A Corpus-Based Study of Korean EFL Learners' Use of Subordinate Conjunctions in Argumentative Writing: With Focus on *if*, *when*, and *because*

Baek, Soondo

Kookmin University

[Abstract]

The use of subordinating conjunctions in L2 writing has been widely studied with regard to frequency of occurrence, grammatical accuracy, and relative ordering of subordinate and main clauses. However, relatively few studies investigated the diverse aspects of the L2 use of subordinating conjunctions within a single study. The present study analyzed three representative subordinating conjunctions, *if*, *when*, *and because* as used in a proficiency-stratified Korean EFL written corpus in terms of their frequency of occurrence, grammatical accuracy, and linear ordering relative to the main clause. The results revealed that the proficiency effect on frequency and grammatical accuracy varied among the three subordinating conjunctions, whereas that on linear ordering of subordinate clauses was uniform across conjunctions. The findings provide implications for a better understanding of L2 adverbial subordinate clauses and its application to teaching practice.

1. Introduction

Subordinate conjunctions such as *if*, *when*, and *because* express different semantic relations between the clauses they combine to make a complex sentence as in *I like apples because they are crunchy*. Subordinate clauses introduced by these conjunctions serve as adverbials adding some semantic information to the content of the main clauses in terms of condition, time, and reason or cause. Subordinate adverbial clauses also serve discourse-level functions by providing natural links between their main clauses and the previous or following sentences in the text (Larsen-Freeman and Celce-Murcia). These syntactic, semantic, and discourse roles of subordinate conjunctions have attracted much attention from researchers in second language (L2) acquisition and pedagogy, especially in regard to L2 writing proficiency development (Crossely, Crossely and McNamara, Y-J Kim, Kwon, Oh, Yoon and Yoo).

Three surface properties of adverbial subordinate clauses identified in L2 English learners' writing have been highlighted as quantitative measures of L2 writing performance. One property is their frequency of occurrence, and it has been extensively investigated as an indicator of the degree of syntactic complexity and cohesiveness of the sentences in a text, with syntactic complexity and cohesiveness in turn taken to reflect overall writing proficiency and development (Crossely and McNamara, Lorenz, Taguchi et al., Y-J Kim, Oh). Findings of the relevant research are mixed yet, with some studies showing a negative correlation (Lorenz, Taguchi et

al.) and others a null or weak positive correlation between the frequency of adverbial subordinate clauses and writing quality (Y-J Kim, Oh).

Another property of subordinate conjunctions as used by L2 learners concerns non-native-like usage patterns (Lorenz, Shin et al., Yoon and Yoo,). It has been well attested that low proficiency L2 writers often show grammatical or punctuational deviations from the native writing norm, using subordinate conjunctions as conjunctive adverbials (e.g., *I like apples.* Because*, they are crunchy.*) or in fragment sentences without a main clause (e.g., *I like apples*. Because they are crunchy.).

A third property of adverbial subordinate clauses of interest to L2 writing researchers is their linear position relative to the main clause. In English, they can either precede or follow the main clause. Multiple studies show that L2 writers tend to position adverbial subordinate clauses in front of their main clause to a greater extent than native English writers do (Y-J Kim, Kwon, Lorenz).

Although the previous research mentioned above provides valuable data for understanding L2 writing proficiency development, two limitations motivated the present study. First, most of the prior studies treated adverbial subordinate clauses as a single category or as a small part of the multitude of syntactic complexity and cohesion measures, leaving under-explored potential differences between individual subordinate conjunctions. Second, there seem to have been more studies focusing on characterization of L2 clausal subordination as compared with first language (L1) writing than the studies focusing on developmental patterns within L2 writing. By addressing these two limitations, the present study aims to contribute to a further understanding of the L2 use and development of English adverbial subordinate clauses.

2. Literature Review

The frequency of subordinate conjunctions or adverbial subordinate clauses has been widely studied as an index of syntactic complexity and cohesiveness of the sentences in the text produced by L2 learners, which in turn have been hypothesized to reflect overall writing quality and its development (See Crossely for a comprehensive review). Previous research suggests that clausal subordination and subordinate conjunctions are relatively weaker or less reliable predictors of writing quality when considered together with lexical sophistication, phrasal complexity (e.g., length and complexity of noun phrases), and/or global cohesion as indexed by overlapping lexical items across paragraphs (Crosseley and McNamara, Kyle and Crossely). The similar patterns were also found with Korean EFL learners (J. Kim, Lee). These findings, however, do not necessarily mean that the development of clausal subordination is independent of writing proficiency. Since the findings were mainly based on statistical procedures to identify variables that uniquely predict writing quality among a large set of diverse linguistic features, the role of clausal subordination, if any, could have been overridden by other correlated linguistic features such as lexical and structural diversity and sophistication.

