Exploring a Diachronic Change in the Use of English Relative Clauses: A Corpus-Based Study and Its Implication for Pedagogy

Muchamad Sholakhuddin Al Fajri and Victoria Okwar

Abstract
This corpus-based diachronic study aims to investigate the change in the use of English relative clauses over a 45-year time span. It does not only focus on change over time but also change between two varieties of English (British and American). The data were taken from the Brown family of corpora. Each corpus in the Brown family corpora consists of 500 texts of approximately 2,000 words of written published standard English. The finding indicates that the overall trend of the use of relative clauses in written texts has largely decreased in both American and British English. The frequency of the relative which has experienced a sharp decline in both English and American varieties but the relative pronoun that has dramatically increased. This trend suggests the move toward colloquialization, meaning that both the English varieties tend to employ less formal or speech-like style in written texts. The pedagogical implication of this research is that it can bring about a change in syllabuses and materials of English language teaching (ELT), particularly for teaching general writing by taking into account colloquialization hypothesis.

Keywords
diachronic study, corpus-based analysis, relative clauses, corpus linguistics, English language teaching

Introduction
Linguists have developed an interest in the relationship between linguistic knowledge and language use (Wiechmann, 2015, p. 1), giving rise to the move toward linguistic studies that focus on usage-based language which involves examination of frequency of occurrences of different patterns using corpora. This approach is likely to overcome the disadvantage of using a qualitative approach. The findings of qualitative approaches are usually based on the researcher’s introspection and may, therefore, not be generalized as factual accounts of language usage. McEnery and Wilson (2001) reiterate that the findings obtained from a qualitative analysis “cannot be tested to discover whether they are statistically significant or more likely to be due to chance” (p. 76). However, both quantitative and qualitative analyses contribute toward corpus studies. Quantitative analysis suggests what is more frequent and what is rarely used, and its findings can be generalized because they are statistically reliable. On the contrary, qualitative analysis provides greater richness and precision in the interpretation of its findings (McEnery & Wilson, 2001, p. 76). In such way, corpus linguistics has since been canonically employed to investigate various areas of linguistics, including translation (e.g., Bernardini, 2015; Čermákrová, 2015), thematic analysis (e.g., Tan & Perudin, 2019), lexicography (e.g., Paquot, 2015; Wild et al., 2013), English language teaching (e.g., Aminui, 2019; Moon & Oh, 2018; Timmis, 2015), and discourse studies (e.g., Berger et al., 2017; Fajri, 2017).

This article seeks to examine the diachronic change in the use of English relative clauses across multiple corpora, more specifically, the Brown Family of corpora and more recent equal corpora: BE06 and AmE06. Huddleston and Pullum (2005) define a relative clause as a “subordinate clause whose primary function as is as a modifier to a noun or nominal” (p. 183). “In addition to the presence of a verb, relative clauses sometimes have a relative pronoun [relativizer] which functions as a boundary between the head noun and relative clause” (Akiyama, 2002, p. 27). Xu and Xiao (2015) sum up the components of a relative clause as: head noun + relativizer + the gap. The relativizer refers back to the head noun

1Universitas Airlangga, Surabaya, Indonesia
2Lancaster University, Lancashire, UK

Corresponding Author:
Muchamad Sholakhuddin Al Fajri, Faculty of Vocational Studies, Universitas Airlangga, Jl. Dharmawangsa Dalam Selatan No.28 - 30, Airlangga, Kec. Gubeng, Kota SBY, Jawa Timur 60286, Indonesia. Email: m-sholakhuddin-al-fajri@vokasi.unair.ac.id
(antecedent), while the gap is a missing element in the relative clause which corresponds in meaning to the antecedent. Relative clauses may be restrictive (also known as defining/integrated) or nonrestrictive (nondefining/supplementary). Biber et al. (1999) distinguish the two types as: “restrictive relative clauses are used to establish the reference of the antecedent, while non-restrictive relatives give additional information which is not required for identification” (p. 195). However, it is difficult to distinguish restrictive and nonrestrictive relatives on the basis of semantic function alone (Akiyama, 2002, p. 214). Indeed, relative clauses have become one of concerns in English language teaching since English language learners often have difficulties in processing and understanding relative clauses (Hamilton, 1994; Izumi, 2003). Thus, a study on the use of English relative clauses over time seems significant to shed light on changes in contemporary English, offering constructive suggestions on the development of English language teaching materials.

