

Readers' Language Experience in Generating Korean *Wh*-Constructions*

On-Soon Lee

(Dong-A University)

Lee, On-Soon. (2019). Readers' language experience in generating Korean *wh*-constructions. *The Linguistic Association of Korea Journal*, 27(3), 149–171. This study investigates the role of language experience, which is shaped by the distributional patterns occurring in input, in structural preferences in language production. In order to accomplish this purpose, a corpus analysis and a sentence completion experiment were conducted. Specifically, thirty-six Korean-speaking adults participated in the experiment, in which they read and completed sentence fragments including either a scrambled or an in-situ *wh*-phrase. When the participants generated questions, they attached the question-marking particle to a verb as soon as possible after encountering the *wh*-phrase, suggesting an active dependency formation mechanism. This finding supports the Active Filler Strategy hypothesis, a major account for the processing of filler-gap dependencies, but does not support any effect of the readers' linguistic experience.

Key Words: filler-gap dependency, active filler strategy, language experience, *which*-question, sentence completion

1. Introduction

Previous studies on sentence processing have been interested in the mechanism for resolving the dependency relationship formed between a filler such as a fronted *wh*-phrase and its gap (its canonical position). For

* This work was supported by the research grant of Dong-A University in 2019.

example, in (1), the dependency formed between the filler, which is the displaced phrase *which student*, and its gap position must be resolved for comprehension, but this process is not always easy for readers during real-time sentence processing.

(1) Which student did John say that Mary read a book to __?

Much of this research has provided evidence for *active dependency formation* (Aoshima, Philips and Weinberg, 2004; Crain and Fodor, 1985; Frazier and Clifton, 1989; Pickering and Traxler, 2003; Staub, 2010; Stowe, 1986). In active dependency formation, the parser will search for a gap position immediately upon encountering a filler (i.e., *which student*). This is called the Active Filler Strategy (e.g., Frazier and Clifton, 1989). Following this view, many researchers have broadly agreed that readers resolve such dependencies through incremental sentence processing guided by active dependency formation.

On the other hand, an emerging body of literature on sentence processing has focused on the effect of readers' language experience in underlying processing mechanisms. A wide variety of relevant theories propose that an incremental processing mechanism is guided by probabilistic expectations for sentences, and such expectations are generated by the distributional patterns occurring in language input (Hale, 2001, 2003, 2006; Levy, 2008). This leads to the assumption that readers' experience of syntactic structures throughout their lives might play an influential role in the processing mechanism. Yet despite the strong research interest in the role of language experience in incremental sentence processing, there is little cross-linguistic research on the topic. In the Korean context, studies in this area are quite scarce (e.g., Yun et al., 2015), and a very few studies have investigated the role of language experience in experiments in which participants must generate *wh*-constructions rather than simply comprehend sentences (e.g., see Atkinson et al., 2016 for a review and discussion). This study, therefore, aims to examine the role of language experience in adult Korean speakers' generation of Korean *wh*-constructions that include filler-gap dependencies.

2. Background

2.1 Processing *Wh*-Constructions

The language processing literature offers two competitive accounts for how the processing mechanism functions during real-time sentence comprehension. One account argues that the parser engages in an active search to resolve dependencies encountered during processing. The revisited example in (2) shows a long-distance dependency between a fronted constituent, the filler (*which student*), and its canonical position, the gap (marked by the underscore), a so-called filler-gap dependency (Frazier, 1987).

(2) Which student does John think that Mary met ___?

Previous studies have demonstrated that the parser actively and immediately seeks the fronted *wh*-phrase's original position (Aoshima, Philips and Weinberg, 2004; Crain and Fodor, 1985; Frazier and Clifton, 1989; Frazier, 1987; Omaki and Schulz, 2011; Pickering and Traxler, 2003; Staub, 2010; Stowe, 1986; Wagers, Borja and Chung, 2015). In particular, Stowe's (1986) reading time analysis formed the basis for her proposal of a Filled Gap Effect. Using an eye-tracking paradigm, she found slower reading times at a direct object position (*us* in [3b]) right after a *wh*-phrase (*who*) relative to a direct object position (*us* in [3a]) that does not follow a *wh*-phrase.

- (3) a. My brother wanted to know if Ruth will bring us home to Mom
at Christmas.
b. My brother wanted to know who Ruth will bring us home to _ _
at Christmas.

The observation of this Filled Gap Effect – the slowdown – indicates that the parser begins actively searching for the gap as soon as the filler is

encountered.

This line of research has been extended to test this hypothesis in head-final languages like Japanese and Korean, *which*, although they canonically keep *wh*-phrases in-situ, also allow them to be scrambled.¹⁾ Japanese has an interesting property in forming *wh*-constructions: the scope of Japanese *wh*-constructions is determined by the position of the question-marking particle (whereas the scope of English *wh*-constructions is determined by the position of the *wh*-phrase). This case-marking system allows the *wh*-phrase to be scrambled. Aoshima et al. (2004) exploited these properties of Japanese in a word-by-word self-paced reading study that manipulated the position of the *wh*-phrase (i.e., in-situ vs. scrambled conditions) and the types of complementizer in the embedded clause verbs (declarative complementizer vs. question-marking particle), as shown in Table 1. In the two word-order conditions, the question particle *-ka* is attached to the embedded clause verb (i.e., *read* in [c] and [d]) to form an indirect question, or attached to the main clause verb (i.e., *said* in [a] and [b]) for a direct question.

