

**VERB USE FOR THE LOCATIVE FUNCTIONS OF THREE ADVERBIAL
POSTPOSITIONS (-EY, -EYSE, AND -(U)LO) IN KOREAN: ANALYSIS OF L1-
KOREAN CORPORA AND L2-KOREAN TEXTBOOKS**

A DISSERTATION SUBMITTED TO THE GRADUATE DIVISION OF
THE UNIVERSITY OF HAWAI'I AT MĀNOA IN PARTIAL FULFILLMENT OF
THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY
IN
EAST ASIAN LANGUAGES AND LITERATURES
AUGUST 2020

By

Boo Kyung Jung

Dissertation Committee:

Mee-Jeong Park, Chairperson

Mary Kim

Kristopher Kyle

Seongah Im

William O'Grady

Keywords: Korean adverbial postpositions, Frequency, Language learning, Textbook analysis

Copyright © 2020

Boo Kyung Jung

All Rights Reserved

For my mother and my daughter

ACKNOWLEDGEMENTS

I am greatly indebted to my professors, colleagues, and family for this project. I first want to thank my academic advisor and dissertation chair, Mee-Jeong Park. She continuously showed her trust in me and my dissertation, which was an invaluable motivator that helped me go through my doctoral program. Moreover, her insightful advice and guidance on language teaching inspired me throughout my graduate years. I would like to express my sincere gratitude to her indefinite support to me.

I want to extend my thanks to all my committee members. Mary Kim offered wonderful opportunities to discuss the use of Korean language in detail with my colleagues through her classes. Her devotion to her research and teaching was a great influence on me. Kristopher Kyle's classes directed me to this interesting world of corpus linguistics and natural language processing. His advice was an integral part of this project and I couldn't have completed my dissertation without his warm support. I would also like to express my gratitude to Seungah Im for assisting me in understanding and applying statistics to my dissertation. William O'Grady's classes opened up new ways of interpreting and understanding languages for me. His unique intuition was truly helpful to my project.

I also feel the need to thank others whom I met at UH and in this field. I want to express my respect to Shinichiro Fukuda who was a trustful role model to me by being passionate about his research and caring deeply about teaching and students. I also thank Sorin Huh who introduced me this path. I thank my friends and colleagues in the EALL, SLS, and Linguistics departments

at UH. Among them, I thank Hyunjung An, Iven Bondoc, Sumi Chang, Hyunsun Chung, Meghan Delaney, Bonnie Fox, Kuni Homma, Dianne Juhn, Inho Jung, Soojin Kang, Thomas Kettig, Hwanhee Kim, Hye Seung Lee, Jee Hyun Lee, Sumire Matsuyama, Tyler Miyashiro, Sooran Pak, Hye Young Smith, and Fred Zenker for overall being supportive and helping me collect data, prepare teaching materials, and proofread my papers. In particular, Gyu-Ho Shin patiently answered all my inquiries and encouraged me to move forward. I am deeply grateful for his ceaseless support.

Last but not least, I thank my family. Without their heartfelt support, I would not have been able to even initiate this project. I thank my father for all his sacrifices and his trust in me. I also want to express my sincere admiration to my mother who has been courageously exploring her life with unquenchable optimism. My sisters and brother, Younsue, Jeah, and Kyungrok, have been wonderful models for me to follow. Their life-long support led me to where I stand now. I also want to deliver my thanks to Junho for coming into my life just in time to help me grow. Finally, I can't thank my daughter Dahyun enough. As a mature and responsible person, she has been a great joy to our family. To me, she always gives me strength to bravely face and happily continue through my journey. My special love goes to her.

ABSTRACT

Due to the multiple form-function relations of adverbial postpositions such as *-ey*, *-eyse*, and *-(u)lo* in Korean, the ways in which learners of Korean acquire these postpositions have been a long-standing research question. This dissertation investigates two main points: 1) the types of linguistic environments that elicit the locative functions of adverbial postpositions such as *-ey*, *-eyse*, and *-(u)lo* in L1 corpora and textbooks for L2 learning, and 2) the ways in which textbooks reflect verb use associated with such adverbial postpositions discovered in L1 corpora. More specifically, I examine the types and frequency of verbs occurring with each postposition and key verbs in relation to the types of corpora (e.g., L1 written-L1 spoken, L1 written-L2 textbook, L2 textbook-L2 textbook). I also analyzed association strengths between each postposition and co-occurring verbs by utilizing written and spoken data from the Sejong corpora as well as two types of textbooks for learners of Korean. The four main implications and findings of this study are:

- 1) In the Sejong corpora, some verbs were overwhelmingly used more frequently with a specific postposition (*iss-* ‘to be/exist’ with *-ey*, *ha-* ‘to do’ with *-eyse*, and *ka-* ‘to go’ with *-(u)lo*) while their token frequencies decreased exponentially, thus following a Zipfian distribution for both written and spoken corpora. Those verbs manifested the highest association strengths with that particular postposition.
- 2) Verb distributions in the two types of textbooks exhibited dominant usage of a few verbs, also following a Zipfian distribution. However, the most frequent verbs with each postposition in the textbooks differed from those in the L1 corpora: *ka-* ‘to go’ with *-ey*, *ha-* ‘to do’ and *sal-* ‘to live’ with *-eyse*, and *ka-* ‘to go’ with *-(u)lo*. The verbs that

exhibited greater association strengths over other verbs with *-ey*, *-eyse*, and *-(u)lo* in the textbooks were *ka-* ‘to go’, *ha-* ‘to do’ and *manna-* ‘to meet’, and *tolaka-* ‘to detour’ respectively.

- 3) In the textbooks, a limited number of verbs was introduced at the beginner level, but this number progressively grew for each postposition as target proficiency levels increased. While there were some dominant verbs for each postposition across all the levels in both textbooks, individual instances of these verbs differed according to level and textbook type.
- 4) Keyness analysis for each corpus yielded a number of distinctive verbs that occur with *-ey*, *-eyse*, and *-(u)lo* respectively, thus revealing unique characteristics for each corpus.

Findings from this dissertation lead to considerations about setting and building instructional goals and directions in two ways. First, frequency lists extracted from L1 corpora are useful not only in designing textbooks but also for class materials. Second, combinatorial use of these postpositions and verbs (which follow Zipfian distribution) suggests the need for a pattern-based instruction of postpositions. Considering that one important aim of second/foreign language learning is to expand repertoires of language use attested in the target language, learning frequent patterns of each postposition and verbs in the L1 corpus as a first step followed by purposefully expanded uses of each postposition with other linguistic items will ensure efficiency in learning.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS	iv
ABSTRACT	vi
LIST OF TABLES	xi
LIST OF FIGURES	xiv
LIST OF ABBREVIATIONS	xv
I. INTRODUCTION	1
II. INPUT AND LANGUAGE LEARNING	5
2.1. Usage-based approaches to language development	5
2.2. The effect of frequency and linguistic environment on language development	9
2.3. Implications of usage-based approaches to L2 development	14
III. THREE POSTPOSITIOSNS IN KOREAN: <i>-EY</i>, <i>-EYSE</i>, AND <i>-(U)LO</i>	17
3.1. General characteristics of <i>-ey</i> , <i>-eyse</i> , and <i>-(u)lo</i>	20
3.1.1. <i>-ey</i>	20
3.1.2. <i>-eyse</i>	23
3.1.3. <i>-(u)lo</i>	25
3.2. Many-to-one mapping: <i>-ey</i> , <i>-eyse</i> , and <i>-(u)lo</i> for the locational function	28
3.2.1. <i>-ey</i> and <i>-eyse</i>	28
3.2.2. <i>-ey</i> and <i>-(u)lo</i>	32
3.3. Acquisition of <i>-ey</i> , <i>-eyse</i> , and <i>-(u)lo</i>	36
IV. TEXTBOOKS FOR L2 LEARNING	40
V. METHODS	48

VI. ANALYSIS OF L1 CORPORA: USE OF <i>-EY</i>, <i>-EYSE</i>, AND <i>-(U)LO</i> AS LOCATIONAL POSTPOSITIONS WITH A FOCUS ON VERB USE.....	60
6.1. Verb use with each postposition.....	61
6.1.1. <i>-ey</i>	61
6.1.2. <i>-eyse</i>	63
6.1.3. <i>-(u)lo</i>	65
6.2. Relations between postpositions and verbs	68
6.2.1. Frequency distribution	68
6.2.2. Keyness analysis	73
6.2.3. Association strength	77
6.3. Summary of findings	83
VII. ANALYSIS OF L2 TEXTBOOKS: USE OF <i>-EY</i>, <i>-EYSE</i>, AND <i>-(U)LO</i> AS LOCATIONAL POSTPOSITIONS WITH A FOCUS ON VERB USE.....	85
7.1. Verb use with each postposition	86
7.1.1. <i>-ey</i>	86
7.1.2. <i>-eyse</i>	89
7.1.3. <i>-(u)lo</i>	91
7.2. Relations between postpositions and verbs	93
7.2.1. Frequency distribution.....	93
7.2.2. Keyness analysis	99
7.2.3. Association strength	101
7.2.3.1. General relations	101
7.2.3.2. Association strength by textbook level.....	105

7.3. Summary of findings	111
VIII. COMPARISON BETWEEN L1 CORPORA AND L2 TEXTBOOKS.....	114
8.1. Proportions across proficiency level	114
8.1.1. <i>-ey</i>	115
8.1.2. <i>-eyse</i>	117
8.1.3. <i>-(u)lo</i>	119
8.2. Keyness analysis	122
8.3. Association strength.....	129
8.4. Summary of findings	131
IX. CONCLUSION: PEDAGOGICAL IMPLICATIONS AND LIMITATIONS	133
9.1. Summary of findings	134
9.1.1. Three locational postpositions in L1 corpora	134
9.1.2. Three locational postpositions in L2 textbooks	137
9.1.3. Comparison of L1 corpora and L2 textbooks	138
9.2. Implications	141
9.2.1. Use of corpora in language education	141
9.2.2. Utilizing linguistic patterns in language instruction	143
9.3. Limitations and future studies	146
REFERENCES	150
APPENDICES.....	167

LIST OF TABLES

Table 3.1. Frequency percentages of postpositions in the Sejong written corpus	19
Table 5. 1. Locative concepts and example of verb types in Nam (1993).....	50
Table 5. 2. Categories and example of verb types used with the locative <i>-ey</i> in Nam (1993).....	51
Table 5. 3. Categories and example of verb types used with the locative <i>-(u)lo</i> in Nam (1993) .	52
Table 5. 4. Data amount breakdowns for the SK and IK textbooks	54
Table 5. 5. Contingency for log likelihood.....	56
Table 5. 6. Contingency for association strength.....	58
Table 6. 1. Frequencies of the three locational postpositions in the Sejong corpora.....	60
Table 6. 2. Frequencies and proportions of verbs used with <i>-ey</i> in the Sejong written corpus	62
Table 6. 3. Frequencies and proportions of verbs used with <i>-ey</i> in the Sejong spoken corpus	62
Table 6. 4. Frequencies and proportions of verbs used with <i>-eyse</i> in the Sejong written corpus .	63
Table 6. 5. Frequencies and proportions of verbs used with <i>-eyse</i> in the Sejong spoken corpus .	64
Table 6. 6. Frequencies and proportions of verbs used with <i>-(u)lo</i> in the Sejong written corpus	65
Table 6. 7. Frequencies and proportions of verbs used with <i>-(u)lo</i> in the Sejong spoken corpus	66
Table 6. 8. CKI verbs used with <i>-ey</i> in the Sejong corpora	75
Table 6. 9. CKI verbs used with <i>-eyse</i> in the Sejong corpora.....	76
Table 6. 10. CKI verbs used with <i>-(u)lo</i> in the Sejong corpora.....	76
Table 6. 11. Top five $\Delta P_{(\text{verb} \text{postposition})}$ values for verbs used with each postposition in the Sejong corpora	78
Table 6. 12. Verbs with high $\Delta P_{(\text{postposition} \text{verb})}$ values for each postposition over the other two postpositions in the Sejong corpora	80
Table 6. 13. $\Delta P_{(\text{postposition} \text{verb})}$ values for <i>iss-</i> , <i>ha-</i> and <i>ka-</i> with all three postpositions in the Sejong corpora	81
Table 6. 14. $\Delta P_{(\text{postposition} \text{verb})}$ values of <i>nathana-</i> and <i>sal-</i> in the Sejong corpora.....	82
Table 6. 15. Differences in ΔP values between all three postpositions and their dominant verbs in the Sejong corpora	82
Table 7. 1. Token numbers of locative postpositions in the SK textbook series (#).....	85
Table 7. 2. Token numbers of locative postpositions in the IK textbook series (#)	86
Table 7. 3. Frequency of verbs occurring with <i>-ey</i> in the SK textbook series by level	88
Table 7. 4. Frequency of verbs occurring with <i>-ey</i> in the IK textbook series by level	88
Table 7. 5. Frequency of verbs occurring with <i>-eyse</i> in the SK textbook series by level.....	90
Table 7. 6. Frequency of verbs occurring with <i>-eyse</i> in the IK textbook series by level.....	90

Table 7. 7. Frequency of verbs occurring with <i>-(u)lo</i> in the SK textbook series by level	92
Table 7. 8. Frequency of verbs occurring with <i>-(u)lo</i> in the IK textbook series by level	92
Table 7. 9. Five most frequent verbs occurring with each postposition across all volumes of the SK and IK textbook series	94
Table 7. 10. CKI verbs used with <i>-ey</i> in the SK and IK textbooks.....	100
Table 7. 11. CKI verbs used with <i>-eyse</i> in the SK and IK textbooks.....	100
Table 7. 12. Top three $\Delta P_{(\text{verb} \text{postposition})}$ values for verbs used with each postposition throughout the entire SK and IK textbook series	102
Table 7. 13. $\Delta P_{(\text{postposition} \text{verb})}$ values for the four verbs <i>ka-</i> , <i>ha-</i> , <i>manna-</i> and <i>tolaka-</i> with all three postpositions in the SK and IK textbook series	104
Table 7. 14. Differences in ΔP values between all three postpositions and their strongly associated verbs in the SK textbook series	104
Table 7. 15. Differences in ΔP values between all three postpositions and their strongly associated verbs in the IK textbook series	104
Table 7. 16. Top three $\Delta P_{(\text{verb} \text{postposition})}$ values for verbs used with <i>-ey</i> at each level in the SK and IK textbook series	106
Table 7. 17. Differences in ΔP between <i>-ey</i> and verbs at each level in the SK textbook series.	107
Table 7. 18. Differences in ΔP between <i>-ey</i> and verbs at each level in the IK textbook series..	107
Table 7. 19. Top three $\Delta P_{(\text{verb} \text{postposition})}$ values for verbs used with <i>-eyse</i> at each level in the SK and IK textbook series	108
Table 7. 20. Differences in ΔP between <i>-eyse</i> and verbs at each level in the SK textbook series	109
Table 7. 21. Differences in ΔP between <i>-eyse</i> and verbs at each level in the IK textbook series	109
Table 7. 22. Top three $\Delta P_{(\text{verb} \text{postposition})}$ values for verbs used with <i>-(u)lo</i> at each level in the SK and IK textbook series	110
Table 7. 23. Differences in ΔP between <i>-(u)lo</i> and verbs at each level in the SK textbook series	111
Table 7. 24. Differences in ΔP between <i>-(u)lo</i> and verbs at each level in the IK textbook series	111
Table 8. 1. CKI verbs used with <i>-ey</i> in the Sejong corpora and SK textbook series.....	123
Table 8. 2. CKI verbs used with <i>-ey</i> in the Sejong corpora and IK textbook series	124
Table 8. 3. CKI verbs used with <i>-eyse</i> in the Sejong corpora and SK textbook series	126
Table 8. 4. CKI verbs used with <i>-eyse</i> in the Sejong corpora and IK textbook series.....	127
Table 8. 5. CKI verbs used with <i>-(u)lo</i> in the Sejong corpora and SK textbook series	128
Table 8. 6. CKI verbs used with <i>-(u)lo</i> in the Sejong corpora and IK textbook series.....	129

Table 8. 7. $\Delta P_{(\text{postposition} \mid \text{verb})}$ value for <i>nathana-</i> and <i>sal-</i> in the SK and IK textbooks.....	130
Table A- 1. Functions and examples of postposition <i>-ey</i>	167
Table A- 2. Functions and examples of postposition <i>-eyse</i>	168
Table A- 3. Functions and examples of postposition <i>-(u)lo</i>	169

LIST OF FIGURES

Figure 5. 1. Scrambled order of sentences with locative postpositions.	53
Figure 6. 1. Frequency distributions of the top 30 verbs occurring with <i>-ey</i> in the Sejong corpora	70
Figure 6. 2. Frequency distributions of the top 30 verbs occurring with <i>-eyse</i> in the Sejong corpora	71
Figure 6. 3. Frequency distribution of the top 30 verbs occurring with <i>-(u)lo</i> in the Sejong corpora.	72
Figure 7. 1. Frequency distributions of the top 15 verbs occurring with <i>-ey</i> across all volumes of the SK and IK textbook series.	96
Figure 7. 2. Frequency distributions of the top 15 verbs occurring with <i>-eyse</i> across all volumes of the SK and IK textbook series	97
Figure 7. 3. Frequency distributions of the top 15 verbs occurring with <i>-(u)lo</i> across all volumes of the SK and IK textbook series.	98
Figure 8. 1. Proportion change of verbs used with <i>-ey</i> in the SK and IK textbook series.	116
Figure 8. 2. Proportion change of verbs used with <i>-eyse</i> in the SK and IK textbook series.	118
Figure 8. 3. Proportion change of verbs used with <i>-(u)lo</i> in the SK and IK textbook series.	121

LIST OF ABBREVIATIONS

ACC = accusative marker

ADN = adnominal modifier suffix

CAU = causative marker

GEN = genitive marker

LOC = locational marker

NOM = nominative marker

POL = polite speech level

PR = propositive marker

PRO = prospective marker

PST = past marker

PSV = passive marker

Q = question marker

SE = sentence ender

TOP = topic marker

I. INTRODUCTION

As an agglutinative SOV language, Korean utilizes several types of postpositions. Postpositions deliver important grammatical and semantic information about the noun phrase to which they are attached (Sohn, 1999), which has caused them to emerge as one of the main topics of research within the field of Korean linguistics. In terms of language learning, Korean speakers obtain essential information from postpositions to construe argument information for an event representation (Kim, 1999). Previous studies of postpositions have mostly investigated multiple functions of a postposition and/or the differences between postpositions that share similar functions (e.g., Nam, 1993; Yi, 1981). Frequent omission in colloquial settings and the many-to-many mappings between form and function for postpositions (National Institute of Korean Language, 2005; Choo & Kwak, 2008) make research on postpositions even more challenging.

Due to the complex nature of such postpositions, the question of how learners of Korean acquire and use them has also emerged as a major research area within disciplines that examine the acquisition of Korean as a second or foreign language (L2). It is reported that postposition errors are one of the most common errors committed by Korean learners, even those at the most advanced levels (Ko, Kim, Kim, Seo, Jung, & Han, 2004; Lee, 2003). Previous L2 Korean studies have either put special emphasis on textbooks' isolated explanation about the functions of individual postpositions or addressed learners' production errors (e.g., Han, 2014; Lee, 2003).

Under usage-based approaches, recent studies have been attesting the importance of frequency-based language input in language acquisition (e.g., Ellis & Ferreira-Junior, 2009a). Usage-based approaches to language development broadly investigate the process and enhancement of language acquisition. These approaches assume that actual experiences with

language affect the formation of conceptual structures of language (Bybee, 2013; Langacker, 2008; Tomasello, 2003). Similar to any other types of learning, a massive amount of input is stored in learners' memory in language acquisition. While several cognitive properties affect the perception of language items in the process of storing input, the relative frequency of linguistic items and the environments in which they occur often influence speakers' recognition (and abstraction) of the individual items as exemplars (Bybee, 2006; Goldberg, 2006; Verhagen, 2005). On the basis of the exemplars, a form-function correspondence emerges and is then registered as a constructional chunk, which allows language users to expand their linguistic knowledge to novel language items (Abbot-Smith & Tomasello, 2006; Bybee, 2013; Ellis, 2002; Tomasello, 2000).

Given the critical role of input frequency and linguistic environments within language acquisition, there exists a need to investigate how frequently and with which language elements individual postpositions co-occur in authentic usages of language. Although previous research has made many advances through the qualitative analysis of Korean postpositions' various functions, the similarities/differences between Korean postpositions, and common errors committed by L2 learners of Korean, little attention has been paid to the distributional properties of each postposition in both L1 and L2 contexts. Thus, my dissertation aims at filling these gaps by analyzing the distributional properties of postpositions in L1 resources such as the Sejong written and spoken corpora as well as examining the functional characteristics of two types of L2 learner textbooks. Textbooks were selected as a data source due to the fact that they serve as a primary source of linguistic input for language learners that influences their initial perceptions of Korean postpositions.

Of the various postpositions in Korean, this project focuses on three adverbial postpositions: *-ey*, *-eyse*, and *-(u)lo*. While they each have their own distinctive functions due to an overarching polysemous semantic network, they overlap in terms of the locative function: the three postpositions are attached to locational noun phrases and represent how they work in composing sentential structure. Conventionally, the meaning of co-occurring verbs has been considered to play a key role in distinguishing between locative postpositions (Kang, 2012). In this regard, the present project examines what types of verbs are frequently used with each postposition when they are used to express a locative function. The analysis of this study consists of two main components:

1) An investigation into the frequency and distributional properties of co-occurring verbs with *-ey*, *-eyse*, and *-(u)lo* within two types of data sources (authentic L1 data contained in the Sejong corpora provided by the National Institute of Korean Language (NIKL) and two types of language textbooks for L2 learners of Korean)

2) An examination of the extent to which tendencies found in the L1 data are reflected in the Korean L2 textbooks. I expect this project to serve as an initial step towards understanding the properties of input provided to language learners as well as finding pedagogical implications.

I also pose the following three research questions:

- i) What types of locative functions for each postposition are expressed in L1 corpora? Is there a relationship between locative postpositions and the environments within which they are used (e.g., co-occurrence with the verb)? If so, what are the characteristics of the relationship?
- ii) How is the locative function of each postposition used in textbooks for L2 learners? Is there a relationship between locative postpositions and the environments within

- which they are used (e.g., co-occurrence with the verb)? If so, what are the characteristics of the relationship?
- iii) What are the similarities and differences between L1 corpora and L2 textbooks regarding the use of these three locational postpositions?

This dissertation is composed of nine chapters. In Chapters 2 through 4, previous studies on human cognition and its implications for language learning, the use of Korean postpositions, and the role of textbooks will be reviewed. Chapter 5 describes the corpora and methods used in this paper, with results reported in Chapter 6 (L1) and Chapter 7 (L2). Chapter 8 compares the results from the L1 and L2 corpora. Pedagogical implications and conclusions are discussed in Chapter 9.

II. INPUT AND LANGUAGE LEARNING

2.1. Usage-based approaches to language development

Usage-based approaches have tried to address questions surrounding the process of language learning itself. The basic assumption of these approaches is that actual experience with language shapes conceptual representations of language (Behrens, 2009; Bybee, 2013; Langacker, 2008). Accumulated language use establishes a link between linguistic forms and corresponding meanings (and vice versa), which constitutes language knowledge in a systematic way. Since knowledge is *grounded in the usage event*, both concrete language items and generalization occur by way of individual instances of language use (Kemmer & Barlow, 2000). Therefore, these approaches stress the importance of linguistic instances in the course of learning.

Usage-based approaches are inspired by the basic insights from cognitive linguistics (Behrens, 2009). Cognitive linguistics investigates the relations between language and domain-general cognitive factors, which are independent of language, that affect the establishment of linguistic patterns (Langacker, 1987, 2008). Those factors include, but are not limited to, processing cost, embodiment from experience, abstraction, metaphorical expansion, and entrenchment. To illustrate, when a new linguistic item is perceived, information processing about the item is affected by our cognitive systems. Our cognitive system utilizes all those factors to process the new information using pre-existing knowledge about the item. In other words, knowledge about a certain entity is established in relation to previous understanding of the entity through cognitive mechanisms (Ellis, Römer, & O'Donnell, 2016). This way of understanding language views the underlying bases of a linguistic structure as something that emerges and is enhanced as a function of language use and properties of human cognition.

Usage-based approaches see language learning processes as something similar to processes of other types of learning: language acquisition also requires input from the outer world. Cognitive abilities work while processing the input, and all such information as form, function, distributional property, co-occurring elements, and context involving each linguistic instance is registered in the language user's memory. Repeated language use establishes form-function mapping.

Form-function mapping is not fully predictable from its components (Goldberg, 1995). To illustrate, a word *table* has a series of sounds (form) and a meaning that roughly equates to something like *a piece of furniture that has a flat top supported by one or more legs*. The form and meaning of a word such as *table* are associated with each other. In addition, the mapping is not limited to words but ranges from a morpheme (e.g., *-ing*) to an idiomatic expression and a syntactic pattern (e.g., Subject-Verb-Object-Object for ditransitives in English) (Goldberg, 2013). This concept of mapping changes the traditional view on grammar, which distinguishes grammar from words. Instead of considering a sentence as a sequence of words that follows grammatical rules, for instance, a sentence becomes a combination of mappings that have varying levels of complexity. From this approach, language is acquired through piecemeal learning that is biased to frequency (Ellis et al., 2016). Thus, what we traditionally label as grammar is understood as an aggregation of mappings.

The establishment and development of linguistic form-function mappings have been explained from several perspectives. The concept of prototypicality shows how each linguistic item becomes an exemplar that is used for the development of language structure. Earlier work on categorization studies by Rosch (1975) and Rosch and Mervis (1975) hint at this concept. They argue that language categories have an internal structure in which some members are more

representative than others and the best example of a category becomes a prototype of the category. The prototype then displays the maximum number of attributes that represent that category while the other members become relegated to peripheral groups with differently weighted attributes (Coleman & Kay, 1981; Hampton, 1987; Murphy, 2002). Even though such categorization studies were more focused on word semantics, the basic idea of them is consistent with that of prototypicality.

Exemplar models (Nosofsky, 1991, 1992; Kruschke, 1992) suggest some attributes of prototypicality as well. According to these models, novel instances are categorized by similarities to previously registered exemplars: language users categorize and generalize new items using current knowledge to build up more complex constructions (Bybee, 2006; Goldberg, 2006; Verhagen, 2005; Tomasello, 2000). To illustrate, when a new item from input is identical with a previous one (e.g., meaning, phonetic property, context, etc.), the item is mapped into the existing exemplar while strengthening the previous one. On the other hand, a new item is stored close to the similar exemplar when the degree of variation is small. Form-function relations subsequently emerge from the exemplars that were strengthened by an adequate amount of input. This approach can explain prototypicality in that more typical exemplars are centered as a prototype and share the highest amount of similarities with the other items (Ramscar & Port, 2015; Bybee, 2006). Overall, prototypicality has been studied as a major factor in language acquisition.

What is notable here is that the concept of prototypicality matches the underlining ideas from studies on the frequency of exemplars. When linguistic items with similar forms and meanings are located close to each other, clusters of instances emerge (Bybee, 2013). During the process of accumulation of input, it is highly possible that the most frequent instance in a

category becomes the exemplar and ultimately represents a prototype for the category (Lakoff, 1987; Langacker, 1987; Mandler, 1992; Brugman, 1988). When those clusters are fully generalized through adequate amounts of use, they become constructional chunks.

Several studies on usage-based approaches provide accounts for how low-level constructional chunks, which are established partially through language input, are expanded to more abstract and generalized forms as a result of exposure to language input. In general, what is stressed in the examination of generalization is that high frequency of occurrence intensifies the process and as the process becomes reinforced, differences of representations become distinct (Langacker, 1987; Croft & Cruse, 2004; Bybee, 2006; Blumenthal-Drame, 2012). Goldberg (2019) has shown how frequent use of verbs in argument structure constructions strengthens memory representation. Similar to Nosofsky's (1991) exemplar model, Goldberg argues that a new input strengthens a representation when items are identical or cluster around the originally stored item with an acceptable amount of variation. Eventually, repeated use reduces processing burden through automatization, unitizes patterns and makes them easily accessible by strengthened memory representation (Langacker, 1987; Bybee, 2006). That is, frequency affects every stage of the establishment of language constructions.

A similar approach was found in studies conducted from an emergentist perspective. In a series of papers, O'Grady (2005, 2015) investigates what motivates language users to accept one linguistic form as more natural than others and how people acquire language knowledge. These studies suggested two types of processing pressures that generate acceptable syntactic structures. One is an internal pressure, whose aim is to *minimize the burden of working memory* (i.e., *processing cost*) and to *maximize predictability*. The other is an external pressure that comes from the necessity to align one's language use with others. O'Grady argues that once systemized

language forms are established as a way to reduce processing cost, they become processing routines through repeated use. These routines are entrenched so that language users employ them naturally. While the studies stress the importance of reducing processing burden in shaping acceptable forms, they also acknowledge the role of frequency in the acquisition of language to make those forms become routinized to the point where language users can use them automatically through repeated exposure.

Overall, usage-based accounts have been changing views on language acquisition. With the assumption that language is learned through the experience of language, they investigate how language emerges and develops through language use by examining the properties of input. During all language processes, frequency plays a key role while all the information carried by input (e.g., frequency, association relationship, selection of verb) is registered. This tendency drives learners to acquire lexically specific knowledge and to extend this knowledge up to abstract structures (e.g., Bybee, 2006; Ellis, 2002; Goldberg, Casenhiser, & Sethuraman, 2004). The accumulation of systemized structures subsequently bears a language, which is eventually accepted and conventionalized through language use by the language community.

2.2. The Effect of frequency and linguistic environment on language development

The importance of input calls for an examination of the characteristics and the precise role of input for the growth and development of form-function correspondences that constitute target language systems. Among the features of input that language users receive, frequency of occurrence and distributional properties in the input play a significant role in generating form-function mapping and the formation of exemplars due to humans' innate sensitivity to frequency distribution (Abbot-Smith & Tomasello, 2006; Bybee, 2013; Ellis, 2002; Klein & Manning,

2005; Diessel, 2007). Under the general consensus that input frequency plays a key role in language acquisition from usage-based approaches, a large body of literature addresses the question of frequency effects in first and second/foreign language acquisition.

A number of empirical studies have demonstrated the role of frequencies in every subfield of language acquisition. Frisch, Large, Zawaydeh and Pisoni (2001) created English novel words that exhibited both frequent and infrequent phonotactic patterns found in English words. When native speakers were asked to judge whether those words looked like English words based on a 7-point scale, novel words that matched the characteristics of exemplars were given a high score. This means that humans consider new words as well-formed when they follow highly frequent patterns, which shows that the frequency information stored in cognition works as a standard in judging the degree of familiarity of new items.

Dąbrowska and Szczerbiński's (2006) study shows the effect of input in morphology. They tested 57 Polish-speaking children's productivity of three case markers: genitive, dative, and accusative markers with 12 real words and 12 nonce words. The results showed a strong correlation between input frequency and children's correct production of case markers for nonce words, which supports the usage-based acquisition of children's morphological patterns.

Sentence-level effects of frequency on sentence structure were also tested by Goldberg, Casenhiser and Sethuraman (2004). In their empirical study of 81-undergraduates, learners were found to be able to expand meanings of construction to novel words in a forced choice comprehension task after less than three minutes of SOV word order training with nonce verbs. This implies that learners' early and frequent exposure to certain constructions can aid their comprehension of the syntactic structure of a sentence. In sum, frequency played a key role in establishing knowledge about the target language in every aspect.

As the role of input frequency became emphasized, the properties of input were investigated from several angles. Of special interest is a contrast between ‘token frequency’ and ‘type frequency’ (Ellis, 2002). Token frequency refers to the number of occurrences of specific words or items in an utterance, which allows for an overlap within the same lemma. Type frequency, in contrast, is how often “distinct lexical items that can be substituted in a given slot in a construction” (Ellis, 2002, p .166) occur within the structure of interest. For instance, the token frequency of the verb ‘think’ is three in (1), but the type frequency of the same verb is one.

(1) He thinks that I thought that you would think about me.

Both token and type frequencies are generally considered to affect language acquisition in various ways. The role of token frequency has been manifested as one that facilitates the selection of exemplars and encourages faster processing (entrenchment) through repetition (Ellis, 2012). Bybee (2008) summarizes three effects of token frequency: conserving effect, autonomy, and reducing effect. That is, frequently occurring items become more accessible through repetition because language use strengthens memory representation of form-function mappings. Highly accessible forms are resistant to reform, which results in their conservation. Autonomy, as an extreme case of the preceding effect, takes place when frequently co-occurring strings of elements are merged into a single form element (e.g., *gimme* as one form instead of *give + me*). Finally, sound change is highly associated with degree of frequency. Frequent phrases (e.g., greetings) are more prone to phonetic reduction in comparison to low-frequency items. In regard to type frequency, it is considered to create a category and expedite the productivity of a construction, which is key to language development (Bybee & Hopper, 2001). The reasons why

type frequency is related to productivity are: i) the more lexical items are heard in a construction, the less it will be associated with a specific item, ii) when the construction covers more items, it will be more generalized, easily accepting new members, and iii) high type frequency indicates that the construction is used frequently, which will make its usage more accessible (Thompson, 2007).

Several other studies investigated the influence of ‘skewed’ input versus ‘balanced’ input. Skewed input describes the situation in which one or two types of features (e.g., verb) constitute large portions of all tokens, whilst balanced input is used when several types of tokens are distributed evenly among types. Studies have shown contrasting results: positive effects of skewed input in first language acquisition were proven in some studies (Goldberg et al., 2004; Casenhiser & Goldberg, 2005) while other studies (Year and Gorden, 2009) showed positive effects of balanced input for learning a second language. In either case, they all assume that the role of input is significant for language use. In sum, empirical studies support usage-based approaches that claim language use is a decisive factor for the development of language. Most of all, frequency affects each step of the learning process due to humans’ innate sensitivity to frequency.

In addition to the examination of frequency of input, investigations into the environments within which each token occurs are increasing. Indeed, several scholars consider that analyzing linguistic environment is an effective way to examine contextual diversity, arguing that objects we are interested in are recognized faster and in a more accurate way when contextual information is provided (Biederman, Glass, & Stacy, 1973, recited from Divjak & Caldwell-Harris, 2015). Exploring linguistic environments provides detailed stereoscopic pictures of language use by indicating relations among sentential elements or sentences. In all, analysis of

environment has stronger association power than mere frequency with a token (Jurafsky, 1996; Gries, Hampe, & Schönefeld, 2005) when an appropriate method that captures the distinct features of occurrences is used.

With the development of techniques to deal with large-size corpora, various aspects of linguistic environments can be examined such as in which context a certain word occurs more or less frequently, or with which element a certain entity occurs frequently. One way of examining linguistic environment is to gauge the association strength between linguistic elements (Blumenthal-Drame, 2012). For instance, one can examine mutual association strength between two words: 'Christmas' and 'tree'. The chance that 'tree' occurs after 'Christmas' will be higher than the other way around, which means that the mutual association strength value of 'tree' under the occurrence of 'Christmas' will be higher.

Ellis (2006) discusses the concept of contingency as a way for calculating association strength. Contingency considers four conditions of events regarding cue and outcome: i) outcome under cue, ii) no outcome under cue, iii) outcome under no cue, and iv) no outcome under no cue, with any linguistic feature possibly serving as a cue or an outcome. For instance, a cue can be a verb such as 'give' with 'ditransitive construction' as an outcome. A commonly employed method to consider the directionality of the relationship between cue and outcome is Delta P (ΔP). As several studies adopted ΔP to examine association strength (e.g., Ellis & Ferreira-Junior, 2009b), the current study will also use ΔP to examine the strength of relations between locative postpositions and verbs, which will be further discussed in Chapter 5.

2.3. Implications of usage-based approaches to L2 learning

The applicability of usage-based approaches to Second Language Acquisition (SLA) has been explored by several studies (e.g., Ellis & Ferreira-Junior, 2009a, 2009b; Jiang & Nekrasova, 2007). Findings from SLA studies in general indicated that L2 learners demonstrate knowledge about form-function mappings from very basic units such as co-occurring elements to highly entrenched conventionalized units (Dąbrowska & Divjak, 2015). While several factors such as prototypicality, recency and redundancy influence SLA and some of them are unique in L2 acquisition (e.g., interference from L1), the effect of input frequency is still valid in L2 acquisition (Ellis, 2012; Dabrowsk & Divjak, 2015). Kirsner (1994) notes that this phenomenon can be simply explained by the law of power: more practice yields better performance in every kind of learning that involves cognitive skills (Anderson, 1982; Newell, 1990). For this reason, the essential role of input frequency in SLA has been documented by a number of previous studies as well (e.g., Year, 2009; Larson-Hall & Dewey, 2012; Goldschneider & DeKeyser, 2001; McDonough & Kim, 2009; McDonough & Nekrasova-Becker, 2014).

Upon the development of the concepts and techniques to measure, the effects of input frequency, occasionally along with strength of association, are manifested by recent studies in language acquisition (e.g., Ellis, 2006a, 2006b). Ellis and Ferreira-Junior (2009a) showed that verbs in English constructions (e.g., VOL: verb object locative) followed Zipf's (1935) law, which explains that one single lexical item occupies a large portion of exemplars, in both native and non-native utterances. Also, the frequency order of verbs by learners broadly matched the order of native speakers' data, while the most prototypical and frequent exemplars in each construction drive L2 language learners' use of English verb-argument construction (e.g., *put* in a caused-motion construction, *give* in a ditransitive construction, etc.). In a subsequent study

(Ellis & Ferreira-Junior, 2009b), they calculated contingencies between verbs and each VOL to examine how frequency and frequency distribution affected language acquisition. The results of the study revealed that the first-learned types of verbs by language learners in each construction were the ones that occurred more frequently in the input and those verbs were associated with each construction more distinctively.

Similar results were found in other studies. By using automatic calculation techniques, Kyle and Crossley (2017) analyzed a large-size learner corpus: 480 TOEFL argumentative essays. By deploying a number of indices that reflect usage-based approaches such as frequency, association strength and Verb Argument Construction (VAC), the study examined the relationship between those indices and syntactic sophistication. Statistically significant results showed that novice L2 writers relied on frequently occurring verb-argument constructions, while L2 writers used both frequent as well as infrequent ones as their proficiency developed. The results were the same in the longitudinal data analysis (Kyle, 2016): negative correlations were observed between length of study (i.e., proficiency) and VAC frequency.

Fewer studies have been conducted on input's effect on the acquisition of second/foreign languages other than English. A study by Türker (2017) tried to examine the effect of input on the acquisition order of multiple meanings of the Korean postposition *-ey* by 45 English learners of Korean as a foreign language. The research designed multiple choice (interpretation) and fill-in-the-blanks (production) tests. The study indicated that the functions of *-ey* for which learners showed higher accuracy rates (goal, stative location, contact, time) were the ones that showed high proportions of *-ey* usage in L1 corpora (*CetCon*). On the other hand, the proportional 'unit' function of *-ey* was low in L1 corpora and the accuracy rate of that function by L2 learners was accordingly low. A study by Miyata (2011) examined the effects of frequency on the acquisition

of artificial language. The results suggested that i) there was a significant effect of token frequency when type frequency was controlled, ii) high frequency tokens were learned accurately, and iii) when L1 effects were controlled, skewed input and structured input had significantly positive effects in learning. Overall, the effect of input frequency has been attested in learning second/foreign or artificial languages.

In sum, usage-based approaches emphasize the role of language use and cognition in language learning. Among many cognitive factors that affect language acquisition, input frequency has been considered a key factor for building up knowledge about language structures. In the case of L1 and L2 studies, while specific research results demonstrated that degree of acquisition can vary for each linguistic domain (e.g., phonology, inflectional morphology, word-order, and acquisition of single words and syntactic constructions), the effect of input frequency has been effectively demonstrated. In this vein, the current study will investigate frequency distributions of target linguistic items and the environments within which they occur.

III. THREE POSTPOSITIONS IN KOREAN: *-EY*, *-EYSE*, AND *-(U)LO*

Korean is a verb-final language, which means that the rich amount of information carried by a verb (e.g., tense, aspect, modality) is not accessible until an utterance includes some type of verb. Korean is also a situation-oriented language that allows for the omission of sentential elements deemed to be recoverable from the context itself. Moreover, as an agglutinative language, Korean utilizes more than 600 affixes to manifest diverse functions such as extending the meaning of a word (e.g., *cipwulha-*: to pay, *sen-cipwulha-*: pay in advance), indicating plurality (e.g., *sakwa*: apple, *sakwa-tul*: apples), and expressing politeness (e.g., *ka-*: to go, *ka-si-*: to go (honorific)) (Sohn, 1999). Thus, due to the typological characteristics of Korean, various linguistic factors are engaged in its interpretation and it is essential for Korean language users to obtain information from these sources to fully construe sentential meanings.

One of the sources that language users employ to deliver grammatical information in Korean is a postposition¹, a function word that conveys the syntactic and semantic information of the noun phrase to which it is attached (Sohn, 1999). The information encoded by postpositions is crucial for listeners to infer argument structures in relation to an event representation (e.g., Kim, 1999). For instance, in the sentence *Tom-i Mina-eykey chayk-ul cwu-ta* ‘Tom gave Mina a book’, the sequence of the two postpositions *-eykey* (dative) and *-ul* (accusative) gives Korean language users a sense of the dative construction before a dative verb *cwu-* ‘to give’ is uttered,

¹ In Korean linguistics, different terms have been used and suggested to name function words which are attached to a noun to indicate the relationship of the noun and the sentence. One of the common terms is *cosa*. Other terms such as *kyek-cosa* (meaning ‘*cosa* that indicates its role in a sentence’) and *po-cosa* (meaning ‘*cosa* that supplement the meaning’) are used to specify certain types of function words. However, categorizing a function word into a certain category is not always clear-cut (Kim & Shin, 1987; Ko & Ku, 2008; Choi, 2005). To avoid ambiguity and to follow conventions held within the field of Korean linguistics, the present dissertation uses the term *postposition*, which focuses on the positional attribute of a function word.

thereby establishing a clue about the overall meaning of the sentence. The use of postpositions also allows for scrambling, which results in *Tom-i chayk-ul Mina-eykey cwu-ta*.

Postpositions vary in their functions and frequency of occurrence. To illustrate, *-i/ka* (nominative), *-(l)ul* (accusative), *-uy* (genitive), and *-eykey/-kkey* (dative) are mostly used to indicate the grammatical roles of noun phrases. Other postpositions such as *-ey* (place/goal/others), *-eyse* (place/source/others), and *-(u)lo* (directional/instrumental/others) are used to form an adverbial phrase. Some studies examined the use of each postposition in relation to what types are used frequently in Korean (e.g., Seo, 2006, 2014). Kang and Kim (2009) analyzed about 15 million *eojeols*² in the Sejong written corpus, which is composed of written documents such as newspapers, magazines articles, and books. As Table 3.1 shows, the most common postpositions were *-(l)ul* (accusative, 19.88%), *-i/ka* (nominative, 14.39%), *-(n)un* (topic/contrast marker, 14.1%), and *-uy* (genitive, 12.97%)³. The other four common postpositions were polysemic postpositions such as *-ey*, *-eyse*, and *-(u)lo*, along with the delimiter *-to* (addition). The top eight postpositions occupy more than 85% of the entire proportion of postposition use.

