https://doi.org/10.24303/lakdoi.2023.31.3.195

Exploring the Use of *-ese* and *-nikka* by English-Speaking Learners on the Basis of Corpus-Based Learner Data

Sun Hee Hwang
(Dongguk University)

Hwang, Sun Hee. (2023). Exploring the use of *-ese* and *-nikka* by English-speaking learners on the basis of corpus-based learner data. *The Lintuistic Association of Korea Journal, 31*(3), 195-215. This study investigated how English-speaking learners of Korean use two causal connectives, *-ese* and *-nikka*, on the basis of comparison between the non-native speaker (NNS) corpus and native speaker (NS) corpus, and whether there is a relationship between learners' proficiency level and error rate. A comparison analysis was conducted on the Korean Learners' Corpus (KLC), and a native speaker corpus, the Sejong Corpus. It has been known that English causal connective *because* can be used in all three of Sweetser's (1990) content, epistemic, and speech act domains, but Korean tends to distinguish the use of causal connectives *-ese* and *-nikka* depending on those causal relations. The results indicate that native speakers perform differently in the choice between *-ese* and *-nikka* depending on causal relations, while L2 learners' performance do not reach significance level. Especially for the epistemic causality, unlike native speakers, English-speaking learners use both *-ese* and *-nikka* at a similar rate. With respect to the error rate according to level of proficiency, the results reveal that the error rate decreases as proficiency level increases. This fact could imply that the complete acquisition of causal connectives is ultimately possible in principle despite L1-L2 different combinations of features into lexical items.

Key Words: causal connectives, content domain, epistemic domain, epistemic domain, corpus-based study

1. Introduction

Korean has a range of ways to express causal relations between two events, but two primary causal connectives *-ese* and *-nikka* have received much attention by researchers (Hwang, 2008; Kwon, 2012; Lukoff & Nam, 1982, among others). The reason would be because these two connectives share some distributions, but they show certain semantic differences. The constructions with *-ese* and *-nikka* show the same structure (Cause + [Connective] + Consequence) (Kwon, 2012) and these connectives seem to be used interchangeably in some cases as shown in the following examples.

(1) a. param-i pwul-ese/pwu-nikka siwonhay-yo.
wind-NOM blow-because cool-DECL
'It is cool because wind blows.'
b. kongpwu-lul yelsimhi hay-se/ha-nikka sengcek-i ola-ss-eyo.
study-ACC hard do-because grade-NOM improve-PAST-DECL
'(My) grades improved because (I) studied hard.'

However, in terms of semantics, it has been known that the use of *-ese* is based on the common sense, while the use of *-nikka* is based on the speakers' inference. In addition, the clause after the connective *-nikka* can appear in the form of imperatives and exhortatives, but the use of *-ese* is not compatible with them (e.g., Hwang, 2008; Park, 2015).

(2) a. erum-i nok-ass-unikka/*-ase nayngcangko-ka kocangnan-key
ice-NOM melt-because refrigerator-NOM break-NML
thullimep-ta.
must-DECL
'The refrigerator must be broken because the ice melted.'
b. nayil sihem-i iss-unikka/*ese kongpwu yelsimhi hay/ha-ca.
tomorrow exam-NOM exist-because study hard do-IMP/do-EXH
'Study hard/Let's study because (you) have an exam tomorrow.'

In (2a) where the speaker's belief or judgment is involved, the use of *-nikka* is only allowed. In other words, melted ice causes the speaker to conclude that the refrigerator

must be broken. In (2b), the use of -*ese* is also infelicitous presumably because only -*nikka* clauses can indicate the causal relation for the main clause speech act (Park, 2015).

In English, the causal connective *because* can represent all these causal relationships (Sweetser, 1990). In terms of structure with because, the clause representing consequence precedes the one with cause unlike Korean forward causal construction.

(3) a. John came back, because he loved her.
b. John loved her, because he came back.
c. What are you doing tonight, because there's a good movie on.
(Sweetser, 1990, as cited in Kwon, 2012, p.4)

In terms of L2 acquisition, despite these apparent differences between Korean and English, the acquisition of causal connectives has not yet been much examined. To the best of my knowledge, many previous acquisition studies on the difference between -*ese* and -*nikka* have focused on analyzing the errors made by Chinese-speaking learners (e.g., Jiang, 2020; Lim, 2009). Thus, this study aims to examine the causal constructions produced by native English speakers, comparing with Korean native speakers' data. Since the English causal connective *because* has the ability to operate in Sweetser's (1990) three causal relations (content, epistemic, and speech act), it would be interesting to investigate whether English-speaking learners of Korean can distinguish the use of -*ese* and -*nikka*.