1) I agree with one reviewer's concern that Korean *wh*-scrambling constructions are not an exact structural counterpart to English *wh*-constructions. This study adopted the term "filler-gap dependency" as it is broadly used in psycholinguistics for cross-linguistic comparisons. For the purposes of this study, filler and gap refer to the moved element (filler, *wh*-phrase) and its trace (gap) respectively, in Korean scrambled *wh*-constructions specifically, including the displacement of the *wh*-phrase (see Hahn & Hong, 2014 for discussion). The reviewer also pointed out the lack of a syntactic explanation for these Korean scrambled constructions. Such a discussion is beyond the scope of this study, but I agree that it should be taken up in future research.

1	2	3	4	5	6	7	8
<i>a. Scrambled, Declarative Complementizer</i>							
Dono-seito-ni	tannin-wa	koocyoo-ga	hon-o	yonda-to	tosyositu-de	sisyo-ni	iimasita-ka?
<i>which student-dat</i>	<i>class teacher-top</i>	<i>principal-nom</i>	<i>book-acc</i>	<i>read-DeclC</i>	<i>library-at</i>	<i>librarian-dat</i>	<i>said-Q</i>
<i>b. In situ, Declarative Complementizer</i>							
Tannin-wa	koocyoo-ga	dono-seito-ni	hon-o	yonda-to	tosyositu-de	sisyo-ni	iimasita-ka?
<i>class teacher-top</i>	<i>principal-nom</i>	<i>which student-dat</i>	<i>book-acc</i>	<i>read-DeclC</i>	<i>library-at</i>	<i>librarian-dat</i>	<i>said-Q</i>
‘Which student did the class teacher say to the librarian at the library that the principal read a book for?’							
<i>c. Scrambled, Question Particle</i>							
Dono-seito-ni	tannin-wa	koocyoo-ga	hon-o	yonda-ka	tosyositu-de	sisyo-ni	iimasita.
<i>which student-dat</i>	<i>class teacher-top</i>	<i>principal-nom</i>	<i>book-acc</i>	<i>read-Q</i>	<i>library-at</i>	<i>librarian-dat</i>	<i>said</i>
<i>d. In situ, Question Particle</i>							
Tannin-wa	koocyoo-ga	dono-seito-ni	hon-o	yonda-ka	tosyositu-de	sisyo-ni	iimasita.
<i>class teacher-top</i>	<i>principal-nom</i>	<i>which student-dat</i>	<i>book-acc</i>	<i>read-Q</i>	<i>library-at</i>	<i>librarian-dat</i>	<i>said</i>
‘The class teacher said to the librarian at the library which student the principal read a book for.’							

Figure 1. A sample set of Aoshima et al.’s (2004) experimental sentences²⁾
 (Adopted from Aoshima et al., 2004:30)

The Active Filler Strategy hypothesis would predict that the parser will immediately search for the gap position right after the *wh*-phrase *which student* is identified, with the result that a slowdown would appear at Region 5 (*read*). Aoshima et al.’s findings bear out this hypothesis: in both scrambled (a) and (c) and in-situ (b) and (d) conditions, they found slower reading times with the declarative complementizer (a–b) than with the question particle (c–d). They concluded that the readers expected a question-marking particle right after they encountered either a fronted *wh*-phrase or an in-situ *wh*-phrase, and were therefore surprised to see the declarative complementizer on the embedded verb in (a) and (b). The authors interpreted the study’s results as demonstrating that the Japanese readers immediately searched for the question particle once they had seen the *wh*-phrase, regardless of its position (i.e., in-situ *wh*-phrase vs. scrambled *wh*-phrase). The results indicate that the underlying processing mechanism for *wh*-constructions is guided by active dependency formation in Japanese just as it is in English.

Working in another head-final language, Korean, Hahn and Hong (2014) replicated Aoshima et al.’s (2004) study, with a minor revision of the experimental sentences (Figure 2): Instead of a *which*-phrase, they used *nuwkwu-eykey* ‘whom.’³⁾ In addition, Korean, which has several

2) acc=accusative; dat=dative; decl=declarative; declc=declarative complementizer; nom=nominative; top=topic marker; q=question-marking particle

question-marking particles, uses different form in different clause types (i.e., *-ni* with a main clause verb and *-nunci* with an embedded clause verb). Hahn and Hong found a relative slowdown at the declarative complementizer of the embedded clause verb (Region 5) in the in-situ conditions ([2] and [4] in Figure 2), but not at the declarative complementizer of the embedded clause verb in the scrambled conditions ([1] and [3] in Figure 2). Thus, their findings only partially confirmed the predictions of the Active Filler Strategy hypothesis. However, their use of the different lexical item (*who* instead of *which student*) in the experimental sentences may have led to the apparent difference in results between Japanese and Korea. This issue will be further discussed in section 2.3 and section 5. Therefore, to allow a more precise comparison, the current study uses exact Korean translations of the Japanese experimental sentences in Aoshima et al.'s (2004) study.