**Corpus Linguistics and English Relative Clause**

Some linguists have conducted corpus-based studies on English relative clauses. They are predominantly aimed at identifying semantic and syntactic functions of relative clauses. For examples, Tse and Hyland (2010) examine the role of relative clause constructions in a corpus of 200 journal descriptions in four contrasting disciplines (biology, engineering, applied linguistics, and sociology). The study revealed that relative clause constructions in the genre function as evaluative and persuasive tools to promote the academic journals to both audience and prospective journal writers. Denison and Hundt (2013) use a case study of the science genre from the 17th to the 19th centuries to test the demarcation between restrictive and nonrestrictive relatives. The data from their study was retrieved from the British part of the ARCHER 3.2 corpus and annotated with a parser. Their study reveals that the distinction; that is, restrictive and nonrestrictive, is problematic as it fails to capture the range of relative clauses. They suggest a revised classification with four main bands: restrictive, aspective, nonrestrictive, and continuative (p. 159).

Biber et al. (1999) used the Longman Grammar of Spoken and Written English Corpus (LGSWE) to describe the actual use of grammatical features in different registers. Their study on the use of relativizers (relative pronouns and relative adverbs) focused on different registers which included: conversation, fiction, newspaper and academic prose. It revealed the recurrent use and the importance of relative clauses in these registers, conveying essential information about particular referents in the texts.

Xu and Xiao (2015) explored changes in relative clauses in “spoken” British English on the basis of the Diachronic Corpus of Present—Day Spoken English (DCPSE). Their study was aimed at investigating whether the changes in speech have proceeded in parallel with those in writing over a span of 30 years in the second half of the 20th century. Akiyama (2002) focused on the change in the use of the infinitival relative clause based on the British National Corpus (BNC). These studies reveal an overall trend of using less formal relativizers (that, zero and preposition stranding relatives) as opposed to the formal ones (wh- and pied-piping relatives). Another comparative study is Lu and Chen’s (2008), which investigated the similarities and differences of restrictive relative clauses among three languages (English, Spanish, and Chinese) by comparing and contrasting parallel data extracted from a parts of speech (POS)-tagged multilingual corpus. Their study can be applied to second-language learning.

Furthermore, studies such as by Wiechmann (2015), Takashima (2000), and Cho and Lee (2016) have focused on the structural complexity of the relative clauses. Wiechmann (2015) carried out a comprehensive corpus-based analysis of relative clause constructions with the aim of finding out the difficulty of processing such constructions. He draws on ideas from cognitive psychology, computational linguistics, and artificial intelligence. Data used in his research are drawn from the British component of the International Corpus of English (ICE-GB R2). Takashima (2000) used the Nijmegen Corpus of written data to investigate whether the order of difficulty that Japanese high school and university students perceive with respect to English relative clauses corresponds to the frequency count order from the corpus. He found that there is a frequency order which corresponds to order of difficulty. Cho and Lee (2016) revealed that there is high frequency of relative clauses in science and engineering journals. They advocate for the need to teach relative clause constructions to students who interact with such journals to aid comprehension.

Research on the diachronic use of relative clauses in written texts has also been conducted by Leech et al. (2009). Using the Brown family corpora, they concluded that there was an increase in the use of that-relatives, while Wh-relatives experienced a decline. However, this study does not take into account BE06 and AmE06 corpora so that more recent changes are not revealed. Therefore, this study attempted to fill this gap by investigating the more recent corpora, AmE06 and BE06.

This article focuses on relative clauses that are introduced by relative pronouns: who, whom, whose, which (wh-relatives), and that. The central purpose of the research is to determine whether the use of relative clause constructions has increased or decreased over a 45-year period in British and American varieties of English and whether the rate of change between the two varieties is similar or different. Unlike the related aforementioned studies, this study is bidimensional, it does not only focus on change over time but also change between two varieties of English. The frequency of occurrence of the relative pronouns across all four corpora will be compared. We will attempt to give explanations for
the occurrence of some frequency patterns and changes. Some of the explanations will be based on examination of some concordance lines.