One interesting point with regard to Korean postpositions is the many-to-many mapping relations between form and function: one form carries out multiple functions, and one function can be performed by multiple forms (e.g., Choo & Kwak, 2008; Sohn, 1999). *-ey*, *-eyse*, and *-(u)lo*, the analytical foci of the current study, are representative adverbial postpositions that possess a polysemy of functions (e.g., Kang, 2012): each postposition serves as an indicator of various functions depending on the linguistic environments (e.g., nouns and verbs) within which

² An *eojeol* is a white-space-based unit serving as the minimal unit of sentential components in Korean (Lee, 2011).

³ The study did not consider case marker-omitted instances. Shin (2020) analyzed the use and frequency of nominative, accusative and dative markers in caregivers' speech. The study examined omitted instances as well as overtly used instances of case markers in active transitive and suffixal passive patterns.

it occurs. The multiple functions of *-ey*, *-eyse*, and *-(u)lo* are summarized in Table A-1, A-2 and A-3 in the Appendix section based on the online Standard Korean Dictionary (*Pyocwunkwuketaysacen*) published by the National Institute of Korean Language (NIKL). While these three postpositions each have their own unique functions, they also all share a locative function.

Table 3.1. Frequency percentages of postpositions in the Sejong written corpus

Rank	Form	Function	Frequency (#)	Percentage (%)	Accumulated percentage (%)
1	<i>-(l)ul</i>	accusative	1,181,994	19.88	19.88
2	<i>-i/ka</i>	nominative	855,434	14.39	34.27
3	<i>-(n)un</i>	topic/contrast	838,342	14.10	48.37
4	<i>-uy</i>	genitive	771,005	12.97	61.34
5	<i>-ey</i>	locative/time/others	590,252	9.90	71.27
6	<i>-(u)lo</i>	directional/instrumental/others	386,453	6.50	77.77
7	<i>-to</i>	addition	243,965	4.10	81.87
8	<i>-eyse</i>	locative/source/others	205,389	3.45	85.32
	Others		872,550	14.68	100
Totals			5,945,384	100	

Note: Others include 291 different postpositions (*cosa*) when allomorphs are separately counted (e.g., instrumental/qualification postposition *-lose* after vowel ending nouns and *-ulose* after consonant ending nouns are counted as two forms).

Due to their characteristic of having multiple form-function relations, those three locational postpositions are worthy of investigation. Indeed, *-ey*, *-eyse*, and *-(u)lo* have been widely studied from diverse perspectives: the functions of each postposition type, the relations of those functions, historical changes in them, and the subtle differences in meanings that result from alternations in use. Several studies have focused on directly comparing *-ey*, *-eyse*, and -

(u)lo with each other, which has resulted in explanations of the different ways in which they encode locative functions. In the following sections, multiple functions of the three postpositions and salient differences between them will be reviewed.

3.1. General characteristics of -ey, -eyse, and -(u)lo

3.1.1. -ey

The postposition -ey is productively used in Korean with diverse functions. In general, encoding location is regarded as its most basic function. -ey often co-occurs with stative verbs such as *iss-* ‘to be/exist’, as can be seen in (2), with -ey attaching to the noun that serves as the location of existence.

- (2) *pang-ey iss-ta*
room-ey be-SE
‘(someone) is in (one’s) room.’

When it occurs with motion verbs such as *ka-* ‘to go’ and *o-* ‘to come’, the location is understood as a destination or a goal that the subject is heading towards. Another distinct function of -ey is the marking of time nouns: *yeltwusi-ey cemsim-ul mek-ta* ‘(someone) eats lunch at noon’ and *hwayoil-ey swuep-i iss-ta* ‘(I) have a class on Tuesday’. It is also used to show a causal relationship by indicating a causal noun (*pi-ey tali-ka kkunh-ki-ess-ta* ‘Due to rain, the bridge was disconnected’) or as a fixed form (*ttaymwun/tekwun/thas/param-ey* ‘because of’) as well. Other functions of -ey include *scope, purpose, means/method, environment/condition, unit/standard, comparison*, and so on (see Table A-1 in the Appendix section). Some other studies specify these functions using different terms such as *inanimate recipient* (Kang, 2015) or

dative marker (Sohn, 1999) or *abstract point* (Choo & Kwak, 2008). There are also numerous idiomatic expressions with *-ey*. Among idiomatic expressions in relation with location and time, *tocwung-ey* or *-nun/ten cwung-ey* ‘in the middle of (doing something)’, *sai-ey* ‘among (people/place/time)’, and *swunsikkan-ey* ‘in an instant’ are commonly used.

In order to understand how one postposition performs multiple functions, some recent studies have examined the relations among functions using image schemas, oftentimes by setting one or more central concepts and showing how prototypical meanings (or proto scenes) are expanded to other peripheral meanings (e.g., Kang, 2012; Türker, 2005; Maeng, 2016; Lee, 2002, 2004). For instance, Türker (2005) set four major domains (*Locational*, *Temporal*, *Abstract*, and *Logical*) each of which were comprised of a few numbers of *senses*. Kang (2012) also suggested three clusters (*Location*, *Goal* and *Inclusion*) from which other functions were expanded. In their studies, Türker and Kang indicated how multiple functions of *-ey* were derived from major domains/clusters. The terminologies and categories might be different, but overall, studies that investigated the functions of *-ey* generally agreed that the locational function (including goal) is a distinctive feature of *-ey*.

In addition to such rich discussions about the functions of *-ey*, empirical studies utilizing corpora show how this postposition has been used in terms of syntactic structure, concerning verbs and proportions of functions. Nam’s (1993) research was one of earlier works that analyzed corpora in relation to the use of *-ey*. Using the Yonsei written corpus, which was composed of about one million *eojeols* from fictional writings, essays, newspaper and magazine articles at the time of research, Nam analyzed half of the corpus and extracted 3,875 examples of *-ey*. The study first suggested several functions such as *location*, *about (something)*, *object*, *standard* and *reason*. It also studied the structural characteristics of each function such as

possibility of omission, possibility of substitution with other postpositions, and construction types of sentences based on transitivity. For example, for the locative function, both the intransitive construction NP-*ka* NP-*ey* Verb (*Mina-ka hakkyo-ey ka-ta* ‘Mina goes to school’) and transitive construction NP-*ka* NP-*lul* NP-*ey* Verb (*Minho-ka yelsoy-lul chayksang-ey twu-ta* ‘Minho puts keys on the table’) are allowed. What makes Nam’s study noteworthy is that it examined what types of verbs can be used in each category and listed them, which will be further discussed in Chapter 5.

Türker (2005) showed the proportional relation of functions through her analysis of the *CetCon* corpus created by Korea University, which contained over one million tagged words at the time of her study. After selecting a portion of the data that contained about 4,000 sentences, Türker chose sentences that included *-ey* and analyzed the functions displayed in these instances. The study listed the semantic definitions of *-ey* that made up the highest proportion of sentences in the corpus. The most frequent sense was *proto-goal sense* (23%) followed by *proto-location sense* (18%), *temporal location sense* (14%), and *abstract goal sense* (11%). The results indicated that concrete locational functions (both *proto-goal* and *proto-location*) occupy the highest proportions of usage and inclusion of the abstract goal sense composed more than 50 % of the total usage, which proves the major use of *-ey* as a locational postposition.

A study from Cho (2018) differentiated itself from others in that it analyzed not only the function of *-ey* but also the types of predicates that co-occur with it. The study explored the co-occurrences of *-ey* as a locational postposition with predicates in a corpus containing data from 17th to 18th century Korean literature. Through an analysis of 8,410 usages, Cho found that frequently attested functions (those that had 1,000 tokens or more) were *time* (1,800 tokens), *condition/manner* (1,518 tokens), and *location* (1,449 tokens). Moreover, Cho divided the verbs

into two types based on transitivity and figured out that transitive verbs appeared more frequently than intransitive verbs.

In sum, results of studies based on corpus-based analysis suggest that the locative function occupies a large proportion of usage for *-ey*. In addition, investigations on verb use in those studies showed that the type of verb selected worked as a major feature of the linguistic environment for *-ey*, which helped to more specifically define sentential characteristics.

3.1.2. *-eyse*

The most fundamental function of *-eyse* is to note a place within which dynamic action expressed by a co-occurring predicate takes place. For instance, the library where the action of reading happens is marked by *-eyse* in (3).

- (3) *ku-ka tosekwan-eyse chayk-ul ilk-ta*
he-NOM library-*eyse* book-ACC read-SE
'He reads a book in the library.'

This postposition also indicates a starting point when the action described by a verb involves a change of place such as *ka-* 'to go' or *o-* 'to come'. It is also used to mean *source*, as in *chencang-eyse nao-ta* '(it) comes out from the ceiling', while some studies do not distinguish a starting point from a source (e.g., Yi, 1993; Choo & Kwak, 2008). Starting points and sources can be expanded into more abstract concepts, sometimes adding a meaning of *reason*. In the case of *hokisim-eyse sicakhay-ss-ta* 'It started from/out of curiosity', *hokisim* 'curiosity' is mental status, which is considered to be an abstract starting point/source where 'it' began or a reason for the happening of it. In general, *dynamic location* and *starting point* are considered to be major

functions of *-eyse*. Temporal use of *-eyse* as in *kutul-un yel sal-eyse yel twu sal sai-uy ai-tul-i-ta* ‘their ages range from ten to twelve’ is also considered an expansion of the concept of starting point (Yi, 1993).

Other functions include *comparison/standard* and *indicating subject*. For *comparison/standard*, Ko and Ku (2008) note that these two are identical since comparison is based on standard. Some studies have also termed it as *background* or *scope* (Kang, 2012; Kang, 2015) as it designates the boundaries for cognition limit. For instance, in phrases such as *i pyo-eyse po-tus* ‘as you can see in (from) this table’ and *i sosel-eyse tulena-tus* ‘as can be seen in this novel’, the scope of recognizing ‘seeing/observing/understanding’ is limited to the table and the novel where *-eyse* is attached. The *-eyse* postposition is sometimes used to indicate a subject when a subject refers to a group or an organization. Yi (1993) explains this occurrence as the result of the agent functioning as an energy source from which energy flows. Thus, agent can be understood and marked as the energy source using *-eyse*. Table A-2 in the Appendix section summarizes the functions of *-eyse*.

Similar to the studies on *-ey*, some studies have also analyzed the functions of *-eyse* by setting key clusters/domains and examining the underlying connections among them. Kang (2012) considered *Bounded landmark cluster* (‘landmark’ as a background place where an action takes place) and *Source cluster*, while Türker (2005) set four domains for *-eyse* (*Locational*, *Temporal*, *Abstract* and *Logical*) as major concepts from which other functions of *-eyse* would be derived. In both studies, the locational function of *-eyse* as a place where an action occurs or as a source something starts from is regarded as a distinctive feature of *-eyse*. Türker further analyzed the distribution of *-eyse*’s functions in a study that focused on one part (about 4000 sentences) of the *CetCon* corpus and selected sentences that contained *-eyse*. The results showed

that *locational sense* made up the largest percentage (27%). Other semantic senses that made up more than 15 percent of the entire data set were *proto-source sense* (20%) and *abstract location sense* (18%).

The number of studies on *-eyse* is relatively small compared to *-ey*, possibly because i) the functions of *-eyse* are less complex than *-ey*, and ii) *-eyse* can combine with any predicate that involves movement (e.g., any verb that combines with the light verb *ha-* ‘to do’), thereby making it more flexible than *-ey* in choosing its predicate. Thus, the focus of many studies on *-eyse* has been to compare *-eyse* with *-ey* by examining instances in which both *-ey* and *-eyse* are allowed as well as instances in which only *-ey* or *-eyse* is allowed. These studies will be further discussed in section 3.2.

3.1.3. *-(u)lo*

The functions of *-(u)lo* are summarized in Table A-3 in the Appendix section. Among the multiple functions of *-(u)lo*, *direction/path* and *instrument* are considered to be two of the major functions. As a directional postposition, *-(u)lo* means ‘toward/to/for’ when attached to a noun of location as in (4).

- (4) *honollwulwu-lo ka-ta*
Honolulu-*lo* go-SE
‘(someone) goes to Honolulu.’

Another function of *-(u)lo*, *result of change* is also related to the direction in that the result works as a destination of change. The postposition is often used to mean *path* as well, which differs from direction in that the noun to which *-(u)lo* is attached is not a destination, but just indicates

which path the agent used. As for *-(u)lo*'s role as an instrumental postposition, typical usage involves its co-occurrence with concrete nouns such as *peyn* 'pen' in *peyn-ulo kuli-ta* 'to draw with a pen'. Some studies introduce *means* (*cha-lo ga-ta* 'to go by car'), which is included with the instrumental function in Table A-3 in the Appendix section. *Material* (*namwu-lo mantul-ta* 'build with wood') and *manner* (*culkew-un maum-ulo ilha-ta* 'work with a happy attitude') can also be largely categorized into the instrumental function.

Time nouns also commonly occur with *-(u)lo* as in *sikan-i seysi-lo hyangha-ta* 'time is approaching to three o'clock', *onul-lo samil-i-ta* 'today is the third day', and in fixed expressions such as *ttayttay-lo* 'from time to time' and *taytay-lo* 'generation by generation'. Other functions of *-(u)lo* include *reason*, *qualification*, and *decision*. Two other functions, *object for causative and comitative* and *way of thinking*, are close to idiomatic expressions in that the former mostly occurs with *hayekum* 'making someone to do something' and the latter involves cognitive verbs (such as 'think' or 'regard'). Some other examples of fixed expressions with *-(u)lo* are *-(u)lo inhay/malmiama* 'because of', *pwupwunce-ulo* 'partly', *ilpancek-ulo* 'generally' and *-(u)lo pwase* 'judging from'.

Regarding investigations into the core function of *-(u)lo*, several studies claim the prototypical meaning of *-(u)lo* as *directional*, with the implication of *path* (Park, 1999; Jeong, 2010; Kang, 2012). To illustrate, Park (1999) argued that meanings of *-(u)lo* expanded from the central *locational domain*, which encompasses path, direction and aiming point, into other functions such as *instrument*, *path of time*, *change of state* and *qualification*. Jeong's (2010) simulation model of *-(u)lo* also set *path* as a prototypical meaning since all other functional expansions can be explained from it. Kang (2012) suggested an image of a *proto-scene* of *-(u)lo*

with an arrow indicating direction/path. In sum, studies that examined the polysemic nature of *-(u)lo* generally concluded that ‘directional/path’ was the primary sense of *-(u)lo*.

In relation to the proportion of functions, Jeong (2010) further analyzed 100 instances of *-(u)lo* from the National Korean Language Dictionary and the Sejong corpus managed by NIKL. The most frequent function was *instrument* (26 instances) followed by *target point/direction* (13 instances) and *choice/decision* (13 instances). Jeong added that while the frequency of instances of ‘path’ itself was small (3 instances), the function still plays a critical role as its underlying concept is embedded in other functions.

Nam’s (1993) analysis of the Yonsei written corpus included 7,276 instances of *-(u)lo*. As was the case for *-ey*, the research showed functions of *-(u)lo* that include *aiming point/direction*, *change*, and *materials/cause/instrumental*. It also investigated intransitive and transitive constructions with *-(u)lo*: *Minho-ka hakkyo-lo ka-ta* ‘Minho goes to school’ (NP-*ka* NP-*(u)lo* Verb) and *Minho-ka senmwul-ul hankwuk-ulo ponay-ta* ‘Minho sends the present to Korea’ (NP-*ka* NP-*lul* NP-*(u)lo* Verb) are intransitive and transitive constructions that both contain *-(u)lo*. The verbs that can be used for each function were also suggested in this study, which will be further discussed in Chapter 5.

Previous studies on the three postpositions of *-ey*, *-eyse*, and *-(u)lo* examined their polysemic functions, with some studies suggesting how those functions are related with other functions. What the studies generally agree on is that the locative function works as a prototypical function of all three postpositions. Meanwhile, specific meanings in relation to location involve existence, movement, and direction/path, which show the unique realms of each postposition as locational (Nam, 1993; Sohn, 1999; Choo & Kwak, 2008; Ko & Ku, 2008; Song, 2006; Kang, 2015; Lee, 1981; Park, 1999; Kim, 1998; Lim, 2017). As Nam investigated and

listed the types of verbs that can appear for the realizations of each function, co-occurring verbs for all three postpositions were specified and differentiated from each other (Yi, 1981). For this reason, the current study particularly focuses on the locative functions of *-ey*, *-eyse*, and *-(u)lo* in terms of their co-occurring verbs.

3.2. Many-to-one mapping: *-ey*, *-eyse*, and *-(u)lo* for the locational function

All three postpositions can be combined with nouns to indicate a location, expressing either static/dynamic location or direction. Due to this overlapping function, a number of previous studies compared and examined two or more postpositions together (Lee, 1981; Lee, 2004; Kim, 2009; Chae & Lim, 2013; Jeong, 2011; Lim, 2017). The studies that analyzed all three postpositions together mostly investigated the functions of each postposition, as discussed in 3.1. The main focus of studies that compared two postpositions was to diagnose the roles of each postposition by investigating cases where two postpositions are accepted in the same sentence or only one of the postpositions are allowed in a similar structure. As those instances occur between *-ey* and *-eyse* or *-ey* and *-(u)lo*, the following sections will review research that studied *-ey* and *-eyse* first and then *-ey* and *-(u)lo*.

3.2.1. *-ey* and *-eyse*

A general consensus is that *-ey* is used to denote an existential place while *-eyse* indicates a place for an activity, as an adverbial phrase with *-eyse* (not *-eyse* itself) increases the sense of dynamicity (National Institute of Korean Language, 2005; Chae & Lim, 2013). Yi (1981, 1993)

illustrated the difference between the two postpositions through sentences such as those in (5) and (6)⁴:

(5) Distinction between *-ey* and *-eyse*: *-ey* case

emeni-ka naympi-ey/eyse pap-ul cis-ta*
mother-NOM pot-*ey/eyse** rice-ACC cook-SE
'Mother cooks rice in the pot.'

(6) Distinction between *-ey* and *-eyse*: *-eyse* case

emeni-ka pwuekh-eyse/ey pap-ul cis-ta*
mother-NOM kitchen-*eyse/ey** rice-ACC cook-SE
'Mother cooks rice in the kitchen.'

In (5) where only *-ey* is allowed, *-ey* is attached to the noun *naympi* 'pot', showing the relationship between the entity (rice) and the ground (pot). In (6), on the other hand, *-eyse* occurs along with the noun *pwuekh* 'kitchen', a place where the process of cooking occurs, indicating general and whole background (Yi, 1981). This point of view about *-eyse* which claims its ability to encode general background is in line with other studies (Lim, 2017; An, 1982; Hong, 1983), which also indicate that *-eyse* focuses on the interior of the locational boundary (Lim, 2017; Choo & Kwak, 2008). These viewpoints can explain the general idea that *-eyse* is attached to a place noun where a part (or whole) of the event is occurring within the interior of the place, marking the place as a background. One thing to note here is that this difference does not mean that only dynamic verbs occur with *-eyse*, or stative verbs with *-ey*, as was the case in (5) and (6).

⁴ Adapted from Yi (1981, pp. 27-28). English translation added.

However, there are instances in which both postpositions are accepted in the same structure. For example, both *-ey* and *-eyse* can be used with *sal-* ‘to live’ as in *ku-nun Seoul-ey/eyse sal-ta* ‘He lives in Seoul’. One possible explanation for this is that existence is a necessary condition for an activity to occur. That is, verbs that mean ‘to live’ such as *sal-*, *memwulu-* and *cheylyuha-* have both meanings of ‘existence’ and ‘activity’ (National Institute of Korean Language 2005; Chae & Lim, 2013). Scrutinized examination of such instances in which both postpositions were accepted was conducted by Lim (2017) with an example in (7)⁵.

(7) Acceptance of both *-ey* and *-eyse*

kukcang-ey/eyse pwul-i na-ss-ta
 theater-ey/eyse fire- NOM occur- PST-SE
 ‘A fire occurred in the theater.’

Lim interpreted the place (theater) as background when *-eyse* was attached. That is, in the case of *kukcang-eyse* ‘in the theater’, an incident (fire) occurred within the interior part of the location. On the other hand, when *-ey* is used (*kukcang-ey* ‘in the theater’), the place works as scenery (a non-interior place) and indicates that the fire occurred in the theater and not any other place.

The same interpretation applies to the verb *sal-* ‘to live’. With the postposition *-ey*, *Seoul-ey sal-ta* ‘to live in Seoul’ shows a simple location where a person resides, which is Seoul (not any other place). In the case of *Seoul-eyse sal-ta* ‘to live in Seoul’, it implies ‘living in Seoul while doing some activities within the interior part of Seoul’ (Lim, 2017). This explains the assertion that manner adverbs such as *hayngbokhakey* ‘happily’ changes the felicitousness of the sentence with *-ey*. Since manner adverbs describe how a person behaves, only *-eyse* is acceptable

⁵ Adapted from Lim (2017, p. 112). English translation added.

as in *Seoul-eyse/*ey hayngbokhakey sal-ta* (Kabata & Lee, 2004; National Institute of Korean Language, 2005). This approach supports the generalization that *-ey* puts emphasis on an existential place and *-eyse* on a place for behavior/activity.

Kang's (2012) corpus-based study suggests different perspectives on *-ey* and *-eyse*'s co-occurrence with *sal-* 'to live'. The study randomly selected 2,000 sentences that contained *sal-* 'to live' with different inflectional forms from the *kkokkoma Sejong malmwungchi* (corpus). Out of 2,000 sentences, the number of tokens that occurred with *-ey* along with landmark (place) (i.e., excluding tokens like *caknyen-ey sal-ta* 'to live (in some place) last year') was 336. The number of tokens with *-eyse* with landmark was 190. What Kang found was that language users preferred either *-ey* or *-eyse* depending on the tense: for the present progressive tense of *sal-* 'to live', the ratio was 17.9% (*-ey*) to 9.5 % (*-eyse*), while for the past tense, *-eyse* was more commonly used (20.5%) than *-ey* (5.7%). As the relationship between time and space is generally noted in cognitive linguistics (Lakoff & Johnson, 1980), the attributes of present tense are more 'physical, sensible, and proximate' while past tense is closer to 'a speaker's lack of control, less visual salience, emotional distance, etc.' (Tyler, 2012). Based on this approach, Kang concluded that the use of *-ey* and *-eyse* with the verb *sal-* 'to live' was related to time and spatial recognition. Lee (2004) suggested similar results based on a survey of native Korean speakers that found places indicated by *-ey* had more proximity or contact with the speakers, while *-eyse* was preferred for non-proximity/contact place.

In sum, while locative *-ey* and *-eyse* each have their own unique roles as locative postpositions, both of them are acceptable in some cases. The use of both postpositions with *sal-* 'to live' was one of the most frequently studied cases that accepted both postpositions. Studies on those cases in general reveal that a place with *-eyse* indicated a location where dynamic

actions took place inside of it and worked as background, whereas a place with *-ey* denoted a location as a point or as a scene. The overlapping appearances of *-ey sal-* and *-eyse sal-* ‘to live in’ can be explained with this approach, even though it is generally considered that the difference is negligible with little semantic difference (Choo & Kwak, 2008) or there is nuance of dynamicity in the case of *-eyse* (Sohn, 1999). Recently, corpus-based studies have been showing that cognitive proximity (e.g., tense, the degree of contact) may affect the choice of these two cases. Future investigations that utilize corpora as foci of analysis are expected to provide a more detailed description on this issue.

3.2.2. *-ey* and *-(u)lo*

Another relevant topic is the difference between *-ey* and *-(u)lo*. There are cases in which both postpositions are interchangeable when they are attached to a place noun. For instance, *-(u)lo* can replace *-ey* in *John-i hakkyo-ey ka-ta* ‘John goes to school’. As explained in Section 3.1, some roles of *-ey* and *-(u)lo* overlap (e.g., reason, direction) as well. Ko and Ku (2008) explain that overlapping functions appeared as *-ey* and *-(u)lo* expanded their prototypical meanings. Accordingly, some recent studies have focused on defining the differences embedded in the prototypical concepts of *-ey* or *-(u)lo* in order to characterize them.

Several studies argue that the difference between these two postpositions lies in whether the concept of contact is involved (*-ey*) or not (*-(u)lo*) (Maeng, 2016; Lim, 2017; Kim, 1998). Maeng (2016) examines the prototypical meaning of *-ey* from five concepts of *contact: surface contact, attachment contact, containing contact, neighboring contact, and dynamic contact*.

Those contacts show the positional relationship between trajector (TR) and landmark (LM)⁶. Among them, two specific types of contacts are introduced in the *surface contact: horizontal surface contact* and *vertical surface contact*, whose names originate from the contact line where TR and LM meet. For instance, in example (8), the contact line between TR (Mina) and LM (chair) is horizontal. Likewise, in (9), the contact line between TR (Yumi) and LM (door) is vertical. Maeng explains the functions of *-ey* with these five main contacts and their expanded versions.

(8) *horizontal surface contact of -ey*

Mina-ka uycha-ey anc-ta

Mina-NOM chair-ey sit-SE

‘Mina sits on the chair.’

(9) *vertical surface contact of -ey*

Yumi-ka mwun-ey kitay-ta

Yumi-NOM door-ey lean-SE

‘Yumi leans on the wall.’

Lim (2017) also introduces similar concepts to differentiate *-(u)lo* from *-ey*. That is, *-ey* is used when a notion of contact is involved, while *-(u)lo* does not involve contact. In other words, *-(u)lo* is used to indicate general direction or a place where TR did not exist when the action in

⁶ According to Langacker (1987), there always occurs an asymmetry when profiling participants in a sentence. The “figure within a relational profile” is called the *trajector* while “other salient entities in a relational prediction” are called *landmarks* (p. 217). For instance, in English expressions such as *X equals Y* or *X resembles Y*, X is a trajector because X is the entity that is being assessed while Y is a landmark that indicates the reference point in predicting the relationship (p. 231). This network of relations among participants can be expressed with *image schemas*, which represent bodily movement and have an internal structure that show what they mean and how they are related (Beitel, 1995).

the predicate began to occur (Kim, 1998). Let's consider *hakkyo-lo ka-ta* 'to go to school' to illustrate. Unlike *hakkyo-ey ka-ta* 'to go to school', TR (school) did not exist at the moment when the action of 'going' occurred. As a result, the place with *-(u)lo* has more distance from the agent with no concept of contact, while the place should be conceptually precise (or be in contact or close) for *-ey* to be used. Thus, for sentences such as those in (10)⁷, *-ey* is not acceptable because the precise destination point is not clear with the noun *olunccok* 'right (side)' (10a) and *aphathu* 'apartment' (10b), unless the apartment was contextually specified earlier in 10b).

(10) Acceptance of *-(u)lo* only

a) *olunccok-ulo/*ey ka-ta*

right-ulo/*ey go-SE

'(someone) goes (turns) right (direction).'

b) *aphathu-lo/*ey isaka-ta*

apartment-lo/*ey move-SE

'(someone) moves to an apartment.'

However, the fundamental use of *-(u)lo* with non-specified nouns does not mean that a place used with *-(u)lo* cannot be a destination. When the concept is expanded, *-(u)lo* can be also used when crossing a boundary, as does happen in *cip-ulo tuleka-ta* 'to enter/step in the house'.

Overall, while locative *-ey* and *-(u)lo* are interchangeable in some instances, there is a subtle difference in cognitive representation of their meanings. Studies propose that *-ey* is initiated from the concept of contact, which is the reason for marking a destination with *-ey*, while *-(u)lo* shows general direction. This approach is in accordance with conventional

⁷ Adapted from Choo and Kwak (2008, p. 256).

understanding of the difference between the two postpositions. In relation to the location, it is generally said that starting point and destination are denoted with *-eyse* and *-ey* each, while the ‘path toward the destination’ that connects two points is marked by *-(u)lo* (Yi, 1981; Sohn, 1999; Jeong, 2011).

Studies on many-to-one mappings and each postposition demonstrate that locative usage of these three postpositions reveal cognitive distinctions among them by language users.

Although their approach and implications are remarkable, there are some limitations. First, the number of studies that analyzed corpora is limited, and/or the size of the corpora were relatively small. For instance, Türker (2005) did not specify how many sentences contained *-ey* among the 4,000 sentences she selected and analyzed for her research. Second, frequency, which is a key factor in language acquisition, is not fully considered. Nam’s (1993) research is noteworthy for its extensive listing of possible verbs that can occur with *-ey* and *-(u)lo* but it did not delve into what types of verbs are more frequently used than others. Finally, the investigation about the linguistic environments within which the target postposition occurs has not been the main concern of research. As Lee (1983) and Ko (2011) mentioned, the meanings of locational postpositions do not emerge by themselves, but they are determined as a result of their relations with the preceding noun or the following predicate with which such postpositions are combined. Specifically, the use of predicates is critical in selecting which postposition is to be used. Considering all these together, I examine what types of predicates occur frequently with each postposition in a much larger corpus of contemporary Korean.

3.3. Acquisition of *-ey*, *-eyse*, and *-(u)lo*

As was briefly introduced earlier, L1 Korean users employ a considerable number of postpositions in order to construct grammatical relations in a sentence. L1 children also rely on grammatical information carried by postpositions for sentence comprehension (Shin & Deen, 2019). In relation to the acquisition age of postpositions by L1 users, Lee, Chang-Song, Choi, and Lee (2008) indicated that major postpositions were acquired by the age of three, while Kim and Hong (2001) showed in a story re-telling task that the most frequently used postpositions were locative ones in a broader sense (*-ey*, *-eyse*, *-(u)lo*, *-hanthey*, *-eykey*, *-poko*, *-tele*) by 4-year-old children and that the use of other types of postpositions (e.g., nominative, accusative) increased by the age of eight.

Regarding the acquisition of specific postpositions by L1 speakers, studies mostly investigated general acquisition order rather than the three locative postpositions themselves. Results from these studies are inconclusive, depending on the design or length of longitudinal studies. For example, Cho (1982) suggested comitative (*-lang*, *-hako*, *-kathi*) postpositions were acquired first followed by locative (*-ey*, *-(lo)*, *-hanthey*) and nominative (*-ka*), while Kim's (1997) findings showed that the locative *-ey* and topic/contrast marker *-(n)un* were the first two types of postpositions to be acquired.

Characteristics of errors in postposition usage by L1 children were also examined. Kim and Hong's (2001) results showed that the rates of erroneous use of locative and accusative postpositions were higher than any other types of postpositions, while the most common type of errors was substitution (using another postposition instead of the appropriate one). They indicated that the possible reason for higher rates of error for locative postpositions might be the use of diverse locative forms and the polysemous nature of such forms (i.e., many-to-one

mappings). In general, studies on the acquisition of postpositions by L1 users show that locative postpositions are acquired earlier and have higher error rates.

Due to its complexity, the acquisition of the Korean postposition system poses a significant challenge for L2 Korean learners who have different postposition systems or completely lack any equivalent of a postposition system in their L1s. Studies show that one of the most frequent error types for L2 learners of Korean is postposition errors (Lee, 2003; Ko et al., 2004) and advanced learners still produce considerable error rates for postpositions in both spoken and written tasks (Han, 2014; Min, 2002; Lee, 2003). In this respect, inquiry on how learners of Korean acquire and use postpositions has been one of the main topics of research for acquisition of L2 Korean.

Broadly, main topics of studies on acquisition of postpositions by L2 users have been either acquisition order or error types/rates. Hwang (2002) audio-recorded interviews from 111 adult learners (76 English-speaking and 35 Japanese-speaking learners) to examine their acquisition order of postpositions. The postpositions that both English and Japanese speakers acquired in an early stage included nominative (*-i/ka*), locative (*-ey* and *eyse*) and accusative (*-(l)ul*) markers.

Other studies on postpositions analyzed production errors by L2 learners. Studies indicate that three postpositions *-ey*, *-eyse*, and *-(u)lo*, which are the foci of the current dissertation, show high rates of errors. To illustrate, Kim (2002) collected learners' writing data from three universities in Korea and analyzed them in relation to proficiency levels. From proficiency level one (the lowest) to six (the highest), the postposition category that showed the highest error rate (# of errors in the category /# of total errors) in each level consisted of adverbial postpositions (e.g., *-ey*, *-eyse*, *-(u)lo*, *-eykey*, *-chelem*, *-kkaci*) in all levels except for the 3rd level. Han's (2014)

results were similar to Kim. The research analyzed 23,246 *eojeols* of spoken discourse by advanced learners of Korean. Among 738 postposition errors analyzed, adverbial *-ey* (12.6%) showed the second highest error rate following nominative *-i/ka* (17.2%). Among diverse error types (omission, substitution, addition, etc.), the postposition that showed the highest substitution error rate was *-eyse*, which was mostly replaced with *-ey*. Kim and Lee (2004) also analyzed a composition corpus formed from the production data of learners at three different proficiency levels, whose results also suggested that *-ey* and *-eyse* were most commonly replaced by *-eyse* and *-ey*, respectively.

As error rates and rates of substitution with other similar forms of postpositions were high in regard to locative postpositions, some studies specifically focused on the developmental aspects of them for L2 learners (Lee & Ko, 2013; Kang, 2015). Among them, Kim and Guo (2016) recruited eighty L2 speakers whose native language was Chinese along with twenty L1 speakers as a control group. They designed two speech tasks, describing pictures and summarizing the plot of a silent movie, inducing different functions of *-ey*. The results showed that the rate of correct usage on average in the two tasks was 72.72%, with a 50.52% accuracy rate for ‘location’ and 82.05% for ‘destination’ functions. The study analyzed the results by level as well. As for the ‘locational’ function, the accuracy rate increased from 23.08% (level 3: intermediate) to 58.71% (level 6: advanced), while the ‘destination’ function showed 62.05% (level 3) to 100% (level 6) accuracy rate of usage. The study suggested that even advanced learners experienced difficulties in acquiring competent knowledge about the locative use of *-ey*.

As for the usage errors of the three postpositions *-ey*, *-eyse*, and *-(u)lo*, studies have mainly focused on *-ey* and *-eyse*, with little attention to *-(u)lo*. Two factors might have contributed to this tendency. One possibility is that the production rate of *-(u)lo* itself was low.

Therefore, when researchers conducted an initial experiment to see how L2 learners of Korean understood postpositions, the attention of researchers was most likely focused on other postpositions with high production and error rates. Another complication was that the distinction between two functions of *-(u)lo* (instrumental and locational) was not clearly made in the studies. Learners might have produced a number of utterances containing *-(u)lo* with relatively low error rates (e.g., Han, 2014), but the postposition might have been used more frequently as an instrumental. As a result, the low error rate might not reflect the learners' understanding of *-(u)lo* as a locational. Overall, the investigation about learners' use of *-(u)lo* as a locational postposition needs to be further studied.

While a number of studies examined *-ey*, *-eyse*, and *-(u)lo* in Korean linguistics, the number of studies on the environment of the locative functions is limited, which brings up the necessity of investigation into their unique linguistic environments and speakers' employment of the shared yet divergent functions of such postpositions. Moreover, the results of studies about postposition acquisition by L2 learners showed that although locative postpositions were acquired in an early stage of learning, they occupied a high proportion of postposition errors. The confusion between *-ey* and *-eyse* was observed in many instances and continued with advanced learners (Han, 2014; Kim, 2004; Kim, 2013; Kim & Lee, 2004). While acquisition order and error rates of postpositions by L2 learners of Korean show one aspect of developmental characteristics, little has been investigated as to the effect of input (either frequency or other aspects of input) on the acquisition of Korean as an L2 (Türker, 2017). As properties of input (such as frequency and the environment where the input is manifested) are central to the acquisition of postpositions (Hwang, 2002), the present study aims to further investigate them in both L1 corpora and textbooks for learners of Korean.

IV. TEXTBOOKS FOR L2 LEARNING

The general consensus about the critical role of input in usage-based approaches to language acquisition necessitates the systematic profiling of characteristics of input that second language learners receive. That is, beyond the overall comparison of learner corpora with L1 use, studies are called upon precisely to investigate what types of input are provided to language learners, how these types are manifested in learning contexts, and how they affect learners' understanding of target language knowledge. Although an analysis of all imaginable input is nearly impossible, a textbook for learners is regarded as a major source of learner input (Römer, 2004; Tono, 2004).

The discussion about input is related to the issue of authenticity to some degree, which has been one of the main topics of discussion in SLA literature. An authentic text refers to "such a text which is written for native speakers of the language to be read by other native speakers" (Swaffar, 1985, p.17). It is understood as "a goal that teacher and students have to work towards, consciously and constantly" (Van Lier, 1996, p. 128). Some discussions about authenticity in language learning/teaching contexts include i) to what degree authentic materials should be utilized in the textbook, that is, whether modification of original contents is necessary (e.g., Fox, 1987; Widdowson, 1998), and ii) how to develop, design and use authentic materials pedagogically (e.g., Mishan, 2005). The purpose of the present dissertation is neither of these cases. Rather, by comparing distributional features of linguistic patterns both in L1 and in L2 corpora under the same framework, this project aims to understand the extent to which textbooks reflect L1 use in daily life, and to suggest a better direction for future textbooks to be developed. In the meantime, vigorous studies on authenticity reveal that one core aim of L2 acquisition is to attain usage that approximates L1 users' language behavior while pursuing authenticity as an

ultimate goal of language learning. In this regard, it is not uncommon for textbook analysis to often be combined with and compared to the analysis of target language corpora.

Literature on textbook analysis has been generally focused on either lexicon or sentence structure. In relation to the use of vocabularies in English textbooks for English as a second/foreign language (ESL/EFL) learners, Alsaif and Milton (2012) examined vocabulary size in English textbooks used in Saudi Arabia in comparison with the British National Corpus (BNC). The study showed that about 84% of the most frequent 2,000 words and 55% of the most frequent 5,000 words appeared in the textbooks. In addition, learners' word-list test results showed a positive relationship between word frequency in the textbook and test scores, suggesting that frequency influences the acquisition of lexicon knowledge. In a series of papers, Shin (2009, 2011) examined the use of copula verbs and basic verbs in English textbooks (which were used for official English language education contexts in Korea) and general-purpose ESL coursebooks (which targeted ESL learners and were sold in Korea) in comparison with L1 storybooks. In general, the ratio of copula verbs in textbooks and coursebooks was higher than that of the storybooks. In addition, textbooks and coursebooks showed different tendencies. The ratio of copula verbs increased while the ratio of basic verbs remained nearly the same across the 3rd and 4th levels of the textbook series. In the case of coursebooks, the ratio of copula verbs decreased while the ratio of basic verbs increased at higher level in the series.

Davis and Face (2006) compared vocabulary usage in L1 Spanish resources and textbooks for L2 Spanish learners. What they found was that the coverage rate of the most frequent L1 lexicon decreased as proficiency level went up in the textbook: lower level textbooks introduced more frequent words than upper level ones. In addition, abstract words or words that

were cognates with English words were under-represented in the textbooks, while nouns or words that represented concrete concepts were overused.

Other studies examined sentence/language structures in the textbooks. In general, the studies indicated that certain types of sentence patterns were used more or less frequently in textbooks than in the L1 corpus (for a more detailed analysis, see Gilsan and Drescher (1993) for language learners of Spanish and Römer (2004) for EFL learners). One of them is Tono (2004), which examined the use of verb Subcategorization Frame (SF) patterns (e.g., how many arguments a verb requires) in three corpora: an interlanguage corpus by Japanese-speaking learners of English, an L1 corpus that mostly consisted of data obtained from English newspapers, and EFL textbooks used in Japan as a target language. It was observed that the frequencies of SF in the textbook had an effect on the over/under use of verbs in the interlanguage. The results also suggested that learners' production of verbs in each group was influenced either by frequencies in the textbooks (for the cases of resultative alternations (*bring* and *take*) and *blame* alternations) or by frequencies in both the L1 corpus and the textbooks (dative, locative, and *as* alternations (*make* and *take*)).

Taken together, studies generally agree on the idea that learners' linguistic knowledge is greatly affected by the ways in which they are exposed to the target language. By analyzing which linguistic features were introduced and how they were presented in the textbooks, these studies proved a positive relationship between input and output. They also addressed the usefulness of using an authentic L1 corpus for developing textbooks that allowed for language learners to be more involved in interpreting the property of language use in the L1, such as what types of lexicon or patterns are used in a certain context and how often they appear.

With the recognition of the critical role of input in language learning, studies that examine how input supplies learners with language knowledge have been increasing within the field of Korean language education. Jeong (2011), for example, examined six types of textbooks: four used in universities in Korea and two published by NIKL for immigrants to Korea. Jeong found that all types of textbooks introduced the three postpositions of *-ey*, *-eyse*, and *-(u)lo* at level 1, as they were basic postpositions used to mark time and location. According to the study, the major problem for these textbooks was their lack of a systematic introduction for each postposition's major functions (e.g., locative, time, path, etc.) and grammatical properties (such as allomorphs). Moreover, there was no clear distinction among seemingly similar functions for each postposition, which caused confusion amongst learners. The study proposed an alternative way of providing instruction about postpositions using image schemas that depicted the expansion of functions from prototypical meanings and the differences between such functions.

While Jeong (2011)'s study included only textbooks in its analysis, other studies considered L1 dictionaries as a reference. Kim (2011) examined 5 types of Korean language dictionaries, three of which were designed for language learners of Korean. Taking all dictionaries together, Kim summarized the functions of *-ey* (14 functions), *-eyse* (7 functions), and *-(u)lo* (14 functions). Through the analysis of three types of beginning level textbooks (two were used in China, and one in Korea), the study concluded that all three textbooks introduced only a limited number of functions. What makes Kim's research different from Jeong's study is the inclusion of L1 dictionaries and the examination of how textbooks are different from dictionaries. However, both studies basically aimed to examine how textbooks introduced the functions of each postposition and pointed out a lack of full explanations about these functions.

Another approach by Lee and Ko (2013) included four types of L1 dictionaries and seven types of L2 textbooks as references to gauge learners' understanding of *-ey*. After extracting 15 functions of *-ey* from references, they conducted a survey of 50 language learners of Korean to assess whether they could match the 15 functions of *-ey* with their respective sentences. The results showed that *-ey* as a time postposition displayed the highest rate of success in matching between the postposition and the sentence (64%), followed by addition/enumeration (56%). The locational and the directional *-ey* reached only a correct matching rate of 30%. With these results, the study argued that learners did not fully acquire all of the functions of *-ey* and that even advanced level learners showed high error rates for major functions. Even though this examination of diverse textbooks/dictionaries is remarkable, the study leaves two questions unanswered. One is whether participants could fully understand the description of the functions in the matching task. While more than 80% (41 people) of the participants had less than 2 years of learning experience, the descriptions in the task were all written in Korean with vocabulary and sentence structures that have been indicated as suitable for advanced learners by the NIKL Center for Teaching and Learning Korean (e.g., *inci* 'cognition', *nayeol* 'enumerating'). The descriptions in Korean might have prevented learners from fully understanding every function. The other is the effect of input on language learning. Results obtained by using textbooks and dictionaries as references to extract functions of *-ey* do not indicate whether there were any effects on learners' understanding of the target postposition.

