English relative clauses in science and engineering journal papers: A comparative corpus-based study for pedagogical purposes

Dong Wan Cho [a,*], Kyusong Lee [b]

[a] Division of Humanities and Social Sciences, POSTECH, 77 Cheongam-Ro. Nam-Gu., Pohang, Gyeongbuk, South Korea
[b] Intelligent Software Lab, Department of Computer and Engineering, POSTECH, 77 Cheongam-Ro. Nam-Gu., Pohang, Gyeongbuk, South Korea

HIGHLIGHTS

- We examine how English relative clauses are used in science and engineering journal papers.
- Results show the frequent use of relative clauses, and high frequency of non-restrictive clauses.
- Extremely high use of 'that' over 'which' for restrictive relative clauses is also found.
- Some tips for teaching English relative clauses are provided for pedagogical purposes.

ARTICLE INFO

ABSTRACT

This corpus-based study presents how English relative clauses are used in science and engineering journal papers. Relative clauses ensure semantic clarity and textual variety but they cause difficulty to non-native speakers of English due to their diverse uses and functions. With pedagogical purposes in mind, this research investigates how frequently and in what context relative clauses are employed in three representative science and engineering journals, namely *CELL, Journal of American Chemical Society,* and *IEEE Journal of Solid-State Circuits.* In addition, relative clauses used in papers of *English for Specific Purposes* are investigated and compared with those in the science and engineering journal papers, to reveal the similarities and differences between them. Some unique features of relative clauses used in science and engineering journal papers are identified, such as the frequent use of relative clauses, the high frequency of non-restrictive relative clauses in the papers of *Journal of American Chemical Society* and *Journal of Solid-State Circuits*, the high proportion of 'prepositions + which,' and the extremely high use of 'that' over 'which' for restrictive relative clauses. Pedagogical suggestions are provided to help science and engineering paper authors and ESP/EAP practitioners use and teach relative clauses in an efficient way.

1. Introduction

Writing and publishing academic journal papers has become a significant mission of graduate students and professors in Korea. The situation has been sparked by social pressure and the mass media's evaluation of the universities, leading to harsh competition among them (Cho, 2012). Universities in Korea are now very concerned with this evaluation, one index of which is the research capacity of a university, mainly assessed by the number of papers published in international journals. When it comes to science and

engineering fields, the effort made to publish papers in international journals with a high reputation is doubled as research groups in other parts of the world are likely to conduct similar research. A need for the publication of journal papers is accelerated since some science universities in Korea demand graduate students to publish their papers in international journals with a high impact factor as a graduation requirement. Writing journal papers for publication, however, burdens Korean graduate students and faculty members working in the fields of science and engineering as most prestigious journals in the fields are published in English. While working to publish papers, graduate students and faculty members of the country felt that they were put at a disadvantage, compared to those working in English-speaking countries, where more resources are available to assist journal paper authors (Cho, 2009a). The disadvantage that non-native speakers of English have felt

when writing journal papers was found to be universal across countries. Cases researched in other ESL/EFL settings such as Hong Kong (J. Flowerdew, 1999a, 1999b), China (Li, 2006), Poland (Duszak and Lewkowicz, 2008), Venezuela (Salager-Meyer, 2008), Sudan (El-Malik and Nesi, 2008), and Italy (Giannoni, 2008) have revealed similar findings.

In ESL/EFL settings, many graduate students in science and engineering fields do not have adequate English proficiency and training for paper writing, which may result in the delay of paper publication and decrease the chance of papers being published. Under these circumstances, it is imminent for ESP/EAP practitioners and English teaching faculty to train graduate students with materials developed for them. Part of this task is identifying features targeted for teaching and analyzing their attributes through the investigation of published journal papers. Of the several components comprising journal paper writing, linguistic features such as sentence structure, vocabulary, and grammar need to be addressed more as they are perceived to be more difficult and problematic to non-native graduate students than meta-linguistic features such as overall organization and paragraph structure (Casanave and Hubbard, 1992; Cho, 2009a; Dong, 1998). Research has disclosed common grammatical features of English in science and engineering journals: diachronic evolution of referential behavior in medical articles (Salager-Meyer, 1999), signaling nouns in written biology corpora (J. Flowerdew, 2003), construction of stance through nouns followed by *that* in materials science (Charles, 2007a), verbs in reporting clauses used in citations of materials science (Charles, 2006), word frequency and distribution used in medical research articles (Chen and Ge, 2007), and use of participial and relative clauses in two science journals, namely *Cell* and *Physical Review Letters* (Cho, 2009b; Cho and Kim, 2009).

With pedagogical purposes in mind, this study focuses on English relative clauses used in science and engineering journal papers since they play a key role of ensuring semantic clarity between clauses and promoting syntactic maturity as well as textual variety. Relative clauses, however, are considered one of the most difficult areas of English for non-native speakers of the language to master, due to differences between English and their mother tongue, and the complex grammatical attributes such as restrictive or non-restrictive clauses, human or non-human head nouns, the position of prepositions in relative clauses, the zero-relative pronoun, etc. This study investigates relative clauses used in the papers of three representative science and engineering journals, namely *CELL*, *Journal of American Chemical Society*, and *IEEE Journal of Solid-State Circuits*. The investigation is expected to reveal features of relative clauses unique to science and engineering journals. In addition, relative clauses used in the papers of *English for Specific Purposes*, a journal in a different academic discipline, are probed to disclose differences and similarities of relative clauses used in science and engineering journals and a language research journal.

2. Literature review

2.1. Relative clauses in English

Relativization in languages is a process through which one sentence is embedded in another sentence when the two sentences share the same referential noun or noun phrase (Abdolmanafi and Rahmani, 2012). An English relative clause functioning as an adjective and combining separate clauses modifies a noun or noun phrase in the main clause, helping ensure semantic clarity between clauses. It also promotes syntactic maturity and textual variety, as claimed by Kameen (1978, 1983). He stated that the length of a T-unit, a main clause along with all the other subordinate clauses embedded within it, is one of the decisive factors to differentiate

good and poor writers. In this regard, relative clauses, which increase T-unit length, could be a significant grammatical feature to teach for a better quality of writing. Technical writing textbooks, in fact, recommend the use of complex sentences rather than the repeated use of simple sentences (Markel and Holmes, 1994; Weisman and Collins, 1998), and suggest combining related ideas using relative clauses (Lannon, 1988).

Relative clauses, however, are considered one of the most problematic and difficult areas of English (Marefat and Rahmany, 2009) to non-native speakers due to grammatical differences between their mother tongue and the English language, such as the position of relative clauses with respect to the head noun, the ways for relative clauses to be marked, and the presence of a pronominal reflex, and the complex grammatical features such as restrictive or non-restrictive clauses, human or non-human head nouns, zero-relative pronoun, and position of prepositions in a relative clause. A pioneering study of relative clauses of non-native speakers of English by Schachter (1974) clearly showed the problems that non-native speakers were likely to undergo in their writing. Japanese and Chinese students, whose native languages differed from English in terms of the position of relative clauses, avoided using relative clauses in their writing, thus producing significantly fewer relative clauses, while Persian and Arab students, whose mother tongues had the same postnominal relative clauses as English, produced about the same number of relative clauses in their writing as their native speaker counterparts. Yip and Matthews (1991) also found avoidance strategies of relative clauses of Hong Kong students with a Chinese language background. Chang (2004) was in line with the previous research stating errors made by Chinese ESL learners were caused by L1 transfer. Other language learners of English were found to adopt L1 language transfer and avoidance strategies when producing English relative clauses. The production of English relative clauses of Korean learners of English (Park, 2000), Japanese learners of English (Miura, 1989), Hong Kong learners of English (Bunton, 1979), Thai learners of English (Phoocharoensil and Simargool, 2010), and Persian learners of English (Abdolmanafi and Rahmani, 2012) was affected by their L1.

