NORTHWESTERN UNIVERSITY

Neutralization of Coda Obstruents in Korean: Evidence in Production and Perception

A DISSERTATION

SUBMITTED TO THE GRAUDATE SCHOOL
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

for the degree

DOCTOR OF PHILOSOPHY

Field of Linguistics

By

Kyounghee Lee

EVANSTON, ILLINOIS

June 2016

ABSTRACT

Neutralization of Coda Obstruents in Korean: Evidence in Production and Perception

Kyounghee Lee

The aim of this dissertation is to investigate acoustic and perceptual aspects of coda neutralization in Korean. By extending the Kim and Jongman (1996) study that documents complete neutralization of manner contrasts in coda obstruents, I examine all three types of coda neutralization in Korean (i.e., laryngeal, manner, and palatal) in production and perception. The question of whether coda neutralization is phonetically complete or incomplete is addressed by analyzing acoustic correlates of final obstruents. This study also examines the effect of explicit orthographic cues on coda neutralization by eliciting the stimuli in both spontaneous and read speech and comparing their acoustic properties. A subsequent perception study investigates the perceptual consequences of coda neutralization. In a two-alternative forced-choice identification task, native Korean listeners are asked to identify the correct underlying form of neutralized coda obstruents with or without measurable acoustic differences. The production results reveal that laryngeal neutralization is phonetically incomplete, with its effects stronger in the spontaneous speech than in the read speech. The effects of incomplete neutralization are weaker or absent in the manner or palatal neutralization: the pairs with underlying manner distinctions exhibit acoustic differences in a single measure in one or the other task, whereas the pairs subject to palatal neutralization are completely neutralized in terms of all acoustic parameters in the both speech types. In perception, all three types of coda obstruents are completely neutralized. Regardless of the presence of significant acoustic differences in coda obstruents, the listeners are not able to successfully label neutralized words with the correct underlying forms. These

findings suggest that coda obstruents are not always completely neutralized in production, but exhibit complete perceptual neutralization. The asymmetrical patterns of neutralization across the three types of neutralization as well as across production and perception challenge the current exemplar-based approach to neutralization. I suggest that the current account needs to be modified so that it reflects the fact that the effects of neutralization in production and perception can vary depending on the number of categories neutralized into a single output as well as the dimensions of phonological contrasts that are neutralized.

ACKNOWLEDGMENTS

First and foremost, I would like to express my sincere gratitude to my primary advisor, Ann Bradlow, for inspiring and guiding me during the entire course of my graduate studies. Her research first sparked my interest in speech perception and motivated me to investigate the effect of perceptual training on Korean speakers' production and perception of English vowel contrasts at Hanyang University. I even had the opportunity to meet her at Hanyang and receive critical comments on the project, which I believe paved the way for me to continue my research at NU.

During my doctoral study at Northwestern, I learned an enormous amount from her about how to become an independent researcher as a linguist and scientist. Without her thorough guidance on formulating research questions, designing and running experiments, and becoming a good writer, I would never have been able to complete this dissertation. Besides being my academic advisor, she has also been a good mentor. Thanks to her encouragement, patience, support, and faith in me, I was able to overcome many crises and finish this dissertation. Thank you Ann so many ways.

I am deeply thankful to my co-advisor, Matt Goldrick, who has always been there to listen, give advice, and be open for discussion. Every meeting I had with him greatly helped me to organize my thoughts, sharpen them, and write them out in a logical manner. I am especially grateful to him for teaching me the statistical tools required for the analyses of psycholinguistic data. Thanks to his expertise on statistics and his enthusiasm for teaching it, I was able to tackle the many challenges I faced while analyzing the data.

I would also like to thank Sazzad Nasir for serving as a committee member for my dissertation. His questions and comments helped me seriously consider why I became interested in the topic of neutralization and why this topic is worthy of investigation.

My gratitude is extended to the faculty, staff, and lab members in and outside of the Linguistics department at Northwestern. I am thankful to Masaya Yoshida for encouraging and looking out for me during my graduate career. It was great fun to work with him for sentence processing research. Thank you to the members of NU Speech Communication Lab and Sound Lab groups for the fruitful discussions and feedbacks that greatly improved my dissertation. Special thanks to Chun Liang Chan for his technical support that was always crucial in designing experiments. I am also grateful to Eunmi Lee, program coordinator of the Korean language program at Northwestern, for giving me the invaluable opportunity to work as a teaching assistant in her classes.

Outside of Northwestern, I owe my sincere gratitude to my former advisor, Ki-jeong Lee, Professor of English Language and Literature at Hanyang University in Seoul. He was the one who first inspired me to have great interest in phonology and second language acquisition. Even after I graduated, he continued to support and advise me both in academic and non-academic matters whenever needed.

I am greatly indebted to Taehong Cho, Professor of English Language and Literature and Director of the Phonetics and Psycholinguistics Laboratory at Hanyang. He always willingly provided me with access to lab space and equipment whenever I had to conduct experiments in Korea. I cannot imagine how difficult it would have been to make timely progress in my graduate research without his support.

My special thanks go to my high school English teacher and current principal at Seoul Global high school, Nak Hyun Oh. I was very fortunate to learn from him how fun and exciting it was to learn the English language. I am so grateful to him for helping me discover my language aptitude and fostering in me a genuine interest in language learning.

My friends at NU have greatly helped me to keep motivated throughout my graduate studies. Many thanks to Lauren Ackerman, Ann Burchfield, Susanne Brouwer, Angela Fink, Jordana Heller, Lisa Hesterberg, Nayoun Kim, Jenna Luque, Elizabeth Mazzocco, and Charlotte Vaughn. They have made my graduate life at NU special and unforgettable. Special thanks go to Midam Kim, who has been a wonderful supporter in both academic and non-academic matters from the very beginning of my life in the US. My warmest thanks go to my officemate, Julie Matsubara. She has been a great comfort to me as a good friend and counselor, especially during difficult times. I am also indebted to my old friends in Seoul: Hye Yoon Choi, Intack Im, Jung Mae Eun, and Hwa Sun Lee. Their long-term emotional support and friendship helped me overcome many setbacks and survive my lonely journey.

A big thank you goes to my family in Chicago and Seoul. I am grateful to my parents-in-law for their time and effort for caring for my daughter Bomie while I was working. And great thanks to my two brothers, Hyunsung and Kyounghoon, for understanding my decision to study abroad and encouraging me during my doctoral study.

Most importantly, my deepest gratitude goes to my dad in heaven and my mom, to whom this dissertation is dedicated to. Without their love, support, and confidence in me, I would never have become what I am now or achieved one of my lifetime goals. I am very proud that you are my parents and that I am your daughter. Love you so much.

Finally, I have to acknowledge how lucky I am to have the best supporters in my life during my graduate studies: my husband, Jin Park, and my daughter, Bomie Park, truly a blessing from God. They have made my life more beautiful, exciting, and full of love. Thank you both, and I love you with all my heart.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1. INTRODUCTION

A key assumption of traditional generative phonology is that a word is stored in our mental dictionary as an underlying representation (UR), an abstract form consisting of a string of phonemes or phonological features. When we produce the word, its UR is retrieved and submitted to phonological rules that assign the UR with predictable phonetic properties in a particular phonological environment. Phonological structures at the end of the derivations are then converted into graded articulatory gestures in time via phonetic implementation, which results in the actual spoken form of the word or surface representation (SR).

Within this theoretical framework, the complete merger of a phonological contrast between two URs is accounted for by the assumption that various realizations of a morpheme have the same single UR, and that the SR is derived through a phonological rule of neutralization. One typical case of neutralization is final devoicing. In languages such as German, Dutch, and Polish, all obstruents in a syllable-final position are merged into voiceless ones. For example, the morpheme-final obstruents of the German word *Rad* /rad/ (‘wheel’) is produced with [d] in the dative singular form *Rade* [ra:də], but with [t] in the nominative form *Rad* [ra:t] due to the application of the final devoicing rule. A phonetic implementation rule then converts the [-voice] feature of the final consonant into graded gestures in time to realize voiceless obstruents. Consequently, the word *Rad* has the SR [ra:t] which is identical to that of the word *Rat* (‘advice’).

As shown above, the process involved in the production of neutralized words is strictly modular. That is, information flows from abstract representations at lexical level to a single, discrete, categorically specified phonological representation, and then to phonetic representations required for articulatory execution, without each level looking backwards. This implies that at phonetic implementation stage, the resulting phonological representations that have the same feature specifications are physically realized with the same acoustic-phonetic properties. Therefore, the surface forms of neutralized words derived from the two distinct abstract URs through the application of a rule of neutralization are expected to be identical in their acoustic-phonetic properties. In other words, the process of phonological neutralization is phonetically complete.

Over the last four decades, however, the traditional notion of neutralization has been challenged by a number of experimental studies which have revealed that phonological neutralization is not always phonetically complete (Charles-Luce, 1985; Kleber, John, & Harrington, 2010; Piroth & Janker, 2004; Port & Crawford, 1989; Port & O'Dell, 1985; Roettger, Winter, Grawunder, Kirby, & Grice, 2014 for German; Ernestus & Baayen, 2003, 2006, 2007; Warner, Jongman, Sereno, & Kemps, 2004 for Dutch; Dinnsen & Charles-Luce, 1984; Charles-Luce, 1993 for Catalan; Slowiaczek & Dinnsen, 1985 for Polish; Dmitrieva, Jongman, & Sereno, 2010; Kharlamov, 2012, 2014, 2015; Shrager, 2012 for Russian). Concentrating on final devoicing, these studies have revealed that significant acoustic-phonetic differences are still preserved between underlyingly voiced and voiceless final obstruents. The differences are often observed in temporal-acoustic cues such as preceding vowel duration, voicing into closure duration, and release burst duration, with vowel duration being the most reliable correlate of the underlying voicing contrast. In some of the studies, it has been further shown that acoustic differences serve as perceptual cues for native listeners to identify intended voicing.

It should be noted, however, that the robustness of incomplete neutralization has been contested by researchers arguing that incomplete neutralization may be an experimental artifact (Fourakis & Iverson, 1984; Jassem & Richter, 1989; Port & Crawford, 1989; Warner et al., 2006). A major criticism is that the studies supporting incomplete neutralization have elicited stimuli in a reading task in which speakers read aloud a list of target words in isolation or in a carrier sentence. Therefore, it is inevitable that speakers are exposed to orthographic information and thus highly likely to be encouraged to distinguish or hypercorrect pronunciation of final obstruents to reflect the differences in spelling. Similar effects of hyperarticulation can be triggered by speakers' awareness of minimal pairs and the use of infrequent or outdated words as targets as well, as pointed out by Piroth and Janker (2004). Additionally, it has been documented that incomplete neutralization is also affected by talker- or item-related factors such as prior exposure to L2 languages that preserve the final voicing contrast (Dmitrieva et al., 2010; Jassem & Richer, 1989), phonetic environments and sentential contexts (Charles-Luce, 1985; Slowiaczek & Dinnsen 1985; Piroth & Janker, 2004), and the presence or absence of semantic information (Charles-Luce, 1993).

However, there are challenges to these claims. Particularly with regard to the influence of orthography, it has been demonstrated that the voicing contrast of final obstruents is not completely neutralized even when orthographic interference is controlled (Piroth & Janker, 2004; Port & Crawford, 1989; Roettger et al., 2014). In addition, the mixed findings obtained in Catalan (Dinnsen & Charles-Luce, 1984) and Turkish (Kopkalli, 1993) - where the underlying final voicing distinction is not maintained in the orthography - make it even harder to conclude that orthography is the primary or sole cause of incomplete neutralization.

Besides the fact that the findings from the previous studies are inconsistent and contradictory, research on other types of neutralization has been relatively small in its number and its findings are mixed. For example, manner neutralization of coda obstruents in Korean is reported to be complete in production and perception even in the presence of orthographic inputs

(Kim & Jongman, 1996). In Dutch, the contrast between long and short vowels is neutralized due to the lengthening of underlyingly short vowels in open syllables. Lahiri, Schriefers, and Kuijpers (1987) find that long vowels derived by a lengthening rule (e.g., /dal+en/ → [d**a:**len]) are not acoustically different from underlyingly long vowels (e.g., /ba:l + en/ → [b**a:**len])[1]. Braver and Kawahara (2014) show that Japanese vowels in monomoraic nouns are lengthened when they appear in isolation in a prosodic word, which differ in length from underlyingly long vowels. However, it turns out that the vowel length distinction is completely neutralized in number recitation. English flapping has been considered to be a case of incomplete neutralization (Fox & Terbeek, 1977; Huff, 1980; Braver, 2011), although some researchers claimed that flapping is completely neutralizing, at least in some dialects (Joos, 1942; Port, 1976).

Taken together, it remains an open question whether incomplete neutralization observed in languages with final devoicing would hold for other types of neutralization that are qualitatively different from final devoicing. In order to account for neutralization phenomenon in general and relevant phonological theory more precisely, there should be more empirical research in languages in which other types of neutralization occurs.

In this study, I will investigate neutralization of coda obstruents in Korean in order to extend our understanding about acoustic and perceptual aspects of phonological neutralization. By replicating and extending the Kim and Jongman (1996)'s study, I will examine whether all three types of neutralization of coda obstruents in Korean is phonetically complete or incomplete in production and perception. For this purpose, a production experiment will be conducted in which recordings are made by native Korean talkers and then analyzed in terms of temporal

[1] In this study, minimal pairs were not available for comparison. Instead, the vowel and postvocalic consonant in each set of stimuli were kept constant.

acoustic parameters relevant to underlying contrasts of coda obstruents. In doing so, I will also explore whether coda neutralization is influenced by explicit orthographic inputs by manipulating the type of stimuli elicitation tasks (Question & Answer vs. reading aloud) which vary in the presence of orthographic cues. On the basis of the production results, I will investigate the perceptual consequences of neutralized coda obstruents by conducting an identification task. Specifically, it will be examined whether native Korean listeners are able to identify the correct underlying forms of neutralized words that are auditorily presented.

As the first large-scale investigation of three types of coda neutralization in Korean, this work is expected to contribute to establishing phonetic aspects of coda neutralization and their impacts on perception. Considering that Korean is a language qualitatively different from final devoicing languages, the findings of the present study will significantly extend our understanding of whether neutralization is a phenomena specific to certain language groups or could be genuine phenomenon of languages in general. Furthermore, by adding to the growing body of literature addressing phonetic aspects of phonological neutralization, the findings of this study will shed light on whether phonological processes specify features or articulatory gestures in a categorical or a gradient way. Finally, the findings of this study will have significant implications on psycholinguistic theories of speech production and perception.

The dissertation is organized as follows. Chapter 2 reviews empirical findings from the previous literature on final devoicing along with their theoretical implications, followed by the outline of the current study. Chapter 3 presents the production experiment that involves the recordings and acoustic analysis of final obstruents and discusses its results. Chapter 4 presents the perception experiment that examines native Korean listeners' identification of neutralized final obstruents and discusses its results. Chapter 5 provides the results of a series of regression analysis that reveals the relationship between production and perception of neutralized coda obstruents. Chapter 6 summarizes the findings and discusses their implications. Chapter 7 concludes the dissertation and suggest directions for future work.

CHAPTER 2. LITERATURE REVIEW

2.1. Empirical findings on final devoicing

Previous experimental studies on neutralization have focused on final devoicing in languages such as German, Dutch, Catalan, Polish, and Russian. They have examined voicing-dependent temporal acoustic cues in order to reveal whether the underlying voicing contrast of final obstruents is completely neutralized or not. A major finding is that neutralized obstruents with underlyingly distinct forms exhibit significant differences in acoustic properties relevant to voicing contrast. Some of the studies have further demonstrated that those acoustic differences are salient enough for native listeners to perceive. However, these findings have been often questioned on the grounds that incomplete neutralization can be triggered or influenced by task-, talker-, or item-related factors. In the next section, the major empirical findings from the previous studies are reviewed on the basis of which factor drives incomplete neutralization.

2.1.1. Task-related factor: Orthography

The influence of orthography in combination with experimental tasks and items that emphasize orthographic distinctions has been considered as the most influential factor that triggers incomplete neutralization. In languages such as German, Polish, and Dutch, the underlying voicing contrast in word-final position is orthographically preserved (e.g., *Rad* vs. *Rat*). It has been pointed out that the speakers of these languages are highly likely to hypercorrect the pronunciation of the targets to distinguish spelling difference when asked to read words or sentences in a laboratory setting. Additionally, it has been shown that the exposure to minimal pair stimuli also increases speakers' attention to orthographic representations, which often leads to hyperarticulation.

One of the earliest studies that have been criticized in this aspect is Port and O'Dell (1985)[2]. This study examined final devoicing in German by examining 10 talkers' productions of ten minimal pairs containing final stops (/p, b/, /t, d/, /k, g/). Participants read a list of words in isolation, and their productions were analyzed in terms of vowel durations, closure duration, voicing into closure duration, and burst duration. The results showed that there were significant differences in all of the acoustic parameters except closure duration, and that the direction of the differences was consistent with that observed in a non-neutralizing environment. The strongest effect was observed in aspiration duration, which was about 15 ms[3] shorter before underlyingly voiced stops than before voiceless counterparts. For the other two parameters, underlyingly voiced stops were approximately 15 ms longer in vowel duration and 5 ms longer in voicing into closure duration than underlyingly voiceless stops.

The authors also conducted a perception experiment in order to see whether significant acoustic differences affect listeners' identification of the underlying voicing contrast. Ten native German listeners participated in a two-alternative forced-choice (2AFC hereafter) identification task in which they were asked to identify the intended words produced by 10 talkers. The results revealed that the acoustic differences were large enough for listeners to identify the intended voicing of the obstruents with significantly greater than chance accuracy (59%). The multiple regression analyses further showed that vowel duration and aspiration duration served as

---

[2] The findings reported in this study were based on the earlier unpublished works (O'Dell & Port, 1983; Port, Mitleb, & O'Dell, 1981).

[3] One of difficulties in assessing the magnitude of incomplete neutralization is that most of the previous studies have reported the size of durational differences in absolute terms (milliseconds). For a more accurate interpretation, we calculated and added the proportional terms (the percentage of the differences) if absolute duration values were available. We also compared the size of incomplete neutralization to the baseline, i.e., the size of the contrast in a non-neutralizing condition, whenever available.

significantly reliable predictors of listeners' responses. The findings of the perception study suggested that listeners were able to make use of the acoustic differences as cues to identify the intended voicing, thus supporting the hypothesis of incomplete neutralization in perception. Taken together, the authors concluded that final devoicing in German was not completely neutralized either in production or in perception.

However, Fourakis and Iverson (1984) claimed that incomplete neutralization in production found in the previous studies (O'Dell & Port, 1983; Port, Mitleb & O'Dell, 1981) may be an artifact of the experimental task. They pointed out that as the underlying voicing distinction is orthographically represented in the German writing system, participants may have emphasized the spelling distinction in pronunciation if asked to read out written material. They tested this possibility by eliciting stimuli both from a verb conjugation task as well as from a traditional word-list reading task. In the verb conjugation task, four German speakers were orally prompted with the infinitival form of a verb (e.g., *reiten* 'to ride') and were then asked to produce the conjugated forms of which the final obstruent in the second form was neutralized (*reiten*, ***riet***, *greaten*). The measured acoustic cues were vowel duration, closure duration, release burst duration, and aspiration duration, but only the first two measures that were consistent and reliable were included in the analysis. The results showed that there was no significant acoustic difference between the neutralized obstruents elicited in the conjugation task. In the reading task, however, a significant difference was observed in vowel duration, but only for one out of three minimal pairs analyzed. Based on these findings, the authors suggested that incomplete neutralization occurs due to hyperarticulation induced by orthographic inputs, and that the absence of the same result in the oral task implies complete neutralization of final obstruents.

Although this study has been considered as counter-evidence to incomplete neutralization, it has been challenged by methodological criticisms. For example, the use of a small sample of talkers and items invited the criticism that the low statistical power might have prevented the incomplete neutralization effects from being detected (Frick, 1995). Roettger et al. (2014) also pointed out that the absence of evidence for incomplete neutralization should not be interpreted as the evidence for complete neutralization, since the talkers and the items were not sufficiently large enough to disprove the null hypothesis (i.e., complete neutralization). Dinnsen and Charles-Luce (1984) stated that the use of non-minimal pairs (e.g., *buk* 'baked' - *trag* 'carried') could have compounded the results, since the effect of uncontrolled initial consonant on vowel duration was mixed with voicing-dependent durational differences. Besides, Port and Crawford (1989) revealed that their reanalysis of Fourakis and Iverson (1984)'s data yielded significant differences in vowel duration and closure duration.

The possibility that incomplete neutralization is an artifact of orthography was also examined in Port and Crawford (1989). In this study, five German speakers' productions of 3 minimal pairs were elicited in four different tasks. In the first task, subjects were asked to read the sentences containing the target words in both a non-neutralizing and a neutralizing environments (task 1A) and then produce the sentences from memory after listening to the experimenter's sentence productions (task 1B). The other tasks were to read sentences containing the targets (task 2), to read sentences for the experiments to dictate them (task 3), and to read a list of the target words (task 4). Acoustic measurements included vowel duration, closure duration, release burst duration, nasal closure interval, and voicing into closure duration. The data was then analyzed by using both analyses of variance and discriminant analyses. The latter analyses were used to determine which set of variables discriminate between two or more naturally occurring groups. That is, the aim of the analyses was to examine whether the correct identification of underlyingly voiced vs. voiceless obstruents could be predicted from the combination of the measured acoustic parameters.

The analysis of group data showed that burst duration was the only cue differentiating between underlyingly voiced and voiceless obstruents. However, discriminant analyses revealed that the underlying voicing was identified 55% of the time in the reading task (1A) and 56% in the verbal repetition task (1B). For the other tasks, the overall accuracy rate was 63% (task 3), 78% (task 3), and 62% (task 4). The analyses further showed that all acoustic measures except voicing into closure duration contributed to the discrimination of voicing.

This study also conducted a perception experiment in which the productions of two speakers elicited in the reading tasks (3 and 4) were presented to 5 native German listeners in a 2AFC identification task. Listeners' responses converted into d-prime measures were analyzed using discriminant analyses. The listeners' identification performances reached 69.2%, which was comparable to the results of discriminant analyses (70.8%). The combined results suggested that the listeners successfully make use of temporal acoustic cues when asked to identify the neutralized obstruents.

Taken together, it was concluded that underlyingly voiced vs. voiceless obstruents were incompletely neutralized across all conditions of the experiment, with the degree of voicing contrast varying depending on speech styles. More importantly, the authors claimed that underlying voicing distinctions were maintained in productions even when orthographic information was absent. However, Winter and Röttger (2011) speculated that the role of orthography was not completely eliminated in this study. In the verbal repetition task, the experimenter read the target sentences to the participants who were asked to repeat them from memory. Considering that speakers tend to accommodate their speech to those who interact with them (e.g., Harrington, 2006; Hay, Drager, & Warren, 2009), it is possible that speakers' productions of target stimuli were affected by the speech of the experimenter who was exposed to orthography. Consequently, the effect of orthography could have been carried over from the experimenter to the speakers.

A study by Warner et al. (2004) was the first large-scale acoustic/perceptual investigation of final devoicing in Dutch. Importantly, the authors adopted a new approach to the examination of orthographic influence. Instead of using target word pairs that contrast in their underlying forms, they selected the pairs differing in their spelling but representing the same phoneme string. The benefit of this approach was that any duration difference they observe could be attributed to the influence of orthography.

In the first production experiment, fifteen native Dutch talkers read a list of 36 minimal pairs containing phonemically short and long vowels and ending in underlyingly /t/ or /d/ in both a non-neutralizing and neutralizing conditions. The analyzed acoustic parameters included vowel duration, voicing during the closure, closure duration, and burst duration. The results showed clear evidence for incomplete neutralization. Specifically, vowel duration was 3.5 - 4 ms (2% of long V ~ 3% of short V) longer before underlying /d/ than /t/. The magnitude of difference was 14% (long V) - 20% (short V) of the contrast observed in an environment in which the distinction is preserved (22 ms for long V - 20 ms for short V). Even though the difference in vowel duration was smaller than that in other languages such as German and Polish, it was still significant both by subjects and by items. A significant effect of burst duration was also observed, which was longer by 9 ms (7%) for words with underlying /t/ than for words with underlying /d/, but only in words containing long vowels.

In a perception experiment, thirty listeners participated in a 2AFC identification task in which they were presented with the stimuli from 4 talkers in which vowel and burst durations varied. Listeners' word identification accuracy was significantly above the chance level, but only for the two talkers who produced longer vowels before underlying /d/ than before /t/. In order to confirm which perceptual cues listeners rely on for the underlying voicing distinction, the authors manipulated vowel duration and closure duration in subsequent experiments. The results revealed that listeners were able to make use of vowel duration as a cue to underlying final voicing when vowel duration was the only cue that varied. Interestingly, the same pattern of results was obtained with closure duration: longer closure duration received more 't' (voiceless) than 'd' (voiced) responses. This result clearly showed that listeners were able to utilize closure duration as a cue for the voicing distinction, even though this parameter did not vary with the underlying final voicing, as shown in the acoustic analysis. Overall, the results of the perception study suggested that listeners' perception of final devoicing was highly dependent upon individual speakers' production pattern, and that the most reliable acoustic cues to final devoicing were perceptually salient as well.

Finally, in another production experiment to examine the influence of orthography, fifteen Dutch speakers produced minimal pairs that consisted of infinitive verb spelled with VC sequence and its plural tense form spelled with VVCC sequence (e.g., *heten* /hetən/ 'to fool' vs. *heetten* /hetən/ 'they fooled'). For these productions, vowel duration and closure duration were measured. The results showed that vowel durations were significantly shorter by 3 ms (2%) in words with VVCC spelling than in words with VC spelling, but only for words with medial /t/. This result indicated that durational differences do not necessarily stem from differences in underlying forms, but rather from orthographic differences. Based on the findings, the authors suggested that small durational differences should be referred to as 'sub-phonemic durational differences, rather than incomplete neutralization effects' (p. 274).

In a follow-up study by Warner et al. (2006), the authors pointed out that the pairs used in the Warner et al. (2004)'s study could be regarded as having different morpheme strings (e.g., *heten* /he:tən/ 'to fool' vs. *heetten* /he:t-tən/ 'they fooled'). Therefore, they speculated that the observed incomplete neutralization could be due to the contrast in underlying phonemes rather than orthographic influences. In order to further examine whether incomplete neutralization is caused by the abstract underlying difference or by the orthographic difference, they conducted another production experiment in which they used the target pairs that contrast in their underlying morphological forms but have the same orthography. Since the underlying singleton vs. geminate consonant contrast was created by morpheme concatenation in Dutch, the pair such as *ik heet* /he:t/ ('I am called') vs. *hij hee*t /he:t-t/ ('he is called') contrasts in the underlying forms despite the same orthography. Except the stimuli, the experimental designs and the procedures were consistent with the previous study.

The results revealed that there was no significant difference in vowel duration, closure duration, and consonant duration between underlyingly single and geminate consonant conditions. This result confirmed that incomplete neutralization observed in Warner et al. (2004) was due to the differences in orthography. Combined with the previous findings, this study suggested that orthographic differences can lead to incomplete neutralization, but differences in underlying phoneme structure cannot.

More recently, Roettger et al. (2014) replicated the Fourakis and Iverson (1984) study in order to re-examine the role of orthography in the production and perception of final devoicing in German. They used a sufficient number of subjects and items to increase statistical power and selected pseudowords as target stimuli to minimize the effect of orthography. In the production experiment, sixteen speakers listened to the sentences containing the plural form of pseudowords (e.g., [go:bə]) produced by one male talker. They were then asked to produce the singular form ([go:p] in another sentence. The acoustic measurements focused on the duration of preceding vowel.

Unlike in Fourakis and Iverson (1984) in which the neutralized obstruents elicited in the oral conjugation task did not exhibit any difference in vowel duration, vowels were longer by 8.6 ms before underlyingly voiced stops than before underlyingly voiceless stops. However, it was speculated that incomplete neutralization may be due to the effect of phonetic accommodation. Since the participants heard auditory sentence stimuli produced by the male talker, it was possible that they may have shifted their productions of vowel duration in the targets toward those of the talker. The authors tested this possibility in the two follow-up experiments by manipulating vowel duration.

In the first follow-up experiment, the stimuli were classified into four different categories depending on the magnitude of differences in vowel duration: enhanced (32% longer before voiced stops), original (16% longer before voiced stops), neutralized (no difference), and reversed (16% longer before voiceless stops). They were presented in four different blocks in the same production task: all the stimuli were presented in the first two blocks, but only a subset was repeated in blocks 3 and 4. The results revealed a similar but reduced magnitude of differences in vowel duration in the expected directions across all four conditions. That is, participants produced longer vowel duration before underlyingly voiced stops than before underlyingly voiceless stops even in the neutralized and the reversed conditions where the vowel duration cue was not very informative. This suggested that incomplete neutralization was not merely an artifact of speaker accommodation.

The authors further examined another possibility that the stimuli in one condition may have influenced those in the other conditions. In the second follow-up experiment, they presented only the stimuli from the enhanced and the reversed conditions. They failed to find a significant difference in vowel duration in the pooled data, although the magnitude of the difference in the enhanced condition was similar to those observed in the other two experiments.

Finally, in a perception experiment, sixteen listeners heard the sentences produced by 16 talkers that contained the singular target forms. In a 2AFC identification task, they were asked to decide whether the target words have either underlyingly voiced or voiceless stop. The listeners exhibited greater accuracy in correctly identifying underlyingly voiceless stops (58.7%), but around chance level accuracy when identifying voiced stops (51.4%). Overall, the results showed that listeners' identification accuracy was slightly above chance performance (55%) but statistically significant.

This study confirmed that incomplete neutralization of final devoicing in German is not merely due to orthographic influences. The effects of incomplete neutralization were robust even when the pseudowords that do not have pre-existing orthographic representations were presented in a completely auditory task. It further demonstrated that the observed acoustic differences were perceivable, even though they may not serve a major communication function.

A recent large-scale study on Russian final devoicing by Kharlamov (2012, 2014) showed that incomplete neutralization can be driven by orthography, but may occur even without orthographic influences. This study investigated whether neutralization of final devoicing in Russian is influenced by task-independent phonological/lexical properties of stimuli and/or by task-dependent factors. More specifically, it examined the effects of the type of final obstruents (stops, fricatives), word length (mono- vs. disyllable), lexical types (words with vs. without counterparts that differed only in the voicing of obstruents), and task (the presence vs. absence of orthographic inputs and of minimal pairs in the stimuli). A total of 78 speakers produced 150 targets words either in a word-reading task or in a picture naming/word-guessing task. The measured parameters for stops included vowel duration, voicing into closure duration (as the number of cycles of vocal fold vibration), closure duration, and release burst duration. For fricatives, vowel duration, closure duration, and frication duration were measured.

