SECOND LANGUAGE ACQUISITION OF VARIABLE USE OF THE NOMINATIVE AND ACCUSATIVE CASE MORPHEMES IN KOREAN: A CORPUS STUDY

A Dissertation submitted to the Faculty of the Graduate School of Arts and Sciences of Georgetown University in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Linguistics

By

Go-Eun Chae, M.S.

Washington, DC
May 16, 2018
Copyright 2018 by Go-Eun Chae
All Rights Reserved
SECOND LANGUAGE ACQUISITION OF VARIABLE USE OF THE NOMINATIVE AND ACCUSATIVE CASE MORPHEMES IN KOREAN: A CORPUS STUDY

Go-Eun Chae, M.S.

Thesis Advisor: Donna Lardiere, Ph.D.

ABSTRACT

The goal of this dissertation is to examine the second language (L2) acquisition of variable use of nominative/accusative case morphemes by English- or Japanese-speaking adult learners of Korean.

In Korean, case marking of subjects/objects is optional. Previous research proposes that the naturalness of case (un-)marking is constrained by semantic features for prototypical subject/object-hood (i.e., markedness features) such as animacy and definiteness (Lee, 2006b) and by discourse features such as newness (Sohn, 1999) and various focus subtypes (Lee & Choi, 2010). Regarding the L2 acquisition of these conditioning features, the Interface Hypothesis (henceforth IH, Sorace, 2011) posits persistent non-nativelikeness at the syntax-discourse interface, whereas the Feature Reassembly Hypothesis (henceforth FRH, Lardiere, 2009) predicts that L2 features realized in dissimilar configurations from those of the L1 pose substantial difficulty but are acquirable in principle given evident morphological contrasts.

Using a spoken corpus of media discourse, this study investigates whether learners’ L2 morpheme use displays nativelike sensitivity to these markedness and discourse features. The data were collected from eight advanced adult learners of Korean whose L1 is English or Japanese and four native controls. No effect of the markedness factor was observed among either the L1 controls’ or the L2 speakers’ results. The discourse factor, on the other hand, was found to have significant effects on all three groups’ overall performance, in that overt case marking is
more prevalent among discourse-new or contrastive focus DPs, suggesting that these learners have acquired the interaction between discourse context and case marking. However, further analysis revealed that while the native speakers tend to integrate an exclusivity sub-distinction for contrastive focus in discriminating the two morphemes, neither English nor Japanese speakers recognized this sub-distinction, except for one English speaker. Still, the Japanese speakers, whose L1 has a similar case system to the L2, performed better than the English speakers in overall frequency of case (un)marking and on other sub-patterns.

These findings indicate that adult L2 knowledge of at least some conditioning factors for variability phenomena at the syntax-discourse interface is attainable; the degree of L1-L2 dissimilarities plays a significant role in L2 development, providing additional support for the FRH.
To Young-Sook Kim

This is your work.
ACKNOWLEDGMENTS

One of the luckiest things that ever happened to me was that I met my advisor, Dr. Donna Lardiere, who provided me with truly exceptional support and encouragement throughout my doctoral journey. I was able to complete my studies because of her help. This dissertation would not have been possible at all without her help, from the beginning to the very end. During my first few years at Georgetown, I was mostly silent with my interest in this topic, assuming that a phenomenon of morphological variability may not be considered as a meaningful research topic. Dr. Lardiere was the one who secured my research direction with my interest by encouraging my incipient ideas of looking at the phenomenon from an acquisition perspective. The discussions that I had with her while working on this dissertation were extremely valuable in that they always guided me to balanced research endeavors between theoretical issues and acquisition outcomes. I consider my experience of working with acquisition data under her guidance empirically constructive and theoretically transparent, and therefore intellectually liberating, and I feel deeply honored to complete my dissertation under such positive influences. As my mentor, she provided me with so much support not only for my academic progress but also for my wellbeing at Georgetown. She was warm and kind. She was the type of mentor that would send an email to her student to calm the student’s pre-presentation jitters with cordial advice to eat a few squares of dark chocolate beforehand. She was the type of mentor that would respond to her student’s email greeting with a peaceful landscape photo that she herself took that morning. Because it was so easy to develop a sincere respect for the whole character of my mentor, every single progress that I was making to finish this dissertation was taken as an extra-rewarding experience.

I would like to extend my deepest gratitude to my committee members, who fulfilled crucial roles throughout the whole research process. Dr. Miok Pak helped me to navigate through the maze of real language data that often gave me a difficult time in determining the grammatical roles. My classification of the assorted subtypes of the data greatly benefited from her expertise in the Korean syntax and her generosity and patience to answer my endless questions. With her critical guidance on incorporating the non-subject/non-object data in my analysis, she taught me an indispensable lesson on the importance of presenting linguistic data as faithfully as possible. I
must thank Dr. Paul Portner for his exceptionally helpful guidance on the semantic and pragmatic aspects of this work. Without his invaluable insights into the area and theoretical rigor that he shares with students, accommodating the data with the fine-grained coding schemes would have been plainly impossible. Dr. Jaemyung Goo, although his name does not appear as an officially recognized member of my committee, deserves no less gratitude from me. Dr. Goo was with me at every stage of this dissertation as a valued source of guidance and expert advice on L2 research methods. I owe a debt of gratitude for his generosity with time and effort to send me his precious feedback on my work.

I am deeply grateful to many people in the Linguistics Department at Georgetown. I greatly appreciate the solid theoretical training in syntax and phonology that I received from Dr. Héctor Campos, Dr. Ruth Kramer, and Dr. Elizabeth Zsiga. I would like to thank Dr. Natalie Schilling for her encouraging comments on my sociolinguistic observation of this topic, and Dr. Cynthia Gordon for sharing practical tips for efficient transcription of the media discourse data. I also thank the Graduate Program Coordinator, Ms. Erin Esch Pereira for all the administrative support that facilitated my academic progress here. I thank my fellow students for their help, as well: Hanwool Choi for her careful re-examination of the transcription and coding materials for this dissertation, and Eunji Lee, Bokyung Moon, and Jinsok Lee for sharing their experience and tips for a successful Ph.D. journey.

My special thanks go to my dearest friends/supporters and good neighbors around me. I thank Jihoon Jung for volunteering his time to download my transcription materials to send them to me, and Okjin Kim for her moral support and untainted friendship. I thank Dr. Chang-Won Kim for teaching me a virtue of intellectual humility through our two-year-long Kant seminar, and Dr. Colin Bird for teaching me to embrace the virtue with emotional honesty. I would like to especially thank Deborah, Heather, Tess, and Varsha for extending their values to helping me to persist as an authentic mind. I am particularly thankful for the various resources for students’ wellbeing at Georgetown University that ensured peace and comfort in my everyday routines to finish this dissertation: genuinely pleasant greetings that I received from two security guards as I entered the library, soothing smiles that I received from a night shift cleaning lady on the 1st floor, a thousand moments of delightful pleasure that I received from somebody with perfect blueberry-bagel-toasting-behaviors on the 2nd floor, and caring gestures that I received from two
Georgetown SafeRides drivers as I returned home safely. Wherever there is work done by self-effacing individuals, there are beneficiaries, and I am one of them. I thank these people very much for the quiet kindness that they spread as I needed it. I thank them very much for muffling noises as I too wanted to be quiet.

My heartfelt thanks should go to my family. I thank Young-Sook Kim for being my mother, Hong-Gul Chae for being my father, and Jin-Young Chae for being my only brother. I thank them tremendously for their unwavering trust that I could finish this journey, for their endless support, and for their amazing patience. This dissertation is dedicated to my mother, Young-Sook Kim, who makes her daughter continue to identify the world as a fortunate place.
# TABLE OF CONTENTS

Chapter 1 Introduction ........................................................................................................... 1  
1.1. Variable use of the nominative and accusative case morphemes in Korean .......... 1  
1.2. L2 ultimate attainment ............................................................................................. 3  
1.3. The present study ..................................................................................................... 5  
1.4. Dissertation organization ......................................................................................... 7  

Chapter 2 Variable use of the nominative and accusative case morphemes in Korean ...... 9  
2.1. Nominative and accusative case morphemes in Korean ......................................... 9  
2.2. Conditioning factors for variable case marking of subjects and objects .......... 11  
  2.2.1. Syntactic and phonological factors .................................................................... 11  
    2.2.1.1. Subject-object asymmetry ........................................................................... 11  
    2.2.1.2. Word order ............................................................................................... 13  
    2.2.1.3. Verb adjacency ......................................................................................... 16  
    2.2.1.4. Argument length and complexity .............................................................. 18  
  2.2.2. Semantic and pragmatic factors ......................................................................... 20  
    2.2.2.1. Animacy ................................................................................................... 20  
    2.2.2.2. Definiteness ............................................................................................. 22  
    2.2.2.3. Focus ....................................................................................................... 23  
    2.2.2.4. Discourse newness ................................................................................... 34  
    2.2.2.5. Why do the factors have the effects that they do? ................................... 36  
  2.3. More on the particle marking system in Korean ...................................................... 40  
    2.3.1. Other particle-marked subjects and objects ................................................... 40  
    2.3.2. Nominative and accusative-marked non-subjects and non-objects ............ 48  
      2.3.2.1. Adjectival and verbal predicates in negative sentences ....................... 49  
      2.3.2.2. Nominal complements of linking verbs ............................................... 52  
      2.3.2.3. Multiple nominative constructions ....................................................... 53  
      2.3.2.4. Alternation between accusative and other (case) particles .............. 55  
      2.3.2.5. Nominative and accusative stacking onto other (case) particles .... 56
Chapter 3 Acquisition of variable case marking of subjects and objects in Korean

3.1. L1 acquisition

3.1.1. Background

3.1.2. Chung (2013)

3.1.3. Chae (2017)

3.1.4. Park-Johnson (2017)

3.2. L2 acquisition

3.2.1. Ahn & Herschensohn (2013)

3.2.2. Chung (2013)

3.3. Summary

Chapter 4 L2 ultimate attainment and acquisition at interfaces

4.1. Theoretical approaches to L2 ultimate attainment

4.1.1. The Representational Deficit Hypothesis (RDH)

4.1.2. Where does the acquisition burden lie?

4.1.3. The Feature Reassembly Hypothesis (FRH)

4.1.4. The Interface Hypothesis (IH)

4.1.5. Summary

4.2. The present study

4.2.1. Acquisition at interfaces

4.2.2. The role of the L1

4.2.3. Research questions and predictions
6.4.1.2. Nominative and accusative data ................................................................. 219
6.4.1.3. Where K- and J-groups diverged: Non-subjects and objects .......................... 222
6.4.2. Effects of the discourse factor .......................................................................... 224
6.4.2.1. Effects of the broad discourse categories: [new], [focus], [none] ... 224
6.4.2.2. Effects of the [new] category: Subjects and non-subjects ......................... 226
6.4.2.3. Effects of the [focus] subcategories: Total data ........................................ 228
6.4.2.4. Effects of the [contrast] subcategories: Nominative and accusative data ................................................................. 232
6.5. Summary of the results ....................................................................................... 239
6.5.1. Frequency of case (un)marking ...................................................................... 239
6.5.2. Effects of the markedness factor ...................................................................... 240
6.5.3. Effects of the discourse factor ........................................................................ 241
6.5.4. J-group vs. E-group ....................................................................................... 243
6.5.5. Individual results ............................................................................................ 244

Chapter 7 Conclusion ................................................................................................. 246
7.1. Summary: Answers to the research questions .................................................. 246
7.2. Discussions: For future research ....................................................................... 250

Appendices ............................................................................................................... 258
Appendix A: Sources of L2 speakers’ language background Appendixes .............. 258
Appendix B: Video data used for transcription ....................................................... 259
Appendix C: Supplementary analysis of the markedness factor for subjects and objects .......................................................................................................................... 264
Appendix D: Supplementary analysis of the markedness factor for non-subjects .. 269
Appendix E: Supplementary analysis of individual results .................................... 271

References ................................................................................................................. 273
LIST OF FIGURES

Figure 2.1. Kim’s (2008) findings: Interaction between case marking and number of
modifiers .............................................................. 20

Figure 3.1. Chae’s (2017) findings: Effects of focus on frequency (%) of case
(un)marking.subjects.grouped data ........................................ 76

Figure 3.2. Chae’s (2017) findings: Effects of focus on frequency (%) of case
(un)marking.subjects.individual data ........................................ 77

Figure 6.1. Frequency (%) of case (un)marking: Total data of each L1 group ........ 171
Figure 6.2. Frequency (%) of case (un)marking: NOM and ACC data .................. 172
Figure 6.3. Frequency (%) of case (un)marking: Subjects and non-subjects .......... 173
Figure 6.4. Frequency (%) of case (un)marking: Objects and non-objects .......... 174
Figure 6.5. Frequency (%) of case (un)marking: Subjects/objects and non-subjects/
non-objects ................................................................ 175

Figure 6.6. Effects of the markedness factor on frequency (%) of case (un)marking:
Subjects.K-group ................................................................ 178
Figure 6.7. Effects of the markedness factor on frequency (%) of case (un)marking:
Subjects.K-group.subanalysis of definiteness effects .................... 180
Figure 6.8. Effects of the markedness factor on frequency (%) of case (un)marking:
Subjects.E-group ................................................................ 181
Figure 6.9. Effects of the markedness factor on frequency (%) of case (un)marking:
Figure 6.10. Effects of the discourse factor on frequency (%) of case (un)marking:
Total data of each L1 group .............................................. 190

Figure 6.11. Effects of [focus] on frequency (%) of case (un)marking: Total data
of each L1 group.subcategories ........................................... 197

Figure 6.12. Effects of [focus] on frequency (%) of case (un)marking: Subjects/objects.
subcategories .................................................................. 200

Figure 6.13. Effects of [focus] on frequency (%) of case (un)marking: Subjects.
subcategories .................................................................. 202
Figure 6.14. Effects of [focus] on frequency (%) of case (un)marking: Objects.

Figure 6.15. Effects of [focus] on frequency (%) of case (un)marking: Non-subjects/non-objects.

Figure 6.16. Effects of [focus] on frequency (%) of case (un)marking: Non-subjects.

Figure 6.17. Effects of [focus] on frequency (%) of case (un)marking: Non-objects.

Figure 6.18. Effects of [focus] on frequency (%) of case (un)marking: NOM and ACC data.

Figure C.1. Effects of the markedness factor on frequency (%) of case (un)marking: Subjects.K-group.subanalysis of definiteness.strongly definite vs. low definite.

Figure C.2. Effects of the markedness factor on frequency (%) of case (un)marking: Subjects.K-group.full (sub)categories.

Figure C.3. Effects of the markedness factor on frequency (%) of case (un)marking: Subjects.J-group.

Figure C.4. Effects of the markedness factor on frequency (%) of case (un)marking: Subjects.J-group.full (sub)categories.

Figure C.5. Effects of the markedness factor on frequency (%) of case (un)marking: Subjects.E-group.full (sub)categories.

Figure C.6. Effects of the markedness factor on frequency (%) of case (un)marking: Objects.K-group.subanalysis of definiteness.

Figure C.7. Effects of the markedness factor on frequency (%) of case (un)marking: Objects.K-group.full (sub)categories.

Figure C.8. Effects of the markedness factor on frequency (%) of case (un)marking: Objects.J-group.full (sub)categories.

Figure C.9. Effects of the markedness factor on frequency (%) of case (un)marking: Objects.E-group.full (sub)categories.
Figure D.1. Effects of the markedness factor on frequency (%) of case (un)marking:
    non-subjects.K-group........................................................................................................ 269
Figure D.2. Effects of the markedness factor on frequency (%) of case (un)marking:
    Non-subjects.J-group ........................................................................................................ 269
Figure D.3. Effects of the markedness factor on frequency (%) of case (un)marking:
    Non-subjects.E-group ........................................................................................................ 270
LIST OF TABLES

Table 2.1. Kim’s (2008) findings: Interaction of subject marking and word order .... 14
Table 2.2. Kim’s (2008) findings: Interaction of object marking and word order ..... 15
Table 2.3. Lee’s (2006b) findings: Interaction of case marking and animacy ........ 21
Table 2.4. Lee’s (2006b) findings: Interaction of case marking and definiteness .... 23
Table 2.5. Lee and Choi’s (2010) findings: Mean judgments for each subtype of focus objects ........................................................................................................ 32
Table 2.6. Yang’s (1972) classification of delimiters in Korean (cited from Sohn, 1999) ........................................................................................................................................................................... 41
Table 2.7. Kim’s (2008) findings: Interaction of subject marking and contrast subtype 46
Table 2.8. Kim’s (2008) findings: Interaction of object marking and contrast subtype.. 47
Table 2.9. Conditioning factors for variable case marking of subjects and objects .... 67
Table 4.1. Conditions for overt PF-realization of [case] .............................................. 100
Table 5.1. Overview of the data source: Picengsanghoytam (‘Non-Summit’)......... 120
Table 5.2. L1 and L2 speakers: Preliminary information ........................................... 123
Table 5.3. L2 speakers: Language background.J-group .............................................. 125
Table 5.4. L2 speakers: Language background.E-group ............................................. 126
Table 5.5. Overview of data tokens ............................................................................. 146
Table 5.6. Coding of [focus]: Subcategories ............................................................... 156
Table 5.7. Coding of the markedness factor: Coding categories ............................... 168
Table 5.8. Coding of the discourse factor: Coding categories ................................. 168
Table 6.1. Frequency of case (un)marking: Number of data tokens ...................... 170
Table 6.2. Effects of the markedness factor on case (un)marking: Number of subject tokens ........................................................................................................................ 177
Table 6.3. Effects of the markedness factor on case (un)marking: Number of object tokens ........................................................................................................................ 182
Table 6.4. Effects of the markedness factor on case (un)marking: Number of non-subject tokens .............................................................................................................. 186
Table 6.5. Effects of the markedness factor on case (un)marking: Number of non-object tokens ................................................................. 187
Table 6.6. Coding of the discourse factor: Coding categories.......................... 188
Table 6.7. Effects of the discourse factor on case (un)marking: Number of data tokens .................................................................................. 189
Table 6.8. Effects of the discourse factor on case (un)marking for K-group: Number of data tokens ................................................................. 192
Table 6.9. Effects of the discourse factor on case (un)marking for J-group: Number of data tokens ................................................................. 193
Table 6.10. Effects of the discourse factor on case (un)marking for E-group: Number of data tokens ................................................................. 194
Table 6.11. Subcategories of [contrast] associated with high frequency of case marking: Across data subsets for each L1 group ......................... 213
Table 6.12. Frequency (%) of case marking: Number of data tokens.total data. individual results ........................................................................ 217
Table 6.13. Frequency (%) of case marking: Number of data tokens.NOM data. individual results ................................................................. 220
Table 6.14. Frequency (%) of case marking: Number of data tokens.ACC data. individual results ........................................................................ 220
Table 6.15. Effects of the discourse factor on frequency (%) of case marking: Total data individual results ......................................................... 225
Table 6.16. Effects of [new] on frequency (%) of nominative marking: [new] subjects vs. [new] non-subjects.individual results ......................... 227
Table 6.17. Effects of [focus] on frequency (%) of case marking: [focus] subcategories.total data.individual results ............................................. 230
Table 6.18. Effects of [focus] on frequency (%) of case marking: [contrast] subcategories.NOM data.individual results ............................................. 234
Table 6.19. Effects of [focus] on frequency (%) of case marking: [contrast] subcategories.ACC data.individual results ............................................. 237
Table E.1. Frequency (%) of case marking: Number of data tokens.non-subjects.

individual results ........................................................................................................ 271

Table E.2. Frequency (%) of case marking: Number of data tokens.objects.

individual results ........................................................................................................ 272
LIST OF ABBREVIATIONS

ACC: accusative
ADVZ: adverbializer
COMP: complementizer
CNJ: conjunctive
COP: copula
DAT: dative
DCL: declarative
FUT: future
GEN: genitive
HON: honorific
IN: indicative
INF: infinitive
LOC: locative
NEG: negative
NMLZ: nominalizer
NOM: nominative
PERF: perfective
PL: plural
PR: propositive
PRS: prospective
PRES: present
PST: past
Q: question
QT: quotative
REL: relative
SUP: suppositive
TC: topic-contrast particle
CHAPTER 1
INTRODUCTION

The general objective of this dissertation is to examine the attainment in second language (L2) acquisition of variable use of the nominative and accusative case morphemes by native English- or Japanese-speaking adult learners of Korean and evaluate the predictions of different L2 theories. Using naturalistic corpus data, this study investigates whether adult L2 learners can successfully acquire linguistic knowledge that constrains appropriate (non-)use of the case morphemes in Korean.

1.1. Variable use of the nominative and accusative case morphemes in Korean

Korean is considered as a language in which subject and object arguments are case-marked, but case marking is optional. See (1).

(1) Nabee-ka/-ø kapang-ul/-ø manhi sa-ss-ta.

Nabee-NOM/-ø bag-ACC/-ø a lot buy-PST-DCL

'Nabee bought a lot of bags.'

Although the distribution of case-marked and -unmarked subject/object forms is not governed by categorical rules, a good deal of research on case drop in Japanese and Korean suggests that the variation is not merely idiosyncratic but is constrained by semantic and discourse-pragmatic features.
It has been claimed that the naturalness of case drop is predictable in terms of animacy and definiteness features of subject/object DPs (Fry, 2001; Kim, 2008; Lee, 2006a, 2006b; Minashima, 2001). More specifically, case drop is more prevalent for animate/definite subjects and for inanimate/indefinite objects. The underlying generalization is that animate/definite subjects and inanimate/indefinite objects bear a less prominent need for morphological specification of the grammatical function by virtue of being prototypical agents and patients, respectively (Hawkins, 2004).

Past studies of pragmatics of the case particles indicate that the naturalness of case drop is highly influenced by the discourse-information status of the argument that the particle marks. The general claim is that case drop for subject/object DPs conveying a special discourse function is less natural. Such discourse properties have been considered in terms of two conceptual dimensions: discourse newness and focus. A well-acknowledged observation is that DPs construed as focus or as new information disfavor case ellipsis (Ko 2000; Schütze 1999; Sohn, 1999, among others). Focus, in particular, has been extensively discussed in the previous literature as one of the strongest factors that regulate the variability phenomenon. It has so far been noted that (i) the notion of focus concerned with the phenomenon is a particular type of focus, the so-called ‘contrastive focus’ (Ko, 2000; Lee, 2003; Lee, 2006a; Lee & Ramsey, 2000; Tsutsui, 1984; Yatabe, 1999); (ii) what is likely to be truly accountable for the variation is more fine-grained sub-distinctions of contrastive focus (Kim, 2008; Lee, 2010, 2011; Lee & Choi, 2010); (iii) the sub-distinctions of contrastive focus can also help to define the discourse meaning uniquely associated with the case morphemes in relation to similar discourse uses of delimiting particles in the language (Kim, 2008; Lee, 2003).
While the primary environment in which variable use of the nominative/accusative morphemes is observed is where the morphemes are used to mark subject/object arguments (i.e., variable case marking of subjects/objects), Korean also allows various contexts in which the case morphemes are optionally attached to elements that do not bear the corresponding argument roles (i.e., variable morpheme use for non-subjects and non-objects). These contexts include sentential negation, linking verb usages, multiple nominative/accusative constructions, case alternation, case stacking, different kinds of compound verbs, and psych-predicates. Although instances of such morpheme use are often said to be triggered by a discourse-pragmatic consideration (Sohn, 1999; Yoon, 2004, 2006), no previous research has yet included them in observing the distribution of the (case) particle-marked and -unmarked forms.

1.2. L2 ultimate attainment

Generative L2 acquisition studies have examined possible causes of non-nativelike outcomes in endstate L2 grammars in terms of the role of Universal Grammar (UG) in adult L2 acquisition along with the influence of the first language (L1).

One stream of theoretical thought grouped as the Representational Deficit Hypothesis (RDH) views the issue as an acquisitional consequence of parametric selection of features, predicting ultimate unacquirablility of L2 morphosyntactic features that are not included in the L1 (e.g., Hawkins & Chan, 1997). Another line of approach known as the Feature Reassembly Hypothesis (FRH) associates the target-deviant L2 end results with cross-linguistic differences in configurations of formal features in morpholexical items selected by each language and identifies the primary difficulty as remapping those features onto different morphemes in different
language-specific ways and conditioning environments from the L1 (e.g., Lardiere, 2008, 2009). Based on the Full Transfer/Full Access Hypothesis (Schwartz & Sprouse, 1996), the FRH points to L1-L2 dissimilarities in configurations of morphosyntactic features as the most likely obstacle to L2 acquisition, but unlike the RDH, it predicts that categories and features that are unavailable in the L1 or organized differently from the L1 are all ultimately acquirable as long as feature contrasts in the L2 are detectable. Lastly, recent research on L2 ultimate attainment has paid special attention to acquisition of interface phenomena. The Interface Hypothesis (IH), proposed by Sorace (2011), identifies properties involving the coordination of various linguistic and nonlinguistic modules as the locus of non-nativelikeness. The IH claim refined through its recent developments is that phenomena at internal interfaces such as the syntax-semantics interface are readily acquirable (e.g., Iverson & Rothman, 2008), whereas properties connected to external interfaces such as the syntax-pragmatics interface are more problematic and subject to almost permanent optionality in L2 grammars (e.g., Lozano, 2006).

The three L2 theories make distinct predictions for the L2 acquisition of case variability by advanced adult learners of Korean whose L1 is English or Japanese. Provided that the notion of case that undergoes parametric selection of features is the abstract conception, structural case, case features are considered to be selected by both English and Japanese. This makes the RDH predict no significant learning difficulty for either group of the L2 speakers. On the other hand, the FRH and IH consider that the L2 acquisition of case and case variability in Korean poses nontrivial challenges in mapping the conditioning features onto case morphology in a nativelike manner. The two approaches diverge in terms of the specific predictions, however. The FRH predicts that the extent of persistence of non-target-like properties, which still can be ultimately
overcome, reflects the degrees of the L1-L2 dissimilarities (i.e., English-Korean vs. Japanese-Korean pairings). In contrast, the IH ascribes the degree of non-nativelikeness to the distinct natures of mapping between the features relevant to the acquisition (i.e., semantic vs. discourse features) rather than to the different L1-L2 pairings. Thus, the IH predicts that for both English and Japanese speakers, the discourse features associated with external interfaces are unlikely to be acquired, whereas the semantic features concerned with internal interfaces do not pose any learning difficulty.

1.3. The present study

This dissertation investigated whether native English- or Japanese-speaking adult learners of Korean can develop nativelike sensitivity to the semantic and discourse-pragmatic features that condition the naturalness of variable use of the case particles in Korean. The different types of integration of the conditioning factors involved in the L2 acquisition and the L2 speakers examined in the study provide an intriguing testing ground for the different L2 theories. In particular, the fact that the phenomenon concerns both internal and external interface acquisitions and that the L2 data comes from native speakers of two different L1s enables the investigation to weigh between the conflicting predictions of the FRH and IH.

The present study was based on naturally occurring L1 and L2 spoken corpus of conversational Korean produced in mass media discourse. The data were collected from eight advanced adult learners of Korean whose L1 is English or Japanese and four native controls. The analysis showed no effect of the semantic features among either the L1 controls’ or the L2 speakers’ results. The discourse features, on the other hand, were found to have significant
effects on all three groups’ overall performance, in that overt case marking is more prevalent among discourse-new or contrastive focus DPs, suggesting that these learners have acquired knowledge of the interaction between discourse context and case marking. However, further analysis revealed that while the native speakers tend to integrate an exclusivity sub-distinction for contrastive focus in discriminating the two case morphemes, neither English nor Japanese speakers recognized this sub-distinction, except for one English speaker. Still, the Japanese speakers, whose L1 has a similar case system to the L2, performed better than the English speakers in overall frequency of case (un)marking and on other focus sub-patterns. These findings indicate that adult L2 knowledge of at least some conditioning factors for variability phenomena at the syntax-discourse interface is indeed attainable; rather, the degree of L1-L2 dissimilarities plays a significant role in L2 development, providing additional support for the FRH.

This study is distinguished from the past work on case ellipsis in Korean in that (i) the range of data for the variable properties subsumes not only case marking of subjects/objects but also case morpheme attachment to non-subjects/non-objects, and (ii) the data is examined in terms of fine-grained categories of the semantic and pragmatic features that have been claimed to influence the phenomenon. Overall, by exploring more comprehensive data types and subtle distinctions for possible meanings of the case particles, this dissertation addresses the phenomenon from a broader perspective of the actual morpheme use while investigating the specific extent of the granularity of linguistic knowledge that can be attained in adult L2 acquisition.
1.4. Dissertation organization

The remaining chapters of this dissertation are organized as follows: Chapter 2 provides background information on variable use of the nominative and accusative case morphemes in Korean. Following a comprehensive review of previous research on the conditioning factors for variable case marking of subjects/objects, the chapter discusses other subject/object encoding options available in Korean and optional use of the case morphemes for non-subjects/non-objects.

Chapter 3 addresses previous studies of acquisition of variable case marking of subjects and objects in Korean. Although there has not been much research done on the topic, by discussing the few L1 and L2 acquisition studies that have been carried out, the chapter aims to introduce the recent research stream to which this dissertation can make empirical contributions.

Chapter 4 is devoted to a review of the three different theoretical approaches to L2 ultimate attainment: the RDH, FRH, and IH. After relating these L2 theories to the specific learning tasks of the L2 acquisition investigated in this study, the chapter concludes by articulating the research questions along with the predictions of the L2 theories.

Chapter 5 presents an overview of the corpus data and methodology used in this dissertation. The discussions in the chapter include the data source and collection procedure, methodological criteria to extract the data tokens and to code them for the independent and dependent variables examined in this study.

Chapter 6 reports the results of the corpus analysis including the effects of the semantic features and of the discourse-pragmatic features on each L1 group’s variable use of the case morphemes.
Chapter 7 concludes the dissertation with some answers to the research questions sought in the corpus investigation and additional discussions on the limitations of the study and implications for future research.
CHAPTER 2

VARIABLE USE OF THE NOMINATIVE AND ACCUSATIVE CASE MORPHEMES IN KOREAN

Chapter 2 provides background information about variable use of the nominative and accusative case morphemes in Korean to lead to a better understanding of the research questions pursued in this dissertation. The organization of the chapter is as follows. Section 2.1 introduces the nominative and accusative case morphemes in Korean. Section 2.2 reviews previous research on the conditioning factors for variable case marking of subjects and objects, which include syntactic and phonological factors (section 2.2.1) and semantic and pragmatic factors (section 2.2.2). Section 2.3 looks into the phenomenon within a broader picture of the particle marking system in Korean through a discussion of other subject/object encoding options than nominative/accusative marking available in the language (section 2.3.1) and through an observation of nominative/accusative-marking of elements that are not deemed as a grammatical subject/object in a clause (section 2.3.2). Section 2.4 summarizes the chapter and states the focus of the present study.

2.1. Nominative and accusative case morphemes in Korean

Korean is an agglutinative language in which functional morphemes are attached to nominal expressions or predicate stems. Case on subject and object arguments is thus marked by nominative and accusative case particles, respectively:
(1) Nabee-\textbf{ka} kapang-\textbf{ul} manhi sa-ss-ta.

Nabee-NOM bag-ACC a lot buy-PST-DCL

'Nabee bought a lot of bags.'

The nominative morpheme -\textit{i/ka} and accusative morpheme -(l)\textit{ul}\footnote{When the final syllable of a DP ends with a consonant, nominative (NOM) case is marked with \textit{-i} and accusative (ACC) case is marked with \textit{-ul}. For DPs ending with a vowel, NOM case is marked with \textit{-ka} and ACC case is marked with \textit{-lul}.} are regarded as case markers whose primary function is to indicate the grammatical role of a DP argument that the particle marks. However, subject/object DPs may occur without these case morphemes, especially in informal speech. See (2).

(2) Nabee-\textbf{ø} kapang-\textbf{ø} manhi sa-ss-ta.

Nabee-ø bag-ø a lot buy-PST-DCL

'Nabee bought a lot of bags.'

Case drop is notably frequent in colloquial Korean, and in some contexts, overt case marking is considered less natural than case drop. Past research on the naturalness of case drop in Japanese and Korean has identified several groups of factors that constrain the distribution of case-marked and zero-marked subject/object forms. In the next section, I will introduce major factors that have been well-acknowledged in the literature on case drop in Japanese and Korean. I will also incorporate a brief discussion of previous proposals as to why each factor has the effects that it does on this particular morphosyntactic variation.
2.2. Conditioning factors for variable case marking of subjects and objects

2.2.1. Syntactic and phonological factors

2.2.1.1. Subject-object asymmetry

It is noted that the distribution of zero-marked subjects is more restricted than that of zero-marked objects (Ahn, 1999; Ahn & Cho, 2006, 2007; Hong, 1994; Kim, 2008; Lee & Kim, 2012). Using an extensive corpus of informal Korean conversation, Kim (2008) analyzed 3,900 overt subjects and 2,731 overt objects in terms of distinct morphological realizations: zero-marking, case marking, and other particle-marking (e.g., delimiting particle -to (‘also, too’), topic marker -(n)un, etc.). The results show that while 62% of the objects are zero-marked, only 32% of the subjects are zero-marked. On the other hand, 26% of the objects are marked by the accusative case morpheme, whereas 40% of the subjects are marked by the nominative morpheme.2

Several researchers have made efforts to formulate the subject-object asymmetry in (dis)preference for case drop as a structural requirement. It has been claimed in generative inquiries of Japanese and Korean syntax that a case morpheme is omissible only in a complement position (Ahn, 1999; Ahn & Cho, 2006, 2007; Saito, 1983). Ahn and Cho (2006, 2007) argue that in sentences maintaining the canonical word order, SOV, object DPs sitting in the canonical complement position may occur without the accusative morpheme, whereas subject DPs occupying SpecTP must be marked by the nominative morpheme. According to the analysis, a zero-marked subject in SOV sentences does not occupy the canonical subject

---

2 The rest of the object (28%) and subject (12%) tokens were marked by other particles.
position, SpecTP; rather, it is a left-dislocated nominal associated with a covert resumptive pronoun in the original subject position, as illustrated in (3):

(3) Mary, pro ku chayk ilk-ess-ni?

Mary the book read-PST-Q

'Did Mary read the book?' (Ahn & Cho, 2007: 119)

Ahn (1999) proposes that a left-dislocated nominal in a sentence-initial position receives a generalized theta-role “aboutness”, which in turn makes the nominal the topic of the sentence, rather than the grammatical subject of the sentence. Similarly, Kuno (1972) argues that a preverbal zero-marked nominal in Japanese is not a subject but a topic, and that the particle that undergoes deletion is not the nominative case marker -ga but the topic marker -wa.³

By dissociating a preverbal bare nominal from the canonical subject position and claiming that the nominal in question is in fact a topic, this approach seems to achieve a coherent syntactic explanation as to why zero-marked subjects are far less common than zero-marked objects. However, when it comes to investigating what constrains the actual (non-)use of the particle assumed to undergo deletion and its acquisition, it is unclear what insight can be gained through the analysis. Regardless of whether the missing particle is the nominative marker or the topic marker, what could be more apparent to the language (user’s and) learner’s perception is the fact

³ Although I acknowledge that there is no independent evidence that a missing particle in zero-marked arguments is always a case morpheme, following the majority of the existing literature, I will use the terms ‘case drop’, ‘zero-marking’, and ‘(case) unmarking’ all interchangeably. In addition, this study does not make any theoretical commitment to a specific stance with respect to whether zero-marking is instances of ‘deletion’ or ‘non-occurrences or absence’ (see Fujii & Ono, 2000; Lee & Thompson, 1989).
that there is a morphological specification that is absent sometimes but is present other times. Even if the approach successfully addressed some other signals to indicate a preverbal bare nominal’s topic-hood and rule out its subject-hood simultaneously (i.e., even if the learner were always successfully informed that the missing particle is the topic marker), the question of when to (not) use the marker remains unanswered, unfortunately.

In summary, zero-marked subjects display a more restricted distribution than zero-marked objects. In an attempt to account for the asymmetry, it has been proposed that case drop is allowed only in a complement position and the apparent zero-marking of subjects is not an instance of nominative case drop.

2.2.1.2. Word order

It is well-known that Korean is a head-final language with the unmarked word order, SOV. It is also accepted that Korean permits pre-predicate scrambling (Sohn, 1999): major constituents such as subject and object preceding the verb may be scrambled for emphasis or other discourse purposes (e.g., OSV).4

Scrambled sentences are claimed to involve a high processing load due to their marked status. Psycholinguistic studies find that scrambled sentences in both Japanese and Korean are more difficult to comprehend than canonical counterparts (Jackson, 2008; Miyamoto, 2002). Using a self-paced reading task, Jackson (2008) confirms that native speakers of Korean process OSV sentences more slowly than canonical SOV sentences. Since the default strategy to interpret

---

4 Although in informal speech, non-verbal entities can occur after the verb (e.g., SVO, OVS, etc.), the post-predicate non-verbal element is normally regarded as “after-thought” clarification rather than as an argument consisting of the planned utterance (Kuno, 1978; Sohn, 1999). For this reason, Kuno (1978) contends that post-predicate elements should not be counted for an examination of word order variation.
the first DP as subject and the second DP as direct object cannot be exploited in processing scrambled sentences, in order to identify the grammatical relations, the parser needs to use other linguistic cues. According to Carlson and Tanenhaus (1998), in many verb-final languages, sentence processing can be aided by identifying the properties of particles attached to DP arguments: the information of post-nominal particles facilitates identification of their argument roles in the sentence, which enables the parser to make an early commitment to a syntactic structure of the rest of the sentence before reaching the verb in sentence-final position.

The high processing cost for scrambled sentences leads to a hypothesis that both subjects and objects are more likely to take overt case marking when they appear in non-canonically ordered sentences than in canonical SOV sentences. Kim (2008) examined his corpus data of conversational Korean in terms of the interaction between subject/object forms (i.e., zero-marking, case marking, and other particle-marking) and word order variation (i.e., SOV vs. OSV). The data contained a total of 1,043 transitive clauses, which included 715 tokens of SOV sentence and 44 tokens of OSV sentence. Table 2.1 summarizes the distribution of the different subject forms observed in the SOV and OSV sentences.

(4) Table 2.1. Kim’s (2008) findings: Interaction of subject marking and word order

<table>
<thead>
<tr>
<th>Word order</th>
<th>Subject forms</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>DP-Ø</td>
</tr>
<tr>
<td>SOV</td>
<td>214 (30%)</td>
</tr>
<tr>
<td>OSV</td>
<td>11 (25%)</td>
</tr>
</tbody>
</table>
In SOV sentences, the three subject forms are distributed more evenly than in OSV sentences. Note that while only 36% of subjects in SOV sentences are marked by the nominative morpheme, more than a half of the subjects in OSV sentences are followed by the case morpheme. The result suggests that subject encoding types are in interaction with word ordering, and marked word ordering is more likely to invite overt nominative marking, presumably due to the high processing cost associated with it.

Object DPs in these two types of clauses were also analyzed according to their morphological realizations. Consider Table 2.2.

(5) Table 2.2. Kim’s (2008) findings: Interaction of object marking and word order

<table>
<thead>
<tr>
<th>Word order</th>
<th>Object forms</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>DP-ø</td>
<td>DP-ACC</td>
<td>DP-other</td>
<td>Total</td>
<td></td>
</tr>
<tr>
<td>SOV</td>
<td>457 (64%)</td>
<td>192 (27%)</td>
<td>66 (9%)</td>
<td>715 (100%)</td>
<td></td>
</tr>
<tr>
<td>OSV</td>
<td>22 (50%)</td>
<td>11 (25%)</td>
<td>11 (25%)</td>
<td>44 (100%)</td>
<td></td>
</tr>
</tbody>
</table>

Kim states that although the frequency of zero-marked objects is higher in SOV ordering (64%) than in OSV ordering (50%), the difference was not statistically significant. He points out that in both SOV and OSV orderings, zero-marking is the most dominant object encoding type, whereas case marking is a relatively uncommon option. Thus, unlike subjects, the choice of object marking types was not directly correlated with the word order variation. He proposes that the

---

5 It was reported that the 25% of other particle-marked objects in OSV clauses consist of DP-*(n)un (14%), DP-*to (‘also’) (7%), and DP-etc. (4%).
unexpected low occurrence of case-marked objects in OSV sentences could be ascribed to the relatively high frequency of case-marked subjects in the same clauses, considering that overt marking of either argument can suffice for argument identification. He adds that it could also be interrelated to the frequent use of topic particle -(n)un for the sentence-initial objects since they are typically viewed as a topic, which tends to be marked by -(n)un in the language.\(^6\)

By and large, given case particles are viewed as an effective means for argument identification, it is expected that non-canonical word ordering is likely to employ increased use of nominative and accusative particles. The hypothesis was borne out with respect to subject marking in Kim’s data of SOV vs. OSV clauses. Object marking, however, did not reveal a pattern in which case marking is more frequent in the non-canonical OSV ordering, possibly due to the effects of other interrelated factors on subject and object realizations in Korean.

### 2.2.1.3. Verb adjacency

It is claimed that verb adjacency is a significant predictor for accusative drop in Japanese and Korean. Within the framework of government-binding theory (Chomsky, 1981), Saito (1983) reanalyzes the adjacency requirement (i.e., a complement is assigned accusative case only when it is adjacent to the selecting verb) and proposes that the Japanese accusative case morpheme -o may be omitted only when the object argument immediately precedes the verb (cited in Fry, 2001: 109). A recent study of Korean case drop conducted by Yoon (2012) examines the proposal based on the tenets of the minimalist program (Chomsky, 1995, 2000). Her claim is that

---

\(^6\) In his data, 55% of other particle-marked object DPs are coded with -(n)un, whereas only 30% of other particle-marked subject DPs are followed by -(n)un. See section 2.3.1 for more information of the use of -(n)un in Korean.
overt case marking is one of the means for indicating ‘Phonetic Form (PF) visibility’ of arguments. While defining PF-visibility as a requirement that all arguments be overtly identified at PF to be visible to the sensorimotor system, she assumes that the requirement can be satisfied by one of three means: prosody, morphology, and proximity. It follows that an object argument at PF becomes visible to Logical Form (LF) by prosodic marking or by overt accusative marking or by proximity to the verb. Hence, once PF-visibility of the object is fulfilled by proximity to the verb, overt accusative marking is not necessary.

While the research within Chomskyan syntax has accounted for adjacency effects on accusative case drop in terms of a categorical requirement, other studies have presented empirical evidence for the effects as a probabilistic tendency. Their findings confirm that the accusative particles -o in Japanese and -(l)ul in Korean are more likely to be dropped when the object DP appears immediately adjacent to the selecting verb (Fry, 2001; Kim, 2008). For example, Fry (2001) conducts a corpus analysis of 2,540 tokens of direct object forms contained in his colloquial Japanese data. He reports that the rate of accusative drop is significantly higher when the object is immediately adjacent to the verb (59%) than when it is not (42%). A similar pattern is borne out in Kim’s (2008) conversational Korean data. Kim classified 2,731 object tokens by the number of intervening words between the object and the verb and observed the relative frequency of zero-marked objects. He finds that zero-marking is the most common object encoding type when the object is immediately followed by the verb (67%), and the rate of zero-marking gradually decreases as the number of intervening words goes up: 49% with one word, 42% with two words, and 30% with three words.
Kim proposes a processing-based account for this tendency. Granted that the unmarked word order in Korean is SOV, object arguments are more readily recognized when they appear within the vicinity of the verb, especially when they are immediately followed by the selecting verb. The easy argument identification would reduce processing load, and therefore there should be no further need for morphological specification of the grammatical function of the arguments.

2.2.1.4. Argument length and complexity

Another factor that has been claimed to condition case ellipsis patterns is the length and complexity of the DP argument that the particle marks.

Previous studies of Japanese case morphemes suggest that case drop for monosyllabic DPs is far more restricted than that for multisyllabic DPs (Ono, Thompson, & Suzuki, 2000; Tsutsui, 1984). Ono et al. (2000) contend that in order for a DP to be identified as an argument participating in the state-of-affairs denoted by the sentence, it must be considered to be ‘long enough’ in the language. Pointing out that monosyllabic arguments are uncommon in Japanese, they assume that very short DPs occurring in informal speech (i.e., one-syllable or one-mora DPs) are likely to fail to be identified as an argument, and this may create a need for morphological specifications of their grammatical roles as an argument. This prediction is borne out in the subject tokens of Fry’s (2001) colloquial Japanese data: the frequency of case drop for monosyllabic subjects (21%) is significantly lower than that for multisyllabic subjects (34%).

It is implied in Ono et al. (2000) that two moras may be the minimum length to be construed as being ‘long enough’ in Japanese. They provide interesting evidence that in the Kansai dialect, when one-mora DPs appear without a case morpheme, they are lengthened into two-mora DPs. See Ono et al. (2000) for more information.

Fry (2001) reports that the rate of case drop for monosyllabic objects (47%) was higher than that for multisyllabic objects (53%), but the difference is not statistically significant.
The tendency that monosyllabic DPs resist case drop is not bolstered by Korean data, however. The results from Kim’s (2008) analysis bear no such pattern in which zero-marking is disfavored and overt case marking is more dominant among the monosyllabic Korean DPs. Kim proposes that the cross-linguistic discrepancy has relevance to different frequencies of occurrence of monosyllabic arguments in the two languages: contrary to Japanese, monosyllabic argument DPs are common in Korean\(^9\), and therefore the potential difficulty in argument identification does not arise in Korean.

It has been noted that long complex DPs are more likely to take explicit case marking than short simple DPs (Kim, 2008; Ono et al., 2000; Tsutsui, 1984). In the Japanese data examined by Ono et al. (2000), complex DP subjects, which were defined as nominalizations or DPs with a long modifier, exhibit a strong tendency to be marked by the nominative morpheme -ga (65%), whereas the vast majority of non-complex DP subjects appear as bare DP (79%). Likewise, Kim (2008) claims that argument length measured by the number of modifiers of the pertinent DP is correlated with the frequency of case marking. Below in (6) is the relative frequency of case-marked and -unmarked subject/object forms, taken from his report and translated into graphs.

\(^9\) For example, he points out that while personal pronouns are mostly multisyllabic in Japanese (e.g., *watasi* (‘I’), *karae* (‘he’), etc.), they are mostly monosyllabic in Korean (e.g., *na* (‘I’), *ku* (‘he’), etc.).
As illustrated in Figure 2.1, for both subjects (S) and direct objects (DO), the frequency of case marking is higher when the DP involves more than one modifier. On the other hand, the zero-marking rates mostly display a pattern that is negatively proportional to the number of argument modifiers: they decline as the number of modifiers rises until three modifying phrases.

Kim ascribes the tendency that long complex DPs favor overt case marking to the high processing cost that these DPs often require. That is, “longer entities are informationally more loaded, hence processing load is assumed to be higher than shorter entities, resulting in the tendency for longer entities to be identified with an explicit particle” (Kim, 2008: 117).

2.2.2. Semantic and pragmatic factors

2.2.2.1. Animacy

Previous research suggests that the naturalness of case drop is predictable in terms of semantic
properties of subject/object DPs, one of which is the animacy features. The generalization is that omission of nominative and accusative case morphemes is more frequent for subjects high in animacy (i.e., human and animate subjects) and for objects low in the same featural dimension (i.e., inanimate objects) (Fry, 2001; Kim, 2008; Lee, 2006a, 2006b; Minashima, 2001).

According to Minashima’s (2001) corpus study on the Japanese accusative marker -o, the relative frequency of zero-marked forms is significantly higher among [-animate] objects (25.3%) than [+animate] objects (8.9%). The Korean corpus data analyzed by Lee (2006b) produces similar results in support of the animacy effects on both nominative and accusative case drop. Table 2.3 summarizes her findings.

Table 2.3. Lee’s (2006b) findings: Interaction of case marking and animacy

<table>
<thead>
<tr>
<th>Animacy</th>
<th>Subject forms</th>
<th></th>
<th>Object forms</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>DP-Ø</td>
<td>DP-NOM</td>
<td>Total</td>
<td>DP-Ø</td>
</tr>
<tr>
<td>Human</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DP-Ø</td>
<td>221</td>
<td>224</td>
<td>445</td>
<td>88</td>
</tr>
<tr>
<td></td>
<td>(52.7%)</td>
<td>(47.3%)</td>
<td>(100%)</td>
<td>(54.7%)</td>
</tr>
<tr>
<td>DP-NOM</td>
<td>224</td>
<td>126</td>
<td>194</td>
<td>35</td>
</tr>
<tr>
<td></td>
<td>(47.3%)</td>
<td>(64%)</td>
<td>(100%)</td>
<td>(50.7%)</td>
</tr>
<tr>
<td>Total</td>
<td>445</td>
<td>361</td>
<td>806</td>
<td>123</td>
</tr>
<tr>
<td>Inanimate</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DP-Ø</td>
<td>68</td>
<td>126</td>
<td>194</td>
<td>35</td>
</tr>
<tr>
<td></td>
<td>(36%)</td>
<td>(64%)</td>
<td>(100%)</td>
<td>(50.7%)</td>
</tr>
<tr>
<td>DP-NOM</td>
<td>126</td>
<td>170</td>
<td>236</td>
<td>203</td>
</tr>
<tr>
<td></td>
<td>(64%)</td>
<td>(72%)</td>
<td>(100%)</td>
<td>(74.9%)</td>
</tr>
<tr>
<td>Total</td>
<td>194</td>
<td>296</td>
<td>490</td>
<td>238</td>
</tr>
</tbody>
</table>

Lee reports that the frequency of nominative drop for human and animate subjects is significantly higher than that for inanimate subjects, whereas the accusative drop rate for inanimate objects is significantly higher than that for human and animate objects.

