Morphological Processing of Korean Verb Forms in Visual Word Recognition

D i s s e r t a t i o n

zur

Erlangung des akademischen Grades

Doktor der Philosophie

in der Philosophischen Fakultät

der Eberhard Karls Universität Tübingen

Vorgelegt von

Mihi Kang

aus Seoul, Korea

2025

Gedruckt mit Genehmigung der Philosophischen Fakultät
der Eberhard Karls Universität Tübingen

Dekanin: **Prof. Dr. Angelika Zirker**

Hauptberichterstatter: **Prof. Dr. Harald Baayen**
Mitberichterstatterin: **Apl. Prof. Dr. Inga Hennecke**

Tag der mündlichen Prüfung: **31.10.2025**

Universitätsbibliothek Tübingen, TOBIAS-lib

To my husband, 정민 & my daughter, 나래

Acknowledgment

Throughout the long years of my doctoral journey, I often wondered if a day would come when I could write these very words. Regardless of the outcome, I felt deeply that I must one day find a way to thank those who stood by me in so many ways.

First and foremost, I would like to sincerely thank my supervisor, Professor Harald Baayen. He has supported and encouraged me continuously with genuine academic curiosity and passion, without any prejudice toward my background. Prof. Baayen guided me with trust and kindness, and helped me come this far. Despite my many shortcomings, he listened to my ideas without bias, encouraged me, and gave me the confidence to keep going. Without his support and generosity, I could not have completed this thesis. Whenever I visited the university as a long-distance student, he always welcomed me warmly, took the time to talk with me, and kindly shared meals and thoughtful discussions. I am deeply grateful for those moments. For several years, we had almost weekly online meetings. I was deeply impressed by how patiently and thoroughly he explained things I struggled with, especially in statistics, and how much effort he put into guiding me. All of my experiences with Prof. Baayen left a strong impression on me and have become some of the most valuable memories of my life. I will always remember his cheerful and kind smile, and I know I'll miss it very much.

There is someone who was truly indispensable throughout my academic journey. I sincerely thank Dr.Yu-Ying Chuang, who is my second supervisor, for providing immense practical help and warm encouragement during the first four years of my studies at Tübingen University and continued to do so thereafter. From the moment I first visited Tubingen University, Yu-Ying welcomed me warmly and offered continuous support. Without Yu-Ying's help and consideration, it would have been extremely difficult for me to navigate many of the new challenges on my own. I am truly grateful for Yu-Ying's generous and sincere support. Yu-Ying's genuine warmth, kindness, and friendship left a lasting impression on me, and I will cherish those memories for a long time.

Although I regret not having had more frequent exchanges with other members of the research team, I would like to especially thank Dr.Tino Konstantin for his practical assistance with my experiments. Tino offered crucial help in developing the online experiment platform and, more importantly, generously shared his knowledge of statistics, taking time from his own schedule to support me. I am deeply thankful for this kindness. I am again deeply grateful for his crucial support, without which this experiment could not have been carried out.

And one indispensable person I must mention is Mrs. Baayen, Tineke, who always supported me with a warm smile and kind emails in handling administrative matters. Each time I visited the university, she welcomed me warmly, and I will always cherish the memories of the pleasant meals we shared together with Prof. Baayen. I sincerely thank her for all her kind and warm support. Though she is now retired, I truly hope she enjoys good health, happiness, and plenty of peaceful, fulfilling days ahead.

Finally, I owe my deepest gratitude to my husband, Joungmin, who has supported me with unwavering love and devotion throughout this entire journey. Over the course of nearly ten years, before and during this PhD, he has stood by my side, shouldering the burdens of daily life and offering boundless encouragement through every difficulty. This long journey could never have been accomplished without his devoted sacrifice and generous understanding. No amount of gratitude could ever fully express how thankful I am for his sacrifices and support. I also want to express my love and heartfelt thanks to my daughter Narae, who grew up alongside this journey. Despite the long periods of my distraction and limited care, she has grown up with a bright spirit and endless patience. I am deeply grateful

and, at the same time, sorry for the times I could not be fully present. I'm truly looking forward to finally spending more precious time with our beloved daughter.

Now, I am eager to return the love and time I owe to my family and embrace a new chapter with them.

To all of you —thank you, from the bottom of my heart.

Abstract

This study investigates how native and advanced non-native speakers of Korean process morphologically and semantically complex inflected Korean verb forms through a visual lexical decision task. A total of 4,000 words and 1,000 non-words were used as stimuli, generated by combining 100 verb lexemes with approximately 90 verb endings. The goal was to examine how morphological, semantic, and phonological factors influence the lexical processing of native and non-native Korean speakers.

The results revealed notable differences between native and non-native speakers. Both groups were strongly influenced by specific inflected endings that were morphologically and semantically patterned. However, for non-native speakers, syllable-based word length emerged as the most influential factor in Korean verb recognition, surpassing the effect of specific endings. Frequency also played a major role: ending frequency had the greatest impact across all groups, followed by stem frequency for native speakers, while non-native speakers were more affected by whole-word frequency than by stem frequency. In the native Korean group, although weaker than other frequency effects, a whole-word effect was observed, suggesting that certain verb forms may be processed as full-forms. In most non-native speaker groups, the stem frequency effect appeared weak. However, the Chinese L1 group showed a notably different pattern, where the effects of stem frequency and whole-word frequency appeared to be similar. This is a particularly notable result, as it contrasts sharply with the Vietnamese L1 (first language) group, which is typologically similar to Chinese. It highlights the relevance of typological L1 effects in L2 (second language) processing.

The Korean non-native participants were divided into four groups according to their L1: Japanese, Mongolian, Chinese, and Vietnamese, with Japanese-Mongolian and Chinese-Vietnamese grouped together based on typological similarity. The Chinese group exhibited unique response patterns that set them apart

from the other groups. While clear L1 effects were not observed, some findings could be partially attributed to L1 influence. In the nonword task, native speakers responded most slowly to morphologically and semantically complex forms. The Japanese and Mongolian groups, similar to native speakers, responded more slowly to nonwords manipulated at the ending level, suggesting heightened sensitivity to verb endings. In contrast, the Chinese and Vietnamese groups were more sensitive to manipulations at the stem or syllable level. Error patterns in nonwords further showed that the highest error rates occurred in phonologically manipulated forms and in nonwords generated by incorrect allomorph switching.

This study also explored the applicability of a computational model, the Discriminative Lexicon Model (DLM), to Korean verb processing. Using 59 verb-final inflectional forms, the model achieved high accuracy during both training and testing phases when using bi-syllabic representations. While these results are promising, further research is needed to evaluate the model's performance on a broader range of Korean verb endings, particularly those involving mood and other complex semantic functions.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

The study of word forms and how humans process them in real language use has been a central area of research in psycholinguistics. This topic has been examined through various approaches and perspectives. One of the fundamental and intuitive ways to explore this subject is by considering the diversity of linguistic forms and the potential semantic differences between languages. Word formations and their semantic components vary significantly across languages, which suggests that different processing strategies may be required to accommodate these variations.

This discussion rests on two major questions. The first relates to how we conceptualise the notion of a "word" in terms of form and meaning. In other words, it raises the question of how much meaning is expressed through form and how much form conveys meaning within the concept of a word. Another question is how these diverse meanings and forms are processed by language users in actual language use.

In essence, it comes down to distinguishing between complex and relatively less complex linguistic forms with respect to both their meanings and structural properties. For instance, in Korean, one might compare a complex word such as 드렸을지언정 ("even though (you/she/he/they) may have given (it to a senior person) ...") with a less complex word like 사과들 ("apples"). These forms differ markedly in complexity, both in their structure and meaning. While it seems intuitive that processing these forms involves different levels of complexity, the theoretical frameworks addressing how humans process such intricate linguistic forms remain unclear. There is ongoing debate about how best to approach this question.

This issue becomes particularly relevant when considering second language (L2) processing. Non-native speakers might approach the processing of linguistic forms in their native and non-native languages differently. In many cases, these differences in processing — stemming from variations in linguistic forms and meanings — seem like a natural consequence of the disparities between the two languages. For native speakers, the distinction between simple and complex forms might not play a significant role in comprehension or production. However, for non-native speakers, proficiency in the second language could greatly affect how these forms are processed. Complex forms, in particular, might require distinct strategies for understanding their detailed meanings and forms in both comprehension and production. However, in non-native language processing, there are typically many complex and intertwined variables, making it difficult to draw simple conclusions based on surface-level results.

Examining how non-native speakers process complex language forms could provide valuable clues to understanding human language processing more broadly. Non-native speakers often appear to rely on logical reasoning to interpret forms and meanings in the target language. However, as proficiency in the target language increases, this reliance tends to decrease (Ellis, 2006). Numerous factors, such as the learning environment, individual differences, and other variables, shape this process, making it a highly complex area of study.

Despite the diversity of word forms and semantic ranges across languages, existing theories of language processing often fail to adequately account for this variation, relying instead on simplifying generalisations. While it is reasonable to differentiate word processing based on the complexity of form and meaning, theoretical perspectives remain divided on this issue. The differences in lexicalisation and grammaticalisation observed in cross-linguistic studies(e.g., Frajzyngier, 2008; Xing, 2015; Narrog and Heine, 2018) highlight the variations in the types, number, and modes of expression between meaning and form. This, in turn, underscores the importance of considering language-specific factors alongside universals in meaning-form processing. Particularly in non-native language processing, such cross-linguistic differences raise intriguing questions, such as how these differences affect non-native processing, and how responses may vary between typologically similar or dissimilar languages. In this regard, it is necessary to examine and explore a broader range of linguistic forms.

The Korean verbal[1] morphology examined in this study is highly complex from both morphological and semantic perspectives, differing significantly in many aspects from the languages typically studied in language processing research and theories. Research on Korean verbs has mainly been conducted locally within Korea, often with limited scope, with their morphological and semantic particularities not widely recognised or discussed in the broader field of linguistics.

This study aims to provide an in-depth examination of these morphologically and semantically complex Korean verbs and investigates how they are processed. Specifically, as mentioned earlier, the primary focus is on understanding how these complex forms and meanings are processed by non-native speakers with different first language (L1) backgrounds. Non-native lexical processing involves various factors, and one of the key interests of this study is to explore how these factors shape the processing of complex Korean verbal morphology.

To address this, the current study conducted a visual lexical decision task experiment with native Korean speakers and four groups of non-native speakers from different typologically distinct L1 backgrounds. By including these typologically varied groups, the study seeks to examine the influence of L1 on processing. In addition, a range of linguistic variables were considered and incorporated into the analysis of the experimental

[1] In Korean, both verbs and adjectives appear at the end of a sentence in the predicate position. In this context, they are collectively referred to as '용언'/yoŋʌn/'(predicative words), and they undergo similar grammatical inflectional processes. For convenience, this study will refer to these forms (i.e., verbs and adjectives) simply as 'verbs' or 'verb forms'.

data.

This thesis is structured as follows. Chapter 2 discusses several related topics concerning the fundamental concepts of form and meaning in word formation. Chapter 3 will provide a detailed overview of Korean verb morphology. Chapter 4 will review previous studies on Korean native speakers' processing of Korean verbs, and Chapter 5 will present and discuss the results of a visual lexical decision experiment conducted with native speakers. Chapters 6 and 7 will focus on previous research and the current experimental findings concerning non-native speakers' processing of Korean verbs. In Chapter 8, the experimental results of native and non-native speakers will be compared comprehensively through analyses of reaction times and errors. Finally, Chapter 10 will explore the potential of surface-level word processing by applying a new computational model, the Discriminative Lexicon Model(DLM, Baayen et al., 2019; Heitmeier et al., 2025), to Korean verb processing, focusing on a paradigmatically structured subset of the Korean verb system.

Chapter 2

Meaning and Form

2.1 Inflection for Semantics

One of the core functions of word inflection is to enrich words' semantics. While the semantics realised in inflectional forms are often construed as grammatical functions, they contribute more to the overall meaning of the word in some languages, and agglutinative languages in particular. It is unclear to what extent the richness of meaning in such languages can be adequately captured with theoretical constructs such as the morpheme, and theories positing straightforward one-to-one correspondences between form and meaning. This notion has sparked prolonged discourse regarding the existence of a 'morpheme' and the inconsistency of its conceptualisation as a unit bridging meaning and form (refer to Blevins, 2016).

In some respects, the argument of the reality of morphemes seems to advocate for a rigid one-to-one correspondence between meaning and form solely concerning the inflectional forms of words. The way in which we attribute meaning to a linguistic form varies, encompassing the kind of meaning assigned, such as distinguishing between simple and abstract nouns, and the manner in which complex concepts are denoted by a single term, leading to varied interpretations of the relationship between linguistic form and meaning.

However, despite the multitude of meanings and types of linguistic forms, interpretations of their relationship may vary depending on the criteria used. Particularly in the interpretation of meanings in inflectional forms, opinions diverge regarding the applicability of the concept of 'morpheme', understood as the smallest linguistic form corresponding to a single meaning, for accommodating diverse combinations of meanings. In other words, the notion that the concept of a morpheme is contested arises from several reasons, chief among them being that meanings and forms often do not exhibit a one-to-one correspondence. However, viewing both the concept of a morpheme and its opposing views under the premise that the relationship between meaning and form must adhere to a one-to-one correspondence is contradictory in interpreting the meaning-form relationship. As noted above, interpreting meaning–form relations in inflectional morphology through this lens of one-to-one correspondence may be limiting and does not account for the full range of observed variations and the meaning-form complexities. That is, the diversity of meanings that humans can express through language and the forms that convey those meanings are not realised within a perfectly systematic structure. Rather, the relationship between form and meaning is shaped and interpreted through a more complex interplay with various linguistic and extralinguistic factors.

Setting aside these analyses and interpretations, inflected forms play a crucial role in conveying meaning. Inflectional variation leads to richer semantics and greater morphological complexities. Consequently, a key concern arises regarding how inflectional semantics is expressed in form, and how inflectional variation in form and meaning is handled cognitively.

This study focuses on the Korean verb. In general, of all parts of speech, verbs appear to be the most complex in many languages. Several studies(e.g., Peelen et al., 2012; Bedny and Thompson-Schill, 2006; Hauptman et al., 2022) have observed that verb processing elicits more brain activity, and require more neuronal resources than noun processing. This may be attributed to the fact that verbs have greater syntactic complexity and are more context-dependent. The Korean verb, as will become clear below, can be combined with exponents realising a wide range of syntactic, pragmatic, social, and discourse functions. Furthermore, the combinatorics of the exponents of the Korean verb enable the expression of countless emotional nuances. Since semantic transparency has been identified as one of the significant factors affecting word processing(e.g., Amenta and Crepaldi, 2012; Bozic et al., 2013; Kim et al., 2019; Günther and Marelli, 2019; Libben et al., 2003; Dohmes et al., 2004; Pollatsek and Hyönä, 2005; Marslen-Wilson et al., 1994; Longtin et al., 2003), the question arises of whether there are cognitive differences and specific processes involved in the lexical processing of inflected forms that encode subtle and complex emotions, compared to relatively straightforward inflections such as those marking number or tense.

The semantic complexity of the Korean verb is paralleled by considerable complexity at the level of morphological form. For instance, although Korean is an agglutinative language, some of its exponents are fusional. Furthermore, the exponents of the Korean verb are subject to considerable phonological restructuring, which raises the question of how transparent the morphological system of Korean actually is.

Remarkably, no matter how complex the form and meaning, native speakers appear to encounter little difficulty in processing the complexities of the Korean verb. However, for non-native speakers, it is anticipated that the patterns of processing will vary depending on factors such as language proficiency and the influence of their native language. From this perspective, it is especially intriguing to explore how such complex form and meaning are perceived and processed, particularly by non-native speakers. This study, therefore, seeks to explore how linguistically complex forms, both morphologically and semantically, are actually processed by native and non-native speakers. By comparing these two groups, whose language acquisition backgrounds and processes differ fundamentally, we aim to gain deeper insight into how such complex forms are handled during lexical processing.

2.2 Paradigm in Korean Verbal Morphology

An inflectional paradigm can be defined as 'a set of words that are inflectionally related and share the same lexical stem' (Bybee, 1985). Paradigms are used to clarify how inflectional variants are organized according to their forms and the inflection features that these forms realise. Some generally discussed features (categories) of inflectional forms typically include person, gender, number, case, tense, aspect, and mood, as found in Indo-European languages, which significantly differ from the functional and semantic concepts found in Korean verbal inflection.

There have been several discussions (e.g., Chung, 2015; Kim T., 2015; Kim H., 2006) that the concept of paradigms does not effectively serve as a crucial framework for understanding Korean verbal morphology. It is argued that the agglutinative nature of Korean verb morphology is at odds with the fusional morphology that characterises the inflectional systems of many Indo-European languages. For instance, Chung (2015) commented that various inflected forms of Korean verbs did not change uniformly or simultaneously over the centuries; instead, they underwent individual changes(i.e., involving the attachment of affixes to verb stems), leading to the asymmetric development of paradigms. This raised questions about the applicability of the concept of inflection[2], originally formulated for inflectional languages, to Korean, an agglutinative language, especially in light of the huge number of inflected forms and their grammatical and semantic functions.

This observation underscores the distinction between the elements forming the endings attached to verb stems in Korean, which do not align with the characteristics of inflectional elements found in many other inflectional languages. While the inflectional elements in many Indo-European languages add more grammatical meanings and functions to the stem, Korean verbal inflectional forms express more relational aspects of communicative settings. Specifically, the inflectional categories in Korean verb forms,

[2]Nevertheless, since a lot of phonological adjustments and some fusional forms are also observed in Korean verbal morphology, there is a contradiction in clearly distinguishing and applying these two terms between inflection and agglutination to Korean verbal morphology.

such as subject honorifics, speech level, illocutionary force, and syntactic connective forms embracing various modal nuances (referred to as a mood or mood forms in this study), set the Korean verb apart, unlike typical inflectional languages where inflection primarily encodes grammatical meanings such as tense, aspect, person, number, and gender.

In the intricate semantic relationships inherent in such inflectional forms, the combinatorial possibilities of exponents pose a considerable challenge for the construction of paradigms for Korean verb forms. This particular challenge originates with the exponents for mood, many of which have to be translated into English with modal constructions. Mood forms in Korean play a pivotal role in expanding combinatorial possibilities, expressing subtle nuances in meaning. Furthermore, various semantic and grammatical constraints and extensive phonological readjustments give rise to a vast number of allomorphs and inflectional forms, many of which deviate from the limited standardised patterns typically associated with paradigms.

In this regard, questions have been raised concerning the limitations of defining specific verb forms and the number of forms within the framework of paradigms (e.g., Kim T., 2015; Kim H., 2006). More than 4,000 different types of verbal exponents are attested in 15-million-word corpora[3] (Kang and Kim, 2009) and 5,682 distinct combinations of exponents (henceforth, 'endings') were reported in a 10-million-word corpus of written texts[4] (Choi, 2012). In larger corpora, even more exponents and endings may be expected. Thus, the possible combinations of this vast number of exponents are virtually limitless. Moreover, as a consequence of the interactions of semantic and grammatical constraints, it is not feasible to set up tables of possible inflectional variants, as can be done for verbal and nominal forms in many typical inflectional languages.

From this perspective, investigating how Korean verb forms are processed presents a highly compelling line of inquiry. By examining Korean verb morphology, this study may provide an opportunity to assess the applicability and explanatory power of pre-

[3] The corpora were based on written Korean collected between 2001 and 2007 by the National Institute of Korean Language for the 21st Century Sejong Project.

[4] The modern written Korean part-of-speech tagged corpus(현대 문어 형태 분석 말뭉치) released in 2007 by National Institute of Korean Language for 21C Sejong plan.

viously proposed theoretical frameworks in the context of Korean. The next chapter provides a detailed overview of Korean verb morphology, serving as the foundation for both evaluating theoretical claims in Korean linguistics and conducting the lexical processing experiments reported in this thesis.

Chapter 3

Korean Verbal Morphology

This chapter will closely examine Korean verbal morphology and briefly discuss certain phonological features and orthographic conventions that often pose challenges in lexical processing.

3.1 Structure and Properties of Korean Verbs

Korean is an agglutinative language with rich verbal morphology. Verbs in Korean are inflected for tense, politeness, register, illocutionary force, and various modal meanings. How a Korean verb is inflected depends on its position in the sentence. Figure 3.1 highlights the three main configurations in which verbs occur, depending on specific syntactic structures. For descriptive convenience, we follow Sohn (1994) and distinguish between pre-final exponents (in white) and final exponents (in purple). Korean has SOV word order, and in a simple clause (1) with a final main verb, the verb can be inflected for 7 inflectional categories, with fixed serial positions that all follow the verb stem.

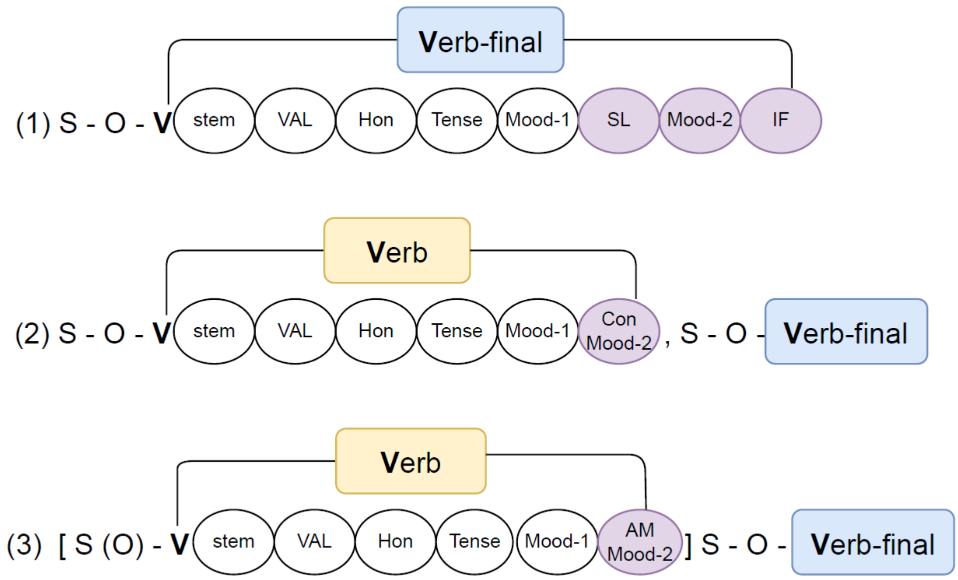


Figure 3.1: Types of Korean verbs according to different syntactic structures: Various inflectional elements form verbs in Korean as a long chain of suffixes. The term 'Verb-final' denotes verb constructions that occur in sentence-final position. The construction (3) shows a relative clause as subject (object cases omitted). Abbreviations: S(subject), O(object), V(verb), VAL(valence : passive/causative), HON(subject honorifics), SL(speech level), IF(illocutionary force), CON(conjunctive), AM(adnominal modifier).

Verbs are inflected for voice, causation, honorifics, tense, and mood. In Korean linguistics (see, Choo, 2003; Lee, 1991; Park, 2004; Jeong, 2018, for detailed discussions), mood (or modal forms) denotes the speaker's attitude or emotions, adding nuance to the expression. In this study, under the broad concept of mood exponents, we include connective endings that typically realise diverse meanings and nuances, as well as noun modifiers, adverbial, and nominal endings that are found in the same morphological position in the verb. Furthermore, in Figure 3.1, the sentence-final main verb (1) is inflected for speech level, mood, and illocutionary force; non-final main verbs in a conjoint sentence (2) only have a slot for conjunction (CON/MOOD-2), and the verb in a relative clause (3) only has a slot for an adnominal modifier (AM/MOOD-2) in the relative clause.

Table 3.1: Inflectional dimensions of the Korean sentence-final verb. Speech levels and illocutionary force are realised only on the sentence-final verb. The two speech levels in parentheses are not frequently used in modern Korean. Exponents for illocutionary force are realised only in the formal and plain speech levels. For the intimate and polite speech levels, intonation is used to express illocutionary force.

type	position	unmarked	marked
valence	(1)	active	passive, causative
honorifics	(2)	plain	honorific
tense	(3)	present (non-past)	past, future (periphrasis)
mood 1	(4)	neutral	volition, conjecture
speech level	(5)	casual (반말, Banmal)	formal, polite, (semi-formal, familiar), intimate, plain
mood 2	(6)	neutral	reasoning, intention, exclamation, retrospective, realisation, etc.
illocutionary force	(7)	realised by intonations in polite/intimate speech levels	declarative, inquisitive, imperative, propositive

Table 3.1 provides a summary of the morphology of the Korean sentence-final verb, and the following sentence provides an example of the sentence-final verb form with exponents for all seven features. Abbreviations : Acc(accusative), Aux(auxiliary), Caus(causative), Hon(honorifics).

고기를 **먹이시었겠습디다.**

kogi-rɯl mʌg-i-çi-ʌt̚-k͈et̚-s͈ɯp̚-t͈i-da

meat-Acc eat-Caus-Hon-Past-Conjecture-Formal-Retrospective-Declarative

(He/she/they) must have fed (someone) meat.

The sentence above represents a verb form in which the number of possible exponents is maximised. In actual usage, however, verb forms typically consist of combinations of only one to three exponents (for further details, see Choi, 2012). The following examples will illustrate how these exponents are actually applied in some practical verb forms.

We first consider the pre-final exponents (white slots in Figure 3.1). If the verb is in the passive, or if it is causativised, the corresponding exponents are realised immediately following the verb stem. Next, in Korean, the concept of honorifics is expressed in various ways. It refers to the linguistic expression of special respect toward others, either the listener or another person mentioned in speech (excluding oneself). This respect is typically conveyed by altering linguistic forms(e.g., verbs, nouns, particles) to either elevate the other person or humble oneself. One common way this is realised is through verb inflection. The so-called 'subject honorifics' (Hon) are a form of honorification that elevate the subject of the verb—the person who performs the action. In conversational speech, this subject will often be the addressee, but in narration, the subject can also be a third person. Subject honorifics are realized by means of the exponent -시-/ɕi/, which follows the verb stem (or the passive or causative exponents, if present), as illustrated in the examples in the formal speech level below.

(1) 신문을 봅니다.
ɕinmun-ɯl po-m-ɲida
newspaper-Acc see-Formal-Declarative
'(any subject) read a newspaper.'

(2) 신문을 보십니다.
ɕinmun-ɯl po-çi-m-ɲida
newspaper-Acc see-Hon-Formal-Declarative
'(any subject except 'I') read a newspaper.'

As seen in the above examples, the subject in the Korean sentence can be omitted and the verb forms do not inflect for person and number. However, there are some semantic restrictions on what possible values of person are to apply the subject honorifics. The form for subject honorifics -시-/ɕi/ cannot combine with a first person as it does not make sense to express particular respect to oneself. On the other hand, volitional mood -겠-/kɛt̚/ is used only for a first person in statements and for second persons in questions, due to semantic constraints.

The exponents for the past and future follow the exponent for the subject honorifics. The exponents for past and future reference are -ㅆ-/t̚/, -았-/at̚/, -었-/ʌt̚/ and -ㄹ-/l/, -을-/ɯl/ respectively. The following sentence (3) shows a past tense verb form with a polite speech style but without a particular illocutionary form. In the polite and intimate speech levels, intonation is varied to convey differences in illocutionary force.

(3) 어제 신문을 봤어요.
ʌdze ɕinmun-ɯl pw-as͈-ʌyo
yesterday newspaper-Acc see-Past-Polite
'(any subject) read a newspaper yesterday.'

The future tense requires a periphrastic construction that involves the nominaliser -것/-거/k͈ʌ/ (meaning 'thing'), as illustrated in the next example (4):

(4) 신문을 볼 겁니다.
ɕinmun-ɯl po-l k͈ʌ-m-ɲida
newspaper-Acc see-Future Nominalizer-Formal-Declarative
'(any subject) will read a newspaper.'

If tense is not overtly marked, the verb typically realises present tense (see sentence (5)), but it can also express habitual activities and the expected future as well:

(5) 매일 신문을 봐요.
mɛ-il ɕinmun-ɯl pw-ayo
everyday newspaper-Acc see-Polite
'(any subject) read a newspaper everyday.'

The last pre-final slot of the verb optionally realises volitional or conjectural mood, as illustrated in the following examples. When a verb is in volitional or conjectural mood, the subject is necessarily understood to be a first person (declarative) or a second person (interrogative).

(6) 신문을 보겠습니다.
ɕinmun-ɯl po-get̚-s͈ɯm-ɲida
newspaper-Acc see-Volitional-Formal-Declarative
'I(We) will read a newspaper.'

(7) 신문을 보겠습니까?
ɕinmun-ɯl po-get̚-s͈ɯm-ɲik͈a
newspaper-Acc see-Volitional-Formal-Inquisitive
'Will you read a newspaper?'

(8) 그 사람이 아마 먼저 그 신문을 보겠어요.
kɯ s^{h}aram-i ama mʌndzʌ kɯ ɕinmun-ɯl po-ges͈-ʌyo
that person-Nom maybe first the newspaper-Acc see-Conjectural-Polite
'I think that person will probably read the newspaper first.'

(9) 그 사람이 그 신문을 보겠어요?
kɯ s^{h}aram-i kɯ ɕinmun-ɯl po-ges͈-ʌyo
that person-Nom the newspaper-Acc see-Conjectural-Polite
'Do you think that person will read the newspaper?'

For Mood-1, the only exponent that is productive is -겠-/ket̚/(volitional or conjectural mood). The other exponents for Mood-1, -으리-/ɯri/(conjectural), -더-/-dʌra/(retrospective), and -으니-/-ɯɲi/(reason) sound poetic or archaic (Park, 2004)[5].

There are three final slots for exponents, one for speech level, one for further exponents realising mood, and one for illocutionary force. The exponent for speech level is described in the literature as realising 'addressee(hearer) honorifics'. Six levels based on the relation between the interlocutors and speech contexts are distinguished: formal, polite, semi-formal, familiar, intimate and plain. In present-day Korean, the semi-formal and familiar levels are rarely used. A key factor determining the choice of speech level is the relative age and the closeness between interlocutors. It is natural that younger people are expected to use the polite forms when addressing older people. For interlocutors

[5]The national standard dictionary of Korean, 국립국어원 표준국어 대사전 (Standard Korean Dictionary of the National Institute of Korean Language, `https://stdict.korean.go.kr/main/main.do`) categorizes the Mood-1 exponents as 'sentence final endings' rather than as 'pre-final' endings.

of the same age, the polite forms are used for strangers, and the intimate/plain forms for close friends, family members and between children. Illocutionary force is expressed through exponents only in combination with the formal, plain, and familiar speech levels. Table 3.2 presents the exponents realising speech level and illocutionary force. In this table, speech levels and illocutionary forces are presented in an integrated form, considering the ambiguity in distinguishing between meaning and form(see some relevant discussions in Chuang et al., 2022). There is also considerable allomorphy, as well as syncretism.

Table 3.2: Exponents realising speech level and illocutionary force. For the polite and intimate speech levels, different intonation patterns are used to express illocutionary force.

Exponent	IPA	Speech level	Illoc. force
-ㅂ니다/-습니다	-mɲida/-s͈ɯmɲida	formal	declarative
-ㅂ니까/-습니까	-mɲik͈a/-s͈ɯmɲik͈a	formal	inquisitive
-십시오/-으십시오	-ɕip̚ɕ͈io/-ɯɕip̚ɕ͈io	formal	imperative
-ㅂ시다/-읍시다/-(으) 십시다	-p̚ɕ͈ida/-ɯp̚ɕ͈ida/-ɯɕip̚ɕ͈ida	formal	propositive
-오/-소	-o/-s^{h}o	semi-formal	dec/inq/imp
-ㅂ시다/-읍시다	-p̚ɕ͈ida/-ɯp̚ɕ͈ida	semi-formal	propositive
-아요/-어요	-ayo/-ʌyo	polite	(intonations)
-아/-어	-a/-ʌ	intimate	(intonations)
-네	-ne	familiar	declarative
-나	-na	familiar	inquisitive
-게	-ge	familiar	imperative
-세	-s^{h}e	familiar	propositive
-ㄴ다/-는다/-다	-nda/-nɯnda/-da	plain	declarative
-니/-냐	-ɲi/-ɲya	plain	inquisitive
-라/-아라/-어라	-ra/-ara/-ʌra	plain	imperative
-자	-dza	plain	propositive

The following examples illustrate commonly used combinations of speech level and illocutionary force.

(10) 신문을 봅니다.
ɕinmun-ɯl po-m-ɲida
newspaper-ACC see-FORMAL-DECLARATIVE
'I am reading a newspaper.'

(11) 신문을 봐요.
çinmun-ɯl pw-ayo
newspaper-ACC see-POLITE
'I am reading a newspaper.'

(12) 신문을 봐.
çinmun-ɯl pw-a
newspaper-ACC see-INTIMATE
'I am reading a newspaper.'

(13) 신문을 본다.
çinmun-ɯl po-nda
newspaper-ACC see-(PLAIN & DECLARATIVE)
'I am reading a newspaper.'

(14) 신문을 봅니까?
çinmun-ɯl po-m-ɲik͈a
newspaper-ACC see-FORMAL-INQUISITIVE
'Do you read a newspaper?'

(15) 신문을 보니?
sinmwun-ul po-ni
newspaper-ACC see(read)- (PLAIN & INQUISITIVE)
'Do (you) see(read) a newspaper?'

(16) 신문을 보십시오.
çinmun-ɯl po-çi-p-ç͈io
newspaper-ACC see(read)-HON-FORMAL-IMPERATIVE
'Please read a newspaper.'

(17) 신문을 보아라 (봐라).
çinmun-ɯl po-a-ra(pw-a-ra)
newspaper-ACC see-PLAIN-IMPERATIVE
'Read the newspaper.'

(18) 신문을 봅시다!
ɕinmun-ɯl po-p̚-ɕ͈ida
newspaper-ACC see-FORMAL-PROPOSITIVE
'Let's read the newspaper!'

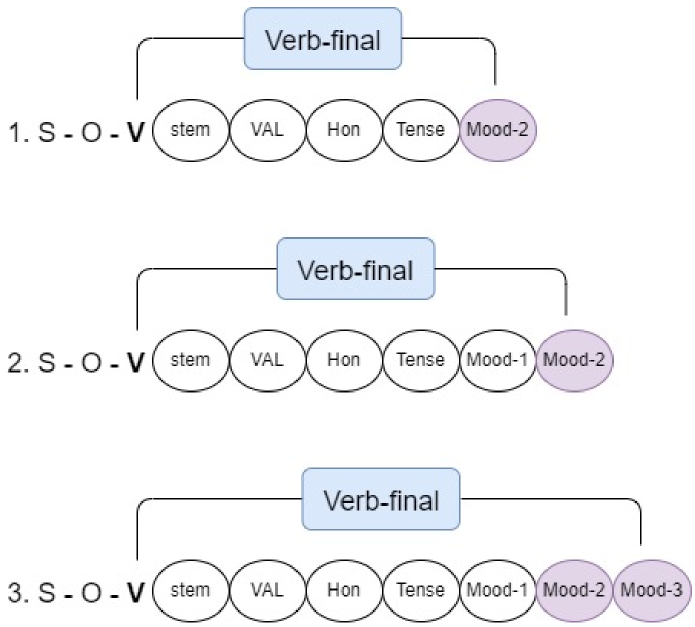
(19) 신문을 보자!
ɕinmun-ɯl po-dza
newspaper-ACC see -(PLAIN & PROPOSITIVE)
'Let's read the newspaper!'

There are restrictions on which feature values of tense and illocutionary force can be felicitously combined. For instance, imperative and propositive illocutionary force are semantically incompatible with future and past tense. For instance, 가라 /kara/ (go!) (PLAIN SPEECH & IMPERATIVE) in present tense makes sense, but the same command for the past tense does not. Furthermore, some conjunctive forms(e.g., -서 /-sʰʌ/ and -(으) 므로 /-ɯmɯro/) also come with restrictions on tense and illocutionary force. Furthermore, it has been argued that -습디다 /s͈ɯp̚t͈ida/ (FORMAL-RETROSPECTIVE-DECLARATIVE) and -습디까/s͈ɯp̚t͈ik͈a/ (FORMAL-RETROSPECTIVE-INTERROGATIVE) are independent exponents in their own right (see, Choi, 2015, for further discussion).

The Mood-2 slot in the final endings accepts a great many exponents for the speaker's attitude towards the contents of an utterance (cf. the definition of mood given by Jespersen, 1924), including perception, reasoning, speculation, intention, boundaries, regret, complaint, emphasis, apathy, concern, concession, explanation, confirmation, admiration, familiarity, origin, promise, and permission, which complement or further narrow down what can be realized with the exponents for honorifics, speech level, and illocutionary force. The first example sentence earlier, '고기를 먹이시었겠습디다' illustrates the simultaneous use of the two mood positions but this realization of Mood-2 in combination with the formal speech level is exceptional.

Most other exponents of Mood-2 require another construction, referred to as 반말 (Banmal, 'half speech', i.e., casual speech). Verb forms in the Banmal construction typically do not have exponents for speech level and illocutionary force (Cha, 1990; Ko, 2001; Park, 1998). Figure 3.2 presents the constructions for the final verb in Banmal. Up to three exponents for mood can be present. Although Banmal is predominantly used for casual speech, it is possible to make such forms more polite by adding the exponent for the polite speech level -요/yo/ to the end of the verb. The following examples illustrate these constructions in their simple, non-polite form.

Figure 3.2: Banmal verb constructions



(20) 그 사람이 신문을 보네!
kɯ s^{h}aram-i ɕinmun-ɯl po-ne!
that person-NOM newspaper-ACC see-EXCLAMATION
'That person is reading a newspaper!' (This is amazing!)

(21) 그 사람이 신문을 보겠네!
kɯ s^{h}aram-i ɕinmun-ɯl po-gen-ne!
that person-NOM newspaper-ACC see-CONJECTURE(MOOD-1)-EXCLAMATION
'That person must be reading a newspaper!' (This is worrisome!)

(22) 많은 사람이 그 신문을 보겠더군.
manɯn s^{h}aram-i kɯ ɕinmun-ɯl po-get̚-t͈ʌ-gun!
many person-NOM that newspaper-ACC see-CONJECTURE-RETROSPECTIVE-EXCLAMATION
'It seemed that many people were going to read that newspaper.'

When the verb is in non-final position, the final exponent realises either a conjunction (e.g., -고/go/(and), -지만/dziman/(but), -서/s^h/(so/ because, ...) or an adnominal modifier for a relative clause functioning as a subordinator in combination with Mood-1 or Mood-2.

(23) 신문을 보지만 정치뉴스는 안 봐요.
ɕinmun-ɯl po-dziman tsʌŋts^{h}i-ɲyus^{h}ɯ-nɯn an bw-ayo
newspaper-Acc see-Conjunctive political news-Topic not see-Polite (Declarative)
'I read newspapers but I don't read political news.'

(24) 제가 본 그 신문을 주세요.
tse-ga po-n kɯ ɕinmun-ɯl tsu-s^{h}e-yo
I-NOMINATIVE see-ADNOMINAL MODIFIER the newspaper-ACC give-HON-POLITE
'Please give me the newspaper that I have read.'

3.2 Phonological Variation in Spoken Forms

In spoken Korean, stems and exponents typically undergo extensive phonological adjustments. The following outlines key phonological phenomena that are likely to influence lexical processing in various ways.

3.2.1 Syllable-final consonants

Some phonological changes are specific to syllable-final consonants.

Syllable-final plosives are **neutralised** and become **unreleased**: ㅋ /kʰ/, ㄲ /k̤/ are simplified to ㄱ /k˺/ (e.g., 부엌 /puʌk˺/ 'kitchen' ; 밖 /pak˺/ 'outside') ; ㅍ /pʰ/ becomes ㅂ /p˺/ (e.g., 짚 /tsip˺/ 'straw'; 집 /tsip˺/ 'house'); ㅌ /tʰ/ is neutralised to ㄷ /t˺/ (e.g., 밭 /pat˺/ 'field'). Furthermore, dental fricatives and affricates also neutralize to ㄷ [t˺]. Thus, for ㅌ /tʰ/, ㅅ /sʰ/, ㅆ /s̤/, ㅎ /h/, ㅈ /ts/, and ㅊ /tsʰ/ we find final unreleased /t˺/ in words such as 곳 /kot˺/ 'place', 히읗 /hiɯt˺/ 'the name of Hangul consonant 'ㅎ', and 낮 /nat˺/ 'daytime'.

VCC and CVCC syllables undergo final **cluster simplification**. In 값 /kap˺t˺/ 'price' the final dental is deleted /kap˺/; in 삶 /sʰalm/ 'life' it is the first consonant of the cluster that is deleted: /sʰam/. However, if a VCC or CVCC syllable is followed by a vowel-initial syllable, the final consonant is realised as the initial syllable of this following syllable (**liaison**) (e.g., 값이 /kap˺ɕ̤i/ 'price + nominative marker'). Exceptionally, syllable-final ㅎ/h/ does not undergo liaison: 많아 'a lot' is pronounced as /mana/.

Some final consonants of verb stems — ㄹ/l/, ㅅ/t˺/, ㄷ/t˺/, ㅂ/p˺/, and ㅎ/h˺/ — may undergo **elision** when inflected. In some irregular verbs, ㄹ/l/, ㅅ/t˺/ may undergo elision (낫-/nat˺/ 'to recover' → 나아/naa/, 살-/sʰal/ 'to live' → 사는/sʰanɯn/). Stem-final ㄷ/t˺/ changes into ㄹ/l/ (걷-/kʌt˺/ 'to walk' → 걸어요/kʌrʌyo/) and ㅂ/p˺/ elision is followed by 워/wʌ/ (춥/tsʰup˺/ 'to be cold' → 추워/tsʰwʌ/) or 와/wa/ (돕/top˺/ 'to help' → 도와/towa/) for some specific inflectional exponents.

3.2.2 Syllable-initial segments

Syllable-initial consonants can be modified by preceding syllable-final consonants, resulting in consonant reinforcement.

In some consonant clusters, ㅂ /p/, ㄷ /t/, ㅅ /s^h/, ㅈ /ts/, and ㄱ /k/ are **tensed** and pronounced as ㅃ /p͈/, ㄸ /t͈/, ㅆ /s͈/, ㅉ /t͈s/, and ㄲ /k͈/ respectively (e.g., 있다/it˺t͈a/ 'to exist', 식사/ɕik˺s͈a/ 'having a meal', 있자/it˺t͈sa/ 'let's stay', 먹고/mʌk˺k͈o/ 'eat and...').

Although ㅎ /h/ is not realised in syllable-final position, it induces **a**spiration of following lax consonants: ㅂ /p/, ㄷ /t/, ㅈ /ts/, and ㄱ /k/ become ㅍ /p^h/, ㅌ /t^h/, ㅊ /c^h/, and ㅋ /k^h/ respectively (e.g., 좋다 /tsot^{h}a/ 'good!', 놓고 /nok^{h}o/ 'put and...', 좋지 /tsots^{h}i/ 'good, isn't it?').

A syllable-final vowel followed by a syllable-initial vowel may merge into a diphthong (**vowel contraction**). This is found in some verbal conjugations, where the stem final vowel is contracted with the following initial vowel of suffixal inflection. For instance, when the verbs 보- /po-/ 'to see' and 오- /o-/ 'to come' combine with the exponent -아요/-ayo/ (for polite speech level in present tense), the resulting forms are 봐요 /pwayo/ and 와요 /wayo/ respectively.

3.2.3 Assimilation

Assimilation is ubiquitous in Korean. Common types of assimilation include:

Regressive nasalization: before ㄴ/n/, ㅁ/m/, and ㄹ/l, ɾ/, ㅂ/p/, ㄷ/t/, and ㄱ/k/ become the corresponding homorganic nasals ㅁ/m/, ㄴ/n/, and ㅇ/ŋ/ (e.g., 갑니다/kamɲida/ 'to go'; 듣는다/tɯnnɯnda/ 'to hear'; 국물/kuŋmul/ 'soup').

Liquidisation: In consonant clusters with ㄴ/n/ and ㄹ/l/, the nasal changes into /l/ (e.g., 신라/ɕilla/ 'Sin-la, a name of ancient country' , 칼날/k^{h}alla/ 'blade')

Palatalisation: when preceding ㅣ/i/ in the final syllable, ㄷ/t/ and ㅌ/t^h/ become palatalized affricates (e.g., 같이/$kats^hi$/ ‘together’)

Vowel harmony: Korean has three classes of vowels, a ‘positive’ class (also referred to as ‘light’), such as ㅏ/a/ and ㅗ/o/, a ‘negative’ class (also referred to as ‘dark’), such as ㅓ /ʌ/ and ㅜ/u/, and a neutral class with ㅣ/i/. When the two syllables have vowels that both fall in one of the first two classes, then the first vowel assimilates to the second vowel, resulting in forms such as 막아/maga/ ‘to prevent’ and 먹어/mʌgʌ/ ‘to eat’.

3.3 Hangul as a Written Form

The basic unit of Hangul, the Korean writing system, bundles phones into syllables. This syllable-based structure of Hangul, in which the final consonant in VC, CVC, or CVCC syllables is positioned beneath the preceding vowels and consonants (e.g., ㅎ /h/ + ㅏ /a/ + ㄴ /n/ = 한 /han/), presents an unfamiliar visual configuration for non-native speakers who are accustomed to predominantly linear writing systems. As a result, some non-native learners may initially struggle with recognising syllable-level combinations and accurately producing the corresponding sounds. However, due to the relatively limited number of phonemes and symbols (i.e., 14 consonants and 10 vowels) in Hangul, its acquisition typically occurs within a short period, and such perceptual challenges tend to disappear for advanced learners.

Nevertheless, due to the phonological variations mentioned above and the syllable-based structure of Hangul, there are cases where the actual phonological form does not align with the written form. In addition, it is not uncommon for inflected verb forms to appear ambiguous in their written representation. For instance, in 매십니다, /mɛçimnida/ (Hon, present, formal, decl), the stem 매-/mɛ/ (to tie) is visible as such. But in 맵니다 /mɛmnida/ (Plain, present, formal, decl), it incorporates an inflectional exponent. Furthermore, phonological changes, such as the nasalisation of the initial consonant of the formal exponent -ㅂ니다/p̚nida/ into -ㅁ니다 /mɲida/, tend not to be written. As a consequence, written Korean does not provide a straightforward window on spoken Korean, nor on the morphological structure of Korean verbs.

Chapter 4

Studies on L1 Korean Verbal Morphological Processing

Previously, the main theoretical discussions on the processing of morphologically complex words(e.g., inflected words) have centred around whether inflected forms are recognised as whole units, whether they are primarily decomposed or whether both mechanisms are involved in processing. Specifically, a key question, widely explored in research on Indo-European languages, is whether lexical processing can occur on the basis of whole word units (e.g., Butterworth, 1983; Caramazza et al., 1988), or conversely, whether lexical processing is (de)compositional (e.g., Taft and Forster, 1975; Taft, 1981; Marslen-Wilson and Tyler, 1997; Taft, 1994; Pinker, 1991). Dual-route accounts, in which whole-word based processes and decompositional processes operate together, have also been proposed(e.g., Schreuder and Baayen, 1995; Kuperman et al., 2009; Baayen et al., 1997; Marslen-Wilson et al., 1994). In addition, connectionist models have also sought to understand word recognition by approaching it in ways similar to neural networks(e.g., Rumelhart et al., 1986; Seidenberg and McClelland, 1989; Harm and Seidenberg, 2004; Rumelhart and McClelland, 1986).

These theoretical perspectives have also served as foundational reference points in many past studies on Korean word recognition. However, since most of the main theories were based on a limited set of languages—mainly Indo-European languages—they are often considered less applicable to the unique features of Korean verb morphology. In this regard, the following section will examine in detail how previous research has approached the processing of Korean verbs and what major issues have been discussed.

4.1 Understanding Korean Verb Processing

There have been several psycholinguistic studies on the processing of Korean word forms. Korean inflectional morphology is exclusively suffixing. Research on the processing of complex words indicates that there are differences depending on parts of speech, due to the characteristics of Korean lexical structure, which is centered around word units called '어절, Eojeol[6]. Although nouns and verbs (including adjectives) both form complex words through the addition of semantically or functionally significant suffixes, the degree of semantic diversity and morphological complexity differs significantly between nouns and verbs in Korean. Inflected nouns form word units (i.e., 어절, Eojeol) typically by attaching one or two grammatical cases[7] to the noun, whereas verb forms combine a wider variety of exponents that carry more complex and nuanced semantic and grammatical functions, resulting in countless possible combinations in verb forms. Given the unique characteristics of Korean verb forms, this section will review some existing studies that focus on the unique morphological aspects of Korean verbs, divided into several key categories.

[6]' 어절, Eojeol' in Korean is a word unit which is separated by spacing between words in a sentence. It can include grammatical particles (e.g., nominative, accusative markers) attached to nouns and verbal inflectional forms.

[7]These nominal suffixes primarily target major case markers such as the nominative, accusative, and topic markers. Various other case markers are also in use, such as genitive, locative, vocative, plural, quotative, instrumental, and directional suffixes. These nominal markers play a significant role in accurately processing the overall meaning of a sentence in relation to its syntactic structure.

Morphological complexity

First, there are some studies that focused on the morphological complexity that is central to research on Korean verb processing. These studies examine how the constituent elements of Korean verbs, i.e., stems and exponents, are perceived and produced during word processing. In conventional Korean verb morphology, the structure of verb forms has typically been proposed to consist of a stem, followed by a pre-final ending and followed by a final ending. The pre-final ending directly follows the stem and includes exponents for honorifics, past tense and mood. The final ending expresses speech level and illocutionary force in sentence-final verb forms. It also includes clausal connectives (the so-called mood forms) that convey complex nuances, as well as adnominal, nominal, and adverbial forms used in non-final sentence positions or in Banmal (casual speech).

