On structural complexity of written English: Relating more through relative clauses

Eduard Drančák

Institute of British and American Studies, Faculty of Arts, University of Prešov, Slovakia
eduard.drancak@smail.unipo.sk

Key words: subordination, subordinate clauses, relative clauses, structural complexity, written language, spoken language

Introduction

Linguistic characteristics of a language product may vary depending on a variety of factors, e.g. purpose, context, recipients of the language product, etc. These factors, determining the way language product is composed, may cause variation in terms of lexis, syntax, structural complexity, or overall organization and style. The present study is focused on syntactic variation in language products produced in different media of production, namely that of spoken and written. It investigates spoken and written discourse of similar formality level for the incidence of relative clauses. The incidence of a particular type of relative clause, namely non-restrictive relative clause, is then interpreted as one of the indices of structural complexity of the discourse.

The aim of the paper is to demonstrate that the incidence of a particular type of subordinate clause might also indicate the level of structural complexity of the text. Studies focused on structural complexity of language commonly consider indices of complexity such as sentence length, complexity of sentences or T-Units, etc. Considering the characteristics of relative clauses, the present paper points out that structural complexity of a language product might be indicated also by the frequency of employment of a particular type of subordinate clause.

The study was carried out on a small-size corpus, i.e. a collection of authentic samples of spoken and written language, represented by *conversation-interviews* and *fiction* respectively. The samples were analyzed for the incidence of relative clauses and the data are presented in the form of tables and figures to better demonstrate the differences in the employment of relative clauses between the registers representing spoken and written language. The paper proceeds from delimiting the key terms through describing the characteristics of the corpus to the presentation and analysis of the obtained data.

1 Relative clauses in spoken and written English

Language products have been investigated for their linguistic characteristics for decades. Linguists attempt to determine how people use language in different contexts and for various purposes. Linguistic studies thus investigate the impact of various variables on linguistic properties of language products. The impact of medium of production on structural complexity of the language product has been one of the central issues investigated by linguists since 20th century. The circumstances that accompany spoken and written language production differ, and this may affect the characteristics of the language products, including their complexity. Multiple studies have led some linguists to conclude that language produced in writing is more complex, while others to conclude the opposite (see Beaman, 1984; Chafe – Tannen, 1987; Biber, 1988; Halliday, 1989; Biber et al., 1999).

Studies of structural complexity generally focus on traditional measures of complexity such as sentence length, complexity of sentences, the ratio of independent and dependent

clauses, etc. It might, however, be useful to consider the incidence of a particular type of dependent clause as an indicator of structural complexity of the language product as well. The results of the study point out that the characteristics of relative clauses in particular allow to draw conclusions regarding the structural complexity of the language product, based on their incidence. The present paper demonstrates how the incidence of a particular type of relative clause might reflect differences in structural complexity between language produced in speech and in writing.

Biber et al. (1999, p. 195) aptly define relative clause as well as the difference between restrictive and non-restrictive type by stating:

"A relative clause is a post-modifier in a noun phrase, introduced by a *wh*-word [or *that*], which has a grammatical role in the relative clause in addition to its linking function. Relative clauses may be either restrictive or non-restrictive based on whether the provided information is required for identification of the noun phrase or not."

Thus, while restrictive relative clause (1) provides information that is essential for the meaning of the sentence, non-restrictive relative clause (2) carries additional information about the post-modified element and is inserted into the sentence as an extra part. See examples:

(1) The man *who is standing in front of the store* is my uncle.
(2) Peter, *who is good at sports*, is my best friend.

Restrictive and non-restrictive relative clauses were, in the present study, distinguished in accordance with the study reported in *LGSWE*, which used punctuation-separation as a working definition of non-restrictive postmodifiers (Biber et al., 1999, p. 602). In addition to punctuation, semantic criterion was considered as well, i.e. whether the relative clause is necessary for the identification of the post-modified noun phrase in terms of the meaning of the sentence.

2 Characteristics of the corpus

The employment of relative clauses in spoken and written English was determined through the investigation of two registers, namely *conversation* (examined via interviews-IW) and *fiction* (FC). Both registers contain language of similar level of formality and might be considered less formal registers. The two registers were selected to represent spoken and written language because of their similar formality level. Potential differences in the incidence of relative clauses should reflect the impact of the medium of production rather than the impact of different level of formality. Such selection of registers thus eliminated the potential influence of unaccounted variables on the results. The characteristics of the corpus, therefore, increase the objectivity of the results of the study.