Kim (2015) attempted to complement these limitations through investigation of the relationship between the characteristics of input and output for postposition acquisition. Kim examined learners' errors involving nominative (*-i/ka*) and accusative (*-(l)ul*) markers, indicating even advanced level learners produced considerable errors. The study examined five types of

textbooks along with questions about postpositions in the Test of Proficiency in Korean (TOPIK). TOPIK questions were included since they have been used for instructional guides. Textbook analysis in the study showed that both postpositions were introduced in beginning level textbooks, albeit the explanations were limited to introducing the allomorphs of each postposition (e.g., for nominative marker, *-i* is used with consonant ending nouns, and *-ka* is used with vowel ending ones). In upper-level textbooks, the target postpositions were used along with advanced-level vocabulary such as *komin-i sayngki-ta* '(a) worry occurred', without further explanation. In relation to the TOPIK questions, the study indicated that postposition questions did not utilize diverse types of verbs. What the study concluded was that textbooks neither explained those markers clearly nor suggested diverse sentence structures. This study is noteworthy in that it investigated how textbooks influenced learners' understanding. Including examinations on how (much) textbooks reflect L1 usage or how TOPIK questions reflect textbooks/L1 usage would have made the results of the study more valid.

Regarding *-ey*, *-eyse*, and *-(u)lo*, Kang (2015) compared their functions in three types of textbooks and seven different reference dictionaries/books. One Korean textbook published in Korea and two textbooks published and used in China were selected. For reference dictionaries and books, two types of dictionaries and five types of grammar books for Korean language learners were chosen. As was the case in the above-mentioned studies, Kang indicated that the textbooks did not introduce all the functions mentioned in dictionaries/books and did not provide enough explanations for each function. Kang also conducted cloze test exercises to understand learners' use of each function and analyzed errors in each item. Through the analysis of these test results, Kang concluded that several factors such as learners' L1, the polysemic nature of each postposition, and overlaps in functions across *-ey*, *-eyse*, and *-(u)lo* seemed to affect L2 learners'

(mis)understanding of these postpositions. While the study did not unequivocally account for a close link between reference corpora and L2 learners' understanding of the postpositions, the study showed an overall picture of instruction and understanding of the three postpositions by encompassing L1 reference sources, L2 textbooks, and learners' production.

Studies for L2 Korean have been raising awareness and stressing the importance of input in language learners' understanding of the target language. Despite the significance of previous studies on textbooks, the implications of these studies are restricted in the following ways. First of all, these studies focused mainly on a qualitative analysis of functions, answering research questions such as how many functions engage in the postpositions, how they are introduced in learners' textbooks, what functions are missing, at which proficiency level the textbooks are introduced, and so on. Consequently, they have not prioritized quantitative analysis of postposition use. As was discussed earlier, frequency plays a key role in reducing processing burden and automatizing production of language. In addition to frequency, all the surrounding properties of tokens (e.g., combinations with a verb, contexts in which they occur, etc.) are also registered in users' minds, which will ease language acquisition. Thus, exploring textbooks from a quantitative approach is a necessary step towards presenting a more holistic picture of the input provided by textbooks, which will show distinctive use for these three postpositions.

One more issue to consider is that textbook-based studies have overlooked the benefits of comparing the use of postpositions in L2 Korean textbooks to that of L1 Korean sources. In previous studies, the scope of utilizing L1 corpora has been limited to generating lists of functions. Davis and Face (2006) hinted at the possible reason for this: textbook analysis for languages other than English was limited until recently due to the lack of representative L1 corpora and frequency data. It is also applicable to the current situation of Korean language

education. As a reference for L2 input, however, L1 corpora must be thoroughly analyzed. Such questions as with which verbs each postposition occurs frequently and how those verbs co-occur with other types of postpositions in L1 will be worth investigating for designing and organizing textbooks for L2 learners.

With all of this in mind, I aim to provide a more detailed picture of how Korean language textbooks use three locational postpositions through quantitative analysis. More specifically, I will be utilizing frequency, keyness analysis and association strength in comparison with L1. These three methods have been particularly considered due to the following reasons: i) properties of L2 input are crucial to establishing L2 learners' cognitive foundation for their knowledge about the target language, and ii) an increasing number of studies have put emphasis on the necessity of teaching postpositions together with other sentential elements that frequently co-occur with them (Woo, 2016; Zhao, 2016; Ko, 2007; Kim, 2011). In addition, a comparison between an L1 corpus and L2 textbooks in terms of these postpositions is expected to show how the characteristics of each postposition in the L1 are reflected in L2 textbooks. For this purpose, four proficiency levels for two types of textbooks are analyzed to see how the use of postpositions change as proficiency level changes in each textbook.

V. METHODS

Studies that adopt usage-based approaches stress the importance of examining authentic instances of language use. The advancement of natural language processing techniques has enabled the analysis of large-scale corpora, which reflect actual language use. As McEnery & Wilson (2001) have stated, a “sufficiently large and representative corpus” (p. 107) leads to an accurate understanding of language use, and the analysis of corpora will provide language users with demonstrations of how various features of languages are employed. From this perspective, this dissertation plans to analyze several types of corpora that exemplify diverse aspects of the Korean language: the Sejong written and spoken Korean L1 corpora as well as two types of textbooks for learners of Korean as a second/foreign language.

The L1 corpora used in this study are based on the Sejong corpora, which were created and publicized by NIKL⁸. Newspaper/magazine articles from different sections and various genres of printed books were included in the Sejong written corpus. In the current study, about 1,124,000 *eojeols* were analyzed, which were randomly selected from *UCorpus* (Lim, Kwon, Kim, & Kim, 2015) provided by the University of Ulsan. *UCorpus* consists of a set of original sentences from the written Sejong corpus and their word-POS tag pairs. I chose *UCorpus* for its tokenized and POS-tagged data, which will simplify the analysis process. For spoken Korean, 200 files were obtained via the NIKL website. These files contain transcriptions of conversations between friends, lectures, presentations, conferences, sermons, television broadcasts of sports events, etc. Close examination of these files revealed that three of them were from children’s books. These three files were removed to prevent extraneous examples of the written properties

⁸ Sejong corpora can be obtained at www.korean.go.kr.

of postpositions. Excluding those three files, a total of 197 files (about 800,000 *eojeols*) were analyzed for spoken Korean data.

Locational phrases in Korean are composed of locational nouns with one of the locational postpositions (*ey*, *-eyse*, and *-(u)lo*) attached. The locational functions of *-ey*, *-eyse*, and *-(u)lo* in the L1 written corpus were first extracted through the following steps. First of all, complex sentences in the L1 written corpus were divided into simple phrases based on the appearance of a verb in the middle of a sentence. For instance, the sentence in (11) is composed of two phrases: *Milo-ka tosekwan-ey ka-ss-* ‘Milo went to the library’ and *keki-eyse swi-ess-ta* ‘(he) took a rest there’.

- (11) *Milo-ka tosekwan-ey ka-ss-ko, keki-eyse swi-ess-ta*
 Milo-NOM library-ey go-PST-and there-eyse rest-PST-SE
 ‘Milo went to the library, and (he) took a rest there.’

In (11), the use of verb *ka-* ‘to go’ marks the formation of a complex sentence. When a POS-tag indicated that an item is a verb within a sentence, the target sentence was divided into two phrases through programming in Python 3.7.

Next, all the phrases containing the three locational postpositions were extracted in a semi-automatic way. More specifically, extraction of the use of three postpositions was completed through Python programming. Adverbial postpositions, such as *-ey*, *-eyse*, *-(u)lo*, and *-eykey/-hanthey* (‘to (a person)’) were POS-tagged as JKB in a POS-tagged data. Thus, phrases that contained JKB tagged words were extracted from all the phrases first. In a group of JKB-included phrases, *-eyse* and *-(u)lo* were searched separately to select instances of use for each postposition respectively. In the case of *-ey*, all tokens that contained the syllable *-ey* (e.g., *-ey*, -

eyse, *-eykey*) were initially chosen and then subsequently narrowed down to only those that contained the specific locational particle *-ey*.

After the extraction of phrases that contained *-ey*, *-eyse*, and *-(u)lo*, manual inspection was conducted to ensure that the tokens were appropriate for the analysis as a locative. The concepts of locatives and verb types were examined in accordance with Nam’s (1993) criteria. Nam broadly suggested six concepts to clarify what ‘locative’ means: *existence*, *action/deed*, *destination*, *arrival*, *contact*, and *source*. To illustrate, a location may include a place where something exists, an action/deed occurs, or someone/something intends to arrive, etc. Nam listed verbs (Table 5.1.) that are commonly used to represent those concepts.

Table 5. 1. Locative concepts and example of verb types in Nam (1993) (translated into English)⁹

	Example of verb types
Existence	<i>iss-</i> ‘to be/exist’, <i>eps-</i> ‘to not be’, <i>wichiha-</i> ‘to locate’, <i>pakhi-</i> ‘to be stuck’
Action/ Deed	<i>sal-</i> ‘to live’, <i>mwuk-</i> ‘to stay’, <i>palsayngha-</i> ‘to occur’, <i>kikeha-</i> ‘to stay’,
Destination	<i>ka-</i> ‘to go’, <i>o-</i> ‘to come’, <i>tani-</i> ‘to attend’, <i>tuleka-</i> ‘to go in’, <i>cipcwungtoy-</i> ‘to be centered’
Arrival	<i>ilu-</i> ‘to reach’, <i>tochakha-</i> ‘to arrive’, <i>totalha-</i> ‘to reach’, <i>michi-</i> ‘to reach’
Contact	<i>tah-</i> ‘to touch’, <i>macwuse-</i> ‘to stand face to’, <i>cepha-</i> ‘to border’, <i>takapwuth-</i> ‘to stick near’
Source	<i>nao-</i> ‘to come out’

Nam’s (1993) locative concepts facilitated further categorization of verb types based on their fundamental meanings. For example, within verbs that co-occur with the locative *-ey*, there are *put*-type verbs (e.g., *cinyelha-*, *pichiha-*, *twu-*, *ssah-*) and *attach*-type verbs (e.g., *tay-*,

⁹ Some verbs in the current study have expanded meanings. For instance, *nao-* ‘to come out’ can be used to mean ‘to quit’ in *ku-ka hoysa-eyse nao-ta* ‘he comes out of the company/he quits the company’. Due to space limit, more generic and commonly used meanings have been utilized in the current study based on the online *Pyocunkwukedaysacen* [Standard Korean Dictionary].

mactay-, *pwuthi-*) (See Table 5.2). The verb lists for each verb type were used as a reference for classifying tokens containing locative functions.

Table 5. 2. Categories and example of verb types used with the locative *-ey* in Nam (1993) (translated into English)

Verb category	Example of verb types
to exist	<i>pwunpho-</i> , <i>camcayha-</i> , <i>sancayha-</i> , <i>se-</i> , <i>cwuceanc-</i>
to stay	<i>memwule-</i> , <i>unkeha-</i> , <i>nathana-</i> , <i>natol-</i>
to fall	<i>ppaci-</i> , <i>camki-</i> , <i>takapwuth-</i> , <i>tteolu-</i> , <i>tteleci-</i> <i>chimmolha-</i>
to arrive	<i>kelchi-</i> , <i>kwihangha-</i> , <i>tatalu-</i> , <i>ilu-</i>
to crash	<i>pwuth-</i> , <i>ieci-</i> , <i>tah-</i> , <i>cepha-</i> , <i>incepha-</i> , <i>kitay-</i>
to appear	<i>kiinha-</i> , <i>kichoha-</i> , <i>pesena-</i> , <i>ipkakha-</i>
to put	<i>ciyelha-</i> , <i>keycayha-</i> , <i>sit-</i> , <i>pichiha-</i> , <i>twu-</i> , <i>seywu-</i>
to define	<i>kyucengha-</i> , <i>kyuhyenha-</i> , <i>silhyenha-</i>
to lift up	<i>olli-</i> , <i>nayli-</i> , <i>titi-</i> , <i>tenci-</i> , <i>tangki-</i>
to attach	<i>tay-</i> , <i>pwuthi-</i> , <i>is-</i> , <i>yenkyelha-</i> , <i>pwuticchi-</i>
to express	<i>nathanay-</i> , <i>palhwiha-</i> , <i>et-</i> , <i>tulenay-</i>
to add	<i>sekk-</i> , <i>tha-</i> , <i>mwuthi-</i> , <i>pipi-</i> , <i>chemkaha-</i> , <i>hapchi-</i>

Nam's (1993) research results were also used as a reference for checking the locational functions of *-(u)lo*. As Table 5.3 indicates, several verb types were recommended for verbs that co-occur with *-(u)lo* as a locative. These included *go*-type verbs (e.g., *ka-*, *o-*, *cinaka-*, *cepetul-*, *kiwul-*, *tteolu-*), *send*-type verbs (e.g., *ponay-*, *olmki-*, *mol-*), *put*-type verbs (e.g., *noh-*, *neh-*, *mwut-*), to express locative *-(u)lo* such as orientation, direction and path.

Of the several functions of *-eyse* (see section 3.1.2 and Table A-2 in the Appendix section), location of an activity and starting point were counted as locative functions. No research has been found that comprehensively examined what types of verbs co-occurred with locational *-eyse*. This is presumably due to the characteristics of locative *-eyse* that allow for

almost any type of action verb to co-occur with it. Thus, the locative concepts and categories in Table 5.1, 5.2 and 5.3 were also applied for diagnosing the locative *-eyse*.

Table 5. 3. Categories and example of verb types used with the locative *-(u)lo* in Nam (1993) (translated into English)

Verb category	Example of verb types
to go	<i>ka-, o-, cepetul-, kiwul-, mollyeka-, tuleka-, kenne-</i>
to be/exist	<i>iss-</i>
to send	<i>ponay-, olmki-, mol-, kacyeke-, tani-, hulli-, mil-, tenci-</i>
to invite	<i>chotayha-, chochengha-, pwulu-</i>
to visit	<i>chacaka-, chacao-</i>
to put	<i>noh-, neh-, mwut-, kamchwu-, kkal-, pye-</i>

Finally, locative verbs examined were collected using Python and lemmatized manually for the analysis. For the present study, only token frequency was counted without consideration of other linguistic markers (e.g., for tense and modality) attached to the verbs. For example, both *ka-* (the present form of ‘to go’) and *ka-ss-* (the past tense form of ‘to go’) were counted as instances of the verb *ka-* ‘to go’. Counting of token numbers were conducted manually.

Extraction of verb types and frequencies in the L1 spoken corpus was mostly conducted manually. As was seen in the case for the L1 written corpus, several stages of manual processes were involved in the extraction of verbs used with the three postpositions (e.g., checking locative functions, lemmatization, etc.). Cases of scrambling were also frequently observed in the spoken corpus as Figure 5.1 (a) shows, which required additional extraction processes. In cases where the sentence was separated into two or more lines in the original corpus as in Figure 5.1 (b), it was not feasible to extract exact verbs for postpositions through automatic extraction.

- (a) <u who="P1"><s n="00003">다시 돌려 앞으로.</s></u> 'Spin again, front-ulo'
- (b) <u who="P1"><s n="00021">항상 이걸 듣다 잠 들어 가지구,</s> 'I always fall asleep while listening to this'
<s n="00022">버스에서 <unclear></s></u> 'bus-eye'

Figure 5. 1. Scrambled order of sentences with locative postpositions (English translation added).

Considering all these conditions, all the extraction procedures of locative verbs in the Sejong spoken corpus such as checking for locative functions, searching for the right verb in sentences with scrambled components, and lemmatization of verbs were conducted manually.

In terms of corpora to be specifically used for language learners of Korean, textbooks were selected as the main subject of analysis. This is due to the fact that textbooks provide an initial understanding of target language systems as a primary source of linguistic input (Kim, 2015; Lee & Ko, 2013; Jeong, 2011; Kang, 2015; Kim, 2011). The current study examines two textbook series: one is commonly used in universities in Korea and some universities in the United States, while the other one is widely used in universities throughout the United States.

The first textbook series is the Sogang Korean (SK) series developed by the Korean Language Education Center at Sogang University. The Korean Language Education Center publishes two volumes of textbooks for each of the first four levels of SK textbooks. From the fifth level, the SK textbooks diversify into components that focus on specific language skills such as reading, speaking, grammar, and vocabulary. The second type of textbook is the Integrated Korean (IK) series published by the University of Hawai'i Press. IK consists of five proficiency levels, for which two volumes have been created. For purposes of consistency in cross-textbook comparisons, the current study analyzes the SK and IK series up to the fourth level, which amounts to eight volumes for each textbook series. A total of 16 volumes was manually converted into a machine-readable format. All of the SK volumes were published in 2018 except for the second volumes of level 3 and level 4 (2017). For the IK textbooks, three

volumes (both volumes for level 1 and the second volume of level 2) were edited in 2015 and the first volume of level 2 was edited in 2012. Two volumes of level 3 were published in 2019 while two volumes of level 4 were published in 2004.

The total amount of data compiled from the textbooks amounted to approximately 90,000~92,000 *eojeols*. Corpus sizes for each level within each textbook series are shown in Table 5.4. SK volumes suggested several listening-oriented practices and activities (e.g., answering True/False questions after listening, summarizing or writing about what they heard) with scripts provided in the appendix for each volume of the textbook. In the current dissertation, those scripts that were provided in the textbook appendixes were included in the analysis, as they served as important material for listening activities in the main lessons. Textbook corpora were tokenized and POS-tagged using *UDpipe* (Straka, Hajič, & Straková, 2016)¹⁰. The extraction of verbs used with the locative functions of the three postpositions in the textbook followed the same procedure as that for the L1 written corpus.

Table 5. 4. Data amount breakdowns for the SK and IK textbooks (Number of *eojeol*)

	Level 1	Level 2	Level 3	Level 4	Total
SK	10,403	22,539	22,761	34,356	90,059
IK	13,708	20,360	19,599	38,977	92,644

The distinction between written and spoken registers was not made in the textbook corpora because of three reasons. First, the criteria for such a distinction may vary depending on the purpose of the analysis. For instance, textbooks can be regarded as written types of input when one considers how the input was originally presented to the readers. On the other hand,

¹⁰ <http://ufal.mff.cuni.cz/udpipe>

conversational type of input in the textbooks can be thought as a spoken register, when the purpose of the context is considered. Second, it often was observed that the topics for both the conversation-type sources (speaking-oriented) and narration-type sources (reading/comprehension-oriented) in one chapter of a textbook were quite similar and utilized similar lexicons, which essentially made them paraphrases of each other. Finally, as will be discussed later, there were cases in which the number of tokens for each postposition in the textbooks was not big enough to divide. For example, the total number of locative *-(u)lo* tokens in the level two and three SK textbooks was 24 and 28, respectively. Dividing such a small number of tokens into written and spoken registers would degrade the overall degree of representativeness and distort interpretations. For these reasons, the textbook corpus was analyzed without distinguishing between distinct registers.

In addition to examining verb types and frequencies within each corpus, keyness analysis was also taken into account. It is a method that analyzes key items in the corpus in order to further identify what the corpus is about. Keyness has been commonly applied to frequency comparison in corpus linguistics (Kilgarriff, 2001, 2005; Leone, 2010; Fraysse-Kim, 2010) and it represents either log-likelihood or Chi-square statistics (Biber, Connor, Upton, Anthony & Gladkov, 2007). The present study uses log-likelihood (*LL*). *LL* is calculated using the observed raw and expected frequency of a target item. In a contingency table (Table 5.5) for instance, ‘a’ and ‘b’ represent the observed number of tokens of a target item (*word*) in Corpus 1 and Corpus 2, respectively. The total number of tokens in Corpus 1 and 2 are ‘c’ and ‘d’ respectively. In this case, the observed numbers of the target item *word* in Corpus1 and Corpus 2 are ‘a’ and ‘b’ ($O_{11}= a$, $O_{12}=b$), while the expected number of the *word* in Corpus 1 is $E_{11}= c*(a+b)/(c+d)$ and $E_{12} = d*(a+b)/(c+d)$ in Corpus 2 (see 12).

Table 5. 5. Contingency for log likelihood (adapted from Kilgarriff, 2001)

	Corpus 1	Corpus 2	Totals
<i>word</i>	<i>a</i>	<i>b</i>	<i>a+b</i>
<i>not word</i>	<i>c-a</i>	<i>d-b</i>	<i>(c+d)-(a+b)</i>
Totals	<i>c</i>	<i>d</i>	<i>c+d</i>

Note. *a, b, c, d* represent frequencies.

$$(12) O_{11}= a, O_{12}=b, E_{11}= c*(a+b)/(c+d), E_{12} = d*(a+b)/(c+d)$$

When *i*: row #, *j*: column #, *O_{ij}* = Observed frequency of *i*th row and *j*th column, *E_{ij}* = Expected frequency of *i*th row and *j*th column

With frequency information in the contingency table, *LL* is calculated using the formula in (13).

The higher the *LL*, the more likely that there is a significant difference between the two frequency values in each corpus. The significance of *LL* (*p*-value) indicates the probability that the keyness of the target item is incidental (Biber et al., 2007). For calculation purposes, zero occurrence was converted to one quadrillionth (0.000000000000000001) (Gabrielatos, 2018).

$$(13) LL = 2 \sum \left(O_{ij} \ln \frac{O_{ij}}{E_{ij}} \right)$$

Note. ln: natural log

As it is likely that the statistical significance shown via *LL* is dependent on the sample size (i.e., weak relationship may reveal statistical significance when the size of corpus is large) (Kilgarriff, 2005; Howell, 2010), effect size was considered as well. Effect size indicates the size of a frequency difference (Rosenfeld & Penrod, 2011). Among the various metrics that can be

used to measure effect size in keyness analysis (e.g., *Ratio*, *Odds Ratio*, *Log Ratio*, *%DIFF*, *Difference Coefficient*) (see Gabrielatos, 2018 for detailed explanation), the present study adopts *%DIFF*, which considers the normalized frequency of items in both the reference corpus and the compared corpus. Normalized frequencies match up with the percentages (%) of each item. The formula is shown below using Normalized Frequency in Corpus 1 (NFC1) and Normalized Frequency in Corpus 2 (NFC2) as in (14) (adapted from Gabrielatos, 2018).

$$(14) \%DIFF = (NFC1 - NFC2) * 100 / NFC2$$

With Corpus 2 (C2) as the reference corpus and Corpus 1 (C1) as the compared corpus (study corpus), a *%DIFF* value of 200 means that C1 employed an item three times as often as C2. For instance, when NFC1 is 300 and NFC2 is 100, the *%DIFF* equals 200 by the formula ($\%DIFF = (300 - 100) * 100 / 100$). There is no limitation for positive effect size (i.e., when there is zero occurrence in C2). The current dissertation marks '> 10,000' in those cases. A negative value represents the underuse of the item in C1 in comparison to C2. For example, a *%DIFF* value of -70 means that the study corpus used the item 70% less frequently than the reference corpus. The maximum negative value is -100 (i.e., when there is no item in C1).

Finally, the present dissertation also considers association strength, which measures how two linguistic items are exclusively related. By examining association strength, this study looked into the degree of co-occurring power between postpositions and verbs. Association strength not only counts the occurrence of one element (outcome) with another element (cue) (e.g., the use of *ka-* 'to go' when a postposition *-ey* occurs), but it also considers the occurrence of one element (outcome) in the absence of another element (cue) (e.g., the use of *ka-* 'to go' with elements

other than the *-ey* postposition). In this way, association strength reveals how a certain item is strongly related with another item. A higher association strength indicates that an outcome has a stronger relationship with its cue.

In order to operationalize association strength, Delta P (ΔP) has been employed as a specific measurement tool. Calculations were made using the contingencies presented in Table 5.6, which shows the number of instances in four cases: i) outcome under cue, ii) no outcome under cue, iii) outcome under no cue, and iv) no outcome under no cue (Ellis, 2006; Ellis & Ferreira-Junior, 2009b). The value can be obtained by deducting the probability of an outcome when there is no cue ($c/(c+d)$) from the probability of an outcome under a cue ($a/(a+b)$), as in (15) (adapted from Ellis & Ferreira-Junior, 2009b). ΔP approaches 1.0 when a cue increases the likelihood of an outcome, showing a positive association between a cue and an outcome. If there is a negative association between a cue and an outcome, it will be closer to -1.0.

Table 5. 6. Contingency for association strength (adapted from Gries et al., 2005)

	Outcome	No outcome
Cue	<i>a</i>	<i>b</i>
No cue	<i>c</i>	<i>d</i>

Note. *a, b, c, d* represent frequencies.

$$(15) \text{ Delta P } (\Delta P) = P(O|C) - P(O|\bar{C}) = a/(a+b) - c/(c+d)$$

Considering the purposes of the current dissertation, the probability of occurrence of verbs (outcome) under cues (postposition) will be the main object of measurement, which will show what types of verbs co-occur with which target postpositions with stronger power. The

reversed case (verb as a cue and postposition as an outcome) will also be measured for certain verbs to examine which postpositions those verbs may prefer.

The scope of examination in this dissertation is limited to the locative functions of three adverbial postpositions *-ey*, *-eyse*, and *-(u)lo*. That is, when a verb *ka-* ‘to go’ occurred with one of the three postpositions as a locative function (e.g., *cip-ey ka-ta* ‘(someone) goes home’), the instance was included in the analysis, while instances such as *Steve-ka ka-ta* ‘Steve goes/leaves’ were not included. Therefore, in the analysis of *-ey*, ‘no cue’ refers to the instances in which *-eyse* and *-(u)lo* are used with a locative function. Examples of tokens that belong to each cell in the contingency table for the measurement of the relationship between *-ey* and *ka-* ‘to go’ are *a: hakkyo-ey ka-ta* ‘(someone) goes to school’, *b: cip-ey iss-ta* ‘(someone) is at home’, *c: cip-eyse ka-ta* ‘(someone) goes/departs from home’, and *d: pang-eyse nolayha-ta* ‘(someone) sings a song in one’s room’. Through such analytical methods, the relationship among the three postpositions and their co-occurring verbs within the context of locative functions will be manifested clearly.

VI. ANALYSIS OF L1 CORPORA: USE OF *-EY*, *-EYSE*, AND *-(U)LO* AS LOCATIONAL POSTPOSITIONS WITH A FOCUS ON VERB USE

As was introduced earlier, the meanings of the verbs that co-occur with those postpositions determine the function of the postpositions: *-ey* for stative location of existence and destination, *-eyse* for a place where an event/action takes place, and *-(u)lo* for path and direction towards the location. Thus, the investigation of verbs with each postposition was conducted to clarify the characteristics of these three locational postpositions.

Table 6.1 shows raw and normed (per million *eojeols*) frequencies of the three locational postpositions (*-ey*, *-eyse*, *-(u)lo*) in the Sejong corpora. Normed frequencies in the table show that the token numbers of three postpositions in the Sejong written corpus was about 2~3 times larger than the spoken corpus. Regarding the numbers of each postposition, *-ey* is used about 30~40% more often than *-eyse*, while the use of *-(u)lo* is comparatively smaller than the other two postpositions in both the written and spoken corpora. More detailed use of the verbs is explained in the following sections.

Table 6. 1. Frequencies of the three locational postpositions in the Sejong corpora

	Raw		Normed (per million <i>eojeols</i>)	
	Written	Spoken	Written	Spoken
<i>-ey</i>	10,024	2,573	8,919	3,217
<i>-eyse</i>	6,724	1,802	5,983	2,253
<i>-(u)lo</i>	2,362	588	2,102	735
Total	19,110	4,963	17,004	6,205

6.1. Verb use with each postposition

6.1.1. -ey

A total of 894 types of verbs were employed for locational *-ey* in the L1 written corpus. Among them, more than 800 verbs occurred less than 20 times, while the most frequent 33 types of verb occupied 50% of total usage. The top 10 verbs most frequently used with locational *-ey* are listed in Table 6.2. The most frequent verbs were *iss-* ‘to be/exist’ (9.46%), followed by *ka-* ‘to go’ (4.37%) and *tuleka-* ‘to enter/to go in’ (2.91%). Two specific functions were salient for the most frequently used verbs: locative-existential and locative-destination. The most frequently used verb *iss-* ‘to be/exist’ represents the former function along with other verbs that characterize behaviors done in a static location including *anc-* ‘to sit’, *se-* ‘to stand’, and *sal-* ‘to live’. Locative-destination was indicated through several movement-oriented verbs such as *ka-* ‘to go’, *tuleka-* ‘to enter/go in’, *ilu-* ‘to arrive’, *o-* ‘to come’, and *nao-* ‘to come out’.

Verbs used with *-ey* in the spoken corpus were also limited to a certain variety. Among the 256 verb types used, only 30 types of verbs occurred more than 10 times. The ten most frequent verbs, which are listed in Table 6.3, composed almost 70% of total usage. The three most frequent verbs were the same as in the written corpus. However, the proportion of each verb in the spoken corpus was larger than the corresponding proportion in the written corpus: *iss-* ‘to be/exist’ was 31.68%, *ka-* ‘to go’ as 14.46%, and 5.36% of *tuleka-* ‘to enter/to go in’. Both the locative-existential and locative-destination functions were also found in the spoken corpus but with slightly different verbs. Aside from *iss-* ‘to be/exist’, verbs like *nam-* ‘to remain’ and *neh-* ‘to put (inside)’ were commonly used to express an existential meaning. For the locative-destination function, similar verbs were used but with differential preferences. For example, *o-* ‘to come’ and *nao-* ‘to come out’ were more preferred in spoken than in written contexts.

Table 6. 2. Frequencies and proportions of verbs used with *-ey* in the Sejong written corpus

Rank	Verb	Token (#)	Proportion (%)	Accumulated proportion (%)
1	<i>iss-</i> ‘to be/exist’	947	9.46	9.46
2	<i>ka-</i> ‘to go’	438	4.37	13.83
3	<i>tuleka-</i> ‘to enter/go in’	291	2.91	16.74
4	<i>anc-</i> ‘to sit’	269	2.69	19.42
5	<i>ilu-</i> ‘to arrive’	251	2.51	21.93
6	<i>se-</i> ‘to stand’	218	2.18	24.11
7	<i>o-</i> ‘to come’	186	1.86	25.96
8	<i>sal-</i> ‘to live’	158	1.58	27.54
9	<i>ppaci-</i> ‘to fall into’	146	1.46	29.00
10	<i>neh-</i> ‘to put (inside)’	144	1.44	30.44
	Others	6,966	69.56	100.00
Total		10,014	100	

Note. Others comprised 884 verb types.

Table 6. 3. Frequencies and proportions of verbs used with *-ey* in the Sejong spoken corpus

Rank	Verb	Token (#)	Proportion (%)	Accumulated proportion (%)
1	<i>iss-</i> ‘to be/exist’	815	31.68	31.68
2	<i>ka-</i> ‘to go’	372	14.46	46.13
3	<i>tuleka-</i> ‘to enter/go in’	138	5.36	51.50
4	<i>o-</i> ‘to come’	132	5.13	56.63
5	<i>nao-</i> ‘to come out’	99	3.85	60.47
6	<i>sal-</i> ‘to live’	52	2.02	62.50
7	<i>naka-</i> ‘to go out’	51	1.98	64.48
8	<i>anc-</i> ‘to sit’	50	1.94	66.42
9	<i>neh-</i> ‘to put (inside)’	49	1.90	68.32
10	<i>nam-</i> ‘to remain’	36	1.40	69.72
	Others	779	30.28	100.00
Total		2,573	100	

Note. Others comprised 246 verb types.

6.1.2. *-eyse*

A total of 1,574 verb types co-occurred with locational *-eyse* in the Sejong written corpus. About 24% of total usage was comprised of the top 15 most frequently occurring verbs, while 1,445 types of verbs occurred less than 10 times. The top ten frequent verbs are listed in Table 6.4. The three most frequent verbs were *ha-* ‘to do’ (4.43%), *nao-* ‘to come out’ (2.94%), and *yelli-* ‘to take place’ (1.98%). As the major function of locational *-eyse* is to denote a place where dynamic activity occurs, the light verb *ha-*, which means ‘to do’, was productively used as it is or in the form of ‘noun-*ha-*’ such as *kongpwu-ha-* ‘to study’ or *il-ha-* ‘to work’ with *-eyse*. Another major function of locational *-eyse* is to indicate a source (starting point), which explains the frequent use of *nao-* ‘to come out’. Other activity related verbs (e.g., *manna-* ‘to meet’ and *chac-* ‘to search’) and verbs that require a source (e.g., *pesena-* ‘to escape/get out of’, *pat-* ‘to receive’, and *na-* ‘to grow/sprout’) were commonly used.

Table 6. 4. Frequencies and proportions of verbs used with *-eyse* in the Sejong written corpus

Rank	Verb	Token (#)	Proportion (%)	Accumulated proportion (%)
1	<i>ha-</i> ‘to do’	298	4.43	4.43
2	<i>nao-</i> ‘to come out’	198	2.94	7.38
3	<i>yelli-</i> ‘to take place’	133	1.98	9.35
4	<i>pesena-</i> ‘to escape/get out of’	109	1.62	10.98
5	<i>manna-</i> ‘to meet’	105	1.56	12.54
6	<i>pat-</i> ‘to receive’	94	1.40	13.94
7	<i>ilena-</i> ‘to occur’	92	1.37	15.30
8	<i>sal-</i> ‘to live’	88	1.31	16.61
9	<i>na-</i> ‘to grow/sprout’	84	1.25	17.86
10	<i>chac-</i> ‘to search’	79	1.17	19.04
	Others	5,444	80.96	100.00
Total		6,724	100.00	

Note. Others comprised 1,565 verb types.

As Table 6.5 shows, the use of locational *-eyse* in a spoken context relied on a few verbs more than in a written one: the top 10 verbs occupied about 36% of total usage in a spoken register compared to about 19% in a written register. A total of 414 types of verbs were used, among which 383 occurred less than 10 times. *ha-* ‘to do’ was the most frequent verb taking up 9.99% of total usage, followed by *nao-* ‘to come out’ (6.99%) and *o-* ‘to come’ (4.55%). Some verbs that describe specific activities such as *yaykiha-* ‘to talk’, *mek-* ‘to eat’, *ca-* ‘to sleep’, *ssu-* ‘to write/use’, and *mantul-* ‘to make’ were ranked highly. The proportion of one verb *iss-* ‘to be/exist’, which was the most frequently occurring verb with *-ey*, was also high. The diverse meanings of *iss-* that include ‘to take place’ seems to have caused its usage with *-eyse*.

Table 6. 5. Frequencies and proportions of verbs used with *-eyse* in the Sejong spoken corpus

Rank	Verb	Token (#)	Proportion (%)	Accumulated proportion (%)
1	<i>ha-</i> ‘to do’	180	9.99	9.99
2	<i>nao-</i> ‘to come out’	126	6.99	16.98
3	<i>o-</i> ‘to come’	82	4.55	21.53
4	<i>iss-</i> ‘to be/exist/take place’	64	3.55	25.08
5	<i>yaykiha-</i> ‘to talk’	41	2.28	27.36
6	<i>mek-</i> ‘to eat’	40	2.22	29.58
7	<i>ssu-</i> ‘to write/use’	34	1.89	31.47
8	<i>ilha-</i> ‘to work’	33	1.83	33.30
9	<i>ka-</i> ‘to go’	32	1.78	35.07
10	<i>sal-</i> ‘to live’	31	1.72	36.79
	Others	1,139	63.21	100.00
Total		1,802	100.00	

Note. Others comprised 404 verb types.

6.1.3. *-(u)lo*

Overall, the number of tokens for locational *-(u)lo* was smaller than other postpositions. A total of 2,362 tokens were found in the Sejong written corpus with a total of 482 verb types. Among them, 434 occurred less than 10 times. As was the case for *-ey* and *-eyse*, a small group of verbs dominated the majority of total usage: the top 10 verbs occupied about 35%, as Table 6.6 shows. *ka-* ‘to go’ (10.63%) and *tuleka-* ‘to enter/to go in’ (4.45%) were the two most frequent verbs.

Table 6. 6. Frequencies and proportions of verbs used with *-(u)lo* in the Sejong written corpus

Rank	Verb	Token (#)	Proportion (%)	Accumulated proportion (%)
1	<i>ka-</i> ‘to go’	251	10.63	10.63
2	<i>tuleka-</i> ‘to enter/go in’	105	4.45	15.07
3	<i>naka-</i> ‘to go out’	83	3.51	18.59
4	<i>nao-</i> ‘to come out’	76	3.22	21.80
5	<i>tolaka-</i> ‘to detour’	73	3.09	24.89
6	<i>tolao-</i> ‘to return’	64	2.71	27.60
7	<i>olmki-</i> ‘to move (sth)’	52	2.20	29.81
8	<i>o-</i> ‘to come’	47	1.99	31.80
9	<i>tulese-</i> ‘to go in/walk in’	37	1.57	33.36
10	<i>ollaka-</i> ‘to climb/go up’	36	1.52	34.89
	Others	1,538	65.11	100.00
Total		2,362	100.00	

Note. Others comprised 472 verb types.

A total of 87 types of verbs were used for the 588 tokens of locational *-(u)lo* in the Sejong spoken corpus. Among these 87 types, only 11 types of verbs occurred more than 10 times. The verb use was similar to that of the written one in that some verbs were used much more frequently used than others, except that the respective proportion of each frequently used verb was larger. As Table 6.7 shows, the most frequent verb *ka-* ‘to go’ took up more than

33.33%, while the second frequent verb *tuleka-* ‘to enter/go in’ took up 7.48%. The top 10 frequent verbs occupied more than 67% of the total usage.

Table 6. 7. Frequencies and proportions of verbs used with *-(u)lo* in the Sejong spoken corpus

Rank	Verb	Token (#)	Proportion (%)	Accumulated proportion (%)
1	<i>ka-</i> ‘to go’	196	33.33	33.33
2	<i>tuleka-</i> ‘to enter/go in’	44	7.48	40.82
3	<i>o-</i> ‘to come’	31	5.27	46.09
	<i>cwu-</i> ‘to give’	31	5.27	51.36
5	<i>naka-</i> ‘to go out’	21	3.57	54.93
6	<i>ppaci-</i> ‘to fall into’	17	2.89	57.82
7	<i>nemeka-</i> ‘to cross’	16	2.72	60.54
8	<i>ponay-</i> ‘to send’	15	2.55	63.10
9	<i>ollaka-</i> ‘to climb/go up’	13	2.21	65.31
10	<i>tuleo-</i> ‘to come in’	12	2.04	67.35
	Others	192	32.65	100.00
Total		588	100.00	

Note. Others comprised 77 verb types.

Verbs co-occurring with *-(u)lo* exhibited two characteristics. The first one was frequent use of serial verbs¹¹ in both written and spoken corpora. Serial verbs are combinations of two verbs. For instance, *tolaka-* ‘to detour’ and *tolao-* ‘to return’ are combinations of *tol-* ‘to turn’ and either *ka-* ‘to go’ or *o-* ‘to come.’ In the written corpus, *naka-* ‘to go out’ (3.51%), *nao-* ‘to

¹¹ Among serial verbs, decisions on which verb would be considered as one distinct type of verb were made based on the online *Pyocunkwukedaysacen* [Standard Korean Dictionary] provided by the NIKL in the current study. When a verb was officially registered as a word in the dictionary, it was considered as one distinct verb type. For instance, *caphye-ka-* ‘to be caught and go’ was counted as one type of verb due to its entry in the dictionary (that is, it was counted as a different verb from either *caphi-* ‘to be caught’ or *ka-* ‘to go’). On the other hand, *caphye-o-* ‘to be caught and come’ was not registered in the dictionary and therefore became classified as a serial verb resulting from the combination of the two separate verbs *caphi-* ‘to be caught’ and *o-* ‘to come’. As Sohn(1999) pointed out, the first verb usually describes the manner and *-(u)lo caphi-** is ungrammatical as a locational. Thus, in the current study, the verb *caphye-o-* ‘to be caught and come’ was counted as an instance of *o-* ‘to come’.

come out' (3.22%), *tolaka-* 'to detour' (3.09%), and *tolao-* 'to return' (2.71%) were ranked from 3rd to 6th in terms of frequency. In the spoken corpus, *naka-* 'to go out' (3.57%), *nemeka-* 'to cross' (2.72%), *ollaka-* 'to climb/go up' (2.21%), and *tuleo-* 'to come in' (2.04%) were commonly used. In general, the first verb in a serial verb describes the manner in which the second verb happens (Sohn, 1999). This attribute fits well with the meaning of *-(u)lo* since a number of studies have pointed out that the prototypical meaning implied by *-(u)lo* is 'path' (Park, 1999; Jeong, 2010; Kang, 2012). Serial verbs can show what happened in the process of passing by the path, which seems to be the reason for the frequent use of serial verbs.

Another interesting aspect of *-(u)lo* involved its use of many verbs that could also occur with *-ey*. To illustrate, most of the verbs listed in Table 6.6 and 6.7 can be used with *-ey* as well. This demonstrates a large amount of functional overlap between *-(u)lo* and *-ey* as locational postpositions. However, not all of the most frequent verbs for *-(u)lo* were frequently used with *-ey*, which also implies that language users preferred *-(u)lo* over *-ey* for certain verbs. There were some less frequent verbs with *-(u)lo* that were not interchangeable with *-ey*. They include *hyangha-* 'to head for' and *ttena-* 'to leave for'. They hardly combine with *-ey* as those verbs imply 'direction' towards a place, which is one of the characteristics of *-(u)lo* as well.

Overall, the use of verb for the three postpositions indicated that some verbs were dominantly used with each postposition, as indicated by noticeably large proportions amongst total usage. Regarding the number of verb types in comparison with total token numbers, *-eyse* employed significantly more verbs than *-ey* or *-(u)lo* among the three postpositions. In terms of the ratio between verb type to token number for *-eyse*, 4.27 and 4.35 were found for the written and spoken corpora, respectively. That is, a new type of verb occurred in every 4 ~ 5 tokens on average. For *-ey*, they were 11.2 (written) and 10.05 (spoken), while *-(u)lo* showed 4.9 and 6.76

for the written and spoken corpora, respectively. The reason for this seems to stem from *-eyse*'s flexibility in terms of combining with verbs. As the postposition *-eyse* can denote any location where an activity occurs, it is open to a diverse group of verbs. For instance, action or emotion-related verbs such as *pis-* 'to comb', *wul-* 'to cry', and *pinanha-* 'to criticize' can occur with *-eyse* unlike *-ey* or *-(u)lo*. In addition, numerous types of verbs derived from light verbs (e.g., *hwanyeng-ha-* 'to welcome') can combine with *-eyse*. Accordingly, the proportions of the verb that most frequently occurred with *-eyse*, which was *ha-* 'to do', were relatively lower (written: 4.43%, spoken: 9.99%) compared to the proportions of verbs that most frequently occurred with *-ey* and *-(u)lo*.