However, it must be noted that other factors, such as the overall English proficiency of learners (Chiang, 1980) and different data elicitation methods (Liu, 1998), affected the avoidance and general production of relative clauses. Baek (2012) investigated the processing of relative clauses by Korean L2 learners and witnessed that the processing behaviors of English relative clauses of Korean L2 learners were quite similar to those of native-speakers, indicating that L1 transfer did not occur in the processing of English relative clauses by Korean learners.

In a different spectrum of research on English relative clauses, investigation on difficulty order has been conducted. Keenan and Comrie (1977) proposed a noun phrase accessibility hierarchy hypothesis, in which a head noun functioning as a subject in the relative clause is most accessible or easiest to process while the object of comparison is least accessible or most difficult to process. They detailed the hierarchy as follows:

SU > DO > IO > OBL > GEN > OCOMP

where SU stands for a subject, DO for a direct object, IO for an indirect object, OBL for a major oblique of a noun phrase, GEN for a genitive and OCOMP for an object of comparison. The hypothesis of accessibility hierarchy has sparked investigation on the issue. Gass (1979a, 1979b), from a language transfer perspective, investigated the hypothesis and claimed that the production of relative clauses by non-native speakers of English was able to be predicted by the hierarchy theory, with the exception of the genitives. The theory, however, as time passed, has lost its universality as other languages

were brought into consideration (Eckman, 2007; Gass et al., 2013; Jeon and Kim, 2007).

More well-established research of relative clauses has been concerned with the position of the head noun in the main clause and its function in the relative clause. Sheldon (1974) categorized sentences with relative clauses into four types, SS, SO, OS, and OO, and found that parallel function relative sentences such as the SS and OO types were easier to understand than non-parallel function relative sentences such as the SO and OS types.

SS: The dog that jumps over the pig bumps into the lion.
SO: The lion that the horse bumps into jumps over the giraffe.
OS: The pig bumps into the horse that jumps over the giraffe.
OO: The dog stands on the horse that the giraffe jumps over.
(Sheldon, p. 275)

Other research, however, did not support Sheldon. On a purely theoretical analysis of several languages, Kuno (1974) claimed that center embedding reduced the comprehensibility of sentences as center-embedded sentences interfered with the natural language process. Thus, OS and OO types which did not include center-embedding were easier to process than SS and SO types with center-embedding relative clauses. Ioup and Kruse (1977), based on data from 87 participants of five different languages, found that OS type was easiest to process, followed by OO, SO, and SS types. Schumann (1980) also was in agreement with Kuno (1974), and Ioup and Kruse (1977). He examined relative clause production data of previous studies and concluded that OS and OO sentence types were easier to process than SS and SO types. Research followed and supported the previous studies. Stauble (1978) investigated the frequency of relative clauses in informal speech, spontaneous writing, and published writing of native speakers of English, reporting that OS structure accounted for 55%, OO 25%, SS 12%, and SO 7%. Schumann's study of ESL learners (1980) presented a quite similar frequency distribution of the four types of structure: 53% of OS structure, 35% of OO structure, 6% of SS structure, and 4% of SO structure. Wong (1991), from 170 compositions of Hong Kong secondary school students, found that the production frequency of the four types of sentences was 48% of OS structure, 26% of OO structure, 17% of SS structure, and 7% of SO structure. Abdolmanafi and Rahmani's study of Iranian students (2012) also showed that the mastery level of the four types of sentences was OS > OO > SS > SO. As to the ratio of errors of relative clauses, OS type sentences contributed to the fewest errors in the construction of relative clauses (Sadighi, 1994).

Large-scale corpus-based investigation also supported the previous research. Cho (2009b) analyzed the use and frequency of relative clauses in two representative science journals, namely *CELL* and *Physical Review Letters,* and discovered a similar frequency distribution: in *CELL* papers, OS structure accounted for 75%, OO 20%, SS 5%, and SO 0.01%, and in *Physical Review Letters*, OS construction accounted for 82%, OO 15%, SS 2%, and SO 0.04%. The four types of relative clause structures described above could be summarized in Table 1.

In terms of the use of relative pronouns, Biber, Johansson, Leech, Conrad, and Finegan's corpus-based investigation (1999) provides a rough estimation of the frequency of relativizers in different registers, such as conversation, fiction, news, and academic prose. Of these, the frequency of relative pronouns in academic prose, which is based on Figure 8.17 in their book, could serve as a guideline for the comparison with other academic corpora. The relative pronoun 'which' has the frequency of about 4950 cases per one million words, followed by 'that' with about 2400 cases, 'who' with about 1300 cases, 'whose' with about 160 cases, and 'whom' with about 130 cases. A relative clause without a relative pronoun has the

frequency of about 1050 cases. In total, about 9990 cases of relativizers are found in the one million word academic corpora. That the restrictive relative pronoun 'which' with a frequency of 3810 cases is more frequently used than 'that' illustrates one unique feature of academic prose; the relative pronoun 'which' is more formal and conservative than 'that.' In contrast, fiction and news show a more frequent use of 'that' than 'which.' As to the frequency of restrictive and non-restrictive clauses in the case of adopting the relative pronoun 'which,' restrictive clauses account for about 77% and non-restrictive clauses about 23%.

2.2. Corpus-based research for pedagogical purposes in the fields of language teaching and learning

In the fields of language teaching and learning, recent years have witnessed book publications based on corpus investigation: *New Trends in Corpora and Language Learning* (2010) by Frankenberg-Garcia, L. Flowerdew and Aston; *Corpus-based Studies in Language Use, Language Learning and Language Documentation* (2011) by Newman, Baayen and Rice; and *Corpora and Language Education* (2012) by L. Flowerdew. The books illustrate how language corpora could be employed for the teaching and learning of languages, in particular, foreign languages. Gavioli's *Exploring Corpora for ESP Learning* (2005) also addresses the same purpose. The interest in language corpora reflects the fact that they are based on massive, authentic, and real language sources. Corpus-based research published in journal papers (Cargill et al., 2012; Chang and Guo, 2011; Charles, 2007b, 2010; J. Flowerdew, 2013; Hafner and Candlin, 2007; Lee and Swales, 2006) has dealt with issues of learning and teaching grammatical features of English such as the definite article, participial phrases and reporting verbs (Lee and Swales, 2006), utilizing law corpora for the teaching of writing tasks for novice lawyers (Hafner and Candlin, 2007), helping learners increase awareness of rhetorical functions of English texts (Charles, 2007b), teaching lexicogrammatical features and discourse patterns of writing in EAP writing classes (Charles, 2010), developing genre-specific teaching materials and applying them in EFL classes (Chang and Guo, 2011), and suggesting ESP oriented and corpus-driven English teaching (Cargill et al., 2012). In contrast to using the pre-existing corpora for the teaching and learning of languages and for the development of teaching materials, some research (Gilquin et al., 2007; L. Flowerdew, 2001) presented how learner corpora could be utilized for pedagogical purposes.

Corpus-based research for language teaching and learning has greatly contributed to the revelation of linguistic and discoursal features of English, illustrating a wide potential for the field. Some research (Lee and Swales, 2006) even claimed that language corpora could replace native speakers when learning some grammatical features of the English language. As Krishnamurthy and Kosem (2007) pointed out, however, language corpora currently available do not suit the various, specific needs of EAP/ESP practitioners and researchers. Thus, a variety of corpus-based studies for pedagogical purposes need to be carried out.