2. Causal connectives in Korean and English

2.1. The differences between -*ese* and -*nikka*

The Korean connectives -*ese*/-*ase*[1)] and -*nikka* are used to indicate the causal relations between two propositions, one event causing another event and the other one induced by the first event, showing that their use can be interchangeable.

Hwang (2004, 2008) shows overall frequencies of causal connectives from the Sejong Corpus. According to the findings, among the causal connectives, -*ese* (raw frequency: 40,485) is identified as the most occurring connective, and -*nikka* (raw frequency: 11,053) is

1) -*Ase* is the allomorph of -*ese*, which occurs after certain vowels such as [a] and [o]. -*Ese* is used as the representative form unless two forms are needed to be indicated.

the second most one in the corpus. Park (2020) also utilized a large corpus consisting of around 16,930,000 eojeols[2], and presents frequencies of *-ese* and *-nikka* as shown in Table 1.

Table 1. Overall Frequencies of *-ese* and *-nikka*

	-ese	*-nikka*
Raw frequency	136,446	28,062
per 10,000 eojeols[3]	80.56	16.57

(Park, 2020)

Along with the results of other previous studies, the results in Table 1 show that the frequency of *-ese* is much higher than the one of *-nikka*. It should be noted that the frequencies include the number of cases encoding other meanings such as preceding events and manners for *-ese* and discovery for *-nikka*.[4] The examples of those meanings are presented in the following sentences.

(4) a. <preceding event>
yengswu-nun mikwuk-ey ka-se umak-ul kongpwuhay-ss-ta.
yengswu-TOP USA-LOC go-and music-ACC study-PAST-DECL
'Yengswu went to the USA and studied music.'
b. <manner>
yelsoy-lul kwumeng-ey ne-ese tolli-ess-ta.
key-ACC hole-in put-and turn around-PAST-DECL
'(I) put the key in the hole and turned it around.'

(5) <discovery>
pak-ey naka po-nikka cengmal tewu-ess-ta.
outside-LOC go out see-when really hot-PAST-DECL
'It was really hot when (I) went outside.'

Park (2020) randomly selected 1200 tokens for each connective to extract only

2) The corpus used in Park (2020) is a collection of data from various sources such as the Sejong Corpus and transcripts of spoken language.
3) An eojeol is a linguistic unit seperated by white spaces. The Sejong corpus and KLC use a segmentation scheme based on eojeols.
4) The subcategories of *-ese* and *-nikka* are presented in Park (2020).

occurrences of causal relations, and found that the use of *-ese* accounts for 38.17% and the use of *-nikka* for 73.17%. His findings suggest that *-nikka* is used in the causal relation in most cases, while *-ese* is used with several meanings.

Based on the previous findings from native speakers' corpora, it should be noted that *-ese* is used much more frequently than *-nikka*. In case of *-ese*, the instances expressing other meanings such as preceding events and manners were found in a considerable rate, while *-nikka* is predominantly used in the meaning of causality.

Sweetser (1990) suggests three domains that the interpretation of causal relations depends on content, epistemic, and speech act domains. Some previous research has applied these three distinctions to the distributions of Korean causal connectives *-ese* and *-nikka* (Kwon, 2012; Park, 2015).

(6) John-i sumathuphon-ul mani sayongha-nikka/-ase nwun-i
John-NOM smartphone-ACC much use-because eyes-NOM
nappaci-ess-eyo.
get worse-PAST-DECL
'John's eyes got worse because he used his smartphone a lot.'

Either *-ese* or *-nikka* can be used to link an objective cause and consequence as in (6) where the content of the main clause—John's eyesight getting worse—is the event caused by his excessive use of a smartphone. The causal connective in (6) operates in the content domain. However, if we reverse the order of two clauses, *-ese* is not allowed as in (7).

(7) nwun-i nappaci-ess-unikka/*ese John-i sumathuphon-ul mani
eyes-NOM get worse-PAST-because John-NOM smartphone-ACC a lot
sayongha-n key thullimup-ta.
use-NML must-DECL
'John must have used his smartphone a lot because his eyes got worse.'

In (7) the clause with *-nikka* expressing 'eyesight deteriotation' works as a basis for the speaker's subjective judgment that John used his smartphone a lot. In other words, it is the speaker's belief or conclusion that John used his smartphone a lot that is caused by the observation or knowledge that John's eyesight got worse. Another distributional difference in the two clausal connectives is related to the types of main clauses (Park,

2015). As shown in (8), only *-nikka* clause is compatible with imperatives (8a), exhortatives (8b), and promissives (8c).