21
2.2.2.2. Definiteness

It has been observed that definiteness has similar effects to animacy on the variable case marking pattern in Japanese and Korean. That is, definite subjects are more likely to elide case morphemes, whereas definite objects display the exact opposite pattern (Fry, 2001; Kim, 2008; Lee, 2006a, 2006b; Minashima, 2001; Ono et al., 2000).

Minashima (2001) examines the object forms in his corpus data according to the definiteness features of the object DPs. He defines definite DPs as “a noun phrase that refers to an entity or a group of entities whose identity is presumably known to addressee” (p. 186) and includes (personal and demonstrative) pronouns, proper nouns, demonstrative-marked nouns, and bare common nouns whose definiteness status is determined by contextual information. The result is that while 53.4% of [-definite] objects are marked by the accusative -o, the overwhelming majority of the [+definite] objects (85%) are followed by the morpheme. Particularly noteworthy is that more than 99% of personal pronoun objects in the data appear with -o.

Lee’s (2006b) Korean data of nominative and accusative drop reveals a similar pattern. See Table 2.4.
(8) Table 2.4. Lee’s (2006b) findings: Interaction of case marking and definiteness\(^{10}\)

<table>
<thead>
<tr>
<th>Definiteness</th>
<th>Subject forms</th>
<th></th>
<th>Object forms</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>DP-Ø</td>
<td>DP-NOM</td>
<td>Total</td>
<td>DP-Ø</td>
</tr>
<tr>
<td>Pronoun</td>
<td>129 (47.6%)</td>
<td>142 (52.4%)</td>
<td>271 (100%)</td>
<td>24 (32.4%)</td>
</tr>
<tr>
<td>Name</td>
<td>36 (46.8%)</td>
<td>41 (53.2%)</td>
<td>77 (100%)</td>
<td>15 (42.9%)</td>
</tr>
<tr>
<td>Other</td>
<td>190 (38.3%)</td>
<td>306 (61.7%)</td>
<td>496 (100%)</td>
<td>287 (73.2%)</td>
</tr>
</tbody>
</table>

The frequency of nominative drop for subjects prominent in definiteness (47.6% for pronoun and 46.8% for name) is higher than that for subjects relatively less prominent in the features (38.3% for other). On the contrary, the accusative drop pattern is formed in the opposite direction: objects with a low value in definiteness (73.2% for other) elide case particles far more frequently than objects with a high value in definiteness (32.4% for pronoun and 42.9% for name).

2.2.2.3. Focus

Previous research on case ellipsis in Japanese and Korean points out that the naturalness of the variation between case marking and unmarking is highly influenced by the discourse-information

\(^{10}\) It is not entirely clear whether Lee’s classification forms adequate evidence of the definiteness effects on case (un)marking in Korean: instead of employing the binary distinction of the definiteness features, she labels ‘pronouns and names’ strongly definite and ‘other’ low definite. This can be considered as a methodological drawback since there are other types of definite DPs besides pronouns and names, and it is not reported how many demonstrative-marked DPs and definite bare DPs were counted into the category of ‘other’.
status of the pertinent DP argument, and it has been claimed that the notion of focus is one of the strongest factors that influence the variation.\textsuperscript{11}

Although the claims vary in their details, previous investigations of pragmatics of nominative and accusative case morphemes in both languages all suggest that omission of a case morpheme sounds far less natural when the DP argument it marks is interpreted as a focus than when the DP carries no such discourse-pragmatic function (Chae, 1999; Kim, 2008; Ko, 2000; Lee, 2006a; Masunaga, 1988; Yatabe, 1999, among others).

Indeed, it is not uncommon to come across literature in which case morphemes are analyzed as a discourse particle that indicates a focus. In a syntactic analysis that examined stacking of nominative and accusative case particles in Korean, for example, Schütze (2001) argues that stacked nominative and accusative particles are not a realization of structural case. He claims that nominative and accusative-marked focused DPs are adjoined to IP and VP, respectively, and that the particles are not a case morpheme at all, but a discourse particle that marks a focus.

Unfortunately, the studies investigating focus effects on the ellipsis phenomenon do not employ a single shared definition of focus, and in fact, not all these studies present a concrete characterization of what is meant by ‘when a DP is focused’. In the following discussion, I will review several of the relevant studies that operationalize the notion of focus in different means.\textsuperscript{12}

\textsuperscript{11} Using a forced-choice task, Lee (2006a) tested the effects of animacy, definiteness, and focus on accusative drop in Korean. Statistical analyses indicated that all three factors simultaneously exert independent influence on the variation of object forms, and further showed that the relative strength of the three predictors is ‘focus > animacy > definiteness’.

\textsuperscript{12} My goal here is to provide a comprehensive review of the past descriptive studies on the role of focus in case drop in Japanese and Korean, while acknowledging that their semantic and pragmatic definitions of focus woven into this review may be considered to lack theoretical uniformity.
Although indirect, one way to inspect the focus effects is to negatively define ‘focused DPs’ by establishing an operational characterization of ‘defocused DPs’. According to Masunaga (1988), a DP is putatively ‘defocused’ or ‘deemphasized’ when the sentence in which the DP occurs contains an element that is marked with an emphatic particle. She states that these defocused DPs may drop a case marker, proposing that occurrence of an emphatic particle in the sentence is a licensing condition for case drop in Japanese. This claim has been tested against Fry’s (2001) data. Fry labels a subject/object DP as defocused when some other item in the sentence is accompanied by an emphatic particle (e.g., -mo (‘even, too’), -dake (‘only’), -yo (sentence-final emphatic particle), etc.). He finds that for both subjects and objects, the case drop rate is significantly higher among defocused DPs than other DPs.

Focus can also be identified in terms of prosodic features. It is generally agreed that in many languages, focus is marked by increasing pitch range, duration, and intensity on focused elements and by lowering pitch range and intensity on post-focus elements (Flemming, 2008; Ardali & Xu, 2012). In Japanese, prosodic marking of focus employs a pitch variation in which focused elements exhibit a raised fundamental frequency ($F_0$) while post-focus elements involve $F_0$ reduction (Lee & Xu, 2012). Fry (2001) examines whether the status of focus, signaled by prosodic prominence, influences the case drop pattern in his Japanese data. In the analysis, a subject/object DP is coded as prosodically focused when the highest peak $F_0$ in the sentence falls onto the argument; otherwise, a DP is classified as not prosodically focused. The result is that for

\[13\] An example sentence from Masunaga (1988) is given in (9):

(9) John-ga/-ø sono hon-o/-ø sanda-mo yonda.
John-NOM/-ø the book-ACC/-ø three-at least read
'John read the book at least three times.' (Masunaga, 1988: 151)
neither subjects nor objects was there a statistically significant pattern in which focused DPs elide a case morpheme less often than other DPs. Thus, prosodic focus, at least the one measured by sentence-maximum peak $F_0$ values, turns out to have no effect on the ellipsis phenomenon.

Finally, there is a good deal of research on focus effects on case drop that directly grounds the investigation on a particular definition of the meaning of focus. The notion of focus implicit in much of previous literature on information structure can be defined as “the information in the sentence that is assumed by the speaker not to be shared by him and the hearer” (Jackendoff, 1972: 230). This broad understanding of focus has been expressed in different ways:

The focus is that portion of a proposition which cannot be taken for granted at the time of speech. It is the unpredictable or pragmatically non-recoverable element in an utterance (Lambrecht, 1994: 207).

What is focal is “new” information; not in the sense that it cannot have been previously mentioned, although it is often the case that it has not been, but in the sense that the speaker presents it as not being recoverable from the preceding discourse (Halliday, 1967: 204 cited in Lambrecht, 1994: 207).

One might still find these definitions to be a bit too general, but what is pertinent to this study is that the notion of focus claimed to resist case drop is a particular focus type, which could be termed as ‘contrastive focus’ (Ko, 2000; Lee, 2003; Lee, 2006a; Lee & Ramsey, 2000; Tsutsui, 1984; Yatabe, 1999). Furthermore, recent studies (e.g., Kim, 2008; Lee, 2010, 2011; Lee
& Choi, 2010) suggest that in order to precisely describe the distribution of different subject/object forms in Korean, researchers should count on a more fine-grained focus distinction that further distinguishes distinct subnotions of contrastive focus.

There are various definitions of contrastive focus proposed in the semantics and pragmatics literature. Moreover, different discourse studies on focus effects on case drop provide different characterizations of the notion of focus that they argue to constrain the phenomenon. Nevertheless, most of the well-acknowledged studies exploit (at least some of) the following conceptual components in defining the pertinent type of focus: ‘exclusive or exhaustive identification’ (Lee & Ramsey, 2000; Tsutsui, 1984), ‘presence of a set of alternatives’ (Ko, 2000; Yatabe, 1999), and ‘counter-presuppositionality’ (Lee, 2006a; Lee & Choi, 2010). How broad the denotation of the pertinent focus meaning claimed by each study is varies depending on what is exactly meant by these conceptual components and which component plays the primary role in their definition. I will explore this point as I review individual studies of discourse-pragmatics of case morphemes in Japanese and Korean.

Several authors propose that the kind of focus meaning influencing the ellipsis pattern involves ‘exclusivity’ or ‘exhaustive listing’ in its essence (Lee & Ramsey, 2000; Sohn, 1999; Tsutsui, 1984, among others). Below in (10) is an example showing that dropping the Japanese nominative morpheme -ga is not natural in a context where the subject argument evokes an exhaustive interpretation.
This type of focus reading can be better understood under a focus distinction proposed by Kiss (1998): ‘identificational focus’ vs. ‘information focus’. Kiss (1998) defines the function of identificational focus as denoting exhaustive identification on a set of individuals present in the domain of discourse (i.e., representing a subset of the set of contextually given elements for which the predicate phrases potentially hold). In contrast, information focus merely conveys new information that does not invite the kind of presupposed information required for exhaustive identification on the members of a given set.\textsuperscript{14} The claim made by Tsutsui and others can be summarized that case omission does not sound natural when the DP argument is deemed as identificational focus in Kiss’s characterization.

Yatabe (1999) contends that the Japanese nominative marker -\textit{ga} cannot be absent for a subject DP conveying a contrastive focus interpretation (i.e., when it is considered as “contrasting with some other object(s) of the same type” (Yatabe, 1999: 90)). However, his notion of contrastive focus diverges from Tsutsui’s generalization in that it does not require exhaustive interpretation. An example from the article is presented in (11).

\textsuperscript{14} Similar distinctions are proposed in different terms such as ‘neutral description’ vs. ‘exhaustive listing’ (Kuno, 1972), and ‘presentational focus’ vs. ‘contrastive focus’ (Drubig, 2000; Herring, 1990).
He writes that the subject of the second sentence, *Tarô to Jirô to Hanako* (‘Tarô, Jirô, and Hanako’) is construed as contrasting with the subject of the first sentence *Tarô*-da(-e) (‘only Tarô’). According to him, the second sentence does not mean that no one other than the three people uttered came, but that as opposed to the addressee’s belief, Jirô and Hanako, in addition to Tarô, came.

Yatabe’s rather informal characterization of contrastive focus is shared and further elaborated by Ko (2000), who adopts the notion of *kontrast* established by Vallduví and Vilkuna (1998). Vallduví and Vilkuna define *kontrast* as the ability of certain linguistic expressions to generate a set of alternatives (i.e., a subset of the elements of the corresponding type including the focused item and at least one additional element) as an additional denotation. Ko (2000) analyzes the Korean accusative case morpheme -(l)ul as a delimiter and claims it to be a *kontrastive* focus marker with an operational function of identification over a set of alternatives evoked by the -(l)ul-marked expression.

The most comprehensive empirical work on focus effects on case drop in Korean has been done by Lee (e.g., Lee, 2006a, 2010, 2011; Lee & Choi, 2010). Her claim is that whether or not...
the case morphemes are overtly realized is highly predictable in terms of a focus type distinction. Lee adopts an extended taxonomy of focus put forward by Dik et al. (1981), who divide focus into two types: ‘completive focus’ and ‘contrastive focus’. The function of completive focus is simply filling in a gap in the pragmatic knowledge between the speaker and the addressee, and a typical example of completive focus is realized as an answer to a wh-question. On the other hand, contrastive focus involves an apparent selection from a set of alternatives either explicitly stated or implicitly presupposed in the context. Using a forced-choice task, Lee (2006a) confirms that native Korean speakers’ choice of object forms is correlated with whether the object DP is [+contrastive focus] or [-contrastive focus] (e.g., completive focus): zero-marked object forms are far more common among the [-contrastive focus] objects (71%) than [+contrastive focus] (29%).

Furthermore, recent studies claim that what is truly responsible for the accusative drop pattern is not merely the binary opposition of [+/-contrastive focus] but a sub-distinction of contrastive focus: ‘selective’, ‘replacing’, ‘expanding’, ‘restricting’, and ‘parallel’ foci, originally proposed by Dik et al. (Lee, 2011; Lee & Choi, 2010). (12) has a summary of the essential meaning of each subtype along with an example presented in Dik et al. (1981: 62-67):

(12) Subdivision of contrastive focus types

a. **Selective focus**: the focus information selects one item from a presupposed set of alternative elements
A: Did John buy coffee or rice?

(presupposition: John bought x; x = coffee or x = rice)

B: He bought COFFEE.

b. Replacing focus: a specific item in the pragmatic knowledge of the addressee is eliminated and substituted by another, the correct item

A: John went to London.

B: No, he went to NEW YORK.

c. Expanding focus: the focus information is added to the previously given presupposed set

Presupposition of A: John bought x; x = coffee

B: Yes, but he also bought RICE.

d. Restricting focus: a previously given presupposed set is restricted to one or more correct elements

Presupposition of A: John bought x; x = coffee and rice

B: No, he only bought COFFEE.

e. Parallel focus: the speaker intends to contrast two pieces of information within one linguistic expression

A: I know that John and Peter bought a Volkswagen and a Toyota. But who bought what?

B: JOHN bought a TOYOTA, and PETER bought a VOLKSWAGEN.

Through an acceptability judgment experiment, Lee and Choi (2010) corroborate that
different subtypes of focus have distinct effects on the choice between case-marked and -unmarked object forms in Korean. In their study, the participants were asked to grade the naturalness of sentences with a case-marked or -unmarked object DP on a five-point rating scale. The sentences were presented in three types of focus contexts: replacing focus, selective focus, and completive focus. The average ratings for both case-marked and -unmarked objects in the three focus conditions are given in (13).

(13) Table 2.5. Lee and Choi’s (2010) findings: Mean judgments for each subtype of focus objects

<table>
<thead>
<tr>
<th>Object forms</th>
<th>[+contrastive]</th>
<th>[-contrastive]</th>
<th>Means</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Replacing</td>
<td>Selective</td>
<td>Completive</td>
</tr>
<tr>
<td>DP-ACC</td>
<td>3.82</td>
<td>2.30</td>
<td>3.17</td>
</tr>
<tr>
<td>DP-Ø</td>
<td>2.97</td>
<td>3.99</td>
<td>3.18</td>
</tr>
</tbody>
</table>

As shown in Table 2.5, accusative-marked DPs are judged to be the most natural when they occur in the replacing focus context, and they were considered the least natural when presented in the selective focus context. The mean judgments for zero-marked DPs reveal the exact opposite pattern: they are rated highest in the selective focus condition and lowest in the replacing focus condition. Noteworthy is that the average rating for case marking of selective focus ([+contrastive]) objects is even lower than that of completive focus ([−contrastive]) objects, which is a compelling piece of evidence against the prediction in the previous literature that contrastive focus DPs may all indiscriminately favor overt case marking. Overall, these findings
indicate that in order to precisely describe variable case marking in Korean, a more subtle distinction of focus types should be exploited.

According to Lee, the essential motivation for contrastive focus is “putting special emphasis on a piece of information judged to be already available to the addressee because that piece of information is opposed to the speaker’s own pragmatic information” (Lee, 2006a: 334), suggesting that ‘counter-presuppositionality’ (or ‘correctiveness’, in Dik et al.’s terms) is the most prototypical meaning of contrastive focus. Replacing focus better fits this characteristic of contrastive focus than selective focus. Dik et al. state that replacing focus involves a specific presupposition entertained by the addressee that is judged to be incorrect by the speaker, and the focus information makes a correction in the addressee’s knowledge state.\(^\text{15}\) Selective focus, on the other hand, does not serve a corrective function since the information entertained by the addressee is considered rather incomplete than incorrect (i.e., the presupposed information is underspecified).

Having reviewed the previous research on focus effects on case drop in Japanese and Korean, I conclude that the term ‘contrastive focus’ may not yield a maximally precise characterization of the claimed notion of focus associated with the phenomenon. Aside from the fact that there is no clear-cut definition of contrastive focus agreed upon in semantics and pragmatics literature, a problem in comparing the existing studies is that the notion of focus that they predict to disfavor case drop covers a rather wide range of focus meanings, in fact. Note that the one adopted in Ko’s (2000) proposal refers to a relatively broad category (i.e., analogous to

\(^{15}\) Within Dik et al.’s subdivision of contrastive focus, restricting and expanding foci are also regarded as corrective. In Korean, they are typically marked with a delimiting particle, the exclusive -만 (‘only, just’) and the concessive -또 (‘also, too’), respectively.
the definition of focus in Rooth’s alternative semantics (Rooth, 1992)) that subsumes completive focus within Dik et al.’s classification. On the contrary, the one attested to be the strongest predictor of variable accusative marking by Lee and Choi (2010) is quite a narrow category, a subtype of contrastive focus that bears relatively stronger prominence in contrastiveness (i.e., counter-presuppositionality or correctiveness). I will base this dissertation on the recent insights into more fine-grained sub-distinctions of contrastive focus in examining the claimed effects of focus on case variability in Korean.

2.2.2.4. Discourse newness

The term ‘newness’ has a twofold meaning: a piece of information can be novel in the sense that it instantiates the value of an underspecified element in a presupposed ‘open proposition’ shared between the interlocutors, but it can also be deemed new with respect to its referential status in the discourse (Birner & Ward, 2009; Prince, 1992). As an illustration, an English cleft sentence “It’s John I saw at the bar last night” can be divided into two parts: the presupposition = ‘I saw X at the bar last night’, and focus = John (Birner & Ward, 2009: 1181), where John is not necessarily previously evoked information in the discourse. Although there is a significant extent of correlation between these two categories, it is not the case that they can always be equated (Lambrecht, 1994). Acknowledging the distinction, I will refer to the former sense of newness as ‘focus’ and the latter as ‘(discourse) newness’.

Following Chafe (1976) and Prince (1992), discourse newness is understood as a dimension that measures how familiar the entity designated by a linguistic expression is assumed to be in the speaker/addressee’s mind. In other words, based on what has been activated in the previous...
discourse, the referent could be deemed more or less familiar with the addressee. Discourse newness is divided into two broad categories: ‘discourse-given’ vs. ‘discourse-new’. Chafe (1976) defines discourse-given information as knowledge which the speaker assumes to be in the consciousness of the addressee at the time of utterance, whereas discourse-new information is construed as what the speaker assumes that she is introducing into the addressee’s knowledge state by the utterance. Discourse-given information is associated with a referent that has previously been mentioned or at least inferable from the preceding discourse; discourse-new information is neither previously mentioned nor inferable from the preceding discourse (Prince, 1992).

Applying discourse newness to variable case marking in Korean, an accepted view is that subject/object DPs denoting new information disfavor case ellipsis (Kim, 2008; Ko, 2000; Lee & Thompson, 1989; Nam, 2000; Sohn, 1999).

The tendency that new information disfavors case ellipsis is more apparent for subject arguments than for object arguments. This becomes plain when we take into account the general characteristics of the two arguments: more often than not subject DPs denote a topic that is already introduced in the discourse (i.e., given information), whereas object DPs are less likely to convey given information. In this regard, it is often noted that a zero-marked subject is awkward when the nominal expression is used to introduce a new entity in a story for the first time (Ahn & Herschensohn, 2013), as shown in (14).
In line with the inappropriateness of omission of the nominative morpheme in the above context, Sohn (1999) presents a generalized statement that “The nominative particle -i/ka is frequently used to introduce new information as a topic or subject into a discourse” (p. 329).

2.2.2.5. Why do the factors have the effects that they do?

*Animacy and definiteness*

Taking the animacy and definiteness effects together, it should be pointed out that the statistical patterns for the Korean speakers’ choice between case-marked and -unmarked forms show a systematic correspondence to cross-linguistic patterns observed in split case marking phenomena.

Split case marking, in which a subject/object argument manifests overt morphological case marking only under certain conditions, has been accounted for in light of a syntactic markedness hierarchy in association with a semantic role hierarchy in which the person, animacy, and definiteness features are welded together. Example (15) provides a classic version of such a hierarchy as proposed in Silverstein’s (1976) discussion of split ergativity.
(15) Markedness hierarchy

Local person > Pronoun 3\textsuperscript{rd} > Proper Noun 3\textsuperscript{rd} > Human 3\textsuperscript{rd} > Animate 3\textsuperscript{rd} > Inanimate 3\textsuperscript{rd}

(Silverstein, 1976 cited in Aissen 1999: 674)

The generalization is that DPs towards the left of the ranking tend to take explicit case marking (i.e., accusative case) when they serve as object; on the contrary, DPs towards the right of the ranking tend to be overtly case-marked (i.e., ergative case) when they serve as subject.

In regard to the animacy and definiteness effects, it follows that in languages that adopt split ergativity based on definiteness (or pronominality) and animacy, grammars specify certain patterns for explicit vs. zero-marking for case as a categorical requirement. For example, in differential subject marking (e.g., in Dyirbal and other Australian languages), inanimate and indefinite subjects are explicitly case-marked while animate and definite subjects bear no morphological case (Aissen, 1999); on the other hand, in differential object marking (e.g., in Persian and Turkish), animate and definite objects are always marked by the accusative morpheme, whereas inanimate and indefinite objects are to be zero-marked (Comrie, 1989 cited in Hawkins, 2004: 47). The empirical evidence provided by Minashima (2001) and Lee (2006b) confirms that in languages like Japanese and Korean in which the distribution of overt vs. covert case marking is left as a matter of the speaker’s choice, the preferred pattern is shaped in the same direction as the categorical grammar specified in the markedness hierarchy for split case marking.

Looking further into the correspondence between the categorical and gradient patterns attested in the morphological case marking phenomena, one may ask why the features of
animacy and definiteness form the patterns the way they are, and not the other way round. According to J. Hawkins (2004), both categorical and gradient grammatical phenomena can be explained in terms of the same underlying processing principles. One of the processing principles relevant to the current discussion is the *Form Minimization* principle, which states that in order to maximize processing efficiency the human processor prefers to minimize the formal complexity of linguistic materials. He states that satisfying the principle in performance is aided by exploiting sentence-external and -internal cues, which he terms *processing enrichments*. With respect to nominative and accusative case assignments, one of the frequently exploited cues is stereotypic correlations between grammatical functions and thematic roles (i.e., correlations between subject/object-hood and agent/patient-hood). The main claim is that by virtue of being a prototypical agent, a nominative case assignment to animate and definite subjects can count less on an explicit morphological cue, whereas the nominative assignment to inanimate and indefinite subjects, which occur much less frequently and therefore are regarded as atypical subjects, relies more on an explicit formal specification. In the same manner, by virtue of being a stereotypical patient, inanimate and indefinite objects are more readily assigned accusative case without the aid of formal marking while animate and definite objects require a formal marking to signal its object role.

In a nutshell, the animacy and definiteness effects on the gradient case (un)marking pattern in Korean bears significant resemblance to the cross-linguistic morphological phenomena of split case marking. Both of the categorical and gradient linguistic patterns can be explained by the markedness hierarchy and a principle of processing efficiency.
**Focus and discourse newness**

Kim (2008) accounts for the effects of focus and discourse newness based on the cognitive complexity of linguistic expressions involving such discourse properties and increased processing difficulties associated with it.

With respect to the effects of discourse newness, he assumes that processing requires an activation cost, which is determined by an activation status of information at the time of utterance. The referent of a nominal expression that is familiar with the interlocutors is necessarily activated in their consciousness, involving a minimum activation cost. In this respect, given information is associated with a lessened processing load for referent identification, and therefore the need for an overt case marking of the argument is less critical. In contrast, new information involves increased processing complexity due to its greater activation cost, and this can be facilitated by an explicit morphological marking of the argument role.

He subsumes the focus effects under a similar explanation. An expression conveying contrastive focus tends to be conceptually more complex and informationally heavier, and the processing of such loaded information calls for identification of the special discourse function as well as the referent. This all may create an increased bias towards an explicit formal marking of the argument role because it can help the parser to allot more processing resources to the identification of the special discourse function.

Lee’s (2006a) account for the focus effects shares the same intuition. According to her, contrastive focus is not only new in that it serves as the value assigned to the variable in the pragmatic presupposition but also counter-presuppositional in that it contradicts a predicted alternative, and therefore informationally more complex than new information types of focus that
lack a counter-presuppositional meaning. Anchored in the principle of iconicity, which states that properties of the concept expressed by a linguistic sign may be reflected in structures of the linguistic expression, she proposes that due to the conceptual complexity involved in contrastive focus, contrastively focused expressions are more likely to receive more complex formal markings.

2.3. More on the particle marking system in Korean

I have addressed variable use of the nominative and accusative morphemes in Korean in terms of case marking and zero-marking of subject/object arguments. In this section, I delineate the morpheme use in connection with the particle marking system in Korean. Section 2.3.1 discusses other particle-marking of subject/object arguments to point to discourse-semantic criteria to distinguish case marking from delimiter marking. Section 2.3.2 introduces occurrences of the nominative/accusative case markers attached to constituents that are not regarded as a grammatical subject/object in a clause. This will serve as background information that leads to the inclusion criteria of the distinct types of data tokens counted for this study.

2.3.1. Other particle-marked subjects and objects

Delimiting particles can be attached to a DP to indicate the discourse function that the DP performs. Sohn (1999) reviews a classification of delimiters proposed by Yang (1972), who suggests three sets of delimiters that take distinct post-DP slots. See Table 2.6.
Table 2.6. Yang’s (1972) classification of delimiters in Korean (cited from Sohn, 1999)

<table>
<thead>
<tr>
<th>DP-</th>
<th>X slot</th>
<th>Y slot</th>
<th>Z slot</th>
</tr>
</thead>
<tbody>
<tr>
<td>-mace ('even, also')</td>
<td>-man ('only, just')</td>
<td>-(n)un ('as for, regarding')</td>
<td></td>
</tr>
<tr>
<td>-mata ('each, every, all')</td>
<td>-cocha ('even, as well')</td>
<td>-to ('also, too')</td>
<td></td>
</tr>
<tr>
<td>-kkaci ('till, up to, even')</td>
<td>-pakkey ('except for')</td>
<td>-(i)ya ('as only for')</td>
<td></td>
</tr>
<tr>
<td>-puthe ('starting from')</td>
<td></td>
<td>-(i)na ('or the line, or so')</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>-(i)lato ('even, even if')</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>-(i)yamallo ('the very, no other than')</td>
<td></td>
</tr>
</tbody>
</table>

A DP can be followed by a series of particles from the three sets in the order of X, Y, and Z, and the maximum number of particles that may appear from each set is one. This observation is demonstrated by (17).

(17) sip pen **kkaci-man-un** ppop-ca.

10 no. till[X]-only[Y]-TC[Z] select-PR

'Let's select only up to no. 10.' (Sohn, 1999: 346)

In particular, Z-delimiters display the same distribution as case morphemes. Thus, co-occurrence of a Z-delimiter and a case morpheme yields ungrammaticality, as shown in (18d).
(18) a. Nabee-nun ne-man salangha-y.
   Nabee-TC you-only[Y] love-DCL
   'Nabee loves only you.'

   Nabee-TC you-only[Y]-ACC love-DCL
   'Nabee loves only you.'

c. Nabee-nun ne-to salangha-y.
   Nabee-TC you-too[Z] love-DCL
   'Nabee loves you, as well.'

d. *Nabee-nun ne-to-lul salangha-y.
   Nabee-TC you-too[Z]-ACC love-DEC
   'Nabee loves you, as well.'

As mentioned in section 2.2.2.3, the nominative -i/ka and accusative -(l)ul are often analyzed
as a discourse particle, as opposed to a case marker (e.g., Ko, 2000; Schütze, 2001). This may
take us to a somewhat puzzling venue, considering that Korean employs productive use of those
delimiting particles to mark various discourse functions. According to Sohn (1999), “While case
particles mark syntactic relations among major constituents, delimiter particles delimit the
meaning of the co-occurring element with little syntactic function” (p. 345); still, he observes
that in constructions such as case stacking, the nominative and accusative morphemes “behave like delimiters with their own semantic content” (p. 345). Furthermore, under the minimalist understanding of Case (Chomsky, 1995, 2000), case features are assumed to be uninterpretable, and nominative and accusative case assignments are considered solely as a result of an Agree relation with (finite) T and v, respectively. Hence, associating a structural case with semantic and pragmatic meanings appears to involve a complex theoretical view with a broad picture of the case and topic/focus marking system in the language (Kramer, personal communication, April 2011). Although I acknowledge the importance of the issue, it is beyond the scope of this study. For this reason, I will provide only necessary background about use of delimiting particles in reference to that of case particles.

Among the members of Z-delimiters, the so-called topic (or topic-contrast) marker, -(n)un and its discourse function seem to deserve more attention. Despite the widespread use of -(n)un in the language, its meaning is not so clear-cut and well-defined. Simply put, -(n)un-marked DPs signal either topicality (‘as for’, ‘regarding’) or contrastiveness. Since it has been pointed out that the nominative and accusative morphemes often also mark a contrastive focused or referentially new element in the discourse, let us consider to what extent the meaning of -(n)un can be discerned from that of the case morphemes. For this comparison purpose, I will rely on the conceptual dimensions utilized in my prior discussion of the two discourse-pragmatic factors for case drop (i.e., contrastiveness and discourse newness), rather than ground the diagnosis on particular definitions of topic and focus.

Typically, the (thematic) topic reading of -(n)un-marking arises when it is attached to a sentence-initial subject with no special prosodic or contextual salience. A difference between -
(n)un vs. nominative markings can be revealed when it is considered that topic referents tend to be discourse-given information. As stated in section 2.2.2.4, one of the frequent functions of nominative marking is introducing a new entity into a discourse. On the other hand, -(n)un-marking for subjects denoting new indefinite information is regarded as inappropriate (Sohn, 1999). Hong (1985) also claims that in both written and spoken Korean, new information tends to be marked by nominative -i/ka, whereas old and inferable information tends to be marked by -(n)un. All in all, one may conclude that discourse newness status is one measure of discernment between the meaning of nominative marking and that of -(n)un topic marking.

Sentence-internal -(n)un-marked DPs predominantly receive a contrastive interpretation. Note that in (19), the first -(n)un-marked element, ne (‘you’) is the topic of the sentence, whereas nwun (‘eye’) bears a contrastive meaning.

(19) Ne-nun nwun-un khu-ta.

you-TC eye-TC big-DCL

'As for you, your eyes are big (in contrast with other body parts).'

(Sohn, 1999: 347)

Thus, an inevitable conclusion is that in Korean, elements carrying contrastive salience are marked not only by case particles but also by -(n)un, and this poses a challenge in distinguishing the discourse meanings achieved by the two formal specifications.

According to Lee (2003), the meanings of case marking and -(n)un-marking diverge in terms of the source or type of the contrastive reading related to each. Her proposal is that while case-
marked DPs are antecedced by a disjunctive question, -(n)un-marked DPs are antecedced by a conjunctive question that involves a scalar implicature. The claim that case marking involves a disjunctive question becomes a plain point as we recall the previous discussion of the focus effects on case marking: in a broad sense, the reviewed studies all associate explicit case marking with an operational function to single out a subset from a set of alternative elements of the equivalent type to the focused element (e.g., exclusivity/exhaustiveness). Lee (2003) argues that -(n)un-marking serves more than this disjunctive selection over an alternative set. Consider (20), an example provided in her discussion.

(20) A: ne-∅ ton-∅ iss-ni?
   you-∅ money-∅ be(have)-Q
   'Do you have money?'

B: na-∅ tongcen-un iss-e.
   I-∅ coin-TC be(have)-DCL
   'I have coins (but not bills).' (Lee, 2003: 2)

According to her analysis, (20A) can be understood as ‘do you have money or do you not have money?’ (disjunctive). In the kinds of worlds in which the respondent has money, (20A) can be conceived to imply ‘do you have (coins and) bills?’ (conjunctive), and this renders (20B) an appropriate answer to (20A). She points out that the set introduced by the conjunctive question, {coins, bills} is hierarchically ordered in the sense that bill is higher than coin on the pragmatic scale in neutral contexts and concludes that exploiting the -(n)un-marking, the
respondent provides an affirmative answer with concessive admission.\textsuperscript{16,17}

A similar disjunctive vs. conjunctive distinction is made by Kim (2008) and tested against his colloquial Korean data of subject and object encoding types. In the study, Kim associates case marking with ‘exhaustive listing’ and -(n)\textit{un}-marking with ‘parallel activities/states’ which takes a format of ‘A does/is X, but B does/is Y’. Below in Table 2.7 is the breakdown of different subject forms by the contrast subtypes.

(23) Table 2.7. Kim’s (2008) findings: Interaction of subject marking and contrast subtype

<table>
<thead>
<tr>
<th>Contrast subtype</th>
<th>Subject forms</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>DP-(\phi)</td>
<td>DP-NOM</td>
<td>DP-(n)\textit{un}</td>
<td>DP-other</td>
<td>Total</td>
</tr>
<tr>
<td>Exhaustive</td>
<td>3 (5%)</td>
<td>53 (95%)</td>
<td>0 (0%)</td>
<td>0 (0%)</td>
<td>56 (100%)</td>
</tr>
<tr>
<td>Parallel</td>
<td>20 (18%)</td>
<td>20 (18%)</td>
<td>70 (63%)</td>
<td>1 (1%)</td>
<td>111 (100%)</td>
</tr>
</tbody>
</table>

\textsuperscript{16} Lee (2003) discusses the following examples to support the presence of the scalar and concessive meaning arising from -(n)\textit{un}. In ordinary usage, (21) and (22) are considered an odd answer to (20A):

(21) ??na-\(\phi\) cicen-\textit{un} iss-e.
    I-\(\phi\) bill-TC be(have)-DCL
    ‘I have bills.’ (Lee, 2003: 3)

(22) ??na-\(\phi\) tongcen-\textit{un} iss-e, kuriko cicen-\textit{un}/-\textit{to} iss-e
    I-\(\phi\) coin-TC be(have)-DCL and bill-TC/-also be(have)-DCL
    ‘I have coins, and I have bills/also have bills.’ (Lee, 2003: 3)

\textsuperscript{17} I agree with Lee’s (2003) analysis in that replacing -(n)\textit{un} in (20B) with the nominative -\(i\) sounds odd (although the article does not have an explicit mention of this). The judgment is shared with five native speakers of Korean through casual consultations.
As seen above, the subjects exhibiting the exhaustive listing-type contrast are predominantly marked by the nominative particle (95%), whereas the majority of the parallel-structured contrastive subjects are followed by -(n)un (63%).

The analysis of the contrastive objects is consistent with the pattern of the contrastive subjects, at least partly. See Table 2.8.

(24) Table 2.8. Kim’s (2008) findings: Interaction of object marking and contrast subtype

<table>
<thead>
<tr>
<th>Contrast subtype</th>
<th>Object forms</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>DP-Ø</td>
</tr>
<tr>
<td>Exhaustive</td>
<td>1 (6%)</td>
</tr>
<tr>
<td>Parallel</td>
<td>16 (36%)</td>
</tr>
</tbody>
</table>

For the parallel-type contrastive object DPs, accusative marking and -(n)un-marking are distributed more evenly. However, when looking at the exhaustive-type contrastive objects, it is found that there is no instance of -(n)un-marking, whereas a significant portion of the objects (41%) is followed by the accusative morpheme.

To summarize, this section has presented a brief overview of delimiter vs. case markings of subjects/objects in Korean. Although the two argument marking types are apparently alike in that they both are sensitive to a discourse-pragmatic meanings that the DP expresses, they are distinguished in terms not only of their distributions but also of the particular discourse functions that they serve. Especially, -(n)un-marking and case marking, despite their shared distribution,
can be discriminated by discourse newness status and a subdivision of contrastiveness. The former marks topical DPs, which are typically discourse-given information, and DPs involving conjunctive parallel-structured contrastive meaning\(^{18}\); the latter frequently occurs with DPs conveying discourse-new information and exclusive disjunctive-type contrastiveness.

2.3.2. Nominative and accusative-marked non-subjects and non-objects

Korean is a language in which subject/object DPs are marked by nominative/accusative case morphemes, which may be omitted or replaced by delimiting particles. Hence, the central environment where variable use of the nominative/accusative morphemes (i.e., case marking vs. zero-marking) occurs is associated with the presence of subject/object arguments in a clause. However, Korean also permits various contexts in which the two case markers are optionally

---

\(^{18}\) Lee (2003) views -(n)\(u\)n-marked elements as ‘contrastive topic’ and case-marked elements as ‘contrastive focus’, which coincides with Repp (2010)’s characterization of the two concepts. In Repp’s terms, contrastive topics involve “two different sets of alternatives out of which elements are contrasted with each other” (p. 1338) and thus are typically realized in a parallel format; they differ from contrastive focus in that they bear neither exhaustiveness nor correctiveness. I suppress using the two terms for this study, however. This is because the primary objective of the dissertation is to precisely report the fine-grained discourse contexts relative to variable use of the nominative and accusative particles in the language without necessarily tying the descriptive facts to a particular formal definition of topic/focus. Still, I consider it very reasonable to view -(n)\(u\)n as a topic marker and case morphemes as a focus marker. As supported by (25), while case-marked DPs can be an answer to a wh-question, -(n)\(u\)n-marked DPs cannot, unless it appears in a negative answer to a positive wh-question (see Lee, 2003).

(25) A: nwu-ka col-ass-ni?
   who-NOM doze.off-PST-Q
   'Who dozed off?'
B: cey-ka/?*ce-nun col-ass-eyo
   I-NOM/I-TC doze.off-PST-DCL
   'I dozed off.'
B’: ce-nun an col-ass-eyo
   I-TC not doze.off-DCL
   'I didn’t doze off.' (Lee, 2003: 5)
attached to linguistic expressions that do not serve the corresponding subject/object argument roles in a clause. Thus, this section provides an overview of nominative/accusative marking of non-subjects/non-objects\textsuperscript{19} in Korean.\textsuperscript{20}

2.3.2.1. Adjectival and verbal predicates in negative sentences

Korean expresses sentential negation in two ways: one is short form negation, and the other is long form negation (Sohn, 1999). Short form negation is implemented by putting a negative adverb immediately before the predicate. Long form involves nominalization of the predicate by the nominalizer suffix -\textit{ci}, and the nominative or accusative morphemes may be attached to the -\textit{ci}-nominalized clause for emphasis (Sohn, 1999). The two types of negation of a sentence with an adjectival predicate are illustrated in (26).

\textsuperscript{19} In this dissertation, the term ‘object’ is used to denote the direct nominal objects of lexical transitive verbs of which the accusative marking is construed as a realization of structural case. All other types of objects are referred to as ‘non-object’. These include indirect objects of ditransitive verbs, verbal nouns selected by various light verbs, and complements selected by other types of (verbal) predicates (e.g., complements of adjectives, non-nominal clausal complements of saying/thinking verbs, and goal arguments of locomotive verbs).

\textsuperscript{20} My goal in this section is to provide a preliminary overview of a variety of the usage contexts that nominative/accusative marking of non-subjects/non-objects is observed in the language and facilitate understanding in the range of the data types counted for the present study. Since these non-subject/non-object data have mostly been neglected in previous literature on case ellipsis, the discussions may often lack adequate descriptions of the specific motivations for particle ellipsis variability pointed out in each subsection. They are in fact a question to which this dissertation hopes to draw researchers’ attention by introducing an incipient set of basic facts about case variability of non-subjects/non-objects to the literature.
(26) a. i cip-i khu-ta.
   this house-NOM large.PRES-DCL
   'This house is large.'

b. i cip-i an khu-ta.
   this house-NOM not large.PRES-DCL
   'This house is not large.' (short form negation)

c. i cip-i [khu-ci](-ka/-lul) an-h21-ta.
   this house-NOM large-NMLZ-NOM/-ACC not-be.PRES-DCL
   'This house is not large.' (long form negation)

Negation of a sentence with an intransitive verbal predicate follows the same pattern as noted in (27).

(27) a. i pay-ka ttu-n-ta.
   this boat-NOM float-house-PRES-DCL
   'This boat floats.'

21 In the long form negation, the negative adverb an ('not') is attached to the contracted form of a light verb ha ('do, be in the state of') (Sohn, 1999: 390). See section 2.3.2.6 for more discussion of the light verb usage of ha.
b. i pay-ka an ttu-n-ta.

this boat-NOM not float-PRES-DCL

'This boat does not float.' (short form negation)

c. i pay-ka [ttu-ci](ka/-lul) an-h-nun-ta.

this boat-NOM float-NMLZ-NOM/-ACC not-do-PRES-DCL

'This boat does not float.' (long form negation)

However, when the verb is transitive, the long form negation permits only the accusative particle, which is demonstrated by (28c).


Meghan-NOM food-ACC throw away-PRES-DCL

'Meghan throws away food.'

b. Meghan-i umsik-ul an peli-n-ta.

Meghan-NOM food-ACC not throw away-PRES-DCL

'Meghan does not throw away food.' (short form negation)

c. Meghan-i [umsik-ul peli-ci](lul/*-ka) an-h-nun-ta.

Meghan-NOM food-ACC throw away-NMLZ-ACC/-NOM not-do-PRES-DCL

'Meghan does not throw away food.' (long form negation)
2.3.2.2. Nominal complements of linking verbs

When the complement is a noun, sentential negation of the copula displays slight deviations from the general pattern: only a short form negation occurs, and the negative adverb is placed before the sentence-ending suffix. In this type of negation, the nominative particle is attached to the nominal complement that is negated, but omission of the particle does not yield ungrammaticality. See (29).

(29) a. Meghan-i haksayng i-ta.
    Meghan-NOM student COP-DCL
    'Meghan is a student.'

b. Meghan-i haksayng-i/-ø an-i-ta.
    Meghan-NOM student-NOM/-ø not-COP-DCL
    'Meghan is not a student.'

Nominative marking of nominal complements of linking verbs is not restricted to negative sentences. As seen in (30), nominal complements of toy (‘become’) in affirmative sentences are normally marked by the nominative case particle, but the particle may be dropped.

(30) Meghan-i yolisa-ka/-ø toy-ess-e.
    Meghan-NOM cook-NOM/-ø become-PST-DCL
    'Meghan became a cook.'
2.3.2.3. Multiple nominative constructions

Korean has multiple nominative constructions in which the nominative-marked sentence-initial DPs are understood as a sentential topic (Hong, 1990; Yoon, 1987, 2004, 2007). According to Yoon (2007), in the multiple nominative constructions, there is more than one subject-like element, the Major Subject, and the grammatical subject. A grammatical subject is the subject of the VP, an unsaturated predicate. A Major Subject is a subject on which the Sentential Predicate, the sentence consisting of the grammatical subject and VP are predicated. He proposes that a Major Subject occupies a position higher than the grammatical subject.

In Korean, Major subjects are frequently marked by the topic particle -(n)un or expressed as a bare nominal. Two examples are provided in (31) and (32).

(31) Meghan-i/-un/-ø son-i cak-ta.

Meghan-NOM/-TC/-ø hand-NOM small.PRES-DCL

'As for Meghan, her hands are small/ Meghan has small hands.'

(32) Meghan-i/-un/-ø pwumonim-i os-ul pha-n-ta.

Meghan-NOM/-TC/-ø parents-NOM clothes-ACC sell-PRES-DCL

'As for Meghan, her parents sell clothes.'

In (31) and (32), Meghan is not a subject but a topic in that it can be marked by the topic marker
The grammatical subject is *son* (‘hand’) and *pwomonim* (‘parents’) which are associated with the predicate *cak-ta* (‘be small’) and *pha-n-ta* (‘sells’), respectively.

The sentence-initial topics in the multiple nominative constructions often take a scene-setting element such as place and time adjuncts. Note that the nominative particle on *D.C.* in (35) and on *owel* (‘May’) in (36) may be replaced with the locative particle *-ey*, which shows that they do not perform a subject argument role.

(35) **D.C.-**<i>ka</i>/<i>-ey</i> nwun-i manhi o-n-ta.
    D.C.-NOM/-LOC snow-NOM a lot come-PRES-DCL
    'As for D.C., it snows a lot/ It snows a lot in D.C.'

(36) **owel-**<i>i</i>/<i>-ey</i> cangmi-ka hwalccak phi-n-ta.
    May-NOM/-LOC rose-NOM wide-open bloom-PRES-DCL
    'As for May, roses are in full bloom/ Roses are in full bloom in May.'

All in all, the first nominative-marked subject-like DPs in the multiple nominative constructions are not a subject but a topic. These instances of nominative marking of non-

---

22 The nominative marker can also be replaced with the genitive case marker *-uy*. See (33) and (34).

(33) **Meghan-**<i>uy</i> son-i cak-ta.
    Meghan-GEN hand-NOM small.PRES-DCL
    'Meghan’s hands are small.'

(34) **Meghan-**<i>uy</i> pwomonim-i os-ul pha-n-ta.
    Meghan-GEN parents-NOM clothes-ACC sell-PRES-DCL
    'Meghan’s parents sell clothes.'
subjects can be analyzed as an outcome of case alternation between nominative and other (case) particles triggered by a discourse-pragmatic consideration (i.e., topicality).

2.3.2.4. Alternation between accusative and other (case) particles

Likewise, accusative marking of some non-objects is often analyzed as an outcome of case alternation between accusative and other (case) particles. According to Sohn (1999), if emphasis for exclusiveness is present, the accusative -(l)ul may replace the locative marker on the goal of a locomotive verb in (37) and the dative marker on the indirect object of a ditransitive verb to yield a multiple accusative construction in (38).

(37) Meghan-i hakkyo-ey/-lul ka-ss-ta.
     Meghan-NOM school-LOC/-ACC go-PST-DCL
     'Meghan went to school.'

(38) na-nun ku ton-ul halapeci-kkey/-lul tuli-ess-ta.
     I-TC that money-ACC grandpa-DAT/-ACC give-PST-DCL
     'I gave the money to grandpa.' (Sohn, 1999: 330)

The emphatic meaning intended by the -(l)ul-marking in (37) is an exclusive designation of hakkyo (‘school’) as the goal argument of ka (‘go’), as opposed to somewhere else evoked in the discourse. Similarly, the -(l)ul-marking on halapeci (‘grandpa’) in (38) gives rise to an additional
meaning that the recipient of ku ton (‘that money’) is not someone else assumed to have received the money in the addressee’s belief.

2.3.2.5. Nominative and accusative stacking onto other (case) particles

It was discussed that the case particles may occur in a sequence of particles as far as the participating particles display distinct distributions (section 2.3.1). Consequently, it is not uncommon in Korean to come across the nominative or accusative case morpheme stacked onto some other case or delimiting particles in a variety of contexts. (39) and (40) provide an example of simple stacking and of multiple stacking, respectively.

(39) Meghan-i hakkyo-ey-lul ka-ss-ta.

Meghan-NOM school-LOC-ACC go-PST-DCL

'Meghan went to school.'

(40) John-i Swunhi-eyke-lul Yehghi-ey tayhayse-lul iyaki hay-ss-ta.23


'John talked to Swunhi about Yehghi.' (Schütze, 2001: 207)

Previous studies suggest that the case stacking phenomenon in Korean involves a discourse-related motivation although the specific claims vary. The aforementioned study, Schütze (2001),

23 Not all native speakers of Korean would find this sentence grammatical or acceptable, and I agree that the judgments of the sentence are less than clear. Yoon (2004) discusses the overall degraded acceptability for case stacking data, and Schütze himself mentions controversial judgments with several examples in the paper.
for example, analyzes stacked case morphemes as a focus marker and contends that they are not a case marker at all since they are optional and appear on non-direct objects that are not associated with the structural accusative case. According to his analysis, multiple accusative stacking in (40) is realized via multiple focus-adjunctions to VP to accommodate multiple foci in the sentence, *Swunhi-eykey* (‘to Swunhi’) and *Yenghi-ey tayhayse* (‘about Yeonghi’). To support the focus analysis, he also provides a nominative stacking example of (41) in which the nominative-stacked dative DP *Chelswu* is the answer to a wh-question.

(41) Q: nwukwu-eykey ton-i kulehkey manh-ni?
    who-DAT money-NOM so have.much-Q

'Who has such a lot of money?'

A: Chelswu-eykey-ka ton-i kulehkey manh-ta.
    Chelswu-DAT-NOM money-NOM so have.much-DCL

'Chelswu has such a lot of money.'