Biber et al. (1999, p. 4) define conversation and fiction as different varieties of English and state that "each of these varieties is termed a register, and each extended sample of language from a register constitutes a text." According to Rafajlovičová (2013, p. 38), "any set of sentences which appear in a sequence and cohere in this way is called a text." The present paper hereby adopts the term "text" to refer to language samples even in the case of spoken language. The results of the present study were obtained through the analysis of empirical data. The study was based on the analysis of "a collection of authentic spoken and written texts, organized by register, which is called a corpus" (Biber et al., 1999, p. 4). Performing the analysis on samples of the actual use of language should ensure valuable results that objectively reveal how language users incorporate relative clauses in spoken and written language production.

The analyzed texts contained language performances of a variety of people of different gender, age, race, social background, nationality, occupation, etc. Each text included in the corpus was of different length, but all the texts in each register combined were of approximately the same length (30,000 words). The whole corpus thus consisted of approximately 60,000 words, equally divided between the register representing the spoken language (conversation) and the written language (fiction). For the summary of the characteristics of the corpus including the selected registers, the number of texts within each register as well as the number of words, see Table 1 below.

Register	Number of texts	Number of words
Conversation	12	30 660
Fiction	3	30 665
Total	**15**	**61 325**

Table 1 Characteristics of the corpus

The spoken language was analyzed for the incidence of relative clauses via the analysis of conversation-interviews. The character of interviews, which are conducted in speech, poses limitations for such an analysis. In order to analyze language samples for the incidence of clauses, the language samples need to be divided into sentences. The absence of sentence signaling marks, when the language is produced in speech, complicates the analysis of relative clauses. The present study overcame this limitation by performing the analysis on written transcripts of the orally conducted interviews. The transcripts of the interviews were obtained from available websites such as www.time.com of *the Time magazine*, *Oprah Winfrey interviews* (www.oprah.com), *Larry King interviews* (www.cnn.com) and from several others (www.foxnews.com, www.asapsports.com, etc.).

Regarding the written language, it was analyzed through samples of contemporary fiction. This sub-corpus consisted of several chapters taken from three books of various genres, written by native English-speaking writers. The books used for the analysis include Mary Stewart's mystery novel *Rose Cottage* (1997), Stephen King's suspense novel *Under the Dome* (2009) and Diane Chamberlain's romance *Reflection* (1997). The books selected for the analysis were written by both British and American writers of both genders. The choice of various gender and nationality writers should secure more objective results by eliminating the potential impact of gender or language variety idiosyncrasies.

3 The distribution of relative clauses in the corpus

Relative clauses were analyzed with regard to their function, i.e. the type of modification. Based on this criterion, they can be divided into *restrictive* and *non-restrictive* relative clauses. Relative clauses can be distributed into these "two groups based on whether they 'identify' the head noun phrase (restrictive) or just give 'additional information' about the head (non-restrictive)" (Rafajlovičová, 2013, p. 125). The investigation of the registers regarding the incidence of restrictive and non-restrictive relative clauses might reveal the complexity of the registers and, hence, the complexity of language produced in speech and writing. Table 2 and Figure 1 below display the incidence of restrictive and non-restrictive relative clauses in the whole corpus as well as within individual registers and express these figures in percentage, too.

Register	Restrictive	Non-restrictive	Total	%
Conversation	291	48	339	51.4%
Fiction	234	86	320	48.6%
Total	525	134	659	100%
%	79.7%	20.3%	100%	

Table 2 The distribution of restrictive vs. non-restrictive relative clauses

There were 659 instances of relative clauses in the whole corpus. Table 2 shows that the distribution of relative clauses among the registers is balanced with only a slight difference between the two registers in the total number of relative clauses. The analyzed conversation-interviews contain 339 relative clauses, which is 51.4% of all relative clauses in the corpus. The fiction texts contain 320 relative clauses, which accounts for 48.6% of all relative clauses. The figures show that the incidence of relative clauses is almost identical in both registers. There is only slight prevalence of relative clauses in interviews, with 19 more instances. The balanced distribution of relative clauses among the registers shows that this type of subordinate clause is not restricted in use in certain medium of language production. Language users employ relative clauses in speaking and writing with similar frequency.