(8) a. nwun-i nappaci-nikka/*ese sumathuphon-ul mani sayongha-ci
eyes-NOM get worse-because smartphone-ACC much use-NML
masey-yo.
NEG-IMP
'Don't use your smartphone too much because your eyes will get worse.'
b. nwun-i nappaci-nikka/*ese wuli sumathuphon-ul sayongha-ci
eyes-NOM get worse-because we smartphone-ACC use-NML
mal-ca.
NEG-EXH
'Let's not use our smarphones because our eyes are getting worse.'
c. sumathuphon-ul sayong an hay-ss-unikka senmwul-ul cwu-ma.
smartphone-ACC use NEG do-PAST-because gift-ACC give-PROM
'(I) will give you a gift because (you) didn't use your smartphone.'

As seen in above examples, the contexts where *-ese* can occur are different from the ones of *-nikka*. Based on the different distributions, Sohn (1993) argues that *-nikka* can appear in all three domains, while the use of *-ese* is limited to the content domain. However, there are some examples where the use of *-nikka* is infelicitous as presented below.

(9) a. os-ul sa-le ka-ss-nundey pissa-se/*nikka kunyang
clothes-ACC buy-to go-PAST-and expensive-because just
wa-ss-eyo.
come-PAST-DECL
'I went to buy clothes, but just came because it was expensive.'
b. chinkwu-tul-i nol-le wa-se/*nikka swukcey-lul mos
friend-PL-NOM come to play-because homework-ACC NEG
hay-ss-eyo.
do-PAST-DECL
'(I) couldn't do my homework because my friends came to play.'

In (9a), the second clause 'not buying clothes' is caused by the *-ese* clause 'being expensive' indicating the assertion of a real-world cause. Thus, it can be suggested that the use of *-ese* is more felicitous to express a direct reason which causes the natural or logical consequence (Ree, 1977).

Based on these kinds of examples, Kwon (2012) claims that the use of *-ese* is more felicitous in the content domain, while *-nikka* is only allowed in the epistemic and speech act domains as in the following examples.

(10) a. khephi-lul masi-ese cam-i o-cianh-ney-yo. (content)
coffee-ACC drink-because sleep-NOM come-NEG-FACTIVE-DECL
'(I) can't sleep because (I) drank some coffee.'
(Kwon, 2012, p. 8)

b. pyengwon-eyse cinryo pat-ass-unikka John-i
hospital-LOC treatment receive-PAST-because John-NOM
aphun-key thillimep-ta. (epistemic)
sick-NML must-DECL
'John must be sick because he saw a doctor at the hospital.'

c. aphu-nikka ellun pyengwon-ey ka. (speech act)
sick-because quickly hospital-LOC go-IMP
'Go to the hospital quickly because you are sick.'

In more recent studies, these three domains have been reclassified into the distinction of subjectivity and objectivity in terms of speaker involvement (Pander Maat & Degand, 2001; Pander Maat & Sanders, 2000, 2001). This approach suggests that the content domain related to cause-consequence is objective, while epistemic and speech act domains where the speaker is clearly involved are subjective (Sanders & Spooren, 2015). Based on this, *-nikka* can be naturally used to express epistemic and speech act domains and thus considered more subjective than *-ese*. On the other hand, *-ese* related to the content domain is more objective because the speaker is not involved. Furthermore, in the content domain volitional and non-volitional relations can be distinguished (Stukker et al., 2008), and thus the four causal relations can be identified from the least subjective one to the most subjective one (Sanders & Spooren, 2015). The order of four relations is given below.

(11) Non-volitional content < Volitional content < Epistemic/Speech Act

Table 2 presents the primary characteristics of *-ese* and *-nikka* based on the discussions so far.

Table 2. Primary characteristics of *-ese* and *-nikka*

	-ese	*-nikka*
Domain	content	epistemic, speech act
Types of causal relations	non-volitional, volitional least subjective	speaker's subjective judgment, belief, or assertion

To summarize so far, the use of *-ese* is closely associated with objective causality and content domain, while *-nikka* predominantly operates in epistemic and speech act domains, even though its use is allowed in the content domain in some cases.

2.2. Causal Connectives in English

In English, the connective *because* can operate in all these three domains (Sweetser, 1990), while Korean uses two different causal connectives (i.e. *-ese* and *-nikka*) to represent them.

(12) a. John went to the store because he needs to buy milk.
b. John went to the store, because his car's not here.
c. Let's take John to the mall, because he needs to get new clothes.
(Hong, 2012, p. 13)

In (12a), the clause after *because* describes an event which causes the event described in the main clause. In other words, the reason why John went to the store is his need to buy milk, showing that *because* is used in the content domain. On the other hand, the *because*-clause in (12b) provides the speaker's evidence for the subjective judgment which he makes in the main clause. This use is related to the epistemic domain. In addition, the *because*-clause provides a cause or reason for an exhortative speech act in the main clause. In (12c), the speaker proposes the listener to take John to the mall, presenting the reason for the act by using *because*. Thus, the *because*-clause is ambiguous because it can be used on these three different types of causation (Sweetser, 1990).