2.3. Research questions

Even though much research has been conducted on English relative clauses, corpus-based investigation of relative clauses used in academic journal papers has not been actively conducted, except for Biber et al. (1999). Facing the lack of research on the issue, the current research explores relative clauses in papers of three representative science and engineering journals, namely *CELL, Journal of American Chemical Society,* and *IEEE Journal of Solid-State Circuits.* The investigation is expected to disclose some features of relative clauses unique to science and engineering journals. In

addition, relative clauses used in papers of *English for Specific Purposes* are examined to identify differences and similarities of the use of relative clauses between science and engineering journals and a language research journal. The comparison, in particular, intends to provide meaningful resources for ESP/EAP practitioners who are engaged in teaching scientific writing to science and engineering journal paper authors. Findings of the research thus will be geared towards addressing pedagogical suggestions for the teaching of English relative clauses. The research questions are detailed as follows:

1) Do science and engineering papers in various journals have a similar frequency and use of relative clauses?
2) Do science and engineering journal papers have unique features of relative clauses which are pedagogically meaningful, compared to relative clauses used in papers of an English language research journal?

3. Journal papers analyzed

Four academic journals were selected for the study: *Cell* (hereafter CELL), *Journal of the American Chemical Society* (hereafter JACS), *IEEE Journal of Solid-State Circuits* (hereafter IEEE), and *English for Specific Purposes* (hereafter ESP). *CELL* is a representative journal of life science and biology, *JACS* is a representative journal of chemistry, and *IEEE* is a representative journal of electrical and electric engineering. The selection of the science and engineering journals was advised by faculty members of a research-oriented science and engineering university in Korea. Fifty papers of each journal were chosen and analyzed: *CELL* from Vol. 156, pp. 84–96 January 16, 2014 to Vol. 156, pp. 1017–1031 February 27, 2014; *JACS* from Vol. 136, pp. 122–129, 2014 to Vol. 136, pp. 671–867, 2014; and *IEEE* from Vol. 49, No. 1, pp. 9–18 January 2014 to Vol. 49, No 3, pp. 708–717, March 2014. Fifty papers of *ESP* were from Vol. 31, pp. 3–13, 2012 to Vol. 34, pp. 48–57, 2014. All the papers chosen for the study were available online in PDF format and thus it was easy to identify relative clauses and relative pronouns by using the 'Find' function. Relative clauses with the relative pronouns 'that,' 'which,' 'who,' 'whom,' and 'whose' were sorted by this 'Find' function and then their frequency was manually counted. Relative clauses with no relative pronoun were manually investigated by looking at every single sentence. Since an error could occur when counting zero-relative pronouns, two researchers were involved in the counting to reduce errors. When counting the number of sentences and words, only those used in complete sentences were considered. Words in the title, captions of figures, graphs, references, bio-data of authors, and acknowledgements were excluded.

4. Results

Table 2 shows descriptive statistics for the papers analyzed. In terms of the total number of words, the *ESP* corpora top with 365,453 words, followed by the *CELL* corpora with 260,199 words,

the *IEEE* corpora with 249,940 words, and the *JACS* corpora with 224,142 words. In terms of the total number of sentences, the *ESP* corpora top with 12,363 sentences, followed by the *IEEE* corpora with 10,901 sentences, the *CELL* corpora with 9,712 sentences, and the *JACS* corpora with 8496 words. In terms of the average number of sentences in a paper, the *ESP* corpora top with 247 sentences, followed by the *IEEE* corpora with 218 sentences, the *CELL* corpora with 194 sentences, and the *JACS* corpora with 170 sentences. As for the average number of words in a sentence, the *ESP* corpora top with 29.6 words, followed by the *CELL* corpora with 26.8 words, the *JACS* corpora with 26.4 words, and the *IEEE* corpora with 22.9 words.

Table 3 displays the number of relative clauses identified and their frequency per 100 sentences in each corpora. The number of relative clauses is equal to that of sentences which contains a relative clause. The percentage of relative clauses per sentence is the ratio of a relative clause used in a sentence. For example, the percentage of the relative clause in the following sentence, 'The other day I got a call from a man who said he was my old school friend,' is 100%.

Among the science corpora, the *CELL* corpora top in terms of the frequency of relative clauses with 1844 cases, followed by the *IEEE* corpora with 1294 cases, and the *JACS* corpora with 1167 cases. The *ESP* corpora show the highest frequency of relative clauses with 3249 cases. As to the frequency of relative clauses per 100 sentences used in each corpora, the *CELL* corpora top with a frequency of 19.0 cases, followed by the *JACS* corpora with 13.7 cases, and the *IEEE* corpora with 11.9 cases. The *ESP* corpora present the highest frequency of 26.3 cases. It should be noted that relative clauses with zero relative pronouns also were counted for the study.

To determine the statistical difference of the frequency of relative clauses in the four journals, an ANCOVA analysis, in which the effects of a covariate (i.e. the number of sentences in each journal in our case) need to be controlled, was performed. The results showed that there are statistically significant differences between the journals, as shown in the post-hoc test results of *ESP* > *CELL* > *JACS* = *IEEE* at the level of 0.001. See Table 4.

Table 5 below summarizes the frequency distribution of restrictive and non-restrictive relative clauses in each corpora. Overall, restrictive clauses are more frequently used than non-restrictive clauses in all four corpora.

Among the science corpora, the *CELL* corpora show the highest frequency of 69.9%, followed by the *IEEE* corpora with 51.8% and the *JACS* corpora with 51.5%. On the other hand, 48.2% of the *JACS* corpora, and 48.5% of the *IEEE* corpora employ non-restrictive relative clauses, accounting for almost half of the relative clauses used in these corpora. The high proportion of non-restrictive relative clauses in the *JACS* and *IEEE* corpora is strikingly different from the *ESP* corpora, which have a frequency of 20.9%.

T-tests were performed to measure statistical differences between restrictive and non-restrictive relative clauses used in the journals and the results show that there is no statistical difference between them in *JACS* and *IEEE*, while there is a statistical difference between them in *CELL* and *ESP* at the level of 0.001.

The sentences below represent a restrictive and non-restrictive relative clause, the difference of which depends on the use of a comma before the relative pronoun 'which.'

1) Sodium dithionite solutions which were used for reducing the enzyme were prepared by dissolving solid sodium dithionite powder into each pH buffer. (*JACS*, 136, 2014, p. 244.)
2) The subsequent rise of LD preserves the data in the other fecap, which has already been written to a "1" during the previous restore operation. (*IEEE*, VOL. 49, NO. 1, JANUARY, 2014, p. 204.)

Table 1
The frequency distribution of four types of relative clauses.

Research types	Stauble(1978)	Schumann(1980)	Wong(1991)	Cho (2009b)	
				CELL	PRL[a]
OS	55%	53%	48%	75%	82%
OO	25%	35%	26%	20%	15%
SS	12%	6%	17%	5%	2%
SO	7%	4%	7%	0.01%	0.04%

[a] PRL stands for Physical Review Letters.

Table 2
Descriptive data for papers analyzed.

	Total words	Total sentences	Average words in a paper	Average sentences in a paper	Average words in a sentence
CELL	260,199	9712	5204	194	26.8
JACS	224,142	8496	4483	170	26.4
IEEE	249,940	10,901	4999	218	22.9
ESP	365,453	12,363	7309	247	29.6

Table 3
Number of relative clauses and frequency of relative clauses per 100 sentences.

	CELL	*JACS*	*IEEE*	*ESP*
Number of relative clauses	1844	1167	1294	3249
Frequency of relative clauses per 100 sentences	19.0	13.7	11.9	26.3

Table 4
ANCOVA results of the frequency of relative clauses in the journals.