**Data and Methodology**

The data for the analysis were retrieved from the Brown Family corpora which consists of Lancaster 1931, LOB corpus (1961), and FLOB corpus (1991), which are representative of written British English, and Brown corpus (1961), and Frown corpus (1992), which are representative of the American variety. In addition, I also employed more recent equal corpora: BE06 (2006) and AmE06 (2006). For the purpose of this study, however, Lancaster 1931 corpus was not used to obtain the same time spans, a 45-year period. Each corpus in the Brown family consists of 500 texts of approximately 2,000 words of written published standard English. The corpora have been used by researchers for comparisons across time and between cultures to examine both synchronic and diachronic change (Baker, 2009). The Brown corpora cover a wide range of texts from different genres; that is, fiction, general prose (nonfiction), learned (academic), and press. The BE06 corpus has the same sampling frame as LOB and FLOB corpora, while the AmE06 corpus uses the same sampling frame as Brown and Frown corpora. For this study, the Brown Family (C8 tags) was used. The C8 tagset includes additional POS tags used for recognizing relative pronouns; that is, PNLS for who, PNLO—whom, DDLGE—whose, DDL—which, and WPR—that. On the contrary, the CLAWS7 tagset was used in the BE06 and AmE06 corpora. It also applies additional POS tags including which_DQ for which, DDQGE—whose, PNQO—whom, PNQS—who, _NN* THAT_CST—that. The C7 and C8 tagset made the search more grammar sensitive; hence, retrieval of only the instances where the words are used as relative pronouns. Hunston (2002, p. 79) refers to this kind of search as a category-based methodology as it involves using the categorized parts of a corpus for the basis for corpus searches and statistical manipulations. All the corpora were accessed through online corpus software “CQPweb” (Hardie, 2012). The software enables us to conduct several corpus linguistic techniques including frequency list and concordance analysis.

The frequency of occurrence for each relative pronoun across the corpora was obtained and listed in Table 1.

The raw frequency counts as opposed to the normalized frequency were used because the corpora are of the same size hence comparable. A normalized frequency is only used if the corpora used are not of the same length and are thus not directly comparable (Biber et al., 1998, p. 263). The analysis of rate of change (the difference in percentage column) was obtained by the following formula:

British English: \( \text{BE06 frequency} - \text{LOB frequency} \times 100 \)

American English: \( \text{AmE06 frequency} - \text{Brown frequency} \times 100 \)

The difference in the rate of change between the two varieties is likely to be comparable because of the same time spans. The frequency counts of all the relative pronouns in both varieties experience change over time.

**Findings**

Considering the above frequency list, it can be said that relative clauses that begin with **which** were the most frequently used in the British varieties, while **that** was the most frequently used in the American corpora. On the contrary, relative clauses that begin with **whom** and **whose** were the least frequently used in both varieties. In general, the British variety used slightly more relatives in 1961 compared to the American variety, but in 2006 the American variety had a higher frequency count of relatives.

The trends in the change of frequency counts of the relatives between the two varieties differ slightly; in that, while in the British English corpora, the frequencies of **wh**-relatives, except **who**, decreased across the 45-year period, in the American variety, the frequencies of **wh**-relatives also declined, but except **who** and **whom**. The relativizer **that**, in addition, rose dramatically in both variety of corpora.
The frequency list of the relative pronouns partly meets the purpose of this study: to determine whether the use of relative clause constructions has increased or decreased over a 45-year period in British and American varieties of English and whether the rate of change between the two varieties is similar or different. The frequency list, however, does not give explanations for these frequency patterns. The study would be incomplete without such explanations. The analysis of frequency lists can be defined as quantitative analysis. Leech et al. (2009) state that “it is characteristic of corpus linguistics to begin with quantitative findings and work towards a qualitative one” (p. 31). Qualitative analysis requires the researcher to make use of his or her introspection to interpret the findings obtained from quantitative research. “It may be the means for classifying examples in particular corpus by their meanings” (Leech et al., 2009, p. 32). The explanations in this section will mostly be compared to Biber et al.’s (1999) LGSWE corpus study that examined the use of relatives across different registers, and Leech et al.’s (2009) analysis on the Brown family corpora.

The Relative Pronoun Which

Although there has been a decrease in the usage of the relative pronoun which over time, its frequency still remains relatively high in both British and American varieties. This corresponds to Biber et al.’s (1999, p. 609) findings. The high-frequency count could be because which can be used with subject, object, and object of the preposition (pied-piping) antecedents. For example:

1. **As subject**: BR adopted a further criterion which was specifically chosen for operation. (FLOB_J74)
2. **As object**: John Wesley has himself left a list of the books which he read with his pupils. (LOB_G18)
3. **As object of preposition (pied-piping)**: Scalfaro must now tackle the same Mafia terror to which he owes his election. (FROWN_A33)

The relative which is used with nonhuman antecedents, for example, criterion, books, classes, as in the above examples. Nonhuman nouns are commonly used in nonfictional and academic genres which make up a huge chunk of the corpora studied; hence, the high-frequency counts of which, while human nouns are frequently used in fictional texts.