It seems that Korean has more differentiated items of causal connectives to indicate causality than English, where *because* can mark all the three domains. In terms of

subjectivity which is present in the nature of relations between events, causal links in the content domain are considered objective, whereas the epistemic and speech act are subjective because the speaker is actively reasoning towards the conclusion or motivating that act on the basis of an observation (Sanders & Spooren, 2015). As shown in the following examples, both types of links can be expressed with a single connective *because* in English.

(13) a. The fields are wet because it has rained a lot this week.
b. Surely all soccer games will be cancelled, because it has rained a lot this week.
(Sanders & Spooren, 2015, p. 54)

Based on what we have discussed so far, the differences in causal connectives between Korean and English can be characterized as follows. Table 3 exhibits the different properties among the lexical items using features.

Table 3. Feature-based differences among three causal connectives

Relevant features for causality	Lexical items	
[objective]		
[content]	*-ese*/(*-nikka*)	
[subjective]		because
[epistemic] [speech act]	*-nikka*	

Turning to L2 acquisition, it is possible to characterize the learning tasks for L1 English-L2 Korean based on the crosslinguistic differences. The feature-reassembly approach proposed by Lardiere (2009) suggests that the different featural configurations of L2 lexical items from their L1 counterparts operate as the source of L2 learning problem. Accordingly, at the initial state of acquisition, it is necessary for English-speaking learners to be aware of the presence of two different lexical items in Korean, and the fact that one causal connective cannot express all three types of causal relations. Then, they are required to split the relevant features and acquire which features are relevant to which causal connective. These developmental predictions based on the feature-reassembly approach could reflect proficiency levels. In other words, if it is ultimately possible to

attain native-like proficiency as argued by the feature-reassembly approach, their performance would improve as their proficiency increases.

2.3. Research Questions

In this study, Korean causal connectives are examined based on the data from two corpora. This study aims to address the following research questions.

1) What differences exist between English-speaking learners and Korean native speakers in terms of use of *-ese* and *-nikka*?
2) Is there a relationship between the proficiency level and the error rate regarding the use of *-ese* and *-nikka*?

3. Methodology

3.1. Corpora

This study used learner data from the Korean Learner Corpus by National Institute of Korean Language (henceforth referred to as the KLC), which contains both spoken and written data. The morph-tagged version of KLC is comprised of 4,013,233 eojeols (1,253,148 eojeols for spoken data; 2,760,085 eojeols for written data), while the error-annotated corpus 1,342,301 eojeols from 7,715 samples (672,718 eojeols for written data; 669,583 eojeols for spoken data). The data to examine the causal connectives within the corpus were obtained from the homepage (https://kcorpus.korean.go.kr/index/goMain.do). On the homepage, learners' attributes such as native language, age and proficiency level can be identified. In this study, the data only from English-speaking learners were selected, and analyzed according to their proficiency levels.[5)]

For comparisons between learners' and native speakers' data, the Sejong Corpus collected by the 21st century Sejong Project was used. Since the Sejong Corpus is one of the largest Korean corpora, considering all instances of two connectives is not available. Thus, 200 occurrences of each connective were randomly selected for analysis.

---

5) In the KLC, there are six proficiency levels for English-speaking learners with level 6 as the highest and level 1 as the lowest. In Section 4.2., these six levels were reclassified into three levels: beginner, intermediate, and advanced.

3.2. Data Collection Procedure

This study used both morph-tagged version and error-annotated versions of KLC. The morph-tagged corpus was used to obtain frequencies of two causal connectives. In the case of the morph-tagged corpus, the causal connectives are tagged as EC (connective ending), and thus we can place a restriction on EC in order to obtain only *-ese* and *-nikka* as connectives. For *-ese*, *-ese* and its allomorph *-ase* were used as a search term one by one, and frequencies of each connective were obtained. In the same way, the frequency of *-nikka* was obtained.

It should be noted that the frequencies for both causal connectives can include instances in which they do not encode causality. For *-ese*, those are the instances where two conjoined events are sequential, and cases where the *-ese* clauses express purpose. In addition, for *-nikka*, some instances can include cases in which the *-nikka* clauses express discovery. In order to exclude those meanings, it is necessary to manually classify them. The remaining data were coded for analysis based on the coding scheme (Lemen et al., 2021). The coding scheme used in this study including examples are presented below.

