ONSET TENSIFICATION IN CONTEMPORARY KOREAN: NOVEL PRONUNCIATIONS AS EVIDENCE OF CONTINUING HISTORICAL PHONOLOGICAL PRESSURES

by

Roderick G. Clare

A Dissertation

Submitted to the Faculty of Purdue University

In Partial Fulfillment of the Requirements for the degree of

Doctor of Philosophy

Department of Linguistics

West Lafayette, Indiana

May 2021

THE PURDUE UNIVERSITY GRADUATE SCHOOL
STATEMENT OF COMMITTEE APPROVAL

Dr. Mariko M. Wei, Chair

Department of Linguistics

Dr. Atsushi Fukada

Department of Linguistics

Dr. Olga Dmitrieva

Department of Linguistics

Dr. Jessica Sturm

School of Languages and Cultures

Approved by:

Dr. Alejandro Cuza-Blanco

For E.K., Erin and 사랑이

"I will build my love a bower
by yon cool crystal fountain
and round it I will pile
all the wildflowers of the mountain"

ACKNOWLEDGMENTS

I would like to thank my committee for their support in making this project possible, with special thanks to my advisor and committee chair, Dr. Mariko M. Wei, to whom I owe a special debt of gratitude. Without her encouragement and guidance, it is likely that my research activities would have stalled during the unprecedented circumstances of the plague year. Each of my committee members deserves recognition for their support, not only during my dissertation year, but throughout my time at Purdue. Dr. Atsushi Fukada provided valuable technical expertise and assistance, as well as his characteristic good humor; Dr. Olga Dmitrieva taught me much of what I know about phonology and has proved an invaluable source of methodological guidance; and it was in the classroom of Dr. Jessica Sturm where I first began to flesh out the topic, although I suspect that she did not anticipate how many pages on Korean historical phonology those first forays would ultimately entail reading. It is to all the members of my committee's credit that they have participated so enthusiastically in a project which may be far removed from their primary interests.

I also want to thank Dr. Alejandro Cuza-Blanco and Dr. Venetria Patton for their help in keeping my project funded while it underwent near-total revision; without their support behind the scenes, this would have been a grim year for my growing family.

This fall will mark seventeen years since I first scratched out my name in shoddy *hangul*, and fifteen since I first set foot in Korea, never to truly leave. I want to thank Kyu-Young Park and Sangmin Bae, who each played a pivotal role in this journey. Likewise, the peripatetic linguist could ask for no more loyal and helpful friend in his wanderings than Paul Delap.

I am indebted to my wife, Eunkyung, for her constant support and for tolerating an endless stream of questions on points of Korean philology; alas, the questions will no doubt continue.

Finally, I would like to thank my mother and father, Trisha and Michael. As scientists they display that keen intellectual curiosity so common to their kind; they temper this with a very uncommon tolerance for the antics of a son. This is as much theirs as it is mine.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

Korean phonology features a cross-linguistically rare tripartite contrast in its stop series between lax, tense, and aspirated segments. Extant evidence suggests this contrast is the result of a fifteenth-century phonological restructuring wherein tense segments, previously an allophone of lax sounds, achieved distinct phonemic status. However, the historical record suggests that almost immediately a pattern of lax segments 'tensifying' began, with words featuring lax onset sounds being realized increasingly with tense sounds until the novel pronunciation was universal. While the action of these shifts is sporadic throughout the lexicon, the resulting changes are unidirectional, with the domain of tense segments expanding at the cost of lax sounds. It has been posited in previous research that such sound changes may suggest a rebalancing of functional load across underutilized segments (Shin & Davis, 2008).

A similar phenomenon in contemporary Korean where speakers exhibit differing pronunciations of onset segments in a number of lexical items is analyzed herein, with the argument that it is best understood as the continuation of these historical processes. Far from an idiosyncratic speaker habit or dialectal quirk, these unexpected tense segments can be interpreted as surface evidence of phonological pressures active since late Middle Korean. The present study explored novel tensified onset pronunciations from a demographic standpoint, aiming to clarify which speaker populations have adopted new variant forms through two experiments. The first featured the elicitation of 'tensification-prone' items by native speakers in a production task, while the second used a combination of acceptability judgments of tensified items and attitudinal surveys regarding the use of novel tense pronunciations.

The results confirm that tensification is active in contemporary Korean, but that a decisive conclusion as to its demographic associations remains elusive. The acceptability judgment experiment suggests that younger speakers and self-affirmed dialect users are more likely to prefer tensified variants, while the production task revealed no significant relationship between these factors and actual pronunciation behavior. Finally, the findings are considered in context of deeper changes in Korean phonology whereby tense and lax segments are increasingly associated with word onset and medial/final position, respectively.

CHAPTER 1.   INTRODUCTION

The Korean language has a typologically rare three-way contrast in its stop series, featuring phonemically distinct voiceless lax (lenis), tense (fortis), and aspirated stops.

/tal/	/t*al/[1]	/tʰal/	/pul/	/p*ul/	/pʰul/
"moon"	"daughter"	"mask"	"fire"	"horn"	"grass"

While many world languages achieve contrasts between as many or more segments through a combinatorial application of voicing and types of phonation such as breathy or creaky voice, Korean contrasts between three sets of voiceless sounds, which bear greater phonetic similarity [2]. This contrast is the product of a restructuring of the Korean phonological system which extant evidence suggests began in the fifteenth century, wherein tense segments— previously an allophonic variant of lax segments— acquired full phonemic status (Kim et al., 2008; Lee & Ramsey, 2011). In the intervening period between this phonological restructuring and the present day, the expansion of tense forms to the detriment of their lax counterparts has been observed in the Korean lexicon (Bae, 2003:256-7; Shin & Davis, 2008). Diachronically, lax segments in word-onset position of select words have become tense, with the change occurring sporadically and idiosyncratically at first, as evidenced by conflicting spelling in period texts (Lee & Ramsey, 2011:133). However, over longer periods, the novel pronunciation ultimately and irreversibly displaces the original, with the change universally diffused throughout the language.

Despite documentation of word-onset tensification as a historically active process, a like phenomenon in modern Korean whereby some speakers replace lax segments with tense ones in word-onset position has been treated until relatively recently as either speaker error, a dialectal variation restricted to regional use, or as added for emphatic purposes arbitrarily by speakers. Only in the past few decades have researchers come to challenge these presumptions, with some arguing

---

[1] The convention used for phonetic representation of tense segments varies from author to author, although the International Phonetic Alphabet's suggested /k̚/ is not widely used in Korean studies. In this paper I will use an asterisk to represent tense segments (for instance, /k*/,) following the precedent of many linguists working with Korean.

[2] For instance, Silva (2006) demonstrates that for some Korean speakers born after the 1960s, the voice onset time (VOT) of lax and aspirated stop segments have come to overlap, with speakers increasingly coming to rely on F_0 as a primary acoustic cue to distinguish these segments.

that word-onset tensification is better understood as part of a greater ongoing process of phonological sound change in Korean (Han, 2011, 2014, 2015). Rather than assume that historical sound changes were conveniently arrested before our time, new experimental research has suggested that the Korean lexicon is still reshaping itself phonologically, with younger generations more likely to use tensified variants (Jeon, 2019).

However, although recent scholarship has clarified the issue and highlighted the need for new perspectives on Korean word-onset tensification, further research is needed to tease out where the sounds of Korean are changing, both within the lexicon and in the literal sense: what speaker populations are adopting novel pronunciations? The relationship between dialectal region and tendencies to tensify remain murky. By identifying the populations where sound changes take root, and through highlighting extralinguistic factors shared by such trend-setting speakers, is it possible to shed light on how phonological changes spread? The present study aims to add to the scholarship on novel tensification habits in Korean through two experiments featuring native-speaker participants.

The first experiment combines a phonological production task featuring tensification-prone lexical items with detailed background surveys, in the hope of teasing out novel pronunciations as well as providing a clearer picture of speaker-specific factors that may lie behind such trendsetting behaviors. The second experiment shifts the focus from participant productions to participant attitudes, with respondents offering judgments on the acceptability of novel tensified pronunciations for given lexical items as well as their perceptions of the use of such forms.

1.1 Research Questions and Hypotheses

Onset tensification in contemporary Korean has received some attention in Korean-language scholarship, both from a theoretical standpoint regarding its wider ramifications for Korean phonology and through a small body of empirical work. However, large-scale data on onset tensification phenomena remains limited to a handful of studies, with scant data from natural language production experiments. The question of demographic or other speaker-specific factors (besides participant age) which may belie adoption of novel tense forms has been little considered. The following research questions and hypotheses were drawn with the consideration of filling

these gaps in the research. This is not to disparage or diminish the work of previous researchers; if anything, all of the studies referenced herein have proven vital to understanding the issue and in creating a robust methodology for approaching onset tensification phenomena. Finally, as the bulk of discussion of onset tensification has appeared in Korean-language sources, it is hoped that the present study will encourage further consideration of the topic among Korean linguists working in an English-language medium.

1.1.1   Experiment 1: Production of Tensified Onsets

RQ1.   Will the findings confirm those of previous investigations into onset tensification, successfully eliciting tensified pronunciations of experimental items?

H1.   Previous empirical studies have succeeded in capturing tensified onsets (Han, 2011; Hong, 2011, 2014; Kim, 2015), sometimes showing that tense variants are produced overwhelmingly by speakers (Han, 2015). Further survey research has suggested participants may prefer tense forms on some lexical items when given a choice (Kim, 2004; Lee, 2009; Hong, 2014). Speakers have been shown using tense forms even against their own judgment (Jeon, 2019). It is predicted that some participants will produce tensified onsets on experimental items, given that many of the lexical items included have been demonstrated in prior studies as prone to tensification (Han, 2013; Jang, 2017).

RQ2.   Will participant ages display a significant relationship with the degree of tense forms produced?

H2.   While no clear consensus emerges from the literature, with some researchers claiming no significant difference in the degree of tensification by age group (Lee, 1999; Hong, 2011), other studies have suggested that tensification is increasingly common in younger generations (Kim, 2004; Jeon, 2019). Silva (2006) has demonstrated that significant phonological change has happened in Korean between speakers born before and after around 1960; although the study in question focused on acoustic cues, it shows that age must be considered in a large-scale analysis of spoken Korean. It is predicted that younger participants will exhibit greater use of tensified forms, following the results of previous studies.

RQ3. Is a significant difference observed in the degree of tensified forms produced between participant groups of different genders?

H3. While some previous studies have suggested a tendency for male speakers to exhibit greater use of novel tense pronunciations (Lee, 1989; Kim, 2004), all empirical studies conducted so far suggest that speakers of both sexes produce tense forms, with many showing no significant difference in production (Hong 2011, 2014; Jeon, 2019). It is noted that some researchers report participants' attitudes of tensified forms as 'masculine' or 'vulgar' (Lee, 1996), although it cannot be concluded from this that male speakers are more likely to actually realize novel tense forms in speech. It is hypothesized that no significant disparity in tensification will be observed across gender lines.

RQ4. Do speakers of different dialects or regional backgrounds demonstrate a marked disparity in their production of novel tense forms?

H4. The assignment of tensification phenomena to a specific dialect has been the greatest point of disagreement in the literature. Some scholars propose that the phenomenon must be best viewed as a peninsula-wide change, occurring across all dialects of the language (Lee, 1999; Jang, 2017), or that it the dispersal of tensified items across different regions inhibits a clear and neat definition of onset tensification as belonging to a specific dialect (Park, 2000). However, many researchers have considered tensification as particularly associated with a specific geographical variety of Korean, albeit with great disagreement as to which dialect it is characteristic of (Sohn, 1999; Lee, 2002; You, 2006; Wee, 2008; Oh, 2011; Yeon, 2012, among others). Although use of tensified variants is attested across all dialect areas, the lack of consensus suggests that there are nevertheless tangible differences between varieties. It is hypothesized that speakers of different dialects will demonstrate different degrees of use of novel tense forms.

RQ5.   Do participants with a self-reported proficiency in a foreign language exhibit greater likelihood of using novel tense pronunciations?

H5.   Observing that onset tensification often happens in foreign loanwords, Kang (2008) proposed a phonological motivation: speakers may be attempting to preserve voiced onsets of the words in question. This is supported by a study by Jeon (2016) which showed broadcast journalists produced tensified onsets in loanwords, a finding all the more interesting given the stricter adherence of such journalists to the standard. Many of the languages from which Korean has borrowed heavily in the modern era utilize voiced onsets, such as English and Japanese. Greater familiarity with these languages, and experience accommodating to novel pronunciations, could plausibly lead a speaker to follow new phonological trends in their own native tongue. It is hypothesized that speakers with high self-reported proficiency in a foreign language will produce more novel tense forms.

Self-assessment of language proficiency is admittedly an imperfect measurement. However, given the constraints of administering both a recording task and a thorough survey including other questions about demographics and language use, and with the complication that participation was open to native adult speakers regardless of L2 background, self-assessment was chosen as an imperfect but hopefully satisfactory measurement for the purposes of this study.

RQ6.   Is a significant relationship observed between participant language attitudes, as measured using questionnaires, and use of tensified pronunciations?

H6.   It is hypothesized that participants who report stronger conservative attitudes towards language change and language use will produce fewer tensified variants. As previous studies have shown that speakers themselves may associate tensification with dialect use (Kang, 2001; Hong, 2014), positive attitudes towards dialect use and preservation of dialects is also hypothesized to correlate with higher use of tensified onsets.

RQ7.   Does use of novel tense forms relate to individual differences in personality, as measured using self-evaluated personality tests?

H7.   Research on individual differences in personality and its relationship to phonology have focused on second- or foreign-language contexts, and the results have not produced a great consensus. A relationship between performance in a phonetic imitation task and openness was suggested by Yu et al. (2013), while Dewaele and Furnham (1999) posit that extraversion may influence L2 speech production, which they confirmed in a follow-up study (Dewaele & Furnham, 2000). Extraversion, agreeableness, and neuroticism have all been suggested to influence second language acquisition and performance during L2 speaking tasks (MacIntyre & Charos, 1996; Oya et al., 2004; Zárate-Sández, 2017). Other researchers hold that evidence for a relationship between personality and L2 proficiency is weak (Busch, 1982; Carrell et al., 1996; Hinton, 2014). However, little has been done to examine the relationship between personality and adoption of novel phonology in the speaker's native language. It is hypothesized that participants with higher openness and extraversion scores on a self-assessed personality questionnaire will produce more tense variant onsets.

1.1.2   Experiment 2: Acceptability of Tensified Onsets

RQ1.   Will participants demonstrate evidence of word-onset tensification by indicating novel tense pronunciations as equally (or more) natural in comparison to lax forms?

H1.   Many previous studies have demonstrated that speakers produce tensified onsets (Han, 2011, 2015; Hong, 2011, 2014; Kim, 2015), and others have established that surveyed participants may prefer tense forms (Kim, 2004; Lee, 2009; Hong, 2014), with further research suggesting that tense forms may be realized even when speakers report themselves not to be using them (Jeon, 2019). The present study includes many of the same stimulus items attested in these previous studies as prone to tensification, so it can be confidently hypothesized that participants will deem some of the tensified forms as more or equally natural to their lax-segment-bearing counterparts.

RQ2.   Will participants of different age groups display significant disparities in the degree of tense forms preferred?

H2.   Researchers in previous studies have drawn different conclusions about the use of onset tensification by different age groups. Some have posited that disparities in use by age group, if at all present, are insignificant (Lee, 1999; Hong, 2011). Others have suggested evidence of a clear

tendency for younger speakers to tensify more (Kim, 2004; Jeon, 2019). It is predicted that younger participants will exhibit greater acceptance of tensified forms, following the results of previous studies.

RQ3. Is a significant difference in the degree of preference for tensified forms demonstrated across gender lines?

H3. While a number of researchers have suggested male speakers use tensified forms more (Lee, 1989; Kim, 2004) and others have posited that speakers perceive tensified forms as more masculine or vulgar (Lee, 1996), the larger share of empirical studies into the phenomenon have suggested that no significant difference is observed in production across gender lines (Hong 2011, 2014; Jeon, 2019). It is hypothesized that no significant difference will be observed in the preference for tensified forms by gender group.

RQ4. Do speakers of different dialects or regional backgrounds demonstrate a marked disparity in their preference of novel tense forms?

H4. Many researchers have associated word-onset tensification with a specific dialect (Sohn, 1999; Lee, 2002; You, 2006; Wee, 2008; Oh, 2011; Yeon, 2012, among others) but no clear consensus has emerged. Previous studies have demonstrated tensified lexical items are dispersed varyingly across dialects (Park, 2000; Han, 2011, 2013). Some scholars have thus proposed that the change is better understood as occurring peninsula-wide, although on different items (Lee, 1999; Jang, 2017). It is hypothesized that speakers of different dialects will demonstrate different degrees of acceptance of novel tense forms.

RQ5. How do participants perceive the use of tensified pronunciations?

H5. A number of studies have suggested that tensified pronunciations are perceived by speakers as less formal (Lee, 1989; Lee, 1996; Kang, 2001; Kim, 2001; Jeon, 2019). Other scholars have suggested that tensification is associated with masculinity (Lee, 1996), perhaps to the point where female speakers' use of it is perceived negatively (Hong, 2014). Other researchers have posited that tensification lends the item a negative connotation (Park, 2000). However, speakers may

overall indicate positive attitudes to use of tense forms (Jeon, 2019). Sociolinguistic research into the phenomenon has suggested speakers may explicitly associate tensified forms with regional dialects (Kang, 2001; Hong, 2014). Further work has indicated word-onset tensification may serve discursive purposes, such as emphasis (Lee, 1996; Park, 2000; Kim, 2001). It is predicted that participants in the present experiment will associate tensified forms as less formal, more masculine, less feminine, more friendly, more vulgar, and more emphatic and expressive. It is further hypothesized that speakers will explicitly associate tensification with regional dialects, with participants having positive attitudes towards tensification more likely to attribute it to their own varieties of Korean.

1.2 Summary

This chapter provided a brief account of onset tensification phenomena in contemporary Korean and possible relationships to documented historical phonological changes in the language. Following a short summary of previous treatments of the topic and how the scholarship has come to view it from new angles in recent studies, remaining gaps in the research were outlined. The research questions and hypotheses for the experiments featured in the present study were introduced with reference to insights provided by previous works. In Chapter 2, a detailed literature review will cover the roots of onset tensification in historical Korean phonology and continue to a discussion of contemporary tensification trends.

CHAPTER 2.   LITERATURE REVIEW

Korean exhibits a cross-linguistically rare three-way contrast between voiceless lax, tense, and aspirated stops. Lax stops (also commonly referred to in research as lenis stops) have been characterized in the literature as "breathy and slightly aspirated" with tense (or fortis) stops as "tense, laryngealized and unaspirated" (Cho et al., 2002:194). The same contrast between lax and tense segments also exists in the Korean fricative and affricate series. This tripartite contrast is the product of relatively recent phonological changes within Korean. Before delving into the specifics, a general overview of how historical linguists in Korean studies divide the history of the language may be helpful.

2.1   History of the Korean language

The ultimate origins of the Korean language and possible genetic relationships to other languages or families remains a matter of debate, although some scholars suggest an affinity to the Altaic language family and possibly Japonic languages (Lee & Ramsey, 2011; Martin, 1966; Miller, 1980.) Fortunately, although evidence of earlier forms of Korean are scant, ample extant written material from the last few centuries allows us to understand how the language has evolved in more recent times. In particular, Korean is noted for using a unique alphabet (*hangul*), developed by a team of scholars appointed to the task by the forward-thinking King Sejong and promulgated in 1446. Even before the development of *hangul*, the ancestors of the modern Korean people were using Chinese characters to represent their language for considerably longer. Chinese characters have been used on the Korean peninsula for over two millennia. The proto-Korean state of Old Joseon was destroyed by the Han empire in the year 108 B.C., with the use of Chinese characters presumably spreading during a successive period of Han domination. Classical Chinese came to serve as the literary and legal language of all the proto-Korean kingdoms that later occupied the peninsula. Virtually all extant prose, poetry, documents, and inscriptions dating before the introduction of *hangul* are in Classical Chinese, although the ancestors of the Koreans also devised ways of using Chinese script to represent their own language, by transcribing the sounds using characters whose Chinese pronunciation was roughly similar to that of the native syllable (in a fashion similar to the Japanese *man'yōgana*, although it should be noted that extant examples are

far fewer in Korea.) These sparse examples— personal or place names recorded on stelae, a few lines of verse preserved in later compilations— comprise the entire body of sources from which we can know Old Korean.

Middle Korean is typically further divided into Early Middle and Late Middle periods. The Early Middle period follows the establishment of the Koryŏ dynasty (from which the English toponym *Korea* derives) and its unification of the peninsula in the late tenth century. Documented evidence of Early Middle Korean is scant, but the period marks the entrenchment of the prestige dialect in the modern Seoul region, where the Koryŏ kings established their capitals. Late Middle Korean refers to the period after the introduction of *hangul* in the fifteenth and sixteenth centuries, from which many documents survive, composed wholly in the native vernacular. Lee and Ramsey (2011) are effusive about the wealth of source material for this period:

> Its texts are consistent and phonologically precise, the textual corpus rich and voluminous. Its transcriptions record segmentals and suprasegmentals; the symbols incorporate articulatory features; spellings are standardized. For both phonological and morphological information, this textual record is unsurpassed anywhere in the premodern world. (p. 7)

Thus, from the Late Middle Korean period onward, it is possible for philologists to surmise about the phonological system of the Korean language with some confidence.

Early Modern Korean refers to the period between the beginning of the seventeenth century and the late nineteenth, when *hangul* came to be used in a wider variety of genres, from personal letters and diaries to poetry, prose, and novels. Predictably, the entrenched aristocracy had recognized the threat posed by *hangul* in promoting mass literacy almost as soon as it had been adopted, with the tyrannical prince Yeonsan going so far as to ban its use in 1504. Fortunately, these efforts proved in vain; although Chinese remained the written language of high culture, the masses continued to use *hangul-* in particular, those who were barred from pursuing education owing to sex or caste. Concurrent with the shift towards use of *hangul* in private and informal contexts, wider public adoption of the script was also served to weaken orthographical standards. The consequent increase in variant spellings, as speakers sought to represent the sounds of vernacular speech, permits a view into how Korean phonology was evolving contemporaneously.

Modern Korean refers to the form of the language from the late nineteenth century to the present time. During the Japanese occupation (1910-1945), Korean-language education was suppressed, with the perhaps predictable result of ensuring that the end of the occupation saw the

rise of an expressly nationalistic fervor for using Korean. Over the twentieth century, increased literacy and mass media have brought a return to stronger adherence to orthographical standards. A language academy tasked with regulating Korean, the National Institute of Korean Language, was formally brought under the direct control of the South Korean Ministry of Culture in 1990. The North Korean government similarly maintains its own standard, but it is worth noting that the foundational (and often underground) linguistic conferences that brought the first standardizing efforts in the 1930s had already taken Seoul-area Korean to serve as the standard (with the Pyongyang dialect also not being dramatically different,) and thus the two "standard Koreans" disagree mostly on various orthographical conventions or on foreign loanwords. Although largely mutually intelligible, a wide dialectal diversity is seen across the peninsula, and one could equally say of the Koreans what is said of the Britons and Americans: a people separated by a common language.