Source	Type III sum of squares	df	Mean square	F	Sig.
Corrected Model	57624.326	4	14406.081	47.931	0.000
Intercept	2735.831	1	2735.831	9.102	0.003
Frequency	3085.666	1	3085.666	10.266	0.002
Journals	34946.094	3	11648.698	38.757	0.000
Error	58609.094	195	300.559		
Total	401548.000	200			
Corrected Total	116233.420	199			

progenitor pool' is the object of the main clause and the relative pronoun 'that' replacing the head noun functions as the subject of the relative clause. It must be noted that sentences with the structure of 'whose + noun(subject) + verb' belong to OS structure, as in sentence 4) below:

4) In addition, one pad contains cells without the repressor construct whose expression measurements serve as the denominator of our fold-change measurements. (*CELL*, 156, March 13, 2014, p. 1321.)

Then, OO type follows with a frequency of 17.8% in the *CELL* corpora, 16.4% in the *JACS* corpora, and 10.6% in the *IEEE* corpora. This structure in the *ESP* corpora presents a relatively high proportion of 31.8%, compared to the science corpora. See below one example of OO structure.

Table 5
Frequency distribution of restrictive and non-restrictive relative clauses.

	CELL (N = 1844)	*JACS* (N = 1167)	*IEEE* (N = 1294)	*ESP* (N = 3249)
Restrictive Relative Clauses	1289 (69.9%)	604 (51.8%)	667 (51.5%)	2571 (79.1%)
Non-restrictive Relative Clauses	555 (30.1%)	563 (48.2%)	627 (48.5%)	678 (20.9%)

Table 6 presents the frequency distribution of the basic four types of restrictive relative clauses. In terms of the frequency distribution of the four types of restrictive relative clauses, all science corpora have the same trend, in which the frequency of OS type is much higher than the other three types, accounting for 84.4% in the *IEEE* corpora, 79.8% in the *JACS* corpora, and 70.1% in the *CELL* corpora. OS type in the *ESP* corpora shows a frequency of 60.7%, which is the highest out of the four types of relative clauses in the corpora but the lowest, when compared to OS types of the other corpora. T-tests were performed to examine the statistical differences of the basic four types of restrictive relative clauses used in each journal. The results show that in all the journals there are statistically significant differences between them with OS > OO > SS > SO types at the level of 0.001. Sentence 3) shows a case of OS type:

3) Stem cell lineages often contain a transit-amplifying (TA) progenitor pool that multiplies the number of differentiating progeny. (*CELL*, 156, March 13, 2014, p. 1259.)

In the sentence, the head noun 'transit-amplifying (TA)

5) Again the larger cubes and octahedra maintain excellent size and shape uniformity that they spontaneously self-assemble into ordered arrays, particularly for octahedra. (*JACS*, 136, 2014, p. 401.)

In the sentence, the head noun 'excellent size and shape uniformity' is the object of the main clause and the relative pronoun 'that' replacing the head noun serves as the object of the relative clause.

Next, SS type follows with a frequency of 10.5% in the *CELL* corpora, 1.8% in the *JACS* corpora, and 4.6% in the *IEEE* corpora. This structure accounts for 4.8% in the *ESP* corpora. Sentence 6) shows one example of SS structure.

6) The differential oscillator structure that has been chosen to implement the temperature-compensated BAW RF and XTAL DCOs was proposed in [17] and is diagrammed in Fig. 7. (*IEEE*, VOL. 49, NO. 1, JANUARY, 2014, p. 217.)

In this structure, the head noun 'the differential oscillator structure' functions as the subject of the main clause as well as that

Table 6
Frequency distribution of the basic four types of restrictive relative clauses.

RC types	*CELL* (N = 1289) Frequency	*JACS* (N = 604) Frequency	*IEEE* (N = 667) Frequency	*ESP* (N = 2571) Frequency
OS	904 (70.1%)	482 (79.8%)	563 (84.4%)	1560 (60.7%)
OO	229 (17.8%)	99 (16.4%)	71 (10.6%)	817 (31.8%)
SS	135 (10.5%)	11 (1.8%)	31 (4.6%)	124 (4.8%)
SO	21 (1.6%)	12 (2.0%)	2 (0.3%)	70 (2.7%)

of the relative clause. The relative clause starting with the relative pronoun 'that' is embedded in the middle of the sentence. The center-embedded relative clause seems to be counter to the natural language process due to its backward modification of the head noun, making the structure difficult to form.

Lastly, SO type relative clauses appear very rarely in all three science corpora: 1.6% in the *CELL* corpora, 2.0% in the *JACS* corpora, and 0.3% in the *IEEE* corpora. The *ESP* corpora also have a very low frequency of 2.7%. Sentence 7) below shows a case of this structure:

7) Another characteristic that these materials presumably share (to a greater or lesser extent) is their publishers' belief that they prepare students for communication in the "real world." (*ESP* 31, 2012, p. 204.)

In the sentence, the head noun 'Another characteristic' is the object of the relative clause and the subject of the main clause. That the relative clause is placed in the middle of the sentence and the head noun functions as the object of the relative clause would make SO structure most difficult to form and its frequency negligible.

Table 7 shows the frequency of restrictive OO type relative clauses. Restrictive OO type relative clauses could be grouped into two types: OO structure with prepositions and ordinary OO structure. In all science corpora, OO type with 'preposition + which' shows a much higher frequency than normal OO type with 'that,' 'which,' or 'zero relative pronoun,' accounting for 73.8% in the *CELL* corpora, 80.8% in the *JACS* corpora, and 84.5% in the *IEEE* corpora. In contrast, the *ESP* corpora present a different frequency distribution in which normal OO type outnumbers OO type with 'preposition + which.' To determine statistical differences between OO structure with prepositions and ordinary OO structure, t-tests were performed and showed that in all the science journals there are statistically critical differences between them with OO structure + prepositions > ordinary OO at the level of 0.001, while *ESP* presented ordinary OO > OO structure + prepositions at the level of 0.01. See below a case of OO type with a preposition:

8) The starting flavin species was assigned as an anionic form in which the N1 position is deprotonated (species 1 in Scheme 3). (*JACS*, 2014, 136, p. 244.)

As in 8), 'in + which' construction combines two separate clauses in an effective way, making the sentence easy to comprehend. In all four corpora, 25 different kinds of prepositions were used. In terms of the frequency of the prepositions employed, 'in' quite outnumbers other prepositions, accounting for about 50%, 52%, 34%, and 56% in the *CELL*, *JACS*, *IEEE*, and *ESP* corpora, respectively. The prepositions 'by,' 'of,' and 'for' follow in the *CELL* corpora; 'for,' 'at,' and 'to' in the *JACS* corpora; and 'at,' 'for,' and 'of' in the *IEEE* corpora. In the *ESP* corpora, the prepositions 'to,' 'of,' and 'for' follow.

Table 8 shows the frequency distribution of non-restrictive relative clauses. Non-restrictive relative clauses have three structures of ', which + verb,' ', preposition + which,' and ', whose + noun + verb.' Of the three types, ', which + verb' far outnumbers the other types, accounting for a frequency of 89.2% in the *CELL* corpora, 86.1% in the *JACS* corpora, and 92.7% in that *IEEE*

corpora. The *ESP* corpora are no exception. T-tests were performed to measure the statistical differences between the three types of non-restrictive relative clauses used in each journal. The results showed that in all the journals there were statistically significant differences between them with ', which + verb' > ', preposition + which' > ', whose + noun + verb' at the level of 0.001. The high frequency of non-restrictive relative clauses is attributable to the easy formation of the structure by adding a comma before the relative pronoun 'which.' The structure is very effective for delivering detailed and additional information about the head noun or the main clause, as illustrated in the sentences below:

9) The CTD of mammalian RNAPII contains 52 repeats of the heptapeptide, which contains the consensus sequence Tyr-Ser-Pro-Thr-Ser-Pro-Ser. (*CELL*, 156, February 13, 2014, p. 685.)
10) Granule cells transmit information in highly divergent connectivity patterns to Purkinje cells, which in turn provide the output of the cerebellum to deep cerebellar nuclei. (*CELL*, 156, January 30, 2014, p. 546.)