Another reason for the high frequency of which is that it can be used to construct both restrictive and nonrestrictive relative clauses. The examples (1), (2), and (3) are all restrictive. Examples of nonrestrictive clauses:

4. Developed that the Soviets had a very different interpretation of democracy, which will be discussed later. (BROWN_J36)
5. BA is liable to pay him the going rate for goods damaged under the Warsaw Convention, which regulates international air travel. (FLOB_A13)
6. They should realize that the mayor won by a big landslide, which indicates that the citizens are supportive of the mayor’s vision. (AmE06_A01)

Turning to the decrease in frequency of which, the decline of the use of which could be due to colloquialization. Biber et al. (1999, p. 616) state that “which has more conservative, academic associations” hence is considered more formal. However, even in learned (academic) texts, the employment of which-relativiser still experienced a decline (see Figures 1 and 2). It could be argued that both varieties of English are becoming less formal although the American variety is leading the way as depicted by the lower frequency counts since the beginning of our period, in 1961.

![Figure 1. Occurrences of relative which in the British corpora.](image-url)
Which is also used to construct complex sentences. A decrease in its frequency could indicate a decrease in the use of highly complex sentences in favor of more simple sentence constructions. Indeed, 100 random samples of concordance lines suggest that only 26 of them seem to use more complex sentences or use nonrestrictive clauses. Examples of complex sentences:

7. Conversely, lower dilutes, like 1X or 6X, which characterize most over-the-counter medicines, contain much higher amount of the active ingredient. (FROWN07)

8. Sydney Larson, a staff representative for the United Steel Workers, which the firm’s 25 workers joined before striking, said the state Labor Relations Board. (BROWN02)

9. They forecast that banks’ prime lending rate, to which many business and consumer loans are tied, would drop from the current 6% to 5.5%. (FROWN03)

The substantial fall in the frequency of the relative which needs to be taken into account alongside the considerable growth in the use of that, which will be discussed in the following section.

**The Relative Pronoun That**

The move toward colloquialization means employing the use of the less formal that-relatives more frequently, hence the increase in the frequency count of that across time and between the two varieties (see Figures 3 and 4). Since the American variety is leading the way, it could be said that British English is becoming more Americanized as it follows the lead. Americanization refers to the American influence on change (Antonio et al., 2019; Leech et al., 2009). As Leech et al. (2009, p. 230) pointed out, the fall of which, and the corresponding rise of that is certainly affected by a strong tradition in the United States of refusing which “as an introducer of restrictive relative clauses,” which has influenced countries across the world, including the United Kingdom. The finding of this study also give evidence on Leech et al.’s (2009) suspicion that there would be a more considerable increase in relative that and decline in relative which in more recent corpora of written texts.

While that may be taking the place of which in most instances, it is unlikely that it can replace it fully as that can only be used with restrictive relative clauses while which can be used with both restrictive and nonrestrictive (Biber et al., 1999, p. 615). Since restrictive relative clauses are less complex than nonrestrictive clauses, the increase in the frequency of that supports the hypothesis that there could be a tendency to use simpler sentences.

Apart from colloquialization, the high frequency of that could also be explained by the fact that it is quite flexible in its use. It can be used to refer to both human and nonhuman antecedents. This means it can replace all the wh-pronouns. Just like which, it can also be used with subject, object, and object of preposition (preposition stranding) antecedents. Examples:

10. **Human subject:** She picked her own Middle-Eastern friends from the flock of ardent Egyptians that buzzed around her. (BROWN09)

11. **Non-human subject:** the stupid things that are still done are exemplified in an account by Brooks. (LOBJ07)

12. **Object:** It would please them to think of me doing the things that normal children in the sixth grade did, like going to movies. (AmE06_G54)
Preposition stranding: tremendous fillip by the Stanley scoop of 1871 and all the publications that flowed from it. (FLOB_G38)

In addition, even though we did not specifically analyze zero relative clauses or bare relatives in detail, it appears necessary to look at them in the discussion of colloquial styles. Leech et al.’s (2009) investigation into zero relatives in the Brown family of corpora (excluding Lancaster 1931 corpus) demonstrated that there is an increase in their frequency. Examples of zero relative clauses:

13. Brandt and Walker, answers Richards. They’re the ones we can expect to do better. (BROWN_B10)

14. Brandt and Walker, answers Richards. They’re the ones we can expect to do better. (BROWN_B10)

15. in his absence England put into the field the most ill-assorted attack I can ever remember. (LOB_A08)

16. The chief challenge we face is to define ourselves as more than servicing institutions that negotiate. (FROWN_G10)

This may again be attributed to colloquialization because zero relatives are more oral in their affinity (Leech et al., 2009).
The Relative Pronoun Who

The relative pronoun *who*, although not as frequent as *which* and *that*, also has relatively high frequencies. The frequency of *who* has increased over time in the American variety, while in the British corpora, it witnessed a fall in 1991 but rose slightly in 2006. This might also support what Leech et al. (2009) call Americanization, in that British English is influenced by changes in American English. The slight increase in the frequency counts in the American and British English variety; however, contradicts the colloquialization hypothesis. A look through a number of concordance lines does not give possible suggestions for this phenomenon. However, this might be because in nonrestrictive relative clause, the relative *who* is more significantly preferred than *that* for human antecedents (Biber et al., 1999). The relative *who*, unlike *which* and *that*, is restricted in its use. This explains its lower frequency counts as compared to the other two. *Who* is only used with human subject antecedents in both restrictive and nonrestrictive relative clause construction. Human antecedents occur more often in fictional writing than in other text genres such as academic prose. For examples:

17. But a particular perspective is shared by all those who take the question radically from the other side. (FLOB_J65)
18. Perez, who was elected to a second term as president in 1988, has carried out an ambitious free-market reform program. (FROWN_A02)
19. Graves in his 1808 History of Cleveland paints an interesting picture of the boy who would change the face of the globe. (BE06_G07)

The Relative Pronouns Whom and Whose

The relative pronouns *whom* and *whose* have the least frequency counts across the corpora. Both *whom* and *whose* are restricted to specific syntactic roles. They are not as flexible as *which* and *that*, hence the low-frequency counts. While their frequency counts in the British English variety decrease over time for both relatives, the frequency counts of *whom* in the American variety slightly increase over the 45-year span. The increase in the frequency of *whom* in American corpora could be because the written texts in American standard English may have a tendency to use more proper and personal nouns which are often antecedents of the *who*, *whom*, and *whose* relatives.

*Whom* is used with human object antecedents. Biber et al. (1999, p. 621) state that subject antecedents are more common than object antecedents so that the frequency counts for *who* is higher compared to *whom*. Examples of relative *whom*:

20. **Human object**: He is Aldo Rostagno, son of the Guglielmo Rostagnos of Florence whom the Burkes met last year in Europe. (BROWN_A16)

21. **Pied-piping preposition**: she found that the last person in the world with whom she would willingly have shared this tense period of waiting was ensconced. (LOB_F01)

Another explanation for low-frequency counts of *whom* is that the use of *that* is often preferred when trying to avoid the formal overtones of *whom* (Biber et al., 1999, p. 615). In addition, the considerable decline of *whom* in the British corpora might indicate colloquialization/informalization, in that the use of relative *whom* is likely to be replaced by *who* or bare relatives. Examples of relative *who* and zero relatives used with human object antecedents:

22. But for me, Devilfish was probably the first proper player who I started to study playing. (BE06_E28)
23. Offered by debt management companies who charge for their services, and who you may have seen advertised on the television. (BE06_F37)
24. Containing the names of thousands of secret police collaborators, informers and the citizens they were monitoring. (BE06_A07)

Biber et al. (1999) and de Haan (2002) stated, *whom* is very rare in spoken registers but is more common in written data, thus suggesting more formal English. However, the most noticeable decline in the use of *whom* in the British corpora is in learner or academic texts, which should be generally more formal.

Turning to the relative *whose*, this relativizer is restricted to the syntactic role of marking possessive relations. The lack of flexibility in its use accounts for its low frequency across the corpora. The use of *whose* in the British English variety has decreased dramatically over time, while it just declined slightly in the American variety. Unlike *who* and *whom* that mostly refer to human antecedents, *whose* can mark a possessive relation with both human and nonhuman antecedents. Examples:

25. **Human possessive relation**: Lord Bridge, whose father was a commander in the Royal Navy, was sent to Marlborough public school. (FLOB_G72)
26. **Non-human possessive relation**: As I wandered among those rich cabinets displaying ornate pistols and rifles whose carved wood stocks were embossed. (FROWN_G04)

Discussion

The major reason for the trend in the changing use of relative clauses is likely to be colloquialization. The dramatic decrease in the use of *which* relative and the significant rise
in relative *that* is one indicator of this phenomenon. Using speech-like style in written texts including the more formal texts may enable readers to better interact and engage with such texts. Although the frequency of the relatives *who* and *whom* in the American English variety has experienced a small increase, this does not necessarily contradict the colloquialization hypothesis. The slight increase may suggest no reason by itself (Leech et al., 2009, p. 229).