In Korean psycholinguistics, lexical processing of Korean verbs has primarily been studied through morpho-structural approaches. In this type of approach, research on lexical processing has typically focused on analysing words into their morphemic components and examining whether each component is processed individually or in an integrated manner with other elements. In this regard, most studies on Korean lexical processing have focused on investigating how the constituent elements of verb forms within the established structural framework (i.e., namely, the stem, pre-final and final endings) are processed during comprehension and production.

Notably, Hwang et al. (2002, 2003, 2005); Hwang (2007) conducted various experiments addressing the lexical processing of Korean verbs (including both verbs and adjectives), within the conventional framework of Korean verb morphology. In the study of Hwang et al. (2002), three types of Korean verb forms were examined: (1) A stem with a pre-final ending realizing the past tense and a final ending for illocutionary force: 'stem + past tense + final ending' (e.g., 감았다 'I/she/he closed (the eyes)'. (2) A stem followed a passive exponent, followed by an exponent for illocutionary force: 'verb root + passive suffix + final ending' (e.g., 감기다 '(eyes) are closed'). (3) A stem followed by only a final ending (a connective): 'stem + final ending' (e.g., 감고 'I/she/he close (eyes) and . . .). Visual primed lexical decision experiments did not provide evidence for priming effects for the structures 'stem+past tense' and 'stem+passive', which led Hwang et al. (2002) to argue that these combinations are processed as whole units, partially supporting a full-listing, whole-word based processing approach. However, the structure 'stem + final ending' (connective or sentence-final endings) revealed priming effects, which was understood to support decompositional processing theories. These findings suggest that the processing mechanisms involved in word recognition may vary and chunk a visual word into major chunks, but not requiring decompositional processes down to the individual underlying exponents. Hwang et al. (2002) concluded that no single existing theoretical model fully accounted for the processing of Korean verb forms, highlighting the need for a new model specifically tailored to the unique morphological features of Korean verbs.

There have also been several experiments focusing on the production of Korean verb forms. Hwang et al. (2003, 2005) registered brain activation using fMRI during the production of Korean verbs. Hwang et al. (2003) investigated whether different cortical areas would be activated for two types of verbal inflectional endings, pre-final endings and final endings, as well as two derivational suffixes (causative and passive suffixes) during speech production. The observations indicated that the areas of brain activation for final verb endings differed from those for pre-final endings, passive, and causative suffixes. This study particularly suggested that pre-final endings might have a mental lexical representation pattern more akin to causative or passive derivational suffixes rather than final inflectional endings, in line with previous research findings (Hwang et al., 2002). Furthermore, in a subsequent study, Hwang et al. (2005) examined brain response priming effects during the production of final endings, specifically focusing on the imperative and interrogative forms. The results showed that both forms activated Broca's area in the frontal lobe of the left hemisphere and the inferior parietal lobule in the right hemisphere, and revealed a facilitative priming effect when the endings served the same function (i.e., the same illocutionary force).

Hwang (2007) also investigated whether verbs and adjectives exhibit different lexical processing patterns. The experimental results from a lexical decision task showed no significant processing differences between verbs and adjectives. However, the results suggested that inflected verbs (and adjectives) are segmented into a 'stem + pre-final ending' unit and a 'final ending' unit. In addition to this, a syllable neighborhood size effect was observed at the 'stem + pre-final ending' level. Words sharing the same syllables for stem and pre-final endings were responded to faster in a visual lexical decision task than words sharing only the syllable of the stem.

Of interest is also a study investigating the morphological distinction between pre-final and final endings and their cognitive reality. Song (2009) carried out a lexical decision experiment aimed to clarify whether the pre-final endings -시-(honorifics), -었-(past tense), -겠-(conjecture/volition), -더-(retrospect) function as independent units in reading. The hypothesis he formulated was that the response time to stimuli with stand-alone pre-final endings (i.e., the particular exponent by itself presented in isolation) would be shorter than the response time to inflected forms containing those pre-final endings. The experiment showed that indeed these stand-alone endings were responded to faster than words with these endings. The statistical analysis of the experimental results showed that stand-alone prefinal endings elicited faster reaction times than conjugated forms, leading to the conclusion that they may have independent psychological representations, whereas such representations may not be present in conjugated forms. However, this interpretation does not clearly address the potential impact of word length differences between single endings (typically one syllable) and full inflected forms. For the above interpretation, exceptions were 먹었다 (ate) and 먹겠다 (will eat)(i.e., faster RTs than the single form of -었- or -겠-), which the study attributed to the high frequency of these particular combinations.

Frequency effects

Several studies have investigated frequency as a key factor in word recognition (e.g., Kim and Nam, 2018; Lee et al., 2019b; Kwon and Lee, 2014; Lee et al., 2018b), building on research from Indo-European languages that examined the effects of usage frequency (see, e.g, Scarborough et al., 1977; Bybee and Hopper, 2001; Baayen et al., 2007).

In a lexical decision task, Kim and Nam (2018) compared noun and verb forms (including adjectives) and found that different frequency-related variables played significant roles in their recognition. For nouns, whole-word (Eojeol) frequency, lexeme frequency, and first-syllable frequency emerged as significant predictors. Across all frequency measures for noun forms, higher frequency was associated with faster reaction times (mean RT: 512 ms). The magnitude of the facilitation effect was largest for whole-word frequency ($\hat{\beta}$ = -0.516), followed by root frequency ($\hat{\beta}$ = -0.475), and first-syllable frequency ($\hat{\beta}$ = -0.277). In contrast, for verb forms (mean RT: 522ms), whole-word frequency ($\hat{\beta}$ = -0.439) and the number of subjective meanings ($\hat{\beta}$ = -0.426) were significant predictors. This study accounted for the facilitative effect of the number of subjective meanings on reaction time through the feedback mechanism proposed by Hino and Lupker (1996). That is, the greater the number of subjective meanings associated with a word, the stronger the feedback from semantic to orthographic and phonological levels, which may lead to faster lexical decisions compared to words with fewer meanings.

In a primed lexical decision experiment, different results were observed depending on stimulus onset asynchrony (SOA) and stem frequency (Lee et al., 2018b). The study targeted the past tense pre-final ending (-었-/-았-), using priming conditions that included identical endings, high- and low-frequency final endings, and unrelated endings. Stimuli were presented with two different stimulus onset asynchronies (SOAs): 43 ms and 230 ms. The results showed that priming facilitated response times in the following conditions: with high-frequency stems at SOA 43 ms, and with both high- and low-frequency stems at SOA 230 ms. These priming effects were interpreted as evidence in support of the decomposition hypothesis, suggesting that morphological exponents can be recognized as separate units in these contexts. However, in the SOA 43 ms condition with low-frequency stems, no priming effect was observed. This result aligns more closely with the full-listing hypothesis and was interpreted as suggesting that decomposition of the stem and ending may be difficult under such a short SOA condition.

Studies on other languages (e.g., Spanish, German and French) typically focused on frequency effects for the first syllable of a word (e.g., Carreiras et al., 1993; Perea and Carreiras, 1998; Hutzler et al., 2004; Conrad et al., 2008). These studies have posited that competition among candidates within the mental lexicon intensifies as initial syllable frequency increases, resulting in slower lexical processing response times and decreased accuracy.

For Korean, inhibitory effects of initial syllable frequency have been reported for uninflected nouns (Kwon, 2012). However, some studies have also reported a facilitative effect of initial syllable frequency in the recognition of inflected nouns (e.g., Kim and Nam, 2018; Kwon et al., 2023; Bae and Yi, 2010) and also verbs (e.g., Lee et al., 2019b, 2023). This syllable effect has been attributed to the convergence of spoken syllables and written syllables of the Hangul orthography: having consistent simultaneous visual and phonological information about syllables would then facilitate word recognition (Bae and Yi, 2010; Lim et al., 2022).

Furthermore, in a study by Lee et al. (2023), a facilitative effect of first-syllable frequency was observed in three-syllable inflected verb words, but not in four-syllable words. Apparently, the complex combinations of stems and endings and the varied and extensive phonological changes affecting stems (compare, e.g., 돌다 and 도는데, or 무겁다 and 무거워), interacts with the number of syllables. Lee et al. (2023) point out that these kind of stem changes and re-syllabification is much more widespread among verb forms than among nouns, and render verbal processing substantially more complex than nominal processing.

Multiple factors

Several studies have discussed various factors that may play a role in specifically visual word recognition. In particular, orthographic and visual properties, such as the length of words, the number of strokes, the number of syllables, and the position of given syllables in the word, have been investigated (e.g., Lee et al., 2019b). Of particular interest for Korean is the syllable, which in the Korean Hangul writing system is represented by a separate orthographic form (henceforth character) that itself is structured internally to represent the consonants and vowels in the syllable.

Hangul has been argued to be a shallow orthography akin to the orthography of Spanish, with highly systematic correspondences between spelling and sound. However, although Hangul provides a much better guide to pronunciation compared to, e.g., the English spelling, there are non-negligible numbers of cases where the actual realization of sounds diverges from what the Hangul characters would lead one to expect. For instance, assimilation between consonants and linking sounds lead to discrepancies between orthographic form and pronunciation (e.g., 먹는다 mʌk̚nɯnda > [멍는다 mʌŋnɯnda] or 국물 kuk̚mul > [궁물 kuŋmul]).

Yi et al. (2005) and Bae and Yi (2010) made use of the primed lexical decision task to assess whether such divergences between writing and sound have adverse consequences for reading. They report that a priming condition where spelling and phonology are congruent (e.g., 숙면 /s^{h}uŋmyʌn/ as prime and 숙녀 /s^{h}uŋɲyʌ/ as target) showed the fastest reaction times. Intermediate reaction times were observed when spelling matched but not sound (e.g., 숙소/s^{h}uk̚s̤o/ as prime and 숙녀/s^{h}uŋɲyʌ/ as target). The longest reaction times were found for the priming condition with only matching phonology (숭배/s^{h}uŋbɛ/ as prime, 숙녀/s^{h}uŋɲyʌ/ as target). These results suggest that orthographic information exerts a greater influence than phonological information in Korean word recognition.

Using a primed lexical decision task, Ahn et al. (2011a) studied the processing of the connective form -고 (non sentence-final ending) and the honorific form -시- (a pre-final ending). A wide range of priming conditions was considered: identical (쏘다 to shoot - 쏘다), morphological (final/pre-final endings) (쏘고 shoot and - 쏘다 / 쏘시다 (subject honorifics) shoot - 쏘다), semantic (권총 gun - 쏘다), phonological (쏘바 (nonword) - 쏘다), and unrelated (계단 stairs - 쏘다). Only morphologically related forms (i.e., final ending priming condition such as 쏘고 - 쏘다 or pre-final ending priming conditions such as 쏘시다 - 쏘다) showed a priming effect, with shorter reaction times for words which included the honorifics exponent '-시-'. Ahn et al. (2011a) proposed that their findings support models positing obligatory decomposition. However, they also noted that similar studies have shown conflicting results and suggested that instead of applying existing dichotomous theories such as full-list or decomposition, a more nuanced theoretical framework that accounts for the specific characteristics of Korean morphology may be required.

Lee et al. (2019b) analyzed various lexical factors involved in the recognition of verbal phrases by examining their relationship with ERP responses and reaction times. This was done through a multiple-meaning decision task, in which participants judged whether an inflected verb form (Eojeol)[8] had a single meaning or more than one meaning. This study examined a range of variables, including whole-word (i.e., Eojeol) frequency, stroke count, the number of consonants and vowels, syllable count, morpheme count, first-syllable-sharing frequency, lexeme frequency[9], as well as the numbers of both dictionary-defined meanings and subjective meanings (i.e., the number of meanings recalled by individual participants)[10]. The results indicated that lexeme and whole-word frequencies

[8] The information regarding the words (Eojoels) used as stimuli in the study is unclear, making interpretation of the results somewhat ambiguous. According to the paper, "A total of 296 verb forms were used as stimuli; of these, 232 were considered to have a single meaning, while 64 were considered to have multiple meanings." However, the criteria used to classify these verb forms as having either a single or multiple meanings are not clearly provided, nor are any examples given. This lack of detail hinders a clear understanding and interpretation of the findings.

[9] While this study focuses on verb inflections, it does not clearly distinguish between "stem" and "root", referring instead to "root frequency" throughout. In this dissertation, we use the term 'lexeme'.

[10] As stated in the original paper(p. 296): "The numbers of subjective meanings were collected by providing participants with a questionnaire immediately following the behavioral experiment, asking them to indicate the number of meanings associated with each word's base form and to describe each meaning."

and the number of dictionary and subjective meanings significantly affected response time. Specifically, higher whole-word frequency and a greater number of dictionary meanings were associated with faster response times. In contrast, increased lexeme frequency and a higher number of subjective meanings led to slower response times. Interestingly, the finding that higher root frequency was associated with slower response times contrasts with previous research. ERP analysis further revealed that the mean amplitude of the P300 component was significantly correlated with first-syllable-sharing frequency and the number of dictionary meanings. Additionally, both the number of dictionary meanings and whole-word frequency had significant effects on the N400 component.

Regularity

Another typical issue in word recognition studies is whether a word's inflected forms are regular or irregular. According to the dual mechanism account (Pinker, 1991; Pinker and Ullman, 2002; Prasada and Pinker, 1993), all regular forms are processed by rule, and all irregular forms are stored in memory. Connectionist theories using early versions of artificial neural networks now known as deep learning networks have argued that regular and irregular forms are not processed by separate systems, but by the same neural network, the 'single mechanism' account (Rumelhart and McClelland, 1986; MacWhinney and Leinbach, 1991; Plunkett and Marchman, 1993).

Irregularity in the verb system of Korean is limited. In some inflected verb forms, the final syllable of the stem changes when followed by certain endings, leading to what is classified as irregular morphological changes. For instance, the stem of the verb 'to thank', 고맙-/komap˺-/ , has an unpredictable inflected form 고마워/komawʌ/. However, these irregular forms can often be categorized into a few patterned types, meaning that although grammatically they are considered irregular due to stem modifications, many words are in many respects quite regular.

In the case of Korean, several studies have reported findings supporting both the single mechanism model (e.g., Kim et al., 2000; Park et al., 2012) and the dual mechanism model (e.g., Yim et al., 2003). Many studies have found that Korean verb processing often occurs at the 'whole-word' level (i.e., at the level of the Eojeol), with no significant differences between regular and irregular forms (e.g., Kim et al., 2000; Park et al., 2012; Lee et al., 2018a; Jo, 2021). Lee et al. (2018a) suggests that for native speakers, sensitivity to irregularities may arise more from differences between basic orthography and actual phonological phenomena rather than from violations of strict grammatical rules for decomposition.

Chapter 5

Experiment 1 : Native Speakers

This chapter investigates how native Korean speakers process Korean verbs through a visual lexical decision task. As discussed in the previous review of related studies, the typical approach in verbal processing studies was to divide Korean verb structure into components such as stem, pre-final ending, and final ending, and then explore how one or two of these particular exponents (e.g., past tense or honorifics forms) are perceived whether they are processed separately or as integrated units with other exponents as posited by conventional theories (e.g., Hwang et al., 2002; Ahn et al., 2011a; Lee et al., 2018b, 2020).

Departing from this trend, the present study aims to take a broader and more integrated approach to Korean verb morphology. In this study, we examined a more diverse and integrated range of Korean verb forms by substantially increasing the number of verb stimuli used in the experiment to 4,000 words, including 1,000 carefully designed nonword verb forms. This approach was intended to yield more comprehensive and statistically reliable results. The following section provides a detailed description of the experiment and the key predictor variables that are the focus of this study.

5.1 Participants

A total of 40 native Korean speakers(17 females and 23 males) participated in the experiment. Participants ranged in age from 18 to 54, with similar numbers in each age group. Participants were born and raised in South Korea and received formal education beyond high school in South Korea. All participants had normal or corrected-to-normal eyesight. Participants were reimbursed for their participation with €60. Participants were recruited through several online short-term job postings. The recruitment posts were written in the formal register.

5.2 Materials

5.2.1 Words

A total of 4000 Korean words were selected for presentation in the experiment. These words were inflected variants of 100 verbs. Each of these verbs was represented by 40 inflectional variants, encompassing 24 sentence-final verb endings and 16 forms with an exponent for mood. Tables 5.1 and 5.2 provide descriptions and examples of these endings. The mood endings selected vary substantially in frequency of use, as indicated in Table 5.2.

The primary predictors addressed in the data analysis of the current study are as follows. First, concerning the effect of frequency of use, we examined the frequency of the lexeme (`lexeme frequency`), the frequency of the ending (`ending frequency`), and the frequency of the inflected variant itself (`whole-Word frequency`). The three frequency counts are based on the NIKL Part-of-Speech Tagged Corpus(version 1.0)[11]. This corpus comprises 2 million words (어절, Eojeol) of written language, and 1 million

[11]국립국어원 형태 분석 말뭉치, NIKL Part-of-Speech Tagged Corpus(version 1.0), National Institute of Korean Language, 2020; `https://kli.korean.go.kr/corpus/main/requestMain.do?lang=ko`.

words of spoken language. Our frequency counts are based on the subcorpus of 1 million-word spoken words. The verbs in the experiment had whole-word form frequencies ranging from 0 to 7127, lexeme frequencies ranging from 1 to 13814, and ranging from 0 to 26168 for ending frequencies. To avoid the adverse effects of outliers on statistical evaluation with regression modeling, all frequency counts were log-transformed.

Predictors for aspects of words' forms are the (`number of syllables`(1 to 8 Hangul characters), the type of `vowel harmony` (light or **dark**, using treatment coding; here and in what follows, reference levels are listed with bold font), and whether the final segment of the stem is a vowel or a consonant: `stem-Final segment` (vowel, **consonant**).

We considered five morphological variables: `honorifics` (**honorific**, non), `tense` (**non**, past), `speech level` (**formal**, polite, intimate, plain), `illocutionary force` (**declarative**, inquisitive, imperative[12], propositive), and `mood`(**mood**, non). A subset of forms contained one out of 16 different exponents for `Mood` (see Table 5.2).

To facilitate statistical analysis, as a first step, we also created a factor for the combination of `speech level` and `illocutionary force`, to which we refer as (`SLIF`), with as levels **formal declarative(for + dec)**, `formal inquisitive(for + inq)`, `honorific declarative/imperative(hon + dec/imp)`, `polite unmarked(pol + unmarked)`, `intimate unmarked(int + unmarked)`,`plain declarative(plain + dec)`, `plain inquisitive(plain + inq)`, and `unmarked`.

An analysis based on `Past Tense`, `Subject Honorifics`, and `SLIF` stays close to a de-compositional, agglutinative analysis of Korean verb morphology. However, as mentioned in Chapter 4, the ending of a Korean verb is widely regarded as of central importance to the language. We therefore also created a random-effect factor, `Ending`, the levels of which are listed in Tables 5.1 and 5.2. This factor brings together combinations of the five above-mentioned morphological variables (past tense, honorifics,

[12]Imperative and propositive forms were applied to nonwords due to the limited number of existing word forms that could accommodate all the different exponents.

speech level, illocutionary force, and mood). Our dataset includes 40 different endings, but we note here that in corpora, the number of different endings is much larger, in the order of 4000 to 5000 forms (Kang and Kim, 2009; Choi, 2012)[13]. Most of these forms arise due to the large number of different mood forms included in clausal connectives and sentence-final endings. If the verb ending is indeed the key to lexical processing in Korean, we expect a GAMM with random intercepts for verb ending to outperform a GAMM with separate predictors for tense, honorifics, mood, and illocutionary force & speech level (SLIF).

code	form(Hangul)	honorifics	tense	mood	speech level	illoc. force
int-1	아/어	unmarked	unmarked	unmarked	intimate	unmarked
for-dec-2	(ㅂ/습) 니다	unmarked	unmarked	unmarked	formal	declarative
pol-3	(아/어) 요	unmarked	unmarked	unmarked	polite	unmarked
for-past-dec-4	(ㅆ/았/었) 습니다	unmarked	past	unmarked	formal	declarative
int-past-5	(ㅆ/았/었) 어	unmarked	past	unmarked	intimate	unmarked
pol-past-7	(ㅆ/았/었) 어요	unmarked	past	unmarked	polite	unmarked
pln-dec-8	(ㄴ/는) 다	unmarked	unmarked	unmarked	plain	declarative
pln-past-dec-9	(ㅆ/았/었) 다	unmarked	past	unmarked	plain	declarative
for-inq-10	(ㅂ/습) 니까	unmarked	unmarked	unmarked	formal	inquisitive
pol-hon-11	(으) 세요	Hon.	unmarked	unmarked	polite	unmarked
for-hon-inq-13	(으) 십니까	Hon.	unmarked	unmarked	formal	inquisitive
pol-hon-past-14	(으) 셨어요	Hon.	past	unmarked	polite	unmarked
pln-inq-15	니	unmarked	unmarked	unmarked	plain	inquisitive
int-hon-16	(으) 셔	Hon.	unmarked	unmarked	intimate	unmarked
for-hon-past-dec-17	(으) 셨습니다	Hon.	past	unmarked	formal	declarative
for-past-inq-18	(ㅆ/았/었) 습니까	unmarked	past	unmarked	formal	inquisitive
for-hon-dec-19	(으) 십니다	Hon.	unmarked	unmarked	formal	declarative
for-hon-past-inq-20	(으) 셨습니까	Hon.	past	unmarked	formal	inquisitive
int-hon-past-22	(으) 셨어	Hon.	past	unmarked	intimate	unmarked
pln-hon-dec-23	(으) 신다	Hon.	unmarked	unmarked	plain	declarative
pln-hon-past-dec-24	(으) 셨다	Hon.	past	unmarked	plain	declarative
pln-past-inq-25	(ㅆ/았/었) 니	unmarked	past	unmarked	plain	inquisitive
pln-hon-inq-26	(으) 시니	Hon.	unmarked	unmarked	plain	inquisitive
pln-hon-past-inq-28	(으) 셨니	Hon.	past	unmarked	plain	inquisitive

Table 5.1: The verb endings represented in the words that do not contain an exponent for mood. (Smaller numbers in the code indicate more frequent forms.)

[13] In the investigation by Kang and Kim(2009), over 4,000 individual exponents were found in a 1.5 million written corpus, including 2,077 sentence-final endings, 1,797 connective endings, approximately 170 nominalization and attributive endings. Choi(2012)'s research estimated that the combinations of these endings amounted to 5,682 distinct forms (i.e., endings) in a corpus analysis of 10 million written texts.

code	form(Hangul)	mood		speech level	illoc. force
		gloss	frequency		
fm1	고	‘and’	HF	unmarked	unmarked
fm2	게	adverbial	HF	unmarked	unmarked
fm3	(아/어) 서	‘so(because)’	HF	unmarked	unmarked
fm4	(으) 면	‘if’	HF	unmarked	unmarked
fm5	(ㄴ/은/는) 데	‘but/so/and’(background)	HF	unmarked	unmarked
fm6	(으) 니까	‘because’	HF	unmarked	unmarked
mm10	지만	‘but’	MF	unmarked	unmarked
mm12	도록	‘so that’	MF	unmarked	unmarked
mm15yo	거든요	‘as you may know...’	MF	polite	unmarked
lm27	(ㄹ/을) 라	‘could happen(worry)’	LF	unmarked	unmarked
lm32	(으) 리라	‘it(I) would’(assumption/determine)	LF	unmarked	unmarked
lm40	던걸	‘(he/she/they) was doing’(witness)	LF	unmarked	unmarked
lm61yo	다마다요	‘of course(sure)’(promise/assurance)	LF	polite	unmarked
rm51	(ㄹ/을) 망정	‘even if’	ZF	unmarked	unmarked
rm58yo	든지요	‘or(does not matter)’	ZF	polite	unmarked
rm59	거들랑	‘you know...’, ‘if’	ZF	unmarked	unmarked

Table 5.2: The verb endings represented in the words that contain an exponent for mood. Mood forms are grouped by their frequency in the corpus (HF: high-frequency mood form, MF: medium-frequency mood form, LF: low-frequency mood form, ZF: zero-frequency mood form). Honorifics and past tense are unmarked in these forms. The categorised frequency indicators (HF to ZF) are provided as rough reference points for the relative frequency of specific forms. Since frequency can vary depending on the size of the corpus and whether it consists of spoken or written data, these indicators are intended to serve only as general references. For instance, based on the corpus used, the frequencies of some representative forms from the highest category downward are as follows: HF ‘-고’(26,168), MF ‘-지만’(926), LF ‘-(ㄹ/을) 라’ (63), and ZF ‘-(ㄹ/을) 망정’(0).

Experimental lists

The 4000 word forms were divided into four groups of 1000 each, ensuring an even distribution of lexemes and endings as well as their frequencies. Each of these four groups was combined with the same set of 1000 nonwords introduced in the following section. The order of the words and nonwords in each of the resulting four lists (2000 items each for 1000 words and 1000 nonwords) was pseudo-randomized.

5.2.2 Nonwords

This study, in particular, used nonword forms to investigate how phonological, morphological, and semantic effects emerge in lexical processing. Given the diverse productivity of Korean verbal morphology, this study also intended to explore a broader range of verb endings by using nonwords, beyond those examined with normal word forms earlier.

We created 1000 nonwords, which were presented together with 1000 words to a given participant. Thus, all participants responded to the same set of 1000 nonwords, but responded to 1/4 of the total set (4000 words) of real words. Table 5.3 provides an overview of how these nonwords were constructed. Each nonword condition presents specific expected patterns based on phonological, morphological, or semantic features in the relationship between lexemes and endings.

The first nonword condition, which includes phonologically similar forms (`nw-phoL/E/LE/-y`)[14] to the original base words, may reveal how phonological and orthographic representations interact in word recognition. These nonwords were created by incorrectly spelling the lexeme(`phoL`), the ending(`phoE`), or both(`phoLE`) the lexeme and the ending.

The second nonword condition (see the second subsection of Table 5.3) is designed to examine sensitivity to the morphological structure of endings by testing cases where allomorphs are swapped (`nw-alloSw/-y`) or the syllables within an ending are rearranged (`nw-sswpE/-y`).

Next, another nonword condition involves examining the interaction between the morphological and semantic composition of endings, focusing on mood-related conditions (`nw-fm/infm`) and variations in the combination of past tense, mood and polite speech endings (`nw-Tmy`). This manipulation was implemented for both frequent (`nw-fm`) and infrequent (`nw-infm`) mood forms. As for `Tmy` condition, in particular, the combination of past tense and mood forms aims to probe morphological and semantic composition.

[14] The codes in parentheses refer to those given in Table 5.3

The final subset in Table 5.3 makes use of completely novel forms in stems (`nw-fagL/Ly`) or in endings (`nw-fagE/Ey`). Similarly, a condition in which only a single syllable is replaced with a novel syllable (`nw-sagL/E/-y`) was added to compare with the above cases where the entire stem or ending takes a novel form (`nw-fag`). These changes are expected to facilitate correct rejections and fast RTs compared to other nonword conditions.

A further observation was implemented for these kinds of nonwords, namely the presence or absence of the exponent -요 /yo/ in the final syllable. In Korean verb endings, -요 always appears at the very end and is a highly frequent and salient form due to its widespread use in various contexts. -요 is not only a fundamental marker of polite speech but is also frequently attached to verb forms that typically end in various connective endings or mood forms, often referred to as '반말체' (Banmal, casual speech). This usage is common in conversational Korean, where -요 is added to these forms to create a polite sentence-final expression from a more informal Banmal form. Given the familiarity and widespread use of -요, nonwords with this exponent are expected to be more word-like, and may therefore either facilitate or inhibit response times.

nonword condition	example	word model	code
phonologically similar lexeme	이러습니다	이렇습니다	nw-phoL
	만니	맞니	
	몰으세요	모르세요	nw-phoLy
phonologically similar ending	이렇씁니다	이렇습니다	nw-phoE
	싫허요	싫어요	nw-phoEy
phonologically similar form	이러씁니다	이렇습니다	nw-phoLE
	맛니이	맞니	
	차즈세요	찾으세요	nw-phoLEy
reversed allomorph	보으십시오	보십시오	nw-alloSw
	높면요	높으면요	nw-alloSwy
swapped order of ending syllables	맛있니습까	맛있습니까	nw-sswpE
	같록도요	같도록요	nw-sswpEy
vowel/consonant change in frequent mood ending	많으느까	많으니까	nw-fm
	힘들저만요	힘들지만요	nw-fmy
vowel/consonant change in infrequent mood ending	보거달랑	보거들랑	nw-infm
	물어보눈지요	물어보는지요	nw-infmy
past tense + vowel/consonant change in mood ending	흘렀을러치면	흘렀을라치면	nw-Tm
	좋아했으느까요	좋아했으니까요	nw-Tmy
novel form in lexeme	석족으셨습니다	괜찮으셨습니다	nw-fagL
	랄아서요	높아서요	nw-fagLy
novel form in ending	좋아추궤	좋아해서	nw-fagE
	맛있호두보	맛있습니까	
novel syllable in lexeme	생곡해라	생각해라	nw-sagL
	리으려고	잡으려고	
	껄려고요	끌려고요	nw-sagLy
novel syllable in ending	괜찮았혀	괜찮았어	nw-sagE
	밝튼지요	밝든지요	nw-sagEy

Table 5.3: Overview of the different kinds of nonwords. A ‘y’ at the end of a code represents the ending ‘요 /yo/’ of the polite speech level.

5.3 Procedure

The 40 Korean native-speaker participants were divided randomly into four groups of 10 subjects each. Each group of 10 subjects was paired with one of the four lists of 2000 stimuli. Each list of 2000 lexical stimuli was divided into six sublists, each consisting of 333-334 forms. These sublists were administered across six experimental sessions, in order to reduce the risk of fatigue and loss of concentration.

The experiment was conducted through the Jatos online platform (Version 3.8.4)[15]. Participants were first familiarized with the overall content and procedure of the experiment. After having signed an informed consent form, they were given a code word and a link to the experiment. At this point, participants were asked to answer several questionnaires, eliciting information about their language background, age, and education level. Each of the six sessions (of 333–334 stimuli) was divided into two blocks, with optional short breaks between blocks. The first session was preceded by a practice session with 20 forms.

A word or nonword was presented for 5000 ms in the center of the screen, between two black parallel horizontal line segments extending along three characters. Participants were requested to press M for words, and Z for nonwords(but reversed keys for left-handed participants). Following a correct response, the line segments briefly changed color from black to green. Following an incorrect response, the line segments briefly turned red. After 5 seconds, the next word automatically appeared. If a participant failed to make a decision within 5 seconds, the response was treated as missing. Participants were encouraged to take a break of at least 30 minutes at the end of a session, and to complete the experiment (6 sessions) within a week.

[15] https://www.jatos.org/

5.4 Results

5.4.1 Statistical analysis

We analysed the experimental data with the generalised additive mixed model (GAMM Wood, 2017b). The generalized additive model is an extension of the linear regression model that makes it possible to predict non-linear effects of one or more predictors on the response variable — in our data, response latency and response accuracy. For modeling the non-linear effect of a numeric predictor, GAMMs make use of smoothing splines. The spline smooths used in this dissertation are thin plate regression splines, which approximate a nonlinear function of a predictor x by means of a weighted sum of k basis functions $f_i(x)$:

$$y = \sum_{i=1}^{k} w_i f_i(x). \tag{5.1}$$

The first two basis functions are straight lines (a horizontal line for the intercept, and a slanted line for linear trends). The user has to specify the number of basis functions k that are to be used. Basis functions $i = 3, 4, \ldots k$ are increasingly wiggly. Thus, choosing a higher value for k makes it possible to approximate more wiggly non-linear effects. GAMMs balance staying faithful to the data (by minimising the mean squared error) and keeping the model simple by means of a penalisation parameter that is estimated from the data. Penalisation ensures that the weights of unnecessary basis functions are shrunk, or even sent to zero. The effective degrees of freedom (edf) in the summary tables reported below are the sum of the proportions of the penalised weights to the unpenalized weights. Larger edfs indicate more wiggliness, whereas edfs close to 1 indicate linearity. Interactions of a numerical predictor and a fixed-effect factor are modeled with by-smooths, which estimate a separate smooth for each level of the fixed-effect factor.

Interactions of a numerical predictor and a random-effect factor are modeled with factor smooths, the non-linear equivalent of what in the linear mixed model would be the combination of random intercepts and random slopes. Factor smooths use a penalisation algorithm that can also take the weight of the slanted linear basis function to zero. As a consequence, if there is no effect of the numeric predictor, the factor smooth will reduce to random intercepts. For factorial predictors, treatment dummy coding was used.

To model interactions of numeric predictors, tensor product smooths, abbreviated as **te** in model summaries, for mathematical details, see Wood (2017b), are used. Standard tensor product smooth incorporates main effects and interactions jointly. In order to tease apart main effects from the interaction itself, one has to specify the main effects with standard spline smooths, and the interaction with a **ti** tensor product smooth. Introductions to GAMMs are provided by Wieling (2018), Chuang et al. (2021), and Baayen and Linke (2020). Detailed discussion of factor smooths is provided by Baayen et al. (2022)

5.4.2 Words

Data points with response latencies below 310 ms or above 2800 ms were excluded (196 datapoints, 0.49% of the total number of datapoints). Response latencies were transformed into scaled negative response rates (-1000/RT) in order to obtain a response variable that approximates normality (and hence can be analysed with a Gaussian model) and that is positively correlated with RT. Trial (the rank of an item in a subject's experimental list) was standardised. Frequency counts were logarithmically transformed.

We fitted a generalized additive mixed model (GAMM) to the data, using the bam function from the **mgcv** package (Wood, 2017b) for R (R Core Team, 2024), predicting reaction times from the inflectional features honorifics, tense, mood, and `SLIF` (which combines speech level, and illocutionary force), from vowel harmony[16], number of syllables, the stem-final segment, the number of exponents, as well as lexeme frequency, whole-word frequency, and ending frequency. The effects of the three frequency measures

[16] In Korean verb conjugation, vowel harmony is typically applied, where the final vowel of the stem influences the vowel of the following ending through assimilation. In this study, the final vowel of the stem was broadly categorized into positive (light) and negative (dark) vowels, and the study investigated whether this distinction has any effect on word processing.

were modeled with thin plate regression splines with 5 basis functions. The model included by-subject factor smooths for trial for individual variation in response times over the course of the experiment, as well as a tensor product smooth for interaction of trial by ending frequency. Autocorrelation in the residual errors was largely removed by including an AR(1) process in the errors with the proportionality parameter ρ set to 0.1. Model criticism revealed a small number of small-valued outliers. Removal of 40 words with a reaction time less than 350 ms (0.1 % of the data) eliminated these outliers.

Table 5.4 presents the summary statistics for the partial effects of parametric (upper part) and smooth (lower part) terms, together with the increase in AIC when a predictor is removed from the model specification, which we use as a measure of variable importance. Past tense was not a significant predictor, and was therefore removed from the model.

Table 5.4: Summary of the partial effects in a GAMM fitted to inverse-transformed (-1000/RT) reaction times for the participants with **Korean as L1** (AIC : 22290). The reference level of SLIF is *formal declarative*.

A. parametric coefficients	Estimate	Std. Error	t-value	p-value	Δ(AIC)
(Intercept)	-1.6725	0.0415	-40.3058	< 0.0001	
honorifics=non	-0.0361	0.0069	-5.2557	< 0.0001	24.6
mood=non	0.1345	0.0123	10.8939	< 0.0001	118.1
SLIF = for+inq	0.0178	0.0074	2.4071	0.0160	422.6
SLIF = hon+dec/imp	-0.0466	0.0140	-3.3373	0.0008	
SLIF = int+unmarked	-0.0495	0.0090	-5.5001	< 0.0001	
SLIF = plain+dec	-0.1108	0.0085	-13.0705	< 0.0001	
SLIF = plain+inq	-0.0905	0.0090	-10.0624	< 0.0001	
SLIF = pol+unmarked	-0.0430	0.0087	-4.9146	< 0.0001	
SLIF = unmarked	0.0798	0.0126	6.3279	< 0.0001	
vowel harmony=light	-0.0439	0.0033	-13.4929	< 0.0001	183.0
number of syllables	0.0386	0.0023	16.5324	< 0.0001	274.0
stem-final segment=vowel	-0.0172	0.0033	-5.2593	< 0.0001	26.8
number of exponents	-0.0235	0.0046	-5.1209	< 0.0001	24.1
B. smooth terms	**edf**	**Ref.df**	**F-value**	**p-value**	**Δ(AIC)**
s(lexeme frequency)	3.9489	3.9983	103.3772	< 0.0001	405.3
s(whole-word frequency)	2.5303	2.9936	35.6300	< 0.0001	104.0
s(ending frequency)	3.9392	3.9975	254.8100	< 0.0001	993.0
s(trial)	7.6760	8.1621	11.6762	< 0.0001	16.9
ti(trial, ending frequency)	7.2516	9.5434	3.5233	0.0002	24.3
s(trial, subject)	279.7344	359.0000	47.2972	< 0.0001	13864.6

Looking at Table 5.4, non-honorific forms were responded to more quickly than honorific forms. Mood forms were responded to more quickly than non-mood forms. The factor SLIF (the combination of speech level and illocutionary force) emerged as the most influential predictor (AIC increase: 422.6) in the parametric part of the model. The shrinkage estimates for the levels of SLIF (obtained by refitting the model with SLIF as random-effect factor) are shown in Figure 5.1.

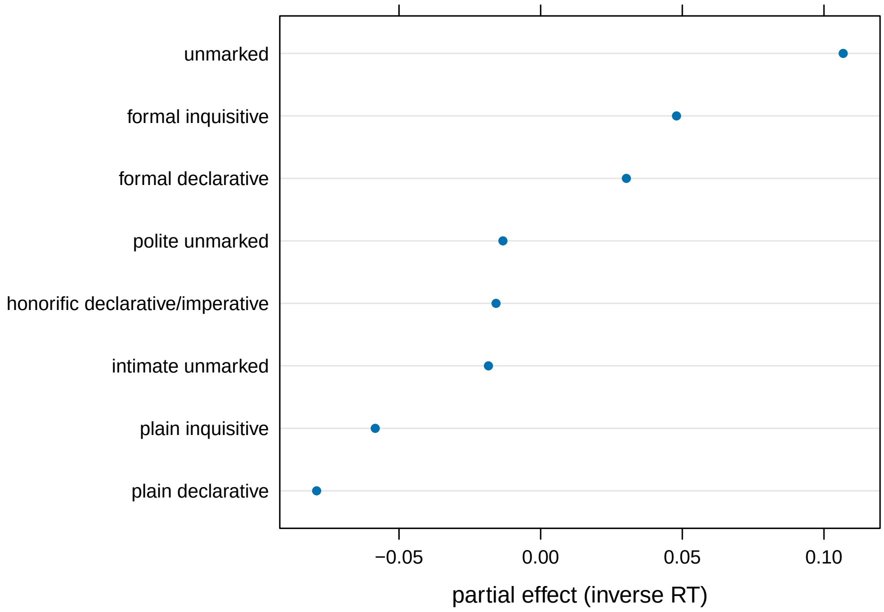


Figure 5.1: Shrinkage estimates for the different combinations of `SLIF` (speech level and illocutionary force). Illocutionary force is unmarked in `polite unmarked` and `intimate unmarked`.

In Figure 5.1, words with a plain speech level and declarative illocutionary force elicited the shortest response latencies, and words that are unmarked for speech level and illocutionary force(i.e., `unmarked`) elicited the longest latencies. As these unmarked forms typically have an exponent for mood, it is conceivable that the difficulty of these words originates from the use of relatively unfamiliar mood forms. `polite unmarked` and `intimate unmarked` forms, which express illocutionary force through intonation, exhibited similar reaction times. `honorific declarative/imperative`[17] characterized by the salient ending '-세요'/-seyo/, also showed comparable reaction times to these forms. The relatively long reaction times for words with `formal inquisitive` and `formal declarative` may be due to the greater length of these endings.

Returning to Table 5.4, words with `light vowel harmony` elicited shorter response latencies than words with `dark vowel harmony`. Response times increased

[17]Strictly speaking, this form is an honorific polite speech form, but due to the high frequency and prominence of its ending, it was categorized separately to observe any specific effects.

linearly with the `number of syllables` and decreased linearly with the `number of exponents`. Although the correlation between `number of syllables` and `number of exponents` is relatively modest ($r = 0.43$), the simple correlation of `number of exponents` with reaction times is positive (albeit small, $r = 0.08$), so it seems likely that the negative coefficient of `number of exponents`in the GAMM is due to collinearity (cf. Friedman and Wall, 2005). Words with vowel-final stems elicited shorter latencies than words with consonant-final stems.

Figure 5.2 presents the non-parametric effects. As expected, response times decrease for increasing lexeme frequency (upper left), form frequency (upper center), and ending frequency (upper right). The three lower panels visualize the effects of trial (experimental time). The lower left panel shows the general effect of trial: as the experiment proceeds, subjects respond more quickly. This could represent a learning effect: over time, subjects get better at doing the lexical decision task. The lower right panel presents the factor smooth for trial by subject. In the course of the experiment, subjects show considerable fluctuations in their responses, which we attribute to the person-specific ebb and flow of attention. The lower center panel presents the interaction of trial by ending frequency. Darker red colors represent shorter response latencies, lighter yellow colors represent longer latencies. Early on in the experiment, a greater ending frequency affords shorter response times. About a third into the experiment, subjects optimize their response behavior so that they respond fastest for the most likely median frequencies, and slowest for the least likely median ending frequencies, i.e., the words with either very infrequent or very frequent endings. However, by the end of the experiment, their strategy is reversed, with the slowest responses generated for the most common ending frequencies. This is likely a consequence of many nonwords having existing endings. As a consequence, an ending of average frequency is not a reliable cue for a word response, and by the end of the experiment, participants appear to have realized this and optimized their behavior accordingly.

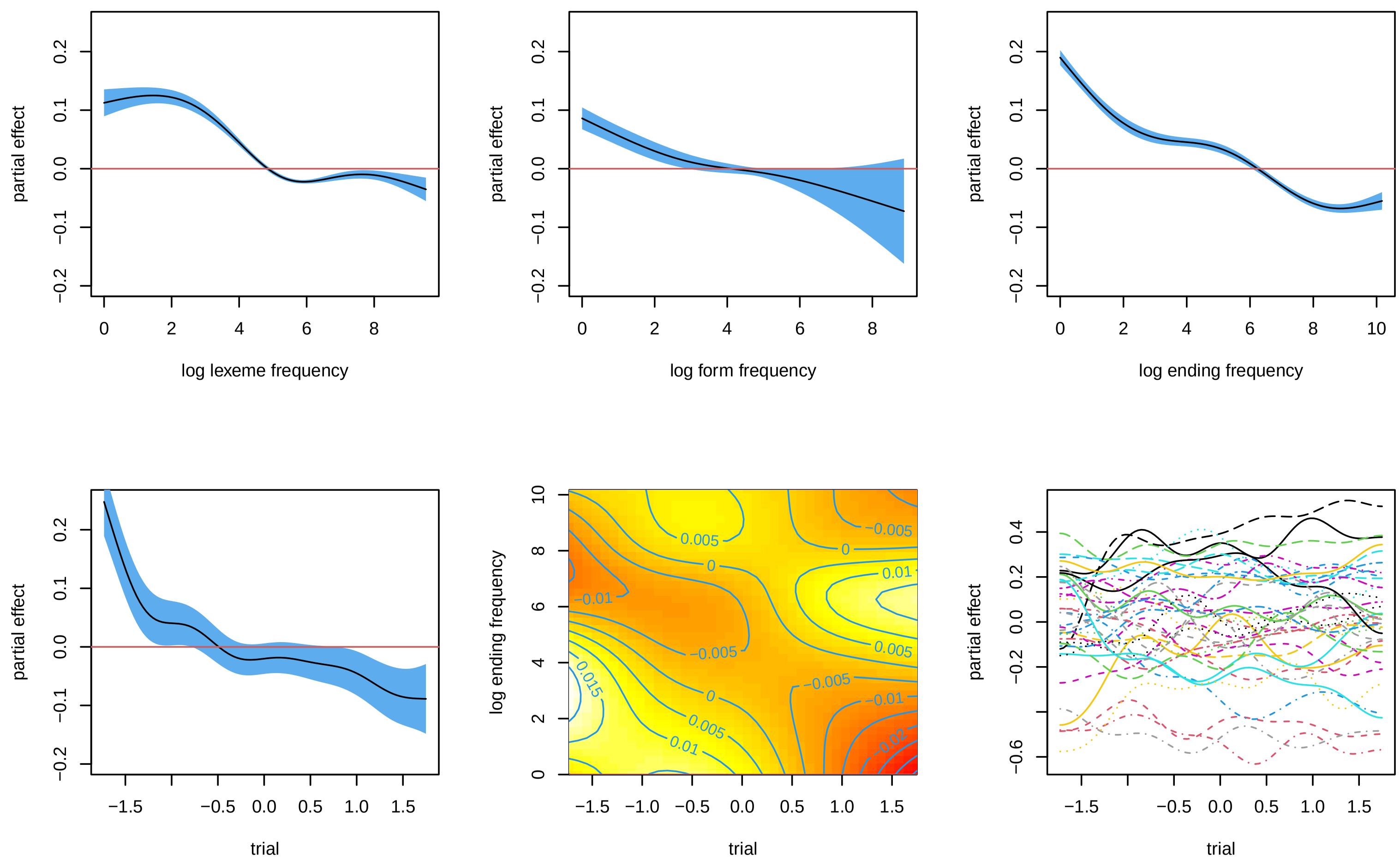


Figure 5.2: Partial effects for log lexeme frequency, log whole-word frequency, log ending frequency, trial, the interaction of trial by log ending frequency, and by-subject factor smooths for trial. The red horizontal lines represent the horizontal axis. Blue regions represent 95% confidence bands. In the plot for the tensor product smooth, lighter shades of yellow represent longer response times, and darker shades of red represent shorter response times.

Given the importance of the verb ending in the theoretical and empirical literature (e.g., Hwang et al., 2002), we fitted a second GAMM to the data in which we replaced the ending component predictors, `honorifics`, `pastTense`, `mood`, and `SLIF` with the (random-effect) factor `Ending` which combines the particular ending components(i.e., exponents) (see Tables 5.1 and 5.2). This factor can be understood not only as representing endings, but also the complex interplay between `honorifics`, `past tense`, `mood`, `speech level`, and `illocutionary force`. As the endings are in part distinguished by the familiarity and frequency of the endings, we did not include a main effect of ending frequency, to avoid problems of interpretation due to collinearity. However, we retained the interaction of trial by ending frequency. As the number of exponents was no longer significant, it was excluded from the model that is reported here. Table 5.5 provides a summary of this model. The AIC of this second model is substantially lower (21675) than that of the previous model (22290), which supports the importance of the `Ending`.

Table 5.5: Summary of the partial effects in a GAMM fitted to inverse-transformed (-1000/RT) reaction times for the participants with Korean as L1, with `Ending` predictor (AIC : 21675).

A. parametric coefficients	Estimate	Std. Error	t-value	p-value	Δ(AIC)
(Intercept)	-1.6056	0.0401	-39.9956	< 0.0001	
vowel harmony=light	-0.0417	0.0032	-12.9001	< 0.0001	165.0
number of syllables	0.0293	0.0025	11.7751	< 0.0001	120.5
stem-final segment=vowel	-0.0152	0.0032	-4.6728	< 0.0001	19.9
B. smooth terms	**edf**	**Ref.df**	**F-value**	**p-value**	**Δ(AIC)**
s(lexeme frequency)	3.9487	3.9983	115.7281	< 0.0001	466.7
s(whole-word frequency)	2.8790	3.3361	30.7671	< 0.0001	89.6
s(trial)	7.5798	8.0838	11.4815	< 0.0001	21.7
ti(trial, ending frequency)	6.5461	8.7494	3.4548	0.0004	21.3
s(trial, subject)	280.5391	359.0000	48.2360	< 0.0001	14067.5
s(Ending)	38.5067	39.0000	57.5782	< 0.0001	2142.8

Figure 5.3 presents a dotplot of the estimated by-ending random intercepts for `Ending`.