(14) a. Content: The clause with a connective expresses a "real-world" cause by explaining the specific cause for the event or state in the main clause.
e.g. nalssi-ka tewu-ese eyekhen-ul khye-ss-ta
weather-NOM hot-because air conditioner-ACC turn on-PAST-DECL
'(I) turned on the AC because it was hot.'

b. Epistemic: The clause with a connective gives an account of the speaker's conclusion expressed in the main clause.
e.g. pang-ey menci-ka mani ssaye iss-unikka chengso-lul
room-LOC dust-NOM a lot gather-because cleaning-ACC
oray an han-key pwunmyengha-ta.
long time NEG do-NML certain-DECL
'(I) am sure that (you) haven't cleaned the room for a long time because there is a lot of dust in the room.'

c. Speech act: The clause with a connective provides an explanation of a speech act (e.g. asking/suggesting/ordering) performed in the main clause.
e.g. os-i terewu-nikka pes-usey-yo.
clothes-NOM dirty-because take off-HON-IMP
'Take off your clothes, because your clothes are dirty.'

The error-annotation generally indicates the process of assigning a tag which identifies and classifies the errors made by L2 learners according to an error-tagging system. In the case of the error-annotated corpus of KLC, the positions where the errors were made are categorized into two types: content words and function words. In the case of causal connectives, thus, they are categorized into FED (connectives) under the category of function words. There are two kinds of search conditions we can use: original error form and correct form. For this study, the first step was to put *-ese*/*-ase* and *-nikka* as original error forms. Second, those connectives were also put as correct forms for errors. The purpose of this process was to extract data which only contain the errors related to those connectives.

4. Results

4.1. Frequencies of Causal Connectives

This section concerns the relative proportion of two causal connectives. Table 4 shows the overall frequencies of *-ese* and *-nikka* from the Sejong Corpus and KLC's morph-tagged corpus. In order to examine the differences between two corpora, Log-likelihood (*LL*)[6] was used.

Table 4. Overall frequencies of *-ese* and *-nikka*

Connectives	KLC (4,013,233)[7]	Sejong (8,058,034)[8]	
	Frequency/%1[9]	Frequency/%2	*LL*
-ese	5026/0.13	102505/1.27-	***53321.91
-nikka	424/0.01	7158/0.09-	***3450.48

****p* < 0.001

6) Log-likelihood (*LL*) was used by the help of a log-likelihood calculator available online (http://ucrel.lancs.ac.kr/llwizard.html).

7) The information about the total number of eojeols of the morph-tagged KLC is from (https://kcorpus.korean.go.kr/service/goMorphAnnotStatistics.do).

8) The information about the total number of eojeols of the morph-tagged Sejong Corpus is from E. Lee (2013).

9) According To The Homepage Of Log-Likelihood Calculator (Http://Ucrel.Lancs.Ac.Kr/Llwizard.Html), "%1 And %2 Values Show Relative Frequencies In The Text."

The (-) in relative frequency indicates underused in the KLC compared to the Sejong Corpus. For both *-ese* and *-nikka*, thus, native speakers tend to use them far more than L2 learners.

Since the frequencies in Table 4 include the cases of *-ese* and *-nikka* encoding other meanings than causality, 200 occurrences of each connective were randomly selected. Then, the cases were manually examined and only instances encoding causal relations were chosen. Table 5 represents the frequencies of causal relations from the Sejong Corpus and KLC. It should be noted that there are some cases where the meanings are not clearly distinct, and thus different results could be made depending on intuition. Moreover, unlike the data from native speakers, there are a few incomplete sentences which are not interpretable, and thus those cases without main clauses were excluded.

Table 5. Frequencies of causal relations from Sejong Corpus and KLC

	-ese /*-ase*		*-nikka*		Total
	Causal	Other	Causal	Other	
Sejong	61 (30.5%)	139 (69.5%)	102 (51%)	98 (49%)	400
KLC	104 (52%)	96 (48%)	106 (53%)	94 (47%)	400

In order to determine whether there is a significant difference between two corpora, a chi-square test was carried out by using R statistics program.[10] For *-ese*, there was a significant different between two corpora (x^2 = 18.197, df=1, $p < 0.01$), while a significant different was not found in *-nikka* (x^2 = 0.090144, df=1, $p > 0.05$). Compared with native speakers, it could be suggested that learners tend to use *-ese* more in the meaning of causal relations, indicating that their use of *-ese* does not vary. It could be because the learners have not acquired all meanings of *-ese* yet.

In addition, these instances encoding only causality were classified into three types of causal relations: content, epistemic, and speech act. Table 6 presents frequencies of connectives according to types of causal relations from the Sejong Corpus.

10) According to the homepage (https://www.r-project.org/about.html), R is a free software program which provides a wide range of statistical techniques. In this study, R version 4.3.1. is used for analysis.