2.2   The emergence of a tripartite stop series in Korean

The paucity of extant Old Korean records makes it difficult to say with certainty how its stop series was structured. Although any placement of Korean in a larger language family is still a matter of debate, evidence of genetic affinities to Altaic languages has been demonstrated, particularly with Manchu and other Tungusic languages and (with somewhat less evidence) to Japonic languages (Lee & Ramsey, 2011; Martin, 1966; Miller, 1980). Proto-Altaic has been reconstructed with a phonemic voicing distinction, although no evidence for this remains in Old Korean, suggesting that such a distinction was already lost. The scholarship is divided as to whether aspiration was used contrastively at this point, although the disagreement lies in the interpretation of the evidence. Lee & Ramsey (2011) argue that aspirated segments were already phonemic, albeit rare (2011:20), with Old Korean contrasting between 'plain' and aspirated obstruents. Kim et al. (2008) posits that at this stage aspiration was not yet phonemic but the result of morphophonological interaction, occurring on coda sounds preceding /x/, and that the development of separate aspirated phonemes happened phoneme-by-phoneme and was not settled until Middle Korean. Either interpretation depends on how to interpret the way Chinese borrowings were "nativized" by Old Korean speakers, as well as an assumption of when such borrowings essentially fossilized. Then as now, aspirated stops were phonemically distinct in Chinese. However, while many characters borrowed during the Old Korean period are pronounced

with aspirated onsets /pʰ/ and /tʰ/, the set of characters pronounced with the velar /kʰ/ is vanishingly small. A typical Korean *okpyeon* (玉篇), as Chinese character dictionaries are often still called in Korea, has but a single entry for characters beginning with this sound (the lonely 快, pronounced /kʰwe/.) Virtually all of the characters featuring aspirated velar stops in Chinese were borrowed without retaining the aspiration and have remained so. Kim and his co-authors suggest that this reflects the lack of a phonemic velar aspirated stop in Old Korean. Lee and Ramsey argue the alternative: if these borrowings allow that Old Korean used aspiration contrastively with bilabial and dental stops, demonstrated aspirated velar onsets in extant Middle Korean words suggest that although dental and bilabial aspirates may have emerged earlier and spread more widely than their velar counterpart, there is no reason to think a corresponding velar aspirated stop was absent in Old Korean. It could merely have been less frequent, a claim bolstered by phonological frequency data of contemporary Korean gathered by Shin (2011) that demonstrates /kʰ/ is less frequent than its bilabial and dental counterparts, although for lax and tense unaspirated stops the reverse is true. Korean linguists agree, however, on the rest of the stop series: Old Korean- and Early Middle Korean beyond- did not have a phonemic distinction between tense and lax sounds. As Lee & Ramsey (2011:90) bluntly put it, "[tense] consonants were almost certainly not phonologically distinctive at an earlier stage of Middle Korean, in the twelfth century. There is no philological evidence for such a consonant series." The possibility remains that tense sounds occurred as an allophone of lax sounds in certain morphophonological contexts, although no extant evidence from Old Korean allows such a conclusion.

Several documented sources of Middle Korean have survived, allowing for greater speculation as to the phonological structure of the language at this time and possible sound changes that had occurred in the intervening period. By this point the phonemic status of aspirated segments was clearly established, while the evidence suggests that tense sounds were not yet phonemic- albeit with several curiosities emerging from the historical record to allow debate. A lexicon of around 350 words is found in transmitted fragments of the *Jilin leishi* (鷄林類事,) a twelfth-century Chinese work on Korea written by a member of a Song dynasty embassy to the kingdom of Koryŏ. Entries in the lexicon consist of a Chinese gloss with an approximation of the Korean pronunciation in Chinese characters, allowing a tentative reconstruction of the original Korean, if filtered through Middle Chinese phonology. Taking the entry for *crow* as an example, following

Baxter's (1992) work on reconstructing the sounds of Middle Chinese (superscript X represents rising tone:)

鴉曰打馬鬼　"*Crow* is pronounced /tæŋ^X.mæ^X.k^{h}woj/ "

(cf. Modern Korean /k*amagwi/)

Several entries in the *Jilin leishi* suggest that between the time of writing in the early twelfth century and the introduction of *hangul*, Korean phonology changed in a way that would be consequential for the eventual emergence of tense sounds. For example, the pronunciation for word *rice* is given as 菩薩, allowing a reconstruction of **posol* (Lee & Ramsey, 2011:89). However, by the fifteenth century the word appears to have lost the vowel of the first syllable, resulting in a consonant cluster. The *Seokbosangjeol* (釋譜詳節, a fifteenth-century biography of the Buddha and one of the first books printed in *hangul* using moveable type) has the word as 뿔, allowing a reconstruction of **psol*, suggesting that vowel syncope resulted in the emergence of consonant clusters (Lee, 1968). Such clusters are disallowed in modern Korean; the modern word for *rice* is /s*al/, suggesting that tense sounds might have developed out of consonant clusters. It is possible that the words were already pronounced with tense onsets rather than clusters in the fifteenth century, but that orthographic convention at the time preserved an archaic pronunciation in form if not practice (as in the fiendish spelling of many English words.) It is also possible that some clusters were an artifact of unclear orthographic conventions wherein a glyph was not being used for its phonetic meaning, but to mark a special kind of articulation. Lee & Ramsey (2011) raise the example of another *Jilin leishi* item: the word for *daughter* is given as 寶姐, permitting a reconstruction of **potol* (cf. modern Korean /t*al/.) Yet intriguingly, in Late Middle Korean texts the word is uniformly written as 똘, which would be **stol,* and never as **ptol*. This leaves open the possibility that the use of the glyph nominally associated with /s/ was not to indicate a complex onset, but rather a tense pronunciation of the following sound. This possibility will be discussed in detail later.

Despite that curious example, it is likely that a cluster was pronounced at *some* point in the development of Korean. We can ascertain this from words in modern Korean where

multimorphemic words have preserved the first sound of the cluster as the coda of a preceding syllable. For example, the modern Korean word for rice made with the millet plant is /tɕop.s*al/, although the name of the plant itself is simply /tɕo/ (keeping in mind as mentioned above that *rice* is /s*al/.) It would be difficult to explain the intervening bilabial without considering the possibility that, at some point, the morpheme for *rice* began with a complex onset.

Another piece of compelling evidence in support of tense sounds as having developed from consonant clusters comes, rather unexpectedly, from one of the earliest documented contacts between Westerners and Koreans— and a fascinating historical episode. In 1653 a Dutch East India Company ship, *De Sperwer,* set off from Batavia (present-day Jakarta, in Indonesia) on a trade mission to the Japanese port of Nagasaki. Unbeknownst to the hapless crew they (or more accurately, the survivors) would not set foot in Nagasaki until some thirteen years later, in 1666. *De Sperwer* was to encounter heavy storms just off the coast of Jeju-do, a large island off the Korean peninsula, and the shipwreck left thirty-six surviving crew stranded in a country no Westerner had ever been known to return from alive. While the worst fears of the survivors were to prove ungrounded- sailors rumored the natives to be cannibals- state policy of the then-ruling Joseon dynasty dictated that foreigners shipwrecked within the country were forbidden to leave, lest they bring valuable strategic information about Korea's defenses back to their native lands. The surviving crew spent thirteen years living in Korea and interacting with all stations of Korean life, for a time even serving as military assistants under the King's command, with some of the sailors taking local wives. However, as famine wracked the country, the surviving Dutchmen took an opportunity to escape to their original destination of Nagasaki in 1666, and from there were able to return back to the Netherlands. The whole adventure is recounted most famously by a Dutch East India Company bookkeeper, Hendrick Hamel, whose manuscript was published by four printers in Amsterdam and Rotterdam over 1668 and 1669 (before Hamel himself had even returned from Asia.) The first English translation[3] appeared in 1704 to great interest, providing the first glimpse into a country then nearly unknown.

Although Hamel is still celebrated in Korea, it was one of his fellow crewmen that left a contribution to Korean philology directly related to the question of how tense sounds evolved. In 1705, while compiling his study of Siberia and East Asia, *Noord en Oost Tartarye*, the Dutch

---

[3] Churchill, J. (1704) *An Account of the shipwreck of a Dutch vessel on the coast of the isle of Quelpaert, together with the description of the Kingdom of Corea.* London.

statesman Nicolaes Witsen interviewed another *De Sperwer* survivor, one Mattheus Eibokken. Although nearly four decades had passed since his escape from Korea, Eibokken remembered enough of the language to provide a small lexicon for Witsen's monograph. Among the items listed are two words which begin with tense sounds today, but whose renderings in Dutch orthography suggest that they began with a consonant cluster when Eibokken learned them (Kang 1995):

Witsen	*Middle Korean*	*Modern Korean*	*Gloss*
spaem	쌈 *spam* (救急方諺解구급방언해, 1446)	뺨 /p*jam/	"cheek"
stock	썩 *stək⁴* (月印釋譜월인석보, 1459)	떡 /t*ək/	"rice cake"

Further evidence suggesting that vowel syncope produced consonant clusters, which then simplified into tense sounds, can be found through reconstructing etymologies of semantically related words. Lee & Ramsey (2011) suggest that the similar phonological form of Late Middle Korean items such as *pskay* ᄢᅢ - "break," *psko* ᄢᅵ - "peel" and *pski* ᄢᅵ - "pierce" allows a reconstruction along the lines of *poso* "break, shatter." Notably, in all three cases the descendant word in contemporary Korean features a tense onset:

Reconstruction	*Late Middle Korean*	*Modern Korean*
poso- "break, shatter"	*pskay-* "break" →	/k*ɛ/ "break"
	psko- "peel" →	/k*a/ "peel"
	pski- "pierce" →	/k*i/ "pierce"

2.3 *Hangul* and early evidence of tensification phenomena

Although Chinese characters had been used by Korean writers for centuries to record their language, the differences between the two languages were and continue to be vast. Mismatch

---

[4] Although purely speculative, this reconstruction makes possible an etymological link with an archaic Japanese word for a similar kind of rice cake, *shitogi*.

between the phonological inventories and syllable structures of each language made it difficult to adapt Chinese characters for their phonetic value neatly to words of Korean. This problem was further compounded by morphosyntactic differences between the two languages, with Chinese largely isolating and Korean highly agglutinative. Nevertheless, complex systems developed such as *idu, hyangchal* and *gugyeol* for representing Korean with Chinese characters, often with a mix of Chinese phrases borrowed into the Korean lexicon used alternately alongside characters chosen for their phonetic or semantic value. Such texts were difficult for inexperienced readers to parse, and would-be readership was limited to elites who enjoyed the prerequisite familiarity with Chinese necessary to learn these conventions. Taking an example from a 1395 translation of the Ming dynasty legal code (adapted from Park, 2003:)

<p align="center">罪狀<u>乙</u>必于施行爲<u>遣</u>勿論爲<u>去乃</u></p>

<p align="center">"<i>Even if the matter of guilt is settled and the [punishment] not decided, or...</i>"</p>

The underlined characters are only being used for their phonetic values, while the bold characters are used for their semantic values. 乙 corresponds to the Korean object marker /ɯl/, and even in contemporary Korean the reading of this character is pronounced the same as the object marker. 遣, which had a Middle Chinese pronunciation presumed similar to the reconstruction by Baxter (1992) of */kʰjien^X/, represents the Korean conjunction *and*, /ko/; likewise 去乃 represents the conjunction *or*, /kəna/, following a reconstruction */kʰjo^X.noj^X/ (Baxter, 1992). 爲 is used for its semantic value, with one shade of its meaning being 'to do,' while 勿 is being used for its semantic value in place of the Korean negating phrase /ani/. In a single clause a reader is already required to correctly parse not only the characters themselves, but to be aware of when characters are used to represent Korean by sound, or even if they represent morphemes or complex words by semantic association. To render the phrase back into coherent Korean requires knowing where to insert the corresponding Korean morpheme(s) for those suggested by a character. A degree of convention developed early on among users of these systems, such as the use of the same character to represent a given syllable of Korean, rather than any character whose phonetic quality was plausibly suitable. While this simplified things to a degree, even contemporary writers realized that such devices were far from ideal (Lee, 2003). If this foray into Korean paleography leaves the reader dazed, then it

will have been successful in demonstrating why the need for orthographic reform was recognized at the highest levels of the Joseon court in the mid-fifteenth century.

To overcome the problems presented by previous conventions for representing Korean with Chinese, and with the stated aims of achieving the social benefits provided by increasing literacy and access to learning among the general population, the forward-thinking King Sejong commissioned a body of scholars to create a new alphabet better suited to the Korean language (Lee, 2003). The unique writing system that resulted, *hangul*[5], has the rare distinction of being a constructed script with well-documented premodern origins. The project to develop a new alphabet to represent Korean culminated in the promulgation of a new script in 1446 with the publication of the *Hunminjeongeum* (訓民正音,) "the correct sounds for instructing the people." Written in Classical Chinese, the document lays out the motivations for creating a new script and discusses the phonological principles behind it. Notable is a division of the syllable into onset, nucleus, and coda, in contrast to traditional Chinese phonology using the onset and rime. The glyphs are explained with explicit reference to place of articulation and voice quality, using Chinese characters bearing the phoneme in question as examples. A later appendix (the *Hunminjeongeum Haeryebon* 訓民正音解例本) provided additional examples using native words.

The creation of an alphabet explicitly aimed at representing the sounds of Korean as spoken in the mid-fifteenth century would plausibly allow us to determine the status of tense sounds at the time. If the sounds are represented by their own glyph, or if reference is made to such articulations, it would allow a conclusion that such sounds were already fully phonemic in the Korean of king Sejong's time. Perhaps if examples are offered, using words which we know from other historical records to contain tense segments, or using words which are realized with tense segments today, it would be permissible to conclude with certainty that the Korean stop series had already coalesced into the tripartite division we see today. It would be fortuitous that the development of *hangul* occurred precisely in a period during which we suspect the aspect of Korean phonology we are interested in to be in flux.

---

[5] I feel compelled to note that the term *hangul* is of modern origin, first coming into use around the turn of the twentieth century. The word is a compound of *han-,* which can be parsed etymologically as either meaning "great" or referring to the Korean people, some of whom have at times used the ethnonym *Han* (itself of a different origin than the Chinese *Han,*) and the word *gul* meaning "script." Those many qualifiers are needed; the modern ethnonym of the Koreans is itself tied to twentieth-century political debates, and different ethnonyms are in use on the peninsula. In North Korea, the term used for the script is *joseongul.* The document promulgating the script called it *hunminjeongeum* and contemporaneous records also refer to it as *eonmun* 諺文, "vernacular script."

Frustratingly, the *Hunminjeongeum* and its appendices make any straightforward affirmation of the status of tense segments difficult, precisely in what is *missing*. Although the glyphs used in modern orthographic convention to represent tense sounds (a doubling of the glyphs used to write their lax counterpart; ㄱ /k/ and ㄲ /k*/, for example,) are found in the *Huminjeongeum*, the document explicitly states that such glyphs are for writing "wholly muddy" (全濁) sounds, in opposition to "partly clear" (次清) and "wholly clear" (全清) sounds. This refers back to traditional Chinese phonology, although scholars disagree as to which phonological quality is referred to as muddy, with the primary candidates being voicing, breathy voice, or a combination of voicing and aspiration (Baxter, 1992). It is notable that Japanese speakers borrowed the Chinese clear/muddy distinction for the voicing distinction which still exists in their language. Conservative traditional tendencies may also mean that the phonological quality referred to may have changed without any emendation made to the term, making it difficult to surmise precisely what articulation is meant by *muddy*. Nevertheless, it was considered necessary for the new alphabet to be capable of representing it. It may seem counterintuitive that an alphabet explicitly designed to capture Korean phonology would feature separate glyphs used to represent sounds of Chinese, but it should be remembered that borrowing from Chinese into Korean has been so intense that fully sixty-five percent of the Korean lexicon can be characterized as Sino-Korean (Sohn, 1999). We cannot alternately conclude that muddy meant tense. Each of the Chinese characters given as an example featuring a wholly muddy sound in the *Hunminjeongeum* are pronounced with lax onsets in modern Korean. In the appendix offering examples for the glyphs using native Korean words, *no entries* are given for the set of glyphs representing wholly muddy sounds— although those glyphs are used today to represent tense sounds. The critical information that would allow us to definitively state that a phonemically tripartite stop system had fully formed in 1446 is tantalizingly lacking.

Kim et al. (2008) suggests that although the wording is vague, it is possible to parse the pertinent sentence of the *Hunminjeongeum* as referring to tense sounds. Forgiving a loose translation:

全清並書則為全濁，以其全清之聲凝則為全濁也。

"Writing the wholly clear [glyph] side-by-side then makes the wholly muddy [glyph], because the hardening of the wholly clear sound makes the wholly muddy sound."

Any interpretation hinges on what was meant by the character 凝, which is used in a variety of contexts representing a semantic field including *congealing, hardening, freezing,* and so forth. If taken as referring to the hardness or tenseness of muscles necessary to articulate the sounds, it can be interpreted as referring to tense segments; it must be noted that characterizing these segments as *tense* in contrast to *lax* sounds is our imperfect attempt at rendering into language minute differences in the extremely complex process of articulation. Such characterizations run the risk of becoming linguistic just-so stories: *tense segments are called so because the muscles are 'more tense,'* However, it has not been studied whether tense segments are perceived by Korean speakers as being more taxing in terms of the effort needed to produce them, although research on the physiological correlates of vocal effort suggest that muscle tension plays a role (McKenna & Stepp, 2018). The primary cue used by Korean speakers to judge whether a stop sound was tense or lax is has shifted from voice onset time to f_0 in speakers born after 1960 (Cho et al., 2002); Johnson (2011) reminds us that speakers use more ways to vary f_0 than muscle contraction, including raising and lowering of the larynx or changing subglottal pressure. Voiced segments require greater expenditure of effort than voiceless sounds; perhaps this leads back to the debate over what *muddy* signified to the team of linguistic scholars working on a new Korean alphabet in 1446. In current Korean phonology the historical issue is skirted by use of newer terms to describe tense segments: either a direct translation of the Latin *fortis,* 硬音경음, or a native coinage 된소리 */twensori/,* using an adjective */tweda/,* which conveys the thickness of too-dry dough or a tightly-wound rope as a modifier for the native word for sound, */sori/*.

Fortunately for the purpose of understanding the position of tense sounds in the Korean phonological inventory at the time of the new script's promulgation, the seeming gap in the *Hunminjeongeum* orthography as it pertains to tense segments can be almost immediately filled with evidence from the earliest publications using the new alphabet. Tense segments are unequivocally represented using two orthographic conventions. The first method was to use a combination of the glyph used to represent /s/ and that of the corresponding lax stop. As complex onsets were still permitted at the time of the *Hunmingjeongeum,* such provisions for representing them are given. We have evidence (such as the testimony of the shipwrecked Eibokken, mentioned above,) that clusters such as /st/ or /sp/ were utilized even in seventeenth-century Korea. However, a number of occurrences of these clusters in early *hangul* works are not justifiable as being

phonetically faithful for etymological reasons (Kim et al., 2008). A form such as 꾀 */k*we/ *'trick, ruse'* featuring this phantom epenthetic /s/ is found in the 1459 collection of Confucian poems *Yongbieocheonga* (龍飛御天歌용비어천가), but can still be found in use centuries later, in the first Korean-French dictionary[6]. Similarly, '*horn'* is glossed as 쓰리 */p*uri/ in the above-mentioned 1447 Buddhist work *Seokbosangjul* (cf. modern Korean /p*uri/.) In neither case can a hypothetical invasive sibilant be explained. It remains unclear why this convention developed, although a distinct possibility emerges from period morphology. Although no longer used productively, Korean previously used a regular epenthetic morpheme /s/ to indicate possession. We can verify that the morpheme triggered tensification of following stop segments, as a number of compound nouns in contemporary Korean are transparently the result of such epenthesis (orthographic convention having preserved the possessive morpheme,) and all feature tensification where predicted:

냇ㅅ곳	→	냇가	旗ㅅ발	→	깃발	썔애ㅅ돓	→	빨랫돌
/neh/+/s/+/kas/		/ne.k*a/	/ki/+/s/+/pal/		/ki.p*al/	/p*al.e/+/s/+/tolh/		/p*alle.t*ol/
'stream side'			'flag'			'(stone) washboard'		

The association of this epenthetic /s/ with tensification processes may have led writers to use the glyph associated with /s/ in a pseudo-cluster to indicate tensification of an onset segment. As we will see below, this morpheme is associated with the expansion of tense forms in later centuries.

The other spelling convention used was identical to that of the present day: the doubling of the glyph pertaining to the lax segment. There are obvious reasons to explain why such a strategy would initially have been avoided: the explicit designation of an identical glyph in the *Hunminjeongeum* for representing 'muddy' sounds suggests that contemporary speakers were aware of a phonological distinction, meaning that extension of these glyphs to tense segments (assuming that, contra the admitted hypothetical possibility raised in Kim et al., 2008 that 'muddy' referred to tensification, tense segments were not represented with these glyphs) would risk

---

[6] Les Missionnaires de Corée de la Société des Missions Etrangères de Paris. (1880). "한불자뎐(韓佛子典, Dictionnaire Coréen- Français)", Yokohama: C. Lévy.

confusion. Nevertheless, we find that late fifteenth-century writers *did* employ these characters to represent tense segments, using them to represent an emergent tensification phenomena which is now universal in Korean. The language has a suffixing morpheme, /l/, used as a future adnominal (for example, 갈 곳 /kal.k*ot/ "the place where [we] will go.") This morpheme triggers tensification on following stop sounds, and evidence from the fifteenth century suggests that the process was already active at the time. Returning to the *Seokbosangjul*, we find two examples:

수믈 ᄠᅵ 업서　→　숨을 데 없어　　　妙法 니를 쏘리롤　→　妙法 이를 소리를

/sumɯl t*ʌi əpsə/　　/sumɯl t*e əpsə/　　/mjobəp nirɯl s*orirɯl/　　/mjobəp irɯl s*orirɯl/
"no place to hide"　　　　　　　　　　"the voice that will recite the Lotus Sutra"

The historical record thus shows that almost immediately following the development of *hangul*, writers were already using the script in novel ways in an attempt to better represent the sounds of their language. As to why they used existing glyphs rather than create new ones, Kim et al. (2008) proposes that as whatever phonological quality was signified by "muddy" in the *Hunminjeongeum* lost its distinction to Korean speakers, the existing set of characters in the alphabet used to represent such variants were naturally suited to represent a *new* emergent phonemic contrast. By using these glyphs to represent tense segments, the transparent relationship between segments that differed in manner but not in place of articulation could be preserved without creating new characters altogether. In an accident of history, Korean had lost a phonological distinction at just the right time for a new one to emerge, and by fortuitous coincidence the orthography had a logical position for the new entrants.

　　　Hangul thus plays a pivotal role in demonstrating that tense sounds were firmly part of Korean phonology by the mid-fifteenth century. The existence of competing conventions to represent the sounds shows that although not explicitly addressed in the *Hunminjeongeum* writers were attempting to convey information about contrasts they perceived as phonemically distinct. For the first time, the evidence of a tripartite contrast is provided by native speakers attempting to represent the sounds of their language to the best of their ability. If a lack of extant evidence had held tense stops to the shadowy fringes of the historical record, following the advent of *hangul* these sounds indisputably came to claim their share of the spotlight. The wealth of works written

in the new script allows us to consider how these sounds came to encroach upon the domains of their contrasting segments, steadily enlarging the set of words using tense sounds over long centuries. As the present study demonstrates, such expansion shows no signs of stopping.