Significant acoustic differences were observed in voicing into closure duration, closure/frication duration, and release burst duration, but not in vowel duration. In detail, underlyingly voiced obstruents were shorter by 1 - 7 ms (1 - 7%) in closure/frication duration, 1 - 6 ms (2 - 8%) in burst duration, and had 0.9 - 2.1 (45 - 167%) more cycles of vocal fold vibration than in underlyingly voiceless counterparts. All stimuli except plosive-final disyllables exhibited significant differences in closure/frication duration or release burst duration, regardless of whether speakers were exposed to orthographic inputs and minimal pairs. On the other hand, differences in voicing into closure duration were significant for all types of items, but only when the targets were elicited in the word-reading task and/or contained minimal pairs.

The author further examined the effects of task-independent vs. task-dependent factors on the perception of the voicing contrast (Kharlamov, 2012; 2015). In a 2AFC identification task, a total of 260 native Russian listeners heard the target words that were produced by 78 speakers in the reading or non-reading task. The results showed that listeners’ identification performances were limited in accuracy (41 - 50%). In detail, the rate of voiced responses was 42 - 44% for underlyingly voiced obstruents and 26 - 30% for underlyingly voiceless ones. This means that voiceless responses were over 50% for each obstruent type, implying an overall response bias toward the voiceless category. It also varied depending on which production context the stimuli were elicited in. For the stimuli elicited in the word-reading task with minimal pairs included, the listeners provided more voiced responses for underlyingly voiced obstruents and more voiceless responses for underlyingly voiceless obstruents. When they heard the stimuli that were elicited in the non-reading task and did not contain minimal pairs, their responses were less consistent with the underlying voicing specification of final obstruents and strongly biased toward voiceless category. In this case, however, the difference in the rate of voiced responses between underlyingly voiced vs. voiceless obstruents was still significant. Finally, a series of regression analyses were conducted to examine the relationship between perceptual identification rate and the acoustic cues present in the stimuli. It was revealed that voicing into closure duration was the only significant predictor of the rate of voiced responses for all types of the stimuli. Other parameters such as release burst duration and vowel duration also contributed to accounting for the identification rate, but to a limited extent.

Overall, this study showed that for some acoustic parameters, incomplete neutralization is driven by phonological and lexical factors in the absence of orthographic inputs or minimal pairs, while for other parameters, it is driven by task-related factors. In perception, the identification of the underlying voicing contrast was strongly affected by acoustic differences present in the stimuli produced when speakers were exposed to orthographic cues and minimal pairs. At the same time, it was revealed that the identification accuracy was also influenced by subtle acoustic differences produced without such exposure.

Another important piece of evidence for incomplete neutralization independently of orthographic influence comes from a study on final devoicing in Catalan. Dinnsen and Charles-

Luce (1984) demonstrated that final obstruents are incompletely neutralized in Catalan, a language in which the underlying voicing distinction is not maintained in the orthography. In this study, five talkers produced 5 minimal pairs in a reading task. The stimuli were embedded in frame sentences and followed by two phonetic contexts (_#V and _#C). The analysis of vowel duration, closure duration, and closure voicing collapsed across all talkers showed that there was no significant effect of the underlying voicing in all of the three acoustic parameters. However, individual subject data revealed considerable between-talker variation. Specifically, one talker maintained the underlying voicing distinction by means of vowel duration which was significantly shorter before underlyingly voiceless obstruents than before voiced counterparts. The other talker exhibited longer closure duration before voiced obstruents, which was against the expected direction (i.e., shorter closure before voiced obstruents). In spite of the group results which support complete neutralization, Dinnsen and Charles-Luce concluded that the word-final voicing contrast was not always completely neutralized in Catalan, at least for some speakers.

Interestingly, the opposite findings were obtained in a study on final devoicing in Turkish which is another language with no orthographic distinction of the final voicing contrast. Kopkalli (1993) demonstrated that final stops in Turkish are completed neutralized both in production and perception. In this study, five Turkish participated in a reading task in which they produced 30 pairs of words ending in stop consonants and embedded in a carrier sentence. Their productions were analyzed with respect to vowel duration, voicing into closure duration, closure duration, and aspiration duration in both a non-neutralizing and neutralizing environments.

The results showed that all of the parameters showed significant differences between voiced and voiceless obstruents in a non-neutralizing environment. Specifically, underlyingly voiced obstruents showed longer vowel duration (by 17 ms, 31%) and voicing into closure duration (by 31ms, 155%) than underlyingly voiceless obstruents. The opposite patterns were observed for closure duration (26 ms, 51%) and aspiration duration (26 ms, 433%). In a neutralizing environment, however, the mean differences in vowel duration (3 ms, 3%), voicing into closure duration (1 ms, 7%), and closure duration (3 m, 3%) were not significant, suggesting that Turkish speakers did not distinguish underlyingly voiced stops from voiceless ones.

The author also examined whether native Turkish listeners perceive voicing cues for word-final stops. Seven listeners participated in both identification and discrimination tasks in which they listened to the stimuli elicited from one female talker who produced the largest mean differences between underlyingly voiced and voiceless stops in three acoustic parameters. The results showed that listeners' performances in the two tasks were at chance level (49% in identification task, 51% in discrimination task), suggesting that the underlying voicing contrast of final stops in Turkish were not perceptually distinct. Taken together, Kopkalli suggested that the lack of acoustic or/and perceptual distinction between underlyingly voiced and voiceless stops is evidence for complete neutralization of the underlying voicing contrast in Turkish.

In summary, the previous literature has suggested that orthographic influences are the main trigger of incomplete neutralization, especially when an experimental task and items emphasize orthographic distinctions. The effects of incomplete neutralization were generally larger and more robust in languages where a target language orthographically maintains the underlying voicing contrast, the stimuli were elicited in a more controlled experimental settings (e.g., a reading task), and minimal pairs were included in the stimuli. On the other hand, the effects of incomplete neutralization were smaller or absent in languages without orthographic distinction of the final voicing contrast, and in an experimental setting where the influence of orthography was minimized.

While most of early studies used a relatively small sample of subjects and items and thus cannot avoid the criticism that the observed effects of incomplete or complete neutralization is not reliable enough, more recent studies (Kharlamov, 2014; 2015; Roettger et al., 2014) increased statistical reliability by adopting a sufficiently large number of subjects and items and also controlled the effect of orthography in a more systematic way. The common finding of these studies is that incomplete neutralization is not merely driven by speakers' exposure to orthography, since robust effects of incomplete neutralization is observed even in the absence of orthographic influences. However, Kharlamov's studies suggested that incomplete neutralization could be also caused by other sources such as phonological and lexical factors for some acoustic parameters. In perception, it was demonstrated that listeners' identification performances showed limited accuracy that was either significant in the Roettger et al.'s study or non-significant in the Kharlamov's study. It thus remains unclear whether listeners can actually make use of small acoustic differences in identifying underlying voicing of final obstruents.

2.1.2. Talker-related factor: L2 experience

It has been claimed that neutralization is affected by speakers' experience in L2 languages that does not have final devoicing. Jassem and Richer (1989) was the first to examine possible influences of prior L2 experiences on neutralization of final obstruents. In criticizing the Slowiaczek and Dinnsen (1985)'s study, they pointed out that their participants reported to be native Polish speakers had been living in English-speaking environments at the time of the experiment. Considering that the participants had been exposed to English that preserves final voicing distinction (e.g., *bat* vs. *bad*), it was probable that participants' phonological knowledge of L2 affected the production of neutralized final obstruents in Polish.

In order to remedy this problem, Jassem and Richer (1989) recruited four monolingual Polish speakers living in Poland that had little knowledge of any other language. This study also minimized the role of orthography by eliciting the stimuli in spontaneous oral dialogue between the experimenter and the participants as one-word answers to the experimenter's oral questions. The analysis of vowel durations, closure durations, and voicing into closure/frication revealed that vowels were shorter by about 4 ms (4%) before underlyingly voiceless obstruents than before voiced counterparts. However, the effects were inconsistent across talkers and not statistically significant, suggesting that the underlying voicing distinction in Polish is not reliably maintained.

In a subsequent perception experiment, ten listeners listened to the target words produced by 4 speakers and were asked to mark one of the paired sentences that contained the target. The results showed that listeners' correct identification rate was 66.8% for underlyingly voiceless obstruents, but only 38.5% for underlyingly voiced obstruents, implying that listeners' overall responses were strongly biased toward voiceless obstruents. Based on this result, the authors concluded that underlying voicing contrast is not preserved in perception, either. As evidence for complete neutralization of final obstruents in Polish, the findings of this study suggested that participants' prior L2 experience and maximum naturalness of the stimuli are crucial factors that should be controlled when investigating neutralization.

Some of later studies also controlled participants' L2 experience as in Jassem and Richer (1989), but often not systematically. Participants were recruited in their countries of origin, but their experience in L2 was either unknown (Kleber et al., 2010; Piroth & Janker, 2004; Roettger et al., 2014) or not very explicitly described (Warner et al., 2004; 2006). In other studies that recruited participants in the US, researchers made an effort to minimize the influence of prior experience in L2 by using the length of stay in L2 environments as a criterion for screening participants (Charles-Luce, 1985; Shrager, 2012). For example, Charles-Luce recruited native German participants who had stayed in the US for less than 2.5 years. In the Shrager's study, all three participants were native Russian speakers who had just come to the US for graduate study at the time of the testing. Still, no information is available as to participants' L2 proficiency in these studies.

More recently, Dmitrieva et al. (2010) systematically investigated the effect of L2 language experience on final devoicing. They tested three groups of participants, monolingual native speakers of Russian (n=7), native speakers of Russian residing in the US (n=4), and native speakers of English learning Russian (n=9). The target words containing final obstruents were produced in a word-list reading task and were analyzed in terms of preceding vowel duration, closure/frication duration, voicing into closure duration, and release duration of final obstruents.

The results showed that underlying voicing contrast was maintained in all four measures, suggesting that final obstruents are incompletely neutralized in Russian. The group analysis revealed that each group of talkers adopted different strategies in making distinctions between the underlyingly different final obstruents. Native Russian speakers with no knowledge of English maintained the underlying voicing contrast in closure/frication duration and release duration, each of which was 16 ms (10%, 17%) longer before underlyingly voiceless obstruents. On the other hand, native Russian speakers with substantial exposure to English showed significant differences in all measures. They produced longer closure/frication duration (15 ms, 9%) and release duration (17 ms, 14%) before underlyingly voiceless obstruents, which was comparable to those produced by monolingual speakers. Vowel duration was longer before underlyingly voiced obstruents than before voiceless counterparts by 8 ms (5%), which was larger than that produced by monolingual Russian speakers (2ms, 1%). Considering that vowel duration is the cue extensively used in English to signal final voicing contrast, this result clearly shows that speakers' knowledge of L2 interferes with L1 phonological processing.

Finally, native English speakers learning Russian distinguished underlyingly voiced from voiceless obstruents using similar strategies as monolingual Russian speakers, thus exhibiting incomplete neutralization. For this group, a significant negative correlation was found between vowel duration and L2 experience. That is, more proficient L2 Russian speakers produced smaller vowel duration differences, suggesting that L2 Russian learners with higher proficiency were more able to suppress the interference of their L1. Interestingly, however, it was found that the most proficient L2 speakers exhibited complete neutralization in all of the parameters. It thus seems that they had not acquired native-like skills of maintaining acoustic distinctions between underlyingly voiced and voiceless obstruents.

In summary, previous studies have suggested that neutralization is affected by prior experience in L2 languages without final devoicing. As demonstrated by Dmitrieva et al. (2010), the amount of exposure to such L2 languages has a significant influence on the use of acoustic cues to final voicing contrast in differentiating between the neutralized words. Despite awareness of the need to minimize L2 interference, many earlier studies lacked systematic control for the effects of L2 proficiency. Even when participants were recruited in their countries of origin, their L2 experience was not taken into account at all or its information was insufficient. It is therefore not clear whether findings of incomplete neutralization from these studies is confounded with L2 language interference.

2.1.3. Item-related factors

2.1.3.1. Phonetic environments

Several studies have shown that incomplete neutralization varies depending on the type of final obstruents and the environments in which final obstruents are placed. For instance, Charles-Luce (1985) examined the effect of phonetic environment and sentential position on final devoicing in German. In a reading task, five native German talkers produced 4 minimal pairs embedded in carrier sentences. The position of the stimuli was manipulated in such a way that the targets appeared either in clause-final or non-clause-final position and were followed either by consonant-initial or vowel-initial word. For the final obstruents /t/-/d/ and /s/-/z/, vowel duration, voicing into closure duration, and closure/frication duration were measured.

The results showed that vowel duration was longer by 10 ms (3%) before underlyingly voiced fricative /z/ than before voiceless /s/ in clause-final position. For stops, significant acoustic differences were observed only in voicing into closure duration (5 ms (14%) longer for /d/ than for /t/) in clause-final position and when followed by a consonant. This study concluded that the voicing contrast in final obstruents in German is incompletely neutralized, but highly restricted by the following phonetic environments and the sentential contexts. However, this finding was questioned on the grounds that the stimuli were not properly controlled. Piroth and Janker (2004) pointed out that some stimuli were embedded in a sentence frame with a very uncommon word ordering (e.g., verb-object-subject (VOS)), deviating from the typical German pattern, OSV or SOV. It was also indicated that the targets followed by a vowel-initial word could be regarded as the ones followed by a consonant-initial word, since words beginning with a vowel are usually preceded by a glottal stop in German.

A study by Slowiaczek and Dinnsen (1985) on final devoicing in Polish also suggested that neutralization of final obstruents may be affected by phonetic environments. In this study, five talkers produced fifteen minimal pairs in which the final obstruents varied in their place and manner of articulation. The target words were placed in two phonetic contexts, one followed by a consonant-initial word and the other followed by a vowel-initial word. The productions were analyzed in terms of vowel durations, closure duration, and voicing into closure/frication duration.

The analysis of group data revealed that vowel duration was approximately 10% longer before underlyingly voiced obstruents than before voiceless ones across all places and manners of articulation. In addition, vowels were 6-7 ms (5%) longer when followed by a vowel-initial word compared to when followed by a consonant-initial word. However, this effect was significant only for dental stops and affricates. Closure duration was 11 ms (16%) longer before underlyingly voiceless obstruents than before voiced counterparts, but only for dental stops and when followed by a vowel-initial word. For voicing into closure duration, the effect of underlying voicing was observed only in labial stops: underlyingly voiced stops were 13 ms (76%) longer than voiceless counterparts. Overall, the findings of this study suggested that significant effects of underlying voicing contrast are observed in Polish, but limited to certain phonetic contexts or certain types of final obstruents.

A later study by Jassem and Richer (1989), however, pointed out that the Slowiaczek and Dinnsen (1985) study failed to control for several important factors. Besides the lack of control of L2 proficiency addressed in the previous section, it was noted that the authors misjudged the domain of final devoicing that should be phrase-final, and placed the target items in word-final position that may not trigger final devoicing.

In Piroth and Janker (2004), the authors tried to deal with methodological issues associated with words and sentential contexts found in the previous studies. They controlled the adequacy of test materials by excluding infrequent or uncommon words but instead selecting minimal pairs from the core of standard German vocabulary. The stimuli were placed within sentence frames to avoid word-list reading effects. Finally, they included various word and phrasal contexts in which final obstruents occur to examine whether final devoicing is position-dependent.

Six native German speakers from three different dialectal groups produced target words containing final stops or fricatives in a non-neutralizing context (intervocalic word-medial) and in neutralizing contexts (syllable-final followed by voiceless vs. voiced segment, word- vs. utterance-final). Their productions were analyzed in terms of vowel duration, voicing into closure/frication duration, closure duration, burst duration, aspiration duration, frication duration, and voicing of burst/frication.

The results showed that in a non-neutralizing condition, the underlying voicing contrast of stops was maintained in vowel durations, voicing into closure duration, and closure plus release duration (coda duration hereafter). Across all six participants, underlyingly voiced stops were longer in vowel duration (by 35 ms, 28%) and voicing into closure duration (by 24 ms, 150%) and shorter in coda duration (by 67 ms, 51%) than voiceless counterparts. For fricatives, only voicing into frication duration and coda durations contributed to the underlying distinction. Voicing into frication duration was longer for underlyingly voiced fricatives across all speakers, while fricative durations were longer for voiceless fricatives only for two speakers.

When neutralized, fricatives became completely neutralized in all acoustic parameters. For stops, however, one dialectal group showed a significant difference in coda duration only in utterance-final position. This results suggested that incomplete neutralization is dependent upon the context the targets is presented in. The authors concluded that the findings would support either complete or incomplete neutralization, since incomplete neutralization was observed only in some of the dialects and the contexts they examined.

2.1.3.2. Phonological frequency and resyllabification

Kleber et al. (2010) examined whether neutralization of final obstruents in German is categorical or gradient in perception and whether the degree of incomplete neutralization is influenced by phonotactics probability and the potential for resyllabification. They used one male speaker's productions to create synthetic voiced to voiceless continua. Each of 4 continua consisted of 4 minimal pair compounds (e.g., /bɪg.lɪn/ *Bigglinn* - /bɪk.lɪn/ *Bicklinn)* in which the ratio of vowel duration to vowel plus closure duration (V/VC) was manipulated. The stimuli were placed in various neutralizing contexts. For example, final obstruents were preceded by lax or tense vowels, with the latter more frequent in occurrence. They were also combined with suffixes, which make either legal onset clusters (e.g., /gl/ - /kl/) or illegal ones in German (e.g., /gʃt/ - /kʃt/). The stimuli were then presented to nineteen listeners in a 2AFC identification task.

The results showed that the proportion of voiceless responses gradually increased as the V/VC duration ratio decreased. This suggested that even under a neutralized environment, the listeners are able to perceive acoustic differences, not in a categorical but continuous manner. This result thus supported incomplete neutralization of final obstruents in perception. It was also found that the listeners' identification of voicing contrast was influenced by the phonotactic frequency of phoneme sequences. Specifically, listeners identified more stops after lax vowels as voiceless than those after tense vowels regardless of the place of articulation of final stops. When preceded by tense vowels, the voicing contrast was more perceptible in alveolars than in velars, presumably because voicing contrast was frequent in the tense vowel plus alveolar stops sequences, but not in the tense vowel plus velar stop sequences. Finally, there was a greater chance of perceiving voicing contrast when the final obstruents plus suffixes formed a legal onset cluster and therefore was considered as not subject to neutralization in domain-final position.

Taken together, the authors concluded that perceptual neutralization of final devoicing in German cannot be fully accounted for by a categorical model of neutralization, but should be modeled with consideration of phonetic context, prosodic position, and phonotactic probability.

2.1.3.3. Availability of semantic information

Previous studies have demonstrated that production and perception of spoken words are affected differently by the presence of semantic or syntactic information. For example, it was found that vowel duration was significantly longer in words embedded in a semantically neutral context (Lieberman, 1963) or in ungrammatical sentences (Charles-Luce & Walker; 1981) than in words produced in semantically biasing context or in a grammatical sentences. It was also shown that the acoustic differences affected listeners' identification of words extracted from the sentences in which they were originally produced: the identification rate was higher for words produced in a semantically neutral context than for words produced in a biasing context (Lieberman, 1963). These findings imply that articulatory precision is reduced when higher levels of linguistic information is present, because speakers tend to produce words with less care when they know that listeners can identify words with the help of contextual information.

Motivated by the previous findings, Charles-Luce (1993) investigated whether neutralization of final obstruents in Catalan is affected by semantic information. The target stimuli were placed in two semantic contexts (biasing vs. neutral). They were also placed in two phonetic environments that trigger regressive voice assimilation (followed by voiceless [s] vs. followed by voiced [r]), which results in the loss of underlying voicing contrast. Five speakers produced paragraphs composed of two sentences containing the target words. Their productions were examined in vowel duration, voicing into closure duration, and closure duration.

In a semantically biasing context, no significant effect of underlying voicing was observed in any of the acoustic parameters. When semantic information was absent, however, the phonemic voicing contrast was preserved in vowel duration. Underlyingly voiced stops were 15 ms (17%) longer in preceding vowel duration than underlyingly voiceless stops in the both assimilatory environments, but only for 3 out of 5 target minimal pairs. These findings suggested that incomplete neutralization is more likely to occur in the absence of higher levels of linguistic information, since speakers may modify their productions in an effort to avoid any ambiguity in the message, thereby helping listeners to recover the intended meaning.

2.1.4. Summary

To sum up the discussion so far, the traditional notion of neutralization has been challenged by numerous experimental studies on final devoicing. These studies have demonstrated that the trace of the underlying voicing contrast of final obstruents remains in various acoustic parameters, and that the acoustic differences are often salient enough for native listeners to perceive. However, the findings of incomplete neutralization has been challenged by the claims that incomplete neutralization is triggered by various factors associated with tasks, talkers, and items. It has been shown that orthographic interference is the most influencing factor, particularly in languages in which the underlying contrast of final obstruents is orthographically represented and in tasks that emphasize orthographic distinctions. However, the role of orthography remains a controversial issue, since some studies have found evidence for incomplete neutralization even when orthographic influence is minimized or when the underlying voicing contrast is not represented in the orthography of a target language.

Another factor triggering incomplete neutralization is speakers' prior experience in L2 languages without final devoicing. Most of the previous studies have failed to control it in a systematic way. In most of the earlier studies conducted in the US, it was inevitable that speakers had already been exposed to English to some extent. Those studies did not provide sufficient information about speakers' L2 proficiency, either. Even when researchers recruited participants in their native countries, they had difficulties in finding purely monolingual speakers from young educated populations who have no prior exposure to any L2 language. So the possibility remains that voicing-dependent effects observed in the previous studies may have been confounded with L2 interference effects.

Besides, it has been shown that neutralization of final obstruents varies depending on item-related factors. Incomplete neutralization is often restricted to certain types of obstruents or certain phonetic contexts, or more likely to occur when higher level linguistic information is absent. In the perceptual domain, it has been demonstrated that listeners show continuous rather than categorical perception of underlying final contrast, thus supporting incomplete neutralization in perception. Additionally, it has been shown that the perception of neutralized final obstruents is influenced by the frequency of phoneme sequences and the potential of final obstruents to be re-syllabified into domain-initial ones.

Taken together, we can draw at least one conclusion from the inconsistent and contradictory findings above that incomplete neutralization is not an across-the-board phenomenon, but rather dependent upon various factors associated with orthography, tasks, talkers, and items.

2.2. Theoretical accounts for incomplete neutralization

2.2.1. Approach of traditional generative phonology

An increasing amount of empirical evidence on incomplete neutralization (mostly in production) has consequently led researchers to cast doubt on the validity and correctness of the traditional generative approach to neutralization. As introduced in the beginning of Chapter 1, the assumption of generative phonology is that the production of neutralized words involves the activation of the abstract UR of a target word, the application of a rule of neutralization that changes the feature specification of final obstruents categorically, and the conversion of the resulting phonological representation into articulatory gestures via phonetic implementation. The feed-forward flow of information throughout the speech production system implies that the output of neutralization rules, i.e., the representations having the identical [-voice] feature specifications, are physically realized with the same acoustic-phonetic properties. Therefore, incomplete neutralization is unexpected and cannot be accounted for within the generative framework.

Nevertheless, several researchers have proposed how to incorporate incomplete neutralization into the traditional models of phonology by reinterpreting phonological rules of neutralization. Port and O'Dell (1985) proposed that the rule of neutralization should be viewed not as a phonological rule but as a phonetic implementation rule which applies after all phonological rules. Particularly, they suggested that final devoicing may be the rule that implements the syllable unit, independently of the rule that implements segments or segmental features. The syllable implementation rule takes linguistic structures at the end of phonological derivations as its input. Thus the syllable endings that contain final obstruents becomes the input, with the [+voice] feature of final obstruents intact. The syllable implementation rule then initiates gestures that resemble the segmental implementation of the [-voice] feature. Consequently, it is supposed that both [-voice] segmental feature and syllable-final [+voice] feature would be pronounced in a similar, but not identical manner. The main problem with this proposal is that final devoicing, which is an aspect of German phonology, becomes indistinguishable from phonetic implementation. The final devoicing rule would not be a phonological rule any more, but more like an allophone rule that cannot capture the generalization that final obstruents alternate in a non-neutralizing environment.

Other researchers have proposed that phonetic implementation rules must be applied before or simultaneously with the phonological rule of final devoicing, since they are sensitive to underlying voicing contrast (Dinnsen & Charles-Luce, 1984; Slowiaczek & Dinnsen, 1985). Slowiaczek and Dinnsen (1985) observed that the closure duration of final obstruents was shortened when followed by a vowel. In this case, the amount of shortening was dependent upon the underlying voicing of the final obstruents. Their speculation was that if a phonological rule of final devoicing that changes [+voice] feature into [-voice] is applied before the phonetic implementation rule, relative shortening of closure duration cannot be accounted for, since all final obstruents are shortened equally when followed by a vowel. In contrast, the opposite order of rule application makes it possible for the [+voice] feature to be available to phonetic implementation and its effect on vowel duration can appear at the surface level. As they admitted, however, the reverse rule ordering goes against the modular and feed-forward architecture of generative phonology (Pierrehumbert, 2001).

Charles-Luce (1985) suggested that final devoicing rule may be considered as feature-deletion rule. Rather than categorically changing the [+voice] feature into [-voice], the rule deletes [+voice] of a word-final voiced obstruents, which results in the segment unspecified for voice and free for phonetic variations. The degree of voicing of the voiced obstruents is then phonetically realized through language-specific, speaker-dependent implementation rules. This proposal has an advantage that it maintains the conventional rule ordering, that is, phonological rules applied before phonetic implementation rules. However, the neutralization rule in this case is no longer a phonological rule but rather an allophonic rule, since it loses linguistic generalization due to its application only to voiced obstruents. Furthermore, final devoicing would not be considered as neutralizing phenomenon because the voicing contrast is still preserved in the word-final position.

More radically, Dinnsen (1985) argued against the existence of true neutralization, claiming that all possible cases of neutralization phonetically examined have turned out to be non-neutralizing[4]. He classified neutralization into four possible types of rules by considering the possibility of distinctions in production and perception. According to his classification, rule Type A produces surface forms that are phonetically identical and not perceivable by listeners, representing the standard notion of neutralization. Both the Type B and C rules produce outputs

[4] It appears that Dinnsen (1985)'s claim need to be modified, since several studies have been documented supporting Type A cases for which no empirical evidence existed at the time of his proposal. Kopkalli (1993)'s study on final devoicing in Turkish and Kim and Jongman (1996)'s study on coda neutralization in Korean demonstrated that the neutralization rule does produce surface forms that cannot be distinguished either in production or perception, thus supporting Type A case.

that are acoustically distinct, but are distinguished by the magnitude of perceptual differences: acoustic differences too small to be perceived (Type B) or large enough to be perceived (Type C). Type B constitutes limited neutralization, whereas Type C is non-neutralization. Type D cases produce outputs that are phonetically identical but still perceivable by listeners, as pointed out by Dinnsen, logically impossible.

His claim suggests that there could be a rule of incomplete neutralization. That is, rather than phonological rules being limited to complete neutralization, they could also specify incomplete neutralization. He proposed that the only way to salvage neutralization as a rule is to better establish Type B cases in which speakers produce acoustic differences that they cannot perceive (i.e., complete neutralization in perception at least). However, this requires fundamental changes in the traditional notion of neutralization in such a way that neutralization is considered only as a perception-oriented phenomenon, independent of the production facts. Such a change would imply that unlike the dominant view that perception reflects production, perception and production are partially independent. It would also imply that phonological theory which is assumed to be neutral with respect to speakers and listeners should admit the distinction between production and perception and thus include both production-based grammars and perception-based grammars. Consequently, Type B cases would be described with two different rules, one describing perceptual facts and the other one describing phonetic differences.

To sum up, various attempts have been made to incorporate incomplete neutralization into the traditional generative phonology. However, the proposed rules do not fit into the basic assumptions of generative grammar or appear to be post-hoc devices implemented differently depending on speakers, contexts, or languages, thus having little explanatory or predictive power. Faced with this problem and with increasing evidence of incomplete neutralization, researchers have shifted their attention to more general cognitive processes as the source of incomplete neutralization. The main suggestion is that incomplete neutralization occurs due to the influence of cognitive properties of the mental lexicon and lexical processing. In the next section, I will discuss a more recent approach to incomplete neutralization based on exemplar-models of speech perception and production.

2.2.2. Approach of exemplar-based models of speech perception and production

Exemplar-based theories were first developed as general models of similarity and category formation in perception. They have been extended to the domain of language to model linguistic categorization in perception (Bybee, 2001; Goldinger, 1996; Johnson, 1997; Klatt, 1979; Pisoni, 1997) and production (e.g., Goldinger, 1998; Pierrehumbert, 2001, 2002). The central idea behind exemplar-based models is that the mental lexicon contains memory traces (or exemplars) for all tokens of words that language users have previously encountered in their production and perception. The exemplars contain detailed acoustic or articulatory information of the tokens, without abstracting away from speaker- or context-specific variability. In these models, a linguistic category is not represented as symbols, but by clouds of the exemplars associated with that category in memory.

The category is incrementally updated as new tokens are encoded and old unused exemplars are gradually faded from memory (Bybee, 2001; Pierrehumbert, 2001, 2002). New speech tokens are categorized on the basis of their similarity to the representations of the existing exemplars stored in memory. Tokens that are highly similar to or whose differences are too fine to be distinguished from the stored exemplars are encoded as identical. On the other hand, dissimilar tokens establish themselves as new exemplars and are stored close to the existing exemplars to constitute a category. The stored exemplars thus show variations within a category, which are exhibited when the category is physically manifested. It is also possible that the stored exemplars show overlap in their distributions across the categories. The continuous updates with new tokens may change the range and activation of the exemplars within a category, which in turn alters the entire category system slightly.

Another important feature of the models is that frequency information is not overtly encoded but implicit in the exemplar database. Given that every token of language experience is stored and categorized in the exemplar space, more frequent categories are assumed to have a larger number of exemplars than less frequent categories whose exemplars are less numerous. It is also assumed that exemplars vary in the resting level of activation. Tokens encoding highly frequent and recent experiences have higher activation levels than those that encode less frequent and temporarily distant experiences.

In these models, the identification/recognition of words involves the calculation of similarity between a new token and the stored exemplars. Once the perceptual encoding of a new token makes it placed in a map of the perceptual space, the token is classified through the computation of the most probable label based on the similarity of the token to each of the stored exemplars and the activation strength of each exemplar in a fixed size neighborhood (Pierrehumbert, 2002). With regard to production, Pierrehumbert (2002) proposes that it involves the same process as in perception, but proceeds in the opposite direction. Specifically, a production target is created based on a randomly selected exemplar in a given category, where the probability of random selection is in proportion to the strength or the base-activation level of the exemplar. Then the actual production target is formed by averaging the phonetic characteristics of exemplars around the randomly selected exemplars.