6.2. Relations between postpositions and verbs

6.2.1. Frequency distribution

The analysis on the types of verbs and their proportions of co-occurrence with *-ey*, *-eyse*, and *u(lo)* demonstrated that there was a strong tendency for employing a limited number of verbs regardless of corpus register and type of postposition. Moreover, the most frequent verbs occupied quite a large portion of the total usage. For example, the most frequent verb with *-ey* was *iss-* 'to be/exist', which occupied 31.68% of the spoken tokens and 9.46% of the written tokens. Such was also the case for the other two postpositions and the verbs with which they frequently co-occurred. This trend was even stronger in spoken contexts. The dominance of these verbs resulted in the skewed distribution of frequency graphs (Figures 6.1, 6.2, and 6.3). The x-axes of the three Figures mark token frequency while the y-axes indicate the top thirty most frequent verbs (from top to bottom) used with *-ey*, *-eyse*, and *-(u)lo* in written and spoken registers, respectively.

This dependency on one (or a few numbers of) item for a certain construction has been observed in a number of studies that examined frequency of language use (e.g., Ellis & Ferreira-Junior, 2009a, 2009b). For example, Ellis and Ferreira-Junior (2009b) manifested that a verb ‘to put’ occupied a high proportion of usage in English VOL constructions [Subj V Obj Obl_{path/loc}] by both L1 and L2 users of English. Zipf’s law (1935) explains this skewedness: the number of tokens exponentially decreases as its frequency rank increases. Highly frequent words occupy most of the linguistic tokens in natural language, showing an inverse relationship between frequency and rank.

A frequency analysis of verbs in the L1 Korean corpora hinted that Zipf’s law can be applied to the verb use with locational postpositions as well. This is due to the fact that one dominant token takes up most of the usage which causes an exponential decrease in token numbers while frequency rank increases, as seen in Figure 6.1, 6.2, and 6.3. To examine whether frequency distribution of verbs follows a Zipfian distribution, log values of token numbers (frequency) and frequency rank of verbs were calculated and plotted. To illustrate the log values, the third most frequent verb with *-ey* in the written Sejong corpus was *tuleka-* ‘to enter/go in’ with token number 291. Then, *log_rank* of the verb was 0.477 ($=\log_{10}(3)$) while *log_frequency* was 2.464 ($=\log_{10}(291)$). The inserted images in Figure 6.1, 6.2 and 6.3 are plots displaying the relationship between *log_rank* and *log_frequency* for each postposition and verbs in each register, with R^2 values in each image. The linear relationship between the *log_rank* (x-axis) of a verb and the *log_frequency* (y-axis) of token numbers in each plot show that frequency distributions of verbs with each postposition follow Zipf’s law.

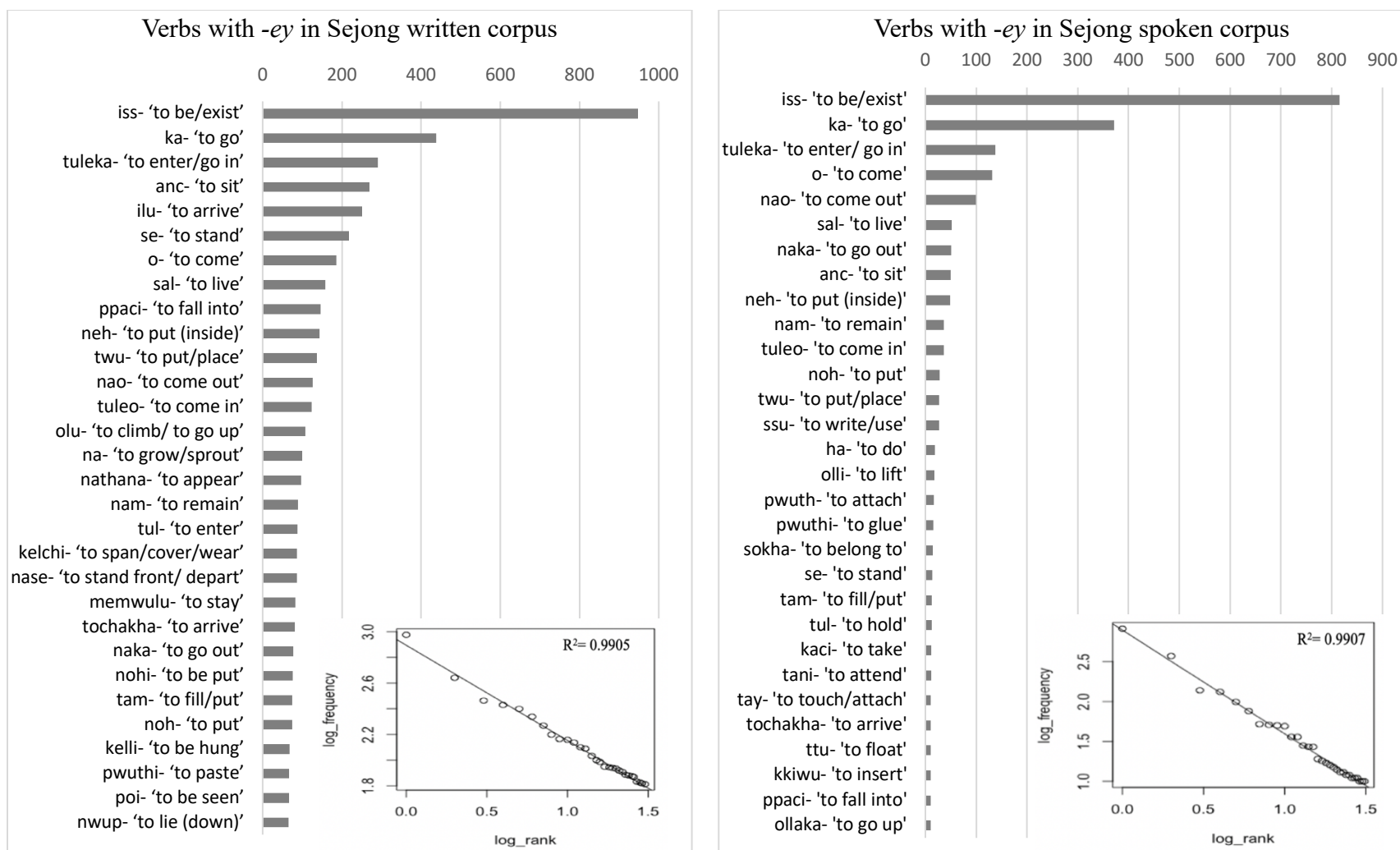


Figure 6. 1. Frequency distributions of the top 30 verbs occurring with *-ey* in the Sejong corpora (x-axis: Frequency, y-axis: Lemma form of verbs).

Note. Inserted images show the linear relationship between *log_rank* and *log_frequency*.

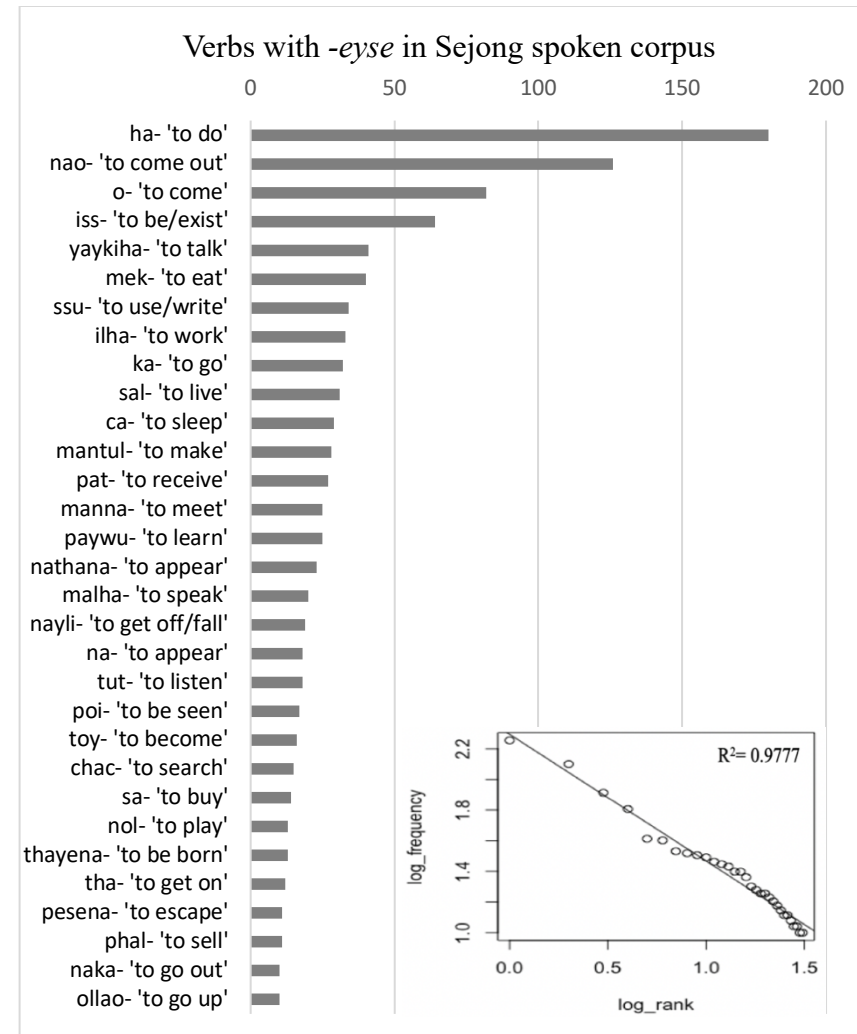
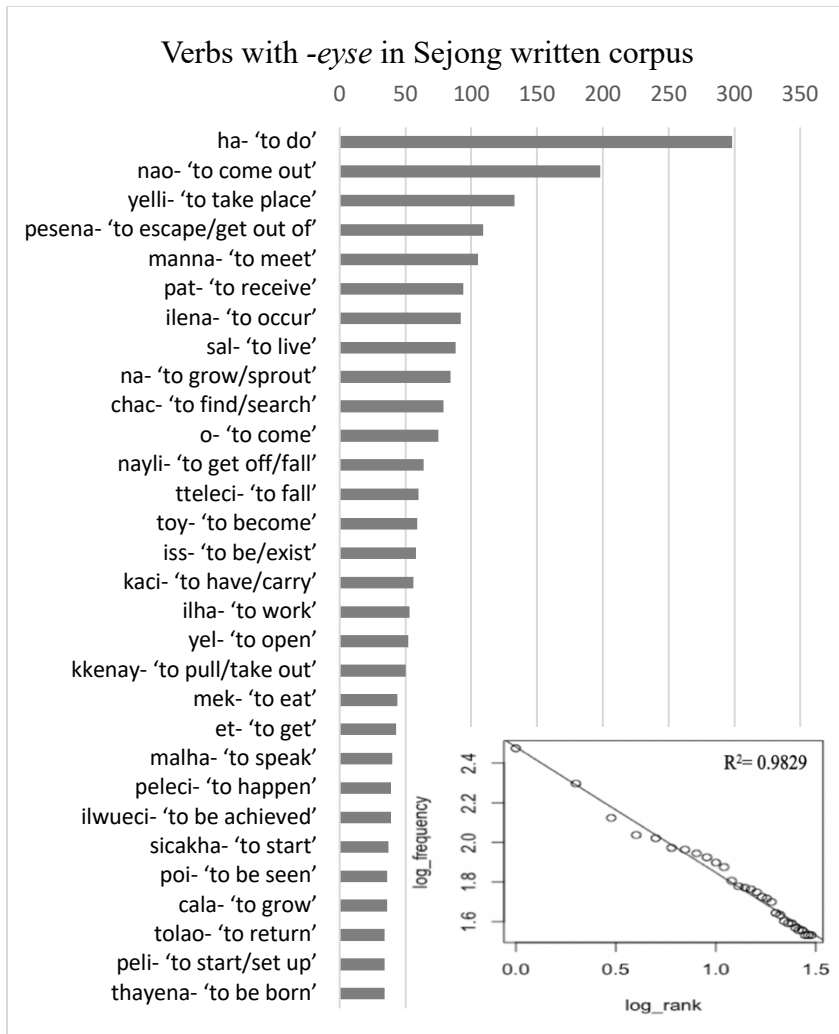


Figure 6. 2. Frequency distributions of the top 30 verbs occurring with *-eyse* in the Sejong corpora (x-axis: Frequency, y-axis: Lemma form of verbs).

Note. Inserted images show the linear relationship between *log_rank* and *log_frequency*.

Note. Number of verbs in the spoken corpus is 31, due to identical token numbers for several types of verbs.

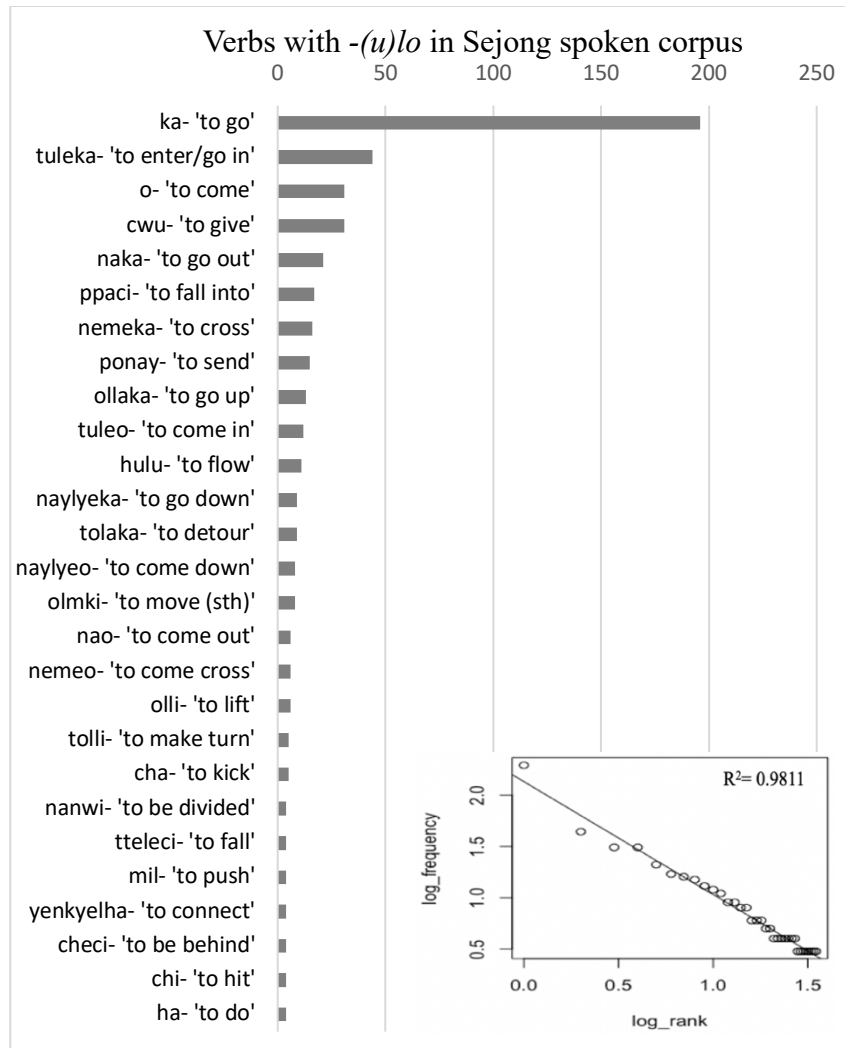
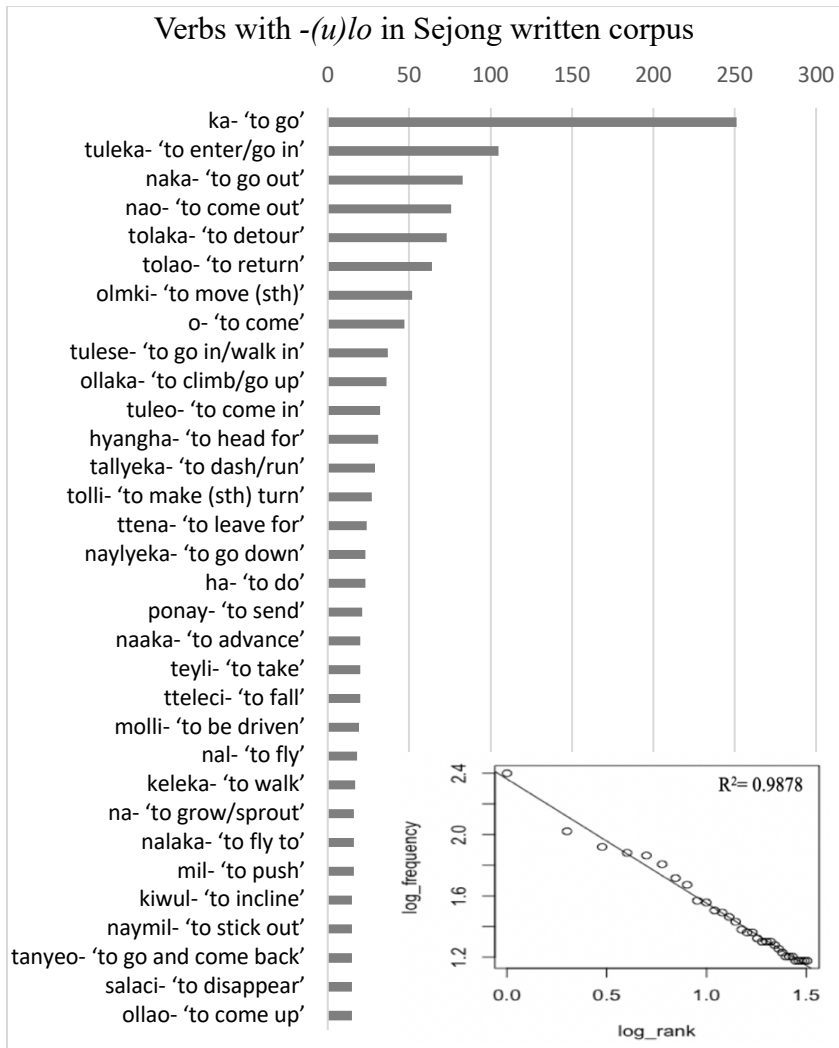


Figure 6. 3. Frequency distribution of the top 30 verbs occurring with -(u)lo in the Sejong corpora (x-axis: Frequency, y-axis: Lemma form of verbs).

Note. Inserted images show the linear relationship between *log_rank* and *log_frequency*.

Note. Number of verbs are 32 (written corpus) and 27 (spoken corpus), due to identical token numbers for several types of verbs.

Zipfian distributions of verbs brings the concept of pathbreaking verbs to the fore (Ninio, 1999). A pathbreaking verb represents the most prototypical meaning of the syntactic structure within which it is used, which will break the path for other verbs to be easily accessible. That is, language users learn combinatorial rules from pathbreaking verbs and generalize the rules to more abstract relations. As pathbreaking verbs are most distinctively associated with the establishment of language patterns, they occur most frequently and tend to be learned first (Ellis & Ferreira-Junior, 2009b). Thus, it is likely that the most frequent verbs occurring with each postposition (*iss-* ‘to be/exist’ for *-ey*, *ha-* ‘to do’ for *-eyse*, and *ka-* ‘to go’ for *-(u)lo*) play a significant role, presumably as pathbreaking verbs, in the acquisition of the locational construction in Korean as an L1, which can be further examined in future studies.

Moreover, the most frequent token is an exemplar that represents the prototypical meaning of the category to which it belongs and tends to be generic and widely applicable, while tokens that carry specific meanings occur less frequently (Goldberg et al., 2004; Ellis & Ferreira-Junior, 2009). The exposure to this low-variance and high-frequency exemplar of the category promotes a ‘fix’ on the characteristics of the category (Elio & Anderson, 1984). The most frequent verb in each postposition subsequently demonstrates prototypical use of each postposition. As previous studies about Korean postpositions have indicated, existential place seems to be the prototypical use of *-ey* as represented by the most frequent verb *iss-* ‘to be/exist’, while place of activity encoded by *ha-* ‘to do’ characterizes a main function of *-eyse*. In the case of *-(u)lo*, the prototypical meaning of ‘path’ seems to be implied in the verb *ka-* ‘to go’.

6.2.2. Keyness analysis

Frequency distributions of both written and spoken corpora show that several verb types take up a considerable proportion of total usage. However, different percentages for different verbs are indicative of discrepancies in verb usage between the two corpora. Thus, the difference between the two corpora in terms of verb preference was examined through keyness analysis. While the statistical significance threshold for keyness analysis varies across studies, it tends to set an extremely low p -value (e.g., $p < .000000001$), especially when the corpus size is large (Gabrielatos, 2018). Otherwise, the results will return a large number of *candidate key items* (CKIs). Considering the size of corpora for the present study and the number of keywords, L1 written and spoken corpora were compared with confidence levels of 99.99% ($p < .0001$, critical value of LL : 15.13). In the present study, an L1 written corpus was used as a reference corpus and an L1 spoken corpus as a study corpus. In the keyness analysis, a higher LL value means a stronger significant difference between two corpora. %DIFF (effect size) shows the effect size using the written corpus as a reference corpus. Also, large values of %DIFF (e.g., more than 10,000) indicate zero occurrence in the reference (L1 written) corpus, whereas -100 indicates zero tokens in the study (L1 spoken) corpus.

Table 6.8 shows CKIs indicative of significant differences between the Sejong written and spoken corpora in relation to the use of verbs with *-ey*. Sixteen key verbs were identified in the comparison. Two verbs *iss-* ‘to be/exist’ and *ka-* ‘to go’ showed the strongest differences. %DIFF indicated that *iss-* ‘to be/exist’ and *ka-* ‘to go’ were used more than three times as often in the spoken corpus than in the written one. The verb *o-* ‘to come’ was used 2.5 times as often in the spoken corpus. In the case of *nase-* ‘to stand front/ depart’, *iluta-* ‘to arrive’, and *kelchi-* ‘to put on/hang’, they were hardly used in the spoken corpus compared to the written one.

Table 6. 8. CKI verbs used with *-ey* in the Sejong corpora

Verb type	<i>LL</i>	%DIFF
<i>iss-</i> ‘to be/exist’	596.79	234.95
<i>ka-</i> ‘to go’	269.78	230.55
<i>ilu-</i> ‘to arrive’	87.63	-93.80
<i>o-</i> ‘to come’	72.60	176.20
<i>nao-</i> ‘to come out’	63.33	205.80
<i>se-</i> ‘to stand’	39.25	-75.01
<i>tuleka-</i> ‘to go in’	33.18	84.57.
<i>nase-</i> ‘to stand front/ depart’	31.74	-95.47
<i>olu-</i> ‘to climb/ go up’	31.58	-89.19
<i>kelchi-</i> ‘to put on/hang’	27.16	-91.05
<i>naka-</i> ‘to go out’	25.05	157.78
<i>ppaci-</i> ‘to fall’	24.44	-73.34
<i>nathana-</i> ‘to appear’	20.45	-79.94
<i>nohi-</i> ‘to be placed’	18.89	-84.64
<i>memwulu-</i> ‘to stay’	18.34	-81.24
<i>ssu-</i> ‘to write/use’	17.52	209.07

Note. The *p*-value of *LL* was < .0001.

Keyness analysis was also conducted to compare verb usage for *-eyse* in both written and spoken corpora. As Table 6.9 shows, a total of 14 CKIs were selected. Among them, *yaykiha-* ‘to talk’ was used more than eight times while two verbs, *o-* ‘to come’ and *ha-* ‘to do’, were used more than two to four times in the spoken corpus than in the written corpus. For *yelli-* ‘to take place’, it did not occur in the spoken corpus but occurred in considerably larger numbers in the written one. Some other verbs that showed different distributions between the written and spoken corpora included *iss-* ‘to be/ exist’, *nao-* ‘to come out’, and *ssu-* ‘to write/ use’.

Table 6.10 shows seven CKI verbs with statistically significant *LL* values for the comparison of verbs used with *-(u)lo* in the two corpora. Keyness results indicate that five verbs

were distinctively used in the spoken corpus: *ka-* ‘to go’, *cwu-* ‘to give’, *ppaci-* ‘to fall into’, *nemeka-* ‘to cross’, and *o-* ‘to come’. Two other verbs, *nemeco-* ‘to cross and come’ and *cha-* ‘to kick’, were not used at all in the written corpus.

Table 6. 9. CKI verbs used with *-eyse* in the Sejong corpora

Verb type	<i>LL</i>	%DIFF
<i>yaykiha-</i> ‘to talk’	112.94	7549.39
<i>o-</i> ‘to come’	74.89	307.97
<i>ha-</i> ‘to do’	72.52	125.39
<i>yelli-</i> ‘to take place’	63.72	-100
<i>iss-</i> ‘to be/ exist’	58.70	311.74
<i>nao-</i> ‘to come out’	55.10	137.45
<i>ssu-</i> ‘to write/ use’	40.07	451.60
<i>ka-</i> ‘to go’	35.92	419.15
<i>tut-</i> ‘to listen’	31.93	1019.42
<i>mek-</i> ‘to eat’	29.33	239.22
<i>ca-</i> ‘to sleep’	28.77	350.88
<i>paywu-</i> ‘to learn’	25.55	366.43
<i>mantul-</i> ‘to make’	19.49	226.50
<i>yel-</i> ‘to open’	17.95	-92.82

Note. The *p*-value of *LL* was < .0001.

Table 6. 10. CKI verbs used with *-(u)lo* in the Sejong corpora

Verb type	<i>LL</i>	%DIFF
<i>ka-</i> ‘to go’	161.28	213.68
<i>cwu-</i> ‘to give’	101.33	301.70
<i>ppaci-</i> ‘to fall into’	29.30	875.56
<i>nemeka-</i> ‘to cross’	24.89	703.40
<i>nemeco-</i> ‘to cross and come’	19.40	> 10,000
<i>o-</i> ‘to come’	16.59	164.95
<i>cha-</i> ‘to kick’	16.16	> 10,000

Note. The *p*-value of *LL* was < .0001.

6.2.3. Association strength

In addition to the investigation of frequency distributions for verbs and CKI verbs in comparison with other corpora (L1 written-L1 spoken) regarding the three locative postpositions, association strength between postpositions and verbs was also examined. While frequency and keyness of verb types show the properties of the corpus in relation to each postposition, association strength will show how verbs were utilized across different forms of postpositions. ΔP , which indicates the degree of association power between two elements, was used. As a number of verbs were used with more than two types of postpositions, ranks of ΔP values do not always match with the frequency ranks. For instance, *ka-* ‘to go’ was the second most frequent verb for *-ey* in the written corpus. Due to its high number of occurrences with *-(u)lo*, however, it ranks sixth for ΔP with *-ey*. A higher ΔP value indicates the stronger the association power.

First, $\Delta P_{(\text{verb} | \text{postposition})}$ was calculated using “postposition” as the cue and “verb” as the outcome. By doing so, this study was able to find out what types of verbs co-occurred with stronger power over other verbs with the target postposition. The top five verbs with the highest association strength values for each postposition in each corpus are listed in Table 6.11. Three verbs, *iss-* ‘to be/exist’, *ha-* ‘to do’, and *ka-* ‘to go’, most frequently occurred with *-ey*, *-eyse*, and *-(u)lo* respectively, showing the highest $\Delta P_{(\text{verb} | \text{postposition})}$ values in both written and spoken registers for each postposition. The spoken corpus showed higher $\Delta P_{(\text{verb} | \text{postposition})}$ scores due to the high token numbers of frequently used verbs for each postposition.

For the postposition of *-ey*, *ka-* ‘to go’ and *tuleka-* ‘to enter/go in’ were the second and third most frequent verbs in the written corpus, but with a low degree of association strength. In a spoken register, the association power of *o-* ‘to come’ was low compared to its frequency rank, whereas some verbs such as *anc-* ‘to sit’ and *neh-* ‘to put (inside)’ had higher association power

ranks compared to their frequency ranks. Among the top ten frequent verbs occurring with *-eyse*, the value of *iss-* ‘to be/exist’ was low due to its high frequency with *-ey*. Two verbs, *nao-* ‘to come out’ in the written corpus and *o-* ‘to come’ in the spoken corpus, were frequently used while still having low association strengths levels. This applies to *-(u)lo* as well: *nao-* ‘to come out’ and *o-* ‘to come’ were frequently used with *-(u)lo*, but association strengths were low due to the fact that they were productively combined with both *-eyse* and *-(u)lo*.

Table 6. 11. Top five $\Delta P_{(\text{verb} | \text{postposition})}$ values for verbs used with each postposition in the Sejong corpora

	Written		Spoken	
	Verb type	ΔP	Verb type	ΔP
<i>-ey</i>				
1	<i>iss-</i> ‘to be/exist’	0.087	<i>iss-</i> ‘to be/exist’	0.290
2	<i>anc-</i> ‘to sit’	0.026	<i>ka-</i> ‘to go’	0.049
3	<i>ilu-</i> ‘to reach’	0.021	<i>tuleka-</i> ‘to enter/go in’	0.032
4	<i>se-</i> ‘to stand’	0.021	<i>anc-</i> ‘to sit’	0.019
5	<i>tuleka-</i> ‘to enter/go in’	0.015	<i>neh-</i> ‘to put (inside)’	0.017
<i>-eyse</i>				
1	<i>ha-</i> ‘to do’	0.039	<i>ha-</i> ‘to do’	0.093
2	<i>yelli-</i> ‘to take place’	0.019	<i>nao-</i> ‘to come out’	0.037
3	<i>pesena-</i> ‘to escape’	0.016	<i>yaykiha-</i> ‘to talk’	0.023
4	<i>manna-</i> ‘to meet’	0.015	<i>mek-</i> ‘to eat’	0.022
5	<i>nao-</i> ‘to come out’	0.013	<i>ilha-</i> ‘to work’	0.018
<i>-(u)lo</i>				
1	<i>ka-</i> ‘to go’	0.079	<i>ka-</i> ‘to go’	0.241
2	<i>tolaka-</i> ‘to detour’	0.030	<i>cwu-</i> ‘to give’	0.051
3	<i>naka-</i> ‘to go out’	0.030	<i>tuleka-</i> ‘to enter/go in’	0.041
4	<i>tuleka-</i> ‘to enter/go in’	0.026	<i>nemeka-</i> ‘to cross/fall’	0.026
5	<i>tolao-</i> ‘to return’	0.023	<i>ppaci-</i> ‘to fall into’	0.025

Association strength was also calculated using “verb” as the cue and “postposition” as the outcome ($\Delta P_{(\text{postposition} | \text{verb})}$). ΔP scores calculated in this manner can show with which

postposition a certain verb enjoys greater association strength. This approach yields higher values of association strength when a verb is used with one type of postposition exclusively. Numerous types of verbs were used with only one locative postposition in the L1 corpora, often times with a small number of tokens. Those verbs demonstrated higher association strength ($\Delta P_{(\text{postposition} | \text{verb})}$) with each postposition. As the purpose of this dissertation is to examine the characteristics of multiple form-function relations rather than what types of verbs were exclusively used with a certain postposition with a limited number of tokens, it will not be useful to list all the verbs that show higher scores for association strength ($\Delta P_{(\text{postposition} | \text{verb})}$). Thus, not all of the results were reported in the current dissertation.

However, examining association power with “verb” as the cue and “postposition” as the outcome ($\Delta P_{(\text{postposition} | \text{verb})}$) is useful when investigating the characteristics of verbs that can be used with multiple postpositions (e.g., *-ey sal-/ -eyse sal-* ‘to live (somewhere)’). Thus, after calculating $\Delta P_{(\text{postposition} | \text{verb})}$ for all verbs in the corpora, a limited number of verb types were examined with particular interest. Two factors were considered in selecting which verbs to examine: the number of postpositions with which a verb combines and token number. First, verbs occurring with at least two locative postpositions were considered. For instance, the $\Delta P_{(\text{postposition} | \text{verb})}$ value of *kyelhonha-* ‘to marry’ was quite high (0.642) for *-eyse* in the Sejong written corpus since it did not occur with *-ey* and *-(u)lo*. While the $\Delta P_{(\text{postposition} | \text{verb})}$ value for *kyelhonha-* ‘to marry’ with *-eyse* was high, it was not considered for examination since it combined with only one type of postposition. Second, verbs that occurred at least 10 times with any of the postpositions were considered. For example, *kwiswunha-* ‘to defect/submit (to enemy country)’ occurred one time with *-ey* and *-(u)lo* respectively in the Sejong written corpus, producing $\Delta P_{(\text{postposition} | \text{verb})}$ values of -0.024 (*-ey*) and 0.376 (*-(u)lo*). Even though the scores

indicated that the verb had stronger association power with *-(u)lo* than *-ey*, this was misleading due to the fact that the total number of tokens for *-(u)lo* was much smaller compared to *-ey*. Thus, rather than concluding that the verb *kwiswunha-* preferred *-(u)lo*, the present study did not include it in the analysis since small token numbers are not appropriate for broad generalizations.

Under these criteria, a number of verbs were selected that had high $\Delta P_{(\text{postposition} | \text{verb})}$ values with one locative postposition over the other two locative postpositions. In Table 6.12, verbs that demonstrated the highest association strengths with *-ey*, *-eyse*, and *-(u)lo* respectively over the other two postpositions in both registers are listed. For example, *neh-* ‘to put (inside)’, *noh-* ‘to put’, and *twu-* ‘to put’ had stronger $\Delta P_{(\text{postposition} | \text{verb})}$ scores with *-ey* compared to *-eyse* and *-(u)lo*. Three other verbs (*nayli-* ‘to get off/fall’, *pat-* ‘to receive’, and *nao-* ‘to come out’) preferred *-eyse* over *-ey* and *-(u)lo*, while *naka-* ‘to go out’, *ponay-* ‘to send’, and *ollaka-* ‘to climb/go up’ demonstrated a stronger association with *-(u)lo* compared to *-ey* and *-eyse*.

Table 6. 12. Verbs with high $\Delta P_{(\text{postposition} | \text{verb})}$ values for each postposition over the other two postpositions in the Sejong corpora

<i>-ey</i>	<i>neh-</i> ‘to put (inside)’, <i>noh-</i> ‘to put’, <i>twu-</i> ‘to put’, <i>tuleka-</i> ‘to enter/go in’, <i>tuleo-</i> ‘to come in’, <i>olli-</i> ‘to lift’, <i>o-</i> ‘to come’
<i>-eyse</i>	<i>nayli-</i> ‘to get off/fall’, <i>pat-</i> ‘to receive’, <i>nao-</i> ‘to come out’
<i>-(u)lo</i>	<i>naka-</i> ‘to go out’, <i>ponay-</i> ‘to send’, <i>ollaka-</i> ‘to climb/go up’

The association strengths ($\Delta P_{(\text{postposition} | \text{verb})}$) of three verbs in particular (*iss-* ‘to be/exist’, *ha-* ‘to do’, and *ka-* ‘to go’) were examined due to the fact that they occurred quite frequently with all three postpositions (see Table 6.13). Regardless of register, these verbs displayed the highest association strength with the postpositions that they combined the most frequently with.

In other words, *iss-* ‘to be/exist’ had the strongest association strength with *-ey* over the other two postpositions, *ha-* ‘to do’ with *-eyse*, and *ka-* ‘to go’ with *-(u)lo*.

Table 6. 13. $\Delta P_{(\text{postposition} | \text{verb})}$ values for *iss-*, *ha-* and *ka-* with all three postpositions in the Sejong corpora

	Written			Spoken		
	<i>-ey</i>	<i>-eyse</i>	<i>-(u)lo</i>	<i>-ey</i>	<i>-eyse</i>	<i>-(u)lo</i>
<i>iss-</i> ‘to be/exist’	0.432	-0.311	-0.120	0.497	-0.353	–
<i>ha-</i> ‘to do’	-0.429	0.490	-0.061	-0.443	0.546	-0.103
<i>ka-</i> ‘to go’	0.094	-0.332	0.238	0.116	-0.352	0.237

While most of the selected verbs showed similar tendencies in both the written and spoken corpora, a few verbs showed different behaviors across registers. Two such verbs were *nathana-* ‘to appear’ and *sal-* ‘to live’. As the $\Delta P_{(\text{postposition} | \text{verb})}$ scores in Table 6.14 show, *nathana-* ‘to appear’ had a higher association strength with *-ey* in a written context while *-eyse* had a higher association strength in a spoken context. The verb *sal-* ‘to live’, which had often been the subject of discussion in previous studies, also showed different association powers in each register. In a written context, *sal-* had a higher association strength with *-eyse* over *-ey* while the ΔP score of *-ey* was higher than *-eyse* in a spoken context. Future investigation on the use of *sal-* ‘to live’ in various registers/contexts will result in deeper insights on its unique characteristics.

Finally, for closer examination into the relationship between each postposition and its dominant verb, the current study compared $\Delta P_{(\text{verb} | \text{postposition})}$ and $\Delta P_{(\text{postposition} | \text{verb})}$ values. This comparison of two types of association strength values ($\Delta P_{(\text{outcome} | \text{cue})}$ and $\Delta P_{(\text{cue} | \text{outcome})}$) allows for a clearer picture of which variable would function as a better predictor over the other. When

the difference between these values ($(\Delta P_{(\text{outcome} | \text{cue})} - \Delta P_{(\text{cue} | \text{outcome})})$) is positive, “cue” is a better predictor than outcome while a negative value would indicate that “outcome” is a better predictor. When the value is zero, neither one is a predictor of the other. The results between the three postpositions and their dominant verbs are shown in Table 6.15. Except for the relation between *-(u)lo* and *ka-* ‘to go’ in the spoken corpus, all of the value differences were negative. That is, each verb proved to be a better predictor of each postposition. For instance, the occurrence of *iss-* ‘to be/exist’ predicted the occurrence of *-ey* better than *-ey*’s occurrence as a predictor of *iss-* ‘to be/exist’. Likewise, the use of *ha-* ‘to do’ predicted the use of *-eyse* better than the other way around. When it comes to *-(u)lo* and *ka-* ‘to go’ in the spoken register, neither one was a convincing predictor of the other.

Table 6. 14. $\Delta P_{(\text{postposition} | \text{verb})}$ values of *nathana-* and *sal-* in the Sejong corpora

	Written			Spoken		
	<i>-ey</i>	<i>-eyse</i>	<i>-(u)lo</i>	<i>-ey</i>	<i>-eyse</i>	<i>-(u)lo</i>
<i>nathana-</i> ‘to appear’	0.196	-0.108	-0.087	-0.342	0.461	–
<i>sal-</i> ‘to live’	0.117	0.249	-0.121	0.110	0.011	–

Table 6. 15. Differences in ΔP values between all three postpositions and their dominant verbs in the Sejong corpora

		X = <i>-ey</i> Y = <i>iss-</i> ‘to be/exist’	X = <i>-eyse</i> Y = <i>ha-</i> ‘to do’	X = <i>-(u)lo</i> Y = <i>ka-</i> ‘to go’
Written	$\Delta P_{(Y X)}$	0.087	0.039	0.079
	$\Delta P_{(X Y)}$	0.432	0.490	0.238
	$\Delta P_{(Y X)} - \Delta P_{(X Y)}$	-0.345	-0.451	-0.159
Spoken	$\Delta P_{(Y X)}$	0.290	0.093	0.241
	$\Delta P_{(X Y)}$	0.497	0.546	0.237
	$\Delta P_{(Y X)} - \Delta P_{(X Y)}$	-0.207	-0.453	0.004

6.3. Summary of findings

The analysis of locative postpositions in the L1 Sejong written and spoken corpora showed that they were used more in the written corpus than in the spoken corpus. Among the three postpositions of *-ey*, *eyse*, and *-(u)lo*, *-ey* was used with the highest level of frequency while *-(u)lo* was used with the lowest level of frequency. Also, *-eyse* utilized the largest number of verb types in relation to the total number of tokens. Verbs that co-occur with the three postpositions were investigated in more detail in terms of three aspects: verb type frequencies, keyness analysis, and association strength measurement.

The results helped to define the characteristics of each postposition as they occur with verbs in a locative sense. A frequency analysis indicated that three verbs composed a large portion of verb usage for each postposition used in the locational sense: *iss-* ‘to be/exist’ for *-ey*, *ha-* ‘to do’ for *-eyse*, and *ka-* ‘to go’ for *-(u)lo* for both written and spoken contexts. In addition, a few verbs dominantly occupied a large proportion of total usage, especially in the case of the spoken data. The frequency distributions for each postposition and its co-occurring verbs followed a Zipfian distribution in each register, with the most frequent items showing an extremely high number of tokens while the token numbers for subsequent items decreased exponentially. Zipfian distributions of verb frequency became indicative of prototypical meanings associated with each locational postposition. That is, such distributions supported findings from previous studies on major differences between these three locative postpositions: *-ey* for existential/stative location, *-eyse* for dynamic location, and *-(u)lo* for direction/path.

Key verbs in each register revealed unique variations in verbs used with each postposition across registers. Keyness analysis identified which verbs were more distinctively used between written and spoken corpora, with special consideration of corpus size and token

numbers. For *-ey*, *iss-* ‘to be/ exist’ and *ka-* ‘to go’ were more commonly used in the spoken corpus while *ilu-* ‘to arrive’ and *se-* ‘to stand’ were more common in the written corpus. For *-eyse*, verbs such as *o-* ‘to come’ and *ha-* ‘to do’ were used significantly more often in spoken data while *yelli-* ‘to take place’ was used more often with *-eyse* in written data. Regarding verbs used with *-(u)lo* in each corpus, all seven key verbs were found more often in the spoken corpus.

Analysis of association strength showed how strongly a certain item is associated with another item. In the present study, the three most frequent verbs used with each of the locative postpositions (*iss-* ‘to be/ exist’, *ha-* ‘to do’, and *ka-* ‘to go’) also had the strongest association strengths with *-ey*, *-eyse*, and *-(u)lo* respectively. Moreover, most verbs used with more than one postposition exhibited similar tendencies: they preferred a certain locative postposition over the other two postpositions in both written and spoken registers. However, *nathana-* ‘to appear’ preferred *-ey* in the written context and *-eyse* in the spoken context while *sal-* ‘to live’ was more commonly used with *-eyse* in the written corpus and with *-ey* in the spoken corpus.

Overall, frequency distributions for verbs co-occurring with the three locative postpositions *-ey*, *-eyse*, and *-(u)lo* followed similar trends in that they were used with a few dominant verbs. However, different verbs were employed more (or less) frequently with each postposition depending on the registers with differing association strengths between postposition and verbs.

VII. ANALYSIS OF L2 TEXTBOOKS: USE OF *-EY*, *-EYSE*, AND *-(U)LO* AS LOCATIONAL POSTPOSITIONS WITH A FOCUS ON VERB USE

Two types of L2 Korean textbooks were analyzed as a major input source for language learners of Korean: Sogang Korean (SK) and Integrated Korean (IK). Eight volumes for each type of textbook (two volumes for four proficiency levels) were examined in terms of frequency distributions of verbs, keyness analysis, and association strength by level and/or in the entire volumes in accordance with the L1 Korean corpora analysis.