(Schütze, 2001: 203)

Yoon (2004) analyses nominative stacking on Major Subjects in the multiple nominative constructions (section 2.3.2.3). Although his view on nominative stacking is against Schütze’s focus treatment, he advocates the context-dependent nature of the interpretive conditions of nominative-stacked nominals and proposes that the nominative-stacked Major Subject must qualify as ‘news-worthy’ entities. As seen in (42a), a multiple nominative sentence is not considered felicitous when formed by stacking the nominative morpheme onto an entity that is
not novel enough to be ‘news-worthy’ to the interlocutor.\(^{24}\)


chair-LOC-NOM Cheli-NOM sit-PST-DCL

'It was in the chair that Cheli sat.'


Austin-LOC-NOM Bill-NOM study-ACC well do-PST-PERF-DCL

'It was while he was in Austin that Bill did well in his studies.'

(Yoon, 2004: 33, 35)

Sohn (1999) also indicates a discourse-related motivation for case stacking. He points out that if emphasis is intended, the goal-oriented particles can be followed by the nominative -ka by attaching the transferentive particle -ta(-ka) which in itself implies an optional use of -ka. See (44) for an example.

\(^{24}\) According to Yoon, (42a) is unacceptable for a non-specific chair is not novel enough to be considered as a ‘news-worthy’ entity by the interlocutors. Pointing out that the same structure is far more acceptable in a certain context that can make a specific chair a ‘news-worthy’ item, he provides an example in (43) uttered by a guest who is curious about prominent display of a specific chair in a restaurant.

(43) ?i uyca-ey-ka Kim taythonglyeng-i anc-usi-ess-ess-ta.

this chair-LOC-NOM Kim president-NOM sit-HON-PST-PERF-DCL

'It was in this chair that President Kim sat.'

See Yoon (2004) for more discussions.
(44) sensayngnim-\textbf{kkey-}(\textbf{ta(ka)}) ku chayk-ul ponay tuly-ess-e.

Teacher-DAT-transferentive-NOM that book-ACC send give-PST-DCL

'(I) sent the book to the teacher.'

(Sohn, 1999: 336)

To recapitulate, Korean permits case stacking, and nominative/accusative-stacking onto other (case) particle-marked elements is found in a variety of contexts. The presence of a discourse-pragmatic intent (e.g., focus, topicality, emphasis) has been proposed as a plausible motivation for the stacking phenomenon.

2.3.2.6. Accusative-marked verbal nouns in compound verbs

Optional accusative marking occurs within a unit of compound verbs. In Korean, verbal nouns are combined with a light verb $ha$ (‘do’) to produce the most predominant form of the so-called light verb constructions. These verbal nouns are often marked by the accusative case morpheme -\textit{(l)}ul.

(45) Meghan-i ecey pam-ey i swukcey-lul ha-ess-ta.

Meghan-NOM yesterday night-LOC this homework-ACC do-PST-DCL

'Meghan did this homework last night.'
a. Meghan-i ecey pam-ey kongpwu(-lul) ha-ess-ta.

Meghan-NOM yesterday night-LOC study-ACC do-PST-DCL

'Meghan studied last night.'

b. Meghan-i ecey pam-ey i kwamok-ul kongpwu(-lul) ha-ess-ta.

Meghan-NOM yesterday night-LOC this subject-ACC study-ACC do-PST-DCL

'Meghan studied this subject last night.'


Meghan-NOM yesterday night-LOC study-ACC do-PST-DCL

'Meghan studied last night.'

(45) is typically analyzed as a heavy usage of a regular verb ha, where ha is the source of information for the subcategorization and theta-role assignment (case assignment, as well) of the referential nominal object, i swukcey (‘this homework’). On the contrary, the ha in (46) combined with the verbal noun kongpwu (‘study’) has a light usage and constitutes a compound intransitive in (46a) and a compound transitive in (46b). It is a standard assumption that unlike regular heavy verbs, light verbs lack (full) ability to assign theta-role to their complements (Grimshaw & Mester, 1988). Accordingly, scholars that adopt the light vs. heavy ha distinction claim that what is responsible for the subcategorization and theta-role assignment is not the light verb ha but the verbal noun that it selects (Chae, 1996, 1997; Pak, 2001). Therefore, in (46b) i kawmok (‘this subject’) is assigned the patient theta-role by kongpwu (‘study’) and is identified as the object of the sentence (i.e., as the object of a lexical transitive verb); and the -(l)ul-marking on the nominal is regarded as a realization of structural case.25 The status of kongpwu (‘study’),

25 There is a controversy as to whether the case assignment to the complement is attributable to the verbal noun, as well. Some researchers argue that although verbal nouns are the source of subcategorization and theta-marking of the complement, they lack the case marking capability (e.g., Grimshaw & Mester, 1988); others propose that verbal nouns can case-mark as well as subcategorize/theta-mark the complement (e.g.,
on the other hand, is not the object of the sentence, and therefore the -(l)ul-marking is to be classified as accusative marking of non-objects.

Accusative marking of verbal nouns occurs with other kinds of compound verbs that incorporate non-\textit{ha}-light verbs, as well. For example, among the distinct types of the causative constructions in Korean\textsuperscript{26}, one type is termed lexical causative constructions or compound causative verbs. When the participating light verb is \textit{sikhi} (‘cause to do, make’), the verbal noun combined with it is marked by the accusative particle. Note the transitivity of the \textit{sikhi}-light verb sentence in (47b) in contrast to the intransitivity of the \textit{ha}-light verb sentence in (47a).

\begin{align*}
(47) & \quad \text{a. } \text{ilon-i} \ [\text{palcen(-ul)}] \text{ha]-n-ta.} \\
& \quad \text{this theory-NOM development-ACC do-PRES-DCL} \\
& \quad '\text{This theory develops.}' \\
& \quad \text{b. Meghan-i} \quad \text{ilon-ul} \ [\text{palcen(-ul)}] \text{sikhi]-ess-ta.} \\
& \quad \text{Meghan-NOM this theory-ACC develop-ACC cause to-PST-DCL} \\
& \quad '\text{Meghan developed this theory.}'
\end{align*}

The other context where accusative-marked verbal nouns appear is compound passive verbs. Korean employs various passivization processes broadly described in a three-way distinction:

\begin{footnotesize}
\begin{footnotes}
\footnotespace
\item Chae, 1996). Since the controversy is beyond the scope of this study, I will limit my discussion to identifying the nominal complement as an object and the verbal noun as a non-object in a clause.
\item According to Sohn (1999), there are two general types of causative constructions in Korean. One type is long form phrasal causatives in which -\textit{key ha(ty)} (‘cause to do/be’) occurs immediately after a predicate verb. The other is called short-form causatives, which are further divided into suffixal and lexical causatives. See Sohn (1999: 374) for the various suffixal and lexical items used for this type of causativization.
\end{footnotes}
\end{footnotesize}
lexical, morphological, and analytic or phrasal (Oshima, 2006; Sohn, 1999).\footnote{Morphological passives, which apply only to a subset of transitive verbs, are formed by attaching one of the four allomorphs of the passive suffix (-i, -hi, -li, and -ki) to a verb stem. Lexical passives include pure lexical passive verbs whose forms are entirely different from the active counterparts (e.g., mac (‘be hit’), chi/ttayli (‘hit’)) and compound passive verbs. Analytic or phrasal passives, which are the most productive, are formed with a verb stem followed by the infinitive suffix (-e, -a) and the inchoative auxiliary -ci. See Oshima (2006) and Sohn (1999) for more information about Korean passives.} One subdivision of the lexical passives is compound passive verbs in which a verbal noun is combined with pat (‘receive’), adversative tangha (‘suffer, undergo’), and toy (‘become’). When the participating verb is pat (‘receive’) or tangha (‘suffer, undergo’), the verbal noun is followed by the accusative morpheme. An example of each verb is provided along with the active counterpart of the ha-light verb sentence.

(48) a. Meghan-i Jordan-ul [conkyeng(-ul) ha]-n-ta.

Meghan-NOM Jordan-ACC respect-ACC do-PRES-DCL

'Meghan respects Jordan.'

b. Jordan-i Meghan-eykey [conkyeng(-ul) pat]-nun-ta.

Jordan-NOM Meghan-DAT respect-ACC receive-PRES-DCL

'Jordan receives respect from Meghan/ Jordan is respected by Meghan.'

(49) a. Meghan-i Jordan-ul [kwuta(-lul) ha]-ess-ta.

Meghan-NOM Jordan-ACC beating-ACC do-PST-DCL

'Meghan beat Jordan.'

Jordan-NOM Meghan-DAT beating-ACC suffer-PST-DCL

'Jordan suffered a beating by Meghan/ Jordan was beaten by Meghan.'

To summarize, I have described three compound verb constructions that incorporate different light verbs: *ha*-light verb constructions, lexical causative constructions with *sikhi*, and lexical passive constructions with *pat* and *tangha*. The verbal nouns that appear in these constructions are not an object of a clause, but they are optionally marked by the accusative case morpheme.

2.3.2.7. Nominative-marked verbal nouns in compound verbs

When a compound passive verb is in the form of a verbal noun followed *toy* (‘become’), it is the nominative morpheme that is attached to the verbal noun as shown in (50b).

(50) a. wuli kwun-un cek-ul [sasal(-ul) ha]-ess-ta.

our army-TC enemy-ACC killing-ACC do-PST-DCL.

'Our soldiers killed the enemy.'

b. cek-i wuli kwun-eykey [sasal(-i) toy]-ess-ta,

enemy-NOM our soldiers-DAT killing-NOM become-PST-DCL

'The enemy was killed by our soldiers.'

(Sohn, 1999: 372, revised)
2.3.2.8. Complements of psych-predicates

In Korean, psych-predicates (e.g., coh (‘like’), silh (‘dislike’), siph (‘want’)) are expressed with distinct syntactic environments: they appear as a transitive verbal predicate or as an intransitive adjectival predicate.

According to Jung (2011), when a psych-predicate occurs as a transitive verb in the form of the stem combined with the light verb ha, as in (51a), the complement is marked by the accusative case morpheme. In contrast, when a psych-predicate appears as an adjective in the stem form as in (51b), the complement is marked by the nominative case morpheme.

(51) a. Transitive (agent-NOM + theme-ACC)

\[ \text{nay-ka paym-ul/-ka} \quad \text{mwusew-e-\textbf{ha}-ess-ta.} \]

I-NOM snake-ACC/-NOM be.fearful-INF-do-PST-DCL

'I feared snakes.'

b. Intransitive (experiencer-DAT\textsuperscript{28} + theme-NOM)

\[ \text{na-ekey paym-i/-ul} \quad \text{mwusew-ess-ta.} \]

I-DAT snake-NOM/-ACC be.fearful-PST-DCL

'I was fearful of snakes.'

(Jung, 2011: 12)

\textsuperscript{28}The experiencer subject may also be marked by the nominative case morpheme.
An analogous behavior is observed in the case marking pattern with the matrix psych-predicate *siph* (‘want’), where the object of the embedded clause may be marked by the nominative particle when *siph* (‘want’) is realized as an intransitive predicate. Thus, in (52b) accusative marking of *halapeci* (‘grandpa’) may alternate with nominative marking.29

(52) a. Mary-ka *halapeci-lul/-ka* po-ko siph-e-**ha**-ess-ta.
Mary-NOM grandpa-ACC/-NOM see-to want-INF-do-PST-DCL
'Mary wanted to see her grandpa.'

b. Mary-ka *halapeci-lul/-ka* po-ko siph-ess-ta.
Mary-NOM grandpa-ACC/-NOM see-to want-PST-DCL
'Mary wanted to see her grandpa.'

(Jung, 2011: 5, 12)

All in all, the complex case marking pattern for psych-predicates in Korean displays mismatches between the nominative/accusative particle use and the grammatical role of the argument that the particle marks. That is, when the complement of a psych-predicate is marked by the accusative particle, it indicates the object function of the particle-marked argument (i.e.,

29 To account for the accusative and nominative alternation on the embedded object, Jung (2010) proposes that *siph* has a peculiar property that allows dual argument structures. According to her analysis, when *siph* takes an agent subject, it always selects for a vP\_DO complement, and the embedded object checks its accusative case in SpecvP. In contrast, when *ship* takes an experiencer subject, it is associated with a phonologically null vBE, which is realized as either a vP or a VP complement. When the complement is in the latter form that does not project a fully-fledged vP layer, the embedded object moves to SpecTP for nominative case checking. See Jung (2011) for more information.
complement of a transitive verb as in (51a), (52a), and (52b)). However, when the complement is marked by the nominative particle, it does not indicate that the particle-marked argument serves a subject role; rather, it is counted as an instance of nominative marking of non-subjects (i.e., complement of an adjective as in (51b) or of a transitive verb as in (52b)).

2.3.2.9. Accusative-marked nominals in intransitive action verbs

The other contexts in which Korean allows accusative marking of non-objects include cognate object constructions and denoting the distance or duration covered by an action (Sohn, 1999).

Cognate object constructions are traditionally defined as constructions in which an intransitive verb occurs with a nominal that appears to be a direct object produced by nominalization of the event or state denoted by the verb (Jones, 1988). In Korean, cognate objects may be followed by the accusative morpheme. Examples include chwum(-ul) chwu (‘dance a dance’), kkwum(-ul) kkwu (‘dream a dream’), and ‘kulim(-ul) kuli (‘paint a painting’).

The accusative case morpheme is often used to mark the duration or distance covered by the action of an intransitive verb. Consider (53) for examples.

(53) Meghan-i se sikan/sam mail(-ul) talli-ess-ta

Meghan-NOM three hour/three mile-ACC run-PST-DCL

‘Meghan ran three miles/or three hours.’

2.4. Summary and focus of the present study

Korean, as an agglutinative language, case-marks subject and object arguments with nominative
and accusative morphemes, respectively.

However, the case morphemes on subjects and objects may be dropped (i.e., variable case marking of subjects and objects). Previous research in Japanese and Korean linguistics points to diverse factors that constrain the naturalness of (non-)use of the case morphemes in marking subjects and objects. The effect of each factor that has been reviewed in section 2.2 is summarized in Table 2.9.

(54) Table 2.9. Conditioning factors for variable case marking of subjects and objects

| Factor                        | <-Zero-marking--| prefers--| Case marking-->
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Syntactic/phonological</td>
<td>Subject-object asymmetry</td>
<td>Object</td>
<td>Subject</td>
</tr>
<tr>
<td></td>
<td>Word order</td>
<td>Canonical order</td>
<td>Non-canonical order</td>
</tr>
<tr>
<td></td>
<td>Verb adjacency</td>
<td>Adjacent to the verb</td>
<td>Not adjacent to the verb</td>
</tr>
<tr>
<td></td>
<td>Argument length/complexity</td>
<td>Short simple</td>
<td>Long complex</td>
</tr>
<tr>
<td>Semantic/pragmatic</td>
<td>Animacy</td>
<td>Subject</td>
<td>High prominence</td>
</tr>
<tr>
<td></td>
<td>Object</td>
<td>Low prominence</td>
<td>High prominence</td>
</tr>
<tr>
<td>Definiteness</td>
<td>Subject</td>
<td>High prominence</td>
<td>Low prominence</td>
</tr>
<tr>
<td></td>
<td>Object</td>
<td>Low prominence</td>
<td>High prominence</td>
</tr>
<tr>
<td>(Contrastive) Focus</td>
<td>Not in focus</td>
<td>Focus</td>
<td></td>
</tr>
<tr>
<td>Discourse newness</td>
<td>Discourse-given</td>
<td>Discourse-new</td>
<td></td>
</tr>
</tbody>
</table>
In Korean, subjects/objects may also be marked by delimiting particles. Delimiter marking and case marking appear to be similar in that they both reflect a discourse function of the argument that the pertinent particle marks. Nevertheless, the two types of subject/object encoding can be discriminated by their distinct distributions. When a delimiter and the case particles display a shared distribution, they could be distinguished by a fine-grained distinction of the discourse-pragmatic meanings of the particle markings. In particular, the meaning of the topic/contrastive -(n)un-marking diverges from that of case marking in terms of discourse newness status and subtypes of contrastiveness.

The nominative/accusative morpheme use in Korean is not restricted to identifying a grammatical role of subject/object DPs since they occur with elements that do not serve the corresponding subject/object roles in a sentence (i.e., variable morpheme use for non-subjects and non-objects). The linguistic environments in which such morpheme use is observed include, but are not limited to, sentential negation, linking verb usages, multiple nominative constructions, case alternation often resulting in multiple nominative/accusative constructions, case stacking, different types of compound verbs, and psych-predicates.

This dissertation examines L2 acquisition of variable use of the nominative and accusative case morphemes by investigating whether the natural pattern of case drop predicted by the semantic and pragmatic factors is acquired by native English- or Japanese-speaking adult learners of Korean. More specifically, using naturalistic corpus data, the present study explores questions as to whether the L2 learners’ use of the case morphemes is sensitive (i) to the features that constitute the notion of prototypical subject/object expressed via the syntactic markedness hierarchy, which serves as the theoretical motivation for the observed animacy and definiteness
effects (section 2.2.2.5), and (ii) to the different discourse-pragmatic functions of the pertinent DPs such as discourse newness and focus (type) distinctions.

This corpus study is distinguished from the previous studies of case drop in Korean in that the range of data included enables the analysis to describe the phenomenon as variable behaviors of both subject/object case marking and case morpheme attachment to non-subjects/non-objects, while the past work disregarded the latter types of data. With respect to (non-)use of the accusative particle for verbal nouns in ha-light verb constructions, for example, Lee (2006b)’s corpus study excluded the verbal nouns from her annotations “in order to avoid unnecessary complications” (p. 74), stating “the use of the accusative case marker after the complement noun of the light verb ha is considered optional in all genres and styles of Korean speech and writing” (p. 74). Kim (2008), in fact, did not exclude such verbal nouns to conduct a separate analysis of the particle marking pattern for the complements of ha-verb. However, he does not base the analysis on a principled distinction of light ha vs. heavy ha and treated both the nominal complements and the verbal nouns equally as object DPs.30

By embracing the intriguing data of non-subjects/non-objects, the present study attempts to describe the variability phenomenon from a broader perspective of the actual morpheme use while providing a finer analysis of the data at the same time. A general prediction of the monolingual Korean speakers’ usage pattern of the case morphemes is that the effect of the semantic features posited for the markedness hierarchy (i.e., person, animacy, definiteness (or

---

30 The result is that the zero-marking rate of the collapsed objects of ha-verb is overall higher than that of the rest of the verbs in his corpus (i.e., non-ha-verb in his annotations). Although this finding may suggest that the involvement of the verbal nouns in the light verb usage would be responsible for the high zero-marking rate with ha-verb, it cannot be considered as conclusive evidence that in Korean verbal nouns predominantly occur with a zero particle, and “it still remains as a question when such verbal nouns are overtly marked with its particle and when they are not” (Kim, 2008: 252).
prominence in definiteness reflected in pronoun-/proper noun-hood, more precisely)) will be borne out only when the morpheme use is associated with subject/object arguments, whereas the presence or absence of the effect of the discourse-pragmatic factor will not be subject to this constraint (i.e., the effect will be borne out in both subjects/objects and non-subjects/non-objects).
CHAPTER 3

ACQUISITION OF VARIABLE CASE MARKING OF SUBJECTS AND OBJECTS IN KOREAN

This chapter reviews previous studies of the acquisition of variable case marking of subjects and objects in Korean. There are relatively few acquisition studies on this phenomenon, but recent research efforts in the acquisition of variable properties have produced a handful of empirical studies. By discussing the L1 and L2 acquisition studies (section 3.1 and 3.2, respectively), Chapter 3 aims to introduce the research stream to which this dissertation can make a contribution.

3.1. L1 acquisition

3.1.1. Background

Prior to the discussion of the L1 studies, this section presents an overview of the developmental path that Korean children undergo in the acquisition of the particle system in Korean and the linguistic features constraining variable use of the nominative and accusative case morphemes.

According to the child data corpus analyzed by Kim (1997), young children do not use case particles at the one-word stage, but when they enter the two-word stage, they begin to produce the nominative case marker -ka (between 1(year);8(months) and 2;0). The other nominative allomorph -i is found to be acquired much later due to lower frequency of input and perceptual salience. Around this time, child sentences are conservative with word order, displaying only the canonical SOV. As the nominative -ka emerges, various kinds of particles are coming along:
delimiting particle -to (‘also’), the comitative marker -lang/-hako (‘with’), and the locative -ey (‘in/at’). Several months later, the topic marker -(n)un emerges. The accusative marker -(l)ul appears much later than the nominative marker -ka (between 2;3 and 2;8), and some children overgeneralize -ka in contexts requiring the accusative particle for a while until they have acquired -(l)ul. Children also drop both nominative and accusative case particles, where the rate of accusative drop is much higher than that of nominative drop. Kim (1997) writes that the asymmetry could be attributed to differences in input frequency in her data since the frequency of accusative drop was higher in her adult data.

Turning to the acquisition of the linguistic features conditioning the naturalness of case drop in Japanese and Korean, empirical evidence suggests that children at the early stages of acquisition are sensitive to distinctions in definiteness, animacy and discourse-pragmatic status of DPs.

In Kim’s data, the children produced proper nouns at the very early acquisition stages. At around two years old, demonstrative pronouns emerged. A few months later, although they still showed noticeable preferences for use of proper nouns, personal pronouns appeared. The predominant use of personal pronouns was na (‘I’), but they also produced pronouns with different values in person feature such as ne (‘you’), yay/kyay (‘he’ or ‘she’) and wuli (‘we’). The fact that the data contained productive use of proper nouns and pronouns suggests that the children at the early developmental stages have acquired definiteness features of DPs.

Evidence that children acquire animacy features very early comes from a study on the acquisition of animacy agreement in Japanese.\textsuperscript{31} Using longitudinal corpora from three Japanese-

\textsuperscript{31} In Japanese, the locational verbs aru (inanimate) and iru (animate) agree in animacy with their nominative phrases. See Sugisaki (2007) for more information.
learning children aged between one and three, Sugisaki (2007) observed that the children showed correct animacy agreement patterns (with errors under 5%) from the earliest observable stages: 2;1, 1;1, and 1;5, respectively. Unless there is independent counterevidence, it seems to be reasonable to assume that Korean children aged between two and three can also tell the difference between animate and inanimate DPs.

As observed by Kim (1997), children begin to produce the topic marker -(n)un several months after they produce the nominative marker -ka. One interesting finding from her data is that children at this stage use -(n)un as a contrastive marker rather than as a neutral topic marker, typically in object naming contexts. Her account for the finding is that since children tend to simply drop sentential topics, they may not have many opportunities to use -(n)un as a neutral topic marker, and this resulted in the predominant use of -(n)un as a contrastive marker.

Concerning the pragmatic development of Korean-learning children, the finding can be taken as evidence that Korean children at the early acquisition stages are sensitive to at least some discourse-pragmatic properties such as contrastive salience that a DP bears in a given context.

To recapitulate, Korean children produce the nominative case morpheme very early as soon as they enter the two-word stage (between 1;8 and 2;0), whereas they begin to use the accusative case morpheme much later (between 2;3 and 2;8). Furthermore, the previous research suggests that children at the early acquisition stages are sensitive to distinctions in definiteness, animacy and discourse-pragmatic features. Therefore, it can be concluded that Korean children aged around three are expected to have acquired the grammatical functions of the nominative and accusative morphemes and have developed linguistic sensitivity to definiteness, animacy and
discourse-pragmatic features, which avail them of the prerequisites to demonstrate the natural patterns of variable use of the case morphemes in their native language.

3.1.2. Chung (2013)

Chung (2013) conducted an experimental study with monolingual Korean children to characterize when and how Korean-learning children acquire variable case marking of subjects and objects.

23 Korean-learning children, ages 5-7 (mean age: 5;4) and 20 monolingual adult controls participated in the study. Using an oral picture description task, the study tested whether the participants display sensitivity to the three conditioning factors for case (un)marking: focus, animacy and definiteness. The combination of the three independent variables yielded eight experimental conditions, and the distribution of case marking vs. zero-marking for each condition was observed.

It was found that for both subjects and objects, the children demonstrated adult-like performance in terms of overall frequency of case (un)marking within each condition and factor. Moreover, although the children omitted the case morphemes more frequently and demonstrated greater ambivalence in their responses, they often produced a more ‘adult-like’ pattern than the adult controls in that they exhibited sensitivity to all the three factors to a greater degree.

In the adults’ results, on the other hand, focus was shown to be the primary predictor of their case (un)marking patterns. For example, different from what was suggested in the previous research on case drop in Japanese and Korean, the effects of animacy and definiteness found in Lee’s (2006a) written task were not replicated in the results (e.g., for both subjects and objects,
zero-marking was more common for inanimate and definite DPs). Chung attributed the unexpected result to the possibly different effects of written vs. oral task modes.

Overall, Chung’s experimental study indicates that Korean children older than five years old have acquired how to integrate the multiple types of conditioning factors for variable case marking in an adult-like manner. Despite the important findings, the child data discussed in the study has no bearing on the development in the earlier acquisition stages for Korean children younger than five years old.

3.1.3. Chae (2017)

Chae (2017) conducted a corpus study to examine how the natural patterns of case (un)marking observed in adult grammar emerge in child grammar. The corpus data came from Korean children younger than those who participated in Chung’s experimental study so that the analysis could reveal an emerging grammar of variable case marking in monolingual children at the early acquisition stages. Using naturalistic child data, she investigated questions as to (i) at what stage of acquisition Korean children display the natural patterns of variable use of the nominative and accusative case morphemes and (ii) in what order children develop sensitivity to the three conditioning factors for the natural patterns (i.e., animacy, definiteness, and discourse focus status of subjects/objects).

The child data was collected from four Korean-learning children (ages: 2;7, 3;8, 4;7, 5;11). Spontaneous interactions between a child and the primary caregiver were audio-recorded and transcribed. All children were monolingual and male, and had been raised in Seoul or Gyeonggi province. The primary caregiver of each child were monolingual Korean parents. A total of
1,211 subject and object tokens extracted from the transcribed texts were coded for the three independent variables of animacy, definiteness and focus features, and the grouped and individual data were analyzed for the frequency of case (un)marking according to each independent variable.

It was found that while animacy effects were not borne out, definiteness had significant effects on the children’s use of the nominative case morpheme. The grouped data showed that the frequency of nominative drop for [+definite] subjects (57%) is significantly higher than that for [-definite] subjects (45.8%) ($\chi^2 = 5.62, p < .05$). The majority of the individual results demonstrated the same pattern, and one child (3;8)’s individual result reached statistical significance ($\chi^2 = 11.2 p < .05$).

With respect to focus effects, both grouped and individual data indicated that the children’s use of the nominative particle is influenced by the focus status of subject DPs. As illustrated in the figure below (Figure 3.1), the frequency of nominative drop was much lower in [+focus] subjects (26.8%) than in [-focus] subjects (65.4%) ($\chi^2 = 79.1, p < .001$).

![Figure 3.1](image)

(1) Figure 3.1. Chae’s (2017) findings: Effects of focus on frequency (%) of case (un)marking, subjects, grouped data
The group result was bolstered by all individual results of the children’s subject marking patterns. The individual results further disclosed an age effect on the relative frequency of case marking for focused subjects. That is, while the youngest child JH (2;7) tends to have focused subjects zero-marked, the three older children are apt to have them case-marked, as shown in Figure 3.2.

(2) Figure 3.2. Chae’s (2017) findings: Effects of focus on frequency (%) of case (un)marking. subjects. individual data

The children’s accusative (un)marking patterns were found to be influenced by the focus status of object DPs, as well. The frequency of case-unmarked objects was significantly lower in [+focus] (68.4%) than in [-focus] (88.5%) ($\chi^2 = 17.5$, $p < .001$). In terms of the individual data,
the three older children displayed a pattern observed in the grouped data, but only one child’s, result (TM (4;7)) was significant. Thus, it was concluded that the data supports the focus effects on the children’s use of the accusative case morpheme, but to a limited extent.

The study further examined the frequency of case (un)marking according to the breakdown by the subcategories of [+focus]: [contrastive focus], [completive focus] (i.e., answer to a wh-question), and [other]. The result is that the frequency of zero-marked subjects is lower both in [contrastive focus] (25%) and in [completive focus] (34.3%) than in [other] (65%) ($\chi^2 = 80.2$, p < .001). The same pattern was borne out in the object data: the rate of accusative drop is significantly lower in [contrastive focus] (65.9%) and in [completive focus] (76.9%) than in [other] (88.5%) ($\chi^2 = 18.2$, p < .001). More importantly, the findings show that for both subjects and objects, [contrastive focus] DPs are associated with zero-marking less frequently than [completive focus] DPs, exhibiting adult-like sensitivity to contrastive salience in young children’s emerging grammar of variable case marking.

Overall, Chae’s findings indicate that Korean children develop sensitivity to the three conditioning factors for case (un)marking in a more gradual fashion. Firstly, the sensitivity arises first for subjects and later for objects. Second, as early as age 2;7 children produce nominative (un)marking patterns predicted by the focus effects, and later when they reach age 3;8, their use of the nominative particle may reflect distinctions in definiteness, but animacy effects do not emerge at least until age 5;11.

Nevertheless, it should be pointed out that the study applied an apparent time interpretation to the cross-sectional data obtained from four children of distinct ages to examining developmental effects in child language acquisition (i.e., instead of using longitudinal data of
each child, the study was based on data of children at different ages collected for less than one month). This implies a gap between the nature of the questions investigated and the type of data used, and it can be considered as a methodological drawback that might have introduced possible confounds to the study.

3.1.4. Park-Johnson (2017)

Park-Johnson (2017) examined the development of variable case marking grammar by young heritage speakers of Korean. The study was based on longitudinal data obtained from three heritage Korean children for over two years during which the variable patterns were assumed to be developing. Using the child corpus data, she explored questions as to whether the young heritage Korean children employ case ellipsis; whether their use of case ellipsis changes over time; and whether canonicality of word order affects the ellipsis patterns.

Three children participated in the study (ages 3;7-5;11, 4;8-6;10, and 5;7-7;11). The children were born in the U.S. and regularly exposed to English around age 3;0. They speak Korean as the home heritage language with their parents who have strongly positive views of maintaining Korean. The data contained naturalistic interactions between a child and the parents and was collected in the children’s home at monthly intervals. The audio-recorded data was transcribed and coded for the presence or absence of the nominative/accusative case morphemes and for canonicality of word order (i.e., canonical (SOV, OV, SV) vs. non-canonical (OSV, SVO, OVS)).

It was found that the heritage Korean children employ case ellipsis, displaying asymmetrical performance in which overt subject marking is more prevalent (61%) than overt
object marking (17%). The children’s use of case ellipsis was reported to not change over time: subject case ellipsis is not correlated with age, but for one child (4;8-6;10), object case ellipsis is correlated with age. Based on the findings, she suggests that as young as 3;7, the heritage Korean children may already be stabilized in their acquisition of the case particles. In regard to the effects of word order variation, she reports that there is no difference between canonical vs. non-canonical word orders. The result showed that word order variation did not affect subject case ellipsis, but for one child (3;7-5;11), overt object marking was found to be more frequent (60%) for non-canonical order.

Although Park-Johnson’s corpus study provides an initial set of descriptive facts for heritage language learning of case (un)marking by young children, the findings do not seem to be conclusive. As acknowledged by the author, non-canonical word order is not used frequently among heritage learners of Korean. It should also be noted that among the three children who took part in the study, one child (4;8-6;10)’s use of the accusative particle changed significantly over time, and one child (3;7-5;11)’s use of the particle was shown to be higher for non-canonical order. More importantly, it is not clear throughout the study what is meant by “employment/use of case ellipsis” by young children when these children were assumed to be still acquiring the variable patterns (and the morphemes). Without further characterization, the interpretation of occurrences of case ellipsis in child data could be ambiguous between truly producing a variable pattern and merely exhibiting incomplete mastery of the morphemes.
3.2. L2 acquisition

3.2.1. Ahn & Herschensohn (2013)

Ahn and Herschensohn (2013) carried out a set of experiments to characterize L2 use of the nominative and accusative case morphemes by English native speakers. The study investigated whether the L2 speakers use the case morphemes correctly and how often they omit them.

The participants for the study were 26 beginning and intermediate L2 learners of Korean whose L1 was English. The average age of the learners was 22, and their initial exposure to Korean was not before college. 15 native speakers of Korean participated in the study as a control group. The tasks were composed of a written task, a translation task as well as two oral tasks, with a picture description task and a short answer task.

The results from the written task showed that the native speakers made no mistake on using the case morphemes, but they were found to omit them frequently (27% for nominative case and 49% for accusative case morphemes). In contrast, the L2 speakers produced incorrect use of either particle (8.9%), but they dropped the particles less frequently (8% for nominative and 4% for accusative particles).

The results from the oral tasks revealed the opposite patterns for both groups. Compared with their written performance, the native speakers omitted the case particles less frequently, but the L2 speakers omitted them more frequently. The rate of incorrect L2 use of the particles increased, as well. For the short answer task, in particular, the L2 speakers’ performance involved 12.6% of erroneous use of either particle. Also, the L2 speakers dropped the particles much more frequently (49.1% for nominative case and 53.2% for accusative case particles), whereas the L1 speakers demonstrated a decreased rate of case drop (8.7% for nominative and
19.3% for accusative particles).

Overall, the intermediate L2 learners exhibited a higher proficiency for using the correct morphemes and a lower rate of case omission in the written task than in the oral tasks. Although the L2 study was not designed to investigate the variable properties of the phenomenon, these findings provide significant insights into the nature of L2 learning of the case particles. By examining both written and spoken data, the study identified the central difficulty involved in the L2 learning with limited processing resources for the overt realization of surface morphology in online tasks, as opposed to defective knowledge of the syntactic feature.

3.2.2. Chung (2013)

Chung (2013) examined early vs. late L2 acquisition of variable case marking of subjects and objects in Korean. Using both written and spoken experimental data, she investigated whether heritage language learners (i.e., early L2ers) and adult L2 leaners (i.e., late L2ers) of Korean can achieve nativelike case drop patterns constrained by focus, definiteness, and animacy.

The participants for the study were 41 heritage language learners and 39 adult L2 learners of Korean. The heritage language learners were native speakers of English. They were exposed to Korean since birth with both parents speaking Korean as their L1. The L2 learners were also native speakers of English. Their mean age of initial exposure to Korean was 21.62 years old (range: 17-33), and the mean length of acquisition was 17.21 months (range: 2-36). In addition to the experimental groups, 43 native speaker controls participated in the study.

The written data was elicited from a forced-choice task employed in Lee (2006a), where the participants were asked to choose between case-marked and zero-marked subject/object DPs
presented in a short conversation. The spoken data was obtained from a picture description task. Both tasks were outfitted with eight experimental conditions from the combination of the three factors, and the frequency of case (un)marking was analyzed for each condition and factor.

The results revealed that the heritage learners’ performance was more nativelike than the L2 learners’. Both the native controls and heritage learners demonstrated gradient preferences significantly influenced by focus effects and overall showed similar results of relative frequency and interaction of factors. Still, certain quantitative differences were observed. Compared with the native speakers, the heritage speakers omitted the case morphemes more frequently and exhibited greater ambivalence in their responses. Also, while zero-marking of contrastive focus subject was almost categorically rejected by the native speakers, the heritage speakers’ judgements were shown to form gradient statistical dispreferences. For example, in the written task, the native speakers case-marked the overwhelming majority of the contrastive focus subjects (95%) while case marking only 56% of the contrastive focus objects. The heritage learners’ results displayed no such sharp asymmetry between subject and object markings: the frequency of case marking among contrastive focus DPs was 53% for objects and 62% for subjects.

The L2 learners’ performance diverged from the native speakers’ and the heritage learners’. Their subject marking showed a significant pattern that is distinct from the nativelike pattern; their object marking did not bear any systematic pattern, resulting in only chance performance. All in all, the L2 learners’ results were found to be qualitatively different from the two other groups’ in that they were primarily guided by the animacy and/or definiteness factors while displaying no nativelike sensitivity to focus effects. Consequently, they produced highly shifting
response patterns which were often opposite from those of the native speakers’ (e.g., they omitted the case particles more frequently for focused DPs than non-focused DPs).

Based on the imbalanced results between the two L2 groups, Chung concluded that heritage learners who begin to acquire Korean in early childhood have an overall advantage over adult L2 learners, which results in more nativelike outcomes. Ultimately, the study suggests that the linguistic knowledge attained in early vs. late L2 acquisitions may involve qualitative differences.

There are a few aspects of the study that remained unexplained, however. First of all, according to the proficiency test conducted on the three groups, the two L2 groups were described to be significantly different in their written proficiency. That is, the heritage language learners gained significantly higher scores (mean: 28.85) than the L2 learners (mean: 21.48). Combined with their relatively short length of acquisition (mean: 17.21 months), the L2 learners’ lower scores indicate that the imbalanced performance between the two L2 groups could actually be accounted for by the overall proficiency mismatch. In this regard, it does not seem to be entirely legitimate to attribute the heritage speakers’ better performance to early exposure to the language.

In addition, it was noted that the native controls’ data, often along with the heritage speakers’ data, did not reproduce the results of previous studies in terms of animacy and definiteness effects (e.g., Lee, 2006a). For example, zero-marking was more prevalent for inanimate subjects than for human subjects, and definiteness did not have a main effect on both subject and object marking patterns. Overall, the study reports that animacy and definiteness
effects were not found to be consistent across the different task modes as well as the argument types.

3.3. Summary

Previous research on L1 acquisition of the case morphemes in Korean indicates that Korean-learning children aged around three produce the nominative and accusative morphemes along with various kinds of other particles, and have developed a certain degree of sensitivity to the conditioning features for variable case marking (Kim, 1997). Chung’s (2013) experimental data showed that Korean children older than five years old are capable of integrating the conditioning features in an adult-like manner. Chae’s (2017) corpus data revealed the developmental path for children at the earlier acquisition stages, where children were found to demonstrate adult-like sensitivity to focus at age 2;7 and to definiteness at age 3;8. Park-Johnson (2017)’s longitudinal data is providing initial findings of heritage Korean children’s use of the case morphemes. In regard to L2 acquisition, Ahn and Herschensohn’s (2013) comparison between written and spoken L2 performance points out that the primary difficulty in L2 acquisition of the Korean case particles resides in overt realization of surface morphology. Finally, Chung’s (2013) experimental results have shown up uneven acquisition outcomes of the variability phenomenon between early and late acquisition conditions, where the former has an overall advantage over the latter. Despite the noted limitations of each study, these studies are making empirical contributions to this relatively underexplored topic in language acquisition.
CHAPTER 4

L2 ULTIMATE ATTAINMENT AND ACQUISITION AT INTERFACES

The goal of this dissertation is to characterize the attainment in L2 acquisition of variable use of the nominative and accusative case morphemes by native English- or Japanese-speaking adult learners of Korean and evaluate the predictions of different L2 theories. Thus, this chapter reviews different theories applicable to L2 acquisition of the case morphemes in Korean and relates them to the research questions investigated in this study.

The chapter is organized as follows. Section 4.1 reviews three different theoretical approaches to L2 ultimate attainment while identifying the learning tasks involved in L2 acquisition of the case morphemes in Korean. Section 4.2 further discusses the nature of the learning tasks (section 4.2.1), addresses the role of the L1 (section 4.2.2), and presents the research questions along with the predictions of the different L2 theories (section 4.2.3).

4.1. Theoretical approaches to L2 ultimate attainment

This section introduces three theoretical approaches to L2 ultimate attainment and compares their general predictions for the L2 acquisition data examined in this study. The comparison will be guided by an L2 learning scenario by L1 English learners of Korean to better explicate the kind of difficulties involved. The discussion will suggest that the coordination of the multiple conditioning factors required for the L2 learning constitutes an intriguing testing ground for the different L2 theories.
4.1.1. The Representational Deficit Hypothesis (RDH)

There is a general consensus that unlike child L1 acquisition, the endstate grammars of adult L2 acquisition do not invariably exhibit target-like convergence: individual learners reach ultimate attainment with varying degrees of non-convergence (Bley-Vroman, 1989), and the extents of persistence of non-target-like properties are not uniform in different domains of L2 grammars (Slabakova, 2009; Sorace & Filiaci, 2006, among others). Faced with this sharp asymmetry between the endstate characteristics of L1 and L2 grammars, a great deal of generative L2 acquisition research has proposed possible causes of the non-target-like outcomes in terms of the role of UG in late L2 acquisition along with L1 influence.

One theoretical stance in this tradition is the Representational Deficit Hypothesis (RDH) (Hawkins & Chan, 1997; Hawkins & Hattori, 2006; Hawkins & Liszka, 2003; Tsimpli & Dimitrakopoulou, 2007; Tsimpli & Mastropavlou, 2007). Proponents of this view assume that post-critical-period L2 learners have only partial access to UG, via features and parameter settings instantiated by the L1, and claim that non-nativelike outcomes are attributed to learners’ inability to project target-like mental representations for certain grammatical properties in the L2. That is, “divergence from native-speaker representations is an effect of the inaccessibility of features of functional categories in second language acquisition” (Hawkins & Chan, 1997: 187). Since maturation is considered to render the universal feature inventory unavailable to a post-childhood learner, what is crucial in determining (non-)convergence on the target grammar is whether the learner’s L1 selects the features required in the L2. Thus, formal features of functional categories which are instantiated in the L2 but were initially unavailable in the L1 are
doomed to be failed to be acquired (the *Failed Functional Features Hypothesis* by Hawkins & Chan, 1997).

An attenuated version of this view is known as the *Interpretability Hypothesis* (Hawkins & Hattori, 2006; Tsimpli & Dimitrakopoulou, 2007; Tsimpli & Mastropavlou, 2007), which restricts the inaccessible functional category features only to uninterpretable features that are not present in the L1. Guided by the minimalist distinction of feature interpretability (Chomsky, 1995), they claim that LF-uninterpretable features that are intrinsic to the language module and concerned exclusively with syntactic computation (e.g., Case, φ-features on verbs) are subject to critical period effects, and therefore inaccessible to late L2 learners unless they are also selected by the L1. On the other hand, LF-interpretable features (e.g., Q-features in interrogatives, φ-features on nouns) do not pose learnability problems due to their semantic import. For example, Tsimpli and Mastropavlou (2007) argue that “Any process of language development will have access to interpretable features, as their LF-related status ensures their mapping onto conceptual representations” (p. 144), and that learners can even benefit from “the ‘compensatory’ role of interpretable features in the analysis of L2 properties that involve uninterpretable feature clusters” (Tsimpli & Dimitrakopoulou, 2007: 218).

Relating the RDH to the acquisition of variable case marking by adult L1 English learners of L2 Korean, it seems unclear how it can account for the learning process. Assuming that both English and Korean equally select uninterpretable case features, the RDH does not predict any significant learning difficulty in their acquisition (i.e., the functional category features are fully accessible to the L2 learners). The apparent explanatory limitation of the RDH may be an unavoidable consequence of its underlying assumption in which the object of learning is
identified with functional categories and formal features within the narrow syntax per se. The research questions pursued in this study are eventually connected to inquiries concerning ultimate attainment in L2 acquisition, which encompasses a plain but important issue as to what is truly meant by complete mastery of L2 grammars. Regarding this question, this dissertation assumes that the object of learning goes far beyond narrow syntax properties and that the presence or absence of L2 knowledge ensuring complete mastery of the target grammar is to be attested in terms of target-like ‘use’, rather than only of target-like ‘(grammatical) representations’. This point will be clearer as we look into the specific learning tasks that English-speaking adult learners of Korean are required to handle in order to display nativelike use of the two case morphemes.

4.1.2. Where does the acquisition burden lie?

Adult second language learners are, in fact, confronted with substantive learning challenges in regard to the morphological realizations of the functional categories that the RDH seemingly does not address. Although the case systems in the L1 and the L2 share a uniform operation regarding feature selection and computation, when the valued case features in syntax are transferred to and spelled out at PF (Legate 2008), they diverge with respect to their PF-requirements. For example, English lacks rich formal inflectional specifications for various internal arguments (e.g., for direct objects, indirect objects/beneficiaries, oblique cases, etc.), whereas the Korean case system distinguishes morphological encodings for them. As for

32 Other learning tasks include figuring out the co-occurrence restrictions (section 2.3.1), the alternation of the two phonologically determined allomorphs of each case morpheme (section 2.1), and the specific morphological strategy to attach a bound morpheme, rather than to decline the DP itself (section 2.1).
nominative and accusative cases, while, in English, the formal distinction is manifested only in the personal pronoun paradigm, in Korean, it is realized with overt morphological exponents (i.e., marked by the case morphemes) that are required for all DP types. What appears to make their acquisition more taxing is the fact that the PF-requirement to indiscriminately mark all DPs is actually optional. The optionality of case marking is likely to pose a new kind of challenge to the learners in that the felicity conditions for a violation of the PF-requirement (i.e., appropriate omission of the nominative and accusative markers) often concern gradient rather than categorical judgments, yet are present in the input as a salient outcome of a complex interplay of multiple factors at different levels.

Consequently, the learners are most likely to face a double burden of acquisition which may have them exploit apparently contradictory learning strategies. On the one hand, they are required to extend the L1 morphological specification of case from pronouns to all nouns in the target grammar. On the other hand, they are supposed to suppress overt case marking in a manner that is constrained by the target grammar. Despite the obscured regularity in the input, can they attain target-like use of the case morphemes by obtaining the kind of L2 knowledge that leads them to appropriately suppress and express case marking? If they indeed can, how can generative theories of L2 acquisition account for the learning mechanism?

Let us consider whether such acquisition can be modeled under the Principles and Parameters (P & P) framework. According to the P & P theory (Chomsky 1981), the biologically determined innate characteristics of the human language faculty (i.e., UG) are defined as a set of principles that are invariant across languages (e.g., subjacency). Principles may involve binary valued parameters to be set either as [+ ] or [- ] (e.g., pro-drop parameter), and varied settings of
each parameter are considered as the source of cross-linguistic variation. The notion of
parametric variation has undergone a significant modification over the development of
generative syntax. Later in Minimalism (Chomsky, 1995, 2000), parametric variation parts with
the idea of ‘variation in principles’; instead, the locus of parametric options is moved to features,
the atomic units of a language bundled in lexical items. Accordingly, parametric variation is
understood as ‘variation in selection of features’ which implies variation in featural organizations
and properties, and the kinds of computation instructed by them (e.g., presence vs. absence of
formal features on functional categories, uninterpretable vs. interpretable features, or strong vs.
weak features). Whichever version is adopted, what the P & P approach implicates for language
acquisition is that learning the grammar of a particular language is equated to setting the
parametric values in the way that is specified for that language. In the case of L2 development,
the process is conceived as parameter ‘resetting’. That is, assuming that adult learners enter into
the starting point of L2 acquisition with parameter values fully-fixed to their L1 settings (L1
parameter transfer, e.g., White, 1985), “to go from the L1 to the L2, learners will often have to
reset existing parameters or reassign values to them” (Haegeman, 1988: 255 cited in Lardiere,
2007: 205) in accordance with the L2 input.

Despite the privileged status of the P & P theory in the field, it seems implausible to
conceive of the L2 learning in question as taking place in a parameter resetting manner. Under
the minimalist understanding of parameterization, it appears that the area which the theory
predicts about typological variation is primarily concerned with the selection of features and
their prescribed computation. In the L2 phenomenon investigated in this study, however, the
participating features (both uninterpretable and interpretable) are selected equally by the L1.
Particularly, the uninterpretable case features, as mentioned in section 4.1.1, undergo the same computational operation in both languages. That being the case, the P & P approach (and by extension, the RDH) would predict virtually no need for parameter resetting among our L2 learners. After all, since the learners’ central difficulty is attributed to cross-linguistic variation at a post-syntax level, namely, the different PF-realization patterns of case features, it is not clear that the parameter resetting view can say anything consequential to alleviate the acquisition burden.

It should be noted that the current discussion is concerned with a two-fold notion of case. The term case is used to refer to abstract features that drive syntactic computation, but it also indicates concrete actualizations of the abstract case features in a given language. The former is often called ‘structural case’ or ‘abstract case’ and the latter ‘morphological case’, and it is generally agreed that languages do not exhibit one-to-one mappings between these two conceptions of case (Legate, 2008; Wierzbicka, 1981). Granted that a crucial aspect of the L2 acquisition of case variability in Korean is linked to the latter understanding of case, I have concluded that the P & P based-account proposed by the RDH is unable to shed light on the L2 learning process. The key assumption to this conclusion was that the so-called ‘parameterized’ case features solely concern the former understanding of case, suggested by the fact that the RDH puts prominent weight on the LF-uninterpretability of case features.

One may challenge this assumption by invoking a logical possibility that the case features subject to parametric selection pertain to a finer dimension (and/or even multiple dimensions) of case and contend that the approach does predict a significant need for parameter resetting among our L2 learners. In other words, one may be tempted to describe the aforementioned PF-level
dissimilarities between English and Korean cases as typological variation triggered by distinct parameter settings of case features: both English and Korean select [-interpretable to LF case], but the two languages differ in terms of parametric selection operated on (presumably) another set of case features, which include but are not limited to [+/-pronoun only case], [+/-identify argument case], and [+/-attach morpheme case] parameters.