	Region							
Conditions	1	2	3	4	5	6	7	8
1 Scrambled/ DeclComp	nwkwu- eykey <i>who-dat</i>	tamim- un <i>teacher- top</i>	[kyocang- i <i>principal- nom</i>	chayk-ul <i>book-acc</i>	ilkecuwess- tako] <i>read- DeclC</i>	<i>yesterday</i>	sase- eykey <i>librarian- dat</i>	malhayss- ni? <i>said-Q</i>
2 In-situ/ DeclComp	tamim- un <i>teacher- top</i>	[kyocang- i <i>principal- nom</i>	nwkwu- eykey <i>who-dat</i>	chayk-ul <i>book-acc</i>	ilkecuwess- tako] <i>read- DeclC</i>	<i>yesterday</i>	sase- eykey <i>librarian- dat</i>	malhayss- ni? <i>said-Q</i>
3 Scrambled/ Q-particle	nwkwu- eykey <i>who-dat</i>	tamim- un <i>teacher- top</i>	[kyocang- i <i>principal- nom</i>	chayk-ul <i>book-acc</i>	ilkecuwess- nunci] <i>read-Q</i>	<i>yesterday</i>	sase- eykey <i>librarian- dat</i>	malhayss- ta <i>said-Decl</i>
4 In-situ/ Q-particle	tamim- un <i>teacher- top</i>	[kyocang- i <i>principal- nom</i>	nwkwu- eykey <i>who-dat</i>	chayk-ul <i>book-acc</i>	ilkecuwess- nunci] <i>read-Q</i>	<i>yesterday</i>	sase- eykey <i>librarian- dat</i>	malhayss- ta <i>said-Decl</i>

Figure 2. A sample set of Hahn and Hong's (2014) experimental sentences (Taken from Hahn and Hong, 2014:70)

Taken together, the results of the existing research thus far suggest that readers' active dependency formation facilitates the resolution of filler-gap

-
- 3) Korean *wh*-phrases can often be interpreted as either indefinites or interrogatives. For example, a translation of Aoshima et al.'s (2004) 'which student' would be *enu haksayng-eykey*, but this can mean either 'to which student' or 'to a certain student' (Sohn, 1994). For this reason, Hahn and Hong (2014) chose *nwkwu-eykey* 'to whom', after finding in a preliminary study that *nwkwu* 'who' shows a strong tendency to be interpreted as the interrogative.

dependencies in real-time sentence processing in English and Japanese but not in Korean. This inconsistency raises a question of why English and Japanese, but not Korean, would share the same processing mechanism. A possible explanation may be provided by a different account, which is that readers' probabilistic predictions, based on their language experience, determine their preferences for structure.

2.2 Language experience

Despite the great interest in incremental sentence processing, only a few studies have examined how readers' experience of the syntactic structures that occur in the input might be tied to their processing (e.g., Hale, 2001; Levy, 2008). If input teaches the parser to expect certain structures, then readers' language experience should affect how they tend to resolve dependencies such as those formed between a fronted *wh*-phrase and its gap position. Among the few studies that have investigated this topic, Roland et al. (2007) conducted a corpus analysis of five adult corpora. They found that subject relative clauses occurred more frequently than other relative clause types (34% of the 196,385 relative clauses). Their findings suggest that, if readers' linguistic experience leads them to build the most frequent relevant structure during their comprehension or production, the lower frequency of object relatives would lead readers to predict subject relatives, leading to slowdown when that prediction was violated. This assumption has been confirmed in a series of studies. For example, Staub (2010) tested the role of readers' expectations in processing object relative clauses, using an eye-tracking paradigm. According to Staub, when readers processed the sentences in (4), the sum of all fixations on *the fireman* as in (4b) was significantly longer in object relative clauses (4a) than in subject relative clauses (4b) due to the violation of readers' expectation for a subject noun phrase right after *that*.

- (4) a. The employees that noticed the fireman hurried across open field.
 b. The employees that the fireman noticed hurried across open field.

Expanding this line of research cross-linguistically, Jäger et al. (2015) recently investigated the role of language experience in a Chinese context. They first conducted a distributional analysis of Chinese relative clauses using a Chinese tree bank, and they then conducted a sentence completion task. They found that, as in English, subject relative clauses were the most frequent in the input, and that L1-Chinese adults preferred to use subject relatives to complete sentences in the task. These findings are consistent with the predictions of probabilistic expectation accounts, but there is still insufficient research to conclude that the effects of probabilistic expectation are similar across different types of syntactic structures.

A recent study by Atkinson et al. (2018), although focusing on children's development, indicates that adult readers' language experience influences their processing of filler-gap dependencies. Using several English adult corpora, they found that, of 546 *what* questions, 474 contained a direct object gap (86.8%). This finding indicates the frequency of direct-object gap *what* questions (i.e., *what are you crawling on ___?*) and suggests that adults' linguistic experience would bias them to expect *what* questions to be direct object questions. The results of Atkinson et al.'s eye-tracking study suggest that such probabilistic predictions do affect readers' processing during language comprehension. Of course, *what* questions have a semantic bias toward being direct object questions, so more research examining other types of question (e.g., *which* and *who* questions) is needed before any generalizations can be made.