This finding is congruent with Hundt and Mair’s (1999) study that language of press was responsive to colloquialization, and Biber and Finegan’s (1997) research that popular written texts, such as fiction, were more oral in style during the 19th and 20th centuries. This study thus gives evidence that such trend still continues. On the contrary, while both the previous studies state that academic writing is more resistant to this trend, our analysis suggests the opposite. This study therefore may echo the finding of more current research by Hyland and Jia (2017) that academic writing tends to be more informal. However, we need to be more cautious in this claim because relative clause is only one of some features that can suggest informality. As Hyland and Jia (2017) pointed out, formalization/colloquialization in academic writing depends on the fields of knowledge and features being considered.

Furthermore, it should be noted that the findings of this study may not be able to be overgeneralized. Corpus-based studies based on the examination of the Brown family corpora including AmE06 and BE06 may be criticized as they contain written, published texts that are edited; hence, do not reflect actual or normal use. Comparisons with studies based on corpora of spoken language can be used to examine differences. Biber et al.’s (1999) study is a good example of such a comparative study. However, Baker (2009, p. 314) points out that a million words is likely to be acceptable for investigating grammatical words, although “only very cautious conclusions can be made about other lexis.” This research, which has relatively similar results to Biber et al.’s (1999) study based on the LGSWE corpus containing more than 40 million words, may lend weight to and confirm the findings of this study that language of press was responsive to colloquialization, suggesting that both the English varieties tend to move toward colloquialization, with slight increase in the use of English relative clauses in the Brown corpus, while *fines, imprisonment, sentences, and deposition* occur more frequently in LOB. They suggest that their finding offers a picture of United States and British culture in 1961. In addition, Li et al. (2020, p. 1) investigate the use of words *gambling* and *gaming* based on the English web corpus enTenTen13. They found that the change in their use corresponded to specific socio-historical events, indicating that “linguistic variations as collective human behavior changes can be leveraged to evidence other collective human behavior changes.”

In addition, the findings of this study can be employed to look critically at existing English teaching syllabuses and materials, particularly teaching English as a foreign language (TEFL). Corpora have been used to inform syllabuses on English language teaching, especially English as first language (EFL) and English as second language (ESL) contexts (Jones & Waller, 2015). This study, for example, may inform teachers about the change in the use of relative clauses so that an adjustment of syllabuses or teaching materials may be made. The trend toward colloquialization in some written genres, which appear to be canonically overlooked, might be considered in the syllabus design and materials development, particularly for teaching general writing.

Colovic-Markovic (2018) points out that ESL/EFL advanced learners often find challenging in understanding the production of reduced relative clauses or zero relatives. Teachers therefore may need to put more emphasis on explaining the processes for relative clause reduction including which relative clauses can and should be reduced, and how to use reduced relative clauses, which are becoming more common in written texts. The use of corpora to teach English relative clauses is also suggested “to make students aware of their errors which ultimately leads them to self-correction” (Sahragard et al., 2013, p. 79). In addition, instructors should be acquainted with the discipline-specific features of relative clause structures to avoid misinforming their students about the choice of relative pronouns appropriate in their discourse community (Cho & Lee, 2016).

**Conclusion**

This study has endeavored to investigate the diachronic change in the use of English relative clauses in the Brown Family of corpora. The finding suggests that the use of relative pronoun *which* has decreased in both English and American varieties, and the relative pronoun *that* has dramatically risen, which is congruent with Biber et al.’s (1999) research. This trend might indicate the move toward colloquialization, suggesting that both the English varieties tend to employ less formal or speech-like style in written texts such as press, general prose, and fiction. In addition, while in British English the use of *whom* and *whose* significantly
declined, the use of whom in American variety increased slightly. This is likely because there is a tendency to use more proper and personal nouns, which are often antecedents of the who, whom, and whose relatives, in the written texts in American standard English. The pedagogical implication of this study may be that it can enrich teachers’ understanding of the use of relative clauses over time, which may bring about a change in TEFL syllabuses and materials, considering colloquialization hypothesis.