Exploiting such meticulous ‘microparameters’ (Kayne, 2005) could indeed capture all trivial and non-trivial cross-linguistic differences between English and Korean cases (and inform our L2 learners of an overwhelming number of parameters to reset). However, it then becomes apparent that the proliferation of the number of (case-related) parameters calls for drastic modifications in our current theoretical understanding concerning not only the notion of case but also features and parameters in general. If any descriptive detail of (case-related) typological variation is reduced to parametric values, the reduction has to be justified with a well-defined notion of certain features. Such notion of features should be reasonably broad enough to delimit the number of parameters for parametric values are supposedly constrained by UG. Otherwise, the seemingly enhanced degree of cross-linguistic comparability for case and case-related phenomena should be misleading in that what is reflected in the resulting parameters would actually be disjointed fragments of multifaceted realization of what was once called case (i.e., they are no longer related under any coherent notion including the very notion, case).

Since there seems to be no independent evidence that the parametric values conceived above are likely to be constrained by UG, I will maintain the assumption that the case features subject to parametric selection is abstract case. In regard to the source of cross-linguistic variation, this dissertation settles on a plain statement that the abstract case features are ‘realized differently’,
Another challenge with the L2 learning that is unlikely to be well-addressed within a P & P model lies in the gradient nature of the grammatical phenomenon. As discussed in Chapter 2, the choices between case marking and zero-marking typically give rise to gradient judgments among native speakers of Korean, whereas parametric choices are widely accepted as binary choices which normally render the resulting forms either grammatical or ungrammatical in a given language. Hence, the variability in the native grammars in itself seems to make it implausible to model the acquisition within the parameter (re)setting approach. Furthermore, the fact that the gradient pattern is produced by a dynamic interaction of multiple factors makes the parameter setting model appear too simplistic to adequately account for the acquisition of case drop. Given that the variation between case-marked and -unmarked forms is not subject to a preset formula for calculating particular feature combinations on DP, it is questionable how a parameter (re)setting model can describe the complex integration of multiple factors solely in terms of presence or absence of certain morphemes.

An inevitable conclusion may be that L2 acquisition of variable case marking in Korean cannot be handled by the RDH or any parameter resetting approach. To properly address the problem, it seems imperative to make a perspective shift in the primary assumption on what needs to be acquired in learning L2 grammars. Once what is to be acquired is simply limited to the functional categories and features associated with them, a learner’s failure to attain target-like use of the case morphemes is subject to a misanalysis of the state of her linguistic knowledge. For instance, if an advanced L1 English learner of L2 Korean invariably expresses the
nominative marking for pronoun subjects, the non-target-like L2 behavior might be erroneously attributed to her defective knowledge about the uninterpretable case features and/or interpretable definiteness features. This analysis is highly problematic not only in that it implies the exact opposite to what is predicted by the hypotheses but also in that it is unlikely to be a faithful representation of the learner’s knowledge state. For these reasons, this dissertation will be couched within alternative views in which the central learning is considered to occur in the acquisition of ‘featural relations’, rather than in the acquisition of ‘featural objects’.

4.1.3. The Feature Reassembly Hypothesis (FRH)

One approach that the present study can benefit from in this respect is known as the Feature Reassembly Hypothesis (FRH) (Lardiere, 2008, 2009). According to Lardiere, L2 learning essentially involves figuring out how to redistribute formal features clustered in the L1 lexical items into new configurations specified in the L2 and the conditions under which they are realized, and it is this reconfiguration or remapping task that poses the foremost grammatical difficulty among adult L2 learners. Embracing the Full Transfer/Full Access Hypothesis (Schwartz & Sprouse, 1996), she assumes that the initial stage of L2 acquisition is constituted by the entire categories and features fully assembled in the L1-specific way (Full Transfer), and that the reselection and reassembly of the features required for the L2 is an input-driven process that is still constrained by UG (Full Access). Therefore, although the FRH attaches significant weight to the role of the L1 as the most likely hindrance for L2 acquisition, unlike the RDH, it predicts that categories and features that are unavailable in the L1 or organized in distinct configurations from the L1 are all ultimately acquirable in principle.
Another aspect in which the FRH crucially differs from the RDH, which this dissertation is particularly anchored in, boils down to its assumption of the loci of cross-linguistic variation and the correct level of analysis for acquisition data. Lardiere (2008) argues that delimiting the source of typological variation to (parametric) selection of certain features, at best, leads to inadequate explanations because numerous cross-linguistic differences are observed to emerge not only from what features are ‘selected’ but also from how the selected features are ‘assembled’ in each language (e.g., both English and Chinese select the [+plural] feature, but it is not unitarily associated with particular interpretable features: while, in English, it is combined with [+/-definite] and [+/-human], in Chinese, it appears restrictively with [+definite] and [+human] (p. 15)). It follows that L2 features that are selected equally by the learner’s L1, regardless of the feature properties (e.g., interpretable vs. uninterpretable), still raises nontrivial acquisition challenges, presumably, in proportion to the extent of the L1-L2 dissimilarities in the feature configurations.

Lardiere (2008) writes that “Formal models of minimalist syntax apparently do not care whether features get spelled out suppletively vs. affixally or perhaps even overtly at all” (p. 25) and emphasizes that the domain in which the language-specific featural organizations are realized is at a distinct post-syntax level. Accordingly, a learner’s success or failure in acquisition of L2 categories and features should not be examined at the level of syntax alone; rather, variability observed in developmental and endstate L2 data must be analyzed at the level of PF (morphology or phonology), as well. Integral to this claim is that the kinds of linguistic knowledge which L2 learners need to develop subsume ‘morphological competence’ (Lardiere,
Achieving morphological competence requires figuring out the appropriate conditioning factors for how, including whether or not, particular features are expressed in a language-specific way, and these conditioning factors are linked to a wide range of distinct linguistic (and non-linguistic) components (i.e., phonological, morphological, syntactic, semantic, and pragmatic, etc.).

Under this view, persistent variability observed in endstate L2 grammars does not entail a learner’s shortfall of nativelike syntactic knowledge of certain features and categories; it can alternatively be ascribed to a paramount difficulty to acquire the interaction of participating features inherent to one or more domains. For example, in her case study that examines production of English past-tense marking by a L1 Chinese near-native speaker of L2 English, Lardiere (2007) contends that the learner’s omission of regular past-tense-marking inflection does not indicate an ineluctable failure to acquire the [+past] or [T] feature, given that the vast majority of the past-tense marking instances (93.85%) reflect correct suppliance of the morpheme. Instead, reporting that the omission rate is lower in the past-tense marking on irregular verbs and in the written production overall, she suggests that a significant aspect of the omission can be accounted for by different structures of phonological features in expressing inflectional morphology between Chinese and English.

Hwang and Lardiere’s (2013) work on the L2 acquisition of the Korean plural marker -tul by native speakers of English provides empirical evidence that L2 features selected equally by the L1 are still difficult to acquire if they are organized differently from the L1, but are eventually

---

33 A similar proposal is the Bottleneck Hypothesis (Slabakova, 2009), which claims the functional morphology as the ‘bottleneck’ of L2 acquisition. It further suggests that mastery of the inflectional morphology implicates successful acquisition in other domains including syntax, semantics, and pragmatics, whereas the FRH makes no such commitment to the relative facility in acquisition of semantic and pragmatic properties.
acquirable in principle. Although both English and Korean have plural morphemes, -s and -tul, the conditioning features for plural marking profoundly differ in the two languages. Appropriate use of -tul in Korean is subject to complex featural co-occurrence restrictions while English lacks the corresponding conditions for the use of -s. In particular, -tul-marking is associated with the [specific] feature on the DP and with the quantifier subtypes (i.e., -tul occurs only with [specific] DPs and with non-numeric quantifiers). When the quantifier is numeric, -tul-marking further interacts with the [human] feature, which allows pluralization of [+human] DPs only. Hence, native English speakers acquiring plural marking in Korean are required to reassemble these features in the Korean-specific way, demonstrating a complex understanding of the hierarchical feature configuration.

To investigate whether L1 English learners of Korean can acquire the various conditioning environments for -tul marking, Hwang and Lardiere carried out an experimental study. 77 native English-speaking adult learners of L2 Korean and 31 native Korean-speaking controls participated in the experiment consisting of five types tasks: an elicitation task, an acceptability judgment task, a preference task, a truth value judgment task, and a multiple-choice translation task. The L2 learners’ proficiency in Korean was measured by their scores on a Korean language proficiency test, and accordingly the learners were divided into four groups: low-intermediate, high-intermediate, low-advanced, and advanced.

It was found that the lower proficiency learners were not sensitive to the feature distinctions for plural marking in Korean, but as the proficiency level increases, the higher proficiency learners were able to use -tul according to the [specific] feature and to the quantifier subtypes, suggesting that these features are gradually acquired, and eventual target-like performance is
attainable in principle. However, the [+human] restriction within the numeral quantifier constructions was not fully integrated into -tul-marking even in the advanced learners’ performance.

Similar results were observed among native Indonesian speakers acquiring the Korean plural marking. Lee and Lardiere (2016) conducted a bidirectional study on the L2 acquisition of the Korean and Indonesian plural number marking. 70 native Indonesian-speaking adult learners of Korean at three different proficiency levels (low intermediate, high intermediate, advanced) completed an experiment employing three types of tasks: a sentence completion task, a grammaticality judgment task, and a multiple-choice task. The results revealed that L2 performance on the quantifier subtype distinction was enhanced with increasing proficiency, but the embedded [+human] restriction was not acquired even among the advanced participants.

The two studies on the L2 acquisition of pluralization in Korean indicate that the most deeply-embedded featural condition is the most difficult and the latest acquired, which is additional support for the FRH claim that what poses the foremost learning difficulty is concerned with featural organization, as opposed to featural selection (and to feature properties such as LF-(un)interpretability).

Let us look at what predictions the FRH makes about the L2 acquisition of variable use of the Korean case morphemes by L1 English learners of Korean. According to the hypothesis, the learners will first need to identify and select both the functional category feature [case] and the interpretable features conditioning its overt PF-realization (e.g., [animacy], [definiteness], [(contrastive) focus], and [discourse newness]). At the same time, they are supposed to figure out whether the overt marking is obligatory or optional, especially given the variability present in the
input. Then, they will need to reassemble these features in a restricted way to appropriately express and suppress the morpheme use. The overall process requires them to grasp the different conditions for overt PF-realization of [case] in the L1 and the L2, which are summarized in Table 4.1.

(1) Table 4.1. Conditions for overt PF-realization of [case]

<table>
<thead>
<tr>
<th>L1-L2 difference</th>
<th>English</th>
<th>Korean</th>
</tr>
</thead>
<tbody>
<tr>
<td>Category applied</td>
<td>Pronoun</td>
<td>All DP types</td>
</tr>
<tr>
<td>Obligatoriness</td>
<td>Obligatory</td>
<td>Optional</td>
</tr>
<tr>
<td>Conditioning features</td>
<td>None</td>
<td>[animacy]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>[definiteness]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>[(contrastive) focus]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>[discourse newness]</td>
</tr>
</tbody>
</table>

Table 4.1 suggests that even after the learners manage to remove the restriction on pronominality and resolve the optionality aspect, they still have to figure out how to constrain the pattern of optionality by associating [case] with [animacy], [definiteness], [(contrastive) focus], and [discourse newness] in the L2-specific way. Furthermore, assuming that the conditioning features are brought to the task of L2 acquisition already being fully-assembled in the L1-specific way, the process is likely to be a remarkably challenging task. Namely, wherever there is a different L1-specific way to express a corresponding feature in the L2, the learners will have to disregard the L1-specification and (un)pack and distribute the feature into the L2 functional
categories. For example, they are supposed to abstract the definiteness features from English articles or the discourse-pragmatic features from particular syntactic structures (e.g., *it*-cleft, expletive *there*-construction) and manage to re-spread them onto Korean case morphemes. In addition, considering that distinct effects of [(contrastive) focus] are found to be associated with different contrast subtypes (e.g., Lee & Choi, 2010; see section 2.2.2.3), the learners also have to comprehend how these subtypes are organized in a hierarchy of [(contrastive) focus] and integrate the particular subtype(s) with their use of the case morphemes. Overall, the FRH predicts that in spite of the L1 availability of the relevant L2 features, the demands involved in the task will be quite formidable, although not necessarily making the acquisition impossible.

Going one step further, the FRH may lead to specific predictions on varying degrees of the learners’ sensitivity to the different conditioning factors by determining the extent of discrepancies in L1-L2 configurations of each feature and that of the detectability of each feature in the L2 input. By way of illustration, providing it is agreed that in English and Korean, [focus] is encoded in a more dissimilar fashion than is [animacy], the learners will be prone to more difficulties in acquiring the focus effects on case marking than the animacy effects (a lapse in reassembly). On the contrary, in light of the relative strength of the two factors (i.e., [focus] > [animacy])[^11], it can be posited that the animacy effects are less prominent in the input than the focus effects. In this scenario, our learners may fail to even identify [animacy] as a conditioning factor and thus develop no sensitivity at all to its role in case marking (a lapse in selection).

In line with the FRH, the present study defines the primary task assigned to our L2 learners as remapping of the participating features from the L1-specific configurations onto the Korean-

[^11]: See footnote 11.
specific ones and will assume the kind of L2 knowledge required for the target-like case marker use as morphological competence.

4.1.4. The Interface Hypothesis (IH)

The term ‘interface’ is attached to the central postulation of generative linguistic theory that the linguistic structure consists of autonomous but interacting submodules such as phonology, syntax, and semantics (Carroll, 2001; Chomsky, 1995, 2000; Jackendoff, 2002, among others). Each submodule of the linguistic structure also interfaces with a non-linguistic module such as factual or encyclopedic knowledge and discourse context (e.g., syntax-pragmatics interface). Therefore, it can be posited that there are distinct types of interface, reflecting whether a particular linguistic component interacts with submodules of language (internal interfaces) or with non-linguistic cognitive modules (external interfaces) (Sorace & Serratrice, 2009).

Based on this fundamental tenet of linguistic modularity, recent studies on L2 ultimate attainment point to properties requiring the integration of various modules as the primary source of the persistent non-nativelike outcomes in endstate L2 grammars (Slabakova, 2009; Sorace & Filiaci, 2006, among others). In particular, the Interface Hypothesis (IH), proposed by Sorace, claims that narrow syntactic properties are all ultimately acquirable, despite the admitted developmental delays, whereas interface properties relating syntax and another cognitive domain may not be completely acquirable (Sorace & Filiaci, 2006). Sorace and Filiaci conducted an experiment on the interpretation of null vs. overt pronominal subjects in Italian by L1 English near-native speakers of Italian and native Italian speakers. The two types of subjects, contained in the subordinate clauses of bi-clausal sentences, were presented in both forward and backward
anaphora contexts. With the null pronouns, the two groups performed very similarly: they both predominantly identified a sentence-internal NP element (either the subject of the matrix clause or the complement of the subordinate clause) as the antecedent, in both the forward and backward anaphora contexts. With the overt pronouns, however, they displayed different response patterns, especially in the backward anaphora condition. That is, while the native controls showed a clear preference for the extralinguistic referent as the antecedent, the L2 learners still favored the reading in which the antecedent is the subject of the matrix clause. The authors interpret the results as an indication that the L2 speakers have attained the correct knowledge of the syntactic licensors for pronominal subjects in Italian, but they have not developed the nativelike pragmatic strategies for the appropriate use of null vs. overt subjects (e.g., minimizing referential ambiguity), exhibiting ‘residual indeterminacy or optionality’.

The dissertation is particularly interested in two separate, but related, central themes of the IH refined through its recent developments: the heavy focus on the vulnerability of the syntax-pragmatics interface in L2 acquisition (compared with other types of interface), and the prediction that optionality at this particular interface is likely to permanently characterize the highest possible near-native endstate of adult L2 acquisition (rather than merely to cause severe developmental delays) (Sorace, 2011, 2012). As reviewed by Sorace (2011), recent research on the IH has effaced the dichotomy between ‘full acquirability of purely syntactic properties’ and ‘less acquirability of interface properties’; instead, the emphasis has been moved onto a question

35 The instability of the syntax-discourse interface has also been attested in L1 attrition (Tsimpili, Sorace, Heycock, & Filiaci, 2004 cited in Sorace 2011: 3) and bilingual L1 acquisition (Sorace, Serratrice, Filiaci, & Baldo, 2009 cited in Sorace 2011: 4). That is, a similar pattern of residual optionality in L2 acquisition (e.g., overgeneralization of overt pronouns in null subject languages) is observed as emerging optionality at incipient stages of native language attrition in language contact situations, and as protracted indeterminacy among young bilingual children (Sorace, 2011; Sorace & Filiaci, 2006).
of whether properties connected to distinct types of interface lead to different degrees of non-convergence at the stage of L2 ultimate attainment. The latest IH claim made through this inquiry is that phenomena at internal interfaces (e.g., syntax-morphology, syntax-semantics) are relatively readily acquirable, whereas properties pertaining to external interfaces such as the syntax-discourse interface are more problematic and subject to almost permanent optionality in L2 grammars (Sorace, 2011; Sorace & Serratrice, 2009; Tsimili & Sorace, 2006).

The claim for successful acquisition at the syntax-semantics interface has gained a substantial amount of empirical support. For instance, Iverson and Rothman (2008) confirm that advanced English learners of L2 Portuguese have acquired a subtle semantic restriction on inflected infinitive complements of epistemic predicates in Portuguese. Their data, obtained from a truth value judgment task, demonstrates that the L2 learners’ interpretation of inflected infinitives is predominantly attached to a generic reading, forming an almost identical pattern with the native controls’. More evidence for ultimate attainment at the syntax-semantics interface is found in the acquisitions of the interpretive differences among French adjectives in variable positions by L1 English learners of L2 French (Anderson, 2008), and of object scrambling of definite DPs over negation in Dutch by L1 English learners of L2 Dutch (Unsworth, 2004).

Likewise, experimental L2 studies continue to produce results that bolster the claim of incomplete mastery of grammatical properties sensitive to discourse-pragmatic factors. Belletti, Bennati, and Sorace (2007) analyze the production and interpretation of Italian postverbal subjects and null and overt pronominal subjects by English near-native L2 speakers of Italian. The results from the L2 speakers illustrate overextension of the overt subjects and unconstrained use of the postverbal subjects, which indicates a failure to coordinate the two subject forms with
the relevant pragmatic features (i.e., topic shift and focus of new information, respectively). Tsimpli and Sorace (2006) also report that similar overuse of overt pronominal subjects is observed among advanced L2 learners of Greek whose L1 is Russian, a null-subject language, but with no such discourse-marking on the overt subjects as in Greek. Lastly, Lozano’s (2006) study on the acquisition of SV/VS word order alternations in Spanish reveals that Greek learners of Spanish at three different proficiency levels all display non-nativelike performance at the syntax-discourse interface which was marked by persistent deficits with encoding focused-subject constituents into VS order.

Nevertheless, there is also ongoing debate over the contemporary claims of the IH. A significant part of the debate concentrates on empirical findings contra the predictions, which suggest non-convergence at the syntax-semantics interface (e.g., Guijarro-Fuentes & Marinis, 2007), on the one hand, and convergence at the syntax-pragmatics interface (e.g., Rothman, 2009), on the other hand. Another noteworthy aspect of the discussion is concerned with determining whether particular properties at issue are regarded as external or internal interface phenomena (Slabakova & Ivanov, 2011).

Guijarro-Fuentes and Marinis’ (2007) experiment on the acquisition of the Spanish personal preposition a poses an interesting question for the promissory prediction on L2 attainment at the syntax-semantics interface. In Spanish, the distribution of a is governed by the specificity/animacy of the object, and the agentivity of the subject along with the telicity of the verb. The results from an acceptability judgment task show that the distribution of a by English L2 learners of Spanish at different stages of proficiency were not sensitive to these semantic features. The advanced learners, however, performed significantly better than the high and low
intermediate learners (still, significantly worse than the native controls) in one condition, [-animate], which was the least complex condition that did not involve more than one semantic feature (i.e., when the object DP is [-animate], a is disallowed regardless of [+/-specificity] of the DP), but in all other conditions in which they were required to coordinate more than one semantic feature (e.g., when the object DP is [+animate], felicitous use of a is determined further by [+/-specificity] of the DP), they did not exhibit any coherent tendency towards the presence or absence of a. The authors ascribed this discrepancy to different degrees of processing complexity and input opacity that may increase with the need to access multiple semantic factors influencing the distribution. Overall, the study suggests that syntax-semantic interface properties also cause noticeable difficulties even among advanced L2 speakers.

Experimental results indicate that target divergence at the syntax-pragmatics interface may not be a predestined endstate characteristic of L2 grammars. For example, Rothman (2009) examines the production and interpretation of Spanish null and overt pronominal subjects by English-native learners of L2 Spanish. According to the results, the intermediate participants have acquired the relevant syntactic properties, but they have not developed the nativelike sensitivity to the pragmatic criteria for the null vs. overt subject distribution (i.e., switch-reference, focus interpretations that encode an answer to a wh-question and contrastiveness). However, individual data from the advanced participants’ performance suggest target-like distribution constrained by the pragmatic features. Although these findings are consistent with the IH claim that the syntax-pragmatics interface is a prominently vulnerable area in late L2 acquisition, they strongly implicate that the difficulty associated with it is present merely as a
source of significant developmental challenges, not as a locus of permanent deficits in endstate L2 grammars.

As cautioned by Slabakova and Ivanov (2011), it may not be the case that the distinction of internal vs. external interfaces and its implications for adult L2 acquisition constitute a clear-cut contrast. Slabakova and Ivanov contend that “It is important to compare available data on external interface acquisition using common measures….and analyses. When subjected to a more careful scrutiny, the external (syntax-discourse) interface does not look so different from the internal interfaces, after all” (p. 650). Sorace argues that focus is a relational feature that designates new information of a sentence in relation to the topic and thus an LF-interpretable property within the language system, whereas topic is a referential feature, and referential givenness-newness is external to linguistic representations (Sorace, 2011, 2012; Tsimpli & Sorace, 2006). In contrast, Slabakova and her colleagues argue that both focus and topic are derived pragmatic concepts by combinations of the primitive features, [+/-anaphor] and [+/-contrast] (Slabakova & Ivanov, 2011; Slabakova, Rothman, Mendez, Campos, & Kempchinsky, 2011), and further maintain that “There is no principled reason why contrastive focus should be easier to acquire than topic shift” (Slabakova & Ivanov, 2011: 639). By stating that the debate “is not whether only one or both concepts involve pragmatic knowledge, but rather whether only one of both concepts involves coordination between linguistic and non-linguistic factors” (p. 212), Sorace’s (2012) response relates the issue to possible cause(s) for the claimed instability of the syntax-discourse interfaces.

It is discussed in Sorace (2011) and Sorace and Serratrice (2009) that there are probably multiple factors that contribute to the learnability issue at interfaces. According to the
representational account, the difficulty is located at the level of L2 speakers’ grammatical knowledge. More specifically, the L2 representations of the interpretable features conditioning certain interface structures may be underspecified because the L1 does not instantiate the corresponding interface settings. Another prime source of difficulty, identified by what is called the processing resources account, is that compared with monolingual native speakers, L2 speakers have limited processing resources insufficient for the coordination of syntactic knowledge with knowledge from other modules. Finally, the fact that L2 speakers often receive defective input, both in terms of quantity (i.e., reduced input) and quality (i.e., input from other L2 speakers or attrited L1 speakers), is also acknowledged as a contributor to non-target-like attainment. The IH suggests that quantity and quality of the input pose challenges for both internal and external interface acquisitions. On the other hand, the hypothesis contends that lack of the corresponding interface condition in the L1 is responsible solely for the difficulty at internal interfaces, and that “the syntax-discourse interface is not affected by language combination” (Sorace & Serratrice: 207) by pointing out optionality at the syntax-pragmatics interface observed among endstate L2 learners whose L1 instantiate similar interface conditions to the L2. Thus, among the three recognized variables, the processing inefficiency is claimed to be particularly associated with the incomplete acquisition at the level of the syntax-discourse interface. This follows from their assumption that properties at the syntax-discourse interface involve “a ‘higher’ level of language use, integrating properties of language and pragmatic processing” (Tsimpli & Sorace, 2006: 653), thereby involving greater processing complexity than mapping between formal properties of the grammar alone.
Investigating the L2 acquisition of variable use of the nominative and accusative case morphemes in Korean can shed light on the current theoretical discussions on the IH in that the acquisition requires mappings between various features connected to multiple interfaces. Evidently, the L2 acquisition is concerned with the syntax-morphology interface, as the learners’ syntactic knowledge of case is to be integrated into its variable morphological realizations. As to the conditioning factors, the acquisition requires the coordination at an internal interface, the syntax-semantics interface because the learners have to manage to choose between overt and covert markings of case in accordance with the semantic features contributing to the notion of prototypical subject-/object-hood. At the same time, the learners need to consider the discourse-pragmatic functions of the DPs in choosing between the two forms, which means that the integration also takes place at an external interface, the syntax-pragmatics interface.

The general prediction of the IH is that advanced L1 English learners of L2 Korean will display target-like knowledge in mapping case onto the internal-interfaced semantic features, whereas they will not exhibit the nativelike sensitivity to the external-interfaced discourse features. Overall, the fact that the learning tasks concern both internal and external interface acquisitions enables this investigation to provide empirical evidence with respect to which type of interface poses learnability problems and whether target-convergence at any particular interface is possible at all.

4.1.5. Summary

While identifying the core burden in L2 learning of variable use of the nominative and accusative case morphemes in Korean, section 4.1 has reviewed different theoretical approaches
applicable to the L2 acquisition. It has been discussed that parametric selection of features adopted in the RDH and parameter resetting approach cannot be an adequate explanatory tool to address the acquisition, and that the learning process can be well-captured under the FRH. Following the FRH, I have defined the major learning task assigned to the L2 learners as remapping of the relevant features and conditioning environments from the L1-specific configurations onto the Korean-specific ones. Finally, by pointing out the fact that the remapping takes place at distinct interfaces, I have suggested that the L2 acquisition can constitute a fruitful testing ground for the controversial predictions of the IH.

4.2. The present study

4.2.1. Acquisition at interfaces

I have identified the L2 acquisition addressed in this study as a learning task that takes place at multiple interfaces. In this section, I further describe the interfaced nature of the learning by discussing some properties involved in the acquisition in reference to the specific claims of the IH.

Concerning the acquisition at the syntax-semantics interface, one interesting aspect of variable use of the case particles in Korean is that the semantic features exert relatively weak effects on the distribution of case-marked and unmarked forms; it is rather the discourse-pragmatic factors that have greater effects on the distribution. If the general prediction of the IH is correct, in spite of the relative strength between the semantic and the pragmatic factors, the L2 speakers’ performance will demonstrate sensitivity to the semantic features subsumed under the markedness hierarchy (i.e., person, animacy, prominence in definiteness) while displaying

---

36 See footnote 11.
no (or less) sensitivity to the discourse-pragmatic features. If the prediction is indeed borne out, the result can be interpreted as evidence that fortifies the IH account for the underlying cause of residual optionality at the external interface. For example, it could be claimed that L2 learners have notably less processing difficulties in integrating syntactic knowledge with formal properties internal to the language module than with non-linguistic knowledge, which can override a high degree of fuzziness in the input.

As for the acquisition at the syntax-pragmatics interface, this study considers focus as a discourse-pragmatic factor, which complies with the classification adopted by the vast majority of previous literature on case drop in Japanese and Korean. In this respect, the prediction of the IH under examination is that the acquisition of the focus and discourse newness conditions will be more difficult or even impossible compared with that of the semantic features. Despite Sorace’s argument that focus is an LF-interpretable feature internal to the language system, this dissertation regards speakers’ sensitivity to the focus effects on case variability in Korean as an external interface phenomenon, thereby investigating the L2 acquisition of focus as an external interface acquisition.

The rationale behind this treatment of focus is as follows. According to the alternative semantics for focus, proposed by Rooth (1992), focus is an LF-represented property associated with a semantic function operated on a set of alternatives, but the semantic value of focus interacts with a pragmatic process, as well. Hence, it does not seem to be illegitimate to view focus as a property internal to the linguistic structure in the sense that focus assigns a value to a variable in a function represented at LF. However, the latter pragmatic aspect of focus implies an indispensable role of contexts in defining a particular focus meaning instantiated by the utterance
since what determines the domain of the alternative set and further defines particular types of focus is the discourse context to which the utterance is linked. For these reasons, it has been accepted as a theoretically sound approach to treat focus as a semantic property that essentially requires a pragmatic input (e.g., via presupposition, anaphoricity, scalar implicature, and question-answer pairs) (Portner, 2005).37

As pointed out in section 2.2.2.3, the focus effects on variable case marking in Korean are characterized by discerning behaviors among different types of focus, primarily concerned with a [+/--contrastive focus] distinction. What renders the semantic value of focus contrastive is most likely to be the discourse context that provides the members of an alternative set via an explicit mention or presupposition; when the context feeds no such information into the semantic value of focus, the value is realized as an information focus (provided that the preceding discourse contains a wh-question). It follows that in order for the L2 speakers of Korean to acquire the focus effects, they are required to process the contextual information and integrate the non-linguistic properties in a nativelike manner. Given that determining whether focus is a semantic or pragmatic property is beyond the scope of the dissertation, I will limit my discussion to the aspect of the debate reframed by Sorace (2012), which solely concerns whether focus involves coordination between linguistic and non-linguistic factors, and state that this study identifies the

37 I agree with the view that focus is not a clear-cut property, which involves a considerable overlapping between semantics and pragmatics. A similar understanding of focus is found in Jackendoff (2002). According to Jackendoff’s modular architecture of the language faculty, which consists of the phonological, syntactic, and semantic structures, the semantic structures are defined as the organization of thoughts expressed by language and described to interact with both propositional and informational properties, suggesting an intrinsic interrelation between semantics and pragmatics. Although these theoretical discussions of focus can pose a challenge against the validity of describing certain features (and by extension, their acquisition) in terms of the internal vs. external interface distinction, it is beyond the scope of this study, and I will leave the inquiry for future research.
L2 acquisition of the focus effects on variable case marking in Korean with an external interface acquisition.

Still, the study will bear in mind the controversial claim that focus is easier to acquire than referential newness. If Sorace and her colleagues are indeed correct, one should expect that the L2 speakers will display less target-like knowledge of case (un)marking conditioned by discourse newness, whereas they will produce a more target-like pattern (or perform significantly better, at least) with respect to focus. While the reviewed studies considered L2 (un)acquirability of these two features by investigating different properties independently conditioned by each (e.g., Tsimpli and Sorace’s (2006) study regarding object fronting by focus and null vs. overt pronouns by topic-shift in Greek) or by examining a phenomenon influenced by both, but without attempting to tease apart the acquisition of each feature (e.g., Rothman’s (2009) study with respect to null vs. overt pronouns by focus and switch-reference), the present study considers the (un)acquirability of the two features separately while analyzing both as conditioning factors for one single phenomenon. Hence, the result can establish more reliable evidence for or against the claim that focus is easier to acquire than discourse newness.

4.2.2. The role of the L1

It was noted that under both the FRH and the IH, the learning tasks involved in L2 acquisition of variable use of the case morphemes in Korean can be identified with remapping of the features into the Korean language-specific way from the L1, although it is apparent that the IH highlights the fact that the remapping occurs at distinct interfaces.
The two approaches crucially differ in two broad aspects, however. Recall that the IH predicts that L2 learners will have more difficulties in acquiring external-interfaced properties and even that those properties are subject to almost permanent optionality as an endstate characteristic of L2 grammars, whereas the FRH makes no commitment to predicting outcomes of L2 acquisition by such sharp distinction of interface types and does not deem certain properties ultimately unacquirable. This in turn provides the FRH adequate room to consider different sources of learning difficulty as the loci of non-nativelikeness. It was discussed that the approach is based upon the Full Transfer Hypothesis (Schwartz & Sprouse, 1996) and thus points to L1-L2 dissimilarities in reconfiguring morphosyntactic features as the primary source of difficulty (section 4.1.3). The IH, on the contrary, does not put significant weight on the role of the L1 in explaining a failure or success in L2 acquisition. As mentioned in section 4.1.4, the IH suggests that (un)availability of the corresponding interface condition in the L1 may account for the difficulty at internal interfaces, at best, but is not relevant to the difficulty at external interfaces (Sorace & Serratrice, 2009).

Considering the contrasting views on the role of the L1, it should be reiterated that the L2 speakers examined in this study belong to two different L1 groups, English and Japanese. While the IH does not anticipate notably varying levels of L2 performance to reflect the different L1-L2 pairings, the FRH predicts that the production pattern by an L2 speaker will be more nativelike if the case-relevant features are organized in a similar fashion to the Korean-specific configuration in her L1. It was described in section 4.1.2 and 4.1.3 that morphological realizations of case in

---

38 L2 acquisition at other types of interface (e.g., the syntax-morphology interface) is reported to be equally problematic to acquisition at the syntax-discourse interface. See White (2011) for a review of L2 studies that pose a challenge to the IH stance that the inherent distinction of interface types can account for defective vs. non-defective L2 performance.
English exhibit notable dissimilarities from those in Korean and the features that condition the naturalness of case (un)marking in Korean are organized in different configurations in English. Conversely, as implied in the review of the conditioning factors for case drop (section 2.2), the case marker use in Japanese bears great resemblance to that in Korean: case drop is frequent in Japanese and is similarly influenced by the semantic and discourse-pragmatic features examined in this study. Furthermore, Japanese has the delimiting particle -wa, which corresponds to the topic particle -(n)un in Korean: -wa cannot co-occur with Japanese case particles, and the discourse function it serves is known to be similar to -(n)un in that the meanings of -wa diverge between thematic and contrast markings (Fry, 2001; Kuno, 1972, 1978). Finally, Japanese also has case-markable verbal nouns that appear in compound verbs: it is claimed that Japanese has a light verb suru (‘do’), which is analogous to ha in Korean (Grimshaw & Mester, 1988; Pak, 2001). Guided by these conspicuous similarities between Japanese and Korean, the FRH predicts significantly better L2 performance among the Japanese native speakers than the English native speakers.

4.2.3. Research questions and predictions

This section presents the research questions along with a brief summary of the predictions of the three theoretical approaches.

The primary object of this dissertation is to examine the attainment in L2 acquisition of variable use of the nominative and accusative case morphemes in Korean. For this purpose, the following research questions are investigated:
**Research Question 1: Testing the markedness factor**

Do English and Japanese L2 speakers of Korean acquire nativelike sensitivity to the semantic features of the markedness hierarchy that condition variable use of the nominative and accusative case morphemes in Korean (i.e., person, animacy, prominence in definiteness)?

**Research Question 2: Testing the discourse factor**

Do English and Japanese L2 speakers of Korean acquire nativelike sensitivity to the different discourse-pragmatic functions that condition variable use of the nominative and accusative case morphemes in Korean (i.e., focus and discourse newness)?

**Research Question 3: Evaluating L2 theories**

To what extent do the predictions of the RDH, FRH, and IH account for the results?

In regard to the first and second research questions, the three L2 theories make different predictions, as summarized below:

The RDH predicts that for both English and Japanese speakers, neither the semantic features nor the discourse-pragmatic features will pose any significant learning difficulty. This prediction is based on the assumption that English and Japanese (and Korean) all select case features in a cross-linguistically comparable sense.\(^{39}\)

---

\(^{39}\) It has been conjectured that the notion of case considered for featural selection by the RDH is the abstract case features. It has also been discussed that the hypothesis views parametrization is sufficient to describe various typological realizations of certain features. See section 4.1.2 for a theoretical quandary of this view.
The FRH predicts substantive learning challenges for both the acquisition of the semantic features and that of the discourse-pragmatic features, without suggesting the acquisitions may not be possible, for the L2 case-relevant features are organized in dissimilar configurations from English and Japanese. However, based on the noticeable similarities between Japanese and Korean case systems, the Japanese speakers are predicted to produce significantly more nativelike performance than the English speakers.

The IH predicts that for both English and Japanese speakers, the semantic features, by virtue of being internal-interfaced properties, will be successfully acquired. In respect of the discourse-pragmatic features, they are unlikely to be acquired for focus and discourse newness require the coordination at external interfaces; despite the similarities between Japanese and Korean case systems, Japanese learners’ performance is not predicted to be significantly better than the English learners’.
Chapter 5 presents an overview of the corpus data explored with the research questions and discusses the methods used to collect and code the data. The chapter is organized as follows. Section 5.1 provides a description of the data source and collection procedure. Section 5.2 discusses the methodological criteria to extract the data tokens (section 5.2.1) and to code them for the independent and dependent variables examined in this study (section 5.2.2).

5.1. Data

5.1.1. Data source

The corpus on which this dissertation is based consists of naturally occurring L1 and L2 spoken data of conversational Korean produced in mass media discourse. The data source selected is a South Korean TV show, *picengsanghoytam*, known as ‘Non-Summit’ or ‘Abnormal Summit’ in English.\(^{40}\)

Implementing a parodic summit talk format, the program features an international panel of young foreign males of various nationalities, residing in South Korea, who are referred to as ‘representatives’ of their own nations. There is also a group of young Korean male hosts who identify themselves as ‘summit leaders’. In each episode, the international panel members and the three Korean summit leaders hold a 75 to 85-minute lively discussion on the topic of the week in the Korean language. Weekly topics are chosen in consideration of the show’s slogan of

\(^{40}\) The ambiguity, which is received as an intended pun, arises from two possible meanings of the noun, *ceongsang* (i.e., *pi* ‘non-’) *cengsang* (‘summit’ or ‘normal’) *hoytam* (‘talk’).
‘The youth group, without borders’ and thus cover a wide range of private and public matters for the world’s youth’s wellbeing, some of which encompass weighty areas such as history and politics. First aired in July 2014, the program gained great popularity as an entertainment debate show in the country and concluded in December 2017 with its 177th episode. Details relevant to this basic description of the data source are presented in Table 5.1.
(1) Table 5.1. Overview of the data source: *Picengsanghoytam* (‘Non-Summit’)

<table>
<thead>
<tr>
<th>Data availability</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Episodes</strong></td>
</tr>
<tr>
<td><strong>Repository</strong></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Show format</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Topics covered</strong></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Show members</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Representatives</strong></td>
</tr>
<tr>
<td><strong>Summit leaders</strong></td>
</tr>
</tbody>
</table>
The Korean speech data uttered by representatives whose L1 is English or Japanese make up the L2 data for this study, and those produced by summit leaders serve as the L1 data referred to for comparison purposes.

The fact that both L1 and L2 groups are composed of male speakers who are all at a relatively young age, ranging in their 20s and 30s, helps the study to readily control potential confounding factors of gender and age. Although there have not been established research findings on the role of gender and age in case (un)marking in Korean, it has been suggested that case drop rates are correlated with the degrees of formality of the extralinguistic context and of familiarity between the interlocutors (Ko, 2000; Lee & Thompson, 1989). Given that in the Asian language culture, gender and age are likely to be significantly associated with these sociolinguistic factors, it should be a sound methodological decision to minimize their effects in the research data.41

The show presents in-depth discussions on weekly topics, and a lot of them are concerned with culture, economics, history, and politics, as demonstrated in Table 5.1. In this respect, the producers’ first priority for casting the international representatives is that their language performance must display a certain high level of proficiency to meet the overall requirement to handle the depth of the discussions and the nature of the topics.

As to the L1 data coming from the three Korean summit leaders, one potential concern about using the data for L1-L2 performance comparison pertains to the distinct roles of the two speaker groups in the show (i.e., L1 speakers as moderator/host vs. L2 speakers as panel/guest), which

41 At times the show invites guests of various backgrounds, including Korean/non-Korean and male/female guests, by the name of ‘visiting intern representative’ or ‘guest representative’. The female guests and the male guests not in their 20s or 30s were not considered for the data collection.
might lead them to display different linguistic behaviors including case drop rates. Despite the validity of the concern, it should be pointed out that compared with other classic talk-show hosts, the moderator/host role performed by these L1 speakers is quite vague and incomplete. That is, instead of limiting their role to opening a discussion topic and appointing a panel member to talk, they actively take part in the discussion by expressing their own opinions (e.g., the host members’ interruption of and immediate response to a guest’s comment are frequent). The L1 data obtained from the hosts consists of utterances that they produce while holding this diluted host position in a discussion (see section 5.1.3).

5.1.2. Speakers

In the pool of international panel members are three Japanese native speakers that appeared in multiple episodes as a regular representative; there is also one Japanese native speaker that appeared in only one episode as a guest representative. Thus, the exhaustive examination of the pool of the show members led this study to identify the group of L1 Japanese L2 Korean speakers (henceforth J-group) with these four Japanese-speaking adult learners of Korean. Accordingly, the group of L1 English L2 Korean speakers (henceforth E-group) is composed of four English native speakers, who appeared in multiple episodes as a regular representative. These L2 speakers are non-heritage learners who were exposed to Korean in post-adolescence. The L1 Korean control group (henceforth K-group) consists of the three Korean hosts referred to as summit leaders and one Korean guest representative that appeared in two episodes. These 12 speakers of the three language groups are all male and in their 20s or 30s. See Table 5.2 for preliminary information of the target speakers whose speech data was collected for the study.
(2) Table 5.2. L1 and L2 speakers: Preliminary information

<table>
<thead>
<tr>
<th>Name</th>
<th>Birth year</th>
<th>Nationality</th>
<th>Occupation</th>
<th>Note</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>J-group (N=4, all male)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HO</td>
<td>1992</td>
<td>Japan</td>
<td>Character designer; Japanese instructor</td>
<td></td>
</tr>
<tr>
<td>HT</td>
<td>1981</td>
<td>Japan</td>
<td>Actor</td>
<td>Guest appearance (1 episode)</td>
</tr>
<tr>
<td>TT</td>
<td>1992</td>
<td>Japan</td>
<td>Singer</td>
<td></td>
</tr>
<tr>
<td>YN</td>
<td>1995</td>
<td>Japan</td>
<td>Singer</td>
<td></td>
</tr>
<tr>
<td><strong>E-group (N=4, all male)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BW</td>
<td>1992</td>
<td>Australia</td>
<td>Digital marketer</td>
<td></td>
</tr>
<tr>
<td>GP</td>
<td>1982</td>
<td>Canada</td>
<td>Video game player</td>
<td></td>
</tr>
<tr>
<td>MT</td>
<td>1980</td>
<td>USA</td>
<td>Businessman</td>
<td></td>
</tr>
<tr>
<td>TR</td>
<td>1988</td>
<td>USA</td>
<td>Graduate student</td>
<td></td>
</tr>
<tr>
<td><strong>K-group (N=4, all male)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HJ</td>
<td>1978</td>
<td>Korea</td>
<td>Announcer</td>
<td></td>
</tr>
<tr>
<td>SC</td>
<td>1981</td>
<td>Korea</td>
<td>Writer</td>
<td>Guest appearance (2 episodes)</td>
</tr>
<tr>
<td>SS</td>
<td>1979</td>
<td>Korea</td>
<td>Singer</td>
<td></td>
</tr>
<tr>
<td>SY</td>
<td>1980</td>
<td>Korea</td>
<td>Comedian</td>
<td></td>
</tr>
</tbody>
</table>

The corpus on which this dissertation is based is mass media discourse coming from one particular TV show. When the data has to be collected from only a limited number of L2 speakers, one important challenge to the corpus-based investigation is to ensure that the L2 data is procured from language groups as homogenous as possible in terms of L2 proficiency and
other language backgrounds that may affect L2 learnability and participants’ eligibility. Such information includes age of exposure to Korean, formal instruction of Korean, age of arrival and duration of stay in Korea, other languages spoken, results of a Korean proficiency/placement test, self-rated proficiency in Korean, and everyday use of Korean and interaction with native Korean speakers. Different from experimental studies, administering a proficiency test or a background questionnaire survey to the L2 speakers examined in this study was not possible. Nevertheless, as the show became popular in the nation, the L2 speakers attracted public attention significant enough to produce written or spoken records of media interviews regarding their backgrounds. Hence, a certain portion of the information is available for public access although there are varying degrees of availability for the individual speakers. To compensate for lack of firsthand reports of the L2 speakers’ language backgrounds, the following tables provide the relevant information that is available online (see Appendix A for the online sources). Table 5.3 and Table 5.4 address the Japanese speakers’ and the English speakers’ language backgrounds, respectively.
Table 5.3. L2 speakers: Language background J-group$^{42, 43}$

<table>
<thead>
<tr>
<th>Name</th>
<th>Language background</th>
</tr>
</thead>
<tbody>
<tr>
<td>HO</td>
<td>(Born in 1992; from Japan) HO studied Korean during his college years at Keio University in Japan. At that time, he was dating a Korean girl, and she made him start learning Korean. He took one year off from a college and came to Korea for the first time in 2014. He spent one year living in Korea and went back to Japan to finish his degree. HO came back to Korea in 2016 to work. He has taught Japanese since 2017 at a private language center in Seoul, while appearing at the show. He said watching Korean TV shows had been helpful for him to learn colloquial Korean.</td>
</tr>
<tr>
<td>HT</td>
<td>(Born in 1981; from Japan; guest appearance in 1 episode) HT is an actor, based in Korea. When he pursued his acting career in Japan, he was interested in Korean movies. He took a 3-month-long language course at Yonsei University in 2006 and then settled down in Seoul in 2007. He has appeared in several Korean movies, where he has maintained close interactions with the Korean cast and crew of a film production on a regular basis.</td>
</tr>
<tr>
<td>TT</td>
<td>(Born in 1992; from Japan) TT first came to Korea in 2012 at the age of 19 to make his debut to Korean media as a member of the Korean pop idol group, called Cross Gene. Ever since, he has lived in Korea. He started learning Korean after he arrived in Korea and passed the level 2 out of 6 levels of the Test of Proficiency in Korean (TOPIK).</td>
</tr>
<tr>
<td>YN</td>
<td>(Born in 1995; from Japan) YN made his debut to the Korean pop music industry as a member of the idol group, called NCT. He has been based in Korea since his first arrival in 2012.</td>
</tr>
</tbody>
</table>

$^{42}$ The test of Proficiency in Korean (TOPIK) (http://www.topik.go.kr) is a nationally standardized test designed to evaluate non-native speakers’ Korean proficiency. The test is administered in two formats. The first format, TOPIK-I consists of multiple-choice questions for listening and reading and has two obtainable grades: level 1-2 (beginner). The other format, TOPIK-II incorporates an additional writing examination and assigns four different levels to the results: level 3-4 (intermediate), and level 5-6 (advanced).

$^{43}$ It is unidentified when TT took TOPIK-I, but in one interview article released in August 2014 (http://osen.mt.co.kr/article/G1109942332), TT mentioned his TOPIK level in reference to two other international panel members’ (level 5 and 6, respectively). He added that he had taken the test long ago, and he would be able to obtain a higher level because his Korean had improved a lot since then.
(4) Table 5.4. L2 speakers: Language background. E-group

<table>
<thead>
<tr>
<th>Name</th>
<th>Language background</th>
</tr>
</thead>
<tbody>
<tr>
<td>BW</td>
<td>(Born in 1992; from Australia) BW attended an international high school, where he met and became friends with Korean students. He took a double major in business administration and the Korean language at a college. He then participated in an exchange program with a scholarship to Korea. After revisiting Australia to complete his degree and graduating in 2013, he came back to Korea, and ever since has worked at a digital marketing company.</td>
</tr>
<tr>
<td>GP</td>
<td>(Born in 1982; from Canada) GP arrived in Korea in 2000 and developed his career as a professional gamer until he retired in 2004. Since then, he has pursued his career in the Korean entertainment industry as a TV personality. Raised in Quebec, he speaks French, as well.</td>
</tr>
<tr>
<td>MT</td>
<td>(Born in 1981; from USA) MT first came to Korea in 2010 and worked for Samsung Electronics for four years. Currently, he is working at TCK Investment Management firm as Partner. He said he was interested in learning Korean while watching Korean movies and focused on learning everyday Korean expressions through natural conversations.</td>
</tr>
<tr>
<td>TR</td>
<td>(Born in 1988; from USA) TR studied international relations at the University of Chicago. During his college years, he taught himself Korean in the summer of 2007. One of his graduation requirements was to choose a foreign language and take an at least 8-week-long language course in a country where the chosen language is spoken. He chose Korean and attended a language school at Ewha Women’s University in 2008 for 3 months. After finishing his undergraduate degree, he worked at the Korea embassy in the U.S. in 2010 and then received a scholarship from the Korean government to attend a language school at Seoul National University in 2011. Upon completing the language course, he entered a Master’s program at the same university in 2012. He passed the level 6 (the highest level) of TOPIK. His command of Korean is considered particularly strong, and he is also known for his keen interest in foreign language learning. He speaks advanced French, and studied Japanese, Portuguese, and Spanish during his college years.</td>
</tr>
</tbody>
</table>
5.1.3. Data transcription and collection

Each episode consists of coherent discussions about a subtopic of the weekly topics. Such discussions were identified as a discourse unit. Discourse units in which a target speaker takes part were transcribed into written texts for further analysis.

Certain portions of a discourse unit were excluded from transcription when they were mainly presented for the purpose of entertaining and not considered part of a coherent discussion about a subtopic (e.g., participants’ singing or dancing and intermittent comments on the performance, an exchange of jokes as a fleeting side conversation held by speakers who are not taking the floor); when the speech production is based on a prepared written script; and when the speech displays an apparent moderator role of the L1 hosts (e.g., introducing a subtopic/guest; opening/closing a discussion).