It is possible that some findings from previous studies could be explained by the probabilistic prediction account. For example, Stowe's (1986) results showed that reading times are slower at the direct object position after a *wh*-phrase when it is not a gap (i.e., *us* in [3b]); the slowdown could be explained by the readers' experience – in which direct object questions are more frequent than other types of questions – which leads the readers to favor the direct object interpretation, and thus to expect a gap instead of *us*. When their expectation is violated, they are surprised, resulting in slower reading. Thus, linguistic experience may play a role in a plausible explanation for readers' processing mechanisms in incremental sentence comprehension. More specifically, the processing of filler-gap dependencies

could be guided by the frequency of different structures in the input that has formed readers' language experience. To explore this possibility, particularly with the Korean scrambled *wh*-constructions of Korean context, the current study will test the predictions of the two accounts with a sentence completion task that will lead participants to generate Korean *wh*-constructions.

2.3 Korean *which*-questions in the language input

For a more precise comparison with the existing findings, the current study uses exact Korean translation (i.e., *enu*-phrase) of the Japanese *which*-questions. Korean, as mentioned, is a head-final language. Korean *wh*-constructions have two well-known syntactic properties. First, Korean *wh*-phrases are in-situ structures, which do not involve overt movement, as in (5a), but the *wh*-phrase can be scrambled, as in (5b), with no change in meaning.⁴⁾ Second, *wh*-scope marking is determined by a question-marking particle (i.e., *-ni*), which is affixed to the verb.⁵⁾ In contrast, in English it is the position of a *wh*-phrase that determines *wh*-scope marking via the overt movement of *wh*-phrases.⁶⁾

- (5) a. Mary-ka sakwa-lul enu namca-eykey cwuess-ni?
 Mary-nom apple-acc which man-dat gave-q
 b. Enu namca-eykey Mary-ga sakwa-lul cwuess-ni?
 which-man-dat Mary-nom apple-acc gave-q
 'Which man did Mary give an apple?'

4) Scrambling is possible due to the case-marking system (i.e., *-(l)ul* for the accusative marker; *-ga/-i* for the nominative marker), which marks the thematic role of each argument.

5) Korean has many question particles. For this study, we use the question particles *-ni* and *-ci* for the test sentences because of their frequency.

6) In English, the scope of *wh*-questions is determined by the position of the *wh*-phrase at the surface level. For a direct question, English *wh*-phrases need to move to the initial position of a sentence, whereas for an indirect question, they need to move to the initial position of the embedded clause. For example:

- a. Who did John say that Mary read a book to?
 b. John said who Mary gave a book to.

The sentences in (6) provide examples. In a direct question, the question-marking particle *-ni* is affixed to the main clause verb (6a), whereas in an indirect question, the question-marking particle *-ci* is attached to the embedded clause verb (6b).

(6)

a. Direct question

Mary-nun John-i enu haksayng-eykey chayk-ul cwetta-ko malhayss-ni?
Mary-top John-nom which student-dat book-acc gave-decl said-q?
'Which student did Mary said that John gave a book to?'

b. Indirect question

Mary-nun John-i enu haksayng-eykey chayk-ul cwennun-ci malhayss-ta.
Mary-top John-nom which student-dat book-acc gave-q said-decl
'Mary said which student John gave a book to.'

Semantically, a Korean *wh*-phrase also can be interpreted as either an interrogative or an indefinite, regardless of its position (i.e., in-situ vs. scrambled; Sohn, 1994; Choi, 2005). This thematic assignment is determined by the sentence context. This usage of Korean *wh*-phrases might lead to inconsistencies in the predictions of the Active Filler Strategy hypothesis in comparison with English.

Considering these properties, this study first conducted a corpus analysis to document frequency patterns for *enu*-phrase constructions to understand how Korean readers' probabilistic expectations might play a role in their production of Korean *wh*-constructions. The tokens of *enu*-phrases in the Sejong Corpus⁷⁾ of adult spoken language (2,050,000 words) were

7) This has produced the *Sejong* Corpus, the largest corpus of Korean language data as of 2004" (Bley-Vroman and Ko 2005: 258). Sejong Corpus includes both written and spoken language. This study used only the spoken language corpus, which consists of the scripts of TV dramas and news. Such spoken data is arguably the closest to the "spontaneous" naturalistic language output that shapes readers' linguistic experience (e.g., Wallis, 2014). However, the corpus analysis on the written data also needed to be included for generalizing the finding of the spoken data, as one reviewer commented.

counted using a corpus concordance program, MonoConc Pro (version 2.2), which extracted a total of 500 bi-clausal sentences including *enu*-phrases. These were then categorized as interrogatives and indefinites, as Table 1 shows. In-situ constructions were much more frequent than scrambled constructions in the corpus.⁸⁾ Of the 496 in-situ constructions, the *enu*-phrase received the interrogative interpretation 29.2% of the time (146 out of 500), whereas it received the indefinite pronoun interpretation 70% of the time (350 out of 500). This difference is statistically significant ($p < .0001$).