The total number of tokens for all three postpositions (*-ey*, *-eyse*, and *-(u)lo*) was 1,108 in the SK textbook series. The number of tokens for each postposition is listed in Table 7.1 according to level. As was the case for the L1 corpora, *-(u)lo* was used the least (164 tokens), while the number of tokens for *-ey* (645 tokens) was nearly four times that of *-(u)lo* and twice that of *-eyse* (299 tokens). Also, as the size of the textbook at each level increased (see Table 5.4), the number of tokens for each postposition increased as proficiency level increased.

Table 7. 1. Token numbers of locative postpositions in the SK textbook series (#)

	Level1	Level2	Level3	Level4	Total
<i>-ey</i>	121	152	169	203	645
<i>-eyse</i>	43	54	73	129	299
<i>-(u)lo</i>	4	24	26	110	164
Total	168	230	268	442	1,108

The number of instances for the three locational postpositions in the IK textbook volumes was 1,040. Token numbers for each level with each postposition are presented in Table 7.2. The IK series also used *-ey* the most (625 tokens), similar to the L1 corpora and the SK series. The total token number of tokens for *-eyse* was 278, which was nearly half that of *-ey* and twice that

of *-(u)lo*. In general, the number of tokens for each postposition increased as level increased in the IK series, except for a decrease in token numbers for *-ey* and *-(u)lo* from level 2 to level 3. This fluctuation may have originated either from the small size of the volume in level 3 compared to level 2 (See Table 5.4) or the types of topics covered in the textbooks.

Table 7. 2. Token numbers of locative postpositions in the IK textbook series (#)

	Level1	Level2	Level3	Level4	Total
<i>-ey</i>	83	148	98	296	625
<i>-eyse</i>	35	59	76	108	278
<i>-(u)lo</i>	8	37	23	69	137
Total	126	244	197	473	1,040

The following sections deals with each analysis one by one. Section 7.1 investigates the frequency of verbs used with each postposition in each type of textbook at each level. In section 7.2, the use of verbs (encompassing all the levels) will be examined in relation to frequency distribution, key verbs, and association strengths. Association strengths between each postposition and their co-occurring verbs for each level will be also included in section 7.2 followed by a summary of findings in section 7.3.

7.1. Verb use with each postposition

7.1.1. *-ey*

The frequencies of each verb type occurring with locational *-ey* at each level in the SK textbook series is shown in Table 7.3. The verb that occurred the most frequently with *-ey* across all levels was *ka-* ‘to go’. The proportion for this verb was as high as 70.25% in level 1, which indicates that the introduction of locative *-ey* in the textbook is heavily dependent on this verb. The proportion subsequently decreased in levels 2 and 3, ultimately amounting to 15.27% in level 4.

Two other verbs that frequently occurred with *-ey* were *iss-* ‘to be/exist’ and *o-* ‘to come’ with slight changes in proportion figures at each level. As might be expected, the number of verb types employed at each level of the textbook series increased from eight types (level 1) to 61 types (level 4). At level 4, the ratio between the number of verb types and token numbers was 3.33 (i.e., a new type of verb occurred for every 3~4 tokens).

Verb types used with locational *-ey* at each level of the IK textbook series are listed in Table 7.4. The most frequent verb was also *ka-* ‘to go’ from level 1 to level 3, whose proportion changed from 45.78% (level 1) to 50% (level 2) and 24.49% (level 3), respectively. Two verbs (*iss-* ‘to be/exist’ and *o-* ‘to come’) were also commonly used at those levels. At the 4th level, the verb occurring most frequently with locational *-ey* was *iss-* ‘to be/exist’ (10.14%). The verb *ssu-* ‘to use/write’ was the second most frequent verb at the 4th level. As was the case for the SK textbook series, the number of verb types also increased. On average, a new type of verb occurred for every 11~12 tokens at level 1 and every 2~3 tokens at level 4.

In both the SK and IK textbooks, it seems that locative-destination was the main trait being featured through locational *-ey* at level 1 volumes, as locational function with postposition *-ey* was mostly introduced with *ka-* ‘to go’. In addition to *ka-* ‘to go’, both *iss-* ‘to be/exist’ and *o-* ‘to come’ were popular verbs in both types.

Table 7. 3. Frequency of verbs occurring with *-ey* in the SK textbook series by level

Rank	Level 1		Level 2		Level 3		Level 4	
	Verb	#	Verb	#	Verb	#	Verb	#
1	<i>ka-</i> 'to go'	85	<i>ka-</i> 'to go'	70	<i>ka-</i> 'to go'	59	<i>ka-</i> 'to go'	31
2	<i>iss-</i> 'to be/exist'	17	<i>o-</i> 'to come'	16	<i>iss-</i> 'to be/exist'	23	<i>iss-</i> 'to be/exist'	29
3	<i>o-</i> 'to come'	8	<i>iss-</i> 'to be/exist'	13	<i>o-</i> 'to come'	12	<i>o-</i> 'to come'	15
4			<i>nao-</i> 'to come out'	11	<i>tuleka-</i> 'to enter/go in'	8	<i>tuleka-</i> 'to enter/go in'	10
5					<i>ppattuli-</i> 'to drop'	7	<i>ppaci-</i> 'to fall into'	9
	Others	11	Others	42	Others	60	Others	109
Total		121		152		169		203

Note. Verb types included in Others by level were as follows: 5 in Level 1; 118 in Level 2; 28 in Level 3; 56 in Level 4.

Table 7. 4. Frequency of verbs occurring with *-ey* in the IK textbook series by level

Rank	Level 1		Level 2		Level 3		Level 4	
	Verb	#	Verb	#	Verb	#	Verb	#
1	<i>ka-</i> 'to go'	38	<i>ka-</i> 'to go'	74	<i>ka-</i> 'to go'	24	<i>iss-</i> 'to be/exist'	30
2	<i>iss-</i> 'to be/exist'	21	<i>iss-</i> 'to be/exist'	11	<i>iss-</i> 'to be/exist'	10	<i>ssu-</i> 'to write/use'	15
3	<i>o-</i> 'to come'	13	<i>o-</i> 'to come'	9	<i>nao-</i> 'to come out'	5	<i>ka-</i> 'to go'	13
4	<i>sal-</i> 'live'	7	<i>sal-</i> 'live'	7	<i>o-</i> 'to come'	5	<i>talli-</i> 'to be hung'	11
5					<i>tuleka-</i> 'to enter/go in'	4	<i>ha-</i> 'to do'	10
6					<i>ssu-</i> 'to write/use'	4		
	Others	4	Others	47	Others	46	Others	217
Total		83		148		98		296

Note. Verb types included in Others by level were as follows: 3 in Level 1; 25 in Level 2; 30 in Level 3; 109 in Level 4.

7.1.2. *-eyse*

Table 7.5 shows the frequency list for verb types occurring with *-eyse* in the SK textbook volumes by level. At level 1, the most frequent verb was *manna-* ‘to meet’ (25.58%) followed by *ha-* ‘to do’ (18.6%). At other levels, *ha-* ‘to do’ occupied the largest proportion, while the proportion of *manna-* ‘to meet’ decreased. In addition to *ha-* ‘to do,’ some verbs such as *sal-* ‘to live’ and *ilha-* ‘to work’ were commonly used. The number of verb types also increased as level increased from 15 (level 1) to 57 (level 4). Compared to *-ey*, the ratio between *-eyse* and the range of co-occurring verb types was relatively low with values of 2.87 (level 1) and 2.26 (level 4): a new type of verb was introduced for every 2~3 tokens on average.

The use of verbs with *-eyse* within the IK textbook volumes by level is summarized in Table 7.6. The most frequent verbs in the IK textbook series varied according to level: *sal-* ‘to live’ (level 1: 20%), *sa-* ‘to buy’ (level 2: 13.56%), *sicaktoy-* ‘to be started’ (level 3: 6.58%), and *yelli-* ‘to take place’ (level 4: 6.48%) were the most popular verbs at each level. In addition, rather than using a limited number of verbs across all levels, various types of verbs were employed at each level. The verb types utilized at each level of the IK textbook series increased from 13 (level 1) to 66 (level 4). Especially at level 4, the ratio between the number of verb types and token numbers was 1.64 (i.e., a new verb was used for every 1~2 tokens).

In both types of textbooks, some of the verbs were frequent at certain levels. For instance, two verbs, *hwaphwuliha-* ‘to vent one’s anger’ (fourth frequent verb in level 4 volumes of SK) and *sicaktoy-* ‘to be started’ (the most frequent one in level 3 of IK), did not appear at other levels of the same type of textbook. As *-eyse* marks a location where an activity takes place and can combine with various types of activity-related verbs, it appears that textbooks utilized various types of verbs depending on the topics being covered at each volume.

Table 7. 5. Frequency of verbs occurring with *-eyse* in the SK textbook series by level

Rank	Level 1		Level 2		Level 3		Level 4	
	Verb	#	Verb	#	Verb	#	Verb	#
1	<i>manna-</i> ‘to meet’	11	<i>ha-</i> ‘to do’	9	<i>ha-</i> ‘to do’	11	<i>ha-</i> ‘to do’	14
2	<i>ha-</i> ‘to do’	8	<i>manna-</i> ‘to meet’	4	<i>sal-</i> ‘to live’	11	<i>sal-</i> ‘to live’	10
3	<i>kongpwuha-</i> ‘to study’	4	<i>ilha-</i> ‘to work’	3	<i>manna-</i> ‘to meet’	5	<i>ilha-</i> ‘to work’	6
4	<i>sal-</i> ‘to live’	4	<i>o-</i> ‘to come’	3	<i>o-</i> ‘to come’	5	<i>hwaphwuliha-</i> ‘to vent one’s anger’	6
5					<i>ilha-</i> ‘to work’	5	<i>nayli-</i> ‘to get off/fall’	5
	Others	16	Others	35	Others	36	Others	88
Total		43		54		73		129

Note. Verb types included in Others by level were as follows: 11 in Level 1; 27 in Level 2; 27 in Level 3; 52 in Level 4.

Table 7. 6. Frequency of verbs occurring with *-eyse* in the IK textbook series by level

Rank	Level 1		Level 2		Level 3		Level 4	
	Verb	#	Verb	#	Verb	#	Verb	#
1	<i>sal-</i> ‘to live’	7	<i>sa-</i> ‘to buy’	8	<i>sicaktoy-</i> ‘to be started’	5	<i>yelli-</i> ‘to take place’	7
2	<i>manna-</i> ‘to meet’	5	<i>manna-</i> ‘to meet’	5	<i>manna-</i> ‘to meet’	3	<i>o-</i> ‘to come’	6
3	<i>mek-</i> ‘to study’	4	<i>ha-</i> ‘to do’	5	<i>sal-</i> ‘to live’	3	<i>kaychoytoy-</i> ‘to be held’	4
4	<i>sa-</i> ‘to buy’	4	<i>ka-</i> ‘to go’	4	<i>o-</i> ‘to come’	3	<i>na-</i> ‘to grow/sprout’	4
	<i>o-</i> ‘to come’	4	<i>nayli-</i> ‘to get off/fall’	4		3	<i>nao-</i> ‘to come out’	4
							<i>panay-</i> ‘to send’	4
							<i>ha-</i> ‘to do’	4
	Others	11	Others	33	Others	59	Others	75
Total		35		59		76		108

Note. Verb types included in Others by level were as follows: 8 in Level 1; 24 in Level 2; 50 in Level 3; 59 in Level 4.

7.1.3. *-(u)lo*

Two types of verbs occurred with *-(u)lo* at the level 1 volumes in the SK textbook series: *kalatha-* ‘to transfer’ (3 tokens) and *isaha-* ‘to move’ (1 token). The most frequent verb from level 2 to level 4 was *ka-* ‘to go’, taking up about 54% (level 2), 50% (level 3), and 18% (level 4), respectively. Some other verbs used with *-(u)lo* in the textbooks included *tolaka-* ‘to detour’, *o-* ‘to come’, and *nao-* ‘to come out’. Verbs frequently used with *-(u)lo* at each level of the SK textbook series are summarized in Table 7.7. Ten types and nine types of verbs were used at level 2 and level 3 each. At the 4th level, 29 types of verbs were used. On average, the ratio between the number of verb types and token numbers at the 4th level was 3.79.

The frequency list for verbs occurring with *-(u)lo* at each level in the IK textbook series is provided in Table 7.8. Two verbs, *isaha-* ‘to move’ (3 tokens) and *ka-* ‘to go’ (2 tokens), occurred more than once at the 1st level volumes. For level 2, the most commonly used verb with *-(u)lo* was *ka-* ‘to go’, whose proportion ranged from 40.54% (level 2) to 11.59% (level 4). *ponay-* ‘to send’ was common at the 3rd and 4th level volumes as well. At the 3rd level, five types of verbs showed identical token numbers. That is, a number of verbs occurred at a similar rate without exhibiting any dominant use. The number of verb types increased as level increased, from five types (level 1) to 36 types (level 4). The ratio between the number of verb types and token numbers was 1.92 at the 4th level of the IK textbook series.

While the total frequency of *-(u)lo* was small compared to *-ey* and *-eyse*, diverse types of verbs were employed, especially in the IK textbook series. In both textbook series, *ka-* ‘to go’ was the most frequently used verb with *-(u)lo* at nearly all levels. As the verb was also the most dominant verb with *-ey* in both textbook series, further inquiries into its use could provide insight into how learners establish form-function mappings in relation to *-ey* and *-(u)lo*.

Table 7. 7. Frequency of verbs occurring with *-(u)lo* in the SK textbook series by level

Rank	Level 1		Level 2		Level 3		Level 4	
	Verb	#	Verb	#	Verb	#	Verb	#
1	<i>kalatha-</i> ‘to transfer’	3	<i>ka-</i> ‘to go’	13	<i>ka-</i> ‘to go’	13	<i>ka-</i> ‘to go’	20
2	<i>isaha-</i> ‘to move’	1	<i>tolaka-</i> ‘to detour’	2	<i>o-</i> ‘to come’	6	<i>nao-</i> ‘to come out’	10
3			<i>tuleka-</i> ‘to enter/go in’	2			<i>tolaka-</i> ‘to detour’	9
4							<i>tol-</i> ‘to turn’	8
5							<i>ollaka-</i> ‘to climb/ go up’	8
			Others	7	Others	7	Others	55
Total		4		24		26		110

Note. Verb types included in Others by level were as follows: 7 in Level 2; 7 in Level 3; 24 in Level 4.

Table 7. 8. Frequency of verbs occurring with *-(u)lo* in the IK textbook series by level

Rank	Level 1		Level 2		Level 3		Level 4	
	Verb	#	Verb	#	Verb	#	Verb	#
1	<i>isaha-</i> ‘to move’	3	<i>ka-</i> ‘to go’	15	<i>ka-</i> ‘to go’	2	<i>ka-</i> ‘to go’	8
2	<i>ka-</i> ‘to go’	2	<i>olmki-</i> ‘to move (sth)’	5	<i>tolaka-</i> ‘to detour’	2	<i>ponay-</i> ‘to send’	4
3			<i>isaha-</i> ‘to move’	3	<i>palcentoy-</i> ‘to be developed’	2	<i>hyangha-</i> ‘to head for’	4
					<i>ponay-</i> ‘to send’	2		
					<i>o-</i> ‘to come’	2		
	Others	3	Others	14	Others	13	Others	53
Total		8		37		23		69

Note. Verb types included in Others by level were as follows: 3 in Level 1; 10 in Level 2; 13 in Level 3; 33 in Level 4.

7.2. Relations between postpositions and verbs

7.2.1. Frequency distribution

The tendencies of verb use for all three postpositions in both textbook series was similar with those of the L1 data. That is, in both the L1 corpora and L2 textbooks, *-ey* was the most common postposition used for the locative construction, while the use of *-(u)lo* was relatively small. In addition, a few dominant verbs occupied a large proportion of the total usage for each postposition in each level.

However, there were slight variations at each level and/or textbook series in terms of the specific verb types that were frequently used. For example, the most frequent verbs co-occurring with *-eyse* in the IK volumes differed among four levels. In addition, those four frequent verbs differed from the most frequent verbs with *-eyse* in the SK volumes. The reasons for these gaps are presumably: i) different topics covered at each level and/or in each textbook, and ii) the need to introduce a variety of vocabulary according to changes in the proficiency level for language learners. Due to this discrepancy, it is not easy to provide a conclusive claim about overall usage tendencies for any possible verb to occur with each postposition in the textbooks.

Thus, in order to attain a thorough understanding about what types of verbs were being introduced in the textbooks for the locative function of the three adverbial postpositions, overall total usage throughout all levels were analyzed. Table 7.9 shows the frequency ranks of each postposition for all of the volumes in both textbook series. Some common verbs were found in both the SK and IK series. The most common verbs occurring with *-ey* was *ka-* ‘to go’ followed by *iss-* ‘to be/exist’ and *o-* ‘to come’. For *-eyse*, the most frequent verb was different (*ha-* ‘to do’ in the SK and *sal-* ‘to live’ in the IK), but these two verbs ranked high in each other’s textbooks. Other verbs such as *manna-* ‘to meet’ also frequently occurred in both textbook series. For -

(u)lo, ka- ‘to go’ was the most frequent verb in both the SK and IK series, with *tolaka-* ‘to detour’ being commonly used as well.

Table 7. 9. Five most frequent verbs occurring with each postposition across all volumes of the SK and IK textbook series

	SK		IK	
	Verb type	Token (#)	Verb type	Token (#)
-ey				
1	<i>ka-</i> ‘to go’	245	<i>ka-</i> ‘to go’	149
2	<i>iss-</i> ‘to be/exist’	82	<i>iss-</i> ‘to be/exist’	72
3	<i>o-</i> ‘to come’	51	<i>o-</i> ‘to come’	36
4	<i>nao-</i> ‘to come out’	23	<i>sal-</i> ‘to live’	25
5	<i>tuleka-</i> ‘to enter/go in’	23	<i>ssu-</i> ‘to write/use’	19
-eyse				
1	<i>ha-</i> ‘to do’	42	<i>sal-</i> ‘to live’	16
2	<i>sal-</i> ‘to live’	27	<i>manna-</i> ‘to meet’	14
3	<i>manna-</i> ‘to meet’	22	<i>o-</i> ‘to come’	14
4	<i>ilha-</i> ‘to work’	15	<i>ha-</i> ‘to do’	13
5	<i>nayli-</i> ‘to get off/fall’	9	<i>sa-</i> ‘to buy’	12
-(u)lo				
1	<i>ka-</i> ‘to go’	46	<i>ka-</i> ‘to go’	27
2	<i>tolaka-</i> ‘to detour’	12	<i>o-</i> ‘to come’	8
3	<i>nao-</i> ‘to come out’	10	<i>tolaka-</i> ‘to detour’	7
4	<i>ollaka-</i> ‘to climb/go up’	9	<i>ponay-</i> ‘to send’	6
5	<i>tol-</i> ‘to turn’	8	<i>olmki-</i> ‘to move (sth)’	6
			<i>isaha-</i> ‘to move’	6

Table 7.9 shows that the most frequent verbs showed high token numbers compared to other verbs with each postposition. Moreover, a small number of verbs were dominantly used with each postposition. To check tendencies more closely, the frequency distributions of verbs were graphed. Frequency distributions for the top 15 most frequent verbs with each postposition in both textbook series are shown in Figures 7.1, 7.2, and 7.3. The x-axes in the Figures encode token frequency while the y-axes mark the frequency order of verbs (from top to bottom) in the

textbooks. Similar to the L1 corpora, Figures clearly show that there were dominant verbs co-occurred with each postposition and token numbers of verbs decreased as their frequency rank decreased.

As this tendency looked similar to the results of the analysis of L1 Korean data, linearity relationships were investigated using log values of frequency rank and token numbers (frequency) of verbs. Inserted images in Figure 7.1, 7.2, and 7.3 show the plots between the *log_rank* (in the x-axis) and *log_frequency* (in the y-axis) of verbs, with R^2 values. All six plots in the Figures show a linear relationship between *log_rank* and *log_frequency*, which proves that verb frequency distributions with each postposition in the textbooks followed a Zipfian distribution. Thus, natural language use in both L1 corpora and textbooks proved to be following Zipf's law. However, some of the verbs that were introduced as presenting the most prototypical meaning for each postposition differed from L1 corpora: *ka-* 'to go' for *-ey* in both types of textbooks and *sal-* 'to live' for *-eyse* in the IK series. It is likely that language learners of Korean who used those textbooks acquired those verbs as pathbreaking ones, which can be further investigated in future studies.

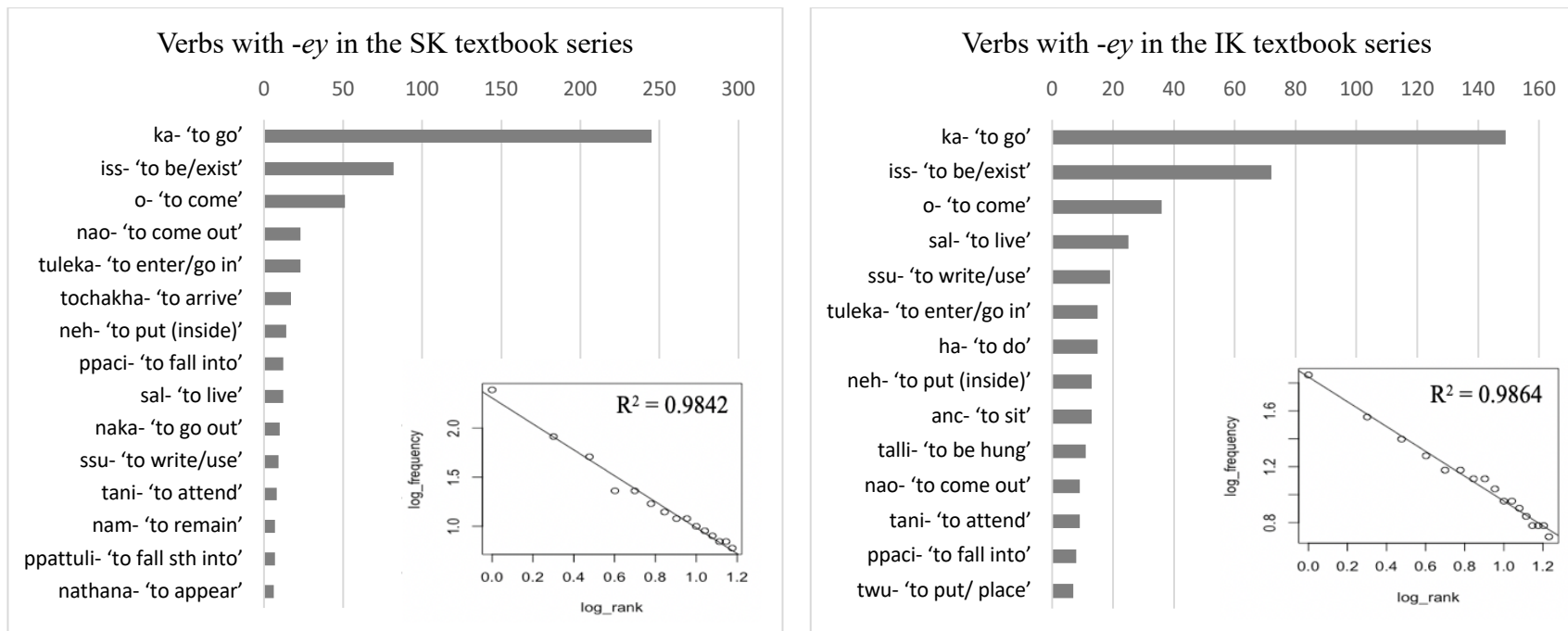


Figure 7. 1. Frequency distributions of the top 15 verbs occurring with -ey across all volumes of the SK and IK textbook series (x-axis: Frequency, y-axis: Lemma form of verbs).

Note. Inserted images show the linear relationship between *log_rank* and *log_frequency*.

Note. Number of verbs in the IK textbook series is 14, due to identical token numbers for several types of verbs.

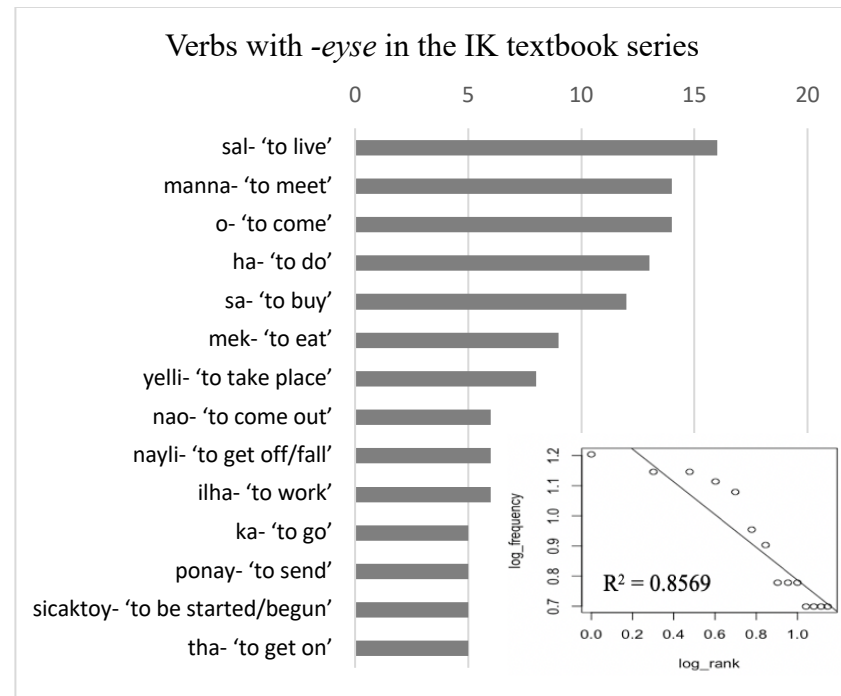
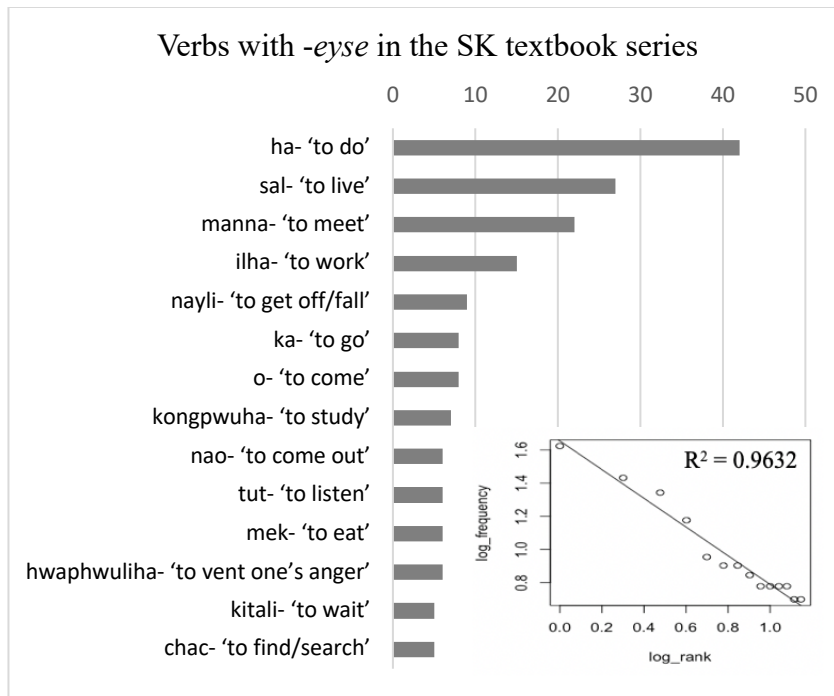


Figure 7. 2. Frequency distributions of the top 15 verbs occurring with *-eyse* across all volumes of the SK and IK textbook series (x-axis: Frequency, y-axis: Lemma form of verbs).

Note. Inserted images show the linear relationship between *log_rank* and *log_frequency*.

Note. Number of verbs in the SK and IK textbook series is 14, due to identical token numbers for several types of verbs.

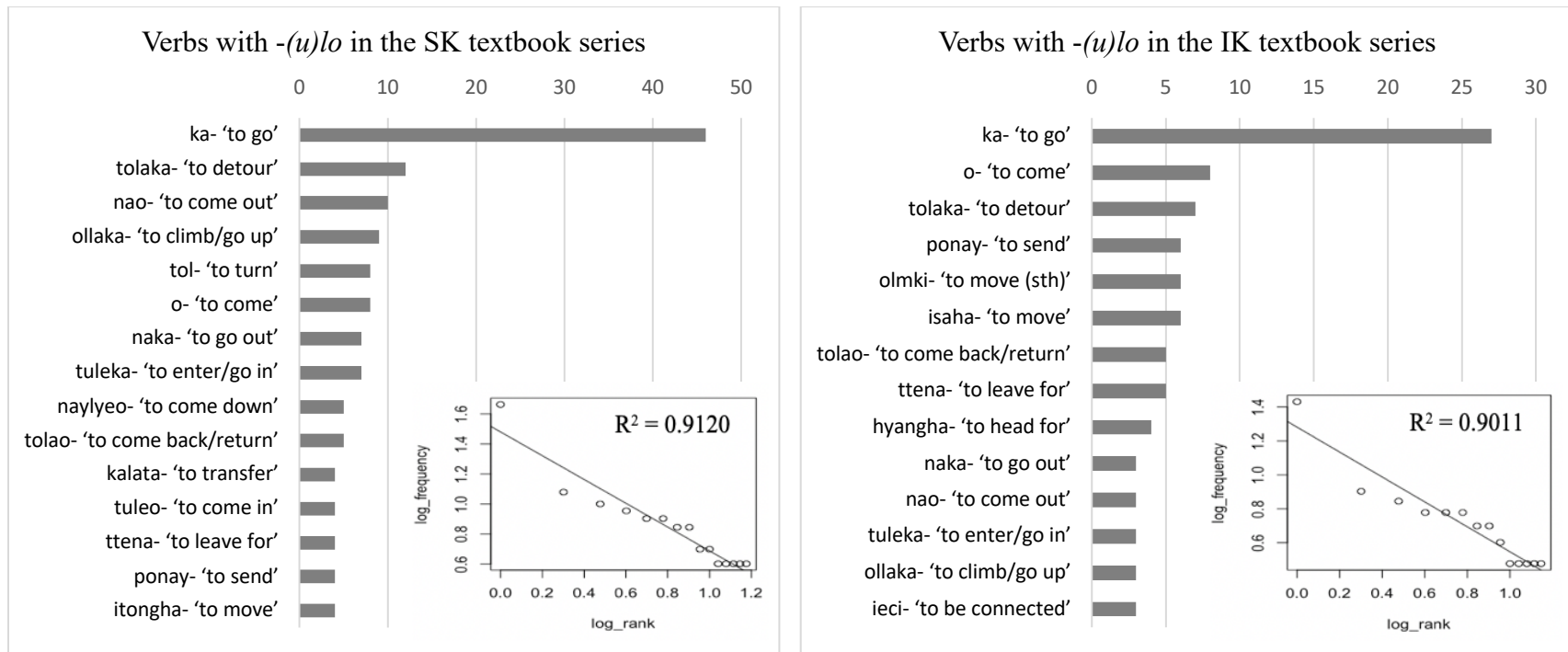


Figure 7. 3. Frequency distributions of the top 15 verbs occurring with *-(u)lo* across all volumes of the SK and IK textbook series (x-axis: Frequency, y-axis: Lemma form of verbs).

Note. Inserted images show the linear relationship between *log_rank* and *log_frequency*.

Note. Number of verbs in the IK textbook series is 14, due to identical token numbers for several types of verbs.

7.2.2. Keyness analysis

To identify key verbs in each of the SK and IK textbooks, keyness analysis was conducted. For keyness analysis of the L1 Sejong corpora, written and spoken corpora were compared with a confidence level of 99.99% in Chapter 6. As for the textbook analysis in the present study, however, a limited number of verb types were employed and the corpus size for each postposition was smaller than the L1 corpora. The number of tokens ranged from 137 tokens for verbs with *-(u)lo* in the IK series and 645 tokens for verbs with *-ey* in the SK series. As a result, setting a lower *p*-value returned extremely limited or no key items. Thus, for textbook comparison purposes, CKIs were examined with 99% confidence levels ($p < 0.01$, critical value of *LL*: 6.63). Higher *LL* means a stronger significant difference between the two textbook types. Effect size was calculated using %DIFF, with IK textbooks used as a reference corpus and SK textbooks used as a study corpus. A high value for %DIFF (>10,000) indicates that a token number for the item in the reference corpus (IK textbooks) is zero, while a %DIFF value of -100 indicates no occurrence of the item in the study corpus (SK textbooks).

In the analysis of verbs for *-ey*, a total of seven CKIs were identified and are listed in Table 7.10. Among them, *ka-* ‘to go’ and *tochakha-* ‘to arrive’ occurred about 1.6 and 5.5 times more frequently in the SK textbooks than in the IK textbooks. On the other hand, *ha-* ‘to do’ was used about five times more frequently in the IK textbooks (i.e., the value of -80 %DIFF means 80% lower frequency in the study corpus). Two other verbs (*naka-* ‘to go out’ and *ppattuli-* ‘to drop’) were not used in the IK volumes, while *talli-* ‘to be hung’ and *moi-* ‘to gather’ did not occur in the SK volumes, but all four of these verbs had considerable token numbers in the other corpus.

Table 7. 10. CKI verbs used with *-ey* in the SK and IK textbooks

Verb type	<i>LL</i>	%DIFF
<i>ka-</i> ‘to go’	29.91	59.33
<i>talli-</i> ‘to be hung’	15.70	-100
<i>naka-</i> ‘to go out’	13.63	> 10,000
<i>tochakha-</i> ‘to arrive’	10.53	449.10
<i>ppattuli-</i> ‘to drop’	9.52	> 10,000
<i>ha-</i> ‘to do’	9.24	-80.62
<i>moi-</i> ‘to gather’	8.54	-100

Note. The *p*-value of *LL* was < .01.

For *-eyse*, five verbs were considered as CKIs, as Table 7.11 shows. The verb *ha-* ‘to do’ was used about three times more in the SK volumes. Two verbs, *yelli-* ‘to take place’ and *sicaktoy-* ‘to get started’ did not occur in the SK textbooks, while the IK textbooks contained considerable numbers of them. The situation is opposite for *hwaphwuliha-* ‘to vent one’s anger’ and *kitali-* ‘to wait’: they did not occur in the IK textbooks while the token numbers for those verbs in the SK volumes appeared at a noteworthy rate of frequency.

Table 7. 11. CKI verbs used with *-eyse* in the SK and IK textbooks

Verb type	<i>LL</i>	%DIFF
<i>ha-</i> ‘to do’	15.45	200.39
<i>yelli-</i> ‘to take place’	11.80	-100
<i>hwaphwuliha-</i> ‘to vent one’s anger’	7.95	> 10,000
<i>sicaktoy-</i> ‘to get started’	7.35	-100
<i>kitali-</i> ‘to wait’	6.61	> 10,000

Note. The *p*-value of *LL* was < .01.

In the case of verbs for *-(u)lo*, no CKIs were found that showed statistically significant differences between the two textbooks series at the level of *p* < 0.01 (critical value of *LL*: 6.63).

The reason for this seems to be that the proportion of each verb with *-(u)lo* in each textbook was quite small, except for a few frequent verbs (e.g., *ka-* ‘to go’) (see Table 7.7 and Table 7.8).

Overall, it was not uncommon to find CKIs that had extremely high %DIFF or -100 (i.e., zero occurrence in one of the two corpora in comparison) in the textbook analysis. Those cases (considerable use in one corpus and no use in the other corpus) seem to have occurred as topics and pedagogical goals for each textbook differed, thereby affecting the deployment of verbs. For instance, *hwaphwuliha-* ‘to vent one’s anger’, which occurred with *-eyse* in the SK textbook series, is a verb commonly associated with an idiomatic expression in a Korean proverb. The verb *hwaphwuliha-* appeared a considerable number of times as a consequence of the specific topic introduced in the SK textbook series, but this was not the case for the IK textbooks. Thus, the selected CKIs in the textbook analysis in Tables 7.10 and 7.11 reflect certain characteristics of both textbook series.

7.2.3. Association strength

7.2.3.1. General relations

In addition to the frequency distribution and keyness analysis, association strengths between each postposition and their co-occurring verbs for the entirety of both textbook series were measured. Like the L1 corpora analysis, ΔP with “postposition” as the cue and “verb” as the outcome ($\Delta P_{(\text{verb} | \text{postposition})}$) was first calculated. Table 7.12 shows the top three verbs that had the highest $\Delta P_{(\text{verb} | \text{postposition})}$ values for each postposition throughout all volumes of both the SK and IK textbook series.

Two verbs, *ka-* ‘to go’ and *iss-* ‘to be/exist’, had the highest association strength with *-ey* for both textbook series. For *-eyse*, *ha-* ‘to do’ was the verb with the highest association strength

in the SK textbooks while *manna-* ‘to meet’ had the highest association strength in the IK textbooks. However, the difference between scores among verbs was not large in the IK textbooks: instead of using *manna-* ‘to meet’ dominantly, some other verbs were used a considerable amount as well. While the most frequently occurring verb with *-(u)lo* was *ka-* ‘to go’, the most strongly associated verb with this postposition was *tolaka-* ‘to detour’ because of a high token frequency for *ka-* ‘to go’ with *-ey* in both the SK and IK textbooks. $\Delta P_{(\text{verb} | \text{postposition})}$ values of verbs with *-(u)lo* in the table do not show a big difference with other verbs within the same textbook series. That is, the token numbers of verbs that were used exclusively or with high association strength with *-(u)lo* were similar. In general, $\Delta P_{(\text{verb} | \text{postposition})}$ values of verbs co-occurring with *-eyse* in the IK textbooks and verbs with *-(u)lo* in both the SK and IK textbooks were small due to either relatively small proportions of the most frequent verbs or overlapping use with other postpositions.

Table 7. 12. Top three $\Delta P_{(\text{verb} | \text{postposition})}$ values for verbs used with each postposition throughout the entire SK and IK textbook series

	SK		IK	
	Verb type	ΔP	Verb type	ΔP
<i>-ey</i>				
1	<i>ka-</i> ‘to go’	0.263	<i>ka-</i> ‘to go’	0.161
2	<i>iss-</i> ‘to be/exist’	0.118	<i>iss-</i> ‘to be/exist’	0.108
3	<i>o-</i> ‘to come’	0.045	<i>ssu-</i> ‘to write/use’	0.023
<i>-eyse</i>				
1	<i>ha-</i> ‘to do’	0.137	<i>manna-</i> ‘to meet’	0.050
2	<i>sal-</i> ‘to live’	0.075	<i>sa-</i> ‘to buy’	0.041
3	<i>manna-</i> ‘to meet’	0.074	<i>mek-</i> ‘to eat’	0.032
<i>-(u)lo</i>				
1	<i>tolaka-</i> ‘to detour’	0.070	<i>tolaka-</i> ‘to detour’	0.049
2	<i>ollaka-</i> ‘to climb/go up’	0.053	<i>olmki-</i> ‘to move (sth)’	0.044
3	<i>tol-</i> ‘to turn’	0.049	<i>isaha-</i> ‘to move’	0.044

$\Delta P_{(\text{postposition} | \text{verb})}$ scores, with “verb” as the cue and “postposition” as the outcome, were calculated in order to find out with which postposition a verb had strong association power. As was mentioned earlier, this metric shows higher values for verbs that combined with only one postposition. Thus, the comparison of $\Delta P_{(\text{postposition} | \text{verb})}$ values among postpositions with low token numbers did not seem to be appropriate to include in the current analysis. To illustrate, in the SK volumes, *phal-* ‘to sell’ occurred once with *-ey* and twice with *-eyse*, resulting in a value of -0.161 with *-ey*, and 0.600 with *-eyse* for $\Delta P_{(\text{postposition} | \text{verb})}$. Even though it indicates that *phal-* ‘to sell’ preferred *-eyse*, the number of tokens were too small to generate meaningful results. Thus, like the analysis of the L1 corpora, verbs that combined with more than two types of postpositions and occurred at least 5 times with one of the postpositions were examined.

Among the verbs that were examined, *tuleka-* ‘to enter/ go in’ and *o-* ‘to come’ showed higher association strength with *-ey* over *-eyse* and *-(u)lo* in the SK textbooks. The verb *iss-* ‘to be/ exist’ had higher association strength with *-ey* over *-eyse*, and *nao-* ‘to come out’ with *-(u)lo* over *-ey* and *-eyse* in the SK textbooks. In the IK textbooks, *iss-* ‘to be/exit’ and *o-* ‘to come’ had higher association strength with *-ey* over *-eyse* and *-(u)lo*. The verbs *tolao-* ‘to return’ and *ponay-* ‘to send’ had strong association with *-(u)lo* over *-ey* and *-eyse*.

$\Delta P_{(\text{postposition} | \text{verb})}$ values for the four verbs that most frequently occurred with each locational postposition in both textbook series were additionally examined. For *-eyse*, the two verbs *ha-* ‘to do’ and *manna-* ‘to meet’ were considered because they were the dominant verbs in each of the SK and IK textbook series (see Table 7.13). The results indicated that *-ey*, *-eyse*, and *-(u)lo* were strongly associated with their dominant verbs: *ka-* ‘to go’ for *-ey*, *ha-* ‘to do’ (SK) and *manna-* ‘to meet’ (IK) for *-eyse*, and *tolaka-* ‘to detour’ for *-(u)lo*, respectively.

Finally, to investigate whether “postposition” or “verb” would be a better predictor, the differences between the two ΔP scores ($\Delta P_{(\text{verb} | \text{postposition})} - \Delta P_{(\text{postposition} | \text{verb})}$) of the most frequent verbs for each postposition were calculated as well. Tables 7.14 and 7.15 show the results for both the SK and IK textbook series. Negative values for differences indicated “verb” as a better predictor of “postposition” rather than the other way around in all cases for both textbook series.