This study utilized short excerpts of the show episodes released on YouTube. The methodological decision was made for enhanced transcription efficiency in consideration of two aspects of the nature of the data. First, the vast majority of the video clips released on YouTube were already created based on a discourse unit, and the search function on the website better facilitates the identification of a discourse unit in which a target speaker participates (vs. conducting an observational search by playing 75 to 85-minute length full episodes). Second, the program is a fast-paced debate show that allows spontaneous interruptions and competitive turn-takings among the participants, which often result in speaker overlaps. Furthermore, the target objects of transcription are occurrences or non-occurrences of functional morphemes, which exhibit far less prosodic salience than lexical/content words, and a great number of the transcription objects are pronounced by L2 speakers. All these properties make the audible data
not always intelligible. Efficient manual manipulation of playback speed control features to
handle such video data was essential for the transcribing procedure, and the features on the
YouTube video player are a labor-saving technical option compared with those on the alternative
players to run full episode videos. In terms of the number of videos available on YouTube,
however, there is an imbalance in data amounts for the individual L2 speakers. Hence, if/when
the short video clips eligible to be used as a particular L2 speaker’s data were insufficient, the
speaker’s data was then sought from a full episode: from the full episode, discourse units in
which the speaker takes part were transcribed.

Using both short video excerpts and full episodes, this dissertation analyzed transcribed texts
of a total of 6.6 hours of discourse units produced in the South Korean TV show,
picengsanghoytam (‘Non-Summit’). The average length of the transcription of each discourse
unit is 2 minutes (see Appendix B for the links for the YouTube videos and the full episodes
used for transcription).

To observe (non-)use of the case morphemes, overt subject/object DPs and case-markable
non-subject/non-object DPs were extracted from the transcribed texts. To identify the target DP
forms, the study first searched for three types of DP forms that correspond to three types of
particle marking of the DP: zero-marking, case marking, and other particle-marking. Due to the
challenge of aural unintelligibility, when multiple attempts of listening to a DP form did not
yield any perceptual bias to one of the three marking types, the token was classified as ‘marking
type unidentifiable’ to be discarded. When the multiple attempts led to a bias towards one of the
three marking types, another native speaker of Korean listened to the data to report her bias.
When a bias produced by this additional rater coincides with the author’s, the DP token was counted for further analysis.44

5.2. Methods
5.2.1. What to count
From the transcribed texts, subject/object tokens and case-markable non-subject/non-object tokens were extracted to observe whether they are zero-marked or nominative/accusative-marked in a given context. This section discusses the inclusion criteria used to extract the data tokens examined in this dissertation.

5.2.1.1. Basic criteria
The transcribed discourse data was initially screened for research eligibility. That is, not all candidate DPs for subject/object and case-markable non-subject/non-object tokens were considered as analyzable data.

A DP produced in an immediate self-repetition of the preceding utterance was excluded from the data when the reproduced form is identical with the form in the previous utterance. In contrast, a DP occurring in a non-immediate self-repetition was counted for further analysis. These repetitions are made immediately after the speaker’s utterance is momentarily interrupted by another speaker (or made after a pause longer than one second (Kim, 2008), when there is no

44 Audible intelligibility of a total of 157 DP tokens was re-examined by the additional rater. 21 tokens were excluded from the data set either because the agreeing bias classified them as other particle-marked DPs or because the author’s perceptual bias was not reliably matched by the other rater’s re-examination.
such interruption), resulting in two utterances of very similar clausal contents sequentially arranged with a significant intervention.

A DP occurring in a repetition of the preceding utterance by another speaker (i.e., other-repetition) was counted regardless of whether the repetition is regarded as immediate or non-immediate. In the discourse data, these repetitions are not merely a simple repetition of the antecedent utterance but are prone to expressing an independent communicative intent of the speaker such as expressing doubt or (dis)agreement and requesting clarification regarding the previous utterance.

As for a DP produced in a(n immediate or non-immediate) self-repetition, instances in which the reproduced form is different from what appears in the prior utterance were separately labeled as self-repair. When a self-repair was observed, it was only the reproduced, repaired form that was included in the data, not the original utterance. When a repair is made by another speaker (i.e., other-repair), both the original and repaired DP forms were counted for the study. When a repair is made by another speaker (i.e., other-repair), both the original and repaired DP forms were counted for the study.

A DP that involves an erroneous use of the case morphemes was disregarded from further analysis. These DPs indicate a misanalysis on a syntactic or morphological aspect of the case morphemes (e.g., accusative marking on the subject of an intransitive clause; nominative marking on the object of a transitive verb; use of an incorrect allomorph or contraction), and are not subsumable under the discussions of nominative/accusative marking on non-subjects/non-objects (section 2.3.2). Although they might constitute intriguing data for the L2 development,
the number of such instances was relatively small. The motivation for the methodological decision is that what is under examination is a variability phenomenon concerning two forms deemed equally grammatical; these erroneously case-marked DPs are judged to have no or little relevance to the effects of the semantic or discourse-pragmatic features tested in the study.

A DP that occurs after the predicate (e.g., SVO, OVS) was included in the data unless there is an interruption by another speaker or a pause longer than one second between the predicate and the post-predicate element.

Another inclusion criterion is identifiability of the predicate with which a DP form can be combined: a DP that lacks the associable predicate uttered by the speaker was not considered fully analyzable with respect to the grammatical function that the DP serves. Such defective data includes a DP in incomplete utterances containing insufficient information to infer the predicate that the speaker has intended to produce, where the lack of the predicate is attributed to the absence of the production or unintelligibility of such information in the transcribed texts. The other type is a DP occurring in co-constructed utterances in which the DP form is uttered by one speaker and all linguistic information to constitute the predicate is provided by another speaker via interruption or competitive turn taking.

Lastly, idiomatic expressions in which case- or zero-marking of the pertinent DP is conventionally fixed were ruled out. Examples of such discarded data include

---

45 The number of tokens bearing an erroneous nominative or accusative marking was 21 for E-group (error rate: 1.1%) and 24 for J-group (error rate: 1.4%).

46 Not all DPs produced without the predicate were excluded, however. When the context provides adequate information to retrieve the predicate (e.g., adjacent question-answer pairs), the DP was reanalyzed as an independent noun phrase that constitutes a complete utterance.
yey (‘example’)-tlul tule (‘take’) (‘for example’), mal (‘speak’)-ø ha (‘do’) camyen (‘if’) (‘so to speak’), pyel (‘unusual’) malssum (‘word’)-ul (‘don’t mention it’).

5.2.1.2. Subject and object tokens

Zero- or case-marked subject/object DPs were pulled out from the transcripts, which make up the majority of the data for the study.

Both transitive and intransitive clauses uttered by the target speakers were taken into consideration. A subject token is defined as an overtly realized grammatical subject of the clause (i.e., the subject of the transitive VP or intransitive VP/AP) that is either zero-marked or nominative-marked. An object token is identified as an overtly realized direct object of lexical transitive verbs\(^{47}\) that is either zero-marked or accusative-marked. (5) and (6) present examples appearing in intransitive and transitive clauses, respectively.


Gyum-ø alone sit-INF-have-CNJ cat-with like this play-CNJ-be-DCL

'Gyum was sitting alone, playing with a cat like this.' (SS, L1=Korean)

\(^{47}\) See footnote 19.
b. maknay tongsayng-un toykey hwalpal-hay-yo. toykey wuntongsinkyeng-i
youngest sister-TC very active-do-DCL very athletic ability-NOM
coh-ketunyo.
good-because.DLC

'My youngest sister is very active. Her athletic ability is great/ She is very athletic.'

(TT, L1=Japanese)

(6) kacok-ø pap-ø mek-ko iss-canhayo.
Family-ø meal-ø eat-CNJ be-you know.DCL

'The family was having a meal.'

(HO, L1=Japanese)

5.2.1.3. Non-subject and non-object tokens

It was discussed in section 2.3.2 that the nominative and accusative case morphemes in Korean may be attached to constituents that do not serve the corresponding subject/object functions in a clause. The linguistic environments described as such optional use of the case morphemes include sentential negation, linking verb usages, multiple nominative constructions, case alternation, case stacking, psych-predicates, subsets of intransitive action verbs, and different types of compound verbs.

The mass media discourse data of L1 and L2 speech involves a significant number of tokens pertaining to nominative/accusative marking of non-subjects/non-objects. Thus, this section introduces examples of the nominative/accusative-markable non-subject/non-object data included in the study. Since the subtypes of the data are mostly in accordance with the variety of
the usage contexts reviewed in section 2.3.2, each example is presented in reference to the corresponding context and example mentioned in that section.

(7) Sentential negation of a -ci-nominalized predicate (cf., (27c) on p. 51)

kitay-ha-ci-lul  an-h-ass-nun-tey,  e,  ikey

expectation-do-NMNL-ACC not-do-PST-IN-CNJ wow this.NOM\textsubscript{contraction}
nongtam-ila-nun ke-ney.

joke-COP-REL  thing-DCL

'I was not expecting a joke, but then I was like “Wow, this is the joke.”'

(TR, L1=English)

(8) Sentential negation of a nominal complement of the copula (cf., (29b) on p. 52)

Trump-nun wenlay  cengchiin-i  an-i-canhayo.

Trump-TC  originally politician-NOM not-COP-you know.DCL

'Originally, Trump was not a politician.'  (TR, L1=English)

(9) Nominal complement of toy (‘become’) (cf., (30) on p. 52)

yayki-lul  ha-taka  malssawum-i  toy-ess-tayyo.

talk-ACC  do-as  argument-NOM  become-PST-I heard.DCL

'As they talked, it became an argument.'  (TT, L1=Japanese)
(10) Multiple nominative constructions (cf., (31) on p. 53)\textsuperscript{48}

kuliko yay-\textit{ka} com khi-ka khu-canhayo.

and she-NOM somewhat height-NOM great-you know.DCL

'And, as for her, her height is somewhat great/she is somewhat tall.'

(TT, L1=Japanese)

\textsuperscript{48}The data contains both L1 and L2 speech that displays a structure equivalent to the multiple nominative constructions. However, the first nominative-markable constituents, which have been claimed to be a topic (see section 2.3.2.3), exhibit characteristics that are not typically noted with the constructions: they are normally realized as a phonologically light pronominal expression (e.g., \textit{ike} (‘this’), \textit{kuke} (‘that, it’)), and the referent associated with the expression is unclear in the context. Two examples are given below:

(11) \textit{ike} pi-ka o-myen swupwun-i com manh-a-ci-canhayo

this.NOM\textsubscript{contraction} rain-NOM come-if moisture-NOM somewhat much-INF get-you know.DCL

'When it rains, the air becomes more humid.'

(TT, L1=Japanese)

(12) \textit{kuke} nacwung-ey colep-ha-nun key

that.NOM\textsubscript{contraction} later-LOC graduation-do-REL thing.NOM\textsubscript{contraction}

incen-gi toy-nu-nya?

acknowledgement-NOM become-IN-Q

'When they graduate later, is it acknowledged (as completion of a formal education)?'

(SS, L1=Korean)

Despite the widely accepted view that Korean does not employ the insertion of a dummy subject, it is speculated that the use of such structure may be justified as a strategy to draw the listeners’ attention in such lively discourse marked by simultaneous involvement of multiple interlocutors. That is, the function of the first nominative-marked pronouns in (11) and (12) may be to pick up the topic currently being discussed in a vague manner (e.g., ‘as for what we are talking about’) to help the speaker to maintain or take the floor. Due to their structural resemblance to the multiple nominative constructions, these examples were preserved in the data: they are classified as non-subject tokens with the unique characteristics noted.
(13) Case alternation - locative to accusative (cf., (37) on p. 55)

\[ \text{wuli-ka yolehkey yelcha-lul tha-ko Europe-ul ka-l swu iss-nun-ke-canha.} \]

we-NOM like this train-ACC ride-CNJ Europe-ACC go-can be-REL-that.be-you

know.DCL

'We can go to Europe by train.' (SS, L1=Korean)

(14) Case stacking - nominative onto locative (cf., (44) on p. 59)

\[ \text{Russia-lang coyak-ul mayc-ess-ul tay ilpon kwukkyeng-ul} \]

\[ \text{Russia-with treaty-ACC conclude-PST-REL time Japan border-ACC} \]

\[ \text{keki-ey-ta-ka cenghay-ss-nun-tey...} \]

\[ \text{there-LOC-transferentive-NOM determine-PST-IN-CNJ} \]

'We when we concluded the treaty with Russia, we determined Japan’s border

on there….' (HO, L1=Japanese)

(15) Psych-predicate - embedded object of siph (‘want’) (cf., (52b) on p. 65)

\[ \text{yeycen-ey hanwu-ka nemwu mek-ko siph-ta-ko…} \]

\[ \text{long ago-LOC Korean beef-NOM so much eat-to want-DCL-QT…} \]

'Long ago (he said that) he wanted to eat Korean beef so much….'

(SY, L1=Korean)
(16) Cognate object (cf., p. 66)

……wuyenhi kulim-ul kuli-key toy-ess-eyo.

by accident painting-ACC paint-AD become-PST-DCL

'…..he started painting by accident.' (HO, L1=Japanese)

(17) Duration covered by the action of an intransitive verb (cf., (53), p. 66)

malathon-un kyeeyesok ney_sikan-ul ttwi-canha.

marathon-TC continuously four hour-ACC run-you know.DCL

'As for the marathon, they keep running for four hours.' (HJ, L1=Korean)

The predominant type of non-subject/non-object data for this study comes from various types of compound verbs.

The discourse data involves accusative-markable verbal nouns in ha (‘do’) - light verb constructions, lexical causative constructions with sikhi (‘cause to do’), and lexical passive constructions with pat (‘receive’) and tangha (‘suffer, undergo’). See (18) for an example of the ha-light verb constructions (cf., (46b), p. 60).

(18) hankwuk-ul chimlyak-ul hay-ss-ess-ta.

Korea-ACC attack-ACC do-PST-PERF-DCL

'They had attacked Korea.' (HJ, L1=Korean)
This study considered that *hankwuk* (‘Korea’) is assigned the patient theta-role by *chimlyak* (‘attack’). Accordingly, *hankwuk* (‘Korea’) is classified as an object token (i.e., the object of the lexical transitive verb expressed in the *ha* (do)-light verb construction), whereas the verbal noun, *chimlyak* (‘attack’) is counted as a non-object token.

When a noun expressed in a foreign language (e.g., the L2 speaker’s L1) occurs with *ha* to denote a particular action, the noun is also considered as a non-object token of verbal nouns in the light verb constructions. An example is provided in (19).

(19) *eynco-la-ko sokay-lul₆ hay-ss-canhayo.*

*eynco-COP-QT introduction-ACC do-PST-you know.DCL*

*kulayse swip-key eynco-Ø ha-ko….*

*so easy-AD cyberbullying-Ø do-CNJ*

'I introduced (the internet phenomenon called) *eynco*. So, the public easily *eynco* (those who express political opinions)….' (HO, L1=Japanese)

The discourse of (19) was concerned with a societal view on celebrities who openly express political opinions. The Japanese speaker uses a Japanese internet neologism *eynco* (‘cyberbullying’), which he previously introduced to the other participants, in combination with *ha*. This token was treated as a non-object in the same manner as the verbal noun in Korean, *sokay* (‘introduction’) was.

When a verbal noun expressed with *ha* carries a modifier phrase, the noun was reanalyzed as a common noun and the combined verb *ha* as a regular heavy verb. Thus, while in (20) *happyeng*
‘annexation’) is counted as a verbal noun non-object, *han il happyeng* (‘the Japanese annexation of Korea’) is classified as an object of the sentence.

(20) happyeng-*ul* ha-lye kule-*l* ttay-ka iss-*ko
annexation-ACC do-to like that-REL time-NOM be-CNJ
cipay-lul ha-lye kulel ttay-ka iss-supnita.
domination-ACC do-to like that-REL time-NOM be-DCL
ilpon-i yey-lul tule *han il ______ happyeng-*ul*
Japan-NOM example-ACC take-INF Korea Japan annexation-ACC
ha-lye kulay-ss-ci-yo.
do-to like that-PST-SUP-DCL
'There are cases when a country tries to annex another country and cases when a
country tries to dominate another country. Japan, for example, tried to do the Korea-
Japan annexation/ tried the Japanese annexation of Korea.' (SC, L1=Korean)

Identification and classification of the data pertaining to the lexical causative and passive constructions follow the same criteria as those for the *ha* (‘do’)-light verb constructions.

Examples of the lexical causative constructions with *sikhi* (‘cause to do’) include (cf., (47b) on p. 61):

(21) koceng(*fixation*)-Ø sikhi (‘cause to fix’) (MT, L1=English)

*myelmang*(*ruination*)-Ø sikhi (‘cause to ruin’) (TR, L1=English)
Examples of the lexical passive constructions with pat (‘receive’) and tangha (‘suffer, undergo’) include (cf., (48b) on p. 62; (49b) on p. 63):

(22) chotay (‘invitation’) - ø pat (‘be invited’) (MT, L1=English)

kongkyek (‘attack’) - ul pat (‘be attacked’) (SC, L1=Korean)

sponsor (‘sponsorship’) - lul pat (‘be sponsored’) (GP, L1=English)

kecel (‘rejection’) - ø tangha (‘be rejected’) (HO, L1=Japanese)

saki (‘fraud’) - lul tangha (‘be swindled’) (HJ, L1=Korean)

foul (‘foul’) - ø tangha (‘receive the referee’s judgment as a foul’) (TT, L1=Japanese)

The data also contains nominative-markable verbal nouns in the lexical passive constructions with toy (‘become’). Examples include (cf., (50b) on p. 63):

(23) thallak (‘elimination’) - ø toy (‘be eliminated’) (MT, L1=English)

huysayng (‘sacrifice’) - i toy (‘be sacrificed’) (BW, L1=English)

palkyen (‘discovery’) - ø toy (‘be discovered’) (SY, L1=Korean)

control (‘control’) - i toy (‘be controlled’) (YN, L1=Japanese)
In addition to the common types discussed above, the data involves various compound verb constructions that utilize other light verbs. The accusative and nominative examples of this new type are provided in (24) and (25), respectively.

(24) salam-tul-i nemwu swipkey [saki-Ø chi]-l swu iss-canhayo. person-PL-NOM too easy-AD fraud-Ø do-can be-you know.DCL 'People can commit fraud very easily then.' (MT, L1=English)

(25) kuntey kuke-nun com [ihay-ka] an ka]-n-ta. by the way that-TC somewhat understanding-NOM not be-IN-DCL 'But, that is not understood fully/But, I don’t fully understand that.' (HJ, L1=Korean)

The meaning of *chi* in (24) as a regular heavy verb is ‘hit’. Applying the meaning to the sentence would make *saki* (‘fraud’) a patient direct object of a lexical transitive *chi* (‘hit’) and fail to engender the correct interpretation of the sentence. Hence, *chi* was reanalyzed as a light verb with a meaning of ‘do’ and *saki* (‘fraud’) as a verbal noun non-object which in itself can subcategorize its complement.\(^\text{49}\) Likewise, the meaning of *ka* in (25) as a regular heavy verb is

\(^{49}\) Other verbs occurring with the similar non-object data include:

(26)  
\begin{align*}
\text{nay (‘submit’): } & \text{caksal} (\text{‘smashing’}) \cdot \phi \text{ nay}_i (\text{‘do’}) (\text{‘smash’}) & \text{(HJ, L1=Korean)} \\
\text{ip (‘wear’): } & \text{phihay} (\text{‘damage’}) \cdot \text{lip}_i (\text{‘receive’}) (\text{‘be damaged’}) & \text{(HO, L1=Japanese)} \\
\text{ssu (‘write’): } & \text{sinjyeng} (\text{‘attention’}) \cdot \text{ul ssu} (\text{‘give’}) (\text{‘pay attention’}) & \text{(BW, L1=Japanese)} \\
\text{mek (‘eat’): } & \text{maum} (\text{‘mind’}) \cdot \text{ul mek} (\text{‘make’}) (\text{‘make one’s mind’}) & \text{(HJ, L1=Korean)}
\end{align*}
‘go’, and a direct application of the meaning to the sentence yields semantic oddness. Thus, *ka* was reanalyzed as a light verb with a passive meaning and *ihay* (‘understanding’) as a verbal noun non-object in compound verbs.\(^5\)

I have discussed the range of examples that show the various subtypes of nominative/accusative-marked non-subjects/non-objects counted for this study. Although the use of the case morphemes is described to be optional for both subjects/objects and non-subjects/non-objects, the nature and degree of the optionality may not be considered to be identical for the two types. Thus, unlike the subject/object data, the non-subject/non-object data does not incorporate all zero-marked tokens of the subtypes. Instead, the study limits the inclusion of zero-marked tokens to the cases where a systematic comparison between case marking vs. zero-marking can be made as in the subject/object data. The process was guided by two principles: the primary guideline was whether the pertinent non-subject/non-object is originally marked by other case morphemes (i.e., when it is marked by a case morpheme, is nominative/accusative marking construed as the default morphological encoding?), and the secondary consideration was the degree of rarity of nominative/accusative marking throughout the data of each subtype.

\(^5\) Other verbs occurring with the similar non-subject data include:

\[(27)\]  
\[tul\text{ (‘lift, come into’): sayngkak}^{\text{‘thought’}}\cdot i\ tul\text{ (‘think’)}\]  
\[na\text{ (‘grow, come out’): kiek}^{\text{‘memory’}}\cdot \emptyset\ na\text{ (‘remember’)}\]  

(YN, L1=Japanese)  
(HJ, L1=Korean)

A subset of the compound verb constructions forms a group of compound adjectival predicates in the data. Examples include:

\[(28)\]  
\[tul\text{ (‘lift, come into’): him}^{\text{‘effort’}}\cdot \emptyset\ tul\text{ (‘effortful’)}\]  
\[na\text{ (‘grow, come out’): kwd}^{\text{‘anger’}}\cdot ka\ na\text{ (‘angry’)}\]  
\[iss\text{ (‘exist, be’: maylyek}^{\text{‘charm’}}\cdot \emptyset\ iss\text{ (‘charming’)}\]  

(TT, L1=Japanese)  
(HJ, L1=Korean)  
(GP, L1=English)
The primary criterion rules out the zero-marking counterpart of nominative/accusative marking via case alternation and stacking. On the other hand, zero-marked tokens are included when the data is concerned with the various compound verb constructions discussed above; the cognate objects constructions; the nominal complements of linking verbs; and the nominal complements involving an intransitive use of psycho predicates.51

A subset of the multiple nominative constructions is analyzed as an outcome of case alternation between nominative and other case particles such as genitive (cf., (31) and (32) on p. 53; footnote 22) and locative (cf., (35) and (36) on p. 54) particles. The data displays another subtype in which the original case of the first nominative-markable element is not readily identifiable or irrecoverable in the structure but may be properly marked by the topic marker - (n)un or other delimiters. Instances of zero-marking of the first nominative-markable element are disregarded if the sentence falls into the first subtype; however, if the sentence falls into the second subtype, the zero-marked instances are not ruled out (i.e., -(n)un-markability is not a criterion to rule them out since subjects/objects can also be marked by the topic marker -(n)un or other delimiters (section 2.3.1)). Case marking and zero-marking examples of this subtype are presented in (29) and (30), respectively.

51 When the predicate occurs in the intransitive form (i.e., adjectival predicate), instances of zero-marking are counted as a zero-marked non-subject (cf., (51b) on p. 64). When the predicate occurs in the transitive form, instances of zero-marking are classified as a zero-marked object (cf., (51a) on p. 64; (52) on p. 65).
(29) chwuek-i meli-ssok-ey nam-a iss-nun key
old memories-NOM heads-inside-LOC remain-INF be-REL thing.NOM
manhi eps-e-se…. much not.exist-INF-because…. 'As for the old memories, there are not many left in my mind, and that’s why/ It’s because I have few memories of my old (childhood) days.' (TT, L1=Japanese)

(30) namca-ø amwuli him-ø sey-to mos, mos naka-canhayo.
man-ø however power-ø strong-even not not come you know.DCL 'As for men, however strong his power is, he cannot get out of it/ However strong a man is, he cannot get out of it.' (GP, L1=Korea)

The study disregarded instances of zero-marking in sentential negation of a nominalized adjectival or verbal predicate. Although they do not violate the primary principle of nominative/accusative marking as a morphological default, when considering the overall pattern of the negative sentences in the data, it is very infrequent to mark them with the case morphemes. In other words, if the zero-marked predicates were taken into consideration, the data for the analysis would contain an overwhelming number of tokens of predicate negation, most of which are irrelevant to the research questions. Likewise, instances of zero-marking of the duration or distance covered by the action of an intransitive verb were not included in the data.
5.2.1.4. Overview of data tokens

A total of 6.6 hours of discourse units presented on a South Korean TV show, *picengsanghoytam* (‘Non-summit’) were transcribed. The transcribed discourse data provides a total of 5,923 tokens of subjects/objects and case-markable non-subjects/non-objects uttered by the target speakers of the three L1 groups: 1,847 tokens by E-group speakers, 1,656 tokens by J-group speakers, and 2,420 tokens by K-group speakers.

For all three groups, the majority of the nominative data consists of subject tokens (82.6% for E-group; 78.4% for J-group; 77.4% for K-group), whereas non-subject tokens take up a relatively small portion of the data. Similarly, for all three groups, more than a half of the accusative dataset is composed of object tokens. However, compared with the portion that non-subjects occupy the nominative set, the accusative set embodies more non-object data, which reflects the high frequency of accusative marking contexts in light verb constructions in Korean (e.g., ha (‘do’) - light verb constructions).

Table 5.5 presents an overview of the subject/object and non-subject/non-object tokens procured from each L1 group and from each speaker.
Table 5.5. Overview of data tokens

<table>
<thead>
<tr>
<th>L1 group</th>
<th>Subj.</th>
<th>Non-subj.</th>
<th>NOM</th>
<th>Obj.</th>
<th>Non-obj.</th>
<th>ACC</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>BW</td>
<td>125</td>
<td>21</td>
<td>146</td>
<td>42</td>
<td>47</td>
<td>89</td>
<td>235</td>
</tr>
<tr>
<td>GP</td>
<td>231</td>
<td>41</td>
<td>272</td>
<td>161</td>
<td>87</td>
<td>248</td>
<td>520</td>
</tr>
<tr>
<td>MT</td>
<td>198</td>
<td>57</td>
<td>255</td>
<td>117</td>
<td>134</td>
<td>251</td>
<td>506</td>
</tr>
<tr>
<td>TR</td>
<td>217</td>
<td>45</td>
<td>262</td>
<td>195</td>
<td>129</td>
<td>324</td>
<td>586</td>
</tr>
<tr>
<td><strong>E-group</strong></td>
<td><strong>771</strong></td>
<td><strong>164</strong></td>
<td><strong>935</strong></td>
<td><strong>515</strong></td>
<td><strong>397</strong></td>
<td><strong>912</strong></td>
<td><strong>1,847</strong></td>
</tr>
<tr>
<td></td>
<td>(82.5%)</td>
<td>(17.5%)</td>
<td>(100%)</td>
<td>(56.5%)</td>
<td>(43.5%)</td>
<td>(100%)</td>
<td></td>
</tr>
<tr>
<td>HO</td>
<td>224</td>
<td>68</td>
<td>292</td>
<td>152</td>
<td>119</td>
<td>271</td>
<td>563</td>
</tr>
<tr>
<td>HT</td>
<td>33</td>
<td>5</td>
<td>38</td>
<td>15</td>
<td>21</td>
<td>36</td>
<td>74</td>
</tr>
<tr>
<td>TT</td>
<td>204</td>
<td>65</td>
<td>269</td>
<td>163</td>
<td>83</td>
<td>246</td>
<td>515</td>
</tr>
<tr>
<td>YN</td>
<td>218</td>
<td>39</td>
<td>257</td>
<td>132</td>
<td>115</td>
<td>247</td>
<td>504</td>
</tr>
<tr>
<td><strong>J-group</strong></td>
<td><strong>679</strong></td>
<td><strong>177</strong></td>
<td><strong>856</strong></td>
<td><strong>462</strong></td>
<td><strong>338</strong></td>
<td><strong>800</strong></td>
<td><strong>1,656</strong></td>
</tr>
<tr>
<td></td>
<td>(79.3%)</td>
<td>(20.7%)</td>
<td>(100%)</td>
<td>(57.7%)</td>
<td>(42.3%)</td>
<td>(100%)</td>
<td></td>
</tr>
<tr>
<td>HJ</td>
<td>274</td>
<td>78</td>
<td>352</td>
<td>191</td>
<td>102</td>
<td>293</td>
<td>645</td>
</tr>
<tr>
<td>SC</td>
<td>258</td>
<td>80</td>
<td>338</td>
<td>177</td>
<td>122</td>
<td>299</td>
<td>637</td>
</tr>
<tr>
<td>SS</td>
<td>248</td>
<td>74</td>
<td>322</td>
<td>170</td>
<td>96</td>
<td>266</td>
<td>588</td>
</tr>
<tr>
<td>SY</td>
<td>224</td>
<td>59</td>
<td>283</td>
<td>169</td>
<td>98</td>
<td>267</td>
<td>550</td>
</tr>
<tr>
<td><strong>K-group</strong></td>
<td><strong>1,004</strong></td>
<td><strong>291</strong></td>
<td><strong>1,295</strong></td>
<td><strong>707</strong></td>
<td><strong>418</strong></td>
<td><strong>1,125</strong></td>
<td><strong>2,420</strong></td>
</tr>
<tr>
<td></td>
<td>(77.50%)</td>
<td>(22.5%)</td>
<td>(100%)</td>
<td>(62.8%)</td>
<td>(37.2%)</td>
<td>(100%)</td>
<td></td>
</tr>
<tr>
<td>All</td>
<td>2,454</td>
<td>632</td>
<td><strong>3,086</strong></td>
<td>1,684</td>
<td>1,153</td>
<td><strong>2,837</strong></td>
<td><strong>5,923</strong></td>
</tr>
<tr>
<td></td>
<td>(79.5%)</td>
<td>(20.5%)</td>
<td>(100%)</td>
<td>(59.4%)</td>
<td>(40.6%)</td>
<td>(100%)</td>
<td></td>
</tr>
</tbody>
</table>
5.2.2. Data coding

This section provides a description of the coding system implemented in this dissertation. The total 5,923 subject/object and non-subject/non-object tokens collected from the 12 L1 and L2 speakers of Korean were coded for markedness and discourse features. The frequency distribution of case marking and zero-marking of the DPs was observed according to the two independent variables.

5.2.2.1. Coding of the markedness factor

The coding categories of the markedness factor were generated based on the classic version of the markedness hierarchy proposed in Silverstein (1976) (section 2.2.2.5), which is reiterated in (32).

(32) Markedness hierarchy

Local person > Pronoun 3rd > Proper Noun 3rd > Human 3rd > Animate 3rd > Inanimate 3rd

(Silverstein, 1976 cited in Aissen 1999: 674)

As discussed in section 2.2.2.5, the hierarchy embodies the semantic features that construct the notion of prototypical subject/object (i.e., person, animacy, prominence in definiteness). The assumption made in this dissertation is that coding the data by the hierarchically organized categories in (32) will lead the analysis into distinct effects of the semantic features on the morphological realization of case in correspondence with the grammatical function that a DP serves. Thus, the present study classified the subject/object and non-subject/non-object tokens
according to the categories expressed in the markedness hierarchy while making a few revisions to indicate additional annotations for potential subcategories.

The data for local person was divided into two categories: [1\textsuperscript{st} person] (e.g., \textit{na/ce} (‘I\textsuperscript{plain/polite}’), \textit{wuli/ceuy} (‘we\textsuperscript{plain/polite}’)) and [2\textsuperscript{nd} person] (e.g., \textit{ne} (‘you’), \textit{neney/yelepwn} (‘you\textsuperscript{plain/polite}’)).

Pronominal expressions that denote a 3\textsuperscript{rd} person referent identified in the discourse are classified as [pronoun 3\textsuperscript{rd}]. These include personal and demonstrative pronouns (e.g., \textit{yay} (‘he’ or ‘she’), \textit{kuke} (‘that, it’), \textit{ike} (‘this’)) as well as reflexive pronouns (e.g., \textit{caki} (‘oneself’), \textit{ponin} (‘oneself’)). The dummy-like first nominative-markable pronouns in the multiple nominative-type structure\textsuperscript{52} were subsumed under the category with the status of referent unidentifiability noted.

[proper noun 3\textsuperscript{rd}] includes nominal expressions that designate a uniquely identifiable 3\textsuperscript{rd} person referent. The study defines a proper noun as a rigid designator with no intrinsic descriptive meaning: the noun in itself is referential, as opposed to needing a DP to be referential (Portner, 2005). Thus, the typical majority of the examples are names of the L1 and L2 participants, of the nations that they are from, and of entities being talked about in the discourse. An example is provided in (33).

(33) tongasia salam-ulose EU-ka ce-hanthey-nun yakkhan huymang-i-ess-eyo.

East Asia person-as EU-NOM I-to-TC a bit hope-COP-PST-DCL

'As an East Asian, the EU was some kind of hope for me.' (SC, L1=Korean)

\textsuperscript{52} See footnote 48.
The transcribed texts include not only the clear-cut examples but also somewhat fuzzy data of proper noun-like tokens that do not readily fit in the above definition. Considering that the markedness hierarchy is deemed to accommodate definiteness as a concomitant feature, these tokens were annotated as a heterogeneous mixture of proper nouns and definite descriptions. A subtype is proper adjective-modified nouns that refer to kinds of objects (i.e., referential uses of definite descriptions) such as ethnic-national groups to which the L1 and L2 speakers belong (e.g., *mikwuk* (‘America’) *salam* (‘person’)*(-tul*{PL}) (‘Americans’)).

The gray-area data also includes instances of descriptive uses of apparent proper nouns. An example is illustrated in (34).

(34) SS: *Italy-*uy **Sung Sikyung-i**   iss-ul-kee-yo,   pwunmyenghi.

Italy-GEN Sung Sikyung-NOM exist-PRS-that.be-DCL definitely

'There must be an Italian Sikyung Sung.'

Italian Speaker: manh-ayo,  manh-ayo.

many-DCL many-DCL

'There are many.'

HJ: *Italy-*ey **Sung Sikyung-i**  manh-ayo?

Italy-LOC Sung Sikyung-NOM many-Q

'Are there many Sikyung Sungs in Italy?'

(SS, L1=Korean; HJ, L1=Korean)

The previous discourse of (34) concerns world’s worst-dressed nations while the Korean host,
"Sung Sikyung (‘Sikyung Sung’) was considered one of the worst dressers in the show. Both occurrences of the name are not referential, as illustrated by the genitive modification in the first occurrence, but provide a descriptive sense for designating indefinite human entities that display the same property as the referent of the name holds.

The multicultural nature of the discussions in the show seemed to provide rich ground to use nominal expressions that display a temporary proper noun status. These are primarily observed when a speaker introduces a unique culture of his country, which is often denoted by a newly coined term in his L1 as well as in Korean. Determining whether the term is a proper noun or not required a significant level of contextual information such as whether the interlocutors have shared knowledge to render the referent uniquely identifiable in the context. Unlike pure proper nouns, they obtain a proper noun status from a given discourse context; and the proper noun status is accepted limitedly to those participating in the specific discourse. For these reasons, examples like (35) were labeled as proper noun-like.

(35) TT: cey-ka cen-ey kapeytong-ul sokay hay-ss-canhayo.
     I-NOM before-LOC kapeytong-ACC introduction do-PST-you know.DCL
     ipen-ey-nun somay cep-ki.
     this time-LOC-TC sleeve fold-NMLZ

'Last time, I introduced (to you) kapeytong (flirting/a flirting skill called kapeytong in Japanese). This time, (I’ll teach you) roll-up-her-sleeves (flirting).'}
HJ: mwe-la-ko-yo?

what-COP-QT-Q?

'What did you say?'

TT: somay-lul cep-e-cwu-nun.
sleeve-ACC fold-INF give-REL

'Rolling up the sleeves.'

SS: ettehkey ha-nun keey-yo somay cepki-ø?

how do-REL that.be-Q sleeve fold-NMLZ-ø

'(Can you teach us) how to do roll-up-her-sleeves?'

(TT, L1=Japanese; HJ, L1=Korean; SS, L1=Korean)

The discourse of (35) concerns flirting tips currently popular among young Japanese men. It is judged that the participants of the discourse have shared knowledge to associate somay cepki ('roll-up-her-sleeves') and a Japanese neologism kapeytong ('thud-into-wall' (i.e., kapey ('wall') tong ('thud'))) with a unique behavior to flirt with the opposite gender in Japan. Therefore, they were classified as proper noun-like tokens while somay ('sleeves') was coded as [inanimate 3rd].

Nominal expressions associated with a 3rd person human referent were coded as [human 3rd]. When the associated 3rd person referent is not a human but a living entity that maintains one and the same life (Yamamoto, 1999), they were classified as [animate 3rd]. Hence, animals and plants treated as a life by the speaker fall into the category. See an example in (36), where chilmyenco ('turkey') is coded as [inanimate 3rd] in (36a) but as [animate 3rd] in (36b).
(36) a. chwuwakamsacel, ney, chilmyenco-lul mek-eyo.

Thanksgivings yes turkey-ACC eat-DCL

'Yes, we eat turkey on Thanksgiving.' (TR, L1=English)

b. etten wencwumin-tul-i nao-a-se chilmyenco-lul
certain native-PL-NOM come out-INF-CNJ turkey-ACC
cap-a-se mek-ul swu iss-ta, ilen ke-lul kaluchi-e-cwu-ko….
catch-INF-CNJ eat-can be-DCL like this thing-ACC teach-INF-give-CNJ

'Some Native Americans came out and taught (the settlers) that they can catch a
turkey to eat.' (TR, L1=English)

A handful of nominal expressions typically associated with an inanimate entity was reanalyzed as [animate 3rd] and annotated as a metonymic human (Zhang, 2015). The reanalysis was motivated by the notion of inferred animacy. Inferred animacy involves semantic properties such as sentience, intentionality, and responsibility to allow a place or an organization to be metonymically construed as standing for humans working in that place or organization (Song, 2011). For example, *hoysa* (‘company’) in (37) is considered as a volitional entity since it is metonymically understood as humans working in the department of human resources, an agent of the action of denoted by the verb *hayko-lul sikhi* (‘cause to dismiss’ or ‘dismiss’).
The rest of the data that does not fall under the above categories was coded as [inanimate 3rd].

5.2.2. Coding of the discourse factor

Generating the coding categories of the discourse factor was guided by the previous claims (i) that (non-)use of the case morphemes is sensitive to focus status of a DP (Ko, 2000; Lee, 2006a; Schütze, 2001); (ii) that the most pertinent notion of focus that resists case drop is a narrow subtype of contrastive focus (Lee, 2011; Lee & Choi, 2010); (iii) that nominative marking is frequently associated with a nominal expression used to introduce new information into a discourse (Sohn, 1999); (iv) that Korean displays a subtle division of labor between case marking and delimiter marking (e.g., -(n)un-marking) in encoding the contrast subtypes (Kim, 2008; Lee, 2003) and the newness distinction (Hong, 1985).

Informed by the previous insights, this study considered the data tokens in terms of the features of focus and newness to divide them into three broad categories: [new], [focus], and [none]. These categories encapsulate a specific discourse function that a DP carries out in a given context. Subcategories were created to accommodate the subtlety of the discourse properties claimed to influence the variability phenomenon.
The [new] category is applied to the nominative data only. The rationale behind the restricted application of [new] is as follows. First, considering that object DPs typically convey new information, the newness distinction is not likely to have a notable impact on the naturalness of (non-)use of the accusative morpheme (section 2.2.2.4). Second, the meaning differences between -(n)un-marking vs. case marking proposed in the past literature were drawn from two separate comparisons. The comparison for the contrast sub-distinction dealt with -(n)un-marking (of sentence-internal element) in relation to both nominative and accusative markings; however, what was compared with -(n)un-marking (of a sentence-initial element) for the newness distinction was restricted to nominative marking (section 2.3.1).

A DP in the nominative data that denotes discourse-new information was classified as [new], where discourse-new information is neither previously mentioned nor inferable from the preceding discourse (Prince, 1992). Below are two examples coded as [new].

(38) hakkyo-eyse applied sciences-Ø kongpwu-ha-ko iss-ess-nun-tey kapcaki
    school-LOC applied sciences-Ø study-do-CNJ be-PST-IN-CNJ suddenly
    StarCraft-Ø nao-n keey-yo.
    StarCraft-Ø come out-REL that.be-DCL
    'When I was studying applied sciences in school, all of a sudden, StarCraft came out.'

(GP, L1=English)
(39) HJ: seykyeysa-ka cenmwun-i-si-n keey-yo?

world history-NOM expertise-COP-HON-REL that.be-Q

'Is world history your expertise?'

SC: yey, amwulayto cey-ka ssu-nun chayk-tul-i keuy taypwupwun

yes somehow I-NOM write-REL book-PL-NOM almost most

yeksa-ey kwanlyen-toy-n chayk-tul-i manh-supnita.

history-LOC connection-become-REL book-PL-NOM many-DCL

'As for the books that I write, somehow they are mostly history books/ Somehow most of the books that I write are history books.'

(HJ, L1=Korean; SC, L1=Korean)

Discourse-new information that occurs with a verb *iss* (‘exist, be’) received an additional annotation. Cases in which the token is a subject and the verb meaning is construed as ‘exist’ were sub-grouped for further analysis since the structure implied in the combination is often associated with information packaging for nominal expressions used to introduce a new entity into a discourse. Two examples are given in (40) and (41).

(40) ilpon mal-lo ha-myen yamanoteyseyn keyim-ila-nun keyim-i

Japan language-in do-if yamanoteysen game-COP-REL game-NOM

iss-nun-ney, cihachel lain-i iss-eyo. kukey ilum-i-n-ney…

exist-IN-CNJ subway line-NOM exist-DCL that.NOM contraction name-COP-IN-CNJ

'There is a game called *yamanoteyseyn* in Japanese, and there is a subway line. That’s the name/ The game is named after the subway line….'

(TT, L1=Japanese)
(41) ilpon-eyse-to caknyen-ey hwacey-ka toyn ke-Ø iss-nun-ney...
Japan-LOC-also last year-LOC issue-NOM become-REL thing-Ø exist-IN-CNJ
There was something, in Japan as well, that became an issue last year…'

(HO, L1=Japanese)

Cases in which the token is not a subject and the verb meaning is not understood as ‘exist’ were determined to form a compound adjectival predicate (e.g., casin(‘confidence’)-i/-Ø iss (‘confident’))\textsuperscript{53} and therefore were not taken into consideration for the sub-label.

The [focus] category encompasses various types of focus phenomena, each of which is accommodated with a subcategory. [focus] is broadly divided into two types. One type that bears contrastive salience via the presence of a contextually invoked set of alternatives is labeled as [contrast]. The other type involving no such alternative set is grouped as [information focus]. [contrast] is further divided into [exclusive contrast] and [non-exclusive contrast]. The [exclusive contrast] type focus is realized either on a DP or at a predicate level, where the former is termed as [focus on DP] and the latter as [focus on polarity]. Table 5.6 illustrates the structural organization of the subcategories subsumed under [focus].

(42) Table 5.6. Coding of [focus]: Subcategories

<table>
<thead>
<tr>
<th>[focus]</th>
<th>--- [information focus]</th>
</tr>
</thead>
<tbody>
<tr>
<td>--- [contrast]</td>
<td>--- [exclusive contrast]</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>--- [non-exclusive contrast]</td>
</tr>
</tbody>
</table>

\textsuperscript{53} See (28) in footnote 52.
Information focus, often referred to as ‘completive focus’ or ‘presentational focus’, is understood as a particular type of focus construction which consists of an open proposition with one or more variables licensed by a wh-operator, and the answer which provides the constituent to fill in the gap in the open proposition (Tomioka et al., 2003). In this study, a DP was classified as [information focus] when (i) the DP assigns a particular value to the variable in an open proposition represented by a wh-question and (ii) the discourse context does not invoke a set of limited members for which the predicate phrase potentially holds. (43) presents an example coded as [information focus].

(43) (talking about Korean products that the international representatives have ever used)

SS: wulinala mwe-ka kwayanchan-ho, wulinala mwe-ka
our country what-NOM fine-CNJ our country what-NOM
an cohu-n-kayo?
not good-IN-Q

'As for Korea, what is fine and what is not good?/ As for Korean products, what do you find good and what do you find not good?'

best-from hear-INF-see-FUT-DCL best

'Let’s first hear about the best.'

GP: han kwuk-eyse-nun navigation-Ø choyko-eyyo.
Korea-LOC-TC navigation-Ø best-DCL

'As for Korea, navigation systems are the best.'

(SS, L1=Korean; HJ, L1=Korean; GP, L1=English)

157
When an element is in focus in contrast with some other element(s) present in the domain of discourse (i.e., the presence of a set of limited alternatives invoked by the preceding discourse), the element was classified as [contrast]. [contrast] could be regarded as either [exclusive contrast] or as [non-exclusive contrast] according to whether the focused element involves exclusive (or exhaustive) identification.

A DP token is coded as [exclusive contrast] when (i) there is a contextually invoked closed set of propositional alternatives (ii) with a presupposed requirement of exclusive denotation. The exclusivity requirement states that choosing one member from the alternative set implies exclusion of the other member(s). That is, the discourse context requires that only one proposition in the set be true. See the examples in (44) and (45) to illustrate the definition.

(44) HJ: Enes-ka yeki-se ceyil hyeng-i-ci-yo?

    Enes-NOM here-LOC most brother-COP-SUP-Q

    'Enes is the oldest here, isn’t he?'

GP: cey-ka ceyl hyeng-i-pnita.

    I-NOM most brother-COP-DCL

    'I’m the oldest.'

**Alternative set:** \{Enes is the oldest, I\textsubscript{GP} am the oldest\}

(HJ, L1=Korean; GP, L1=English)
(45) (talking about top 10 etiquette tips for tourists in Japan)

HO: (listing the tips)…. kil-eyse sukhinswip-ul ha-ki.

street-LOC skinship/intimate touching-ACC do-NMLZ

'…(Couples should) display physical affection on streets.'

SS: e? ha-ci, ha-ci an-h-ki?

uh-Q do-NMLZ do-NMLZ not-do-NMLZ-Q

'Uh? (You mean couples should) not display physical affection?'

SY: Ogi hwanthaci-lul tut-ko siph-ci an-h-ayo, cikum.

Ogi\textsubscript{HO} fantasy-ACC hear-to want-NMLZ not-do-DCL now

'We don’t want to listen to Ogi’s fantasy now/ What we want to listen to now

is not your fantasy.'

Alternative set: \{ we want to listen to the tips, we want to listen to Ogi\textsubscript{HO}’s fantasy \}

(HO, L1=Japanese; SS, L1=Korean, SY, L1=Korean)

More examples include an individual disjunctively chosen from a set of the reciprocal participants in two separate roles and an individual compared with limited alternatives to be associated with a superlative or comparative adjective form. Overall, elements to denote exclusive identification on a set of “immediately relevant alternative referents presupposed in the context” (Lee, 2017: 15) were classified as [exclusive contrast].

The data source of the corpus study is an entertainment debate show. That is, the data comes from a discourse format in which opposing arguments are put forward on a particular topic.
Reflecting the discourse format, the transcribed texts embody copious contexts that a DP is potentially associated with two conflicting predicate contents, one of which typically entails negation of the other. Hence, the alternative propositions are generated by the presence of two incomparable predicates competing to describe one single entity. This means that the exclusive identification is not made on a DP per se. Rather, the identification is performed at the predicate level, which brings forth a type of polarity marking. This is distinguished from the property of the alternative sets involved in the above examples: the alternative propositions in (44) and (45) are generated by the presence of multiple individuals competing to be coupled with one single predicate. Since what undergoes the exclusive designation is an individual denoted by a DP, the focus is deemed to fall onto the DP. In this regard, instances like (44) and (45) were sub-grouped as [focus on DP].

In order to accommodate this difference, the coding was guided by the notion of so-called ‘verum focus’. The meaning of verum focus is described as placing emphasis on the truth of the propositional content of the sentence (Höhle, 1992 cited in Gutzmann & Miró, 2011: 143). Cross-linguistically, there are different ways to instantiate a verum operator: German employs a special kind of focal stress on the auxiliary, and English uses a special kind of lexical insertion (e.g., I did finish his book) (Gutzmann & Miró, 2011). The subtype of [exclusive contrast] under discussion fits in the description of verum focus in that it concerns a discourse context in which only one of the predicates in competition is attached to the DP, thereby yielding a focus with sentence polarity. For this reason, the subgroup was labeled as [focus on polarity], as opposed to [focus on DP]. Since verum focus can be realized in interrogative sentences as well as declarative sentences (Gutzmann & Miró, 2011; Romero & Han, 2004), [focus on polarity] also
includes DPs occurring in questions that can be interpreted as an affirmative or negative assertion: yes/no questions involving an epistemic bias (i.e., bias towards an affirmative proposition in negative yes/no-questions; bias towards a negative proposition in affirmative yes/no-questions) and wh-rhetorical questions conveying a negative assertion. Consider (46) as an example:

(46) SS: phyohyen-uy cayu-i-n-kayo?

expression-GEN freedom-COP-IN-Q

'(Is it considered) free speech?'