On the basis of this theoretical framework, Ernestus and Baayen (2006, 2007) suggest that incomplete neutralization is the reflection of the properties of lexical representations. The premise behind this proposal is that the lexical representations are gradient and retain detailed information about phonetic properties of the segments (Bybee, 2005; Goldinger, 1998; Pisoni, 1997). For example, the lexical representation of the Dutch word for 'basket' (/mant/) directly reflects actual pronunciation of the word and thus contain information regarding the probability that the final obstruent is realized as slightly voiced. Consequently, language users may use this probabilistic information in their production and perception, leading to incomplete neutralization.

An alternative account Ernestus and Baayen propose is that physical manifestations of the neutralized words are affected by the pronunciations of co-activated words in the morphological paradigm. The assumption behind this account is that the mental lexicon contains representations for all words as well as their morphologically related forms (Alegre & Gordon, 1999a; Baayen, McQueen, Dijkstra, & Schreuder, 2003; Bybee, 2001; Stemberger & MacWhinney, 1986). Once a target word is activated, activation spreads to morphologically related words that are tightly connected to the target in the mental lexicon due to their similarity in meaning and form.

When Dutch speakers produce the word 'basket', for example, they activate both the canonical form [mant] and the inflected form [mandən]. Since the co-activated inflected form contains a non-neutralized voiced segment (e.g., [d] in [mandən]), it increases the probability that the final obstruent of a singular form is realized as slightly voiced. In contrast, the inflected form that contain a non-neutralized voiceless segment (e.g., [t] in [kranten] 'newspapers') may not lead to slight voicing of final obstruent in a singular form [krant] (Ernestus & Baayen, 2007).

Listeners' identification of the neutralized words can be also accounted for by similar mechanisms. When a listener is presented with a word as the input, the category of exemplars that best matches the input signal is activated along with its nearest neighbors including inflected forms. The activation level of each exemplar is determined by the similarity of auditory properties between the input and the existing exemplars. Given this, inflected forms containing a voiced obstruent (e.g., [mandən]) are more likely to be activated when listeners perceive a word that is realized with a slightly voiced obstruent (e.g., [mant]) than when they perceive a word realized with a voiceless one (e.g., [krant]). If the final obstruent is realized as slightly voiced as in [mant] ('basket'), listeners may interpret the word as the one with a voiced obstruent (i.e., non-neutralized). In contrast, the final obstruent realized as completely voiceless as in [krant] ('newspaper') may lead to listeners' interpretation of the word as the one with a voiceless obstruent (i.e., neutralized).

The exemplar-based accounts explain why the categorical change of voicing involves gradient variation, especially in production, without relying on the notion of abstract units (e.g., phonemes) and phonological or phonetic implementation rules. The fundamental assumption of the exemplar-based models is that lexical representations are gradient and reflect fine-grained phonetic details of the segments. In other words, lexical representations and phonetic representations are directly associated. Within this framework, therefore, incomplete neutralization can be interpreted as the output of the interaction of lexical representation and detailed phonetic information that are represented in a single level of processing (Ernestus & Baayen, 2006, 2007; Pierrehumbert, 2001, 2002).

This interpretation is consistent with how cascading models of speech production views incomplete neutralization. The cascading models posit that activation flows both forward and backward between levels (Dell, 1986: Stemberger, 1985; Rapp & Goldrick, 2000). Moreover, activation cascades from one level to subsequent processing levels before and after processes at the earlier level is complete (Rapp & Goldrick, 2000). In these models, incomplete neutralization is explained as the interactions via cascading activation between phonological processing and articulatory implementations (Goldrick & Blumstein, 2006). In the German word pair Rad /rad/ ('wheel') vs. Rat ('advice'), for example, the reason for partial voicing of the final obstruent in [rat] for Rad is due to the partial activation of underlying target /rad/ during phonological processing. Since the final obstruent of the word Rad alternates between [t] and [d] depending on phonological environments, its underlying representation would contain the phonetic traces of /d/ in order to reflect this alternation fact. For the counterpart Rat, no phonetic trace of voicing would be found in the underlying representation, since the final obstruent does not alternate. If the partially activated underlying representation spread to articulatory/phonetic processing level, we come to observe incomplete neutralization of final obstruents.

Clearly, this type of interaction effects cannot be accommodated by the discrete models of speech production in which the processing of lexical, phonological, and articulatory information proceeds uni-directionally from higher to lower levels, with information at each level not interacting with each other (e.g., Levelt, 1989; Levelt, Roelofs, & Meyer, 1999). However, the exemplar-based models' assumption of the direct links between lexical and phonetic representations allows us to predict that the degree of phonetic traces of voicing depends on whether the underlying representation of a target contain the phonetic traces of voiced obstruents.

To sum up, the exemplar-based approach to incomplete neutralization is based on the premise that both the previously encountered tokens of stimuli and their inflected forms are stored in our memory and activated when speakers and listeners process the neutralized words. Since the exemplar-based models assume that lexical representations are gradient and reflect fine-grained phonetic details of the segments, it can explain why categorical change of voicing involves gradient variation in production and perception without relying on the notion of abstract units (e.g., phonemes) or speaker normalization.

2.3. Current study

The purpose of the current study is to investigate acoustic and perceptual aspects of coda neutralization in Korean. By replicating and extending the Kim and Jongman (1996)'s study, I will examine whether all three types of neutralization of coda obstruents in Korean is phonetically complete or incomplete in production and perception.

Neutralization of coda consonants in Korean is similar to final devoicing in that it also occurs in word-final position and the underlying contrast is maintained in the orthographic forms, but differs from it qualitatively in several aspects. First of all, coda neutralization involves the merger of two or more phonologically distinct voiceless sounds into a voiceless lenis sound, whereas final devoicing involves the loss of the phonemic distinction between voiced and voiceless sounds. For example, the coda consonants in the words such as /pitʃʰ/ ('light'), /pitʃ/ ('debt'), /pis/ ('comb') have different underlying representations: voiceless aspirated /tʃʰ/, voiceless lenis /tʃ/, and voiceless lenis /s/. The phonemic contrasts among them become neutralized into voiceless lenis [t] at the surface level due to a phonological rule of coda neutralization. Consequently, a single coda neutralization process can result in a case of homophony with more than two underlying representations.

Secondly, coda neutralization is distinct from final devoicing in the phonetic properties of its output. Coda obstruents in Korean become unreleased as a result of neutralization, whereas final obstruents in final devoicing languages becomes released. Coda neutralization in Korean favors lenis over aspirated or tense series, since there is a constraint on syllable structure that obstruents in a syllable-final position are produced as unreleased after complete closure. Therefore, this position involves neither aspiration nor tenseness.

Thirdly, Korean coda neutralization is classified into three different sub-classes (laryngeal, manner, and palatal), depending on what type of phonological contrast is involved in coda neutralization. Laryngeal neutralization involves the neutralization of laryngeal contrasts (i.e., lenis-aspirated (/p-p^h/, /t-t^h/), lenis-fortis (/k-k'/)) into lenis obstruents [p], [t], [k]. Obstruents that differ in their manner of articulation (e.g., fricative vs. stop, affricate vs. stop, fricative vs. affricate) are merged into lenis stop [t]. Palatal neutralization merges the laryngeal contrast among palatal affricates (/tʃ-tʃ^h/) into the lenis stop [t]. This property contrasts with final devoicing in which only the [voice] feature distinction is eliminated. It also means that obstruents subject to coda neutralization are much larger in their numbers than the ones subject to final devoicing.

To our knowledge, there has been only one experimental study (Kim & Jongman, 1996) that examines the acoustic and perceptual aspects of coda neutralization in Korean. In this study, the preliminary acoustic analysis of bisyllabic minimal pairs containing medial /s/ and /t^h/ (e.g., [tsosa] 'investigation' vs. [tsot^{h}a] 'steering') showed that vowel durations were longer before fricatives /s/ than before stops /t^h/ in Korean. Based on this result, it was hypothesized that if Korean manner neutralization is incomplete, there would be differences in vowel durations between the two underlyingly distinct segments. Additionally, it was predicted that neutralized stop [t] derived from underlyingly aspirated stop /t^h/ may be longer than the one derived from underlyingly lenis stop /t/, since previous studies (Han, 1994; Silva, 1992) have shown that closure durations of the aspirated stops are longer than those of the lenis stops in Korean. The results of acoustic analyses, however, revealed that there was no such difference. A subsequent perception experiment was conducted in order to examine the possibility that other acoustic-phonetic correlates such as burst intensity and formant transitions help listeners to distinguish the underlying contrast. The results showed that listeners were not able to correctly identify the intended underlying form, suggesting that neutralized forms did not contain any perceivable acoustic cues to the underlying distinction of manner. Based on these results, the author concluded that manner neutralization in Korean is complete both in production and perception.

Kim and Jongman (1996) is the first and only attempt to explore the acoustic and perceptual aspects of neutralization of coda consonants in Korean. However, their investigation is confined to manner of articulation, one of the three different types of coda neutralization. For a complete picture of neutralization of Korean coda consonants, there should be further investigation of whether the other two types, laryngeal and palatal neutralization, are also complete or not.

Another potential problem of the Kim and Jongman (1996)'s study is that the sample size of the production experiments appears too small to draw a conclusion that manner of articulation in Korean is completely neutralized (n=3 for the analyses of vowel durations preceding fricatives and stops; n=4 for the analyses of vowel and consonant durations of neutralized stops). As in Fourakis and Iverson (1984), this study cannot avoid the criticism that complete neutralization might be due to a small sample size. Moreover, in a perception experiment, it is possible that stimuli elicited from a small number of talkers might have led listeners to be familiarized with each talker's idiosyncratic speech patterns more quickly and easily (e.g., Mullennix, Pisoni, & Martin, 1989; see also Piroth & Janker (2004) for inter-speaker production differences in incomplete neutralization of German final obstruents). A larger sample would help us control for any possible speaker idiosyncracies.

In terms of acoustic correlates of neutralization, they only examined vowel duration and consonant duration. However, previous studies on final devoicing have shown that voicing into closure duration also serves as a reliable cue that contributes to the distinction of underlying voicing contrast (e.g., Kharlamov, 2012, 2014, 2015; Port & Crawford, 1989; Port & O'Dell, 1985; Shrager, 2012). For more accurate assessment of phonetic aspects of coda neutralization, it would be necessary to examine voicing into closure duration as well.

Finally, Kim and Jongman (1996) adopted a reading task in which participants were provided with the sentences containing the target words in Korean orthography. They found that Korean coda obstruents were completely neutralized even in the presence of orthographic information, providing counter-evidence that incomplete neutralization occurs due to orthographic influences. However, the role of orthography remains unclear, since many of the previous studies have revealed that incomplete neutralization is more likely to occur in languages in which underlying contrast is orthographically represented and in tasks that emphasize the orthographic distinctions (Fourakis & Iverson, 1984; Jassem & Richter, 1989; Port & Crawford, 1989; Warner et al., 2006).

In this study, therefore, I will examine all of the three types of coda neutralization in both production and perception aspects by recruiting a sufficiently large number of talkers and listeners and by conducting more thorough acoustic measurements. The possible influences of orthography will be investigated in a more controlled manner. The stimuli will be elicited in two different tasks that vary in the presence of explicit orthographic cues. The target words elicited in these tasks will be compared in their acoustic properties, which will reveal whether the presence or absence of orthographic inputs influences coda neutralization.

The findings of this study will be discussed in terms of the traditional generative account vs. the exemplar-based account of neutralization. If coda neutralization is phonetically complete, it is expected that there will be no systematic acoustic differences between the neutralized words. Consequently, listeners may not be able to distinguish the differences between the neutralized words derived from different underlying forms. This line of results would support traditional generative approach. That is, the production of neutralized words involves the application of a rule of coda neutralization that changes the feature specification of final obstruents categorically (e.g., [-lenis] →[+lenis]), which results in the identical surface forms with no difference in acoustic-phonetic characteristics.

In perception, the identification of completely neutralized words may involve the process assumed in the traditional abstractionist models of spoken word recognition (e.g., Marslen-Wilson, 1987, 1993; McClelland & Elman, 1986; Norris, 1994). That is, listeners take auditory information in the input signals of the word pairs (e.g., 'house' - 'straw') and generate pre-lexical level of representations that are mapped onto lexical representations /tʃip/ and /tʃip^h/, respectively. During this process, all speaker- or context-specific details are abstracted away, making highly variable acoustic signals perceptually equivalent. Once one of the activated lexical candidates that best matches the auditory information is selected, the input signals will be recognized as a word if the activation of the selected candidate reaches a certain threshold level. Considering that completely neutralized words do not exhibit any acoustic difference, it is expected that the two different inputs are recognized as the same (i.e., [tʃip]).

If neutralization is not phonetically complete, on the other hand, it is predicted that the neutralized forms will show systematic acoustic differences in a similar manner to the realization of the underlying distinctions in a non-neutralizing environment. With respect to perception, two predictions can be made. First, it is possible that acoustic differences in production are perceptible enough for listeners to distinguish the differences between the underlyingly distinct forms. This pattern of results would support an exemplar-based account that views incomplete neutralization as the reflection of lexical representations that contain fine-grained phonetic details of final obstruents or the influence of the co-activated inflected forms of a target.

Second, it is possible that acoustic differences observed in production are not perceptible for listeners. In this case, it would be considered that neutralization occurs, but only perceptually. This result could be still explained within exemplar-based models as a phenomenon like near merger (Pierrehumbert, 2001; Yu, 2007). However, this result may challenge the models, since they cannot explain why such a phenomena is observed in coda neutralization in Korean, but not in final devoicing.

The predictions will be tested in the production experiment (Chapter 3) and the perception experiment (Chapter 4), followed by the investigation of the relationship between production and perception (Chapter 5).

CHAPTER 3. PRODUCTION STUDY

3.1. Introduction

The aim of the production study is to investigate whether three types of coda neutralization in Korean are phonetically complete or incomplete. For this purpose, I will replicate and extend Kim and Jongman's (1996) work by adopting its rationale but with a sufficiently large number of talkers/listeners and with more thorough acoustic measurements. Specifically, recordings made by 15 native Korean talkers will be analyzed in terms of three temporal acoustic parameters that are relevant to the underlying contrast of coda obstruents. Data from measurements will be analyzed first in a non-neutralizing environment in order to establish that word pairs subject to neutralization are reliably distinguished from each other in terms of acoustic properties relevant to the coda contrast. Measurements from neutralizing environments will then be analyzed to see whether there is any acoustic evidence for complete or incomplete neutralization.

Another aim of this study is to examine the effect of explicit orthographic cues on coda neutralization. The effects of orthography will be controlled through the manipulation of stimuli elicitation tasks in which orthographic cues are either present or absent. Specifically, in an orthography-absent Q & A task, participants are asked to listen to the questions and provide answers without being exposed to any explicit orthography that could provide incomplete neutralization cues. In a strongly orthography-driven reading task, on the other hand, they will read a list of individual sentences containing the target words. By comparing acoustic properties of the target words elicited from two different experimental tasks, I will be able to demonstrate whether the presence or absence of explicit orthographic cues influences coda neutralization.

If neutralization is phonetically complete, it is expected that there will be no systematic differences in acoustic parameters relevant to coda obstruents between the neutralized words. For instance, the surface forms ([tʃip]) of the minimal pair 집 /tʃip/ (‘house’) and 짚 /tʃipʰ/ (‘straw’) may be acoustically the same without being distinguished by any acoustic cues relevant to coda [p]. If neutralization is not phonetically complete, however, it is expected that the surface forms will show systematic acoustic differences in a similar manner to when their underlying distinctions are maintained. Possible outcomes for each acoustic parameter will be stated in more detail in the next section.

With regard to the effect of orthography on coda neutralization, if Korean orthography interferes with coda neutralization, it is expected that words elicited from the reading task will be more subject to incomplete neutralization, since Korean orthography reflects the underlying contrast of coda. In this case, significant acoustic differences are expected across the Q & A and the reading tasks between the neutralized words with underlyingly distinct codas. If coda neutralization is independent of orthographic influence, on the other hand, no acoustic difference will be found between the neutralized words elicited from the two tasks.

3.2. Coda neutralization in Korean

This section reviews coda neutralization in Korean. It is followed by discussions on acoustic correlates of Korean obstruents and predictions on their behaviors in a non-neutralizing and a neutralizing environment. For the current investigation, I will mainly address temporal-acoustic correlates of the obstruents in an intervocalic position, since they will serve as a reference to predict the direction of contrasts in those acoustic parameters in a neutralizing environment.

3.2.1. Three types of coda neutralization in Korean

The obstruents in Korean consist of stops, affricates, and fricatives, which are all voiceless (See Table 1). Both stops and affricates exhibit a three-way laryngeal contrast (lenis, aspirated, and fortis: Cho, Jun, & Ladefoged, 2002; C.-W. Kim, 1965), whereas fricatives are either lenis or fortis.

			Place				
			bilabial	alveolar	palatal	velar	glottal
Manner	stops	lenis	ㅂ /p/	ㄷ /t/		ㄱ /k/	
		aspirated	ㅍ /p^h/	ㅌ /t^h/		ㅋ /k^h/	
		fortis	ㅃ /p’/	ㄸ /t’/		ㄲ /k’/	
	affricates	lenis			ㅈ /tʃ/		
		aspirated			ㅊ /tʃ^h/		
		fortis			ㅉ /tʃ’/		
	fricative	lenis		ㅅ /s/[5]			ㅎ /h/
		fortis		ㅆ /s’/			

Table 1. Obstruents in Korean

All obstruents are allowed in word- or phrase-initial position. In syllable-final position, however, neither aspirated nor fortis stops appear in this position, but only three lenis stop

[5] Cho *et al*. (2002) pointed out that /s/ in Korean has been categorized either as lenis or as aspirated (Park, 1999; Yoon, 1998). In order to determine which categorization is more appropriate, they investigated acoustic-phonetic properties of fricative by measuring duration of frication and aspiration, centroid of the fricative noise, amplitude difference between first harmonic (H1)-second harmonic (H2) and between H1-second formant (F2), and fundamental frequency (F0). Their findings suggested that fricative /s/ has more properties of lenis category than those of aspirated. In this study, therefore, I adopt the naming convention of /s/ as lenis.

consonants ('ㅂ' /p/, 'ㄷ' /t/, 'ㄱ' /k/) surface in coda. This is due to the constraint on syllable structure that obstruents in syllable-final position are unreleased after complete closure.

Since aspiration and tenseness are the major features necessary for distinctions among Korean obstruents (e.g., /p, pʰ, p'/, /s, s'/, /tʃ, tʃʰ, tʃ'/), their absence in a syllable-final position leads to neutralization of underlying phonological distinction among the obstruents. This is what is termed as coda neutralization (Kim-Renaud 1974; C.-W. Kim, 1979; Chung, 1980; Kim & Jongman, 1996). In detail, bilabial stops (/p, pʰ/) are neutralized into [p], alveolar stops, palatal affricates, and fricatives (/t, tʰ, s, s', tʃ, tʃʰ, h/) into [t], and velar stops (/k, kʰ, k'/) into [k][6]. Neutralization also occurs before a consonant and before a word/compound boundary when the onset of the next syllable is a consonant (Sohn, 1999), as shown in (1), a phonological rule of coda neutralization in Korean.

(1) Coda neutralization (Sohn 1999; 165)

/p, pʰ/	→ [p]
/t, tʰ, s, s', h, tʃ, tʃʰ/	→ [t]
/k, kʰ, k'/	→ [k]

when unreleased (i.e., in the environment of __] σ or __ C, #, +, where # is a word boundary and + is a compound boundary)

Depending on what type of phonological contrast is involved in coda neutralization, neutralization can be classified into three types: laryngeal, manner, and palatal neutralization

[6] Korean allows /p', t', tʃ'/ only in word- or phrase-initial position. As they are absent from underlying forms in syllable-final position, they are not subject to coda neutralization.

(Kim & Jongman, 1996). For the purpose of the present study, I adopted the rule of laryngeal neutralization and palatal neutralization from Kim and Jongman, and the rule of manner neutralization from Sohn (1999).

Laryngeal neutralization refers to a process that laryngeal contrasts (i.e., lenis-aspirated (/p-pʰ/, /t-tʰ/), lenis-fortis (/k-k'/)) are neutralized into lenis obstruents [p], [t], [k]. In manner neutralization, the obstruents that contrast in their manner of articulation are merged into lenis stop [t]. For the current study, I only examined three pairs of coda with manner distinctions (fricative-stop, affricate-stop, and fricative-affricate pairs). Finally, palatal neutralization merges the laryngeal distinction between palatal affricates (/tʃ-tʃʰ/) into the lenis stop [t]. The rules below represent the three types of coda neutralization in Korean the current study investigates:

(1a) Laryngeal neutralization: /p, pʰ/ → [p]

/t, tʰ/ → [t]

/k, kʰ, k'/ → [k]

(1b) Manner neutralization: /t, tʰ, s, s', h, tʃ, tʃʰ/ → [t]

(1c) Palatal neutralization: /tʃ, tʃʰ/ → [t]

(1a), (1b), and (1c) occurs when unreleased (i.e., in the environment of __] σ or __ C, #, +)

When coda obstruents are followed by a vowel-initial suffix and/or case markers (e.g., 에 (locative particle) /-e/), they are syllabified into the onset of the following syllable without being subject to neutralization. In this environment, the phonemic distinction is maintained as in the onset position. However, voiceless lenis obstruents except /s/ become voiced intervocalically due to a phonological rule of Korean that obstruents get voiced between vowels.

Some examples of coda neutralization are presented in Table 2. As shown in the right two columns under phonological environments, coda obstruents become neutralized into lenis consonants when they appear in syllable-final position or followed by consonant-initial suffix (e.g., /-to/ (‘also’) preceded by nouns or /-kwa/ (‘and’) preceded by nouns or verbs). As a result of neutralization, word pairs or multiple words with distinct underlying representations become homophones at the surface level.

	Word pairs	Phonological environments					
		Non-neutralizing		Neutralizing			
		__ V		__] σ		__ C	
1a)	집 ‘house’	/tʃip + ɨl/	→ [tʃi.bɨl]	/tʃip/	→ [tʃip]	/tʃip + to/	→ [tʃip.t’o]
	짚 ‘straw’	/tʃip^h + ɨl /	→ [tʃi.p^{h}ɨl]	/tʃip^h/		/tʃip^h + to/	
	믿 ‘believe’	/mit + ɨl/	→ [mi.dɨl]	/mit/	→ [mit]	/mit + to/	→ [mit.t’o]
	밑 ‘bottom’	/mit^h + ɨl/	→ [mi.t^{h}ɨl]	/mit^h/		/mit^h + to/	
	박 ‘gourd’	/pak + ɨl/	→ [pa.gɨl]	/pak/	→ [pak]	/pak + kwa/	→ [pak.k’wa]
	밖 ‘outside’	/pak’ + ɨl/	→ [pa.k’ɨl]	/pak’/		/pak’ + kwa/	
1b)	굿 ‘exorcism’	/kus + ɨl/	→ [ku.sɨl]	/kus/	→ [kut]	/kus + to/	→ [kut.t’o]
	굳 ‘harden’	/kut + ɨl/	→ [ku.dɨl]	/kut/		/kut + to/	
	돛 ‘sail’	/totʃ^h + ɨl/	→ [to.tʃ^{h}ɨl]	/totʃ^h/	→ [tot]	/totʃ^h + to/	→ [tot.t’o]
	돋 ‘break out’	/tot + ɨl/	→ [to.dɨl]	/tot/		/tot + to/	
	맛 ‘taste’	/mas + ɨl/	→ [ma.sɨl]	/mas/	→ [mat]	/mas + to/	→ [mat.t’o]
	맞 ‘be hit’	/matʃ + ɨl/	→ [ma.dʒɨl]	/matʃ/		/matʃ + to/	
1c)	낮 ‘day’	/natʃ + ɨl/	→ [na.dʒɨl]	/natʃ/	→ [nat]	/natʃ + to/	→ [nat.t’o]
	낯 ‘face’	/natʃ^h + ɨl/	→ [na.tʃ^{h}ɨl]	/natʃ^h/		/natʃ^h + to/	

Table 2. Examples of three types of coda neutralization in Korean

3.2.2. Temporal-acoustic correlates of obstruents in Korean

In this section, I discuss how Korean obstruents contrast in their temporal-acoustic parameters. Based on previous findings, I will make predictions on the durational pattern of each acoustic parameter. The predictions are restricted to a non-neutralizing environment, since we expect that contrasts in acoustic parameters observed in a neutralizing environment may be in the same direction as in a non-neutralizing environment, but with the magnitude of contrasts considerably smaller. Therefore, the discussions on acoustic characteristics of obstruents are focused on durational patterns in an intervocalic position.

3.2.2.1. Vowel duration

As explained in section 3.2.1, coda obstruents are syllabified into the onset of the following syllable when followed by vowel-initial suffix and/or case markers (e.g., /-ɨl/: objective case marker). In this environment, voiceless lenis stops (/p, t, k/) become voiced (/b, d, g/)[7], whereas aspirated (/p^h, t^h, k^h/) and fortis stops (/p’, t’, k’/) are realized as they are.

According to previous studies on Korean stops (e.g., Chang, 2007; Chen, 1970; Sin, Shin, Kiaer, & Cha, 2012), vowels followed by lenis stops are longer before than those followed by aspirated or fortis stops. The duration of preceding vowel also varies depending on the place of articulation of the following consonant. Specifically, vowel duration is longest before velar stops and shortest before bilabial stops (Park, Shin, & Yang, 2002; T. Kim, 2014). Even though the

[7] Recent experimental studies (Jun, 1994; Kim, Lee, & Lotto, 2003) have argued that this is not always the case. For example, it has been shown that intervocalic lenis stops can be realized with voiceless stops when preceding high vowel is devoiced, even though they are mostly fully or partially voiced in between voiced sounds. Since voicing of intervocalic lenis stops was found to vary across- and within-talkers, they concluded that lenis stop voicing is not categorical, but gradient in nature.

effect of lenis vs. aspirated obstruents on preceding vowel has been established, it is mostly for stop consonants. Considering that affricates exhibit the same laryngeal distinctions as stops, however, we expect that vowels may be longer before lenis affricates than before aspirated affricates.

For relative duration of preceding vowel depending on the manner of articulation of coda, it has been shown that vowels are longer before fricatives than before stops or affricates (D.W. Kim, 2002; Park, 2013). This pattern is similar to that in other languages (Dutch: Ernestus & Baayen, 2007; English: House & Fairbanks, 1953; Peterson & Lehiste, 1960; French: Delattre, 1965). However, few studies made comparisons between stops and affricates in an intervocalic environment. Thus, little is known about whether vowel duration varies before stops vs. affricates.

3.2.2.1.1. Predictions

Laryngeal/Palatal neutralization. Based on the observations above, it is expected that vowel duration before lenis stops/affricates will be longer than those before non-lenis stops/affricates in a non-neutralizing environment. In terms of the place of articulation of coda, vowel duration is expected to be longest before velar, intermediate before alveolar, and shortest before bilabial stops.

Manner neutralization. For codas with manner distinctions, vowel duration is expected to be longer before fricatives than before stops or affricates. For affricate-stop pairs, the pattern of vowel duration in a non-neutralizing environment will serve as a reference to predict the pattern of vowel duration in a neutralizing environment.

3.2.2.2. Closure duration

Along with vowel duration, closure duration has been considered as an acoustically distinctive feature associated with Korean obstruents. In general, closure duration of tense obstruents in Korean is considerably longer than that of lenis or aspirated counterparts (Ahn, 1999b). This is consistent with the observation in other languages that closure duration has an inverse relation with vowel duration (e.g., German: Jessen, 2001). Specifically, for both stops and affricates, closure duration is longest for the tense, intermediate for the aspirated, and shortest for the lenis series in an intervocalic position (stops: Han, 1994; Kim & Lotto, 2002; M.-R.C. Kim, 1994; Silva, 1992; affricates: Kim, Honda, & Maeda, 2005). Closure duration also varies depending on the place of articulation of following stops. It has been shown that closure duration is generally ranked in the opposite order shown in vowel durations, that is, labials have longer closure duration than alveolars and velars (Repp, 1984).

In terms of the manner of following obstruents, stop-affricate pairs are the only ones that can be compared in closure duration in an intervocalic and coda positions. However, few studies have provided evidence for differences in closure duration between stops and affricates in those environments. Therefore, the durational pattern in a non-neutralizing environment will serve as a reference to predict the pattern of contrasts in a neutralizing environment.

3.2.2.2.1. Predictions

Laryngeal/Palatal neutralization. It is expected that closure durations will be longer before non-lenis stops/affricates than before lenis stops/affricates in a non-neutralizing environment. In terms of the place of articulation of coda, closure duration is expected to be longest before bilabial, shorter before alveolar, and shortest before velar stops.

Manner neutralization. The behavior of closure duration in affricate vs. stop pairs in a non-neutralizing environment will serve as a reference to predict the pattern of contrast in a neutralizing environment.

3.2.2.3. Voicing into closure duration

Compared with vowel duration and closure duration, relatively less attention has been given to voicing into closure duration as a relevant cue to voicing distinction in stops in an intervocalic or final position. Since the literature on Korean obstruents provides little information on the durational pattern of voicing into closure, I will base my predictions on previous studies that have examined voicing into closure duration either in an intervocalic or in final position (German: Port & O'Dell, 1985; Smith, Hayes-Harb, Bruss, & Harker, 2009; Hungarian: Gósy & Ringen, 2009; Russian: Dmitrieva et al., 2010; Kharlamov, 2014).

In an intervocalic position, it has been shown that in German voicing into closure is longer before lenis stops than before fortis stops (Fischer-Jorgensen, 1976; Mitleb, 1982). Across different places of articulation of stops, it is longest for bilabial, intermediate for velar, and shortest for alveolar (Mitleb, 1982). This pattern is also observed in Hungarian (Gósy & Ringen, 2009).

Durational patterns of voicing into closure in final position are similar to those in an intervocalic environment. Studies on final devoicing have demonstrated that underlyingly voiced obstruents have longer voicing into closure duration than voiceless counterparts (Dmitrieva et al., 2010; Kharlamov, 2014; Port & O'Dell, 1985). The same pattern is observed in Smith et al. (2009) and Jessen (1998). When the place of articulation is taken into account, it is longest for bilabials, intermediate for alveolars, and shortest for velars (Port & O'Dell, 1985). However, this behavior may not be reliable, since it is based on a single talker's productions.