2.4   Historical expansion of the tense phoneme

Tense sounds in attested documents remain comparatively rare until a period of dramatic increase around the end of the 16th century (Kim et al., 2008:178). The sudden increase of tensification has been attributed to several factors. Tense segments began either as allophonic variants of lax segments, occurring due to predictable phonological environments, or as substitutes for lax segments in optional and idiosyncratic free variation. Tense forms began to expand across four fronts; the following examples are adapted from Kim et al. (2008:178-9):

1. A now-disused morpheme, an epenthetic /s/ used as a possessive marker, triggered tensification on following voiceless consonants and was deleted:
   蓮ㅅ불휘 [jǝn.s.pulhwi] *"lotus root"* → 蓮뿔휘 [jǝn.p*ulhwi]

2. Following /s/ in coda position, voiceless consonants in onset position became tensified (and again, often the triggering segment was deleted:)
   곳고리 [koskori] *"nightingale"* → 고꼬리 [kok*ori]

3. Tense forms were optionally used for emphasis:
   사호다 [sahota] "to fight" → 싸호다 [s*ahota]

4. In a minimal number of exceedingly rare cases, the pronunciation of borrowed Chinese characters became tensified:
   雙 [saŋ] → [s*aŋ]   喫 [kik] → [k*ik]

Perhaps because these segments started out as allophonic variants of lax phonemes, only becoming reanalyzed as phonemically distinct within the last six centuries, the functional load of these segments is quite low in Modern Korean. This is confirmed by a frequency analysis of Korean phonemes calculated from spoken corpus data provided by Shin (2011:189):

Phoneme	Frequency (%)	Phoneme	Frequency (%)	Phoneme	Frequency (%)
/k/	7.68	/kʰ/	0.81	/k*/	0.96
/t/	2.98	/tʰ/	0.90	/t*/	0.63
/p/	3.50	/pʰ/	1.13	/p*/	0.42

2.5   Tensification phenomena in contemporary Korean

Given philological evidence of lax segments shifting to tense pronunciations in earlier periods, it is perhaps less surprising that a variety of tensification phenomena (or phenomena wherein tense forms are preferred over their lax counterparts) have been observed across *modern* Korean. Use of novel tensified-onset pronunciations has been observed in all South Korean dialects in empirical studies (Jang, 2017). Shin & Davis (2008) provide an overview of several different tensification phenomena taking place in the contemporary era, to which I have appended better examples when possible:

1) Optional tensification of lax stops in word-initial position of some words in spoken Korean:
닦다 /takt*a/ "to polish" → 딲다 /t*akt*a/ (Shin & Davis, 2008:72)

2) Relexicalization of lax stops as tense in regional dialects:
두꺼비 /tuk*əbi/ "toad" → 뚜꺼비 /t*uk*əbi/ (Gyeongsang-do dialect) (Shin & Davis 2008:73)

3) Avoidance of lax stops in foreign borrowings in favor of tense forms:
뚜레쥬르 /t*uredʑjurɯ/ "Tous Les Jours" (a ubiquitous South Korean bakery franchise, with deliberately French-styled branding)

4) Early acquisition and production of tense stops compared to lax stops by children acquiring Korean natively: my own daughter is bilingual, speaking both Korean and English, and many of her idiosyncratic Korean baby talk items featured tense stops, such as /k*ak*a/ "bird" and /t*abit*abi/ "walking"

Jeon (2016) observed unexpected use of tense onsets in loanwords by broadcast journalists, a phenomenon to which Kang (2008) had earlier attempted to provide an acoustic motivation: speakers may be trying to preserve the voiced onsets of the original foreign word, substituting an

acoustically similar if not alike segment from the Korean phonological toolbox. Voiced onsets are not permitted in Korean phonology, although voicing occurs intervocalically (Bae, 2003; Shin, 2011). However, this is contradicted by Shin and Davis (2008), who posit that English voiced stops in onset position *should* be substituted by Korean lax stops, owing to their acoustic similarity in terms of voice onset time- yet there are a "rather large number" of loanwords featuring tense onsets where lax onsets would be expected, such as 빠 [b*a] for the English *bar* (p. 75). A simple explanation along acoustic lines is further complicated by the findings of a further investigation by Jeon (2019), wherein use of tense onsets in some loanwords was shown to be markedly *increasing* from one generation to the next. In absence of an acoustic motivation capable of explaining all these examples, the question remains: why should lax segments sporadically become tense at all?

2.6 Tensification and the possible restructuring of the Korean phoneme inventory

This question brings us to the crux of the matter. The weight of evidence from extant historical record and recent empirical research demonstrates that onset tensification is a robust phenomenon in Korean phonology, both from a diachronic perspective and in investigating current articulatory behaviors of Korean speakers. But having shown conclusively that such changes are occurring, it is not immediately clear what motivates such shifts. Appeals to acoustics based on neighboring segments proves inadequate. The application of tensification processes seems finicky, with tensification picking and choosing randomly among the lexicon as to which items will fall under its influence. Most perplexingly, its efforts are never reciprocated by other phonemes: onset tense segments never become lax. Yet from this an appeal to deeper principles of language may be possible, allowing a furtive guess as to invisible but mighty pressures which shape the entirety of phonology. John Whitman, as cited in Shin and Davis (2008), offers one potential explanation for onset tensification phenomena:

> In Korean, the tense stops had a low functional load, mainly developing historically out of syllable-initial consonant clusters. The current expansion of tense stops as documented [here] can be viewed as a means by which the language evens out of the functional load of the different stops. (p. 78)

Whether tense stops developed chiefly out of consonant clusters is contested by other scholars (for detailed discussions on this, see Kim et al., 2008:158-61, Lee & Ramsey, 2011:131-

35). Notwithstanding, Whitman makes an intriguing point, especially given the low frequency of tense segments as observed above. The spread of functional load from lax onto tense segments could be a way that Korean is economizing; by shifting the pronunciation of some lexical items from lax to tense onsets, the language is getting more use out of those segments, justifying the added expenditure of a larger phonemic inventory and the cost to listeners who must distinguish between the phonemes. It could also reduce homophony and ambiguity, further aiding both speakers and listeners in preserving the intended message.

Appeals to such economizing principles are well-established in corpus linguistics, most famously enshrined by Zipf's law, which has been robustly applied to show how the rank and frequency of words in natural language corpora are inversely proportional:

$$f(r) \propto \frac{1}{r^{\alpha}}$$

That is, the frequency of a word is proportional to its rank, r, where α is close to 1 (Zipf, 1949). Thus, the word ranked second will appear roughly half as often as the most frequent word, the third around a third as often, and so forth. Zipf motivates this through an appeal to a constant tension between the economies of the speaker and listener, best expressed in his seminal *Human Behavior and the Principle of Least Effort* (1949):

> The speaker's economy will tend to reduce the size of the vocabulary to a single word by unifying all meanings behind a single word … the auditor's economy will tend to increase the size of a vocabulary to a point where there will be a distinctly different word for each different meaning. … From this it follows that whenever a person uses words to convey meanings he will automatically try to get his ideas across most efficiently by seeking a balance. (p. 21-22)

While Zipf was writing of lexical items, the same principles could be applied to phonemic inventories of natural languages. There is an unending tension between complexity and economy in phonology: adding to the phonological inventory has benefits but comes at the cost of increased effort on the part of speaker and listener alike. If maintaining a tripartite distinction in the stop series of a language is costly to its users, who must shoulder the added burden of producing and distinguishing them, perhaps there is a kind of evolutionary pressure to *use* the segments.

Perhaps surprisingly given the wide application of power laws across the sciences, including exhaustive use in linguistic studies of word frequencies, considerably fewer researchers have applied power laws such as Zipf's law to observed phoneme frequencies in world languages. Scholars that have made tentative explorations in this direction have pursued different approaches

with regards to the power laws applied, citing key differences between small phoneme inventories and the inventories of natural language phenomena where Zipf's law has proved robust, such as in studies of large-scale corpora where tens of thousands of individual words are represented. Tambovtsev & Martindale (2007) compare the fit of the Zipf equation and several alternatives to phoneme frequencies from 95 world languages across varied language families. Although their conclusion is that another equation (the Yule equation) proves a better fit, their figures for the fit of the Zipf equation across a diverse range of languages allows a rudimentary comparison with the Korean phoneme inventory. While Tambovtsev & Martindale included a wide variety of world languages in their study, Korean was not represented. Similarly, Macklin-Cordes and Round (2020) examined the phonemic inventories of 166 Australasian languages and concluded that they displayed a Zipfian-like distribution among the most frequent phonemes, but an exponential distribution among less-frequent outliers. Leaving more complex equations to future consideration, we can perform a simple fit of Korean data to see how well it matches a Zipfian distribution. This will clarify whether the frequency of Korean phonemes differs markedly from other world languages where sufficient data on phoneme frequency has been provided by previous research.

While numerous frequency studies based on dictionary entries have been conducted, studies featuring analysis of actual speaker productions remain scarce, and a number of scholars have attempted to fill this gap in recent years. At the time of writing, the largest study of Korean phoneme frequency in natural speech was conducted by Shin (2008), who collected just over twenty-three hours of free conversation produced by fifty-seven native speakers aged 19-32, all of whom considered themselves speaking a variety close to standard Korean. A total of 858,512 phonemes were isolated in the following analysis. The resulting phoneme frequencies are given on the following page (Shin 2008, 2011). Using the data, the rudimentary fit of Korean phoneme frequencies to the Zipf equation can be observed, plotting the rank and frequency of the phonemes on a log-log graph:
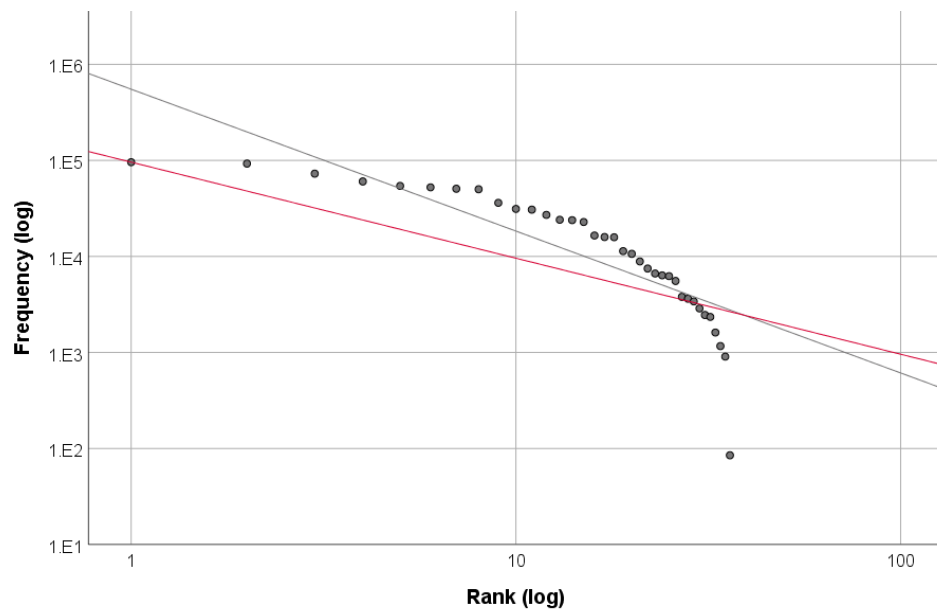
Figure 1. Fit of Zipf equation (red) to Korean phoneme frequency data as provided by Shin (2008, 2011).

Table 1. Frequencies of Korean phonemes derived from a spoken corpus study by Shin (2008, 2011).

Rank	Phoneme	Frequency	Frequency (%)		Rank	Phoneme	Frequency	Frequency (%)
1	ɑ	95853	11.17		19	k*	11355	1.32
2	n	92621	10.79		20	s*	10638	1.24
3	k	72756	8.47		21	t*	8873	1.03
4	i	60318	7.03		22	tɕ*	7480	0.87
5	l	54253	6.32		23	jɑ	6640	0.77
6	ʌ	52414	6.11		24	tɕʰ	6349	0.74
7	ɯ	50704	5.91		25	tʰ	6229	0.73
8	ɛ	50024	5.83		26	kʰ	5553	0.65
9	m	36132	4.21		27	wɛ	3792	0.44
10	o	31268	3.64		28	wɑ	3628	0.42
11	t	30704	3.58		29	pʰ	3399	0.40
12	tɕ	27076	3.15		30	jo	2874	0.33
13	u	24061	2.80		31	p*	2457	0.29
14	ŋ	23874	2.78		32	wʌ	2345	0.27
15	s	22827	2.66		33	jɛ	1614	0.19
16	h	16494	1.92		34	ju	1165	0.14
17	jʌ	15915	1.85		35	wi	905	0.11
18	p	15837	1.84		36	ɯi[7]	85	0.01

---

[7] The case of the diphthong /ɯi/, an extreme outlier by any measure, demonstrates how Korean phonology is in flux beyond the changes the present study is concerned with. This phoneme is now pronounced by speakers almost solely in word-initial position; in words historically featuring it in other positions it is typically pronounced /i/, or in the case of the possessive morpheme, now pronounced /ɛ/. That a phoneme associated with a morpheme performing such an essential role as grammatical marking of possession should fall so dramatically out of fashion presents intriguing questions in itself.

The fit of the equation as described by R^2 for the given Korean phoneme frequencies is .801, considerably lower than the mean of .90 for 95 languages found by Tambovtsev and Martindale (2007). One possible interpretation is that Zipf's equation (and perhaps other power laws) cannot be applied so easily to phoneme inventories of natural language, or perhaps that good fit in these instances is a statistical artifact unrelated to researchers' assumptions about its relationship to language (a possibility discussed in detail in Piantadosi, 2014.) Alternatively, another possibility is that if phoneme inventories roughly follow a Zipfian distribution, the notably lower fit found in the case of Korean is evidence that phonemes of contemporary Korean exhibit an atypical distribution compared to other languages. The skewed frequencies of Korean phonemes could support Whitman's guess as to the underlying motivations of tensification: each lax segment shifted towards a tense articulation further "balances" the distribution towards one perhaps more typical of (or at least, more well-observed in) other world languages. If true, it is perhaps expected that the encroachment of one phoneme on the domain of another would be unreciprocated elsewhere in the language's phonology; this is borne out by the evidence from Korean.

Shin & Davis (2008) note the unidirectionality of these changes, wherein only lax stops seem to fall victim to tensification. Considering the markedness of the tripartite contrast in the first place, they speculate that one possibility is that modern Korean is in the process of restructuring the three-way stop system into a two-way system, preserving only the distinction between tense and aspirated stops. However, they conclude that "the expansion is probably unlikely to lead to the restructuring of the Korean stop system," due to its restriction to word-initial position (Shin & Davis, 2008:79). It is also notable that cases of lax stops shifting to *aspirated* pronunciations diachronically are limited to a small number of lexical items, each of them featuring voiceless glottal fricatives in coda position which were lost in the process (Lee & Ramsey, 2011). However, some limited empirical evidence has been observed in favor of restructuring arguments. The weakening of tense stops into lax ones in medial positions is attested by Han & Park (2012), who demonstrated that at morpheme boundaries in compound nouns, speakers may overwhelmingly prefer pronunciations featuring lax stops where tense are predicted under rules of Korean morphophonology.

2.7 Contemporary tensification phenomena as the continuation of historical processes

While a complete restructuring of Korean stops is perhaps too extreme, at the lexical level the results of tensifying pressures is readily apparent given insights from Korean historical linguistics. A look at the history of individual lexical items in Middle and Early Modern Korean and a comparison with present forms confirms that tensification has been active for centuries. What presumably began as optional, idiosyncratic tensification eventually consolidated, with the result being that use of the lax segment in word-initial position is now *erroneous*. The following is adapted from Bae (2003) and Han (2013:)

Archaic Form	Modern Korean	English	Period of Change
구짖다 [kudʒitta]	꾸짖다 [k*udʒitta]	'to scold'	16th century
곡지 [koktʃi]	꼭지 [k*oktʃi]	'knob'	17th century
감작이다 [kamdʒagida]	깜작이다 [k*amdʒagida]	'to blink'	18th century
가마괴 [kamagwe]	까마귀 [k*amagwi]	'crow'	19th century

Not only has word-initial tensification been active in reshaping the Korean lexicon for centuries, it appears to apply capriciously but with great impact. Once a word falls into the onset-tensification trap, so to speak, it cannot go back. Word-onset weakening of tense stops into lax is not attested, although Han & Park (2012) suggests that this may not be true of tense stops in *medial* positions. The results of that study showed for some lexical items, speakers were sometimes overwhelmingly producing lax stops at morpheme boundaries within compound words where rules of Korean phonology should produce a tense stop. While this behavior is still restricted to a fairly particular and specific morphophonological environment, it does offer a tantalizing counterpart to word-onset tensification, suggesting perhaps pressures to simply the tripartite contrast back into two phonemes, with one bearing two positional allophones. These findings will be investigated further in the second experiment of the present project.

Interestingly, although onset tensification is well-established as a historical phenomenon, evidence from present-day native speaker productions would seem to suggest it is far from being

an archaic process relegated to works of Korean philology. A cursory foray onto the Korean internet, via Google searches of known tensification-prone words, is enough to confirm that native speakers use tense segments where lax ones are expected (and furthermore, that aspirated segments are never substituted in the same contexts:)

Original Korean Text	Phonetic Transcription	English
눈물 **딲**아 주세요	[nunmul **d***ak*a tʃusejo]	"Please **w**ipe away my tears"
머리 **깜**고 싶어	[məri **k***amk*o ʃipʰə]	"I want to **w**ash my hair"
뻔데기 칼로리 알아보기	[**p***əndegi kʰallori arabogi]	"Find out how many calories are in ***b**eondaegi* [a popular snack made from steamed silkworm cocoons]"
지훈이 **꿘**투한다!	[tʃihuni **k***wəntʰuhanda]	"Jihoon is **b**oxing!"

In light of this, it seems reasonable to consider the possibility that word-initial tensification as demonstrated by Korean netizens is a glimpse into the future of the Korean language. However, tensified forms such as those in the above example are considered nonstandard or incorrect by some speakers, even when confronted with their *own* tensified pronunciations (Jeon 2019).

2.8 Explaining tensification: different perspectives

Earlier treatments of word-initial tensification considered the phenomenon from extralinguistic perspectives: perhaps tensification is added for emphatic purpose by speakers, or as the product of idiosyncratic speech error, or used purposefully as an expression of regional or social membership. Kim (2001) attributes tensification to emphasis in relaxed discourse situations. Lee & Ramsey (2011) note that many words which have undergone tensification are semantically open to emphasis: *rub, beat, pound, scatter, chew, chop,* and so forth. The words all refer to actions that can be performed with varying degrees of intensity, perhaps meaning speakers are varying their pronunciations to imbue a specific intensity to the action. The authors make the claim that use of tensified forms signals a *qualitatively* different item: "a [/ko*tɕʰu/] is a spicier 'chili pepper' than a [/kotɕʰu/]" (Lee & Ramsey, 2011:133). But even taking this to be true, whither the crow,

the silkworm cocoon, levees, and dried persimmons? An explanation based on emphatic use alone would need to accommodate such cases, while also explaining why tensification is not *universally* exhibited, however occasionally, where the lexical item is as plausibly open to emphasis as *beat* or *rub*. Park (2000:191) points out that for two homophonous verbs /ka:mt'a/ *to wash (hair)* and *to wind (around)*, tensification is *only* permitted on the former. If tense onsets are the product of reinforced emphatic use over long periods, is it plausible that such restrictions are so strong that native-speaker researchers deem one permissible but outright reject the other? Perhaps Korean speakers have yet to be as emphatic about winding as hair-washing, while being emphatic about no fruits but persimmons; this explanation does not strike me as parsimonious.

Other investigators have explored sociolinguistic factors related to use of tensified forms and other novel phonological developments. Lee (1996) found that speakers in a survey associated onset tensification with masculinity and vulgarity, and that younger, male, and working-class speakers were more likely to use these forms in daily life. The class aspect of tensification use was explored by a different Lee (1999), but he found that contrary to expectations, no significant disparity was observed in use of tense onset variants between members of different social classes. Hong (2014) notes that the most marked use of tensified onsets is by speakers in their teens and early twenties. An analysis of use of tense forms as an expression of youth would be premature; Jeon (2019) explored tensification by a wider range of speakers and demonstrated that tensification increased progressively between participants aged sixty and over, those between their thirties and sixties, those in their twenties, and teenage participants. If tensification was originally spurred by social or class factors, it would seem to be spreading beyond these bounds with the passage of time.

Beyond the realm of the speaker-specific and personal, some have tried to go so far as to ground explanations of tensified onsets in extra-linguistic sociopolitical phenomena. Park (2000:180-1) offers several examples: one commenter suggesting tensification as a reflection of psychological pressure induced by the rapid growth of material civilization and the corresponding increase of desires; another positing that pronunciation grows tense as society becomes more complicated and competition grows fierce. It comes as no surprise, perhaps, that both comments were penned in the 1980s, a time of political upheaval in South Korea. These suggested explanations for word-onset tensification tell us more about the climate on university campuses in the 1980s than they do about the Korean language. Lee (1999) suggests that tensification reflects

downward trends in education owing to the increasingly hopeless economic climate of the time: the Asian financial crisis of 1997 heavily affected Korea, and questions about economic hardships posed by the crisis were used as one measure to gauge the class of respondents. However, given that we have established how tensification began acting before the modern era, often centuries before, can we reasonably attribute this to the pressures of modern industrial society, or periods of economic downturn? Even if such sociopolitical climates accelerate linguistic change, such analyses are inadequate on their own, as the philological evidence shows.

Explanations of tensification as a purely dialectal phenomenon, separate from other factors, also prove explanatorily inadequate. Any relationship between tensification and regional dialects is confused in general literature on Korean varieties. Tensification has been associated with Chungcheong-do and Jeolla-do dialects by some researchers (Sohn, 1999:70, 74). Wee (2008) observes that Jeolla-do sees a "high frequency" of tensification, while other researchers assert that North Korean dialects tensify more than Southern varieties (Yeon, 2012:15). Some scholars associate onset tensification with Seoul dialects (Lee, 2002; You, 2006). One researcher's in-depth empirical study showed the *most* tensification in the Gyeongsang-do dialect region (Jang, 2017, in line with the observations made on optional tensification in this dialect area in Shin & Davis, 2008), although she noted tensification of different lexical items in differing dialectal regions. Gyeongsang-do as the source of novel tense forms is advanced by Oh (2011), who suggests that the change promulgates outward from Gyeongsang-do to Jeolla-do, and thence onward. An attempt to clarify this on a map for readers less familiar with Korean geography would be less helpful than pointing out here that over this short paragraph *all* mainland Korean dialects have been advanced as hotspots of tensification phenomena by one source or another.

Tensification would seem to be happening everywhere and anywhere, all at once, depending on who is asked. A summary of evidence from sociolinguistic research would suggest, perhaps, that such a view is correct: tensification is happening concurrently across South Korea[8] (Lee, 1999; Jang, 2017). However, that still leaves open the question of whether such novel

---

[8] I feel compelled here to note that owing to the unique political situation, Northern dialects of Korean are underrepresented in research- but also to remind the reader that as national borders (and particularly the borders established by the 1953 armistice ending the Korean War) can be quite arbitrary, they may not align neatly with the boundaries of dialect areas. The Gangwon-do dialectal region is bisected by the demilitarized zone, and the dialect of Korean spoken in the far northeast of the People's Republic of China, in the environs of the North Korean border, is a Hamgyeong-do dialect. Thus, for at least some Northern dialects, a reasonable attempt can be made at research outside the borders of the Democratic People's Republic of Korea. (See Oh & Yang, 2013, for an example using Yanbian Korean.)

x

pronunciations arise independently across the peninsula, or whether perhaps the diffusion from one variety to the next can be modelled.