HJ: ikey phyohyen-uy cayu-eyyo?

this.NOM contraction expression-GEN freedom-Q

'Is this free speech?'

Unrecognizable Speaker 1: phyohyen-uy cayu-c-yo.

expression-GEN freedom-SUP-DCL

'It is free speech.'

TR: ce-nun phyohyen-uy cayu-la-ko sayngkak hay-yo.

I-TC expression-GEN freedom-COP-QT thought do-DCL

'I think it is free speech.'

Unrecognizable Speaker 2: way-yo?

why-Q

'How come?'
HJ: ikey-yo?

this.NOM\_contraction-Q?

'Is this?'

Alternative set: \{this is free speech, this is not free speech\}

(SS, L1=Korean; HJ, L1=Korean, TR, L1=English)

The discourse of (46) is about limits of free speech. The referent denoted by ikey (‘this’) is an example of making derogatory comments on one’s appearance and was introduced in the preceding discourse to challenge TR’s stance on the subject by HJ. Therefore, the two occurrences of the yes/no-question uttered by HJ in (46) are deemed as a question biased to a negative answer, which stands as a presupposed proposition in the domain of the discourse. This completes the set of propositional alternatives, where HJ picks the negative alternative over the positive alternative.

Although negations admittedly presuppose their positive counterparts, negative sentences uttered out of the blue were not considered for coding of [focus on polarity] as a null context does not provide adequate grounds for putting a special emphasis on sentence polarity. Accordingly, coding [focus on polarity] in negative sentences was restricted to contexts in which the positive counterparts are introduced as an opposite proposition in the previous discourse. As an illustration, in the earlier example (45), Ogi hwanthaci (‘Ogi’s fantasy’) occurs in a negative sentence, but it is not coded as [focus on polarity]. This is because the positive counterpart, we want to listen to Ogi’s fantasy is not introduced in the discourse as the opposite proposition of the negative sentence. In other words, the preceding discourse does not address Ogi hwanthaci
(‘Ogi’s fantasy’), as an entity potentially associated with the positive predicate and thus fails to generate a set of alternatives felicitous for [focus on polarity] (i.e., \{we want to listen to Ogi’s fantasy, we don’t want to listen to Ogi’s fantasy\}).

A DP token is classified as [non-exclusive contrast] when (i) there is a contextually invoked closed set of alternative propositions (ii) with no presupposed requirement of exclusive denotation. Since choosing one member from the alternative set does not imply exclusion of the other member(s), the discourse context allows more than one proposition in the set to be true. Two examples are given in (47) and (48).

(47) (talking about the international education market in Korea)

kunikka ehakyenswu saep-i cwuletul-kwuyo kuliko seksa,

so language program-NOM decrease-CNJ-be-DCL and master’s degree

paksa haksa-ø ta nulena-kwuyo supnita….

bachelor’s degree doctoral degree bachelor’s degree all increase-CNJ-be-DCL

'I mean language programs are contracting, and master’s, doctoral, and bachelor’s degree programs are all expanding….'

Alternative set: in the Korean market \{language programs are contracting, degree programs are expanding\} (TR, L1=English)

(48) (talking about a figure skating rivalry between Yuna Kim from Korea and Mao Asada from Japan, and Koreans’ support for Yuna in the 2010 Winter Olympics)
SS: na-n Mao-to ungwen hay-ss-ess-nun-tye.

I-TC Mao-also support do-PST-PERF-IN-DCL

'I supported Mao, as well.'

HJ: na-to Mao-ø ungwen hay-ss-eyo.

I-also Mao-ø support do-PST-DCL

'I also supported Mao (as well).'

Alternative set: \{I_{Korean} supported Yuna, I_{Korean} supported Mao\}

(SS, L1=Korean; HJ, L1=Korean)

Finally, there are some subtle data tokens coded as [exclusive contrast] (i.e., either as [focus on DP] or as [focus on polarity]). They were judged to be embedded in a hierarchical structure of multiple foci in a sentence and received additional annotation as ‘parallel under higher focus’.

(49) is an example coded as [focus on polarity].

(49) (talking about a Japanese couple that fights because the wife likes zucchini in miso soup while the husband finds it unpalatable)

….way toyncangkwu-key hopak-ul nehnya? wuli cip-eyse-nun

why miso soup-LOC zucchini-ACC put-Q we home-LOC-TC

hopak-ul celtay an neh-nun-ta….

zucchini-ACC never not put-IN-DCL

'Why do you put zucchini in miso soup? In my family, we never put zucchini in it.'

Alternative set: in a family \{the cook puts zucchini in miso soup, the cook does not put zucchini in miso soup\}

(TT, L1=Japanese)
In (49), whether to put hopak (‘zucchini’) in miso soup or not constitutes a disjunctive set, which renders the DP eligible for a subcategory of [exclusive contrast], [focus on polarity]. Still, the sentence evokes an additional contrastive reading arising at a higher level. The alternative set responsible for the higher focus is \{in my family, the cook does not put zucchini in miso soup, in your family the cook puts zucchini in miso soup\}. Since the two propositions are not mutually exclusive, the higher focus reading falls into [non-exclusive contrast]. The study analyzed hopak (‘zucchini’) as [focus on polarity] provided that it is parallelly embedded under higher focus of [non-exclusive contrast].

Another example provided in (50) represents [focus on DP] under higher focus of [focus on polarity]. Note that the alternative propositions involving cinap (‘extinguishment’) and pwul (‘fire’) stand under higher negation of the propositional content of the relative clause.

(50) cinap-ul sikhi-nun-key an-ila pwu-lul ciphy-eyo.
    extinguishment-ACC cause to-REL-thing not-COP.CNJ fire-ACC burn-DCL
    'You’re not extinguishing a fire but burning a fire (meaning: you’re not helping to cool down the heated argument but making it more heated).'
    Alternative set: \{you are extinguishing a fire, you are accelerating a fire\}
    (HJ, L1=Korean)

The rest of the data that falls neither into [new] nor into [focus] was assigned to [none]. Below are examples coded as [none].
(51) HJ: cikum ha-nun il-un mwe-yeyyo?

    now do-REL work-TC what-Q

'What’s your current occupation?'

BW: cey-ka makheything eyicenssi-eyse il-ø hako iss-nun-tey…
    I-NOM marketing agency-LOC work-ø do-CNJ be-IN-CNJ

'I’m working at a marketing agency….'

(HJ, L1=Korean; BW, L1=English)

(52) ney, ilpon-eyse-nun yosay yok-ul tul-e-cwu-nun

    well Japan-LOC-TC these days swear word-ACC hear-INF-give-REL
    alpa-ka sayngky-ess-eyo.
    part-time job-NOM arise-PST-DCL

'Well, these days in Japan we have a part-time job to listen to someone swearing…..'

(YN, L1=Japanese)

The procedure of coding the discourse factor was generally guided by two separate discourse
notions claimed to condition the naturalness of variable use of the case morphemes: focus and
discourse newness. What feeds into the coding categories, however, is different discourse
functions that a DP can serve in a given context. Since the categories of [new], [focus], and
[none] are generated by specific discourse functions, not by specific values for each discourse
notion, one may find that a DP may simultaneously fall into both [new] and [focus]. This
dissertation assumed that a DP serves one single discourse function that is the most salient in the
context (i.e., the coding decisions were made with the assumption that the categories are
mutually exclusive). For example, although an answer to a subject wh-question that denotes a discourse-new referent could be classified as [new], the study determined that the salient discourse function of the DP is assigning a particular value to the variable in an open proposition represented by a wh-question rather than introducing the new referent into a discourse, thereby coding it as [focus], not as [new]. This principle of [focus] over [new] was applied to other possible overlaps between categories.\textsuperscript{54}

5.2.2.3. Overview of coding system

The subject/object and non-subject/non-object tokens collected from the L1 and L2 speakers of Korean were coded for markedness and discourse features.

The procedure of coding the markedness factor was led by the markedness hierarchy that encompasses the semantic features for prototypical subject-/object-hood. Table 5.7 summarizes the coding categories stemming from the hierarchy.

\textsuperscript{54} The subcategories of [focus] do not produce such overlap since they correspond to specific values of one single discourse dimension. Thus, the definition of each subcategory renders them mutually exclusive.
The coding categories of the discourse factor represent distinct discourse functions that a DP serves in regard to the focus and discourse newness dimensions. Table 5.8 recapitulates the categories and subcategories used to code for the discourse factor.

(53) Table 5.7. Coding of the markedness factor: Coding categories

<table>
<thead>
<tr>
<th>Category</th>
<th>Additional annotation</th>
</tr>
</thead>
<tbody>
<tr>
<td>[1st person]</td>
<td></td>
</tr>
<tr>
<td>[2nd person]</td>
<td></td>
</tr>
<tr>
<td>[pronoun 3rd]</td>
<td></td>
</tr>
<tr>
<td>[proper noun 3rd]</td>
<td>proper noun-like</td>
</tr>
<tr>
<td>[human 3rd]</td>
<td>metonymic human</td>
</tr>
<tr>
<td>[animate 3rd]</td>
<td></td>
</tr>
<tr>
<td>[inanimate 3rd]</td>
<td></td>
</tr>
</tbody>
</table>

(54) Table 5.8. Coding of the discourse factor: Coding categories

<table>
<thead>
<tr>
<th>Category</th>
<th>Subcategory</th>
<th>Additional annotation</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>[new]</strong> (NOM only)</td>
<td></td>
<td>Co-occurring verb, <em>iss</em> (<em>exist, be</em>)</td>
</tr>
<tr>
<td>[focus]</td>
<td>[information focus]</td>
<td></td>
</tr>
<tr>
<td>[contrast]</td>
<td>[exclusive contrast]</td>
<td>[focus on DP]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>[focus on polarity]</td>
</tr>
<tr>
<td></td>
<td>[non-exclusive contrast]</td>
<td></td>
</tr>
<tr>
<td><strong>[none]</strong></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
CHAPTER 6

RESULTS: CORPUS STUDY

In this corpus study, 6.6 hours of discourse units presented in a South Korean TV show, *picengsanghooytam* (‘Non-summit’) were transcribed. The transcribed texts provided a total of 5,923 subject/object and non-subject/non-object tokens produced by four L1 English L2 Korean speakers (E-group), four L1 Japanese L2 Korean speakers (J-group), and four L1 Korean controls (K-group). The data tokens were coded for markedness and discourse features. In order to establish consistency of coding for the two independent variables, an independent coder evaluated the criteria by coding approximately 10% of the total data (600 tokens): the inter-coder reliability calculated with percent agreement was 95% for the markedness factor and 86.7% for the discourse factor. The coded tokens were then submitted to frequency analysis to observe the distribution of case marking and zero-marking according to the two independent variables.

The results of the corpus analysis are discussed in Chapter 6. The chapter is organized as follows. Section 6.1 reports the frequency distributions of case marking vs. zero-marking produced by the three L1 groups. Section 6.2 and section 6.3, respectively, discuss effects of the markedness factor and of the discourse factor on each group’s variable use of the case morphemes. Section 6.4 re-examines the results by exploring individual data. Section 6.5 summarizes the key results of the corpus study.

6.1. Frequency of case (un)marking

The total 2,420 tokens of K-group, 1,656 tokens of J-group, and 1,847 tokens of E-group were
classified according to whether they are case-marked or -unmarked. The breakdown of the subject/object and non-subject/non-object forms by the morphological realizations is given in Table 6.1.

(1) Table 6.1. Frequency of case (un)marking: Number of data tokens

<table>
<thead>
<tr>
<th>L1 group</th>
<th>Case</th>
<th>Subj.</th>
<th>Non-subj.</th>
<th>NOM</th>
<th>Obj.</th>
<th>Non-obj.</th>
<th>ACC</th>
<th>Row total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Marked</td>
<td>787</td>
<td>165</td>
<td>952</td>
<td>419</td>
<td>126</td>
<td>545</td>
<td>1497</td>
</tr>
<tr>
<td></td>
<td>Unmarked</td>
<td>217</td>
<td>126</td>
<td>343</td>
<td>298</td>
<td>292</td>
<td>580</td>
<td>923</td>
</tr>
<tr>
<td><strong>K-group</strong></td>
<td>Column total</td>
<td>1004</td>
<td>292</td>
<td>1295</td>
<td>707</td>
<td>418</td>
<td>1125</td>
<td>2420</td>
</tr>
<tr>
<td></td>
<td>Marked</td>
<td>518</td>
<td>123</td>
<td>641</td>
<td>316</td>
<td>94</td>
<td>410</td>
<td>1051</td>
</tr>
<tr>
<td></td>
<td>Unmarked</td>
<td>161</td>
<td>54</td>
<td>215</td>
<td>146</td>
<td>244</td>
<td>390</td>
<td>605</td>
</tr>
<tr>
<td><strong>J-group</strong></td>
<td>Column total</td>
<td>679</td>
<td>177</td>
<td>856</td>
<td>462</td>
<td>338</td>
<td>800</td>
<td>1656</td>
</tr>
<tr>
<td></td>
<td>Marked</td>
<td>458</td>
<td>75</td>
<td>533</td>
<td>231</td>
<td>68</td>
<td>299</td>
<td>832</td>
</tr>
<tr>
<td></td>
<td>Unmarked</td>
<td>313</td>
<td>89</td>
<td>402</td>
<td>284</td>
<td>329</td>
<td>613</td>
<td>1015</td>
</tr>
<tr>
<td><strong>E-group</strong></td>
<td>Column total</td>
<td>771</td>
<td>164</td>
<td>935</td>
<td>515</td>
<td>397</td>
<td>912</td>
<td>1847</td>
</tr>
</tbody>
</table>

The frequency analysis of the total number of case-marked and -unmarked of data tokens in each L1 group reveals that K-group and J-group produce similar distribution patterns, which E-group pattern diverges from. See Figure 6.1.
For both K- and J-groups, more than 60% of the data is marked by the nominative or accusative case morpheme, whereas less than 40% of the data is zero-marked. The distribution is not reproduced in E-group’s performance in that the rate of case marking (45%) is lower than that of zero marking (55%).

The distribution of case marking vs. zero-marking was further examined across the subsets of the data analyzed for this study. The general observation is that the patterns produced by K- and J-groups bear conspicuous resemblance to each other. That is, the rates of case marking in the two groups display close percentage values (for most subsets) and are significantly higher than the rate of case marking in E-group (for all subsets).

The distributional similarities between K-group and J-group are reproduced in both NOM and ACC data sets. The results are illustrated in Figure 6.2.
For all three groups, the rate of case marking is higher in the NOM data than in the ACC data. Regarding the NOM set, both K- and J-groups demonstrate a distribution in which more than 70% of the data is marked by the nominative case morpheme while less than 30% of the data is zero-marked. Note that the rate of nominative marking in E-group is lower (57.0%) than those for the two other groups. A similar result was obtained on the ACC data. For both K- and J-groups, accusative-marked and zero-marked forms are evenly distributed, whereas, for E-group, zero-marking is higher (67.2%) than accusative marking (32.8%).

The frequency of case (un)marking within the NOM set was further analyzed according to the subdivision of subjects vs. non-subjects. The findings are summarized in Figure 6.3.
As the majority of the NOM data for this study consists of subjects (77.5% for K-group, 79.3% for K-group, and 82.5% for E-group; see section 5.2.1.4), the case (un)marking pattern of subjects for each group is notably similar to the pattern of the entire NOM data while accompanying a slightly higher rate of case marking. For all three groups, the frequency of case marking in non-subjects is lower than in subjects. K- and J-groups attach the nominative morpheme to more than a half of the non-subjects, and it is noted that J-group involves a higher rate of the morpheme attachment (69.5%) than K-group does (56.7%). On the contrary, E-group attaches the nominative morpheme to non-subjects less frequently (45.7%) than the two other groups do.

The frequency of case (un)marking within the ACC set was further observed in terms of the sub-distinction of objects vs. non-objects. Figure 6.4 presents the results.
For all three groups, the frequency of accusative marking in non-objects is far lower than in objects. The ACC data for this study contains a significant number of non-object tokens (37.2% for K-group, 32.3% for K-group, and 43.5% for E-group; see section 5.2.1.4), which tend to be zero-marked. Thus, the frequency of accusative marking is found to be higher in objects than in the entire ACC data. Concerning the object marking pattern, K- and J-groups display an overall preference of case marking over zero-marking, and it is observed that J-group involves a higher rate of case marking (68.4%) than K-group (59.3%). On the other hand, E-group case-marks only 44.9% of the objects, leaving more than a half of the objects case-unmarked (55.1%). In regard to non-objects, although all three groups exhibit a noticeable bias towards zero-marking, the bias is less prominent in K- and J-groups than in E-group, as shown by the higher rates of accusative marking in K- and J-groups (30.1% and 27.8%, respectively) than that in E-group (17.0%).

Lastly, the study classified the case-marked and -unmarked data in terms of the distinction of variable case marking of subject/object arguments vs. variable morpheme attachment to non-subject/non-object items. The results are provided in Figure 6.5.
As expected, the frequency of case marking is higher in subjects/objects than in non-subjects/non-objects across the three groups. Both in subjects/objects and in non-subjects/non-objects, the general resemblance between K-group and J-group is replicated: the rates of case marking in K- and J-groups are similar and higher than that in E-group.

To summarize the results, for all three groups, the frequency of case marking is higher in NOM data than in ACC data (in non-subjects than in non-objects as well as in subjects than in objects); it is also higher in subjects/objects than in non-subjects/non-objects. The data reveals that the patterns produced by K- and J-groups bear distributional similarities. The analysis of the data subsets confirms that K- and J-groups use the case morphemes more frequently than E-group does. Although J-group occasionally demonstrates more frequent use of the morphemes than K-group does (e.g., for non-subjects, for objects), the two groups mostly show similar rates of case marking across the subsets.
6.2. Effects of the markedness factor

The subject/object and non-subject/non-object tokens were coded for the markedness factor, where each token received one of the following labels: [1st person], [2nd person], [pronoun 3rd], [proper noun 3rd], [proper noun 3rd]-like, [animate 3rd], [animate 3rd]-metonymic human, [inanimate 3rd]. The categories are arranged in line with the semantic features responsible for stereotypic correlations between grammatical functions and thematic roles manifested in the markedness hierarchy.

If the markedness factor has significant effects on the case (un)marking patterns produced by each L1 group, the group should demonstrate (i) a subject (un)marking pattern in which zero-marking is more frequent for DPs towards the left of the hierarchy; (ii) an object (un)marking pattern in which zero-marking is more frequent for DPs towards the right of the hierarchy, and (iii) such directional patterns are observed limitedly to the subsets of the argument data (i.e., only when the case morpheme use is associated with marking of subject/object argument roles).

Frequency analysis of case (un)marking was conducted against the four subsets of the data (i.e., subjects, objects, non-subjects, and non-objects) for each L1 group. The analysis shows that no group, across the subsets, produced a distribution that supports a significant effect of the markedness factor on the naturalness of case (un)marking. The following discussion deals with the results obtained from each data subset across the three L1 groups.

6.2.1. Subjects

The breakdown of the case-marked and -unmarked subject forms by the categories for the markedness factor is provided in Table 6.2.
Table 6.2. Effects of the markedness factor on case (un)marking: Number of subject tokens

<table>
<thead>
<tr>
<th>L1 group</th>
<th>Case</th>
<th>[1ˢᵗ]</th>
<th>[2ⁿᵈ]</th>
<th>[p3ʳᵈ]</th>
<th>[pn3ʳᵈ]</th>
<th>[pn3ʳᵈ]-like</th>
<th>[h3ʳᵈ]</th>
<th>[a3ʳᵈ]</th>
<th>[a3ʳᵈ]-mh</th>
<th>[i3ʳᵈ]</th>
<th>Row total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Marked</td>
<td></td>
<td>77</td>
<td>4</td>
<td>60</td>
<td>102</td>
<td>24</td>
<td>91</td>
<td>5</td>
<td>11</td>
<td>413</td>
<td>787</td>
</tr>
<tr>
<td>Unmarked</td>
<td></td>
<td>28</td>
<td>5</td>
<td>29</td>
<td>22</td>
<td>2</td>
<td>27</td>
<td>1</td>
<td>4</td>
<td>99</td>
<td>217</td>
</tr>
<tr>
<td><strong>K-group</strong></td>
<td>Column total</td>
<td>105</td>
<td>9</td>
<td>89</td>
<td>124</td>
<td>26</td>
<td>118</td>
<td>6</td>
<td>15</td>
<td>512</td>
<td>1004</td>
</tr>
<tr>
<td>Marked</td>
<td></td>
<td>50</td>
<td>2</td>
<td>30</td>
<td>22</td>
<td>13</td>
<td>140</td>
<td>3</td>
<td>7</td>
<td>251</td>
<td>518</td>
</tr>
<tr>
<td>Unmarked</td>
<td></td>
<td>11</td>
<td>2</td>
<td>14</td>
<td>5</td>
<td>9</td>
<td>30</td>
<td>2</td>
<td>1</td>
<td>87</td>
<td>161</td>
</tr>
<tr>
<td><strong>J-group</strong></td>
<td>Column total</td>
<td>61</td>
<td>4</td>
<td>44</td>
<td>27</td>
<td>22</td>
<td>170</td>
<td>5</td>
<td>8</td>
<td>338</td>
<td>679</td>
</tr>
<tr>
<td>Marked</td>
<td></td>
<td>77</td>
<td>3</td>
<td>30</td>
<td>36</td>
<td>10</td>
<td>134</td>
<td>10</td>
<td>9</td>
<td>149</td>
<td>458</td>
</tr>
<tr>
<td>Unmarked</td>
<td></td>
<td>37</td>
<td>3</td>
<td>34</td>
<td>29</td>
<td>5</td>
<td>39</td>
<td>1</td>
<td>1</td>
<td>164</td>
<td>313</td>
</tr>
<tr>
<td><strong>E-group</strong></td>
<td>Column total</td>
<td>114</td>
<td>6</td>
<td>64</td>
<td>65</td>
<td>15</td>
<td>173</td>
<td>11</td>
<td>10</td>
<td>313</td>
<td>771</td>
</tr>
</tbody>
</table>

Note. [1ˢᵗ] = [1ˢᵗ person]; [2ⁿᵈ] = [2ⁿᵈ person]; [p3ʳᵈ] = [pronoun 3ʳᵈ]; [pn3ʳᵈ] = [proper noun 3ʳᵈ]; [pn3ʳᵈ]-like = [proper noun 3ʳᵈ]-like; [a3ʳᵈ] = [animate 3ʳᵈ]; [a3ʳᵈ]-mh = [animate 3ʳᵈ]-metonymic human; [i3ʳᵈ] = [inanimate 3ʳᵈ]

The frequency distribution for each group was observed based on the rate (%) of case (un)marking. For this purpose, certain (sub)categories were collapsed into one category. [1ˢᵗ] and [2ⁿᵈ] were collapsed into a broader category, which is equivalent to ‘local person’ originally expressed in the markedness hierarchy (cf., (31) in section 5.2.2.1). [pn3ʳᵈ] and [pn3ʳᵈ]-like were treated as one category that can be subsumed under Lee’s (2006b) label of ‘strongly definite’.  

55 See footnote 10.
non-pronominal DPs. [a3rd] and [a3rd]-mh were collapsed together, as well. This was because these (sub)categorizes involve a rather small number of data tokens, and the analysis conducted with the full (sub)categorizes does not reveal a pattern that supports a significant effect of the factor, either (see Appendix C for supplementary analysis of the markedness factor on case (un)marking of subjects and objects).

K-group’s performance did not exhibit the previously-suggested nativelike sensitivity to the markedness factor that predicts higher frequency of zero-marking among subjects towards the left of the hierarchy. The results are illustrated in Figure 6.6.

(8)  Figure 6.6. Effects of the markedness factor on frequency (%) of case (un)marking: Subjects.K-group

(χ² = 14.47, p < .05)
The results revealed that K-group produced a statistically significant distribution that does not match with a pattern predicted by the proposed hierarchy.\textsuperscript{56} As seen above, the rates of zero-marking across the categories do not display any conclusive pattern in which zero-marking is more frequent in categories towards the left of the x-axis than in those on the right. Note that [p3\textsuperscript{rd}]+[p3\textsuperscript{rd}]-like involves the lowest rate of zero-marking, even lower than the rightmost category, [i3\textsuperscript{rd}]. It overall appears that the frequency of case marking is far higher than that of zero-marking in all categories.

Recall that Lee’s (2006b) corpus analysis of the definiteness effects on case (un)marking involves three categories of definiteness: [pronoun] and [name] as ‘strongly definite’, and [other] as ‘low definite’ (cf., Table 2.4 in section 2.2.2.2). In her study, the frequency of zero-marking is higher in the two strongly definite categories (47.6\% for [pronoun] and 46.8\% for [name]) than in the low definite category (38.3\% for [other]). As implied by the lowest rate of zero-marking in [p3\textsuperscript{rd}]+[p3\textsuperscript{rd}]-like, Lee’s (2006b) result is not replicated in this study. Figure 6.7 shows the frequency of case (un)marking when the subject data was reclassified in an analogous way to Lee’s three-way distinction: [1\textsuperscript{st}]+[2\textsuperscript{nd}]+[p3\textsuperscript{rd}] and [p3\textsuperscript{rd}]+[p3\textsuperscript{rd}]-like as ‘strongly definite’ and the rest of [h3\textsuperscript{rd}]+[a3\textsuperscript{rd}]+[a3\textsuperscript{rd}]-mh+[i3\textsuperscript{rd}]-mh as ‘low definite’.

\textsuperscript{56} Each group performed differently on the markedness factor throughout the subanalyses, where they were found to produce either a random distribution, or a statistically significant distribution that does not conform to the tested hierarchy. Both types of distributions were interpreted as evidence to disconfirm the markedness effects on the data examined in this study.
(9) Figure 6.7. Effects of the markedness factor on frequency (%) of case (un)marking:

Subjects.K-group.subanalysis of definiteness effects

Only in one category of strongly definite DPs (i.e., [1st]+[2nd]+[3rd]) did zero-marking occur significantly more frequently than in low definite DPs. Contrary to Lee’s result, the zero-marking rate in the other strongly definite category, [pn3rd]+[pn3rd]-like was not significantly higher than that in low definite DPs.

Likewise, part of the subject data can be compared with Lee’s (2006b) analysis of the animacy effects on subject case (un)marking (cf., Table 2.3 in section 2.2.2.1), where the frequency of zero-marking for [human] and [animate] subjects is higher than that for [inanimate] subjects. Although Figure 6.6 shows that the frequency of zero-marking for [h3rd] subjects (22.9%) and [a3rd]+[a3rd]-mh subjects (23.8%) is higher than for [i3rd] subjects (19.3%), the differences were not significant ($\chi^2 = .93, p = .627$).
Neither E-group nor J-group produced a subject case (un)marking pattern predicted by the effects of the markedness factor. The frequency analysis for J-group was not significant ($\chi^2 = 6.96, p = .223$). See Figure 6.8 for the subject (un)marking pattern by E-group.

(10) Figure 6.8. Effects of the markedness factor on frequency (%) of case (un)marking:

Subject.E-group

E-group’s case (un)marking pattern exhibits significant differences across the categories that do not comply with the hierarchy. That is, relatively high frequency of zero-marking is linked to [p$^{3^{rd}}$], [p$^{3^{rd}}$]+[p$^{3^{rd}}$]-like, and [i$^{3^{rd}}$] subjects. These 3$^{rd}$ person subject DPs were previously claimed to involve lower frequency of case drop than [1$^{st}$]+[2$^{nd}$] DPs; [i$^{3^{rd}}$] DPs were predicted to display lower frequency of case drop than [a$^{3^{rd}}$]+[a$^{3^{rd}}$]-mh.
All in all, it is concluded that markedness effects on the naturalness of subject case (un)marking in Korean were not borne out in this study: the effects were absent not only in the L2 speakers of E- and J-groups but also in the L1 controls of K-group.

6.2.2. Objects

The classification of the case-marked and -unmarked object forms according to the markedness categories is given in Table 6.3.

(11) Table 6.3. Effects of the markedness factor on case (un)marking: Number of object tokens

<table>
<thead>
<tr>
<th>L1 group</th>
<th>Case</th>
<th>[1st]</th>
<th>[2nd]</th>
<th>[p3rd]</th>
<th>[pn3rd]</th>
<th>[pn3rd]-like</th>
<th>[h3rd]</th>
<th>[a3rd]</th>
<th>[a3rd]-mh</th>
<th>[i3rd]</th>
<th>Row total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Marked</td>
<td>4</td>
<td>0</td>
<td>24</td>
<td>25</td>
<td>16</td>
<td>18</td>
<td>2</td>
<td>0</td>
<td>330</td>
<td>419</td>
</tr>
<tr>
<td></td>
<td>Unmarked</td>
<td>2</td>
<td>0</td>
<td>21</td>
<td>14</td>
<td>12</td>
<td>11</td>
<td>7</td>
<td>0</td>
<td>221</td>
<td>288</td>
</tr>
</tbody>
</table>

| K-group  | Column total | 6     | 0     | 45     | 39      | 29          | 29     | 9      | 0         | 551    | 771       |
| Marked   | 4 2 37 5 16 23 5 0 224 316|
| Unmarked | 0 0 5 2 5 6 1 0 127 146|

| J-group  | Column total | 4     | 2     | 42     | 7       | 21          | 29     | 6      | 0         | 351    | 462       |
| Marked   | 6 2 21 2 5 23 4 0 168 231|
| Unmarked | 2 0 18 15 3 14 3 0 229 284|

| E-group  | Column total | 8     | 2     | 39     | 17      | 8           | 37     | 7      | 0         | 397    | 515       |

Frequency analysis of case (un)marking across the three L1 groups provides no evidence that the naturalness of object case (un)marking in Korean can be described in terms of the
hierarchically organized markedness categories. The frequency of object case (un)marking for each group is provided in Figure 6.9. As seen below, each group displayed a statistically significant pattern, but none of them involve relatively higher frequency of zero-marking in categories towards the right of the x-axis than in those on the left.
(χ² = 14.47, p < .05 for K-group; χ² = 16.50, p < .01 for J-group; χ² = 15.08, p < .05 for E-group)

(12) Figure 6.9. Effects of the markedness factor on frequency (%) of case (un)marking:

6.2.3. Non-subjects and non-objects

Should markedness effects on subject/object case (un)marking be borne out, frequency analysis of (non-)occurrence of the case morphemes for non-subjects/non-objects is conducive to bolstering the presence of the effects in the results. In other words, the absence of the effects on the subsets of non-subject/non-object data would indicate that the choice of subject/object case (un)marking is indeed associated with a mapping of grammatical functions and thematic roles.

In this corpus investigation, the results obtained from the subsets of subject/object data render the frequency analysis of (non-)occurrence of the case morphemes for non-subjects/non-objects ineffectual: in whichever case, it is claimed that the markedness factor had no significant effect on variable use of the nominative and accusative case morphemes across the three groups.

As observational facts from the research, the breakdown of the nominative-marked and -unmarked non-subject forms is reported in Table 6.4.
Table 6.4. Effects of the markedness factor on case (un)marking: Number of non-subject tokens

<table>
<thead>
<tr>
<th>L1 group</th>
<th>Case</th>
<th>[1ˢʰ]</th>
<th>[2ᵗʰ]</th>
<th>[p³ʰ]</th>
<th>[pn³ʰ]</th>
<th>[p³ʰ]-like</th>
<th>[h³ⁿ]</th>
<th>[a³ʰ]</th>
<th>[a³ʰ]-mh</th>
<th>[i³ⁿ]</th>
<th>Row total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Marked</td>
<td>21</td>
<td>1</td>
<td>29</td>
<td>10</td>
<td>3</td>
<td>15</td>
<td>0</td>
<td>1</td>
<td>104</td>
<td>165</td>
</tr>
<tr>
<td></td>
<td>Unmarked</td>
<td>8</td>
<td>4</td>
<td>6</td>
<td>3</td>
<td>4</td>
<td>12</td>
<td>0</td>
<td>0</td>
<td>89</td>
<td>126</td>
</tr>
<tr>
<td>K-group</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Column total</td>
<td>10</td>
<td>5</td>
<td>35</td>
<td>13</td>
<td>7</td>
<td>27</td>
<td>0</td>
<td>1</td>
<td>193</td>
<td>291</td>
</tr>
<tr>
<td></td>
<td>Marked</td>
<td>5</td>
<td>0</td>
<td>26</td>
<td>0</td>
<td>6</td>
<td>9</td>
<td>0</td>
<td>0</td>
<td>77</td>
<td>123</td>
</tr>
<tr>
<td></td>
<td>Unmarked</td>
<td>7</td>
<td>0</td>
<td>7</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>37</td>
<td>54</td>
</tr>
<tr>
<td>J-group</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Column total</td>
<td>12</td>
<td>0</td>
<td>33</td>
<td>0</td>
<td>7</td>
<td>11</td>
<td>0</td>
<td>0</td>
<td>114</td>
<td>177</td>
</tr>
<tr>
<td></td>
<td>Marked</td>
<td>1</td>
<td>0</td>
<td>6</td>
<td>3</td>
<td>1</td>
<td>11</td>
<td>0</td>
<td>0</td>
<td>53</td>
<td>75</td>
</tr>
<tr>
<td></td>
<td>Unmarked</td>
<td>3</td>
<td>0</td>
<td>6</td>
<td>5</td>
<td>1</td>
<td>7</td>
<td>1</td>
<td>0</td>
<td>66</td>
<td>89</td>
</tr>
<tr>
<td>E-group</td>
<td>Column total</td>
<td>4</td>
<td>0</td>
<td>12</td>
<td>2</td>
<td>2</td>
<td>18</td>
<td>1</td>
<td>0</td>
<td>119</td>
<td>164</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The numbers in parentheses indicate Ns obtained when the dummy-like first nominative-markable pronouns in the multiple nominative-type structure are excluded (see section 5.2.2.1). Using the parenthesized Ns of [p³ʰ], the frequency distribution between nominative-marked and -unmarked non-subject forms was observed according to the markedness categories. The analysis revealed no significant difference ($\chi^2 = 9.28$, $p = .098$ for K-group; $\chi^2 = 5.75$, $p = .331$ for J-
group; $\chi^2 = 3.65, p = .601$ for E-group) (see Appendix D for supplementary analysis of the markedness factor on (non-)occurrence of the case morpheme for non-subjects).

The classification of the accusative-marked and -unmarked non-objects is also provided in Table 6.5.

(14) Table 6.5. Effects of the markedness factor on case (un)marking: Number of non-object tokens

<table>
<thead>
<tr>
<th>L1 group</th>
<th>Case</th>
<th>[1st]</th>
<th>[2nd]</th>
<th>[p3rd]</th>
<th>[pn3rd]</th>
<th>[pn3rd]-like</th>
<th>[h3rd]</th>
<th>[a3rd]</th>
<th>[a3rd]-mh</th>
<th>[i3rd]</th>
<th>Row total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Marked</td>
<td></td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>123</td>
</tr>
<tr>
<td>Unmarked</td>
<td></td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>292</td>
</tr>
<tr>
<td>K-group</td>
<td>Column total</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>415</td>
</tr>
<tr>
<td>Marked</td>
<td></td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>93</td>
</tr>
<tr>
<td>Unmarked</td>
<td></td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>243</td>
</tr>
<tr>
<td>J-group</td>
<td>Column total</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>336</td>
</tr>
<tr>
<td>Marked</td>
<td></td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>68</td>
</tr>
<tr>
<td>Unmarked</td>
<td></td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>329</td>
</tr>
<tr>
<td>E-group</td>
<td>Column total</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>397</td>
</tr>
</tbody>
</table>

No frequency analysis was conducted against the non-object data for almost all tokens in the subset turned out to be [i3rd].

---

57 This is attributed to the prevalence of verbal nouns in compound verbs in the subset.
6.3. Effects of the discourse factor

The subject/object and non-subject/non-object tokens were coded for the other independent variable, the discourse factor. The coding categories represent different discourse functions that a DP serves in a given context. Table 6.6 reiterates the (sub)cATEGORIES and additional documentation implemented in the coding procedure.

(15) Table 6.6. Coding of the discourse factor: Coding categories

<table>
<thead>
<tr>
<th>Category</th>
<th>Subcategory</th>
<th>Additional annotation</th>
</tr>
</thead>
<tbody>
<tr>
<td>[new] (NOM only)</td>
<td></td>
<td>Co-occurring verb, <em>iss</em> (<em>exist, be</em>)</td>
</tr>
<tr>
<td>[focus]</td>
<td>[information focus]</td>
<td></td>
</tr>
<tr>
<td></td>
<td>[contrast]</td>
<td></td>
</tr>
<tr>
<td></td>
<td>[exclusive contrast]</td>
<td></td>
</tr>
<tr>
<td></td>
<td>[focus on DP]</td>
<td>Parallel under higher focus</td>
</tr>
<tr>
<td></td>
<td>[focus on polarity]</td>
<td></td>
</tr>
<tr>
<td></td>
<td>[non-exclusive contrast]</td>
<td></td>
</tr>
<tr>
<td>[none]</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

If the discourse factor has significant effects on the case (un)marking patterns of each L1 group, the group should demonstrate a distribution in which case marking is more frequent for [new] and [focus] DPs than for [none] DPs. Unlike the markedness factor, effects of the discourse factor are not to be confined to the subsets of subject/object data; they may equally extend to the subsets of non-subject/non-object data since the discourse motivations do not imply a particular correlation with distinct grammatical functions.
6.3.1. Overall analysis

The total tokens of each L1 group were classified according to the three-way broad distinction of [new], [focus], and [none], and the frequency distribution of case marking and zero-marking was observed. See Table 6.7 for the breakdown of case-marked and -unmarked tokens by the discourse categories.

(16) Table 6.7. Effects of the discourse factor on case (un)marking: Number of data tokens

<table>
<thead>
<tr>
<th>L1 group</th>
<th>Case</th>
<th>[new]</th>
<th>[focus]</th>
<th>[none]</th>
<th>Row total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Marked</td>
<td>267</td>
<td>368</td>
<td>862</td>
<td>1497</td>
</tr>
<tr>
<td></td>
<td>Unmarked</td>
<td>87</td>
<td>128</td>
<td>708</td>
<td>923</td>
</tr>
<tr>
<td><strong>K-group</strong></td>
<td>Column total</td>
<td>354</td>
<td>496</td>
<td>1570</td>
<td>2420</td>
</tr>
<tr>
<td></td>
<td>Marked</td>
<td>230</td>
<td>206</td>
<td>615</td>
<td>1051</td>
</tr>
<tr>
<td></td>
<td>Unmarked</td>
<td>89</td>
<td>77</td>
<td>439</td>
<td>605</td>
</tr>
<tr>
<td><strong>J-group</strong></td>
<td>Column total</td>
<td>319</td>
<td>283</td>
<td>1054</td>
<td>1656</td>
</tr>
<tr>
<td></td>
<td>Marked</td>
<td>150</td>
<td>204</td>
<td>478</td>
<td>832</td>
</tr>
<tr>
<td></td>
<td>Unmarked</td>
<td>146</td>
<td>180</td>
<td>689</td>
<td>1015</td>
</tr>
<tr>
<td><strong>E-group</strong></td>
<td>Column total</td>
<td>296</td>
<td>384</td>
<td>1167</td>
<td>1847</td>
</tr>
<tr>
<td></td>
<td>(16.0%)</td>
<td>(20.8%)</td>
<td>(63.2%)</td>
<td>(100%)</td>
<td></td>
</tr>
</tbody>
</table>
While [none] DPs take up the greater part of the data, at least 35% of the total data for each group turns out to be [new] and [focus] DPs, which provides an adequate basis for further analysis.

Frequency analysis of case (un)marking was conducted against the total tokens of subjects/objects and non-subjects/non-objects for each L1 group. The overall result is that the rate of case marking is higher for [new] and [focus] DPs than for [none] DPs, which entails that case drop is more frequent for [none] DPs than for [new] or [focus] DPs. The result is consistent across the three groups, and the distributional differences reached statistical significance ($\chi^2 = 91.78$, $p < .001$ for K-group; $\chi^2 = 32.77$, $p < .001$ for J-group; $\chi^2 = 21.79$, $p < .001$ for E-group). Figure 6.10 summarizes the findings.

(17) Figure 6.10. Effects of the discourse factor on frequency (%) of case (un)marking:

Total data of each L1 group
Figure 6.10 illustrates the frequency of case marking in the three discourse conditions by each group while implying the reverse frequency distribution of zero-marking. For all three groups, the rates of case marking for [new] and [focus] DPs are higher than for [none] DPs and display close percentage values. It is noted that K- and J-groups’ patterns are similar in that they exhibit a higher rate of case marking in all three conditions than E-group does, which is consistent with the comparability between the two groups in the overall frequency of case marking (section 6.1).

As for E-group, their overall preference of zero-marking over case marking (55% vs. 45%, cf., Figure 6.1) seems to be reflected in the result in that they are not clearly shown to be more inclined to case-mark [new] and [focus] DPs than to zero-mark them. However, the bias towards zero-marking is found to be reduced for [new] and [focus] DPs and enhanced for [none] DPs, thereby producing a pattern that conforms to the generalization that [new] and [focus] DPs are more likely to be case-marked than [none] DPs.

6.3.2. Further analysis

The data was further analyzed both in terms of the coding subcategories and the subsets of the data. Prior to the discussion of the analysis, the tables below (Table 6.8 for K-group; Table 6.9 for J-group; Table 6.10 for E-group) are provided to report the number of case-marked and -unmarked tokens found in each subcategory in each data subset.
Table 6.8. Effects of the discourse factor on case (un)marking for K-group: Number of data tokens

<table>
<thead>
<tr>
<th>K-group</th>
<th>Subset</th>
<th>Case</th>
<th>[new]</th>
<th>[focus]</th>
<th>[none]</th>
<th>Row total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>[new]</td>
<td>[info f]</td>
<td>[cont]</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>[excl cont]</td>
<td>[n-excl cont]</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>[f on DP]</td>
<td>[f on pol]</td>
<td></td>
</tr>
<tr>
<td>NOM</td>
<td>Subj.</td>
<td>Marked</td>
<td>57</td>
<td>1</td>
<td>84</td>
<td>48</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>172</td>
<td></td>
<td>75</td>
<td>350</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>73</td>
<td>5</td>
<td>87</td>
<td>476</td>
</tr>
<tr>
<td></td>
<td>Unmarked</td>
<td></td>
<td>16</td>
<td>4</td>
<td>3</td>
<td>12</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>38</td>
<td></td>
<td>12</td>
<td>18</td>
</tr>
<tr>
<td></td>
<td>Column total</td>
<td></td>
<td>73</td>
<td>5</td>
<td>87</td>
<td>66</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Non-subj</td>
<td>Marked</td>
<td>0</td>
<td>0</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>0</td>
<td></td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>71</td>
<td>1</td>
<td>3</td>
<td>11</td>
</tr>
<tr>
<td></td>
<td>Unmarked</td>
<td></td>
<td>0</td>
<td>0</td>
<td>25</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>33</td>
<td></td>
<td>25</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td>Column total</td>
<td></td>
<td>0</td>
<td>0</td>
<td>70</td>
<td>11</td>
</tr>
<tr>
<td>ACC</td>
<td>Obj.</td>
<td>Marked</td>
<td>0</td>
<td>0</td>
<td>39</td>
<td>32</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>0</td>
<td></td>
<td>39</td>
<td>32</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>71</td>
<td>1</td>
<td>70</td>
<td>11</td>
</tr>
<tr>
<td></td>
<td>Unmarked</td>
<td></td>
<td>N/A</td>
<td>4</td>
<td>18</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td></td>
<td>18</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td>Column total</td>
<td></td>
<td>N/A</td>
<td>4</td>
<td>57</td>
<td>40</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td></td>
<td>57</td>
<td>40</td>
</tr>
<tr>
<td></td>
<td>Non-obj.</td>
<td>Marked</td>
<td>0</td>
<td>0</td>
<td>13</td>
<td>11</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>0</td>
<td></td>
<td>13</td>
<td>11</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>18</td>
<td>29</td>
<td>18</td>
<td>14</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>0</td>
<td></td>
<td>29</td>
<td>14</td>
</tr>
<tr>
<td></td>
<td>Column total</td>
<td></td>
<td>0</td>
<td>0</td>
<td>29</td>
<td>29</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>0</td>
<td></td>
<td>29</td>
<td>14</td>
</tr>
</tbody>
</table>

Note. [new]-iss = [new] with the co-occurring verb iss (‘exist, be’); [new]-rest = [new] that is not [new]-iss; [info f] = [information focus]; [cont] = [contrast]; [excl cont] = [exclusive contrast]; [n-excl cont] = [non-exclusive contrast]; [f on DP] = [focus on DP]; [f on pol] = [focus on polarity]
Table 6.9. Effects of the discourse factor on case (un)marking for J-group: Number of data tokens

<table>
<thead>
<tr>
<th>J-group</th>
<th>Subset</th>
<th>Case</th>
<th>[new]</th>
<th>[focus]</th>
<th>[none]</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>[new]</td>
<td>[cont]</td>
<td>[none]</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>[excl cont]</td>
<td>[n-excl cont]</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>[fon DP]</td>
<td>[fon pol]</td>
</tr>
<tr>
<td>NOM</td>
<td>Subj.</td>
<td>Marked</td>
<td>59</td>
<td>5</td>
<td>21</td>
</tr>
<tr>
<td></td>
<td>Unmarked</td>
<td>36</td>
<td>3</td>
<td>12</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td>Column total</td>
<td>95</td>
<td>8</td>
<td>33</td>
<td>36</td>
</tr>
<tr>
<td>Non-subj.</td>
<td>Marked</td>
<td>0</td>
<td>5</td>
<td>18</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>Unmarked</td>
<td>0</td>
<td>6</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>Column total</td>
<td>0</td>
<td>61</td>
<td>24</td>
<td>2</td>
</tr>
<tr>
<td>ACC</td>
<td>Obj.</td>
<td>Marked</td>
<td>0</td>
<td>20</td>
<td>22</td>
</tr>
<tr>
<td></td>
<td>Unmarked</td>
<td>N/A</td>
<td>1</td>
<td>5</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td>Column total</td>
<td>1</td>
<td>27</td>
<td>27</td>
<td>19</td>
</tr>
<tr>
<td>Non-obj.</td>
<td>Marked</td>
<td>0</td>
<td>6</td>
<td>9</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>Unmarked</td>
<td>N/A</td>
<td>3</td>
<td>20</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Column total</td>
<td>0</td>
<td>9</td>
<td>29</td>
<td>3</td>
</tr>
</tbody>
</table>

Parallel under higher focus: 64
(20) Table 6.10. Effects of the discourse factor on case (un)marking for E-group: Number of data tokens

<table>
<thead>
<tr>
<th>E-group</th>
<th>Case</th>
<th>[new]</th>
<th>[focus]</th>
<th>[none]</th>
<th>Row total</th>
</tr>
</thead>
<tbody>
<tr>
<td>NOM</td>
<td>Marked</td>
<td>27</td>
<td>42</td>
<td>206</td>
<td>458</td>
</tr>
<tr>
<td>Subj.</td>
<td>Unmarked</td>
<td>36</td>
<td>18</td>
<td>122</td>
<td>313</td>
</tr>
<tr>
<td>Column total</td>
<td>63</td>
<td>60</td>
<td>65</td>
<td>328</td>
<td>771</td>
</tr>
<tr>
<td>Non-subj.</td>
<td>Marked</td>
<td>0</td>
<td>5</td>
<td>28</td>
<td>75</td>
</tr>
<tr>
<td></td>
<td>Unmarked</td>
<td>0</td>
<td>2</td>
<td>31</td>
<td>89</td>
</tr>
<tr>
<td>Column total</td>
<td>0</td>
<td>7</td>
<td>2</td>
<td>59</td>
<td>164</td>
</tr>
<tr>
<td>ACC Obj.</td>
<td>Marked</td>
<td>N/A</td>
<td>24</td>
<td>191</td>
<td>231</td>
</tr>
<tr>
<td>Unmarked</td>
<td>4</td>
<td>21</td>
<td>8</td>
<td>431</td>
<td>515</td>
</tr>
<tr>
<td>Column total</td>
<td>4</td>
<td>35</td>
<td>10</td>
<td>431</td>
<td>515</td>
</tr>
<tr>
<td>Non-obj.</td>
<td>Marked</td>
<td>N/A</td>
<td>6</td>
<td>53</td>
<td>68</td>
</tr>
<tr>
<td>Unmarked</td>
<td>0</td>
<td>26</td>
<td>3</td>
<td>296</td>
<td>329</td>
</tr>
<tr>
<td>Column total</td>
<td>0</td>
<td>33</td>
<td>5</td>
<td>349</td>
<td>397</td>
</tr>
</tbody>
</table>

Based on the breakdowns presented in Table 6.8 to 6.10, the DPs in [new] and in [focus] conditions for each L1 group were further examined in terms of the respective sub-distinctions across the data subsets.

6.3.2.1. Effects of [new] on case (un)marking

As discussed in section 5.2.2.2, coding of the [new] category concerned the NOM data only. The
general finding shown by Figure 6.10 is that case marking is more frequent for [new] DPs than for [none] DPs in all three L1 groups (75.4% > 58.3% in K-group; 72.1% > 54.9% in J-group; 50.7% > 41.0% in E-group).