In sentence 9), additional information about the head noun 'heptapeptide' is provided in the relative clause, and in sentence 10), a consequence induced by the head noun 'Purkinje cells' is described in the relative clause. A non-restrictive relative clause of ', which + verb' structure sometimes has no clear head noun, as in Sentence 11), where the whole main clause functions as the head noun and the relative clause elucidates an outcome from an experiment mentioned in the main clause. This type of relative clause appears in about 22% of the *CELL* corpora, 39% of the *JACS* corpora, and 23% of the *IEEE* corpora. This structure is found in about 17% of the *ESP* corpora.

11) In contrast, in the case of LUCA-HisH, only 21 out of 226 residues are strictly conserved, which makes the reconstruction much more prone to uncertainties. (*JACS*, 136, 2014, p. 218.)

Another type of a non-restrictive relative clause, ', prepositions + which,' accounts for 8.8% in the *CELL* corpora, 13.0% in the *JACS* corpora, and 6.1% in the *IEEE* corpora. This structure appears in 11.2% in the *ESP* corpora. This structure is employed when the relative clause with the help of a preposition plays a particular role of inducing a result, or describes a particular situation generated, as illustrated in sentence 12):

12) Node clock is distributed across all the uncore modules, in which greater modularity and lower skew requirement are somewhat conflicting in such a large area. (IEEE, VOL. 49, NO. 1, JANUARY, 2014, p. 46.)

In this structure, 16 different kinds of prepositions are used. Of these, the prepositions 'in' and 'of' account for 60% in the *CELL* corpora, 63% in the *JACS* corpora, 85% in the *IEEE* corpora, and 85% in the *ESP* corpora.

The ', whose + noun + verb' structure is minimally used in all three science corpora: 2.0% in the *CELL* corpora, 0.9% in the *JACS*

Table 7
Frequency of restrictive OO type relative clauses.

	CELL (N = 229)	*JACS* (N = 99)	*IEEE* (N = 71)	*ESP* (N = 817)
preposition + which	169 (73.8%)	80 (80.8%)	60 (84.5%)	336 (41.1%)
ordinary OO	60 (26.2%)	19 (19.2%)	11 (15.5%)	481 (58.9%)

Table 8
Frequency distribution of non-restrictive relative clauses.

	CELL (N = 555)	*JACS* (N = 563)	*IEEE* (N = 627)	*ESP* (N = 678)
, which + verb	495 (89.2%)	485 (86.1%)	581 (92.7%)	583 (86.0%)
, preposition + which	49 (8.8%)	73 (13.0%)	38 (6.1%)	76 (11.2%)
, whose + noun + verb	11 (2.0%)	5 (0.9%)	8 (1.3%)	19 (2.8%)

corpora, and 1.3% in the *IEEE* corpora. Even though the frequency of this structure is very low, it is one of most compelling ways to give additional information about the head noun, illuminating semantic clarity between the main and relative clauses. Sentence 13) shows an example of ', whose + noun + verb' structure.

13) However, the main current variation in the voltage sensor is due to the sub-threshold transistor string, whose bias current goes from 17 nA to 42 nA considering process tolerances and temperature variation. (*IEEE*, VOL. 49, NO. 2, FEBRUARY, 2014, p. 338.)

Table 9 shows the frequency of 'that,' 'which,' and 'zero relative pronoun' in restrictive relative clauses. The frequency of relative pronouns was investigated to make sure the general belief that 'which' is more favored in academic texts than 'that' holds true in science and engineering journal papers. This belief was statistically validated by Biber et al. (1999), who stated that 'which' for restrictive relative clauses is more commonly used than 'that' in 70% of academic texts that they investigated. However, in all three science corpora, 'that' is more commonly used than 'which' with a frequency of 97.3% in the *CELL* corpora, 84.3% in the *JACS* corpora, and 71.2% in the *IEEE* corpora. T-tests were performed to measure the statistical differences between the relative pronouns 'that,' 'which,' and zero relative pronoun used in each journal. The results showed that in all the journals there were statistically significant differences between them with that > which > zero relative pronoun at the level of 0.001. Sentence 14) shows one typical example:

14) Given the absence of NK cells in the NSG mice, these results strongly implicate macrophages as the effector cells that mediate the antitumor effect of alemtuzumab in this model. (*CELL*, 156, January 30, 2014 p. 591.)

The proportion of 'zero relative pronoun' is minimal, accounting for 1.7% in the *CELL* corpora, 2.1% in the *JACS* corpora, and 0.5% in the *IEEE* corpora. It is worth noting that the dominant subject of the relative clause in this structure is 'we,' occupying 39 cases out of 40 across all three science corpora, with only one exception: the use of 'this.'

15) The operation principle of the proposed ringing killer looks similar to that proposed earlier [16], but the proposed one was implemented without a comparator, which consumes more perating current than that of the inverter chain we used. (*IEEE*, VOL. 49, NO. 2, FEBRUARY, 2014, p. 507.)

In contrast, the *ESP* corpora show a relatively high ratio of zero relative pronouns, 11.2%. This is probably due to the use of a variety of pronouns as the subject and human subjects in the relative clause, which are usually found in speech registers. For example, in an *ESP* paper where 28 cases of zero relative pronoun appear, 'they,' 'she,' 'he,' 'you,' and 'I' are used as the subject of the relative clause and human subjects such as 'students' and 'participants' also are adopted.

5. Discussion

This comparative corpus-based investigation of the frequency and use of relative clauses in science and engineering journal papers provides valuable insights for the learning and teaching of them to science and engineering paper authors and ESP/EAP practitioners. Of the attributes identified, the first noteworthy point is the high frequency of relative clauses: 19.0 cases in the *CELL* corpora, 13.7 cases in the *JACS* corpora, and 11.9 cases in the *IEEE* corpora per 100 sentences. The high frequency is apparent when compared to the frequency of supplementive participle clauses, another significant sentence-combining device of English, 6.2 cases and 3.8 cases per 100 sentences in *CELL* and *Physical Review Letters*, respectively (Cho and Kim, 2009). This high frequency of relative clauses used in science and engineering journal papers reflects the fact that there are various types of relative clauses with different uses and functions and that relative clauses play significant roles in science and engineering journal papers, such as giving semantic clarity between clauses and promoting textual variety. This also implies the need for relative clauses to be taught with more emphasis and attention.

The frequency of relative clauses in academic journal papers investigated in this research, however, is lower than that of relative clause structures found in Biber et al. (1999). In their research, about 9900 cases of relative clauses, including zero-relativizers, per million words were found. If one million words apply to the corpora of this study, about 7087 cases of relative clauses are assumed to be used in the *CELL* corpora, about 5207 cases in the *JACS* corpora, and about 5177 cases in the *IEEE* corpora. The *ESP* corpora with the frequency of 8890 cases per million words present a bit less but similar frequency of relative clauses to Biber et al. (1999). This difference is probably from the difference of academic texts investigated. The corpora for this study belong to academic journals of science and engineering, whereas academic prose in Biber et al. (1999) include book extracts as well as research articles, and a wide range of academic disciplines, including sciences, social sciences, and humanities. Another interpretation of the less frequent use of relative clauses found in this study is attributable to science and

Table 9
Frequency of 'that,' 'which' and 'zero relative pronoun' in restrictive relative clauses.