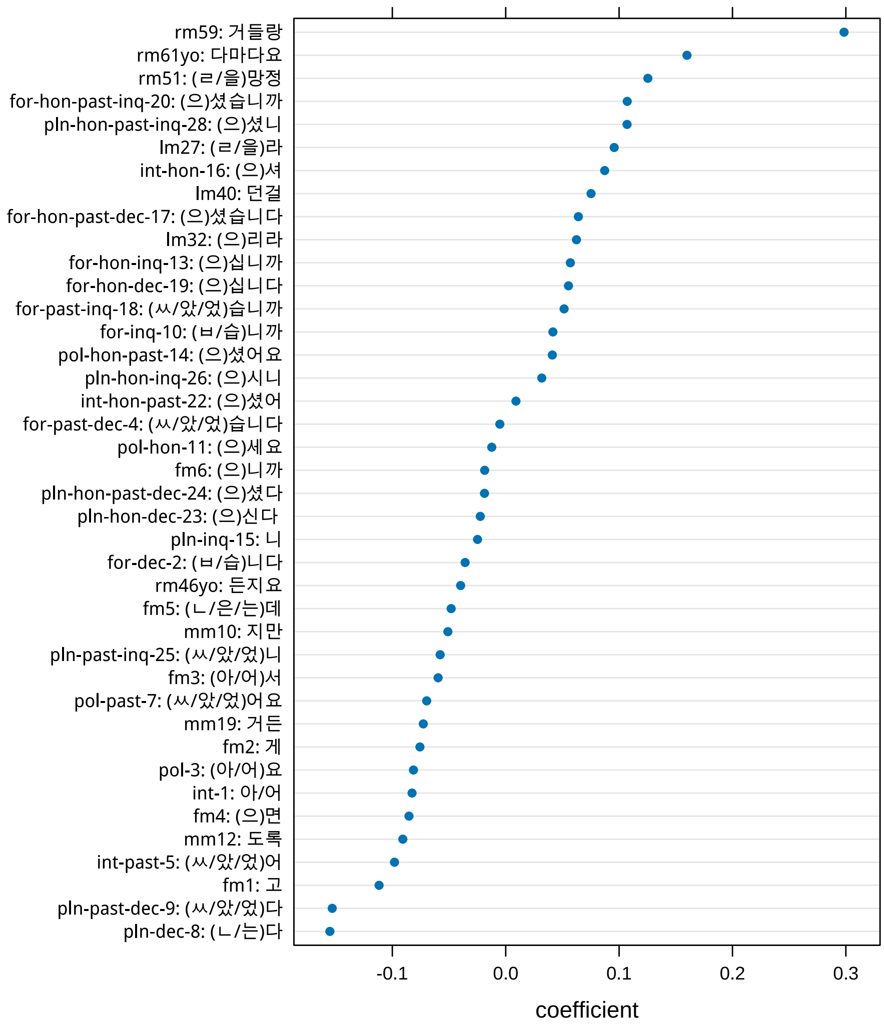


Figure 5.3: Random intercepts for `Ending` in a GAMM fitted to the reaction times for words for **Korean native speakers**. For an explanation of the abbreviations, see Tables 5.1 and 5.2. (NB: Within both the mood-added and mood-absent combinations in the code, smaller numbers indicate more frequent forms.)

.

As expected, more frequent mood forms (e.g., fm1: 고, fm4: (으) 면, fm2: 게) are recognized faster than low frequent mood forms (e.g., rm59: 거들랑, rm61yo: 다마다요). In non-mood forms, plain and intimate speech-level declarative forms are recognised quickly (e.g., pln-past-dec-9: (ㅆ/았/었) 다, int-past-5: (ㅆ/았/었) 어), while formal speech forms (e.g., for-hon-past-inq-20: (으) 셨습니까) generally show slower response times. Overall, the results reflect the effect of ending frequency; however, some forms deviate from this pattern. For example, for-past-dec-4: (ㅆ/았/었) 습니다 is a highly frequent sentence-final ending (with the lower numerical code 4 indicating higher frequency), yet it shows a relatively slower reaction time compared to less frequent forms such as pln-past-inq-25: (ㅆ/았/었) 니, pln-hon-dec-23: (으) 신다, and pln-hon-past-dec-24: (으) 셨다. Similarly, in mood markers, forms like rm46yo: 든지요 and mm12: 도록 exhibit faster reaction times than other, relatively more frequent mood forms. These findings may reflect the perceptual salience of certain mood forms.

The strong effect of `Ending`, or alternatively, `ending frequency`, with a clear effect of `lexeme frequency`, partially aligns with proposals in the literature on Korean lexical processing, which suggest that lexical access in Korean is mediated either by the stem and the final ending (Hwang et al., 2002; Hwang, 2008), or by the stem and either the pre-final or the final ending (Ahn et al., 2011a). However, the presence of the effect of whole-word frequency suggests that form-specific knowledge is also exploited by Korean native readers (Kim and Nam, 2018). Below, we will discuss the consequences of these results for theories of lexical processing in Korean.

5.4.3 Nonwords

For the analysis of the nonwords, we followed a similar modeling strategy as for the word forms. For the nonword analysis, our primary focus is on the nonword condition variable (`nwCondition`). That is, by manipulating nonword forms based on various morphological, phonological, and semantic conditions, we aim to improve our understanding of how these factors affect word processing. The specified nonword conditions were laid out above in Table 5.3 (see section 5.2.2). Table 5.6 presents a summary of the GAMM fitted to the response latencies to nonwords.

Table 5.6: Summary of the partial effects in a GAMM fitted to inverse-transformed response latencies (-1000/RT) of **Korean native speakers** to nonwords (AIC : 20201).

A. parametric coefficients	Estimate	Std. Error	t-value	p-value	Δ(AIC)
(Intercept)	-1.5173	0.0427	-35.5629	< 0.0001	
number of syllables	0.0368	0.0018	20.8422	< 0.0001	431
B. smooth terms	**edf**	**Ref.df**	**F-value**	**p-value**	**Δ(AIC)**
s(lexeme frequency)	3.7061	3.9486	11.5732	< 0.0001	40
s(ending frequency)	3.4802	3.8451	4.7515	0.0050	11
s(nwCondition)	21.6704	22.0000	119.0685	< 0.0001	2533
ti(trial, ending frequency)	11.7804	14.0792	3.9779	< 0.0001	40
s(trial, subject)	295.4314	359.0000	56.4902	< 0.0001	15788

Looking at the Table 5.6, as for the `number of syllables`, words with more syllables elicited longer reaction times, probably reflecting the costs of increased visual complexity.

The effect of `nwCondition` is very well-supported (Δ AIC: 2533). The estimated random intercepts by each level of `nwCondition` (i.e., nonword type) are presented in Figure 5.4, ordered by magnitude.

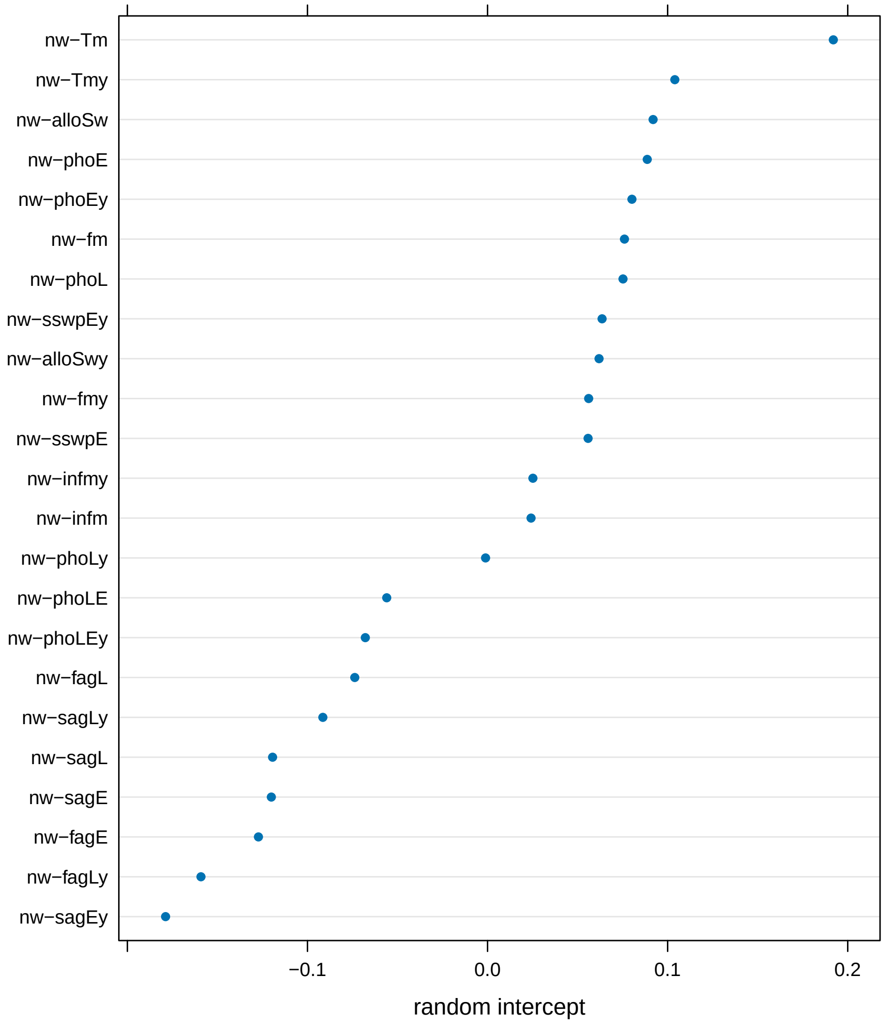


Figure 5.4: Random intercepts for `nwCondition` according to a GAMM fitted to the response latencies of **Korean native speakers** to nonwords.

In order to further assess the processing of the different nonword types, we carried out a series of follow-up analyses in which we replaced the random effect of `nwCondition` by a simple factorial contrast, where necessary reducing the dataset by discarding data points for which a given contrast is irrelevant.

The nonwords with exponents for tense and mood exponent (`Tm`[18], `Tmy`) elicited

[18]Although `Tm` showed the slowest reaction times in the results, `Tmy` was initially created as nonword stimuli and later `Tm`(i.e., absence of '-요') was regrouped based on the presence or absence of '-요'. However, due to the small number of `Tm` stimuli, we focus more on `Tmy` in the analyses of the data.

the longest reaction times. This nonword type was implemented to probe the effects of morphological and semantic complexity. The results indicate that lexical decision-making is heavily affected by the relationships between exponents and their complex semantic combinations. These nonwords are semantically rich and hence, are difficult to tell apart from real words, resulting in elongated response times.

Replacing an exponent by an inappropriate allomorph (`alloSw`) also led to long response latencies. This condition showed the second slowest response time, following the complex form-meaning combination condition (`Tm/Tmy`), which had the slowest response. Since most allomorphs are determined by the phonological conditions of the stem, it can be inferred that even in visually presented forms, word judgment is affected by phonological sensitivity. These results suggest that native speakers are somewhat sensitive to allomorphic variations.

A similar verb-ending modification involves a condition `sswpE/y`, where the syllables forming the ending are swapped. This condition showed a similar response time to those results from phonologically altered forms (`pho`) and frequent mood (`fm`) conditions that follow `alloSw` in the ranked dotplot (Figure 5.4). Similar to the results of `alloSw`, the results from the syllable-swapped condition suggest that phonological influences lead speakers to become aware of forms that deviate from expected normal patterns.

The next nonword type (pho) is included to test for phonological and orthographic effects in Korean visual word recognition. For the subset of nonwords with phonologically similar conditions but changed spellings, we considered a factor distinguishing between whether the manipulation affected the lexeme (`phoL/Ly`), the ending(`phoE/Ey`), or both (`phoLE`), with the ending manipulation as the reference level. The GAMM for this subset of the data revealed that compared to nonwords with a manipulated ending (`phoE`), nonwords with only a lexeme manipulation (`phoL`) elicited somewhat shorter reaction times ($\hat{\beta} = -0.02, p = 0.0026$) and that the manipulations in both lexeme and ending (`phoLE/LEy`) resulted in substantially shorter reaction times

($\hat{\beta} = -0.146, p < 0.0001$). This is, nonwords that involve changes in both the stem and the ending are the most clearly nonword-like in this set. This suggests that changes in verb endings are less easily recognisable than changes in other parts, such as the lexeme or both the lexeme and ending. On the other hand, the fastest response time was observed in the condition where both the lexeme and the ending were altered can be attributed to the fact that the overall modification facilitated the recognition of the form as clearly deviating from a normal word.

Next, zooming in on forms with mood exponents, interestingly, nonwords with less frequent (altered) mood exponents (`infm, infmy`) exhibited slightly faster response times ($\hat{\beta} = -0.037, p = 0.0006$) compared to nonwords with more frequent (altered) mood forms (`fm, fmy`). One possible explanation is that modifications to less frequent mood forms, particularly those with unique three-syllable structures familiar to native speakers (e.g., -거들랑, -답시고), may be more salient. In contrast, alterations to monosyllabic forms(e.g., -고, -서) within more frequent mood forms may be less noticeable, making the nonwords harder to detect.

As expected, the conditions that produced the fastest response times were those with phonologically impossible or unused forms (`sag/fag`), indicating these were least similar to words. These include conditions where either a single syllable of the verb ending or stem was altered (`sag`) and those where the entire stem or ending was modified (`fag`). Among them, the condition where only one syllable of the ending was changed with the addition of -요/-yo/ (`sagEy`) showed the fastest response time, while the condition where the entire stem was altered (`fagL`) had relatively slower response times than other `sag/fag` conditions. In the case of the results from `fagL`, the delay in recognition can be attributed to the fact that many verb stems consist of a single syllable, thus, the combination of a modified single-syllable stem with a regular ending may have caused a slight processing delay. However, the slightly faster response time in the -요/yo/-added `fagLy` condition suggests that the presence of -요 facilitates word recognition.

Lastly, regarding the last ending form, -요/yo/, here, we assume the effect of

-요 as a salient effect. That is, it starts from the recognition of a familiar form, which leads to closer scrutiny of the preceding segments before -요 /yo/, influencing lexical decision and ultimately affecting response times. We assume that this prioritised and rapid recognition of the form will generally facilitate processing. However, under conditions where stronger overriding factors are at play, this effect may instead result in delayed responses. Comparing conditions with and without -요 (indicated by `y` at the end of the code) across various nonword conditions shows that the presence of -요 generally led to faster responses (`i.e., Tmy, alloSwy, phoEy, fmy, phoLy, phoLEy, fagLy, sagEy`). This suggests that the prominent -요 at the very end of the verb ending facilitates word recognition. However, a few exceptions did not show these facilitation effects. This may be due to recognition effects related to syllable structure (e.g.,`sswpEy, sagLy`) or certain less frequent unique three-syllable mood forms (e.g.,`infmy`).

For the frequency effects, we focused on lexeme frequency and ending frequency. Whole-word frequency is zero for all nonwords, and hence is not taken into consideration. The partial effects of the two frequency effects are presented in Figure 5.5.

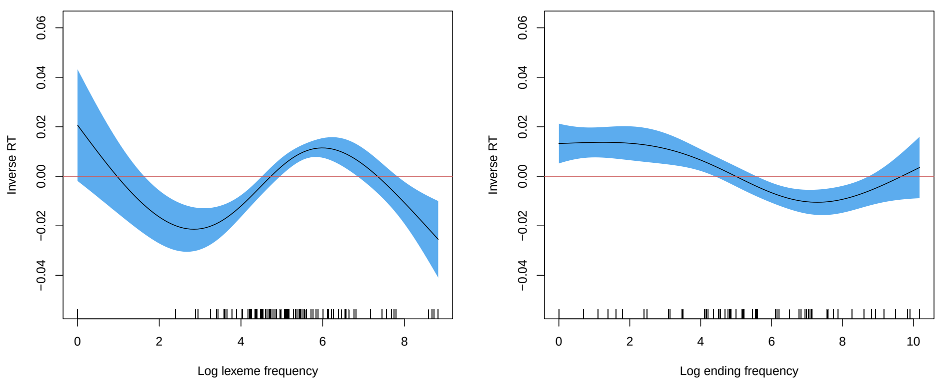


Figure 5.5: Partial effect of lexeme frequency (left) and ending frequency (right) for inverse-transformed response latencies to nonwords.

The lexeme frequency (left panel) in Figure 5.5 clarifies that the single nonword created from a lexeme with only one attested token in the corpus (log lexeme frequency 0), has a long reaction time. Setting this outlier aside, we observe an inverse U-shaped effect, with the longest response times for nonwords with intermediate lexeme frequencies. Nonwords constructed from higher frequency lexemes elicited the shortest reaction times. Apparently, for such words, the nonword manipulation resulted in forms that were clearly wrong, enabling fast nonword responses. At the same time, responses were also reasonably fast for the nonwords created from lower-frequency lexemes. In the right panel showing ending frequency, a small effect is observed in the frequency range of 2 to 7, within which higher ending frequency tends to predict faster response times.

On top of this main effect rides an interaction of ending frequency by trial, shown in Figure 5.6. As for words, participants changed their response strategy over the course of the experiment. Instead of responding more quickly to stimuli with a high ending frequency, they responded more slowly. Participants realized that a high ending frequency is not a reliable cue to word status, and started to respond to such stimuli with caution. Additionally, considering that the nonwords were created by manipulating stems and endings separately across various conditions, it is reasonable to assume that there may be interactive effects between ending frequency and the nonword conditions.

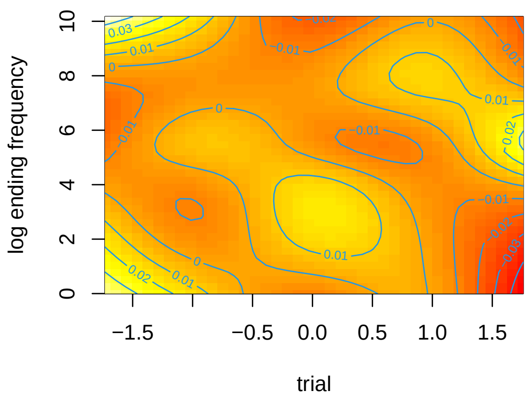


Figure 5.6: Partial effect of the interaction of trial by log ending frequency in the GAMM fitted to the response latencies of L1 speakers to nonwords.

5.5 Discussion

This study examined how native Korean speakers process Korean verbs in a visual lexical decision task. This section summarises several key patterns revealed in the study and discusses them in more detail.

First, excluding individual difference variables, the `Ending` variable showed a strikingly high importance score based on the AIC results. This was followed by the ending frequency, which also had a substantial weight, and then by lexeme frequency and the `SLIF` variable, which showed similar levels of importance. These results suggest that verb endings play a particularly significant role in Korean verb processing. Specifically, when comparing endings and stems, it appears that endings exert a greater impact — a finding that reflects the morphological characteristics of Korean verbs, where endings are often more complex both morphologically and semantically than stems.

The finding that `Ending` had a strong impact on Korean verb processing is highly significant. Although the effects of individual verb components, such as verb stems and the various elements constituting endings, were also examined, the impact of `Ending` far surpassed them. This suggests that Korean verbs may be recognised as larger, patterned structural morphological and semantic units rather than as collections of smaller parts. For native speakers, when replacing the `Ending` variable with ending frequency in the statistical model, the increase in AIC suggests that the effect of ending frequency is approximately twice as large as that of stem frequency, while whole-word frequency shows the smallest effect, about one-fourth that of stem frequency. Based on these results, a general picture of the processing structure suggests that the stem and a larger unit of the ending may be recognised separately. In this case, individual exponents that make up the ending may be integrated into a single "`Ending`" unit, partially aligning with the full-listing hypothesis, while the whole word is still processed by separating the stem and the `Ending`, consistent with the decomposition hypothesis. This duality corresponds to the parallel dual-route model (Baayen et al., 1997), which assumes that both routes may operate simultaneously. The strong influence of ending frequency also implies that processing might begin with a holistic scan of the ending, with the ending playing a more dominant role than the stem in lexical access. Additionally, although whole-word frequency showed a weaker effect compared to ending frequency, its strong presence in certain cases may lead to recognition of the entire verb form as a single lexical unit.

In this regard, the results of this study partly align with some previous findings. First, when we look at native speakers, for example, several earlier studies (e.g., Lee et al., 2018b; Ahn et al., 2011a; Hwang et al., 2002) suggested the possibility that verb stems and prefinal or final endings (e.g., conjunctive endings and sentence-final speech-level and illocutionary forms) are processed separately. In this sense, the general processing structure of 'stem' + 'Ending' found in the current study is consistent with those earlier claims. However, the interpretation from Hwang et al. (2002), which suggested that prefinal endings (e.g., past tense or honorific exponents) are processed together with the stem—based on experiments that separated prefinal and final endings according to traditional structural classifications—does not align with the current findings. Moreover, most earlier studies on Korean verbal processing have treated endings by dividing them into prefinal and final exponents, focusing on whether a specific exponent is processed independently or together with the stem, rather than examining the combination of the multiple exponents within verb forms. In this regard, it is not straightforward to directly compare those results with the present study. Additionally, some studies have supported the full-listing hypothesis, suggesting that Korean verb forms are recognised as an Eojeol unit (i.e., whole-word form) (e.g., Song, 2009; Kim T., 2015; Kim et al., 2022, 2018b; Choi, 2012; Kim et al., 2000). These views differ from the 'stem'+ 'Ending' processing structure proposed in the current study. However, the interpretation of this study is compatible with the possibility that whole-word recognition may selectively occur when whole-word frequency has a significantly dominant influence.

Another notable feature is the influence of `SLIF`, which appeared at a similar level to lexeme frequency. This suggests that, during verb processing, particular attention is paid to the final sentence-ending (i.e., speech level and illocutionary force).

A closer look at SLIF reveals that unmarked elicited the slowest responses, while plain declarative showed the fastest. The unmarked forms largely consisted of mood forms and typically ended in conjunctive endings, which may have contributed to their slower processing. This was followed by the formal inquisitive and formal declarative forms, which likely resulted in slower responses due to their noticeably longer word length compared to other forms. The remaining three forms (polite unmarked, honorific declarative/imperative, and intimate unmarked) showed relatively similar response times, with no clear shared features standing out. The fastest response times were observed for the plain declarative and plain inquisitive forms. These were shorter in length and may have functioned as strong cues for word-final position, featuring characteristic endings such as -는다 or -니 that likely served as reliable lexical signals. In particular, the plain declarative form '-는다' is not only a common feature of spoken Korean but also one of the most representative forms in written language, and as such, it is likely to be highly competitive in terms of frequency effects compared to other forms.

Compared to SLIF, the influence of other ending components, particularly honorific endings, was minimal, ranking near the bottom in terms of importance. Response times were faster when honorific endings were absent. The past tense marker, meanwhile, did not show a statistically significant effect. These results are intriguing, but they raise questions about how authentically honorific or past tense forms were perceived in the lexical decision task, where no contextual information was provided. This issue also applies to the SLIF making the interpretation somewhat unclear.

Mood forms, defined broadly in this study as all forms that end in structures other than SLIF, typically including various nuanced conjunctive endings, had an influence similar in magnitude to overall whole-word frequency. Surprisingly, response times were faster for mood forms than for non-mood forms. This suggests that the semantic complexity introduced by mood may have less impact on L1 speakers than other variables, such as word length, which appear to exert a stronger influence in cases not involving mood.

The variables with the least impact were stem-final segment, honorifics, and number of exponents, while past tense did not yield statistically significant effects.

In terms of frequency effects, although the pattern was non-linear, there was a general trend showing faster response times for higher-frequency items. For whole-word frequency, a wider distribution of response times was observed in the high-frequency range (range 6–8). The stem and ending frequency effects are nonlinear compared to the whole-word frequency effect. Reaction times showed facilitation up to about a frequency of 6 for stems, but no greater effect was observed in the high stem frequency range (between 6 and 8). In contrast, ending frequency showed a steady decrease in reaction times up to a frequency of 8.

A closer look at the structure of Ending in the results reveals that, while frequency effects were generally present, certain exceptions emerged (e.g., for-past-dec-4: (ㅆ/았/었) 습니다, m46yo: 든지요). These exceptions were mostly found in mood forms, which often involve longer forms or low-frequency but salient forms (e.g., m46yo: 든지요, mm12: 도록) that are more likely to be perceived as independent words rather than typical verb endings. Such forms tended to deviate from the expected frequency-based pattern.

This study also aimed to examine morphological, semantic, and phonological effects through nonword manipulations. Excluding speaker-specific differences—the most influential factor in the GAMM statistical model—the nonword condition(`nwCondition`) showed a remarkably strong effect. These nonword conditions were designed to reflect specific morphological, semantic, and phonological characteristics. In the detailed results, the `Tm/Tmy` conditions, which were manipulated to incorporate a higher level of morphological and semantic complexity, yielded the slowest response times. In contrast, the `sag` and `fag` conditions, which involved unattested or ill-formed syllables, showed the fastest response times.

Conditions that followed the `Tm/Tmy` pattern and exhibited slower response times included: `alloSw`, which involved switching allomorphs; `phoE` and `phoEy`, which manipulated ending part with similar phonological effects to standard forms; `fm`, a frequent mood form; and `sswpEy`, which involved rearranging the syllables of the endings.

In the case of `alloSw` condition, both morphological and phonological aspects are considered. The manipulation involving the reversal of allomorphs led to the slowest response times and the highest error rate (see Figure 8.11), suggesting that visual word processing is highly sensitive to both morphological and phonological disruptions. A similar pattern was observed in condition `sswpEy`, where the syllables comprising the ending were rearranged. This condition also showed relatively slow response times, which may indicate that morphological alterations of the ending can produce phonological impacts that interfere with visual lexical judgment, thereby slowing down processing.

The subsequent `phoE/Ey` results can be interpreted in a similar vein. Among two types of manipulations(i.e., ending manipulation(`phoE`) or stem manipulation(`phoL`)), the ending manipulations resulted in longer response times than stem manipulations such as `phoL` and `phoLy`. This implies that recognising endings might be more challenging than recognising stems in lexical processing. Another possibility[19] is that lexical judgment may have taken longer when manipulating longer endings, compared to stem changes that were often monosyllabic and occurred at the beginning of the word.

Interestingly, the frequent mood forms (`fm`) yielded slower response times and higher error rates compared to the less frequent mood forms (`infm`). One possible interpretation is that, from a morphological standpoint, the manipulated forms of the more frequent moods—often consisting of common monosyllabic conjunctive endings (-고, -서, -며)—may experience delayed recognition because their altered forms provide less

[19]The analysis of lexical factors through these nonword conditions also calls for consideration of several non-lexical factors known to affect nonword judgment. That is, previous studies have pointed out characteristic influences that arise in the decision process between word and nonword choices (e.g., Hendrix and Sun, 2021), as well as features specific to different types of nonwords (e.g., Bae, 2018; Frisch et al., 2000). However, in this study, those complex variables were not incorporated in the interpretation of the results. One point to note is that most of the stimuli used in this study were only minimally modified from real words, making it difficult to directly apply the above-mentioned complex factors in a straightforward manner. A similar nonword approach to this study is found in the study by Csaba and Levente (1995), where the entire stimulus forms are treated as either 'correct' words or 'incorrect' ('spoiled') words (For further discussion, see Csaba and Levente, 1995).

informative cues. In contrast, the less frequent `infm` mood forms (-답시고, -더라손, -거들랑), though lower in frequency, tend to be longer (2–3 syllables) and have more distinctive single word-like forms, which may have facilitated quicker and more accurate recognition.

Among the `sag/fag` conditions manipulated into impossible forms, condition `sagEy`, which involved ending manipulation, elicited the slowest response times, while `fagL` showed the fastest. Overall, with the exception of `fagLy`, forms involving ending manipulations tended to produce faster responses compared to those involving stem manipulations. This may be due to the fact that ending typically consist of more syllables than stems, which are often only one or two syllables long. As a result, when these longer endings were manipulated into novel forms, participants may have more readily recognized them as nonwords. In the same vein, the faster responses in `fagLy` and `fagE` may be attributed to the fact that both the entire stem or ending was altered, making recognition as a nonword more immediate. On the other hand, the slower response times in `fagL` and `sagLy` may be explained by the manipulation of short, monosyllabic stems alone, which could have made the nonword status less immediately apparent compared to other conditions.

Lastly, to examine the effect of -요 /-yo/ (polite speech ending), one of the most frequent and salient final syllables in Korean verbs, we manipulated the presence or absence of -요 in the nonword conditions. Within each specific nonword manipulation, response times were generally faster when -요 was present compared to when it was absent. This suggests that the presence of -요 may facilitate lexical decision processing. However, there were exceptions—specifically in conditions `sswpEy, infmy` and `sagLy`, where the presence of -요 actually led to slower response times than when it was absent. In these cases, the presence of -요 appeared to interfere with processing rather than enhance it. It is not easy to determine the exact cause of these results. However, one possible explanation is that the morphological and phonological alterations caused by switching the ending syllables (`sswpEy`) may have had a stronger effect than the facilitating effect of the ending -요 in lexical decision. In addition, in the case of changes made only to the stem syllables (`sagLy`), many of the stems were monosyllabic, and the alteration might have conflicted with the -요 cue that typically makes the form appear word-like, thereby causing delayed response times. This issue will be revisited when discussing the results for non-native speakers.

Chapter 6

Studies on L2 Lexical Processing

In most cases, unlike in L1 processing, L2 lexical processing is expected to involve more complex and multifaceted factors. Theories of L2 lexical processing commonly focus on how non-native speakers' processing differs from that of native speakers, or to what extent advanced L2 learners can approximate native-like processing. Within this framework, numerous lexical and non-lexical factors are intricately involved, including L1 effects, semantic transparency, and word frequency as lexical variables, and individual differences such as working memory and the L2 learning environment as non-lexical factors.

One of the most fundamental questions in L2 lexical processing is how it differs from or resembles L1 lexical processing. There is an ongoing debate on this issue. Some scholars argue that L2 processing is qualitatively different from L1 processing (e.g., Clahsen and Felser, 2006b,a; Silva and Clahsen, 2008; Neubauer and Clahsen, 2009; Ullman, 2001, 2004; Lim et al., 2020), while others claim that L2 learners process language in ways that are, to some extent, similar to native speakers (e.g., Diependaele et al., 2011; Feldman et al., 2010; Hwang et al., 2008). More specifically, it has been proposed that native speakers tend to decompose words into morphemes during processing, whereas non-native speakers are more likely to recognise words as whole forms (e.g., Ahn et al., 2011b; Kim et al., 2018b; Clahsen and Neubauer, 2010; Clahsen et al., 2013; Silva and Clahsen, 2008; Ullman, 2001; VanPatten, 2004). However, other studies suggest that non-native speakers may also engage in morphological decomposition (e.g., Li et al., 2017). Additionally, some studies provide support for a hybrid account (e.g., Kirkici and Clahsen, 2013; Ahn et al., 2013; Lee et al., 2020; Silva and Clahsen, 2008; Neubauer and Clahsen, 2009; Feldman et al., 2010).

Focusing on Korean verb processing, the present study aims to examine how advanced L2 learners of Korean process Korean verbs, with an emphasis on lexical factors. From this perspective, a key issue in L2 morphological processing is how L1 morphological processing patterns influence the processing of morphologically similar or dissimilar structures in the L2 (e.g., Kroll and De Groot, 2014; MacWhinney et al., 2005; Clahsen and Felser, 2006c; Portin et al., 2008; Vainio et al., 2014). Based on this view, a widely assumed hypothesis is that advanced L2 learners whose L1 is typologically similar to the L2 are more likely to show native-like morphological processing patterns in the L2, while those with typologically different L1s will exhibit divergent patterns (e.g., Hwang et al., 2008; Portin et al., 2008; Vainio et al., 2014). For example, Portin et al. (2008) found that advanced L2 learners of Swedish with L1 backgrounds in Hungarian (an agglutinative language) and Chinese (an isolating language) showed different responses in a visual lexical decision task involving inflected Swedish nouns. The L1 Hungarian learners were interpreted as recognising high-frequency inflected forms as whole words, while decomposing low-frequency forms; in contrast, the L1 Chinese learners were found to rely on whole-word recognition regardless of frequency. Another important issue is semantic transparency. In L2 lexical processing, the morphological and semantic complexity of a word raises the question of how clearly its meaning is conveyed to non-native speakers. This concern is particularly relevant for Korean verbs, where multiple meanings are often encoded in inflectional endings, and a wide range of mood forms—many of which express subtle emotional nuances—are intertwined. It remains unclear to what extent such complex forms are semantically transparent and accessible to L2 learners. In this context, discussions of semantic transparency have also drawn significant attention in research on L2 lexical processing (see Gor et al., 2021, for detailed discussions). In addition, various factors are likely to interact with the linguistic characteristics of L2 and L1, contributing in complex ways to L2 lexical processing.

6.1 Studies on L2 Korean verb processing

While research on Korean as a second language has been extensive since the early 1990s, it appears that studies focusing on L2 Korean from a psycholinguistic perspective have not been as prevalent compared to other lines of research, such as those focusing on research undertaken with pedagogical intent. In fact, the limited research on Korean from psycholinguistic perspectives extends not only to L2 Korean but also to Korean as L1. Nonetheless, there exist several studies that provide a point of departure for the present thesis. As briefly mentioned earlier, there are important variables to consider for non-native lexical processing research that do not play a role for native speakers. Specifically, factors such as a non-native speaker's proficiency in the target L2, their L1 background, and how they have developed and acquired their L2 vary significantly across individuals. These variables can greatly impact experimental research outcomes. In what follows, a summary is provided of several relevant studies from the perspective of these key variables.

Lee et al. (2020) examined highly proficient learners (level 6 or above on TOPIK) of Korean, targeting whether they recognise the pre-final past tense ending of Korean verbs. Primed lexical decision tasks, including a cross-modal visual-auditory priming task, suggested that there was a stem frequency effect, but that there was no significant difference in RTs between low and high frequent final-endings which follow the pre-final past tense ending. Accordingly, this result was interpreted to mean that those highly proficient learners process Korean verbal inflectional form as a whole-word form rather than parsing the form into a sequence of morphemes. However, in this study, the L1 backgrounds of the 72 non-native participants were highly diverse. While 50 participants had a Chinese L1 background, the remaining participants represented various L1 backgrounds, with only one to three speakers per language. For this reason, there are limitations to interpreting the results solely based on the characteristics of the participants' L1s.

Lim et al. (2020) investigated the effects of semantic context and word frequency on L2 Korean learners' word[20] processing. 72 Korean non-native speakers with highly advanced Korean levels (TOPIK 6) had visual word recognition tests in three conditions for presenting stimuli with two SOAs (43ms, 230ms) in a cross-modal primed lexical decision task. The stimuli were classified into four conditions according to two factors: the lexical frequency of the target word (high-frequency vs. low-frequency) and the semantic relatedness between the prime and the target word (semantically related, e.g., 친구 (friend)–사람 (person), vs. semantically unrelated, e.g., 첨가 (addition)–비극 (tragedy)). The results showed that there were word frequency effects (the size of the effects: cross-modal > SOA 230ms > SOA 43ms) but no semantic priming effects in all priming conditions (i.e., two SOAs and cross-modal) and also no interactions between word frequency and measures of semantic association. This result contrasts with findings from native speaker studies (e.g., Lee et al., 2019a) that support the theory of automatic spreading activation (Neely and Keefe, 1989), suggesting that even advanced-level non-native speakers process lexical items differently from native speakers in Korean.

Kim et al. (2018a) carried out a lexical decision task with 45 non-native speakers of Korean from various L1 backgrounds to examine the impact of several lexical variables on the recognition of Korean verb forms. The experiment was designed with 300 Korean verb forms and non-words as stimuli, focusing on eight lexical variables: stroke count, number of letters, syllable count, morpheme count, root frequency, whole word frequency, first-syllable-sharing frequency, and the number of dictionary meanings. The experiment indicated that in the non-native speaker groups, which differed in their level of proficiency (TOPIK Level 5 and 6), the group with slightly lower proficiency (TOPIK Level 5) exhibited significant effects from syllable count and whole word frequency. In contrast, the more proficient group (TOPIK Level 6) showed the effects (ordered by effect size) of

[20] The part-of-speech information for the lexical stimuli used in the experiment is not clearly specified. However, based on the examples provided, it appears that both verbal and nominal forms were included among the stimuli.

whole word frequency, number of letters, syllable count, and root frequency. Additionally, the native Korean speaker group demonstrated only a weak effect for whole word frequency. Given that whole word frequency was consistently significant across all groups, the authors interpreted the results to support the full-listing hypothesis. The study also provides clear evidence that the variables affecting word recognition may differ depending on the proficiency level of non-native speakers.

The studies reviewed thus far did not focus on how the L1 of non-native speakers might affect lexical processing in Korean as a second language. This issue is addressed by the study of Hwang et al. (2008), which considered four sets of participants, cross-classified by L1 (Chinese and Japanese) and proficiency in Korean (beginner and advanced). The experiment reported in this study was a grammaticality judgment task of Korean verb forms, including complex verb forms with two or three pre-final endings. The results revealed distinct differences based on L1. Advanced-level participants with a Japanese L1 background, which is morphologically similar to Korean, made judgments comparable to those of native speakers. On the other hand, advanced-level participants with a Chinese L1 background made significantly more errors, showing results similar to those of beginner-level Chinese and Japanese participants. As Chinese has no inflectional morphology, this result is unsurprising. Hwang et al. (2008) argue that for Japanese participants, L1 transfer is beneficial, whereas for Chinese L1 speakers, their native language stands in the way of learning Korean.

Ahn (2015) reported a study with three groups of non-native speakers, each with different L1 backgrounds (Chinese, English, and Japanese). This study also made use of a grammaticality judgement task using nouns with nominative/accusative case. The study reported that the Japanese L1 group achieved the highest accuracy in their judgments of the case particles, while the Chinese L1 group showed the lowest accuracy.

Ahn et al. (2013) conducted a visual priming lexical decision experiment with advanced non-native speakers of Korean with a Chinese L1 background. The aim of this study was to investigate the processing of both final and pre-final verb endings for honorific form -시- and past tense -었/았- in Korean. Their primed lexical decision task included seven priming conditions: identitical (사다/s^{h}ada/'to buy' – 사다), morphological (사고/s^{h}ago/ 'buy-CONJ'– 사다), semantic (시장/ɕidzaŋ/'market' – 사다), phonological (사다/s^{h}ada/ – 사다), past (샀다/s^{h}at̚t͈a/'buy-PAST' – 사다), honorific (사시다/s^{h}aɕida/ 'buy-HON' – 사다), and unrelated (하늘/hanɯl/'sky' – 사다). The study reports a strong priming effect for past tense endings, but no priming for honorific endings and reduced priming effects for the final ending (i.e., CONJ -고/-go/in morphological condition). This finding also contrasts with previous studies (e.g., Ahn et al., 2011a) on native speakers, suggesting that non-native speakers are more likely to perceive words as unanalysed units rather than through morphological decomposition.

Several studies report analyses of the errors made by L2 learners, and discuss developmental patterns in the non-native learners' productions of Korean verb forms(e.g., Yune, 2007; Seong, 2006; Eom, 2011; Chae, 2012; Choi, 2013; Han, 2018; Choi et al., 2020). In particular, Chae (2012) analyzed the usage patterns of Korean sentence-final verb endings in everyday conversations by L1 Chinese learners with advanced Korean proficiency[21], comparing them to native Korean speakers using data from the 21st Century Sejong Project corpus. The study found that while advanced non-native speakers displayed a level of naturalness in their use of verb endings that was nearly indistinguishable from that of native speakers, they differed in their range of endings used. Unlike native speakers, who employed a variety of verb endings[22], non-native speakers overwhelmingly relied on the most frequent endings such as -아요/-어요 (polite speech level, 78.4%), -지요 (-죠) (mood ending 'isn't it?', 5.59%) and -아/-어 (intimate speech level, 5.26%). In contrast, the analysis of native speaker data revealed that the most frequently used endings were -아/-어 (intimate speech level, 44.97%), followed by -지 (mood ending, 'isn't it?'(casual speech), 9.46%), and -아요/-어요 (polite speech level, 8.51%). This indicates that informal, casual sentence endings were predominantly used in spoken language by

[21] Chinese students enrolled in master's and doctoral programs in Korean Language Education at graduate schools in Korea, with an average of 5 to 6 years of Korean language learning experience.

[22] Although the number of ending forms identified in the corpus differed significantly between native speakers (481 types) and non-native speakers (57 types), the size of the corpora also varied greatly (native speakers: 369,859 words; non-native speakers: 13,379 words). Therefore, it is difficult to make a definitive comparison regarding the diversity of ending forms solely based on these raw numbers.

native speakers. This pattern can be interpreted in light of the broader range of interlocutors native speakers communicate with, including friends and family, for whom informal speech is common. On the other hand, non-native speakers tend to use more polite forms such as -아요/-어요, as they primarily interact with Korean speakers to whom honorifics are expected. Moreover, -아요/-어요 endings are typically introduced early in language instruction and are thus more frequently used by non-native speakers. And more importantly, in the data of native speakers, approximately 40% of the remaining cases use different endings, clearly indicating that native speakers use a much wider range of those endings compared to non-native speakers.

Han (2018) also analysed an L2 learner corpus[23], focusing on the use of Korean verb sentence final-endings, and examining patterns across different proficiency levels and L1 backgrounds in both written and spoken language. Notably, the study found that a wider variety of verb endings appeared in spoken language compared to written language. In the analysis, the types of the final endings that accounted for 99% of cumulative frequency were found to be 7 types in written language (e.g., -다 [58%], -습니다 [20%]) and 24 types in spoken language (e.g., -어요 [55%], -습니다 [33%]). Nevertheless, non-native speakers exhibited a higher error rate in the use of sentence-final endings in written language compared to spoken language. As expected, error rates declined as proficiency levels increased from beginner to advanced.

[23] 'Korean Learners' Corpus' based on 240,000 written and 170,000 spoken words collected from 2015 to 2017 by the National Institute of Korean Language, https://kcorpus.korean.go.kr/.

Chapter 7

Experiment 2: Non-native Speakers

This chapter reports an unprimed visual lexical decision experiment carried out with 4 groups of non-native speakers of Korean, using the same 4000-word and 1000-nonword stimuli that were presented to native speakers in Experiment 1. As for Experiment 1, we made use of a regression design analysed with the generalised additive mixed model.

The non-native speakers of Korean had one of four Asian languages as their L1: Japanese, Mongolian, Chinese, and Vietnamese. Japanese was selected as a language with a morphological system of verb morphology that is similar to that of Korean. Both Korean and Japanese share the concept of speech levels and honorifics, which are determined by the relationship between the speaker and the listener. However, the grammatical system of honorifics in Japanese is less complex and less elaborately developed compared to Korean. Unlike Korean, standard Japanese does not have vowel harmony.

Mongolian is a language with vowel harmony (as does Korean) and a complex syllabic structure with clusters of up to three consonants in syllable-final position (unlike Korean). It is also an agglutinative language, but with more extensive case marking on nouns, and with verbs that are marked for one of five voices, and in addition are marked for aspect, tense and epistemic modality and evidentiality. Unlike Korean, speech levels are not applied to verb morphology. Instead, sentence types (i.e., illocutionary force) can include forms for a concept of honorifics written separately from the main verb by a space.

Chinese[24] and Vietnamese are tonal languages with extensive compounding but no inflectional morphology. Chinese has a few derivational suffixes (see, e.g., Shen and Baayen, 2022). A wide range of particles and auxiliary verbs carry the burden of expressing temporal, aspectual, and epistemic semantics. The tone system of Vietnamese is more complex than that of Chinese, and Vietnamese also has a larger inventory of syllables (both with and without taking tone into consideration).

The writing systems of these languages are markedly different. Japanese combines Kanji (i.e., Chinese characters) with the simpler Hiragana and Katakana scripts. Chinese use their own Chinese characters, but Pinyin is widely used for typing. Vietnamese uses a Roman alphabet with many diacritics distinguishing between consonants and tones. Mongolian is written with the Cyrillic alphabet in Russia, with its own alphabetic script in the PRC, but for informal internet communication, the Latin alphabet is widely used.

In light of the differences between the four languages, the following predictions for lexical processing differences can be formulated:

- Japanese: because the verbal morphology of Japanese is most similar to that of Korean, lexical processing for Japanese L1 speakers is expected to be most similar to that of Korean L1 speakers.
- Mongolian: since Mongolian has vowel harmony, the advantage of light vowel harmony for Korean speakers is expected to be easier to learn for Mongolian L1 speakers.
- Chinese and Vietnamese: as these are languages without inflectional morphology, learning Korean is expected to be less accurate compared to speakers of Mongolian and Japanese.

[24]In this study, since the participants'first languages include both Mandarin and Cantonese, and given the morphological similarities between the two, we will collectively refer to them as "Chinese".

- Korean L1 speakers respond more slowly for words with more syllables. For the other two agglutinative languages (Japanese, Mongolian), this length effect may be similar, but for the isolating languages (Vietnamese and Mandarin Chinese), the cost of having more syllables might be larger or shorter than the above language groups.
- For all non-native speaker groups, all frequency effects are expected to be present, with possibly a larger whole word frequency effect compared to Korean L1 speakers (Clahsen and Felser, 2006b), and possibly a weaker effect of ending frequency.
- The response optimisation for ending frequency (the trial by ending frequency interaction) may be available only for native speakers.
- In relation to the morphological and semantic complexity of Korean verbs, non-native speakers are likely to be more consciously aware during processing these particular forms than native speakers. That is, in variables such as tense, honorifics, and mood forms, response delays or facilitation effects are expected to be more pronounced in non-native speakers than in native speakers.

7.1 Methods

7.2 Participants

Native speakers of four Asian L1 languages, Japanese, Mongolian, Chinese, and Vietnamese, were recruited to participate in the visual lexical decision experiment. Each group consisted of 20 to 21 participants. Recruitment was conducted both online (via posts in foreign students' or professional communities for expatriates residing in Korea) and offline (through postings on university language institute bulletin boards). Most participants were highly proficient non-native Korean speakers (having reached the highest level (level 6) of the Test of Proficiency in Korean (TOPIK). The reason for limiting participants to those with advanced proficiency is to examine how closely non-native speakers with near-native fluency process Korean verb morphology—how detailed or in what manner this processing occurs. While highly proficient non-native speakers generally have no difficulty with everyday communication, it remains unclear how they process complex Korean verb forms. Participants ranged in age from 19 to 37, with the gender distribution detailed in Table 7.1.

	Japanese	Mongolian	Chinese	Vietnamese
male	0	4	2	2
female	21	16	18	18

Table 7.1: Gender distribution of the participants from the L1 speaker groups.

7.3 Materials and procedure

The materials and experiment procedure were the same as those in Experiment 1.

7.4 Results from Words

In what follows, we first report the GAMMs fitted to the datasets of the four non-native groups separately. We then present analyses comparing the reaction times of these four speaker groups to those of the Korean native speakers. For each language, we carried out two parallel analyses, one with detailed ending components, `honorifics, past tense, mood, SLIF` as predictors, and another in which `Ending` was the key morphological and semantic predictor for combined ending forms but without the above ending components.

7.4.1 Participants with Japanese as L1

Table 7.2 presents the model summary when detailed ending components, `honorifics, past tense, mood, SLIF` are included as predictors. Words that were unmarked for honorifics were responded to more quickly than words overtly realising honorifics. Words in the past were responded to more slowly, contrasting with the Korean native speakers, for whom there was no significant difference. Words without an exponent for `mood` showed a small facilitation effect, but the increase in AIC (5.1) when this predictor is withheld is small compared to the other predictors. Regarding `SLIF`, many contrasts are supported compared to the reference level (formal declarative). The increase in AIC(61.0) when `SLIF` is removed from the model supports this predictor as an important variable. Words with light `vowel harmony` were responded to more quickly than words with dark/neutral vowel harmony, with strong support from AIC (89.4). In terms of the increase in AIC when a predictor is withheld, the `number of syllables` shows the strongest variable (1191.6) importance of all the parametric terms in the model. The `number of syllables` outperforms the `number of phones` as a predictor (model

Table 7.2: Summary of the partial effects in a GAMM fitted to inverse-transformed(-1000/RT) reaction times for the participants with **Japanese as L1**, final exponent as fixed-effect factor (AIC : 2077). The reference level of `SLIF` is *formal declarative*.

A. parametric coefficients	Estimate	Std. Error	t-value	p-value	Δ(AIC)
(Intercept)	-1.0540	0.0573	-18.3851	< 0.0001	
honorifics=non	-0.1099	0.0155	-7.0709	< 0.0001	46.9
tense=past	0.0771	0.0139	5.5285	< 0.0001	27.0
mood=non	-0.0348	0.0135	-2.5738	0.0101	5.1
SLIF = for+inq	-0.0064	0.0089	-0.7184	0.4725	61.0
SLIF = hon + dec/imp	-0.1288	0.0221	-5.8295	< 0.0001	
SLIF = int + unmarked	-0.1095	0.0189	-5.8007	< 0.0001	
SLIF = plain + dec	-0.0426	0.0097	-4.3754	< 0.0001	
SLIF = plain + inq	-0.0006	0.0103	-0.0627	0.9500	
SLIF = pol + unmarked	-0.1039	0.0188	-5.5361	< 0.0001	
SLIF = unmarked	-0.1041	0.0213	-4.8935	< 0.0001	
vowel harmony = light	-0.0338	0.0036	-9.4830	< 0.0001	89.4
number of syllables	0.0899	0.0026	35.1077	< 0.0001	1191.6
stem-final segment = vowel	-0.0156	0.0036	-4.3602	< 0.0001	15.9
number of exponents	-0.0921	0.0138	-6.6711	< 0.0001	40.8
B. smooth terms	**edf**	**Ref.df**	**F-value**	**p-value**	**Δ(AIC)**
s(lexeme frequency)	3.5768	3.8980	15.4493	< 0.0001	55.4
s(whole-word frequency)	1.0013	1.0026	146.5512	< 0.0001	144.4
s(ending frequency)	3.7784	3.9687	39.9361	< 0.0001	145.0
s(trial)	4.6196	5.1696	9.4087	< 0.0001	8.3
s(trial, subject)	136.7067	188.0000	29.6552	< 0.0001	4809.0

not shown), which is perhaps unsurprising as Hangul is a syllable-based script.

Predictors related to frequency generally exhibit frequency effects across all three components (stem, ending, and whole-word). Figure 7.1 presents the partial effects of these predictors, which main trends showing the expected facilitation for higher frequencies. To judge from the ΔAIC values, the variable importance of `lexeme frequency` lags behind that of `ending frequency` and `whole-word frequency`. For Korean native speakers, on the other hand, the importance of `whole-word frequency` was the lowest. The statistical analysis did not support an `interaction of trial by ending frequency`.