Further analysis of nominative (un)marking for [new] DPs was conducted against the subsets of the subject and non-subject data. The result from K-group indicates that nominative (un)marking pattern for [new] DPs discriminates between subjects and non-subjects: the rate of nominative marking for [new] subjects (80.9%) is significantly higher than for [new] non-subjects (53.5%) ($\chi^2 = 22.99$, $p < .001$). The result suggests that K-group is more sensitive to the appropriateness of case marking discourse [new] entities when they occur as the subject of a sentence than when they do not. In contrast, J- and E-groups’ performance across the two subsets bears no significant difference: J-group nominative-marks 73.3% of [new] subjects and 67.2% of [new] non-subjects ($\chi^2 = .90$, $p = .344$), and E-group nominative-marks 51.4% of [new] subjects and 47.5% of [new] non-subjects ($\chi^2 = .31$, $p = .581$).

Part of the [new] DPs were annotated as [new]-iss, which represents discourse-new information that occurs with the verb *iss* (‘exist, be’). As seen in Table 6.8 to 6.10, instances of [new]-iss were found only in the subset of subject data. The subject data was examined in terms of the frequency of case (un)marking for [new]-iss vs. [new]-rest. The results across the three groups display no tendency that case marking is more frequent for [new]-iss subjects than for [new]-rest subjects. No significant difference was found in K- and E-groups’ performance: the rate of case marking for [new]-iss vs. [new]-rest is 78.1% vs. 81.9% in K-group ($\chi^2 = .51$, $p = .474$), and 42.9% vs. 54.5% in E-group ($\chi^2 = 2.55$, $p = .110$). J-group’s pattern (62.1% vs. 79.8%) was shown to be significant ($\chi^2 = 9.54$, $p < .01$).
In summary, the discourse function captured by the [new] category had a significant effect on variable production of the nominative case particle by the three L1 groups in that the particle occurs more frequently with [new] DPs than with [none] DPs. K-group’s performance involved a higher rate of nominative marking for [new] DPs when they function as the subject of a sentence than when they occur as a non-subject, whereas the tendency was not observed in J- and E-groups’ performance. In addition, it turned out that no group produced a pattern in which [new]-iss subjects are more inclined towards case marking than [new]-rest subjects. A significant difference in the opposite direction was noted in J-group, however.

6.3.2.2. Effects of [focus] on case (un)marking

The total data classified as [focus] and as [none] for each L1 group (2,066 tokens for K-group; 1,337 tokens for J-group; 1,551 tokens for E-group) underwent a further examination to determine whether the various subtypes of focus have distinct effects on the case (un)marking patterns. The [focus] DPs were divided into the subcategories of [information focus], [focus on DP], [focus on polarity], and [non-exclusive contrast]. Then, the distribution of case marking and zero-marking across five discourse conditions was observed: [information focus], [focus on DP], [focus on polarity], [non-exclusive contrast], and [none].

The analysis of the subject/object and non-subject/non-object data counted altogether indicated that for all three groups, the subtypes of [focus] exert different effects on the case (un)marking patterns. That is, all three groups revealed varying degrees of bias towards case marking according to the subcategories of [focus] while the specific patterns of variation of the degrees are not uniform between the L1 speakers of K-group and the L2 speakers of J- and E-
groups. The findings are illustrated in Figure 6.11, below which is presented a ranking of the different degrees of bias towards case marking across the five discourse conditions for each group.

![Figure 6.11. Effects of [focus] on frequency (%) of case (un)marking: Total data of each L1 group subcategories](image)

Bias towards case marking by the approximate percentage values:

- K-group: [f on DP] > [f on pol] > [n-excl cont] > [none] > [info f]
- J-group: [f on DP] > [n-excl cont] > [f on pol] > [none] > [info f]
- E-group: [f on DP] > [n-excl cont] > [f on pol] > [none] > [info f]

(21) Figure 6.11. Effects of [focus] on frequency (%) of case (un)marking:

For all three groups, the frequency of case marking is the highest in [f on DP]. It is also observed that K- and J-groups show similarly higher rates of case marking in [f on DP] than E-group.
All three groups display the lowest frequency of case marking in [info f]. Despite its apparent focus status, case marking in [info f] is even less frequent than in [none] as well as in the other subtypes of focus subsumed under [contrast]. It is noted that K-group case-marks only 10% of [info f] DPs, for example. Although this could be interpreted as consistent with the previous claim that the notion of focus pertinent to variable use of the case morphemes in Korean is a particular focus type that conveys contrastive salience (section 2.2.2.3), the result may not be considered conclusive. A caution can be brought to the fact that the number of [info f] tokens examined is too small (10 tokens for K-group; 9 tokens for J-group; 12 tokens for E-group) to make a confident generalization.

K-group’s distribution reveals that the frequency of case marking is higher in [f on pol] than in [n-excl cont], whereas the reverse pattern is found in J- and E-groups’ distributions. More specifically, the distribution of case marking for K-group involves a gradual decrease in the sequence of [f on DP], [f on pol], [n-excl cont], and [none], resulting in a pattern in which case marking is more frequent in both subcategories of [exclusive contrast] (i.e., [f on DP] and [f on pol]) than in [n-excl cont]. J- and E-groups’ distributions, on the other hand, bear no such decrease describable in terms of the exclusivity distinction within [contrast]. Note that for both J- and E-groups, while the rate of case marking in [n-excl cont] is approximately 15% higher than that in [none], the rate of case marking in [f on pol] appears to be similar to that in [none].

These results indicate that all three groups’ variable use of the case morphemes is sensitive to the [contrast] distinction, supported by the far higher frequency of case marking in [f on DP] than in [info f], but the L1 and L2 performance diverges in terms of the sub-distinctions of [contrast]. The divergence appears to lie in the discrimination between [f on pol] and [n-excl
cont], which further suggests that the L1 and L2 groups’ case morpheme use may differ with respect to the specific application of the exclusivity distinction within the [contrast] category to the different subsets of the data.

Guided by the overall findings, frequency analysis of case marking was conducted against the subsets of the data. Due to the insufficient number of tokens for [info f] DPs, the subanalysis of focus effects was restricted to the subcategories of [contrast]. Accordingly, the distribution of case marking and zero-marking was observed across four discourse conditions: [f on DP], [f on pol], [n-excl cont], and [none].

**Subjects and objects**

The focus effects were first re-examined in terms of variable case marking of subjects/objects, as opposed to variable occurrence of the morphemes with non-subject/non-objects. Figure 6.12 summarizes the results obtained from the analysis of the subject and object data collapsed together.
Bias towards case marking by the approximate percentage values:

- K-group: \([f \text{ on DP}] > [f \text{ on pol}] > [n\text{-excl cont}] > [\text{none}]\)
- J-group: \([f \text{ on DP}] > [n\text{-excl cont}] > [f \text{ on pol}] > [\text{none}]\)
- E-group: \([f \text{ on DP}] > [n\text{-excl cont}] > [f \text{ on pol}] > [\text{none}]\)

(22) Figure 6.12. Effects of [focus] on frequency (%) of case (un)marking:

The results of the subject/object data were shown to be similar to those of the total data in that (i) for all three groups, the rate of case marking in \([f \text{ on DP}]\) is the highest, and (ii) K-group’s distribution involves a decrease in the sequence of \([f \text{ on pol}], [n\text{-excl cont}], \) and \([\text{none}]\), thereby revealing a pattern in which case marking is more frequent in both subcategories of [exclusive contrast] (i.e., \([f \text{ on DP}]\) and \([f \text{ on pol}]\)) than in \([n\text{-excl cont}]\), whereas no such description can be made for J- or E-group’s distributions.
Still, certain differences are evident between the subject/object and total data. For K-group, the frequency of case marking for [f on pol] DPs is higher in the subject/object data (84.3%) than in the total data (72.1%); the case marking rate for [f on pol] subjects/objects is as high as that for [f on DP] subjects/objects (85.4%). With respect to J- and E-groups, the results of the total data showed that the frequency of case marking for [n-excl cont] DPs was clearly higher than that for [f on pol] DPs, but the contrast is less prominent in the subject/object data. This ends up making the frequency of case marking appear to be similar across the [f on pol], [n-excl cont], and [none] conditions, especially in J-group.

Overall, the analysis of the subject and object arguments collapsed together indicates that K-group employs variable case marking strategies of [f on pol] over [n-excl cont] and of [n-excl cont] over [none], whereas these strategies are not clearly present in J- and E-groups’ performance (i.e., the effects of [f on pol] over [n-excl cont] were absent; the effects of [n-excl cont] over [none] appeared to be weak(er than in K-group)).

To further examine the L1-L2 divergence pattern, the frequency of case marking within each argument subset was analyzed. Figure 6.13 provides the results of the analysis of the subject data, most of which are found to be compatible with the preceding results.
\( \chi^2 = 27.28, p < .001 \) for K-group; \( \chi^2 = 12.72, p < .01 \) for J-group; \( \chi^2 = 1.53, p = .676 \) for E-group

Bias towards case marking by the approximate percentage values:
- K-group: [f on DP] > [f on pol] > [none] > [n-excl cont]
- J-group: [f on DP] > [n-excl cont] > [none] > [f on pol]
- E-group: [f on DP] > [n-excl cont] > [none] > [f on pol]

(23) Figure 6.13. Effects of [focus] on frequency (%) of case (un)marking:

The analysis replicates the highest frequency of case marking in [f on DP] across the three groups, but the differences in the relative frequency for E-group were not significant.

Noteworthy is that K- and J- groups case-mark the vast majority of [f on DP] subjects (96.6% and 93.3%, respectively) while they were reported to case-mark approximately 85% of [f on DP] data in the previous examinations of the collapsed data. E-group’s results contain no noticeable increase in the case marking rates for [f on DP] tokens across the examinations: 68.8% in the total data set, 69.5% in the subject/object set, and 70% in the subject set.
Only K-group demonstrates a higher frequency of case marking in [f on pol] than in [n-excl cont], and the reverse order is observed in J-group’s distribution. Unlike what appeared in the subject/object data, K-group’s subject marking pattern does not reflect the ranking of [n-excl cont] over [none].

The results from the object data revealed certain dissimilarities between subject and object case marking patterns produced by each group. See Figure 6.14 for the results.

(Bias towards case marking by the approximate percentage values:
- K-group: [f on pol] > [n-excl cont] > [f on DP] > [none]
- J-group: [f on pol] > [f on DP] > [none] > [n-excl cont]
- E-group: [f on DP] > [none] ≥ [f on pol] > [n-excl cont])

(24) Figure 6.14. Effects of [focus] on frequency (%) of case (un)marking: Objects.subcategories

203
For K-group, it was the [f on pol] condition that was associated with the highest frequency of case marking (80%), not the [f on DP] condition. Following [f on pol], the rates of case marking in [f on DP] (68.4%) and in [n-excl cont] (71.4%) are similarly higher than in [none] (57%).

These findings contrast with the results obtained from the subject data. First, although both for subjects and objects, the frequency of case marking is significantly higher in [f on DP] than in [none], while 96.6% of [f on DP] subjects are case-marked, only 68.4% of [f on DP] objects are case-marked. Second, the [n-excl cont] condition is shown to have no effect on subject case marking, whereas object case marking is significantly more frequent in [n-excl cont] (71.4%) than in [none] (57.1%).

J-group’s performance with the [f on DP] and [f on pol] objects is similar to K-group’s. That is, the frequency of case marking in the two conditions is higher than in [none], and [f on pol] is linked to the highest case marking rate, as well. Unlike K-group, J-group does not case-mark [n-excl cont] objects more frequently than [none] objects. The differences in J-group did not reach statistical significance, however.

E-group case-marks [f on DP] objects as frequently as K-group does. Unlike what was observed in K- and J-groups’ results, there is no noticeable asymmetry between subjects and objects in the focus condition: the frequency of case marking for the [f on DP] objects (68.6%) is similar to that for [f on DP] subjects (70%). More importantly, neither [f on pol] objects nor [n-excl cont] objects are associated with a higher rate of case marking than [none] objects.

As mentioned earlier, the analysis of the subject/object data suggests that K-group’s use of the nominative and accusative morphemes can be described by a ranking of [f on pol] over [n-excl cont], on the one hand, and by a ranking of [n-excl cont] over [none], on the other hand. In
contrast, it was observed that J- and E-groups’ use of the morphemes did not bear a pattern describable by these rankings.

This line of thought is supported by the results obtained from the data for each type of argument. The results from the subject data revealed that K-group’s pattern matches with the ranking of [f on pol] over [n-excl cont], whereas J- or E-group’s pattern does not match with the ranking. The ranking of [n-excl cont] over [none], on the other hand, is reflected in the object marking pattern by K-group, but neither J- nor E-group’s object marking pattern appears to bear out this ranking.

Accordingly, the discourse conditions associated with higher frequency of case marking relative to the [none] condition diverge between the L1 speakers of K-group and the L2 speakers of J- and E-groups (and across the subsets of the data). With respect to subject case marking, these conditions for K-group were identified with [f on DP] and [f on pol] (i.e., K-group’s bias towards explicit case marking for subjects is restricted to the [exclusive contrast] category and is concerned with both of the subcategories). However, J- and E-groups’ bias was mainly concerned with one of the subcategories, [f on DP]. In terms of object case marking, K-group’s bias towards explicit case marking was detected in [n-excl cont] as well as in [exclusive contrast] since [f on DP], [f on pol], and [n-excl cont] objects all involved a higher rate of case marking than [none] objects. On the contrary, J- and E-groups’ results exhibited the lowest frequency of case marking for [n-excl cont] objects while the [f on DP] condition continued to be the primary environment that favors explicit case marking.

These results are consistent with what was suggested by the prior analysis of the total data. That is, all three groups’ variable use of the case morphemes is sensitive to the [contrast]
distinction, but the L1 and L2 groups’ case morpheme use may diverge in terms of applying the exclusivity distinction within the [contrast] category to the different subsets of the data: K-group employs the exclusivity distinction in an asymmetrical manner between subject and object case markings, whereas J- and E-groups’ L2 knowledge may lack such fine-grained mappings between case marking and argument-specific application of the exclusivity distinction within [contrast].

**Non-subjects and non-objects**

The focus effects were reanalyzed in terms of variable occurrence of the morphemes with non-subject/non-objects, as well. Non-subject/non-object tokens coded as [f on DP], [f on pol], [n-excl cont], and [none] for each group (637 tokens for K-group; 454 tokens for J-group; 502 tokens for E-group) were sent to a further examination. Figure 6.15 summarizes the results of the analysis of the non-subject and non-object data collapsed together.
Bias towards case marking by the approximate percentage values:
- K-group: [f on DP] > [f on pol] > [n-excl cont] > [none]
- J-group: [n-excl cont] > [f on DP] > [f on pol] > [none]
- E-group: [f on DP] > [n-excl cont] > [f on pol] > [none]

(χ² = 29.10, p < .001 for K-group; χ² = 18.43 p < .05 for J-group; χ² = 10.62, p < .05 for E-group)

(25) Figure 6.15. Effects of [focus] on frequency (%) of case (un)marking:

Non-subjects/non-objects.subcategories

The three groups’ frequency distributions of non-subject/non-object data maintain the primary characteristics of the patterns of the subject/object data (Figure 6.12). First, the conspicuously high frequency of case marking in [f on DP] was replicated by all three groups’ results. Second, while the frequency of case marking is higher in [f on pol] than in [n-excl cont] for K-group, the opposite pattern was found for J- and E-groups. The three groups’ non-subject/non-object patterns bear great resemblance to the total data patterns (Figure 6.11) in that only K-group’s graph involves a gradual decrease in the sequence of [f on DP], [f on pol], [n-excl cont], and
and shows that case marking is more frequent in both subcategories of [exclusive contrast] (i.e., [f on DP] and [f on pol]) than in [n-excl cont].

A further frequency observation was made within each data subset. Figure 6.16 presents the results of the non-subject data.

Bias towards case marking by the approximate percentage values:
- K-group: [f on DP] > [f on pol] > [none] > [n-excl cont]
- J-group: [f on DP] = [n-excl cont] > [f on pol] > [none]
- E-group: [f on DP] > [n-excl cont] > [none] > [f on pol]

(26) Figure 6.16. Effects of [focus] on frequency (%) of case (un)marking:

Non-subjects subcategories

As shown by the graphs and the bias rankings, J- and E-groups’ patterns remain the same as those of the collapsed non-subject and non-object data except for the increased rates of case marking across all four conditions. The differences in the relative frequency were not statistically significant, however.
K-group’s non-subject marking pattern displays a noticeable dissimilarity from the collapsed data pattern. That is, the frequency of case marking in [n-excl cont] no longer appears to be higher than in [none]. The demotion of the effects of [n-excl cont] on bias towards explicit case marking was detected by the analysis of the group’s subject marking pattern in which the overall bias towards case marking was restrictively concerned with [exclusive contrast]. As seen in Figure 6.16, the general L1 tendency was emerging in the non-subject data, as well, although the differences in the L1 group did not reach statistical significance, either.

The non-object data of each group also underwent a further examination. See Figure 6.17 for the results.
Bias towards case marking by the approximate percentage values:

- K-group: \([n\text{-excl cont}] \geq [f \text{ on DP}] > [f \text{ on pol}] > [\text{none}]\)
- J-group: \([n\text{-excl cont}] = [f \text{ on DP}] > [f \text{ on pol}] > [\text{none}]\)
- E-group: \([f \text{ on DP}] > [n\text{-excl cont}] > [f \text{ on pol}] > [\text{none}]\)

(27) Figure 6.17. Effects of [focus] on frequency (%) of case (un)marking:

Non-objects.subcategories

According to the analysis, the effects of the subtypes of [focus] are nearly uniform across the three groups in that the frequency of case marking is high in both [f on DP] and [n-excl cont] but is low in [f on pol].

K-group’s non-object marking pattern is distinguished from their object marking pattern in that the frequency of case marking for [f on pol] non-objects (37.9%) is much lower than that for [f on pol] objects (80%). Nevertheless, the non-object marking pattern is aligned with the general tendency observed in their object marking pattern since all three subcategories of [contrast] (i.e., [n-excl contrast] as well as both subcategories of [exclusive contrast]) are linked to higher
frequency of case marking than the [none] category. This is further supported by the peculiarly high rate of case marking in [n-excl cont], a characteristic absent from K-group’s nominative marking of subjects and non-subjects both.

Although the results were not statistically significant, J-group also demonstrated distinctly high frequency of case marking for [f on pol] objects (81.5%). However, as shown in Figure 6.17, the [f on pol] category no longer displays such high frequency of case marking for non-objects. Another observation is that while J-group’s object marking pattern tied the lowest frequency of case marking to the [n-excl cont] category, their non-object marking pattern shows relatively high frequency of case marking in [n-excl cont]. Overall, all these characteristics render J-group’s non-object marking pattern notably similar to K-group’s.

With regard to E-group, it was pointed out that unlike K- and J-groups the L2 group did not case-mark [f on pol] objects more frequently than [none] objects. Their non-object marking pattern maintains the overall low frequency of case marking in [f on pol] although the case marking rate for [f on pol] non-objects is slightly higher than [none] non-objects. Similar to J-group, the [n-excl cont] category, which also showed the lowest case marking rate in E-group’s object marking pattern, is associated with relatively high frequency of case marking for non-objects.

In a nutshell, the further analysis conducted to test the effects of the various subtypes of focus showed that the three groups’ performance bears varying degrees of bias towards case marking across the subtypes. However, each group revealed distinct patterns of variation of the degrees.
More specifically, K-group’s use of the nominative morpheme exhibits bias towards explicit marking when the DP that the morpheme marks is construed as [exclusive contrast] focus, and both subcategories, [f on DP] and [f on pol] are found to form the bias. The tendency was observed in both subject and non-subject data although the latter data, which involves a relatively small number of tokens, did not reach statistical significance. On the contrary, the L1 group’s accusative marking patterns do not incorporate the exclusivity distinction in contrastive salience, displaying relatively high frequency of case marking across all three subcategories of [contrast] (i.e., [n-excl cont] as well as [exclusive contrast]). The tendency was borne out in both object and non-object data.

Despite their overall preference of case marking for [contrast] focus, J- and E-groups’ performance was shown to be different from the L1 performance in terms of the sensitivity to the subcategories of [contrast]. That is, the L2 performance does not bear a pattern that can be accounted for by the dichotomy between [exclusive contrast] and [n-excl cont]. Accordingly, the subcategories linked to relatively high frequency of case marking diverge between the L1 and L2 speakers. The L2 speakers appeared to identify [f on DP] as the primary environment that favors explicit case marking throughout the data in different subsets, especially in E-group. To illustrate the contrasts, Table 6.11 recapitulates the discourse conditions associated with a higher case marking rate than [none] in each data subset of the three groups.
Table 6.11. Subcategories of [contrast] associated with high frequency of case marking:

Across data subsets for each L1 group

<table>
<thead>
<tr>
<th>L1 group</th>
<th>Subcategory</th>
<th>NOM</th>
<th>ACC</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>K-group</strong></td>
<td>[f on DP]</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td></td>
<td>[f on pol]</td>
<td>X</td>
<td>(X, 8.7%)</td>
</tr>
<tr>
<td></td>
<td>[n-excl cont]</td>
<td></td>
<td>X</td>
</tr>
<tr>
<td><strong>J-group</strong></td>
<td>[f on DP]</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td></td>
<td>[f on pol]</td>
<td></td>
<td>(X, 7.9%)</td>
</tr>
<tr>
<td></td>
<td>[n-excl cont]</td>
<td>(X, 3.8%)</td>
<td>X</td>
</tr>
<tr>
<td><strong>E-group</strong></td>
<td>[f on DP]</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td></td>
<td>[f on pol]</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>[n-excl cont]</td>
<td>(X, 1.8%)</td>
<td>(X, 2.5%)</td>
</tr>
</tbody>
</table>

Note 1. The use of parentheses indicates that the rate difference is not greater than 10%.

The actual differences are given in the parentheses.

Note 2. The gray background indicates that the frequency differences in the data subset were statistically significant.

With respect to the L2 use of the nominative morpheme, the table shows that the case marking-friendly conditions are not restricted to [exclusive contrast], and often keep out the [f on pol].
category. Concerning the L2 accusative marking patterns, neither J- nor E-group connects [n-excl cont] objects to relatively high frequency of case marking.

To confirm these findings, a final observation was made in terms of each morpheme use (i.e., NOM data vs. ACC data). As presented in Figure 6.18, the results were overall consistent with the general descriptions of the L1-L2 performance discrepancies.

![Figure 6.18](image)

(29) Figure 6.18. Effects of [focus] on frequency (%) of case (un)marking: NOM and ACC data, subcategories

Again, K-group’s results from the NOM data confirm that the frequency of case marking is high in [f on pol] but low in [n-excl cont]. It is found that J- and E-groups’ results exhibit the opposite pattern, although E-group’s differences were not significant. Turning to the ACC data, the observed L1 pattern of object and non-object marking, where the frequency of case marking was high across all three subcategories of [contrast], reappears with [n-excl cont] bearing the highest
case marking rate. E-group produced a remarkably dissimilar ACC marking pattern from K-groups’, demonstrating bias towards explicit case marking only in the [f on DP] condition. The differences between J- and K-groups appear to be less prominent in that the rate of accusative marking in J-group is also higher for all three subcategories than for the [none] category. This is because the high case marking rates for [f on pol] objects (Figure 6.14) and for [n-excl cont] non-objects (Figure 6.17) were counted altogether in the analysis of the ACC data. Still, when comparing J-group’s ACC marking pattern with their NOM counterpart, the effect of [n-excl cont] is shown to be slightly demoted, contra the essential L1 characteristic that distinguishes NOM marking (i.e., feeble effects of [n-excl cont]) from ACC marking (i.e., manifest effects of [n-excl cont]).

6.4. Individual results

The mass media discourse data for this study had to be procured from only a limited number of L2 speakers when administering a proficiency test or a background questionnaire survey to the L2 speakers was not feasible (section 5.1.2). In light of the variability present in each L1 group, this section presents further analysis of the individual data based on the number of data tokens collected from each speaker (Table 5.4 in section 5.2.1.4). Due to the limited number of data tokens, the primary discussion of the individual results is restricted to the total, NOM, and ACC data.
6.4.1. Frequency of case (un)marking

6.4.1.1. Total data

Recall that the frequency of case marking of the total data for each L1 group was 61.9% for K-group, 63.5% for J-group, and 45% for E-group, suggesting overall similarities between K- and J-groups (Figure 6.1).

The individual results were mainly found to conform to the group patterns in which K- and J- groups displayed an overall preference of case marking over zero-marking, whereas E-group exhibited a reversed preference. However, it was also noted that certain individuals’ case marking patterns involve varying degrees of deviation from their group patterns in terms of not only case marking rates but also preferred directions of case (un)marking. See Table 6.12 for overall frequency of case marking for each speaker.
All Korean speakers case-mark more than a half of the DPs. SS and SY display lower case marking rates than the group average (by 7.9% and by 8.9%, respectively). SC’s preference of overt case marking appears to be particularly prominent, having 80% of the total DPs case-
marked. Despite the shared preference of case marking, the varying rates of case marking among the native Korean speakers suggest that the phenomenon involves a certain degree of stylistic variation, as well.

All Japanese speakers, except for HT, case-mark more than a half of the DPs, revealing a similar bias towards case marking to the Korean speakers’. TT demonstrates the highest case marking rate that is 11.5% higher than the group average. HT’s result strongly deviates from the group pattern in that he case-marks only 38% of the DPs, exhibiting an overall preference of zero-marking over case marking.

All English speakers, except for TR, case-mark less than a half the DPs. BW’s case marking rate is close to 50%, but in GP’s and MT’s results, the majority of the DPs are zero-marked (70% and 77%, respectively). It is noteworthy that TR’s result does not fit in his group pattern: unlike all the other English speakers, he case-marks the majority of the DPs (76%), manifesting a remarkable preference of case marking over zero-marking.

Overall, HT’s and TR’s results were formed in the opposite direction to the overall preference of case (un)marking of his group, suggesting the differences among the individuals within each L2 speaker group are likely to be qualitative as well as quantitative. The native Korean speakers’ results also indicate some quantitative differences among the group members.

It is unclear why particular speakers’ performance exhibited such evident deviations from their group results. With respect to TR, it may be speculated that such a high case marking rate found in the L2 performance can be ascribed to the fact that his command of Korean is regarded especially strong with a level of proficiency required to obtain the highest level of TOPIK and to complete a higher education at one of the most prestigious universities in the L2-speaking
country. It may also be related to his special interest in foreign language learning and his knowledge of languages other than English and Korean (cf., Table 5.4 in section 5.1.2). A similar account may apply to the prominently high frequency of case marking in SC’s L1 performance. SC is a renowned author of numerous world history books, who completed his undergraduate education in the U.S. and pursued his graduate studies in France. He speaks fluent English, French, and Italian and has at least some basic knowledge of Chinese, German, and Latin. In one episode of the show, he described one of his hobbies as learning foreign languages through his world history studies. It is difficult to make any confident statement concerning HT’s result due to the notably small number of data tokens that came from only one episode of his guest appearance (74 tokens in total). Perhaps, the low frequency of case marking might be considered in reference to the higher rate of his case marking errors compared with the other Japanese speakers’. That is, the number of erroneously case-marked tokens was 24 for J-group (error rate: 1.4%); among the 24 tokens, five tokens were attributed to HT’s production, which makes his error rate higher than the group average (i.e., 6.3% = 5/(74+5)*100, calculated as the ratio of erroneously case-marked DPs to the total DPs produced).

6.4.1.2. Nominative and accusative data

In the group results, the frequency of NOM marking for each group was 73.5% for K-group, 74.9% for J-group, and 57.0% for E-group, and that of ACC marking was 48.6% for K-group, 51.2% for J-group, and 32.8% for E-group groups (Figure 6.2). It was discussed that for all three groups,

59 See footnote 45.
case marking is more prevalent in the NOM data than in the ACC data. It was also pointed out that the overall case marking rates in J- and K-groups were comparably higher than in E-group for both NOM and ACC data.

The individual results were mostly found to be in line with the group results. Table 6.13 and Table 6.14, respectively, provide frequency of NOM marking and of ACC marking for each individual.

![Table 6.13. Frequency (%) of case marking: Number of data tokens. NOM data. individual results](image)

<table>
<thead>
<tr>
<th></th>
<th>Marked</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>HJ</td>
<td>242 (69%)</td>
<td>352 (100%)</td>
</tr>
<tr>
<td>SC</td>
<td>294 (87%)</td>
<td>338 (100%)</td>
</tr>
<tr>
<td>SS</td>
<td>218 (68%)</td>
<td>322 (100%)</td>
</tr>
<tr>
<td>SY</td>
<td>198 (70%)</td>
<td>283 (100%)</td>
</tr>
<tr>
<td><strong>K-group</strong></td>
<td>952 (73.5%)</td>
<td>1295 (100%)</td>
</tr>
<tr>
<td>HO</td>
<td>203 (70%)</td>
<td>292 (100%)</td>
</tr>
<tr>
<td>HT</td>
<td>24 (53%)</td>
<td>38 (100%)</td>
</tr>
<tr>
<td>TT</td>
<td>229 (85%)</td>
<td>269 (100%)</td>
</tr>
<tr>
<td>YN</td>
<td>185 (72%)</td>
<td>257 (100%)</td>
</tr>
<tr>
<td><strong>J-group</strong></td>
<td>641 (74.5%)</td>
<td>856 (100%)</td>
</tr>
<tr>
<td>BW</td>
<td>95 (65%)</td>
<td>146 (100%)</td>
</tr>
<tr>
<td>GP</td>
<td>102 (33%)</td>
<td>272 (100%)</td>
</tr>
<tr>
<td>MT</td>
<td>107 (42%)</td>
<td>255 (100%)</td>
</tr>
<tr>
<td>TR</td>
<td>229 (87%)</td>
<td>262 (100%)</td>
</tr>
<tr>
<td><strong>E-group</strong></td>
<td>533 (57%)</td>
<td>935 (100%)</td>
</tr>
</tbody>
</table>

![Table 6.14. Frequency (%) of case marking: Number of data tokens. ACC data. individual results](image)

<table>
<thead>
<tr>
<th></th>
<th>Marked</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>HJ</td>
<td>137 (47%)</td>
<td>293 (100%)</td>
</tr>
<tr>
<td>SC</td>
<td>215 (72%)</td>
<td>299 (100%)</td>
</tr>
<tr>
<td>SS</td>
<td>98 (37%)</td>
<td>266 (100%)</td>
</tr>
<tr>
<td>SY</td>
<td>95 (36%)</td>
<td>267 (100%)</td>
</tr>
<tr>
<td><strong>K-group</strong></td>
<td>545 (48.4%)</td>
<td>1125 (100%)</td>
</tr>
<tr>
<td>HO</td>
<td>148 (55%)</td>
<td>271 (100%)</td>
</tr>
<tr>
<td>HT</td>
<td>4 (11%)</td>
<td>36 (100%)</td>
</tr>
<tr>
<td>TT</td>
<td>155 (63%)</td>
<td>246 (100%)</td>
</tr>
<tr>
<td>YN</td>
<td>103 (42%)</td>
<td>247 (100%)</td>
</tr>
<tr>
<td><strong>J-group</strong></td>
<td>410 (51.2%)</td>
<td>800 (100%)</td>
</tr>
<tr>
<td>BW</td>
<td>18 (20%)</td>
<td>89 (100%)</td>
</tr>
<tr>
<td>GP</td>
<td>55 (22%)</td>
<td>248 (100%)</td>
</tr>
<tr>
<td>MT</td>
<td>8 (3%)</td>
<td>251 (100%)</td>
</tr>
<tr>
<td>TR</td>
<td>218 (67%)</td>
<td>324 (100%)</td>
</tr>
<tr>
<td><strong>E-group</strong></td>
<td>299 (32.8%)</td>
<td>911 (100%)</td>
</tr>
</tbody>
</table>
First, for all individuals across the three groups, the frequency of NOM marking is higher than that of ACC marking. Second, when disregarding HT’s and TR’s results, in both NOM and ACC data, case marking is more frequent among all individuals in K- and J-groups than among those in E-group. In these respects, it appears that the individual results bolster the reported group patterns to a major extent.

However, as the individual differences noted in the total data reentered the NOM and ACC marking patterns, individuals within each group were often found to diverge in terms of their preferred directions of case (un)marking, especially for ACC marking.

As for K-group, all four speakers case-mark the majority of the NOM data. For ACC marking, three speakers (HJ, SS, and SY) case-mark less than a half of the ACC data. SS and SY, who revealed relatively lower case marking rates in the total data, case-mark slightly more than a third of their ACC data (37% and 36%, respectively). Finally, SC’s strong preference of case marking reappears for both NOM (87%) and ACC (72%) markings.

Looking at J-group, TT is identified with the highest frequency of case marking in both NOM (85%) and ACC (63%) data, replicating his result for the total data. With the lowest frequency of case marking in the group, HT’s result seems to deviate from the group pattern, especially with his ACC data. His NOM marking rate (63%) is slightly lower than an English speaker, BW’s (65%), and his ACC marking rate is notably low (11%) and is even lower than an English speaker, GP’s (22%).

Unlike K- and J-groups’ individuals, the English speakers diverge in terms of preferences of case (un)marking of the NOM data. Two speakers (BW and TR) are more likely to case-mark them than to zero-mark them, whereas the other two speakers (GP and MT) exhibit a reversed
bias. However, with the ACC data, the majority of the speakers (BW, GP, and MT) demonstrate a strong preference of zero marking. BW, who was shown to case-mark almost a half of his total data, maintains the relatively high frequency of case marking in his NOM data (65%). Although his NOM marking rate is high enough to be compared to some Korean and Japanese speakers’ results (e.g., SS, HT), his ACC marking rate is as low as 20%, which seems to fall into neither Korean nor Japanese speakers’ pattern. GP’s and MT’s NOM marking rates are similarly low (38% and 42%, respectively), but MT’s ACC marking rate involves a more dramatic drop from his NOM marking rate (from 42% to 3%) than GP’s ACC marking rate does (39% to 22%), which leaves the vast majority of his ACC tokens zero-marked. TR continues to exhibit the highest frequency of case marking in the group for NOM and ACC data. When compared with all other speakers’ performance in the three L1 groups, his NOM marking rate (87%) is the highest, which turns out to be as high as SC’s, and his ACC marking rate (67%) is the second highest, following SC’s (72%).

6.4.1.3. Where K- and J-groups diverged: Non-subjects and objects

According to the group results, although K- and J-groups mostly showed comparable rates of case marking across the data subsets, J-group was noted to demonstrate higher frequency of case marking than K-group in non-subjects and objects.

In regard to non-subjects, the frequency of case marking for J-group was 69.5%, which was higher than that for K-group (56.7%) (Figure 6.3). One possible explanation for this divergence may be sought in the individual data. Among the 123 nominative-marked non-subject tokens of J-group, 55 tokens were produced by TT, who manifested the highest frequency of case marking.
in the group. Within this particular data subset, his case marking rate was found to be 85%, which is as high as in his NOM data. Turning to the Korean speakers’ results, among the 165 nominative-marked non-subject tokens, 57 tokens were attributed to SC, who also exhibited the highest frequency of case marking in the group, but it was observed that his case marking rate in the non-subject data (71%) was lower than that in the NOM data (87%). It appeared that although both speakers revealed a pronounced preference of the overt use of the nominative morpheme, the L1 speaker uses it less frequently for non-subjects than subjects while the L2 speaker uses it equally frequently for both non-subjects and subjects. The relatively lower frequency of nominative marking for non-subjects was reproduced by the other three native speakers’ data, whereas the tendency was less prominent in the Japanese speakers’ results (see Table E.1 in Appendix E: supplementary analysis of individual results).

As for the object data, while K-group case-marked 59.3% of the objects, J-group case-marked 68.4% of the objects (Figure 6.4). Looking at the individual patterns, there was an overall increase in all Korean and Japanese speakers’ case marking rates from the ACC data to the object data. When not considering HT’s results that involved only 15 objects, for all speakers in both groups, except for YN, the increase was confined to 10% to 13%. However, YN’s data revealed a greater increase (28%), which lifted his object case marking rate to 70%. Overall, the individual results appear to support that J-group is more likely to case-mark objects than K-group does, which should be pointed out as a discrepancy between the two groups’ performance (see Table E.2 in Appendix E: supplementary analysis of individual results).
6.4.2. Effects of the discourse factor

6.4.2.1. Effects of the broad discourse categories: [new], [focus], [none]

According to the group analysis of the total data, the frequency of case marking was significantly higher for [new] and [focus] DPs than for [none] DPs. The pattern was consistent across the three L1 groups’ results while reflecting the relatively lower frequency of case marking found in E-group’s overall performance (Figure 6.10).

To re-examine the group result, frequency analysis was conducted against the total data of each speaker. Table 6.15 presents the frequency of case marking for the three discourse categories observed in each individual’s performance.
Table 6.15. Effects of the discourse factor on frequency (%) of case marking:

<table>
<thead>
<tr>
<th></th>
<th>[new]</th>
<th>[focus]</th>
<th>[none]</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Total data</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HJ***</td>
<td>68% (54)</td>
<td>71% (111)</td>
<td>52% (214)</td>
</tr>
<tr>
<td>SC**</td>
<td>83% (100)</td>
<td>90% (85)</td>
<td>77% (324)</td>
</tr>
<tr>
<td>SS***</td>
<td>68% (58)</td>
<td>70% (90)</td>
<td>45% (168)</td>
</tr>
<tr>
<td>SY***</td>
<td>80% (55)</td>
<td>69% (82)</td>
<td>43% (156)</td>
</tr>
<tr>
<td><strong>K-group</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HO*</td>
<td>68% (78)</td>
<td>72% (70)</td>
<td>58% (203)</td>
</tr>
<tr>
<td>HT**</td>
<td>58% (7)</td>
<td>73% (8)</td>
<td>25% (13)</td>
</tr>
<tr>
<td>TT***</td>
<td>86% (80)</td>
<td>81% (72)</td>
<td>70% (232)</td>
</tr>
<tr>
<td>YN*</td>
<td>65% (65)</td>
<td>65% (56)</td>
<td>53% (167)</td>
</tr>
<tr>
<td><strong>J-group</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BW*</td>
<td>64% (27)</td>
<td>56% (22)</td>
<td>42% (64)</td>
</tr>
<tr>
<td>GP**</td>
<td>28% (27)</td>
<td>41% (53)</td>
<td>26% (77)</td>
</tr>
<tr>
<td>MT***</td>
<td>38% (31)</td>
<td>27% (25)</td>
<td>18% (59)</td>
</tr>
<tr>
<td>TR**</td>
<td>86% (65)</td>
<td>84% (104)</td>
<td>72% (278)</td>
</tr>
<tr>
<td><strong>E-group</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>50.7% (150)</td>
<td>53.1% (204)</td>
<td>41% (478)</td>
</tr>
</tbody>
</table>

Note 1. The percentage values represent the ratio of case-marked DPs to the total DPs in each discourse condition, and the parenthesized numbers indicate the number of case-marked DPs found in each discourse condition.

Note 2. * p < .05, ** p < .01, *** p < .001
For all speakers, except for SC and GP, case marking occurs significantly more frequently (by at least 9%) for [new] and [focus] DPs than for [none] DPs. The effects of the [new] category on two speakers’ results were less apparent or negligible. For SC, the frequency of case marking is higher in [new] than in [none] only by 6%. GP’s case marking rate in [new] involves only a 2% increase from that in [none], which serves no or little evidence for the effects of [new] on this individual’s use of the nominative morpheme.

In addition, it is noted that the frequency of case marking in each discourse condition varies across the individuals of each L1 group in accordance with their overall frequency and preference of case (un)marking (Table 6.12). For instance, across the three discourse conditions, speakers who showed a pronounced preference of case marking (SC, TT, and TR) demonstrate the highest frequency of case marking in the L1 group, whereas the most conservative case-marking individual (MT) continues to display relatively lower case marking rates.

Overall, the individual analysis provides additional support for the group result in that for all 12 speakers, case marking occurred significantly more frequently in [focus] than in [none] (by at least 9%); for 10 out of 12 speakers, case marking was significantly more frequent in [new] than in [none] (by at least 10%).

6.4.2.2. Effects of the [new] category: Subjects and non-subjects

According to the group analysis concerning the effects of the [new] category on the subject vs. non-subject distinction, only K-group was found to be more sensitive to the appropriateness of case marking of [new] subjects than to that of [new] non-subjects. That is, for K-group, nominative marking of discourse [new] entities was significantly more frequent in the subject
data than in the non-subject data, but no such tendency was observed in J- and E-groups’ results (section 6.3.2.1). This result was supported by the majority of the individual speakers’ performance. See Table 6.16.

(34) Table 6.16. Effects of [new] on frequency (%) of nominative marking: [new] subjects vs. [new] non-subjects. Individual results

<table>
<thead>
<tr>
<th></th>
<th>[new] subjects</th>
<th>[new] non-subjects</th>
</tr>
</thead>
<tbody>
<tr>
<td>HJ*</td>
<td>74% (46)</td>
<td>44% (8)</td>
</tr>
<tr>
<td>SC**</td>
<td>89% (79)</td>
<td>68% (21)</td>
</tr>
<tr>
<td>SS**</td>
<td>74% (52)</td>
<td>40% (6)</td>
</tr>
<tr>
<td>SY*</td>
<td>84% (52)</td>
<td>43% (3)</td>
</tr>
<tr>
<td><strong>K-group</strong></td>
<td>80.9% (229)</td>
<td>53.5% (38)</td>
</tr>
<tr>
<td>HO</td>
<td>69% (63)</td>
<td>65% (15)</td>
</tr>
<tr>
<td>HT</td>
<td>55% (6)</td>
<td>100% (1)</td>
</tr>
<tr>
<td>TT</td>
<td>85% (63)</td>
<td>89% (17)</td>
</tr>
<tr>
<td>YN*</td>
<td>70% (57)</td>
<td>44% (8)</td>
</tr>
<tr>
<td><strong>J-group</strong></td>
<td>73.3% (189)</td>
<td>67.2% (41)</td>
</tr>
<tr>
<td>BW</td>
<td>65% (24)</td>
<td>60% (3)</td>
</tr>
<tr>
<td>GP</td>
<td>27% (21)</td>
<td>32% (6)</td>
</tr>
<tr>
<td>MT</td>
<td>34% (20)</td>
<td>48% (11)</td>
</tr>
<tr>
<td><strong>TR</strong></td>
<td>89% (57)</td>
<td>67% (8)</td>
</tr>
<tr>
<td><strong>E-group</strong></td>
<td>51.4% (122)</td>
<td>47.5% (28)</td>
</tr>
</tbody>
</table>
For all Korean speakers, case marking occurred significantly more frequently for [new] subjects than for [new] non-subjects. In addition, SC’s result revealed that the less apparent effect of [new] on his NOM data (Table 6.13) was due to his relative reluctance to case-mark [new] non-subjects compared with his particularly strong preference of case marking in his overall performance.

In respect of J-group, two speakers (HT and TT) produced a reversed distribution in which case marking is more frequent for [new] non-subjects than for [new] subject. The other two speakers’ results (HO and YN) were shown to match with the nativelike production pattern, and YN’s result reached statistical significant ($\chi^2 = 4.08$, $p < .05$).

As for E-group, two individuals (GP and MT) displayed a non-native like pattern, but the other two individuals’ results (BW and TR) were found to fit in the nativelike pattern, where TR’s result was significant ($\chi^2 = 4.10$, $p < .05$).

Hence, although the majority of the L2 speakers did not incorporate the subject vs non-subject distinction in nominative marking of discourse [new] entities, unlike what was informed by the grouped data, the individual analysis indicates a nativelike employment of the distinction by one native Japanese speaker (YN) and one native English speaker (TR).

6.4.2.3. Effects of the [focus] subcategories: Total data

It was discussed that although the three groups exhibited different degrees of sensitivity to the various subcategories of [focus], they all demonstrated higher frequency of case marking for [contrast] DPs (i.e., [focus on DP], [focus on polarity], and [non-exclusive contrast]) than for [none] DPs. Despite the limited number of [information focus] tokens, it was pointed out that no
group case-marked [information focus] DPs than [none] DPs, suggesting that a [contrast] distinction within the [focus] category was integrated into all three groups’ use of the case morphemes (Figure 6.11).

Individual analysis was conducted against each speaker’s total data tokens that fall into the [focus] and [none] categories. Frequency of case marking in each subcategory of [focus] (i.e., [information focus], [focus on DP], [focus on polarity], and [non-exclusive contrast]) was observed and compared with that in the [none] category. The results are summarized in Table 6.17.
Table 6.17. Effects of [focus] on frequency (%) of case marking: [focus] subcategories.

<table>
<thead>
<tr>
<th></th>
<th>[f on DP]</th>
<th>[f on pol]</th>
<th>[n-excl cont]</th>
<th>[info f]</th>
<th>[none]</th>
</tr>
</thead>
<tbody>
<tr>
<td>HJ***</td>
<td>86% (30)</td>
<td>70% (62)</td>
<td>67% (18)</td>
<td>17% (1)</td>
<td>52% (214)</td>
</tr>
<tr>
<td>SC*</td>
<td>100% (34)</td>
<td>84% (31)</td>
<td>87% (20)</td>
<td>N/A</td>
<td>77% (324)</td>
</tr>
<tr>
<td>SS***</td>
<td>73% (37)</td>
<td>72% (36)</td>
<td>63% (17)</td>
<td>N/A</td>
<td>45% (168)</td>
</tr>
<tr>
<td>SY***</td>
<td>83% (38)</td>
<td>67% (34)</td>
<td>56% (10)</td>
<td>0% (0)</td>
<td>43% (156)</td>
</tr>
<tr>
<td><strong>K-group</strong></td>
<td>84.2% (139)</td>
<td>72.1% (163)</td>
<td>61.9% (65)</td>
<td>10% (10)</td>
<td>54.9% (862)</td>
</tr>
<tr>
<td>HO*</td>
<td>86% (24)</td>
<td>68% (30)</td>
<td>57% (12)</td>
<td>100% (4)</td>
<td>58% (203)</td>
</tr>
<tr>
<td>HT*</td>
<td>100% (2)</td>
<td>67% (2)</td>
<td>80% (4)</td>
<td>0% (0)</td>
<td>25% (13)</td>
</tr>
<tr>
<td>TT***</td>
<td>85% (34)</td>
<td>77% (24)</td>
<td>82% (14)</td>
<td>0% (0)</td>
<td>70% (232)</td>
</tr>
<tr>
<td>YN***</td>
<td>87% (27)</td>
<td>40% (14)</td>
<td>82% (14)</td>
<td>33% (1)</td>
<td>53% (167)</td>
</tr>
<tr>
<td><strong>J-group</strong></td>
<td>86.1% (87)</td>
<td>61.9% (70)</td>
<td>73.3% (44)</td>
<td>55.6% (9)</td>
<td>58.3% (615)</td>
</tr>
<tr>
<td>BW</td>
<td>75% (3)</td>
<td>43% (10)</td>
<td>75% (9)</td>
<td>N/A</td>
<td>42% (64)</td>
</tr>
<tr>
<td>GP**</td>
<td>53% (25)</td>
<td>29% (13)</td>
<td>45% (14)</td>
<td>17% (1)</td>
<td>26% (77)</td>
</tr>
<tr>
<td>MT***</td>
<td>44% (7)</td>
<td>12% (6)</td>
<td>50% (11)</td>
<td>N/A</td>
<td>18% (59)</td>
</tr>
<tr>
<td>TR**</td>
<td>93% (42)</td>
<td>80% (49)</td>
<td>76% (13)</td>
<td>0% (0)</td>
<td>72% (278)</td>
</tr>
<tr>
<td><strong>E-group</strong></td>
<td>68.8% (77)</td>
<td>43.8% (78)</td>
<td>57.3% (47)</td>
<td>16.7% (12)</td>
<td>41% (478)</td>
</tr>
</tbody>
</table>

230
For all Korean speakers, case marking was more frequent in all three subcategories of [contrast] than in [none] (by at least 7%), and there was no individual whose case marking rate in [info f] is higher than in [none]. It was noted that the effects of [f on pol] on SC’s result were smaller than on the other three native controls’ results in that his case marking rate in [f on pol] was only 7% higher than in [none].

In regard to the Japanese speakers, two speakers’ results (HT and TT) were comparable the Korean speakers’ performance characterized by relatively higher frequency of case marking in all three [contrast] subcategories but not in the [info f] category. For the other two speakers (HO and YN), case marking occurred more frequently only in two subcategories of [contrast]: YN did not case-mark more DPs in [f on pol] than in [none], and HO did not case-mark more DPs in [n-excl cont] than in [none]. Additionally, HO turned out to case-mark all his [info f] DP tokens. Four of the nine [info f] tokens for J-group were attributed to his production, which ended up lifting the group’s case marking rate in [info f] to 54.9%, which was found to be notably higher than K- and E-groups’ (10% and 16.7%, respectively) (Figure 6.11). It remains unclear why he case-marked all his [info f] tokens. Although one could interpret his result as evidence for lack of L2 sensitivity to the [contrast] distinction within the [focus] category, it should be also considered that the number of tokens of [info f] is too small to make any conclusive statement.