Studies with a specific focus on adverbial subordinate clauses in L2 writing indeed suggest that there may be some meaningful relationship between clausal subordination and writing proficiency development, although mixed findings suggest a complicated relationship. Some studies suggest that higher-proficiency L2 writers use adverbial subordinate clauses *less* frequently than lower-proficiency writers (Lorenz, Taguchi et al.). Taguchi et al. conclude from these findings that increased clausal subordination may not necessarily be a reliable indicator of higher writing proficiency in L2 academic writing (426). This result is consistent with the findings that in L1 English

writing, clausal subordination is more characteristic of speech than of formal written language (Biber et al.) and better writers use fewer finite subordinate clauses (Myhill). L2 writers in Lorenz and Taguchi et al. thus seem to be learning and using more features of written register as proficiency develops.

However, the negative correlation between the frequency of adverbial subordinate clauses and writing proficiency level does not seem to be generalized to different populations of L2 writers, because studies on Korean EFL learners' argumentative essays show an opposite trend (Y-J Kim, Oh). Y-J Kim, for example, shows that higher-proficiency writers use *more* subordinate conjunctions than lower-proficiency writers. Oh's study also shows that higher proficiency writers use *because* more frequently than lower-proficiency writers. The overall finding may suggest a non-linear relationship between proficiency and the frequency of clausal subordination such that clausal subordination initially increases and then decreases later as L2 writers move toward the native norm (Norris and Ortega). More empirical research is needed to clarify the proficiency effects on clausal subordination in L2 writing.

Grammatical and punctuational errors that L2 writers often make in using subordinate conjunctions have also received attention from researchers (Y-J. Kim, S. Kim, Lorenz, Shin et al., Yoon and Yoo). Subordinate conjunctions connect syntactically dependent adverbial clauses to a main clause, making a complex sentence. It has been shown that Korean EFL writers often misuse subordinate conjunctions, especially *because*, as conjunctive adverbials, which do not syntactically subordinate a clause. This misuse is indicated by a comma after the sentence-initial subordinating conjunction as shown in (1). Sentence fragments are also frequent in Korean EFL learner writing, where a subordinate conjunction introduces an independent sentence as in (2).

(1) However, today's students don't want physical activities. **Because,** facility of playgrounds is too old. (Yoon and Yoo 226)

(2) If we could change one important think about my home town, I would change "Communication." **Because** my hometown's communication is poor. (Yoon and Yoo 236)

Yoon and Yoo attribute these errors to L1 influence as *because* is often translated to a conjunctive adverbial *waynya-hamyen* in Korean (e.g., *Na-nun sakwa-ka joha. Waynya-hamyen asakasak-haki ttaymuwniya.* 'I like apples because they are crunchy.'). This view is consistent with the fact that the other most frequent subordinating conjunctions identified in Yoon and Yoo's study (i.e., *if* and *when*), which are rarely translated to conjunctive adverbials in Korean, were never used in sentence fragments in the study. Another possibility is that L2 learners are not sensitive enough to the stylistic differences between spoken and formal written language. In L1 English, sentence fragments introduced by subordinate conjunctions are common in casual conversation unlike in written register (Biber et al.). Therefore, the frequent use of sentence fragments by L2 learners might result from the generalization of their knowledge about spoken register to written register (Lorenz). Although it is plausible to expect that grammatical errors in the use of subordinate conjunctions will gradually drop out as L2 writing proficiency increases, there seems to be little empirical research on when and how these grammatical errors are overcome along proficiency development.

The linear order of adverbial subordinate clauses relative to their main clause has also been studied in L1 and L2 English writing. In English, adverbial subordinate clauses can either precede or follow their main clauses, although the likelihood of pre- or postposing them varies across different conjunctions. Diessel's analysis of L1

English corpora shows that the conjunctions expressing condition (e.g., *if*) tend to precede main clauses while those expressing cause/result (e.g., *because*) tend to follow main clauses, with temporal subordinate clauses (e.g., *when*) featuring in between. Diessel attributes these ordering tendencies to the inherent semantic roles of the different types of subordinate clauses on the one hand and discourse-organizing motivations on the other hand. Conditional clauses provide the conceptual frame in which the content of the main clause is interpreted. Therefore, they should precede the main clause for the latter to be interpreted as is intended by the speaker/writer. Temporal clauses do not affect the interpretation of the main clause in the way that conditional clauses do. Instead, they either provide temporal background for the subsequent main clause at initial position or modify the meaning of the main clause at final position, with their position determined by discourse organization needs at the moment. Causal clauses, in contrast, rarely serve discourse-organizing functions at initial position but in most cases add new information which typically occurs toward the end of a sentence (Gundel).