Table 7. 13. $\Delta P_{(\text{postposition} | \text{verb})}$ values for the four verbs *ka-*, *ha-*, *manna-* and *tolaka-* with all three postpositions in the SK and IK textbook series

	SK			IK		
	<i>-ey</i>	<i>-eyse</i>	<i>-(u)lo</i>	<i>-ey</i>	<i>-eyse</i>	<i>-(u)lo</i>
<i>ka-</i> ‘to go’	0.325	-0.333	0.008	0.161	-0.213	-0.005
<i>ha-</i> ‘to do’	-0.537	0.616	–	-0.007	0.027	–
<i>manna-</i> ‘to meet’	–	0.745	–	–	0.050	–
<i>tolaka-</i> ‘to detour’	-0.387	–	0.654	-0.014	–	0.049

Table 7. 14. Differences in ΔP values between all three postpositions and their strongly associated verbs in the SK textbook series

	X = <i>-ey</i> Y = <i>ka-</i> ‘to go’	X = <i>-eyse</i> Y = <i>ha-</i> ‘to do’	X = <i>-(u)lo</i> Y = <i>tolaka-</i> ‘to detour’
$\Delta P_{(Y X)}$	0.263	0.137	0.070
$\Delta P_{(X Y)}$	0.325	0.616	0.654
$\Delta P_{(Y X)} - \Delta P_{(X Y)}$	-0.062	-0.479	-0.584

Table 7. 15. Differences in ΔP values between all three postpositions and their strongly associated verbs in the IK textbook series

	X = <i>-ey</i> Y = <i>ka-</i> ‘to go’	X = <i>-eyse</i> Y = <i>manna-</i> ‘to meet’	X = <i>-(u)lo</i> Y = <i>tolaka-</i> ‘to detour’
$\Delta P_{(Y X)}$	0.161	0.050	0.049
$\Delta P_{(X Y)}$	0.269	0.743	0.652
$\Delta P_{(Y X)} - \Delta P_{(X Y)}$	-0.108	-0.693	-0.603

In all, a cumulative analysis of all tokens within every textbook volume (from level 1 to level 4) for the three locational postpositions showed that a small number of verbs were dominantly used with each postposition (section 7.2.1): *ka-* ‘to go’ with *-ey*, *ha-* ‘to do’ (SK) and *sal-* ‘to live’ (IK) with *-eyse*, and *ka-* ‘to go’ with *-(u)lo*. Among those verbs, *ka-* ‘to go’ had high association strength ($\Delta P_{(\text{verb} | \text{postposition})}$) with *-ey* in both textbook series and *ha-* ‘to do’ was combined with a high degree of association strength with *-eyse* in the SK textbook series. The verb that showed the highest ΔP value for *-eyse* in the IK textbook series was *manna-* ‘to meet’, due to frequent appearance of the most frequent verb *sal-* ‘to live’ with *-ey*. In the same fashion, *-(u)lo* had high association strength with *tolaka-* ‘to detour’ because of the frequent use of *ka-* ‘to go’ with *-ey*. Those verbs that showed higher ΔP values for each postposition were the better predictor of the postposition than the other way around. In the case of the association strength for the other way around ($\Delta P_{(\text{postposition} | \text{verb})}$), each postposition had high association strength values for their dominant verbs.

7.2.3.2. Association strength by textbook level

Next, an analysis was conducted of what types of verbs had high association strength with each postposition according to level and how association strengths changed between proficiency levels for each textbook series. Association strengths in both directions ($\Delta P_{(\text{verb} | \text{postposition})}$ and $\Delta P_{(\text{postposition} | \text{verb})}$), as well as the difference between them, were calculated and compared for each postposition. In calculating $\Delta P_{(\text{postposition} | \text{verb})}$, only verbs that had higher association strengths with each postposition ($\Delta P_{(\text{verb} | \text{postposition})}$) at each level were considered.

Table 7.16 shows three verbs that ranked highly in terms of association strength ($\Delta P_{(\text{verb} | \text{postposition})}$): “postposition” as the cue and “verb” as the outcome) between *-ey* and verbs at each

level of both textbook series. As a limited number of verbs were introduced in the textbooks without overlapping use with other postpositions, the rank of highest $\Delta P_{(\text{verb} | \text{postposition})}$ scores was generally similar to the order of frequency. In both SK and IK textbooks, the verbs that had the highest association power were *ka-* ‘to go’ followed by *iss-* ‘to be/exist’ at levels 1 to 3. For level 4, *iss-* ‘to be/exist’ had the highest association power for both textbook series.

Table 7. 16. Top three $\Delta P_{(\text{verb} | \text{postposition})}$ values for verbs used with *-ey* at each level in the SK and IK textbook series

		SK		IK	
		Verb type	ΔP	Verb type	ΔP
Level 1					
1	<i>ka-</i> ‘to go’		0.681	<i>ka-</i> ‘to go’	0.411
2	<i>iss-</i> ‘to be/exist’		0.140	<i>iss-</i> ‘to be/exist’	0.253
3	<i>o-</i> ‘to come’		0.066	<i>o-</i> ‘to come’	0.040
Level 2					
1	<i>ka-</i> ‘to go’		0.281	<i>ka-</i> ‘to go’	0.302
2	<i>iss-</i> ‘to be/exist’		0.073	<i>iss-</i> ‘to be/exist’	0.064
3	<i>o-</i> ‘to come’		0.067	<i>tani-</i> ‘to attend’	0.034
				<i>tuleka-</i> ‘to enter/go in’	0.034
Level 3					
1	<i>ka-</i> ‘to go’		0.198	<i>ka-</i> ‘to go’	0.225
2	<i>iss-</i> ‘to be/exist’		0.116	<i>iss-</i> ‘to be/exist’	0.082
3	<i>tuleka-</i> ‘to enter/go in’		0.047	<i>o-</i> ‘to come’	0.051
Level 4					
1	<i>iss-</i> ‘to be/exist’		0.139	<i>iss-</i> ‘to be/exist’	0.101
2	<i>o-</i> ‘to come’		0.066	<i>ssu-</i> ‘to write/use’	0.045
3	<i>ppaci-</i> ‘to fall into’		0.040	<i>talli-</i> ‘to be hung’	0.037

Differences between both directions of ΔP scores for *-ey* and verbs at each level of the SK and IK textbook series are shown in Table 7.17 and Table 7.18. At the level 1 volumes for the SK series, *-ey* was a better predictor of *ka-* ‘to go’: *-ey* had stronger predicting power over *ka-* ‘to go’ than *ka-* ‘to go’ had over *-ey*. Neither one was a better predictor of each other at level

2. At level 3 and 4, the verbs *ka-* ‘to go’ and *iss-* ‘to be/exist’ were the better predictors of *-ey*. In the IK volumes, verbs were the better predictors of *-ey* at all levels. That is, *ka-* ‘to go’ and *iss-* ‘to be/exist’ attracted *-ey* more than *-ey* necessarily attracting those verbs at the 3rd and 4th level of the SK textbook series and at all levels of the IK textbooks series.

Table 7. 17. Differences in ΔP between *-ey* and verbs at each level in the SK textbook series

	Level 1	Level 2	Level 3	Level 4
	X = <i>-ey</i>	X = <i>-ey</i>	X = <i>-ey</i>	X = <i>-ey</i>
	Y = <i>ka-</i> ‘to go’	Y = <i>ka-</i> ‘to go’	Y = <i>ka-</i> ‘to go’	Y = <i>iss-</i> ‘to be/exist’
$\Delta P_{(Y X)}$	0.681	0.281	0.198	0.139
$\Delta P_{(X Y)}$	0.549	0.272	0.230	0.544
$\Delta P_{(Y X)} - \Delta P_{(X Y)}$	0.132	0.009	-0.032	-0.405

Table 7. 18. Differences in ΔP between *-ey* and verbs at each level in the IK textbook series

	Level 1	Level 2	Level 3	Level 4
	X = <i>-ey</i>	X = <i>-ey</i>	X = <i>-ey</i>	X = <i>-ey</i>
	Y = <i>ka-</i> ‘to go’	Y = <i>ka-</i> ‘to go’	Y = <i>ka-</i> ‘to go’	Y = <i>iss-</i> ‘to be/exist’
$\Delta P_{(Y X)}$	0.411	0.302	0.225	0.101
$\Delta P_{(X Y)}$	0.427	0.306	0.490	0.400
$\Delta P_{(Y X)} - \Delta P_{(X Y)}$	-0.026	-0.004	-0.265	-0.299

Table 7.19 shows the top three verbs that ranked highest for association strength ($\Delta P_{(\text{verb} | \text{postposition})}$) with *-eyse* in both textbook series. In the SK volumes, *manna-* ‘to meet’ showed the highest score at the 1st level, while *ha-* ‘to do’ scored highest at the other levels. In the IK volumes, the verbs that showed the highest association power differed at each level: *manna-* ‘to meet’ at level 1, *sa-* ‘to buy’ at level 2, *sicaktoy-* ‘to be started’ at level 3, and *yelli-* ‘to take place’ at level 4. While some verbs such as *sal-* ‘to live’ and *manna-* ‘to meet’ overlapped across levels or textbook series, *sicaktoy-* ‘to be started’, *yelli-* ‘to take place’, and *kaychoytoy-* ‘to be held’ at the 3rd and 4th level of IK volumes were unique, possibly due to the topics covered at

those levels. Also, as a number of different types of verb were used in the IK textbook series, the differences between ΔP values was small amongst verbs co-occurring with each postposition within the same textbook series.

Table 7. 19. Top three $\Delta P_{(\text{verb} | \text{postposition})}$ values for verbs used with *-eyse* at each level in the SK and IK textbook series

	SK		IK	
	Verb type	ΔP	Verb type	ΔP
Level 1				
1	<i>manna-</i> ‘to meet’	0.256	<i>manna-</i> ‘to meet’	0.143
2	<i>ha-</i> ‘to do’	0.186	<i>sal-</i> ‘to live’	0.123
3	<i>kongpwuha-</i> ‘to study’	0.093	<i>mek-</i> ‘to eat’	0.114
	<i>sal-</i> ‘to live’	0.093	<i>sa-</i> ‘to buy’	0.114
Level 2				
1	<i>ha-</i> ‘to do’	0.161	<i>sa-</i> ‘to buy’	0.136
2	<i>manna-</i> ‘to meet’	0.074	<i>manna-</i> ‘to meet’	0.085
3	<i>ilha-</i> ‘to work’	0.056	<i>ha-</i> ‘to do’	0.074
Level 3				
1	<i>ha-</i> ‘to do’	0.140	<i>sicaktoy-</i> ‘to be started’	0.058
2	<i>sal-</i> ‘to live’	0.125	<i>manna-</i> ‘to meet’	0.039
3	<i>manna-</i> ‘to meet’	0.068	<i>o-</i> ‘to come’	0.039
	<i>ilha-</i> ‘to work’	0.068		
Level 4				
1	<i>ha-</i> ‘to do’	0.109	<i>yelli-</i> ‘to take place’	0.065
2	<i>sal-</i> ‘to live’	0.062	<i>kaychoytoy-</i> ‘to be held’	0.037
3	<i>ilha-</i> ‘to work’	0.047	<i>na-</i> ‘to appear/sprout’	0.034
	<i>hwaphwuliha-</i> ‘to vent one’s anger’	0.047		

Reversed direction in association strengths ($\Delta P_{(\text{postposition} | \text{verb})}$) between *-eyse* and verbs with higher $\Delta P_{(\text{verb} | \text{postposition})}$ was also calculated. Tables 7.20 and Table 7.21 summarize the results for the SK and IK volumes, respectively. All the differences between these two values at each level for both textbook series were negative: verbs were the better predictors than post

positions. That is, verbs such as *ha-* ‘to do’, *manna-* ‘to meet’, *sa-* ‘to buy’, *sicaktoy-* ‘to be started’, and *yelli-* ‘to take place’ were stronger predictors of *-eyse* than vice versa.

Table 7. 20. Differences in ΔP between *-eyse* and verbs at each level in the SK textbook series

	Level 1	Level 2	Level 3	Level 4
	X = <i>-eyse</i>	X = <i>-eyse</i>	X = <i>-eyse</i>	X = <i>-eyse</i>
	Y = <i>manna-</i> ‘to meet’	Y = <i>ha-</i> ‘to do’	Y = <i>ha-</i> ‘to do’	Y = <i>ha-</i> ‘to do’
$\Delta P_{(Y X)}$	0.256	0.161	0.140	0.100
$\Delta P_{(X Y)}$	0.796	0.695	0.603	0.731
$\Delta P_{(Y X)} - \Delta P_{(X Y)}$	-0.540	-0.534	-0.463	-0.631

Table 7. 21. Differences in ΔP between *-eyse* and verbs at each level in the IK textbook series

	Level 1	Level 2	Level 3	Level 4
	X = <i>-eyse</i>	X = <i>-eyse</i>	X = <i>-eyse</i>	X = <i>-eyse</i>
	Y = <i>manna-</i> ‘to meet’	Y = <i>sa-</i> ‘to buy’	Y = <i>sicaktoy-</i> ‘to be started’	Y = <i>yelli-</i> ‘to take place’
$\Delta P_{(Y X)}$	0.143	0.136	0.058	0.065
$\Delta P_{(X Y)}$	0.752	0.784	0.462	0.783
$\Delta P_{(Y X)} - \Delta P_{(X Y)}$	-0.609	-0.648	-0.404	-0.718

Table 7.22 shows a list of verbs that had high association strengths with *-(u)lo* ($\Delta P_{(\text{verb} | \text{postposition})}$) at each level of both the SK and IK textbook series. In the SK volumes, *kalatha-* ‘to transfer’ (level 1), *ka-* ‘to go’ (level 2 to level 3), and *tolaka-* ‘to detour’ (level 4) had high association strength with *-(u)lo*. At level 4 of the SK series, differences in $\Delta P_{(\text{verb} | \text{postposition})}$ values among verbs were small. In the IK volumes, the three verbs *isaha-* ‘to move’, *olmki-* ‘to move (sth)’, and *ka-* ‘to go’ were highly associated with *-(u)lo* at level 1, level 2, and level 4, respectively. For level 3, four verbs (*tolaka-* ‘to detour’, *palcentoy-* ‘to be developed’, *ponay-* ‘to send’, and *o-* ‘to come’) each occurred twice, and they had the same association strength with *-(u)lo*.

Table 7. 22. Top three $\Delta P_{(\text{verb} | \text{postposition})}$ values for verbs used with *-(u)lo* at each level in the SK and IK textbook series

	SK		IK	
	Verb type	ΔP	Verb type	ΔP
Level 1				
1	<i>kalatha-</i> ‘to transfer’	0.732	<i>isaha-</i> ‘to move’	0.375
2	<i>isaha-</i> ‘to move’	0.244	<i>tolaka-</i> ‘to detour’	0.125
3			<i>tuleka-</i> ‘to enter/go in’	0.116
Level 2				
1	<i>ka-</i> ‘to go’	0.197	<i>olmki-</i> ‘to move (sth)’	0.135
2	<i>tolaka-</i> ‘to detour’	0.078	<i>isaha-</i> ‘to move’	0.081
3	<i>tuleka-</i> ‘to enter/go in’	0.054	<i>chotayha-</i> ‘to invite’	0.054
Level 3				
1	<i>ka-</i> ‘to go’	0.248	<i>tolaka-</i> ‘to detour’	0.087
2	<i>o-</i> ‘to come’	0.161	<i>palcentoy-</i> ‘to be developed’	0.087
3	<i>keleka-</i> ‘to go (by walking)’	0.038	<i>ponay-</i> ‘to send’	0.087
	<i>ttena-</i> ‘to leave for’	0.038	<i>o-</i> ‘to come’	0.087
	<i>annayha-</i> ‘to guide’	0.038		
Level 4				
1	<i>tolaka-</i> ‘to detour’	0.079	<i>ka-</i> ‘to go’	0.081
2	<i>ka-</i> ‘to go’	0.076	<i>hyangha-</i> ‘to head for’	0.058
3	<i>tol-</i> ‘to turn’	0.073	<i>ttena-</i> ‘to leave for’	0.043
			<i>ollaka-</i> ‘to climb/ go up’	0.043

Association strength with verbs that had higher $\Delta P_{(\text{verb} | \text{postposition})}$ as the cue and postposition *-(u)lo* as the outcome ($\Delta P_{(\text{postposition} | \text{verb})}$) was also calculated. Table 7.23 shows the differences in ΔP values between postpositions and verbs at each level in the SK textbook series. For the first three levels of the SK volumes, the postposition *-(u)lo* was a better predictor of verbs: *kalatha-* ‘to transfer’ (level 1) and *ka-* ‘to go’ (level 2 and 3) attracted *-(u)lo* more than *-(u)lo* did for those verbs. At the 4th level, *tolaka-* ‘to detour’ was a better predictor of *-(u)lo* rather than vice versa.

Table 7.24 shows the results from the IK volumes. Verbs were better predictors of the postposition *-(u)lo* at all proficiency levels in the IK textbooks. Especially for the 3rd level, four

verbs had the same $\Delta P_{(\text{verb} | \text{postposition})}$ value. Among them, three verbs (*tolaka-* ‘to detour’, *palcentoy-* ‘to be developed’, and *ponay-* ‘to send’) showed the same value for $\Delta P_{(\text{postposition} | \text{verb})}$ at 0.894. Thus, all three verbs were better predictors of *-eyse* than the other way around. The remaining verb was *o-* ‘to come’, whose $\Delta P_{(\text{postposition} | \text{verb})}$ value was 0.090. The difference between $\Delta P_{(\text{verb} | \text{postposition})}$ and $\Delta P_{(\text{postposition} | \text{verb})}$ for *o-* ‘to come’ and *-eyse* was -0.003 (=0.087-0.090), which does not indicate either variable as being a strong predictor of the other.

Table 7. 23. Differences in ΔP between *-(u)lo* and verbs at each level in the SK textbook series

	Level 1	Level 2	Level 3	Level 4
	X = <i>-(u)lo</i>	X = <i>-(u)lo</i>	X = <i>-(u)lo</i>	X = <i>-(u)lo</i>
	Y = <i>kalatha-</i> ‘to transfer’	Y = <i>ka-</i> ‘to go’	Y = <i>ka-</i> ‘to go’	Y = <i>tolaka-</i> ‘to detour’
$\Delta P_{(Y X)}$	0.732	0.197	0.248	0.079
$\Delta P_{(X Y)}$	0.494	0.079	0.109	0.666
$\Delta P_{(Y X)} - \Delta P_{(X Y)}$	0.238	0.118	0.139	-0.587

Table 7. 24. Differences in ΔP between *-(u)lo* and verbs at each level in the IK textbook series

	Level 1	Level 2	Level 3	Level 4
	X = <i>-(u)lo</i>	X = <i>-(u)lo</i>	X = <i>-(u)lo</i>	X = <i>-(u)lo</i>
	Y = <i>isaha-</i> ‘to move’	Y = <i>olmki-</i> ‘to move (sth)’	Y = <i>tolaka-</i> ‘to detour’	Y = <i>ka-</i> ‘to go’
$\Delta P_{(Y X)}$	0.375	0.135	0.087	0.081
$\Delta P_{(X Y)}$	0.959	0.866	0.892	0.228
$\Delta P_{(Y X)} - \Delta P_{(X Y)}$	-0.584	-0.731	-0.805	-0.147

Note. For level 3, three verbs (*tolaka-* ‘to detour’, *palcentoy-* ‘to be developed’, and *ponay-* ‘to send’) showed the same $\Delta P_{(\text{postposition} | \text{verb})}$ of 0.892.

7.3. Summary of findings

Some similarities and differences were observed between the SK and IK textbooks. Both textbook series incorporated a variety of verbs for forming locative phrases with the three

postpositions across proficiency levels. Among various verbs, however, there was one dominant verb for each postposition when all of the volumes were considered together: *ka-* ‘to go’ for *-ey*, *ha-* ‘to do’ (SK) and *sal-* ‘to live’ (IK) for *-eyse*, and *ka-* ‘to go’ for *-(u)lo*. In addition to those dominant verbs, some verbs were used more frequently than others while the token numbers proportionally decreased. The analysis of linearity between *log_frequency* and *log_rank* indicated that the distribution of verb frequency for each postposition in each type of textbook followed a Zipfian distribution.

Also, both types of textbooks utilized slightly different types of verbs for the same postposition. Keyness analysis showed that, for *-ey*, verbs that had significant differences in frequency between the two types of textbooks include *ka-* ‘to go’, *talli-* ‘to be hung’, and *naka-* ‘to go out’. For *-eyse*, some verbs such as *ha-* ‘to do’, *yelli-* ‘to take place’, and *hwaphwuliha-* ‘to vent one’s anger’ showed statistically significant differences in the keyness analysis. The difference between the two types of textbooks with *-ey* and *-eyse* seems to have resulted from the particular topics introduced each textbook series. No verbs that occurred with *-(u)lo* were found to be significantly different, possibly because of their small token numbers.

In the SK textbooks, association strengths of verbs ($\Delta P_{(\text{verb} | \text{postposition})}$) with each postposition in the total usage (from levels 1 to 4) was similar to the frequency rank for *-ey* and *-eyse*. That is, verbs used with each postposition generally occurred exclusively with those postpositions. For *-(u)lo*, instead of *ka-* ‘to go’ which was the most frequent verb with the postposition, *tolaka-* ‘to detour’ had the strongest association power with *-(u)lo*. Regarding association strength in the IK volumes, the most frequent verb *ka-* ‘to go’ showed a higher $\Delta P_{(\text{verb} | \text{postposition})}$ value with *-ey*. *manna-* ‘to meet’ and *tolaka-* ‘to detour’ were strongly associated with *-eyse* and *-(u)lo* respectively in the IK textbooks. The verb *o-* ‘to come’ was used

with all three postpositions frequently and it did not show high association power with the three postpositions in the IK textbook series. Reversed direction of $\Delta P_{(\text{postposition} | \text{verb})}$ scores in both textbook series revealed that dominant verbs with each postposition (*ka-* ‘to go’ for *-ey*, *ha-* ‘to do’ and *manna-* ‘to meet’ for *-eyse*, and *tolaka-* ‘to detour’ for *-(u)lo*) were better predictors of the postpositions with which they occurred.

Overall, verb distributions followed Zipf’s law in the textbooks and some dominant verbs with each postposition were quite similar in both textbook series. Both series introduced various types of verbs with each of the three locative postpositions as part of the inherently unique nature of topics and themes utilized as units of instruction.

VIII. COMPARISON BETWEEN L1 CORPORA AND L2 TEXTBOOKS

The purpose of learning a second/foreign language is to acquire knowledge of conventional language use by target L1 users. Language textbooks work as a main input source for language learners. Thus, examining language use in textbooks with reference to L1 usage data will clearly show the properties of the target language in each corpus. More specifically, this way of comparison will explicate how language has been presented to language learners, how and in what aspects textbooks (dis)resemble the target language, and how to discover pedagogical strategies.

In Chapter 6 and 7, the use of three locational postpositions *-ey*, *-eyse*, and *-(u)lo* was analyzed in terms of the types and proportions of verbs that co-occur with these postpositions in the L1 Sejong written/spoken corpora and textbooks for language learners of Korean, both within a single proficiency level and across all levels. In this chapter, how these postpositions are presented in the both textbook series will be examined using L1 corpora as references. As the sizes differ between the L1 corpora and L2 textbooks, I utilized the percentages for each verb, which work as a normalized frequency.

8.1. Proportions across proficiency level

In this section, the proportions of the two most frequent verbs with each postposition in the L1 corpora were examined. The reasons for comparing only two verbs are that L1 written and spoken corpora deployed different verbs starting from the third most frequent ranking, and more importantly, textbooks did not fully utilize the other frequently used verbs in L1 corpora.

8.1.1. -ey

In both the L1 written and spoken corpora, the most frequent verb that occurred with *-ey* was *iss-* ‘to be/exist’ (written: 9.45%, spoken: 30.77%), followed by *ka-* ‘to go’ (written: 4.37%, spoken: 14.04%). The proportion of *iss-* ‘to be/exist’ was about twice that of *ka-* ‘to go’ in each register. These two verbs represent two fundamental functions of locative *-ey*: locational-existence and locational-destination. They were also the top two verbs that were most commonly throughout all volumes (from the 1st to 4th level) for each textbook series

In the SK textbook, *ka-* ‘to go’ was the dominant verb used to introduce locative *-ey* during first level courses, while *iss-* ‘to be/exist’ was the second most commonly used verb. Although *ka-* ‘to go’ made up the largest proportion across all levels, this decreased as the level went up. The proportions of the two verbs *ka-* ‘to go’ and *iss-* ‘to be/exist’ were nearly the same in the 4th level volumes (Figure 8.1). The percentages of the two verbs were larger compared to the L1 written corpus, particularly *ka-* ‘to go’. In relation to the L1 spoken corpus, *iss-* ‘to be/exist’ was not used frequently enough to resemble its use in the Sejong corpus.

In the IK volumes, an emphasis was also placed on the locational-destination function at the first three levels, with dominant usage of *ka-* ‘to go’. The proportion also decreased at higher levels, ultimately resulting in a 4~5% proportion of the total usage at the 4th level (see Figure 8.1). In the 4th level volumes, the usage rate of *ka-* ‘to go’ was about half of the verb *iss-* ‘to be/exist’. The proportion and ratio of the two verbs at the 4th level resembled those of the L1 Sejong written corpus: the two verbs took up about 15% of total usage while *iss-* ‘to be/exist’ was used twice as often as *ka-* ‘to go’. Compared to the Sejong spoken corpus, the combined percentages of these two verbs composed about one third of the total usage.

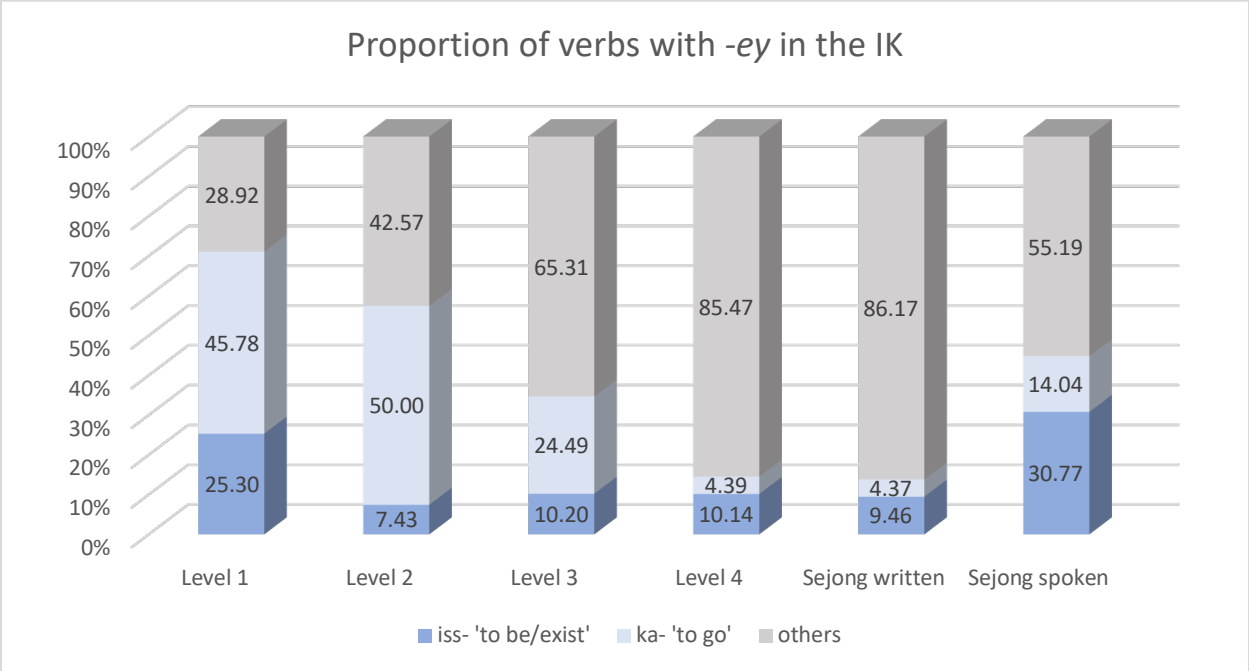
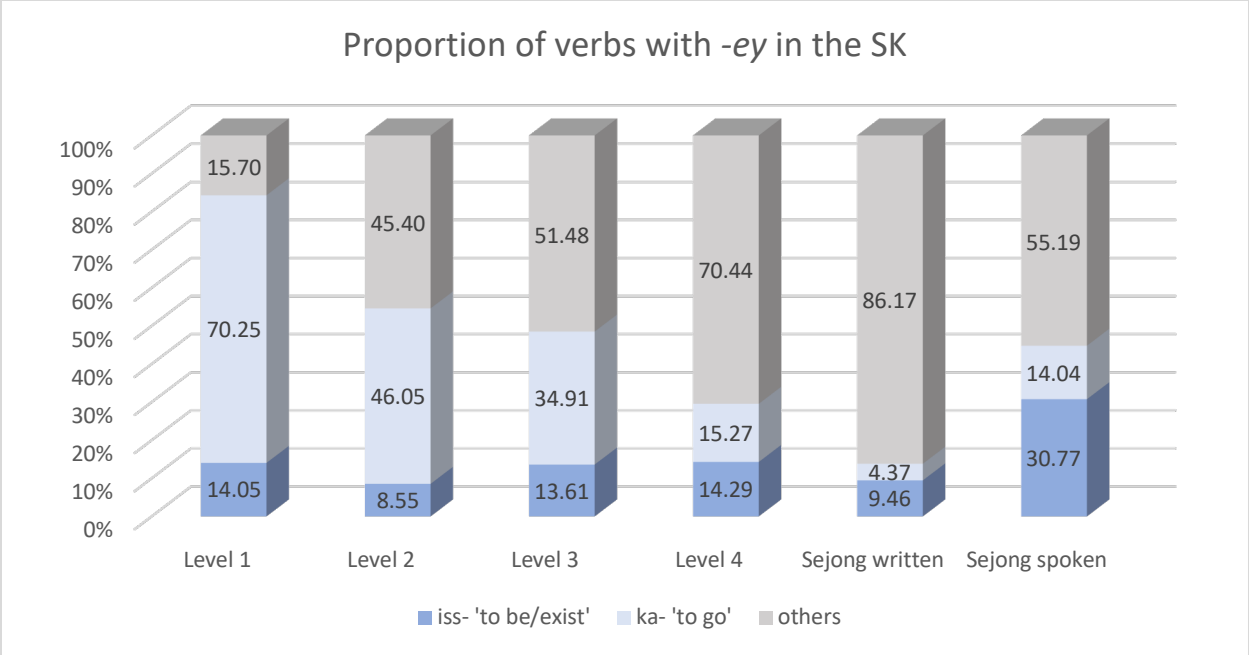


Figure 8. 1. Proportion change of verbs used with -ey in the SK and IK textbook series

It is clear that both textbook series were heavily dependent on *ka-* 'to go' to introduce the locative function of -ey at the beginning. Presumably, this verb serves as a pathbreaker for the

locative construction with *-ey* in Korean for language learners. After the introduction of the target postposition with *ka-* ‘to go’, both textbooks expanded to using other verbs. The proportion of *iss-* ‘to be/exist’, which was the most dominant verb with *-ey* in the L1 corpora, became similar to either the L1 spoken (SK) or written (IK) corpora.

8.1.2. *-eyse*

The most frequent verb used with *-eyse* as a locative was *ha-* ‘to do’ in both the L1 written and spoken corpora. The spoken proportion (9.99%) was about two times higher than that of the written one (4.43%). The second most frequent verb was *nao-* ‘to come out’. This section examines the proportions of these two verbs. Figure 8.2 shows changes in the proportions of these verbs in both textbook series as compared to L1 proportions.

The SK textbooks introduced *ha-* ‘to do’ at quite a high rate across all levels, even though the proportion decreased slightly as level increased. In the case of *nao-* ‘to come out’, it did not appear in the level 1 and level 3 volumes. Both verbs in level 4 were more frequently used than in the L1 written corpus. The ratio of *ha-* ‘to do’ at level 4 approached close to that of the L1 spoken corpus, taking up about 10%. The ratio of *nao-* ‘to come out’ was about half of the L1 spoken proportion.

The proportion of *ha-* ‘to do’ at the 1st level of IK textbooks was 5.71%. It fluctuated for the 2nd and 3rd level volumes, finally taking up about 3.7% in the 4th level. *nao-* ‘to come out’ was used at the 3rd and 4th level. The proportions of these two verbs in the IK volumes were quite close to the L1 written corpus. In relation to the spoken corpus, *ha-* ‘to do’ and *nao-* ‘to come out’ were underused, reaching about half the amount of usage in the Sejong spoken data.

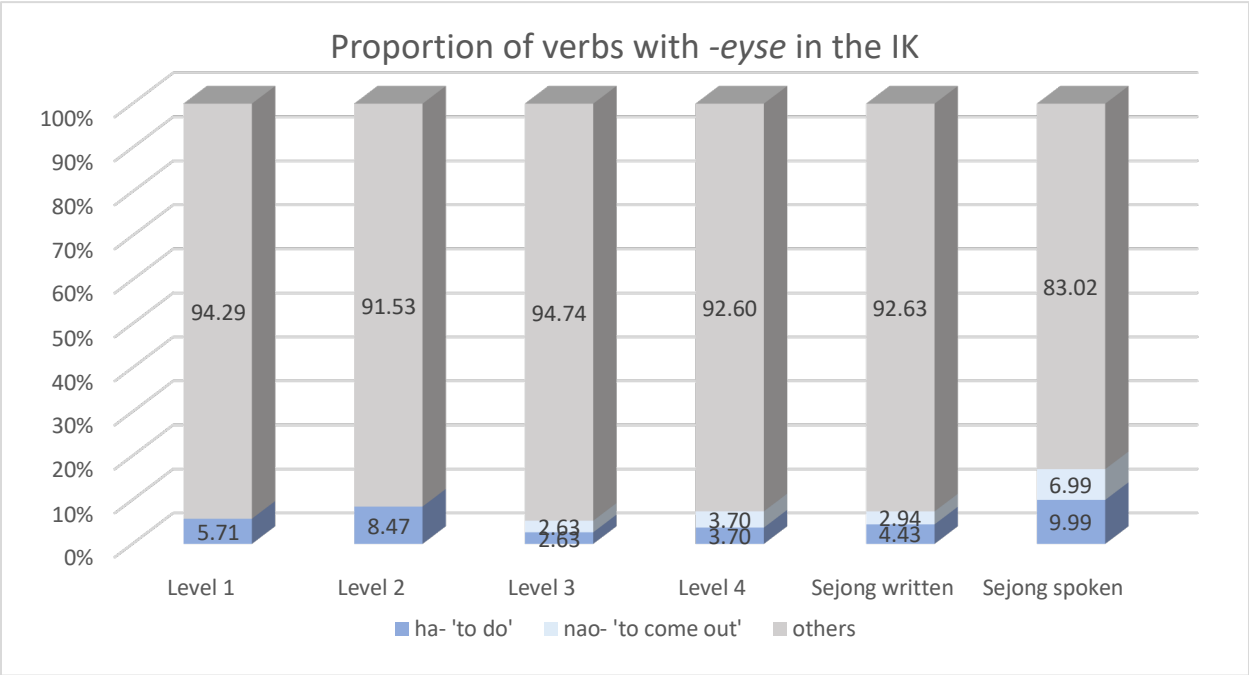
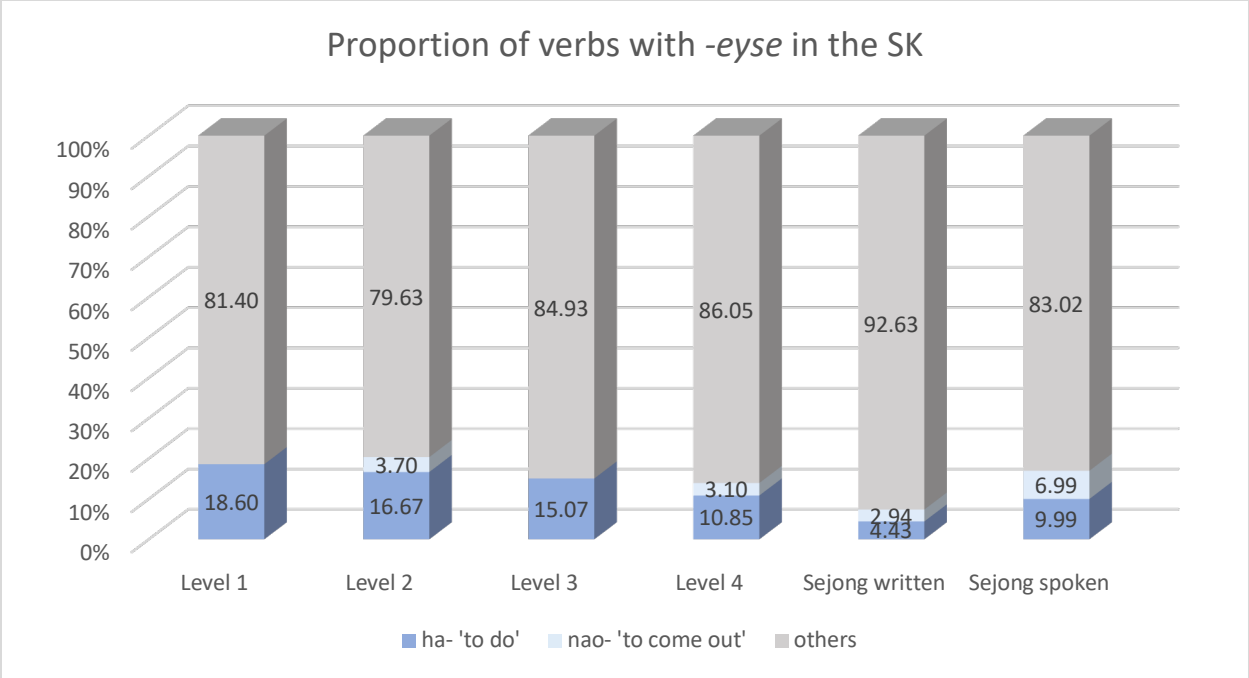


Figure 8. 2. Proportion change of verbs used with *-eyse* in the SK and IK textbook series

In general, the use of verbs with *-eyse* at the 4th level volumes of the SK textbooks resembled the use of verbs in the L1 spoken corpus, with *ha-* ‘to do’ having the highest and *nao-*

‘to come out’ having the second highest proportions. Meanwhile, the proportion of *nao-* ‘to come out’ was a little low. In contrast, verb usage for *-eyse* at the 4th level volumes of the IK textbook series resembled the written L1 data with nearly equal proportions for *ha-* ‘to do’ and *nao-* ‘to come out’.

8.1.3. *-(u)lo*

In the L1 Sejong corpus, *ka-* ‘to go’ was the most frequently used verb with *-(u)lo* in both the written (10% of total usage) and spoken (33% of total usage) corpora while *tuleka-* ‘to enter/go in’ ranked second. Changes in the proportions for these two verbs at each level of both the SK and IK textbook series are displayed in Figure 8.3, along with their L1 corpora proportions.

The SK textbooks introduced locative *-(u)lo* with other verbs besides *ka-* ‘to go’ and *tuleka-* ‘to enter/go in’ at the 1st level. At the 2nd and 3rd levels, the proportion of *ka-* ‘to go’ was at or even higher than 50%, eventually dropping to about 18% at the 4th level. *tuleka-* ‘to enter/go in’ also appeared at the 2nd level and then reappeared at the 4th level. Compared to the L1 Sejong written corpus, *ka-* ‘to go’ was overused while the proportion of *tuleka-* ‘to enter/go in’ was nearly the same. In relation to the L1 spoken corpus, *ka-* and *tuleka-* were used half as much. Although the percentages for both verbs were halved, this change throughout the SK volumes seemed quite moderate in relation to the L1 spoken data. The SK series introduced both verbs from level 2 with high proportions, but gradually used them less frequently at the higher levels. Also, the ratio of verb types occurring with *-(u)lo* was 4.97 in the SK and 6.76 in the L1 spoken corpora: the SK volumes introduced a new type of verb more frequently (about every 5 tokens) compared to the L1 spoken corpus (about every 6~7 tokens). The noticeable decrease in the

usage proportions for those two verbs at the 4th level can be at least partially attributed to the use of a wider variety of verbs in the SK textbooks.

In the IK series, the proportion of *ka-* ‘to go’ was relatively high compared to other verbs at the first two levels but decreased to 11.59% at the 4th level. Meanwhile, *tuleka-* ‘to enter/go in’ was introduced at the 1st level and reappeared in the 4th level. In general, the usage of the two verbs for *-(u)lo* in the IK volumes resembled that of the L1 written corpora at the 4th level. Proportions of the two verbs were identical to those of the L1 written corpus. The ratio of the two verbs was similar to the L1 spoken corpus as well, which was around a 2.5 time decrease in the proportions for each verb.

While there were fluctuations in the proportions of verbs for *-(u)lo* throughout the four proficiency levels for each textbook series, the usage of two verbs *ka-* ‘to go’ and *tuleka-* ‘to enter/go in’ at the 4th level showed very similar trends to that of the L1 usage data. The proportions for both verbs in the SK textbook series were half of those from the L1 spoken corpus. In the case of the IK textbook series, the percentages for both verbs were quite similar to those of the L1 written corpus and about 1/3 of those in the L1 spoken corpus.