3.2.2.3.1. Predictions

Laryngeal/Palatal neutralization. Based on limited observations on durational patterns, it is expected that voicing into closure duration will be longer before lenis obstruents than before non-lenis obstruents in a non-neutralizing environment. In terms of the place of articulation of following consonants, voicing into closure duration may be longer before bilabial than before alveolar and velar codas.

Manner neutralization. For codas with manner distinctions, affricate-stop pairs are the only ones that can be compared in terms of voicing into closure duration in an intervocalic and coda positions. Since little information on durational patterns is available, the behavior of voicing into closure duration in affricate-stop pairs in a non-neutralizing environment will serve as a reference to predict the pattern of contrasts in a neutralizing environment.

3.3. Methods

3.3.1. Stimuli

The target stimuli consisted of 28 minimal pair words that differed only in their coda obstruents (See Appendix 1). Twelve pairs contained lenis vs. non-lenis voiceless stops with laryngeal distinctions (/p-p^h/, /t-t^h/, /k-k'/). Another 16 pairs consisted of fricative-stop (/s-t/, /s-t^h/; n=4), affricate-stop (/$tʃ^h$-t/, /$tʃ^h$-t^h/; n=4) and fricative-affricate (/s-tʃ/, /s-$tʃ^h$/; n=6) pairs that are neutralized in the manner of articulation of coda. The remaining two pairs contained lenis vs. non-lenis palatal affricates (/tʃ-$tʃ^h$/) subject to palatal neutralization. Among the target items, there were three triplets (/nas-natʃ-natʃʰ/, /pis-pitʃ-pitʃʰ/, /mas-matʰ-matʃ/), each of which was split into 3 pairs (e.g., /nas-natʃ/, /nas-natʃʰ/, /natʃ-natʃʰ/).

All stimuli had CVC structure except two pairs having VC (/ip-ipʰ/, /up-upʰ/). The lexical category of each minimal pair was noun-noun (N-N), noun-verb (N-V), or verb-verb (V-V). In addition to the targets, seven minimal pairs unrelated to coda neutralization were included as fillers to disguise the purpose of the experiment. The total number of stimuli was 118 (47 target words x 2 phonological environments + 24 fillers).

Target items were placed both in a non-neutralizing and neutralizing environments. They were followed by a vowel-initial accusative case marker /ɨl/ in a non-neutralizing environment, and by a consonant-initial case marker (e.g., /kwa/ or /to/ for noun targets, /ko/ or /teon/ for verb targets) in a neutralizing environment. All of N-N and V-V pairs were followed by the same case marker, but the N-V pairs did not, since they cannot be used with the same suffix due to the difference in their lexical categories. In this case, however, the case marker still shared the same initial consonants (e.g., /mit + tən/, /mitʰ + to/). Additionally, the place of articulation of a following consonant was matched to that of the preceding coda in order to prevent assimilation of the coda to the following consonant. For the pairs with /p-pʰ/, however, the suffix beginning with /t/ was used, since non-lenis /pʰ/ is never followed by a consonant with the same place of articulation.

The stimuli were elicited from both a Q &A task and a reading task. In the Q & A task, participants were asked to listen to the questions recorded by the investigator who is a native Korean speaker and then provide answers that contain target words. Each question consisted of the dictionary definition of a target, an example of the context the target was used in, and the question followed by the format of the answer. All of the target words appeared in a sentence initial position. An example of the oral prompts (Q) and intended response (A) is presented below in English translation (See Appendix B for more examples). In the reading task, participants read a total of 118 sentences (94 targets, 24 fillers) that were identical to the answers in the Q & A task.

a) When followed by a vowel (= non-neutralizing environment)

Q: This is the time between sunrise and sunset. (Dictionary definition of the target)
The moon represents the night. Then what does the sun represent? (Question)
Please say "(The sun) represents this". (Instructions on how to answer)

A: (The sun) represents the day[8].

b) When followed by a consonant (= neutralizing environment)

Q: (Same definition as the above)
During the winter, it is very cold and this gets shorter too, since the sun sets earlier.
What does it get shorter, too?
Please say "This gets shorter, too".

A: The day gets shorter, too.

3.3.2. Participants

Fifteen native speakers of Korean (8 males, 7 females) participated in this study. In order to minimize influences of prior L2 experience on coda neutralization, participants who had stayed more than 1 year in English-speaking environment were excluded. All participants were undergraduate students at Hanyang University in Korea, ranging in age from 21 to 29 (average = 24.9). They were born and had lived in Seoul or Gyeonggi province in Korea and had used

[8] The word "The sun" is in the parenthesis, since it is typical that Korean language omits a subject or an object when it is already introduced in the previous context. By making use of this grammatical feature of Korean, we are able to control the position of stimuli in a way that all of them appear in a sentence-initial position.

standard Seoul Korean as their first language. They began English language study between 6 and 12 years old and had learned English as a foreign language for an average of 13.1 years in a school environment. The length of stay in English-speaking countries ranged from 0 to 4 months with a mean of 0.9 month. None of participants had any known history of speech and hearing impairments at the time of testing.

3.3.3. Procedure

The production experiment was conducted in a sound-attenuated booth in Hanyang Phonetics and Psycholinguistics Laboratory in Korea. The participants performed the Q & A task followed by the reading task, with each task repeated twice. All participants visited the lab twice, once for the first session of the Q & A task and once again for the remaining tasks. These two visits could occur on the same or separate days. In case they visited the lab on the same day, they were asked to visit with at least 2-hour interval in order to minimize learning effects and to avoid fatigue.

In the Q & A task, they were seated in front of a computer and instructed to listen to the questions presented via headphones and provide the answer by clicking the ‘ANSWER’ button on the screen. The button was programmed to stay inactive while the questions were being played in order to make sure that participants press it only after they finish listening to the questions. If they missed a question or were not sure of their answer, they replayed the question by clicking ‘REPEAT’ button or requested a hint from the experimenter. This strategy was effective enough to elicit 100% correct answers from all of the participants. Once participants provided the right answer, they were presented with the next trial by the experimenter. Prior to the first test session, participants were familiarized to the task in 8 practice trials. Each session of the Q & A task lasted for 40-50 minutes including a 5-minute mandatory break in between the blocks. In their second visit, participants took another 5-minute break after completing the second session of the Q & A task and before proceeding to the reading task.

Stimulus presentation was controlled with MAX/MSP software patches. The stimuli were presented in two blocks, each of which consisted of 59 trials (47 targets + 12 fillers). They were pseudo-randomized within and across the blocks in a way that the members of a minimal pair never appeared next to each other, as well as target items subject to the same type of neutralization did not appear more than 2 times in a sequence.

In the reading task, participants were asked to read a list of sentences that were identical to the answers in the Q & A task at a comfortable rate. The sentences were presented one at a time on a computer screen and their presentation was controlled by the participants. If they made any mistake while reading a sentence (e.g., speech error, pause, cough, etc.), they repeated it without interruption from the experimenter and proceeded to the next trial by pressing the 'NEXT' button located under the stimulus.

The stimuli were presented using List Presentation Application developed in the Department of Linguistics at Northwestern University, which presented each sentence in the uploaded text file in a sequence on the screen. As in the Q & A task, the sentence stimuli were pseudo-randomized and presented in two blocks. Each session of the reading task lasted for less than 10 minutes. The participants took a 1-minute mandatory break in between the blocks and another 5-minute break in between the sessions. The total duration of the entire experiment was 2 hours and 10 minutes in average.

For the two tasks, the recordings were made at a sampling rate of 44.1 kHz using a SHURE KSN44 dynamic microphone connected to a portable digital recorder (Tascam HD-P2).

The recorded speech files were segmented into individual sentences using Trigger Wave, an automatic audio file segmentation tool developed in the Department of Linguistics at Northwestern University. The sentence files were then leveled to equate root mean-square (RMS) amplitude using Praat software (Boersma & Weenink, 2014) for the acoustic measurements.

3.3.4. Acoustic measurements

Acoustic measurements were performed using Praat software. The measured acoustic parameters included preceding vowel duration, voicing into closure duration, and closure duration (Dinnsen & Charles-Luce, 1984; Dmitrieva et al., 2010; Jassem & Richter, 1989; Kharlamov, 2014; Kim & Jongman, 1996; Kopkalli, 1993; Piroth & Janker, 2004; Port & Crawford, 1989; Port & O'Dell, 1985; Shrager, 2012; Slowiaczek & Dinnsen 1985; Warner et al., 2004). For each parameter, the beginning and end points were first marked by hand by referring to both spectrogram and waveforms. The durations were then extracted by using Praat scripts.

Criteria for acoustic measurements were as follows: (1) Vowel duration: The interval from the onset of the periodicity in the waveform to the offset of the vowel which is signaled by a drop in amplitude and the end of F2 in the spectrogram. (2) Voicing into closure duration: The interval from the offset of the vowel to the end of glottal pulsing as reflected in the waveform and/or spectrogram. (3) Closure duration: The interval from the offset of the vowel where the glottal pulses become invisible to the onset of the release burst signaled by a sudden increase in the amplitude. Since Korean coda obstruents are usually unreleased, closure duration in this study was considered as the interval between the start of the closure of coda to the release burst of a following consonant. The examples of measurements are shown in Figure 1 and Figure 2.

(Note: vdur= vowel duration, voicing= voicing into closure duration, cdur= closure duration)

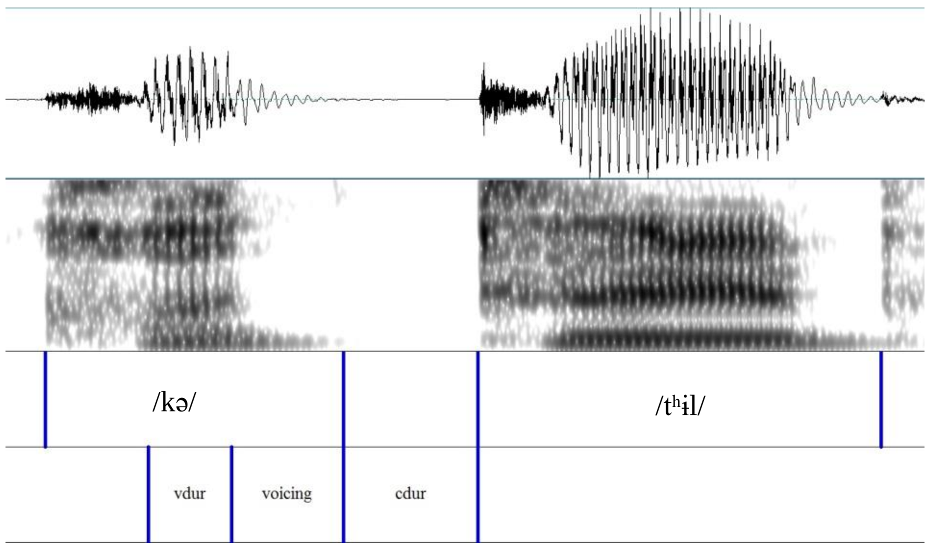

Figure 1. An example of measurements in a non-neutralizing environment. Boundaries of each acoustic parameter for the word /kətʰ/ followed by a vowel-initial suffix /ɨl/, which is re-syllabified into [kə.tʰɨl]

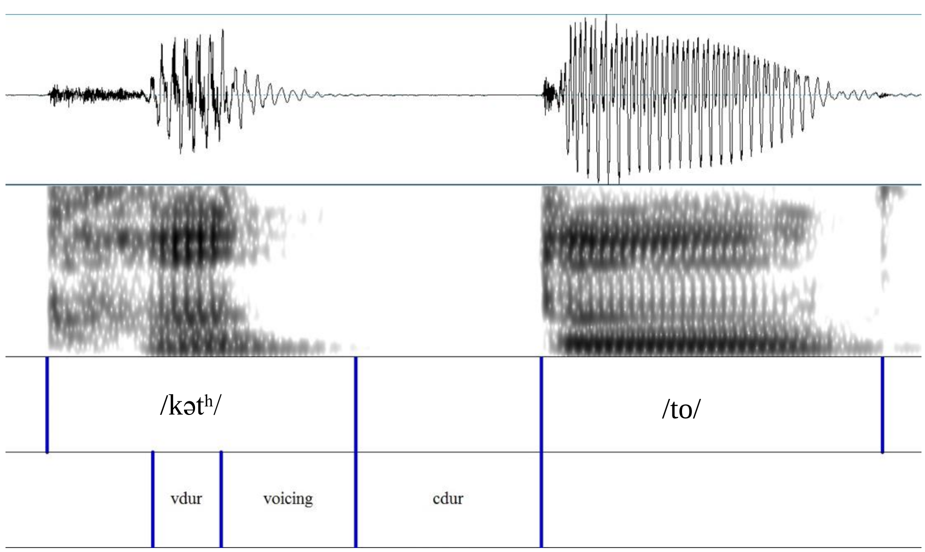

Figure 2. An example of measurements in a neutralizing environment. Boundaries of each acoustic parameter for the word /kətʰ/ followed by a consonant-initial suffix /to/

There are several important points to note with regard to measurements and analyses. In a non-neutralizing environment, /s/ is syllabified into the onset of the following syllable when followed by a vowel. For this reason, frication voicing instead of voicing into closure duration was measured for the fricative coda /s/. Consequently, the pairs containing /s/ subject to manner neutralization were not compared in terms of voicing into closure duration and closure duration.

The total number of stimuli submitted to acoustic measurements were 6691, which corresponded to 99.9% of the total stimuli elicited from the two tasks (28 pairs x 2 phonological environments (non-neutralizing vs. neutralizing) x 15 talkers x 2 tasks x 2 repetitions = 6720). Excluded trials (n=29) included cases where targets were produced with variant forms (n=12) or lacked certain acoustic cues e.g., no voicing into closure duration, n=17).

An example of productions with variant forms is that a talker produced a variant form /sus.eul/ for the canonical form /sut^{h}.eul/ ('charcoal' + objective case marker) in the first trial in the Q & A task, but not in the second trial in the Q & A task or in the reading task. This pattern of variant form productions suggests that the talker may not have the correct underlying form, but have multiple underlying forms of this particular word item in his/her lexicon (i.e., /sut^{h}/, /sus/). Therefore, I excluded the productions with variant forms (e.g., /sut^{h}/, n=6) along with the productions from their paired counterparts (n=6, e.g., /sut/) from the analysis. When doing cross-task comparisons (and only for those comparisons), the trials from the reading task matched with those variant productions in the Q & A task were excluded as well.

3.3.5. Statistical analyses

The measurement data was analyzed using linear mixed effects regression models (Baayen, Davidson, & Bates, 2008; Pinheiro & Bates, 2000). Three types of neutralization were analyzed separately, since they differed from each other in terms of how underlying coda obstruents were classified into sub-groups. For example, coda obstruents subject to laryngeal neutralization consist of underlyingly lenis and non-lenis coda across three places of articulation (bilabial, alveolar, velar), whereas those subject to manner neutralization are classified into three different types: fricative-stop (FS), affricate-stop (AS), and fricative-affricate (FA). Consequently, the fixed effects included in the regressions varied across three types of neutralization, although the two variables Task and Repetition were examined for all three types of neutralization. Table 3 shows the fixed effects examined in the regression models for the within- and across-task analyses for each type of neutralization.

<table>
<tr><th colspan="2" rowspan="2"></th><th colspan="2">Fixed effects</th></tr>
<tr><th>(a) Within-task analyses</th><th>(b) Across-task analyses</th></tr>
<tr><td colspan="2">Laryngeal</td><td>Underlying Coda (lenis, non-lenis)
POA (place of articulation of underlying coda: bilabial, alveolar, velar)</td><td rowspan="5">Task (Q & A, reading)
Repetition (1^{st}, 2^{nd})</td></tr>
<tr><td rowspan="3">Manner</td><td>FS</td><td>Underlying Coda (fricative, stop)</td></tr>
<tr><td>AS</td><td>Underlying Coda (affricate, stop)</td></tr>
<tr><td>FA</td><td>Underlying Coda (fricative, affricate)</td></tr>
<tr><td colspan="2">Palatal</td><td>Underlying Coda (lenis, non-lenis)</td></tr>
</table>

Table 3. Fixed effects tested in the two sets of regression analyses for each type of neutralization

Each of the fixed effects variables was contrast-coded. For the variable POA that has three levels (bilabial, alveolar, velar), it was divided into two contrasts in a way that the first contrast compares bilabial with alveolar and velar as a group (i.e., lingual), and the second compares alveolar with velar. All regression models included the maximal random effects structure justified by the experimental design (Bar, Levy, Scheepers, & Tily, 2013). Since the models varied in fixed effects and random effects, details of model structures for each of three types of neutralization will be discussed in the next section.

In cases where the analysis algorithm had difficulty fitting regression parameters, the model was selected by following backward stepwise elimination. Starting with the model with the maximal random effects structure, I constructed a series of reduced models by eliminating variables in an order of the correlations between random effects, random slopes for 3-way interactions, random slopes for 2-way interactions, all random slopes, and random intercept of item (where item was each minimal pair). Once the model converged, significance for fixed effects was assessed via model comparison.

3.3.5.1. Model structure

The first regression analyses were conducted to answer the question of whether there is any difference in durations of each acoustic parameter between the words with underlyingly different codas. The analyses were conducted separately for each of the two tasks (i.e., Q & A and reading) in each of the two phonological environments (i.e., non-neutralizing vs. neutralizing).

For laryngeal neutralization, the pairs with lenis and non-lenis coda obstruents were analyzed across the three places of articulation of coda (bilabial, alveolar, velar) for each acoustic parameter. The regression models included one of the acoustic parameters as a dependent variable. The fixed effects were Underlying Coda, POA, and the interactions between the two. The models also included random intercepts for talkers and items (as pairs) along with a random slope by talkers for the fixed effects.

The pairs subject to manner neutralization were analyzed in separate regressions, since they were composed of three different types of coda pair (FS, AS, FA). For example, the model to assess FS pairs included one of the acoustic parameters as a dependent variable, Underlying Coda (fricative vs. stop) as a fixed effect, and a random intercept for talkers along with a random slope by talkers for the fixed effect. The random effect of items was not included in any model, since the number of pair for each type was too small (n=4 for FS; n=4 for AS, n=6 for FA).

The targets subject to palatal neutralization were analyzed in a regression model which included one of the acoustic parameters as a dependent variable, Underlying Coda as a fixed effect, and a random intercept for talkers along with a random slope by talkers for the fixed effect. The random effect of items was excluded due to the small number of pairs (n=2).

The second regression analyses were aimed at examining whether durational differences in acoustic parameters between coda obstruents are reliably different across the two tasks or across the two production trials. These analyses were conducted separately, since the first analyses described above examined the effects within each task, while the second ones focused on revealing the effects within each phonological environment.

For all three types of neutralization, the regression model to assess the effect of task or repetition included one of the acoustic parameters as a dependent variable. The fixed effects were Underlying Coda (for laryngeal or palatal neutralization) or one of the three coda pairs (FS, AS, FA, for manner neutralization), Task, and the interaction between the two. Random intercepts for talkers and word pairs were also included along with random slopes by talkers and word pairs for the fixed effects. For the pairs subject to palatal neutralization, the random effect of items was excluded for the same reason as stated above. The model structure of regressions to assess the effect of repetition was exactly the same as the above, except that Task was replaced with Repetition.

3.4. Results

The results of regressions were reported for each acoustic parameter across three types of neutralization for both a non-neutralizing and neutralizing environments. In each section, descriptive statistics based on by-participant mean and standard error were followed by the results of regression analyses.

3.4.1. Vowel duration

3.4.1.1. Laryngeal neutralization

Table 4 shows the by-participant mean and standard error (in parentheses) of vowel duration before underlyingly lenis vs. non-lenis codas subject to laryngeal neutralization across three places of articulation, two phonological environments, and two tasks. It also presents the difference in mean vowel duration between lenis and non-lenis codas.

				Non-neutralizing environment			Neutralizing environment		
Task	**POA**	**Underlying Coda**		Lenis	Non-lenis	Difference	Lenis	Non-lenis	Difference
QA	Bilabial	/p/	/pʰ/	68.4 (2.8)	50.0 (2.2)	18.4 (1.8)	47.9 (2.0)	43.0 (1.9)	4.9 (1.1)
	Alveolar	/t/	/tʰ/	72.1 (2.0)	53.2 (1.8)	18.9 (1.9)	50.3 (1.5)	48.9 (1.6)	1.4 (1.0)
	Velar	/k/	/k'/	84.3 (2.8)	56.5 (2.0)	27.8 (2.0)	64.7 (2.1)	56.4 (2.0)	8.3 (1.0)
	Mean			76.0 (1.7)	53.5 (1.2)	**22.5** (1.2)***	55.5 (1.3)	50.0 (1.2)	**5.5** (0.7)**
RD	Bilabial	/p/	/pʰ/	65.7 (2.9)	47.9 (2.2)	17.8 (1.8)	48.4 (2.2)	47.9 (2.1)	0.5 (1.1)
	Alveolar	/t/	/tʰ/	68.3 (2.1)	53.3 (1.8)	15.0 (1.7)	51.3 (1.8)	54.0 (1.8)	-2.7 (1.3)
	Velar	/k/	/k'/	83.8 (2.9)	55.3 (2.2)	28.5 (1.9)	63.2 (2.2)	61.7 (2.2)	1.5 (1.0)
	Mean			73.9 (1.7)	52.3 (1.3)	**21.6** (1.1)***	55.3 (1.4)	55.2 (1.3)	**0.1** (0.6)

Table 4. By-participant mean and standard error of vowel duration before underlyingly lenis vs. non-lenis codas subject to laryngeal neutralization across three places of articulation (POA), two phonological environments, and two tasks. The difference in mean vowel duration highlighted in bold. (Note: significance level = * $p < 0.5$, ** $p < 0.01$, *** $p < 0.001$)

Analyses within task in each phonological environment. The results of the regression analyses revealed that the main effect of Underlying Coda was significant in a non-neutralizing environment in each task, as presented in the left panel of Figure 3. Vowels followed by lenis codas /p, t, k/ were significantly longer than those followed by non-lenis codas /pʰ, tʰ, k'/ in both tasks (QA: $\beta = 21.579$, SE = 2.996, $\chi^2(1) = 22.61$, $p < 0.001$; RD: $\beta = 20.424$, SE = 62.825, $\chi^2(1) = 23.45$, $p < 0.001$).

When neutralized, the underlying contrast between lenis and non-lenis codas was still differentiated by vowel duration in the same manner as in a non-neutralizing environment. This pattern of result was observed in the Q & A task ($\beta = 4.827$, SE = 0.829, $\chi^2(1) = 18.20$, $p < 0.001$), but not in the reading task ($\chi^2(1) = 0.08$, $p = 0.77$), as presented in the right panel of Figure 3.

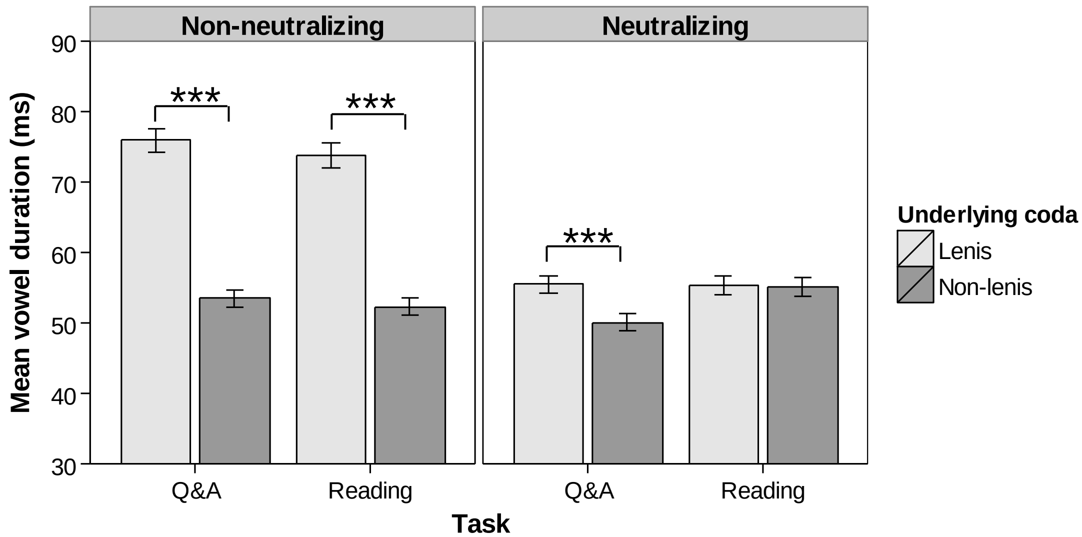

Figure 3. Mean vowel duration before underlyingly lenis vs. non-lenis codas subject to laryngeal neutralization in the Q & A and the reading tasks in a non-neutralizing (left) and neutralizing environment (right)

Figure 4 shows how vowel duration before lenis vs. non-lenis codas changes depending on the places of articulation of coda in a non-neutralizing environment. The interaction between Underlying Coda and POA was not significant in the Q & A task, suggesting that the difference in vowel duration between lenis and non-lenis codas did not vary across bilabial and lingual pairs (QA: χ^2 (1) = 0.78, p = 0.37) or across alveolar and velar pairs (QA: χ^2 (1) = 1.93, p = 0.16). In the reading task, the interaction was not significant across bilabial and lingual pairs (RD: χ^2 (1) = 0.56, p = 0.45), but reached significance across alveolar and velar pairs (RD: β = 13.551, SE = 15.204, χ^2 (1) = 4.48, p < 0.05). In order to examine the source of the interaction effect, follow-up regressions were performed for data subset separated by POA. The results showed that the difference in vowel duration was larger in velar pairs /k-k'/ than in alveolar pairs /t-t^h/ (alveolar: β = 14.954, SE = 2.061, χ^2(1) = 10.44, p < 0.01; velar: β = 28.505, SE = 4.967, χ^2(1) = 10.71, p < 0.01).

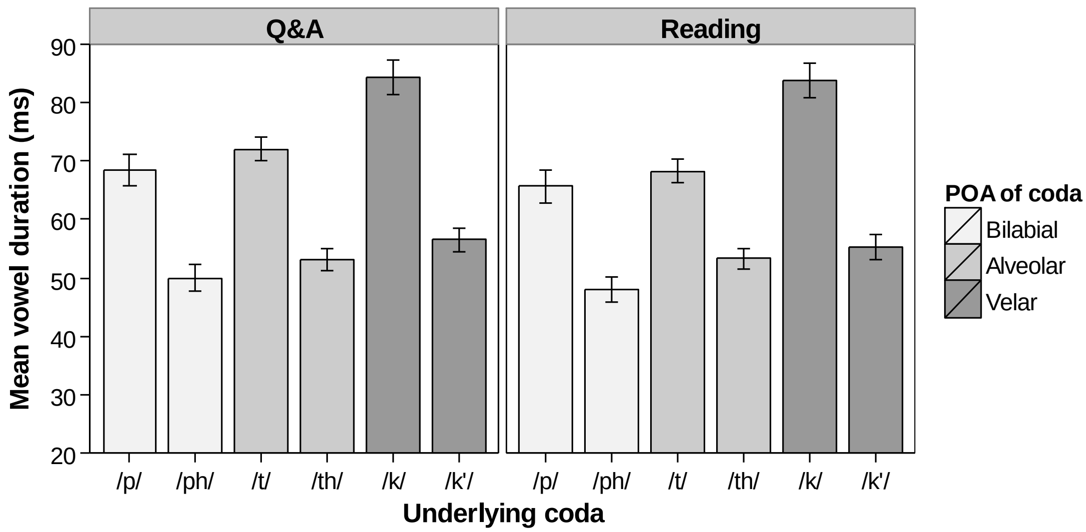

Figure 4. Mean vowel duration before underlyingly lenis vs. non-lenis codas subject to laryngeal neutralization across three places of articulation in the Q & A and the reading tasks (non-neutralizing environment)

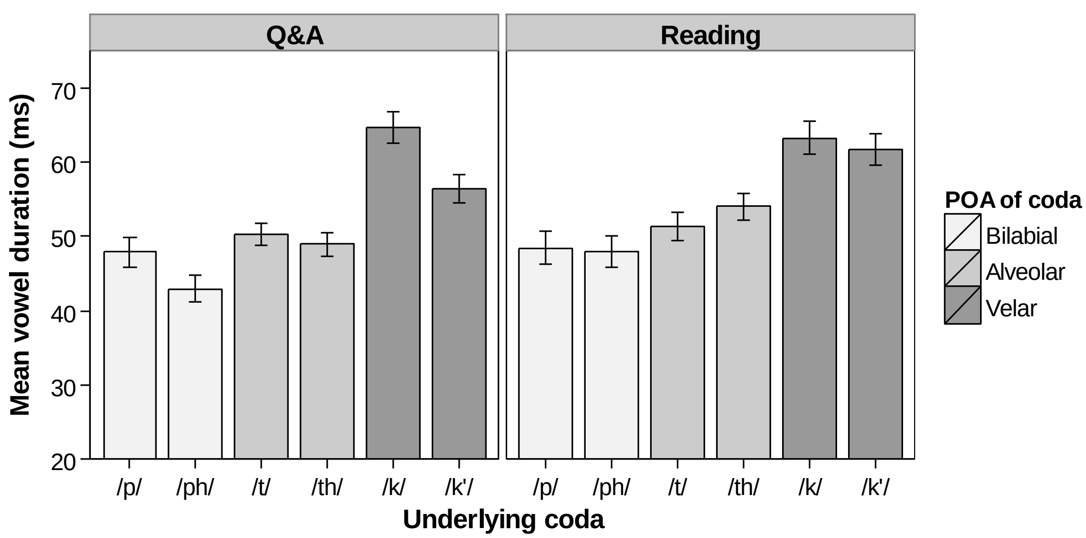

Figure 5. Mean vowel duration before underlyingly lenis vs. non-lenis codas subject to laryngeal neutralization across three places of articulation in the Q & A and the reading tasks (neutralizing environment)

(Note: ph = p^h, th = t^h)

In a neutralizing environment, the main effect of POA was not significant as shown in Figure 5 above,: vowel duration did not vary across bilabial and lingual pairs in either task (QA: $\chi^2(1) = 1.74$, $p = 0.18$; RD: $\chi^2(1) = 1.36$, $p = 0.24$) or across alveolar and velars (QA: $\chi^2(1) = 1.65$, $p = 0.19$; RD: $\chi^2(1) = 1.11$, $p = 0.29$). However, a significant interaction was observed between Underlying Coda and POA across alveolar and velar pairs in the both tasks (QA: $\beta = 6.996$, SE = 2.197, $\chi^2(1) = 7.77$, $p < 0.01$; RD: $\beta = 4.211$, SE = 1.854, $\chi^2(1) = 4.76$, $p < 0.05$). In the Q & A task, follow-up regressions for data subset separated by place of articulation yielded a significant main effect of Underlying Coda in velar pairs /k-k'/ ($\beta = 8.297$, SE = 1.709, $\chi^2(1) = 9.83$, $p < 0.01$) but not in alveolar pairs /t-tʰ/ ($\chi^2(1) = 1.28$, $p = 0.25$). This implies that the interaction effect originated from a significant difference in vowel duration between /k/ and /k'/. In the reading task, the difference in vowel duration was marginally significant in alveolar pairs ($\beta = -2.682$, SE = 1.285, $\chi^2(1) = 3.58$, $p = 0.05$), but not significant in velar pairs (velar: $\chi^2(1) = 1.07$, $p = 0.29$).