In Kwon's study on Korean EFL learners' argumentative essays, the same positioning trend is identified in the use of subordinate conjunctions *if*, *when*, and *because*, which represent conditional, temporal, and casual adverbial clauses, respectively. The finding suggests that the semantic and discoursal motivations behind the linear ordering of English subordinate clauses may be universal and accessible to L2 learners whose L1 allows preposed adverbial clauses only like Korean. At the same time, Kwon shows that Korean EFL writers generally prefer sentence-initial over sentence-final position for subordinate clauses, suggesting a possible L1 transfer. Y-J. Kim further shows proficiency effects on adverbial clause ordering such that higher-proficiency Korean EFL writers produce *less* sentence-initial subordinate clauses than lower-proficiency writers, approaching the native-like ordering pattern.

These findings are yet incomplete, however, because Kwon did not consider proficiency effects and Y-J. Kim did not distinguish between different subordinate conjunctions.

As reviewed above, although the use of English adverbial subordinate clauses has been extensively studied in L2 writing research from multiple perspectives, relatively few studies comprehensively addressed the diverse aspects of the usage pattern and development of individual subordinating conjunctions. The present study aims to fill this research gap by closely investigating adverbial clauses of three semantic types as represented by *if*, *when*, and *because* found in argumentative essays written by Korean EFL learners. The findings will be able to contribute to a better understanding of the development of L2 writing proficiency as well as to more informed teaching practice.

3. Method

3.1. Data

In the current study, a Korean EFL learner corpus and a native English corpus were analyzed. The Korean learner corpus consisted of argumentative essays drawn from the Yonsei English Learner Corpus (YELC) (Rhee and Jung). It is a large collection of writings submitted by Korean high school graduates admitted to Yonsei University in 2011 as part of a timed placement test. Two types of essays are included in the YELC: descriptive and argumentative. Only argumentative essays were analyzed in the current study to render the current findings comparable to the previous findings which were mostly based on argumentative essays (Y-J Kim, Kwon, Lorenz, Oh,

Taguchi et al.). To explore proficiency effects, the Korean learner corpus was divided into three sub-corpora using the 9-point grade scales provided in the YELC, which were assigned to each essay based on the Common European Framework of Reference for Language (Council of Europe). Since the number of essays at each grade level varied widely, sub-corpora were constructed so that the number of essays in each sub-corpus was maximally comparable. The three sub-corpora were each labeled HP (High Proficiency: B1+ and above), IP (Intermediate Proficiency: B1), and LP (Low Proficiency: A2 and below).

The native English corpus was drawn from the Lovain Corpus of Native Speaker English (LOCNESS), which consists of argumentative and literary essays written by British and American university students. Only American students' argumentative essays were used for analysis and labeled as the NE (Native English) corpus. The NE corpus served as the native norm regarding the use of *if-*, *when-* and *because-*adverbial clauses. Note that argumentative essays in the YELC and those in the LOCNESS are different in various aspects including grade levels of writers, length of text, and writing conditions (e.g., timed vs. untimed). These factors have been shown to affect many aspects of written production including syntactic complexity and cohesiveness (Crosseley). Therefore, the LOCNESS does not represent the ideal native norm for the writings in the YELC. The NE corpus thus should be taken as an approximate native norm at best. But this limitation may not undermine the main findings of the current study because its focus was on the relationship between Korean learners' use of subordinate clauses and their proficiency levels.

Table 1 presents the basic information about the characteristics of the texts in the three Korean sub-corpora and the NE corpus. Number of words and sentences was counted using the SiNLP, a computational tool for automatic extraction of linguistic features from text (Crossley et al.).

⟨Table 1: Characteristics of the Corpora.⟩

	LP	IP	HP	NE
Number of texts	910	1,173	1,203	175
Number of word tokens	156,964	275,430	335,508	149,979
Number of sentences	13,410	21,030	22,785	8,270
Number of words per text	172	235	279	857
Number of sentences per text	14.7	17.9	18.9	47.3
Number of words per sentence	11.7	13.1	14.7	18.1

Table 1 shows that the Korean writers used a much smaller number of words and sentences in a text as compared with the native writers. This large difference may be due to the different nature of writing tasks between the two corpora (e.g., timed vs. untimed writing) as well as proficiency difference. Within the Korean writers, who wrote the texts under the same condition, advancement in writing proficiency was associated with a greater number of words per text, sentences per text, and words per sentence.