	CELL (N = 1099[a])	*JACS* (N = 515[a])	*IEEE* (N = 591[a])	*ESP* (N = 2188[a])
that	1069 (97.3%)	434 (84.3%)	421 (71.2%)	1178 (53.8%)
which	11 (1.0%)	70 (13.6%)	167 (28.3%)	766 (35.0%)
zero relative pronoun	19 (1.7%)	11 (2.1%)	3 (0.5%)	244 (11.2%)

[a] Please note that 'whose + noun + verb' of OS type and 'preposition + which' of OO type were excluded in the distribution frequency count since 'whose' and 'which' cannot be replaced by 'that'.

engineering journal paper authors' preference for simple, straightforward sentences. Even though sentences with relative clauses ensure semantic clarity between phrases and sentences, the preferable use of simple sentences may result in the less frequent use of relative clauses in science and engineering journal papers.

Next, the high ratio of non-restrictive relative clauses found in the *JACS* and *IEEE* corpora, reaching 48, is worth discussing from a pedagogical perspective. It stands out, compared to 30.1% in the *CELL* corpora and 20.9% in the *ESP* corpora, as well as 23% in Biber et al. (1999). This is probably because *JACS* and *IEEE* papers, representing chemistry and electrical and electric engineering, are likely to include many unfamiliar, new concepts and findings, which need to be accurately and extensively explained in a relative clause. Non-restrictive clauses could successfully fulfill the role. As illustrated below, the implication of the head noun '30–40 Å' is provided in a non-restrictive relative clause:

16) Using the results from self-crowding, we estimate that the distance between protein surfaces at which this transition occurs is 30–40 Å, which is a striking manifestation of the collective and coordinated behavior of strongly hydrogen bonding environments. (*JACS*, 136, 2014, p. 193.)

An *IEEE* paper also illustrates the typical use of a non-restrictive relative clause, in which an attribute of the head noun '70 A cm' is well accounted for:

17) Consequently, the maximum current density is 70 A cm, which satisfies the safety standard <100 A cm [16]. (*IEEE*, VOL. 49, NO. 1, JANUARY, 2014, p. 170.)

The high frequency of non-restrictive clauses across all three science corpora could be better explained by the fact that they all are from academic journals. Since there is no data available concerning the frequency of non-restrictive relative clauses in other text types, such as fiction and newspapers, it cannot be confirmed that the high frequency is solely attributable to academic journal papers. However, we can conjecture that some characteristics of academic journals, such as the need to provide additional, detailed information about new concepts and findings, led to the high use of non-restrictive relative clauses in the science and engineering journal papers. The ease with which non-restrictive relative clauses can be constructed also likely contributed to their frequent use in the science and engineering journal papers.

The frequency distribution of the four basic types of restrictive relative clauses is in line with the previous research (Abdolmanafi and Rahmani, 2012; Cho, 2009b; Sadighi, 1994; Stauble, 1978; Wong, 1991). In all three science corpora, OS type tops, accounting for 70.1% in the *CELL* corpora, 79.8% in the *JACS* corpora, and 84.4% in the *IEEE* corpora. OO type follows with a frequency of 17.8% in the *CELL* corpora, 16.4% in the *JACS* corpora, and 10.6% in the *IEEE* corpora. Then, SS construction shows a frequency of 10.5% in the *CELL* corpora, 1.8% in the *JACS* corpora, and 4.6% in the *IEEE* corpora. SO type is minimally used in all three science corpora: 1.6% in the *CELL* corpora, 2.0% in the *JACS* corpora, and 0.3% in the *IEEE* corpora.

A careful investigation of the frequency distribution of restrictive relative clauses, however, reveals differences between the current research and previous research. In this research, the frequency of OS structure in the *JACS* and *IEEE* corpora far outnumbers the other types while that of SO type in the corpora is pretty low. These extreme differences are not found in other research. In Stauble (1978), OS structure accounts for 55% and SO 7%; in Schumann (1980), OS structure has a frequency of 53% and SO structure 4%; and in Wong' study (1991), OS structure presents only 48% and SO 7%. The differences between the current research and

the previous studies probably stem from the different registers analyzed in each study. Stauble's study is based on informal speech, spontaneous writing and published writing altogether and Wong' study utilizes 170 compositions of Hong Kong secondary school students. On the other hand, the data of this study are from published papers of science and engineering journals. The very high frequency of OS structure and extremely low frequency of SO structure identified in the *JACS* and *IEEE* corpora is characterized as one of the prominent features of science and engineering journal paper writing, reflecting the limited use of relative clauses by science and engineering journal paper authors. The trait is also manifested by comparing the relatively low frequency of OO structure found in the *JACS* and *IEEE* corpora, 16.4% and 10.6%, respectively, with the high frequency of the same structure found in the *ESP* corpora, 31.8%.

Next, 'preposition + which' in restrictive OO structure presents a high proportion of 84.5% in the *IEEE* corpora, 80.8% in the *JACS* corpora, and 73.8% in the *CELL* corpora. This could be also considered a prominent feature of relative clauses used in science and engineering journal papers. The high ratio is compared to the relatively low frequency of the same structure in the *ESP* corpora, which is 41.1%. The high frequency of the structure in the science and engineering journal papers resonates that it is an efficient way to deliver the meanings intended and is easy to construct. Sentences with the structure read well, and they are easy to comprehend without the interference of natural language processes between the main and relative clauses. A variety of prepositions with different functions and usages also may have led to the high frequency of the structure. Out of 25 prepositions used in the structure, 'in' is most frequently used in all three science and engineering corpora, followed by 'by,' 'of,' 'for,' 'at,' and 'to.'

Next, the relative pronoun 'that' is highly favored over 'which' in restrictive relative clauses. This is particularly true in the *CELL* and *JACS* corpora, where 'that' accounts for 97.3% and 84.3%, respectively. These findings contradict the general belief that 'which' is closely associated with academic writing and thus is preferred in the register. The preference of 'that' found in this study also does not coincide with the corpus-based investigation of Biber et al. (1999), who discovered that 70% of the academic texts in the corpus analyzed for their study adopt 'which' more commonly than 'that' for restrictive clauses, while 75% of fiction writings use 'that' more frequently than 'which' for the same case. The unproportionally high use of 'that' found in this study manifests a current trend of using 'that' over 'which' in some science journal papers. As far as *CELL* papers are concerned, the relative pronoun 'that' seems to have become the norm for restrictive relative clauses, and 'which' is no longer recommended for restrictive relative clauses in the discourse community. The *JACS* papers seem to follow the trend of the *CELL* papers. The clear-cut distinction concerning the use of the two relative pronouns in *CELL* papers provide guidelines for how the relative pronouns 'that' and 'which' are used in the journal.

The findings and implications of the use of relative clauses in science and engineering journal papers remind us of Hyland (2006, 2011) and Hyand and Tse (2009) investigating differences and variations in academic disciplines. This research has found that some features of relative clauses in *CELL*, *JACS*, and *IEEE* papers, such as the high proportion of non-restrictive clauses and preference of 'that' as a relative pronoun in restrictive relative clauses, are apparently different from those in *ESP* papers. Hyland (2006) argues that different disciplines have different ways to represent their arguments in discourse level. For example, humanities and social sciences differ from science and engineering in how they illustrate their research: science and engineering are more concerned with conventions of research such as generalizations, methods, and procedures, while humanities and social sciences are

likely to take more personal positions than science and engineering in order to engage readers. Differences in disciplines are also found in interactional markers such as stance and engagement between soft fields of disciplines of philosophy, marketing, sociology and applied linguistics, and science and engineering fields (Hyland, 2011). Even in the area of vocabulary (Hyand and Tse, 2009), different disciplines, such as sciences, engineering, and social sciences have shown a different frequency distribution of vocabulary used in academic journals, textbooks, and book reviews, which suggests the need for the development and teaching of a more disciplinary-oriented lexical list. The discourse and lexical differences in academic disciplines and their pedagogical implications witnessed in Hyland (2006, 2011) and Hyland and Tse (2009) foreshadow the different features of relative clauses we observed between science and engineering journals, and a language research journal.