The more interesting matter for this analysis is the frequency of question-marking particles affixed to the embedded clause verb and to the main clause verb. In the 146 interrogative sentences⁹⁾ with the question-marking particle, it appeared at the embedded clause verb 52.7% of the time (77 out of 146), and at the main clause verb 46.3% of the time (69 out of 146). This difference, however, is not statistically significant ($p > 1$).¹⁰⁾

8) As Table 1 shows, very few scrambled sentences appeared. One of the reviewers expressed a concern that scrambled sentences rarely occur in speech due to the prosodic features (e.g., stress, pause, and intonation) of spoken language. The relevant discussion remains a question for further research.

9) The corpus data sample size (i.e., 146 sentences) might not be enough to generate readers' linguistic experience of *wh*-constructions, so one reviewer suggested including various other *wh*-phrase constructions from the corpus. However, this study aimed to provide evidence for cross-linguistic comparisons by using *wh*-phrases similar to those used in previous studies (i.e., *which*).

10) A reviewer raised the question of why the spoken corpus data showed no significant difference in the frequency of question-marking particles affixed to embedded clause verbs and to main clause verbs. Although a clear explanation cannot be provided in this study, I agree that this must be explored in further research.

Table 1. Distribution of Korean *enu*-phrase constructions

	Indefinite		Interrogative		Total
		Total	Embedded clause	Main clause	
Scrambled	0	0 (0%)	1	3	4 (0.8%)
In-situ	350	350 (70%)	77	69	146 (29.2%)

This finding implies that *enu*-phrases preferentially receive the indefinite pronoun interpretation in in-situ constructions even though *enu*-phrases are frequently interpreted as the interrogative. Due to the linguistic experience of Korean-speaking adults implied by the corpus findings, this study makes two different predictions for the participants' generation of Korean *wh*-constructions in the study's sentence completion task. Examples of task items appear in Table 2.

Table 2. A sample set of experimental items

Condition	Sentence fragment		
In-situ	John-un	Mary-ka	<i>enu</i> haksayng -eykey...
	John-top	Mary-nom	which student -dat
Scrambled	Enu haksayng -eykey	John-un	Mary-ka...
	which student -dat	John-top	Mary-nom

The specific predictions are as follows. If the processing mechanism is guided by the readers' probabilistic expectations from the language input, the participants will generate more declarative sentences, in which the *enu*-phrase receives the indefinite pronoun interpretation, than interrogative sentences. For example, the declarative complementizer will be attached to the embedded clause verb or the matrix clause verb right after the *enu*-phrase is identified, regardless of whether the *enu*-phrase is in-situ or scrambled. However, if the *enu*-phrases are treated as interrogatives, the participants will display no preference for attaching the question-marking particle to either the embedded clause verb or the main clause verb. On the other hand, if the processing mechanism is guided by the Active Filler

Strategy, once the *enu*-phrase is identified, regardless of whether it is in-situ or scrambled, the participants will produce an indirect question by immediately attaching a question-marking particle to the embedded clause verb in order to resolve the scope ambiguity of the *enu*-phrase. Table 3 summarizes the predictions for where the question-marking particle will occur based on the two accounts.

Table 3. Predictions based on the two accounts

	In-situ condition (control)		Scrambled condition (experimental)	
	Embedded clause	Main clause	Embedded clause	Main clause
Active Filler Strategy	√	–	√	–
Probabilistic prediction	√	√	–	–

Note. If the theory makes a prediction, the cell is marked √; if the theory cannot make a prediction, the cell is marked –.

3. Methods

3.1 Participants

Thirty-six Korean college students (19 women and 17 men; mean age: 22.27, $SD = 1.3$, range: 19–26) participated in the experiment. They were asked to complete sentences after reading given sentence fragments. Twelve experimental items (6 for in-situ conditions; 6 for scrambled conditions) were presented in a randomized order with 12 filler items, so that participants completed a total of 24 sentences.

3.2 Materials and Procedures

The sentence completion task utilized two experimental conditions. Experimental items were sentence fragments with a sequence of three phrases. The materials were similar to those used by Aoshima et al. (2004),

to enable a cross-linguistic comparison. As the revisited experimental items in Table 4 show, in the two conditions, the experimental items included two subject NPs and one dative-marked NP, which was always an *enu*-phrase (i.e., *John-un Mary-ka enu haksayng-eykey...*). The two subject NPs made the sentence bi-clausal, and, along with the presence of an *enu*-phrase, allowed the participants to decide whether to complete it as a direct question with a main clause question-marking particle, or as an indirect question with an embedded clause question-marking particle.

Table 4. A sample set of experimental items

Condition		Sentence fragment	
In-situ	John-un	Mary-ka	enu haksayng -eykey...
	John-top	Mary-nom	which student -dat
Scrambled	Enu haksayng -eykey	John-un	Mary-ka...
	which student -dat	John-top	Mary-nom

The experiment was conducted as a pen-and-paper task. Participants were instructed to first read the fragment and then to complete the sentence, and to create sentences that were grammatically and semantically acceptable. The task took 25 minutes.

