Restrictive Relative Clauses in EFL Learners' Writing

Fangzhou Zhu

Department of Linguistics and English Language (LAEL), Lancaster University, UK.

This research aims to examine how Chinese EFL learners use restrictive relative clauses (RRCs) in English essay writing, based on corpus analysis. The study designs the query syntax to facilitate the investigation of learners' preferences in using different types of RRCs and to analyse the errors that Chinese EFL learners frequently make when they try to apply the knowledge of English RRC into their writing.

One of the most difficult English syntactic structures for EFL learners to learn is the relative clause (RC) (Celce-Murcia & Larsen-Freeman, 1999), thus exploring its use and typical errors has benefits for improving learners' language proficiency. English restrictive relative clauses (RRCs) are defined as realising the function of restricting or defining the meaning of a noun or noun phrase or a pronoun and providing necessary information in the sentence (Cowan, 2008). The *wh*-words, as well as *that*, are major relative markers used to introduce RRCs. There are two main kinds of relative markers: relative pronoun and relative adverb.

Three areas of literature informed the current study: second language RC research, error analysis, and corpus investigation on relevant topics. First, the theory Noun Phrase Accessibility Hierarchy, which is one of the hypotheses of this study, predicts the degree of difficulty of relativisation as a function of the grammatical role of the head noun phrase modified by the RC in a wide range of languages, i.e., (from easiest to hardest order), subject > direct object > indirect object > oblique > genitive (Keenan & Comrie, 1977). Second, error analysis shows that the misuse of relative pronouns, the misuse of resumptive noun or

pronoun, the agreement in number and tense, the absence of prepositions are the main errors that learners make (Chang, 2004; Yip & Matthews, 1991). Corpus linguistics plays an important role in providing a significant quantity of genuine data produced by second language learners as direct and specific evidence. Lu and Chen's study (2010) investigated Chinese and English restrictive clauses in the Bible parallel corpora and found English follows the noun phrase accessibility hierarchy while Chinese does not. Chu (2015), based on a local English language proficiency test, extracted five major problems from a learner corpus: general structure errors, omission of relative subjects, resumptive pronouns, missing prepositions and confusion about relative pronouns. In terms of relative adverbs investigation, Crompton (2005) reported the overuse of the relative adverb *where* in the Brunei Learner Corpus and Phoocharoensil (2014) investigated the English RC in the Thai Learner Corpus and found that learners often misuse *where* as a relative pronoun.

However, these studies or hypotheses focused on clauses which only used certain relative pronouns or adverbs. The fact that learners, especially more advanced learners, may use all types of English RRCs with many possible errors, should raise our awareness about the value of conducting more research on the use preference and error analysis.

The current study aims to investigate mainland Chinese EFL learners' use of RRCs in English essay writing. In China, the RC is one of the major grammatical constructions in English language teaching, starting from high school education (Ministry of Education of the People's Republic of China, 2017). Due to its importance and difficulty, it is assessed in a broad scope of tests, in which RCs are used as a tool to test whether a test taker can apply advanced sentence constructions in their writing (Yi, 2017). One of the examples can be seen in Jiangsu Province, China. The knowledge of RRCs is systematically taught in high school, where RRCs with relative pronouns are the first ones to teach, followed by relative adverbs (Kent, 2010). This knowledge is frequently tested in the College Entrance Exam (Jiangsu Education and Examination Institute, 2019), and later in the university; similar assessments are still emphasised in the College English Test (National College English Testing Committee, 2016)

and other international language proficiency tests (Lynch & Anderson, 2013). Therefore, it is of great value to explore the learners' preferences and typical errors when they have already learnt the English RRC and try to produce it in English writing.

Research Questions

The research shall answer the following two questions:

1. Do Chinese EFL learners have preferences for using certain types of RRCs in English essay writing? If so, which types are most/least produced?
2. What are the errors that Chinese learners often make in producing RRCs in English essay writing?

Methods

The current study collects data from Longman Learner Corpus (LLC). As this corpus is annotated by Oxford Simplified Tag, it is possible to design the appropriate query syntax to search RRCs in the database.

Previous research often used relative markers as node words to explore their left and right sides to decide whether the clause is relative one or not. However, this approach has limitations in terms of time and manual investment. As this study only focuses on the sentences already showing RC structure due to the mentioned context, query syntax can be used as an efficient and suitable tool. Query syntax refers to a specific programming language that requests and retrieves data from corpora by sending structured queries. Following the Oxford Simplified Tag, this study tried to "simulate" the structure of RRC in English based on grammatical rules.

A small-scale pilot study was conducted to test the query syntax design. The result reported that using query syntax for searching for English RRCs shows high accuracy for targeted concordance lines, but object RCs and adverbial RCs were not ideally searched in the corpus due to the similar structure of objective clauses, attributive clauses, and adverbial clauses. The following examples demonstrate the final query syntax design for the current study.