Analyses across tasks within each phonological environment. Follow-up regressions then compared performance across tasks within neutralizing and non-neutralizing environments. The interaction between Underlying Coda and Task was non-significant in a non-neutralizing environment ($\chi^2(1) = 0.30$, $p = 0.58$), suggesting that differences in vowel duration between lenis and non-lenis codas did not vary depending on the task. However, it reached significance in a neutralizing environment ($\beta = 5.327$, SE = 1.069, $\chi^2(1) = 15.16$, $p < 0.001$). Follow-up regressions performed for data subset separated by the task yielded a significant main effect of Underlying Coda in the Q & A task ($\beta = 5.439$, SE = 1.162, $\chi^2(1) = 13.44$, $p < 0.001$), but not in the reading task ($\chi^2(1) = 0.01$, $p = 0.89$). This implies that the observed interaction effect stemmed from a significant difference in vowel duration between lenis and non-lenis codas in the Q & A task.

Analyses across repetitions within each phonological environment. Finally, regressions examined whether there was an effect of repetition within each phonological environment. Across both environments, no interaction was found between Underlying Coda and Repetition, suggesting that differences in vowel duration between lenis and non-lenis codas did not vary across the first and the second productions (non-neutralizing: $\chi^2(1) < 0.01$, $p = 0.95$; neutralizing: $\chi^2(1) = 0.03$, $p = 0.86$).

3.4.1.2. Manner neutralization

Table 5 shows the by-participant mean and standard error (in parentheses) of vowel duration before each member of three types of coda pairs subject to manner neutralization across two phonological environments and two tasks. It also presents the difference in mean vowel duration between the members of each type of coda pairs.

	Underlying Coda (AB)	**Non-neutralizing environment**			**Neutralizing environment**		
		A	B	Difference	A	B	Difference
QA	Fricative - Stop (FS)	67.9 (2.7)	45.4 (2.3)	**22.5** (1.6)***	48.8 (2.4)	43.9 (2.2)	**4.9** (0.8)*
	Affricate - Stop (AS)	57.3 (3.1)	54.6 (2.0)	**2.7** (2.0)	50.4 (2.1)	48.8 (2.1)	**1.6** (1.0)
	Fricative - Affricate (FA)	89.2 (2.6)	79.2 (2.8)	**10.0** (1.8)**	64.0 (2.1)	61.6 (2.1)	**2.4** (0.8)
RD	Fricative - Stop (FS)	66.7 (2.8)	45.2 (2.4)	**21.5** (1.5)***	49.6 (2.5)	46.8 (2.4)	**2.8** (1.0)
	Affricate - Stop (AS)	60.4 (3.3)	55.8 (2.3)	**4.6** (2.1)	54.1 (2.2)	53.8 (2.2)	**0.3** (0.8)
	Fricative - Affricate (FA)	89.5 (2.6)	81.1 (2.8)	**8.4** (1.6)**	66.9 (2.2)	65.7 (1.9)	**1.2** (0.8)

Table 5. By-participant mean and standard error of vowel duration before each member of three types of coda pairs subject to manner neutralization across two phonological environments and two tasks. The differences in mean vowel duration highlighted in bold.

Analyses within task in each phonological environment. The results of regression analyses showed that the main effect of MOA was significant in FS and FA pairs in a non-neutralizing environment. Specifically, vowel duration was longer before fricative /s/ than before stops /t, tʰ/ (QA: $\beta = 22.514$, $SE = 2.357$, $\chi^2(1) = 33.19$, $p < 0.001$; RD: $\beta = 21.449$, $SE = 2.539$, $\chi^2(1) = 34.02$, $p < 0.001$), and longer before fricative /s/ than before affricates /tʃ, tʃʰ/ (QA: $\beta = 9.988$, $SE = 2.788$, $\chi^2(1) = 10.60$, $p < 0.01$; RD: $\beta = 8.409$, $SE = 2.726$, $\chi^2(1) = 8.50$, $p < 0.01$). Thee effect did not reach significance in AS pairs (QA: $\chi^2(1) = 0.76$, $p = 0.38$; RD: $\chi^2(1) = 2.70$, $p = 0.10$). These results are summarized in Figure 6.

In a neutralizing environment, the main effect of MOA was still significant in FS pairs, but only in the Q & A task, as shown in Figure 7 (QA: $\beta = 4.898$, $SE = 2.114$, $\chi^2(1) = 5.28$, $p < 0.05$; RD: $\chi^2(1) = 1.44$, $p = 0.22$). For AS and FA pairs, there was no significant difference in vowel duration in either task (AS_QA: $\chi^2(1) = 0.30$, $p = 0.58$; AS_RD: $\chi^2(1) = 0.01$, $p = 0.90$; FA_QA: $\chi^2(1) = 1.41$, $p = 0.23$; FA_RD: $\chi^2(1) = 0.40$, $p = 0.52$).

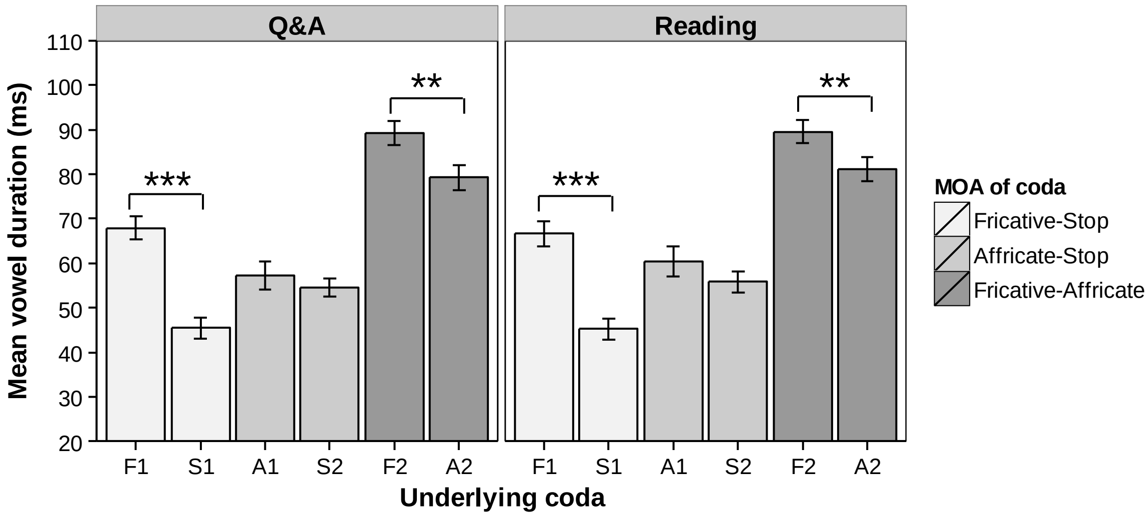

Figure 6. Mean vowel duration before each member of three types of coda pairs subject to manner neutralization in the Q & A and the reading tasks (non-neutralizing environment) (Note: F1/F2 = fricatives, S1/S2 = stops, A1/A2 = affricates)

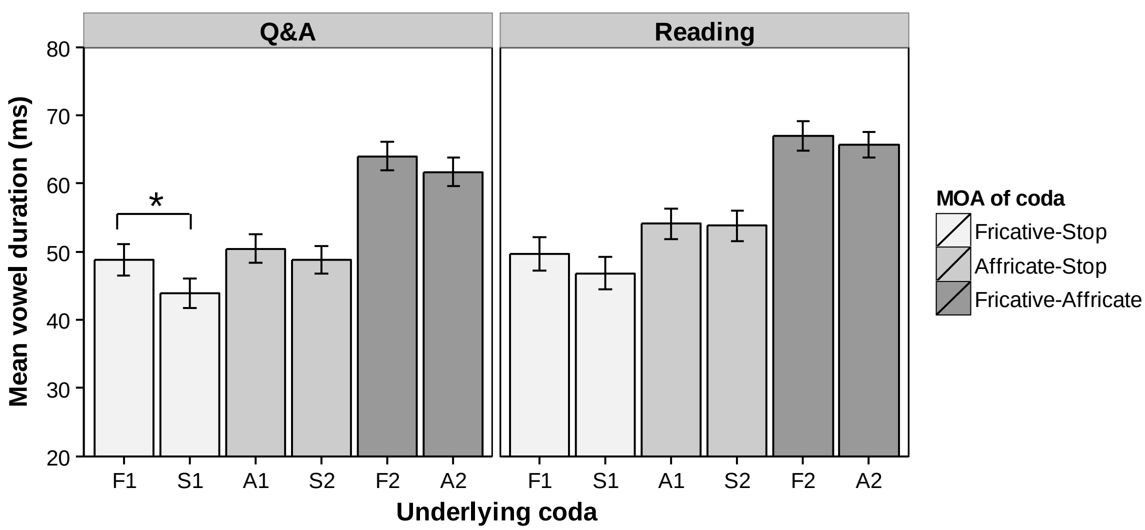

Figure 7. Mean vowel duration before each member of three types of coda pairs subject to manner neutralization in the Q & A and the reading tasks (neutralizing environment)

<u>Analyses across tasks/repetitions within each phonological environment</u>. The interaction between MOA and Task was not significant. For each member of three types of coda pairs,

differences in vowel duration did not vary across the two tasks either in a non-neutralizing environment (FS: $\chi^2(1) = 0.09$, p = 0.75; AS: $\chi^2(1) = 0.41$, p = 0.52; FA: $\chi^2(1) = 0.19$, p = 0.66) or in a neutralizing environment (FS: $\chi^2(1) = 0.46$, p = 0.49; AS: $\chi^2(1) = 0.08$, p = 0.77; FA: $\chi^2(1) = 0.15$, p = 0.69). The interaction between MOA and Repetition did not reach significance, either. There was no significant difference in vowel duration across the first and the second productions either in a non-neutralizing environment (FS: $\chi^2(1) = 0.20$, p = 0.64; AS: $\chi^2(1) = 0.08$, p = 0.77; FA: $\chi^2(1) < 0.01$, p = 0.95) or in a neutralizing environment (FS: $\chi^2(1) = 0.20$, p = 0.64; AS: $\chi^2(1) = 0.01$, p = 0.91; FA: $\chi^2(1) = 0.04$, p = 0.83).

3.4.1.3. Palatal neutralization

Table 6 shows the by-participant mean and standard error (in parentheses) of vowel duration before lenis vs. non-lenis affricate codas subject to palatal neutralization across two phonological conditions and two tasks. It also presents the difference in vowel duration between lenis and non-lenis codas.

	Underlying Coda		**Non-neutralizing environment**			**Neutralizing environment**		
	Lenis	Non-lenis	Lenis	Non-lenis	Difference	Lenis	Non-lenis	Difference
QA	/tʃ/	/tʃʰ/	86.8 (5.0)	58.4 (4.2)	**28.4** (2.5)***	58.2 (3.8)	57.7 (4.2)	**0.5** (1.6)
RD	/tʃ/	/tʃʰ/	88.0 (5.5)	61.4 (3.9)	**26.6** (2.3)***	65.3 (3.9)	62.1 (3.9)	**3.2** (1.2)

Table 6. By-participant mean and standard error of vowel duration before lenis vs. non-lenis codas subject to palatal neutralization across two phonological environments and two tasks. The difference in mean vowel duration highlighted in bold.

Analyses within task in each phonological environment. The regression results showed that the main effect of Underlying Coda was significant in a non-neutralizing environment;

vowels were longer before lenis affricate /tʃ/ than before aspirated affricate /tʃʰ/ in the both tasks (QA: $\beta = 28.453$, SE = 4.599, $\chi^2(1) = 25.27$, $p < 0.001$; RD: $\beta = 26.567$, SE = 4.786, $\chi2(1) = 22.69$, $p < 0.001$). In a neutralizing environment, however, the effect was non-significant in either task (QA: $\chi^2(1) = 0.01$, $p = 0.89$; RD: $\chi^2(1) = 0.70$, $p = 0.40$), suggesting that the underlying contrast between /tʃ/ and /tʃʰ/ was not differentiated by vowel duration. These results are summarized in Figure 8 and Figure 9.

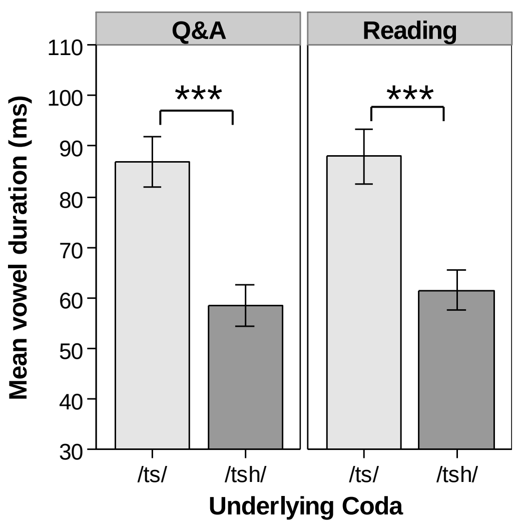

Figure 8. Mean vowel duration before lenis vs. non-lenis affricates subject to palatal neutralization in the Q & A and the reading tasks (non-neutralizing environment)

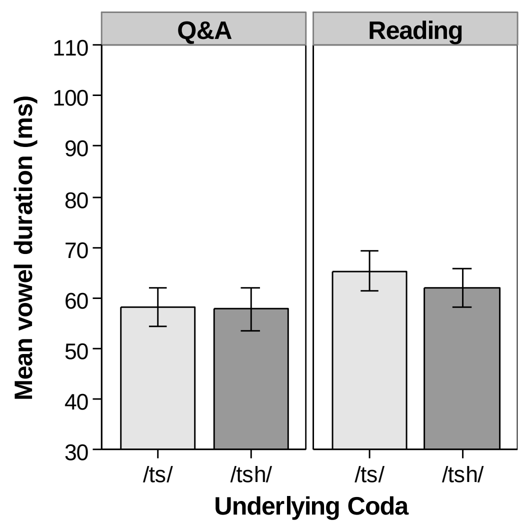

Figure 9. Mean vowel duration before lenis vs. non-lenis affricates subject to palatal neutralization in the Q & A and the reading tasks (neutralizing environment)

(Note: ts=tʃ, tsh=tʃʰ)

<u>Analyses across tasks, repetitions within each phonological environment.</u> The interaction between Underlying Coda and Task was not significant in either environment, indicating that difference in vowel duration between lenis and non-lenis codas was not significantly different

across the two tasks (non-neutralizing: $\chi^2(1) = 0.31$, $p = 0.57$; neutralizing: $\chi^2(1) = 1.72$, $p = 0.18$). Finally, there was no interaction between Underlying Coda and Repetition in either environment (non-neutralizing: $\chi^2(1) = 0.21$, $p = 0.64$; neutralizing: $\chi^2(1) = 0.12$, $p = 0.72$).

3.4.1.4. Discussion

Across all three types of coda neutralization, vowel duration served as a reliable acoustic cue to distinguish underlying different coda obstruents in a non-neutralizing environment in the both tasks. For codas with laryngeal contrasts, vowels before lenis stops /p, t, k/ were longer by 21.6 ms (42% in RD) ~ 22.5 ms (41% in QA) than those before non-lenis counterparts /p^h, t^h, k'/. This result replicates the finding of previous studies that vowels are longer before lenis stops than before aspirated or fortis stops in Korean (Chang, 2007; Chen, 1970; Shin, Kiaer, & Cha, 2013).

Vowel duration was shortest before bilabial stops /p, p^h/ and longest before velar stops /k, k'/, as predicted based on previous observations (Park, Shin, & Yang, 2002; T. Kim, 2014). Differences in vowel duration between lenis and non-lenis codas were still largest in velar pairs /k-k'/ (QA: 27.8 ms ~ RD: 28.5 ms). However, there was little difference between bilabial /p-p^h/and alveolar pairs /t-t^h/ in the Q & A task (18.4 ms vs. 18.9 ms), although bilabial pairs showed a slightly larger difference than alveolar pairs in the reading task (17.8 ms vs. 15.0 ms). Consequently, the interaction between Underlying Coda and POA was only significant across alveolar-velar contrast in the reading task, due to a larger difference in vowel duration in velar pairs /k-k'/.

For codas subject to manner neutralization, it was established that vowels preceding fricative /s/ were longer by 21.5 ms (50% in RD) ~ 22.5 ms (48% in QA) than those preceding stops /t, t^h/, which is consistent with previous findings (D.W. Kim, 2002; Park, 2013). Part of these findings is also in line with Kim and Jongman (1996) which showed that vowels preceding /s/ were longer than those preceding /t^h/ by 33 ms. However, no significant difference in vowel duration was observed in affricate-stop pairs (/tʃ-t/, /tʃ-t^h/, /tʃ^h-t^h/). Another finding in a non-neutralizing environment was that lenis and non-lenis affricate pairs /tʃ-tʃ^h/ subject to palatal neutralization showed a significant difference in vowel duration, such that vowels were longer by 26.6 ms (43% in RD) ~ 28.4 ms (49% in QA) before lenis than before non-lenis codas.

Overall, the observed differences in vowel duration across three types of coda neutralization were not restricted to one experimental task. Therefore, it seems that vowel duration is a robust cue to distinguish underlying contrast of coda obstruents in a non-neutralizing environment.

In a neutralizing environment, differences in vowel duration were still significant between codas with laryngeal contrasts. The direction of the difference between lenis and non-lenis codas was the same as in a non-neutralizing environment: vowels followed by lenis coda were longer than those followed by non-lenis. The magnitude of the difference was 24% (5.5 ms) of the contrast in a non-neutralizing environment. It was slightly larger than reported in Warner et al. (2004), which was 14% - 20%.

Within a neutralizing environment, underlyingly lenis codas were 10% longer in vowel duration than non-lenis codas. This falls well within the previously reported range 3.5 - 15 ms (2% - 17%) for incomplete neutralization (e.g., 15 ms in Port & O'Dell (1985); 15 ms (17%) in Charles-Luce (1993); 8 ms (5%) in Dmitrieva et al. (2010); 8.6 ms in Roettger et al. (2014); 6-7 ms (10%) in Slowiaczek & Dinnsen (1985); 3.5 - 4 ms (2 - 3%) in Warner et al. (2004)). Considering that most of the previous studies have observed significant effects of incomplete neutralization with a relatively small sample size, our results are encouraging in that we obtain statistically significant effects with a sufficiently large number of subjects and items.

As in a non-neutralizing environment, vowels preceding bilabial stops /p, p^h/ were shortest and those preceding velar stops /k, k'/ were longest. However, the magnitude of the difference in vowel duration varied depending on the task. In the Q & A task, the difference in vowel duration was smallest in alveolar pairs /t-t^h/ (1.4 ms) and largest in velar pairs /k-k'/ (8.3 ms). In the reading task, on the other hand, bilabial pairs /p-p^h/ showed the smallest difference (0.5 ms) and alveolar pairs /t-t^h/ exhibited largest difference with the direction of difference opposite to the Q & A task (i.e., non-lenis longer by -2.7 ms than lenis codas). Consequently, the interaction between Underlying Coda and POA across alveolar-velar contrast stemmed from different sources in each task. In the Q & A task, the effect was due to larger difference in vowel duration in velar pairs, whereas it originated form larger difference in alveolar pairs in the reading task.

Another critical finding of this study is that vowel duration differences in the pairs with laryngeal contrasts were observed only in the Q & A task in which explicit orthographic cues were not present. This shows that as far as laryngeal neutralization is concerned, incomplete neutralization is influenced by the experimental task. It appears that incomplete neutralization is more likely to occur in a communicative-driven task than in a reading task that explicitly exposes the participants to the underlying coda contrast as represented in the orthography.

With regard to manner neutralization, we obtained mixed results: FS pairs were incompletely neutralized, while AS and FA pairs were completely neutralized in terms of vowel duration. Specifically, vowels before fricative /s/ were significantly longer by 4.9 ms (11%) than those before stops /t, t^h/. When FS pairs were analyzed separately, the difference in vowel duration was still significant in /s-t/ pairs ($\beta = 6.207$, SE = 1.516, $\chi^2(1) = 11.25$, $p < 0.001$), and marginally significant in /s-tʰ/ pairs ($\beta = 4.462$, SE = 2.515, $\chi^2(1) = 3.11$, $p = 0.07$). The finding on /s-t/ pair is not consistent with Kim and Jongman (1996) in which this pair was completely neutralized in vowel duration. However, it should be noted that the number of /s-t/ pair was very small in this study (n=1), while there were 11 pairs in Kim and Jongman. It is therefore possible that incomplete neutralization observed in this particular pair could be item-specific effect.

Nevertheless, the combined results suggest that some items subject to manner neutralization show a trend toward incomplete neutralization. For the AS and FA pairs and the pairs subject to palatal neutralization, we did not observe a significant difference in vowel duration. The mixed results may be attributed to the small number of target pairs that are not enough to yield statistically reliable durational patterns (AS: n=4, FA: n=6, Palatal: n=2). Therefore, the interpretation of these results should be exercised with caution.

In sum, it is revealed that all types of coda pairs with underlying laryngeal or manner contrast are distinguished by vowel duration in a non-neutralizing environment. When neutralized, the underlying contrast is still maintained in some of the pairs. It is clearly shown that a significant difference in vowel duration exists in coda pairs with laryngeal contrasts, particularly in the Q & A task in which explicit orthographic cues are absent. Some pairs with manner contrast also show a trend toward incomplete neutralization. However, the rest of the pairs subject to manner/palatal neutralization is completely neutralized, exhibiting no significant difference in vowel duration.

3.4.2. Voicing into closure duration

3.4.2.1. Laryngeal neutralization

Table 7 shows the by-participant mean and standard error (in parentheses) of voicing into closure duration before lenis vs. non-lenis codas subject to laryngeal neutralization across three places of articulation, two phonological environments, and two tasks. It also presents the difference in mean voicing into closure duration between lenis and non-lenis codas.

				Non-neutralizing environment			Neutralizing environment		
Task	**POA**	**Coda**		Lenis	Non-lenis	Difference	Lenis	Non-lenis	Difference
QA	Bilabial	/p/	/pʰ/	50.1 (1.7)	45.5 (2.1)	4.6 (2.3)	49.9 (2.4)	50.0 (2.7)	0.1 (1.4)
	Alveolar	/t/	/tʰ/	47.1 (1.9)	45.6 (2.1)	1.5 (2.0)	40.3 (2.8)	53.4 (2.1)	-13.1 (2.2)
	Velar	/k/	/k'/	43.4 (1.2)	38.4 (1.8)	5.0 (2.0)	41.7 (1.9)	49.9 (3.3)	-8.2 (3.0)
	Mean			46.6 (0.9)	42.6 (1.2)	**4.0** (1.2)	44.1 (1.3)	49.6 (1.3)	**-5.5** (1.0)***
RD	Bilabial	/p/	/pʰ/	52.0 (2.0)	45.1 (2.3)	6.9 (2.5)	45.3 (2.2)	47.7 (2.2)	-2.4 (1.4)
	Alveolar	/t/	/tʰ/	49.3 (2.1)	41.6 (2.5)	7.7 (2.9)	41.3 (2.6)	44.6 (2.7)	-3.3 (2.0)
	Velar	/k/	/k'/	45.0 (1.3)	39.7 (1.9)	5.3 (2.1)	38.5 (1.8)	41.9 (1.9)	-3.4 (1.4)
	Mean			48.4 (1.0)	42.0 (1.3)	**6.4** (1.4)	41.4 (1.2)	44.5 (1.3)	**-3.1** (0.9)

Table 7. By-participant mean and standard error of voicing into closure duration before underlyingly lenis vs. non-lenis codas subject to laryngeal neutralization across three places of articulation (POA), two phonological environments, and two tasks. The difference in mean voicing into closure duration highlighted in bold.

Analyses within task in each phonological environment. In a non-neutralizing environment, the main effect of Underlying Coda was not significant in the Q & A task (QA: $\chi^2(1) = 2.30$, $p = 0.12$) or marginally significant in the reading task (RD: $\chi^2(1) = 3.14$, $p = 0.07$). In a neutralizing environment, on the other hand, the effect was significant in the Q & A task ($\beta = -6.021$, $SE = 1.439$, $\chi^2(1) = 11.47$, $p < 0.001$): voicing into closure duration before lenis coda was significantly longer than those before non-lenis coda. In the reading task, the effect was marginally significant ($\chi^2(1) = 3.43$, $p = 0.06$). The results are summarized in Figure 10.

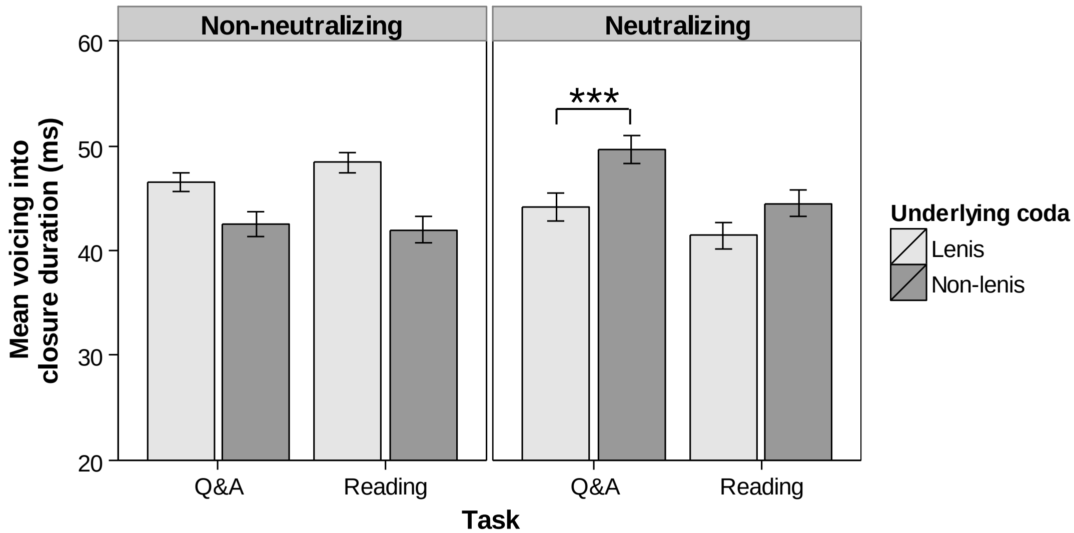

Figure 10. Mean voicing into closure duration before lenis vs. non-lenis codas subject to laryngeal neutralization in the Q & A and the reading tasks in a non-neutralizing (left) and neutralizing environment (right)

Figure 11 presents how voicing into closure duration between lenis and non-lenis codas changes across three places of articulation of coda in a non-neutralizing environment. The main effect of POA did not reach significance in either task. Voicing into closure duration did not significantly vary across bilabial and lingual pairs (QA: $\chi^2(1) = 1.75$, p = 0.18; RD: $\chi^2(1) = 1.29$, p = 0.25) or across alveolar and velar pairs (QA: $\chi^2(1) = 1.97$, p = 0.16; RD: $\chi^2(1) = 0.42$, p = 0.51). The interaction between Underlying Coda and POA did not reach significance in either task. This indicates that differences in voicing into closure duration between lenis and non-lenis codas do not vary across bilabial and lingual pairs (QA: $\chi^2(1) = 0.33$, p = 0.56; RD: $\chi^2 (1) = 0.01$, p = 0.90) or across alveolar and velar pairs (QA: $\chi^2 (1) = 1.05$, p = 0.30; RD: $\chi^2 (1) = 0.25$, p = 0.61).

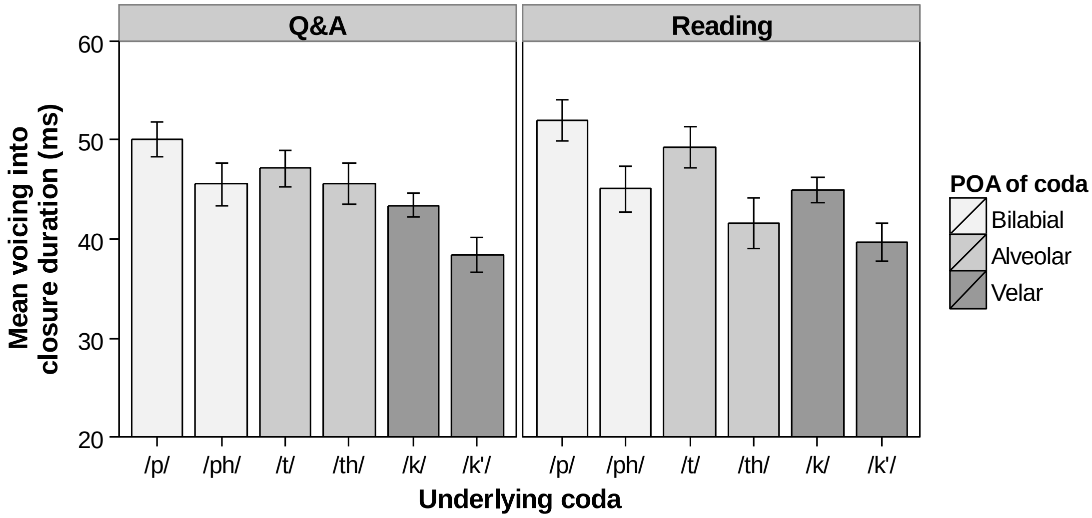

Figure 11. Mean voicing into closure duration before lenis vs. non-lenis codas subject to laryngeal neutralization across three places of articulation in the Q & A and the reading tasks (non-neutralizing environment)

In a neutralizing environment, the main effect of POA was non-significant (See Figure 12). Difference in voicing into closure duration did not significantly differ across bilabial and lingual pairs (QA: $\chi^2(1) = 0.65$, $p = 0.41$; RD: $\chi^2(1) = 0.80$, $p = 0.36$) or across alveolar and velar pairs (QA: $\chi^2(1) = 0.21$, $p = 0.64$; RD: $\chi^2(1) = 0.19$, $p = 0.65$).

However, a significant interaction between Underlying Coda and POA was observed across bilabial and lingual pairs in the Q & A task ($\beta = -12.059$, $SE = 3.924$, $\chi^2(1) = 7.19$, $p < 0.01$). Follow-up regressions for data subset separated by place of articulation yielded a significant effect of Underlying Coda in lingual pairs ($\beta = -8.135$, $SE = 2.256$, $\chi^2(1) = 7.82$, $p < 0.01$), but not in bilabial pairs ($\chi^2(1) < 0.001$, $p = 0.98$). Therefore, the source of interaction effect was a larger difference in voicing into closure duration in lingual codas. In the reading task, the interaction between Underlying Coda and POA did not reach significance (bilabial-lingual: $\chi^2(1) = 0.06$, $p = 0.80$; alveolar-velar: $\chi^2(1) < 0.001$, $p = 0.97$).