4. Results

Clause number

First, monoclausal sentence completions (0.5%; 24 out of 432) were removed from the dataset. The remaining sentence completions were categorized as bi-clausal. Only bi-clausal completions were used for the data analysis, because such sentences are relevant for testing the hypotheses described in the previous section. In 113 of the 408 bi-clausal sentences generated in the task, the *wh*-phrase had an existential interpretation as in (7); these were also excluded, because *enu kaswu-eykey* is interpreted

as the argument of the matrix verb, but not that of the embedded verb.¹¹⁾ The remaining 295 sentences were included for further analysis.

- (7) Enukaswu-eykey umpan ceycakca-nun cakkokka-ka hithukok-ul
 which singer-dat music producer-top composer-nom hit song-acc
 cal mantunta-nun sasil-ul malhayss-ta.
 well made-that fact-acc said-decl
 'A music producer said to a certain singer the fact that a composer
 made a hit song well.'

Question-marking particle

Next, the sentence completions were divided into those that used the *enu*-phrase as an indefinite and those that used it as an interrogative. Figure 3 summarizes the frequency of the semantic meanings by the position of the *enu*-phrase. Sentence completions that included declarative complementizers with either the embedded clause or the main clause verb occurred 71.8% of the time (210 out of 295), whereas sentences including at least one question-marking particle were produced 28.2% of the time (85 out of 295). That is, the participants preferred to use the *enu*-phrase as an indefinite word rather than an interrogative word.

11) Although the existential interpretation as in (7) was excluded, the indefinite interpretation was included for further analysis, with the goal of providing a specific description of how *wh*-phrases are used in order to more specifically describe the role of readers' linguistic experience in their generation of *wh*-constructions.

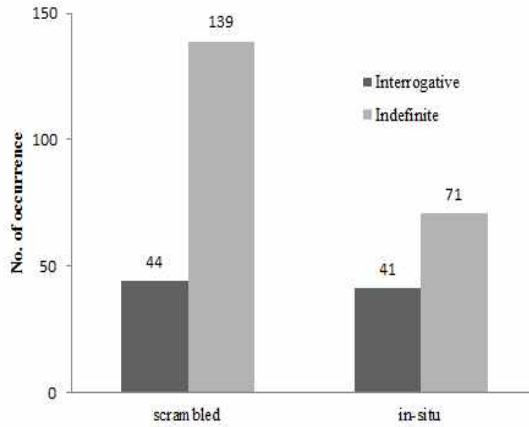


Figure 3. Interpretation of *enu*-phrases in scrambled and in-situ conditions

As in Figure 3, the *enu*-phrase was mostly assigned the indefinite word interpretation – and more so in the scrambled condition than in the in-situ condition. These results mirror the distributional pattern found in the Sejong spoken corpus, in that they were mainly used as indefinites regardless of the location of the *enu*-phrases, but the indefinite interpretation was even more frequent for scrambled than in-situ phrases, as exemplified in (8). On the other hand, as exemplified in (9), the *enu*-phrase was often associated with the embedded clause via the affixation of *-ci*, thus forming an indirect question including the dependency between the *enu*-phrase and the question-marking particle at the embedded clause verb.

- (8) Enu tokca-eykey coswu-nun cakka-ka cakphwum-ul keyu
 which-reader-dat assistant-top writer-nom a novel-acc almost
 wansenghayssta-ko malhayss-ta.
 finished-decl said-decl
 ‘The assistant said to a certain reader that a writer almost finished a novel.’

- (9) Enu hwanca-eykey uysa-nun kanhosa-ka sathang-ul cwennun-ci
 which patient-dat doctor-top nurse-nom a candy-acc gave-q
 kwungkumhayss-ta.
 wondered-decl
 'A doctor wondered which patient a nurse gave a candy to.'

This finding is itself of interest, but this study aims to answer the more important question of whether the underlying processing mechanism is guided by the readers' linguistic experience. Therefore, a more detailed analysis was conducted of only the 85 interrogative interpretations. As Table 5 shows, in the 85 interrogative sentences,¹²⁾ the question-marking particle appeared at the embedded clause verb 48.24% of the time (41 out of 85) in the in-situ conditions, and 51.78% of the time (44 out of 85) in the scrambled conditions. This difference was not significant ($p > 1$). This finding shows that once the *enu*-phrase was identified as an interrogative, regardless of its position (in-situ vs. scrambled), the readers tended to affix the question-marking particle to a verb as soon as possible, that is, to the embedded clause verb, to resolve the scope ambiguity of the interrogative.

Table 5. Occurrences and rates of question-marking particles in interrogative sentences:
 In-situ and scrambled conditions

	In-situ condition		Scrambled condition	
	Embedded clause	Main clause	Embedded clause	Main clause
No. of occurrences	41	0	44	0
%	48.24%	0	51.78%	0

(10) shows examples of such sentences, in which the question-marking particle is associated with the embedded clause verb (*gave*) in both

12) One of the reviewers questioned whether all 85 sentences were truly interrogatives, noting that, for example, "Na-nun ku-ka mwues-ul mekennun-ci kwungkum-hata" can be interpreted as *I wonder what he ate* or *I wonder whether he ate something*. Such sentences, however, were already excluded due to the ambiguity in the interpretations.

conditions. The Active Filler Strategy hypothesis would predict that the readers would immediately search for the question-marking particle as soon as they had identified an *enu*-phrase. This prediction was confirmed by the data from Korean adult participants' generation of *wh*-constructions, in which, when they produced questions, they preferred to affix the question-marking particle as soon as possible.