An in-depth study on the geographical dispersion of tensified forms by Jang (2017) suggests that the Gyeongsang-do dialect is particularly innovative, perhaps even permitting a hypothesis that these forms diffuse from Gyeongsang-do into other regional dialects. However, even if we assume a model of diffusion wherein novel tense forms originate in Gyeongsang-do and spread throughout the peninsula, relegating words with tensified onsets as nonstandard, dialectal variants seems at odds with observed native speaker production in experimental settings. Native speakers of different dialectal backgrounds may produce or not produce tensified forms; speakers of the standard Seoul dialect may produce tense forms while speakers of the Gyeongsang dialect *do not,* and different lexical items show different patterns of dispersion geographically (Park, 2000; Han, 2011, 2013; Jang, 2017). It has been observed Seoul speakers are becoming *more like* speakers of different dialects in terms of *acoustic* cues they use to discern between tense, lax and aspirate segments. Silva (2006) shows that younger Korean speakers are shifting from using voice onset time to F_0 as the primary cue used to distinguish between these segments, and that in this case the *standard* is changing under outside influence. The picture of how one dialect influences (and even the question of how dominant the prescriptive standard is) becomes murkier when we consider how Korean speakers of different regions are becoming more like one another, with accommodations being made in both directions.

Perhaps most challenging to any assumption that tensification is purely dialectal comes the evidence from Han (2013:409-414) with an experimental study on native-speaker raters' judgments of the acceptability of tensified forms. This large-scale study, involving 355 native-speaker judges, showed that in many cases the form with the tense segment was *overwhelmingly preferred* to that featuring the lax counterpart. It becomes difficult to motivate the case that a given item is a dialectal variant when most speakers, many of whom use a standard dialect, prefer it to the prescriptive norm. Even if research suggests that some dialects prove more innovative, the suggestion remains that speakers of all dialect backgrounds, including those closest to the standard, are open to adopting novel tensified forms. Furthermore, the consistency demonstrated across large populations of speakers (as evidenced by widespread acceptability) suggests that attributing tensified productions uniformly to speaker error is misguided (and if most speakers produce and prefer a given pronunciation, can it be characterized as *wrong?*)

Over the past few decades, new approaches to word-onset tensification have sought alternative explanations. Park (2000) and Kim (2004) examined tensification from a language-internal standpoint, searching for phonological explanations. Pointing out that educated speakers of Seoul dialect consistently produce some words with tensified onsets, while other words with lax onsets are *not* subject to tensification by these speakers, Park suggests that tensification is governed (or at least, influenced) by regular principles (Park, 2000:181):

홍부는 머리를 감았다 / 깜았다.
[hoŋbunɯn mərirɯl kamatta] / [k*amatta]
"Hongbu washed his hair."

놀부는 연필을 잡았다. / *짭았다.
[nolbunɯn jənpʰirɯl tʃabatta] / *[tʃ*abatta]
"Nolbu grasped the pencil."

In this example, tensifying the onset of *wash* is acceptable to his Seoul dialect speakers, but not the onset of *grasp*. He demonstrates that any simple explanation of the phenomenon as motivated by phonological environments or lexical category is difficult (Park, 2000:191). However, given the speakers' uniform distaste for some tensified forms and amenability to others, this behavior must ultimately find explanation somewhere.

To try and elucidate the matter, Park sifted through the *Yonsei Dictionary of Korean,* a dictionary of 22,166 high-frequency Korean words (Park, 2000:185). He collected 812 lexical items he believed were open to tensification, then conducted a survey task among five native speakers of Seoul dialect (ranging in age from their teens to their forties) as to whether the words were realized with tensified onsets. His participants identified 233 of these candidates as tensified, with items not demonstrating a marked skew towards specific lexical classes (100 nouns, 69 verbs, 47 adjectives, and 13 adverbs), allowing this to be discounted as the cause of tensification. However, Park noted that a clear asymmetry emerges when the *etymological* origin of the word is considered. The Korean lexicon includes not only 'native' words with no discernible non-Korean origin, but also many words borrowed from classical Chinese (often referred to as Sino-Korean words), as well as modern loanwords from Western and other world languages. Although Sino-

Korean forms comprised 88.6% of the Yonsei dictionary (owing to the heavy use of Sino-Korean words for technical subjects, even if these words are less common in usage), only twenty-six Sino-Korean words were found among the 233 tensified items. This is confirmed by Han (2011), who notes that 'native' words are far more susceptible to tensification than Sino-Korean lexical items. Native speakers are, of course, *implicitly* aware of the different 'categories;' English speakers know when using the morpheme *-ic* to form an adjective from a noun is permissible, even if they are blissfully unaware that they are pulling from established categories of English words of Greek and Latin origin.

Park also notes that among the tensified items identified by his Seoul speakers, one finds many words definitively tensified in regional dialects, a finding confirmed by the results of Jang (2017). He also observed that the degree of tensification differs among the items identified, suggesting that while for some of the words, onset tensification is optional, for others tensification has become almost obligatory. This aligns with the explanation of the phenomenon as a manifestation of longer diachronic sound changes, which are applying to individual lexical items at different times.

The largest investigation into the phenomenon from an experimental standpoint has been conducted by Han Myeong-suk over the span of several experimental studies, ranging from analyses of natural speech, orthographic investigations into whether participants use graphemes representing tensified forms in response to dictation tasks, and attitudinal studies on listeners' acceptability ratings of forms featuring tensified onsets (Han, 2011, 2013).

Han (2011) considers word-initial tensification in natural language settings as produced by college students. In this study, five-minute class presentations by speakers of both genders were recorded and then analyzed for the presence of tensified segments. A second experiment tested whether words already identified as undergoing widespread tensification would be orthographically realized as such by participants performing a dictation test. The results showed that 40.6% of the participants (119 of 293 speakers) produced at least one tensified form; the author notes that although the presentation task did elicit natural language, the setting may have caused speakers to use more careful speech, echoing an observation by Kim (2004) that more formal settings may suppress the appearance of word-initial tensification.

Addressing the issue of individual words demonstrating disparate degrees of tensification, Han (2013) attempts to create an index to categorize words according to the level of tensification

observed. Seven 'levels' are identified, ranging from words with a complete prohibition on tensification to words where tensification has become *universal* (for instance, those offered earlier as examples of the historical tensification process, where tensified forms had supplanted lax onset forms already by the fifteenth or sixteenth centuries).

Her results, based on several factors (phonological evidence of tensification from real productions, substitution of graphemes representing tense segments in writing tasks, and attitudinal studies of native-speaker listeners presented with tensified forms) are summarized as follows, adapted from Han (2013):

Table 2. Index of tensification as described by Han (2013).

Tensification Index	Level of Realization of Tensification	Examples
6	Tensification has completed diachronically, and the segment is realized universally as tense in modern Korean.	/p*uri/ 'root' /k*ot/ 'flower' /k*amagwi/ 'crow'
5	Tensification is produced by over 50% of speakers and represented orthographically by over 50% of speakers when writing.	/tʃ*uk*umi/ 'webfoot octopus' /s*uŋmek/ 'fool' /k*omdʒaŋə/ 'sea eel'
4	Tensification is produced by over 50% of speakers, and the lax form is considered awkard or incorrect by over 50% of raters.	/s*eda/ 'to be strong' /tʃ*oktʃipke/ 'tweezers' /s*odʒu/ 'soju (a type of distilled spirits)'
3	Both lax and tensified forms are considered equally acceptable by raters.	/k*aʃi/ 'thorn' /t*akta/ 'to wipe' /tʃ*arɯda/ 'to cut'
2	Lax forms are overwhelmingly preferred by raters.	/k*wəntu/ 'boxing' /tʃ*uŋguk/ 'China'
1	Tensification is markedly idiosyncratic, or otherwise attributable to speaker error.	/k*jesok/ 'continuing' /t*adɯmda/ 'to trim'
0	Tensification never occurs.	*/k*abaŋ/ 'bag' */t*ari/ 'leg'

Han's efforts to clarify tensification at the level of the individual lexeme provide a critical insight into the phenomenon. She points out that previous inquiries into word-onset tensification have investigated the matter from sociolinguistic or morphosemantic standpoints, theorizing that differences in regional background or gender of the speaker, or larger morphological or semantic principles, have been responsible for the realization of previously lax onsets as tense. Through her investigation, we find that each of these proffered explanations cannot accommodate observed linguistic behavior.

Furthermore, soliciting acceptability judgments from native-speaker raters in combination with eliciting productions in natural speech tasks has permitted a finer distinction between items where tensification is optional and where it remains marginal. A clearer pattern emerges where we can possibly see which lexical items are being 'drawn into the black hole of tensification,' so to speak. It seems plausible that in future centuries, *beondaegi* or *komjangeo* will have joined *k*ot* and *k*amagwi* as words realized universally with tensified onsets, despite historically starting out with lax segments. Certainly, such a prediction seems validated by what has been observed experimentally.

The suggestion of a link between speaker attitudes and prevalence of tense forms over lax (as suggested by Han, 2013) seems further borne out by work by Hong (2014) in a study on word-onset tensification within a specific geographical dialect (more specifically, speakers from Daegu, a large city in the Gyeongsang-do region) and speaker perceptions of such variants as positive or negative. Her results suggested that if a form is demonstrated to have greater positive judgment by native speakers, the form can be expected to be realized at greater rates in speech. This study contains some caveats in terms of how widely we may apply it to Korean as a whole; Jang (2017) identified innovative use of tensification as characteristic of this dialect, opening the possibility that speakers are aware of this and consider such pronunciations to be an expression of regional pride or a means of asserting identity. These remarks are echoed by participants in a survey by Jeon (2019) who offer that use of tensified variants among friends establishes a feeling of closeness in contrast to the distance expressed by use of standard. These studies demonstrate how a novel form may reinforce itself (and ultimately, take prevalence) through a kind of feedback loop. Whether tense forms are judged more positively by speakers *or* simply accepted as valid alternative articulations (as per Han, 2013), the result could be increased usage- again, reinforcing these pronunciations until they become dominant. This could also explain why the sound changes

appear to apply capriciously but *precipitously* over longer periods of time: once a lexical item has begun the process, such change could come about rapidly.

Explanations of word-onset tensification as the manifestation of longer diachronic sound changes still active in Korean are bolstered by evidence provided by cross-generational sociolinguistic studies. Kang (2008) offers that "sociolinguistic studies show that tensification is more frequently found in the speech of younger generational speakers," a claim echoed by Lee (1999) and observed empirically by Jeon (2019). If tensification were idiosyncratic and simply the result of speaker preferences or quirks, it would not be observed to display such generational tendencies. To the contrary, such increases are *predicted* under a hypothesis that tensified forms represent ongoing diachronic sound change.

Shifting pronunciations seem all the more plausible, given that modern Korean is in a state of flux and generational phonological change has been confirmed in other domains: over numerous studies, it has been shown that speakers born after the 1960s use different acoustic cues to categorize between the tripartite stop series (Cho et al., 2002; Silva, 2006). Within the span of a single generation, the primary acoustic means speakers are using- both in production and perception- to distinguish between these segments has shifted from voice onset time to F_0. Kang (2008) reminds us that multiple cues are utilized in the tense-lax distinction, citing breathiness and the quality of the vowel onset in addition to F_0, echoing the seminal work by Cho et al. (2002) which established the acoustic cues used by Korean speakers to discern these phonemes. The generational shift from using one primary acoustic cue to another serves as an interesting parallel to that of generational changes in pronouncing individual words with tense or lax onsets.

2.9   Summary

This chapter provided a review of pertinent literature on onset tensification, both from a historical perspective and as an issue of immediate study. Beginning with an overview of the history of the Korean language and the origins of the current stop system, it proceeded through a summary of historical evidence of tensification phenomena and the expansion of the tense phoneme into previously lax domains. The possibility that Korean is undergoing a rebalancing of its phonemic inventory due to deeper selection pressures was introduced. The continuing action of onset tensification in contemporary Korean was described, along with the perspectives of previous

researchers as to what conditions and motivates such novel phonological behavior. The following chapter will describe the methodology followed in the experiments featured in the present study.

CHAPTER 3.   METHODS

3.1   Experiment 1: Production of Tensified Onsets

3.1.1   Stimuli

To test whether tensified onsets would be produced by native Korean speakers, a task was designed where participants would be shown sixty-three pairings of images with gapped sentences aiming to elicit a target word, given the context of the image. For example, to elicit the target word /toŋgurami/ "circle," the stimulus provided was an image of a calendar with a circle drawn around a date, alongside the sentence *My sibling drew ____ around her birthday on the calendar.* An example of what participants would see during the task is provided below:
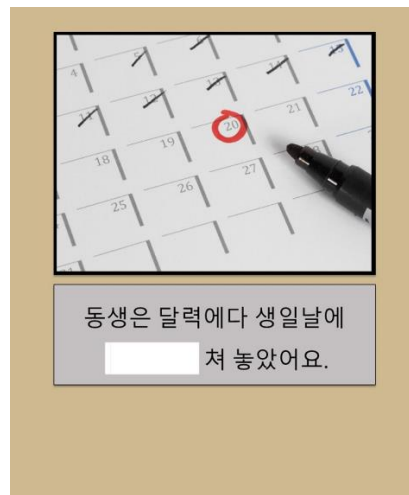
Figure 2. Sample stimulus slide from the from the production task.

Of the sixty-three stimuli presented, twenty-one were designed to elicit words where onset tensification is attested by previous researchers (Bae 2003, Shin & Davis 2008, and Han 2013). Each of the lax stop sounds subject to tensification was represented seven times. Words were selected with an eye to balancing the parts of speech featured between target and control items as well as the following vowel after the onset sound. Twenty-one stimuli featured words with lax or tense onsets where it was assumed tensification (or the unattested alternative, weakening of a tense stop onset into its lax counterpart) is impermissible, to serve as a control. Defining any word as

being closed to tensification is perhaps contentious, given that if lax onsets tensifying is an ongoing phonological change in Korean, it could theoretically happen *anywhere*. However, as demonstrated in Park (2000), native speakers will strongly reject tensification on some items. For this project, high-frequency words that were mentioned in previous work as closed to or otherwise dis-preferred as tensified were chosen (Oh, 2011; Han, 2013; Jang, 2017). Another twenty-one items featured non-stop onsets, intended as distractors. A list of the stimuli items is provided in Table 23 in Appendix B.

3.1.2 Participants

Seventy-seven participants were recruited for the project, with fifty-four female and twenty-three male speakers represented. Participant ages ranged from 24 to 58, with a mean age of 34.2 years. All were self-identified native speakers of Korean with no reported speech or hearing disorders. Participants were recruited over the last week of June 2020 through word of mouth and social media, particularly on *Momsholic*[9], a Korean 'internet café' or discussion forum for mothers and expecting mothers. The forum is divided into subforums for specific countries; recruitment took place on the American subforum, although participation was open to both American and South Korean users. Accordingly, two methods of providing compensation were utilized: participants were offered the choice of either a $10 Amazon gift card or a ₩10,000 gift card for CU (a ubiquitous South Korean convenience store franchise.) Response to recruitment efforts was enthusiastic, and all participants completed the experimental task within one week. Of the seventy-six participants in the present study, sixteen were resident in South Korea at the time of recording, with the remaining sixty-one scattered across the United States.

---

[9] https://cafe.naver.com/imsanbu

Table 3. Descriptive statistics of participants in the production experiment.

	N	Age (Range)	Age (Mean)
Male	23	28-41	33.7
Female	54	24-58	34.5
Total	77	24-58	34.2
Residency			
United States	61		
South Korea	16		

A further four native speakers were recruited to serve as raters. The native speaker raters were a thirty-five-year-old female college graduate with several years' experience teaching preschool and kindergarten, resident in the United States, a thirty-six-year-old female homemaker with a four-year university degree, resident in South Korea, and two twenty-six-year-old tradesmen with two-year technical degrees, resident in South Korea. All four are of Gyeongsang-do area origin, although each rater did not self-identify as using a regional dialect often.

3.1.3   Procedure

Participants performed the tasks at their own convenience using their own personal computers. Upon completing a background questionnaire formatted as a Qualtrics survey, each participant was provided a link to the Speak Everywhere portal and unique login credentials. The questionnaire featured items on participant demographics, foreign and native language background, use of regional dialect, attitudes towards language, and a short personality test. The personality inventory, based off of the NEO Personality Inventory (Costa & McCrae, 1985), used questions from a Korean adaption of the NEO-PI by Ahn and Chang (1996). A translation of the questionnaire is presented in Appendix A.

For the recording task, stimuli were presented using Speak Everywhere, an online oral practice/instruction/assessment platform originally developed for foreign language teaching by the

Center for Technology-Enhanced Learning at Purdue University[10]. Although not specifically designed for the purpose of collecting data for phonological studies, the platform was selected as it allowed data collection from participants during the unique circumstances following the outbreak of COVID-19 and ensuing restrictions on human subject research. Through Speak Everywhere, participants were able to record their own productions at their leisure, while the researcher was able to access them remotely without either party meeting.

After logging into Speak Everywhere and entering the task page, participants were presented with brief instructions on how to use the platform and a walkthrough on recording and playback. Participants were told to read each sentence presented aloud, filling the gap as they wished. Including training, it typically took participants between twenty to thirty minutes to complete the recording task, spending between ten and fifteen seconds on each stimulus. Combined with the preceding questionnaire, participation typically took forty minutes.

3.1.4 Measurements

The resulting audio was reviewed by the researcher to verify whether target words were produced; in each successful instance, the syllable subject to tensification was segmented in Audacity. A small group of control syllables from items where tensification is believed by the researcher to be inactive or universal was also segmented to test the reliability of the rating task. These audio clips were then presented to native Korean-speaking raters in a series of forced-choice identification task through Qualtrics. Each token was presented to three raters; two of the four raters performed all of the tasks, while the other two handled half of the tokens each. Raters were presented each item once in random order, having the options of identifying what they heard as having a lax onset, a tense onset, or that the sample is too inaudible for identification. This last option was offered owing to the inconsistent and sometimes low quality of audio provided by participants, many of whom can be assumed to have used onboard microphones on their personal computers. Fortunately, despite the less-than-ideal recording environment, only twenty-two of 1,191 audio samples featured in identification tasks were deemed by a rater as too difficult to identify, and in these instances always by a single rater.

---

[10] For a detailed look at how Speak Everywhere can facilitate projects like the present one, see Fukada (2013.)

3.2 Experiment 2: Acceptability of Tensified Onsets

3.2.1 Stimuli

The stimuli selected for the present study consisted of a hundred words, divided between ninety experimental items and ten control words (high-frequency words where the onset is either lax and assumed not prone to tensification in previous research, or words realized universally with tense onsets) to test whether participant assessments could be relied upon. Of the experimental items, sixty-seven consisted of words identified in previous studies as prone to tensification, including the twenty-one of the first study in this series. The remainder were gathered from Han (2013) and Jang (2017). A further nineteen were compound nouns where medial tense segments are attested, and four where *medial* tensification attested, in Han & Park (2012). These items were included in an attempt to confirm the findings of previous researchers on word-medial tensification phenomena and to clarify greater questions of whether Korean is restructuring its phonological system along the lines suggested in Shin & Davis (2008). A list of stimuli is provided in Table 26 in the Appendix.

3.2.2 Participants

Eighty-four participants were recruited for the present study, numbering 57 female and 27 male participants with a mean age of 35 years. All were self-identified native speakers of Korean. Initial recruitment took place through an email addressed to participants of previous research who indicated a willingness to participate in future projects. These participants were encouraged to share the survey link to friends and family. Participants were either resident in the United States or South Korea. Accordingly, two methods of providing compensation were utilized: participants were offered the choice of either a $5 Amazon gift card or a ₩5,000 gift card for CU (a ubiquitous South Korean convenience store franchise.) Recruitment was completed over five days. Of the thirty participants in the present study, thirty-eight were resident in South Korea at the time of recording, with the other forty-six scattered across the United States.

Table 4: Descriptive statistics of participants in the acceptability judgment experiment.

	N	Age (Range)	Age (Mean)
Male	27	25-64	34.26
Female	57	22-62	35.39
Total	84	22-62	35.02
Residency			
United States	46		
South Korea	38		
Regional Background		*Use of Dialect in Daily Life? (Yes)*	*Use of Dialect in Daily Life? (No)*
Seoul area/Gyeonggi-do	39	1	38
Gyeongsang-do	32	32	-
Chungcheong-do	6	6	-
Jeolla-do	4	2	2
America[11]	1	-	1
Unsure/No Response	2	2	-
Total	84	43	41

3.2.3 Procedure

Participants completed a Qualtrics questionnaire consisting of three phases. The first phase asked participants about their dialect background and day-to-day use of dialect; participants who indicated that they did not use dialect were asked about where they grew up. A translation of this

---

[11] As this participant's questionnaire was filled out perfectly well in Korean, it may be assumed that the participant is a Korean-American who speaks the language but does not identify with a particular dialectal grouping, a reasonable conclusion as the participant indicated they do not consider themselves a dialect user in daily life.

questionnaire can be found in Figure 4 in the Appendix. The second phase presented the hundred stimuli, directing participants to select the pronunciation which seemed natural to them, and reminding them not simply to follow orthographic convention. The stimuli list can be found in Table 26 in the Appendix. Each stimulus was presented as a set of three options: the word with a lax onset, a selection for 'both sound natural,' and the word with a tense onset. The third phase explicitly discussed tensification, offered examples of words commonly tensified, and presented fifteen statements about attitudes towards onset tensification (for example: "tense onset pronunciations are easier to pronounce".) Participants provided a rating for how much they agreed or disagreed with each attitude statement using a five-point Likert scale (from 1: 'completely disagree' to 5: 'agree completely'.) A translation of this set of questions can be found in the second half of Figure 4 in the Appendix. The task was optimized for both desktop and mobile, with most participants completing the entire task in ten minutes.

3.3 Summary

This chapter described the methodology followed in the experiments featured in the present study: a production task aimed at eliciting productions of words deemed prone to tensification and a survey of speaker judgments as to the acceptability of tensified pronunciations on experimental items and attitudes towards use of novel tense forms. The stimuli and delivery system were introduced, with a discussion of how current public health measures influenced the adoption of the present methodology. The next chapter will present the results, addressing each research question and hypothesis in turn with the relevant findings.

CHAPTER 4.   RESULTS

This chapter presents the results of the two experiments conducted during the present study, beginning with the findings of the production task, moving on to those of the acceptability judgment task. Before introducing the results, a brief assessment of how well the methodology succeeded in eliciting the desired phenomena will be provided.

4.1   Efficacy of the methodology

The target word was successfully solicited in just over two-thirds of all instances. Out of a total of 1,617 attempts, the desired word was produced by participants in 1,091 instances, with 505 occurrences wherein participants produced unintended words, and a further twenty-one items judged unintelligible by the researcher owing to poor audio quality. These unintelligible items were discarded, with the remaining items delivered to the native-speaker raters for the rating task. As not all stimuli were equally successful in eliciting the target word, the total number of items passed on to raters differed from one target word to the next.

By measure of how many participants responded with the target word, the stimuli performed across a broad range. One stimulus pairing of image and sentence only elicited the target word four times out of seventy-seven attempts, while one pairing succeeded in every instance *but* one. The median success rate of the stimuli was sixty-one successful elicitations of the target word out of seventy-seven attempts.

In an attempt to verify the reliability of the rating task, included among the items from successful elicitations of target words were a further hundred items garnered from the pool of non-target words. These additional control samples came from solicitations of five words with lax onsets where tensification is believed inactive and five words already realized universally with tense onsets. Table 25 in Appendix D displays the results for these control items. Out of a total of 300 tokens, the unpredicted segment was given by raters in twenty instances (6.67%.) It cannot be determined whether in these instances participants produced the sought-after segment, which was then misjudged by raters, or if the opposite occurred; it merely suggests that the results delivered following the methodology largely align with predictions.