Table 7.3 presents the summary of a GAMM with `Ending` as predictor, instead of `honorifics, past tense, mood, and SLIF`. Whereas for Korean L1 speakers, the ΔAIC for `Ending` was substantially larger than that for `number of syllables` (2142.8 vs. 120.5), for the Japanese L2 speakers of Korean, the variable importance for `number of syllables` was greater than `Ending` (1235.5 vs. 568.8). The random intercepts for `Ending` are presented in Figure 7.2. Overall, the results seem to reflect frequency effects; however, similar to native speakers, some forms exhibit patterns that deviate from the frequency effect. Among the forms that generally show faster reaction times, for-hon-past-dec-16:-(으) 셨습니다 and for-hon-past-inq-19: -(으) 셨습니까 are particularly interesting. These forms, which had slower reaction times among native speakers, displayed faster reaction times in the L1 Japanese group. As with native speakers, certain mood forms showed reaction times that were not aligned with their frequency. For instance, rm50:-(ㄹ/을) 망정 and rm58:-거들랑 were recognised more quickly than other, relatively more frequent mood forms such as mm20:-거든 and mm12:-도록.

Table 7.3: Summary of the partial effects in a GAMM fitted to inverse-transformed (-1000/RT) reaction times for the participants with **Japanese as L1**, with random intercepts for `Ending` (AIC : 1829).

A. parametric coefficients	Estimate	Std. Error	t-value	p-value	Δ(AIC)
(Intercept)	-1.4150	0.0336	-42.0897	< 0.0001	
vowel harmony = light	-0.0360	0.0035	-10.1534	< 0.0001	103.5
number of syllables	0.0994	0.0026	37.6875	< 0.0001	1253.5
stem-final segment = vowel	-0.0176	0.0036	-4.9481	< 0.0001	21.8
B. smooth terms	edf	Ref.df	F-value	p-value	Δ(AIC)
s(lexeme frequency)	3.5828	3.9006	11.6687	< 0.0001	43.3
s(whole-word frequency)	1.0007	1.0014	148.0116	< 0.0001	122.9
s(trial)	4.6672	5.2204	9.2657	< 0.0001	8.6
s(trial, subject)	137.3074	188.0000	30.1211	< 0.0001	4873.2
s(Ending)	36.9853	39.0000	16.4593	< 0.0001	568.8

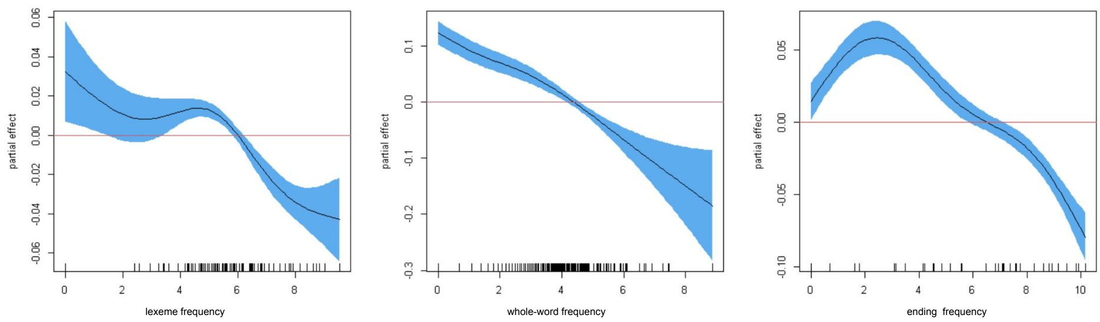


Figure 7.1: Partial effects of lexeme frequency, whole-word frequency, and ending frequency, on the inverse-transformed response latencies of **Japanese** L1 participants. The solid horizontal line represents zero partial effect and the dashed lines represent 95% confidence bands.

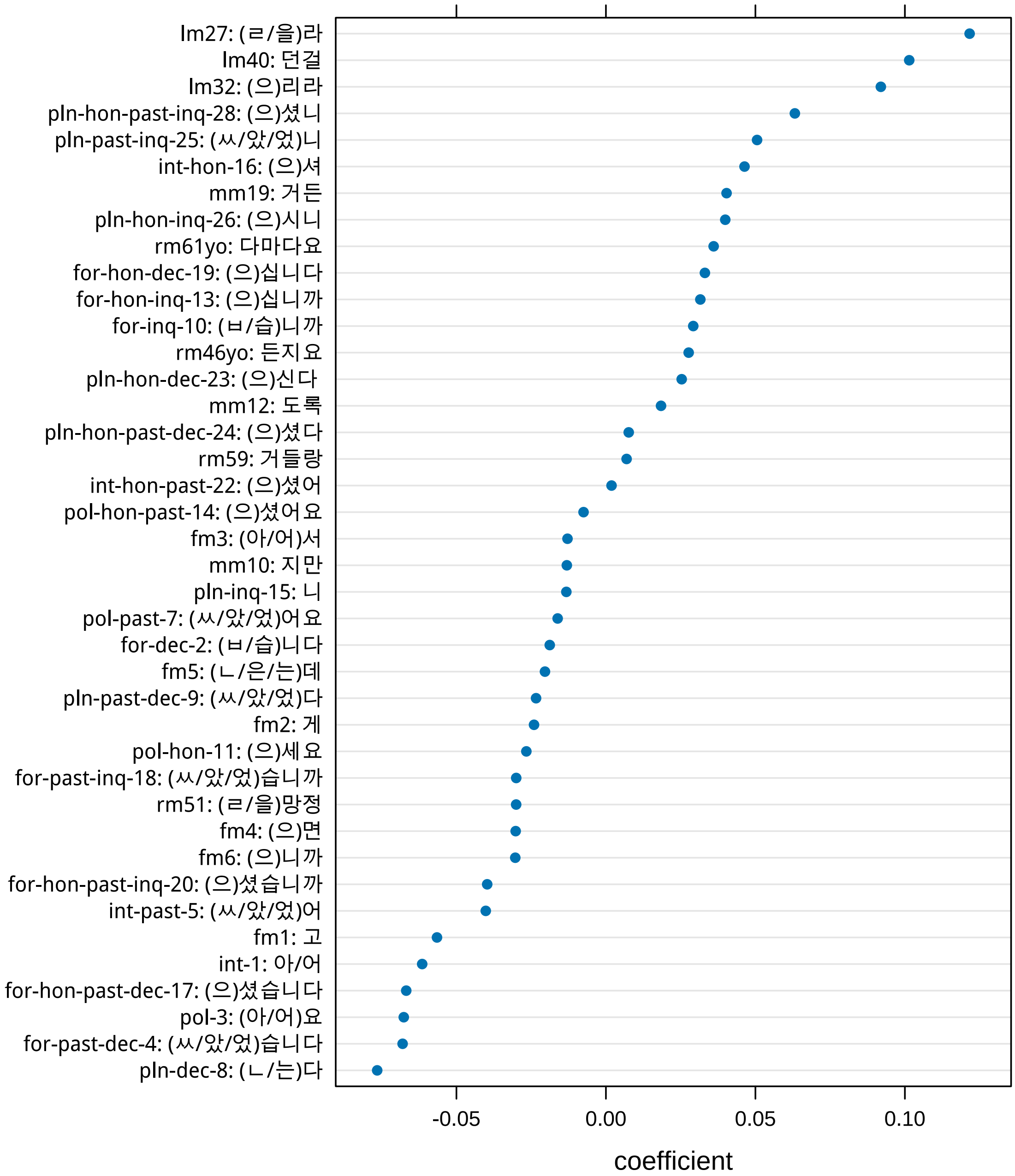


Figure 7.2: Estimates of the random effects of `Ending` for **Japanese** L1 participants. For the abbreviations used, see Tables 5.1 and 5.2.

7.4.2 Participants with Mongolian as L1

The data analysis results from the L1 Mongolian group revealed similar trends in the key variables observed for the L1 Japanese group. The most important variable affecting the lexical decisions for Korean verb forms is word length based on the `number of syllables`, followed by `Ending` and `whole-word frequency`. Tables 7.4 and 7.5 present the model summaries for the analyses with detailed ending components and `Ending`, respectively. Figure 7.4 presents the random intercepts for `Ending`. The endings that elicited the shortest reaction times are `pln-past-dec-9`:-(ㅆ/았/었) 다 and `pol-hon-11`:-(으) 세요, whereas `lm40:-던걸` and `pln-hon-past-inq-28`:-(으) 셨니 elicited the longest reaction times. With respect to `SLIF`, `formal inquisitive` verbs elicited the longest reaction times (compared to the `formal declarative` baseline). The effects of `honorifics` and `past tense` were similar to those observed for the Japanese and Korean participants. The statistical analysis did not support an `interaction of trial by ending frequency`.

Focusing on the ΔAIC values for the frequency effects, `ending frequency` had the greatest impact, followed by `whole-word frequency`, and at a distance `stem frequency`. This pattern is consistent with the results observed in the L1 Japanese group. The most notable difference with the Korean native speakers is the low importance of `lexeme frequency` for the Mongolian L1 participants. Figure 7.3 presents the partial effects of the three frequency measures. Compared to both Korean and Japanese speakers, the Mongolian curves approximate linearity to a much greater extent.

Overall, variable importances (evaluated with ΔAIC) are greatest for the `number of syllables` (740/755), followed by `Ending` (423). `Whole-word frequency` (101/115), `vowel harmony` (90/97), `ending frequency` (90) and `SLIF` (80) have intermediate variable importance. The lowest scores are for `lexeme frequency` (25/40), `stem-final segment` (13/18) and `number of exponents` (9).

The random intercepts for `Ending` are presented in Figure 7.4. Interestingly, in the L1 Mongolian group, `pln-past-dec-9`: -(ㅆ/았/었) 다 was the most quickly recognised form. This pattern is similar to that of native speakers and the Chinese group, whereas the Japanese and Vietnamese groups showed slightly slower reaction times for this form. Another notable finding is that the `pol-hon-11`: -(으) 세요 form was recognised more quickly in the Mongolian group compared to other groups. Similarly, `for-hon-past-dec-17`: -(으) 셨습니다 and `for-hon-past-inq-20`: -(으) 셨습니까 were processed faster than by native speakers, aligning with the pattern observed in the Japanese group. Despite being a highly frequent form, `int-1`: -아/어 exhibited relatively slower reaction times, similar to other groups. This may be due to the difficulty in processing forms that end with a vowel in a short form. Mood forms generally reflected frequency effects, but the effects of particularly salient forms, such as `rm59`: -거들랑, also seemed to play a role.

Table 7.4: Summary of the partial effects in a GAMM fitted to inverse-transformed(-1000/RT) reaction times for the participants with **Mongolian as L1**, final exponent as fixed-effect factor(AIC : 7067). The reference level of `SLIF` is *formal declarative*.

A. parametric coefficients	Estimate	Std. Error	t-value	p-value	Δ(AIC)
(Intercept)	-1.2947	0.0686	-18.8722	< 0.0001	
honorifics=non	-0.0776	0.0179	-4.3356	< 0.0001	17.2
tense=past	0.0374	0.0159	2.3551	0.0185	4.2
SLIF = for+inq	0.0296	0.0102	2.9167	0.0035	95.1
SLIF = hon + dec/imp	-0.1212	0.0254	-4.7680	< 0.0001	
SLIF = int + unmarked	-0.0304	0.0218	-1.3922	0.1639	
SLIF = plain + dec	-0.0337	0.0112	-3.0035	0.0027	
SLIF = plain + inq	0.0056	0.0119	0.4686	0.6394	
SLIF = pol + unmarked	-0.0603	0.0212	-2.8413	0.0045	
SLIF = unmarked	-0.0140	0.0207	-0.6762	0.4989	
vowel harmony = light	-0.0391	0.0042	-9.3853	< 0.0001	88.2
number of syllables	0.0825	0.0029	28.0457	< 0.0001	748.6
stem-final segment = vowel	-0.0163	0.0042	-3.9152	0.0001	14.1
number of exponents	-0.0503	0.0159	-3.1710	0.0015	9.2
B. smooth terms	edf	Ref.df	F-value	p-value	Δ(AIC)
s(lexeme frequency)	2.2651	2.7967	15.5153	< 0.0001	35.1
s(whole-word frequency)	1.8964	2.3062	60.0220	< 0.0001	117.1
s(ending frequency)	1.9544	2.4204	64.0866	< 0.0001	135.7
s(trial)	4.4824	4.9560	3.0850	0.0083	5.6
s(trial, subject)	140.1849	179.0000	42.1505	< 0.0001	6244.9

Table 7.5: Summary of the partial effects in a GAMM fitted to inverse-transformed (-1000/RT) reaction times for the participants with **Mongolian as L1**, with random intercepts for `Ending` (AIC : 6984).

A. parametric coefficients	Estimate	Std. Error	t-value	p-value	Δ(AIC)
(Intercept)	-1.4839	0.0457	-32.4610	< 0.0001	
vowel harmony = light	-0.0409	0.0042	-9.8259	< 0.0001	96.9
number of syllables	0.0911	0.0031	29.7741	< 0.0001	755.0
stem-final segment = vowel	-0.0181	0.0042	-4.3447	< 0.0001	17.7
B. smooth terms	edf	Ref.df	F-value	p-value	Δ(AIC)
s(lexeme frequency)	2.2168	2.7439	10.3478	< 0.0001	25.0
s(whole-word frequency)	1.9075	2.3157	59.7320	< 0.0001	100.5
s(trial)	4.6378	5.1299	3.1162	0.0102	5.8
s(trial, subject)	140.3605	179.0000	42.4749	< 0.0001	6287.0
s(Ending)	36.3639	39.0000	12.4187	< 0.0001	423.2

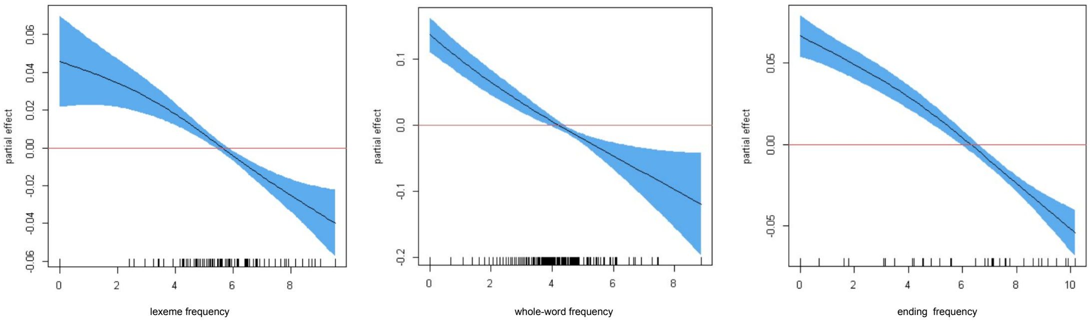


Figure 7.3: Partial effects of lexeme frequency, whole-word frequency, and ending frequency, on the inverse-transformed response latencies of **Mongolian** L1 participants. The solid horizontal line represents zero partial effect and the dashed lines represent 95% confidence bands.

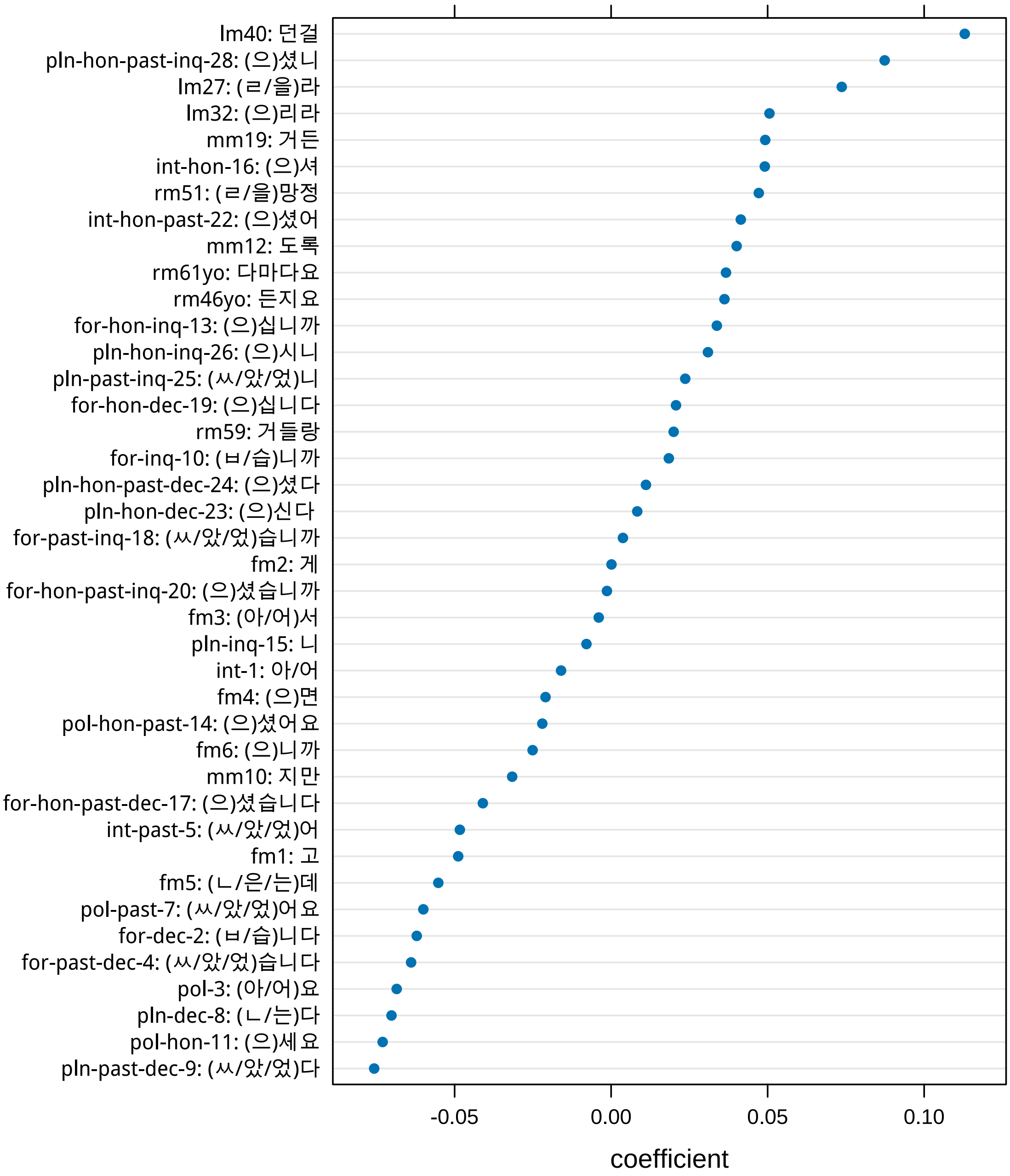


Figure 7.4: Estimates of the random effects of `Ending` for **Mongolian** L1 participants. For the abbreviations used, see Tables 5.1 and 5.2.

7.4.3 Participants with Chinese as L1

Similar to other non-native groups mentioned earlier (L1 Japanese and Mongolian), the L1 Chinese group also struggled with longer words, i.e., words with more syllables (and more Hangul characters): the `number of syllables` has the highest variable importance. When examining the summary of the statistical model in Table 7.6, the variable importance (evaluated with ΔAIC) of the `number of syllables` (532) is followed by that of the effect of `ending frequency` (252). This contrasts with the L1 Japanese and Mongolian groups, where `whole-word frequency` was the second most influential factor after the syllable count. However, in the L1 Chinese group, `whole-word frequency` (88) showed a much smaller AIC difference compared to `ending frequency` (252), yet like other non-native groups, `lexeme frequency` (69) had an even smaller variable importance. Additionally, the absence of honorific forms and `past tense` (27) facilitated response times. `Vowel harmony` (Δ AIC 89) also affected response times with similar variable importance as that of `honorifics` (65) and the `number of exponents` (51).

Figure 7.7 shows the partial effects of the three frequency measures for the L1 Chinese group. For all three measures, as expected, response times decrease as frequency is increased. The most influential variable for the L1 Chinese group was verb `ending frequency` (252), variable importances for the other two measures were much smaller, at 69 for `lexeme frequency`, and at 88 for `whole-word frequency`. This constellation of variable importances differs from the L1 Japanese group, where `whole-word frequency` and `ending frequency` had a similar impact, and from the L1 Mongolian group, where `whole-word frequency` had the greatest impact.

Table 7.7 below presents the results of the statistical model that includes `Ending` but excludes verb ending-related variables. Based on the AIC difference, `Ending` (619) and the `number of syllables` (493) emerged as the most influential lexical variables. Following these, `vowel harmony` (91), `whole-word frequency` (78), and `lexeme frequency` (61) were well-supported by the GAMM. The fact that the variable `Ending` has a greater influence than the `number of syllables` contrasts with previous findings where the syllable count was the most influential variable across all other non-native speaker groups.

The random intercepts for `Ending` are presented in Figure 7.6. The fastest recognised ending type was `for-past-dec-4`:-(ㅆ/았/었) 습니다, which showed a pattern similar to the Japanese and Vietnamese groups. Like other non-native speakers, `for-hon-past-dec-17`:-(으) 셨습니다 was processed more quickly than by native speakers. On the other hand, `pol-hon-11`:-(으) 세요 was recognised more slowly compared to other non-native groups, displaying a pattern similar to that of native speakers. Overall, both mood and non-mood forms exhibited frequency effects. Among sentence-final (i.e., non-mood) forms, the slowest recognised form was `int-hon-16`:-(으) 서, which followed the same pattern as the Vietnamese group. The next slowest form was `pln-hon-past-inq-28`:-(으) 셨니, which showed the slowest reaction times in the native speaker, Japanese, and Mongolian groups. In summary, these two sentence-final forms elicited the slowest reaction times across all groups.

Table 7.6: Summary of the partial effects in a GAMM fitted to inverse-transformed(-1000/RT) reaction times for the participants with **Chinese as L1**, final exponent as fixed-effect factor(AIC : 9040). The reference level of `SLIF` is *formal declarative*.

A. parametric coefficients	Estimate	Std. Error	t-value	p-value	
(Intercept)	-1.0441	0.0781	-13.3693	< 0.0001	Δ(AIC)
honorifics=non	-0.1523	0.0187	-8.1326	< 0.0001	64.7
tense=past	0.0851	0.0166	5.1163	< 0.0001	26.8
SLIF = for+inq	-0.0102	0.0107	-0.9582	0.3380	27.1
SLIF = hon + dec/imp	-0.1037	0.0267	-3.8832	0.0001	
SLIF = int + unmarked	-0.0903	0.0229	-3.9531	0.0001	
SLIF = plain + dec	-0.0374	0.0117	-3.1840	0.0015	
SLIF = plain + inq	-0.0092	0.0125	-0.7362	0.4616	
SLIF = pol + unmarked	-0.1045	0.0222	-4.7005	< 0.0001	
SLIF = unmarked	-0.0790	0.0217	-3.6372	0.0003	
vowel harmony = light	-0.0413	0.0043	-9.5200	< 0.0001	88.9
number of syllables	0.0724	0.0031	23.6914	< 0.0001	531.5
stem-final segment = vowel	-0.1167	0.0166	-7.0211	< 0.0001	51.1
B. smooth terms	edf	Ref.df	F-value	p-value	Δ(AIC)
s(lexeme frequency)	1.9627	2.4533	27.6263	< 0.0001	69.2
s(whole-word frequency)	1.0003	1.0007	101.1732	< 0.0001	87.5
s(ending frequency)	2.3627	2.8743	104.0855	< 0.0001	252.0
s(trial)	3.5097	3.8449	5.5852	0.0002	5.7
s(trial, suject)	141.4369	179.0000	59.3149	< 0.0001	8366.4

Table 7.7: Summary of the partial effects in a GAMM fitted to inverse-transformed (-1000/RT) reaction times for the participants with **Chinese as L1**, with random intercepts for `Ending` (AIC : 8985).

A. parametric coefficients	Estimate	Std. Error	t-value	p-value	Δ(AIC)
(Intercept)	-1.4233	0.0571	-24.9055	< 0.0001	
vowel harmony = light	-0.0418	0.0043	-9.6270	< 0.0001	90.8
number of syllables	0.0758	0.0032	23.5320	< 0.0001	493.4
B. smooth terms	edf	Ref.df	F-value	p-value	Δ(AIC)
s(lexeme frequency)	1.9002	2.3787	24.2533	< 0.0001	60.8
s(whole-word frequency)	1.0003	1.0005	102.8890	< 0.0001	77.8
s(trial)	3.5294	3.8677	5.6594	0.0002	6.2
s(trial, subject)	141.4038	179.0000	59.6739	< 0.0001	8403.3
s(Ending)	37.2102	39.0000	18.5348	< 0.0001	618.6

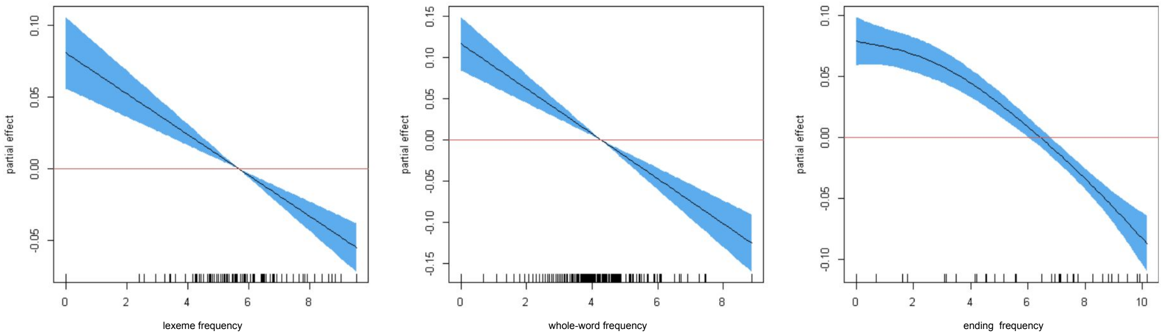


Figure 7.5: Partial effects of lexeme frequency, whole-word frequency, and ending frequency, on the inverse-transformed response latencies of Chinese L1 participants. The solid horizontal line represents zero partial effect and the dashed lines represent 95% confidence bands.

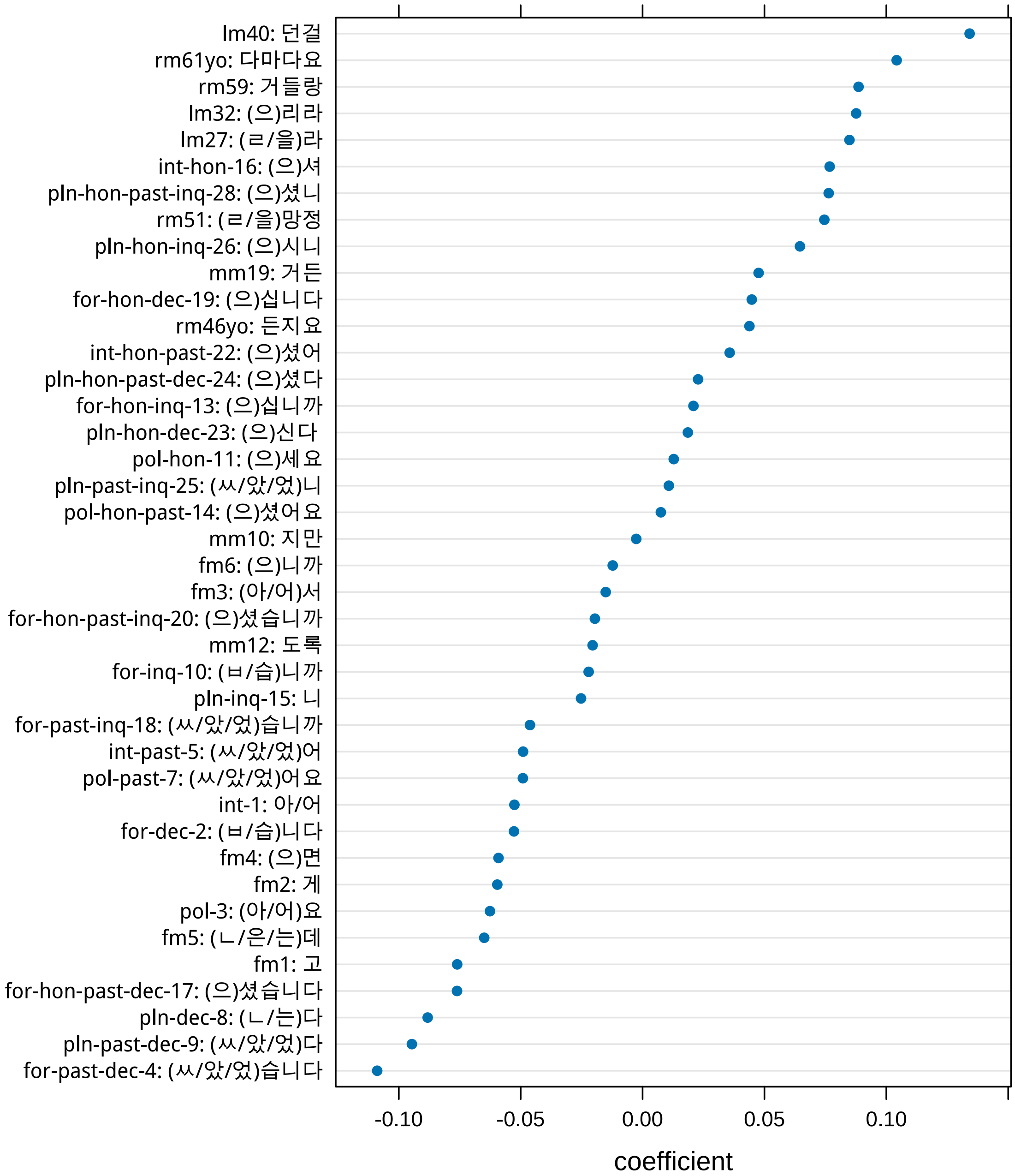


Figure 7.6: Estimates of the random effects of `Ending` for **Chinese** L1 participants. For the abbreviations used, see Tables 5.1 and 5.2.

7.4.4 Participants with Vietnamese as L1

Table 7.8 presents the summary of the GAMM model fitted to the transformed response latencies and shows the results of L1 Vietnamese group. The `number of syllables` has the greatest variable importance by far (1093): longer words were more difficult to respond to. The next most important predictors are `SLIF` (102) and `vowel harmony` (99). Non-`mood` forms elicited somewhat longer reaction times, but the variable importance of this predictor (7) was the lowest of all among the parametric predictors.

Figure 7.7 shows the effects of the three frequency measures. As expected, response times decrease for words with higher frequencies. The frequency effects for lexeme and whole-word are linear, the frequency effect for `ending frequency` is nearly linear, but levels off slightly for the lowest `ending frequency`. For Vietnamese, the `interaction of trial by ending frequency` received good support. This interaction is similar to that observed for native speakers of Korean, but weaker: by the end of the experiment, the facilitatory effect of `ending frequency` is somewhat attenuated.

The summary of a GAMM with `Ending` as predictor, replacing the standard morphological predictors, but without a main effect for ending frequency (to avoid collinearity) is available in Table 7.9. The random intercepts for `Ending` are presented in Figure 7.8. The fastest recognised ending type is `for-past-dec-4`:-(ㅆ/았/었)습니다, the ending type that was also responded to most quickly by the Chinese L1 group. Unlike for native speakers, but apparently characteristic for all non-native groups, `for-hon-past -dec-17`:-(으) 셨습니다 and `for-hon-past-inq-20`:-(으) 셨습니까 also elicited short response latencies. This is particularly surprising because these are relatively infrequent forms. The word type that elicited the longest reaction times was the low-frequency mood type `lm27`:-(ㄹ/을) 라, an ending type that also challenged the Japanese L1 participants. Other ending types that all L2 learners found difficult are `rm59`:-거들랑 and `lm40`:-던걸.

Although an effect of `ending frequency` is well-supported (see Table 7.8), some fairly frequent endings show opposite effects. For instance, words with the ending `fm3`: -(아/어) 서, despite being a highly frequent mood form, were recognised more slowly than words with many other ending types. Additionally, words with the low-frequency ending `rm46yo`:-든지요 and words with the highly frequent mood form `fm5`:-(ㄴ/은/는) 데 exhibited similar reaction times. These results suggest that beyond frequency effects, other factors, such as the distinctiveness of form shapes or semantic ambiguity, may affect the recognition of Korean verb endings.

Table 7.8: Summary of the partial effects in a GAMM fitted to inverse-transformed(-1000/RT) reaction times for the participants with **Vietnamese as L1**, final exponent as fixed-effect factor(AIC : -17.4). The reference level of `SLIF` is *formal declarative*.

A. parametric coefficients	Estimate	Std. Error	t-value	p-value	Δ(AIC)
(Intercept)	-1.0917	0.0567	-19.2507	< 0.0001	
honorifics=non	-0.1088	0.0150	-7.2398	< 0.0001	47.3
tense=past	0.0594	0.0134	4.4430	< 0.0001	19.5
mood=non	0.0407	0.0130	3.1342	0.0017	6.8
SLIF = for+inq	0.0021	0.0085	0.2526	0.8006	101.5
SLIF = hon + dec/imp	-0.1115	0.0213	-5.2426	< 0.0001	
SLIF = int + unmarked	-0.0614	0.0182	-3.3721	0.0007	
SLIF = plain + dec	-0.0091	0.0093	-0.9751	0.3295	
SLIF = plain + inq	0.0129	0.0100	1.2930	0.1960	
SLIF = pol + unmarked	-0.0785	0.0181	-4.3344	< 0.0001	
SLIF = unmarked	-0.0008	0.0206	-0.0372	0.9703	
vowel harmony = light	-0.0345	0.0034	-9.9989	< 0.0001	98.8
number of syllables	0.0827	0.0024	34.0339	< 0.0001	1092.5
number of exponents	-0.0887	0.0132	-6.6937	< 0.0001	47.9
B. smooth terms	edf	Ref.df	F-value	p-value	Δ(AIC)
s(lexeme frequency)	1.0059	1.0118	40.1741	< 0.0001	37.6
s(whole-word frequency)	1.0003	1.0006	213.5006	< 0.0001	189.0
s(ending frequency)	2.4563	2.9778	69.9706	< 0.0001	192.0
s(trial)	6.5373	7.1350	6.8695	< 0.0001	12.0
ti(trial, ending frequency)	4.1549	5.7122	2.9786	0.0081	10.7
s(trial, subject)	139.8022	179.0000	35.9995	< 0.0001	5448.5

Table 7.9: Summary of the partial effects in a GAMM fitted to inverse-transformed (-1000/RT) reaction times for the participants with **Vietnamese as L1**, with random intercepts for `Ending`(AIC : -271.4).

A. parametric coefficients	Estimate	Std. Error	t-value	p-value	Δ(AIC)
(Intercept)	-1.3433	0.0352	-38.1917	< 0.0001	
vowel harmony = light	-0.0363	0.0034	-10.5769	< 0.0001	105.5
number of syllables	0.0912	0.0026	35.5351	< 0.0001	1157.7
B. smooth terms	edf	Ref.df	F-value	p-value	Δ(AIC)
s(lexeme frequency)	1.4162	1.7263	15.1308	< 0.0001	18.4
s(whole-word frequency)	1.0002	1.0004	192.5285	< 0.0001	155.9
s(trial)	5.9809	6.5773	6.6089	< 0.0001	8.5
ti(trial, ending frequency)	4.0924	5.6465	3.0380	0.0076	11.5
s(trial, subject)	140.4476	179.0000	36.6348	< 0.0001	5521.9
s(Ending)	37.5337	39.0000	60.9317	< 0.0001	840.0

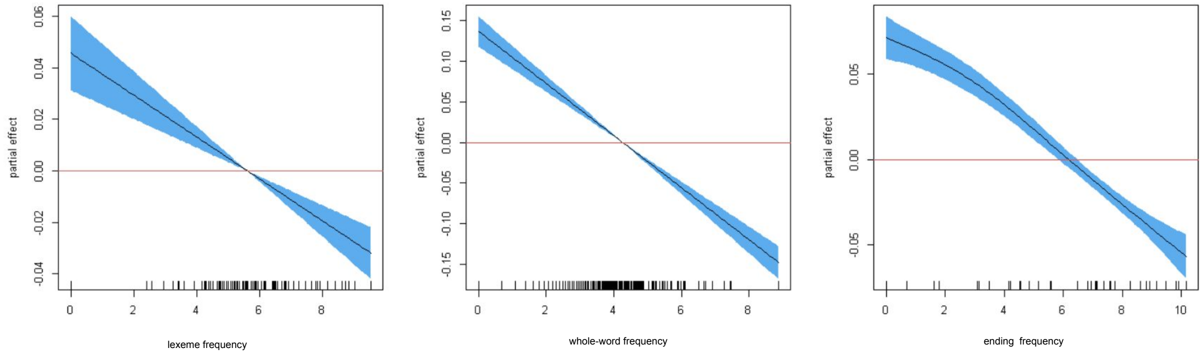


Figure 7.7: Partial effects of lexeme frequency, whole-word frequency, and ending frequency, on the inverse-transformed response latencies of Vietnamese L1 participants. The solid horizontal line represents zero partial effect and the dashed lines represent 95% confidence bands.

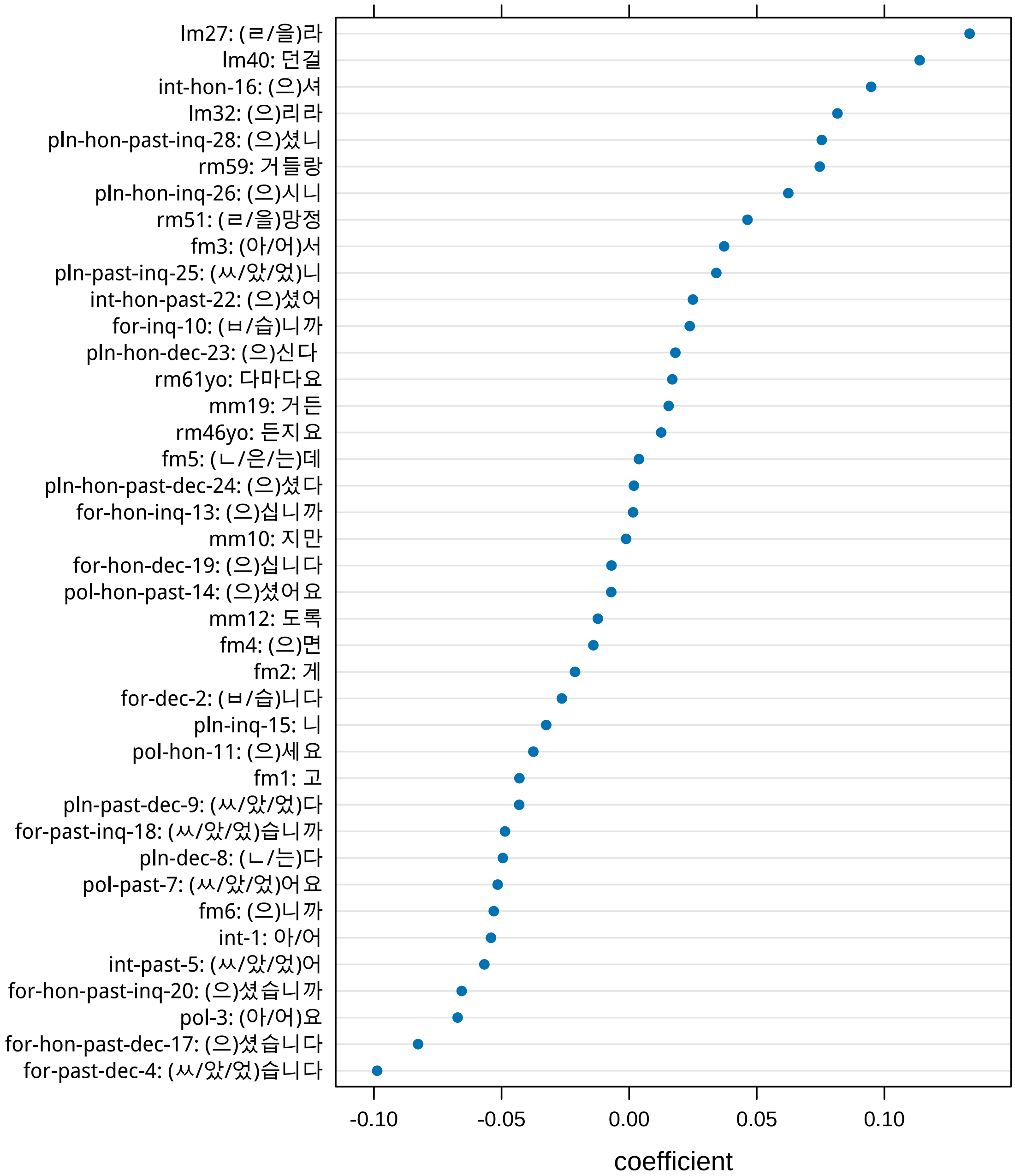


Figure 7.8: Estimates of the random effects of `Ending` for **Vietnamese** L1 participants. For the abbreviations used, see Tables 5.1 and 5.2.

Figure 7.9 provides an overview of the relative importance of the predictors, for each speaker group, including the native speakers, assessed by means of the increase in AIC when a predictor is withheld from the model fitted to the data of a given group. Variable importances for the first 11 predictors in this figure are taken from models that do not include `Ending` as the predictor. The final rows in the dotplot panels present the effect of `Ending` in the models that include the `Ending`, but exclude the parameters for `past tense, mood, honorifics, and SLIF`.

For native speakers, the predictor with the greatest variable importance was by far the `Ending`(marked in lower case `ending` in the figure). The variable importance of the `Ending` compared to other variables for the non-native speakers was strongest for the Chinese group(i.e., `Ending` is bigger than the `number of syllables`), followed by the Japanese, with the Mongolian and Vietnamese group showing the similar impacts. For the L2 speakers, the `number of syllables` was the most important predictor, except for the Chinese L2 group, where the `Ending` was slightly more important. But for the other groups, `Ending` was the second-most important predictor.

Another notable distinction is that native speakers are more strongly tuned into `lexeme frequency`, compared to all non-native L2 groups. Additionally, the final ending form `SLIF` (speech level + illocutionary force) had a pronounced impact on native speakers but had a more modest effect for non-native groups.

Overall, the variation in variable importances appears to be much greater for Korean natives than for the non-native groups, which suggests that, in spite of being very advanced non-native speakers, they have not tuned into the language in the same way as the native speakers.

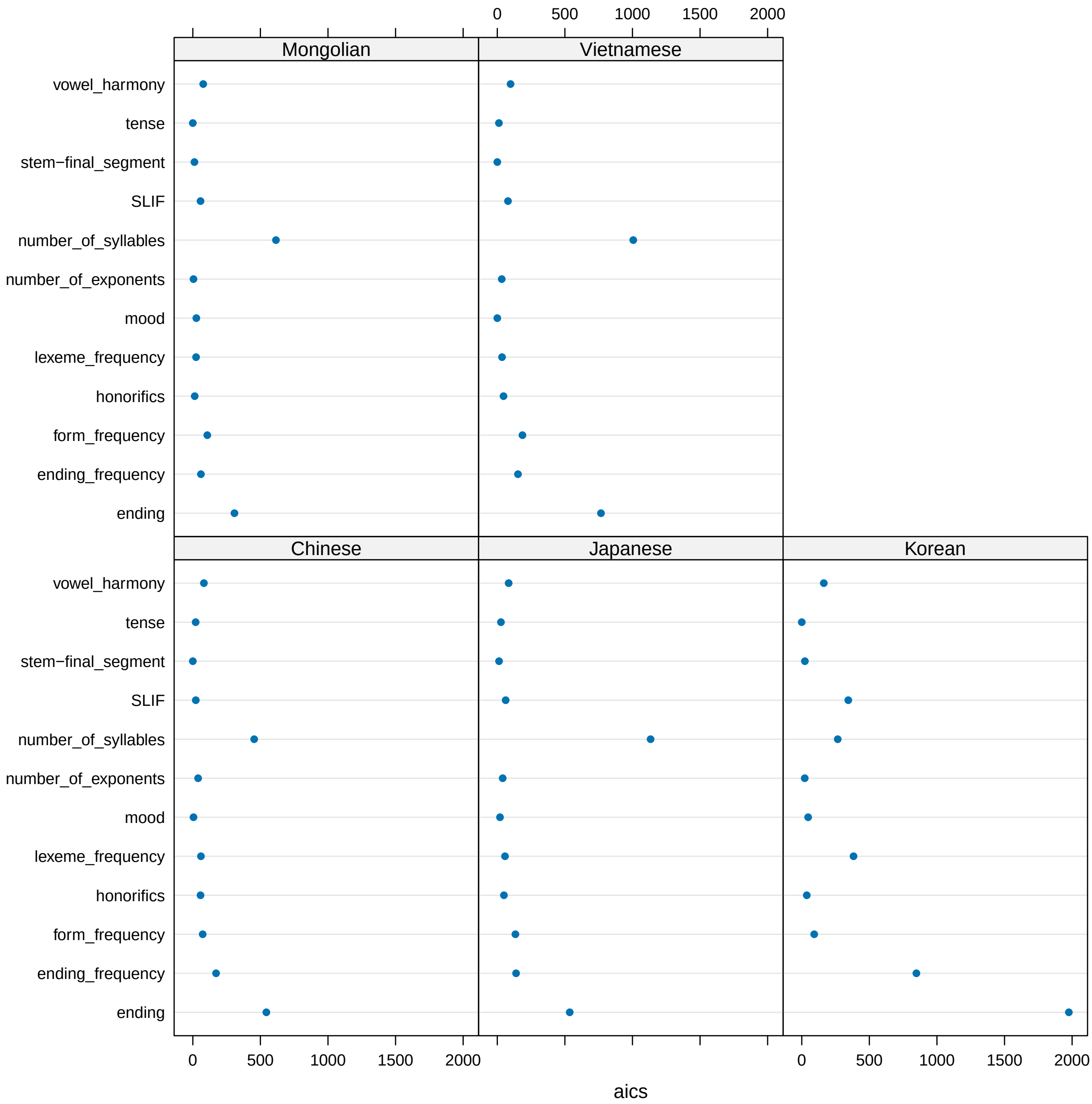


Figure 7.9: Variable importance of selected predictors, gauged with the increase in AIC when a predictor is removed from a language-specific GAMM. The predictor `Ending` is marked as 'ending' in this Figure.

7.5 Results from Nonwords

In this study, nonwords were manipulated with morphological, semantic, and phonological effects in mind. Along with this objective, verb lexemes and endings were manipulated in various ways, distinguishing cases where only one component (lexeme or ending) was altered from cases where both were modified. This allowed for an examination of differences in the recognition of stems and endings. To this end, the statistical model was simplified to focus on the results of the nonword condition (i.e., `nwCondition`), which provides the most critical information in the analysis of nonword data. In doing so, word length (`number of syllables`), which had the strongest effect on Korean verb recognition among non-native speakers, was retained, while other less transparent manipulated variables such as ending components, (`honorifics, past tense, vowel harmony, number of exponents,` and `stem-final segment`), were excluded. The statistical model, therefore, consisted of `lexeme frequency, ending frequency and nwCondition`. As expected, across all non-native speaker groups, the most influential variable, aside from individual differences (i.e., `trial and subject`), was `nwCondition`, with word length (i.e., `number of syllables`) also playing a significant role. The following tables 7.10, 7.11, 7.12, 7.13 present the statistical results for each group. In all groups, a similar pattern emerged: excluding individual differences, the `nwCondition` had the greatest impact, followed by `number of syllables`, as evidenced by AIC differences.

Table 7.10: Summary of the partial effects in GAMMs fitted to inverse-transformed(-1000/RT) for reaction times in nonword forms (**L1 Japanese group**, AIC : 5034)

A. parametric coefficients	Estimate	Std. Error	t-value	p-value	Δ(AIC)
(Intercept)	-1.1938	0.0390	-30.6388	< 0.0001	
number of syllables	0.0611	0.0020	30.0078	< 0.0001	880
B. smooth terms	**edf**	**Ref.df**	**F-value**	**p-value**	**Δ(AIC)**
s(lexeme frequency)	3.8894	3.9923	25.1169	< 0.0001	91
s(ending frequency)	3.8309	3.9821	17.3668	< 0.0001	61
s(nwCondition)	21.4574	22.0000	101.5936	< 0.0001	1630
s(trial, subject)	151.6856	188.0000	41.6993	< 0.0001	6255

Table 7.11: Summary of the partial effects in GAMMs fitted to inverse-transformed(-1000/RT) for reaction times in nonword forms (**L1 Mongolian group**, AIC : 8809)

A. parametric coefficients	Estimate	Std. Error	t-value	p-value	Δ(AIC)
(Intercept)	-1.2923	0.0470	-27.4698	< 0.0001	
number of syllables	0.0619	0.0023	26.8263	< 0.0001	702
B. smooth terms	edf	Ref.df	F-value	p-value	Δ(AIC)
s(lexeme frequency)	3.8063	3.9771	18.3364	< 0.0001	63
s(ending frequency)	3.6373	3.9213	7.6544	< 0.0001	22
s(nwCondition)	20.9813	22.0000	42.0729	< 0.0001	900
s(trial, subject)	149.3279	179.0000	46.0215	< 0.0001	6688

Table 7.12: Summary of the partial effects in GAMMs fitted to inverse-transformed(-1000/RT) for reaction times in nonword forms (**L1 Chinese group**, AIC : 8741)

A. parametric coefficients	Estimate	Std. Error	t-value	p-value	Δ(AIC)
(Intercept)	-1.1951	0.0496	-24.1118	< 0.0001	
number of syllables	0.0471	0.0023	20.3638	< 0.0001	401
B. smooth terms	edf	Ref.df	F-value	p-value	Δ(AIC)
s(lexeme frequency)	3.5845	3.9010	8.2449	< 0.0001	30
s(ending frequency)	3.3249	3.7484	6.1697	0.0005	15
s(nwCondition)	20.9343	22.0000	36.0369	< 0.0001	766
s(trial, subject)	151.4908	179.0000	52.3529	< 0.0001	7404

Table 7.13: Summary of the partial effects in GAMMs fitted to inverse-transformed(-1000/RT) for reaction times in nonword forms (**L1 Vietnamese group**, AIC : -277)

A. parametric coefficients	Estimate	Std. Error	t-value	p-value	Δ(AIC)
(Intercept)	-1.0804	0.0334	-32.2997	< 0.0001	
number of syllables	0.0550	0.0019	29.6325	< 0.0001	860
B. smooth terms	edf	Ref.df	F-value	p-value	Δ(AIC)
s(lexeme frequency)	3.8600	3.9878	14.8714	< 0.0001	51
s(ending frequency)	3.7830	3.9708	9.0754	< 0.0001	30
s(nwCondition)	21.2322	22.0000	65.1858	< 0.0001	1211
s(trial, subject)	146.7907	179.0000	39.8479	< 0.0001	5715

Figure 7.10, 7.11, 7.12, 7.13 presents the distributions of response times broken down by nonword type, for each of the non-native groups. Figure 7.14 combines the previous plots onto a single page to allow for direct comparison.