3.2. Data Analysis

For data retrieval, the AntConc concordance tool was used (Anthony). As an attempt to ensure the representativeness of the three target subordinate conjunctions (*becuase, if, when*), a full list of subordinating conjunctions presented in Larsen-Freeman and Celece-Murcia (543) were extracted from the four corpora under analysis. Simple frequency counts of subordinate conjunctions without excluding non-adverbial uses (e.g., complement clauses and wh-questions) suggested that *if*, *when*, and *because* are the most frequent subordinators accounting for about 70% of all the queried items in the leaner corpora and about 64% in the NE corpus. This may justify, to a certain

extent, the selection of the three subordinating conjunctions as the representatives of three semantic types of subordinate clauses (i.e., conditional, temporal, and causal).

In the first stage of data analysis, all the relevant instances of the three subordinators were extracted with non-adverbial uses excluded through manual inspection. Because the number of the identified concordance lines reached several thousands for each subordinator, disambiguation was performed based on a few heuristic guidelines focusing on surface features such as punctuations. First, sentence-initial subordinators immediately following a sentence-ending punctuation or a sentence-initial adverb were included for the analysis whether or not they were followed by a comma. Second, sentence-middle subordinating conjunctions following a main clause ending with a comma were included. Third, for the subordinators used in the middle of a sentence without a comma in front of them, a potential verb that can take them as its complement was searched in the preceding context. When there was such a verb in the vicinity, the sentence was closely analyzed to determine the syntactic role of the subordinate clause. Non-adverbial subordinate clauses such as clausal complements were excluded at this step. In the case of *if*, concessive clauses were also excluded because the current study concerns its conditional use. Examples of the excluded tokens of subordinators are shown in (3)-(5).

(3) They don't care **if** they lie or cheat the customer ...

(4) On driving, who knows **when** accident happens?

(5) This is **because** two years of youth is so much important ...

The finiteness of the subordinate clauses and ungrammatical features within sentences were not considered at this step since they were irrelevant for the present purposes. The resulting list of the target subordinate clauses thus included finite,

non-finite, and verb-less clauses (e.g., *when talking about our culture*, *when necessary*) as well as subordinating conjunctions used as adverbial conjunctions and in fragment sentences. This list was used to compare the frequency of each subordinator across the corpora.

In the next stage, adverbial conjunctions and fragment sentences were identified, mainly based on surface punctuation patterns. First, when sentence-initial subordinators (as identified by a sentence-ending punctuation before them) were followed by a comma, they were coded as adverbial conjunctions (e.g., *Because, animals are important being.*). Second, when the sentence-initial subordinators were not followed by a comma and there was no main clause, those were coded as sentence fragments (e.g., *Because my school has counseling room.*). Lastly, all the remaining instances of subordinators were coded as correctly used subordinating conjunctions. The list resulting from this stage of coding was analyzed to capture the distribution of deviant uses of subordinate conjunctions across different proficiency levels in the learner corpora as compared to the NE corpora.

The final stage of data analysis concerned the ordering of adverbial clauses relative to their main clause. In this stage, only finite adverbial clauses were included to make the results comparable to the earlier studies on the ordering pattern of adverbial clauses in L1 English (Diessel) and Korean EFL writing (Kwon). First, conjunctive adverbials and fragment sentences were excluded from the list of subordinate conjunctions identified in the previous stage. Second, the relative order of adverbial clauses was coded (e.g., preposed: Because *the students are very young, they don't know what they did. postposed: They will just follow the rules because they have to.*) Third, the finiteness of each adverbial clause was manually checked to exclude non-finite and verb-less subordinate clauses (e.g., *Because school not be allowed physical punishment ..., Because only mistake, ...*). Lastly, the remaining

preposed and postposed adverbial clauses were tallied to be compared across corpora.

Manual inspection of several thousands of concordance lines is subject to coding errors, so it is a convention that a second corder checks the reliability of coding in this kind of study. However, reliability checking was deemed to be dispensable in this study on the basis of two considerations. First, coding was mainly based on objective surface features such as punctuation marks, the presence of finite verbs, and surface order. This should have minimized subjective human interpretation, which is often the source of unreliable coding. Second, the size of the corpora under current analysis was deemed to be large enough to faithfully reflect the general patterns of interest despite some coding errors. As shown in the following section, the patterns of usage and development of the three subordinating conjunctions across corpora seemed to be differentiating enough to override potential noises from coding errors.

4. Results and Discussion

4.1 Frequency of Subordinate Conjunctions across Corpora

In this section, the frequency of occurrence of the target subordinating conjunctions are compared across corpora to explore whether and how proficiency effects manifest for each of the conjunctions. A usual convention in quantitative corpus research is to normalize token frequency to the total number of word tokens in a corpus when comparing corpora of different sizes (Hunston). As Bolton rightly points out, however, normalization based on word tokens might introduce some unfounded bias in interpreting the frequency of special lexical items such as subordinating conjunctions. Suppose we compare two texts; text 1 contains one complex sentence

with 10 words and text 2 contains one complex sentence with 20 words. In this case, word-based normalization would lead to a conclusion that text 1 contains more subordinate clauses than text 2 (i.e., 1 vs. 0.5 subordinate clauses per 10 words), whereas sentence-based normalization would suggest that there is no difference between the two texts in that regard (i.e., 1 subordinate clause per sentence for both texts). Given that the number of words per sentence fast increases with proficiency advancement (see Table 1), word-based normalization might exaggerate low-proficiency writers' clausal subordination or conversely underestimate high-proficiency writers' performance. Considering this possibility, the present study presents both word-based and sentence-based normalized frequency.