As mentioned in the previous section, this study cannot make any general claims about the change in the use of relative clauses between the British and American varieties and across the time period represented in the corpus. The study briefly illustrates the kind of grammatical or even lexical analysis possible using a diachronic corpus. The frequency counts give statistical evidence that there have been some changes, but more research needs to be done about how these changes are affected. More research could be carried out to establish if colloquialization and Americanization could affect significant changes in different varieties of English. Research could also be carried out to investigate the lexical or grammatical items that are taking the place of relative pronouns; for example, more comprehensive studies on zero relativizers.

In addition, Xu and Xiao (2015) state the importance of exploring spoken language to determine changes in language use. They claim such changes are likely to originate from spoken language. This argument can be supported by the fact that colloquialization, which is a speech-like style, is taking root in written standard English. Analyses of both spoken and written registers could be carried out to determine the influence that spoken registers may have on written registers.

**Declaration of Conflicting Interests**

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

**Funding**

The author(s) received no financial support for the research, authorship, and/or publication of this article.

**ORCID iD**

Muchamad Sholakuddin Al Fajri https://orcid.org/0000-0002-5628-065X

**References**

Akiyama, T. (2002). *The infinitival relative clause in English: An analysis based on the British National Corpus* [Unpublished doctoral thesis]. Lancaster University, Lancaster.

Anmuai, W. (2019). Analyses of rhetorical moves and linguistic realizations in accounting research article abstracts published in International and Thai-Based Journals. *SAGE Open*, 9(1), 1–9. https://doi.org/10.1177/215824018822384

Antonio, A. F., Bacang, B. G., Rillo, R. M., Alieto, E. O., & Caspillo, W. D. (2019). American or British? A corpus-based analysis of Asian Englishes’ orthographical norms. *Asian EFL Journal*, 26(6.2), 180–204.

Baker, P. (2009). The BE06 Corpus of British English and recent language change. *International Journal of Corpus Linguistics*, 14(3), 312–337. https://doi.org/10.1075/ijcl.14.3.02bak

Baker, P. (2011). Times may change, but we will always have money: A diachronic variation in recent British English. *Journal of English Linguistics*, 39(1), 65–88. https://doi.org/10.1177/075424210368368

Berger, C., Friginal, E., & Roberts, J. (2017). Representations of immigrants and refugees in US K-12 school-to-home correspondence: An exploratory corpus-assisted discourse study. *Corpora*, 12(2), 153–179. https://doi.org/10.3366/cor.2017.0115

Bernardini, S. (2015). Translation. In D. Biber & R. Reppen (Eds.), *The Cambridge handbook of English corpus linguistics* (pp. 515–536). Cambridge University Press.

Biber, D., Conrad, S., & Reppen, R. (1998). *Corpus linguistics: Investigating language structure and use*. Cambridge University Press.

Biber, D., & Finegan, E. (1997). Diachronic relations among speech-based and written registers in English. In T. Nevalainen & L. Kahlas-Tarkka (Eds.), *To explain the present: Studies in the changing English language in honour of Matti Rissanen* (pp. 253–275). Société Néophilologique.

Biber, D., Johansson, S., Leech, G., Conrad, S., & Finegan, E. (1999). *Longman Grammar of Spoken and Written English*. Pearson Education Limited.

Čermáková, A. (2015). Repetition in John Irving’s novel a widow for one year: A corpus stylistics approach to literary translation. *International Journal of Corpus Linguistics*, 20(3), 355–377. https://doi.org/10.1075/ijcl.20.3.04cer

Cho, D. W., & Lee, K. (2016). English relative clauses in science and engineering journal papers: A comparative corpus-based study for pedagogical purposes. *Ampersand*, 3, 61–70. https://doi.org/10.1016/j.amper.2016.03.002

Colovic-Markovic, J. (2018). Teaching relative clauses. In J. I. Liontas, M. DelliCarpini, & S. Abrar-ul-Hassan (Eds.), *The TESOL Encyclopedia of English Language Teaching* (pp. 1–8). Wiley Blackwell.

de Haan, P. (2002). Whom is not dead?! In P. Peters, P. Collins, & A. Smith (Eds.), *New frontiers of corpus research* (pp. 215–228). Brill Rodopi.