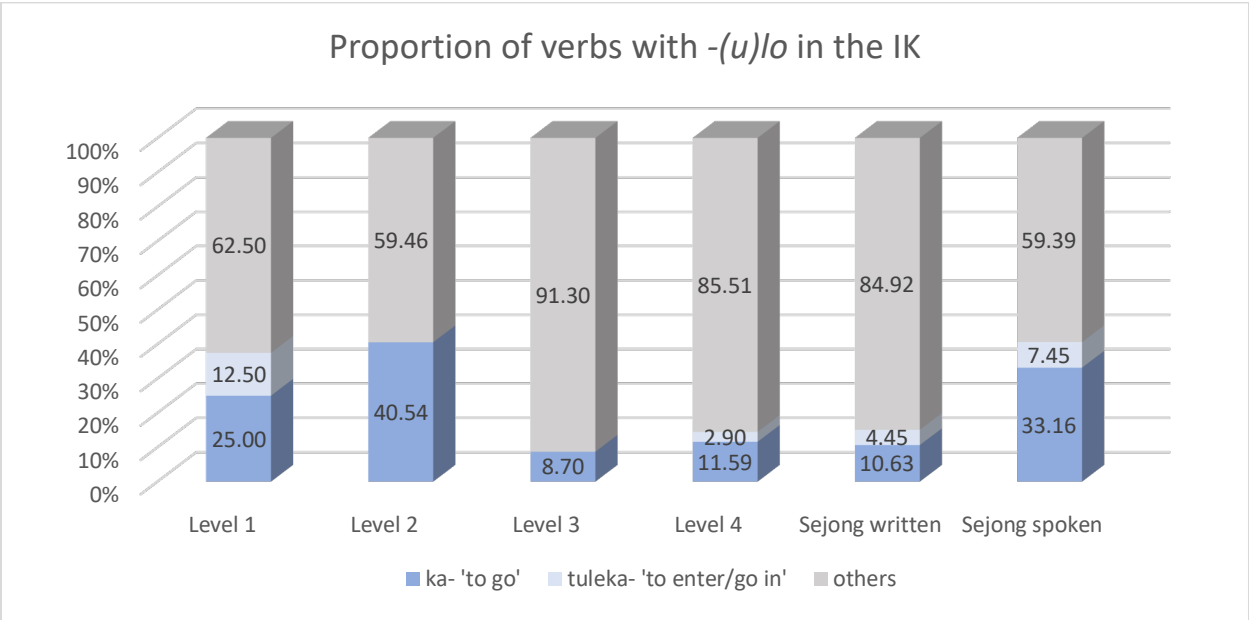
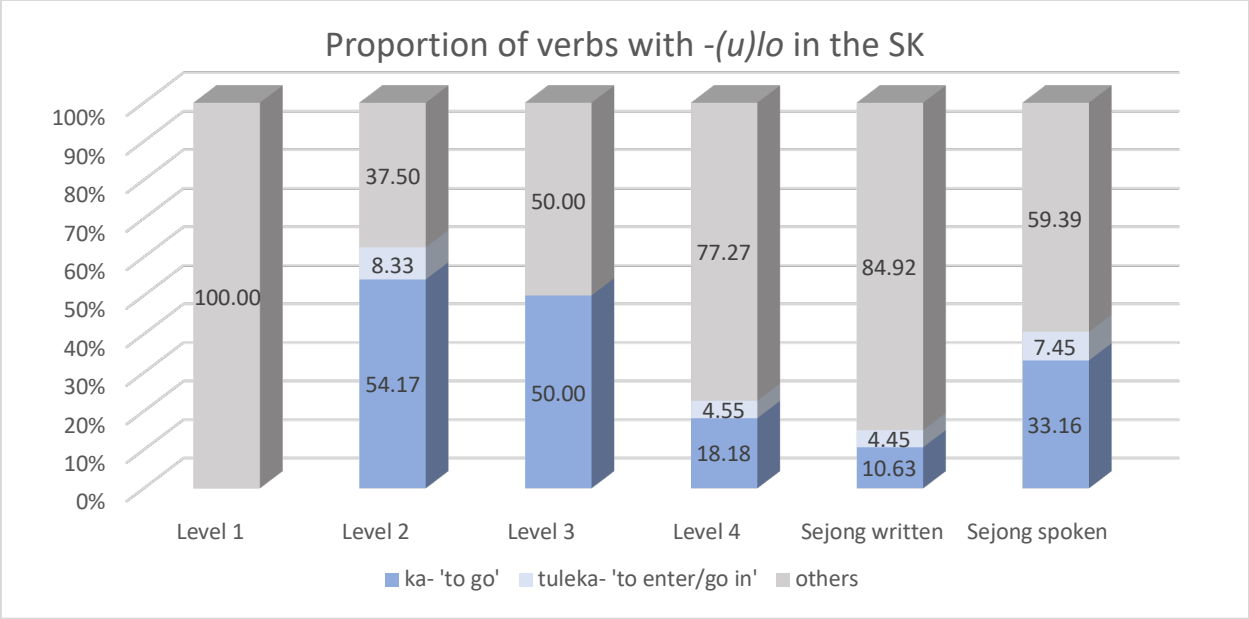


Figure 8. 3. Proportion change of verbs used with *-(u)lo* in the SK and IK textbook series

The three locational postpositions *-ey*, *-eyse*, and *-(u)lo* displayed their own combinatorial characteristics with certain verbs in the L1 corpora and textbooks. For *-ey*, it often combined with *iss-* ‘to be/exist’ in the L1 written/spoken corpora and *ka-* ‘to go’ in both textbook series. The use of *iss-* ‘to be/exist’ increased at the 4th level textbooks. SK textbooks showed overuse of

ka- ‘to go’ in comparison with the L1 spoken data while the IK series’ use of the verb was similar to that of the L1 written corpus. For *-eyse*, diverse types of verbs were utilized in both the L1 corpora and textbooks. While *ha-* ‘to do’ was the most frequently used in the L1 corpora and the SK textbooks across all levels except for level 1, several types of verb were used in the IK textbooks. Regarding the proportion of *ha-* ‘to do’, the SK textbook was appropriate in comparison to the Sejong spoken corpus while the IK textbook resembled proportion in the Sejong written corpus. In the case of *-(u)lo*, the most frequent verb was *ka-* ‘to go’ in both the L1 corpora and textbooks across all levels except for the level 1 volumes. While various types of verbs were introduced with *-(u)lo* at the first three levels of both textbook series, the use of *ka-* ‘to go’ and *tuleka-* ‘to enter/ go in’ resembled the L1 corpus to a certain degree: their percentages in the SK series were half of those in the Sejong spoken corpus while the IK series’ distributions were similar to those of the Sejong written corpus.

8.2. Keyness analysis

In order to examine which verbs demonstrated statistically significant occurrences in the textbooks in comparison to the L1 corpora, a keyness analysis was conducted with a confidence level of 99.9% ($p < 0.001$, critical value of *LL*: 10.83). To calculate the effect size (%DIFF), the L1 corpus was used as reference corpus (marked as Ref. in the tables in the current section) in relation to the textbook corpus. High % DIFF (>10,000) indicates no token was found in the reference (L1 Sejong) corpus, while -100 indicates zero usage in the study (textbook) corpus.

Table 8.1 shows a total of 16 CKIs with *-ey* in comparison to the L1 Sejong corpora and SK textbook. Compared to the L1 corpora, the SK textbooks utilized some verbs more frequently: *ka-* ‘to go’ was used eight times and two time more frequently than in the L1 written

and spoken corpora each. The verb *ppattuli-* ‘to drop’ was used more than eight and 14 times more frequently than in the L1 written and spoken corpora, which was presumably due to the topic of Korean traditional fables introduced in the SK textbooks. Other verbs such as *o-* ‘to come’, *nao-* ‘to come out’, *tochakha-* ‘to arrive’, and *ppaci-* ‘to fall’ were also used more in the SK textbook data than in the L1 corpora. On the other hand, some verbs were used less frequently compared to the L1 corpora. The frequencies of *anc-* ‘to sit’, *twu-* ‘to put/place’, and *iss-* ‘to be/exist’ were relatively low, whereas those verbs had considerable usage in the L1 corpora. *ilu-* ‘to arrive’, *olu-* ‘to climb/go up’, and *kelchi-* ‘to put on/hang’ did not appear in the SK textbooks at all.

Table 8. 1. CKI verbs used with *-ey* in the Sejong corpora and SK textbook series

Ref.	Verb type	<i>LL</i>	%DIFF
Sejong written	<i>ka-</i> ‘to go’	620.02	768.44
	<i>o-</i> ‘to come’	64.80	325.70
	<i>ilu-</i> ‘to arrive’	31.72	-100
	<i>nao-</i> ‘to come out’	16.89	183.40
	<i>tochakha-</i> ‘to arrive’	15.26	225.85
	<i>ppattuli-</i> ‘to drop’	15.05	735.99
	<i>anc-</i> ‘to sit’	14.56	-76.91
	<i>olu-</i> ‘to climb/ go up’	13.55	-100
	<i>twu-</i> ‘to put/place’	10.96	-88.67
	<i>kelchi-</i> ‘to put on/hang’	10.91	-100
Sejong spoken	<i>ka-</i> ‘to go’	178.13	162.73
	<i>iss-</i> ‘to be/exist’	84.63	-59.87
	<i>tochakha-</i> ‘to arrive’	22.25	516.50
	<i>ppattuli-</i> ‘to drop’	13.90	1296.20
	<i>ppaci-</i> ‘to fall’	12.83	378.70
	<i>pat-</i> ‘to receive’	11.13	1894.57

Note. The *p*-value of *LL* was < .001.

Some of the verbs that appeared in the IK textbooks also displayed some distinctive differences compared to the L1 data. Table 8.2 shows 16 CKI verbs used with *-ey* in the IK textbooks in comparison to the L1 corpora. Among those verbs, *talli-* ‘to be hung’ and *alli-* ‘to announce’ were used almost 10 times more frequently than in the L1 corpora. Some verbs such as *ssu-* ‘to write/use’, *o-* ‘to come’, *ha-* ‘to do’, and *sal-* ‘to live’ were more frequently used in the textbooks than in the L1 corpora. In the case of *ka-* ‘to go’, which was the most common verb occurring with *-ey* in the textbooks and the second most common verb in the Sejong corpora, it occurred five times and 1.7 times more frequently than in the L1 written and spoken corpora each. Verbs that were used significantly less frequently or did not appear at all compared to L1 data included *ilu-* ‘to arrive’, *iss-* ‘to be/exist’, *naka-* ‘to go out’, and *tuleka-* ‘to go in’.

Table 8. 2. CKI verbs used with *-ey* in the Sejong corpora and IK textbook series

Ref.	Verb type	<i>LL</i>	%DIFF
Sejong written	<i>ka-</i> ‘to go’	257.83	445.05
	<i>ssu-</i> ‘to write/use’	43.09	795.40
	<i>ilu-</i> ‘to arrive’	30.77	-100
	<i>o-</i> ‘to come’	30.74	210.11
	<i>talli-</i> ‘to be hung’	29.91	1158.90
	<i>ha-</i> ‘to do’	27.28	549.56
	<i>sal-</i> ‘to live’	15.26	153.52
	<i>alli-</i> ‘to announce’	12.02	901.40
Sejong spoken	<i>iss-</i> ‘to be/exist’	93.67	-63.63
	<i>ka-</i> ‘to go’	36.80	64.89
	<i>talli-</i> ‘to be hung’	22.83	1409.49
	<i>naka-</i> ‘to go out’	22.19	-100
	<i>alli-</i> ‘to announce’	16.35	> 10,000
	<i>ssu-</i> ‘to write/use’	11.78	189.70
	<i>tuleka-</i> ‘to go in’	10.96	-55.25
	<i>ha-</i> ‘to do’	10.84	225.01

Note. The *p*-value of *LL* was < .001.

Table 8.3 shows 18 CKI verbs that occurred with *-eyse* in the SK textbook series compared to the Sejong corpora. Among them, the verb that had the strongest statistical significance in comparison with the Sejong corpora was *sal-* ‘to live’ with *LL* values of 54.54 (written) and 36.11 (spoken). *sal-* was used 5~6 times more frequently in the SK textbook series. Other verbs such as *tut-* ‘to listen’, *yaykiha-* ‘to talk’, *ka-* ‘to go’, and *kyenghemha-* ‘to experience’ were used 22 to 44 times more frequently than in the L1 written corpus. Some other verbs that were used more frequently than in the L1 corpus included *ha-* ‘to do’, *manna-* ‘to meet’, *hwaphwuliha-* ‘to vent one’s anger’, *ilha-* ‘to work’, *kalatha-* ‘to transfer’, and *kongpwuha-* ‘to study’. *nao-* ‘to come out’ was used considerably less frequently than in the L1 spoken corpus while *yelli-* ‘to take place’ showed a zero occurrence in the textbook corpus in contrast to a considerable number of tokens in the L1 written corpus.

Table 8. 3. CKI verbs used with *-eyse* in the Sejong corpora and SK textbook series

Ref.	Verb type	<i>LL</i>	%DIFF
Sejong written	<i>sal-</i> ‘to live’	54.54	589.98
	<i>ha-</i> ‘to do’	39.70	216.95
	<i>hwaphwuliha-</i> ‘to vent one’s anger’	37.99	> 10,000
	<i>manna-</i> ‘to meet’	31.93	371.18
	<i>ilha-</i> ‘to work’	28.07	536.46
	<i>tut-</i> ‘to listen’	21.87	2148.83
	<i>yaykiha-</i> ‘to talk’	17.84	4397.66
	<i>ka-</i> ‘to go’	17.26	2837.58
	<i>kalatha-</i> ‘to transfer’	16.00	682.20
	<i>kyenghemha-</i> ‘to experience’	16.00	2898.44
	<i>kongpwuha-</i> ‘to study’	14.65	649.61
	<i>yelli-</i> ‘to take place’	11.69	-100
Sejong spoken	<i>sal-</i> ‘to live’	36.11	424.91
	<i>manna-</i> ‘to meet’	29.45	430.35
	<i>hwaphwuliha-</i> ‘to vent one’s anger’	23.50	> 10,000
	<i>kyenghemha-</i> ‘to experience’	15.64	775.67
	<i>nao-</i> ‘to come out’	13.94	-71.30
	<i>kongpwuha-</i> ‘to study’	12.64	743.75

Note. The *p*-value of *LL* was < .001.

Fifteen verbs, identified as CKIs for verb use with *-eyse* in the IK textbook series compared to the Sejong corpora, are listed in Table 8.4. Except for one verb *nao-* ‘to come out’, which occurred about 70% less frequently than in the L1 spoken corpus, all other verbs occurred more frequently in the IK textbooks than in the L1 corpora. To illustrate, *sa-* ‘to buy’ occurred 12 times and 5 times more than in the L1 written and spoken corpora, respectively. Verbs such as *memwul-* ‘to stay’, *yelli-* ‘to take place’, *kaychoytoy-* ‘to be held’, and *salaci-* ‘to disappear’ were extremely frequent in the IK textbook series. Other verbs such as *sal-* ‘to live’, *o-* ‘to come’, *mek-* ‘to eat’, *manna-* ‘to meet’, and *sa-* ‘to buy’ demonstrated frequencies that were three to five times higher than those in the L1 corpora. As was the case for the SK textbooks, verbs that

indicated specific activities such as eating, buying, or meeting (with someone) were frequent in the IK textbooks, which is not unusual considering language textbooks introduce and repeat vocabulary that have concrete meanings and are frequently used in contexts of daily activities.

Table 8. 4. CKI verbs used with *-eyse* in the Sejong corpora and IK textbook series

Ref.	Verb type	<i>LL</i>	%DIFF
Sejong written	<i>sa-</i> ‘to buy’	34.72	1161.93
	<i>sal-</i> ‘to live’	21.62	326.83
	<i>memwul-</i> ‘to stay’	19.39	> 10,000
	<i>o-</i> ‘to come’	19.37	351.49
	<i>kaychoytoy-</i> ‘to be held’	13.90	1834.96
	<i>mek-</i> ‘to eat’	13.53	394.74
	<i>manna-</i> ‘to meet’	11.95	222.49
Sejong spoken	<i>yelli-</i> ‘to take place’	32.40	> 10,000
	<i>sa-</i> ‘to buy’	16.74	455.60
	<i>kaychoytoy-</i> ‘to be held’	16.15	> 10,000
	<i>sal-</i> ‘to live’	13.42	234.56
	<i>manna-</i> ‘to meet’	12.93	262.99
	<i>memwul-</i> ‘to stay’	12.10	> 10,000
	<i>salaci-</i> ‘to disappear’	12.10	> 10,000
	<i>nao-</i> ‘to come out’	12.08	-69.13

Note. The *p*-value of *LL* was < .001.

The results for the keyness analysis of verbs used with *-(u)lo* in the SK textbook series compared to the L1 corpora are listed in Table 8.5. One of the characteristics of verbs classified as CKIs in comparison to L1 corpora and the SK textbooks can be seen through effect size values (%DIFF). As Table 8.5 shows, many of the verbs did not occur (*kalatha-* ‘to transfer’, *tayphiha-* ‘to evacuate’, *olmkyeci-* ‘to be moved’, and *ttena-* ‘to leave’) or had limited use (*tol-* ‘to turn’ and *itongha-* ‘to move’) in the L1 corpora compared to the SK textbooks. Three other verbs (*ka-* ‘to go’, *tolaka-* ‘to detour’, and *nao-* ‘to come out’) were used two to six times more

frequently in the textbooks than in the L1 corpora. Only one verb (*cwu-* ‘to give’) had zero tokens in the textbooks and a considerable number of tokens in the Sejong spoken corpus.

Table 8. 5. CKI verbs used with *-(u)lo* in the Sejong corpora and SK textbook series

Ref.	Verb type	<i>LL</i>	%DIFF
Sejong written	<i>ka-</i> ‘to go’	34.82	163.95
	<i>kalatha-</i> ‘to transfer’	21.97	> 10,000
	<i>tol-</i> ‘to turn’	20.67	1052.20
	<i>itongha-</i> ‘to move’	11.40	1340.24
	<i>tayphiha-</i> ‘to evacuate’	10.96	> 10,000
	<i>olmkyeci-</i> ‘to be moved’	10.96	> 10,000
Sejong spoken	<i>tol-</i> ‘to turn’	18.87	2768.29
	<i>cwu-</i> ‘to give’	15.62	-100
	<i>tolaka-</i> ‘to detour’	12.74	378.05
	<i>nao-</i> ‘to come out’	12.58	497.56
	<i>kalatha-</i> ‘to transfer’	12.26	> 10,000
	<i>ttena-</i> ‘to leave’	12.26	> 10,000

Note. The *p*-value of *LL* was < .001.

Table 8.6 contains comparative data for seven CKI verbs that occurred with *-(u)lo* in both the IK textbooks and the Sejong corpora. These verbs had relatively small *LL* values compared to the verbs that occurred with *-ey* or *-eyse*. Compared to the written corpus, *isaha-* ‘to move’, *molaney-* ‘to expel/remove’, *palcentoy-* ‘to be developed’, and *paysoktoy-* ‘to be assigned to’ were identified as CKIs while *ttena-* ‘to leave’, *isaha-* ‘to move’, and *cwu-* ‘to give’ were listed as CKIs for the spoken corpus. All of these verbs had little to no tokens in the L1 corpora, which would reasonably allow for the conclusion that higher token numbers for those verbs in the IK textbooks were skewed by topic-specificity that produced a more conducive environment for using *-(u)lo* with a particular variety of verbs.

Table 8. 6. CKI verbs used with *-(u)lo* in the Sejong corpora and IK textbook series

Ref.	Verb type	<i>LL</i>	%DIFF
Sejong written	<i>isaha-</i> ‘to move’	23.96	3348.18
	<i>molanay-</i> ‘to expel/remove’	11.64	> 10,000
	<i>palcentoy-</i> ‘to be developed’	11.64	> 10,000
	<i>paysoktoy-</i> ‘to be assigned to’	11.64	> 10,000
Sejong spoken	<i>ttena-</i> ‘to leave’	16.81	> 10,000
	<i>isaha-</i> ‘to move’	14.87	2475.18
	<i>cwu-</i> ‘to give’	13.30	-100

Note. The *p*-value of *LL* was < .001.

The keyness analysis was able to reveal characteristics for verbs co-occurring with each locational postposition. In general, such verbs were either over- or underused in the textbooks compared to the L1 data. To illustrate, for verbs with *-ey*, *ka-* ‘to go’ was overused and *iss-* ‘to be/exist’ was underused compared to the L1 corpora. Moreover, for verbs used with *-(u)lo*, all of the CKI verbs in the IK textbook and half of the CKIs in the SK textbook had an effect size of -100 or an extremely high value, which indicates that those verbs were used in neither the L1 corpora nor the textbooks. Altogether, findings suggest that while the SK and IK textbooks commonly used some dominant verbs for each postposition, they also utilized specific verbs depending on the types of topics and learning objectives being introduced.

8.3. Association strength

In relation to the association strength between each postposition and verb, the three most frequently used verbs with *-ey*, *-eyse*, and *-(u)lo* in the L1 corpora showed the strongest association with each postposition. Throughout all of the volumes for both textbook series, *ka-* ‘to go’, which was the most frequently used verb with *-ey*, showed high association strength. In the case of *-eyse*, *ha-* ‘to do’ and *manna-* ‘to meet’ showed the highest association strengths in

the SK and IK textbook series, respectively. Due to overlapping usages of *ka-* ‘to go’ with *-ey* and *-(u)lo* in the textbooks, the verb that showed the highest association strength with *-(u)lo* in both textbooks series was *tolaka-* ‘to detour’. However, textbook volumes at each level showed fluctuations in the frequencies and association strengths of verbs.

As there were limitations in the employment of diverse verb types in the textbooks (e.g., a number of frequent verbs in the L1 data were not used in the textbooks, and not all the verbs that occurred with more than one locative postposition in the L1 data were utilized in the textbooks), it was not feasible to compare tendencies of association strength values of other verbs consistently between L1 corpora and textbooks. Thus, only the two verbs that showed different behaviors with two locative postpositions across the L1 register were examined in the textbooks.

While most of the verbs showed similar tendencies across registers in the L1 corpora, *nathana-* ‘to appear’ and *sal-* ‘to live’ were exceptions (see Table 6.14). In the case of *nathana-* ‘to appear’, it had higher association strength ($\Delta P_{(\text{postposition} | \text{verb})}$) with *-ey* over the other postpositions in the written corpus, while the score was high for *-eyse* in the spoken corpus. For *sal-* ‘to live’, the association strength was high with *-eyse* in the written and *-ey* in the spoken corpora. Thus, these two verbs that showed unique behavior in the L1 corpora were examined in the textbooks and their results are summarized in Table 8.7.

Table 8. 7. $\Delta P_{(\text{postposition} | \text{verb})}$ value for *nathana-* and *sal-* in the SK and IK textbooks

	SK		IK	
	<i>-ey</i>	<i>-eyse</i>	<i>-ey</i>	<i>-eyse</i>
<i>nathana-</i> ‘to appear’	0.420	–	-0.101	0.233
<i>sal-</i> ‘to live’	-0.284	0.438	0.009	0.128

In the textbook data, association strengths for *nathana-* and *sal-* showed their own characteristics. For *nathana-* ‘to appear’, it was exclusively used with *-ey* for a total of six times in the SK volumes. It resulted a higher association strength with *-ey*, which was similar to the L1 written corpus. In the IK volumes, two instances were found: one of them was used with *-ey* and the other one with *-eyse*. As a result, the verb had higher association strength with *-eyse*, which was similar to the L1 spoken corpus. However, token numbers were relatively too small to make any meaningful generalizations for *nathana-* ‘to appear’, which require further investigations.

For *sal-* ‘to live’, it occurred with *-ey* a total of 12 times and 27 times with *-eyse* in the SK volumes, which resulted in a higher association strength with *-eyse*. In the IK volumes, the total number of tokens for *sal-* was 25 with *-ey* and 16 with *-eyse*. Although $\Delta P_{(\text{postposition} | \text{verb})}$ values for *sal-* were positive for both *-ey* and *-eyse* in the IK volumes, the latter was still higher. Thus, *sal-* ‘to live’ had a higher $\Delta P_{(\text{postposition} | \text{verb})}$ value with *-eyse* compared to *-ey* in both textbook series, which resembled the L1 written corpus.

8.4. Summary of findings

The proportions of verbs that were frequently used with each postposition in the textbooks were following usage trends found in the L1 corpora as proficiency level increased. Regarding the two verbs that were used most frequently with each postposition, the frequency distributions in the SK textbook series were generally similar to those in the L1 spoken corpus or in between the levels found within the written and spoken corpora. As for the IK textbook series, verb proportions were overall similar to those found in the L1 written corpus.

As for CKIs, a keyness analysis showed that each textbook employed a number of key verbs with *-ey*, *-eyse*, and *-(u)lo*, respectively. While there were some verbs that showed common

tendencies in both textbook series (e.g., overuse of *ka-* ‘to go’ with *-ey* and *sal-* ‘to live’ with *-eyse* in comparison with the L1 corpora), a number of verbs that were identified as CKI were unique to the SK and IK textbooks. For instance, *hwaphwuliha-* ‘to vent one’s anger’ was used quite frequently in the SK textbooks, while there were no occurrences in the L1 corpora and IK textbooks, showing a concentrated effort by the SK textbooks to introduce Korean proverbs. Overall, CKI verbs in the textbooks in comparison with the L1 corpora revealed characteristics that were special to each series.

In the both L1 and textbook corpora, highly frequently used verbs with each postposition exhibited the higher association strengths in general. However, in the case of *-eyse* (IK textbook series) and *-(u)lo* (both SK and IK textbook series), the most frequent verbs with each postposition showed lower ΔP values due to their overlapping use with other locative postpositions (e.g., *-ey*).

IX. CONCLUSION: PEDAGOGICAL IMPLICATIONS AND LIMITATIONS

Previous studies on the three postpositions of *-ey*, *-eyse*, and *-(u)lo* in L1 Korean contexts have mostly examined their uses as indications of multiple functions. It is these multiple form-function relations that have been the central topic of research for a number of studies. While the three postpositions are productively used for their own functions, they all share one function as an indicator of a location. General consensus on the differences between these three postpositions claims that *-ey* denotes existential/stative location, whereas *-eyse* indicates dynamic location. The use of *-(u)lo* is acknowledged as the implication of path/direction.

In general, there have been two approaches toward the investigation of these postpositions in studies of L2 Korean: one using production data from learners of Korean, and the other using textbooks for learners of Korean. In studies on learner production, some factors were considered such as error rate, postposition replacement, and individual functions less preferred in production. Textbook-based studies mostly considered which functions were (not) introduced at each level. As the interest in input for language acquisition increased, some L2 studies on postpositions attempted to connect textbooks and learner production, such as comparing what is listed in the textbooks to what learners acquire (or do not acquire) (e.g., Lee & Ko, 2013).

As the polysemy of such postpositions has been a main area of concern, many L1 and L2 studies have not directly examined the linguistic environments within which each function is realized, especially from a quantitative framework of analysis. Motivated by this gap that lies between the consideration of the linguistic environment and the use of postpositions, my dissertation investigated how each postposition was used with verbs. This project specifically

focused on the locative function because of its prototypicality in the meanings and uses of verbs, with the most salient differences observed between *-ey* (stative location) and *-eyse* (dynamic location). The L1 Sejong corpora and textbooks were examined with three research questions in mind that are re-stated below:

- i) What types of locative functions for each postposition are expressed in L1 corpora? Is there a relationship between locative postpositions and the environments within which they are used (e.g., co-occurrence with the verb)? If so, what are the characteristics of the relationship?
- ii) How is the locative function of each postposition used in textbooks for L2 learners? Is there a relationship between locative postpositions and the environments within which they are used (e.g., co-occurrence with the verb)? If so, what are the characteristics of the relationship?
- iii) What are the similarities and differences between L1 corpora and textbooks regarding the use of these three locational postpositions?

With L1 findings serving as a base, I analyzed two types of textbooks: the Sogang Korean series and the Integrated Korean series, with the expectation of finding pedagogical implications for teaching these three locational postpositions.

9.1. Summary of findings

9.1.1. Three locational postpositions in L1 corpora

Regarding the first research question, a total of 19,110 instances were found in the Sejong written corpora while 4,963 instances were found in the spoken corpora, respectively. That is, locative postpositions were more frequently used in the written data. Among the three postpositions, *-ey* was the most frequently used in both the written and spoken corpora, registering about 1.4 times more frequently than *-eyse* and four times more frequently than *-(u)lo*, respectively. Meanwhile, *-eyse* utilized a more diverse group of verbs compared to the other two postpositions. The ratio between the number of verb types and the total number of instances for each postposition shows that, on average, a new verb occurred for every four to five instances with *-eyse* in both the written and spoken registers. A new verb occurred in every ten to eleven instances with *-ey* and every four to seven instances for *-(u)lo*. It seems that *-eyse*'s flexible use with any dynamic verb resulted in a higher occurrence of new verbs.

In the main analysis, I calculated the token frequencies of co-occurring verbs, conducted keyness analysis, and computed association strength values between all three postpositions and their co-occurring verbs. One of the major characteristics in terms of verb usage was that the proportions of the most frequent verbs were noticeably high for each postposition type. In both the written and spoken corpora, the most frequent verb was *iss-* 'to be/ exist' with *-ey*, *ha-* 'to do' with *-eyse*, and *ka-* 'to go' with *-(u)lo*. Those verbs, possibly working as pathbreaking verbs, clearly demonstrate prototypical meanings for each postposition, that is, static location via *-ey*, dynamic location via *-eyse*, and path via *-(u)lo*. In addition, a small number of verbs occupied a large proportion of total usage. To illustrate, the ten most frequent verbs that occurred with *-ey* occupied 30.44% and 69.72% of the written and spoken corpora, respectively. For *-eyse*, the percentages for the ten most frequent verbs were 19.04% (written) and 36.79% (spoken) of total usage, while they ranged from 34.89% (written) to 67.35% (spoken) of total usage for *-(u)lo*.

Also, while their proportions were still large, the frequency of those verbs decreased at a nearly exponential rate, which resulted in a frequency distribution that followed Zipf's law. In other words, the most common item dominantly took a large proportion of total usage while less frequently used verbs had their token numbers decrease in a proportional manner.

Keyness analysis revealed that a certain number of verbs were more (or less) distinctively used in relation to the other register of comparison. Results showed that 16 verbs were identified as key verbs occurring with *-ey*, including more frequent usage with *iss-* 'to be/exist' and *ka-* 'to go' in the spoken corpus, and *ihu-* 'to arrive' and *se-* 'to stand' in the written corpus. 14 key verbs used with *-eyse* include *ha-* 'to do' (significant amount of usage in the spoken corpus) and *yelli-* 'to take place' (significant amount of usage in the written corpus). A total of seven key verbs with *-(u)lo* have been used more frequently in the spoken corpus than in the written one, including *ka-* 'to go', *cwu-* 'to give', and *ppaci-* 'to fall into'.

Association strength analysis also indicated that some verbs demonstrated greater association strength with a postposition due to exclusive or highly frequent use with each corresponding postposition. The three verbs *iss-* 'to be/ exist', *ha-* 'to do', and *ka-* 'to go' showed higher association strength with *-ey*, *-eyse*, and *-(u)lo*, respectively. It was also investigated whether a verb was a better predictor of a postposition or if a postposition was a better predictor of a postposition. The three most frequent verbs were better predictors of the postpositions with which they co-occurred for both the written and spoken registers, except for *-(u)lo* and *ka-* 'to go' in the spoken corpus, where neither one was found to be a better predictor of the other.

The association strength measurement also revealed that some verbs exhibited different behaviors in the written and spoken corpora: *sal-* 'to live' and *nathana-* 'to appear'. The verb

sal- ‘to live’, one of the most commonly discussed verbs in the studies that compared *-ey* and *-eyse*, was used with *-ey* more commonly in the spoken corpus than in the written corpus, while it was used with *-eyse* more frequently in the written corpus than in the spoken corpus. On the other hand, *nathana-* ‘to appear’ was frequently used with *-ey* in the written corpus and *-eyse* in the spoken corpus.

9.1.2. Three locational postpositions in L2 textbooks

The second research question aimed to examine verb use for all three locative postpositions in L2 learner textbooks. A total of sixteen volumes (eight volumes from both textbook series) were examined. Overall, the number of locative instances increased as the proficiency level increased in both series. *-ey* was the most commonly used postposition followed by *-eyse* and *-(u)lo*.

Throughout all of the volumes for both textbook series, dominant verbs emerged for each postposition: *ka-* ‘to go’ with *-ey*, *ha-* ‘to do’ (SK) and *sal-* ‘to live’ (IK) with *-eyse*, and *ka-* ‘to go’ with *-(u)lo*. Along with these verbs, some verbs that occurred frequently occupied a large proportion of total usage: the top five frequent verbs made up 65.74% (with *-ey* in SK), 48.16% (with *-ey* in IK), 38.46% (with *-eyse* in SK), 24.82% (with *-eyse* in IK), 51.83% (with *-(u)lo* in SK), and 39.42% (with *-(u)lo* in IK), respectively. The frequency distributions of verbs in the textbooks followed Zipf’s law, showing almost proportional decreases in token numbers as frequency ranks increased. While the dominant verbs and the distribution of verbs showed similar trends throughout all volumes for both textbook series, usage of individual verbs also differed between levels and/or textbook series.

The SK and IK textbook series showed different key items for *-ey* and *-eyse*. For *-ey*, *naka-* ‘to go out’ and *ppattuli-* ‘to drop’ were exclusively used in the SK textbooks while *talli-*

‘to be hung’ and *moi-* ‘to gather’ were exclusively used in the IK textbooks. For *-eyse, kitali-* ‘to wait’ did not appear in the IK series while *yelli-* ‘to take place’ and *sicaktoy-* ‘to get started’ did not appear in the SK series. Different topics or instructional goals may have caused particular verbs to be used. No verb was identified as a key verb with *-(u)lo*, presumably due to low token numbers.

Analysis of association strength between postpositions and verbs indicated that the verbs most frequently occurring with *-ey* (*ka-* ‘to go’) and *-eyse* (*ha-* ‘to do’) also showed the highest association strengths with each postposition in the SK textbooks. For *-(u)lo*, *tolaka-* ‘to detour’ demonstrated the highest association strength with this postposition, while the most dominant verb occurring with *-(u)lo* was *ka-* ‘to go’. In the IK textbooks, *ka-* ‘to go’ manifested the highest association strength. For *-eyse* and *-(u)lo*, *manna-* ‘to meet’ and *tolaka-* ‘to detour’ showed high association strengths, respectively. In general, frequent verbs in the textbooks exhibited a higher degree of association strength, which indicates that each textbook deployed a verb exclusively in combination with each postposition. However, verbs that showed high association values with each postposition were different at each textbook level depending on the topics being introduced.

9.1.3. Comparison of L1 corpora and L2 textbooks

As for the third research question, the verb use in the L2 textbooks was generally similar to those found in the L1 corpora. Of the three postpositions, *-ey* was the most commonly used locative postposition, with *-eyse* the second and *-(u)lo* the least used. Each corpus showed that some verbs occurred dominantly with each postposition. Also, the frequency distributions of verbs in the L1 corpora and L2 textbooks (all volumes) followed Zipf’s law.

While both the L1 and L2 corpora showed that a number of verbs occupied a considerable amount of total usage, some discrepancies were found to exist between the L1 corpora and the L2 textbooks regarding which verb was used the most. Most notably, the most frequent verbs with *-ey*, *-eyse*, and *-(u)lo* were *iss-* ‘to be’, *ha-* ‘to do’, and *ka-* ‘to go’, respectively, in both the written and spoken Sejong corpora. In the textbooks, *ka-* ‘to go’ was the most frequent verb for both *-ey* and *-(u)lo*. For *-eyse*, the most frequent verb was *ha-* ‘to do’ in the SK textbooks and *sal-* ‘to live’ in the IK textbooks.

As the proficiency level changed, however, the proportions of frequent verbs with each postposition generally followed the trend of L1 corpora. Regarding verb use with *-ey*, the SK textbooks showed a gradual decrease in the use of *ka-* ‘to go’ and an increase in the use of *iss-* ‘to be/exist’, ultimately resulting in a nearly identical number of tokens for both at the 4th level. The proportions of the two verbs in the SK textbooks were higher than those in the L1 written corpus, while *iss-* ‘to be/ exist’ registered a lower proportion in the textbook data in comparison to the L1 spoken corpora. The proportions of *ka-* ‘to go’ and *iss-* ‘to be/exist’ at the 4th level IK textbook volumes were nearly identical to those in the L1 written corpus. For verbs used with *-eyse*, the 4th level SK textbooks used *ha-* ‘to do’ and *nao-* ‘to come out’ more often compared to the L1 written corpus while *nao-* ‘to come out’ appeared less frequently than in the L1 spoken corpus. The IK Level 4 volumes had a similar distribution for the two verbs with the L1 written corpus. As for verbs for *-(u)lo*, *ka-* ‘to go’ was used more frequently in the SK textbooks than in the L1 written corpus, although both *ka-* ‘to go’ and *tuleka-* ‘to enter/go in’ were used less frequently in the 4th level volumes of the SK textbook series. Meanwhile at the 4th level of the IK textbook series, the frequency levels for both verbs were similar to those of the L1 written corpus at the 4th level and lower than those of the L1 spoken corpus. In general, the percentages of verbs

used frequently with each postposition for all volumes of the SK textbook series were higher than those of the L1 written corpus, but similar or somewhat lower than those of the L1 spoken corpus. Throughout all of the volumes in the IK textbook series, the percentages of verbs used with each postposition resembled the L1 written corpus but were smaller in comparison to the L1 spoken corpus.

Some verbs showed significant differences in their frequencies among the two L1 corpora and the two types of L2 textbooks. While some CKIs showed that certain verbs were more (or less) significantly used in comparison with other corpora (e.g., more use in the L1 spoken corpus than in the L1 written corpus or more use in the textbooks compared to the L1 spoken corpus), there were also some extreme cases (i.e., zero occurrences in either the L1 or textbook corpora). Among these cases, the verbs that did not appear in the textbooks included *ilu-* ‘to arrive’, *olu-* ‘to climb/go up’, and *naka-* ‘to go out’ for *-ey*, *yelli-* ‘to take place’ for *-eyse*, and *cwu-* ‘to give’ for *-(u)lo*. The differences seem to be due to pedagogical considerations when it came to creating the language textbooks, such as pronunciation and definiteness of the meaning of a certain word along with the inclusion/exclusion of topics.

Finally, the analysis of association strength between verbs and locative postpositions in the both L1 and textbook corpora showed that dominantly used verbs with each locative postposition showed highest ΔP values with their corresponding postpositions in general. However, ΔP value of *ka-* ‘to go’, which was the most frequent verb with *-(u)lo* in both of the textbook types, was low due to the overlapping use of the verb with other locative postpositions. Two verbs that exhibited different behaviors in the L1 corpora were *nathana-* ‘to appear’ and *sal-* ‘to live’. *nathana-* ‘to appear’ preferred *-ey* in the written register and *-eyse* in the spoken register. Meanwhile, *sal-* ‘to live’ had a stronger association with *-eyse* in the written corpora and

-*ey* in the spoken corpora. In the textbooks, *nathana-* ‘to appear’ was preferably used with *-ey* in the SK series and with *-eyse* in the IK series, while token numbers were too small to make any meaningful generalizations. For *sal-* ‘to live’, it had strong association power with *-eyse* in both textbook series.

9.2. Implications

The findings of this dissertation suggest two pedagogical implications: the utilization of corpus data for postposition instruction and a more systematic form of instruction that is based on frequency patterns found in corpus data.

9.2.1. Use of corpora in language education

The first suggestion for teaching these three postpositions involves the utilization of corpus data in language instruction. The usefulness of corpus/data-driven learning in second/foreign language acquisition has been suggested by several studies (Shin, Velázquez, Swatek, Staples, & Partridge, 2018; Cotos, 2014) and indicated through diverse methods (Lenko-Szymanska & Boulton, 2015).

The necessity of utilizing frequency information from L1 corpora for postposition instruction has been raised by several studies (e.g., Lee, 2005; Ro, 2008) as well. The main idea of those studies was to teach the most frequently occurring language items in the L1 environment at the beginning level, after which less frequently used language items would be introduced at higher target proficiency levels. This creates opportunities for more effective education by ensuring that learners can steadily expand their usage of postpositions. Both the SK and IK textbooks introduced a limited number of verbs at early stages of learning and introduced other

usages at the upper levels for all three postpositions. However, the verbs used in the textbooks did not reflect L1 usage thoroughly. To illustrate, verbs that were frequently used with *-(u)lo* were not introduced at the beginning level of the SK textbook while the dominant verbs used with *-eyse* in the L1 were not frequently used throughout the different levels of the IK textbook series. This shows that learners of Korean whose learning materials consisted of either the SK or IK textbook series may not have been fully exposed to how the target language is composed of in terms of these three locative postpositions.

Frequency information for each verb obtained from the present study can be useful in designing future language textbooks. While introducing and displaying input that demonstrates the polysemous nature of all three postpositions in textbooks for L2 learners, suggesting a distinctive one-to-one form-function mapping may be feasible for language learners of Korean to have a more solid orientation to their usage. So far, the distributional properties of co-occurring verbs for each locative postposition have not been fully captured in L1 studies. However, frequency lists of the verbs from this analysis will ease the process of composing such arrangements for verbs in future textbooks. The most frequently used verbs can be considered to be introduced in a more focused manner for the beginning stages of instruction followed by further extension of their usage at the upper levels.

Corpora can also be used more practically as teaching materials. Some beginning level textbooks address the fact that the dynamicity of a verb often determines differences between *-ey* and *-eyse* (Jeong, 2011; Kim, 2011). As mentioned earlier, however, the analysis of L1 corpora showed that some verbs co-occurred with both *-ey* and *-eyse* (e.g., *na-* ‘to appear/sprout’ in *yeki-ey/eyse na-ta* ‘(something) appears (from) here’). Among them, verbs such as *sal-* ‘to live’ were introduced, but most of the other verbs were not introduced, even at the upper level textbook

volumes. Consequently, as language users learn the difference between these two postpositions in terms of degrees of dynamicity for different verb types, encountering those verbs that do not prioritize a clear presentation of stative versus dynamic locations may confuse learners. The results from the corpus analysis will further assist the development of instructional materials as a resolution to this issue: overlapping verbs that occur with more than two postpositions can be summarized and discussed in the class.

The difference between written and spoken registers can be a source for instruction as well. Results from the analysis of L1 corpora show representative characteristics of each text. For example, language use in written Korean and spoken Korean can be used as instructional guides to improve target skills such as writing academic essays or composing spoken conversations, respectively. Different ΔP values for each verb in each register can also be used as a way for suggesting what types of verbs are distinctively used in those registers. As techniques for effectively dealing with large size corpora continue to be developed, textbook developers and language instructors will be provided with more opportunities to establish systematic approaches to developing language textbooks/materials that utilize L1 corpora.

9.2.2. Utilizing linguistic patterns in language instruction

Verb distributions discussed in the current dissertation call for discussion on the feasibility of implementing pattern-driven postposition instruction. Indeed, the need for instructing postpositions in terms of patterns has been suggested by several studies (Ko, 2007; Woo, 2016). Those studies specifically emphasize the role of verbs in understanding the functions of each postposition.

The results from L1 written and spoken corpora in the current study showed that some verbs dominantly co-occurred with each postposition: *iss-* ‘to be/exist’ for *-ey*, *ha-* ‘to do’ for *-eyse*, and *ka-* ‘to go’ for *-(u)lo*. These verbs not only had the highest association strengths with each postposition, but they also provided evidence to support the general consensus regarding the locational functions of these three postpositions. That is, the meanings of those verbs represented the prototypical functions of each postposition: - stative/existential location for *-ey*, dynamic location for *-eyse*, and path/directional for *-(u)lo*. Thus, instruction for each postposition that incorporates its respectively dominant verb will enhance understanding of their unique locative functions of each postposition. Also, those verbs presumably could work as pathbreaking ones for the acquisition of other dominant usage for each postposition. Such pathbreaking verbs aid in the reinforcement of prototypical usages for each postposition.

Besides dominant verbs, the token numbers of other verbs decreased at a nearly exponential rate. The analysis of L1 corpora indicated that the distribution of verbs co-occurring with each postposition followed a Zipfian distribution. That is, frequency graphs (Figure 6.1, 6.2, and 6.3) showed downward curves following Zipf’s law. As a result, a number of the most frequent verbs occupied large portions of total usage for all three postpositions. This is an indication that knowledge about each postposition can be better acquired through strategic usage of frequently occurring verbs, which will be advantageous towards assisting language learners’ greater understanding of a larger percentage of L1 contexts.

Despite these salient features of verb usage in the L1 corpora, textbooks had some limitations in reflecting them. While verb distributions for each postposition in the textbooks also followed Zipf’s law, the frequency order of verb types was different from that of the L1 data. For *-ey*, the dominant verb was *ka-* ‘to go’, which indicates an emphasis on the locative-destination

function of *-ey* in the textbooks. While the proportion of *iss-* ‘to be/exist’ (the most frequently occurring verb with *-ey* in the L1 corpora) increased as the proficiency level changed and showed a higher rate of use in the 4th level volumes for both textbook series, the locative-existential aspect of *iss-* ‘to be/exist’ was generally underrepresented in both textbook series.

