

**Selecting complex sentence structures to teach in academic writing:  
A corpus-based innovation**

**KHOA DANG TRUONG & NHAN THANH DO**

Gia Viet English Language Center, Vietnam

***Introduction***

In most academic writing programs worldwide, instruction aims to develop students' awareness of essay types, written language use, and test-taking strategies. However, teachers' instructional practice is often dictated by their knowledge base, perception, and the intuitions they have about language use (Borg, 2006). Such intuitions are not always correct (Cobb & Boulton, 2015) and may not fully reflect how language is used in a particular type of academic writing by native speakers. Thus, the grammatical structures they teach may not correspond well to what is frequently and naturally used in that type of text. On the part of students, because of their limited knowledge of writing conventions, they might end up using a variety of grammatical forms while not considering how appropriate these are.

Thus, we feel that it is imperative for writing teachers to allow for 'data-driven learning' (Johns, 1991, p. 17) whereby students explore language constructions through consulting a purpose-built corpus. Such a corpus exposes students to language features peculiar to a particular written genre, but it can also help them identify which are more frequently and appropriately used than others in the genre. In what follows, we will describe how to identify frequently used complex sentence forms in a corpus of essays written by experts to inform the design and implementation of teaching materials for writing.

***Complex Structures***

To achieve a high score for grammatical range and accuracy (a common criterion for writing assessment) students often need to demonstrate a high level of grammatical

complexity in their writing. They are also asked to diversify the complex structures they use in their essays. Although complexity can be at both phrasal and clausal levels, our pedagogical innovation concentrates only on clausal complexity, aiming to explore which clauses are most frequently used by expert writers to construct complex sentences. Understanding this will encourage teachers to focus their material design and instruction on these forms.

According to Hewings and Haines (2015), complex sentences in English can be formed with adverbial clauses such as reason (*because, since, as, now that*), concession (*although, though, even though*), contrast (*while, whilst, whereas*), condition (*if, unless, given that, provided that, supposed that, as long as*), time (*when, whenever, once, before, after, as soon as*), and place (*where, wherever*). Complex sentences can also include noun clauses – a group of words with a subject and a verb that can be a subject, an object, or an object of a preposition. Common noun clause markers are *that, if, whether*, and *wh*-question words. Relative clauses, introduced by relative pronouns such as *that, who, whom, which*, and *whose*, are another type. In the following sections, we will examine the frequency of use of these clauses in a corpus of expert IELTS writings and provide related pedagogical implications.

### ***Collecting and Analysing Samples***

We collected from printed textbooks 162 IELTS discursive essays written by native-English-speaking experts, constituting a corpus size of 49,102 words, which is large enough for classroom application. To identify complex sentences in the corpus, we used AntConc, a software created by Anthony (2019) for analysing written corpora. First, we converted the corpus into a plain text before loading it onto the application. Finally, we extracted subordinate clauses by keying individual subordinators in the search box. For instance, to find the adverbial clauses of reason, we used keywords such as *because, since, as*, and *now that*. However, the subordinator *as*, similar to *when*, can be used to form an adverbial clause of time, and it can also be part of other lexico-grammatical constructions such as comparatives. We addressed this problem by manually removing the non-adverbial constructions from our output.

### ***Results***

For adverbial clauses, we found the most frequently occurring adverbial clauses were created with *because* (84), *although* (45), *while* (67), *if* (145), *when* (61), and *where* (44). Noun clauses with *that* stand out, with 667 occurrences. In addition, relative clauses with *that* rank second, including 254 occurrences, while those with *who* and *which* are also relatively frequent, with 153 and 104 occurrences respectively. The other less frequently used clauses are shown in Tables 1 and 2 below.