The frequency of all extracted subordinating conjunctions is shown in Table 2. Figure 1 visualizes the changes in normalized frequencies across corpora per 1,000 words and per 100 sentences, respectively.

⟨Table 2: Frequency of Subordinating Conjunctions (per 1,000 words/100 sentences).⟩

	LP	IP	HP	NE
If	1,141(7.3/8.5)	1,792(6.5/8.5)	1,897(5.7/8.3)	474(3.2/5.7)
When	607(3.9/4.5)	1,079(3.9/5.1)	1,271(3.8/5.6)	346(2.3/4.2)
Because	737(4.7/5.5)	1,093(4.0/5.2)	1,240(3.7/5.4)	375(2.5/4.5)

⟨Figure 1. The Normalized Frequency of Subordinating Conjunctions.⟩
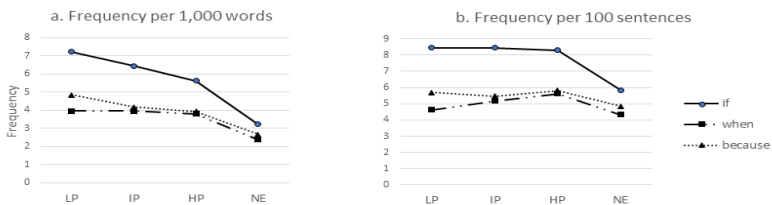
As shown in Table 2, the frequency ranks of the three conjunctions are the same across the four corpora, with *if* standing out with much higher frequency over the other two. This may be partly because there are more alternative lexical and structural means to express causal and temporal relations as compared to conditional relations. Or it may also partly be due to the characteristics of argumentative writing, in which conditionals are widely used to describe conditions or hypothetical situations where a particular argument could better be supported or countered. Another finding to note is the overall overuse of the subordinating conjunctions in the learner corpora as compared with the NE corpus, which is consistent with widely attested overuse of limited set of lexical items by L2 learners (Hesselgren).

Of greater interest for the present purposes is the change in frequency across proficiency levels in the Korena corpora. According to word-based normalization, *if* and *because* show a decreasing pattern whereas *when* is used at a more or less similar frequency across proficiency levels. Sentence-based normalization shows that the use of *if* and *because* is rather constant across proficiency levels, while *when* is used increasingly more often as proficiency advances. The increasing use of *when* with proficiency development manifest in sentence-based normalization could be attributed to its semantic property. Adding temporal information would mainly contribute to the elaboration of arguments, whereas giving reason or setting conditions are integral part of argumentation. Under this view, high proficiency learners seem to be achieving the ability to elaborate their argumentation moving beyond simply providing the necessary reasons or conditions.

Another notable pattern of the findings concerns how we should interpret clausal complexity as a measure of writing proficiency. As was reviewed in the literature review section, clausal complexity has been shown to be a relatively weak predictor of writing quality and development (Kyle and Crossely). The current findings suggest

that reasons behind the weak relationship between clasual complexity and writing proficiency may have to do with such factors as the prevalent subordinate conjunctions in a particular corpus and the writing proficiency level represented by that corpus. If the distribution of subordinate conjunctions in a particular corpus resembled that shown in Figure 1a, a negative correlation would be drawn between clausal subordination and writing proficiency mainly due to the overriding effect of the most frequent and fast decreasing use of *if*. In contrast, a distributional pattern similar to that shown in Figure 1b would lead to a null or weak positive correlation due to different directions of frequency change across proficiency levels on the one hand and the consistently increasing use of *when* on the other hand. One lesson that can be drawn from this finding is that the relationship between clausal subordination and writing proficiency should be interpreted in combination with specific subordinating conjunctions prevalent in a particular text.

Table 3 presents the proportion of deviant and native-like uses of the three subordinating conjunctions. The overall pattern is self-evident, in which *because* is most subject to deviant uses, whereas *if* and *when* are rarely used as conjunctive adverbials or in fragment sentences. This verifies the well established finding of error-proneness of *because* (Shin et al., Yoon and Yoo). A novel finding is the fast decrease of deviant uses of *because* with proficiency development, which is graphically displayed in Figure 2.