Raters listened to all the tokens presented in random order, with the aforementioned reliability-assessment tokens mixed in among experimental items. Raters were presented two competing orthographic representations of the syllable (one with a lax onset and one with a tense onset) and instructed to pick the one they felt represented the sound. A third option was offered to indicate that a given item was inaudible or indiscernible.

4.2   Results of the production experiment

RQ1.   Will participants demonstrate word-onset tensification in their productions of experimental items?

H1.   It is predicted that some participants will produce tensified onsets on experimental items.

The hypothesis was roundly confirmed, with most participants demonstrated as producing at least some tense forms, and some participants utilizing a novel tensified pronunciation in over eighty percent of all instances. The results of the rating task showed that tensification was indicated by at least one rater in one instance for all target words. Of 1,091 pronunciations of target words by participants, 297 were judged as having tense onsets by at least one rater. Notably, the raters displayed disagreement, with only 96 of the tokens deemed as tense by all three raters. A further 66 were judged as tense by two raters, and a further 135 deemed tense by only one rater. The results for the lax and universally tense words included as a test of the rating task suggest that the method is valid, although slight disparities emerged again between raters.

Table 5. Results of the rating task.

Target Word	IPA	N	Tense Onsets (N)	Tense Onsets (%)
동그라미	/toŋgɯrami/	225	15	6.8
번데기	/pəndegi/	222	115	51.8
닭다	/takt*a/	132	27	20.5
당기다	/taŋgida/	63	0	0
구부리다	/kuburida/	45	3	6.7
감다	/ka:mt*a/	228	18	7.9
곶감	/kotk*am/	147	46	31.5
밖	/pak/	12	1	8.3
둥글다	/tuŋgɯlda/	144	14	9.7
볶다	/pokt*a/	165	16	9.8
두드리다	/tudɯrida/	198	20	10.1
부수다	/pusuda/	192	17	9.1
가시	/kaʃi/	228	22	9.6
둑	/tuk/	105	70	66.7
굽다	/kupt*a/	219	21	9.6
베끼다	/pek*ida/	213	23	10.8
고추	/kotɕʰu/	213	51	24.1
구기다	/kugida/	213	20	9.5
부서지다	/pusədʑida/	63	6	10.2
두껍다	/tuk*əpt*a/	63	22	35.5
불리다	/pullida/	183	28	15.3
Totals:		3273	555	17.1

Note that as three raters issued judgments on each sample, each token is included three times. As raters at times differed in their judgments, the numbers here represent the number of instances a single rater judged a pronunciation of the onset as tense.

Given evidence of disagreement between the raters, an inter-rater reliability analysis was performed following the methodological recommendations for computing inter-rater reliability with nominal values as provided by Hallgren (2012). Discarding tokens where one rater deemed the audio quality too low to distinguish the sound[12], Cohen's κ was computed for each pair of raters in both rater groups. The resulting kappas were then averaged to provide a single index of inter-rater reliability for each team of raters; the averaged kappas of 0.564 and 0.673 reflect 'moderate' and 'substantial' strength of agreement according to work by Landis & Koch (1977) on interpreting kappa values for observer reliability data.

Table 6. Kappas for each pair of raters.

Group 1

Rater	1	2	3
1		0.659	0.439
2	0.659		0.595
3	0.439	0.595	

Average κ: 0.564

Group 2

Rater	1	2	3
1		0.725	0.699
2	0.725		0.595
3	0.699	0.595	

Average κ: 0.673

---

[12] Happily, such instances proved rare, given the provenance of the recordings; only twenty-two samples out of 1,191 presented (including experimental and control items) were deemed by a rater as too difficult to discern.

For each participant, an index of their individual inclination towards tensification was calculated, using the number of tokens marked by raters as tense out of the total number of tokens.

Table 7: Tensified onsets as produced by individual participants, broken down by gender and use of dialect.

		N	*Mean*	*SD*	Range	
					Min.	Max.
Total		77	17.14	13.38	0	82.05
Sex	F	54	14.96	10.47	0	52.08
	M	23	22.26	17.75	3.7	82.05
Dialect Use	Y	40	15.72	14.60	0	82.05
	N	37	18.68	11.92	0	51.28

Percent of Tense Onsets

RQ2. Will participant ages display a significant relationship with the degree of tense forms produced?

H2. It is predicted that younger participants will exhibit greater use of tensified forms, following the results of previous studies.

The hypothesis was proved incorrect, with no discernible disparity in use of novel tense forms by younger speakers in the participant pool. A Spearman's rank-order correlation was run to determine whether a statistically significant relationship exists between participant age and ratio of tense segments produced. No significant correlation was observed between the two variables (r_s = -.076, n = 77, p = .511).

RQ3. Is a significant difference observed in the degree of tensified forms produced between participant groups of different genders?

H3. It is hypothesized that no significant disparity in tensification will be observed across gender lines.

The hypothesis was not supported by the present findings. In contrast to previous empirical studies of tensified form use, male speakers were shown to use such novel forms more. However, the findings are in line with studies on the *perception* of tensification which suggest associations with masculinity. Participants were divided into two groups for each gender, and the resulting mean tensification ratios for each group calculated.

Table 8: Percent of tense onsets by gender group.

		N	*Mean*	*SD*	Range	
					Min.	Max.
Gender	F	54	14.96	10.47	0	52.08
	M	23	22.26	17.75	3.7	82.51
Total		77	17.14	13.38	0	82.51

To determine whether a statistically significant relationship exists between gender group and tensification ratio, an independent-samples t-test was carried out. Before running the t-test, the homogeneity of variances was checked using Levene's test. The results showed the distribution of variances to be equal: $F(1,75) = 2.079$, $p = .154$.

Table 9: Results of independent t-test comparing degree of tensification between gender groups.

	N	Mean	SD	Sig. (2-tailed)
F	54	14.96	10.47	
M	23	22.26	17.75	
				.028

A significant difference was observed in use of tensification between male speakers (Mean = 22.26, SD = 17.75) and female speakers (Mean = 14.96, SD = 10.47); t(75) = -2.248, *p* = .028.

RQ4. Do speakers of different dialects demonstrate a marked disparity in their production of novel tense forms?

H4. It is hypothesized that speakers of different dialects will demonstrate different degrees of use of novel tense forms.

The results suggest that the hypothesis is incorrect, with speakers self-identifying as dialect or nondialect users tensifying to similar degrees. Before discussing participant dialects, it may be helpful to go through an overview of Korean dialects, their distribution, and the share of the Korean population using each dialect.

방언에 의한 지역 구분

동북 방언
서북 방언
중부 방언
동남 방언
서남 방언
제주 방언
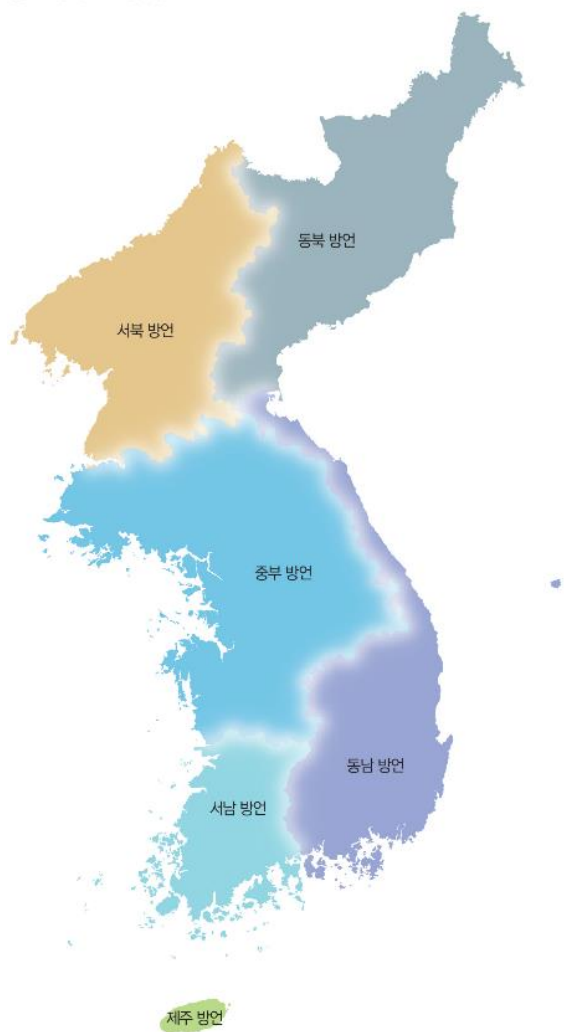
Korean is largely divided into six dialect areas [13], although further distinctions are often made within these larger varieties.

Northwestern and **Northeastern** varieties are spoken in their respective parts of North Korea and will not be discussed in the present study. The **Central** dialect area comprises both the **Seoul** and **Pyongyang** regions, as well as the region of **Chungcheong-do** south of Seoul. The **Southwestern** dialect is spoken in the heavily populated **Gyeongsang-do** region, while the **Southeastern** dialect is associated with the **Jeolla-do** region. The island of **Jeju**, off the southern coast of the Korean peninsula, also has a unique dialect. The region of **Gangwon-do** is divided between speakers of Central dialects on the western side of the Taebaek mountains and speakers of Southwestern dialects on the coast.

---

[13] The maps on this the following page are taken from the National Atlas produced by the National Geographic Information Institute, a South Korean governmental body. The full atlas can be accessed at http://nationalatlas.ngii.go.kr/pages/page_65.php

국내 방언 사용 인구 비율
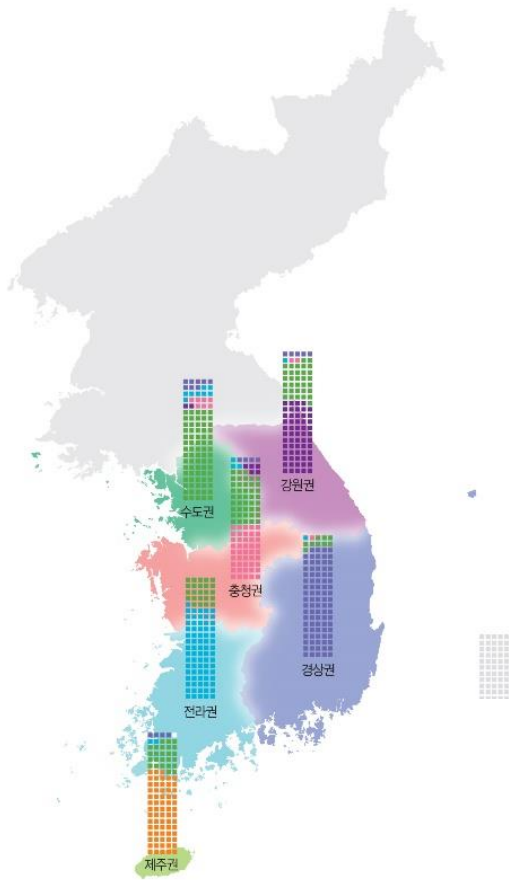
In practice, South Koreans typically draw finer distinctions: **Standard Korean** speakers in the Seoul region, **Gangwon-do** dialect in the mountainous northeast, **Chungcheong-do** dialect in the middle of the country, **Gyeongsang-do** dialect in the southwest, **Jeolla-do** dialect in the southeast, and **Jeju** dialect off on its island. The map attached at left also shows the distribution of dialect speakers *within* the regions, with each square representing 1% of the population.

The speaker population of each regional dialect varies considerably:
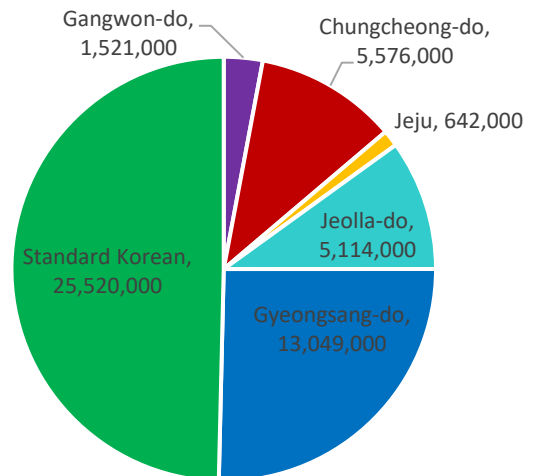
Eighteen participants claimed not to use a regional dialect, and the same participants' responses to a query about their home regions confirmed that all are speakers from Seoul and its immediate environs. While Korean linguists would perhaps rightfully point out that Seoul has its own particular dialect, the standard dialect of Korean is based on the speech of this region. Established in 1394 as Hanseong, the capital of the kingdom of Joseon, Seoul remained the political center of Korean civilization until the collapse of Joseon in the late nineteenth-century and an ensuing period of subjugation under imperial Japan. The city has regained its former

prominence as the capital of the modern South Korean state. The position of Seoul speech as a prestige dialect has been formally recognized by state actors as early as 1912 (Lee & Ramsey 2000:307). Meeting in 1933 in an attempt to standardize spelling, Korean linguists took middle-class Seoul speech as the 'standard' spoken Korean (Sohn, 1999:59). The position of Seoul speech as a prestige dialect, with greater influence on speech norms, was confirmed in 1988 when South Korea's Ministry of Education released a set of regulations attempting to define a 'standard' Korean. The very first clause sets the stage: "*As a principle, the standard language is based on modern Seoul speech as spoken among the educated*" (South Korean Ministry of Education 1988). Since democratization, the government has perhaps become more inclined to take regional sensitivities into consideration; this clause is absent from the most recent regulations (South Korean Ministry of Culture, Sports, and Tourism 2017). Nevertheless, the Seoul variety of speech still wields outsized influence on Korean as the prestige dialect.

Of the forty participants who indicated use of dialect in daily life, twenty-six reported use of Gyeongsang-do dialect, five of Jeolla-do dialect, four of Seoul area dialect, and three of Chungcheong-do dialect. One participant responded that they use both Jeolla-do and Gyeongsang-do dialect, while another provided no information on which dialect they use. Even among participants that indicated use of dialect in day-to-day life, the results of the survey suggest that such behavior is limited. Just over half of participants reporting dialect use do so 'often' or 'always' in the most intimate social context offered, speaking with family members. Participants reported using dialect more in speaking situations than when writing, which has the benefit of somewhat alleviating the potential issue of orthography influencing pronunciation.

Table 10. Participant self-reported use of dialect by social situation.
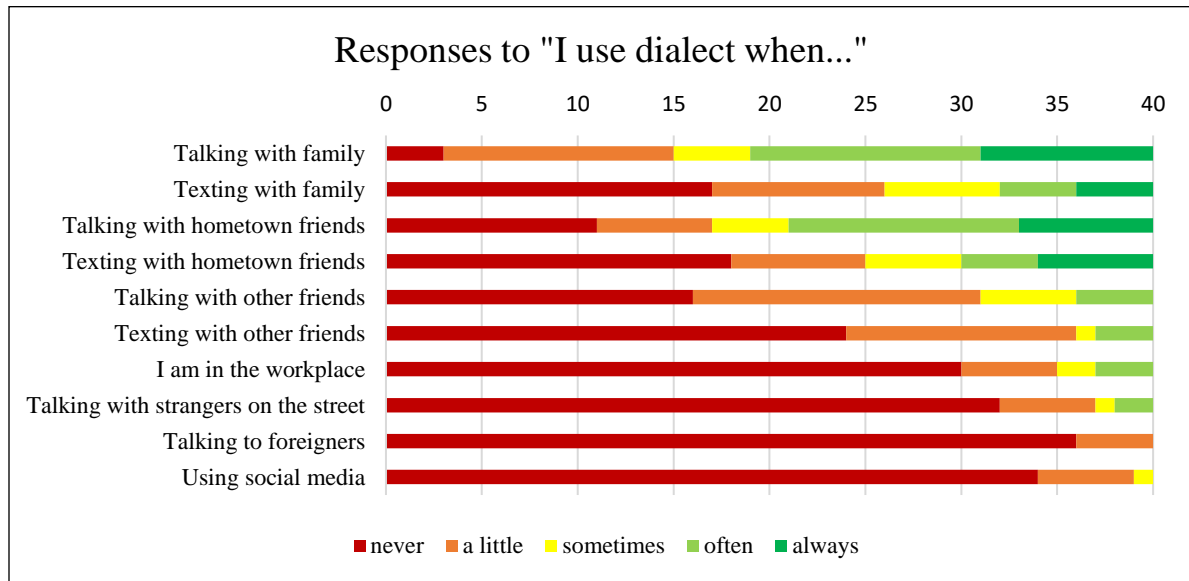
Responses to "I use dialect when..."

Talking with family
Texting with family
Talking with hometown friends
Texting with hometown friends
Talking with other friends
Texting with other friends
I am in the workplace
Talking with strangers on the street
Talking to foreigners
Using social media

■ never ■ a little ■ sometimes ■ often ■ always

The participants were divided into two groups: self-reported dialect users and non-dialect users. The percentage of tense onsets produced was calculated for each group.

Table 11. Summary of tensified onsets produced by dialect group.

		N	*Mean*	*SD*	**Range**	
					Min.	Max.
Dialect Use	Y	40	15.72	14.60	0	82.05
	N	37	18.68	11.92	0	51.28
Total		77	17.14	13.38	0	82.05

To determine whether a statistically significant relationship exists between use of regional dialect and tensification ratios, an independent-samples t-test was carried out.

Table 12. Results of independent t-tests comparing degree of tensification between dialect groups.

	N	*Mean*	*SD*	*Sig. (2-tailed)*
Y	40	15.72	14.60	
N	37	18.68	11.92	
				.336

No significant difference in the use of tensification was demonstrated between dialect speakers (N = 40, 15.72 ±14.60) and speakers who claimed no use of dialect (N = 37, 18.68 ± 11.92;) t(75) = -.969, *p* = .336.

RQ5. Do participants with a self-reported proficiency in a foreign language exhibit greater likelihood of using novel tense pronunciations?

H5. It is hypothesized that speakers with high self-reported proficiency in a foreign language will produce more novel tense forms.

The results do not support the hypothesis. Foreign language skills were not demonstrably associated with increased use of novel phonological changes in participant native-language speech in the present experiment. Two participants indicated use of a foreign language other than Korean in the home during their upbringing, Japanese in one household and both English and Chinese in the other. All participants but four indicated that they had studied at least one foreign language. All participants except four had studied English. Participants self-assessing as having high proficiency in a foreign language, taken here as scoring themselves at seven or higher on the scale given, only did so in the cases of English (thirteen participants) and Japanese (two).

With the exception of English, no foreign language reported by participants exhibited a mean proficiency at the intermediate level. Based on this, only proficiency in English was selected as a potential correlate for later analysis.

Table 13. Descriptive statistics for participant self-assessment of foreign language proficiency.

Language	N	Mean	SD	Range	
				Min.	Max.
English	73	4.78	1.89	1	9
Japanese	27	2.26	2.16	1	9
Chinese	11	1.64	1.29	1	5
Spanish	7	1.43	0.53	1	2
French	9	2.00	1.32	1	5
German	5	1.00	0.00	1	1

Participants assessed their abilities on a nine-point Likert scale, values given from 1: Absolute beginner to 5: intermediate and 9: near-native.

A Pearson's correlation coefficient was computed to determine whether a statistically significant relationship exists between self-reported English proficiency and speaker tensification ratio. Four participants who claimed no English proficiency were excluded from this analysis. No significant correlation was observed between the two variables (N = 73, r = -.090, p = .450).

RQ6. Is a significant relationship observed between participant language attitudes, as measured using questionnaires, and use of tensified pronunciations?

H6. It is hypothesized that participants who report stronger conservative attitudes towards language change and language use will produce fewer tensified variants.

The hypothesis is not supported by the findings of the experiment. Participant attitudes towards language use, both in terms of linguistic conservativism and positive attitudes towards dialect use, had little relationship to the degree of novel tense forms exhibited during the production task. Participants who expressed a sentiment that upon hearing an unfamiliar language, they wished the speaker used Korean instead were *more* likely to use novel tense pronunciations. This is contrary to the hypothesis, but it suggests either that the sentiment is not a clear indicator of conservative attitudes toward language use or (weakly) that bearers of such sentiments are also quicker to adopt novel tense pronunciations.

Table 14. Descriptive statistics for participant responses to each item on language attitudes.

	N	Mean	SD	Range	
				Min.	Max.
Foreign languages have too much influence on Korean.	77	3.36	0.90	1	5
When learning a foreign language, one should aim at achieving a native-like pronunciation.	77	3.45	1.10	1	5
When I hear native speakers of a foreign language, I wish I could speak it as well as them.	77	4.32	0.82	2	5
When I hear a foreign language I do not know, I wish the speakers would just speak Korean instead.	77	2.65	1.23	1	5
Young people these days speak Korean differently than I do.	77	2.83	0.92	1	5
The previous generation speaks Korean differently than mine does.	77	2.65	1.01	1	5
Korean is changing rapidly.	77	3.68	0.80	2	5
It would be a shame if Korea loses its dialectal diversity.	77	3.81	1.00	2	5
I wish the whole world would speak the same language.	77	3.05	1.40	1	5
When I study a foreign language, I find it more difficult to speak and listen than to read or write.	77	3.49	1.08	1	5

Responses were given on a Likert scale ranging from 1: Disagree completely to 5: Agree completely.

For each statement, a one-way analysis of variance (ANOVA) was computed to determine whether participant groups who responded differently to each item displayed significant differences in the degree of tense pronunciations exhibited during the production task.

Table 15. Significance values for tense onsets produced and responses to attitudinal questions.

	Sig.
Foreign languages have too much influence on Korean.	.192
When learning a foreign language, one should aim at achieving a native-like pronunciation.	.072
When I hear native speakers of a foreign language, I wish I could speak it as well as them.	.461
When I hear a foreign language I don't know, I wish the speakers would just speak Korean instead.	.042
Young people these days speak Korean differently than I do.	.022
The previous generation speaks Korean differently than mine does.	.190
Korean is changing rapidly.	.893
It would be a shame if Korea loses its dialectal diversity.	.080
I wish the whole world would speak the same language.	.080
When I study a foreign language, I find it more difficult to speak and listen than to read or write.	.303

A statistically significant difference between groups was observed with the attitude statement *"When I hear a foreign language I don't know, I wish the speakers would just speak Korean instead;"* a Tukey post-hoc test revealed that participants who answered *agree completely* produced significantly more tense segments (N = 8, 30.14 ± 22.93) than participants who answered *disagree* (N = 35, 15.67 ± 11.95, *p* = .041) or *neither agree nor disagree* (N = 9, 11.89 ± 11.53, *p* = .036). No significant difference was observed with the groups answering *disagree completely* or *agree.*

A statistically significant difference between groups was observed with the attitude statement *"Young people these days speak Korean differently than I do;"* however, a Tukey post-hoc test showed that this difference was exhibited between participants who answered *agree completely* (N = 2, 44.39 ± 55.30) and those who answered *disagree* (N = 26, 17.32 ± 11.58, p = .038) or *neither agree nor disagree* (N = 28, 13.41 ± 9.31, *p* = .012). Any interpretation of this

result is hindered by the low size of the group answering *agree completely* and that this group contains the participant who exhibited the highest proportion of tense pronunciations.

A statistically significant difference between groups was observed with the attitude statement *"It would be a shame if Korea loses its dialectal diversity;"* however, a Tukey post-hoc test did not suggest significant differences between any two participant groups in direct comparison, although the difference between groups answering *disagree* (N = 10, 26.81 ± 23.48) and *agree* (N = 28, 14.21 ± 9.54) was close to significance (p = .051).

A statistically significant difference between groups was observed with the attitude statement *"I wish the whole world would speak the same language;"* however, a Tukey post-hoc test showed that no differences between means of specific groups proved significant.

RQ7. Does use of novel tense forms relate to individual differences in personality, as measured using self-evaluated personality tests?