**Table 1**  
**Frequency of adverbial clauses**

Clause Types	Subordinators	Frequency (No. of occurrences)
Reason	<i>because</i>	84
	<i>as</i>	51
	<i>since</i>	8
	<i>now that</i>	1
Concession	<i>although</i>	45
	<i>though</i>	17
	<i>even though</i>	6
Contrast	<i>while</i>	67
	<i>whereas</i>	7
	<i>if</i>	145
Condition	<i>as long as</i>	8
	<i>unless</i>	3
	<i>when</i>	61
	<i>before</i>	15
	<i>as</i>	12
Time	<i>once</i>	8
	<i>whenever</i>	2
	<i>as soon as</i>	1
	<i>after</i>	2
Place	<i>where</i>	44
	<i>wherever</i>	3

**Table 2**  
**Frequency of noun clauses and relative clauses**

Clause Types	Markers	Frequency (No. of occurrences)
Noun Clause	that	667
	what	54
	how	37
	why	36
	whether	26
	how much	5
	how many	0
	if	0
	that	254
	who	153
Relative Clause	which	104
	whom	1
	whose	3

## Pedagogical Suggestions

Our results show the frequency of some clause types in a corpus of IELTS writing models. In what follows, we suggest some ideas and activities teachers can execute in teaching writing on the basis of these results.

### Material and Lesson Design

We recommend that teachers consider complex sentence forms and associated markers with the highest level of frequency an important part of their material and lesson design. For example, materials can focus on clauses of reason (*because*), concession (*although*), contrast (*while*), time (*when*), place (*where*), condition (*if*), noun clauses (*that*), and relative clauses (*that*) and include multiple exercises that require students to write complex sentences. This allows for their ample practice, and thus greater proficiency, appropriateness, and naturalness in using these structures over time. However, teachers' choices should be sensitive to essay types and topics. For example, in an essay that discusses a social issue, students are likely to frequently use *if* to evaluate their proposed measures, while in a discussion essay, they tend to use *while* to present opposing ideas.

### Consciousness Raising

An activity that teachers could carry out in class to raise students' grammatical consciousness is to have them examine a frequently used type of clause shown by a purpose-built corpus (see Figure 1). They can, for example, look into the different constructions of *if* clauses in some model paragraphs to identify tense uses and related meanings (e.g., indicative mood or subjunctive mood). They can also be asked to discuss the thematic reason behind placing the conditional clauses at the beginning of or inside a sentence by juxtaposing them in the context of the paragraphs.

1	believe that young people would lead happier lives	if they had a more \x91old-fashioned\x92	Corpus 162 IELTS essays.txt
2	ties and connecting provinces within the country.	If someone has a family, faster transport allows the	Corpus 162 IELTS essays.txt
3	regarding how and where to find that knowledge.	If you are a French national, you can apply	Corpus 162 IELTS essays.txt
4	ryone, including the poorest. On the other hand,	if too high a percentage of salaries is taken	Corpus 162 IELTS essays.txt
5	worthwhile, but I think that it is better	if people are able to give money directly to	Corpus 162 IELTS essays.txt
6	xtricably bound up in its manufactured artefacts.	If the relentless advance of international brands in	Corpus 162 IELTS essays.txt
7	by private car more difficult is only viable	if reliable and affordable alternatives are in place	Corpus 162 IELTS essays.txt
8	ds will likely become dependent on their parents.	If your parents always pick out your clothes then	Corpus 162 IELTS essays.txt
9	to look after their child during the day.	If more meetings and training took place online, thi	Corpus 162 IELTS essays.txt
10	However, there would be better long-term effects	if international organisations and governments redir	Corpus 162 IELTS essays.txt
11	study. In my opinion, society will benefit more	if our students are passionate about what they are	Corpus 162 IELTS essays.txt
12	part of mice or rats can be justified	if human lives are saved. They argue that opponents	Corpus 162 IELTS essays.txt
13	our society need this type of aid and	If these costs are covered, at least partially, by	Corpus 162 IELTS essays.txt
14	est medical problems these days is dementia. Even	if old people are still fit, they may not	Corpus 162 IELTS essays.txt
15	such products in moderation, but some messages,	if indeed there are any, only appear as minor	Corpus 162 IELTS essays.txt
16	our grandparents\x92 generation. The same is true	if we look at the increasing tendency for young	Corpus 162 IELTS essays.txt
17	ound the world in various countries, particularly	if they are at war or experiencing a severe	Corpus 162 IELTS essays.txt
18	o education indirectly and poverty more directly.	If you look at countries with extremely low poverty	Corpus 162 IELTS essays.txt
19	computing industry in today\x92s world. However,	if this is at the expense of having new	Corpus 162 IELTS essays.txt
20	a short time they often do not care	if they behave badly or litter the streets. Secondly	Corpus 162 IELTS essays.txt
21	of compulsory training for parents. Besides, even	if parents could be forced to attend, I doubt	Corpus 162 IELTS essays.txt
22	n existence, because our lives would be diminished	if fascinating and beautiful creatures such as tiger	Corpus 162 IELTS essays.txt
23	cultural identity. It would be a real pity	if pop music became so predominant that these nation	Corpus 162 IELTS essays.txt
24	ittle to disrupt the underlying causes of hunger.	If people are better educated then they can get	Corpus 162 IELTS essays.txt
25	is required to return what they have stolen	if they are caught. I think that there is	Corpus 162 IELTS essays.txt