It also should be mentioned we could not verify whether the journal papers chosen for this study were composed by native speakers or not. This is because a science paper has several authors and it is not possible to confirm whether they are native speakers of English by their names and/or their institutes. It is likely that non-native speaking authors of science and engineering journal papers would avoid using sentences with relative clauses, as found in the previous research of non-native speakers of English (Chang, 2004; Schachter, 1974; Yip and Matthews, 1991). However, it is also plausible to claim that non-native authors of academic journals tend to have high English proficiency, which would not constrain the use of complex sentences with relative clauses.

6. Conclusion and pedagogical implications

Relative clauses of English, despite their diverse, effective functions of giving semantic clarity and promoting syntactic maturity and textual variety, have not received adequate attention in ESP/EAP fields. This corpus-based research was intended to inquire into relative clauses used in science and engineering journal papers for pedagogical purposes. In particular, it was designed to reveal unique features of relative clauses employed in them through a comparison of papers of science and engineering journals and those of an English language research journal, *English for Specific Purposes*. It was expected that findings from this research could offer valuable insights for the teaching and learning of the grammatical feature to science and engineering paper authors and ESP/EAP practitioners. In fact, this research revealed some distinctive traits of science and engineering journal papers, which could be taken into account when teaching scientific writing and composing science and engineering journal papers. Some of them, such as the high frequency of non-restrictive relative clauses, the high proportion of 'prepositions + which' in the restrictive OO structure, and the extremely high use of 'that' over 'which' for restrictive relative clauses, would constitute valuable resources for the teaching of relative clauses for science and engineering journal paper writing.

Teaching relative clause structures with high frequency needs to be conducted with the emphasis on usages in a discourse community, and their functions and effectiveness. For example, the high frequency of the non-restrictive relative clauses in *JACS* and *IEEE* papers is highlighted as a specific, unique feature of science and engineering journals, compared to those used in the *ESP* corpus. As discussed, the proportion of the non-restrictive relative clauses in *JACS* and *IEEE* papers accounts for 48.2% and 48.5%, respectively, which far outnumbers the proportion of non-restrictive relative clauses used in *ESP* papers, 20%. Thus, while teaching relative clauses to science and engineering students, the use of non-restrictive relative clauses should be highly

recommended along with an explanation of the easy construction of the structure as an effective way to provide additional, detailed information about new concepts and findings.

This suggestion implies that instructors, in particular, ones with an ESP background, should be familiar with the discipline-specific features of relative clause structures. Otherwise, they would not offer effective instruction. This is particularly true in the case of teaching relative pronouns for restrictive relative clauses. It has been widely known and taught that the relative pronoun 'that' is less favored than 'which' in academic writing (Biber et al., 1999). Science and engineering journals, as found in this research, however, prefer 'that' to 'which' as a proper relative pronoun for restrictive relative clauses. Not being familiar with this fact, ESP/EAP practitioners would misinform their students about the choice of relative pronouns proper in their discourse community.

When teaching relative clauses to science and engineering journal paper authors, mainly PhD students in the fields of science and engineering, it would be advisable to provide them with well-written sample papers showing representative features of relative clauses and to have them probe the traits. This way, rather than giving them a detailed account of relative clauses, will help the authors quickly comprehend in what context and how often relative clauses could be used in science and engineering journal papers. The self-directed and autonomous approach would bring about a better, more direct understanding of relative clauses, as revealed in Hafner and Candlin (2007) and Lee and Swales (2006).

Lastly, it should be noted that the extent to which the findings of the current study can be generalized is somewhat limited. This limitation therefore calls for further corpus-based research of other science and engineering journals and journals of different academic disciplines, which is expected to highlight a more comprehensive picture of relative clauses used in academic journal papers.

References

Abdolmanafi, S.J., Rahmani, Z., 2012. An investigation of the learnability of relative clauses by EFL learners. World J. Engl. Lang. 2, 29–37.

Baek, S., 2012. Processing of English Relative Clauses by Adult L2 Learners. Unpublished doctoral dissertation. University of Illinois, Urbana-Champaign.

Biber, D., Johansson, S., Leech, G., Conrad, S., Finegan, E., 1999. Grammar of Spoken and Written English. Pearson Education Limited, Essex.

Bunton, D., 1979. Common English Errors in Hong Kong. Longman, London.

Cargill, M., O'Connor, P., Li, Y., 2012. Educating Chinese scientists to write for international journals: addressing the divide between science and technology education and English language teaching. Engl. Specif. Purp. 31, 60–69.

Casanave, C.P., Hubbard, P., 1992. The writing assignments and writing problems of doctoral students: faculty perceptions, pedagogical issues and needed research. Engl. Specif. Purp. 11, 33–49.

Chang, C., Guo, C., 2011. A corpus approach to online materials development for writing research articles. Engl. Specif. Purp. 30, 222–234.

Chang, Y., 2004. Second language relative clause acquisition: an examination of cross-linguistic influences. In: Paper Presented at the Annual Meeting of the American Association for Applied Linguistics, Portland, OR.

Charles, M., 2006. Phraseological patterns in reporting clauses used in citation: a corpus-based study of theses in two disciplines. Engl. Specif. Purp. 25, 5–24.

Charles, M., 2007a. Argument or evidence? Disciplinary variation in the use of the noun *that* pattern in stance construction. Engl. Specif. Purp. 26, 203–218.

Charles, M., 2007b. Reconciling top-down and bottom-up approaches to graduate writing: using a corpus to teach rhetorical patterns. J. Engl. Acad. Purp. 6, 289–302.

Charles, M., 2010. Using concordancing to teach rhetorical functions. In: Frankenberg-Garcia, A., Flowerdew, L., Aston, G. (Eds.), New Trends in Corpora and Language Learning. Continuum, New York, pp. 26–48.

Chen, Q., Ge, C., 2007. A corpus-based lexical study on frequency and distribution of Coxhead's AWL word families in medical research articles (RAs). Engl. Specif. Purp. 26, 502–514.

Chiang, D., 1980. Predictors of relative clause production. In: Scarcella, R., Krashen, S. (Eds.), Research in Second Language Acquisition. Newbury House, Rowley, MA, pp. 142–145.

Cho, D.W., 2009a. Science journal paper writing in an EFL context: the case of Korea. Engl. Specif. Purp. 28, 230–239.

Cho, D.W., 2009b. A corpus-based study on the frequency and usage of English relative clauses in two science journals. Foreign Lang. Educ. 16 (3), 1–20.

Cho, D.W., 2012. English-medium instruction in the university context of Korea:

tradeoff between teaching outcomes and media-initiated university ranking. J. ASIA TEFL 9 (4), 135–163.

Cho, D.W., Kim, J., 2009. A corpus-based study on the frequency and usage of participial clauses in two science journals. Engl. Teach. 64 (4), 25–42.

Dong, Y.R., 1998. Non-native graduate students' thesis/dissertation writing in sciences: self-reports by students and their advisors from two US institutes. Engl. Specif. Purp. 17, 369–390.

Duszak, A., Lewkowicz, J., 2008. Publishing academic texts in English: a Polish perspective. J. Engl. Acad. Purp. 7, 108–120.

Eckman, F., 2007. Hypotheses and methods in second language acquisition: testing the noun phrase accessibility hierarchy on relative clauses. Stud. Second Lang. Acquis. 29, 321–327.

El-Malik, A., Nesi, H., 2008. Publishing a research in a second language: the case of Sudanese contributors to international medical journals. J. Engl. Acad. Purp. 7, 87–96.

Flowerdew, J., 1999a. Writing for scholarly publication in English: the case of Hong Kong. J. Second Lang. Writ. 8 (2), 123–145.