Table 6. Type of relation by connective from Sejong Corpus

Connective	Causal Relations			Total
	Content	Epistemic	Speech act	
-ese/-ase	56	4	1	61
-nikka	48	30	24	102
total	104	34	25	163

Table 6 shows differences in frequencies between the two Korean causal connectives in terms of type of relations: the connective *-ese* is predominantly used in the content causal relation, while the connective *-nikka* shows some preference to the epistemic causal relation even though it is still frequently used in the content domain. As discussed in Section 2, *-nikka* can be used in the content domain, but its predominant use by native speakers is somewhat unexpected because the epistemic and speech act causal relations of *-nikka* have been emphasized in the previous literature.

A chi-square test was computed by using the R statistics program to determine whether the choice of connectives is associated with causal relations, and the association was statistically significant (x^2=33.462, df=2, $p < .001$). This result from native speakers' data indicates that the types of causal relations affect the choice of a connective. Table 7 presents frequencies of each connective according to types of causal relations from the KLC.

Table 7. Type of relation by connective from KLC

Connective	Causal Relations			Total
	Content	Epistemic	Speech act	
-ese/-ase	84	16	4	104
-nikka	75	14	17	106
total	159	30	21	210

As shown in Table 7, unlike the data from native speakers, the data from learners indicate that both *-ese* and *-nikka* are predominantly used in content causal relations. A chi-square test was used, and the difference between two connectives was not statistically significant (x^2=5.8926, df=2, $p > .05$), indicating that there was no relationship between the choice of a connective and the types of causal relations.

In order to identify which causal relation influences the different result, chi-square tests were carried out for each domain. For content and speech act domains, the differences between the Sejong Corpus and KLC were not significant (x^2=0.0012305, *df*=1, $p > .05$ for content domain; x^2=1.3404, *df*=1, $p > .05$). For the epistemic domain, on the other hand, the difference between two corpora was significant (x^2=10.956, *df*=1, $p < .01$). As a result, in the corpus data of native speakers, *-nikka* is much more used in the epistemic domain than *-ese*, while *-ese* and *-nikka* occur at a similar rate in the epistemic domain in the learners' data.

It should be noted that coding the data was not based on objective scores but the researcher's subjective judgment. Even though the coding scheme was used presented in (14), it was difficult to determine the types of causal relations for some sentences, and thus the choices made by the researcher could be different from other researchers. In order to redeem this issue, the KLC's error annotated corpus was utilized in the next section.

4.2. Analysis of Error Patterns According to Learners' Proficiency Levels

This section analyzes the errors related to causal connectives from the error-annotated corpus of KLC. First, we consider the instances with the incorrect use of *-ese* and *-nikka* in which other connectives should have been used.

(15) a. ...oyrowu-l ttay hayngpokha-l ttay umak-ul tul-*ese(-umyen)
..lonely-when happy-when music-ACC listen-*because(if)
maum-i phyenhayci-nta. (level 3)
mind-NOM comfortable-DECL
'When I am lonely and happy, listening to music makes me comfortable.'

b. ce-nun cinan cwumal-ey chinkwu-lul manna-*nikka(-ryeko)
I-TOP last weekend-in friend-ACC meet-*because(in order to)
cihachel-ul tha-ss-eyo. (level 1)
subway-ACC take-PAST-DECL
'I took the subway to meet my friends last weekend.'

This section also concerns those instances with errors in which the causal connectives under examination (i.e. *-ese* and *-nikka*) should have been used despite the occurrence of other connectives. A couple of examples are given below.

(16) a. onul chinkwu-hako syophing-ul hako siph-*untey(-ese)
today friend-with shopping-ACC do want-*and(because)
paykhwacem-ey ka-yo. (level 2)
department store-to go-DECL
'I will go to the department store today because I want to go shopping with my friend.'

b. taum cwumal-ey chinkwu kyelhonsik-ey ka-*ryeko(-nikka)
next weekend-in friend wedding-to go-because
cengcang-ul saya hay. (level 2)
suit-ACC buy-have to.
'I have to buy a suit because I go to my friend's wedding next weekend.'

Table 8 shows the frequencies of two kinds of errors mentioned above. It should be noted that the frequencies can include cases in which *-ese/-ase* and *-nikka* express other meanings rather than causality. Moreover, the extracted errors can contain misspellings and verb conjugation errors. Those errors are included for analysis in this section because excluding some cases from the data can influence the results.

Table 8. frequencies of two kinds of errors

Connectives	*-ese/-ase* or *-nikka* → other connective	other connective → *-ese/-ase* or *-nikka*	Total
-ese/-ase	264	425	689
-nikka	58	48	106
total	322	473	795

Table 9 exhibits the frequencies of errors and norm frequencies per one million eojeols according to proficiency levels. The six proficiency levels provided in the KLC were reclassified into three levels: beginner (level 1 and 2), intermediate (level 3 and 4), and advanced (level 5 and 6). Since the data size from each proficiency level is different from each other, normalized frequencies are presented along with the raw frequencies.