〈Table 3: The Percentage of Deviant and Native-like Uses of *if*, *when*, and *because* (Raw Frequency).〉

	LP	IP	HP	NE
If				
Conjunctive adverbials	0.2(2)	0(0)	0.1(2)	0(0)
Fragments	1.5(17)	1.3(24)	0.7(14)	0.6(3)
Subordinators	98.3(1122)	98.7(1768)	99.2(1881)	99.4(471)
Total	100(1141)	100(1792)	100(1897)	100(474)
When				
Conjunctive adverbials	0.2(1)	0(0)	0(0)	0(0)
Fragments	1.7(10)	1.0(11)	0.4(5)	0.6(2)
Subordinators	98.2(596)	98.9(1068)	99.6(1266)	99.4(344)
Total	100(607)	100(1079)	100(1271)	100(346)
Because				
Conjunctive adverbials	3.5(26)	1.6(17)	1.3(16)	0(0)
Fragments	49.8(367)	37.2(407)	16.1(200)	2.4(9)
Subordinators	46.7(344)	61.2(669)	82.6(1024)	97.6(366)
Total	100(737)	100(1093)	100(1240)	100(375)

〈Figure 2. The Percentage of Deviant and Native-like Uses of *because*〉
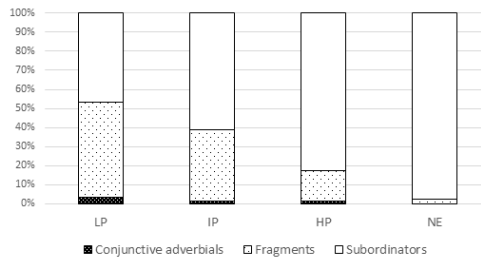
As can be seen in Figure 2, the erroneous use of *because* as a conjunctive adverbial is rare even from low proficiency. In contrast, its use in sentence fragments is quite frequent in the LP, accounting for about half of all the instances of *because*. As proficiency advances, however, sentence fragments decrease fast, accounting for

about 16.1% of the *because*-clauses in the HP. The result suggests that the two deviant usage patterns of *because* might not be as problematic for developing learners as their saliency may suggest. Given that the writers of the texts under analysis were high school graduates who might not had received systematic instruction on academic writing, they seem to have learned the grammatical properties of *because* as well as its register-sensitivity through input. We may expect that they could learn, without much difficulty, to use *because* in native-like ways with appropriate instruction on its grammar and genre-related constraints.

The last aspect of the use of adverbial subordinate clauses of current interest is their linear ordering relative to the main clause. Table 4 and Figure 3 show the proportion of preposed adverbial clauses for each of the three subordinating conjunctions across the corpora. The distribution of postposed clauses is not presented because it is complementary to that of preposed clauses.

〈Table 4: The Percentage of Preposed Adverbial Clauses〉

	LP	IP	HP	NE
If				
# of all finite clauses	1,060	1,707	1,822	454
# of preposed clauses	990	1,522	1,578	347
% of preposed clauses	93.4	90.5	86.6	76.4
When				
# of all finite clauses	525	974	1160	287
# of preposed clauses	349	606	625	152
% of preposed clauses	66.5	62.2	53.9	53.0
Because				
# of all finite clauses	323	648	1012	355
# of preposed clauses	74	112	123	26
% of preposed clauses	23.0	17.3	12.2	7.3

〈Figure 3. The Percentage of Preposed Adverbial Clauses〉
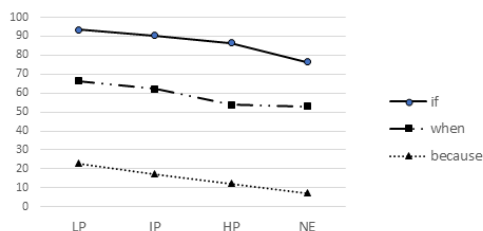
Different ordering preferences of the three types of subordinate clauses show the same pattern in the four corpora, wherein *if* prefers initial position and *because* final position with *when* in between. This pattern closely replicates the findings in L1 English (Diessel) and Korean EFL writing (Kwon), suggesting that the semantic and discourse demands on adverbial clause placement may be universal across L1 and L2 and robust enough to exert an consistent effect along proficiency levels.

A new finding is the uniform effect of proficiency on the change in positioning preference across the three subordinating conjunctions, with the likelihood of preposing an adverbial subordinate clause decreasing as proficiency advances. This finding is interesting because as presented earlier, the proficiency-associated changes in overall frequency and deviant uses in the Korean learner corpora show a different pattern for each of the three subordinating conjunctions. To recapitulate, the overall frequency change across proficiency levels is most salient with *if* when counted in word-based normalized frequency, while sentence-based frequency normalization highlights the increasing frequency of *when*. As for deviant use, only *because* is problematic while grammatical errors rarely occurr with *if* and *when*. This suggests that ordering patterns may be the more robust indicator for the development of adverbial subordinate clauses in L2 writing as compared with their frequency and

grammatical accuracy.