Flowerdew, J., 1999b. Problems in writing for scholarly publication in English: the case of Hong Kong. J. Second Lang. Writ. 8 (3), 243–264.

Flowerdew, J., 2003. Signaling nouns in discourse. Engl. Specif. Purp. 22, 329–346.

Flowerdew, J., 2013. Corpus-based Research and Pedagogy in EAP: from Lexis to Genre. *Language Teaching*. http://dx.doi.org/10.1017/S0261444813000037 available on CJO2013.

Flowerdew, L., 2001. The exploitation of small learner corpora in EAP materials design. In: Ghadessy, M., Henry, A., Roseberry, R.L. (Eds.), Small Corpus Studies and ELT: Theory and Practice. John Benjamins Publishing Company, Philadelphia, pp. 363–379.

Flowerdew, L., 2012. Corpora and Language Education. Palgrave Macmillan, New York.

Frankenberg-Garcia, A., Flowerdew, L., Aston, G., 2010. New Trends in Corpora and Language Learning. Continuum, New York.

Gass, S., 1979a. An Investigation of Syntactic Transfer in Adult Second Language Acquisition. Unpublished doctoral dissertation. Indiana University, Bloomington.

Gass, S., 1979b. Language transfer and universal grammatical relations. Lang. Learn. 29, 327–344.

Gass, S., Behney, J., Plonsky, L., 2013. Second Language Acquisition: an Introductory Course. Routledge, New York and London.

Gavioli, L., 2005. Exploring Corpora for ESP Learning. John Benjamins Publishing Company, Amsterdam.

Giannoni, D.S., 2008. Medical writing at the periphery: the case of Italian journal editorials. J. Engl. Acad. Purp. 7, 97–107.

Gilquin, G., Granger, S., Paquot, M., 2007. Learner corpora: the missing link in EAP pedagogy. J. Engl. Acad. Purp. 6, 319–335.

Hafner, C.A., Candlin, C.N., 2007. Corpus tools as an affordance to learning in professional legal education. J. Engl. Acad. Purp. 6, 303–318.

Hyland, K., 2006. Disciplinary differences: language variation in academic discourses. In: Hyland, K., Bondi, M. (Eds.), Academic Discourse across Disciplines. Peter Lang<x>, Frankfort, pp. 17–45.

Hyland, K., 2011. Disciplines and discourses: social interactions in the construction of knowledge. In: Starke-Meyerring, D., Paré, A., Artemeva, N., Horne, M., Yousoubova, L. (Eds.), Writing in the Knowledge Society. Parlor Press and The WAC Clearinghouse, West Lafayette, IN, pp. 193–214.

Hyand, K., Tse, P., 2009. Academic lexis and disciplinary practice: corpus evidence for specificity. Int. J. Engl. Stud. 9 (2), 111–129.

Ioup, G., Kruse, A., 1977. Interference versus structural complexity in second language acquisition: language universals as a basis for natural sequencing. In: Brown, R.D., Yorio, C.A., Crymes, R.R. (Eds.), On TESOL 77: Teaching and Learning English as a Second Language: Trends in Research and Practice. TESOL, Washington, DC, pp. 159–171.

Jeon, K.S., Kim, H.Y., 2007. Development of relativization in Korean as a foreign language: noun Phrase accessibility hierarchy in a head-internal and head-external relative clauses. Stud. Second Lang. Acquis. 29, 253–276.

Kameen, P., 1978. A mechanical, meaningful, and communicative framework for ESL sentence combining exercises. TESOL Q. 12, 395–401.

Kameen, P., 1983. Syntactic skills and ESL writing quality. In: Freedman, A., Pringle, I., Yalden, J. (Eds.), Learning to Write First Language/second Language. Longman, New York, pp. 162–170.

Keenan, E.L., Comrie, C., 1977. Noun phrase accessibility and universal grammar. Linguist. Inq. 8, 63–99.

Krishnamurthy, R., Kosem, I., 2007. Issues in creating a corpus for EAP pedagogy and research. J. Engl. Acad. Purp. 6, 356–373.

Kuno, S., 1974. The position of relative clauses and conjunctions. Linguist. Inq. 5, 117–136.

Lannon, J.M., 1988. Technical Writing. Scott, Foresman and Company, Glenview.

Lee, D., Swales, J., 2006. A corpus-based ESP course for NNS doctoral student: moving from available specialized corpora to self-compiled corpora. Engl. Specif. Purp. 25, 56–75.

Li, Y., 2006. A doctoral student of physics writing for international publication: a socio-politically-oriented case study. Engl. Specif. Purp. 25, 456–478.

Liu, M., 1998. An Experimental Study of English Relative Clauses by Chinese Junior High School Students in Taiwan. Unpublished MA thesis. National Taiwan Normal University, Taipei.

Marefat, H., Rahmany, R., 2009. Acquisition of English relative clauses by Persian EFL learners. J. Lang. Linguistic Stud. 5, 22–48.

Markel, M.H., Holmes, H., 1994. Technical Writing: Situations and Strategies (Ontario: Nelson Canada).

Miura, S., 1989. Hiroshima English Learners' Corpus. Department of English Education, Hiroshima University.

Newman, J., Baayen, H., Rice, S., 2011. Corpus-based Studies in Language Use, Language Learning and Language Documentation. Rodopi<x>, Amsterdam.

Park, C., 2000. The Syntactic Difficulty of Relative Clauses for Korean Students: an Analysis of Errors and Avoidance Strategies. Unpublished MA thesis. Arizona State University, Arizona.

Phoocharoensil, S., Simargool, N., 2010. English relativization and learners' problems. J. Pan Pac. Assoc. Appl. Linguist. 14 (1), 109–129.

Sadighi, F., 1994. The acquisition of English restrictive relative clauses by Chinese, Japanese, and Korean adult native speakers. IRAL(Int. Rev. Appl. Linguist.) 32, 141–153.

Salager-Meyer, F., 1999. Referential behavior in scientific writing: a diachronic study (1810-1995). Engl. Specif. Purp. 18, 279–305.

Salager-Meyer, F., 2008. Scientific publishing in developing countries: challenges for the future. J. Engl. Acad. Purp. 7, 121–132.

Schachter, J., 1974. An error in error analysis. Lang. Learn. 24, 205–214.

Schumann, J., 1980. The acquisition of English relative clauses by second language learners. In: Scarcella, R., Krashen, S. (Eds.), Research in Second Language Acquisition. Newbury House, Rowley, MA, pp. 118–131.

Sheldon, A., 1974. The role of parallel function in the acquisition of relative clauses in English. J. Verbal Learn. Verbal Behav. 13, 272–281.

Stauble, A., 1978. A Frequency Study of Restrictive Relative Clause Types and Relative Pronoun Usage in English. Unpublished English 215 paper. UCLA.

Weisman, H.M., Collins, C.E., 1998. Basic Technical Reporting. Prentice Hall Canada Inc, Ontario.

Wong, J., 1991. Learnability of relative clauses: a Hong Kong Case. Perspectives 3 (1), 108–117.

Yip, V., Matthews, S., 1991. Relative complexity: beyond avoidance. CUHK Pap. Linguist. 3, 112–124.

Dong Wan Cho is Professor in the Division of Humanities and Social Sciences of the Pohang University of Science and Technology, Pohang, Korea. He earned his doctoral degree in language testing from the University of Illinois at Urbana-Champaign in 1995 and published many journal articles on ESL/EFL writing assessment, ESP/EAP writing for publication purposes, and ESL/EFL curriculum development in international and domestic journals. He now serves as the Vice President of the Korea English Language Testing Association, KELTA and editorial board of English Language Assessment, Language Testing in ASIA, the Journal of ASIA TEFL, and Studies in English Education.