Table 9. Error frequencies per one million eojeols according to proficiency

Proficiency	-*ese*/-*ase*		-*nikka*		Total	
	Raw F	Norm F	Raw F	Norm F	Raw F	Norm F
Beginner (82,134)	313	3,810.8457	56	681.8126	369	4,492.6583
Intermediate (76,532)	256	3,345.006	44	574.9229	300	3,919.9289
Advanced (54,370)	120	2,207.0995	6	110.355	126	2,317.4545
Total (213,036)	689	3234.1952	106	497.5685	795	3731.7636

As shown in Table 9, the number of errors decreases as proficiency level increases. In order to determine whether those differences among groups are significantly different, Log-Likelihood (*LL*) was carried out by making comparisons between the beginner level and the rest two proficiency levels one by one. The statistical results are presented in Table 10.

Table 10. Log-likelihood results between beginner and each of the other levels

Connectives		Proficiency levels		
		Beginner	Intermediate	Advanced
-*ese*/-*ase*	*F*	313	256	120
	LL		+2.40	+27.80***
-*nikka*	*F*	56	44	6
	LL		+0.72	+28.52***
total	*F*	369	300	126
	LL		+3.09	+45.28***

*** $p < .001$

In Table 10, the (+) symbol in LL indicates an overuse of errors at the beginner level compared to one of the other proficiency groups. On the other hand, the (-) in LL indicates an underuse of errors at the beginner level relative to the other proficiency groups being compared. In other words, since Table 10 presents only (+) marks, we can say that the learners in the lowest proficiency makes more errors than the higher-proficiency learners

when they produce two causal connectives. Based on the fact that the frequency of errors decreases as proficiency level increases, it can be said that L2 learners gradually acquire the meanings and functions of each causal connective, implying that the complete acquisition is ultimately possible.

5. Discussions and Conclusions

This study investigated how English-speaking learners of Korean use two causal connectives, *-ese* and *-nikka*, on the basis of comparison between the NNS corpus and NS corpus, and whether there is a relationship between learners' proficiency level and error rate.

The two connectives, *-ese* and *-nikka*, have in common that they encode causal relationships between two events. However, they show different distributions depending on types of causal relations. In the choice of causal connectives between *-ese* and *-nikka*, native speakers perform differently depending on causal relations that the causal connectives express, while English-speaking learners of Korean do not show significant differences. The difference between the NS and NNS corpora was not significant for the distribution of *-ese* (x^2=3.642, df=2, $p > .05$)[11], while the difference was significant for the distribution of *-nikka* (x^2=12.868, df=2, $p < .05$). Regarding the order of acquisition, based on the result, it can be suggested that English-speaking learners acquire *-ese* first, and then *-nikka*.

Especially for expressing the epistemic causality, native speakers predominantly use *-nikka*, while L2 learners use both *-ese* and *-nikka* at a similar rate. That could be attributed to the difficulty of acquiring epistemic modality. Previous L2 studies as well as L1 studies have shown that learning expressions of epistemic modality is not easy for L2 learners (e.g. Hyland & Milton, 1997). Evers-Vermeul and Sanders (2011) argue that expressing epistemic relations requires more abstract knowledge than other relations in the content and speech-act domains. Furthermore, not only the abstractness but also the complexity of syntactic constructions could be associated with the difficulty that English-speaking learners have. Unlike English which does not require the use of explicit modal expressions, Korean requires a combination of a main verb and an overt epistemic modal expression

11) This statistical result from the chi-square test should be treated with caution, because there are three cells with expected counts less than 5.

such as *pwunmyengha-* 'certain' or *thullimeps-* 'must' on the surface in order for *-nikka* to express the epistemic causality in most cases (Park, 2015). Thus, they need to acquire how to construct the predicates of epistemic modality, and be aware of the fact that only *-nikka* is compatible with the circumstance expressing the epistemic causality. Accordingly, the correct use of *-nikka* in the epistemic domain should be a complicated process, and as a result, English-speaking learners tend to use both *-ese* and *-nikka* in the epistemic relation without the correct domain distinctions.

However, it does not mean that successful acquisition is not possible. The fact that the error rates decrease as proficiency level increases could imply that the complete acquisition of causal connectives requires considerable proficiency. According to the feature-reassembly approach (Lardiere, 2009), since the causal connectives *-ese* and *-nikka* are different from English counterpart *because* in terms of the division and assembly of features, the distinction between *-ese* and *-nikka* would be difficult for English-speaking learners. Even though it takes some time to successfully accomplish the feature-reassembly process, it should be ultimately possible in principle (Lardiere, 2009).

References

Evers-Vermeul, J., & Sanders, T. (2011). Discovering domains – On the acquisition of causal connectives. *Journal of Pragmatics 43*(6), 1645-1662.