5. Conclusions and Suggestions for Future Research

The present corpus-based study conducted an in-depth quantitative analysis of the usage pattern of English adverbial subordinate clauses in a large corpus of Korean EFL learner writings, with focus on three representative subordinating conjunctions, *if*, *when*, and *because*.

The results based on the overall frequency of subordinating conjunctions suggest that Korean EFL learners learn to use them quite early and heavily rely on them to mark relevant semantic relations. Frequency-based measures were also shown to have weak or complex relationships with proficiency development, with the developmental patterns differing across individual subordinating conjunctions.

Deviant uses of subordinating conjunctions as adverbial conjunctions or in fragment sentences were shown to be almost exclusively restricted to *because*. This suggests that grammatical accuracy in using subrodinators develops differentially across individual conjunctions, indicating that instructional effort could be focused on problematic ones.

Lastly, the linear ordering preferences of different types of adverbial subordinate clauses were shown to be consistent with the earlier findings in L1 and L2 (Diessel, Kwon) and progressively approaching the native norm as proficiency advanced. This finding suggests that the ordering of adverbial subordinate clauses relative to their main clause could be a rich source of information about L2 writing proficiency development and the way in which L2 writers of different proficiency levels arrange clauses to organize their ideas in writing. In pedagogy, it will be beneficial to raise

learners' awareness of the target-like subordinate clause ordering patterns.

This study leaves several questions open to future research, two of which are discussed here. First, the current results are based on only three subordinating conjunctions although they are the most frequent ones and each represent different semantic relations typically marked by subordinating conjunctions. Because some aspects of clausal subordination may develop in different patterns for different conjunctions as was shown in the current study, investigation of other subordinating conjunctions will be able to offer more fruitful implications for L2 acquisition research and teaching practice. Second, due to the strictly quantitative approach, the current findings about the ordering pattern of adverbial clauses could not address the role of discourse factors affecting ordering decisions made by L2 writers. An investigation into the extent to which discourse-organizing motivations determine ordering choice for L2 writers at different proficiency levels may contribute valuable insights into the development of cohesiveness and coherence of L2 writing.

Works Cited

Anthony, Laurence. "AntConc" (Version 4.0.5) [Computer Software]. Tokyo, Japan: Waseda University. Available from https://www.laurenceanthony.net/software

Biber, Douglas, et al. "Should We Use Characteristics of Conversation to Measure Grammatical Complexity in L2 Writing Development?" *TESOL Quarterly*, vol. 45, no. 1, 2001, pp. 5-35.

Council of Europe (CEFR). *Common European Framework of Reference for Languages: Learning, Teaching, Assessment.* Cambridge University Press, 2001.

Crossely, Scott A. "Linguistic Features in Writing Quality and Development: An Overview." *Journal of Writing Resarch*, vol. 11, no. 3, 2020, pp. 415-43.

Crossely, Scott A, et al. "Analyzing Discourse Processing Using a Simple Natural Language Processing Tool." *Discourse Processes*, vol. 51, no. 5-6, 2014, pp. 511-34.

Crossely, Scott A, and Danielle S. MacNamara. "Predicting Second Language Writing Proficiency: The Roles of Cohesion and Linguistic Sophistication." *Journal of Research in Reading*, vol. 35, no. 2, 2012, pp. 115-35.

Diessel, Holger. "Competing Motivations for the Ordering of Main and Adverbial Clauses." *Linguistics*, vol. 43, no. 3, 2005, pp. 449-70.

Gundel, Jeanette K. "Universals of topic-comment structure." *Studies in Syntactic Typology*, edited by Michael Hammond et al., John Benjamins, 1988, pp. 209-39.

Hesselgren, Angela. "Lexical Teddy Bears and Advanced Learners: A Study into the Ways Norwegian Students Cope with English Vocabulary." *International Journal of Applied Linguistics*, vol. 4, no. 2, 1994, pp. 237-58.

Hunston, Susan. *Corpora in Applied Linguistics*. Cambridge University Press, 2002.

Kim, Jungyeon. "The Use of Cohesive Device in Korean EFL Writing across Different Proficiency Levels." *Korean Journal of English Language and Literature*, vol 22, 2022, pp. 1078-100.

Kim, Sugene. "Revisiting Causes of Grammar Errors: Commonly Confused Lexical Category Dyads in Korean EFL Student Writing." *English Teaching*, vol. 71, no. 2, 2016, pp. 61-85.