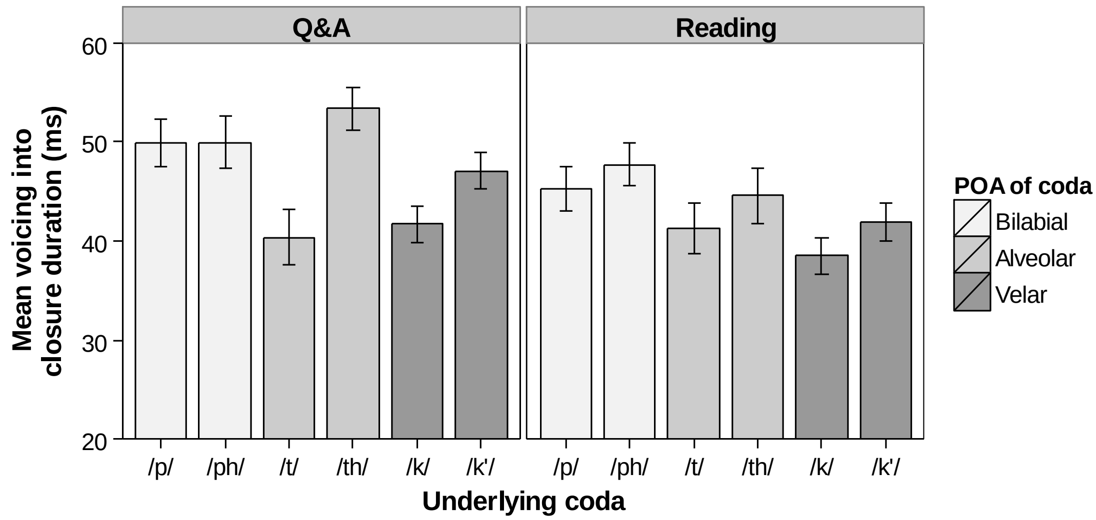

Figure 12. Mean voicing into closure duration before lenis vs. non-lenis codas subject to laryngeal neutralization across three places of articulation in the Q & A and the reading tasks (neutralizing environment)

<u>Analyses across tasks, repetitions within each phonological environment.</u> The interaction between Underlying Coda and Task was not significant in either environment (non-neutralizing: $\chi^2(1) = 1.82$, $p = 0.17$; neutralizing: $\chi^2(1) = 1.38$, $p = 0.23$), suggesting that difference in voicing into closure duration between lenis and non-lenis codas did not vary across the two tasks either in a non-neutralizing or in a neutralizing environment. The interaction between Underlying Coda and Repetition was not significant, either (non-neutralizing: $\chi^2(1) = 0.25$, $p = 0.61$; neutralizing: $\chi^2(1) = 0.17$, $p = 0.67$).

3.4.2.2. Manner neutralization

Table 8 shows the by-participant mean and standard error (in parentheses) of voicing into closure duration before each member of three types of coda pairs subject to manner neutralization across two phonological environments and two tasks. It also presents the difference in mean voicing into closure duration between the members of each type of coda pairs. As stated in section 3.3, voicing into closure duration in a non-neutralizing environment was not compared in the pairs with fricative /s/ in which frication voicing instead of voicing into closure duration was available for measurement.

	Underlying Coda (AB)	**Non-neutralizing environment**			**Neutralizing environment**		
		A	B	Difference	A	B	Difference
QA	Fricative - Stop (FS)	N/A	N/A	N/A	47.6 (2.1)	46.1 (2.0)	**1.5** (1.5)
	Affricate - Stop (AS)	43.1 (2.0)	45.2 (1.7)	**-2.1** (2.2)	44.7 (2.4)	44.1 (2.1)	**0.6** (1.8)
	Fricative - Affricate (FA)	N/A	N/A	N/A	44.1 (1.8)	43.1 (1.9)	**1.0** (1.4)
RD	Fricative - Stop (FS)	N/A	N/A	N/A	50.3 (2.0)	46.5 (2.1)	**3.8** (1.4)
	Affricate - Stop (AS)	41.1 (2.2)	44.9 (1.8)	**-3.8** (2.1)	44.7 (2.1)	42.1 (2.2)	**2.6** (1.4)
	Fricative - Affricate (FA)	N/A	N/A	N/A	40.7 (1.8)	40.6 (1.7)	**0.1** (1.4)

Table 8. By-participant mean and standard error of voicing into closure duration before each member of three types of coda pairs subject to manner neutralization across two phonological environments and two tasks. The difference in mean voicing into closure duration highlighted in bold.

<u>Analyses within task in each phonological environment.</u> In a non-neutralizing environment, the main effect of MOA in AS pairs was not significant in the Q & A task ($\chi^2(1)$ = 0.30, p = 0.58) or marginally significant in the reading task ($\chi^2(1)$ = 2.70, p = 0.09), suggesting that there was no significant difference in voicing into closure duration between /tʰ/ and /tʃʰ/.

In a neutralizing environment, the main effect of MOA was not significant in most of pairs, as shown in Figure 13. In FS pairs, differences in voicing into closure duration were not significant in the Q & A task ($\chi^2(1)$ = 0.70, p = 0.40) and marginally significant in the reading task ($\chi^2(1)$ = 3.72, p = 0.05). For AS and FA pairs, there was no difference in voicing into closure duration between each member of the pairs (AS_QA: $\chi^2(1) = 0.12$, $p = 0.72$; AS_RD: $\chi^2(1) = 1.51$, $p = 0.21$; FA_QA: $\chi^2(1) = 0.29$, $p = 0.58$; FA_RD: $\chi^2(1) = 1.51$, $p = 0.21$).

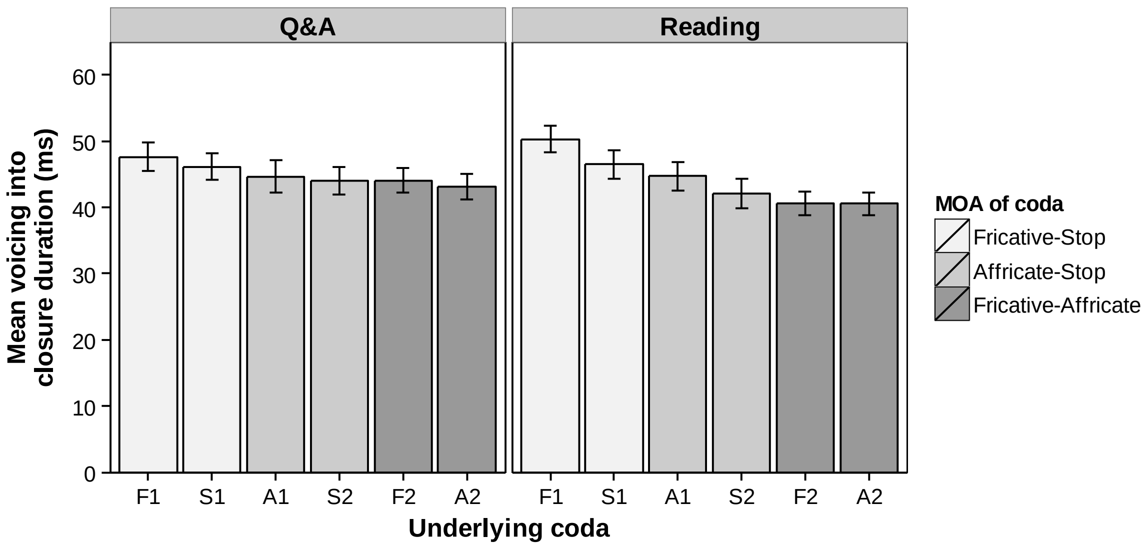

Figure 13. Mean voicing into closure durations before lenis vs. non-lenis member of three types of coda pairs subject to manner neutralization in the Q & A and the reading tasks (neutralizing environment)

<u>Analyses across tasks, repetitions within each phonological environment.</u> In a non-neutralizing environment, differences in voicing into closure duration in AS pairs did not vary across the two tasks ($\chi^2(1) = 0.30$, $p = 0.58$) or across the two productions ($\chi^2(1) = 1.55$, $p = 0.21$). In a neutralizing environment, the interaction between MOA and Task was not significant in each type of coda pairs (FS: $\chi^2(1) = 0.70$, $p = 0.40$; AS: $\chi^2(1) = 0.33$, $p = 0.55$; FA: $\chi^2(1) = 0.17$, $p = 0.67$). The interaction between MOA and Repetition was not significant, either (FS: $\chi^2(1) = 0.04$, $p = 0.82$; AS: $\chi^2(1) = 0.01$, $p = 0.90$; FA: $\chi^2(1) = 0.44$, $p = 0.50$).

3.4.2.3. Palatal neutralization

Table 9 shows the by-participant mean and standard error (in parentheses) of voicing into closure duration before lenis vs. non-lenis affricate codas subject to palatal neutralization across two phonological conditions and two tasks. It also presents the difference in mean voicing into closure duration between lenis and non-lenis codas.

	Underlying Coda		**Non-neutralizing environment**			**Neutralizing environment**		
	Lenis	Non-lenis	Lenis	Non-lenis	Difference	Lenis	Non-lenis	Difference
QA	/tʃ/	/tʃʰ/	38.6 (2.5)	44.2 (2.6)	**-5.6** (2.8)	44.8 (3.2)	46.5 (3.1)	**-1.7** (2.4)
RD	/tʃ/	/tʃʰ/	34.6 (2.1)	39.1 (2.8)	**-4.5** (2.8)	40.4 (3.4)	41.7 (2.8)	**-1.3** (2.5)

Table 9. By-participant mean and standard error of voicing into closure duration before lenis vs. non-lenis codas subject to palatal neutralization across two phonological environments and two tasks. The difference in mean voicing into closure duration highlighted in bold.

<u>Analyses within task in each phonological environment.</u> In a non-neutralizing environment, the effect of Underlying Coda was marginally significant in the Q & A task (β = -5.635, SE = 2.944, $\chi^2(1) = 3.27$, p = 0.07) or non-significant in the reading task ($\chi^2(1) = 1.97$, p = 0.16). In a neutralizing environment, there was no significant difference in voicing into closure duration between /tʃ/ and /tʃʰ/ in each task (QA: $\chi^2(1) = 0.29$, p = 0.58; RD: $\chi^2(1) = 0.15$, p = 0.69). These results are summarized in Figure 14 and Figure 15.

<u>Analyses across tasks, repetitions within each phonological environment.</u> The interaction between Underlying Coda and Task was not significant in either environment (non-neutralizing: $\chi^2(1) = 0.09$, p = 0.75; neutralizing: $\chi^2(1) < 0.01$, p = 0.92). The interaction between Underlying Coda and Repetition was not significant, either (non-neutralizing: $\chi^2(1) = 0.01$, p = 0.92; neutralizing: $\chi^2(1) = 0.01$, p = 0.90).

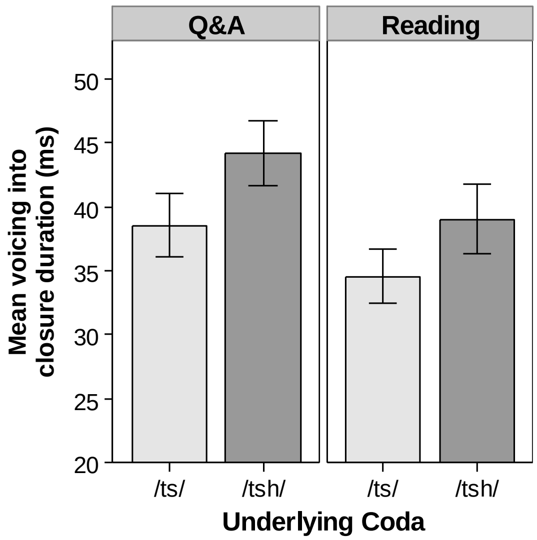

Figure 14. Mean voicing into closure duration before lenis vs. non-lenis affricates subject to palatal neutralization in the Q & A and the reading tasks (non-neutralizing environment)

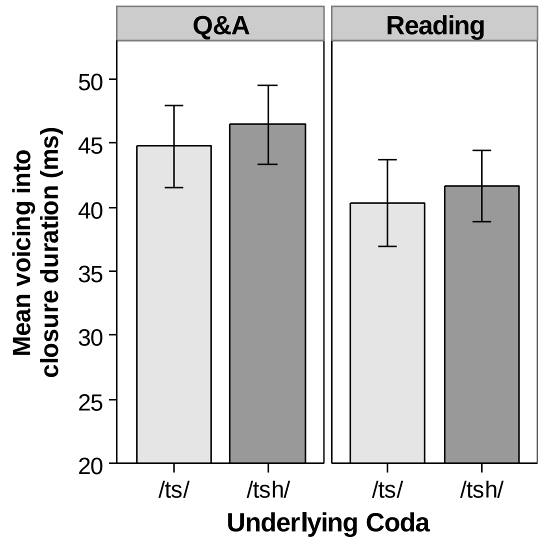

Figure 15. Mean voicing into closure duration before lenis vs. non-lenis affricates subject to palatal neutralization in the Q & A and the reading tasks (neutralizing environment)

3.4.2.4. Discussion

It was shown that voicing into closure duration did not make a clear distinction between underlyingly different codas in a non-neutralizing environment. This was observed across all three types of neutralization regardless of the tasks. Specifically, lenis and non-lenis codas subject to laryngeal neutralization did not show any difference in voicing into closure duration across three places of articulation and across the task. For codas subject to manner or palatal neutralization, no difference in voicing into closure duration was observed in either task.

In a neutralizing environment, most of coda obstruents across three types of neutralization were completely neutralized in terms of voicing into closure duration. However, coda pairs with laryngeal contrasts showed a significant difference in voicing into closure duration. Talkers produced 5.5 ms (12%) longer voicing into closure before non-lenis codas than before lenis counterparts. The magnitude of difference was within the range reported in previous studies (e.g., 16 ms (9%) in Dmitrieva et al., 2010; 0.9 - 2.1 mean cycles of glottal pulsing (31 - 62%) in Kharlamov, 2014; 5 ms in Port & O'dell, 1985).

Contrary to the prediction, however, the direction of contrasts in voicing into closure duration was reversed in a neutralizing environment. That is, voicing into closure duration was significantly longer before non-lenis than before lenis codas. Due to the unexpected durational contrasts in a neutralizing environment, it may appear not as reliable as the other cues in differentiating neutralized codas. However, the same, but non-significant difference was also observed in the reading task. This implies that the unexpected pattern may not be due to errors or problems associated with the measurement itself. The opposite direction of duration contrasts could be due to talker or item effects. As shown in Figure 24 and 25 in Appendix 2, however, this appears not necessarily the case. For alveolar (/t-t^h/) and velar (/k-k'/) pairs in which durational contrasts were evident, most of the talkers produced non-lenis codas with longer voicing into closure duration than lenis counterparts. In terms of the items, all alveolar pairs and 60% of velar pairs showed longer voicing into closure duration in non-lenis than in lenis codas. Therefore, the overall pattern of contrast appears fairly consistent both in talkers and items.

One possible reason for the opposite direction of the contrast is high variability in the duration and/or the position of glottal pulsing. As pointed out by Fourakis and Iverson (1984), glottal pulsing in some tokens would end on the onset of closure, but reappear in the middle of the closure for a short period of time and disappear. We observed the same patterns in measuring voicing into closure. In such a case, we only measured the duration up to the point of the first cessation of periodicity and discarded the discontinuous portion in order to ensure the consistency of the measurements. Some other studies have dealt with this variability by adopting a more conservative measurement approach that they count the number of glottal pulses rather than measuring the absolute duration of voicing in milliseconds (Kharlamov, 2014; Port & Crawford, 1989). However, it is unknown whether both approaches yield similar results or whether one is more accurate than the other. This question may need to be addressed in future research.

A significant durational difference was observed only in the Q & A task, suggesting that incomplete neutralization of laryngeal codas may not be related to orthographic influences but rather more affected by the nature of experimental task. It was also shown that difference in voicing into closure duration varied across bilabial and lingual pairs. Voicing into closure duration before non-lenis lingual codas /t^h, k'/ was longer than those before lenis counterparts /t, k/, whereas it did not vary across bilabial codas /p/ and /p^h/.

Overall, it seems that voicing into closure duration is not a reliable indicator that distinguishes underlyingly different coda obstruents, especially in a non-neutralizing environment. This finding is not surprising, since many of previous studies on final devoicing failed to show that voicing into closure duration serves as a cue to the underlying voicing contrast of final obstruents (Dinnsen & Charles-Luce, 1984; Jassem & Richer, 1989; Kopkalli, 1993; Piroth & Janker, 2004; Port & Crawford, 1989; Slowiaczek & Dinnsen, 1985, Warner et al., 2004). Nevertheless, this study shows that a significant difference in voicing into closure duration exists between codas subject to laryngeal neutralization, especially in the Q & A task. This provides additional supporting evidence that incomplete neutralization occurs in coda obstruents with laryngeal contrasts that are produced in a more communicative-oriented task.

3.4.3. Closure duration

3.4.3.1. Laryngeal neutralization

Table 10 shows the by-participant mean and standard error (in parentheses) of closure duration for lenis vs. non-lenis codas subject to laryngeal neutralization across three places of articulation, two phonological environments, and two tasks. It also presents the difference in mean closure duration between lenis and non-lenis codas.

				Non-neutralizing environment			Neutralizing environment		
Task	**POA**	**Coda**		Lenis	Non-lenis	Difference	Lenis	Non-lenis	Difference
QA	Bilabial	/p/	/p^h/	55.7 (1.8)	140.7 (2.7)	-85.0 (2.3)	140.4 (4.0)	144.2 (4.4)	-3.8 (3.6)
	Alveolar	/t/	/t^h/	57.2 (1.5)	133.0 (3.6)	-75.8 (3.3)	164.8 (3.7)	172.7 (4.0)	-7.9 (2.4)
	Velar	/k/	/k'/	53.5 (1.8)	130.4 (2.6)	-76.9 (2.4)	180.3 (3.3)	180.7 (3.1)	-0.4 (2.7)
	Mean			55.2 (1.0)	134.5 (1.7)	**-79.3** (1.5)***	163.1 (2.5)	166.5 (2.5)	**-3.4** (1.8)
RD	Bilabial	/p/	/p^h/	56.7 (1.7)	136.0 (2.7)	-79.3 (2.7)	136.0 (4.0)	140.3 (4.4)	-4.3 (2.4)
	Alveolar	/t/	/t^h/	57.9 (1.6)	139.6 (3.7)	-81.7 (3.5)	158.2 (4.0)	166.1 (4.2)	-7.9 (2.6)
	Velar	/k/	/k'/	52.2 (1.8)	133.8 (3.0)	-81.6 (3.1)	172.3 (3.1)	175.2 (2.9)	-2.9 (2.6)
	Mean			55.1 (1.0)	136.0 (1.8)	**-80.9** (1.8)***	156.7 (2.4)	161.3 (2.4)	**-4.6** (1.5)*

Table 10. By-participant mean and standard error of closure duration before underlyingly lenis vs. non-lenis codas subject to laryngeal neutralization across three places of articulation (POA), two phonological environments, and two tasks. The difference in mean closure duration highlighted in bold.

<u>Analyses within task in each phonological environment.</u> In a non-neutralizing environment, the main effect of Underlying Coda was significant in the both tasks (QA: β = -79.202, SE = 3.501, $\chi^2(1) = 58.90$, $p < 0.001$; RD: β = -80.826, SE = 5.091, $\chi^2(1) = 48.18$, $p < 0.001$). Closure duration was significantly longer for non-lenis codas /p^h, t^h, k'/ than for lenis counterparts /p, t, k/. When neutralized, lenis and non-lenis codas were still differentiated by closure duration, but only in the reading task (β = -5.198, SE = 2.198, $\chi^2(1) = 4.74$, $p < 0.05$).

The effect did not reach significance in the Q & A task ($\chi^2(1) = 1.36$, $p = 0.24$). The results are summarized in Figure 16.

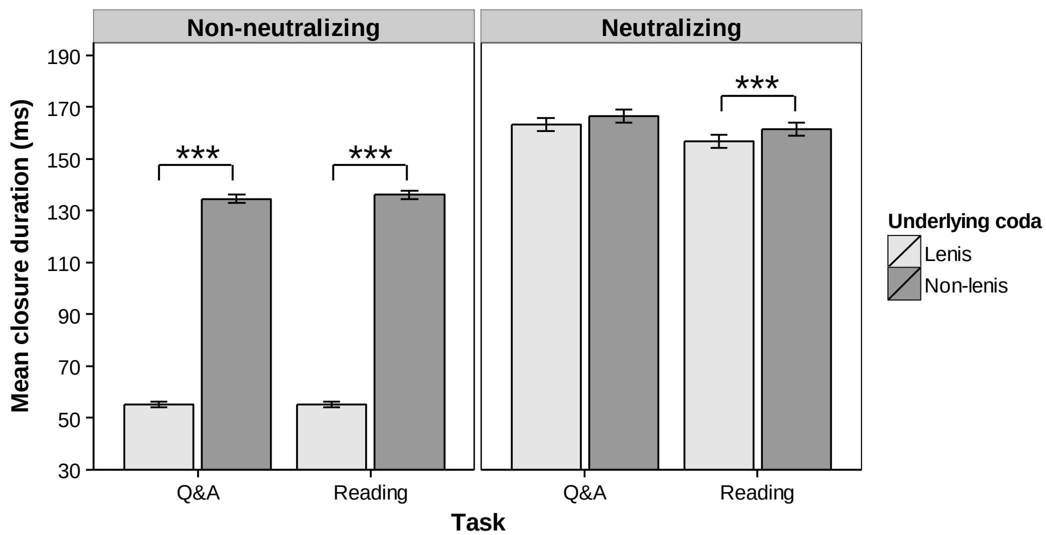

Figure 16. Mean closure duration for lenis vs. non-lenis codas subject to laryngeal neutralization in the Q & A and the reading tasks in a non-neutralizing (left) and neutralizing environment (right)

In a non-neutralizing environment, closure duration did not significantly vary across bilabial and lingual pairs (QA: $\chi^2(1) = 1.39$, $p = 0.23$; RD: $\chi^2(1) = 0.01$, $p = 0.91$) or across alveolar and velar pairs (QA: $\chi^2(1) = 0.45$, $p = 0.49$; RD: $\chi^2(1) = 1.02$, $p = 0.31$), as shown in Figure 17. No interaction was found between Underlying Coda and POA across bilabial and lingual pairs (QA: $\chi^2(1) = 2.21$, $p = 0.13$; RD: $\chi^2(1) = 0.12$, $p = 0.72$) or across alveolar and velar pairs (QA: $\chi^2(1) = 0.02$, $p = 0.88$; RD: $\chi^2(1) < 0.001$, $p = 0.99$).

In a neutralizing environment, the main effect of POA was significant across bilabial and lingual pairs in both tasks (QA: $\beta = 42.784$, SE = 10.406, $\chi^2(1) = 11.94$, $p < 0.001$; RD: $\beta = 39.786$, SE = 12.324, $\chi^2(1) = 8.61$, $p < 0.01$). Specifically, closure durations for lingual pairs were significantly longer than those for bilabial pairs, as shown in Figure 18. However, the effect was non-significant across alveolar and velar pairs (QA: $\chi^2(1) = 2.05$, $p = 0.151$; RD: $\chi^2(1) = 1.87$, $p = 0.17$), indicating that closure duration for alveolar pairs was not significantly different from that for velar pairs.

The interaction between Underlying Coda and POA was not significant in either task. Closure duration difference between lenis and non-lenis codas did not vary across bilabial and lingual pairs ($\chi^2(1) < 0.01$, $p = 0.95$; RD: $\chi^2(1) = 0.02$, $p = 0.87$) or across alveolar and velar pairs (QA: $\chi^2(1) = 0.81$, $p = 0.36$; RD: $\chi^2(1) = 1.00$, $p = 0.31$).

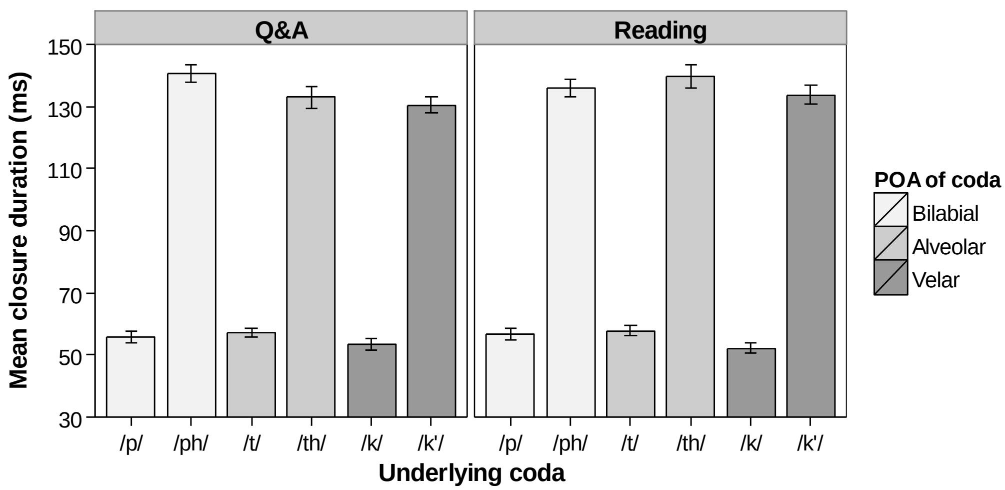

Figure 17. Mean closure duration for lenis vs. non-lenis codas subject to laryngeal neutralization across three places of articulation in the Q & A and the reading tasks (non-neutralizing environment)

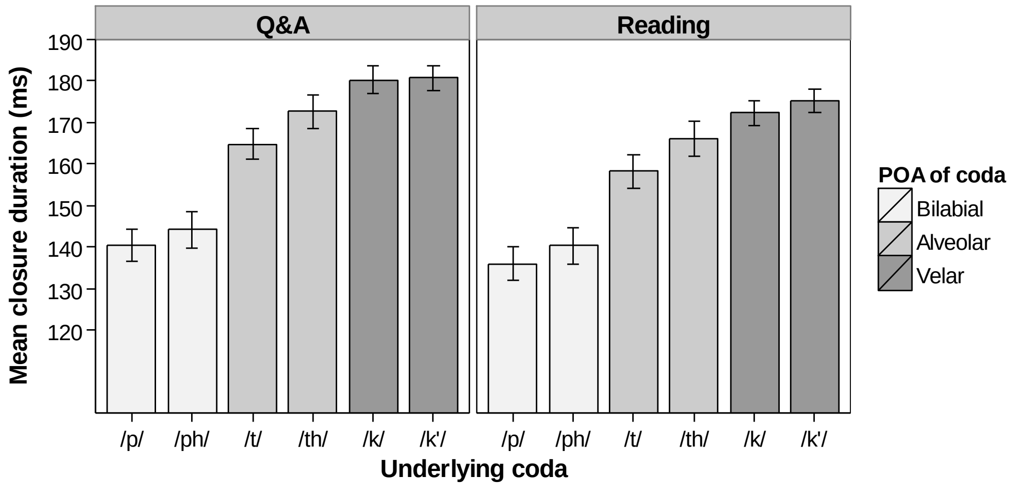

Figure 18. Mean closure duration for lenis vs. non-lenis codas subject to laryngeal neutralization across three places of articulation in the Q & A and the reading tasks (neutralizing environment)

3.4.3.2. Manner neutralization

Table 11 shows the by-participant mean and standard error (in parentheses) of closure duration for each member of three types of coda pairs subject to manner neutralization across two phonological environments and two tasks. It also presents the difference in mean closure duration between the members of each type of coda pairs. As stated in section 3.3, closure duration in a non-neutralizing environment was not compared in the pairs with fricative /s/, for the same reason stated in section 3.4.2.2.

	Underlying Coda (AB)	Non-neutralizing environment			Neutralizing environment		
		A	B	Difference	A	B	Difference
QA	Fricative - Stop (FS)	N/A	N/A	N/A	182.4 (3.1)	185.1 (3)	**-2.7** (2.3)
	Affricate - Stop (AS)	92.7 (5)	90.9 (5.3)	**1.8** (8.8)	185.2 (3.5)	185.3 (3.2)	**0.0** (3.1)
	Fricative - Affricate (FA)	N/A	N/A	N/A	189.6 (2.6)	184.4 (2.5)	**5.2** (1.8)*
RD	Fricative - Stop (FS)	N/A	N/A	N/A	177.9 (3.1)	179.3 (2.7)	**-1.4** (2.9)
	Affricate - Stop (AS)	92.9 (5)	89.6 (5.5)	**3.3** (8.9)	182.4 (3.6)	172.7 (3.0)	**9.7** (2.4)***
	Fricative - Affricate (FA)	N/A	N/A	N/A	182.1 (3.0)	177.8 (2.7)	**4.3** (2.6)

Table 11. By-participant mean and standard error of closure duration for each member of three types of coda pairs subject to manner neutralization across two phonological environments and two tasks. The difference in mean closure duration highlighted in bold.

<u>Analyses within task in each phonological environment.</u> In a non-neutralizing environment, the main effect of MOA in AS pairs was not significant in either task (QA: $\chi^2(1) = 12.18$, $p < 0.001$; RD: $\chi^2(1) = 11.84$, $p < 0.001$), suggesting that there was no significant difference in closure duration between /tʃʰ/ and /tʰ/.

In a neutralizing environment, the main effect of MOA was significant in AS and FA pairs, as shown in Figure 19. Specifically, closure duration was longer for affricate codas /tʃ, tʃʰ/ than for stop codas /t, tʰ/ ($\beta = 9.714$, SE = 2.313, $\chi^2(1) = 14.43$, $p < 0.001$), but only in the Q &A task (RD: $\chi^2(1) = 0.01$, $p = 0.90$). For FA pairs, fricative codas /s/ showed longer closure duration than affricate codas /tʃʰ/ in the reading task ($\beta = 5.180$, SE = 2.110, $\chi^2(1) = 5.06$, $p < 0.05$), but not in the Q & A task ($\chi^2(1) = 2.47$, $p = 0.11$). No significant difference in closure duration was observed in FA pairs ($\chi^2(1) = 0.35$, $p = 0.55$).