Looking at the English speakers, no individual’s result demonstrated a higher case marking rate in [info f] than in [none]. As to the [contrast] subcategories, it appears that all four individuals primarily connect relatively higher frequency of case marking to two of the subcategories, [f on DP] and [n-excl cont]. Except for TR’s result, the effects of [f on pol] were either absent or feeble in their results (although there was a 3% increase in GP’s case marking
rates between [none] and [f on pol], which may be comparable to SC’s result in K-group). TR, one the other hand, seems to prefer [f on pol] to [n-excl cont] as one of his case-friendly conditions. In his data, relatively higher case marking rates are majorly linked to [f on DP] and to [f on pol], while [n-excl cont] still exhibits a slightly higher case marking rate than in [none] (by 4%). To a certain extent, this pattern is comparable to three native speakers’ results (HJ, SS, and SY) in which case marking is more frequent in [f on pol] (and in f on DP) than in [n-excl cont].

By and large, the individual results agree with the statement made out of the group results that the [contrast] distinction within the [focus] category played a significant role in all three groups’ use of the case morphemes. 11 speakers’ results displayed significantly higher frequency of case marking for at least two subcategories of [contrast], and 10 speakers’ results did not bear a significantly higher case marking rate in [info f] than in [none]. Furthermore, the varied preferences between [f on pol] and [n-excl cont] among the individual L2 speakers seem to be the source of the different rankings of the [contrast] subcategories with respect to forming a strong or weak bias towards case marking for each L2 speaker group (i.e., [n-excl cont] > [f on pol] for J-group, and [f on pol] > [n-excl cont] for E-group, cf. Figure 6.11)

6.4.2.4. Effects of the [contrast] subcategories: Nominative and accusative data

According to the group analysis of the effects of the distinct subcategories of [contrast], it was suggested that while the native controls’ use of the nominative morpheme employs an exclusivity distinction within [contrast], their use of the accusative morpheme is not constrained by that sub-distinction. More specifically, for K-group, relatively higher frequency of nominative marking was associated with [focus on DP] and [focus on polarity] DPs (i.e., with the [exclusive
contrast] subtypes), but not with [non-exclusive contrast] DPs, but relatively higher frequency of accusative marking was found across all three subcategories of [contrast]. On the other hand, neither J- nor E-group incorporated the exclusivity sub-condition to distinguish their use of the nominative morpheme from that of the accusative morpheme (Figure 6.18).

These finding was re-evaluated against the individual data by looking at each speaker’s case marking rates across the four discourse categories (i.e., [focus on DP], [focus on polarity], [non-exclusive contrast], and [none]) in his NOM and ACC data. Table 6.18 reports the individual results of the NOM data.
(36) Table 6.18. Effects of [focus] on frequency (%) of case marking: [contrast] subcategories. NOM data. Individual results

<table>
<thead>
<tr>
<th></th>
<th>[f on DP]</th>
<th>[f on pol]</th>
<th>[n-excl cont]</th>
<th>[none]</th>
</tr>
</thead>
<tbody>
<tr>
<td>HJ*</td>
<td>100% (13)</td>
<td>78% (49)</td>
<td>64% (14)</td>
<td>66% (111)</td>
</tr>
<tr>
<td>SC</td>
<td>100% (21)</td>
<td>82% (18)</td>
<td>88% (14)</td>
<td>89% (141)</td>
</tr>
<tr>
<td>SS*</td>
<td>92% (23)</td>
<td>74% (29)</td>
<td>63% (15)</td>
<td>62% (93)</td>
</tr>
<tr>
<td>SY***</td>
<td>97% (30)</td>
<td>73% (24)</td>
<td>60% (9)</td>
<td>60% (80)</td>
</tr>
</tbody>
</table>

**K-group** 96.7% (87) 76.4% (120) 67.5% (52) 69.6% (425)

<table>
<thead>
<tr>
<th></th>
<th>[f on DP]</th>
<th>[f on pol]</th>
<th>[n-excl cont]</th>
<th>[none]</th>
</tr>
</thead>
<tbody>
<tr>
<td>HO</td>
<td>89% (16)</td>
<td>61% (14)</td>
<td>64% (7)</td>
<td>69% (84)</td>
</tr>
<tr>
<td>HT</td>
<td>100% (2)</td>
<td>100% (2)</td>
<td>100% (3)</td>
<td>56% (10)</td>
</tr>
<tr>
<td>TT</td>
<td>93% (25)</td>
<td>83% (15)</td>
<td>91% (10)</td>
<td>83% (99)</td>
</tr>
<tr>
<td>YN(*)</td>
<td>100% (18)</td>
<td>57% (8)</td>
<td>85% (11)</td>
<td>75% (82)</td>
</tr>
</tbody>
</table>

**J-group** 93.8% (61) 68.4% (39) 81.6% (31) 74.5% (275)

<table>
<thead>
<tr>
<th></th>
<th>[f on DP]</th>
<th>[f on pol]</th>
<th>[n-excl cont]</th>
<th>[none]</th>
</tr>
</thead>
<tbody>
<tr>
<td>BW</td>
<td>67% (2)</td>
<td>57% (8)</td>
<td>75% (9)</td>
<td>65% (49)</td>
</tr>
<tr>
<td>GP</td>
<td>56% (15)</td>
<td>32% (11)</td>
<td>50% (13)</td>
<td>41% (35)</td>
</tr>
<tr>
<td>MT(*)</td>
<td>50% (7)</td>
<td>15% (4)</td>
<td>73% (11)</td>
<td>47% (53)</td>
</tr>
<tr>
<td>TR*</td>
<td>100% (23)</td>
<td>94% (34)</td>
<td>71% (10)</td>
<td>86% (97)</td>
</tr>
</tbody>
</table>

**E-group** 70.1% (47) 51.8% (57) 64.2% (43) 60.5% (234)

Note. The use of parenthesized asterisks indicates that the L2 speaker’s non-target-like use of each morpheme still involved significantly higher frequency of case marking in at least one of the [contrast] subcategories.
As for K-group, three speakers’ results (HJ, SS, and SY) were statistically significant. It is shown that for these three individuals, frequency of case marking is higher for both subcategories of [exclusive contrast] DPs than for [none] DPs: case marking occurs more frequently in [f on DP] (by at least 30%) and in [f on pol] (by at least 12%) than in [none].

Looking at their case marking rates for the [n-excl cont] category, they mostly remain the same or close to those for the [none] category. SC’s result was not significant ($\chi^2 = 3.81, p = .283$) and deviates from the other native controls’ pattern in that he case-marks all [f on DP] DPs, but his case marking rate in [f on pol] is in fact lower than in [none]. It appears that the frequency of his NOM marking remains consistently higher than 80% across the other three categories than the [f on DP] category, reflecting the conspicuous bias towards overt case marking in his NOM data.

Although this is a result expected by the smaller effects of [f on pol] in his total data (Table 6.17), it is still an unanswered question why his production does not conform to the other Korean speakers’ performance on [f on pol]. Overall, the majority of the individual results can be considered to support the group result in that three out of four speakers produced a significant distribution in which relatively higher frequency of case marking is linked to both [f on DP] and [f on pol] DPs but not to [n-excl cont] DPs.

Concerning J-group, for all four speakers, case marking is more frequent in [f on DP] than in [none], and their preference of case marking for the [f on DP] category appears to be strong enough to be compared to the Korean speakers’. When disregarding HT’s result, no individual connects a relatively stronger bias towards case marking with the [f on pol] category; two individuals (TT and YN) associate the bias with the [n-excl cont] DP, instead. HT, whose result came from only five [contrast] tokens, was found to case-mark all DPs that fall under the
category. In any case, it can be concluded that no individual in J-group produced a nativelike pattern with respect to the exclusivity distinction in their use of the nominative morpheme.

Turning to E-group, when disregarding TR’s result, for two speakers (BW and MT), case marking occurs more frequently primarily in the [n-excl cont] condition, but in neither of the two [exclusive contrast] subcategories. For GP, case marking is more frequent in [f on DP] as well as in [n-excl cont]. Unlike the other English speakers, relatively higher case marking rates are found both in [f on DP] and [f on pol], but not in [n-excl cont], indicating that his production was guided by the exclusivity distinction within [contrast].

Frequency of ACC marking across the same discourse categories was also observed in each individual’s data. The results are illustrated in Table 6.19.
Table 6.1. Effects of [focus] on frequency (%) of case marking: [contrast]

<table>
<thead>
<tr>
<th></th>
<th>[f on DP]</th>
<th>[f on pol]</th>
<th>[n-excl cont]</th>
<th>[none]</th>
</tr>
</thead>
<tbody>
<tr>
<td>HJ**</td>
<td>77% (17)</td>
<td>52% (13)</td>
<td>80% (4)</td>
<td>43% (103)</td>
</tr>
<tr>
<td>SC*</td>
<td>100% (13)</td>
<td>87% (13)</td>
<td>86% (6)</td>
<td>69% (183)</td>
</tr>
<tr>
<td>SS*</td>
<td>54% (14)</td>
<td>64% (7)</td>
<td>67% (2)</td>
<td>33% (75)</td>
</tr>
<tr>
<td>SY</td>
<td>53% (8)</td>
<td>56% (10)</td>
<td>33% (1)</td>
<td>33% (76)</td>
</tr>
<tr>
<td>K-group</td>
<td>69.3% (52)</td>
<td>62.3% (43)</td>
<td>72.7% (13)</td>
<td>45.6% (437)</td>
</tr>
<tr>
<td>HO(*)</td>
<td>80% (8)</td>
<td>76% (16)</td>
<td>50% (5)</td>
<td>48% (111)</td>
</tr>
<tr>
<td>HT</td>
<td>N/A</td>
<td>0% (0)</td>
<td>50% (1)</td>
<td>9% (3)</td>
</tr>
<tr>
<td>TT</td>
<td>69% (9)</td>
<td>69% (9)</td>
<td>67% (4)</td>
<td>62% (133)</td>
</tr>
<tr>
<td>YN</td>
<td>69% (9)</td>
<td>29% (6)</td>
<td>75% (3)</td>
<td>41% (85)</td>
</tr>
<tr>
<td>J-group</td>
<td>72.2% (26)</td>
<td>55.4% (31)</td>
<td>59.1% (13)</td>
<td>49.6% (340)</td>
</tr>
<tr>
<td>BW</td>
<td>100% (1)</td>
<td>22% (2)</td>
<td>N/A</td>
<td>19% (15)</td>
</tr>
<tr>
<td>GP(*)</td>
<td>50% (10)</td>
<td>18% (2)</td>
<td>20% (1)</td>
<td>20% (42)</td>
</tr>
<tr>
<td>MT</td>
<td>0% (0)</td>
<td>9% (2)</td>
<td>0% (0)</td>
<td>3% (6)</td>
</tr>
<tr>
<td>TR</td>
<td>86% (19)</td>
<td>60% (15)</td>
<td>100% (3)</td>
<td>66% (181)</td>
</tr>
<tr>
<td>E-group</td>
<td>66.7% (30)</td>
<td>30.9% (21)</td>
<td>26.7% (4)</td>
<td>31.3% (244)</td>
</tr>
</tbody>
</table>

The Korean speakers’ ACC data revealed that all individuals maintained the observed bias towards case marking both in [f on DP] and [f on pol] while the bias was extended to the [n-excl cont] category, as well. It should be pointed out that no such extension is found in SY’s result: he
case-marked two of his three [n-excl cont] ACC tokens, thereby displaying equally low case marking rates in both [n-excl cont] and [none]. Although the issue remains admittedly obscure, the deviation might also be considered in relation to the small number of [n-excl cont] data tokens. The number of ACC tokens for each [contrast] subcategory, especially for [n-excl cont]60, is smaller than in the NOM data, and only HJ’s, SC’s, and SS’s results reached statistical significance.

Looking at the L2 speakers’ results, due to lack of data tokens in the compared discourse categories, descriptions of certain individuals’ results (HT, BW, and MT) are not likely to generate any meaningful statement. Among the rest of the L2 speakers, there is one individual (TT) whose case marking rates are higher in all three [contrast] subcategories than in [none]. Nevertheless, this result should not be interpreted as evidence for his nativelike knowledge of ACC marking. First, the differences across the four categories were not statistically significant. More importantly, considering his NOM marking pattern (Table 6.18), which was also not significant, there is no indication at all that his bias towards case marking in [n-excl cont] increased from NOM marking to ACC marking. Since his NOM data lacks any evidence to support that the bias was restricted to [exclusive contrast] in NOM marking, in the first place, it cannot be claimed that it was extended to the [n-excl cont] in ACC marking. Conservatively speaking, no individual in J- or E-group is found to exhibit sensitivity to the exclusivity distinction in their use of the accusative morpheme.

---

60 One of the downsides of having a limited number of [n-excl cont] tokens was also found in the group analysis of non-subjects and non-objects. For example, there appeared diverging performance of J-group from K- and E-groups in the non-subject data (Figure 6.16). Although the group result was not significant, the divergence was conspicuous since all [n-excl cont] non-subjects were reported to be case-marked. The individual analysis revealed that the total two [n-excl cont] non-subject tokens of J-group (cf. Table 6.9) were produced by one individual (YN).
All things considered, three out of four native speakers’ performance on each case morpheme was consistent with the group results, and seven out of eight L2 speakers did not integrate the exclusivity distinction with either of the case morphemes. In this regard, the individual results can be considered to support the group results, but to a limited extent. Although the individual data do not seem to disconfirm the group results, they revealed two new facets of the variability phenomenon and its L2 acquisition. First, the native controls’ judgments on NOM marking in [f on pol] and on ACC marking in [n-excl cont] were not found to be as robust as expected: one individual’s performance (SC) did not conform to the group’s NOM marking pattern, and one individual’s result (SY) did not match with the group’s ACC marking pattern. Second, unlike what was suggested by the group results, one native English-speaking adult learner of Korean (TR) demonstrated target-like convergence on integrating the exclusivity distinction with his NOM marking pattern, suggesting that the L2 acquisition involving such exacting tasks is indeed possible.

6.5. Summary of the results

The subject/object and non-subject/non-object tokens extracted from the transcribed texts were coded for markedness and discourse features, and the frequency distribution of case marking and zero-marking for each L1 group was observed according to the two independent variables. This section recapitulates the key results of the corpus analysis.

6.5.1. Frequency of case (un)marking

The total data of each L1 group was first examined in terms of the frequency of case
(un)marking. The result (Figure 6.1) is that while less than a half the data in E-group (45%) is case-marked, more than 60% of the data in K- and J-groups is found to be case-marked, indicating distributional likeness between K- and J groups’ patterns.

The distribution of case marking vs. zero-marking was further observed across the subsets of the data. In all three groups’ results, the frequency of case marking is higher in NOM data than in ACC data (in non-subjects than in non-objects as well as in subjects than in objects); it is also higher in subjects/object arguments than in non-subjects/non-object items. The analysis of each data subset supported that K- and J-groups’ patterns bear great resemblance to each other. Across all four data subsets (i.e., subjects, objects, non-subjects, and non-objects), the frequency of case marking in the two groups is shown to be significantly higher than in E-group; the two groups were primarily found to display similar rates of case marking across the subsets.

6.5.2. Effects of the markedness factor

The subject/object and non-subject/non-object tokens were coded as one of the following categories: [1st person], [2nd person], [pronoun 3rd], [proper noun 3rd], [proper noun 3rd]-like, [animate 3rd], [animate 3rd]-metonymic human, [inanimate 3rd].

As the analysis conducted with the full (sub)categories showed no significant effect of the markedness factor, certain (sub)categories involving an insufficient number of tokens (e.g., [2nd person], [animate 3rd], [animate 3rd]-metonymic human, [proper noun 3rd]-like) were collapsed into a broader category. The frequency analysis re-conducted against the four data subsets confirmed that no group produced a pattern that manifests a significant effect of the markedness factor on variable case marking in Korean.

240
The frequency of case (un)marking was also reconsidered according to the three-way distinctions adopted in Lee’s (2006b) analysis of the definiteness and animacy effects (i.e., [pronoun] vs. [name] vs. [other] for definiteness; [human] vs. [animate] vs. [inanimate] for animacy). No group’s data reproduced Lee’s (2006b) results of the definiteness and animacy effects.

In this dissertation, unlike what was predicted by the previous research on case drop in Japanese and Korean, the subject and object case (un)marking patterns produced by the L1 speakers does not reveal the markedness effects. Overall, the markedness factor was found to have no significant effect on variable use of the nominative and accusative case particles across the three groups.

6.5.3. Effects of the discourse factor

The total data of each L1 group was classified according to the three broad discourse categories of [new], [focus], and [none], and the frequency distribution of case marking and zero-marking was observed. The result (Figure 6.10) is that in all three groups, the frequency of case marking is higher for [new] and [focus] DPs than for [none] DPs ($\chi^2 = 91.78, p < .001$ for K-group; $\chi^2 = 32.77, p < .001$ for J-group; $\chi^2 = 21.79, p < .001$ for E-group). It was also noted that K- and J-groups’ results are similar in that they display a higher rate of case marking than E-group across all three discourse conditions.

The frequency of case (un)marking for the [new] category, which has a bearing on the NOM data only, was further analyzed within each data subset (i.e., subjects vs. non-subjects). K-group’s performance demonstrates a higher rate of case marking for [new] DPs when they serve as the subject of a sentence than when they do not, whereas the sensitivity was not detected in J-
and E-groups’ results. As to the additional annotation distinguishing between [new]-iss (i.e., discourse-new information that occurs with the verb iss (‘exist, be’) and [new]-rest, no group’s subject marking pattern demonstrated a higher frequency of case marking in [new]-iss than in [new]-rest.

The total data coded as [focus] and as [none] was reanalyzed to investigate whether the different subcategories of focus (i.e., [information focus], [focus on DP], [focus on polarity], and [non-exclusive contrast]) exerted distinct effects on the three groups’ variable use of the case morphemes. The findings (Figure 6.11) revealed that in all three groups, the frequency of case marking in [focus on DP] was the highest, whereas [information focus] was linked to the lowest frequency of case marking ($\chi^2 = 37.53, p < .001$ for K-group; $\chi^2 = 19.20, p < .001$ for J-group; $\chi^2 = 31.41, p < .001$ for E-group). This result can be interpreted as evidence that both L1 and L2 knowledge of the case morphemes bears sensitivity to the [contrast] distinction within [focus].

The focus effects were further examined within each subset of the data. Due to the limited number of data tokens for [information focus], the inspection was restricted to the subtypes of [contrast]. Not all comparisons made in the data subsets were significant (Note 2 on Table 6.11), but the overall result is that the three groups’ patterns suggest varying degrees of bias towards case marking according to the subtypes of [contrast] while the specific modes of variation of the degrees diverge between the L1 speakers of K-group and the L2 speakers of J- and E-groups.

In essence, K-group appears to utilize the notion of exclusivity in marking [contrast] DPs with the case morphemes, but the application is asymmetrical between nominative and accusative markings. K-group’s bias towards explicit nominative marking is shown to be confined to [exclusive contrast], whereas their bias towards explicit accusative marking does not
discriminate between [exclusive contrast] and [non-exclusive contrast]. The tendency was borne out in K-groups’ results that revealed a relatively high frequency of nominative marking in both [focus on DP] and [focus on polarity], but not in [non-exclusive contrast], and that of accusative marking in all three subcategories of [contrast] (Table 6.11). On the contrary, J- or E-group’s performance did not reflect the specific mappings between the case morphemes and the exclusive distinction within [contrast] (Table 6.11).

**6.5.4. J-group vs. E-group**

The two groups’ results were comparable in that they both demonstrated the overall sensitivity to the broad three-way discourse distinction of [new], [focus], and [none], and to the [contrast] distinction within [focus]. In terms of the further analysis performed on the data subsets, neither group exhibited the nativelike sensitivity to the discrimination between [new] subjects vs. [new] non-subjects, and to the notion of exclusivity involving [contrast]. All these properties make the two L2 groups’ performance appear to be similar to each other.

However, of the various sets of frequency analysis provided several pieces of evidence to support that J-group’s production patterns are more nativelike than E-group’s. First, the overall case marking rates in J- and K-groups were shown to be similarly higher than in E-group across the subsets of the data. Although J-group occasionally exhibited more frequent use of the case morphemes than K-group did (e.g., for non-subjects, for objects), the two groups mostly revealed close case marking rates across the subsets. In regard to case marking of the [focus on DP] tokens, both K- and J-groups demonstrated asymmetrical patterns between subjects and objects in which [focus on DP] subjects are case-marked far more frequently than [focus on DP] objects.
whereas such subject-object asymmetry was not found in E-group’s production patterns (Figure 6.13 and 6.14). Furthermore, although neither group’s results for the effects of the subtypes of [contrast] can be claimed to be nativelike, J-group’s results occasionally exhibited nativelike performance as an emerging pattern. The examples include the relatively higher frequency of case marking for [focus on polarity] non-subjects and objects (Table 6.11), and the promotion of the effects of [non-exclusive contrast] on non-object marking (Figure 6.17), part of which made J-group’s ACC pattern appear less dissimilar from K-group’s (Figure 6.18).

6.5.5. Individual results

Going one step further, the major findings of the group analysis were re-evaluated against the individual data.

The individual results primarily complied with the group results of the overall frequency of case (un)marking. Certain individuals’ patterns revealed varying degrees of deviation from their group patterns, however. One Japanese speaker (HT) displayed lower frequency of case marking than the other individuals of J-group, whereas one English speaker (TR) demonstrated higher frequency of case marking than the other individuals of E-group, which often ended up making their results diverge from their group’s preferred direction of case (un)marking. One native control’s result (SC) involved remarkably high case marking rates, but his pronounced preference of case marking was confined to quantitative differences from the other individuals of K-group.

In respect of the effects of the discourse features, the analysis conducted against the total data confirmed that for at least 10 out of 12 speakers, case marking was significantly more
frequent among [new] and [focus] DPs than among [none] DPs. Further analysis regarding the effects of the [new] category on the subject vs. non-subject distinction showed that the majority of the individual results conformed to the group results. However, one native Japanese speaker (YN) and one native English speaker (TR) were found to involve a nativelike employment of the distinction in their nominative marking of [new] discourse entities. As for the effects of the various subcategories of [focus], 11 speakers’ results demonstrated significantly higher frequency of case marking in at least two subcategories of [contrast], and 10 speakers’ results did not involve a significantly higher case marking rate in [info f] than in [none], providing overall support for the group results that all three groups’ use of the case morphemes was guided by the [contrast] distinction within the [focus] category. Lastly, the analysis of each individual’s NOM and ACC data revealed that three out of four native speakers’ performance on each morpheme matched with the claimed nativelike pattern, and seven out of eight L2 speakers did not integrate the exclusivity distinction with either of the case morphemes, providing significant support for the group results. However, the native controls’ judgments on NOM marking in [f on pol] and on ACC marking in [n-excl cont] were not observed to be as consistent as expected: one individual’s performance (SC) did not comply with the group’s NOM marking pattern, and one individual’s result (SY) did not conform to the group’s ACC marking pattern. Moreover, unlike what was informed by the group results, one English speaker (TR) exhibited target-like knowledge of the integration of the exclusivity distinction into his use of the nominative morpheme, which indicates that the L2 learning requiring such an intricate understanding of the discourse feature structure is indeed possible in adult L2 acquisition.
CHAPTER 7
CONCLUSION

This chapter concludes the dissertation. Section 2.1 presents a summary of the corpus investigation along with answers to the research questions found in the analysis. Section 2 provides additional discussions on the findings to address the shortcomings of the present study and potential directions for future research.

7.1. Summary: Answers to the research questions

This dissertation examined the L2 acquisition of variable use of the nominative and accusative case morphemes by English and Japanese L2 speakers of Korean with an aim of evaluating different approaches to adult L2 acquisition of morphosyntactic features. Using naturalistic corpus data, this study investigated whether adult L2 learners can attain nativelike knowledge of the semantic and discourse-pragmatic features that regulate the naturalness of (non-)use of the case morphemes.

The L2 theories considered were the Representational Deficit Hypothesis (RDH), the Feature Reassembly Hypothesis (FRH), and the Interface Hypothesis (IH). The specific learning tasks of the L2 acquisition and the L2 data analyzed in this study were considered suitable to assess the different theoretical approaches. The internal and external interfaced-properties involved in the L2 learning and the L2 data obtained from different L1-L2 pairings provided empirical evidence to compare the conflicting predictions of the FRH and IH, in particular.
The study was based upon naturally occurring L1 and L2 spoken data of conversational Korean produced in a South Korean TV show, *picengsanghoityam* (‘Non-summit’). A total of 5,923 DP tokens extracted from the transcribed discourse data were coded for markedness and discourse features, and the frequency distribution of case marking and zero-marking was analyzed according to the two independent variables.

The investigation was guided by three research questions. The remainder of this section discusses some answers to the research questions sought in the corpus analysis.

**Research Question 1: Testing the markedness factor**

Do English and Japanese L2 speakers of Korean acquire nativelike sensitivity to the semantic features of the markedness hierarchy that condition variable use of the nominative and accusative case morphemes in Korean (i.e., person, animacy, prominence in definiteness)?

Across the three groups, the markedness factor was found to have no significant effect on variable use of the case morphemes. The L1 group’s performance does not exhibit the previously-claimed nativelike sensitivity to the semantic features of person, animacy, definiteness and thereby nullifies the validity of affirming or disaffirming the presence of L2 sensitivity to them.

**Research Question 2: Testing the discourse factor**

Do English and Japanese L2 speakers of Korean acquire nativelike sensitivity to the different discourse-pragmatic functions that condition variable use of the nominative and accusative case
According to the analysis of the total data, both English and Japanese L2 speakers of Korean exhibit nativelike sensitivity not only to the broad discourse functions of [new] and [focus] but also to the further distinction of [contrast] within the [focus] feature. However, the analysis of the data subsets shows that neither group displays nativelike sensitivity to more subtle discourse conditions. That is, unlike the L1 speakers of Korean, the L2 speakers’ sensitivity to [new] did not discriminate between subjects and non-subjects; the L2 speakers did not integrate the distinction between [exclusive contrast] and [non-exclusive contrast] with their use of the two case morphemes. However, a close observation of the individual data indicated nativelike performance on the distinction of [new] subjects and [new] non-subjects by one Japanese speaker and one English speaker, and on the exclusivity distinction within [contrast] by the same English speaker. All things considered, it can be concluded that (i) the L2 speakers do acquire the broad discourse functions of the case morphemes; (ii) the L2 knowledge of the more fine-grained mappings between the relevant discourse properties and the morphemes is significantly more difficult to attain; (iii) such fine-grained mappings are still acquirable in adult L2 acquisition.

Research Question 3: Evaluating L2 theories

To what extent do the predictions of the RDH, FRH, and IH account for the results?
Given that neither the L2 groups’ nor the L1 control group’s results revealed markedness effects, the assessment is based solely on the results of the discourse factor.

The RDH, which predicted no learning difficulty for either L2 group, has little bearing on the findings that displayed varying degrees of (non-)target-like properties across the different sets of the analysis. The findings can be interpreted as empirical evidence not merely disconfirming the prediction but also calling into question the adequacy and coherency of the leading assumptions to the prediction.

The results of this study are consistent with the predictions of the FRH. First, the non-nativelike outcomes regarding the fine-grained mappings between the discourse subcategories and each morpheme (or argument) are coherent with assuming significant learning difficulties in remapping of the relevant features into the L2-specific way from the L1. Furthermore, given that these mappings are guided by an intricate understanding of the hierarchy of focus subtypes incorporated in the coding, these findings also indicate that the most deeply-embedded condition is likely to be the most taxing to acquire in adult L2 acquisition. Second, the effects of the broad discourse categories were found to be present in both L2 groups’ performance, and one individual’s performance exhibited target-like convergence on integrating the more fine-grained conditions of both [new] and [focus] into his case morpheme use. These findings are in line with ultimate acquirability of the discourse features. Lastly, the corpus analysis provided several pieces of evidence to consider J-group’s performance more nativelike than E-groups’, which suggests the inevitable influence of the L1. The overall analysis of the frequency of case (un)marking displayed great resemblance between J- and K-groups. Moreover, when further analysis of the sub-categories of [contrast] was conducted on the different data subsets, J-group’s
results revealed a nativelike pattern that was not present in E-group’s (e.g., subject-object asymmetry in preferences of case marking for the [focus on DP] category).

Contrary to original expectation, this study did not engender empirical evidence to evaluate the IH prediction on the disparity between the internal vs. external interface acquisitions (i.e., the semantic features associated with the internal interface acquisition turned out to have no significant effect on the native controls’ performance, either). Nevertheless, considering that the overall results of the total data for both L2 groups exhibited target-like convergence on the discourse features, the data does not support the prediction of ultimate unacquirability at the syntax-discourse interface. Furthermore, the imbalanced performance between J- and E-groups seems to undermine the validity of explaining language acquisition in terms of a sharp contrast of interface types without considering a significant role of the learners’ L1. In addition, the findings of this study do not seem to support Sorace’s (2011, 2012) claim that focus (as an internal-interfaced property) is easier to acquire than discourse newness (as an external-interfaced property) since both L2 groups produced nativelike performance in which case marking is significantly more frequent for [new] DPs.

7.2. Discussions: For future research

This dissertation explored comprehensive data of both variable case marking of subjects/objects and variable morpheme attachment to non-subjects/objects to gain a broader perspective on the variability phenomenon. It should be highlighted that the latter type of data have not been addressed in the past work on case drop in Korean, and incorporating such intriguing data helped the present study to implement a finer analysis of the data into a broader picture of the actual
morpheme use. For instance, generating the coding category of [focus on polarity] was aided by everyday observation of frequent co-occurrences of the nominative morpheme with various syntactic categories in sentential negation in the language (section 2.3.2.1. and section 2.3.2.2). Having this subcategory of [exclusive contrast] in addition to [focus on DP] was conducive to securing the dichotomic division of [contrast] DPs according to the exclusivity sub-distinction, which led the study to the novel findings on a significant role of that distinction in forming the L1 and L2 usage patterns of the two case morphemes. In this regard, despite the assorted subtypes of non-subject/non-object data, future research is encouraged to embrace these data in the examination of case drop in Korean and confirm whether (and to what extent) the effects of the discourse conditions discussed in the previous literature and this dissertation hold for non-subjects/non-objects.

Although this study provided detailed descriptions of the various subtypes of case-markable non-subjects/non-objects (section 2.3.2), the analysis conducted was admittedly premature due to the limited number of data tokens and to the predetermined scope of the study, and this should be noted as a shortcoming of the analysis. Nevertheless, it might be added that some sporadic observations of the data suggested that the non-subject/non-object data may be better described by the discourse factor than by the markedness factor. More specifically, nominative marking on certain subtypes of non-subjects is likely to be linked to the effects of discourse [new] (e.g., nominal complements of *to* *y* (‘become’) (section 2.3.2.2) tend to be new information due to the semantic aspect of the verb). With respect to verbal nouns in *ha*-light verb constructions (section 2.3.2.6), the notion of [contrast] (and possibly the further exclusivity distinction, as well) may shed light on the data of accusative marking of verbal nouns. For example, [focus on polarity]
could play a role in the phenomenon if it is assumed that the matrix predicate is saturated with the focus property of its complement. As an illustration, when the clausal complement of a verbal noun, sayngkak (‘thought’) stands with an opposite proposition, the [focus on polarity] in the embedded clause may be realized on the verbal noun that selects the clausal complement, thereby serving a discourse function equivalent to the emphatic do in English (i.e., I do think that S V). These are some areas that future research could pay attention to while escorted by solid theoretical understanding in semantics and pragmatics. As non-subjects/non-objects also co-occur with the topic/contrastive marker -(n)un, which tends to mark discourse-given or [non-exclusive contrast] information (section 2.3.1), it is not unlikely to bear fruitful results to exploit the newness and exclusivity distinctions in examining variable case morpheme attachment to some subtypes of non-subject/non-object data.

Another point to be made concerns the fact that the previously-claimed nativelike sensitivity to the semantic features of person, animacy, and definiteness were found to have no significant effect on the native controls’ performance. Considering that the data of this study was classified according to quite a fine-grained coding scheme and was submitted to multiple sets of analysis, this result could pose a non-trivial challenge to the existing literature on the phenomenon. On the one hand, this result was not entirely unexpected given that in the aforementioned L2 acquisition study by Chung (2013) (section 3.2.2), the native controls’ results, sometimes along with the heritage speakers’ results, did not replicate the effects of animacy and definiteness. As suggested by the previous research (e.g., Lee, 2006a; footnote 11), the effects of animacy and definiteness on case drop are relatively weak compared with those of focus. Taking this further, it is not entirely improbable to conceive that the effects of the semantic features may actually be too
feeble to observe in the language. That is, perhaps in Korean the discourse-marking function of the case morphemes is so prominent that it may overwhelm the function to differentiate grammatical roles, to an even greater extent than what was proposed in the previous literature. This line of thought appears to be consistent with Chae’s (2017) L1 acquisition study in that her Korean-learning children were not found to fully exhibit the supposedly adult-like sensitivity to animacy and definiteness even by the age of 5:11 (section 3.1.3). Another speculation is that the markedness hierarchy introduced into the coding system (section 2.2.2.5) in itself may not be an adequate tool to describe all various case-related cross-linguistic phenomena. As noted by the need to create a subcategory or an extra annotation for some gray-area data (e.g., proper noun-like, metonymic human in section 5.2.2.1), the actual language data did not always seem to fit in the discrete categories presented in the hierarchy. Simultaneously, it was shown that not all markedness categories carry purely semantic properties given the need for contextual knowledge to properly classify the gray-area data. Overall, the result that indicated no significant effect of the markedness factor in the L1 data gives rise to questions as to some salient theoretical aspects of the variability phenomenon.

In contrast, the discourse factor was found to have significant effects on the case-drop variability phenomenon. In particular, the finding regarding the exclusivity distinction within [contrast] is a novel one in the literature of case drop in Korean, which can be added to the previously-observed properties of the subject-object asymmetry in case drop. The finding is consistent with the uneven results on contrast subtypes between subjects and objects in Kim (2008) (i.e., nominative marking was significantly associated with the exhaustive listing-type contrast and -(n)un-marking with the parallel-type contrast, whereas no such division of labor
between case marking and -(n)un marking was observed for the accusative morpheme (Table 2.7 and Table 2.8 in section 2.3.1). It appears that the results are compatible with an assumption that the nominative morpheme is responsible for marking [exclusive contrast] DPs, whereas the accusative morpheme is not. Furthermore, the previous claim that ‘focused’ subjects resist case drop more strongly than ‘focused’ objects may also be considered in light of this assumption. Given that what has been mostly denoted by ‘focused’ in the previous literature is equivalent to the subcategory of [focus on DP] in this study, the claim might also be examined under this assumption which underlines a tighter connection of the ‘focused’ DPs to the nominative morpheme than to the accusative morpheme.

However, it should be pointed out as an important limitation that the native controls’ judgments on this exclusivity distinction were not found to be as robust as expected. Even with the observed individual deviations, an attenuated statement of a distinguishing group characteristic of the native controls could still be made. That is, all native speakers of Korean confine their bias towards explicit nominative marking to [exclusive contrast], whereas no such restriction holds for their bias towards explicit accusative marking. There are two implications of this conservative generalization. First, the nativelike nominative marking pattern requires a bias towards case marking not be linked to [non-exclusive contrast]. Second, the nativelike accusative marking pattern does not require a bias towards case marking be linked to [non-exclusive contrast]. It then appears that these implications put clearly more restrictions on appropriate (non-)use of the nominative morpheme than on that of the accusative morpheme. The uneven restrictions on the two case morphemes do seem to be compatible with the subject-object asymmetry and with the previously-observed acquisition order of the nominative morpheme.
before the accusative morpheme). Nevertheless, since such modifications on the native controls’ performance in an L2 acquisition study might bring forth a potential concern about fairness of interpretation between L1 and L2 data (e.g., Table 6.19 in section 6.4.2), this dissertation leaves examining the validity of the attenuated generalization about the nativelike performance for future research on case drop in Korean while ascribing the limitation to an inadequate amount of data to avoid making any ad hoc modification.

It was noted that not all analyses of the data subsets revealed statistically significant differences or led to a fully faithful representation of group performance (e.g., analysis of non-subjects/non-objects; of [non-exclusive contrast]; of [information focus]). The issue of having an inadequate amount of data seems to be closely related to the source of the data collected. For example, the debate format of the show could have contributed to the lack of [information focus] DPs. That is, although there were copious instances of wh-questions issued by the Korean hosts, since the conversations were held among multiple participants divided to represent two opposite stances on a given topic, DPs that could have been classified as [information focus] ended up being realized as [contrastive focus] as the different L2 speakers were exchanging their conflicting views in the proceeding discourse. In this respect, data from paired conversations would balance the number of data tokens between [contrast] DPs and [information focus] DPs. One of the most evident drawbacks of using the particular data source could be that the study was restricted only to the L2 speakers available in the show while being unable to carefully screen them in terms of proficiency matches and participants’ eligibility (section 5.1.2). Although the close observation of the individual results may be understood as a remedy for this disadvantage, future L2 acquisition research is encouraged to further (dis)confirm the primary
findings of this study. Lastly, the overall frequency of case marking observed among the native controls appeared to be higher than what had been reported in the previous studies that used non-media conversational data (e.g., Lee’s (2006b) corpus data of paired telephone conversations). Despite the casual format of the show and the participants’ informal speech style, it is not impossible that the media language data may be imperceptibly tainted by the effects of an extralinguistic factor, which may have caused an increased bias towards explicit case marking in the overall results. These are some shortcomings of the present study that can be better addressed by future studies. As an example, more reliable L2 corpus data could be gained from recruiting L2 speakers to whom administering a valid proficiency test or a language background survey would be possible and recording their natural interactions. Considering that case drop is identified with a characteristic property of colloquial conversation held between familiar interlocutors, it would be a good idea to recruit the L2 speakers in pairs when each pair has a reasonably intimate personal relationship.

By incorporating fine-grained categories of the conditioning semantic and pragmatic factors, this study examined the specific extent of the granularity of linguistic knowledge that can be achieved in adult L2 acquisition, suggesting that such granularity may be one of the most elusive properties to acquire. The overall findings of this study were found to support the predictions of the FRH. One of the aspects of the data that led to this conclusion was that the Japanese speakers’ performance was often observed to be more nativelike than the English speakers’, which required the basic assumption that Japanese has a similar case system to Korean. Going one step further, future L2 research may provide a more precise description of similarities and dissimilarities of the two languages’ case systems to pinpoint where exactly the L2 performance
diverges from the L1 performance. This will bolster the overall findings of this study and the predictions of the FRH.

This dissertation was guided by a view that acquisition studies should be based on a solid theoretical framework, and acquisition data can examine the validity of different theoretical approaches (and vice versa). Perhaps, the fact that the RDH and the IH were not able to successfully handle the L2 acquisition data reported in this study might be associated with their theoretical assumptions regarding the locus of cross-linguistic variation (for the RDH) and regarding the structure of linguistic knowledge (for the IH). The findings of this L2 acquisition study can be used to enhance our current understanding of case variability in Korean and various case-related cross-linguistic phenomena.
APPENDICES

Appendix A: Sources of L2 speakers’ language background

http://tvdaily.asiae.co.kr/read.php3?aid=1427154737873896002
http://sports.khan.co.kr/entertainment/sk_index.html?cat=view&art_id=201407081211253&sec_id=540101&pt=ny
http://tvdaily.asiae.co.kr/read.php3?aid=1407487349744087002
http://www.tvreport.co.kr/?c=news&m=newsview&idx=568225
http://m.kocis.go.kr/koreanet/view.do?seq=6072
https://github.com/bskim45/enha/blob/master/mirror/%ED%83%80%EC%9D%BC%EB%9F%AC%20%EB%9D%BC%EC%89%AC.md
http://www.sisain.co.kr/?mod=news&act=articleView&idxno=24423
http://osen.mt.co.kr/article/G1109942332
Appendix B: Video data used for transcription

Excerpts on YouTube:

https://www.youtube.com/watch?v=Wetd3ZPqBKl
https://www.youtube.com/watch?v=4W-o-qyh1KY
https://www.youtube.com/watch?v=bRsW5G8e2Jo
https://www.youtube.com/watch?v=fjpqomJSm2I
https://www.youtube.com/watch?v=KKJlcZpGlBw
https://www.youtube.com/watch?v=9VCAKYIEuGg
https://www.youtube.com/watch?v=NUtiheFkEAQ
https://www.youtube.com/watch?v=vQtlNNmxOjc
https://www.youtube.com/watch?v=m_gi84dh4rk
https://www.youtube.com/watch?v=4mEqSbGom3I
https://www.youtube.com/watch?v=5k1xXbaDQk4
https://www.youtube.com/watch?v=9pMaOJDnJ30
https://www.youtube.com/watch?v=D0BHVLPaQ_Q
https://www.youtube.com/watch?v=9FTL1hmIDt8
https://www.youtube.com/watch?v=sKuTjARgG7o
https://www.youtube.com/watch?v=DDqkWxa1qmA
https://www.youtube.com/watch?v=ScLiQzHtdy0
https://www.youtube.com/watch?v=sHHRwiy7ju4
https://www.youtube.com/watch?v=dx5qeulrzJE
https://www.youtube.com/watch?v=rAOW3MeQeII
https://www.youtube.com/watch?v=eBDbz8tELBM
https://www.youtube.com/watch?v=5uD4dM8bdM
https://www.youtube.com/watch?v=0cn-Nnxnigs
https://www.youtube.com/watch?v=d1AdrGVHzh4
https://www.youtube.com/watch?v=ReH4pnuUsoE
https://www.youtube.com/watch?v=Gi1zKgNo7Is
https://www.youtube.com/watch?v=Ye24nBeXluU
https://www.youtube.com/watch?v=9hsQKPOpQsQ
https://www.youtube.com/watch?v=E_qUxNSXQY4
https://www.youtube.com/watch?v=hV373b0oHyk
https://www.youtube.com/watch?v=yoFO79Dljck
https://www.youtube.com/watch?v=49lCQ0Vr4d0
https://www.youtube.com/watch?v=3g80K8DNN2g
https://www.youtube.com/watch?v=tbKzsWNqQnw
https://www.youtube.com/watch?v=9x2Jc9MGLNY
https://www.youtube.com/watch?v=uiTYCAOAVH4
https://www.youtube.com/watch?v=jOklKwzjXUU
https://www.youtube.com/watch?v=XY-b163FQgU
https://www.youtube.com/watch?v=kqGeoU4pGgc
https://www.youtube.com/watch?v=ao42o7PiBDo
https://www.youtube.com/watch?v=pfmP4n7YMkI
https://www.youtube.com/watch?v=oPDM8K-BZ7s
https://www.youtube.com/watch?v=8Z3OWaKM-wY
https://www.youtube.com/watch?v=roqPw-DoHIM
https://www.youtube.com/watch?v=zTjwILpKxe0
https://www.youtube.com/watch?v=56lzJ2KC-uU

Full Episodes:

Episodes 17; 30; 46; 47; 50; 53; 55; 60; 61; 65; 66; 68; 69; 70; 71; 72; 73; 74; 75; 76; 77; 78; 111; 126; 127; 128; 129; 131 (The complete collection of the episodes is available on the TV channel homepage (http://tv.jtbc.joins.com/nonsummit))
Appendix C: Supplementary analysis of the markedness factor for subjects and objects

(1) Figure C.1. Effects of the markedness factor on frequency (%) of case (un)marking:

Subjects.K-group.subanalysis of definiteness.strongly definite vs. low definite

(χ² = 2.43, p = .119)

(2) Figure C.2. Effects of the markedness factor on frequency (%) of case (un)marking:

Subjects.K-group.full (sub)categorues

(χ² = 20.09, p < .05)
(3) Figure C.3. Effects of the markedness factor on frequency (%) of case (un)marking:

Subjects.J-group

(\chi^2 = 6.96, p = .223)

(4) Figure C.4. Effects of the markedness factor on frequency (%) of case (un)marking:

Subjects.J-group.full (sub)categories

(\chi^2 = 13.73, p = .089)
(χ² = 58.14, p < .001)

(5) Figure C.5. Effects of the markedness factor on frequency (%) of case (un)marking:
Subjects.E-group.full (sub)categories

(χ² = 0.511, p = .774)

(6) Figure C.6. Effects of the markedness factor on frequency (%) of case (un)marking:
Objects.K-group.subanalysis of definiteness
(7) Figure C.7. Effects of the markedness factor on frequency (%) of case (un)marking:

Objects.K-group.full (sub)categories

(χ² = 6.52, p = .589)

(8) Figure C.8. Effects of the markedness factor on frequency (%) of case (un)marking:

Objects.J-group.full (sub)categories

(χ² = 16.55, p < .05)
(χ² = 21.15, p < .01)

(9) Figure C.9. Effects of the markedness factor on frequency (%) of case (un)marking:

Objects.E-group.full (sub)categories
Appendix D: Supplementary analysis of the markedness factor for non-subjects

(1) Figure D.1. Effects of the markedness factor on frequency (%) of case (un)marking:

Non-subjects.K-group

(χ² = 9.28, p = .098)

(2) Figure D.2. Effects of the markedness factor on frequency (%) of case (un)marking:

Non-subjects.J-group

(χ² = 5.75, p < .331)
(3) Figure D.3. Effects of the markedness factor on frequency (%) of case (un)marking:

Non-subjects.E-group

\( \chi^2 = 3.65, p < .001 \)
**Appendix E: Supplementary analysis of individual results**

(1) Table E.1. Frequency (%) of case marking: Number of data tokens

<table>
<thead>
<tr>
<th>Group</th>
<th>Marked</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>HJ</td>
<td>38 (49%)</td>
<td>78 (100%)</td>
</tr>
<tr>
<td>SC</td>
<td>57 (71%)</td>
<td>80 (100%)</td>
</tr>
<tr>
<td>SS</td>
<td>40 (54%)</td>
<td>74 (100%)</td>
</tr>
<tr>
<td>SY</td>
<td>30 (51%)</td>
<td>59 (100%)</td>
</tr>
<tr>
<td><strong>K-group</strong></td>
<td>165 (56.7%)</td>
<td>126 (100%)</td>
</tr>
<tr>
<td>HO</td>
<td>40 (59%)</td>
<td>68 (100%)</td>
</tr>
<tr>
<td>HT</td>
<td>4 (80%)</td>
<td>5 (100%)</td>
</tr>
<tr>
<td>TT</td>
<td>55 (85%)</td>
<td>65 (100%)</td>
</tr>
<tr>
<td>YN</td>
<td>24 (62%)</td>
<td>39 (100%)</td>
</tr>
<tr>
<td><strong>J-group</strong></td>
<td>123 (69.5%)</td>
<td>177 (100%)</td>
</tr>
<tr>
<td>BW</td>
<td>8 (38%)</td>
<td>21 (100%)</td>
</tr>
<tr>
<td>GP</td>
<td>13 (32%)</td>
<td>41 (100%)</td>
</tr>
<tr>
<td>MT</td>
<td>20 (35%)</td>
<td>57 (100%)</td>
</tr>
<tr>
<td>TR</td>
<td>34 (76%)</td>
<td>45 (100%)</td>
</tr>
<tr>
<td><strong>E-group</strong></td>
<td>75 (45%)</td>
<td>164 (100%)</td>
</tr>
</tbody>
</table>
(2) Table E.2. Frequency (%) of case marking: Number of data tokens.

<table>
<thead>
<tr>
<th></th>
<th>Marked</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>HJ</td>
<td>110 (58%)</td>
<td>191 (100%)</td>
</tr>
<tr>
<td>SC</td>
<td>151 (85%)</td>
<td>177 (100%)</td>
</tr>
<tr>
<td>SS</td>
<td>80 (47%)</td>
<td>170 (100%)</td>
</tr>
<tr>
<td>SY</td>
<td>78 (46%)</td>
<td>169 (100%)</td>
</tr>
<tr>
<td><strong>K-group</strong></td>
<td><strong>419 (59.3%)</strong></td>
<td><strong>707 (100%)</strong></td>
</tr>
<tr>
<td>HO</td>
<td>103 (68%)</td>
<td>152 (100%)</td>
</tr>
<tr>
<td>HT</td>
<td>4 (27%)</td>
<td>15 (100%)</td>
</tr>
<tr>
<td>TT</td>
<td>117 (72%)</td>
<td>163 (100%)</td>
</tr>
<tr>
<td>YN</td>
<td>92 (70%)</td>
<td>132 (100%)</td>
</tr>
<tr>
<td><strong>J-group</strong></td>
<td><strong>316 (68.4%)</strong></td>
<td><strong>462 (100%)</strong></td>
</tr>
<tr>
<td>BW</td>
<td>13 (31%)</td>
<td>42 (100%)</td>
</tr>
<tr>
<td>GP</td>
<td>42 (26%)</td>
<td>161 (100%)</td>
</tr>
<tr>
<td>MT</td>
<td>8 (7%)</td>
<td>117 (100%)</td>
</tr>
<tr>
<td>TR</td>
<td>168 (86%)</td>
<td>195 (100%)</td>
</tr>
<tr>
<td><strong>E-group</strong></td>
<td><strong>231 (44.9%)</strong></td>
<td><strong>515 (100%)</strong></td>
</tr>
</tbody>
</table>
REFERENCES


Generative Grammar, 4(1).


acquisition, 4, 335-443.


Studies in Language, 13(1), 105-128.


Australian languages.


282