Kim, Yoon-Jung. "A Study on the Use of Conjunctions by Korean English Learners." *English Language Teaching*, vol. 25, no. 2, 2013, pp. 47-62.

Kwon, Soo-Ok. "A Corpus-Based Analysis of the Ordering of Main and Adverbial Clauses in Korean Learners of English." *Korean Journal of Applied Linguistics*, vol. 24, no. 1, pp. 159-76.

Kyle, Kristopher, and Scott A. Crosseley. "Measuring Syntactic Complexity in L2 Writing Using Fine-Grained Clausal and Phrasal Indices." The Modern Language Journal, vol. 102, no. 2, 2018, pp. 333-49.

Larson-Freeman, Diane, and Marianne Celce-Murcia. *The Grammar Book*. Heinle ELT, 2015.

Lee, Young-Ju. "The Relationship between English Proficiency and Syntactic Complexity for Korean College Students." *The Journal of the Convergence on Culture Technology*, vol. 7, no. 3, 2021, pp. 439-44.

Lorenz, Gunter. "Learning to Cohere: Causal Links in Native vs. Non-Native Argumentative Writing." *Coherence in Spoken and Written Discourse: How to Create It and How to Describe It. Selected Papers from the International Workshop on Coherence, Augsburg, 24-27 April 1997*, edited by Wolfram Bublitz et al., John Benjamins, 1999, pp. 55-75.

Myhill, Debra. "Towards a Linguistic Model of Sentence Development in Writing."

Language and Education, vol. 22, no. 5, 2008, pp. 271-88.

Norris, John M., and Lourdes Ortega. "Towards and Organic Approach to Investigating CAF in Instructed SLA: The Case of Complexity." *Applied Linguistics*, vol. 30, no. 4, 2009, pp. 555-578.

Oh, Sun-Young. "Functions of *Because* in Native English and Korean Learner Corpora." *Korean Journal of Applied Linguistics*, vol. 25, no. 3, 2009, pp. 443-70.

Rhee, Seok-Chae, and Chae Kwan Jung. "Compilation of the Yonsei English Learner Corpus (YELC) 2011 and Its Use for Understanding Current Usage of English by Korena Pre-University Students." *The Journal of the Korea Contents Association*, vol. 14, no. 1, 2014, pp. 1019-29.

Shin, Yu Kyung, et al. "English Conjunctives Adjuncts Used by Native and Nonnative Freshman University Students." *Modern English Education*, vol. 20, no. 4, 2019, pp. 54-62.

Taguchi, Naoko, et al. "What Linguistic Features Are Indicative of Writing Quality? A Case of Argumentative Essays in a College Composition Program." *TESOL Quarterly*, vol. 17, no. 2, 2013, pp. 420-30.

Yoon, Jung Wan, and Isaiah W. Yoo. "An Error Analysis of English Conjunctive Adjuncts in Korean College Students' Writing." *English Teaching*, vol. 66, no. 1, 2011, pp. 225-44.

국문초록

한국인 영어 학습자의 종속접속사 사용에 대한
코퍼스 기반 연구: *if, when, because*를 중심으로

백 순 도

단독 / 국민대학교

제2언어 작문 연구에서 종속접속사 사용은 다양한 관점에서 연구되었는데, 통사 복잡도 지표로서의 사용빈도, 문법 적합성, 종속절과 주절의 상대적인 어순 등에 대해 집중적인 연구가 이루어졌다. 그러나 종속접속사 사용의 다양한 측면을 종합적으로 고찰한 단일 연구가 부족하여 이전 연구 결과들의 관계에 대한 이해와 개별 종속접속사 사용에 대한 구체적인 정보는 부족한 실정이다. 이 연구는 제2언어 작문에 나타나는 종속접속사의 사용과 발달에 대한 종합적인 이해를 위해 대규모 한국인 영어학습자 작문 코퍼스를 이용해 대표적인 종속접속사인 *if, when, because*를 사용빈도, 문법적 정확성, 주절-종속절 순서 등의 측면에서 세밀하게 관찰하였다. 이 연구를 통해 사용빈도나 문법적 정확성은 능숙도 발달에 따라 변화하는 양상이 개별 접속사에 따라 다르지만, 주절-종속절 순서 결정은 능숙도에 따른 변화가 모든 종속접속사에서 동일한 양상을 보인다는 것을 알 수 있었다. 이러한 발견에서 제2언어 습득의 이론적 이해와 효과적인 제2언어 교수에 유용한 함의를 찾을 수 있다.

주제어: 한국인 영어 학습자, 논증적 글쓰기, 작문 능숙도, 종속접속사, 종속절 순서

이름: 백순도
소속: 국민대학교 영어영문학부 부교수
이메일: sdbaek@kookmin.ac.kr