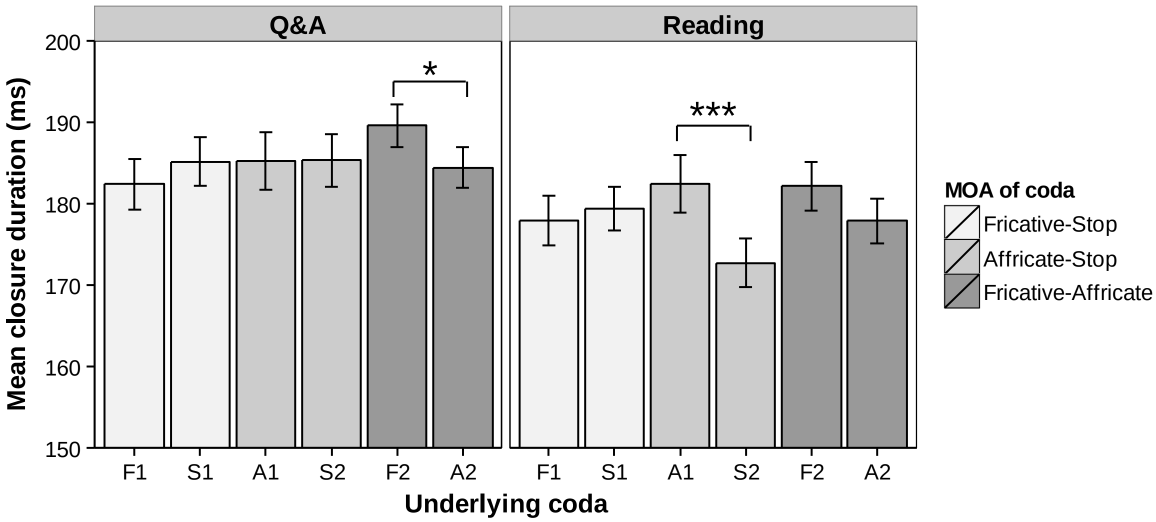

Figure 19. Mean closure duration for each member of three types of coda pairs subject to manner neutralization in the Q & A and the reading tasks (neutralizing environment)

Analyses across task, repetition in each environment. In a non-neutralizing environment, AS pairs showed no interaction between MOA and Task ($\chi^2(1) < 0.001$, $p = 0.98$) as well as the interaction between MOA and Repetition ($\chi^2(1) = 0.08$, $p = 0.77$). In a neutralizing environment, the interaction between MOA and Task was not significant in FS pairs ($\chi^2(1) = 0.13$, $p = 0.71$) and FA pairs ($\chi^2(1) = 0.04$, $p = 0.76$), but reached significance in AS pairs ($\beta = 9.373$, SE = 3.496, $\chi^2(1) = 7.04$, $p < 0.01$).

Follow-up regressions performed for data subset separated by the task yielded a significant main effect of Underlying Coda in the reading task ($\beta = 9.714$, SE = 2.313, $\chi^2(1) = 14.43$, $p < 0.001$), but not in the Q & A task ($\chi^2(1) = 0.01$, $p = 0.90$). This suggests that the interaction effect stemmed from a significant difference in closure duration between affricate and stop codas in the Q & A task. Finally, the interaction between MOA and Repetition was not significant in any of the three types of coda pairs (FS: $\chi^2(1) = 0.01$, $p = 0.90$; AS: $\chi^2(1) < 0.01$, $p = 0.92$; FA: $\chi^2(1) = 1.83$, $p = 0.17$).

3.4.3.3. Palatal neutralization

Table 12 shows the by-participant mean and standard error (in parentheses) of closure duration for lenis vs. non-lenis affricate codas subject to palatal neutralization across two phonological conditions and two tasks. It also presents the difference in closure duration between lenis and non-lenis codas.

	Underlying Coda		**Non-neutralizing environment**			**Neutralizing environment**		
	Lenis	Non-lenis	Lenis	Non-lenis	Difference	Lenis	Non-lenis	Difference
QA	/tʃ/	/tʃʰ/	46.2 (2.8)	108.1 (4.1)	**-61.9** (2.9)***	188.2 (4.8)	184.7 (4.4)	**3.5** (3.3)
RD	/tʃ/	/tʃʰ/	44.8 (2.8)	109.2 (4.1)	**-64.4** (3.2)***	177.4 (4.8)	180.5 (5.1)	**-3.1** (2.8)

Table 12. By-participant mean and standard error of closure duration before lenis vs. non-lenis codas subject to palatal neutralization across two phonological environments and two tasks. The difference in mean closure duration highlighted in bold.

Analyses within task in each phonological environment. In a non-neutralizing environment, the main effect of Underlying Coda was significant in the both tasks (QA: $\beta = -61.897$, $SE = 3.590$, $\chi^2(1) = 45.53$, $p < 0.001$; RD: $\beta = -64.355$, $SE = 4.201$, $\chi^2(1) = 42.18$, $p < 0.001$). Specifically, closure duration was significantly longer for non-lenis affricate /tʃʰ/ than for lenis affricate /tʃ/. When neutralized, lenis and non-lenis affricate codas did not show any difference in closure duration in either task (QA: $\chi^2(1) = 0.89$, $p = 0.34$; RD: $\chi^2(1) = 0.76$, $p = 0.38$). The results are summarized in Figure 20 and Figure 21.

Analyses across task, repetition in each environment. In each environment, no significant interaction was found between Underlying Coda and Task (non-neutralizing: $\chi^2(1) = 0.48$, $p = 0.48$; neutralizing: $\chi^2(1) = 1.54$, $p = 0.21$) or between Underlying Coda and Repetition (non-neutralizing: $\chi^2(1) = 0.04$, $p = 0.83$; neutralizing: $\chi^2(1) = 0.17$, $p = 0.67$).

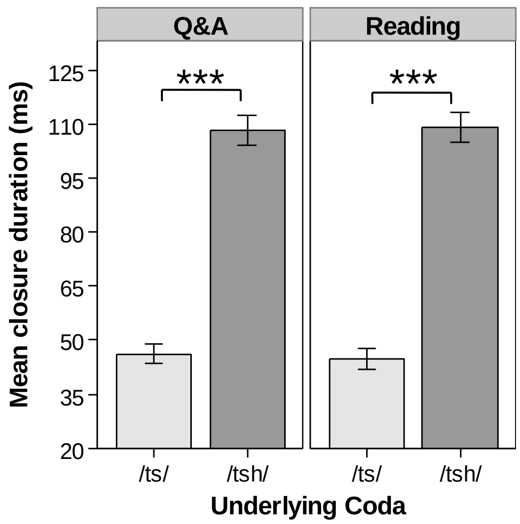

Figure 20. Mean closure duration for lenis vs. non-lenis affricates subject to palatal neutralization in the Q & A and the reading tasks (non-neutralizing environment)

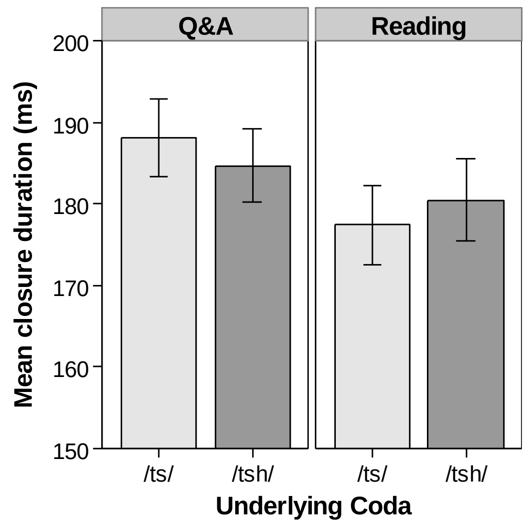

Figure 21. Mean closure duration for lenis vs. non-lenis affricates subject to palatal neutralization in the Q & A and the reading tasks (neutralizing environment)

3.4.3.4. Discussion

In a non-neutralizing environment, closure duration served as a robust cue to distinguish between underlyingly distinct codas subject to laryngeal or palatal neutralization. For codas with laryngeal contrasts, talkers produced 79.3 ms (144% in QA) ~ 80.9 ms (147% in RD) longer closure duration for non-lenis codas than for lenis counterparts. This result is consistent with previous findings that closure duration in Korean is shortest for lenis and longest for fortis stops

(Han, 1994; M.-R.C. Kim, 1994; Kim & Lotto, 2002; Silva, 1992). Across the three places of articulation of coda, differences in closure duration did not vary between lenis and non-lenis codas. For codas subject to palatal neutralization, it was established that closure duration was longer for non-lenis affricate /tʃʰ/ than for lenis /tʃ/ by 61.9 ms (134% in QA) ~ 64.4 ms (144% in RD).

In a neutralizing environment, differences in closure duration were significant in coda pairs with laryngeal contrasts and FA and AS pairs with manner contrasts. For lenis vs. non-lenis pairs, talkers produced 4.6 ms (3%) longer closure duration for non-lenis codas than for their lenis counterparts, but only in the reading task. Across the three places of articulation, closure duration was longest for velar and shortest for bilabial codas. Consequently, there was a significant interaction between Underlying Coda and POA across bilabial vs. lingual pairs due to a larger difference in closure duration in lingual pairs. Interestingly, a significant difference in closure duration was observed only in the reading task. This result is inconsistent with the finding that significant differences in vowel duration and voicing into closure duration were observed only in the Q & A task.

In the pairs with manner contrasts, closure duration was 5.2 ms (3%) longer for fricative /s/ than for affricates /tʃ, tʃʰ/ only in the Q & A task, and 9.7 ms (6%) longer for affricates than for stops /t, tʰ/ only in the reading task. Although this finding may serve as evidence for incomplete neutralization of manner, it should be interpreted with caution. Besides the fact that the findings were based on a small number of items, the magnitude of the contrasts in AS pairs was larger than in a non-neutralizing environment. For example, the difference in closure duration between affricate and stop codas (9.7 ms) was larger than those observed in non-neutralizing environment (3.3 ms). Considering that magnitude of contrasts in a neutralizing environment is expected to be smaller than the full contrast observed in a non-neutralizing condition, this unexpected result seems not reliable.

Overall, closure duration served as a robust cue to distinguish underlyingly distinct codas subject to laryngeal or palatal neutralization in a non-neutralizing environment. When neutralized, codas with laryngeal contrasts were still differentiated by closure duration, but only in the reading task. Although the results are mixed and inconclusive, some of codas with manner distinction also showed a significant trend toward incomplete neutralization in terms of closure duration.

CHAPTER 4. PERCEPTION STUDY

4.1. Introduction

The production experiment revealed that coda neutralization in Korean is not always complete in production. The pairs with laryngeal contrasts showed significant differences in vowel duration, voicing into closure duration, or closure duration. The effects of incomplete neutralization were stronger in the spontaneous speech than in the read speech. For coda pairs subject to manner or palatal neutralization, the effects of incomplete neutralization were inconsistent or absent. Coda pairs with manner contrasts were incompletely neutralized in terms of vowel duration or closure duration, but only in one of the two speech types. For the pairs subject to palatal neutralization, they were completely neutralized in terms of all measured temporal acoustic parameters in the both types of speech. However, the findings regarding manner and palatal neutralization should be interpreted with caution, since the sample size was too small to be a representative of each type of neutralization.

Now a question that arises is whether there are perceptual consequences of complete or incomplete neutralization of coda obstruents in Korean. Are subtle but significant acoustic differences between neutralized codas perceptible to listeners to such an extent that listeners can identify intended coda by making use of those cues? Or are listeners able to pick up on other unmeasured acoustic differences even when there is no significant difference between neutralized codas along the measured acoustic dimensions? In order to address these questions, we conducted a 2AFC identification experiment in which native Korean listeners were asked to identify the correct underlying form of neutralized words in the presence or absence of the subtle acoustic differences reported in Chapter 3.

With regard to the impact of complete or incomplete neutralization in production on speech perception, two predictions can be made. First, if the measured acoustic parameters serve as salient perceptual cues for identifying the intended coda, it is expected that listeners may be able to successfully distinguish underlyingly different codas for which reliable acoustic differences exist. However, such a pattern of results is not expected in coda pairs with no measured acoustic differences.

Second, it is possible that acoustic differences are physically present but not large enough to be perceived by listeners. Consequently, listeners may not be able to distinguish the underlying differences between the neutralized codas, regardless of whether coda neutralization is complete or incomplete.

Thirdly, even with the possibility that acoustical differences are too small to be perceived, listeners may be still able to identify the intended codas by attending to unmeasured cues for discrimination. For example, it is possible that listeners make use of unmeasured spectral cues or a combination of temporal and spectral cues for identification. In this case, it is expected that listeners may be still able to differentiate neutralized codas that exhibit no difference in the measured acoustic parameters.

4.2. Methods

In order to limit the size of the experiment, a subset of stimuli was selected from a list of target words recorded in the production study. The stimuli were selected based on several criteria. First of all, they were selected by focusing on one contrast (lenis vs. non-lenis) and one task (Q & A). As introduced in section 3.2.1, coda pairs subject to laryngeal or palatal neutralization consist of lenis and non-lenis obstruents. As for codas subject to manner neutralization, all pairs contrast in the manner of articulation of the coda, some of which (/s-tʰ/, /tʃ-tʰ/, /tʃʰ-t/, /s-tʃʰ/) also exhibit a lenis/non-lenis contrast. By making use of the contrast common across all three types of neutralization, we were able to compare laryngeal neutralization for which acoustic measurements indicated incomplete neutralization with manner/palatal neutralization of which the results were unclear. With regard to the task, it was the Q & A task that showed a clearer difference between laryngeal and manner/palatal neutralization.

Secondly, the target words were selected from the first out of two production trials. In case that the item from the first productions was not usable for any reason (e.g., noise, imprecise articulation of coda obstruent, etc.), it was replaced with the same item from the second productions. A total of 178 items (92.7%) were selected from the first trials, and 14 items (7.3%) from the second.

Third, all of the three measured acoustic parameters were considered for the perception experiment. The target stimuli selected based on the above criteria were compared with the original production stimuli in terms of regression results for each acoustic parameter (See Appendix 2). For all parameters, the subset stimuli did not show the same level of significance as in the production stimuli, since they consisted of only the first production out of two repetitions. Still, they reflected the overall pattern of results observed in the original data set.

4.2.1. Stimuli

The target word pairs selected from the production stimuli are presented in Appendix 3. They consisted of 22 minimal pairs with lenis vs. non-lenis contrast produced by 15 talkers. Twelve pairs of coda obstruents were subject to laryngeal neutralization, 8 pairs manner neutralization, and 2 pairs palatal neutralization. All words had a CVC syllable structure except two pairs with a VC structure (/ip-ip^h/, /up-up^h/). The total number of stimuli presented in the identification task was 660 (44 words x 15 subjects).

The stimuli were cut from the sentences recorded in the production study and saved as individual files. All of the files were then leveled to equate root mean-square (RMS) amplitude. The comparison between the original full set of production stimuli with the subset for the perception experiment is presented in Table 13.

	# of subjects	Task	Trial	Contrast	# of stimuli	Total # of trials
Original	15	Q & A, reading	1, 2	L*: LN, POA M: FS, AS, FA P: LN	L: 24 (12mp) M: 28 (14mp) P: 4 (2mp)	56 words x 15 subjects x 2 repetitions = 1680 (QA only)
Subset	15	Q & A	1	LN	L: 24 (12mp) M: 16 (8mp) P: 4 (2mp)	44 words x 15 subjects = 660 trials

Table 13. Comparison of subset stimuli to be used in the perception experiment with the original set of stimuli used in the production experiment

Note: L = Laryngeal, M = Manner, P = Palatal
LN = Lenis vs. Non-lenis
FS = Fricative vs. Stop
AS = Affricate vs. Stop
FA = Fricative vs. Affricate

4.2.2. Participants

Participants in the perception experiment consisted of twenty six native Korean listeners who did not participate in the production study (14 males, 12 females). In order to avoid influence of prior L2 experience on identification performance, participants with more than a year of stay in English-speaking environment were excluded.

All participants were undergraduate students at Hanyang University in Korea, ranging in age from 20 to 28 (average = 23.3). They were born and had lived in Seoul or in Gyeonggi province in Korea and had used standard Seoul Korean as their first language. They began English language study between 3 and 13 years old and had learned English as a foreign language for an average of 10.8 years in a school environment. The length of stay in English-speaking countries ranged from 0 to 12 months with a mean of 1.7 months. None of participants had any known history of speech and hearing impairments at the time of testing.

4.2.3. Procedure

The experiment was conducted in a sound-attenuated booth in Hanyang Phonetics and Psycholinguistics Laboratory in Korea. Participants were seated in front of a computer screen and wore headphones. Prior to the experiment, participants were asked to fill out a language-background questionnaire. They were then asked to read step-by-step written instructions carefully and then complete 8 practice trials in order to get familiarized with a 2AFC identification task. The practice stimuli consisted of the same-same or the same-different pairs from each of 3 minimal pairs (/kus-kut/, /məs-mətʃ/, /sut^{h}-sutʃ^{h}/), none of them involves lenis vs. non-lenis contrast.

After completing practice trials, they proceeded to the actual experiment by pressing ‘Start’ button. In each trial, participants were presented with a single auditory stimulus at a comfortable listening level via the headphones and asked to identify it by selecting one of the two response alternatives. The response alternatives as a minimal pair were presented on the computer screen with Korean orthography for 500 ms; one member of a minimal pair (e.g., 집

meaning ‘house’) appeared on the lower left side of the screen and the other (e.g., 짚 meaning ‘straw’) appeared on the lower right side of the screen. Participants identified the token by clicking the button labeled “1” for the word on the left side and “2” for the word on the right side. For half the trials, the target words with lenis coda were placed on the left and words with non-lenis coda on the right. The order of position was reversed for the other half trials. Once the participants provided their response, the next trial was played after 500 ms.

Stimuli were presented using Superlab presentation software. The order of stimuli was pseudo-randomized in such a way that the members of a pair never appeared next to each other and that target items in the same type of neutralization did not appear more than 2 times sequentially. Stimuli were divided into 5 blocks, each of which consisted of 132 trials. Participants took a 30-second mandatory break after completing half the trials within a block and 2-minute break between the blocks. Test data for each participant was saved automatically as a plain text file. The experiment lasted for 60-70 minutes. Participants were compensated for their participation at the rate of 10,000 won (equivalent to $10) per hour.

4.2.4. Analyses

Data obtained from the identification task consisted of 17160 binary responses (26 listeners x 660 trials) that were either lenis or non-lenis. According to signal detection theory (Green & Swets, 1966; Macmillan & Creelman, 1991; Macmillan, 2002), a listener’s binary response is influenced by discriminability and response bias. Discriminability, also called as sensitivity, is the ability to reflect a stimulus-response correspondence manipulated by the experimenter, whereas response bias is the systematic tendency to favor a particular response over others.

In order to disentangle listener's discrimination (i.e., the ability to select non-lenis response when it is the right or wrong answer) from response bias (i.e., the tendency to identify the coda as non-lenis regardless of what the intended coda is), response variables were coded in a following way. The responses were first coded with 1 for non-lenis responses and with 0 for lenis ones in one column named Non-lenis Response and then coded with 0.5 when non-lenis is the correct response and with -0.5 when non-lenis is the wrong response in a separate column named Intended Coda. The variable Type of Neutralization, which has three levels (laryngeal, manner, palatal), was divided into two contrasts: the first contrast compared laryngeal with manner and palatal neutralization as a group and the second contrast compared manner with palatal neutralization.

Following Macmillan (2002), these variables were then entered into a logistic regression predicting the odds of non-lenis response. The outcome of regressions was interpreted in the following way: the intercept represents response bias, that is, the odds a listener identifies the target as non-lenis regardless of what the intended coda is. Therefore, it informs us of the degree to which listeners' responses are biased toward either lenis or non-lenis. The main effect of Intended Coda represents the odds a listener identifies the target as non-lenis when non-lenis is the right or wrong answer. The main effects of Type of Neutralization indicates how listeners' response bias varies across three types of neutralization. Finally, interaction effect between Intended Coda and Type of Neutralization shows how listeners' discrimination varies across three types of neutralization.

A logistic mixed effects regression analysis (Jaeger, 2008) following this structure was performed. The models had the maximal random effects structure for this design (Barr et al., 2013). They included random intercept for talkers, listeners, and items as pairs. Random slopes by talkers and listeners were also included for all fixed effects and their interactions. As in the analysis of production data, the best regression model structure was selected by following the backward model selection procedure. I first started with the full model and then reduced the model through a step-wise elimination of the correlations between random effects, random slopes for 3-way interactions, random slopes for 2-way interactions, all random slopes, and random intercept for items. Once the final model converged, significance for fixed effects was assessed via model comparison.

4.3. Results and Discussion

The results of identification task are summarized in Table 14 and Figure 22. They show how listeners' non-lenis response rate for underlyingly lenis vs. non-lenis coda varies as a function of the type of neutralization. The difference in the non-lenis response rate between lenis and non-lenis codas was 2.1% in laryngeal, 1% in manner, and 2.3% in palatal neutralization. Overall, underlyingly non-lenis coda was identified as non-lenis 1.7% more than underlyingly lenis coda was identified as non-lenis across three types of coda neutralization.

Underlying coda / Type of Neutralization	Lenis	Non-lenis	Difference
Laryngeal	53.4 (1.6)	55.5 (1.4)	-2.1 (0.9)
Manner	52.6 (1.7)	53.6 (1.7)	-1.0 (1.0)
Palatal	50.7 (2.5)	53.0 (3.3)	-2.3 (2.2)
Total	52.9 (1.4)	54.6 (1.3)	-1.7 (0.8)

Table 14. Percentage non-lenis response rate with standard error (in parentheses) for underlyingly lenis vs. non-lenis coda by three types of coda neutralization

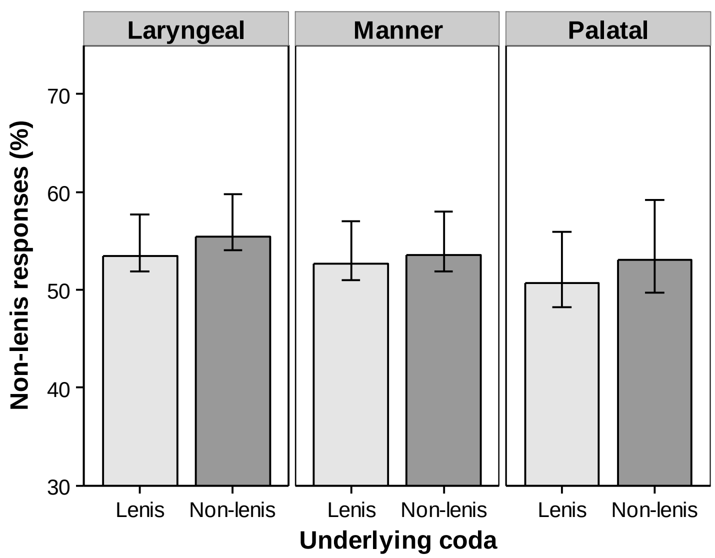

Figure 22. Percentage non-lenis responses to underlyingly lenis vs. non-lenis coda obstruents subject to laryngeal, manner, or palatal neutralization

The results of regressions revealed that response bias was not significant ($\beta = 0.13$, SE = 0.15, $\chi^2(1) = 0.83$, $p = 0.36$), suggesting that listeners' responses were not biased toward either lenis or non-lenis. Listener's non-lenis response did not vary either across laryngeal and manner/palatal neutralization ($\chi^2(1) < 0.001$, $p = 0.98$) or across manner and palatal neutralization ($\chi^2(1) = 0.08$, $p = 0.77$).

The main effect of Intended Coda was marginally significant ($\beta = 0.078$, SE = 0.045, $\chi^2(1) = 2.87$, $p = 0.08$). This indicates that there was an overall non-significant trend that listeners identify the underlying form as non-lenis when non-lenis is the right or wrong answer. There was no interaction between Intended Coda and the type of neutralization (laryngeal vs. manner/palatal: $\chi^2(1) = 0.44$, $p = 0.50$; manner vs. palatal: $\chi^2(1) = 0.01$, $p = 0.91$), suggesting that listeners' discrimination performance did not vary across three types of neutralization.

Overall, the results of perception experiment suggest that native Korean listeners' identification of neutralized coda obstruents is at chance level. Even though the listeners showed a slightly higher non-lenis response rate for underlyingly non-lenis coda across all three types of coda neutralization, they were not able to successfully label most of neutralized codas with the correct underlying forms regardless of the presence or absence of acoustic differences.

In the next section, I will further explore whether there is any relationship between production and perception. Specifically, it will be examined whether listeners' discrimination performance can be predicted by any of acoustic parameters measured for neutralized coda obstruents.

CHAPTER 5. RELATIONSHIP BETWEEN PRODUCTION AND PERCEPTION

5.1. Introduction

The results of the perception study showed that native Korean listeners were not able to distinguish underlyingly lenis coda from non-lenis counterparts in a consistent manner, even when there was a significant durational difference between them. Even though listeners' discrimination was around chance level, it appears still important to know which acoustic cue made a greater contribution to listeners' identification of the intended coda. By revealing it, we may be able to show that listeners did not respond to the stimuli in a random manner, but in fact attended to the stimuli, which implies that their judgments systematically reflect properties of the stimuli.

In this chapter, it will be further explored whether listeners' discrimination of non-lenis coda can be predicted by single or multiple acoustic parameters measured in the production study. Based on prior findings on durational patterns of acoustic cues addressed in section 3.2.2, we predict that there will be more lenis responses for codas with shorter vowel duration, longer voicing into closure duration, and shorter closure duration.

5.2. Methods

The relationship between listeners' discrimination performance and acoustic cues was examined by conducting a series of regressions. Data used for the analysis was both listeners' identification data obtained for 22 minimal pairs produced by 15 talkers and acoustic measurement data for the same stimuli (i.e., vowel duration, voicing into closure duration, and closure duration).

In the analysis using mixed effect regressions, the models with three acoustic parameters and their interactions as fixed effects did not converge. When each acoustic parameter was assessed in a separate model, the models assessing closure duration did not still reach convergence. For this reason, I fitted logistic regression models to the data using the glm (generalized linear model) function.

The models included Non-lenis Response as a dependent variable. The fixed effects were Intended Coda (lenis, non-lenis), Type of Neutralization (laryngeal, manner, palatal), three acoustic parameters, and the interactions between all fixed effects. The binary fixed effects (Intended Coda, Type of Neutralization) were contrast-coded. The variable Type of Neutralization, which has three levels (laryngeal, manner, palatal), was divided into two contrasts: the first contrast compares laryngeal with manner and palatal neutralization as a group and the second contrast compares manner with palatal neutralization. Since this full model converged, likelihood ratio tests were conducted to assess the significance of the fixed effects. In each test, the full model containing a fixed effect was compared to a model without the fixed effect.

Secondly, logistic mixed effects regressions (Jaeger, 2008) were conducted to further examine whether the results from logistic regressions changes when any possible influences of random effects on fixed effects are taken into account. The models contained the same dependent variable and the fixed effects as used in the logistic regression models. They also had maximal random effects structure (Barr et al., 2013). Random intercept for talkers, listeners, and items as pairs were included as well as random slopes by talkers and listeners for all fixed effects and their interactions.

The best regression model structure was selected by following backward model selection procedure, as explained in the previous section 4.2.4. Since the final model failed to converge, the fixed effects structure was reduced in subsequent regression models. Specifically, the variable Type of Neutralization was removed from the structure, since it did not improve the model fit. Besides, each of the three acoustic parameters was assessed in separate models.

The next section summarizes the results of logistic regressions and logistic mixed effect regressions.

5.3. Results

5.3.1. Logistic regressions

Figure 23 shows how non-lenis responses to underlyingly lenis or non-lenis coda vary depending on preceding vowel duration, voicing into closure duration, and closure duration. The bins in each figure are evenly distributed between minimum and maximum values of the acoustic parameter. For example, all vowel durations are ranked, and then divided into 5 intervals, ranging from the minimum (14.8 ms) to the maximum value (117.3 ms), as presented in Table 15.

Bin / **Parameter**	Shortest duration →				Longest duration
	1	2	3	4	5
Vowel duration	14.8 ~ 35.3	35.4 ~ 55.8	55.9 ~ 76.3	76.4 ~ 96.8	96.9 ~ 117.3
Voicing into closure duration	3.3 ~ 19.6	19.7 ~ 36.0	36.1 ~ 52.4	52.5 ~ 68.8	68.9 ~ 85.2
Closure duration	37 ~ 87	87.1 ~ 137.1	137.2 ~ 187.1	187.2 ~ 237.2	237.3 ~ 287.2

Table 15. The five intervals of each acoustic parameter (in ms)

Figure 23(a) shows that going from left to right, the percentage of non-lenis decreases as vowel duration increases. That is, listeners provided less non-lenis responses as the preceding vowel duration gets longer regardless of the underlying status of coda as lenis or non-lenis. One noticeable response pattern observed in Figure 23(b) is that listeners provided more non-lenis response to neutralized codas with palatal distinction as voicing into closure duration becomes longer. For other coda pairs, the impact of voicing into closure duration appears less consistent. As for closure duration, it seems to have least influence on listener's identification across the three types of coda neutralization.

These observations were confirmed by a series of logistic regressions. The analyses revealed that the intercept was significant ($\beta = 1.412$, SE = 0.183, $z = 7.69$, $p < 0.001$), indicating that listeners' response biased toward non-lenis coda regardless of the underlying status of coda. There was a significant main effect of vowel duration ($\beta = -0.022$, SE = 0.001, $\chi^2(1) = 376.81$, $p < 0.001$) and voicing into closure duration ($\beta = -0.005$, SE = 0.001, $\chi^2(1) = 12.63$, $p < 0.001$), but marginally non-significant effect of closure duration ($\chi^2(1) = 2.84$, $p = 0.09$). This result suggests that listeners' non-lenis response was influenced by vowel duration and voicing into closure duration, but not by closure duration.