Table 2, however, reveals considerable differences between the incidence of restrictive and non-restrictive relative clauses on corpus as well as register level. On the corpus level, the data show a vast prevalence of restrictive as opposed to non-restrictive relative clauses. The former were identified 525 times, which accounts for 79.7% of all relative clauses in the corpus, while the latter occurred only 134 times and form 20.3% of the identified relative clauses in the corpus. The data thus reveal that there are almost four times as many restrictive relative clauses as there are non-restrictive relative clauses (see Figure 1 below for better illustration).
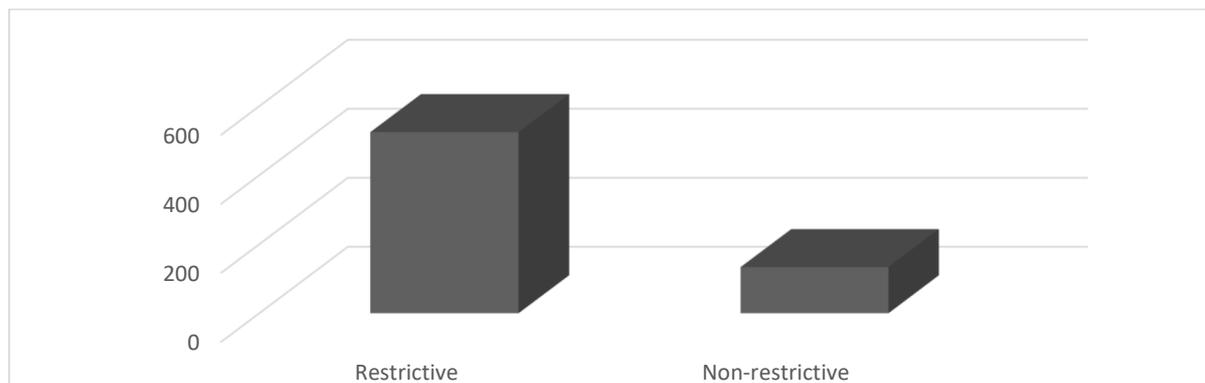

Figure 1 The distribution of restrictive vs. non-restrictive relative clauses

The considerable prevalence of restrictive relative clauses appears to be a logical consequence of their function. As the aforementioned definition states, they are used for the identification of the head of the noun phrase, which means they are necessary for the meaning of the sentence. Non-restrictive relative clauses, on the other hand, provide information that is only additional, i.e. unnecessary. The data reveal that it is more common to specify that which requires specification as opposed to that which need not be specified. The data thus support the logical option that the type of relative clause which is essential is employed more frequently than the type which is non-essential.

As for the distribution across the registers, a number of contrasts can be drawn. First, both registers contain more restrictive than non-restrictive relative clauses. However, the

differences vary between the registers. The analyzed conversation-interviews contain 291 restrictive [3.1] and only 48 non-restrictive [3.2] relative clauses. These figures reveal a considerable difference in the use of individual types of relative clauses in the conversation register. In this register, there are almost six times as many restrictive relative clauses as non-restrictive relative clauses. On the other hand, the fiction register does not exhibit such a vast difference between the frequency of employment of restrictive and non-restrictive relative clauses. This is mostly due to considerably more common incidence of non-restrictive relative clauses in the fiction texts. The analyzed samples of fiction contain 234 restrictive [3.3] and 86 non-restrictive [3.4] relative clauses. In this register, there are approximately only three times as many restrictive relative clauses as non-restrictive relative clauses. If we compare the obtained data, the conversation-interviews contain more restrictive clauses than fiction, while it is vice versa with non-restrictive clauses. However, the difference in the incidence of non-restrictive relative clauses is much more considerable. There are almost twice as many non-restrictive relative clauses in fiction as in the conversation-interviews.

[3.1] The hit is a song *that most people feel familiar with*. [R-IW]
[3.2] I love to do my "real" work, *which is developing new games and content for Funbrain*. [R-IW]
[3.3] Jack dropped the bowl *he´d been holding*. [R-FC]
[3.4] … and after a few more songs, finally the CD, *which was called 'Sound of Silver'*, ended. [R-FC]

As for the individual types of relative clauses, non-restrictive relative clauses also include a specific type of post-modifying clause "in which the antecedent is not a noun, but a whole clause" (Rafajlovičová, 2013, p. 126). These clauses are referred to as *sentential relative clauses*. In the investigated corpus, there were 14 non-restrictive relative clauses that were classified as sentential relative clauses [3.5], [3.6]. The interviews contain 11 instances, while the fiction texts only 3 instances of sentential clauses. These figures demonstrate the rareness of the occasions when language users specify the content of the whole preceding clause.