Nonwords are manipulated such as forms with phonologically similar effects (`pho`), allomorphic substitutions (`alloSw`), syllable-reversed endings (`sswp`), the combination of past tense and mood with polite speech ending -'-요'/yo/(`Tmy`) and frequent mood manipulations (`fm`) were found to slow response times. In contrast, nonwords with novel syllables (`sag`) or full stems/endings (`fag`) being unlike real words, elicited faster nonword response times.

In the analysis of lexeme (L) and ending (E) manipulations under different conditions, native Korean speakers, as well as the L1 Japanese and L1 Mongolian groups, demonstrated slower response times for ending manipulations (`phoE/Ey`) than for lexeme manipulations (`phoL/Ly`) under phonological similarity conditions (`pho`). However, in the L1 Chinese and L1 Vietnamese groups, lexeme manipulations generally resulted in slower response times than ending manipulations. Across all groups, full lexeme-ending manipulations (`phoLE`) consistently led to the fastest response times under similar phonological conditions. Additionally, the slower response times for ending manipulations compared to lexeme manipulations persisted in the non-existent syllable condition (`sag`) for all non-native groups. On the whole, the generally longer response times in many cases involving manipulation of endings indicate that participants tend to experience more difficulty when processing endings.

The effect of the presence or absence of -요/yo/ was notable in most conditions, with the presence of -요 generally facilitating faster response times. The facilitative -요 effect exhibited similar patterns across most groups[25]. In this regard, it is noteworthy that `sswpEy` and `sagLy` do not exhibit a facilitation effect of -요 consistently across

[25] There is an exception that L1 Japanese and L1 Mongolian groups in the `phoEy` condition did not show the facilitative effect observed in other groups.

all groups including the group of native speakers as well. It is an intriguing finding that both native and non-native speakers showed similar response patterns in the `sswpEy` and `sagLy` nonword conditions when combined with the -요 form. This may suggest that even in visual lexical decision tasks, participants internally rehearse the phonological form of the word during recognition. The familiar ending -요 may create a word-like illusion, which could conflict with the altered forms in the `sswpEy` condition and result in slower responses. Similarly, in the `sagLy` condition, when monosyllabic stems were distorted, the remaining segment combined with -요 may have formed a longer, word-like cue that interfered with the judgment process and further delayed response times.

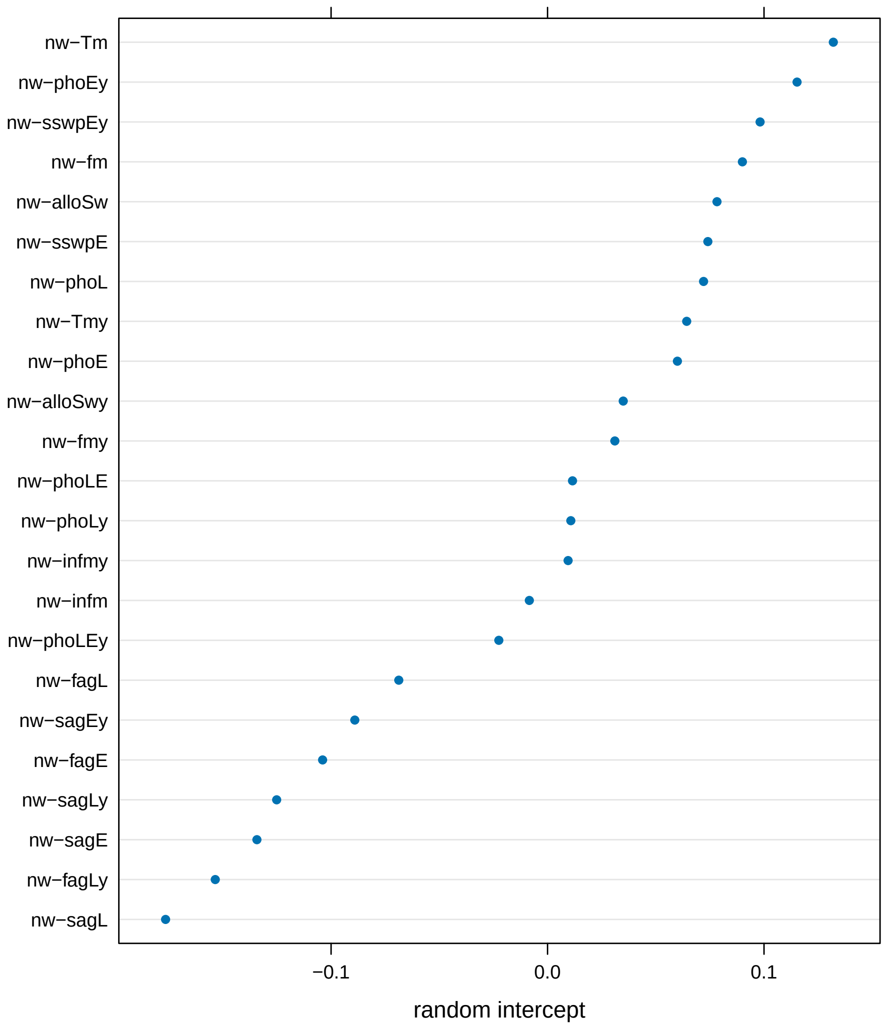


Figure 7.10: Random intercepts for nonword type (`nwCondition`) in L1 Japanese group.

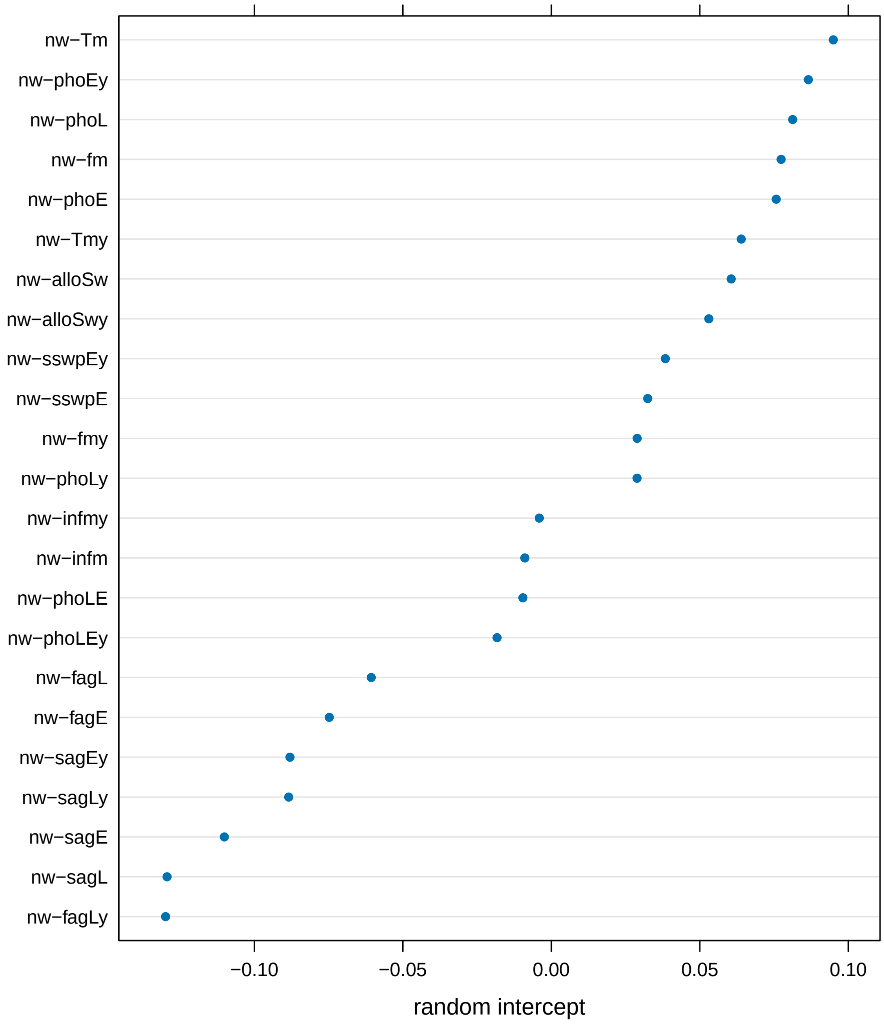


Figure 7.11: Random intercepts for nonword type (`nwCondition`) in L1 Mongolian group.

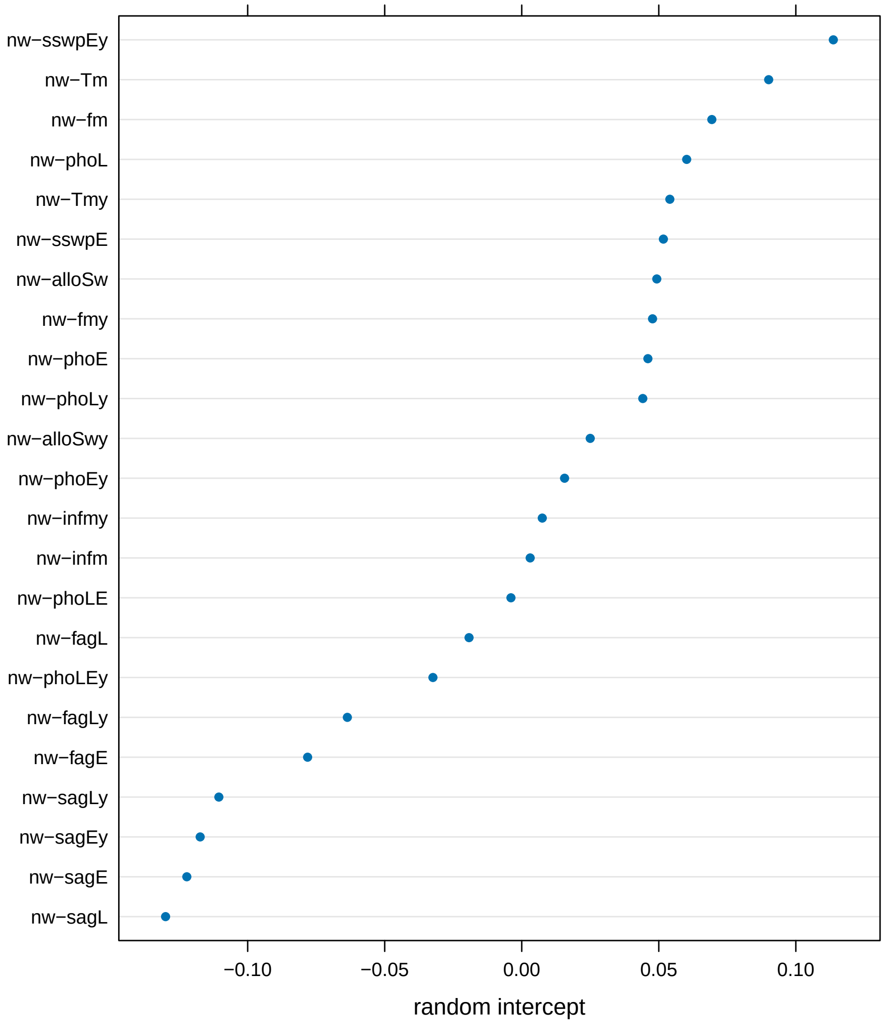


Figure 7.12: Random intercepts for nonword type (`nwCondition`) in L1 Chinese group.

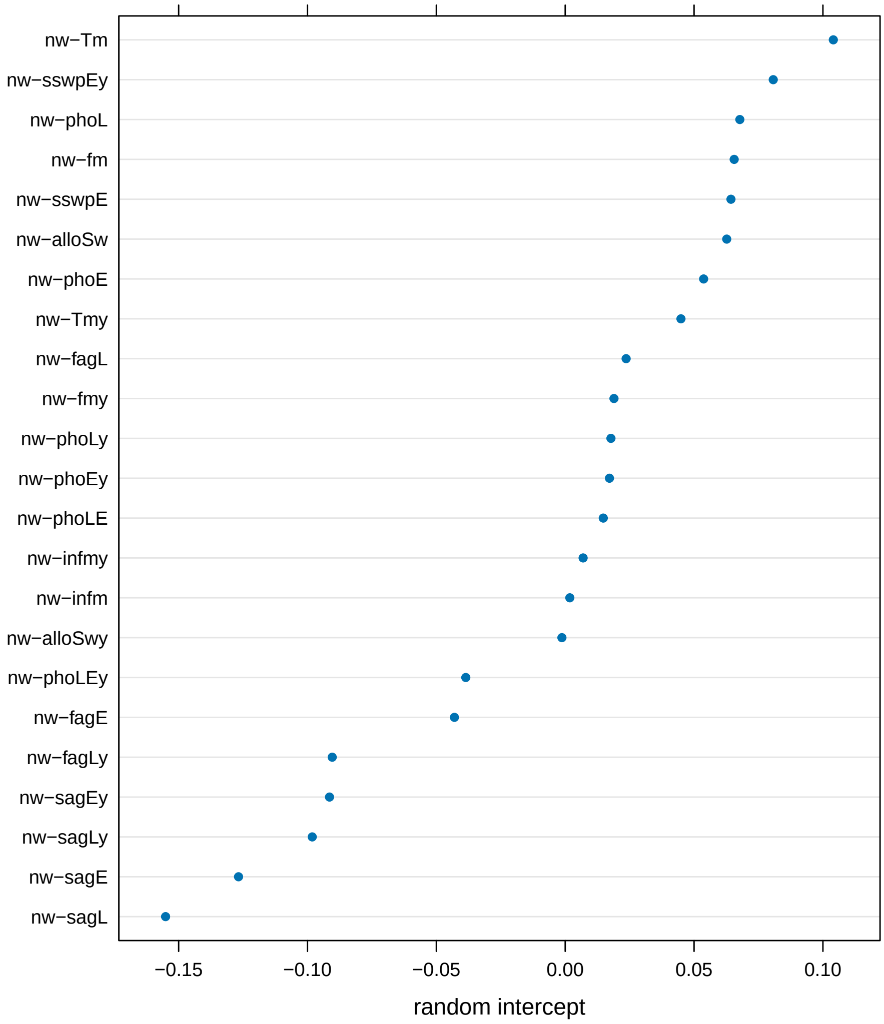


Figure 7.13: Random intercepts for nonword type (`nwCondition`) in L1 Vietnamese group.

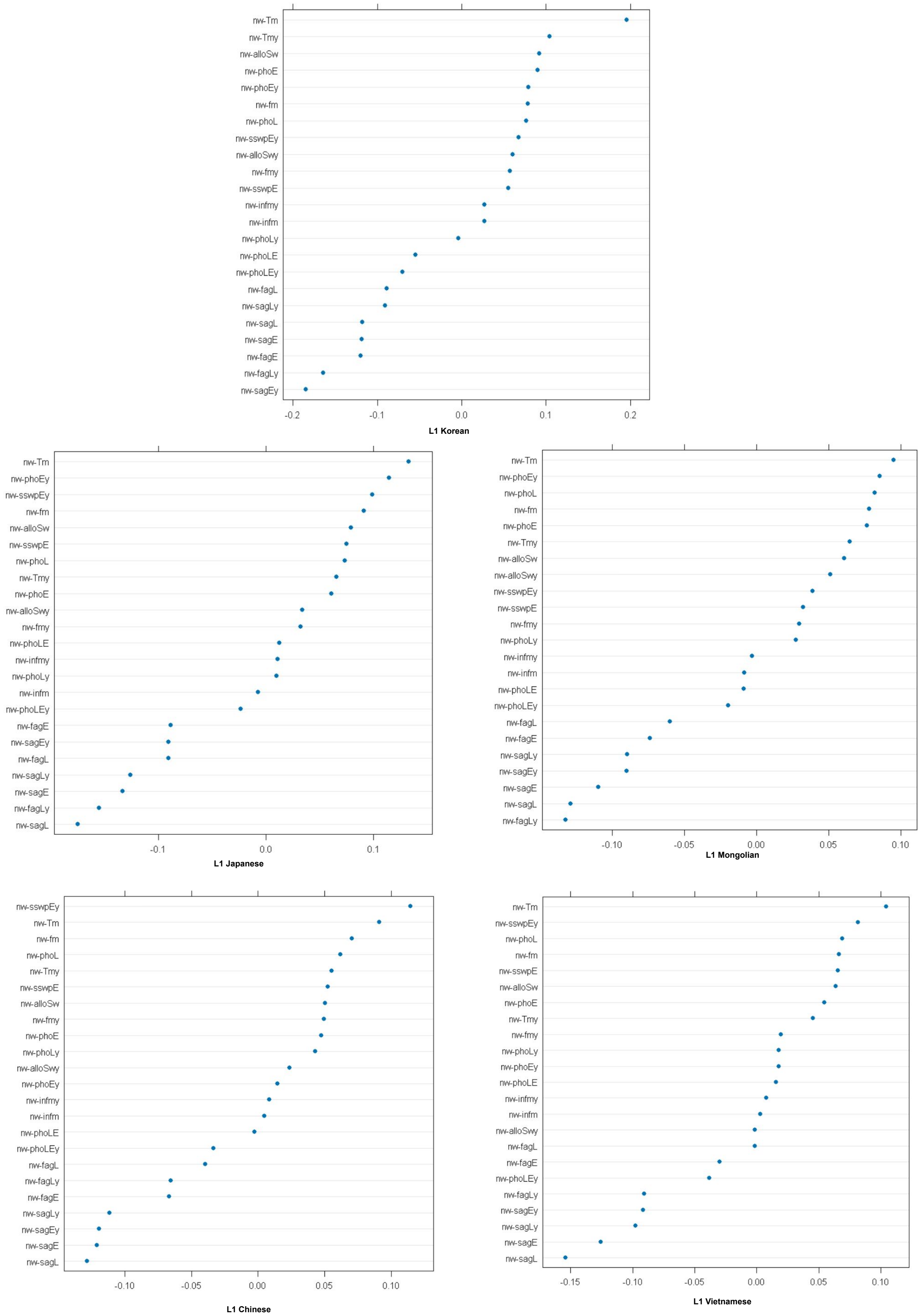


Figure 7.14: Random intercepts for nonword type (`nwCondition`) in All groups

Chapter 8

Joint Models for All Language Groups

Thus far, we have discussed the experimental results for each language group separately. In this chapter, we present omnibus analyses in which all language groups are considered jointly. This will make it possible to provide more precise assessments of similarities and differences between groups.

8.1 Reaction times

8.1.1 Words

We fitted a generalised additive mixed model to the data of all L2 groups together with the data from the Korean native speakers. We used treatment coding for language, with Korean as the reference level. All coefficients for the interactions of language by factorial predictors thus present the differences in effects with respect to Korean. Figure 8.1 presents the effects for the parametric terms of the GAMM. Significant effects ($p < 0.01$) are shown in blue, non-significant effects are presented in red. The lower set of panels visualises the contrasts for the levels of `SLIF`, and the upper set of panels show the effects of the other lexical predictors (e.g., past tense, number of syllables) as well as language.

The panel for `language` (upper figure, lower left panel) shows that mean reaction times were consistently longer for L2 speakers compared to native Korean participants, as expected. Of the four L2 groups, the Mongolian participants responded more quickly than the other three groups of speakers. Among the non-native groups, Mongolian speakers showed the fastest reaction times but also the highest error rates (see Figure 8.5). In contrast, the Japanese group exhibited the slowest response times. Considering that both Mongolian and Japanese, like Korean, are agglutinative languages, this stark contrast is particularly intriguing. One possible explanation may lie in non-linguistic factors, such as behavioural patterns shaped by broader cultural tendencies. That is, cultural traits and shared cognitive styles within a society might also influence language processing (e.g., Hedden et al., 2008). In fact, during interactions with participants, many Mongolian speakers tended to communicate in a quick and concise manner, whereas Japanese participants generally used longer and more detailed expressions. These differences in communication style may also contribute to variations in cognitive processing during language tasks.

The next panel presents the differences (with respect to Korean L1 speakers) in the effect size of the `honorifics` (using the treatment coefficients contrasting non-honorific with the honorific reference level). For the Korean speakers, non-honorific forms were responded to more quickly ($\hat{\beta} = -0.0496, p = 0.0001$). For the Mongolian L2 speakers, the effect size did not differ much from the native speakers, but the Chinese, Japanese, and Vietnamese L2 speakers all had significantly more negative coefficients compared to the Korean group. In other words, these non-native groups revealed a larger difference in response time between non-honorific and honorific forms, compared to the Korean group. For the Korean L1 speakers, non-honorific forms were responded to more quickly than honorific forms. For the non-native groups, except the Mongolian speakers, the advantage of non-honorific forms observed for Korean was even stronger.

The panel in the lower right of the upper trellis plot presents the differences in the effect sizes for `mood` forms. For the native speakers, non-mood forms elicited longer reaction times ($\hat{\beta} = 0.133, p < 0.0001$). For the Chinese and Vietnamese speakers, this effect was substantially reduced, and for the Japanese and Mongolian speakers, this effect even reversed into a small facilitatory effect for non-mood forms as compared to mood forms. This suggests that for non-native speakers, words with mood endings are comparatively more difficult to process.

Regarding the `past tense`, the native speakers responded slightly more slowly to past-tense forms ($\hat{\beta} = 0.0140$), but this difference was not reliable ($p = 0.2247$). All non-native speaker groups responded more slowly to past-tense forms, compared to the L1 group, suggesting that non-native speakers are more sensitive to the past tense forms compared to native speakers. However, this effect was significant only for the L1 Chinese and Japanese groups.

As the `number of syllables` increases, native speakers exhibited slower reaction times. ($\hat{\beta} = 0.0388, p < 0.0001$). All non-native speaker groups responded even more slowly to words with more syllables. This indicates that non-native speakers are much more affected by word length than native speakers.

As for the `number of exponents`, for the Korean native speakers, the estimated effect was $\hat{\beta} = -0.0364, p = 0.0015$. Except for the Mongolian group, the non-native groups had a slope for this predictor that was somewhat more negative, indicating that more exponents made a stimulus more word-like, facilitating a yes response. Interpreting the results in two different ways, the effect might arise from the perception of the word form as a typically patterned physical structure (i.e., a patterned `Ending`) by increasing the number of exponents. Alternatively, a greater number of exponents might come with increased semantic content, making stimuli more word-like and facilitating yes responses. In contrast to these interpretations, as mentioned earlier, the facilitative effect of the `number of exponents` on reaction time appears to coincide with the inhibitory effect of word length based on the `number of syllables`. This raises some concerns about potential collinearity; however, the correlation ($r = 0.43$) between the two appears to be modest.

For `vowel harmony`, the omnibus analysis did not support any significant differences between native and non-native speakers. All groups consistently showed the same facilitation effect in response times for light vowel stem-final sounds, just as native speakers. Thus, the anticipated greater sensitivity of Mongolian learners to the vowel harmony of Korean was not supported. Why words with light vowel stem-final sounds facilitate word recognition for both Korean native speakers ($\hat{\beta} = -0.0439, p < 0.0001$) and all learner groups is unclear.

When the `stem-final segment` was a vowel, Korean speakers responded more quickly ($\hat{\beta} = -0.0172, p < 0.0001$). The Japanese and Mongolian L2 learners had a very similar effect, but for the Chinese and Vietnamese learner groups, this effect reversed.

The lower set of panels in Figure 8.1 presents the differences between native and non-native speakers for the levels of `SLIF`, i.e., the combinations of speech level and illocutionary force. The reference level of the factor capturing the interaction of these two morphosyntactic features is `formal declarative`. For the Korean speakers, positive contrasts were detected for `Formal inquisitive` ($\hat{\beta} = 0.0138, p = 0.0637$) and `Unmarked` ($\hat{\beta} = 0.0636, p = 0.0003$). For the `Formal inquisitive`, the non-native groups responded similarly (lower left panel). For the `Unmarked` forms, with the exception of the Vietnamese, coefficients had substantial negative values, indicating that for these speaker groups, responses were faster for the `unmarked` forms (compared to the formal declarative), instead of slower.

All other contrasts of SLIF for the Korean native speakers were negative, indicating facilitation compared to the `Formal declarative`. The only levels of SLIF for which the non-native groups revealed significant differences compared to Korean are `Plain declarative` and `Plain inquisitive`. For Korean, the treatment contrasts were estimated as $\hat{\beta} = -0.1136, p < 0.0001$ and $\hat{\beta} = -0.0934, p < 0.0001$ respectively. The absolute values of the difference coefficients of the non-native groups are similar in magnitude, indicating that for these groups, there were hardly any effects of these two types of illocutionary force at the plain speech level. In other words, this suggests that most forms of the `plain speech level` are more difficult to recognise for non-native speakers as compared to native Korean speakers. `Plain inquisitive` -니 (-냐) /-ɲi(-ɲya)/ and `plain declarative` -ㄴ다 (-는다/-다)/-nda(-nɯnda/-da)/ words tend to be short, with one or two syllables only, contrasting with the relatively long syllables such as the `formal speech level` -(스) ㅂ니다 /-(s̤ɯ) mɲida/ and the distinct, `polite speech` ending -아 (어) 요/-a(ʌ)yo/ and for `honorific polite speech`, -(으) 세요/-(ɯ)s^{h}eyo/. The shorter forms with each distinct illocutionary force (i.e., declarative, inquisitive, imperative and propositive) of plain speech level combined with verb stems and undergoing phonological variation may be more difficult for non-native speakers.

The differences between native and non-native speakers for `polite unmarked` were not significant. However, all native and non-native groups showed facilitatory effects with the polite ending '요'/yo/, which can be a distinctive final form in the word recognition. The detailed analysis of this phenomenon will be explored further in the next section 8.2.2 for error analysis.

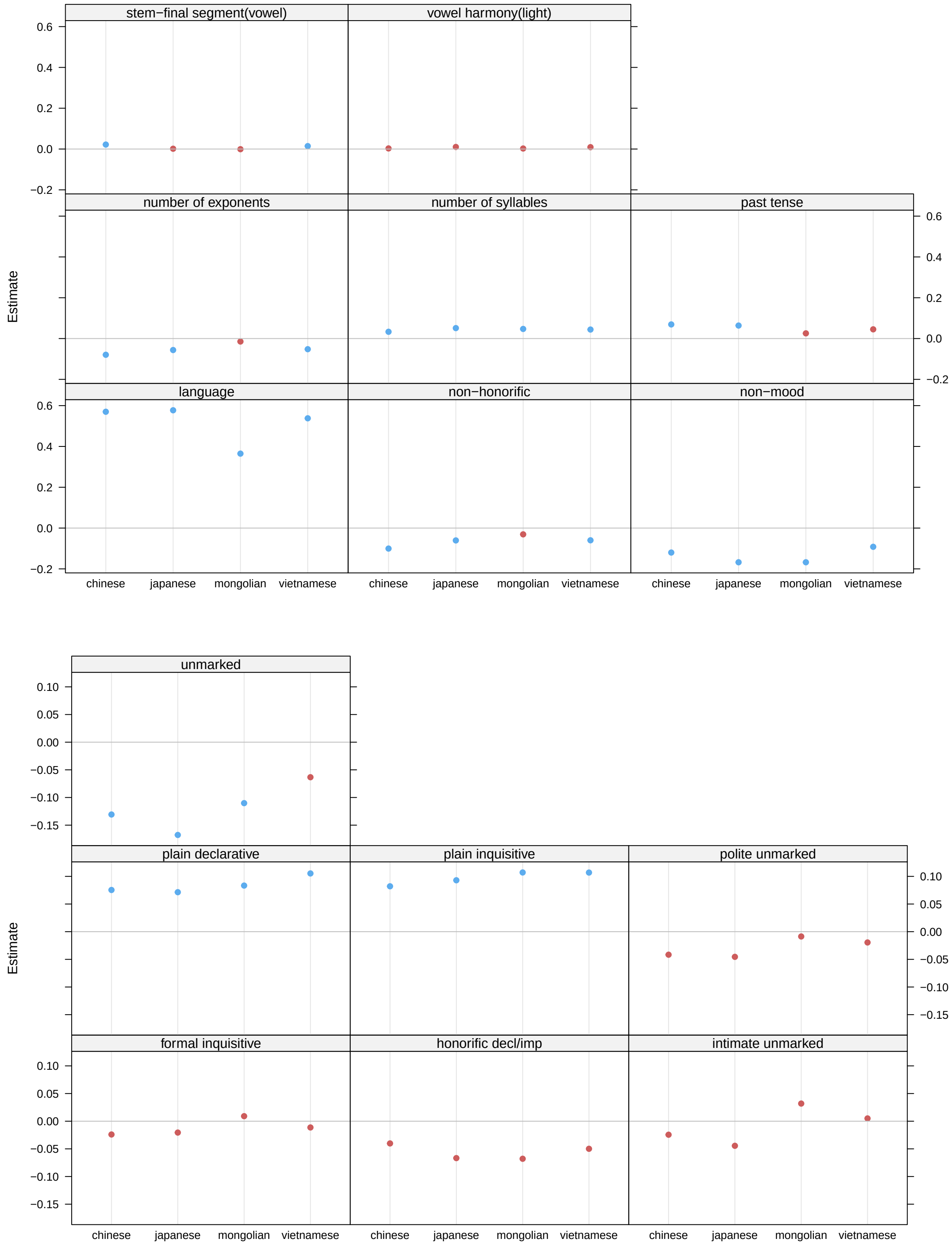


Figure 8.1: Parametric effects for lexical predictors in the GAMM fitted to the reaction times of the data of all language groups jointly. Effects (treatment coefficients with Korean as reference level) with $p < 0.01$ are shown in blue, non-significant effects are shown in red. The lower set of panels show the contrasts for the levels of *SLIF*, the upper set of panels show the effects of the other lexical predictors.

Figure 8.2 presents the smooth terms of the model. The rows of panels in this figure pertain to predictors, the columns pertain to language. For most predictors, p-values (displayed with each panel) are extremely small, indicating that these frequency effects have a solid chance of replicating in future experiments.

The frequency of the lexeme showed a facilitation effect across all groups. This effect was particularly pronounced for the native speaker and Chinese groups, with native speakers displaying a facilitation effect at (log) frequency levels around 2-6, after which a plateau was observed. In contrast, the lexeme frequency effect for other non-native groups tended to be more gradual overall.

For the whole-word form frequency effect, non-native speakers exhibited a steeper facilitation effect compared to the lexeme frequency effect, with response times sharply decreasing as whole-word frequency increased. For non-native speakers, experience with individual word forms is clearly reflected in their response behavior. For native speakers, a whole-word frequency effect was visible only for the lower form frequencies, with no effect for the higher form frequencies. Of the non-native groups, only Mongolian participants revealed a leveling off of the word frequency effect, with wider confidence intervals reflecting more uncertainty about the effect. This tentative qualitative similarity dovetails well with the Mongolian speakers having the overall fastest response latencies of all learner groups.

The effect of ending frequency is well-supported across all five groups. For Mongolian and Vietnamese, it is linear, for Japanese and Korean, it is non-linear with a general downward trend. For the lowest-frequency endings, Korean evidences the strongest effect, Chinese, Mongolian and Vietnamese show an intermediate effect, and the Japanese have no effect, or even a small effect in the opposite direction. For the highest-frequencies of the endings, the effect for Korean levels off, but remains stable and facilitatory for the L2 groups.

A longitudinal effect of trial was also present across all groups. Notably, the native speaker group displayed a more pronounced non-linear fluctuation in response patterns, suggesting that additional complex factors may affect native groups beyond experience-related variables. These factors may include enhanced sensitivity to linguistic factors, but also less stability in attention over time, due to the task being more boring and less challenging compared to L2 learners. Furthermore, their greater proficiency may have enabled the Korean native speakers to engage in response optimization. The interaction of trial by ending frequency supports this possibility. It is well-supported for Korean native speakers ($p < 0.0001$) and possibly for Vietnamese L2 speakers ($p = 0.0416$), but not for the other three languages.[26]

[26] For completeness, we also fitted a separate GAMM to the data using different curves with respect to Korean. All difference curves for the frequency effects were well supported (all $p < 0.0001$).

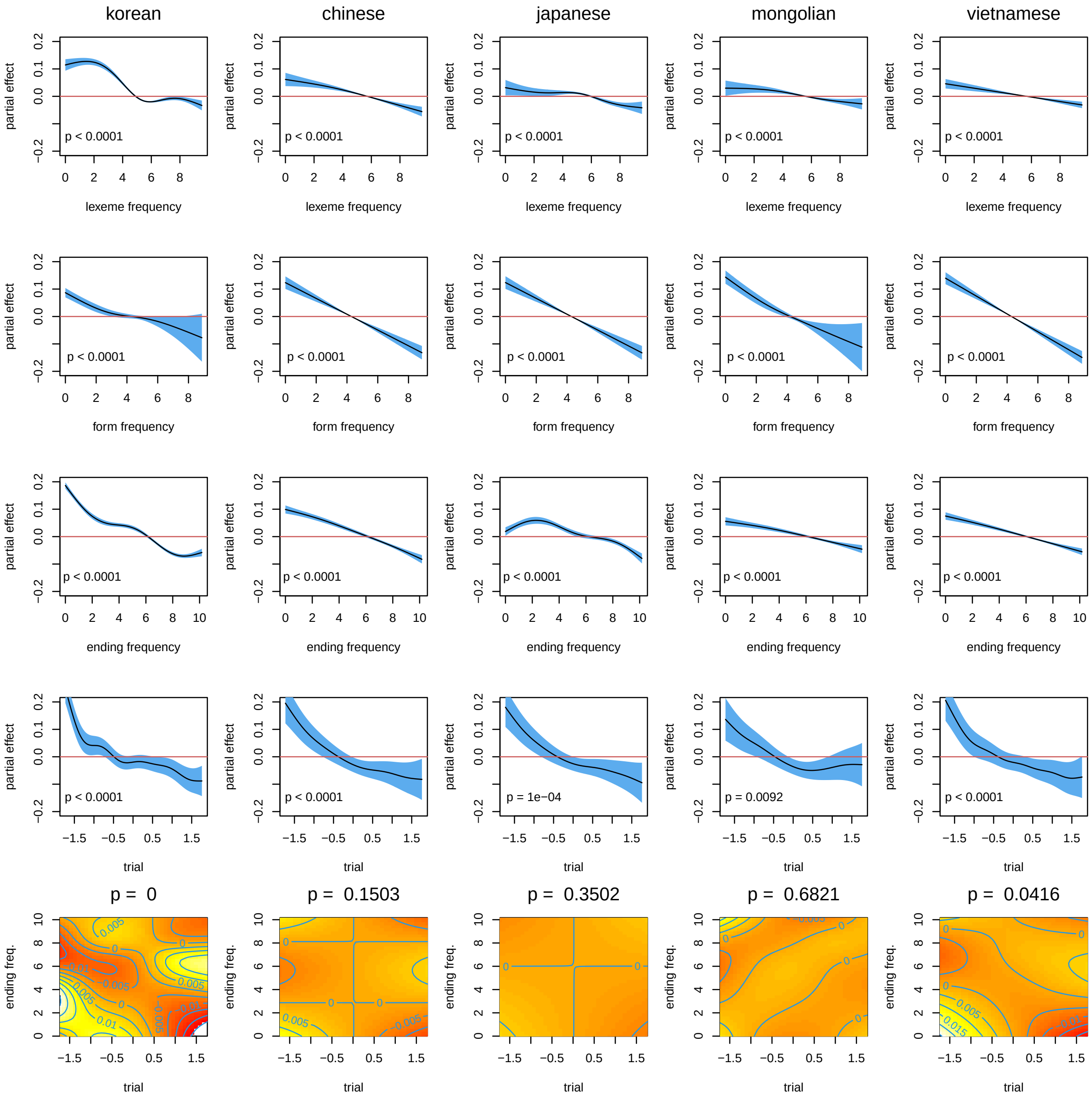


Figure 8.2: Partial effects for the non-linear terms in the GAMM fitted to the reaction times to the combined dataset of all speaker groups. P-values for the univariate smooths are added in the lower left corner, and above the panels for the tensor product smooths (bottom row). Panels are organized by predictor (rows) and language (columns).

8.1.2 Nonwords

Table 8.1 provides a summary of the GAMM fitted to the inverse-transformed reaction times to the nonwords of the participants in all five language groups. As before, treatment coding was used for factor coding. For language, Korean was the reference level. All non-native speakers had longer reaction times to nonwords compared to the native speakers, with the longest times being present for the Vietnamese speakers, and the shorter for Mongolian (compared to Korean). For all non-native groups, the effect of the `number of syllables` was stronger, leading to longer reaction times compared to native speakers. A robust effect of the `number of exponents` was present only for the Mongolian speakers, for whom the slope was substantially more positive compared to the slope of the native speakers.

Table 8.1: Summary of a Gaussian GAMM fitted to the inverse-transformed RTs to the nonwords of the speakers of all five language groups. The reference level for language is Korean.

A. parametric coefficients	Estimate	Std. Error	t-value	p-value
(Intercept)	-1.5244	0.0369	-41.2912	< 0.0001
language=Chinese	0.3123	0.0612	5.1064	< 0.0001
language=Japanese	0.3049	0.0604	5.0518	< 0.0001
language=Mongolian	0.2119	0.0612	3.4640	0.0005
language=Vietnamese	0.4121	0.0612	6.7380	< 0.0001
number of syllables	0.0362	0.0017	20.7987	< 0.0001
number of exponents	0.0024	0.0021	1.1842	0.2363
language=Chinese:number of syllables	0.0108	0.0030	3.6565	0.0003
language=Japanese:number of syllables	0.0249	0.0029	8.5543	< 0.0001
language=Mongolian:number of syllables	0.0215	0.0030	7.2893	< 0.0001
language=Vietnamese:number of syllables	0.0221	0.0030	7.4882	< 0.0001
language=Chinese:number of exponents	0.0029	0.0036	0.8132	0.4161
language=Japanese:number of exponents	0.0081	0.0035	2.3088	0.0210
language=Mongolian:number of exponents	0.0153	0.0036	4.2631	< 0.0001
language=Vietnamese:number of exponents	0.0002	0.0036	0.0629	0.9499
B. smooth terms	**edf**	**Ref.df**	**F-value**	**p-value**
s(log lexeme frequency):language=Korean	3.7624	3.9662	14.3754	< 0.0001
s(log lexeme frequency):language=Chinese	3.5668	3.8899	8.3710	< 0.0001
s(log lexeme frequency):language=Japanese	3.8925	3.9925	22.1149	< 0.0001
s(log lexeme frequency):language=Mongolian	3.8374	3.9832	17.1000	< 0.0001
s(log lexeme frequency):language=Vietnamese	3.8418	3.9841	10.9434	< 0.0001
s(trial):language=Korean	7.4165	7.8898	10.4785	< 0.0001
s(trial):language=Chinese	3.5908	3.9248	6.9702	< 0.0001
s(trial):language=Japanese	4.8570	5.3459	6.9103	< 0.0001
s(trial):language=Mongolian	8.0692	8.4044	6.0596	< 0.0001
s(trial):language=Vietnamese	3.9651	4.3488	9.1783	< 0.0001
random intercepts type × language	106.6411	110.0000	64.8762	< 0.0001
s(trial,participant)	885.7782	1084.0000	55.3614	< 0.0001

Figure 8.3 presents the partial effect of log lexeme frequency, with the differences in the intercept added in. Korean speakers responded fastest, with the Mongolian speakers as runner-up. Reaction times were the longest for the Vietnamese speakers. All speaker groups revealed a non-linear effect of lexeme frequency that was somewhat more symmetrical for the non-native speakers. The overall trend, although small, is for nonwords created from more frequent lexemes to elicit somewhat shorter reaction times. However, non-native speakers appear to have been especially cautious in responding to nonwords created from words with roughly average lexeme frequency. This may reflect a task-specific response strategy optimizing for the words with more common frequencies of use.

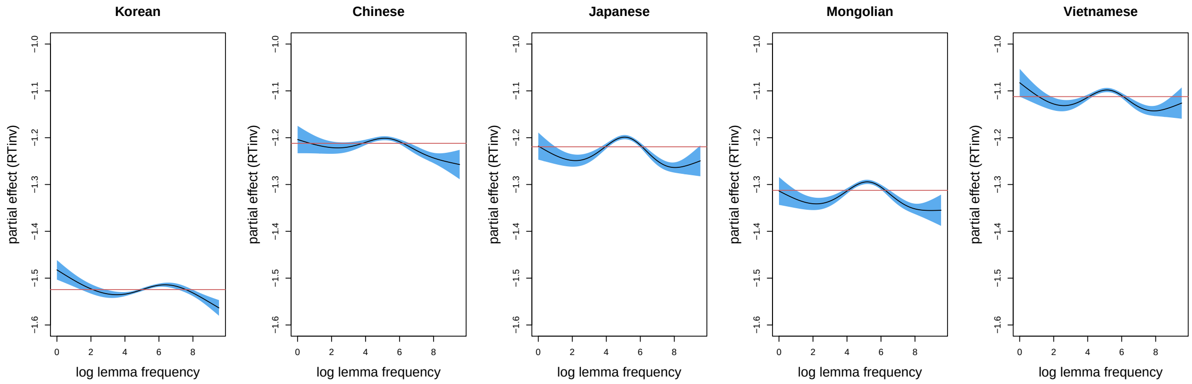


Figure 8.3: Partial effects (including shifts in the intercept) in the GAMM fitted to the inverse transformed reaction times to nonwords, for the smooths for log lexeme frequency, for each of the five language groups.

Figure 8.4 presents the random intercepts for the interaction of nonword type and language, panel-wise presenting the differences between languages for a given nonword type. For detailed descriptions, see Section 7.5.

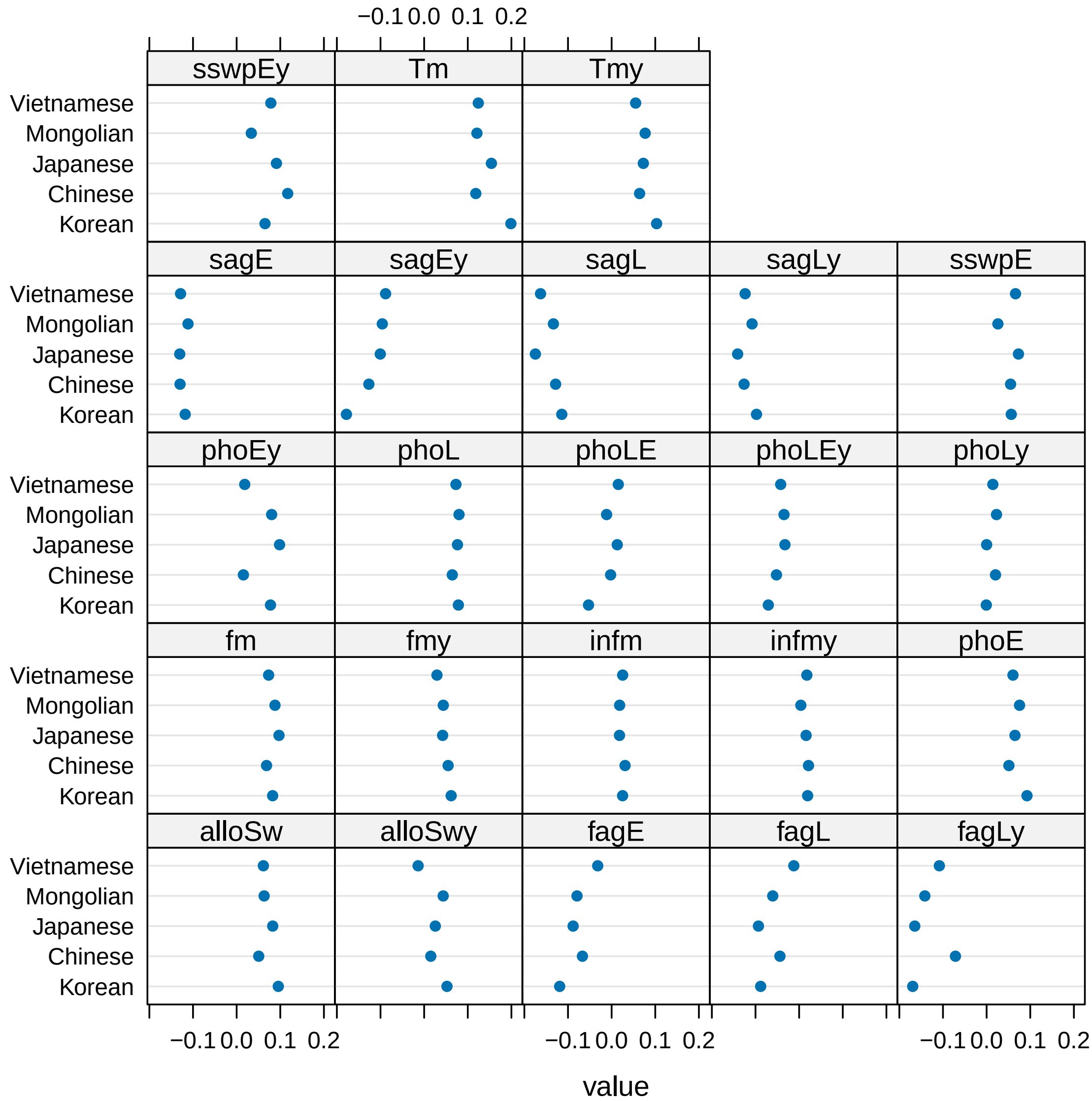


Figure 8.4: Estimated random intercepts for the interaction of nonword type × language in the model fitted to the inverse-transformed RTs to nonwords, for all language groups.

8.2 Error analyses

8.2.1 Errors from Words

Previously, we primarily analysed the results of the data, focusing on reaction times. In this section, we examine the accuracy of the results by analysing error patterns and relating them to key issues previously discussed in Korean verbal processing.

The overall error rate for word forms in each group is shown in Figure 8.5. As expected, native speakers showed the lowest error rate. Among the non-native groups, the Vietnamese group had a relatively lower error rate compared to other non-native groups. The highest error rate was observed in the Mongolian group, followed by the Japanese and Chinese groups, in that order.

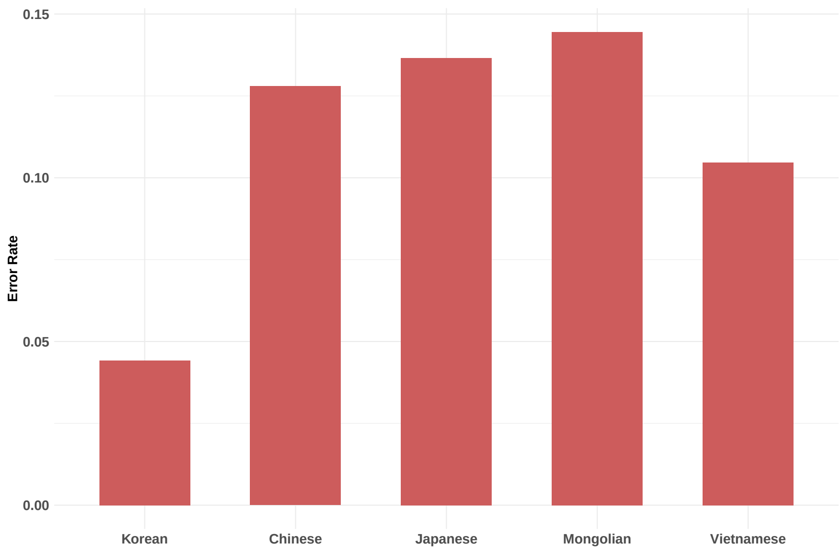


Figure 8.5: Overall Error Rates by Group for Word Forms

Figure 8.6 visualises the error rates related to honorifics, past tense, and mood forms across different language groups. The following Figure 8.7 illustrates how error rates vary depending on the presence or absence of each variable. The light blue bars represent the error rates when the variable is present, while the light red bars indicate the error rates when the variable is absent.

