z., & Rousseau, R. (2021). COVID-19, the Yule-Simpson paradox and research evaluation. *Scientometrics*, *126*(4), 3501-3511.

Westin, I. (2016). Language change in English newspaper editorials (Vol. 44). Brill.

Zamanian, M., & Heydari, P. (2012). Readability of texts: State of the art. *Theory & Practice in Language Studies*, *2*(1), 43-53.