H7. It is hypothesized that participants with higher openness and extraversion scores on a self-assessed personality questionnaire will produce more tense variant onsets.

The results did not indicate any statistically significant relationship between personality factors such as openness and extraversion with the degree of tensification displayed by participants. The results cannot be offered in support the hypothesis, although as noted below, internal consistency issues related to the personality inventory hinder any confident conclusion based on those findings. In the hope of making interpretation more intuitive, participant responses to the Big Five inventory were transformed into percentage of maximum possible scores (Cohen et al, 1999), a linear transformation of the test scale into a zero-to-one-hundred scale where zero represents the minimum and 100 the maximum possible score.

Table 16. Descriptive statistics for participant responses on subcomponents of personality.

Subcomponent	N	Mean	SD	Range	
				Min.	Max.
Agreeableness	77	59.35	12.42	30	90
Conscientiousness	77	69.03	11.95	35	90
Extraversion	77	65.06	14.08	30	95
Neuroticism	77	35.26	16.38	0	75
Openness	77	66.95	11.45	45	95

The internal consistency of inventory items for each subcomponent of personality, using Cronbach's alpha as a measure of reliability, is .440 for Openness, .612 for Conscientiousness, .762 for Extraversion, .568 for Agreeableness, and .741 for Neuroticism.

Excepting extraversion and neuroticism, these values are lower than desired, with openness— a trait identified in the literature review as potentially influencing phonetic imitation— falling well below standards. Further analysis determined that removing a single statement from the set of openness items, "*I am keenly sensitive by nature*," would raise the measure Cronbach's alpha for the remaining set to .577, yet still below desired values.

The Pearson's correlation coefficient was calculated for each pairing of participant tensification ratios and Big Five personality subcomponents (extraversion, agreeableness, conscientiousness, neuroticism, and openness.)

Table 17. Pearson's correlation coefficients for percent of tense onsets produced and personality trait.

Personality Trait	r	Sig
Agreeableness	-.137	.235
Conscientiousness	.137	.235
Extraversion	.205	.074
Neuroticism	-.008	.946
Openness	-.031	.786

4.3   Results of the acceptability judgment experiment

A multiple regression was run to predict participant acceptance of novel tense pronunciations, expressed as the ratio of tense pronunciations deemed natural of the set of stimuli, from age, gender, self-reported use of dialect in daily life, whether the participant viewed tensification as characteristic of their regional speech, whether the participant viewed tensification positively, and whether the participant identified as using tense pronunciations in their own speech. The model weakly predicted the ratio of novel tense pronunciations deemed acceptable to the participants, with $F(6,77) = 2.623$, $p = .023$, $R^2 = .170$. However, none of the regression coefficients for the individual variables in the model proved statistically significant. To determine the effect size of each predictor variable, squared semi-partial correlation coefficients were calculated. The results suggest that the effect size is low for each variable.

Table 18. Results of the multiple regression analysis.

	β	*Sig.*	sr^2
Age	-.006	.097	.030
Gender	.035	.564	.004
Self-reported dialect use	-.079	.212	.017
Tensification as characteristic in own regional speech	.075	.322	.011
Positive view of tensification	-.026	.493	.005
Self-reported use of tensified forms	.035	.303	.012

Given that the model proved significant, the failure to demonstrate significance among the variables may suggest that one or more of the predictors are correlated. It could also simply be the result of adding many predictor variables to the model. To try and clarify this, a series of independent-sample t-tests was run with each predictor variable, with the exception of age, for which a linear regression was performed, and positive view of tensification, for which a one-way analysis of variance was performed. The results of each are presented under the pertinent research question.

RQ1. Will participants demonstrate evidence of word-onset tensification by indicating novel tense pronunciations as equally (or more) natural in comparison to lax forms?

H1. It is hypothesized that participants will deem some of the tensified forms as more or equally natural to their lax-segment-bearing counterparts.

The hypothesis was readily confirmed, with participants demonstrating a clear preference for pronunciations using novel tense forms in many cases. Out of the sixty-seven stimuli featuring tensification-prone onsets, at least half of all participants indicated a tense pronunciation was natural for thirty-seven words. For fourteen words, the tense pronunciation was preferred by participants over the lax. An item-by-item breakdown of the preferred pronunciations for each stimulus item is provided in Table 27 in the Appendix.

RQ2. Will participants of different age groups display significant disparities in the degree of tense forms preferred?

H2. It is hypothesized that younger participants will display greater acceptance of novel tense forms.

The results support the hypothesis, with younger participants displaying a positive difference in the degree of tensified variants preferred. A simple linear regression was computed to predict the ratio of tensified onsets deemed as natural based on participant age. A significant correlation was observed between the two variables: $F(1,82) = 6.939$, $p = 0.10$, R = .078.

RQ3. Is a significant difference in the degree of preference for tensified forms demonstrated across gender lines?

H3. It is hypothesized that no significant difference will be observed in the preference for tensified forms by gender group.

The results support the hypothesis, with the findings suggesting no significant difference across gender lines. Participants were divided into two groups for each gender, and the resulting mean percentage of tense onsets deemed natural sounding calculated for each group.

Table 19. Percent of tense onsets deemed natural by gender group.

		N	Mean	SD	Range	
					Min.	Max.
Gender	F	57	48.05	26.85	4.48	88.06
	M	27	52.63	28.03	2.99	90.90
Total		84	49.52	27.15	2.99	90.90

To determine whether a statistically significant relationship exists between gender group and tensification ratio, an independent-samples t-test was carried out. There was no significant difference in use of tensification between the two gender groups: $t(82) = -.720$, $p = .474$.

RQ4. Do speakers of different dialects demonstrate a marked disparity in their preference of novel tense forms?

H4. It is predicted that participants reporting dialect use will demonstrate different levels of tensification than self-affirmed nonusers.

The results suggest that the hypothesis was correct, with dialect and nondialect groups demonstrating markedly different acceptance and preference of novel tense forms. Self-affirmed nondialect users were shown less likely to prefer tense forms compared to speakers who claimed daily use of dialect.

The participants were divided into two groups: self-reported dialect users and non-dialect users. All but three of the self-reported 'non-dialect' users are from the Seoul region, as reflected in their responses to the questionnaire. The percentage of pronunciations with tense onsets deemed natural was calculated for each group.

Table 20. Percentage of words with tensified onsets deemed natural by self-identified participant dialect use.

		N	*Mean*	*SD*	Range	
					Min.	Max.
Dialect Use	Y	43	56.99	26.63	2.99	90.90
	N	41	41.68	25.73	4.48	86.57
Total		84	49.52	27.15	2.99	90.90

To determine whether a statistically significant relationship exists between use of a regional dialect and the percent of novel tense pronunciations preferred, an independent-samples t-test was carried out.

A significant difference in preference of tensified forms was demonstrated between dialect speakers and speakers who claimed no use of dialect: t(82) = -2.678, *p* = .009.

RQ5. How do participants perceive the use of tensified pronunciations?

H5. It is predicted that participants in the present experiment will associate tensified forms as less formal, more masculine, less feminine, more friendly, more vulgar, and more emphatic and expressive. It is further hypothesized that speakers will explicitly associate tensification with regional dialects, with participants having positive attitudes towards tensification more likely to attribute it to their own varieties of Korean.

Table 21. Participant attitudes to novel tensified pronunciations.

Tensified onset pronunciations are…	N	Mean	SD	Median	Range	
					Min.	Max.
More expressive	84	3.99	1.04	4	1	5
Easier to pronounce	84	3.73	0.95	4	1	5
Emphatic	84	3.67	1.00	4	1	5
Friendlier	84	3.55	0.95	4	1	5
Better at communicating what I mean	84	3.35	1.05	4	1	5
Used by me	84	3.24	1.15	3	1	5
Vulgar	84	2.77	1.09	3	1	5
Positive	84	2.46	0.86	2	1	5
Negative	84	2.31	1.01	2	1	5
Awkward	84	2.30	0.86	2	1	5
Masculine	84	2.23	1.18	2	1	5
Cute	84	2.14	1.02	2	1	5
Formal	84	2.11	0.96	2	1	5
Used more by younger people	84	1.92	0.82	2	1	4
Feminine	84	1.81	0.74	2	1	4

Responses were given on a Likert scale ranging from 1: Disagree completely to 5: Agree completely.

Several parts of the hypothesis were confirmed, with participants considering tensified pronunciations as less formal, less feminine, more friendly, (slightly) more vulgar, more emphatic, and more expressive. Speakers surveyed did not turn out to associate such pronunciations with masculinity, contrary to the hypothesis. However, this seems to contrast both with the low association given with femininity (with only a single participant assigning it a rating of 'agree.')

It was hypothesized that speakers who view themselves or their own dialect speakers as using tense pronunciations will have a greater positive attitude towards the phenomenon. A one-

way analysis of variance (ANOVA) was carried out to determine whether a statistically significant relationship exists between reported self-use of tense forms and a positive view of onset tensification. A significant difference was observed between the two groups, with participants who indicated that their dialect or regional speech uses tensified pronunciations more likely to indicate agreement with the statement "I use tensified pronunciations," $F(1,82) = 10.516$, $p = .002$.

Likewise, participants who indicated agreement with the statement "I view tensified pronunciations positively" (a looser translation of the questionnaire item to make it sound more natural in this context) were more likely to report self-use of such novel forms. A one-way analysis of variance found the difference to be significant: $F(4,79) = 6.165$, $p = .000$. Interestingly, this is not reflected in the raw data of preferred tense forms from the questionnaire: a one-way analysis of variance found that participants grouped together by degree of positive attitude towards tensification were *not* more likely to prefer a greater number of tensified forms: $F(4,79) = .541$, $p = .706$.

To determine whether a statistically significant relationship exists between belief that one's own regional dialect uses tense pronunciations and a positive attitude towards the phenomenon, an independent-samples t-test was carried out between speakers who answered yes and no to a query about whether such pronunciations are characteristic of their dialect. There was a significant difference between speakers who saw tense pronunciations as characteristic of speakers of their region or not: t(82) = -2.424, $p = .018$. This aspect of the hypothesis was proved correct.

Participants disagreed as to whether tensified pronunciations are characteristic of their own dialects and in the assignment of said pronunciations to a certain regional dialect.

Table 22. Participant responses to "These pronunciations are characteristic of speakers from my region."

Region	N	Yes	No	If *no*, which region do you associate them with?
Seoul area	39	20	19	Gyeongsang-do (5), Jeolla-do (2), Daegu (2), Ulsan (1), Busan (1)[14]
Gyeongsang-do	32	28	4	Jeolla-do (1), Gyeongsang-do (2)[15]
Chungcheong-do	6	4	2	Gyeongsang-do (3), Busan (1), Gangwon-do (1)
Jeolla-do	4	3	1	Gyeongsang-do (1)
Don't Know/NR	2	2	-	
America	1	-	1	
Total		55	27	

4.4 Summary

This chapter presented the results of the experiments performed during this study, confirming that onset tensification is active in contemporary Korean and can be captured using the present methodology. The findings were inconclusive with regards to hypotheses about facets of participant demographic, attitudinal, and experiential backgrounds and use of or preference for novel tensified pronunciations. The results did not support the hypotheses regarding participant personality traits, foreign language learning background, or language attitudes. However, the results of each experiment stood in opposition as to the significance of participant age and dialectal background. The following chapter will discuss these results, providing a possible interpretation of the data in light of previous research on onset tensification phenomena.

---

[14] Daegu, Ulsan, and Busan are all large cities in Gyeongsang-do with distinctive accents, even within the region.
[15] It is possible that these participants were attempting to convey something like "*My* variety of Gyeongsang-do dialect doesn't do this, but *over there* they do!" It is unclear, however.

CHAPTER 5.   DISCUSSION

5.1   Interpreting the results

The results of the present study suggest that tensification is alive and active in modern Korean; participant productions of tensification-prone items during the production tasks were judged by raters as produced with tense onsets in up to 71.1% of instances. During the acceptability judgment task, tensified onset pronunciations were deemed acceptable by over half of all participants for 37 out of a total of 67 experimental items; for fourteen of those words, the tensified form was deemed by participants as *more* acceptable than with a lax pronunciation. While wide variability in degree of tensification was observed from item to item, no tense onsets were observed for only one word out of twenty-one in the sample. The results confirm the findings of previous studies, suggesting that with few exceptions, the experimental items featured are susceptible to tensification. Based on both participant productions and judgments of acceptable pronunciations, some of the featured words are already deep within the 'tensification trap,' largely being produced and preferred with tense onsets.

5.2   Onset tensification and its association with demographic factors

Whether tensification can be tied to demographic factors such as age, gender, or regional dialect background remains less clear, owing to conflicting results between the production and acceptability judgment experiments. No statistically significant effect of participant age, gender, or dialectal background on the degree of tense forms produced was identified following the production task (although, as will be noted later, the participant pool may have been too homogeneous to permit an interpretation in either direction.) In contrast, the acceptability judgment task revealed a significant effect for age, with younger speakers preferring more tensified onset forms. Similarly, self-identified dialect users similarly displayed a significantly higher inclination to novel tensified pronunciations. No gender effect was found in the acceptability judgment task, but participants were demonstrated to believe that use of such forms is not associated with femininity. The lack of a gender gap in the data observed is in line with other studies on onset tensification which suggest no significant difference between genders in its application (Jeon, 2019), contrasting interestingly with earlier studies suggesting tensification is

associated with masculinity, at least in younger speakers (Lee, 1996). Given the two-and-a-half decades between Lee (1996) and the present study, an interpretation is possible that although tensification may have started among male speech communities, it has now spread to the larger population, evidenced by the disappearance of a significant gender gap as well as increasing use of tense forms from one generation to the next as shown by Jeon (2019) and in the present acceptability judgment experiment, wherein younger participants were shown to display a higher preference for tensified onset pronunciations.

The apparent disparity observed between the production and acceptability judgment task, wherein many speakers affirmed that tensified forms are acceptable or even preferred— yet *did not themselves* utilize novel tense pronunciations— could possibly be attributed to the experimental setup. Even though participants were performing the experimental tasks at their own convenience, often presumably from their own homes, the awareness that they were participating in an academic study concerned with speech may have led participants to be more careful of their productions. Likewise, if onset tensification is associated with less formality (as participants indicated in the attitudinal survey,) it could reasonably be expected that participant productions in such a task would perhaps exhibit less tensification than in informal settings. This is in line with findings by Han (2011) on the realization of tensified pronunciations by university students, wherein only 40.6% of participants used a tensified pronunciation at least once during a one-to-five-minute presentation given as part of coursework.

5.3   Onset tensification as a regional dialectal trait

Any assignment of tensification phenomena to a specific dialectal region remains tentative, with participants from diverse dialectal backgrounds showing wide individual differences in the production of novel tensified forms. While no significant effect of regional background was demonstrated during the production task, these findings stand in sharp contrast to the results of the acceptability judgment task. For the two participant populations represented in large number in the present study, speakers of a Seoul/Gyeonggi-do dialect and those of a Gyeongsang-do dialect, a notable difference was observed in their affirmed preference for tensified onset variants, despite this contrast not being reflected in the participant groups' own productions. Of 67 experimental items assumed prone to onset tensification, participants from Gyeongsang-do were shown to display a greater acceptance of the tensified variant than Seoul region speakers in 62 cases. The

small participant pool representing Chungcheong-do and Jeolla-do dialects, combined with the absence of speakers representing other varieties of Korean such as those spoken in Gangwon-do, Jeju-do, or farther-flung Northern regions, makes definitive conclusions as to the status of tensification in these dialects impossible. Nevertheless, the large samples of Seoul area and Gyeongsang-do variety speakers in the present study permit some tentative conclusions. The results of the present study support the findings of Oh (2011) and Jang (2017), suggesting that Gyeongsang-do variety speakers are more accepting of tensified onset variations, in line with the suggestion in Shin and Davis (2008) that this dialect exhibits marked tensification phenomena. Notably, while Seoul area participants were almost evenly divided as to whether onset tensification was a feature of their regional variety (with twenty participants affirming as such, and nineteen responding in the negative,) out of thirty-two participants using Gyeongsang-do varieties, all but four indicated a belief that onset tensification was utilized in their dialect. The findings of the acceptability judgment study also stand in contrast to proposals such as in Jang (2017) that onset tensification is a word-by-word phenomena wherein novel variant pronunciations arise concurrently across the peninsula, with multitudinous regional variants then spreading among speakers of other dialects. Such a conclusion would support the findings of Han (2013), who argued that onset tensification is best characterized as an ongoing phonological change happening across Korean. A larger participant pool drawn from a wider variety of regional backgrounds would help greatly to clarify the matter. Although their representation in the present study was small, speakers of Chungcheong-do and Jeolla-do varieties were also divided (4:2 and 3:1, respectively) as to whether tensification was a feature of their dialect. As observed in the literature review, tensification would appear to be active across the peninsula, although in this instance the spread of novel tense forms may be from the periphery to the core.

Such findings dovetail with another curious phenomenon observed in contemporary Korean phonology. Seoul speakers are becoming *more like* nonstandard variety users in terms of the acoustic cues used to distinguish between stops, shifting from voice onset time to f_0 (Silva, 2006). The possibility remains that speakers of 'nonstandard' dialects such as Gyeongsang-do varieties are quicker to adopt novel phonological behaviors, and that although these changes ultimately diffuse throughout the speaker population (as observed in the historical data demonstrating onset tensification over previous centuries,) the 'standard' varieties closely associated with the Seoul region are more resistant to change.

CHAPTER 6.   CONCLUSION

The results of the present study confirm the findings of previous research into onset tensification phenomena, demonstrating that native speakers both produce and prefer tense onsets on items realized in earlier stages of Korean with lax onsets. In light of the historical record, such behavior can be interpreted as the continuation of sound change processes that have been active in Korean for centuries. Although the present study examined only a limited number of lexical items, it should be noted that many of the words featured are fixtures of day-to-day life: basic kitchen terminology such as *to fry* and *to grill*, the *chili pepper* used ubiquitously in Korean cuisine, common adjectives such as *round, thick* and *wrinkled*, and verbs such as *to wash one's hair, to knock,* and *to smash*. These are not comparatively infrequent or archaic terms which speakers might hesitate before assigning a pronunciation to, but rather, common words necessary for conducting daily life. That the pronunciations of such items are demonstrably in flux reflects the scale of sound change occurring in Korean: where participants were divided in their articulations, this reflects the lapping of a great and centuries-old wave against a shore itself increasingly eroded with each blow.

The ultimate causes of such processes remain murky, to borrow again a word which features so prominently among controversies in Korean phonology, yet so perfectly describes the mechanisms which drive language change. A significant body of work has been assembled, including (if modestly) the present study, to attempt to explain where and when sound changes are manifest, which populations set phonological trends, and how new pronunciations spread through speech communities. Yet even the weight of evidence set forth only allows the most tentative of guesses as to what motivates the sweeping and comprehensive changes we observe in language. In the case of changes involving wholesale restructuring of phonemic inventories, an appeal to evolutionary principles is possible: namely, that strong and invisible selective pressures shape the inventories of each language. The results of such pressures, and a consequent and continuous balancing of complexity and simplicity, can be inferred from linguistic universals and implicational laws showing how phonological inventories are built up as well as the relationship between inventory size, allowed syllable structures, and word length (Jakobson, 1990; Maddieson, 2016). Researchers such as Flemming (1997) and Boersma (1998) have modelled speech

production as the result of compromise between constraints which stand in mutual opposition, between innate desires to reduce articulatory effort and the need to maintain perceptual distinctiveness and the fidelity of the signal, between the interests of speaker and listener. Following such reasoning, given the rarity of tripartite stop contrasts such as that of Korean in other world languages, it is possible that such systems are selected against by internal pressures. Where such a contrast emerges, invisible forces may be acting to simplify them in a way that makes the phonemic inventory wieldier to speakers and convenient to listeners. Alternatively, the spread of tense segments throughout previously lax domains may be part of a great unseen balancing act, a continuous repackaging of phonemes to achieve a distribution which better fits an unspoken ideal. It has been observed above that phoneme frequencies in world languages fit distributions described by power laws (Tambovtsev & Martindale, 2007; Macklin-Cordes & Round, 2020). If the frequency of Korean phonemes displays a demonstrably poorer fit to these distributions, perhaps the unidirectional action of lax segments becoming irredeemably tense with time is a way that the language remedies an inventory out of balance. Such mechanisms remain shrouded in mystery, and we are yet far from deriving the laws governing them. But just as the scale of language change is vast, so is our time to study it. Physicists have solved a few of their easier problems; for linguists interested in the deeper operating principles of language, there is time yet for a hundred apples to strike their Newtons. Let us sit under the tree.

6.1 Is Korean restructuring back towards a bipartite contrast?

Previous discussions of onset tensification have mentioned the possibility, albeit as a remote one, that Korean is in the middle of a larger restructuring of its phonology that will reduce the tripartite phonemic contrast to one between tense and aspirated segments in onset position, with the lax serving as a positional allophone (Shin & Davis, 2008). Already, both aspirated and tense segments are realized as lax in coda position:

밖	밖이	부엌	부엌이
/pak/	/pa.k*i/	/pu.ək/	/pu.ə.kʰi/
"outside"	"outside"+SUBJECT	"kitchen"	"kitchen"+SUBJECT

A shift within Korean towards preference of tense and aspirated segments in word-onset position and lax segments elsewhere is suggested by the findings of several researchers on the *weakening* of word-medial tensification in positions where it is predicted by Korean morphophonological rules (Oh, 2006; Han & Park, 2012). In compound words, the presence of a segment in coda position at the morpheme boundary triggers tensification of a segment in onset position of the following morpheme:

봄	바람	봄바람
/pom/	/param/	/pom.p*aram/
"spring"	"wind"	"spring wind"

However, experimental studies of actual native-speaker productions and the results of acceptability judgment studies suggest that this rule of Korean morphophonology is less binding than previously believed, with predicted tense segments in these medial positions being realized as lax. Oh (2006) found that for eight experimental items, participants produced the prescribed medial tense sound in only 10.3 to 35.8% of all instances. These results were confirmed by Lee (2009), who also found that younger participants expressed a greater preference for pronunciations with such 'weakened' medial segments. Han & Park (2012) found in a survey of native-speaker judgments as to how 44 compound words should be pronounced that participants opted for lax medial segments in 81.82% of instances. In an attempt to confirm these findings, the acceptability judgment experiment in the present study included in its stimuli nineteen of the experimental items from Han & Park (2012) where word-medial tensification is prescribed. The results, presented in Table 29 in the Appendix, showed that participants overwhelmingly preferred pronunciations using lax medial segments for these items, to an even greater degree than that observed in previous studies.

Although the present study included a largely homogeneous cohort in terms of age, it should be noted that the great preference for weakened word-medial pronunciations among participants aligns with Lee (2009), suggesting that in addition to preferring more novel tense onset pronunciations as found in the same experiment, younger generations of Korean speakers are increasingly coming to prefer lax segments in word-medial positions where Korean morphophonological rules predict tensification. Further research is needed to clarify whether this

phenomenon is yet restricted to a limited number of items and applies in degree, as in the case of onset tensification, or whether this is evident of larger shifts in Korean morphophonology. Any guess as to the latter would remain purely speculative from the current standpoint, although the suggestion is intriguing in light of the evidence documented in the acceptability judgment experiment. Given the overwhelming numerical dominance of lax segments, it is perhaps strange to imagine that they would find themselves relegated to the status of a positional allophone, but it is likewise true that pure numerical frequency is no indicator of what phoneme is accepted as underlying (compare, for instance, the case of the English voiceless alveolar stop and its many flavors depending on dialect: aspirated, unreleased, glottalized, flapped, palatalized.) Despite a wealth of research into the frequency of phonemes in dictionary, written, and spoken corpora, a surprising paucity of work on *allophone* frequencies leaves this question opaque.