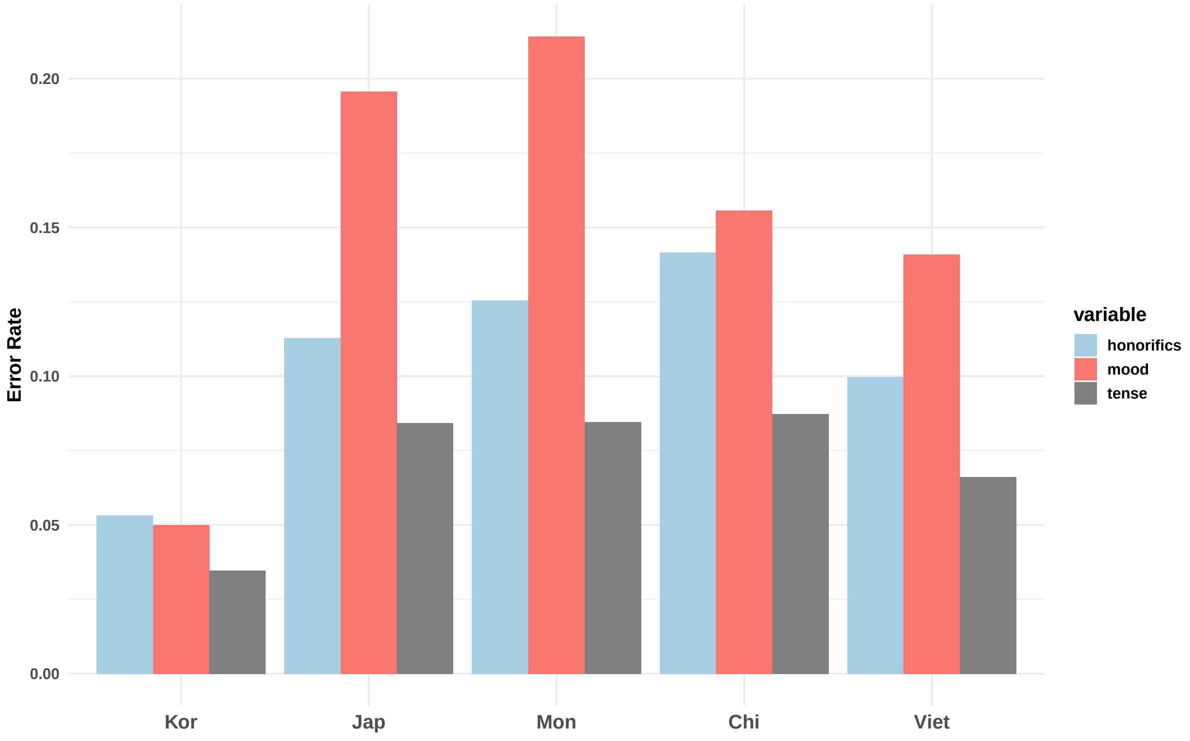


Figure 8.6: Error Rate Comparison by Language Across Levels of Honorifics, Past Tense, and Mood

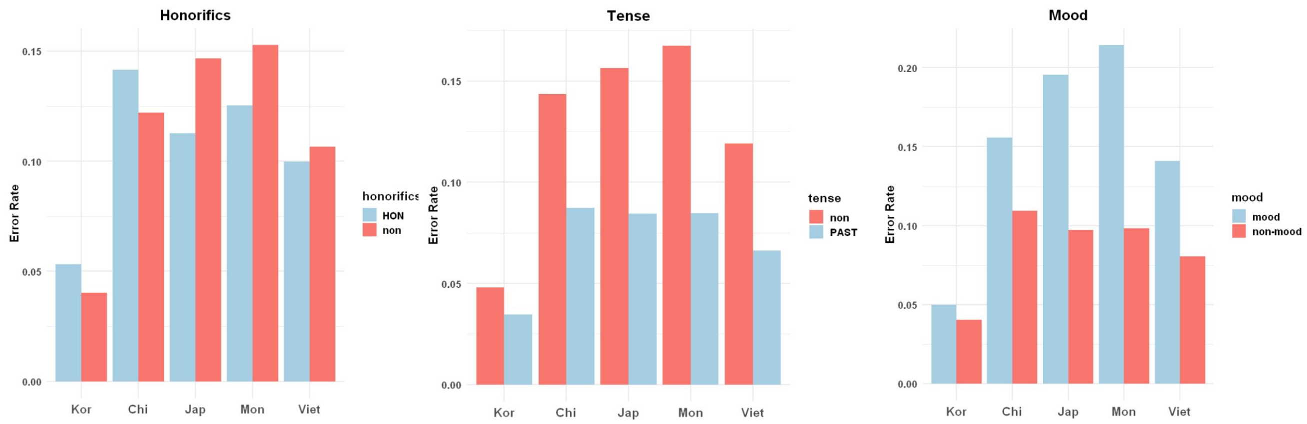


Figure 8.7: Error Rate Comparison Across Groups by Past tense, Honorific, and Mood Markings

Focusing on Figure 8.6, as expected, native Korean speakers exhibited significantly lower error rates across all forms. Interestingly, the Japanese and Mongolian groups showed similar error rates and patterns in the use of honorifics, past tense, and mood variables. Similarly, the Chinese and Vietnamese groups also appeared to follow a comparable pattern.

Among native speakers, honorifics had the highest error rate, followed by mood and past tense. In contrast, for all non-native speaker groups, mood forms had the highest error rates, followed by honorifics, while past tense consistently showed the lowest error rates.

Mood forms, in particular, exhibited higher error rates in the Japanese and Mongolian groups compared to other honorific and past tense forms, which is a notable contrast to the Chinese and Vietnamese groups. In the Chinese group, the difference in error rates between honorifics and mood forms was relatively small, whereas in the Vietnamese group, mood forms showed slightly higher error rates than honorifics.

Overall, the past tense had the lowest error rate across all groups, while non-native speakers showed the highest error rates for mood forms. Interestingly, the Japanese and Mongolian groups, both agglutinative language speakers, exhibited higher error rates than the Chinese and Vietnamese groups, which speak analytic languages. This result is particularly intriguing, as Japanese and Mongolian, which share the same linguistic typology(i.e., agglutinative languages) as Korean, displayed higher error rates than the more typologically distant Chinese and Vietnamese.

As shown in the Figure 8.7 below, one noticeable pattern is that the error rate is higher when the past tense exponent is absent than when it is present. This aligns with the earlier results in Figure 8.6, where the past tense showed relatively low error rates. That is, forms consisting only of a verb stem and SLIF without past tense or a stem with another exponent (excluding the past tense), such as honorific or mood forms, tend to result in more errors.

In contrast to the past tense, the presence of mood forms leads to a significant increase in error rates among the non-native speaker groups. However, for honorific forms, the pattern differs: in the Japanese, Mongolian, and Vietnamese groups, the error rate was lower when honorifics were absent, whereas the Chinese group and the native speaker group showed the opposite trend.

The following Figure 8.8 presents both the errors related to SLIF and the response times for SLIF side by side, to provide a rough overview of how errors in SLIF relate to reaction times. As expected, native speakers responded significantly faster compared to all non-native groups. Among the non-native groups, the Mongolian group(Mon) exhibited generally faster reaction times than the others. However, similar to the Chinese group(Chi), the Mongolian group showed a wider interquartile range compared to the Japanese(Jap) and Vietnamese(Vie) groups. An interesting feature is that the Mongolian group shows generally high error rates, yet relatively short response times, with few noticeable outliers. This pattern is particularly intriguing. In contrast, the Vietnamese group exhibits the lowest overall error rates, and also shows a relatively narrow interquartile range compared to other groups. This pattern may suggest that the Vietnamese group has a more stable processing basis for SLIF.

Regarding the overall error rates for SLIF (the bar plot), native speakers, as expected, had significantly fewer errors than non-native speakers. Within the native Korean group, plain speech inquisitive (`plain_inq`) forms had a higher error rate than other forms, while declarative (`plain_dec`) forms, despite sharing the same speech level, had the lowest error rate. As expected, the `unmarked` mood-related forms exhibited higher error rates across all L2 groups compared to other SLIF combinations. The Chinese group showed particularly high error rates in plain speech inquisitive and intimate speech (`int_unmark- ed`) forms compared to other groups. The Mongolian group generally showed high error rates, especially for `unmarked` forms (i.e., mood exponents), with the Japanese group following closely behind. A similar trend was observed in polite speech (`pol_unmarked`) forms, although in this case, the Japanese group showed the highest error rate, followed by the Mongolian group. The Vietnamese group showed the lowest error rates across most SLIF, except for the intimate speech level. This pattern is consistent with the overall error rate trends by group shown in Figure 8.5. Overall, the Mongolian group exhibited high error rates across various levels. Notably, the Chinese group showed distinct error patterns at certain levels, particularly in the plain speech inquisitive form and intimate speech level, compared to the other groups. In the formal speech, the Mongolian group exhibited the highest overall error rate, followed by the Chinese group in the `for_dec` condition and the Japanese group in the `for_inq` condition.

We also employed a logistic GAMM on error data to further investigate error patterns across the groups. Figure 8.9 presents the partial effects of the smooth terms for the five languages. The five languages show the same kind of lexeme frequency effect, which is least wiggly for the Mongolian speakers and which has a general negative trend. As expected, more frequent lexemes are recognised more accurately(i.e., low errors). The whole-word(form) frequency effect is also present for all groups of languages. Whereas the effect levels off for Korean, Chinese, and Japanese, it is linear for Mongolian and Vietnamese. The Mongolian L1 speakers show the strongest whole-word frequency effect of all. The effect of ending frequency is similar and similarly wiggly for all groups, with the Chinese group being the exception. As the experiment proceeded, errors decreased slightly across all groups. The Japanese, Mongolian, and Vietnamese groups showed difficulties with low-frequency endings, but in the course of the experiment, they became more accurate. For the Vietnamese speakers, an opposite trend is visible for the high-frequency endings.

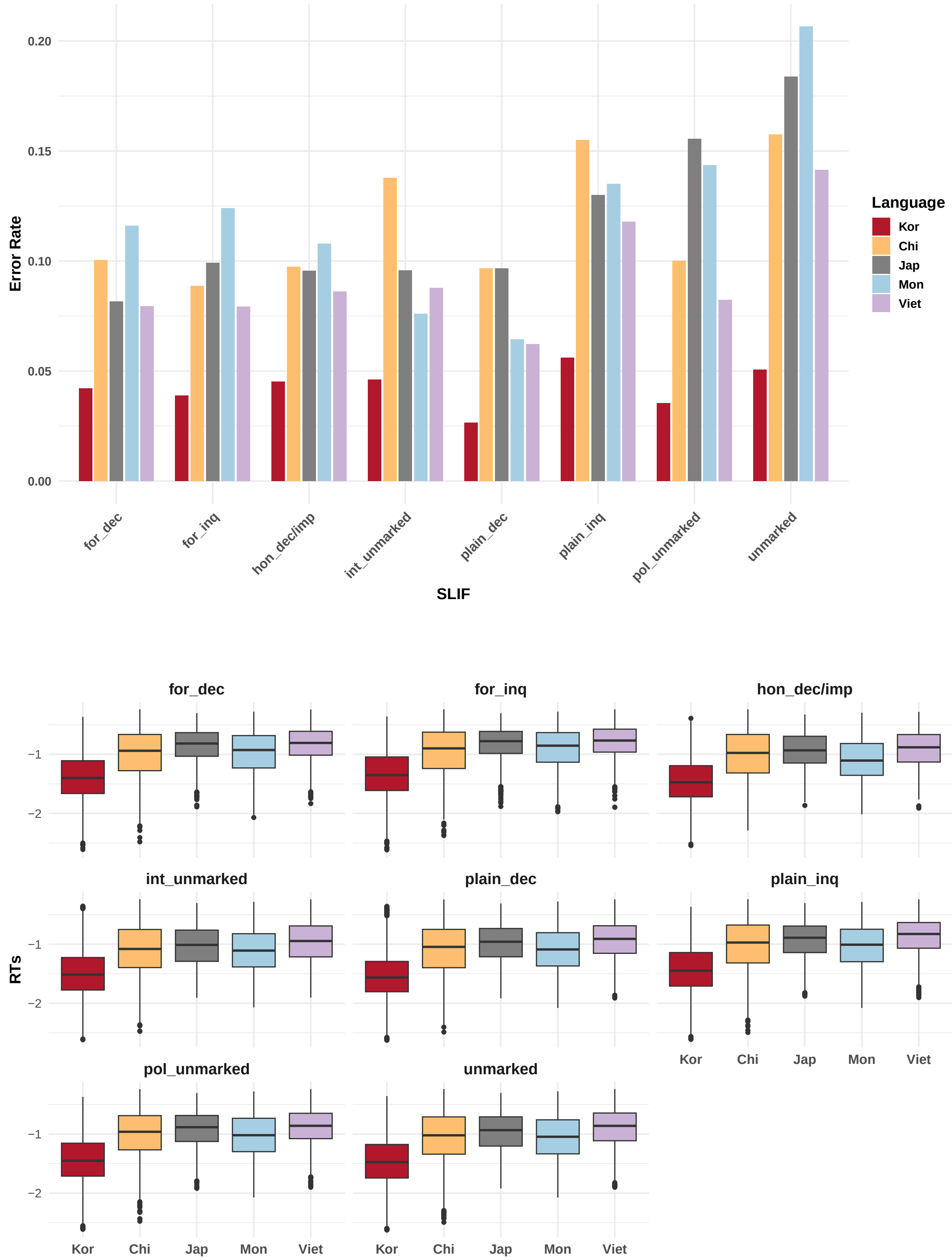


Figure 8.8: Comparison of Error Rates and Response Times by SLIF Across All Groups

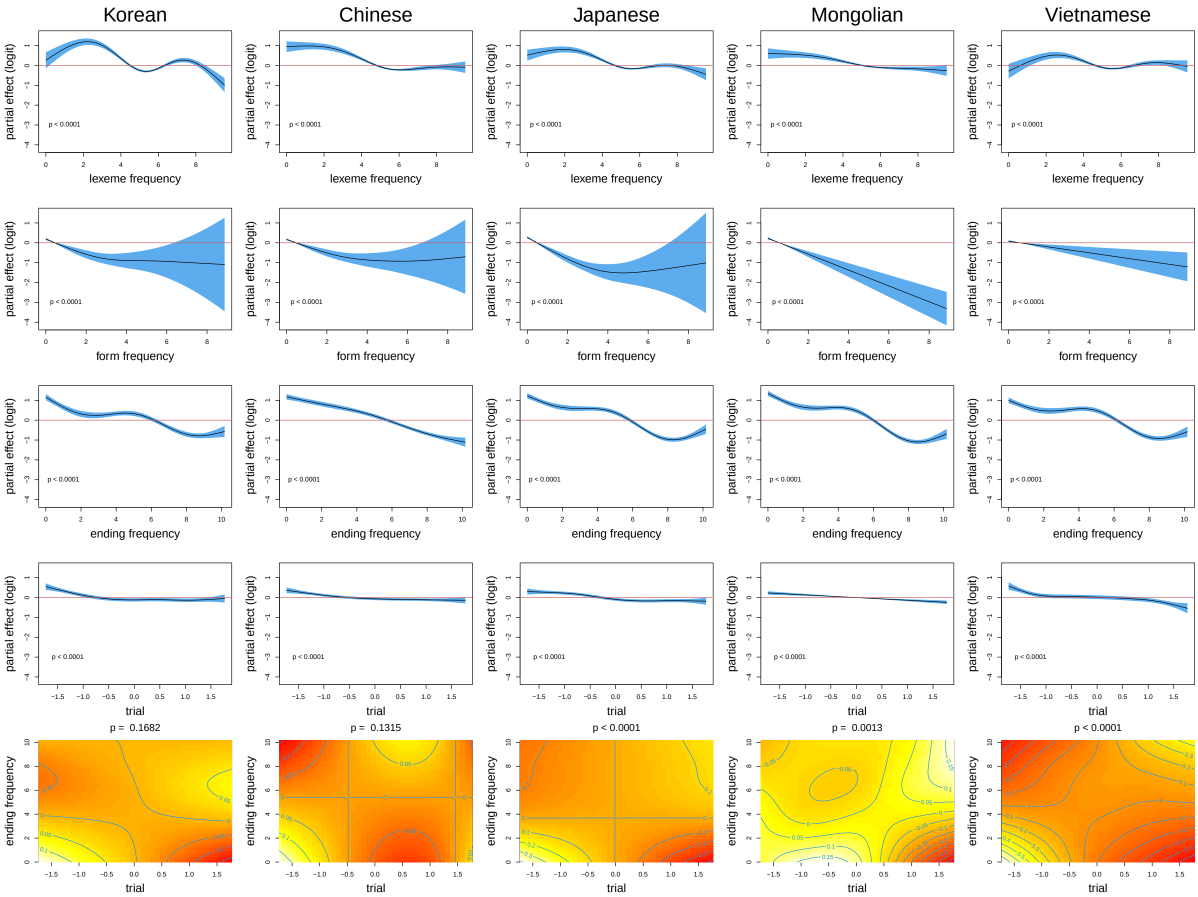


Figure 8.9: Partial effects of the smooths in a logistic GAMM fitted to the accuracy data for words, for each of the five language groups.

8.2.2 Errors from Nonwords

Figure 8.10 presents the nonword error rates observed across all groups. As expected, the Korean native speaker group showed the lowest nonword error rate, while the L1 Chinese group exhibited the highest. This was followed by the L1 Vietnamese group, which also showed a relatively high error rate, whereas the L1 Mongolian and Japanese groups demonstrated similarly low error rates. Interestingly, this pattern contrasts with the results for real words, where the Mongolian group had the highest error rate, followed by the Japanese group, suggesting a reversed trend.

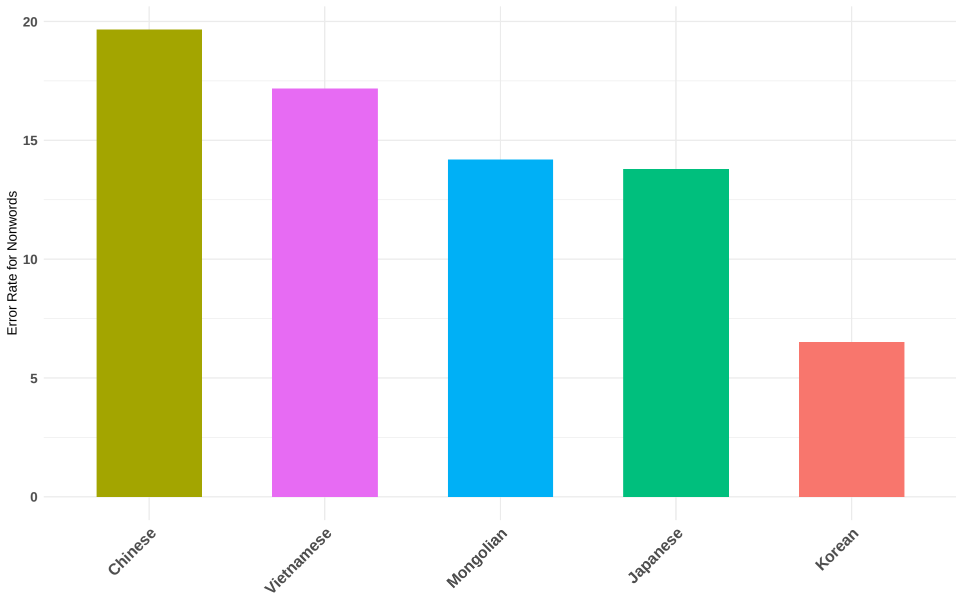


Figure 8.10: Nonword Error Rates Across Language Groups

Figure 8.11 presents the estimated coefficients for the random effect of nonword type by language. Nonword types with values more to the right in a panel are the nonword types that are more likely to elicit errors (e.g., `alloSw`). Conversely, types with values more to the left in a panel elicited fewer errors (e.g., `sagE`). The following bar plot 8.12 presents the error rates by nonword condition in a more intuitive manner.

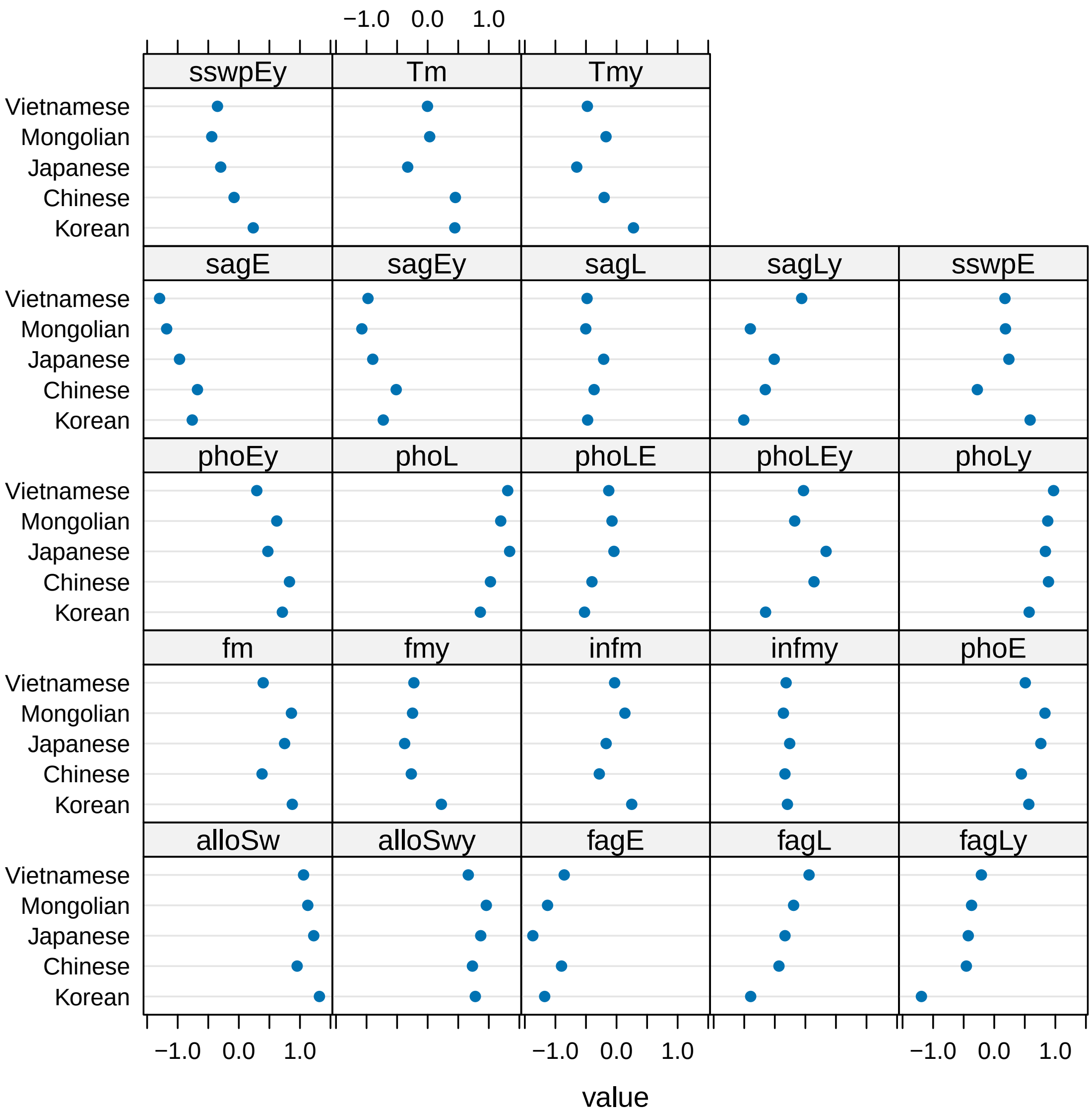


Figure 8.11: Random effect estimates predicting nonword error rates by language and nonword type

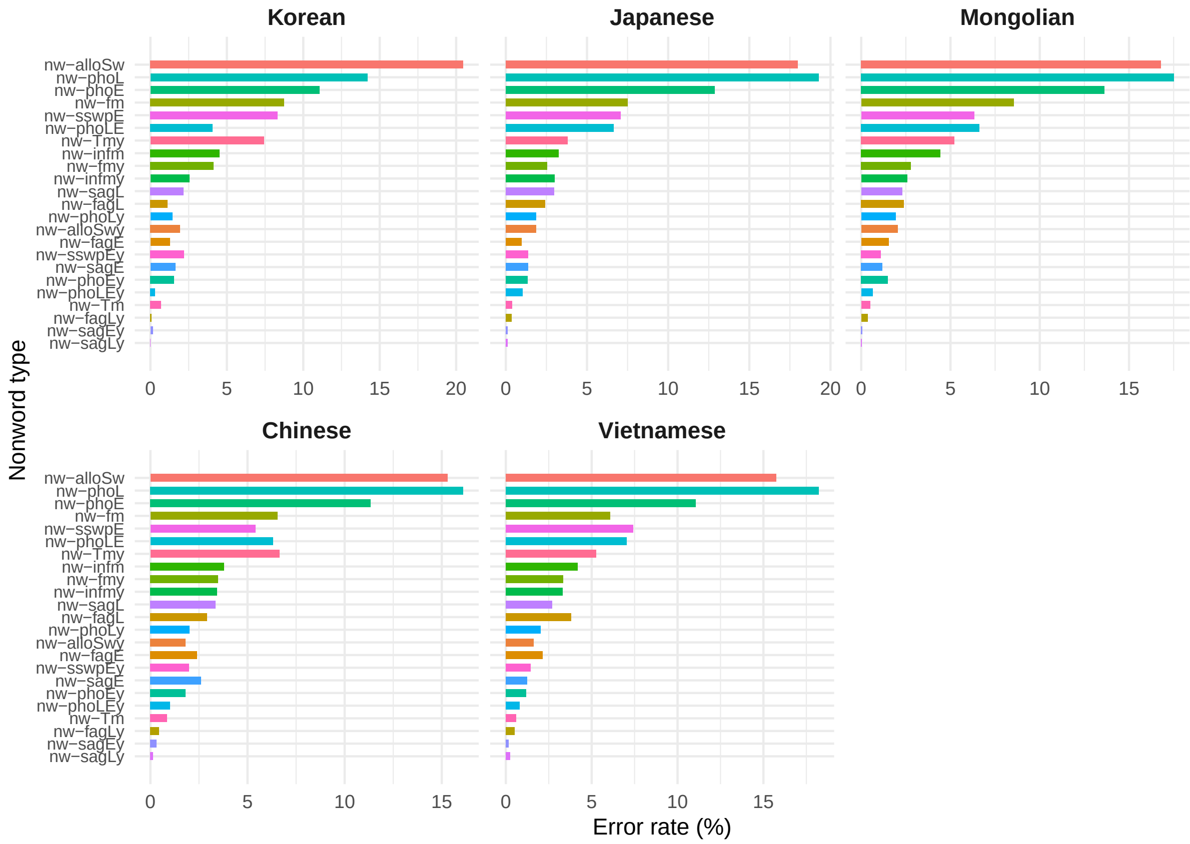


Figure 8.12: Error rates by nonword condition across language groups

The nonword condition with the highest error rate across all groups was the `alloSw` condition, where allomorphs were swapped. One possible interpretation of these results is that allomorphic variations arising from phonological processes may not be highly salient to the participants. Alternatively, the incorrect addition of such allomorphs may make nonwords appear more word-like. In the `alloSwy` condition, where the -요/yo/ form was added, there was a general decrease in error rates.

Conversely, the nonword condition `fagL/E/y`, which exhibited the lowest error rate among native speakers, involved phonologically impossible elements in Korean orthography with both the stem and ending altered. This likely resulted in faster recognition and higher accuracy, as the form was easily identifiable as non-Korean. Additionally, in the `fagL/y` condition, where only the stem was manipulated, non-native speakers showed significantly more errors compared to native speakers. This can be interpreted as non-native speakers having more difficulty recognising unfamiliar orthographic patterns in the stem, leading them to mistakenly perceive such forms as new, unknown words.

Yet, interestingly, in the `fagE` condition, where only the ending was manipulated, the error rate decreased compared to the stem manipulation condition. This reduction was most pronounced in the Japanese group, followed by the Mongolian, Vietnamese, and Chinese groups.

Two possible explanations can be suggested: first, most speakers may place greater emphasis on processing the ending rather than the stem; second, endings often consist of more syllables than stems, making them more salient novel forms and thus easier to process accurately.

Similar to the fag condition, the `sagL/E/y` condition involved phonologically impossible elements with only one syllable manipulated in the word. In the `sagE` condition, where a single syllable in the ending was altered, the error rate was the lowest across all groups. Interestingly, the addition of -요/yo/ slightly increased the error rate, possibly because the presence of -요 led participants to perceive the form as more word-like in the particular novel form condition for endings. As observed in the fag condition, non-native groups exhibited higher error rates when the manipulation was applied to the stem (`sagL/y`) compared to the ending (`sagE/y`). Given that many Korean verb stems are monosyllabic, nonwords with one syllable altered stems, but relatively long, well-formed endings may be interpreted by non-native speakers as unfamiliar but possible real words.

The nonword condition `sswpE`, where the two syllables of the ending were swapped, showed the highest error rate among native speakers and the lowest error rate in the Chinese group. The relatively low error rate in the Chinese group may be due to the influence of Chinese orthographic or phonological features, which could lead to stronger syllable-level processing, making them more aware of the syllable reordering. However, in cases where -요 was added (`sswpEy`), the error rate in the Chinese group was actually higher than in other non-native groups. This may be attributed to a word-like effect triggered by the presence of -요, which likely led to increased judgment errors. However, across all other groups except the Chinese group, the error rate decreased in the `sswpEy` condition, indicating the word-like effect of the -요 form.

In conditions `Tmy`[27], native speakers showed higher error rates, while the Japanese group exhibited the lowest. This may be interpreted as the Japanese speakers responding more sensitively to morphologically complex forms involving combinations of tense and mood, possibly due to similarities in verbal inflection patterns between Japanese and Korean. In contrast, the relatively higher error rates among native speakers stand out as a noteworthy finding.

In the pho series forms, which were manipulated to create similar phonological effects, `phoL`, where only the stem was manipulated, showed the highest error rate. When the -요/yo/ form was added to the same manipulation (`phoLy`), the error rate decreased slightly. `phoE`, where only the ending was manipulated, had a slightly lower error rate than `phoL`. Interestingly, for native speakers and the Chinese group, the -요 form had no noticeable effect, while other groups showed a slight reduction in errors with the -요 form. Notably, `phoLE`, where both the stem and the ending were manipulated, resulted in the fewest errors overall. Interestingly, in this condition, the presence of the -요 form led to more errors for all groups except native speakers and the Mongolian group. Overall, the error rate followed the general pattern: `phoL/y > phoE/y > phoLE/y`, indicating a clear difference in the magnitude of errors across conditions.

Examining mood-related nonword forms (`fm/y, infm/y`), the frequent mood form without -요 (`fm`) exhibited the highest error rate. However, when -요 was added (`fmy`), the error rate significantly decreased, suggesting a -요/yo/ positive effect. Interestingly, the less frequent mood form (`infm`) showed fewer errors compared to the more frequent `fm`. As discussed in previous analyses, this can be interpreted as an effect where less frequent mood forms with distinctive features, such as having more than two syllables, actually facilitate word recognition. Similarly, in the less frequent mood form

[27] As mentioned earlier, we focus on `Tmy` rather than `Tm`, given the limited data for `Tm`.

with -요 (`infmy`), the error rate was slightly lower than in the form without -요 `infm`). Notably, this pattern was consistent across all groups, showing nearly similar error rates.

Chapter 9

Key Findings and General Discussion

In this study on Korean verb processing among non-native speakers, the primary focus was to identify which variables have the greatest impact on word processing and how these effects differ depending on the speakers' first languages. This section will summarise the results and discuss the implications based on several related subtopics,.

9.1 Key factors by AIC

To begin with, we examined the impacts of predictors using AIC values (i.e., the increase[28] when a predictor is removed from the model). The `number of syllables` (i.e., word length) emerged as the most influential factor in most L1 groups, Japanese, Mongolian, and Vietnamese, but L1 Chinese group showed `Ending` (i.e., the combination of the ending exponents) as a greater impact than the `number of syllables`. This primary trend, which centres on `number of syllables` contrasts with that of native Korean speakers, where the `Ending` variable had the most substantial effect, followed

[28] The greater the increase in AIC, the more influential the variable is considered to be. The figures in the bracket in the following description represent the AIC increases.

by the `ending frequency`. The next most influential predictor for Korean non-native speakers was the `Ending`. In sum, while both native and non-native speakers are most affected by the `Ending` and the `number of syllables`, the order of their impacts is reversed between native and non-native Korean speakers. These results suggest that word length, represented by the `number of syllables`, plays a more significant role in the visual processing of Korean verbs for L2 speakers than it does for native speakers. Kim et al. (2018b) found that advanced (TOPIK 6) non-native speakers[29] were more influenced by whole-word frequency than word length (whole-word frequency > number of letters > number of syllables > stem frequency), which contrasts with the findings of this study. However, since this study did not take into account whole-ending frequency, the results—excluding this factor—can be considered somewhat consistent with the findings of the present study.

The next most impactful predictors were frequency-related variables. The results showed that for the Japanese and Vietnamese L1 groups, the AIC increases of the `ending frequency` and `whole-word frequency` show relatively similar impacts. For the Mongolian group, the AIC increase of the `ending frequency` was slightly higher than that of the `whole-word frequency`. Interestingly, the Chinese group stood out from the other non-native groups, as the influence of `ending frequency` (252) was substantially greater than any other frequency measures (`lexeme frequency (69), whole-word frequency (88)`). It is also interesting that `lexeme frequency` had relatively little influence across all non-native groups. This contrasts sharply with the native speaker results, where `lexeme frequency` was the second most influential factor after `ending frequency`, and `whole-word frequency` had a much smaller effect than the other two. In summary, while for native speakers, both `lexeme and ending frequencies` had stronger effects than `whole-word frequency`, non-native speakers showed minimal sensitivity to `lexeme frequency`. Instead, most non-native groups demonstrated `whole-word frequency` effects comparable in magnitude to `ending frequency`.

[29] Of the 45 participants, 31 were L1 Chinese speakers, and the rest had various other language backgrounds.

Overall, based on frequency effects, the general structure of Korean verb processing in non-native speakers in this study appears broadly similar to that of native speakers (i.e., 'stem' + '`Ending`'), but the relative weighting of endings and stems differs significantly. That is, `ending frequency` accounted for the largest proportion of all the frequency effects across all non-native speaker groups. Moreover, except for the L1 Chinese group, the L1 Japanese, Mongolian, and Vietnamese groups also showed a relatively high `whole-word frequency` effect—slightly smaller than the `ending frequency` effect—while `stem frequency` had only a minimal effect across all non-native groups. These results suggest that non-native speakers may primarily scan the `Ending` when recognising Korean verbs, referring to the stem only secondarily. In this sense, their processing architecture may resemble that of native speakers in being based on both 'stem' + '`Ending`', but with a heavier reliance on `Ending`, which sets them apart from native speakers who show a noticeable sensitivity to `stem frequency` as well. In addition, the generally strong `whole-word frequency` effects observed in non-native groups imply that learners may recognise some words (e.g., high-frequency words) holistically. Accordingly, these results may be interpreted in light of both the parallel dual-route model (Baayen et al., 1997)—which assumes simultaneous processing of stem and `Ending`—and the hybrid model (e.g., Kirkici and Clahsen, 2013; Ahn et al., 2013; Lee et al., 2020; Silva and Clahsen, 2008; Neubauer and Clahsen, 2009; Feldman et al., 2010; Clahsen and Felser, 2006b), which suggests a possible shift to whole-word recognition depending on the context. Given that non-native speakers exhibit a strong `whole-word frequency` effect, it may be misleading to interpret this simply as a result of high-frequency words alone. It may be possible that, for non-native speakers, the morphological and semantic distinction between stem and ending is unclear, and they perceive Korean verbs as single whole-word units. In doing so, they may focus more on the ending rather than the initial part (i.e., the stem) when recognising Korean verbs. In other words, the most plausible interpretation is that non-native speakers are more likely to process Korean verbs based on larger, holistic units rather than through an analysis of individual exponents. Within this context, it may be even possible that the distinction between stem and ending may not be clearly recognised in their processing of Korean verb forms. However, given the substantial morphological and semantic load typically carried by verb endings in Korean, it is likely that non-native speakers rely heavily on endings during lexical decision tasks. That said, lexical decision tasks inherently lack contextual information and are designed to assess wordhood only, which may have led to a reduced sensitivity to stem-related meaning. Since the stem typically carries the core semantic content in actual usage, it is important that future research takes this into account by examining verb processing within more naturalistic, contextualised settings.

The other prominent factors were `vowel harmony`, `SLIF`, and `honorifics`, with the order of their impacts varying by L1 background (bigger AIC increase > smaller AIC increase) : `vowel harmony > SLIF` (Japanese), `SLIF > vowel harmnoy` (Mongolian), `vowel harmony > honorifics` (Chinese), `SLIF > vowel harmony` (Vietnamese). A noteworthy finding is that, unlike other groups, the Chinese group showed a particularly small influence of `SLIF` (27), but a relatively high influence of `honorifics` (65). This pattern distinctly differentiates the Chinese group from the others.

The remaining variables showed differences depending on the group. In the Vietnamese and Chinese groups, `honorifics` had a relatively greater influence. Other variables such as `past tense, mood, and the number of exponents` showed only minor or variable effects across groups. Mood forms were mostly weak or not statistically significant, especially in the Chinese and Mongolian groups. This also contrasts with Ahn et al. (2013), which found that Chinese L1 speakers showed strong priming effects for past tense and mood (conjunctive) endings in a Korean verb priming task, but no priming effect at all for honorific endings.

These results reveal a striking contrast between native and non-native speakers. While the `Ending` variable had the strongest influence for native speakers, word length (i.e., `the number of syllables`) emerged as the most influential factor for non-native speakers. This suggests that native speakers rely more on familiar morphological and semantic patterns (i.e., `Ending`) in verb endings than being affected by word length, whereas non-native speakers are more affected by the physical length of the word during the processing of Korean verbs.

In summary, native speakers were primarily influenced by `Ending`, as well as `ending frequency` and `lexeme frequency` effects. In contrast, non-native speakers were more affected by word length (`the number of syllables), Ending, ending frequency and whole-word frequency`.

Among the non-native speaker groups, the Chinese group displayed several notable distinctions that set them apart from the others. Unlike the other groups, the `Ending` variable had a greater impact than the `number of syllables`, and the effect of `ending frequency` (252) was significantly stronger than that of `lexeme frequency` (69) and `whole-word frequency` (88). Moreover, the influence of `SLIF` (the final syllable identity factor) was minimal in this group, whereas the effect of `honorifics` was relatively large - an opposite trend compared to other non-native groups. There may be several interpretations for why the Chinese group showed such unique patterns, but no single explanation fits clearly within an established framework. The results of this study differ from those of the previously mentioned Kim et al. (2018b) and Ahn et al. (2013), though the exact reasons for these differences remain unclear. Focusing on the results of the current study, in general, the strong effects of the `Ending` variable and `ending frequency` suggest that, the L1 Chinese speakers tend to place more weight on verb endings during Korean verb processing and are more sensitive to `honorific` forms than to final syllables (`SLIF`) or `past tense` markers.

9.2 Ending patterns

The `Ending` (i.e., the combination of ending exponents) variable showed the strongest effects in native speakers and also had a significant impact on non-native speakers. When examining how reaction times were affected by the structure of these exponent combinations, the results generally reflected frequency effects. However, all groups also showed some patterns that deviated from frequency-based predictions, suggesting additional processing factors at play.

A few examples related to the pattern reveal that forms with slower reaction times than expected based on their frequency are mostly : int-hon-16:-(으) 셔 (all groups), fm3:-(아/어) 서, fm5:-(ㄴ/은/는) 데 (notable in the Vietnamese group). These forms tend to be either monosyllabic and thus relatively short, or involve connective endings that may be semantically ambiguous or complex (in particular, fm5:-(ㄴ/은/는) 데). It can be interpreted that the ambiguity in either form or meaning may have led to delayed processing, regardless of their high frequency.

Another type includes forms that showed faster reaction times than expected based on their frequency, such as rm51: -(ㄹ/을) 망정 and rm59: -거들랑. These forms are typically 2–3 syllables long and have salient orthographic features. It is possible that they are processed not as patterned verb endings, but rather as individually recognised and stored single-word-like forms.

Another notable pattern is that all non-native speaker groups showed generally faster reaction times for 'for-hon-past-dec-17: -(으) 셨습니다 and for-hon-past-inq-20: -(으) 셨습니까 endings. This might be because these two forms are frequently presented during Korean language instruction, such as in textbooks, which could have led to faster recognition across all non-native participants.

For non-native speakers, excluding Chinese participants, both ending frequency and whole-word frequency showed similar levels of influence, with ending frequency being slightly stronger overall. Stem frequency, on the other hand, had only a minimal effect across the non-native groups. Based on this pattern, it can be inferred that non-native speakers tend to focus more on verbal `Ending` while simultaneously scanning the whole word to judge its familiarity. However, among Chinese participants, the effect of ending frequency was three times greater than that of whole-word frequency, suggesting a strong reliance on `Ending` in lexical processing. Since stem and whole-word frequencies had comparatively weaker effects, this implies a primary focus on `Ending`. It is particularly interesting that Chinese and Vietnamese, despite being morphologically similar, do not show similar patterns in L2 processing. In particular, L1 Chinese speakers, whose native language is not rich in inflection, appear to focus more on inflectional elements(e.g., honorifics) when processing an inflectionally rich L2 Korean.

Considering the large number of possible `Endings` in Korean, this study selected a set of 92 Endings (including nonword forms) by taking into account various factors such as form, meaning, and frequency. These Endings were then combined with 100 verb stems to create 4000 real words and 1000 nonwords as experimental stimuli. Future research using a wider and more diverse range of lexical items will be important in further exploring the patterns and generalizability of these findings.

9.3 Findings from Nonwords

We aimed to explore morphological, semantic, and phonological effects using nonwords in Korean verb processing. While the results across different nonword conditions showed generally similar patterns in reaction times and error tendencies among non-native groups, there were also some language-specific characteristics observed.

A comparison of the non-native speaker groups shows that Japanese and Mongolian speakers exhibited the slowest reaction times in condition `phoEy`[30], while Chinese and Vietnamese speakers showed the slowest responses in condition `sswpEy`. This may be due to the influence of L1 orthography or phonology—specifically, in conditions `sswpEy`, where the syllables of verb endings were reversed, Chinese and Vietnamese speakers may have been more sensitive to these changes. Interestingly, Japanese speakers also showed relatively slow responses in condition `sswpEy`, following condition `phoEy`, possibly due to the shared use of Chinese characters, which could suggest a similar L1 influence.

[30] As shown in Figure 7.10-7.13, the `Tm` condition elicited the slowest response time. However, since `Tm` is a subsidiary condition of `Tmy`, which includes only a small number of stimuli, `Tm` was excluded from the analysis.

Interestingly, in condition `pho`, the Japanese and Mongolian groups showed slower reaction times for the manipulated endings (i.e., `phoEy`), whereas the Chinese and Vietnamese groups responded more slowly to the manipulated stems (i.e., `phoL`). This pattern may suggest that speakers of agglutinative languages, where verb endings are more morphologically developed, are more sensitive to phonological changes in the endings and therefore tend to focus more on the ending portion than the stem. In contrast, participants with Chinese or Vietnamese as their L1—languages that may rely more heavily on syllable-based processing—likely experienced greater difficulty with nonwords that had manipulated stems, especially when those stems consisted of a single syllable. Compared to manipulated endings, which tend to have more syllables and offer more cues, these distorted stem-based nonwords may have provided fewer clues for lexical decision, making them harder to process. However, further research is needed to confirm whether this reflects a broader typological influence.

Unlike native speakers, who showed the slowest reaction times in condition `Tmy`, including manipulation `Tm`, non-native speakers showed generally faster reaction times for this condition compared to other nonword conditions like `pho, sswpEy, alloSw` and `fm`. This is an interesting finding and may suggest that complex morphological or semantic elements have less of an influence on non-native processing than phonological or syllabic cues.

Interestingly, in all non-native groups, nonwords in condition `fm` (frequent mood) showed slower reaction times, while those in condition `infm` (infrequent mood) were processed significantly faster. Assuming that factors beyond frequency are at play, one likely explanation is the influence of physical forms from orthographic and phonological aspects. In the `infm` condition, many of the items had distinctive 2- to 3-syllable structures, which may have facilitated quicker recognition. In contrast, although the `fm` conditions are frequent mood forms, many were short, monosyllabic items that may have been more difficult to recognise due to their brevity and their manipulated forms. This aligns with the earlier word form analysis, where some high- or low-frequency mood forms were recognised more slowly or quickly than expected compared to other forms of opposite frequency. That is, regardless of frequency effects, it is assumed that morphologically distinctive or short monosyllabic forms—when manipulated in nonword conditions—may have led to such recognition patterns.

In the nonword manipulations using impossible forms, `sag` (one syllable manipulated) and `fag` (full stem/ending manipulated) condition, the `sag` manipulation consistently elicited the fastest responses in both the Chinese and Vietnamese groups. This suggests that their heightened syllable awareness, likely influenced by their L1, may account for this pattern. In contrast, the Japanese and Mongolian groups showed a mix of fastest responses between `sag` and `fag` conditions. Overall, the `sag` and `fag` conditions produced the fastest reaction times across all groups. Following these, condition `phoLEy` consistently showed the next fastest responses in every group. This suggests that among the conditions not involving impossible forms (i.e., `fag` and `sag` manipulations), the `phoLEy` condition, where both stem and ending were manipulated in phonologically similar ways and the polite speech -요/yo/ was added, was the most easily recognised among the manipulated nonword conditions.

We also examined participants' responses depending on the presence or absence of -요/yo/, a very common sentence-final form in Korean. Interestingly, in all nonword conditions—except for the `sswpEy` and `sagLy` conditions—both native and non-native groups showed a facilitation effect in response times when -요 was present. In the case of the `sswpEy` nonword condition, the result may indicate that even in a visual lexical decision task, participants engaged in phonological rehearsal, which made them sensitive to the altered syllable order.

9.4 L1 effects

The four non-native speaker groups can be divided into two categories based on morphological similarity to Korean: Japanese and Mongolian, which are morphologically similar, and Chinese and Vietnamese, which are morphologically different. By comparing these distinct L1 backgrounds, we aimed to investigate the similarities and differences in how learners process Korean verbs.

Looking at the responses to specific ending components, it was generally observed that the addition of `honorific` or `past tense` forms tended to slow down response times. As for `mood` forms, they did not show statistically significant effects in the Mongolian and Chinese groups, while the Vietnamese group responded faster when `mood` forms were used.

When examining the impacts of `SLIF` (i.e., Speech Level + Illocutionary Force), Japanese, which is typologically the most similar to Korean among the L1 groups, shares similar concepts regarding `honorifics` and `SLIF`. This raises the question of how such influence differs across groups. In Korean native speakers, when comparing the AIC weights of all predictors, `SLIF` emerged as the third most influential factor, following `Ending` and `ending frequency`. In the case of the Japanese group, however, the AIC weight pattern was similar to those of the other non-native groups, showing no particularly distinctive pattern. The AIC weight for `honorifics` did not stand out among native Korean speakers, and similarly, in the Japanese group, it was at a comparable level to other variables such as `lexeme frequency`, `stem-final segment`, and the `number of exponents`, showing a pattern similar to that of the other non-native groups.

The Chinese group, on the other hand, showed several distinctive patterns. Notably, it was the only non-native group in which the `Ending` (619) had a higher AIC weight than the `number of syllables` (493)—a pattern shared with native Korean speakers. Furthermore, the weight of `honorifics` (65) was higher than in other non-native groups, while the weight of `SLIF` (27) was markedly lower compared to all other groups. Moreover, and more importantly, the frequency effects for `whole-word` (88) and `stem` (69) frequencies were similarly low, while the effect of `ending frequency`(252) was notably high. This indicates a strong concentration of processing focus on the verb endings. The findings of the present study differ in several respects from previous research on L2 lexical processing involving native Chinese speakers. For example, in Kim et al. (2018b), where the majority of participants were L1 Chinese speakers, the most prominent predictor was Eojeol (i.e., whole-word) frequency, followed by word length. In Hwang et al. (2008), while advanced Japanese learners demonstrated grammatical judgment abilities for Korean verbs comparable to native Korean speakers, advanced Chinese learners did not perform as well as the Japanese group—an outcome interpreted as reflecting potential L1 effects. Similarly, Portin et al. (2008) found that L1 Hungarian participants showed processing differences between monomorphemic and inflected L2 Swedish nouns, whereas L1 Chinese participants did not—suggesting reduced sensitivity to inflection, possibly due to L1 background. These results stand in contrast to the response patterns observed among the Chinese group in the current study.

In summary, although there are slight differences, the three non-native groups (Japanese, Mongolian, and Vietnamese) generally showed similar patterns in the AIC weight distribution across variables. The Chinese group, in contrast, exhibited a distinct pattern. One notable feature of the Vietnamese group was that, unlike the others, the `number of exponents` had the second highest AIC weight after `vowel harmony`. In contrast, the `number of exponents` was not statistically significant in the Chinese group.

When narrowing the analysis to two typologically similar pairs (Japanese- Mongolian and Chinese-Vietnamese), we find that while all groups differ from native speakers, the Japanese- Mongolian pair shows somewhat similar patterns in AIC weight distribution across variables. In contrast, the Chinese-Vietnamese pair does not exhibit such similarity. From a morphological perspective, therefore, it still remains unclear whether typologically similar languages share characteristic patterns in lexical processing. For instance, while the results observed in the Japanese and Mongolian groups regarding verb endings and those in the Chinese and Vietnamese groups related to syllabic features may suggest some degree of L1 effects, the overall AIC results across variables show divergent patterns. In particular, the Chinese and Vietnamese groups—despite both being isolating languages—show substantial differences in AIC outcomes. This suggests that even typologically similar languages can exhibit distinct patterns in L2 lexical processing, likely influenced by a wide range of interacting variables.