[3.5] I love to play volleyball, *which I do twice a week*. [R-IW]
[3.6] She had tracked mud onto the floor, *which was very unlike her*. [R-FC]

The incidence of non-restrictive relative clauses might also be interpreted as an indicator of the effect the medium of production has on the complexity of the language product. If we consider the character of non-restrictive relative clauses, the frequency of their use might reveal the complexity of the discourse. As mentioned above, restrictive clauses provide information that is necessary for identifying the head of the noun phrase while non-restrictive clauses provide only additional information that need not be mentioned. In the present study, the register representing written language is characteristic for considerably more frequent incidence of non-restrictive relative clauses. This might be explained by the characteristics of written medium of language production. The detachment between the production and delivery of written language product offers time and space for planning as well as revising. These characteristics present better opportunity to include more information within a single sentence by means of additional subclauses. This is reflected in more frequent employment of non-restrictive relative clauses, which carry additional information, in the written register. Frequent employment of non-restrictive relative clauses means frequent inclusion of additional information, which makes the sentences longer, the information more densely packed and it is realized via subordinate clauses. All these three factors increase the

structural complexity of the language product. More frequent incidence of non-restrictive relative clauses in fiction might thus be considered as the evidence of greater structural complexity of written language.

Conclusion

The present paper describes the use of a particular type of subordinate clause, namely relative clause, in spoken and written English. The quantitative data presented in the paper were obtained through the analysis of authentic samples of spoken and written English. The samples were analyzed for the incidence of relative clauses to determine whether the medium of production affects the use of this type of subordinate clause. Furthermore, the identified relative clauses were classified as restrictive or non-restrictive, according to their function. Besides describing the use of relative clauses, the aim of the paper is to draw conclusions regarding the structural complexity of the investigated discourses, building on the characteristics of a particular functional type of relative clause, namely non-restrictive relative clause.

The analysis has revealed 659 instances of relative clauses in the corpus. There is only insignificant difference in their incidence between spoken and written register, with 19 more instances in the spoken register. The analysis has thus showed that the medium of production does not have considerable impact on the use of relative clauses in general. The data, however, reveal considerable differences between the incidence of restrictive and non-restrictive relative clauses on corpus as well as register level. On the corpus level, the data show a strong dominance of restrictive as opposed to non-restrictive relative clauses with the former representing almost 80% and the latter only 20% of all relative clauses. This result, however, might be considered logical considering the character of restrictive clauses, which are necessary for the meaning, as opposed to non-restrictive clauses, which are merely additional.

Perhaps the most intriguing finding of the analysis stems from the comparison of the two registers in terms of the incidence of non-restrictive relative clauses. While the incidence of restrictive clauses is relatively balanced between the registers, the difference in the incidence of non-restrictive relative clauses is much more considerable. According to the data, there are almost twice as many non-restrictive relative clauses in fiction as in the conversation register. This result, however, does not only describe the use of individual types of relative clauses in spoken and written English. If we consider the character of non-restrictive clauses, we can interpret their incidence as an index of the structural complexity of the analyzed discourses. With non-restrictive clauses being additional parts of the sentence, their use increases sentence length, the density of information packaging as well as the amount of subordination in the language product. Frequent use of this type of relative clause, therefore, increases the structural complexity of the whole language product. Hence, considerably more frequent use of non-restrictive relative clauses in fiction points to greater structural complexity of the investigated written discourse. Producing language in the mentioned written medium might, therefore, increase the complexity of the final product as evidenced by more frequent inclusion of additional subclauses in this register of written English. The present paper thus indicates that the incidence of a particular type of subordinate clause might also indicate the level of structural complexity of the discourse.

Bibliography:

AXELROD, D.: Full Transcript: David Axelrod Interviews President Barack Obama for the Axe Files. [Quot. 2019-08-18.] Available at: <http://edition.cnn.com/2016/12/26/politics/axe-files-obama-transcript/>

BEAMAN, K. (1984): Coordination and subordination revisited: Syntactic complexity in spoken and written narrative discourse. In: D. Tannen (ed.): *Coherence in spoken and written discourse.* New Jersey: ABLEX, p. 45–80.