(10) a. In-situ condition

John-un Mary-ka enu chinkwu-eykey sathang-ul cwuessnun-ci
 John-top Mary-nom which friend-dat candy-acc gave-q
 malhayss-ta.
 said-decl.
 'John said which friend Mary gave a candy to.'

b. Scrambled condition

Enu sungkayk-eykey sungmwuwen-un coongsa-ka sathang-ul
 which-passenger-dat crew-top pilot-nom candy-acc
 cwessnun-ci malhayss-ta.
 gave-q said-decl
 'A crew said which passenger a pilot gave a candy to.'

5. Discussion

This study investigated the predictions of two competing accounts of how Korean *wh*-constructions are processed. It focused on whether the underlying processing mechanism was guided by active dependency formation or by readers' probabilistic expectations based on the frequency of alternate syntactic structures in the language input. In the corpus data, Korean speakers used the *enu*-phrase as an interrogative word 28.2% of the time (85 out of 295 cases). This finding indicates that while Korean speakers prefer to interpret the *enu*-phrase as an indefinite word, it is also frequently used as an interrogative word in the case.

A more interesting finding of this study is that in interrogative sentences, both in-situ and scrambled *enu*-phrases were associated with embedded clause verbs. As exemplified in (10), the Korean-speaking participants completed the sentences by immediately attaching the question-marking particle at the most adjacent verb (i.e., the embedded clause verb) once they had identified the *enu*-phrase. This finding is in line with the prediction of the Active Filler Strategy hypothesis, indicating that the underlying processing mechanism is guided by active dependency formation.

To summarize, the major finding is that Korean-speaking adults immediately formed filler-gap dependencies in generating *wh*-constructions, confirming the finding in Aoshima et al.'s (2004) comprehension study during real-time processing.¹³⁾ However, no significant effect of readers' linguistic experience was found in this study. While native Korean-speaking adults most frequently use an *enu*-phrase as an indefinite pronoun (i.e., *a certain student*), once they interpreted an *enu*-phrase as an interrogative word, they immediately tried to form the filler-gap dependency. They completed the indirect questions including the question-marking particle at the embedded clause verb regardless of its position, as shown in (10). This finding confirmed the prediction of the Active Filler Strategy hypothesis, and furthermore, showed that Korean *which*-questions and Japanese *which*-questions might share the same mechanism for resolving scope ambiguity in the processing of *wh*-constructions.

The results of this study raise two questions. The first question is why readers' linguistic experience did not affect their processing of Korean *wh*-constructions. This finding was unexpected, in light of the observation that almost equal numbers of question-marking particles appeared at embedded clause verbs and main clause verbs (77 vs. 69) for the resolution of scope marking in the corpus data (Table 1). Therefore, language experience cannot explain why all the question-marking particles were

13) One reviewer concerned that the finding in both comprehension and production studies cannot be exactly comparable, but this comparison is worth to identify the effect of reader's linguistic experience on the underlying mechanism operated in processing *wh*-constructions.

attached to the embedded clause verbs in the interrogative sentences generated in this study's experiment. A plausible explanation might be that the resolution of scope-marking in Korean might be related not only to the semantic meaning of the *enu*-phrase but also to its syntactic properties. Korean is a discourse-oriented language, hence a language comprehender's interpretation of an *enu*-phrase will be inconclusive until s/he understands its context. In addition, a question-marking particle is not syntactically obligatory following an *enu*-phrase, due to its different meanings (i.e., indefinite vs. interrogative), in contrast to the case of the Japanese *wh*-phrase used in Aoshima et al.'s (2004) study, which requires a co-occurring question-marking particle. The importance of context, the nonobligatory nature of the question-marking particle, and the *enu*-phrase's potentially different meanings together might lead to readers' linguistic experience having no effect on their generating of scrambled *wh*-constructions. This in turn would suggest that the amount of input might not be enough to generate readers' probabilistic expectations, a possibility suggested by Atkinson et al. (2018), who also pointed out that children's ability to use linguistic information that they have learned from frequency patterns in the input might be limited when children comprehend or produce *wh*-constructions. This question might be investigated in the Korean context for further research.

The second question is why the prediction of the Active Filler Strategy hypothesis was confirmed in the current study's production task (i.e., sentence completion), but not in Hahn and Hong's (2014) online comprehension task. Hahn and Hong found only partial confirmation of the prediction of the Active Filler Strategy hypothesis, in the in-situ condition but not in the scrambled condition. This inconsistency in Korean contexts might be due to the studies' different task types (i.e., online comprehension vs. offline sentence completion tasks). If working memory plays a role in the processing of Korean *wh*-constructions, then differences in participants' working memory capacity might influence the operation of the Active Filler Strategy. For example, readers with more limited working memory capacity might have more difficulty in online sentence processing, when they would need to hold in mind the information they have already encountered while

processing upcoming words. This suggests directions for further research, which should examine the role of working memory capacity in processing such filler-gap dependencies in Korean contexts, as well as in adults' ability to use the linguistic information they have learned from frequent patterns in language input.