Chapter 10

Computational Modeling

This chapter investigates whether a recent computational theory of the mental lexicon, the Discriminative Lexicon Model (DLM, Baayen et al., 2019; Heitmeier et al., 2025) can learn to understand and produce Korean verb forms. The DLM provides a modeling framework in which stems and exponents do not play a role. The model does not attempt to segment a word form into constituents, and likewise, it does not assemble words from constituent parts for production. The DLM has been applied successfully to a range of typologically different languages, including Estonian (Finno-Ugric Chuang and Baayen, 2020), Maltese (Semitic, Nieder et al., 2022), Kinyarwanda (Bantu, van de Vijver et al., 2021), Mandarin Chinese (Sino-Tibetan, Chuang et al., 2022), Indonesian (Austronesian, Denistia and Baayen, 2022), as well as to the Germanic languages English (Baayen et al., 2019; Heitmeier and Baayen, 2021) and German (Heitmeier et al., 2021).

The DLM is a model of the mental lexicon that posits end-to-end mappings between words' forms (represented by numeric vectors derived from words' spoken or written forms) and their meanings (also represented by numeric vectors, taking inspiration from distributional semantics, see, e.g., Landauer and Dumais, 1997; Mitchell and Lapata, 2008; Mikolov et al., 2013). The DLM research programme attempts to integrate insights from distributional semantics with Word and Paradigm Morphology (Matthews, 1974; Blevins, 2016). An introduction to the DLM is provided by Heitmeier et al. (2025).

Comprehension is modeled with mappings from form to meaning, and production with mappings from meaning to form. These mappings can be set up in several ways. First, the learning rules of Widrow and Hoff (1960) can be used to learn a mapping word by word, using simple networks with input and output units and no hidden layers. Second, instead of incremental learning, one can use the mathematics of multivariate multiple regression to estimate the mapping (now represented by the matrix of beta coefficients) that is expected for the endstate of learning. This endstate of learning (EOL) describes the mapping that would be obtained if incremental learning would be repeated an infinite number times on the training data. EOL is not sensitive to word frequency (Heitmeier et al., 2021, 2023). To bring usage frequency into the model, frequency-informed learning (FIL) is available. This method for calculating mappings introduces weights that are proportional to words' frequencies of use. As shown by Heitmeier et al. (2023), FIL performs very similar to incremental learning with the Widrow-Hoff learning rule.[31]

In this study, we make use of straightforward linear mappings, using both EOL and FIL. We do not make use of incremental learning, for two reasons. First, step-by-step incremental learning is currently prohibitively expensive for large datasets. Second, and more importantly, for properly modeling incremental word-by-word learning, we need detailed longitudinal corpora that document language use from early childhood into young adulthood and beyond. In the absence of such corpora, we decided to study the theoretical learnability of Korean verbal morphology with EOL and FIL. For detailed discussions of incremental learning versus endstate learning, the reader is referred to Heitmeier et al. (2023).

This study first evaluated the DLM's performance as a memory for previously encountered word forms. It then investigated the productivity of the Korean verb system based on corpus-based analyses, and assessed whether the model's predictions account for processing complexity in a visual lexical decision task, using a small dataset from native speakers in chapter 5.

[31] It is also possible to set up mappings with deep neural networks. Such mappings, explored by Heitmeier et al. (2025), are beyond the scope of this study.

10.1 A DLM-based theory of Korean verbal morphology

The modeling of Korean morphology reported in this section has as its overarching goal to clarify whether the DLM provides a framework that is sufficiently powerful to understand and produce Korean inflected verbs. Given the complexity of Korean verb forms(see the details in capter 3.1), which combine a rich morphology with a rich variety of phonological processes, the task faced by the DLM model is not a trivial one.

More specific goals are the following. First, we are interested in which sublexical unit of representation for Korean verb forms is optimal for obtaining reasonably accurate mappings between form and meaning.

Second, although all kinds of verb forms are not attested in Korean corpora, ‘unseen’ forms are readily interpretable for Korean speakers. Clearly, the verb system is productive. This raises the question of whether training the model on forms attested in written or in spoken corpora is sufficient to obtain good generalization performance for the forms that are not attested in these corpora. In other words, the question we address here is whether the DLM captures the productivity of the Korean verb system[32].

Third, we compare model performance using simulated semantic representations that maximally distinguish between inflectional features, with model performance using empirical embeddings calculated from corpora with the FastText algorithm (Bojanowski et al., 2017). This allows us to contrast a model implementing linguistic contrasts with a model that is informed by usage (as gauged from written texts). In section 10.2, the two models will be used as alternative theories for predicting lexical decision latencies.

[32]Specifically, this study only examines sentence-final verb forms, excluding other types of Korean verb forms that represent various types of mood (see the definition in the relevant chapter 2.2).

10.1.1 Materials

We selected 462 general verbs used in spoken and written Korean. These verbs are supposed to have been acquired by Korean learners taking the TOPIK (Test of Proficiency in Korean) level 3-4. For each of these verbs, we constructed their basic paradigm with 59 types of sentence-final ending forms. The total number of paradigm forms in our dataset therefore is $59 \times 462 = 27,258$. In this study, we do not consider the many different clausal connectives and modifying forms with various moods[33] that can be expressed in the Korean verb. Even with these restrictions in place, as we shall see below, large numbers of possible forms in our dataset are not attested even in large corpora of Korean.

More than 4000 possible verb ending exponents have been observed in a written corpus with 15 million words[34] (Kang and Kim, 2009). Choi (2012) analysed another written corpus (10 million Eojeol of the 21 Century Sejong corpus) and observed 5,682 different verb endings (i.e., distinct surface forms of sequences of exponents). Interestingly, Choi (2012) observed a maximum of four exponents realised in a single verb form, with a relative frequency equal to 0.0009%. On the other hand, verb forms with a single exponent account for around 86.7% of word tokens, and words with two exponents represent 13.2% of the tokens. In other words, only about 0.1% of verb forms have three or four exponents.

Han (2020) analysed a corpus of spoken Korean (21C Sejong)[35], which comprises 14,418 Eojeols (words) of in all 1,991 verb forms (only sentence-final ending verb forms)[36]. Four exponents (어 /ʌ/, and its allomorphs 아 /a/ 야 /ya/; -지 /dzi/; 구 /gu/; and 다 /da/ and its allomorph ㄴ다 /nda/) accounted for more than 70% of all the endings[37].

[33] See the definition of mood in chapter 2.2

[34] Word tokens are defined following the spacing conventions of Korean 어절 (Eojeol). 'Eojeol' in Korean is a word unit which is separated by spaces from other words.

[35] In the literature, it refers to Contemporary Korean Colloquial Corpus of 21st Century Sejong Plan by the National Institute of the Korean Language(released 2010).

[36] This study estimates only the sentence-final endings, so the total number of full verb endings combined with prefinal exponents remains uncertain.

[37] In this regard, Kim (2013) found that, in an analysis of 5,899 sentences of quasi-spoken Korean, 14.5% ended with non-final verb endings (i.e., 반말체 (Bnamal, casual speech) with connective or nominalizing endings), and about 9% ended with non-verbal elements such as nouns or adverbs.

The study identified a total of 81 distinct sentence-final verb endings.

These studies show that the space of possible Korean verbs is very sparsely populated in actual use. Korean is not the only language characterised by usage sparsity, Karlsson (1986) discusses the same phenomenon for Finnish. Loo et al. (2018b,a) reports for Estonian that the observed number of forms inflected for case and number that are actually attested for a noun is a good predictor for the processing cost of that noun as gauged by several psycholinguistic tasks.

Thus, a substantial part of the forms in our dataset are likely to be very low-frequency. To assess this possibility more precisely, we inspected two corpora, koTenTen18 and NIKL_SPOKEN (v1.0), henceforth referred to as koTenTen and NIKL. The koTenTen corpus is based on web internet texts crawled by SpiderLing from December 2017 to April 2018.[38] The number of words in koTenTen is more than 1.6 billion. NIKL is a spoken corpus collected by the Korean National Institute of Korean Language.[39] The texts collected by NIKL include daily life conversations, speeches and lectures, as well as broadcasts and quasi-spoken data such as TV drama scripts. According to the documentation of the two corpora, we may assume that NIKL samples registers of spoken Korean, whereas koTenTen samples registers of written Korean.

We annotated each verb form in our dataset for whether it is present in these corpora. About 52% of the words in our dataset are present in both corpora, and 33.6% are absent in both corpora. Figure 10.1 presents the distributions of the inflectional features of the words that are present in both corpora. The features used most commonly are PLAIN HONORIFICS, PRESENT TENSE, POLITE SPEECH LEVEL, and DECLARATIVE ILLOCUTIONARY FORCE. Overall, the proportions of use are remarkably similar across the two corpora. There are slightly more tokens of PLAIN SPEECH and of DECLARATIVE SPEECH in the koTenTen corpus, as expected given that koTenTen contains more written

[38] https://www.sketchengine.eu/kotenten-korean-corpus/

[39] According to information provided by the corpus development organisation, the corpus size comprises 15,052 hours of spoken language(approx. 75 million words) (such as lectures, debates, interviews, and conversations) and 15,591,197 words of a quasi-spoken language (such as drama and play scripts). https://corpus.korean.go.kr/main.do

Korean, a register which makes heavy use of the PLAIN SPEECH LEVEL.

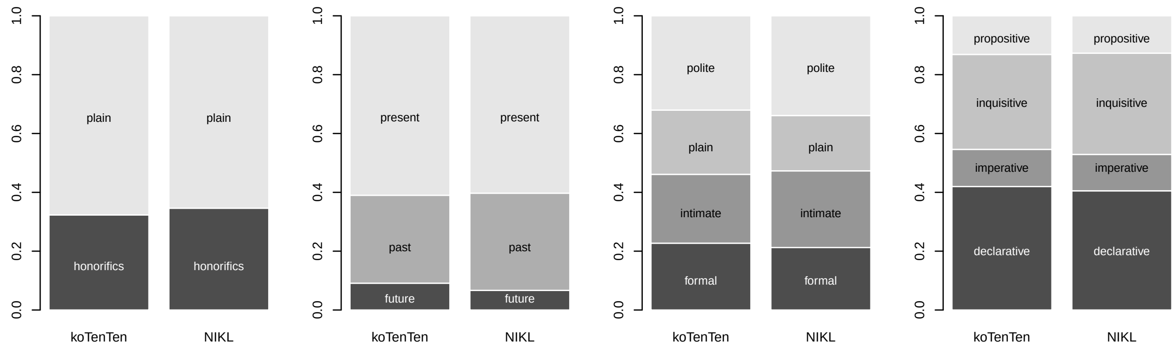


Figure 10.1: Distributions of inflectional features for honorifics, tense, speech level, and illocutionary force for words that are present in both the koTenTen and NIKL corpora.

Having introduced the dataset that we used for modeling the Korean verb, we next explain how we represented words' forms and their meanings.

10.1.2 Representing word form

We represent words' forms by means of numeric vectors that specify, using 0/1 encoding, which form features jointly define a word's form. As elementary form features, we consider pairs of Hangul characters, triplets of phones, and pairs of syllables. For a given choice of elementary feature, the form vectors of words are brought together as the row vectors of a matrix $\boldsymbol{C}$, which for our dataset is a 27,258 $\times$ 4,222 matrix when using bi-Hangul form units, a 27,258 $\times$ 2,211 matrix when using triphones, and a 27,258 $\times$ 4,770 matrix when using bi-syllables.

10.1.3 Representing word meaning

Each word form is associated with a numeric vector of real-valued numbers that together represent its meaning. These vectors are brought together into a semantic matrix $\boldsymbol{S}$. In this study, we compare two kinds of semantic vectors, simulated vectors that reflect inflectional contrasts, and empirical vectors calculated from corpora.

Simulated semantic vectors

When simulating semantic vectors, we set the length of these vectors to be identical to the lengths of the form vectors. Thus, we considered three simulated $\boldsymbol{S}$ matrices, each with the same dimensions as the three form matrices $\boldsymbol{C}$.

The simulated semantic vector of a Korean verb is constructed by summation of the simulated vectors of the verbal lexeme and its inflectional features.[40] Thus, the verb form 꾸미셨습니다 /k͈umisʰyʌt̚s͈ɯmɲida/ is constructed as follows:

$$\overrightarrow{\text{꾸미셨습니다}} = \overrightarrow{\text{decorate}} + \overrightarrow{\text{hon}} + \overrightarrow{\text{past}} + \overrightarrow{\text{for}} + \overrightarrow{\text{dec}} + \overrightarrow{\boldsymbol{\varepsilon}},$$

with $\overrightarrow{\boldsymbol{\varepsilon}}$ as a random vector with a small amount of noise representing word-specific usage (see Nikolaev et al., 2023, for discussion of word-specific error vectors).

Empirical semantic vectors

For the empirical vectors, we collected FastText (Bojanowski et al., 2017) embeddings for our words.[41] FastText vectors are available only for a relatively small subset of 8881 words. Among this set of words, 1680 are homographs (i.e., words with the same Hangul forms, but different meanings). The length of the FastText vectors is 300.

10.1.4 Mappings between form and meaning

For comprehension, we calculated a mapping $\boldsymbol{F}$ (using either EOL or FIL) such that $\boldsymbol{CF} = \boldsymbol{S}$. For production, we calculate a mapping $\boldsymbol{G}$ such that $\boldsymbol{SG} = \boldsymbol{C}$ (for technical details, see Heitmeier et al., 2025). Given the estimated mappings $\boldsymbol{F}$ and $\boldsymbol{G}$, we obtain predicted semantic vectors $\hat{\boldsymbol{S}} = \boldsymbol{CF}$ and predicted form vectors $\hat{\boldsymbol{C}} = \boldsymbol{SG}$.

[40] A semantic vector is simulated with random numbers that are sampled from a normal distribution.
[41] `https://dl.fbaipublicfiles.com/fasttext/vectors-crawl/cc.ko.300.vec.gz`

For comprehension, model accuracy is evaluated by comparing the predicted vectors $\hat{\boldsymbol{S}}$ with the gold standard vectors $\boldsymbol{S}$. The i-th predicted row vector $\hat{\boldsymbol{s}}_i$ of $\hat{\boldsymbol{S}}$ is taken to be correct if the correlation of $\hat{\boldsymbol{s}}_i$ with $\boldsymbol{s}_i$ is greater than any of the correlations of $\hat{\boldsymbol{s}}_i$ with any other row vector $\boldsymbol{s}_j, j \neq i$. In other words, a mapping $\boldsymbol{F}$ takes a point in form space and places it in semantic space. We take this new point in semantic space to be correct if it is closer to the point in space representing the word's actual meaning than to any of the other points in space that represent other words'meanings.

For production, a predicted form vector $\hat{\boldsymbol{c}}$ specifies the amount of support that each of the elementary form features receives from the semantics. This vector, however, does not provide information on how to properly order the features that are appropriate for a given word. In other words, the vector specifies which Hangul pairs, triphones, or bi-syllables are present, but it does not provide information about where in a word they occur. A separate algorithm is needed that places the form features in the appropriate order.

The **JudiLing** package (Luo et al., 2021; Heitmeier et al., 2025) provides an algorithm for doing this. This algorithm first learns what the most probable features are for the different positions in the word, and then weaves the most likely features at the different positions together into proper word forms. From the resulting set of word form candidates, it selects that word form for articulation that best matches the meaning ($\boldsymbol{s}$) that is to be expressed. The word form predicted by the model is evaluated as correct if and only if it exactly matches the gold-standard word form.

Two technical details that we need to introduce are the following. When weaving features together into word forms, it is often necessary to consider features with low semantic support, for instance, for novel features that are absent during training. A threshold parameter $\boldsymbol{\theta}$ specifies the minimum support that a feature should have for it to be taken into consideration. Lower values of $\boldsymbol{\theta}$ lead to larger sets of form candidates, and very low values of $\boldsymbol{\theta}$ come with considerable computational cost. In situations where several target cues of a given word form do not receive sufficient semantic support, it is also possible to also specify the number of weakly supported features (with supports smaller than $\boldsymbol{\theta}$) that one would like the algorithm to consider during path-building. The parameter that specifies the number of features to be included despite having supports less than $\boldsymbol{\theta}$ is referred to as the 'tolerance' parameter.

10.1.5 Computational experiments

Experiment 1

Experiment 1 investigates the accuracy of EOL linear mappings between form and meaning, using simulated semantic vectors, for comprehension and production, across three different form vectors: vectors specifying which pairs of adjacent Hangul characters are present (bi-Hangul), which triplets of adjacent phones (triphones) are present, and which pairs of adjacent syllables (bi-syllable) are present. Table 10.1 summarizes model accuracies, for models trained and evaluated on the full dataset. All accuracies are high, and greater than 97%.

Table 10.1: Model performance, evaluated for the full dataset, of comprehension and production, with bi-Hangul symbols, triphones, and bi-syllables as form cues, and simulated semantic vectors representing meanings.

	comprehension	production	number of cues
bi-Hangul	**99.4**%	97.9%	4222
triphone	98.4%	**98.9**%	2210
bi-syllable	99.2%	97.8%	4790

For comprehension, the models with bi-Hangul and bi-syllable form vectors slightly outperform the model with triphone vectors. The slightly better performance of bi-Hangul and bi-syllable representations is perhaps unsurprising, given that the number of different units for these form representations is roughly twice that of the triphone representations (see the right-most column of Table 10.1). Interestingly, the number of bi-syllable cues exceeds that of Hangul digraphs by more than 500, reflecting that the Hangul writing conventions neutralize many differences in pronunciation that are due to the phonological adjustment rules discussed above. For production, on the other hand, the triphone model is slightly more accurate than the other two models, even though the number of triphone cues is much smaller.

Experiment 2

The goal of Experiment 2 is to clarify whether our linear mappings properly capture the productivity of Korean verb inflection. Experiment 1 showed that these mappings provide good 'memories' for how to produce meanings from forms, or forms from meanings. But are these mappings refined enough to also predict, with reasonable accuracy, how to understand novel forms, and how to produce forms given novel conceptualizations?

Experiment 1 clarified that models trained and tested on the full dataset perform with high accuracies, irrespective of which form encoding is selected for the form vectors. As bi-syllable representations come closer to the representations driving articulation in speech production (for the syllable as the unit driving speech production, see, e.g., Levelt et al., 1999) and perform well, in the following experiments, we selected bi-syllables as the form features for the simulation experiments reported in the remainder of this study.

In Experiment 2, as in Experiment 1, we used simulated semantic vectors for words' meanings. To clarify whether our model is capable of recognizing and producing previously unseen forms, we held out a subset of the verb forms in our dataset as testing (out-of-bag) data, trained the model on the remaining (in-bag) data, and evaluated model performance on both the in-bag and out-of-bag data. For this cross-validation procedure, rather than randomly splitting the data, we sought to better approximate the real learning experience of native Korean speakers by basing our splitting on whether word forms are attested in the two Korean corpora introduced in section 10.1.1.

We ran two cross-validation studies for our dataset. For the first cross validation study, we assigned verb forms to the in-bag training set if they occurred in the koTenTen corpus, and to the out-of-bag test set otherwise. For the second cross-validation study, the criterion for assignment to the in-bag training set was occurrence in the NIKL corpus.

Table 10.2 clarifies that the sizes of the in-bag training and out-of-bag held-out data are similar for the two simulations. Accuracy on the in-bag training data is above 99% for comprehension and above 98% for production. As expected, for the held-out test data, accuracy drops, for comprehension to around 90% and for production to around 80% and 70% for koTenTen and NIKL respectively. Overall, the testing results are better with koTenTen than with NIKL, possibly due to koTenTen's larger training dataset. In addition, the model is more successful in recognizing than producing unseen forms, an asymmetry that is commonly observed in acquisition and modeling studies (Gershkoff-Stowe and Hahn, 2013; Chuang et al., 2020b,a).

Table 10.2: Cross validation results using bi-syllables as sublexical cues, and with simulated semantic vectors.

	Size		Comprehension		Production	
	train	test	train	test	train	test
koTenTen	16,398	10,860	99.6%	91.7%	98.2%	80.4%
NIKL	15,231	12,027	99.5%	88.7%	98.5%	71.0%

Experiment 3

So far, the mappings we used were all estimated with endstate learning without taking the different frequencies of occurrence of individual words into account. That is, the mappings between form and meaning that we estimated above are agnostic with respect to words' frequencies of occurrence. Word frequency, however, has a well-known robust effect on lexical processing, and has been discussed in a number of word recognition models (e.g., McClelland and Rumelhart, 1981; Norris, 2006). For the models that we implemented above, however, frequency has no effect on learning whatsoever. These models are purely type-based.

Nevertheless, it is to be expected that actually frequency of use does leave traces on the mappings, affecting connection weights between cues and outcomes, making these very robust for high-frequency words at the cost of low-frequency words. When incremental learning is used, and applied to data reflecting actual language use, mappings are updated one word token at a time. In this way, high frequency words will receive more training than low frequency words, analogous to the fact that native speakers encounter high frequency words more often than low frequency words, and learn these words better.

However, given that incremental learning is computationally costly, and given that detailed information on the order of word occurrences encountered by L1-learners of Korean is unavailable to us, a more efficient way of estimating mappings while taking frequency into account is required. Here, we therefore use frequency-informed learning (Heitmeier et al., 2023). Given that about 40% of the words in our dataset are absent in the two corpora, we added one to the frequency of every word to ensure that all zero-frequency words are encountered for training, albeit with the lowest-possible non-zero frequency. Model performances of FIL are presented in Table 10.3.

Model accuracy is evaluated in two ways. First, we can inspect how many word types are correctly predicted by the model, without taking the frequencies of these types into account. We refer to this as 'type accuracy'. However, language users know higher-frequency words better than lower-frequency words. We therefore also evaluate accuracy over tokens, effectively weighting type accuracy for the number of tokens of the types. We refer to this as 'token accuracy'. For frequency-informed learning, we do not evaluate on held-out data, as inspection of the model for lower-frequency words is already informative about how the model will generalize to unseen types. Table 10.3 summarizes the results we obtained.

Table 10.3: Model performance of comprehension and production when using frequency-informed learning, in combination with simulated semantic vectors. The column `Type acc` lists accuracies that are calculated over word types; token-based accuracies are presented in column `Token acc`.

	Comprehension		Production	
	Type acc	Token acc	Type acc	Token acc
koTenTen	40.3%	98.4%	41.5%	99.0%
NIKL	45.2%	98.3%	47.6%	99.6%

Type-based accuracies of FIL are substantially worse than the accuracies of models that have no access to frequency (cf. Table 10.1 and 10.2). The lower accuracies of FIL to a large extent result from the fact that low frequency words are not well learned, unsurprisingly.

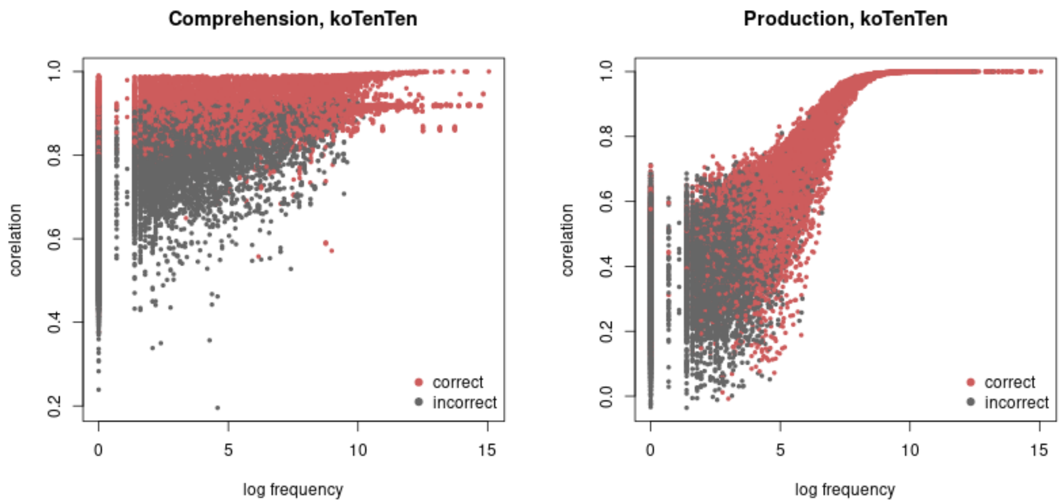


Figure 10.2: The relationship between frequency (on logarithmic scales) and model performance for comprehension (left) and production (right). The vertical axis represents the correlation value between predicted and targeted semantic vectors. Dots in red are words that are recognized/produced correctly, and those in dark gray are words that are recognized/produced incorrectly.

As shown in Figure 10.2, higher frequency words (based on koTenTen frequency) tend to be predicted with greater accuracy, indicated by higher correlation values on the vertical axis. The relation between frequency and correlation is especially strong for production, suggesting that the mapping from meaning to form is modulated more strongly by frequency of use. (The results based on the NIKL corpus are very similar and are therefore not presented here.) The spearman correlation between log frequency and correlation values is 0.63 and 0.77 for comprehension and production respectively. Moreover, the predicted semantic vectors of correctly recognized word forms (red dots) are more highly correlated with the gold standard vectors than those of incorrectly recognized word forms (gray dots), as the former has significantly higher correlation values than the latter according to a Wilcoxon signed-rank test ($p < .0001$). The same holds for production. This pattern of results corresponds well with the effect of frequency, as low frequency words are usually difficult words that native speakers fail to recognize or produce. Furthermore, as there are a lot of low frequency words — often, in a text or corpus, about half of the word types have a token frequency equal to 1 — accuracy thus suffers when it is calculated over word types. When calculated over word tokens (column "Token acc", Table 10.3), accuracies are very high: the recognition and production of high-frequency words is often quite accurate.

Experiment 4

In the preceding experiments, we simulated semantic vectors. The way in which we created semantic vectors for inflected forms assumes that content words are all unrelated, and that any pair of semantic functions is also equally uncorrelated. In other words, what we have done is set up a system of contrasts that mirrors the way in which featural contrasts are set up in realizational morphology and related structuralist approaches to morphology. For instance, instead of having forms that are [+honorific] and [-honorific], we have semantic vectors for honorific and plain forms that are orthogonal. How the specifications for honorifics relate to those for tense or speech level is left unspecified in realizational theories – typically, it is assumed that features are independent and have to be realized jointly. However, to our knowledge, whether feature values make independent contributions that do not depend on the feature values of other inflectional features has not received systematic reflection.

Distributional semantics (Landauer and Dumais, 1997; Mikolov et al., 2013; Westbury et al., 2015) makes it possible to investigate to what extent the way we constructed semantic vectors for inflected words is in accordance with actual language use, conditional on the corpus-based embeddings being sufficiently accurate. In the preceding simulations, we have set up a maximally orthogonal, distinctive, set of contrasts in a hypothetical semantic space. But the realities of the inflectional semantics of Korean may be much more intricate. For instance, we have seen that certain combinations of inflectional features are not possible—for example, propositive illocutionary force combined with past tense. In our dataset, such impossible combinations do not occur. However, there is nothing in the way we set up semantic vectors mathematically that would stand in the way of building vectors for such combinations of features. The way in which we constructed simulated vectors is blind to semantically infelicitous combinations of inflectional vectors.

Importantly, the actual semantics of inflection can be much more nuanced and intricate (e.g., the case of Korean verbal inflection, in particular) than suggested by simple featural oppositions. For instance, in the high-dimensional embedding space of English, the vectors that when added to a singular result in the corresponding plural show clear differences by semantic class (Shafaei-Bajestan et al., 2022a,b). A model that knows only about featural oppositions may therefore be less cognitively valid than a model that builds on empirical embeddings (see also Nikolaev et al., 2023; Chuang et al., 2023).

To explore model performance with empirical semantic vectors, we collected the FastText (Bojanowski et al., 2017) semantic vectors that are available for our words.[42] Given that many of our verb forms do not occur in the two corpora studied above, it is unsurprising that FastText vectors are available for a relatively small subset comprising 8881 words. Among this set of words, 1680 are homographs.

How do the linear mappings perform when we replace the simulated semantic vectors with fastText word embeddings? Evaluation of model performance on the subset of 8881 words for which fastText embeddings are available, using bi-syllables and end-state learning, revealed accuracies at 65.7% for comprehension and 82.4% for production. Cross-validation on a held-out dataset consisting of 800 word forms chosen such that no unseen form features or unseen lexemes were included, yielded worse accuracies as compared to the ones with simulated vectors. Comprehension accuracy for the held-out data was at 44.6% and production accuracy was at 58.5%.

There are several reasons for why model performance for held-out data is better with simulated vectors than with empirical FastText vectors. First, the dataset on which we evaluate the model is much smaller, due to FastText vectors being available only for a subset of our words. Second, FastText embeddings are derived from text types

[42] `https://dl.fbaipublicfiles.com/fasttext/vectors-crawl/cc.ko.300.vec.gz`

(e.g., newspapers, Wikipedia) that predominantly make use of written verb forms rather than spoken forms. But it is primarily the spoken forms that most clearly encode the social dimensions of communication that Korean realizes by means of speech levels and subject honorifics. As a consequence, algorithms such as FastText likely may not get sufficient exposure to examples of the more complex Korean verb forms that are found in conversational Korean. For instance, FastText may have encountered more forms from the plain speech level, and fewer forms from the formal, polite, and intimate speech levels. By contrast, our simulated vectors encode maximally orthogonal contrasts for all speech levels.

A third consideration is that Korean tightly packs the social dimension of communication into the verb, without leaving any traces in the context of the verb. By way of example, consider the following three sentences.

(1) 날씨가 정말 안 좋았다.

nalɕi-ga tsʌŋmal an tso-at̚-t̤a
weather-NOM really not good-PAST-PLAIN&DECLARATIVE
The weather was really bad.

(2) 태풍에 그 나무가 쓰러지는 것을 SEE.

tʰɛpʰuŋ-e kɯ namu-ga s̤ɯrʌdzi-nɯn kʌsʰ-ɯl SEE(one of the verbs in the list below).
typhoon-by(INS) the tree-NOM-falldown-NM thing(scene)-ACC SEE.
(subject omitted) saw the tree falling down by the storm.

(3) 태풍이 많은 피해를 가져왔다.

tʰɛpʰuŋ-i man-ɯn pʰihɛ-rɯl kadzʌwa-t̚-t̤a.
typhoon-NOM alot-NM demage-ACC bring-PAST-PLAIN&DECLARATIVE.
The storm caused(brought) a lot of damage.

In the above sentences, the final verbs(i.e., the last word before the full stop in each sentence) can be expressed using different verb forms. For example, in sentence (2) below, the verb 'see' ('to see', 'saw' in the context) may appear in various forms depending on speech levels and honorifics. Similarly, the verbs in the other two sentences (1) and (3) can also be transformed into comparable alternative forms, just like in sentence (2).

보았다	poat̚t͈a	plain speech style	non-honorific (written form)
봤다	pwat̚t͈a	plain speech style	non-honorific (contracted speech form)
보았어	poas͈ʌ	intimate speech style	non-honorific (written form)
봤어	pwas͈ʌ	intimate speech style	non-honorific (contracted speech form)
보셨어	pos^{h}yʌs͈ʌ	intimate speech style	subject honorific
봤어요	p^{w}as͈ʌyo	polite speech style	non-honorific
보셨어요	pos^{h}yʌs͈ʌyo	polite speech style	subject honorific
보았습니다	poat̚s͈ɯmɲida	formal speech style	non-honorific (written form)
봤습니다	pwat̚s͈ɯmɲida	formal speech style	non-honorific (contracted speech form)
보셨습니다	pos^{h}yʌt̚s͈ɯmɲida	formal speech style	subject honorific

Although the larger discourse in which words are embedded may contain more information about the social dimensions (e.g., honorific marker and speech level) of Korean communication, the distributional hypothesis — you shall know a word by the company that it keeps (Firth, 1968) — is at a disadvantage when it comes to incorporating the subtleties of the social determinants of the Korean verb.

Although undoubtedly the FastText embeddings for Korean are far from perfect, they may nevertheless capture important aspects of the inflectional semantics of Korean verbs. Following studies on Mandarin Chinese and Russian (Shen and Baayen, 2022; Chuang et al., 2023), we carried out a cluster analysis using t-distributed stochastic neighborhood embedding (t-SNE Maaten and Hinton, 2008). This analysis reveals that some structure is present even when we consider the embeddings in only two dimensions. The input to this analysis of FastText vectors was the set of all unique word forms for which we have a FastText embedding. For homographs, since all different meanings of a given form share the same form frequency in corpora, it is not possible for us to determine the dominant meanings. We therefore randomly selected one of the possible feature sets defining one of the meanings of a given form to represent that word form.

The result of the t-SNE clustering is shown in Figure 10.3. The points in each of the four panels represents the same data points, representing the words in our dataset. The color coding differs between the panels. The upper left panel differentiates between words with subject honorifics (red) and words without subject honorifics (black). The words with honorifics cluster more to the left mid-center of the scatterplot. The upper right panel uses color-coding to highlight the tense. As there are very few future forms, the main distinction is between present-tense forms, more prevalent in the upper half of the plot, and past-tense forms, which dominate in the lower half of the plot.

The lower-left panel highlights the speech levels, and the lower-right panel the illocutionary force. With respect to speech level, plain forms show a wide scatter, which may be due to the multifunctional use of plain forms in both speech and writing. Formal forms tend to have more negative values on the vertical axis, polite forms show intermediate values centred around zero, and the intimate forms predominate for larger positive values on this axis. Turning to illocutionary force, declarative forms dominate in the lowest part of the plot region, whereas inquisitives and imperatives are more prevalent in the centre of the plot.

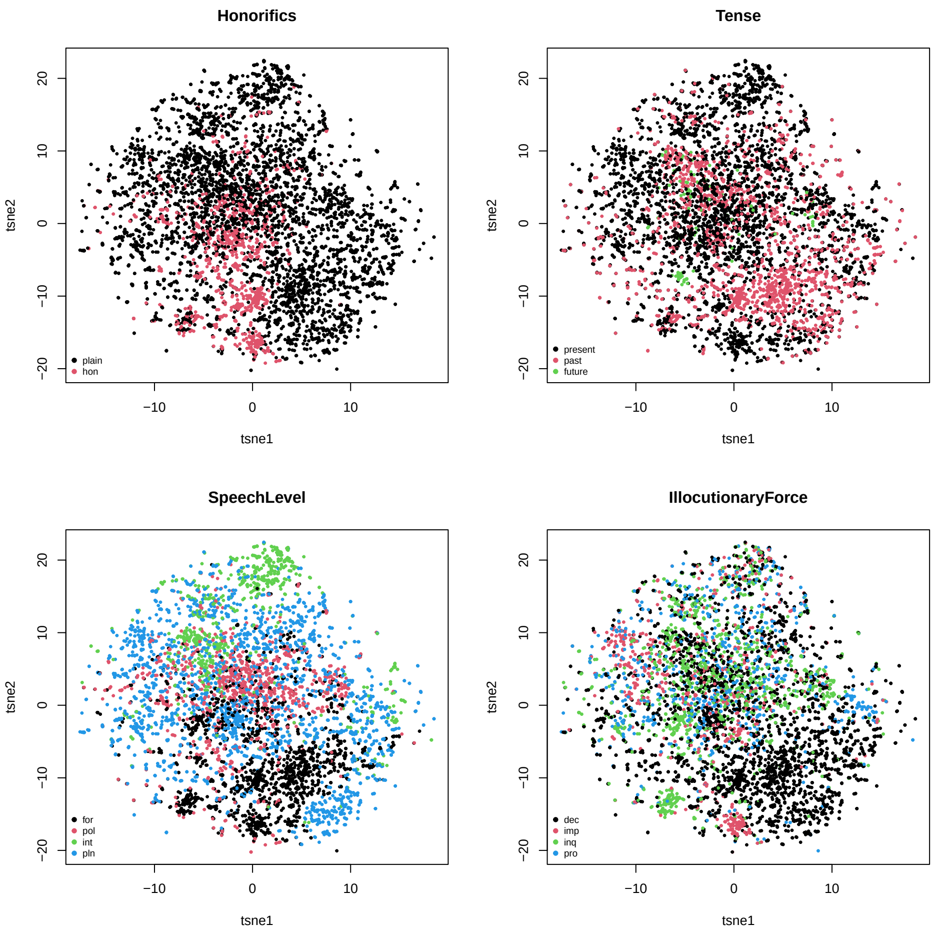


Figure 10.3: The t-SNE scatter plots that presents the projection of the high-dimensional semantic space constructed by fastText to a 2D plane.

The dimensions of the tSNE analysis are only partially interpretable. The upper left panel suggests that honorific forms have negative values on the first dimension, and the lower left panel suggest that intimate forms have higher values on the second dimension. Nevertheless, if there is a clearly distinguishable clustering, it is likely that a specific inflectional ending corresponding to that combination can be inferred. For instance, in the same lower central region across all panels, form clusters appear for honorifics, present tense, formal speech and imperative(request) illocutionary force, suggesting that a specific verb ending representing this form, such as -(으) 십시오, can be inferred.

However, since subtle similarities in a high-dimensional space may be invisible to clustering with t-SNE (see, e.g., Stupak and Baayen, 2022), we also examined how well the values of an inflectional dimension can be predicted from the FastText embeddings, using linear discriminant analysis (LDA) with leave-one-out cross-validation. Interestingly, except for illocutionary force, the FastText embeddings can predict the value of a semantic class with above 80% accuracy, well above a majority class baseline classifier (see Table 10.4).

Table 10.4: LDA leave-one-out cross-validation accuracy and majority baselines. Except for illocutionary force, FastText embeddings are surprisingly good predictors of inflectional features.

inflectional type	accuracy LDA	majority baseline
honorifics	0.89	0.79
tense	0.88	0.73
speech level	0.83	0.32
illocutionary force	0.47	0.42

To obtain further insight into the relation between endings and semantics, we investigated how well, for the present dataset, a linear discriminant analyses with leave-one-out cross-validation can predict verb endings given words' fasttext embeddings. Classification accuracy turns out to be at 48.5% for the word types, which far exceeds the majority baseline (9.2%). Given the classification accuracies in Table 10.4, and using these as independent probabilities, the probability of getting all four inflectional dimensions correct is $0.89 \times 0.88 \times 0.83 \times 0.47 = 0.31$. This is substantially lower than the LDA accuracy (0.485).

The LDA analysis clarifies that the FastText vectors successfully capture important aspects of the semantics and pragmatics of honorifics, tense, and speech level. In section 10.2, we therefore make use of the FastText vectors when modeling Korean lexical decision latencies.

Experiment 5

In the preceding experiment, we used FastText semantic vectors in combination with endstate learning. How is model performance with FastText embeddings when we replace endstate learning with frequency-informed learning?

Table 10.5 shows that type accuracies are much lower for models using empirical embeddings, but that token-based accuracies are very similar. The good token-based performance of models with FastText semantic vectors suggests that these models may also be a good choice for predicting lexical processing costs as experienced by native speakers. This possibility is addressed in the next section.

Table 10.5: Type and token accuracy for models using simulated semantic vectors and FastText empirical vectors, with frequency-informed learning, using the subset of 8881 verb forms for which FastText vectors are available. Token accuracies are based on the koTenTen frequencies.

	comprehension		production	
	type	token	type	token
Simulated	0.72	0.98	0.86	0.95
FastText	0.24	0.93	0.34	0.96

10.2 Predicting reaction times with the DLM

In this section, we consider whether measures derived from the DLM model that was trained with FastText vectors and Frequency-Informed Learning (cf. Table 10.5) are predictive for lexical processing as gauged with the visual lexical decision task. One question of interest to us is how these model-based predictors perform compared to measures based on frequency: (log-transformed) lexeme frequency, whole-word frequency, and ending frequency. A second question is whether how much can be accomplished with a model that is not explicitly told about inflectional features for tense or illocutionary force, as this kind of information is ideally available in words' embeddings. In what follows, we first introduce the dataset, we then discuss a generalized additive model (GAMM Wood, 2017a) using frequency counts as predictors, and finally discuss a GAMM that is informed by model-based predictors. As the computational model is trained on only a small dataset, with imperfect embeddings, we do not expect the second GAMM to provide more precise predictions than the GAMM with all standard predictors in place. Our interest, therefore, is in how much traction can be obtained with the computational model.

10.2.1 Materials

From the visual lexical decision experiment (see chapter 5) carried out with 40 native speakers of Korean and 4000 Korean verb forms, we selected the 577 forms that are included in the dataset considered in the present paper, and for which FastText embeddings are available. Response latencies were inverse transformed (RTinv=-1000/RT) in order to obtain an approximately Gaussian response variable. We included a factor smooth for the interaction of experimental trial and speaker, in order to account for the ebb and flow of attention in the course of the experiment, as well as for possible effects of learning and fatigue. Response latencies showed autocorrelation, which was brought into the model by means of an AR(1) process in the errors with proportionality parameter $\rho = 0.5$.

10.2.2 Analysis with frequency counts

Table 10.6 summarises a GAMM fitted to the response latencies, using two frequency counts (for lexeme and whole-word) and an array of additional predictors describing words' phonological and morphological properties. Of these measures, vowel harmony, stem-final segment, honorifics, past tense and SLIF are well supported.

Figure 10.4 presents the partial effects of lexeme frequency and whole-word frequency; the effect of ending frequency is not included as it is not significant for the present small dataset (but see previous chapters for evidence that for a much larger dataset with improved statistical power, ending frequency is well-supported). For lexeme frequency, low frequencies predict long RTs, high frequencies short RTs, with not much of an effect for intermediate frequencies. Whole-word frequency shows an effect only for the lowest-frequency words, which ilicited longer RTs. The AIC of this model was 1505.5.

Table 10.6: Summary of a GAMM fitted to the response latencies of 577 Korean word forms, using classical predictors. Treatment dummy coding was used, with the following reference levels(*vowel harmony*: dark (/i/ is grouped with dark set), *stem-final segment*: consonant), SLIF(Speech Level + Illocutionary Force), (AIC:1505.5)

A. parametric coefficients	Estimate	Std. Error	t-value	p-value
(Intercept)	-1.3798	0.0689	-20.0208	< 0.0001
honorifics : unmarked	-0.1075	0.0215	-4.9949	< 0.0001
past tense : marked	0.0462	0.0173	2.6666	0.0077
SLIF : for + inq	-0.0199	0.0143	-1.3955	0.1629
SLIF : hon + dec/imp	-0.1593	0.0257	-6.1977	< 0.0001
SLIF : pol + unmarked	-0.1100	0.0210	-5.2446	< 0.0001
SLIF : int + unmarked	-0.1559	0.0220	-7.0833	< 0.0001
SLIF : plain + dec	-0.1537	0.0121	-12.6554	< 0.0001
SLIF : plain + inq	-0.0540	0.0146	-3.6969	0.0002
vowel harmony : light	-0.0355	0.0061	-5.7864	< 0.0001
number of syllables	0.0018	0.0054	0.3292	0.7420
stem-final segment : vowel	0.0288	0.0064	4.5043	< 0.0001
number of exponents	-0.0268	0.0178	-1.5096	0.1312
B. smooth terms	edf	Ref.df	F-value	p-value
s(log lexeme frequency)	3.8224	3.9805	24.1586	< 0.0001
s(log whole-word frequency)	2.7244	3.1976	5.5870	0.0006
s(trial, subject)	159.7342	360.0000	7.5966	< 0.0001

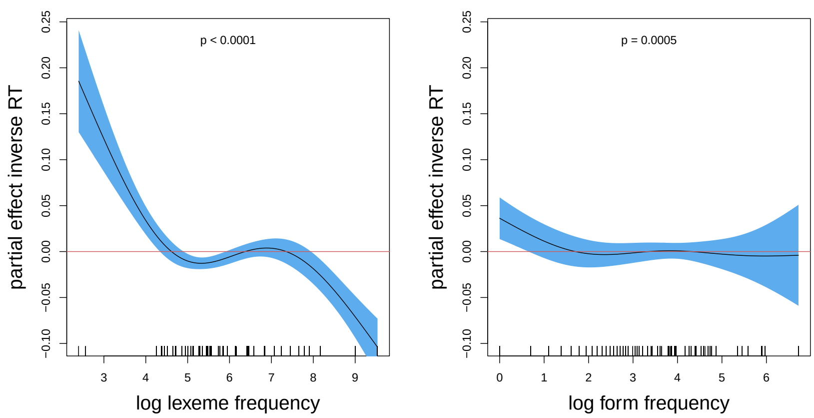


Figure 10.4: Partial effects of log-transformed lexeme frequency and log-transformed form(whole-word) frequency on inverse-transformed lexical decision latencies.

10.2.3 Analysis with model-based predictors

Two model-based measures emerged as well-supported predictors. The first measure, `Target Correlation`, quantifies how well the meaning of a word is understood by the model given the word's form. This measure is the Pearson correlation of the word's gold standard semantic vector (its FastText embedding), and the semantic vector predicted by the model. The closer to 1 the `Target Correlation` is, the more precisely the model has learned to predict a word's meaning. This measure was observed to be highly predictive of reaction times in the study of Heitmeier et al. (2023). The second measure, `Semantic Density`, evaluates how densely the semantic vectors of orthographic neighbours of the target word cluster around a word's predicted semantic vector. These model-based measures were complemented with word length (number of Hangul characters) as a control for the number of fixations. Table 10.7 and Figure 10.5 clarify that for where data are dense, a greater semantic density appears to predict shorter RTs, suggesting that greater general 'lexicality' (cf. Grainger and Jacobs, 1996) afforded faster responses. The effect of target correlation shows an overall negative trend. This main trend suggests that words can be recognized more precisely are responded to more quickly. The AIC of this model was 1810.6.

Table 10.7: Summary of a GAMM fitted to the response latencies of 577 Korean word forms, using model-based predictors, with word length in syllables as orthographic control. (AIC : 1810.6)

A. parametric coefficients	Estimate	Std. Error	t-value	p-value
Intercept	-1.7169	0.0388	-44.2859	< 0.0001
number of syllables	0.0344	0.0030	11.5306	< 0.0001
B. smooth terms	edf	Ref.df	F-value	p-value
factor smooth for trial by speaker	159.7812	360.0000	7.4184	< 0.0001
s(target correlation)	3.7528	3.9629	13.5632	< 0.0001
s(semantic density)	3.6367	3.9277	7.8441	< 0.0001

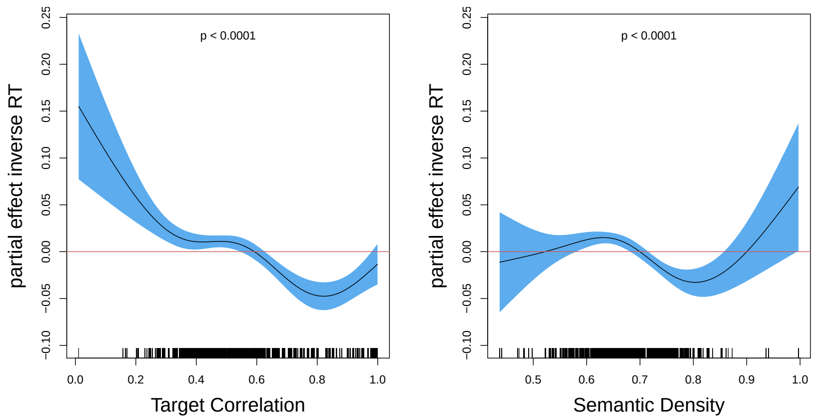


Figure 10.5: Partial effects of target correlation and semantic density on inverse-transformed lexical decision latencies.

10.2.4 The functional load of verb endings

Although the inflectional morphology of Korean verbs is agglutinative, due to the many phonological adjustments that operate on sequences of exponents, the resulting spoken forms tend to have certain patterned endings that are recurrent in the vocabulary. This suggests that endings(i.e., `Ending` as referred to in previous chapters), rather than individual exponents, are well-distinguished in semantic space.

To explore this possibility, we first defined an ending to comprise the final two bi-syllable cues. For instance, the ending for 꾸몄습니다 (k͈umyʌt˺s͈ɯmɲida) comprises the bi-syllables `ɲi-da` and `da_#`. For our dataset, there are in total 1172 unique endings. For each ending, we created a corresponding form vector $\boldsymbol{c}_{\text{ending}}$ with ones for the two bi-syllables of the ending, and zeroes for all other bi-syllables.

Next, we created a semantic vector for each of the 59 paradigm cells, using simulated vectors, to represent the inflectional meanings realised in that cell. For 꾸몄습니다, the inflectional meanings are PLAIN (HONORFICS), PAST, FORMAL, DECLARATIVE and hence the inflectional vector $\boldsymbol{s}_{\text{cell}}$ for this paradigm cell is given by

$$\boldsymbol{s}_{\text{cell}} = \overrightarrow{\text{plain}} + \overrightarrow{\text{past}} + \overrightarrow{\text{formal}} + \overrightarrow{\text{declarative}}.$$

We also calculated a predicted semantic vector $\hat{\boldsymbol{s}}_{\text{ending}}$ for each final bi-syllable by multiplying the endings' form vector with the comprehension mapping $\boldsymbol{F}$ obtained with an endstate of learning mapping:

$$\hat{\boldsymbol{s}}_{\text{ending}} = \boldsymbol{c}_{\text{ending}}\boldsymbol{F}.$$

For assessing the quality of a predicted semantic vector $\hat{\boldsymbol{s}}_{\text{ending}}$, we calculated its correlation with all the gold standard semantic vectors $\boldsymbol{s}_{\text{cell}}$, and selected as the meaning of $\boldsymbol{c}_{\text{ending}}$ the gold standard cell vector $\boldsymbol{s}_{\text{cell,max}}$ for which $r(\hat{\boldsymbol{s}}_{\text{ending}}, \boldsymbol{s}_{\text{cell}})$ was the largest. The vector $\boldsymbol{s}_{\text{cell,max}}$ encoded the targeted inflectional meaning for 54 of the 59 cells, an accuracy of 91.5%.