1. *General query of RRC*

    (_{N}|_{PRON}((_{ADV})*_{A})*)(_{PREP})?(who|whom|which|th

at|whose|where|when|why) ((_{ART})? ((_{ADV})* _{A})* (_{N})+|(_{PRON}) ((_{ADV})* _{A})*)* (_{ADV})* _{V}

2. *As the subject of clause*

(_{N}|_{PRON} ((_{ADV})* _{A})*) (who|that|which) (_{ADV})* _{V}

3. *As the object of clause*

(_{N}|_{PRON} ((_{ADV})* _{A})*) (who|that|whom|which) ((_{ART})? ((_{ADV})* _{A})* (_{N})+|(_{PRON})+ ((_{ADV})* _{A})*) (_{ADV})* _{V}

4. *As the object of preposition*

((_{N})|(_{PRON}) ((_{ADV})* _{A})*) _{PREP} (whom|which)

5. *Possessive*

(_{N}|_{PRON} ((_{ADV})* _{A})*) whose ((_{ADV})* _{A})* _{N}

6. *Time, place, and reason*

_{N} (where|when|why) ((_{PRON}) + ((_{ADV})* _{A})*|(_{ART})* ((_{ADV})* _{A})* (_{N})+) (_{ADV})*_{V}

All queries were used in the restricted query interface with three filters, namely learner nationality, learner mother tongue, and writing genre, and these were Chinese mainland learners, Chinese native speakers, and the essay writing tasks (free essay, set essay and project essay), respectively. They were thinned to 100 lines by unproduceable randomisation.

Results and Discussion

The general query retrieved 2,370 concordance lines from 1,088 different texts. After non-reproducible thinning, the first 100 lines of RRCs were filtered and labelled by the detailed categories. As shown in Figure 1, among 100 concordance lines, Chinese EFL learners show the tendency to use more subjective and objective English RRCs than any other types. This phenomenon is generally consistent with the previous hypothesis of the accessibility hierarchy. Additionally, the study found that Chinese learners use adverbial RRCs less frequently than other types of RRC, which can be considered as a new piece of evidence in RC studies.
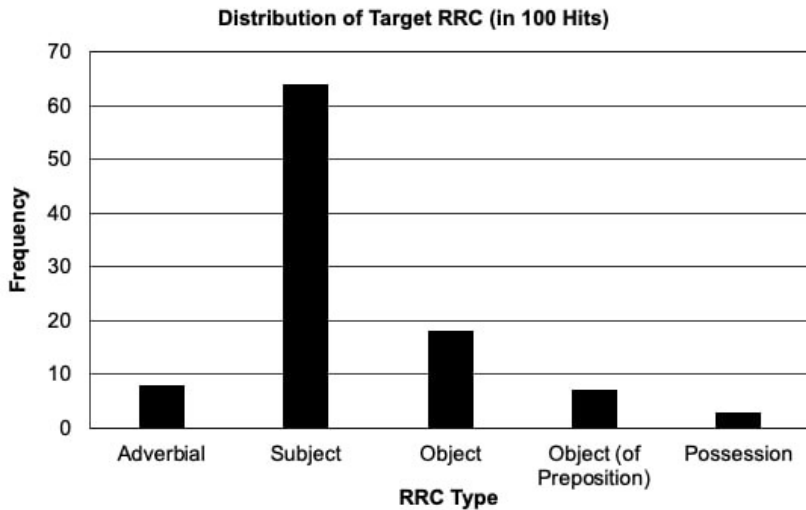
Distribution of Target RRC (in 100 Hits)

Figure 1. Distribution of RRCs in different clause types per 100 hits.

Learner avoidance and the transfer of training may also contribute to the phenomenon. As mentioned, the knowledge of RRCs is taught by a specific order. Selinker (1992) believes if the teacher or the textbook fails to treat the language knowledge appropriately, such as emphasis on certain structures at the expense of others, learners may develop, in a limited manner, the knowledge of that grammar point in a second language. Ellis (1994), from another perspective, pointed out that learners avoid using linguistic structures which they consider difficult due to differences between their native language and the target language. Accordingly, students tend to use those RC structures they are familiar with. Such avoidance and transfer of teacher instruction, with the linguistic fact that the RC as post-modifier in English is very different from those in Asian languages (Phoocharoensil & Simargool, 2010), may explain the high proportion of subjective RRCs in corpus data.

Further error analysis revealed that the poor awareness of head noun function in clauses is the main reason for generating relevant errors. Based on the reading of 100 lines in each RRC category, this study summarises the following typical errors among target learners:

1. *Subject-verb mismatch*

   Example: *She wears <u>the suitable clothes</u> which <u>makes</u> her more beautiful.*

2. *Wrong relative markers*

   Example: *You can get some special education on <u>TV</u> <u>which</u> channels show some programme of education.*

3. *Redundant pronoun/adverb in the clause*

   Example: *People say that stories like that are just in <u>a magazine</u> that you continue to buy <u>it</u>.*

   I would pay a wonderful visit to <u>the Lin Yin Temple</u> where I could enjoy the exquisite ancient architecture <u>there</u>.