Figure 1. Example of the 'if' and 'unless' output from the corpus

### *Writing and Comparing*

For written production, teachers can use some of the essay topics from the corpus, provide students with relevant ideas, and then have them write sentences or even a paragraph using these ideas (e.g., a relative clause with *that*). Subsequently, they can make structural and semantic comparisons between their own sentences with those from the corpus and make ensuing adjustments. This activity, if done frequently, may help to take their use of the clause to a higher level thanks to consulting expert models.

### *Written Error Correction*

Some teachers may be interested in correcting grammar errors for their students, but from a learning-by-doing perspective, it is advisable that students scrutinise and correct their faulty sentences themselves. Rather than directly correcting their wrong use of a noun phrase after *although*, for instance, teachers can just highlight the error and ask students to consult a corpus in terms of the differences between *although* and *despite*. By experimenting with these concession clauses while being informed by the corpus, they can better understand, acquire, and retain how to use the constructions.

## CONCLUSION

As demonstrated above, a corpus software like AntConc is highly efficient in directing students' attention to how grammatical features such as complex sentences are used in academic writing. However, teachers need to be aware of how much time a corpus-based classroom activity on a syntactic structure may take, so it is advisable to turn it into an extramural project where students examine the structure collaboratively. Although it may be time-consuming, such an activity is instrumental in deepening students' understanding of written language features and also in boosting their autonomous learning and analytical skills, so it is worth the investment.

## REFERENCES

- Anthony, L. (2019). *AntConc* (Version 3.5.8) [Computer Software]. Waseda University. Retrieved from <http://www.laurenceanthony.net/software.html>
- Borg, S. (2006). *Teacher cognition and language education*. Continuum.
- Cobb T., & Boulton A. (2015). Classroom applications of corpus analysis. In D. Biber & R. Reppen (Eds), *The Cambridge Handbook of Corpus Linguistics*. Cambridge University Press.

Hewings, M., & Haines, S. (2015). *Grammar and vocabulary for advanced*. Cambridge University Press.

Johns, T. (1991). Should you be persuaded: Two samples of data-driven learning materials. *English Language Research Journal*, 4, 1–16.

**Khoa Dang Truong** has 10 years' teaching experience. He is interested in how to help students develop their academic writing. His current role is as academic manager and TESOL trainer for Gia Viet English Language Center, Vietnam.

khoafulbrighter@gmail.com

**Nhan Thanh Do** has taught IELTS at Gia Viet English Language Center for eight years now. He has a particular interest in teaching academic writing and reading.

wilson.nhan1991@gmail.com