All incorrectly predicted cell meanings were meanings failing to realise the plain speech level while correctly realising honorifics, tense, and illocutionary force. For instance, the form vector of -셨니 (sʰyʌnɲi, as found in, e.g., 꾸미셨니, k̤umisʰyʌnɲi, HONORIFIC, PAST, PLAIN, INQUISITIVE) generated a semantic vector that was incorrectly paired with the gold standard cell meaning that realizes HONORIFIC, PAST, FORMAL, INQUISITIVE, for which the appropriate ending is -니까 (ɲik̤a)(FORMAL, INQUISITIVE, as found in words such as 꾸미셨습니까, k̤umisʰyʌt̚s̤ɯmɲik̤a). One likely reason for this mistake is that the incorrect form has two additional exponents -셨습- before the final bi-syllable -니까. Another possible reason is the use of word-final -니 in completely fusional forms such as 꾸미니 (k̤umiɲi), where it realizes NON-HONORIFIC, PRESENT, PLAIN, INQUISITIVE where it realizes non-honorific, present, plain, inquisitive renders this exponent vulnerable to inaccuracies in the mapping from form to meaning, and shifts the functional load for the inflectional semantics more to the bi-syllables that straddle the final syllable of the stem.

Figure 10.6 presents a t-SNE clustering analysis of the predicted semantic vectors of the bi-syllable endings, color-coded by the last syllables. Most of the ending vectors are very clearly distinguished, supporting the hypothesis that they provide robust (albeit not perfect) cues to the inflectional meanings associated with words' forms. In other words, words' endings provide robust information about inflectional semantics, which makes it possible for our model to pick up the isomorphies between these mean-

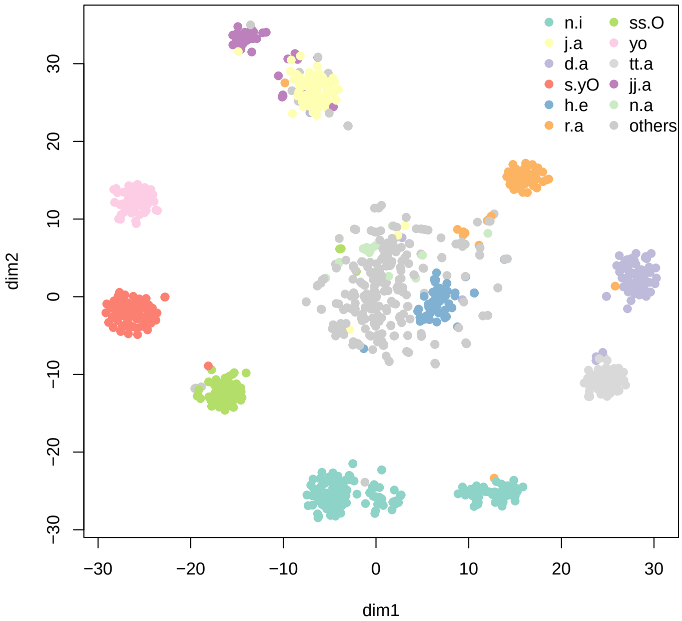


Figure 10.6: Projection of the predicted semantic vectors of the final bi-syllable endings onto a 2D plane by using tSNE, color-coded by the last syllables.

ings and the corresponding forms — without ever having to identify individual exponents and without having to untangle the effects of assimilation and stem allomorphy that are ubiquitous Korean morphology, and that constitute substantial computational challenges for strictly decompositional theories of morphological processing.

10.3 Discussion

In this chapter, we addressed the question of whether the complex morphological systems of the Korean verbs can be approximated in an insightful way using a computational model of the mental lexicon, the discriminative lexicon model (DLM) that has been applied successfully to a range of other languages (Chuang and Baayen, 2021; Chuang et al., 2022; Heitmeier et al., 2025). This study restricted the dataset to the Korean verbal morphology of the sentence-final verb forms, excluding those forms that express mood. We put together a dataset with some 27 thousand forms of 462 general Korean verbs.

The first set of computational modeling experiments illustrates the various decisions about how to represent form and meaning, and how to set up adequate mappings between form and meaning, that have to be addressed. Among three types of representations(bi-Hangul, bi-syllables and triphones), bi-syllables are especially useful(the accuracies of 99.2% in comprehension and 97.8% in production) for capturing the effects of liaison and other between-syllable phonological adjustments in Korean.

Another expeiment examined to what extent a model with bi-syllable form vectors and simulated semantic vectors is productive. In general, our experiments indicate that the DLM performs with high accuracy on data it has been trained on(above 98%), and performs with reasonable accuracy on held-out test data(in the range of 70%–92%).

We also explored the extent to which Korean verb inflectional forms can be accounted for within a distributional hypothesis. The concepts such as speech level and subject honorifics, frequently marked in Korean verb endings, pose limitations for this approach. This is because these features are not determined by the inherent meanings of individual words in a sentence, but rather by the social relationship between language users. As a result, they do not align well with the basic assumption of distributional semantics that words with similar usage contexts appear in similar semantic space. Nevertheless, the morphological and semantic distribution of Korean verb endings as visualized using t-SNE reveals some intriguing patterns that allow for limited yet meaningful interpretation. In particular, the distribution of speech levels is noteworthy: the plain speech style, which is widely used across both written and spoken registers, displays a broad dispersion, while other speech levels tend to cluster together. In addition, if the distribution is distinctive, it may also be possible to infer a specific form (e.g., -(으)십시오) as seen in some distinctive clusters.

We also investigated whether predictors derived from DLM model, using empirical embeddings and frequency-informed learning, are predictive for lexical decision latencies of native speakers of Korean, using the data collected in the experiment reported in chapter 5. Due to data sparsity, modeling was possible for only a subset of 577 verb forms. Two model-based measures, `Target Correlation` and `Semantic Density` emerged as well-supported predictors. As for `Target Correlation`, reaction times were shorter the more precisely the model predicts a word's meaning from its form. The second measure, `Semantic Density` indicated that greater semantic density based on orthographic neighbours afforded shorter reaction times, suggesting that more evidence for lexicality speeded correct lexical decision making, in line with the multiple read-out hypothesis of Grainger and Jacobs (1996).

We have compared two models, one with a wide range of classical predictors, and one with predictors based on the DLM. The model with the classical predictors provides a more precise fit to the data (AIC 1506) compared to the model with DLM-based predictors (AIC 1811). Given that the DLM model was based on less than 9000 forms from 462 verbs, it is notable that the DLM-based GAMM performs as well as it does. Korean speakers have much more comprehensive and variegated experiences with their language. It remains noteworthy that measures based on a model that sidesteps identifying stems and individual exponents, and that is given semantic representations that, unfortunately, are unlikely to be optimal, are fairly predictive for lexical processing in Korean. This result supports the possibility that theories positing obligatory decomposition into stems and exponents may be unnecessarily complex (see Heitmeier et al., 2025, for detailed discussion).

Finally, to further examine the semantic vectors of verb endings, this study also examined the extent to which the final two syllables of a verb ending are semantically related to the overall semantic vector of the full ending. The results showed that, for 54 out of the 59 selected endings (91.5%), the meaning vector of the final two syllables closely matched that of the full ending. This finding aligns with results from previous lexical decision experiments involving both native and non-native speakers (see chapters 5 and 7), where the SLIF (Speech Level and Illocutionary Force) factor ranked as the third most influential variable for native speakers, after the variable Ending and stem frequency, based on AIC values. In non-native groups except L1 Chinese group, SLIF also ranked higher than other ending components, such as honorifics and past tense. That is, since SLIF is located at the final part of the verb ending and is at least partially reflected in the last two syllables, this result suggests that the final segment of the ending may carry substantial weight in the visual lexical processing of Korean verbs.

The DLM offers an alternative perspective on the processing of complex verbs in Korean. This theoretical framework avoids working with discrete sublexical units such as morphemes or endings, and thus avoids all the hotly-debated questions about what the morphological units in Korean might be (see, e.g., Choi, 1961; Martin, 1960; Lee, 1989; Yang, 1994; Song, 2005). DLM model, which utilizes bi-syllables, is designed to accommodate the widespread effects of phonological adjustments. Rather than attempting to reconstruct 'underlying' discrete exponents and stems, it takes 'surface forms' as given and analyzes the discriminative contributions of individual bi-syllables to the meanings of words. This approach offers a straightforward and effective framework for lexical processing. It is simpler compared to theories that assume mandatory decompositional processing for complex words, which would involve addressing the extensive phonological adjustments during comprehension and determining how to reorganize concatenated exponents into pronounceable words during production.

Several questions remain open. One is whether DLM model scales to comprehensive Korean datasets, including mood markers, clausal connectives, and nominal or gerundive forms. It is also unclear whether text-based embeddings adequately capture the semantic and pragmatic nuances of mood, honorifics, and speech levels. Additionally, the model does not address how L2 learners acquire Korean morphology, which may involve more explicit, cognitively driven processes.

To conclude, while Korean verb morphology is highly productive, native speakers can interpret and produce novel forms, and the model shows similar capabilities within the set of 59 sentence-final endings, with higher accuracy in comprehension than production. A modeling experiment using simulated semantic vectors showed that the meanings predicted from the final two syllables of verb endings were highly correlated with the full semantic vectors of the endings, accounting for 54 out of 59 endings. This result aligns with the findings from the earlier visual lexical decision task with native speakers, where the AIC increase of SLIF, which occupies the final part of the sentence-final ending, was found to be notably high. The experiments with frequency-informed learning clarify that the low-frequency forms are harder to acquire, aligning with longer lexical decision times. Model accuracy for frequency-informed learning, evaluated over tokens, remains high (above 98%), indicating that the model effectively acquires frequently used lexical items.

Chapter 11

Conclusion

This study originated from two simple questions. The most fundamental and simple question is how humans process morphologically and semantically complex language forms, and furthermore, how this is handled by non-native speakers who do not speak the language as their first language. The other is, then, along with the above question, how the complex mixture of various elements (i.e., form, meaning, phonology) is processed in the complex inflected Korean verb forms, which are clearly distinct from other types of word forms in Korean.

The recognition that Korean verbs represent a highly intricate combination of morphological, semantic, and phonological complexity, distinguishing them clearly from other word types, serves as an important starting point for this study. Furthermore, this study also recognises the diversity of such complex forms that share a single verb stem and the vast number of forms that can be generated from them, examining how both native and non-native speakers process these various forms. In addition, this study poses the question of how non-native speakers of Korean are influenced by their first language when processing Korean verbs.

In the results of the visual lexical decision tasks, the groups of native and non-native Korean speakers showed several distinct differences. For native speakers, the

Ending (i.e., the combination of the exponents) was the biggest impactful variable when processing Korean verbs, with related variables such as ending frequency, SLIF, and the lexeme frequency also playing significant roles. In contrast, for most non-native speakers with advanced proficiency in Korean, the primary influencing factor was word length based on the number of syllables, followed by Ending (only the Chinese group, however, the order was reversed—Ending followed by the number of syllables.) Following these, ending frequency and whole-word frequency were similarly influential variables. Notably, the Chinese group showed a distinctive pattern, with ending frequency resulting in an AIC increase approximately three times greater than that of whole-word frequency, setting them apart from the other groups.

Taken together, these findings suggest that while both native and non-native speakers are most strongly influenced by Ending and ending frequency, a key difference lies in the role of stem and whole-word frequency: native speakers are more influenced by stem frequency, whereas non-native speakers rely more heavily on whole-word frequency.

In conclusion, it appears that Korean verb processing is likely to involve a separation between the morphological units of the stem and the ending part(i.e., Ending). That is, the fact that Ending emerges as a major unit suggests that processing may not occur at the level of individual exponents but rather at the level of larger, frequently encountered ending patterns. Additionally, the influence of sentence-final endings (SLIF) indicates that the perception of final syllables may also play an important role in processing.

Among the non-native groups, the Chinese group showed a distinct pattern: they were more influenced by ending frequency than by whole-word frequency, and the weight of sentence-final endings (SLIF) was notably lower compared to other groups, while the influence of honorific forms was significantly higher. The distinct characteristics of Chinese—its use of ideogram script and, like Vietnamese, its status as an analytic language without inflectional morphology—are speculated to be possible sources of these differences, though this remains uncertain.

Regarding L1 effects, typologically similar languages such as Japanese and Mongolian showed roughly comparable weights in AIC increases of the predictors, whereas Chinese and Vietnamese exhibited notable differences across several aspects. Observations based on non-word results suggest some effects attributable to L1. Specifically, the Korean, Japanese, and Mongolian groups—languages categorised as agglutinative—showed longer reaction times for manipulations of verb endings than for stems, indicating a greater sensitivity to endings. In contrast, the Chinese and Vietnamese groups responded more sensitively(slower RTs) to syllable manipulations, suggesting a syllable-based processing influence stemming from their L1 phonologies. These findings imply that even among typologically similar languages, lexical processing is shaped by a complex interplay of language-specific features and multiple interacting variables.

We also examined how well Korean verb processing aligns with the newly proposed computational model, DLM. The 59 Korean verb endings (including homophones) used in the model showed a good level of feasibility in both the training and held-out datasets. The model does not present in detail how it processes such complex interactions with discrete sublexical units, such as morphemes or endings between form and meaning. Even if a word may be morphologically and semantically complex, it cannot be assumed that, in actual lexical processing, especially for native speakers, simple words are always processed simply and complex words through more complex mechanisms. In this regard, although the model was applied using a limited set of Korean verb endings, the findings are meaningful in various respects. Nevertheless, since the form vectors in the model rely primarily on words' physical cues, such as orthographic or phonological features in word processing, it remains necessary to further test how this operates in form-meaning mapping using a wider variety of Korean verb endings. Given that approximately 5,700 types of Korean verb endings appear in the corpus (Choi, 2012), the use of 59 endings may reflect a very limited scope. Many of the numerous verb endings found in the corpus are derived from mood forms that express a wide range of complex and subtle nuances. One of the key reasons for the large number of verb endings is that many mood forms, which individually have low frequency, often appear as standalone endings together with the verb stem, functioning almost like independent lexical items. Given this, it raises the question of how the simulated semantic vector, formed by the combination of the verb stem and inflectional ending, can capture the complex and varied meanings associated with mood. This highlights the need for further research on how to incorporate such cases into computational models.

Alongside all these findings, this study also has several limitations. Given that human processing of complex words involves a range of interacting factors, there are important variables that were not addressed in sufficient depth. In particular, considering that Korean verbs convey subtle shades of meaning through a wide array of mood forms within their complex form-meaning structures, the absence of observations and discussion related to semantic transparency is a notable limitation. Although mood forms were treated in this study as a broad category, primarily based on features of connective endings or sentence-final connective endings, there is a need for more refined categorisation and analysis from the perspective of semantic transparency. It is anticipated that semantic transparency may pose greater challenges for non-native speakers than for native speakers in lexical processing. Furthermore, the exponential productivity of Korean verb forms is closely tied to the wide variety of mood forms, and this suggests that future research on Korean verb processing should place greater emphasis on diverse morphological and semantic approaches to these structures.

In this study, a total of 92 types of verb endings—including both real words and non-words—were combined with 100 verb stems and examined through experimental observation. It is hoped that future research on Korean verb processing, designed with diverse morphological and semantic approaches, will incorporate a wider range of verb endings and yield more extensive findings.

Bibliography

Ahn, H. (2015). *Second language acquisition of Korean case by learners with different first languages*. PhD thesis, University of Washington.

Ahn, H., An, D., Choi, J., Hwang, J.-B., Kim, J.-H., and Jeon, M. (2011a). 한국어 어미와 선어말어미의 형태론적 처리: 점화실험을 중심으로 (Morphological processing of final endings and pre-final endings in Korean verbs: Evidence from priming experiments). 언어과학연구 *(The Journal of Linguistic Science)*, 59:181–202.

Ahn, H.-D., An, D.-H., Choi, J.-Y., Jeon, M., Kim, J.-H., and Hwang, J.-B. (2011b). Chinese learners'processing of nominal suffixes in Korean. 한국외국어교육학회 *(Foreign Languages Education)*, 18:57–71.

Ahn, H.-D., Hwang, J.-B., Jeon, M., Cai, C., and Kim, J.-H. (2013). Chinese learners' L2 morphological processing of final and prefinal endings in Korean verbs. *Foreign Languages Education*, 20(3):1–16.

Amenta, S. and Crepaldi, D. (2012). Morphological processing as we know it: An analytical review of morphological effects in visual word identification. *Frontiers in psychology*, 3:232.

Baayen, R. and Linke, M. (2020). An introduction to the generalized additive model. In Gries, S. and Paquot, M., editors, *Practical Handbook of Corpus Linguistics*, page in press. Springer.

Baayen, R. H., Chuang, Y.-Y., Shafaei-Bajestan, E., and Blevins, J. (2019). The discriminative lexicon: A unified computational model for the lexicon and lexical processing in comprehension and production grounded not in (de)composition but in linear discriminative learning. *Complexity*.

Baayen, R. H., Dijkstra, T., and Schreuder, R. (1997). Singulars and plurals in Dutch: Evidence for a parallel dual route model. *Journal of Memory and Language*, 36:94–117.

Baayen, R. H., Fasiolo, M., Wood, S., and Chuang, Y.-Y. (2022). A note on the modeling of the effects of experimental time in psycholinguistic experiments. *The Mental Lexicon*, 17(2):178–212.

Baayen, R. H., Wurm, L. H., and Aycock, J. (2007). Lexical dynamics for low-frequency complex words. a regression study across tasks and modalities. *The Mental Lexicon*, 2:419–463.

Bae, S. (2018). Effects of nonword type and proportion on lexical decision in hangul word recognition. *한국심리학회지: 인지 및 생물 (The Korean Journal of Cognitive and Biological Psychology)*, 30(3):301–307.

Bae, S. and Yi, K. (2010). Processing of orthography and phonology in korean word recognition. *The Korean Journal of Cognitive and Biological Psychology*, 22(3):369–385.

Bedny, M. and Thompson-Schill, S. L. (2006). Neuroanatomically separable effects of imageability and grammatical class during single-word comprehension. *Brain and language*, 98(2):127–139.

Blevins, J. P. (2016). *Word and paradigm morphology*. Oxford University Press.

Bojanowski, P., Grave, E., Joulin, A., and Mikolov, T. (2017). Enriching word vectors with subword information. *Transactions of the Association for Computational Linguistics*, 5:135–146.

Bozic, M., Szlachta, Z., and Marslen-Wilson, W. D. (2013). Cross-linguistic parallels in processing derivational morphology: Evidence from polish. *Brain and language*, 127(3):533–538.

Butterworth, B., editor (1983). *Language Production (Vol. II): Development, Writing and Other Language Processes*. Academic Press, London.

Bybee, J. L. (1985). *Morphology: A Study of the Relation between Meaning and Form*. Philadelphia: J. Benjamin.

Bybee, J. L. and Hopper, P. J. (2001). *Frequency and the Emergence of Linguistic Structure*, volume 45. John Benjamins Publishing.

Caramazza, A., Laudanna, A., and Romani, C. (1988). Lexical access and inflectional morphology. *Cognition*, 28(3):297–332.

Carreiras, M., Álvarez, C. J., and Devega, M. (1993). Syllable frequency and visual word recognition in spanish. *Journal of memory and language*, 32(6):766–780.

Cha, H. S. (1990). 반말체의 구성과 반말체 어미의 문법적 기능에 대하여 (About the composition of informal speech and the grammatical function of informal speech ending). *Journal of Ehwa Korean Language and Literature*, 11:5–26.

Chae, S. (2012). 고급학습자의 일상대화에서의 종결어미 사용 양상 연구-중국어권 학습자의 경우 (An analysis of final-ending usage patterns in daily conversation produced by advanced learners: The case of Chinese learners of Korean). 한국어 교육 *(Journal of Korean Language Education)*, 23(4):371–389.

Choi, E. J. L., Lee, S. Y., Yoo, k., and Lee, S. (2020). 초급 한국어 학습자의 종결어미 사용 양상에 대한 종적 연구 (A longitudinal study of sentence endings of beginner level Korean language learners). 한국어 교육 *(Journal of Korean Language Education)*, 31(3):309–338.

Choi, H.-B. (1937, reprinted in 1961). 우리 말본 *(Korean Grammar), 3rd edition.* 정음문화사, Chong-Um-Munwha-Sa, Seoul.

Choi, j. (2013). 한국어 고급 학습자의 상대 높임의 등급 전환 양상 연구 (Research on the aspects of the advanced Korean language learners' code-switching in the relative honorification). 한국언어문화학 *(Korean Language Literature)*, 10(1):229–251.

Choi, K. S. (2015). 융합 형식을 포함한 용언 종결형의 분석 (2)(An Analysis of Verbal Forms with Fusion Forms in Korean(2)). 우리말연구 *Urimal(Our language)*, 42:35–60.

Choi, S. J. (2012). 결합형 어미의 양상 연구 (The Study of Combined Endings' Condition). *Journal of Korean Linguistics*, 63:275–311.

Choo, Y.-J. (2003). *A synchronic and diachronic study of Korean modal suffixes.* University of London, School of Oriental and African Studies (United Kingdom).

Chuang, Y., Brown, D., Evans, R., and Baayen, R. H. (2023). Paradigm gaps are associated with weird "distributional semantics" properties: Russian defective nouns and their case and number paradigms. *The Mental Lexicon*.

Chuang, Y. Y. and Baayen, R. H. (2020). Discriminative learning and the lexicon: Ndl and ldl. In *Oxford Research Encyclopedia of Linguistics*, page under review. Oxford University Press.

Chuang, Y. Y. and Baayen, R. H. (2021). Discriminative learning and the lexicon: NDL and LDL. In *Oxford Research Encyclopedia of Linguistics*. Oxford University Press.

Chuang, Y.-Y., Bell, M. J., Banke, I., and Baayen, R. H. (2020a). Bilingual and multilingual mental lexicon: a modeling study with linear discriminative learning. *Language Learning*, pages 1–73.

Chuang, Y. Y., Fon, J., Papakyritsis, I., and Baayen, R. H. (2021). Analyzing phonetic data with generalized additive mixed models. In Ball, M. J., editor, *Handbook of Clinical Phonetics*, page 108. Routledge.

Chuang, Y. Y., Kang, M., Luo, X. F., and Baayen, R. H. (2022). Vector space morphology with linear discriminative learning. In Crepaldi, D., editor, *Linguistic morphology in the mind and brain*. Routledge.

Chuang, Y.-Y., Loo, K., Blevins, J. P., and Baayen, R. H. (2020b). Estonian case inflection made simple. A case study in Word and Paradigm morphology with Linear Discriminative Learning. In Körtvélyessy, L. and Štekauer, P., editors, *Advances in Morphology*, pages 119–141. Cambridge University Press. Preprint available at https://psyarxiv.com/hdftz/.

Chung, K. (2015). 한국어 용언 활용 체계의 통시적 변화 *(Historical change of the Korean conjugation system)*. PhD thesis, Korea University.

Clahsen, H., Balkhair, L., Schutter, J.-S., and Cunnings, I. (2013). The time course of morphological processing in a second language. *Second Language Research*, 29(1):7–31.

Clahsen, H. and Felser, C. (2006a). Continuity and shallow structures in language processing. *Applied Psycholinguistics*, 27(1):107–126.

Clahsen, H. and Felser, C. (2006b). Grammatical processing in language learners. *Applied psycholinguistics*, 27(1):3–42.

Clahsen, H. and Felser, C. (2006c). How native-like is non-native language processing? *Trends in cognitive sciences*, 10(12):564–570.

Clahsen, H. and Neubauer, K. (2010). Morphology, frequency, and the processing of derived words in native and non-native speakers. *Lingua*, 120(11):2627–2637.

Conrad, M., Carreiras, M., and Jacobs, A. M. (2008). Contrasting effects of token and type syllable frequency in lexical decision. *Language and Cognitive Processes*, 23(2):296–326.

Csaba, P. and Levente, J. (1995). Processing of multimorphemic words in hungarian. *Acta Linguistica Hungarica*, 43:1–2.

Denistia, K. and Baayen, R. H. (2022). The morphology of indonesian: Data and quantitative modeling. In Shei, C., editor, *The Routledge Handbook of Asian Linguistics*. Routledge.

Diependaele, K., Duñabeitia, J. A., Morris, J., and Keuleers, E. (2011). Fast morphological effects in first and second language word recognition. *Journal of Memory and Language*, 64(4):344–358.

Dohmes, P., Zwitserlood, P., and Bölte, J. (2004). The impact of semantic transparency of morphologically complex words on picture naming. *Brain and language*, 90(1-3):203–212.

Ellis, R. (2006). *Understanding second language acquisition*. Oxford university press Oxford.

Eom, J. S. (2011). 한국어 학습자의 해라체 평서문 습득 양상 (A study on the acquisition of a declarative sentence of Hera style by learners of Korean). 한국어 교육 *(Korean Language Education)*, 22(4):193–214.

Feldman, L. B., Kostić, A., Basnight-Brown, D. M., Đurđević, D. F., and Pastizzo, M. J. (2010). Morphological facilitation for regular and irregular verb formations in native and non-native speakers: Little evidence for two distinct mechanisms. *Bilingualism: Language and cognition*, 13(2):119–135.

Firth, J. R. (1968). *Selected papers of J R Firth, 1952-59.* Indiana University Press.

Frajzyngier, Z. (2008). Grammaticalization, typology and semantics: Expanding the agenda. In *Rethinking grammaticalization: New perspectives*, pages 61–102. John Benjamins Publishing Company.

Friedman, L. and Wall, M. (2005). Graphical views of suppression and multicollinearity in multiple regression. *The American Statistician*, 59:127–136.

Frisch, S. A., Large, N. R., and Pisoni, D. B. (2000). Perception of wordlikeness: Effects of segment probability and length on the processing of nonwords. *Journal of memory and language*, 42(4):481–496.

Gershkoff-Stowe, L. and Hahn, E. R. (2013). Word comprehension and production asymmetries in children and adults. *Journal of experimental child psychology*, 114(4):489–509.

Gor, K., Cook, S., Bordag, D., Chrabaszcz, A., and Opitz, A. (2021). Fuzzy lexical representations in adult second language speakers. *Frontiers in Psychology*, 12:732030.

Grainger, J. and Jacobs, A. M. (1996). Orthographic processing in visual word recognition: A multiple read-out model. *Psychological Review*, 103:518–565.

Günther, F. and Marelli, M. (2019). Enter sandman: Compound processing and semantic transparency in a compositional perspective. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 45(10):1872.

Han, S. (2018). 한국어 학습자의 종결어미 사용 양상과 오류 연구 (The study of the usage and errors of Korean final endings). 문법교육 *(Grammar Education)*, 33:165–210.

Han, S. (2020). 한국어 구어 대화의 종결어미 사용 양상 연구 (A Study on the Use of Final Endings in Korean Language Conversation). *The Journal of Humanities and Social Sciences*, 11:2315–2328.

Harm, M. W. and Seidenberg, M. S. (2004). Computing the meanings of words in reading: cooperative division of labor between visual and phonological processes. *Psychological review*, 111(3):662.

Hauptman, M., Blanco-Elorrieta, E., and Pylkkänen, L. (2022). Inflection across categories: Tracking abstract morphological processing in language production with meg. *Cerebral Cortex*, 32(8):1721–1736.

Hedden, T., Ketay, S., Aron, A., Markus, H. R., and Gabrieli, J. D. (2008). Cultural influences on neural substrates of attentional control. *Psychological science*, 19(1):12–17.

Heitmeier, M. and Baayen, R. H. (2021). Simulating phonological and semantic impairment of English tense inflection with Linear Discriminative Learning. *The Mental Lexicon*, 15:385–421.

Heitmeier, M., Chuang, Y., Axen, S., and Baayen, R. H. (2023). Frequency-informed linear discriminative learning. *Front. Hum. Neurosci., Sec. Speech and Language*, 17.

Heitmeier, M., Chuang, Y.-Y., and Baayen, R. H. (2021). Modeling morphology with linear discriminative learning: considerations and design choices. *Frontiers in Psychology: Language Sciences*.

Heitmeier, M., Chuang, Y.-Y., and Baayen, R. H. (2025). *The Discriminative Lexicon: Theory and implementation in the Jul ia package JudiLing*. Cambridge University Press, Cambridge. in press.

Hendrix, P. and Sun, C. C. (2021). A word or two about nonwords: Frequency, semantic neighborhood density, and orthography-to-semantics consistency effects for nonwords in the lexical decision task. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 47(1):157.

Hino, Y. and Lupker, S. J. (1996). Effects of polysemy in lexical decision and naming: An alternative to lexical access accounts. *Journal of Experimental Psychology: Human Perception and Performance*, 22(6):1331.

Hutzler, F., Bergmann, J., Conrad, M., Kronbichler, M., Stenneken, P., and Jacobs, A. M. (2004). Inhibitory effects of first syllable-frequency in lexical decision: an event-related potential study. *Neuroscience Letters*, 372(3):179–184.

Hwang, J.-B., Ahn, H.-D., Han, J.-I., Lim, I., Hong, J., and Shim, C. (2008). 외국인 학습자의 한국어 선어말어미-시-,-았/었-,-겠-의 중첩 습득 연구 (A study of foreign learners' acquisition of some combinations of Korean prefinal endings). 언어과학 *(Journal of Language Sciences)*, 15(2):183–199.

Hwang, Y. (2007). The representation of Korean verb and adjective Eojeols in the mental lexicon (doctor dissertation). *Korea University, Seoul, South Korea.*

Hwang, Y. (2008). 한국어 용언의 심성어휘 표상연구 *(The representation of Korean verb and adjective Eojeols in the mental lexicon)*. PhD thesis, Korea University.

Hwang, Y., Kichun, N., and Myung-Yoon, K. (2003). Cerebral activation in production of Korean inflectional and derivational affixes. In *Proceedings of the KSPS conference*, pages 97–100. The Korean Society Of Phonetic Sciences And Speech Technology.

Hwang, Y., Kwon, Y., and Lim, H. (2002). 한국어 용언의 형태소 정보처리 특성 (A model of Korean Verb Processing). 대한음성학회 *(The Korean Society Of Phonetic Sciences And Speech Technology)*, pages 101–104.

Hwang, Y.-M., Shin, J.-M., Lim, S.-M., Ryu, K.-T., Khang, H.-S., Yi, K.-O., and Nam, K.-C. (2005). Cerebral activation related with morphological priming effect in production of Korean endings. In *Proceedings of the Korean Society for Cognitive Science Conference*, pages 273–277. The Korean Society for Cognitive Science.

Jeong, W. H. (2018). 한국어 양태의 표현수단 체계의 대한 고찰 (A review of the expression system of the Korean modality). *Gwanak study of Korean*, 43:361–381.

Jespersen, O. (1924). *The philosophy of grammar.* GEORGE ALLEN UNWIN LTD.

Jo, S. (2021). 규칙성에 따른 한국어 동사 활용형의 심성어휘집 표상 방식: 어절판단과제를 중심으로 (Representation of Korean verbal inflections in mental lexicon: evidence from verbal regularity, Eojeol frequency and stem frequency). Master's thesis, Korea University.

Kang, B. M. and Kim, H. K. (2009). 한국어 사용빈도 *(The frequency of using Korean)*. Hankook Munhwasa.

Karlsson, F. (1986). Frequency considerations in morphology. *STUF-Language Typology and Universals*, 39(1-4):19–28.

Kim, J., Kang, J., Kim, J., and Nam, K. (2022). Temporal dynamics of form and meaning in morphologically complex word processing: An erp study on korean inflected verbs. *Journal of Neurolinguistics*, 64:101098.

Kim, J. and Nam, K. (2018). Lexical factors that influence the Korean Eojeol recognition. *The Korean Journal of Cognitive and Biological Psychology*, 30(4):373–390.

Kim, S. (2013). 한국어 대화문 종결 형태의 실현 양상 연구 - 한국어 교육 교재 대화문의 실제성 확보를 위하여 (A Study on Aspect of Sentence Ending Realization in Korean Dialogue- For Securing Authenticity of Dialogue in Korean Language Materials). 문법교육 *(Grammar Education)*, 19:57–79.

Kim, S. Y., Yap, M. J., and Goh, W. D. (2019). The role of semantic transparency in visual word recognition of compound words: A megastudy approach. *Behavior research methods*, 51:2722–2732.

Kim, Y., Jung, J., and Nam, K. (2000). 한국어 불규칙 용언의 형태 정 (The Processing of Irregular Verbals in Korean Shown in Ahasics, Communication Sciences and Disorders). 한국언어청각임상학회 *(The Korean Academy Of Speech-Language Pathology And Audiology)*, 5(2):1–16.

Kim, Y., Lee, S., Lee, E., Nam, K., Jun, H., and Lee, S. (2018a). 한국어 용언 어절 재인에 미치는 어휘 변인의 영향 -모어 화자와 고급 학습자의 예 (The influence of lexical factors on verbal Eojeol recognition: Evidence from L1 Korean speakers and L2 Korean learners). 한국어교육 *(The Journal of Korean Language Education)*, 29-3:25–53.

Kim, Y., Lee, S., Lee, E.-H., Nam, K., Jun, H., and Lee, S.-Y. (2018b). The influence of lexical factors on verbal eojeol recognition: Evidence from l1 korean speakers and l2 korean learners. *Journal of Korean language education*, 29(3):25–53.

Kim H., H. (2006). 활용의 형태음운론적 변화 *(Morphophonological changes in verbal conjugation)*. 국어학회 (The Society of Korean Linguistics).

Kim T., T. (2015). 활용 어간과 어미의 어휘 표상 (The mental representation of inflected form). 형태론 *(Morphology)*, 17-2:233–263.

Kirkici, B. and Clahsen, H. (2013). Inflection and derivation in native and non-native language processing: Masked priming experiments on turkish. *Bilingualism: Language and Cognition*, 16(4):776–791.

Ko, K. M. (2001). 반말체의 등급과 반말체 어미의 발달에 대하여 (The Level of the hae Style and the Development of the hae Style Endings in Korean). *The Linguistic Society of Korea*, 30:3–27.

Kroll, J. F. and De Groot, A. M. (2014). Lexical and conceptual memory in the bilingual: Mapping form to meaning in two languages. In *Tutorials in bilingualism*, pages 169–199. Psychology Press.

Kuperman, V., Schreuder, R., Bertram, R., and Baayen, R. H. (2009). Reading of multimorphemic Dutch compounds: Towards a multiple route model of lexical processing. *Journal of Experimental Psychology: HPP*, 35:876–895.

Kwon, S., Kim, J., Lee, S., and Nam, K. (2023). The facilitative effect of first syllable frequency during visual recognition of Korean noun Eojeols. *The Korean Journal of Cognitive and Biological Psychology*, 35(2):93–106.

Kwon, Y. (2012). The dissociation of syllabic token and type frequency effect in lexical decision task. *The Korean Journal of Cognitive and Biological Psychology*, 24(4):315–333.

Kwon, Y. and Lee, Y. (2014). Time course of word frequency and word length effect in visual word recognition: Evidence from event-related brain potential study. *The Journal of Linguistic Science*, 69:43–62.

Landauer, T. and Dumais, S. (1997). A solution to Plato's problem: The latent semantic analysis theory of acquisition, induction and representation of knowledge. *Psychological Review*, 104(2):211–240.

Lee, E.-h., Lee, S.-j., and Kim, Y.-j. (2018a). 한국어 규칙 용언과 불규칙 용언의 처리기제: 점화실험 연구: 점화실험 연구 (The mechanism for processing regular and irregular inflections of Korean verbs — A priming study —). 국어국문학, (183):67–124.

Lee, E.-H., Lee, S.-j., Lim, W.-y., and Kim, Y.-j. (2019a). 의미적 맥락과 단어빈도가 한국어 시각단어 재인에 미치는 영향: 점화연구: 점화연구 (The effects of semantic context and word frequency on visual word recognition in Korean). 국어국문학 *(The Society of Korean Language and Literature)*, (186):65–109.

Lee, H. B. (1989). *Korean grammar*. Oxford University Press, New York.

Lee, H. S. (1991). *Tense, aspect, and modality: A discourse-pragmatic analysis of verbal affixes in Korean from a typological perspective*. University of California, Los Angeles.

Lee, S., Kim, S., Kim, J., Kwon, S., Lee, E.-H., and Nam, K. (2023). The facilitative effect of the first syllable token frequency in visual recognition of Korean predicate eojeols. *The Korean Journal of Cognitive and Biological Psychology*, 35(4):337–345.

Lee, S., Lee, E.-h., Lim, W., and Kim, Y. (2018b). 과거시제선어말어미 결합 용언의 처리 기제: 점화실험 연구: 점화실험 연구 (The processing mechanism of past tense pre-final endings -A priming study —). 국어국문학, (185):105–142.

Lee, S., Lim, W., and Kim, Y. (2020). 한국어 고급 학습자의 과거시제 선어말어미 결합 용언 처리에 대한 점화 실험 연구 (A Priming Study for Korean Advanced Learners' Processing of Verbal Eojeols Combined with Past-Tense Pre-Final Endings). 새국어교육 *(The Academy for Korean Language Education)*, 122:335–378.

Lee, S., Nam, K., Lee, S., Jeon, H., and Kim, Y. (2019b). 한국어 용언 어절 처리에 미치는 어휘 변인의 영향: 사건관련전위 연구 (The influence of lexical factors on verbal Eojeol recognition: An erp study). *The Journal of Linguistic Science*, 91:289–314.

Levelt, W., Roelofs, A., and Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences*, 22:1–38.

Li, M., Jiang, N., and Gor, K. (2017). L1 and l2 processing of compound words: Evidence from masked priming experiments in english. *Bilingualism: Language and Cognition*, 20(2):384–402.

Libben, G., Gibson, M., Yoon, Y. B., and Sandra, D. (2003). Compound fracture: The role of semantic transparency and morphological headedness. *Brain and language*, 84(1):50–64.

Lim, C., Baek, H., Kim, T. H., and Choi, W. (2022). Activation of phonological and orthographic information during Korean visual word recognition: Evidence from a meta-analysis and a priming study. *The Korean Journal of Cognitive and Biological Psychology*, 34(4):221–236.

Lim, W., Lee, S., Lee, S., Jeon, H., Nam, K., and Kim, Y. (2020). 시각적 단어 재인에 미치는 단어 빈도와 의미적 연관성의 영향: 한국어 고급 학습자의 예 (The Effects of Word Frequency and Semantic Context on Visual Word Recognition: Evidence from Highly Proficient Korean L2 Learners). 국어국문학 *(The society of Korean language and literature)*, 192:459–494.

Longtin, C.-M., Segui, J., and Hallé, P. A. (2003). Morphological priming without morphological relationship. *Language and cognitive processes*, 18(3):313–334.

Loo, K., Jaervikivi, J., and Baayen, R. (2018a). Whole-word frequency and inflectional paradigm size facilitate Estonian case-inflected noun processing. *Cognition*, 175:20–25.

Loo, K., Jaervikivi, J., Tomaschek, F., Tucker, B., and Baayen, R. (2018b). Production of Estonian case-inflected nouns shows whole-word frequency and paradigmatic effects. *Morphology*, 1(28):71–97.

Luo, X., Chuang, Y.-Y., and Baayen, R. H. (2021). Judiling: an implementation in Julia of Linear Discriminative Learning algorithms for language modeling. Available at `https://megamindhenry.github.io/JudiLing.jl/stable/`.

Maaten, L. v. d. and Hinton, G. (2008). Visualizing data using t-sne. *Journal of machine learning research*, 9(Nov):2579–2605.

MacWhinney, B., Kroll, J. F., and De Groot, A. (2005). A unified model of language acquisition. *Handbook of bilingualism: Psycholinguistic approaches*, 4967:50–70.

MacWhinney, B. and Leinbach, J. (1991). Implementations are not conceptualizations: Revising the verb learning model. *Cognition*, 40(1-2):121–157.

Marslen-Wilson, W., Tyler, L. K., Waksler, R., and Older, L. (1994). Morphology and meaning in the english mental lexicon. *Psychological review*, 101(1):3.

Marslen-Wilson, W. D. and Tyler, L. K. (1997). Dissociating types of mental computation. *Nature*, 387(6633):592–594.

Martin, S. E. (1960). *Korean reference grammar.* American Council of Learned Societies.

Matthews, P. H. (1974). *Morphology. An Introduction to the Theory of Word Structure.* Cambridge University Press, Cambridge.

McClelland, J. L. and Rumelhart, D. E. (1981). An interactive activation model of context effects in letter perception: Part I. An account of the basic findings. *Psychological Review*, 88:375–407.

Mikolov, T., Sutskever, I., Chen, K., Corrado, G. S., and Dean, J. (2013). Distributed representations of words and phrases and their compositionality. In *Advances in neural information processing systems*, pages 3111–3119.

Mitchell, J. and Lapata, M. (2008). Vector-based models of semantic composition. In *ACL*, pages 236–244.

Narrog, H. and Heine, B. (2018). *Grammaticalization from a typological perspective*, volume 31. Oxford University Press.

Neely, J. H. and Keefe, D. E. (1989). Semantic context effects on visual word processing: A hybrid prospective-retrospective processing theory. In *Psychology of learning and motivation*, volume 24, pages 207–248. Elsevier.

Neubauer, K. and Clahsen, H. (2009). Decomposition of inflected words in a second language: An experimental study of german participles. *Studies in Second Language Acquisition*, 31(3):403–435.

Nieder, J., Chuang, Y.-Y., van de Vijver, R., and Baayen, R. H. (2022). Comprehension, production and processing of Maltese plurals in the discriminative lexicon. *submitted.*

Nikolaev, A., Chuang, Y.-Y., and Baayen, R. H. (2023). A generating model for finnish nominal inflection using distributional semantics. *The Mental Lexicon.*

Norris, D. (2006). The Bayesian reader: Explaining word recognition as an optimal Bayesian decision process. *Psychological Review*, 113(2):327–357.

Park, H.-J., Koo, M.-M., and Nam, K.-C. (2012). 한국어 규칙동사와 불규칙 동사의 심성 어휘집 접근 과정 (The lexical access of regular and irregular Korean verbs in the mental lexicon). *Korean Journal of Cognitive Science*, 23(1):1–23.

Park, J. Y. (1998). 현대국어 반말체 종결어미 연구 (A study on Sentence final endings of Panmal style in Contemporary Korean). Seoul National University.

Park, J. Y. (2004). 한국어 양태 어미 연구 *(A study on Modal Endings in Korean)*. PhD thesis, Seoul National University.

Peelen, M. V., Romagno, D., and Caramazza, A. (2012). Independent representations of verbs and actions in left lateral temporal cortex. *Journal of cognitive neuroscience*, 24(10):2096–2107.

Perea, M. and Carreiras, M. (1998). Effects of syllable frequency and syllable neighborhood frequency in visual word recognition. *Journal of Experimental Psychology: Human perception and performance*, 24(1):134.

Pinker, S. (1991). Rules of language. *Science*, 253(5019):530–535.

Pinker, S. and Ullman, M. T. (2002). The past and future of the past tense. *Trends in cognitive sciences*, 6(11):456–463.

Plunkett, K. and Marchman, V. (1993). From rote learning to system building: Acquiring verb morphology in children and connectionist nets. *Cognition*, 48(1):21–69.

Pollatsek, A. and Hyönä, J. (2005). The role of semantic transparency in the processing of finnish compound words. *Language and Cognitive processes*, 20(1-2):261–290.

Portin, M., Lehtonen, M., Harrer, G., Wande, E., Niemi, J., and Laine, M. (2008). L1 effects on the processing of inflected nouns in l2. *Acta Psychologica*, 128(3):452–465.

Prasada, S. and Pinker, S. (1993). Generalisation of regular and irregular morphological patterns. *Language and cognitive processes*, 8(1):1–56.

R Core Team (2024). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.

Rumelhart, D. E., Hinton, G. E., McClelland, J. L., et al. (1986). A general framework for parallel distributed processing. *Parallel distributed processing: Explorations in the microstructure of cognition*, 1(45-76):26.

Rumelhart, D. E. and McClelland, J. L. (1986). On learning the past tenses of english verbs. *Psycholinguistics: critical concepts in psychology*, 4:216–271.

Scarborough, D. L., Cortese, C., and Scarborough, H. S. (1977). Frequency and repetition effects in lexical memory. *Journal of Experimental Psychology: Human Perception and Performance*, 3:1–17.

Schreuder, R. and Baayen, R. H. (1995). Modeling morphological processing. *Morphological aspects of language processing*, 2:257–294.

Seidenberg, M. S. and McClelland, J. L. (1989). A distributed, developmental model of word recognition and naming. *Psycholgical Review*, 96:523–568.

Seong, J.-y. (2006). 초급 한국어 학습자의 어미 사용 양상-어말어미의 사용을 중심으로 (Patterns of using endings of adult Korean language learners of elementary level). 어문연구 *(The Society for Korean Language Literary Research)*, 34(4):497–516.

Shafaei-Bajestan, E., Moradipour-Tari, M., Uhrig, P., and Baayen, R. H. (2022a). Semantic properties of english nominal pluralization: Insights from word embeddings. *arXiv preprint arXiv:2203.15424*.

Shafaei-Bajestan, E., Uhrig, P., and Baayen, R. H. (2022b). Making sense of spoken plurals. *The Mental Lexicon*, 17(3):337–367.

Shen, T. and Baayen, H. R. (2022). Productivity and semantic transparency: An exploration of word formation in Mandarin Chinese. *The Mental Lexicon*.

Silva, R. and Clahsen, H. (2008). Morphologically complex words in L1 and L2 processing: Evidence from masked priming experiments in English. *Bilingualism: Language and cognition*, 11(2):245–260.

Sohn, H. M. (1994). *Korean*. Routledge.

Song, J. J. (2005). *The Korean language: Structure, use and context*. Routledge.

Song, W. (2009). 국어 선어말 어미의 심리적 실재성 검증 (A Verification of Psychological Realities of Prefinal Endings in Korean). 한국어문학회, 104:83–102.

Stupak, I. V. and Baayen, R. H. (2022). An inquiry into the semantic transparency and productivity of german particle verbs and derivational affixation. *The Mental Lexicon*, 17(3):422–457.

Taft, M. (1981). Prefix stripping revisited. *Journal of Verbal Learing and Verbal Behavior*, 20:289–297.

Taft, M. (1994). Interactive-activation as a framework for understanding morphological processing. *Language and cognitive processes*, 9(3):271–294.

Taft, M. and Forster, K. I. (1975). Lexical storage and retrieval of prefixed words. *Journal of Verbal Learning and Verbal Behavior*, 14:638–647.

Ullman, M. (2004). Contributions of memory circuits to language: The declarative/procedural model. *Cognition*, 92:231–270.

Ullman, M. T. (2001). The neural basis of lexicon and grammar in first and second language: The declarative/procedural model. *Bilingualism: Language and cognition*, 4(2):105–122.

Vainio, S., Pajunen, A., and Hyönä, J. (2014). L1 and l2 word recognition in finnish: Examining l1 effects on l2 processing of morphological complexity and morphophonological transparency. *Studies in second language acquisition*, 36(1):133–162.

van de Vijver, R., Uwambayinema, E., and Chuang, Y.-Y. (2021). Comprehension and production of kinyarwanda verbs in the discriminative lexicon. Manuscript, University of Düsseldorf/Tübingen.

VanPatten, B. (2004). *Processing instruction: Theory, research, and commentary*. Routledge.

Westbury, C., Keith, J., Briesemeister, B. B., Hofmann, M. J., and Jacobs, A. M. (2015). Avoid violence, rioting, and outrage; approach celebration, delight, and strength: Using large text corpora to compute valence, arousal, and the basic emotions. *The Quarterly Journal of Experimental Psychology*, 68(8):1599–1622.

Widrow, B. and Hoff, M. E. (1960). Adaptive switching circuits. *1960 WESCON Convention Record Part IV*, pages 96–104.

Wieling, M. (2018). Analyzing dynamic phonetic data using generalized additive mixed modeling: A tutorial focusing on articulatory differences between L1 and L2 speakers of english. *Journal of Phonetics*, 70:86–116.

Wood, S. N. (2017a). *Generalized Additive Models*. Chapman & Hall/CRC, New York.

Wood, S. N. (2017b). *Generalized additive models: an introduction with R*. chapman and hall/CRC.

Xing, J. Z. (2015). A comparative study of semantic change in grammaticalization and lexicalization in Chinese and germanic languages. *Studies in Language. International Journal sponsored by the Foundation "Foundations of Language"*, 39(3):593–633.

Yang, B. S. (1994). *Morphosyntactic phenomena of Korean in role and reference grammar: psych-verb constructions, inflectional verb morphemes, complex sentences, and relative clauses*. PhD thesis, State University of New York, Buffalo.

Yi, K.-O., Bae, S.-B., and Nam, K.-C. (2005). Processing of orthography and phonology in reading sino-Korean words. In *Annual Conference on Human and Language Technology*, pages 63–72. Human and Language Technology.

Yim, H., Lim, H., and Nam, K. (2003). 한국어 용언과 체언어절의 형태소 정보처리 특성 (Representations of morphological information processing in Korean verb and noun Eojeols). 한국인지과학학회 춘계학술대회 발표 논문집 *(The Korean Society for Cognitive Science)*, pages 192–194.

Yune, Y. (2007). 중국인 학습자의 한국어 어미오류 분석: 연결어미와 종결어미를 중심으로: 연결어미와 종결어미를 중심으로 (A study on the verb ending errors of Chinese KFL learners). 코기토 *(Cogito)*, (61):31–58.