Hong, J. (2012). *Connective polysemy and clause linkage typology in Korean.* Unpublished doctoral dissertation, University of California at Berkeley.

Hwang, H. (2004). *Hankuke Censan Hyengthayron*. Seoul: Worin Publishing.

Hwang, H. (2008). Yenkyelemy '-ese, -nikka' uy uymy kinungkwa hwuhayngcel yuhyeng ('The semantic function of conjunctive ending '-eoseo, -nikka' in Korean'). *Kukehak 51*, 57-89.

Hyland, K. & Milton, J. (1997). Qualification and certainty in L1 and L2 students' writing. *Journal of Second Language Writing, 6*(2), 183-205.

Jiang, Q. (2020). Hankwuke inkwakwankye yenkyelemi '-ese'wa '-nikka'uy sayong mich olyu yangsang yenkwu: cwungkwukekwen haksupca-lul taysangulo ('A research for the usage and error patterns in Korean conjunctive ending '-ese' and '-nikka': Targeting at Chinese Korean learners'). Unpublished master's thesis. Yonsei University.

Lim, J. (2009). Hankwuke haksupca-uy yenkyelemi '-ase/ese'ey tayhan olyu yenkwu (A research for the Korean learner's error in the connective ending '-ase/ese'). Unpublished Master's thesis. Chongju University.

Kwon, I. (2012). Forward causality in basic communicative spaces networks: the –ese and -nikka constructions in Korean. *Discourse and Cognition, 19*(1), 1–27.

Lardiere, D. (2009). Some thoughts on the contrastive analysis of features in second language acquisition. *Second Language Research 25*(2), 173-227.

Lee, E-R. (2013). 'Kyelkuk'uy uymywa hwayongcek kinung ('Semantic and pragmatic functions of 'kyeolguk'). *Hanmincokemwunhakhoy 64*, 5-36.

Lemen, H., Lieven, E., & Theakston, A. (2021). A comparison of the pragmatic patterns in the spontaneous because- and if-sentences produced by children and their caregivers. *Journal of Pragmatics 185*, 15-34.

Lukoff, F., & Nam, K. (1982). Constructions in -nikka and -ese as logical formulations. In The Linguistic Society of Korea (Ed.), *Linguistics in the morning calm: Selected papers from SICOL-1981* (pp. 559–583). Seoul: Hanshin.

Pander Maat, H., & Degang, L. (2001). Scaling causal relations and connectives in terms of speaker involvement. *Cognitive Linguistics 12*(3), 211–245.

Pander Maat, H., & Sanders, T. (2000). Domains of use or subjectivity? The distribution of three Dutch causal connectives explained. In E. Couper- Kuhlen & B. Kortmann (Eds.), *Cause, condition, concession and contrast: Cognitive and discourse perspectives* (pp. 57–81). Berlin & New York: Mouton de Gruyter.

Pander Maat, H., & Sanders, T. (2001). Subjectivity in causal connectives: An empirical study of language in use. *Cognitive Linguistics 12*(3), 247–273.

Park, J. (2020). *A study on causal connective ending '-eseo' and '-nikka'.* Unpublished doctoral dissertation, Seoul National University.

Park, Y. (2015). *A grammar of mood and clausal adjunction in Korean.* Unpublished doctoral dissertation. University of Delaware.

Ree, J. (1977). A reanalysis of (u)ni and (e)se. In Kim, Chin-W. (Ed.), Papers in Korean linguistics: *Proceedings of the symposium on Korean linguistics* (pp. 177–184). Columbia: Hornbeam Press.

Sanders, T. J., & Spooren, W. P. (2015). Causality and subjectivity in discourse: The meaning and use of causal connectives in spontaneous conversation, chat interactions and written text. *Linguistics, 53*(1), 53-92.

Sohn, S.-O. (1993). Cognition, affect, and topicality of the causal particle -nikka in

Korean. In Clancey, Patricia M. (Ed.), *Japanese/Korean linguistics, vol. 2* (pp. 82-97). Stanford: CSLI Publications.

Stukker, N., Sanders, T., & Verhagen, A. (2008). Causality in verbs and in discourse connectives. Converging evidence of cross-level parallels in Dutch linguistic categorization. *Journal of Pragmatics, 40*(7), 1296-1322.

Sweetser, E. 1990. *From etymology to pragmatics.* Cambridge: Cambridge University Press.

Sun Hee Hwang
Lecturer
Department of English Language and Literature, Dongguk University
30 Phildongro 1-gil, Jung-gu, Seoul 04620, Korea
Email: sunhee.hwang38@gmail.com

Received on September 1, 2023
Revised version received on September 30, 2023
Accepted on September 30, 2023