6. Concluding Remarks

This paper tested the role of readers' language experience in generating Korean *wh*-constructions in comparison with the role of the Active Filler Strategy, a major account for the processing of filler-gap dependencies. The findings showed no effect of language experience; rather the readers preferred to immediately form a filler-gap dependency, in accord with the predictions of the Active Filler Strategy hypothesis, when generating Korean *wh*-constructions. Therefore, this study suggests that English, Japanese, and Korean share a similar mechanism for the processing of scrambled *wh*-constructions. The study's results further help resolve an issue raised by Hahn and Hong's (2014) replication study, which partially confirmed the prediction of the Active Filler Strategy hypothesis. However, further research should be conducted to fully understand the role of several factors that might be involved in processing filler-gap dependencies, including readers' linguistic experience.

References

- Aoshima, S., Philips, C., & Weinberg, A. (2004). Processing filler-gap dependencies in a head-final language. *Journal of Memory and Language, 51*, 23-54.
- Atkinson, E., Wagers, M., Lidz, J., Philips, C., & Omaki, A. (2018). Developing incrementality in filler-gap dependency processing. *Cognition, 179*, 132-149.
- Choi, J. (2005). A corpus-based approach to the Korean *wh*-question words *mwusun* and *enu*: A pedagogical perspective. In R. Bley-vroman, & H. Ko (Eds.) *Corpus*

- Linguistics for Korean Language Learning and Teaching* (pp. 107–147). Honolulu, HI: University of Hawai'i Second Language Teaching & Curriculum Center.
- Crain, S., & Fodor, J. D. (1985). How can grammars help parsers? In D. R. Dowty, & L. Karttunen, & Zwicky, M. (Eds.). *Natural Language Parsing: Psychological, Computational and Theoretical Perspectives* (pp. 92–128). Cambridge: Cambridge University Press.
- Frazier, L. (1987). Syntactic processing: Evidence from Dutch. *Natural Language and Linguistics Theory*, 5, 519–560.
- Frazier, L., & Clifton, C. (1989). Successive cyclicity in the grammar and the parser. *Language and Cognitive Processes*, 4, 93–126.
- Hahn, H., & Hong, S. (2014). Processing scrambled Wh-Constructions in head-final languages: Dependency resolution and feature checking. *Language and Information*, 18(2), 59–79.
- Hale, J. (2001). A probabilistic early parser as a psycholinguistic model. In *Proceedings of the 2nd Meeting of the North American Chapter of the Association for Computational Linguistics on Language Technologies* (pp. 159–166). Pittsburgh, PA: Association for Computational Linguistics.
- Hale, J. (2003). The information conveyed by words in sentences. *Journal of Psycholinguistic Research*, 32(2), 101–123.
- Hale, J. (2006). Uncertainty about the rest of the sentence. *Cognitive Science*, 30(4), 643–672.
- Jäger, L., Chen, Z., Li, Q., Lin, C., & Vasishth, S. (2015). The subject–relative advantage in Chinese: Evidence for expectation–based processing. *Journal of Memory and Language*, 80(1), 97–120.
- Levy, R. (2008). Expectation–based syntactic comprehension. *Cognition*, 106, 1126–1177.
- Omaki, A., & Schulz, B. (2011). Filler–gap dependencies and island constraints in second language sentence processing. *Studies in Second Language Acquisition*, 33, 563–588.
- Pickering, M. J., & Traxler, M. J. (2003). Evidence against the use of subcategorisation frequency in the processing of unbounded dependencies. *Language and Cognitive Processes*, 18(4), 469–503.
- Roland, D., Dick, F., & Elman, J. L. (2007). Frequency of English grammatical structures: A corpus analysis. *Journal of Memory and Language*, 57(3), 348–379.
- Staub, A. (2010). Eye movements and processing difficulty in object relative clauses. *Cognition*, 116(1), 71–86.
- Sohn, H. (1994). *The Korean language*. UK: Cambridge University Press.
- Stowe, L. E. (1986). Parsing wh–constructions: Evidence for on–line gap location. *Language and Cognitive Processes*, 1(3), 227–245.
- Wagers, M., Borja, M. F., & Chung, S. (2015). The real–time comprehension of *wh*

- dependencies in a *wh*-agreement language. *Language*, 91(1), 109–144.
- Wallis, S. (2014). What might a corpus of parsed spoken data tell us about language? *Proceedings of Olinco 2014* (pp.1–14). Czech Republic: Palacký University,
- Yun, H., Nam, Y., Yoo, D., & Hong, U. (2015). The effect of role shifting and expectation in the processing of center-embedded relative clauses in Korean. *Linguistic Research*, 32(2), 313–353.

On-Soon Lee

Assistant Professor

College of General Education

Dong-A University

37 550beon-gil, Nakdongdae-ro, Saha-gu, Busan, South Korea, 49315

Phone: 051-200-0924

Email: oslee@dau.ac.kr

Received on July 31, 2019

Revised version received on September 25, 2019

Accepted on September 30, 2019