Based on the retrieved corpus data, the common reason for the errors above is that the learners fail to realise the grammar function of the head noun in the clause. Thus, the implication for pedagogy here is that, based on the existing instruction, the teacher should raise students' awareness to revise the English RRCs they have produced by emphasising two aspects. First, the students should realise that the head noun must play a missing grammatical role in the RRC. Second, how the head noun is restored into the RRC (to make it complete) determines the choice of relative markers.

One contribution of the study is the development of data-driven learning (DDL), initially proposed by Johns (1994), which argued that the use of computer-generated concordance lines can help students explore regularities of patterning in a target language. The authentic English corpora, such as the British National Corpus, can provide rich examples of language use for learners. However, learners may find it difficult to compare what they produce with the authentic examples. Still, learner corpora, where the common errors can be identified, provide examples of inappropriate use in comparison with the native English corpora, and thus show learners what is wrong and what is right. Facilitated with the contextualised query syntax searching, the DDL material can further reduce the negative effect of a large corpus overwhelming the learner with too many examples (Chen & Flowerdew, 2018).

Conclusion

The purpose of this study was to investigate Chinese learners' preferences and errors of RRCs in English essay writing. It revealed that Chinese learners use more subjective and objective RRCs than any other types. Error analysis found that the misuse of relative markers, the subject-verb mismatch, and the redundant pronoun or adverb are the main errors. Most of these happened due to the lack of analysis of the grammar function of head noun in the clauses, hence the solution for teachers is to ask students to revise the role the head noun plays in the relative clause. This research also showed that query syntax searching can enhance the development of DDL material in class.

References

Celce-Murcia, M., & Larsen-Freeman, D. (1999). *The grammar book: An ESL/ EFL teacher's course* (2nd ed.). Heinle & Heinle.

Chang, Y. (2004, May). Second language relative clause acquisition: An examination of cross-linguistic influences. Paper presented at the *Annual Meeting of the American Association for Applied Linguistics*, Portland, Oregon, USA.

Chen, M., & Flowerdew, J. (2018). A critical review of research and practice in data-driven learning (DDL) in the academic writing classroom. *International Journal of Corpus Linguistics*, *23*(3), 335–369. https://doi. org/10.1075/ijcl.16130.che

Chu, H. (2015, April). *A learner corpus-based study on the GEPT test-takers' errors of English relativization.* Paper presented at the LTTC International Conference 2015, Taipei, Taiwan.

Cowan, R. (2008). *The teacher's grammar of English: A course book and reference guide.* Cambridge University Press.

Crompton, P. (2005). "Where", "in which", and "in that": A corpus-based approach to error analysis. *RELC Journal*, *36*(2), 157–176. https://doi. org/10.1177/0033688205055570

Ellis, R. (1994). *The study of second language acquisition.* Oxford University Press.

Jiangsu Education and Examination Institute. (2019). *Jiangsu College Entrance Exam Syllabus: English* (2019 ed.). Jiangsu Education and Examination Institute.

Johns, T. (1994). From printout to handout: Grammar and vocabulary teaching in the context of data-driven learning. In T. Odlin (Ed.), *Perspectives on pedagogical grammar* (pp. 293–313). Cambridge University Press.

Keenan, E. L., & Comrie, B. (1977). Noun phrase accessibility and universal grammar. *Linguistic Inquiry*, *8*(1), 63–99.

Kent, J.C. (2010). *Advanced with English* (3rd ed.). Oxford University Press (China).

Lu, H. C., & Chen, Y. H. (2010). A corpus-based study of restrictive relative clauses. In *Using corpora in contrastive and translation studies* (pp. 473–487). Cambridge Scholars Publishing.

Lynch, T., & Anderson, K. (2013). *Grammar for academic writing.* English Language Teaching Centre, University of Edinburgh.

Ministry of Education of the People's Republic of China. (2017). *New English curriculum for Chinese primary schools and junior/senior middle schools (experimental version).* People's Education Press.

National College English Testing Committee. (2016). *Syllabus for college English test* (2016 ed.). Shanghai Language Education Press.

Phoocharoensil, S. (2014). Errors on the relative marker WHERE: Evidence from an EFL learner corpus. *3L: Language, Linguistics, Literature*, *20*(1), 1–20. http://dx.doi.org/10.17576/3L-2014-2001-01

Phoocharoensil, S., & Simargool, N. (2010). English relativization and learners' problems. *Journal of Pan-Pacific Association of Applied Linguistics*, *14*(1), 109–129.

Selinker, L. (1992). *Rediscovering interlanguage*. London, England: Longman.

Yi, D. (2017). Teaching relative clause in secondary school English classroom. *International Journal of Liberal Arts and Social Science*, *5*(5), 1–6.

Yip, V., & Matthews, S. (1991). Relative complexity: Beyond avoidance. *CUHK Papers in Linguistics*, *3*, 112–124.

Author bio

Fangzhou Zhu is a PhD student in applied linguistics at the Department of Linguistics and English Language, Lancaster University, UK. His research interest is the effective application of language corpora in EFL writing instruction. His current project is the investigation of data-driven learning in EFL writing error correction in China. <f.zhu@lancaster.ac.uk>