In the case of *-eyse*, verbs used frequently in the L1 corpora were not fully reflected in the textbooks: the most frequent verb was the same for both the L1 corpora and the SK textbook series, but not the IK series. The second and third most frequent verbs in the L1 written/spoken corpora also had limited usage in both textbook series. In addition, verbs that co-occurred with *-eyse* in the textbooks had nearly identical token numbers, with no dominant usage of a certain verb. This feature of *-eyse* revisits the debate on how to compose and provide input. There has been a debate as to the effectiveness of skewed or balanced input in second/foreign language acquisition (McDonough & Nekrasova-Becker, 2014; Nakamura, 2012). A number of studies have argued that early presentation of a single highly frequent item, which also shows prototypical meaning, facilitates learners’ understanding of the exemplar (Casenhiser & Goldberg, 2005; Ellis & Ferreira-Junior, 2009b; Goldberg, 2019). For *-eyse* then, pattern-based instruction of *-eyse ha-* ‘to do (at a place)’ at an early stage of learning may provide a stable orientation toward the item. Also, other frequently occurring verbs in the L1 can be utilized as part of patterns to increase the producibility of *-eyse* in upper levels.

Previous studies regarding *-(u)lo* in textbooks showed that while the locative functions of *-ey* and *-eyse* were emphasized at the 1st level of textbooks (e.g., Jeong, 2011), various other functions for *-(u)lo* such as *by means of transportation, qualification, and cause* were often introduced before its locative function (Lee, 2005). Moreover, the results in Chapter 7 also indicate that textbooks avoided confusion between *-ey* and *-(u)lo* by using verbs that were not

commonly used with *-ey*. Explicit exposure to the target exemplar along with proper intervention might be useful for language instruction, especially when there is another linguistic feature that seemingly looks similar. Introducing *-(u)lo ka-* ‘to go (to somewhere)’ as a meaningful unit rather than explaining the locative functions via other verbs can encourage a higher degree of proper usage for *-(u)lo*’s locative function by language learners. As the difference between *-ey* and *-(u)lo* is not explicitly recognizable when the same verb is utilized, it will be useful to also utilize image schema that cognitively present the use of *-(u)lo* (see Chapter 3). Indeed, several studies argue that the use of images will facilitate learners’ understanding of *-(u)lo* (e.g., Lee, 2002; Jeong, 2011). A comprehensive pedagogy will incorporate pattern-based instruction of *-(u)lo* with supplemental explanation that utilizes image schema in order to improve learners’ recognition of the differences between *-ey* and *-(u)lo*.

Overall, exposure to Korean language is limited to formal classroom settings in the case of beginners. Thus, presentation of the target features should be designed in a way that clearly shows the prototypical meaning of each postposition while enhancing learning efficiency. The introduction of postposition-verb patterns such as *-ey iss-* ‘to be/exist (in a place)’, *-eyse ha-* ‘to do (something in a place)’ and *-(u)lo ka-* ‘to go (to a place)’ will not only provide opportunities for being exposed to co-occurring elements that represent prototypical usage of *-ey*, *-eyse*, and *-(u)lo*, but also ensure usage combinations that resemble those in L1 corpora.

9.3. Limitations and future studies

The present dissertation investigated the three locative functions of *-ey*, *-eyse*, and *-(u)lo* with a focus on verb use in the L1 Sejong written and spoken corpora as well as two types of textbook series for language learners of Korean. It examined what types of verbs combined with each

postposition frequently, how exclusively each verb was used in each corpus, and how language textbooks reflected the combinatorial characteristics of use for these three postpositions in such L1 corpora. While the current study explored ways to quantitatively analyze corpora for L1 and L2 Korean studies, there are still limitations and remaining tasks for future studies.

First, due to the lack of a sizable L1 corpus that could completely represent each genre, the current dissertation analyzed the Sejong written and spoken corpora as a whole. By analyzing each corpus as a whole, this dissertation presents a general picture of postposition use in the Korean language. However, it did not suggest the specific features of each genre. Both written and spoken data in the L1 Sejong corpora are composed of diverse genres for each register. Fiction writing, articles from magazines, and newspapers have all been included in the written Sejong corpus. For the spoken corpus, transcripts from conference presentations, conversations between friends, sermons at church gatherings, and even sports broadcasts were included. The properties of each text that composed each register may be different. For instance, spoken conversation between friends and presentations in a formal context may show distinct linguistic features. Building up L1 corpora that reveal diverse aspects of the Korean language will benefit future studies.

The second limitation is related to the extraction of phrases that contain the three target postpositions. In the current study, complex sentences were divided into two or more phrases based on the existence of verbs in sentence-medial position. However, there exist some cases in which a complex sentence was not properly divided into different phrases. For example, in a sentence like *Tom-i ttwi-ess-ta hakkyo-lo, kuliko ca-ss-ta* ‘Tom ran to the school, and (there he) slept/ took a nap’, the first phrase’s components have been scrambled: the canonical order is *Tom-i hakkyo-lo ttwi-ess-ta* ‘Tom ran to the school’. The division of sentences depends on the position

of verbs, which results in two phrases ‘Tom ran’ and ‘to the school and took a nap’.

Subsequently, *ca-* ‘to sleep/take a nap’ instead of *ttwi-* ‘to run’ will be detected as a co-occurring verb with *-(u)lo*. Such tokens that contained mismatches of adverbial phrases and verbs were removed during a manual checking process for the locative function, but there still may be some instances in which some data from the L1 corpora was inadvertently excluded from the analysis.

Third, the current dissertation focused on the locative function, which is a prototypical and overlapping function across the three postpositions. However, as polysemous postpositions, *-ey*, *-eyse*, and *-(u)lo* are productively used for other functions, including *time*, *source*, *instrument*, etc. Due to the amount of manual work necessary to categorize each function, other functions were not examined in the present dissertation. Future work that explores the linguistic environments of other functions will bring thorough understanding about the relations between the functions of each postposition. In addition, the development of more corpora that mark each function would ease the process and expand the scope of research on Korean postpositions.

Fourth, only general purpose L2 textbooks were selected for the analysis, whose results provided a glimpse into the overall use of the three postpositions in textbooks. There are other types of textbooks that have been designed for specific purposes such as those published for immigrants to Korea and fifth level textbooks in the SK series that have been tailored for the development of reading and speaking skills. Future investigation on those textbooks will show how language input is shaped for learners with specific interests or different registers such as spoken and written contexts.

Finally, the examination on how input from textbooks affects learners’ production would be an area of further interest that lies beyond the scope of the current dissertation. In consideration of the few studies that have quantitatively analyzed L1 corpora and learner

textbooks, this dissertation has focused on how postpositions were used in consideration of verbs and to what extent postposition use in the textbooks reflected L1 usage. I expect that follow-up studies will unveil how characteristics of L1 corpora and textbooks influence learners' understanding of postpositions.

Under the major tenet of usage-based approaches, the present dissertation investigated how locative functions of the three postpositions *-ey*, *-eyse*, and *-(u)lo* in Korean were realized with a focus on verb use using L1 corpora as references along with language textbooks, which are considered to be the major input source affecting learners' understanding and linguistic knowledge. With the support of qualitative research results from previous studies, the present dissertation initiated a more quantitative approach to examine how input can be shaped and potentially influence learners' understanding of Korean as a second/foreign language while drawing a more thorough picture of postposition use. Despite these limitations, I believe that the results and implications of the present dissertation will open the window to further considerations of input effect on learners' development in their Korean language proficiency.

REFERENCES

- Abbot-Smith, K., & Tomasello, M. (2006). Exemplar-learning and schematization in a usage-based account of syntactic acquisition. *The Linguistic Review*, 23(3), 275–290.
<https://doi.org/10.1515/TLR.2006.011>
- Alsaif, A. & Milton, J. (2012). Vocabulary input from school textbooks as a potential contributor to the small uptake gained by English as a foreign language learners in Saudi Arabia. *The Language Learning Journal*, 40(1), 21-33.
<https://doi.org/10.1080/09571736.2012.658221>
- An, M. C. (1982). Chekyek ‘ey’uy uymi [The meaning of locative ‘ey’]. *Kwanakemwunyenkwu*, 7, 245-268.
- Anderson, J. R. (1982). Acquisition of cognitive skill. *Psychological Review*, 89, 369–406.
<https://doi.org/10.1037/0033-295X.89.4.369>
- Behrens, H. (2009). Usage-based and emergentist approaches to language acquisition. *Linguistics*, 47(2), 383-411. <https://doi.org/10.1515/LING.2009.014>
- Beitel, D. A. (1995). *The embodied image schematic approach to the polysemy of spatial prepositions*. Unpublished doctoral dissertation. University of California, Santa Cruz.
- Biber, D., Connor, U., Upton, A., Anthony, M., & Gladkov.(2007). Rhetorical appeals in fundraising. In D. Biber, U. Connor & A. Upton, *Discourse on the move: Using corpus analysis to describe discourse structure* (pp. 121-151). Amsterdam: John Benjamins.
- Biederman, I., Glass, A., & Stacy, E. (1973). Searching for objects in real-world scenes. *Journal of Experimental Psychology*, 97(1), 22-27. <https://doi.org/10.1037/h0033776>

- Blumenthal-Drame, A. (2012). *Entrenchment in usage-based theories: What corpus data do and do not reveal about the mind*. Berlin: De Gruyter.
- Brugman, C. (1988). *The story of “over”: Polysemy, semantics and the structure of the lexicon*. New York: Garland.
- Bybee, J. L. (2006). From usage to grammar: The mind's response to repetition. *Language*, 82(4), 711–733. <https://doi.org/10.1353/lan.2006.0186>
- Bybee, J. L. (2008). Usage-based grammar and second language acquisition. In P. Robinson & N. C. Ellis (Eds.), *Handbook of cognitive linguistics and second language acquisition* (pp. 216–236). New York: Routledge.
- Bybee, J. L. (2013). Usage-based theory and exemplar representations of constructions. In G. Trousdale & T. Hoffmann (Eds.), *Oxford handbook of construction grammar* (pp. 49–69). Oxford: Oxford University Press.
- Bybee, J. L., & Hopper, P. (2001). *Frequency and the emergence of linguistic structure*. Amsterdam: Benjamins.
- Casenhiser, D., & Goldberg, A. E. (2005). Fast mapping between a phrasal form and meaning. *Developmental Science*, 8(6), 500-508. <https://doi.org/10.1111/j.1467-7687.2005.00441.x>
- Chae, H-R., & Lim, E. (2013). An analysis of locative expressions [NP-ey] and [NP-eyse] in Korean. *Korean Journal of Linguistics*, 38(4), 997-1026. <https://doi.org/10.18855/lisoko.2013.38.4.010>
- Cho, C-h. (2018). A study for the relationship between the adverbial case marker ‘-e’ and the predicate on the modern Korean. *Yongponginnwnonchong*, 52, 225-252.
- Cho, M. H. (1982). *Language acquisition strategy of Korean children*. Seoul, Korea: Seoul National University Press.

- Choi, U-H. (2005). A study on classification and function of particles in Korean. *The Journal of Linguistics Science*, 33, 31-348.
- Choo, M., & Kwak, H-Y. (2008). *Using Korean*. Cambridge: Cambridge University Press.
- Coleman, L., & Kay, P. (1981). Prototype semantics: The English word “lie”, *Language*, 57(1), 19-42. <https://doi.org/10.2307/414285>
- Cotos, E. (2014). Enhancing writing pedagogy with learner corpus data. *ReCALL*, 26(2), 202-224. <https://doi.org/10.1017/S0958344014000019>
- Croft, W., & Cruse, A. (2004). *Cognitive linguistics*. Cambridge: Cambridge University Press.
- Dąbrowska, E. & Divjak, D. (2015). *Handbook of cognitive linguistics (Eds)*. De Gruyter Mouton.
- Dąbrowska, E., & Szczerbiński, M. (2006). Polish children’s productivity with case marking: The role of regularity, type frequency, and phonological diversity. *Journal of Child Language*, 33(3), 559-597. <https://doi.org/10.1017/S0305000906007471>
- Davis, M., & Face, T. (2006). Vocabulary coverage in Spanish textbooks: How representative is it? In N. Sagarra & A. J. Toribio (Eds.), *Selected Proceedings of the 9th Hispanic Linguistics Symposium* (pp. 132-143). Somerville, MA: Cascadilla Proceedings Project.
- Diessel, H. (2007). Frequency effects in language acquisition, language use, and diachronic change. *New Ideas in Psychology*, 25(2), 108–127. <https://doi.org/10.1016/j.newideapsych.2007.02.002>
- Divjak, D., & Caldwell-Harris, C. (2015). Frequency and entrenchment. In E. Dąbrowska & D. Divjak (Eds.), *Handbook of cognitive linguistics (Vol. 39)* (pp. 53-75). De Gruyter Mouton.

- Elio, R., & Anderson, J. (1984). The effects of information order and learning mode on schema abstraction. *Memory and Cognition*, 12(1), 20-30. <https://doi.org/10.3758/BF03196994>
- Ellis, N. C. (2002). Frequency effects in language processing: A review with implications for theories of implicit and explicit language acquisition. *Studies in Second Language Acquisition*, 24(2), 143–188. <https://doi.org/10.1017/S0272263102002024>
- Ellis, N. C. (2006a). Language acquisition as rational contingency learning. *Applied Linguistics*, 27(1), 1-24. <https://doi.org/10.1093/applin/ami038>
- Ellis, N. C. (2006b). Selective attention and transfer phenomena in L2 acquisition: Contingency, cue competition, salience, interference, overshadowing, blocking, and perceptual learning. *Applied Linguistics*, 27(2), 164-194. <https://doi.org/10.1093/applin/aml015>
- Ellis, N. C. (2012). Frequency-based accounts of second language acquisition. In S. Gass & A. Mackey (Eds.) *The Routledge handbook of second language acquisition* (pp. 193–210). Routledge.
- Ellis, N. C., & Ferreira-Junior, F. (2009a). Construction learning as a function of frequency, frequency distribution, and function. *The Modern Language Journal*, 93(3), 370–385. <https://doi.org/10.1111/j.1540-4781.2009.00896.x>
- Ellis, N. C., & Ferreira-Junior, F. (2009b). Constructions and their acquisition: Islands and the distinctiveness of their occupancy. *Annual Review of Cognitive Linguistics*, 7(1), 187–220. <https://doi.org/10.1075/arcl.7.08ell>
- Ellis, N. C., Römer, U., & O'Donnell, M. B. (2016). Constructions and usage-based language approaches to language acquisition. *Language Learning*, 66(S1), 23-44. https://doi.org/10.1111/lang.1_12177

- Fox, G. (1987). The case for examples. In J. McH. Sinclair (Eds.) *Looking up* (pp. 137-149), London: Harper Collins.
- Fraysse-Kim, S. H. (2010). Keywords in Korean national consciousness: A corpus-based analysis of school textbooks. In Bondi, M. & Scott, M. (eds.) *Keyness in text: Corpus linguistic investigations* (pp. 219-233). Amsterdam: John Benjamins.
- Frisch, S. F., Large, N. R., Zawaydeh, B., & Pisoni, D. B. (2001). Emergent phonotactic generalizations in English and Arabic. In J. Bybee & P. Hopper (Eds.), *Proceedings of the Symposium on Frequency Effects and Emergent Grammar* (pp. 159-180). Amsterdam: Benjamins.
- Gabrielatos, C. (2018). Keyness analysis: nature, metrics and techniques. In C. Taylor & A. Marchi (Eds.), *Corpus approaches to discourse: A critical review* (pp. 225-258). Oxford: Routledge.
- Glisan, E. W., & Drescher, V. (1993). Textbook grammar: Does it reflect native speaker speech? *The Modern Language Journal*, 77(1), 23-33. <https://doi.org/10.1111/j.1540-4781.1993.tb01941.x>
- Goldberg, A. E. (1995). *Constructions: A construction grammar approach to argument structure*. Chicago: University of Chicago Press.
- Goldberg, A. E. (2006). *Constructions at work: The nature of generalization in language*. Oxford: Oxford University Press.
- Goldberg, A. E. (2013b). Constructionist approaches. In G. Trousdale & T. Hoffmann (Eds.), *Oxford handbook of Construction Grammar* (pp. 15–31). Oxford: Oxford University Press.
- Goldberg, A. E. (2019). *Explain me this*. Princeton University Press.

- Goldberg, A. E., Casenhiser, D. M., & Sethuraman, N. (2004). Learning argument structure generalizations. *Cognitive Linguistics*, 15(3), 289–316.
<http://search.proquest.com/docview/37955224/>
- Goldschneider, J. M., & DeKeyser, R. M. (2001). Explaining the “natural order of L2 morpheme acquisition” in English: A meta-analysis of multiple determinants. *Language Learning*, 51(1), 1-50. <https://doi.org/10.1111/1467-9922.00147>
- Gries, S., Hampe, B., & Schönefeld, D. (2005). Converging evidence: Bringing together experimental and corpus data on the association of verbs and constructions. *Cognitive Linguistics*, 16(4), 635-676. <https://doi.org/10.1515/cogl.2005.16.4.635>
- Hampton, J. (1987). Inheritance of attributes in natural concept conjunctions. *Memory and Cognition*, 15(1), 55-71. <https://doi.org/10.3758/BF03197712>
- Han, S-M. (2014). An analysis of errors on the usage of postpositions in the discussions of advanced Korean language learners. *Bilingual research*, 57, 223-255.
<https://doi.org/10.17296/korbil.2014..57.223>
- Hong, C. S. (1983). Idongdongsawa haynglouy poe [Movement verbs and complements of path]. *Oykwukeloseuy hankwukekyoyuk*, 8, 151-167.
- Howell, D. C. (2010). *Statistical methods for psychology* (7th ed.). Belmont, CA: Cengage Wadsworth.
- Hwang, J. (2002). *Acquisition hierarchy of Korean as a foreign language*. Unpublished doctoral dissertation. University of Hawai'i.
- Jeong, B-c. (2010). An integrated study of the postposition ‘-ro’ based on simulation model. *Korea Semantics*, 32, 215-243.

- Jeong, S-j. (2011). A study on the description methods of the adverbial case postpositions for Korean education based on cognitive linguistics. *The Korean Language and Literature*, 112, 79-110. <http://www.dbpia.co.kr/Article/NODE01650599>
- Jiang, N. A. N., & Nekrasova, T. M. (2007). The processing of formulaic sequences by second language speakers. *The Modern Language Journal*, 91(3), 433-445. <https://doi.org/10.1111/j.1540-4781.2007.00589.x>
- Jurafsky, D. (1996). A probabilistic model of lexical and syntactic access and disambiguation. *Cognitive Science*, 20(2), 137-194. https://doi.org/10.1207/s15516709cog2002_1
- Kabata, K., & Lee, J-H. (2004). Factors underlying spatial particle distributions in Japanese and Korean. In M. Archard & S. Kemmer (Eds.), *Language, culture and mind* (pp. 167-180). Stanford, California: CSLI Publications.
- Kang, B. M., & Kim, H. G. (2009). *Hankwuke sayong pindo [Usage Frequency of Korean language]*. Seoul, Korea: Hankwuk Mwonhwsa.
- Kang, Y. (2012). *Cognitive linguistics approach to semantics of spatial relations in Korean*. Unpublished doctoral dissertation. Georgetown University.
- Kang, Y. (2015). *A study on education of Korean postposition 'e', 'eseo', 'ro' for Chinese Korean learners*. Unpublished master dissertation. Seoul National University.
- Kemmer, S., & Barlow, M. (2000). Introduction: A usage-based conception of language. In M. Barlow & S. Kemmer (Eds.), *Usage-based models of language* (pp. vii–xxviii). Stanford: CSLI Publications.
- Kilgarriff, A. (2001). Comparing corpora. *International Journal of Corpus Linguistics*, 6(1), 97-133. <https://doi.org/10.1075/ijcl.6.1.05kil>

- Kilgarriff, A. (2005). Language is never ever ever random. *Corpus Linguistics and Linguistic Theory*, 1(2), 263-276. <http://search.proquest.com/docview/85643706/>
- Kim, J., & Lee, S. (2004). Interlanguage study of Korean learners about using ‘ul/lul, i/ka, ey, eyse’. *Bilingual Research*, 24, 87-108.
- Kim, J-e. (1998). Meaning and use of the case marker ‘-lo’. *Journal of Korea Linguistics*, 31, 361-393.
- Kim, J-E. (2004). The Japanese learners misuse aspect for Korean particles. *Journal of Korean Language Education*, 15(1), 1–31.
- Kim, J-E. (2015). Analysis on interrelation between error of postpositions by learners of Korean language and education contents. *Journal of Korean Language Education*, 26(4), 1-38. <https://doi.org/10.18209/iakle.2015.26.4.1>
- Kim, J. (2009). Comparative consideration on the appearance environment of ‘-e’ and ‘-eseo’ of the adverbial case postposition. *Emwunyenkwu*, 62, 49-76. <https://doi.org/10.17297/rsll.2009.62..003>
- Kim, M-J. (2013). Cwungkwukin yuhaksaynguy hankwuke cosa olyu pwunsek [Analysis of error of particles in Korean language among Chinese students in Korea]. *Yengcwuemwun*, 26, 61–89.
- Kim, S-K. (2011). Education method for the adverb postpositions of ‘ey’, ‘eyse’, ‘lo’ in the Korean language. *Kwukhakyenkwulonchong*, 8, 199-236.
- Kim, S. Y. (1997). The early acquisition of Korean morphology: A case study. *Korean Journal of Communication Disorders*, 2, 89-118.
- Kim, Y-B., & Shin, H-S. (1987). *Hyuntay hankwuke mwunpep [Modern Korean Grammar]*. Seoul: Hansin mwunhwasa.

- Kim, Y-j., & Guo, J. (2016). A study on the acquisition of Korean adverbial case marker *ey* in spoken production by Chinese Korean L2 learners. *Kwukekyoyukyenkwa [Korean Education Research]*, 38, 1-26. <https://doi.org/10.17313/jkorle.2016..38.1>
- Kim, Y. (1999). The effects of case marking information on Korean sentence processing. *Language and Cognitive Processes*, 14(5/6), 687–714. <https://doi.org/10.1080/016909699386239>
- Kim, Y. M. (2002). A study of error analysis of Korean learners by using learner corpus. *Teaching Korean as a Foreign Language*, 27, 141-168.
- Kim, Y. T., & Hong, G. (2001). A study of the acquisition of grammatical markers used by Korean children in a story retelling task: Usage and error types. *Journal of Early Childhood Education*, 21(2), 111-131.
- Kirsner, K. (1994). Implicit processes in second language learning. In N. Ellis (Eds.), *Implicit and explicit learning of languages* (pp. 283–312). San Diego, CA: Academic Press.
- Klein, D., & Manning, C. D. (2005). Natural language grammar induction with a generative constituent-context model. *Pattern Recognition*, 38(9), 1407–1419. <https://doi.org/10.1016/j.patcog.2004.03.023>
- Ko, K. (2007). A study on the necessity of the teaching ‘case marker+verb’ construction as a pattern. *Mwunpekyoyuk [Grammar Education]*, 6, 1-21.
- Ko, S-J. (2011). A reconsideration of Korean postposition ‘e’s meaning. *Journal of Korean Linguistics*, 61, 93-115. <https://doi.org/10.15811/jkl.2011..61.004>
- Ko, S-J., Kim, M. O., Kim, J. Y., Seo, S. G., Jung, H. J., & Han, S. (2004). ‘*Hankwuke haksupca*’ *malmwungchiwa olyu pwunsek* [‘Korean learners’ Corpus and error analysis]. Seoul, Korea: Hankwukmwunhwasa.

- Ko, Y. G., & Ku, P. G. (2008). *Wuli-mal mwunpep-lon [A grammar theory of Korean]*. Seoul: Cipsmwuntang.
- Kruschke, J. (1992). ALCOVE: An exemplar-based connectionist model of category learning. *Psychological Review*, 99(1), 22-44. <https://doi.org/10.1037/0033-295X.99.1.22>
- Kyle, K. (2016). *Measuring syntactic development in L2 writing: Fine grained indices of syntactic complexity and usage-based indices of syntactic sophistication*. Unpublished doctoral dissertation. Georgia State University.
- Kyle, K., & Crossley, S. (2017). Assessing syntactic sophistication in L2 writing: A usage-based approach. *Language Testing*, 34(4), 513–535. <https://doi.org/10.1177/0265532217712554>
- Lakoff, G. (1987). *Women, fire and dangerous things: What categories reveal about the mind*. Chicago: University of Chicago Press.
- Lakoff, G., & Johnson, M. (1980). *Metaphors we live by*. Chicago: University of Chicago Press.
- Langacker, R. (1987). *Foundations of cognitive grammar*. Stanford, California: Stanford University Press.
- Langacker, R. (2008). *Cognitive grammar: A basic introduction*. Oxford University Press.
- Larson-Hall, J., & Dewey, D. (2012). An examination of the effects of input, aptitude, and motivation on the language proficiency of missionaries learning Japanese as a second language. In Hansen, L. (Eds.), *Second language acquisition abroad: the LDS missionary experience* (pp. 51-88). Amsterdam/Philadelphia: John Benjamin Publishing Company.
- Lee, C-k., & Ko, Y-j. (2013). The degree of advanced Korean learners' recognition of postposition. *Emwunnoncip*, 56, 485-511.
<http://www.dbpia.co.kr/Article/NODE02341090>

- Lee, H. R., Chang-Song, Y-K., Choi, Y-L., & Lee, S. (2008). Acquisition of Grammatical function words in Korean children. *Korean Journal of Communication Disorders*, 13(2), 159-173.
- Lee, I. (2011). *Kwukehakkaysel [Introduction to Korean linguistics]*. Seoul: Hakyensa.
- Lee, J-H. (2002). A cognitive-semantic analysis of the Korean postpositions -ey and -(u)lo. *Discourse and Cognition*, 9(2), 169-190.
- Lee, J-H. (2004). A cognitive account of the Korean locative postpositions -ey and -eyse. *Discourse and Cognition*, 11(1), 237-251.
- Lee, J. H. (2003). *Hankwuke haksupcauy olyu yenkwu [Error analysis of Korean learners]*. Seoul, Korea: Pakiceng.
- Lee, K. D. (1981). The meaning of the postpositions ey and eyse. *Hangul*, 173, 174. 9-34.
- Lee, N. S. (1983). 'ey'wa 'lo'uy thongsawa uymi [Syntax and semantics of 'ey' and 'lo']. *Ene [Language]*, 8(2), 338-360.
- Lee, Y. (2005). A teaching of the Korean case-marker 'ro' for foreign learners. *Bilingual Research*, 29, 269-295.
- Lenko-Szymansk, A., & Boulton, A. (2015). *Multiple affordances of language corpora for data-driven learning (Eds)*. Amsterdam: John Benjamins.
- Leone, P. (2010). General spoken language and school language. In M. Bondi & M. Scott (Eds.), *Keyness in text: Corpus linguistic investigations* (pp. 235-248). Amsterdam: John Benjamins.
- Lim, D-H. (2017). How to express local concepts in Korean. *Journal of Korean Linguistics*, 82, 101-125. <https://doi.org/10.15811/jkl.2017..82.004>

- Lim, S., Kwon, M., Kim, J., & Kim, H. (2015). Korean proposition bank guidelines for ExoBrain. In *Proceeding of the 27th Annual Conference on Human & Cognitive Language Technology* (pp. 250–254). Human and Language Technology.
- Maeng, K. H. (2016). Hyentay hankwuke cosa ‘ey’uy inciuymilon [Cognitive understanding of modern Korean postposition ‘ey’]. *The Journal of Korean Studies*, 41, 325-366.
<http://www.dbpia.co.kr/Article/NODE06693339>
- Mandler, J. M. (1992). How to build a baby II: Conceptual primitives. *Psychological Review*, 99(4), 587-604. <https://doi.org/10.1037/0033-295X.99.4.587>
- McDonough, K., & Kim, Y. (2009). Syntactic priming, type frequency, and EFL learners' production of wh-questions. *The Modern Language Journal*, 93(3), 386–398.
<https://doi.org/10.1111/j.1540-4781.2009.00897.x>
- McDonough, K., & Nekrasova-Becker, T. (2014). Comparing the effect of skewed and balanced input on English as a foreign language learners' comprehension of the double-object dative construction. *Applied Psycholinguistics*, 35(2), 419–442.
<https://doi.org/10.1017/S0142716412000446>
- McEnery, T., & Wilson, A. (2001). *Corpus linguistics: An introduction (2nd ed.)*. Edinburgh: Edinburgh University Press.
- Min, J. Y. (2002). *Hankwuke kokup haksupcauy cosa olyu pwunsek [Postposition error analysis of advanced Korean learners]*. Unpublished master dissertation. Yonsei University, Korea.
- Mishan, F. (2005). *Designing authenticity into language learning materials*. Bristol, UK; Portland, OR: Intellect.

- Miyata, M. (2011). *The effects of frequency, distribution, mode of presentation, and first language on learning an artificial language*. Unpublished doctoral dissertation. University of Hawai'i.
- Murphy, G. (2002). *The big book of concepts*. Cambridge, MA: MIT Press.
- Nakamura, D. (2012). Input skewedness, consistency, and order of frequent verbs in frequency-driven second language construction learning: A replication and extension of Casenhiser and Goldberg (2005) to adult second language acquisition. *International Review of Applied Linguistics in Language Teaching*, 50(1), 1–37. <https://doi.org/10.1515/iral-2012-0001>
- Nam, K. S. (1993). *Kwuke cosauy yenkwu: 'ey'wa 'ro'lul cwungsimulo [Study on Korean postpositions: focused on 'ey' and 'lo']*. Seoul, Pakiceng.
- National Institute of Korean Language. (2005). *Oykwukinul wihan hankwuke mwunpep [Korean grammar for foreigners]*. Seoul, Korea: National Institute of Korean Language.
- Newell, A. (1990). *Unified theories of cognition*. Cambridge, MA: Harvard University Press.
- Ninio, A. (1999). Pathbreaking verbs in syntactic development and the question of prototypical transitivity. *Journal of Child Language*, 26(3), 619-653. <https://doi.org/10.1017/S0305000999003931>
- Nosofsky, R. (1991). Tests of an exemplar model for relating perceptual classification and recognition memory. *Journal of Experimental Psychology: Human Perception and Performance*, 17(3), 3-27. <https://doi.org/10.1037/0096-1523.17.1.3>
- Nosofsky, R. (1992). Similarity scaling and cognitive process models. *Annual Review of Psychology*, 43(1), 25-53.

- O'Grady, W. (2005). *Syntactic carpentry: An emergentist approach to syntax*. Mahwah, N.J.: Lawrence Erlbaum.
- O'Grady, W. (2015). Processing determinism. *Language Learning*, 65(1), 6-32.
<https://doi.org/10.1111/lang.12091>
- Park, J-W. (1999). A polysemy network of the Korean instrumental case. *Korean Journal of Linguistics*, 24(3), 405-425. <http://www.dbpia.co.kr/Article/NODE02132361>
- Ramscar, M., & Port, R. (2015). Categorization (without category). In E. Dąbrowska & D. Divjak (Eds.), *Handbook of cognitive linguistics* (Vol. 39) (pp. 75-99). De Gruyter Mouton.
- Ro, K. (2008). Hankwuke cosa olyu yangsang kwa cosa kyoyuk pangan [Aspects of Korean postposition errors and ideas of instruction], *Hankwuke Kyoyuk Yeonkwu [Research on Korean education]*, 6, 234-249.
- Römer, U. (2004). Comparing real and ideal language learner input: The use of an EFL textbook corpus in corpus linguistics and language teaching. In G. Aston, S. Bernardini, and D. Steward (Eds.), *Corpora and language learners* (pp. 151-168). Amsterdam; Philadelphia: John Benjamins Pub.
- Rosch, E. (1975). Cognitive reference points. *Cognitive Psychology*, 7(4), 532-547.
[https://doi.org/10.1016/0010-0285\(75\)90021-3](https://doi.org/10.1016/0010-0285(75)90021-3)
- Rosch, E., & Mervis, C.B. (1975). Family resemblances: Studies in the internal structure of categories. *Cognitive Psychology*, 7(4), 573-605. [https://doi.org/10.1016/0010-0285\(75\)90024-9](https://doi.org/10.1016/0010-0285(75)90024-9)
- Rosenfeld, B., & Penrod, S. D. (2011). *Research methods in forensic psychology*. John Wiley & Sons Inc.

- Seo, S. G. (2006). *Oykwukinul wihan hankwuke haksup sacen [Learner's dictionary of Korean]*. Seoul, Korea: Sinwon Prime.
- Seo, S. G. (2014). *Hankwuke kiponehwi uymi pinto sacen [Frequency dictionary of Korean basic lexicon and meaning]*. Seoul, Korea: Hankwuk Mwunhwasa.
- Shin, G-H. (2009). Basic verbs with constructions as an effective primary English learning material. *Primary English Education*, 15(2), 151-178.
- Shin, G-H. (2011, January). On the matter of input frequency in Korean primary English education: Basic verbs and argument structure constructions in English textbooks. *Paper presented at 2011 KAPEE International Conference*, Korea National University of Education.
- Shin, G-H. (2020). *Connecting input to comprehension: First language acquisition of active transitives and suffixal passives by Korean-speaking preschool children*. Unpublished doctoral dissertation. University of Hawai'i.
- Shin, G-H., & Deen, K. (2019). Word order, case marking, and verbal morphology in children's comprehension of suffixal passives in Korean. In M. M. Brown & B. Dailey (Eds.), *Proceedings of the 43rd Annual Boston University Conference on Language Development*. Boston, MA: Cascadilla Press.
- Shin, J-Y., Velázquez, A., Swatek, A., Staples, S., & Patridge, R.S. (2018). Examining the effectiveness of corpus-informed instruction of reporting verbs in L2 first-year college writing. *L2 Journal*, 10(3), 31-46. <https://doi.org/10.5070/L210337022>
- Sohn, H. M. (1999). *The Korean language*. Cambridge University Press.
- Song, J-J. (2006). *The Korean language: Structure, use and context*. Routledge.

- Straka, M., Hajič, J., & Straková, J. (2016). UDPipe: Trainable pipeline for processing CoNLL-U files performing tokenization, morphological analysis, POS tagging and parsing. *Proceedings of the Tenth International Conference on Language Resources and Evaluation (LREC 2016)*, Portorož, Slovenia, May 2016.
- Swaffar, J. (1985). Reading authentic texts in a foreign language: A cognitive model. *The Modern Language Journal*, 69(1), 16-32. <https://doi.org/10.1111/j.1540-4781.1985.tb02521.x>
- Thompson, S. (2007). Three Frequency Effects in Syntax. In J. Bybee, *Frequency of Use and the Organization of Language*, (pp. 269-278). Oxford University Press.
- Tomasello, M. (2000). First steps toward a usage-based theory of language acquisition. *Cognitive Linguistics*, 11(1), 61–82. <https://doi.org/10.1515/cogl.2001.012>
- Tomasello, M. (2003). *Constructing a language: A usage-based theory of language acquisition*. Cambridge, MA: Harvard University Press.
- Tono, Y. (2004). Multiple comparisons of IL, L1 and TL corpora: The case of L2 acquisition of verb subcategorization patterns by Japanese learners of English. In G. Aston, S. Bernardini, & D. Stewart (Eds). *Corpora and language learners* (pp. 45-66). Amsterdam/Philadelphia: John Benjamins Publishing Company.
- Türker, E. (2005). *Locative expressions in Korean and Turkish: A cognitive grammar approach*. Unpublished doctoral dissertation. University of Hawaii.
- Türker, E. (2017). The acquisition and development of the Korean adverbial particle -ey by L1 English learners of Korean. *Journal of Korean Language Education*, 28(4), 337-366.
- Tyler, A. (2012). *Cognitive linguistics and second language learning: Theoretical basics and experimental evidence*. UK: Routledge.

- Van Lier, L. (1996). *Interaction in the language curriculum: Awareness, autonomy and authenticity*. Essex: Longman.
- Verhagen, A. (2005). *Constructions of intersubjectivity: Discourse, syntax, and cognition*. Oxford: Oxford University Press.
- Widdowson, H. G. (1998). Context, community, and authentic language. *TESOL Quarterly*, 32(4), 705-716. <https://doi.org/10.2307/3588001>
- Woo, C-H. (2016). A methodology on the education of Korean case marker considering the syntactic context: On the focus of adverbial case marker ‘-(eu)ro’. *Kyeleyemwukhak*, 57, 209-231.
- Year, J. (2009). *Korean speakers’ acquisition of the English ditransitive construction: The role of input frequency and distribution*. Unpublished doctoral dissertation. Colombia University.
- Year, J., & Gordon, P. (2009). Korean speakers’ acquisition of the English ditransitive construction: the role of verb prototype, input distribution, and frequency. *Modern Language Journal*, 93, 399-417. <https://doi.org/10.1111/j.1540-4781.2009.00898.x>
- Yi, K. (1981). Cosa ‘ey’wa ‘eyse’uy kiponuymi [Basic meanings of postposition ‘ey’ and ‘eyse’]. *Hangul [Korean letter]*, 73, 9-33.
- Yi, K. (1993). *A Korean grammar on semantic-pragmatic principles*. Seoul, Korea: Hankwukmwunhwasa.
- Zhao, M. (2016). A study for the adverb postpositions of ‘ey’, ‘eyse’ in the Korean language: Focused on ‘ey’, ‘eyse’. *Studies in Linguistics*, 40, 337-359.
- Zipf, G. K. (1935). *The psycho-biology of language*. Boston: Houghton Mifflin.

APPENDICES

Table A- 1. Functions and examples of postposition *-ey*

Function	Example
Location	<i>Nari-ga cip-ey iss-ta.</i> Nari-NOM home-ey be- SE 'Nari is at home.'
Time	<i>seysi-ey swuep-i iss-ta.</i> three-ey class-NOM have-SE '(I) have a class at three.'
Destination/Goal	<i>Mina-ga hakkyo-ey ka-ta.</i> Mina-NOM school-ey go-SE 'Mina goes to school.'
Cause	<i>pi-ey tali-ka kkunh-ki-ess-ta.</i> rain-ey bridge-NOM disconnect-PSV-PST-SE 'Due to rain, the bridge was disconnected.'
Causing certain movement	<i>ku-uy uykyen-ey chansengha-ta.</i> he-GEN opinion-ey agree-SE '(I) agree with his opinion.'
Scope	<i>kkochpath-ey mwul-ul cwu-ta.</i> garden-ey water-ACC give-SE '(I) gave water to the flowers ((I) watered the garden).'
Aim/Purpose	<i>i-key-ul eti-ey ssu-na?</i> this-thing-ACC where-ey use-Q? 'Where can (we) use this one?'
Means/Method	<i>khal-ey son-ul pey-i-ess-ta.</i> knife-ey hand-ACC cut-CAU-PST-SE 'I had my hand cut with the knife.'
Environment/Condition	<i>i nalssi-ey pakk-ey ettehkey naka-ci?</i> this weather-ey outside-LOC how go out-Q 'How could (we) go out in this weather?'
Unit/Standard	<i>sakwa-ga twu-kay-ey man-won-i-ta.</i> apple-NOM two-counter-ey ten hundred-currency-to be-SE 'For the apples, (the price) is 10 dollars for two.'
Comparison	<i>ku apeci-ey ku atul.</i> the father-ey the son 'Like father, like son.'

Qualification	<i>ku-nun taythonglyeng-ey chwiimhay-ss-ta.</i> he-NOM president-ey inaugurate-PST-SE 'He was inaugurated as president.'
Limiting scope	<i>phoyulyu-ey mwe-ka iss-ci?</i> <i>mammal-ey what- NOM be-Q?</i> 'What belongs to (the category of) mammal?'
Object (mostly with <i>tayhay, uyhay</i> 'about')	<i>i mwuncey-ey tayhay iyakiha-ca.</i> this issue-ey about talk-PR 'Let's talk about this issue.'
Addition / Enumeration	<i>sa-ey o-lul teha-myen ku-i-ta.</i> four-ey five-ACC add-once nine-to be-SE 'Once you add five to four, it is nine.'

Table A- 2. Functions and examples of postposition *-eyse*

Function	Example
Location for activity	<i>cip-eyse picha-lul mek-ess-ta.</i> house-eyse pizza-ACC eat-PST-SE. '(I) ate pizza at home.'
Starting point	<i>cip-eyse hakkyo-kkaci ka-ta.</i> house-eyse school-to go-SE '(I) go from home to school.'
Source	<i>ku-nun hoysa-eyse senmwul-ul pat-ass-ta.</i> he-TOP company-eyse present-ACC receive-PST-SE 'He received a present from the company.'
Reason	<i>komawun maum-eyse kulehkey hay-ss-ta</i> thankful heart-eyse like that do-PST-SE '(I) did that with (my) thankful heart (to express my thanks).'
Comparison/Standard	<i>i ceyphwum-i mas-eyse ceyil-i-ta.</i> this product-NOM taste-eyse the best-be-SE 'This product is the most delicious one.'
	<i>i sanghwang-eyse ettehkey te naaci-l-kka?</i> this situation-eyse how more become better-PRO-Q 'How can things get better from this situation?'
Indicating subject	<i>wuli haykyo-eyse wusungha-yess-ta.</i> our school-eyse win-PST-SE 'Our school won (the game).'

Table A- 3. Functions and examples of postposition *-(u)lo*

Function	Example
Direction	<i>John-i hakkyo-lo ka-yo.</i> John-NOM school-lo go-SE. 'John goes to(wards) school.'
Path	<i>ku-nun twiskil-lo naka-ss-ta</i> he-TOP backstreet-lo go out-PST-SE 'He went out (through) the backstreet.'
Result of change	<i>mwul-i elum-ulo pyenha-ss-ta</i> water-NOM ice-ulo change-PST-SE 'The water changed into ice.'
Ingredient/Material	<i>namwu-lo cip-ul ci-ess-ta</i> wood-lo house-ACC make-PST-SE '(I) made the house with wood.'
Instrument	<i>yenphil-lo ssu-ca</i> pencil-lo write-PR '(Let's) write with a pencil.'
Method/Manner	<i>culke-wun maum-ulo ilha-ta</i> happy-ADN mind-ulo work-SE '(Someone) works with a happy attitude.'
Reason	<i>am-ulo cwuk-ta</i> cancer-ulo die-SE '(Someone) dies of cancer.'
Qualification	<i>hankwuk-eyse kica-lo ilnyen-kan iss-ess-ta</i> Korea-LOC report-lo one year-for be-PST-SE 'I spent one year in Korea as a reporter.'
Time	<i>ku-nun pom-kaul-lo pyengwen-ey ka-ss-ta</i> he-TOP spring-fall-lo hospital-LOC go-PST-SE 'He visited the hospital every spring and fall.'
Counting time	<i>olhay-lo sip nyen-i toy-ess-ta.</i> this-year-lo ten year-NOM become-PST-SE 'This is the tenth year.'
Object for causative or comitative	<i>na-lo hayekum ku il-ul ha-key hay-ss-ta.</i> I-lo making the work-ACC do-CAU-PST-SE '(S/he) made me do the work.'
Decision/Promise	<i>wuli-nun nayil mannaki-lo hay-ss-ta.</i> we-TOP tomorrow meeting-lo do-PST-SE 'We decided to meet tomorrow.'

Way of thinking

ku-ka na-lul papo-lo yeki-ess-ta.
he-NOM I- ACC fool-lo regard-PST-SE
'He regarded me as a fool.'