Whether a concurrent preference for 'weakened' medial segments along with increased acceptance and use of 'tensified' onset segments is coincidence or evidence of larger trends associating phonemes with specific positions in Korean words, and thus possibly even a dramatic restructuring of Korean phonology in the future, the results of the present study demonstrate that both changes are active. Future treatments of either phenomenon would do well to consider the other. The findings of the present study also bolster those of Han & Park (2012) regarding mismatch in the phonetic transcription of word-medial tense segments in prescriptive works on language use such as dictionaries. More research is necessary into how Korean speakers of our time are actually articulating such words; a revision of prescriptive standards may be long overdue.

6.2   Limitations of the present study

The results failed to demonstrate a statistically significant relationship between tensification and the 'Big Five' subcomponents of personality, foreign language proficiency, and language attitudes. While these factors were included to explore an as-yet little studied juncture between speaker experiences or backgrounds and native-language pronunciation, it would appear that adoption of novel pronunciations is not markedly tied to these characteristics. Another possibility is that these factors are correlated to the degree that teasing their effects apart is not feasible under the current methodology; if attitudes towards language use are shaped by experience, both learned experience and that acquired with age, then a high degree of correlation between these

factors may be suspected from the outset. The results of the linear regression performed upon the acceptability judgment results would that this remains a distinct possibility, with the model weakly predicting a tendency to tensify, but without each of these facets of participant backgrounds and attitudes proving a significant predictor in themselves.

The experimental design proved successful at eliciting the desired phenomena. The ratio of tensified forms observed in the data is in line with previous studies; in the most similar recent study, Jeon (2019), participants tensified between six and forty percent of the time depending on age bracket and word. Some points of failure are inescapable under the present experimental design: the participant may produce a different word than the target, for instance, or the sample may contain noise (such as a passing vacuum cleaner, in one case) at the critical juncture where the tensification-prone segment occurs. However, despite concerns about the quality of participants' microphones, the greatest issue observed in the elicitation task proved to be a failure to draft the right pairing of sentence and image. A wide disparity was observed from stimulus to stimulus in efficacy at soliciting the desired word. This can be remedied by testing potential stimuli with greater numbers of participants during the pilot phase, although performance in this regard is impossible to assure; two of three native speakers who tested the present stimuli in the pilot produced the target item for the stimulus pairing that ultimately succeeded *once* among the participant population. A combination of greater pre-testing could alleviate these issues, but one benefit of the current design is that if the item works at all, simply recruiting more participants is also feasible (funds permitting).

The disagreement observed among raters is notable. Even given that inter-rater reliability was demonstrably within the bounds of acceptability suggested in the methodological literature, as noted above, it is perhaps curious that native speakers should disagree to the extent displayed. It is possible that given the degraded condition of the audio, disagreement may arise owing to the low sonority of stop sounds. However, this explanation is perhaps less convincing when considering the highly degraded signals native speakers parse successfully in daily life: a cheap onboard microphone recording a speaker in a quiet room could certainly be more faithful than that of a cellular phone in a subway station as a train comes in. Why should participants disagree so much on the classification of sounds which are phonemic in their own shared language?

While it proved possible to recruit a large number of participants and collect data remotely, upon consideration of the results the participant pool is more homogeneous than is desirable for

the present project. The speakers who participated came from diverse enough backgrounds geographically to permit consideration of dialect as a factor, which proved decisive. However, all but one of the participants are largely of the same generation, and presumably of similar educational attainment. English was only introduced as a mandatory school subject in Korea in 1997, but the median self-assessed proficiency of the participants who spoke some English was at the intermediate level. It would seem that the participants represent a fairly skewed group of native Korean speakers in terms of their experiences- and thus, possibly, their attitudes. In an investigation where attitudes were hypothesized to play a part, this is a potential confound that would require recruitment of more (and more diverse) participants to ameliorate.

Tests of internal consistency for the personality inventory produced poor results, suggesting that for some subcomponents of personality (most notably, openness) the questions used cannot be safely assumed to test for the same features. While Ahn and Chang (1996) propose that the abbreviated test, given here in the hopes of reducing participant fatigue, can capture the desired measures adequately, the results suggest that future iterations of this method should utilize a different inventory. While a smaller complaint in the grander scheme of the present project, low internal consistency between these questions reduces faith in any conclusion that subcomponents of personality are not tied to adoption of novel pronunciations.

6.3   Avenues for future research

The limitations imposed by current pandemic conditions on potential avenues of phonological research have greatly affected both the methodology chosen for the present study as well as the general thematic direction of it. In those heady pre-coronavirus days, the present project originally aimed to collect data directly from participants in a well-equipped recording studio, allowing samples to be analyzed using more detailed methods. The hope was that a detailed acoustic investigation might show whether tensification-prone onsets were markedly different from their lax or universally-tense counterparts (*even* in cases where native-speaker raters classified the result as lax). The results of the present study- in this case, the observation that native-speaker raters *disagreed* over the data- suggests that the original line of thinking remains worth pursuing, when (or if) the world ever returns to normal. It is hoped that the present researcher

will have the opportunity while still in academe; if not, I leave these lines as a trail of breadcrumbs leading off into unknown territory for the next Hansel or Hong Gil-dong[16].

Likewise, collecting acceptability judgments on the pronunciation of a given item can be a quick and easy substitute for soliciting actual production data. This is particularly beneficial in the case of words which might be much more difficult to elicit in a task such as that of the present production experiment; a gapped sentence paired with an image proved quite adept at soliciting *circle* or *red pepper*, familiar and concrete concepts with few synonyms to get in the way. However, given the difficulty of eliciting some fairly frequent monomorphemic words, whether such a technique would work as easily with often rarer multimorphemic (and therefore necessarily more complex) compound words featured in the discussion on word-medial weakening is another matter. The present project specifically aimed to avoid orthographic influence on pronunciation in the production task by using a gapped-sentence presentation. Whether such a step is necessary may also prove a fruitful avenue of research, as is whether the assumption is justified that Korean speakers are aware enough of prescriptive pronunciations for it to override tensification habits. Certainly, for many of the words which demonstrated high acceptability with tensified onsets, a cursory Internet search of the prescriptive standard and the tensified variant is enough to prove that native speakers are often puzzled as to which is the 'correct' version. Such netizens may be comforted to know that with the passage of time, the standard is bound to catch up with their own pronunciation.

---

[16] Korean folklore's answer to Robin Hood.

APPENDIX A: PRE-RECORDING QUESTIONNAIRE

Pre-Recording Questionnaire

What year were you born?[17]

1. Sex (F/M)

2. Besides Korean, did you use any other languages in the home growing up? (Y/N)
   3a. If *Yes*, please write the names of the language(s) used in the space provided.

3. Have you ever studied a foreign language, including English? (Y/N)
   4a. If *Yes,* please write the name of the language(s) in the space provided and indicate your proficiency on the scale provided. (Likert scale, from 1 *absolute beginner* to 7 *near-native*).

4. *However little,* do you believe you use a regional dialect in daily life?
   5a. If *Yes,* which dialect do you believe you use the most?

   5b. If *Yes,* please indicate how often you use dialect in the following contexts using the scales provided. (Likert scale for each item, from 1 *never* to 5 *always*:)

   - When speaking with family members.
   - When texting with family members.
   - When speaking with friends from my hometown or home region.
   - When texting with friends from my hometown or home region.
   - When speaking with friends from outside my hometown or home region.
   - When texting with friends from outside my hometown or home region.
   - In the workplace.

---

[17] Koreans commonly follow traditional East Asian age reckoning practices, now widely disused *outside* Korea, wherein a newborn is *one* at the moment of birth and becomes a year older at either the lunar *or* solar new year. However, the method of age calculation familiar to Westerners, where newborns are zero at birth and age increments annually on anniversaries, is *also* used in South Korean government and legal contexts (for instance, in determining eligibility to vote or purchase tobacco products.) Thus, simply asking a population of Korean participants to provide their *age* will produce answers that may be up to *two years* divergent from researchers' assumptions, regardless of whether traditional or international ages are assumed. I offer this verbose note as a cautionary tale.

●When speaking with strangers I meet in everyday situations.

 ●When speaking with foreigners.

 ●When writing on the internet or on social media.

5c. If *No,* please indicate the place where you lived the longest while growing up.

5. The following are questions about your personality and attitudes toward language use. Please indicate on the scale provided how much you agree with each statement. (Likert scale for each item, from 1 *firmly disagree* to 5 *firmly agree*).
   a. I always treat people warmly.
   b. I get along well with others.
   c. I make efforts to think about everything positively.
   d. I try my best at everything in a positive manner.
   e. I am good at leading others.
   f. I tend to trust other people.
   g. I am honest in everything.
   h. I think of others before myself.
   i. I try to accept every situation.
   j. I respect others and do not raise myself.
   k. I consider myself competent.
   l. I think I organize and manage others well.
   m. I strive to realize my goals.
   n. I work to restrain myself.
   o. I handle given responsibilities surely.
   p. I am always chased after by some thing or other.
   q. I am agitated easily, even by small things.
   r. I always feel sad and powerless.
   s. My feelings of inferiority and constraint are strong.
   t. I am sensitive even to tiny issues.
   u. I always think of new things.
   v. I enjoy cultural life and leisure.
   w. I am keenly sensitive by nature.
   x. I enjoy new experiences.
   y. I respect the opinions of others, even when they differ from my own.
   z. Foreign languages have too much influence on Korean.
   aa. When learning a foreign language, one should aim at achieving a native-like pronunciation.
   bb. When I hear native speakers of a foreign language, I wish I could speak it as well as them.
   cc. When I hear a foreign language I do not know, I wish the speakers would just speak Korean instead.
   dd. Young people these days speak Korean differently than I do.

ee. The previous generation speaks Korean differently than mine does.

ff. Korean is changing rapidly.

gg. It would be a shame if Korea loses its dialectal diversity.

hh. I wish the whole world would speak the same language.

ii. When I study a foreign language, I find it more difficult to speak and listen than to read or write.

Figure 3. Pre-recording questionnaire.

APPENDIX B: STIMULI FOR THE RECORDING TASK.

Table 23. Stimuli for the recording task.

	Word	*IPA*	*English translation*
"Tensified" onsets (lax → tense attested) (*Target words*)	동그라미	/toŋgurami/	'circle'
	번데기	/pəndegi/	'pupa, chrysalis'
	닦다	/takt*a/	'to wipe'
	당기다	/taŋgida/	'to pull'
	구부리다	/kuburida/	'to stoop over'
	감다	/ka:mt*a/	'to wash (hair)'
	곶감	/kotk*am/	'dried persimmon'
	밖	/pak/	'outside'
	둥글다	/tuŋgɯlda/	'to be round'
	볶다	/pokt*a/	'to fry'
	두드리다	/tudɯrida/	'to knock (on)'
	부수다	/pusuda/	'to smash'
	가시	/kaɕi/	'thorn'
	둑	/tuk/	'dyke, levee'
	굽다	/kupt*a/	'to roast'
	베끼다	/pek*ida/	'to copy'

	고추	/kotɕʰu/	'red pepper'
	구기다	/kugida/	'to crumple'
	부서지다	/pusədʑida/	'to break'
	두껍다	/tuk*əpt*a/	'to be thick'
	불리다	/pullida/	'to steep (in water)'
Lax onsets (tensification unattested)	번개	/pəngɛ/	'lightning'
	베다	/peda/	'to cut'
	부르다	/purɯda/	'to call'
	다치다	/tatɕʰida/	'to injure'
	가방	/kabaŋ/	'bag'
	다리미	/tarimi/	'(clothes) iron'
	돌다	/tolda/	'to turn'
	다리	/tari/	'leg'
	바지	/padʑi/	'pants'
	고기	/kogi/	'meat'
	가누다	/kanuda/	'to balance'
	구르다	/kurɯda/	'to roll'
Established tense onsets (tensification *universal*)	뚜껑	/t*uk*əŋ/	'lid'
	꽃	/k*ot/	'flower'

	따다	/t*ada/	'to pluck'
	깎다	/k*akta/	'to peel'
	꼬리	/k*ori/	'tail'
	뿌리	/p*uri/	'root'
	뺨	/p*jam/	'cheek'
	빠르다	/p*arɯda/	'to be fast'
	때리다	/t*ɛrida/	'to hit'
	까맣다	/k*amatʰa/	'to be black'
	뚫다	/t*ultʰa/	'to pierce'
	뿌리다	/p*urida/	'to scatter'
Distractors (words without lax or tense stop onsets)	피다	/pʰida/	'to bloom'
	타다	/tʰada/	'to ride'
	머리	/məri/	'head'
	세다	/seda/	'to count'
	차리다	/tɕʰarida/	'to prepare'
	내리다	/nerida/	'to descend'
	젖다	/tɕətt*a/	'to be wet'
	풀다	/pʰulda/	'to untie'
	켜다	/kʰjəda/	'to turn on'

	코	/kʰo/	'nose'
	털다	/tʰəlda/	'to dust'
	마르다	/marɯda/	'to dry'
	새다	/seda/	'to leak'
	높다	/nopt*a/	'to be high'
	지다	/tɕida/	'to fall'
	울다	/ulda/	'to cry'
	차다	/tɕʰada/	'to kick'
	읽다	/ikt*a/	'to read'

APPENDIX C: RAW RESULTS BY TOKEN FOR THE PRODUCTION TASK.

Table 24. Raw results by token for the production task.

Target Word	IPA	Tokens	Lax (all raters)	Lax (2 raters), Tense (1 rater)	Tense (2 raters), Lax (1 rater)	Tense (all raters)
동그라미	/toŋgɯrami/	30	28	2	0	0
번데기	/pəndegi/	30	14	2	3	11
닦다	/takt*a/	15	10	1	2	2
당기다	/taŋgida/	8	8	0	0	0
구부리다	/kuburida/	8	7	0	1	0
감다	/ka:mt*a/	30	27	2	1	0
곶감	/kotɕ*am/	17	7	3	6	1
밖	/pak/	1	0	1	0	0
둥글다	/tuŋgɯlda/	17	14	3	0	0
볶다	/pokt*a/	22	19	3	0	0
두드리다	/tudɯrida/	26	17	7	2	0
부수다	/pusuda/	22	17	5	0	0
가시	/kaɕi/	30	28	1	0	1
둑	/tuk/	15	3	1	2	9
굽다	/kupt*a/	30	23	6	1	0
베끼다	/pek*ida/	26	19	6	1	0
고추	/kotɕʰu/	27	16	4	5	2
구기다	/kugida/	28	19	7	1	1
부서지다	/pusədʑida/	8	6	1	1	0

두껍다	/tuk*əpt*a/	5	3	2	0	0
불리다	/pullida/	27	18	3	3	3
Totals:		462	321	60	35	46

APPENDIX D: RAW RESULTS FOR CONTROL ITEMS IN THE RATING TASK.

Table 25. Raw results for control items in the rating task.

Target Word	IPA	Tokens	Lax (all raters)	Lax (2 raters), Tense (1 rater)	Tense (2 raters), Lax (1 rater)	Tense (all raters)
바지	/padʑi/	10	10	0	0	0
가방	/kabaŋ/	10	8	0	1	1
달력	/talljək/	10	9	0	1	0
곳	/kot/	10	10	0	0	0
방	/paŋ/	10	10	0	0	0
Totals:		*50*	*47*	*0*	*2*	*1*
꼬리	/k*ori/	10	0	0	4	6
땅	/t*aŋ/	10	0	0	1	9
꺼내다	/k*əneda/	10	0	0	3	7
빨갛다	/p*algatʰa/	10	1	1	0	8
떨어지다	/t*ərədʑida/	10	0	0	0	10
Totals:		*50*	*1*	*1*	*8*	*40*

APPENDIX E: PARTICIPANT QUESTIONNAIRES ISSUED BEFORE THE ACCEPTABILITY JUDGMENT TASK.

1. What year were you born?

2. Sex (F/M)

3. *However little,* do you believe you use a regional dialect in daily life?
   3a. If *Yes,* which dialect do you believe you use the most?

   3b. If *Yes,* please indicate how often you use dialect in the following contexts using the scales provided. (Likert scale for each item, from 1 *never* to 5 *always*:)

   - ●When speaking with family members.
   - ●When texting with family members.
   - ●When speaking with friends from my hometown or home region.
   - ●When texting with friends from my hometown or home region.
   - ●When speaking with friends from outside my hometown or home region.
   - ●When texting with friends from outside my hometown or home region.
   - ●In the workplace.
   - ●When speaking with strangers I meet in everyday situations.
   - ●When speaking with foreigners.
   - ●When writing on the internet or on social media.

   3c. If *No,* please indicate the place where you lived the longest while growing up.

Attitudes towards Onset Tensification:

4. The following are statements about these pronunciations. Please indicate on the scale provided how much you agree with each statement. (Likert scale for each item, from 1 *firmly disagree* to 5 *firmly agree*). ***Tensified onset pronunciations are…***
   a. Masculine
   b. Vulgar
   c. Easier to pronounce
   d. Better at communicating what I mean
   e. More expressive
   f. Awkward
   g. Emphatic

h. Friendlier
i. Formal
j. Used more by younger people
k. Negative
l. Used by me
m. Feminine
n. Cute
o. Positive

5. People from my dialect use these pronunciations:
   a. YES / NO
   b. *If NO,* do you associate these pronunciations with a certain dialect? Write the dialect in the space provided.

Figure 4. Participant questionnaires issued before the acceptability judgment task.

APPENDIX F: STIMULI FOR THE ACCEPTABILITY JUDGMENT TASK.

Table 26. Stimuli for the acceptability judgment task.

	Word	IPA	English translation
"Tensified" onsets **(lax → tense attested)** *(Target words from the first experiment)*	동그라미	/toŋgurami/	'circle'
	번데기	/pəndegi/	'pupa, chrysalis'
	닦다	/takt*a/	'to wipe'
	당기다	/taŋgida/	'to pull'
	구부리다	/kuburida/	'to stoop over'
	감다	/ka:mt*a/	'to wash (hair)'
	곶감	/kotk*am/	'dried persimmon'
	밖	/pak/	'outside'
	둥글다	/tuŋgɯlda/	'to be round'
	볶다	/pokt*a/	'to fry'
	두드리다	/tudɯrida/	'to knock (on)'
	부수다	/pusuda/	'to smash'
	가시	/kaɕi/	'thorn'
	둑	/tuk/	'dyke, levee'
	굽다	/kupt*a/	'to roast'

	베끼다	/pek*ida/	'to copy'
	고추	/kotɕʰu/	'red pepper'
	구기다	/kugida/	'to crumple'
	부서지다	/pusədʑida/	'to break'
	두껍다	/tuk*əpt*a/	'to be thick'
	불리다	/pullida/	'to steep (in water)'
"Tensified" onsets **(lax → tense attested in Han, 2013 and Jang, 2017)**	주꾸미	/tʃuk*umi/	"webfoot octopus"
	곰장어	/komdʒaŋə/	"sea eel"
	숙맥	/suŋmɛk/	"fool"
	골초	/koltʃho/	"chain smoker"
	족집게	/tʃoktʃipke/	"tweezers"
	소나기	/sonagi/	"rain shower"
	공짜	/koɲtʃ*a/	"free, no cost"
	감방	/kambaŋ/	"prison"
	수세미	/susemi/	"loofah"
	소주	/sodʒu/	"soju (a type of distilled spirits)"
	거꾸로	/kək*uro/	"backwards"
	다르다	/tarɯda/	"to be different"
	고물	/komul/	"junk"
	생맥주	/sɛŋmɛktʃu/	"draft beer"
	쇠	/swe/	"iron (metal)"

	족제비	/tʃoktʃebi/	"weasel (animal)"
	진하다	/tʃinhada/	"to be thick, deep"
	자르다	/tʃarɯda/	"to cut"
	작다	/tʃakt*a/	"to be small"
	질기다	/tʃilkida/	"to be tough"
	좁다	/tʃopt*a/	"to be narrow"
	고소하다	/kosohada/	"to be savory"
	줄다	/tʃulda/	"to lessen"
	세다	/seda/	"to be strong"
	세련되다	/serjəndweda/	"to be refined"
	중국	/tʃuŋguk/	"China"
	권투	/kwəntu/	"boxing"
	공돈	/koŋton/	"easy money"
	사이비	/saibi/	"pseudo-"
	사모님	/samonim/	"Madam (address)"
	조금	/tʃogɯm/	"a little"
	조끼	/tʃok*i/	"vest"
	갓난아이	/kannanai/	"newborn baby"
	강철	/kaŋtʃhəl/	"steel"
	개구리	/kɛguri/	"frog"
	도랑	/toraŋ/	"ditch"
	두꺼비	/tuk*əbi/	"toad"

	붉어지다	/pulkədʒida/	"to redden"
	삯	/sak/	"wages"
	수수	/susu/	"sorghum"
	시다	/ʃida/	"to be sour"
	지린내	/tʃirinnɛ/	"stench of urine"
	계속	/kjesok/	"continuing"
	다듬다	/tadɯmta/	"to trim"
	졸병	/tʃolpjəŋ/	"army private"
	삶다	/samt*a/	"to boil"
Controls (words with lax onsets where tense forms are unattested; words realized universally with tense onsets)	번개	/pəngɛ/	'lightning'
	다리	/tari/	'leg'
	바지	/padʑi/	'pants'
	고기	/kogi/	'meat'
	가방	/kabaŋ/	'bag'
	깎다	/k*akta/	'to peel'
	꼬리	/k*ori/	'tail'
	때리다	/t*ɛrida/	'to hit'
	뺨	/p*jam/	'cheek'
	땅	/t*aŋ/	'soil'
	효과	/hjogwa/	"effect"

Words where tensification is attested medially in Han & Park (2012)	김밥	/kimbap/	"rice rolled in seaweed"
	신기다	/ʃingida/	"to put footwear on (someone)"
	간단하다	/kandanhada/	"to be simple"
Words where weakening is attested medially in Han & Park (2012)	강비탈	/kaŋ.p*ital/	"river slope"
	인감도장	/ingam.t*odʑaŋ/	"registered seal"
	밤새	/pam.s*ɛ/	"all night long"
	안간힘	/an.k*anhim/	"straining, restraint"
	신뒤축	/ʃin.t*witɕʰuk/	"shoe heel"
	발자취	/pal.tɕ*atɕʰwi/	"footprint"
	손도끼	/son.t*ok*i/	"hand-axe"
	가을장마	/kaɯl.tɕ*aŋma/	"autumn rainy spell"
	해장술	/hɛdʑaŋ.s*ul/	"hangover-chaser, hair-of-the-dog"
	물갈래	/mul.k*allɛ/	"river fork"
	눈정신	/nun.tɕ*əŋʃin/	"perspicacity" [lit. eye spirit]
	눈송이	/nun.s*oŋi/	"snowflake"
	발뒤꿈치	/pal.t*wik*umtɕʰi/	"(foot) heel"
	발그림자	/pal.k*ɯrimtɕa/	"footprint, trace"
	어림수	/ərim.s*u/	"rough estimate"

	인기척	/in.k*itɕʰək/	"presence, feeling of a person being around"
	한통속	/hantʰoŋ.s*ok/	"party, gang, co-conspirators"
	쌀집	/s*al.tɕ*ip/	"rice store"
	길벗	/kil.p*ət/	"travelling companion"

Table 27. Results of the acceptability judgment task.