Denison, D., & Hundt, M. (2013). Defining relatives. *Journal of English Linguistics*, 41(2), 135–167. https://doi.org/10.1177/075424213483572

Fajri, M. S. A. (2017). Hegemonic and minority discourses around immigrants: A corpus-based critical discourse analysis. *Indonesian Journal of Applied Linguistics*, 7(2), 381–390. https://doi.org/10.17509/ijal.v7i2.8349

Hamilton, R. (1994). Is implicational generalization unidirectional and maximal? Evidence from relativization instruction in a second language. *Language Learning*, 44(1), 123–157.

Hardie, A. (2012). CQPweb: Combining power, flexibility and usability in a corpus analysis tool. *International Journal of Corpus Linguistics*, 17(3), 380–409. https://doi.org/10.1075/ijcl.17.3.04har

Huddleston, R. D., & Pullum, G. K. (2005). *A student’s introduction to English grammar*. Cambridge University Press.
Hundt, M., & Mair, C. (1999). “Agile” and “uptight” genres: The corpus-based approach to language change in progress. *International Journal of Corpus Linguistics, 4*(2), 221–242. https://doi.org/10.1075/ijcl.4.2.02hun

Hunston, S. (2002). *Corpora in applied linguistics*. Cambridge University Press.

Hyland, K., & Jiang, F. (2017). Is academic writing becoming more informal? *English for Specific Purposes, 45*, 40–51. https://doi.org/10.1016/j.esp.2016.09.001

Izumi, S. (2003). Processing difficulty in comprehension and production of relative clauses by learners of English as a second language, 53(2), 285–323. https://doi.org/10.1111/1467-9922.00218

Jones, C., & Waller, D. (2015). *Corpus linguistics for grammar: A guide for research*. Routledge.

Leech, G., & Fallon, R. (1992). Computer corpora—What do they tell us about culture. *ICAME Journal, 16*, 29–50.

Leech, G., Hundt, M., Mair, C., & Smith, N. I. (2009). *Change in contemporary English: A grammatical study*. Cambridge University Press.

Li, L., Huang, C.-R., & Wang, V. X. (2020). Lexical competition and change: A corpus-assisted investigation of gambling and gaming in the past centuries. *SAGE Open, 10*(3), Article 2158244020951272. https://doi.org/10.1177%2F2158244020951272

Lu, H.-C., & Chen, Y.-H. (2008). *A corpus-based study of restrictive relative clauses* [Paper presentation]. The International Symposium on Using Corpora in Contrastive and Translation Studies, Hangzhou, China.

McEnery, T., & Wilson, A. (2001). *Corpus linguistics: An introduction* (2nd ed.). Edinburgh University Press.

Moon, S., & Oh, S. Y. (2018). Unlearning overgenerated “Be” through data-driven learning in the secondary EFL classroom. *ReCALL, 30*(1), 48–67. https://doi.org/10.1017/S0958344017000246

Paquot, M. (2015). Lexicography and phraseology. In D. Biber & R. Reppen (Eds.), *The Cambridge handbook of English corpus linguistics* (pp. 460–477). Cambridge University Press.

Sahragard, R., Kushki, A., & Ansaripour, E. (2013). The application of corpora in teaching grammar: The case of English relative clause. *Journal of Pan-Pacific Association of Applied Linguistics, 17*(1), 79–93.

Takashima, H. (2000). Another look at the order of difficulty of relative clauses from corpus linguistics—Statistical procedures, analysis and results. *International Review of Applied Linguistics in Language Teaching, 38*(3), 313–330. https://doi.org/10.1515/iral.2000.38.3-4.313

Tan, K. H., & Perudin, A. (2019). The “Geopolitical” factor in the Syrian Civil War: A corpus-based thematic analysis. *SAGE Open, 9*(2), 1–15. https://doi.org/10.1177/2158244019856729

Timmis, I. (2015). *Corpus linguistics for ELT*. Routledge.

Tse, P., & Hyland, K. (2010). Claiming a territory: Relative clauses in journal descriptions. *Journal of Pragmatics, 42*(7), 1880–1889. https://doi.org/10.1016/j.pragma.2009.12.025

Wiechmann, D. (2015). *Understanding relative clauses: A usage-based view on the processing of complex constructions*. De Gruyter Mouton.

Wild, K., Kilgarriff, A., & Tugwell, D. (2013). The Oxford children’s corpus: Using a children’s corpus in lexicography. *International Journal of Lexicography, 26*(2), 190–218. https://doi.org/10.1093/ijl/ecv017

Xu, X., & Xiao, R. Z. (2015). Recent changes in relative clauses in spoken British English. *English Studies, 96*(7), 1–21. https://doi.org/10.1080/0013838X.2015.1051874