BIBER, D. (1988): *Variation across speech and writing.* Cambridge: Cambridge University Press.

1.   BIBER, D. – JOHANSSON, S. – LEECH, G. – CONRAD, S. – FINEGAN, E. (1999): *Longman Grammar of Spoken and Written English.* Harlow: Pearson Education Limited.

CHAFE, W. – TANNEN, D. (1987): The Relation Between Written and Spoken Language. In: *Annual Review of Anthropology*, 16/1, p. 383–407.

2.   CHAMBERLAIN, D. (1996): *Reflection.* London: Pan Macmillan.

3.   COHEN, R.: Sidney Crosby on Slapshots, Fights, and the Best Hockey Movies. [Quot. 2019-08-25.] Available at: <http://www.vanityfair.com/news/2010/12/sidney-crosby-on-slapshots-fights-and-the-best-hockey-movies>

HALLIDAY, M. A. K. (1989): *Spoken and Written language.* Oxford: Oxford University Press.

4.   KING, L.: Interview with Jon Bon Jovi. [Quot. 2019-08-23.]
Available at: <http://transcripts.cnn.com/TRANSCRIPTS/1012/09/lkl.01.html>

5.   KING, S. (2009): *Under the Dome.* London: Hodder and Stoughton.

6.   LUSCOMBE, B.: Kelly Clarkson: "Not a Feminist". [Quot. 2019-8-22.] Available at: <http://entertainment.time.com/2013/10/30/kelly-clarkson-not-a-feminist/?iid=ent-main-lead>

MARY, W. – DARINA, S.: Singer Alicia Keys. [Quot. 2019-08-22.] Available at: <http://www.teenink.com/nonfiction/celebrity_interviews/article/5341/Singer-Alicia-Keys/?page=5>

7.   O'LEARY, D.: Hugh Jackman. [Quot. 2019-08-21.] Available at: <https://www.bafta.org/sites/default/files/uploads/transcripts/red-carpet-hugh-jackman-1693.pdf>

RAFAJLOVIČOVÁ, R. (2013): *Subordination in Different Text Types.* Prešov: Filozofická Fakulta Prešovskej Univerzity v Prešove.

8.   STEWART, M. (1997): *Rose Cottage.* London: Hodder and Stoughton.

9.   UNKNOWN: 10 Questions for Emma Watson. [Quot. 2019-08-25.]
Available at: <http://content.time.com/time/magazine/article/0,9171,2032143,00.html>

UNKNOWN: An Interview with Jeff Kinney. [Quot. 2019-08-22.]
Available at: <https://www.teachervision.com/authors/interview-jeff-kinney>

10.   UNKNOWN: 10 Questions for Jim Parsons. [Quot. 2019-08-30.]
Available at: <http://content.time.com/time/magazine/article/0,9171,2048322,00.html>

11.   WINFREY, O.: Oprah Talks to Bono. [Quot. 2019-08-19.] Available at: <http://www.oprah.com/omagazine/Oprahs-Interview-with-Bono-U2-and-AIDS-Activism>

WINFREY, O.: Oprah Talks to Janet Fitch. [Quot. 2019-08-20.] Available at: <http://www.oprah.com/omagazine/Oprah-Interviews-White-Oleander-Author-Janet-Fitch>

WINFREY, O.: Oprah talks to Denzel Washington. [Quot. 2019-08-21.]
Available at: <http://www.oprah.com/omagazine/Oprahs-Interview-with-Denzel-Washington>

Summary

On structural complexity of written English: Relating more through relative clauses

The present research paper describes the use of relative clauses in spoken and written English. The focus is placed on the frequency of incidence of relative clauses and their particular functional types. Quantitative analysis was performed to determine the frequency of incidence of relative clauses in the selected registers of spoken and written English. Comparative analysis was applied to determine differences between spoken and written English in the use of relative clauses. The present study also applied the principles of corpus linguistics by performing the analysis on a small-size corpus of authentic samples of language. The aim of the paper is to demonstrate the correlation between the incidence of relative clauses and the structural complexity of the discourse. The results of the study point to greater structural complexity of the written discourse by considering the characteristics of non-restrictive relative clauses. The paper provides an alternative view upon the structural complexity of discourse by determining the complexity using a non-conventional index, namely the incidence of a particular type of subordinate clause.