Word	IPA pronunciation	Natural (%)	Unnatural (%)
졸병 /tʃolpjəŋ/ '(army) private'	/tʃolpjəŋ/	20.7	79.3
	/tʃ*olpjəŋ/	100	0
주꾸미 /tʃuk*umi/ 'webfoot octopus'	/tʃuk*umi/	26.7	73.3
	/tʃ*uk*umi/	96.7	3.3
숙맥 /suŋmɛk/ 'fool'	/suŋmɛk/	44.8	55.2
	/s*uŋmɛk/	96.6	3.4
세다 /seda/ 'to be strong'	/seda/	63.3	36.7
	/s*eda/	93.3	6.7
곰장어 /komdʒaŋə/ 'sea eel'	/komdʒaŋə/	33.3	66.7
	/k*omdʒaŋə/	90	10
공돈 /koŋton/ 'easy money'	/koŋton/	30	70
	/k*oŋton/	90	10
족집게 /tʃoktʃipke/ 'tweezers'	/tʃoktʃipke/	51.7	48.3
	/tʃ*oktʃipke/	89.7	10.3
지린내 /tʃirinnɛ/ 'stench of urine'	/tʃirinnɛ/	70	30
	/tʃ*irinnɛ/	86.7	13.3
	/koltʃʰo/	63.3	36.7

골초 /koltʃʰo/ 'chain smoker'	/k*oltʃʰo/	83.3	16.7
번데기 /pəndegi/ 'pupa, chrysalis'	/pəndegi/	65.5	34.5
	/p*əndegi/	82.8	17.2
진하다 /tʃinhada/ 'to be thick, deep'	/tʃinhada/	82.8	17.2
	/tʃ*inhada/	82.8	17.2
자르다 /tʃarɯda/ 'to cut'	/tʃarɯda/	83.3	16.7
	/tʃ*arɯda/	76.7	23.3
구부리다 /kuburida/ 'to bend over'	/kuburida/	83.3	16.7
	/k*uburida/	73.3	26.7
둑 /tuk/ 'embankment'	/tuk/	66.7	33.3
	/t*uk/	73.3	26.7
구기다 /kugida/ 'to wrinkle'	/kugida/	86.7	13.3
	/k*ugida/	73.3	26.7
공짜 /koŋtʃ*a/ 'free, no cost'	/koŋtʃ*a/	83.3	16.7
	/k*oŋtʃ*a/	73.3	26.7
감방 /kambaŋ/ 'prison'	/kambaŋ/	73.3	26.7
	/k*ambaŋ/	73.3	26.7
	/sodʒu/	70	30

소주 /soʤu/ 'soju (a type of distilled spirits)'	/s*oʤu/	73.3	26.7
고소하다 /kosohada/ 'to be savory'	/kosohada/	83.3	16.7
	/k*osohada/	73.3	26.7
불리다 /pullida/ 'to steep, soak'	/pullida/	86.7	13.3
	/p*ullida/	70	30
조금 /tʃogɯm/ 'a little'	/tʃogɯm/	86.7	13.3
	/tʃ*ogɯm/	70	30
세련되다 /serjəndweda/ 'to be refined'	/serjəndweda/	70	30
	/s*erjəndweda/	70	30
도랑 /toraŋ/ 'ditch'	/toraŋ/	80	20
	/t*oraŋ/	70	30
거꾸로 /kək*uro/ 'backwards'	/kək*uro/	93.1	6.9
	/k*ək*uro/	69	31
곶감 /kotkam/ 'dried persimmon'	/kotkam/	79.3	20.7
	/k*otkam/	69	31
감다 /kamt*a/ 'to wash'	/kamt*a/	86.7	13.3
	/k*amt*a/	66.7	33.3

조끼 /tʃok*i/ 'vest'	/tʃok*i/	90	10
	/tʃ*ok*i/	66.7	33.3
사이비 /saibi/ 'pseudo-'	/saibi/	70	30
	/s*aibi/	66.7	33.3
족제비 /tʃoktʃebi/ 'weasel (animal)'	/tʃoktʃebi/	86.7	13.3
	/tʃ*oktʃebi/	66.7	33.3
수세미 /susemi/ 'loofah'	/susemi/	76.7	23.3
	/s*usemi/	66.7	33.3
좁다 /tʃopt*a/ 'to be narrow'	/tʃopt*a/	93.3	6.7
	/tʃ*opt*a/	63.3	36.7
줄다 /tʃulda/ 'to lessen'	/tʃulda/	86.7	13.3
	/tʃ*ulda/	63.3	36.7
고추 /kotʃʰu/ 'chili pepper'	/kotʃʰu/	93.3	6.7
	/k*otʃʰu/	63.3	36.7
가시 /kaʃi/ 'thorn'	/kaʃi/	90	10
	/k*aʃi/	63.3	36.7
닦다 /takta/ 'to wipe'	/takta/	86.7	13.3
	/t*akta/	63.3	36.7
	/sonagi/	82.8	17.2

소나기 /sonagi/ 'rain shower'	/s*onagi/	62.1	37.9
당기다 /taŋkida/ 'to pull'	/taŋkida/	96.7	3.3
	/t*aŋkida/	60	40
생맥주 /sɛŋmɛktʃu/ 'draft beer'	/sɛŋmɛktʃu/	80	20
	/s*ɛŋmɛktʃu/	60	40
쇠 /swe/ 'iron (metal)'	/swe/	80	20
	/s*we/	60	40
질기다 /tʃilkida/ 'to be tough'	/tʃilkida/	93.3	6.7
	/tʃ*ilkida/	60	40
두드리다 /tutɯrida/ 'to knock'	/tutɯrida/	96.7	3.3
	/t*utɯrida/	56.7	43.3
작다 /tʃakt*a/ 'to be small'	/tʃakt*a/	90	10
	/tʃ*akt*a/	56.7	43.3
동그라미 /toŋkɯrami/ 'circle'	/toŋkɯrami/	96.7	3.3
	/t*oŋkɯrami/	56.7	43.3
삶다 /samt*a/ 'to boil'	/samt*a/	80	20
	/s*amt*a/	53.3	46.7
	/pokt*a/	90	10

볶다 /pokt*a/ 'to fry'	/p*okt*a/	53.3	46.7
베끼다 /pek*ida/ 'to copy'	/pek*ida/	86.2	13.8
	/p*ek*ida/	51.7	48.3
부수다 /pusuda/ 'to smash'	/pusuda/	93.3	6.7
	/p*usuda/	50	50
부서지다 /pusədʒida/ 'to be broken'	/pusədʒida/	90	10
	/p*usədʒida/	50	50
굽다 /kupt*a/ 'to roast'	/kupt*a/	89.7	10.3
	/k*upt*a/	44.8	55.2
두껍다 /tuk*əpt*a/ 'to be thick'	/tuk*əpt*a/	96.6	3.4
	/t*uk*əpt*a/	44.8	55.2
사모님 /samonim/ 'Madam (address)'	/samonim/	100	0
	/s*amonim/	43.3	56.7
갓난아이 /kannanai/ 'newborn baby'	/kannanai/	90	10
	/k*annanai/	40	60
중국 /tʃuŋguk/ 'China'	/tʃuŋguk/	92.9	7.1
	/tʃ*uŋguk/	39.3	60.7
	/sak/	96.6	3.4

삯 /sak/ 'wages'	/s*ak/	37.9	62.1
개구리 /kɛguri/ 'frog'	/kɛguri/	100	0
	/k*ɛguri/	26.7	73.3
둥글다 /tuŋkɯlta/ 'to be round'	/tuŋkɯlta/	100	0
	/t*uŋkɯlta/	23.3	76.7
고물 /komul/ 'junk'	/komul/	96.7	3.3
	/k*omul/	23.3	76.7
권투 /kwəntu/ 'boxing'	/kwəntu/	96.7	3.3
	/k*wəntu/	20	80
다른 /tarɯn/ 'different	/tarɯn/	100	0
	/t*arɯn/	20	80
두꺼비 /tuk*əbi/ 'toad'	/tuk*əbi/	100	0
	/t*uk*əbi/	16.7	83.3
수수 /susu/ 'sorghum'	/susu/	100	0
	/su*su/	13.3	86.7
계속 /kjesok/ 'continuing'	/kjesok/	100	0
	/k*jesok/	13.3	86.7
	/ʃida/	93.3	6.7

시다 /ʃida/ 'to be sour'	/ʃ*ida/	10	90
다듬다 /tadɯmta/ 'to trim'	/tadɯmta/	100	0
	/t*adɯmta/	10	90
강철 /kaŋtʃʰəl/ 'steel'	/kaŋtʃʰəl/	100	0
	/k*aŋtʃʰəl/	6.7	93.3
붉어지다 /pulkədʒida/ 'to redden'	/pulkədʒida/	100	0
	/p*ulkədʒida/	6.7	93.3
밖 /pak/ 'outside'	/pak/	100	0
	/p*ak/	0	100

APPENDIX G: RESULTS FOR CONTROL ITEMS IN THE ACCEPTABILITY JUDGMENT TASK.

Table 28: Results for control items in the acceptability judgment task.

Target Word	IPA	N	Tense Onsets Accepted (N)	% Tense
가방	/kabaŋ/	30	0	0.0
곳	/kot/	30	0	0.0
달력	/talljək/	30	2	6.7
바지	/padʑi/	30	5	16.7
방	/paŋ/	30	0	0.0
꼬리	/k*ori/	30	26	86.7
꺼내다	/k*əneda/	30	27	90.0
땅	/t*aŋ/	30	29	96.7
떨어지다	/t*ərədʑida/	30	30	100.0
빨갛다	/p*algatha/	30	25	83.3

APPENDIX H: RESULTS FOR ITEMS FEATURING WORD-MEDIAL TENSE SEGMENTS IN THE ACCEPTABILITY JUDGMENT TASK.

Table 29. Results for items featuring word-medial tense segments in the acceptability judgment task.

Word	*IPA pronunciation*	*Natural (%)*	*Unnatural (%)*
강비탈 /kaŋ.p*ital/ "river slope"	/kaŋ.pital/	100	0
	/kaŋ.p*ital/	3.6	96.4
인감도장 /ingam.t*odʑaŋ/ "registered seal"	/ingam.todʑaŋ/	98.8	1.2
	/ingam.t*odʑaŋ/	3.6	96.4
밤새 /pam.s*ɛ/ "all night long"	/pam.sɛ/	100	0
	/pam.s*ɛ/	0	100
안간힘 /an.k*anhim/ "straining, restraint"	/an.kanhim/	98.8	1.2
	/an.k*anhim/	2.4	97.6
신뒤축 /ʃin.t*witɕʰuk/ "shoe heel"	/ʃin.twitɕʰuk/	96.4	3.6
	/ʃin.t*witɕʰuk/	7.1	92.9
발자취 /pal.tɕ*atɕʰwi/ "footprint"	/pal.tɕatɕʰwi/	98.8	1.2
	/pal.tɕ*atɕʰwi/	2.4	97.6
손도끼 /son.t*ok*i/ "hand-axe"	/son.tok*i/	100	0
	/son.t*ok*i/	1.2	98.8

가을장마 /kaɯl.tɕ*aŋma/ "autumn rainy spell"	/kaɯl.tɕaŋma/	97.6	2.4
	/kaɯl.tɕ*aŋma/	3.6	96.4
해장술 /hɛdʑaŋ.s*ul/ "hangover-chaser, hair-of-the-dog"	/hɛdʑaŋ.sul/	97.6	2.4
	/hɛdʑaŋ.s*ul/	3.6	96.4
물갈래 /mul.k*allɛ/ "river fork"	/mul.kallɛ/	98.8	1.2
	/mul.k*allɛ/	2.5	97.5
눈정신 /nun.tɕ*əŋʃin/ "perspicacity" [lit. *eye spirit*]	/nun.tɕəŋʃin/	97.5	2.5
	/nun.tɕ*əŋʃin/	3.7	96.3
눈송이 /nun.s*oŋi/ "snowflake"	/nun.soŋi/	94	6
	/nun.s*oŋi/	16.7	83.3
발뒤꿈치 /pal.t*wik*umtɕʰi/ "(foot) heel"	/pal.twik*umtɕʰi/	95.2	4.8
	/pal.t*wik*umtɕʰi/	10.8	89.2
발그림자 /pal.k*ɯrimtɕa/ "footprint, trace"	/pal.kɯrimtɕa/	97.6	2.4
	/pal.k*ɯrimtɕa/	6	94
어림수 /ərim.s*u/ "rough estimate"	/ərim.su/	95.1	4.9
	/ərim.s*u/	6.2	93.8
	/in.kitɕʰək/	94	6

인기척 /in.k*iɕʰək/ "presence, feeling of a person being around"	/in.k*iɕʰək/	17.9	82.1
한통속 /hantʰoŋ.s*ok/ "party, gang, co-conspirators"	/hantʰoŋ.sok/	93.9	6.1
	/hantʰoŋ.s*ok/	15.9	84.1
쌀집 /s*al.tɕ*ip/ "rice store"	/s*al.tɕip/	73.5	26.5
	/s*al.tɕ*ip/	33.7	66.3
길벗 /kil.p*ət/ "travelling companion"	/kil.pət/	94	6
	/kil.p*ət/	8.3	91.7

REFERENCES

Ahn, C. G., & Lee, K.-I. (1996). Clinical application of the revised NEO Personality Inventory. *The Korean Journal of Clinical Psychology*, *8*(1), 65-79.

Bae, J. (2003). 한국어의 발음. *[The pronunciation of Korean.]* Seoul, South Korea: 삼경문화사. [Samgyeong Munhwasa.]

Baxter, W. H. (1992). *A handbook of Old Chinese phonology*. Berlin: Mouton de Gruyter.

Boersma, P. (1998). *Functional phonology: formalizing the interactions between articulatory and perceptual drives.* The Hague: Holland Academic Graphics.

Busch, D. (1982). Introversion-extraversion and the EFL proficiency of Japanese students. *Language Learning, 32*(1), 109–132.

Carrell, P. L., Prince, M. S., & Astika, G. G. (1996). Personality types and language learning in an EFL context. *Language Learning, 46*(1), 75–99.

Cho, T., Jun, S.-A., and Ladefoged, P. (2002). Acoustic and aerodynamic correlates of Korean stops and fricatives. *Journal of Phonetics, 30*, 193–228.

Cohen, P, Cohen, J, Aiken, L. S., and West, S. G. (1999) The problem of units and the circumstance for POMP. *Multivariate behavioral research, 34*(3), 315-46.

Costa, P. T, Jr., & McCrae, R. R. (1985). *The NEO personality inventory manual.* Odessa, FL: Psychological assessment resources.

Dewaele, J-M., and Furnham, A. (1999). Extraversion: the unloved variable in applied linguistic research. *Language learning, 49*(3), 509-44.

Dewaele, J-M, and Furnham, A. (2000) Personality and speech production: a pilot study of second language learners. *Personality and individual differences, 28*(2), 355-65.

Flemming, E. (1997). Phonetic optimization: compromise in speech production. *University of Maryland Working Papers in Linguistics 5: Selected phonology papers from H-OT-97*.

Fukada, A. (2013). An online oral practice/assessment platform: Speak Everywhere. *The IALLT journal, 43*(1), 64-77.

Hallgren, K. A. (2012). Computing inter-rater reliability for observational data: an overview and tutorial. *Tutorials in quantitative methods for psychology, 8*(1), 23–34.

Han, M. (2011). College students' word-initial glottalization and perception. 겨레어문학 *[Journal of Korean National Language and Literature,] 46,* 329-370.

Han, M., and Park, D. (2012). A study on the weakening phenomenon of word-mid glottalization reflected in modern Korean. 한말연구 *[Korean Language Research,] 30,* 351-386.

Han, M. (2013). A basic research on the numerical development for word-initial glottalization. 한말연구 *[Korean Language Research,] 32*, 394-424.

Han, M. (2015). A study on the word-initial glottalization of Sino-Korean words in modern Korean language. 우리말연구 *[Urimal research,] 41*, 62-94.

Hinton, M. (2014). The bold and the beautiful: How aspects of personality affect foreign language pronunciation. *Research in language, 12*(3), 217–232.

Hong, M. (2014). A study of phonological variation and language attitude of word-initial tensification. *The Sociolinguistic Journal of Korea, 22*(1), 281-307.

Jakobson, R. (1990). *On language* (Waugh, L. R. and Monville-Burston, M., Eds.). Cambridge, MA: Harvard University Press.

Jang, S. (2017). *현대 한국어의 어두 경음화 연구. [Word-onset tensification in Modern Korean].* Unpublished dissertation. Seoul National University, Seoul, South Korea.

Jeon, J. (2016). *방송언어의 서구 차용어 고찰 [A study on loanwords from European languages in broadcasting].* Unpublished dissertation. Kangwon National University, Chuncheon, South Korea.

Jeon, J. (2019). 경음화에 대한 언어태도의 사회언어학적 조사 및 분석 [A study on language attitudes about glottalization pronunciation]. *어문론집 [Journal of language and literature], 77,* 341-74.

Johnson, K (2011). *Acoustic and auditory phonetics.* Chichester: Wiley-Blackwell.

Kang, J.-S. (1995) *다시 읽는 하멜표류기 [Revisiting Hamel's shipwreck diary.]* Seoul, South Korea: 웅진닷컴. [Woongjin.]

Kang, Y. (2008). Tensification of voiced stops in English loanwords in Korean. *Harvard Studies in Korean Linguistics, XII.*

Kim, S.-K. (2001). 음성 언어 층위와 문자 언어 층위의 위상에 대한 연구 [Strategies of Spoken and Written Language]. *언어학 [Linguistics], 30*, 65-88.

Kim, S. (2016). *Phonological trends in Seoul Korean compound tensification.* Proceedings of the Annual Meeting on Phonology 2016.

Kim, Y.-G. (2004). 경음화 현상에 대한 사회언어학적 연구 [Sociolinguistic research on tensification phenomena]. *겨레어문학 [Journal of Korean National Language and Literature], 32*, 1-23.

Kim, J.-H., Park, Y.-S., Park, D.-G., Kim, T.-G., and Kim, J.-H. (2008) *한국어의 역사. [History of the Korean language.]* Paju, South Korea: 집문당. [Jipmoondang.]

Landis, J. R., and Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics, 33*(1), 159.

Lee, H.-Y. (2009). 서울 토박이들의 경음화 선호도 [Tensification preference of native Seoul speakers of Korean.] *말소리와 음성과학 [Journal of the Korean society of speech sciences], 1*(2), 151-162.

Lee, I., and Ramsey, S. R. (2000). *The Korean language.* Albany: State University of New York.

Lee, K.-M. (1968). 鷄林類事의 再檢討: 주로 音韻史의 觀點에서. [Notes on the Chih-lin Lei-shih]. *東亞文化 [Eastern Culture], 8,* 205-248.

Lee, K.-M., and Ramsey, S. R. (2011). *A history of the Korean language.* Cambridge: Cambridge University Press.

Lee, J.-H. (1999). 한국 사회계층별 언어 특성에 관한 연구 [Research on characteristics of Korean class idiolect]. *사회언어학 [Sociolinguistics], 7*(1), 51-76.

Lee, J.-H. (2002). 서울 방언에 대한 연구 [Seoul dialect research]. *국어 교육 [Korean language education], 107*, 283-318.

Lee, M.-J. (1996). *언어 변화에 관한 사회 언어학적 연구 [A sociolinguistic investigation of language change].* Unpublished dissertation. Seoul National University, Seoul, South Korea.

Lee, P. (2003). *A history of Korean literature*. Cambridge: Cambridge University Press.

MacIntyre, P. D., and Charos, C. (1996). Personality, attitudes, and affect as predictors of second language communication. *Journal of language and social psychology, 15*(1), 3-26.

Macklin-Cordes, J. L. and Round, E. R. (2020). Re-evaluating phoneme frequencies. *Frontiers in psychology*, *11*:570895.

Maddieson, I. (2016). Word length is (in part) predicted by phoneme inventory size and syllable structure. *The Journal of the acoustical society of America, 139*(4), 2218.

Martin, S. E. (1966). Lexical evidence relating Korean to Japanese. *Language, 42*, 185–251.

McKenna, V. S., & Stepp, C. E. (2018). The relationship between acoustical and perceptual measures of vocal effort. The Journal of the Acoustical Society of America, 144(3), 1643.

Miller, R. A. (1980). *Origins of the Japanese language: lectures in Japan during the academic year, 1977-78.* Seattle: University of Washington Press.

Oh, S.-N. (2006). 현대 국어의 형태음운론적 변이 현상에 대한 사회언어학적 연구 *[A sociolinguistic study on the morphological variation of the Seoul Metropolitan urban dialect.]* Unpublished dissertation. Korea University, Seoul, South Korea.

Oh, J.-G. (2011). 국어 방언에 반영된 어두경음화 [The glottalization of initial consonants reflected in Korean dialects]. 한민족어문학*[Language and literature of the Koreans], 58*, 239-71.

Oh, M. and Yang, H. (2013). The production of stops by Seoul and Yanbian Korean speakers. *Journal of the Korean Society of Speech Sciences, 5*(4), 185-193.

Oya, T., Manalo, E., and Greenwood, J. (2004). The influence of personality and anxiety on the oral performance of Japanese speakers of English. *Applied cognitive psychology, 18*(7), 841-55.

Park, C.-J. (2003). 大明律直解의 構文 研究 *[A Study on Sentence Structure in Daemyongryuljikhae].* Unpublished dissertation. Sogang University, Seoul, South Korea.

Park, D. (2000). Word-initial tensification in Seoul Korean. 언어학 *[Linguistics,] 27*, 179-200.

Piantadosi, S. (2014). Zipf's word frequency law in natural language: A critical review and future directions. *Psychonomic Bulletin & Review, 21*(5), 1112-1130.

Shin, J. (2008). Phoneme and syllable frequencies of Korean based on the analysis of spontaneous speech data. *Korean Journal of Communication Disorders, 13,* 193-215.

Shin, J. (2011). 한국어의 말소리. *[Korean phonetics and phonology.]* Seoul, South Korea: 박이정. [Pagijong Press.]

Shin, S. and Davis, S. (2008). Where have all the lax stops gone? On the possible restructuring of the Korean stop system. In *Japanese/Korean Linguistics, 13*, 69–80. Stanford: CSLI Publications.

Silva, D. (2006). Acoustic evidence for the emergence of tonal contrast in contemporary Korean. *Phonology, 23*(2), 287-308.

Sohn, H.-M. (1999). *The Korean language.* Cambridge, UK: Cambridge University Press.

문교부 [South Korean Ministry of Education] (1988). *표준어 규정 [Regulations on standard language.] 88 § 2.*

문화체육관광부 [South Korean Ministry of Culture, Sports, and Tourism] (2017). *표준어 규정 [Regulations on standard language.] 2017 § 13.*

Tambovtsev, Y., and Martindale, C. (2007). Phoneme frequencies follow a Yule

distribution. *SKASE Journal of Theoretical Linguistics*, *4*, 1–11.

Wee, J. (2008). 전남방언에 나타난 어두경음화 [The Word-initial tensification of Jeollanamdo dialect]. *한국언어어문학 [Korean language and literature], 65*, 111-135.

Yeon, J. (2012). Korean dialects: a general survey. In *The languages of Japan and Korea*, ed. Tranter, N. Abingdon: Routledge.

You, P.-J. (2006). *서울방언의 음운론 [Phonology of Seoul dialect].* Seoul, South Korea: 월인[Weolin].

Yu, A. C. L., Abrego-Collier, C., and Sonderegger, M. (2013). Phonetic imitation from an individual-difference perspective: subjective attitude, personality and "autistic" traits. *PloS One 8*(9), e74746.

Zárate-Sández, G. (2017). Reexamining L2 accent: How much can personality explain?. *Ilha Do Desterro, 70*(3), 227-244.

Zipf, G. K. (1949). *Human behavior and the principle of least effort*. Cambridge, MA: Addison-Wesley Press.