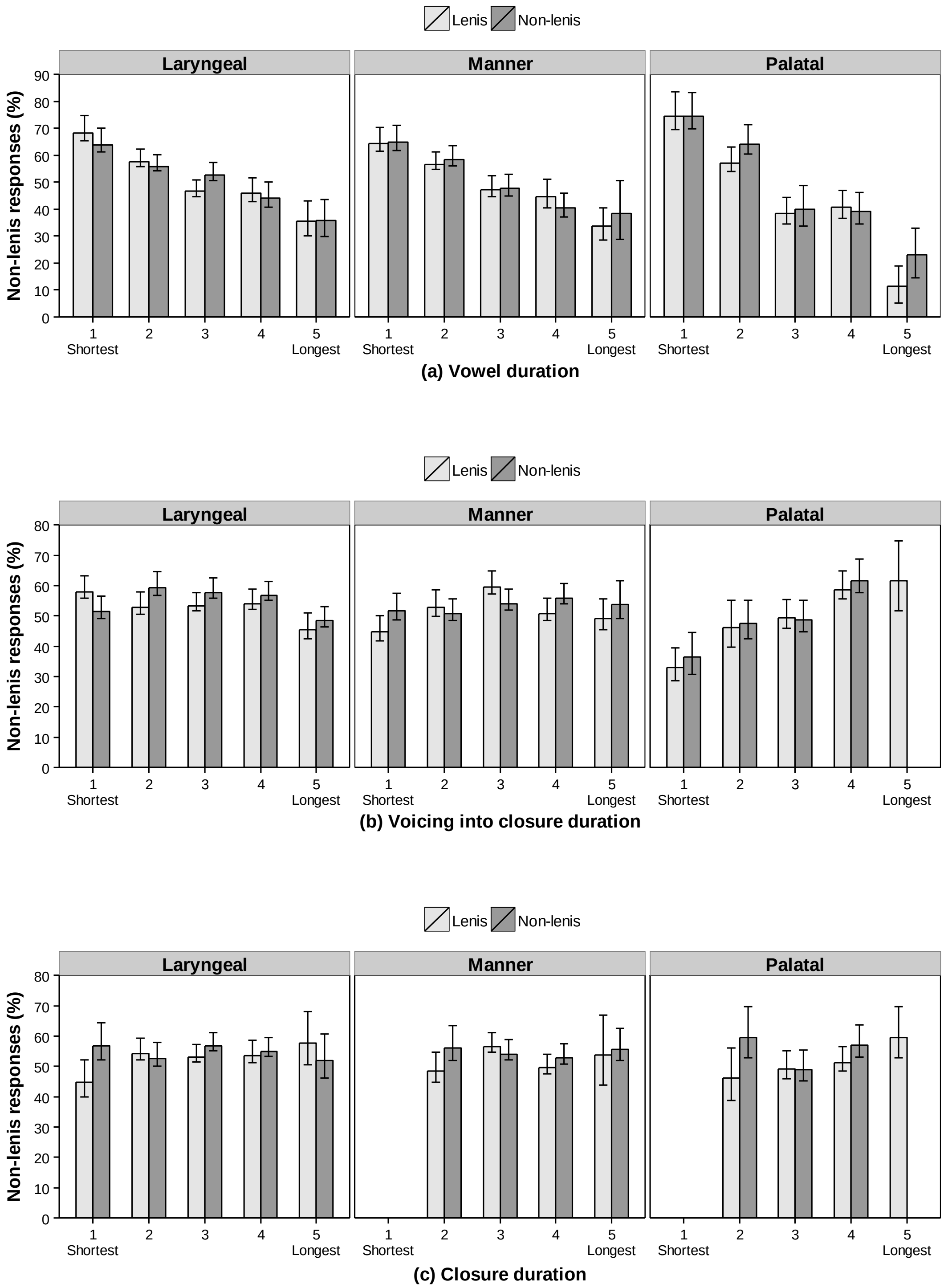

Figure 23. Percentage non-lenis responses to underlyingly lenis vs. non-lenis coda varying by the range of preceding vowel duration (top), voicing into closure duration (middle), and closure duration (bottom)

A series of follow-up regressions were conducted for data subset separated by the type of neutralization. The results showed that when vowel duration became longer, listener provided less non-lenis responses for codas subject to manner or palatal neutralization ($\beta = -0.019$, $SE = 0.001$, $\chi^2(1) = 301.48$, $p < 0.001$) than for codas with laryngeal contrasts ($\beta = -0.016$, $SE = 0.001$, $\chi^2(1) = 193.40$, $p < 0.001$). Across manner and palatal neutralization, listeners provided less non-lenis responses for codas subject to palatal neutralization ($\beta = -0.027$, $SE = 0.002$, $\chi^2(1) = 142.69$, $p < 0.001$) than for codas subject to manner neutralization ($\beta = -0.017$, $SE = 0.001$, $\chi^2(1) = 170.93$, $p < 0.001$). Similarly, with longer closure duration, non-lenis response rate became higher for codas subject to manner or palatal neutralization ($\beta = 0.004$, $SE = 0.002$, $\chi^2(1) = 6.01$, $p < 0.05$) than for codas with laryngeal contrasts ($\beta = 0.001$, $SE = 0.000$, $\chi^2(1) = 3.30$, $p = 0.06$).

The main effect of Intended Coda was marginally significant ($\beta = 0.694$, $SE = 0.367$, $\chi^2(1) = 3.598$, $p = 0.06$), confirming that there was a non-significant trend towards discriminating non-lenis from lenis coda correctly. Across three types of neutralization, listeners' discrimination did not vary across laryngeal and manner/palatal neutralization ($\chi^2(1) = 0.37$, $p = 0.71$), but varied across manner and palatal neutralization ($\beta = -2.517$, $SE = 1.026$, $\chi^2(1) = 6.06$, $p < 0.05$). Follow-up regressions revealed that the main effect of Intended Coda was non-significant in either type of neutralization (manner: $\chi^2(1) = 0.54$, $p = 0.46$; palatal: $\chi^2(1) = 0.83$, $p = 0.36$), although the effect was stronger in palatal neutralization.

The interaction between Intended Coda and acoustic cues were non-significant (vowel duration: $\chi^2(1) = 0.11$, $p = 0.73$; voicing into closure duration: $\chi^2(1) = 1.16$, $p = 0.27$; closure duration: $\chi^2(1) = 2.02$, $p = 0.15$). This indicates that any of the acoustic cues did not make a significant contribution to listeners' discrimination. The interaction did not vary across laryngeal and manner/palatal neutralization (vowel duration: $\chi^2(1) = 0.34$, $p = 0.55$; voicing into closure duration: $\chi^2(1) = 0.47$, $p = 0.48$; closure duration: $\chi^2(1) = 0.07$, $p = 0.77$). Across manner and palatal neutralization discrimination was affected by both vowel duration ($\beta = 0.016$, SE = 0.006, $\chi^2(1) = 5.79$, $p < 0.05$) and voicing into closure duration ($\beta = 0.019$, SE = 0.008, $\chi^2(1) = 5.73$, $p < 0.05$). Follow-up regressions showed that the effect of vowel duration on discrimination was not significant either manner ($\chi^2(1) = 0.68$, $p = 0.40$) or palatal neutralization ($\chi^2(1) = 0.04$, $p = 0.82$), although the effects was stronger in the latter. The same pattern of results was observed for voicing into closure duration (manner: $\chi^2(1) = 0.03$, $p = 0.85$; palatal: $\chi^2(1) < 0.001$, $p = 0.99$).

5.3.2. Logistic mixed effects regressions

5.3.2.1. Vowel duration

The final regression model assessing vowel duration included Intended Coda, vowel duration, and the interaction between them as fixed effects and random intercept for talkers, listeners, and items as pairs. The results showed that listeners' responses were still biased toward non-lenis coda ($\beta = 1.116$, SE = 0.136, $\chi^2(1) = 49.09$, $p < 0.001$). The main effect of vowel duration was significant, indicating that listeners non-lenis response was affected by vowel duration ($\beta = -17.671$, SE = 1.709, $\chi^2(1) = 101.96$, $p < 0.001$).

5.3.2.2. Voicing into closure duration

The final regression model assessing voicing into closure duration included Intended Coda, voicing into closure duration, and the interaction between them as fixed effects. It also included random intercept for talkers, listeners, and items (as a pair) along with random slopes for the fixed effects and their interactions by talkers and listeners. The results showed that listeners' responses were not biased toward either lenis or non-lenis ($\chi^2(1) = 1.57$, $p = 0.11$).

Unlike in glm model, the main effect of voicing into closure was not significant, suggesting that listeners' non-lenis response was not affected by voicing into closure duration ($\chi^2(1) = 0.17$, $p = 0.67$). Overall, it seems that voicing into closure duration did not serve as a perceptual cue that helped listeners to identify either non-lenis coda in general or a correct underlying form of neutralized coda.

5.3.2.3. Closure duration

The models assessing closure did not converge even after eliminating random intercept for items. They were thus excluded from the final report.

5.4. Discussion

The purpose of regression modeling was to examine whether acoustic cues made any contribution to listeners' identification of underlying coda. When regression models took into account the effect of acoustic cues, it turned out that listeners' overall responses were biased toward non-lenis coda and influenced by vowel duration and voicing into closure duration, but not by closure duration. That is, when vowel duration was shorter and voicing into closure duration was longer, listeners showed a strong tendency to favor non-lenis responses regardless of the underlying status of coda.

This effect varied across three types of neutralization. The odds listeners provide less non-lenis responses when vowel duration gets longer were greater in manner or palatal neutralization (with the latter stronger in the effect size) than in laryngeal neutralization. The odds listeners provide more non-lenis response when closure duration gets longer were greater for codas subject to manner or palatal neutralization than for codas with laryngeal contrasts.

When the model included random effects such as talkers, listeners, and items, the effect of voicing into closure duration on listeners' non-lenis response did not reach significance. However, the effect of vowel duration was still significant. Therefore, it seems that vowel duration serves as a predictor of listeners' response bias toward non-lenis.

On the other hand, it turned out that listeners' discrimination was marginally significant. Listeners showed a non-significant trend toward correctly identifying non-lenis when the target was non-lenis. It was also revealed that any of acoustic cues did not make a significant contribution to facilitating listeners' discrimination performance. However, the influence of vowel duration and voicing into closure duration on discrimination varied across manner and palatal neutralization. Each acoustic parameter's influence on discrimination was stronger in palatal neutralization, although the main effect was not significant in either type of neutralization.

Overall, a series of regression analyses confirmed that listeners' discrimination was around chance level. Although listeners seem to have relied on vowel duration in guessing their responses, they did not appear to benefit from any acoustic cues in correctly identifying non-lenis coda when the target was non-lenis. This finding implies that listeners were at least sensitive to the difference in vowel duration but not able to make an effective use of it in discriminating non-lenis from lenis.

CHAPTER 6. GENERAL DISCUSSION

6.1. Summary of the findings

The aim of the present study was to investigate acoustic and perceptual aspects of coda neutralization in Korean. By extending the Kim and Jongman (1996)'s study that documented complete neutralization of manner contrasts in coda obstruents, we examined all three types of coda neutralization in Korean (i.e., laryngeal, manner, and palatal) in production and perception. The first goal of the production study was to reveal whether neutralization of final obstruents is phonetically complete or incomplete. This was addressed through the analysis of acoustic correlates of final obstruents that included preceding vowel duration, voicing into closure duration, and closure duration both in a non-neutralizing and neutralizing environments. Another goal was to examine the effect of explicit orthographic cues on coda neutralization. For this purpose, we elicited stimuli in both spontaneous and read speech in which orthography was either absent or present and compared acoustic properties of the target words elicited in each task.

A subsequent perception study aimed to investigate the perceptual consequences of neutralization of coda obstruents. We conducted a 2AFC identification experiment in order to reveal whether native Korean listeners are able to identify the correct underlying form of neutralized codas by utilizing acoustic cues. The listeners were auditorily presented with the neutralized words containing lenis or non-lenis coda obstruents in which the subtle acoustic differences were either present or absent. They were then asked to select the correct underlying form of the target out of the two alternatives presented in Korean orthography.

The results of the production study showed that three different types of coda pairs in a non-neutralizing environment were differentiated by vowel duration and/or closure duration.

When neutralized, significant acoustic differences were still observed in some of the coda pairs. Specifically, coda pairs with laryngeal contrasts were incompletely neutralized in terms of vowel duration, voicing into closure duration, or closure duration. Interestingly, the effects of incomplete neutralization were stronger in the spontaneous speech than in the read speech. For the pairs subject to manner or palatal neutralization, we obtained mixed findings: the pairs with manner distinctions were differentiated by vowel duration or closure duration in one or the other speech, whereas the pairs subject to palatal neutralization were completely neutralized in terms of all acoustic parameters in the both speech types.

In perception, however, we found that all three types of coda obstruents were completely neutralized. Regardless of the presence of significant acoustic differences in coda obstruents, the listeners were not able to successfully label neutralized words with the correct underlying forms. In a subsequent regression analysis, it was revealed that listeners relied on vowel duration in guessing their responses, but were not able to make effective use of the cue in correctly identifying targets.

Taken together, the findings of this study suggest that coda obstruents in Korean are not always completely neutralized in production, but exhibit complete perceptual neutralization.

6.2. Discussion

6.2.1. Asymmetrical pattern of neutralization across the three types of coda neutralization

The current study has newly established that the effects of neutralization vary across the three types of coda neutralization. It is shown that the pairs with laryngeal contrasts are incompletely neutralized, with its effects stronger in the spontaneous speech. The most reliable acoustic correlate of coda obstruents is the duration of preceding vowel, as in the previous studies (Charles-Luce, 1985; 1993; Dmitrieva et al., 2010; Fourakis & Iverson, 1984; Port & O'Dell, 1985; Roettger et al., 2014; Slowiaczek & Dinnsen, 1985; Warner et al., 2004). For manner or palatal neutralization, however, the effects of incomplete neutralization are confined to a single acoustic cue and one speech type or absent in the both types of speech.

One reason for the asymmetrical pattern of neutralization may be the different number of underlying representations that induce a homophone as a consequence of neutralization. For 75% (9 out of 12 pairs) of the stimuli with laryngeal contrasts, only two underlyingly distinct coda obstruents are merged and produce a homophone (e.g., /tʃip/, /tʃipʰ/ → [tʃip]). On the other hand, most of coda obstruents (14 out of 16 pairs, 88%) subject to manner or palatal neutralization have multiple competitors that are neutralized into a single homophone (e.g., /mit/, /mitʰ/, /mis/, /mitʃʰ/ → [mit]).

The exemplar-based models assume that acoustically similar categories overlap in their phonetic distributions. Since neutralizing competitors are highly similar in their acoustic characteristics, they may exhibit extensively overlapped phonetic distributions. The exemplar space for a category would then include a number of exemplars of its neutralizing competitors. Given that the production target is created based on the averaged phonetic characteristics of randomly selected exemplars within the exemplar space, the production targets and their associated articulatory implementation processes may share great similarity across the competing categories. This high degree of similarity will make it more likely that productions of each category will be more similar. It is thus likely to be more challenging for speakers to produce one category distinctively from multiple neutralizing competitors, as in the cases of manner and palatal neutralization, than to separate one from just one other, as in the case of laryngeal neutralization.

Another possibility is that the asymmetry is related to the pattern of free variation observed in stem-final obstruents in Korean (Kang, Lee, & Kim, 2004; Jun & Lee, 2007; Oh & Shin, 2007). Jun and Lee (2007) suggest that underlyingly coronal stem-final obstruents exhibit a wider range of coronal variants ([s ~ t ~ t^h ~ tʃ ~ $tʃ^h$])[9] than underlyingly non-coronal obstruents whose variation is binary (bilabial: [p ~ p^h], velar: [k ~ k^h], [k ~k']). For example, /sot^h.e/ ('pot' + dative case marker) supposed to be pronounced as [so.t^he] are often produced with multiple variants such as [so.se], [so.$tʃ^h$e], and [so.te] by some speakers (Kang et al., 2004).

In the current study, less than 1% of the productions show free variations, all of which are the variants of the targets with manner contrasts. In spite of its negligible effects, free variations more prevalent in codas subject to manner/palatal neutralization could have affected the retrieval of the underlying representation of a coda obstruent and the production of its neutralized form. Assuming that all experienced tokens including morphologically related forms of a word are stored as exemplars and affect the process by which the production target is constructed (Ernestus & Baayen, 2006, 2007; Pierrehumbert, 2000), it is possible that the presence of multiple variants may diminish the expression of the neutralized cue.

Even though these speculations may account for the current findings, it is important to confirm in future research whether similar asymmetry in neutralization pattern still holds when the number of target items is controlled across the three types of neutralization. As discussed in Chapter 3, the number of target pairs subject to manner/palatal neutralization is too small to yield statistically reliable durational patterns, compared with that of the pairs with laryngeal contrasts.

[9] The degree of preference for the stem-final variants ([s] (most preferred) > [$tʃ^h$] > [t^h] > [t], [tʃ]) is determined by the relative frequency of nouns containing a given stem-final obstruents (Albright, 2008; Jun, 2007).

Therefore, it is hard to conclude that the current findings on complete neutralization are representative of manner/palatal neutralization in Korean. If we observe a similar asymmetrical pattern of neutralization in a more controlled study, we would be able to propose that the degree of neutralization is closely related to the number of competitors or their variants that are neutralized into a single homophony.

6.2.2. Incomplete laryngeal neutralization in the Q & A task

At the onset of the study, we made two predictions on the role of explicit orthographic cues in coda neutralization. Based on the previous finding that the influence of orthography is the main trigger of incomplete neutralization (Fourakis & Iverson, 1984; Jassem & Richter, 1989; Warner et al., 2004, 2006), we predicted that word pairs elicited from the reading task in which orthographic information is available would be more subject to incomplete neutralization. Another prediction was that if coda neutralization is independent of orthographic influences, there would be no acoustic difference between the neutralized obstruents elicited from the two tasks.

Our results show that laryngeal neutralization is incomplete, but contradicting the first prediction, its effect is stronger and more consistent in the items elicited in the Q & A task than in the reading task. Second, this result confirms the previous finding that incomplete neutralization does take place even in the absence of explicit orthographic cues (Kharlamov, 2014, 2015; Port & Crawford, 1989; Piroth & Janker, 2004; Roettger et al., 2014). It thus casts doubt on the claim that incomplete neutralization is a result of orthographic interference associated with a reading task (Fourakis & Iverson, 1984; Jassem & Richter, 1989; Warner et al., 2004, 2006).

However, it should be noted that this finding does not necessarily rule out the possibility that orthographic information still influences the production of neutralized words even when talkers are not accessible to written form of the words. A number of empirical studies have revealed that orthographic information is processed even during auditory spoken word recognition in which the spelling of a word is not presented. For example, Seidenberb and Tanenhaus (1979) find that participants show faster rhyme decision latencies for orthographically similar cue-target pairs (e.g., pie-lie) than orthographically dissimilar pairs (e.g., rye-lie) presented auditorily. Similarly, auditory lexical decision times are longer for words containing rimes that could be spelled in multiple ways (e.g., [i:p] could be spelled '-eep' or '-eap') than for words whose rimes are always spelled only one way (e.g., [ʌk] '-uck') (Stone, Vanhoy, & Van Orden, 1997; Ziegler & Ferrand, 1998; Ziegler, Petrova, & Ferrand, 2008).

These studies provide converging evidence that orthographic representation is activated during spoken word processing and that mapping between orthography and phonology takes place. Given this evidence, it could be the case that once the talkers in the present study listen to the question prompt and process it, they may still activate orthographic representation of a target word they retrieve along with its phonological information. It thus appears more correct to say that we cannot entirely eliminate the influence of orthography even in an auditory task but rather can only manipulate the presence or absence of explicit orthographic cues.

The result that stronger effects of incomplete neutralization are observed in the productions elicited from the Q & A task suggests that incomplete neutralization is affected by a communicative context. This is in line with the findings of Port and Crawford (1989) that a stronger effect of incomplete neutralization is observed in a task in which speakers read sentences for the experiments to dictate them than in a non-communicative task. Specifically, the participants in this study produce neutralized words with a larger difference in vowel duration when encouraged by a communicative context to convey the underlying phonological contrast than when asked to read a list of words.

An implication of the current finding is that incomplete neutralization is more likely to occur when talkers are encouraged to pay more attention to the goal of communication, i.e., conveying their intended messages effectively to listeners. In the current study, the main task of the talkers in the Q & A task is to listen to the oral prompts recorded by the experimenter (i.e., the listener) and complete the sentences by filling in the right target word. When the talkers provide a wrong answer, they are given another hint on site by the experimenter. Even though it is under a highly controlled experimental setting, the Q & A task involves an interaction between the talker and the listener, which lacks in the reading task. Therefore, it is possible that talkers have produced the neutralized target pairs with more temporal contrasts in the preceding vowel in order to ensure that the listener discriminates the target word from the other.

This explanation is supported by Lindblom's Hyper- and Hypo-articulation (H & H) theory (1990). It postulates that talkers vary their speech productions along a continuum from hyper- to hypo-articulation by taking into account the goal of communication and listeners' needs so that listeners can successfully discriminate talkers' intended messages from others. Talkers tend to produce words with greater clarity when listeners do not have sufficient signal-independent information to use and thus require maximum acoustic information (e.g., when words are not predictable from the context). On the other hand, they reduce or minimize articulatory efforts when listeners have enough access to signal-independent sources, resulting in hypo-articulated speech.

Here, an important point relevant to the current finding is that hyper-articulated speech is guided by contrast enhancement strategies, one of which is the enhancement of vowel duration contrast as a way of maximizing the distance between contrastive sound categories. For example, Uchanski (1988, 1992) showed that the temporal contrast between tense and lax vowels in English is increased in a hyper-articulated speech through more lengthening of tense vowel than lax vowel.

Even though the current study demonstrates that incomplete neutralization is more likely to occur in a communicative-driven task, we need to further investigate the possibility that incomplete neutralization is driven by confounding factors related to experimental design. First, it is possible that a stronger effect of incomplete neutralization in the Q & A task may have stemmed from the focus marking of target words. In the Q & A task, the oral prompts auditorily presented to talkers contain the question followed by the format of the answer[10]. The task of the talkers is to complete the format sentence by substituting the demonstrative pronoun ‘this’ with the target word. In this context, the target word is marked with contrastive focus/emphasis, since it is the new information relative to the context in the format sentence (Nadig & Shaw, 2015). Therefore, it is possible that talkers hyper-articulate the target words to emphasize it in contrast to other lexical items.

---

[10] Below is the one of the trials in the Q & A tasks.

Q: This is the time between sunrise and sunset. (Dictionary definition of the target)
During the winter, it is very cold and this also gets shorter, since the sun sets earlier. (Example context)
What does it get shorter, too? (Question)
Please say “<u>This</u> gets shorter, too”. (Instructions on how to answer)

A: <u>The day</u> gets shorter, too.

Previous studies have shown that focused elements exhibit acoustic characteristics such as longer duration, higher F0 or pitch, and greater intensity (Bolinger, 1961; Watson, 2010). Particularly, studies on the focal lengthening (i.e., the lengthening of duration as a temporal effect of focus) have revealed that the durational contrast between long and short segments is enhanced in focus (Bannert, 1979; Heldner & Strangert, 2001; de Jong & Zawaydeh, 2002). For example, Heldner and Strangert (2001) showed that focal lengthening often occurs within the stressed syllable, and that phonologically long segments in this focal domain are lengthened more than phonologically short segments.

Similarly, de Jong and Zawaydeh (2002) found that the durational difference of English vowels preceding voiced vs. voiceless coda consonants increase when the target words are marked with focus, which consequently enhances voicing distinction. The findings of the both studies suggest that the non-linear lengthening leads to the increased durational contrast between long and short vowels. In this light, the current findings could be understood as a similar strategy to enhance the contrast in a focused context. That is, the durational contrast of preceding vowels is enhanced through asymmetrical lengthening of vowels before underlyingly lenis vs. non-lenis obstruents.

The second possibility is that significant durational differences in vowel duration observed in the Q & A task may have been influenced by the order of the task. We conducted the Q & A task first in order to prevent participants from being exposed to orthographic cues to the target stimuli that is available in the reading task. In an effort to minimize possible learning effects, we also kept the interval between the tasks ranging from at least 2 hours to a day. However, it is still possible that the performance of the reading task may have been affected by that of the Q & A task.

The distinction of lenis vs. non-lenis obstruents involves an overall difference in 'force of articulation' (Belasco, 1953). Non-lenis (aspirated or fortis) obstruents take more force to produce than lenis obstruents, and the anticipation of greater articulatory efforts to produce non-lenis obstruents shortens the preceding vowel more. Even if this distinction is supposed to be neutralized in coda position, we find that the effect of vowel shortening is still observed in the Q & A task. The reason for the absence of the same effect in the reading task could be that learning effects lead talkers to put less articulatory efforts in producing underlying non-lenis codas, which results in longer vowel duration than in the Q & A task. Future work may need to examine these possibilities in order to confirm that incomplete neutralization is not an artifact of experimental compounds.

6.2.3. Complete neutralization in perception

Another important finding of this study is that three types of coda obstruents are completely neutralized in perception. To the best of our knowledge, this is the first reported case that incomplete neutralization in production is completely neutralized in perception. First of all, the failure of the identification of coda pairs with no acoustic differences suggests that listeners do not make substantial use of unmeasured acoustic cues to underlying coda contrast for identification. Second, the failure to identify a correct underlying form of the pairs with significant acoustic differences could be attributed to several factors.

First, listeners' failure to correctly identify the intended coda may be due to the small magnitude of acoustic differences. That is, systematic acoustic differences in vowel duration and voicing into closure duration in pairs with laryngeal contrasts may not be large enough to be perceived by native Korean listeners. According to Klatt (1976), rule-governed systematic changes in vowel duration less than one JND (just noticeable difference; e.g., 25 ms in vowel duration in English) are perceptually less important than changes that exceed one JND. If the magnitude of differences in each of the two parameters is below one JND threshold for Korean, it is highly probable that listeners do not even make use of these durational cues for the identification of the underlying coda.

Second, it may be related to listeners' insensitivity to the temporal acoustic information distinguishing neutralized codas. When processing neutralized codas in a normal speech communication situations, it is possible that native Korean listeners solely rely on context and not the acoustic signal to retrieve their underlying forms. As one of acoustic correlates to coda obstruents, vowel length had been a contrastive feature of Korean vowels, but lost in standard Seoul Korean (Magen & Blumstein 1993). As a result, most of younger speakers either do not distinguish vowel length in a consistent manner or cannot distinguish it at all (e.g., Ingram & Park, 1997; Park, 1994). Given that all of the participants use Seoul dialect as their native language and have had little allophonic experience with vowel duration, it is not surprising that they are not sensitive to subtle differences in vowel duration in discriminating between neutralized codas, especially in an experimental setting in which no contextual cues are provided.

Thirdly, the identification failure may be more related to listeners' failure to direct their perceptual attention to acoustic cues than their perceptual insensitivity. In the current study, listeners' task was to classify the targets into the two categories. The task requires the listeners to focus their attention on phonemic category-level contrast rather than on surface phonetic differences. It is therefore possible that the listeners might have perceived subtle differences in temporal-acoustic cues but not paid enough attention to them. It is also probable that high variability task may have led to listeners' failure to direct their attention to a particular acoustic cue. In the task, a wide range of items and multiple talkers were presented from trial to trial in a pseudo-random manner. Since listeners had to quickly readjust and retune their perceptual criteria from trial to trial, which requires extra processing resources and attentional capacity (Goldinger, Pisoni, & Logan, 1991; Rabbit, 1966), they may have experienced difficulty attending to one particular cue in making a classification.

If the identification failure is closely related to the hard nature of the task, then Korean listeners might be able to perceive subtle differences in temporal acoustic cues in different experimental settings. Warner et al. (2004) provides evidence that testing conditions can affect listeners' identification performances. They show that under high-variability condition, listeners' identification accuracy is above the chance level only for the productions of 2 out of 4 talkers. However, listeners' performances significantly improve in low-variability condition in which only a single acoustic cue (i.e., vowel duration) is available, stimuli are blocked by speaker and word, and listeners are familiarized with the target items before the test. This implies that it may be relatively easier for listeners to pick up on an acoustic cue under the best experimental condition for directing their perceptual attention. Considering this evidence, it is possible that complete neutralization in perception is not due to small magnitude of acoustic cues or listeners' insensitivity, but due to the nature of the task. Therefore, it remains a question for future research whether Korean listeners' identification would be affected by the type of experimental tasks.

6.2.4. Theoretical implications

6.2.4.1. Traditional generative approach vs. Exemplar-based approach

The present findings have important consequences for theories of speech production and perception. In production, the findings on manner/palatal neutralization are compatible with the traditional generative approach. Within this approach, the phonological computation makes use of discrete categories (i.e., abstract representations and rules) and produces outputs that exhibit identical feature values for the neutralized categories. Since the contrast between the underlyingly distinct words (e.g., /sot/ vs. /sot^h/) is eliminated in this stage, the phonetic mechanisms receive the identical output of phonological derivation as the input (i.e., /sot/). Consequently, the two words are physically realized with the same acoustic-phonetic properties.

However, the generative framework cannot account for incomplete laryngeal neutralization. Since the phonetic mechanisms only compute the degree and timing of intended articulatory gestures for the single input, they cannot reflect a gradient nature of neutralization process. Rather, laryngeal neutralization can be nicely explained by the account couched within exemplar-based models (See for the detailed explanation section 2.2.2). According to the Ernestus and Baayen's proposal (2006, 2007), incomplete neutralization could be viewed as the reflection of lexical representations that contain detailed acoustic-phonetic properties of final obstruents. The lexical representation of the Korean word 'house' (/tʃip/), for example, may contain the probabilistic information that the final obstruent /p/ is produced with longer vowel duration than the obstruent /p^h/ in the word 'straw' (/tʃip^h/). Since speakers make use of this information in the production of the word 'house' vs. 'straw', they make subtle acoustic distinctions between the two words.

Ernestus and Baayen (2006, 2007) also suggest that the physical realization of final obstruents may be influenced by the pronunciations of co-activated morphological neighbors. For the current data, it is expected that the activation of the Korean word 'house' (/tʃip/), for example, may co-activate inflected forms that contain a non-neutralized obstruent (e.g., [tʃi.be], [tʃi.bi], [tʃi.bɨl], [tʃi.bɨro]). Since the voicing feature of the stem-final obstruents in the inflected form is [+voice], it may affect the probability that the final obstruent of a singular form is realized as slightly voiced. Comparable voicing effects are not expected to occur in the item 'straw' (/tʃipʰ/) for which all inflected forms contain non-neutralized voiceless obstruents. This prediction is confirmed by the current finding that vowel duration before underlyingly voiced laryngeal coda obstruents is significantly longer than before underlyingly voiceless counterparts.

Now, we see that the Ernestus and Baayen's lexical analogy account is challenged by Korean data, since it only predicts incomplete neutralization. The question is, then, how can the exemplar-based models explain the asymmetrical pattern of neutralization between laryngeal and manner/palatal neutralization? In addition, the current account predicts that the co-activation of inflected forms affects listeners' identification of neutralized words and leads to incomplete neutralization in perception. Contrary to the prediction, however, we observe complete neutralization in perception even when acoustic differences are present. Then another question is how the exemplar-based account can explain Korean data that exhibits the asymmetry across production and perception.

6.2.4.2. Asymmetry in production across laryngeal vs. manner/palatal neutralization

I suggest that the asymmetry between laryngeal and manner/palatal neutralization may stem from the different number of the categories that are neutralized into a homophone. For laryngeal neutralization, most of the pairs involves the merge of two categories (e.g., /tʃip/, /tʃipʰ/ → [tʃip]). On the other hand, a majority of coda obstruents subject to manner or palatal neutralization have multiple competitors that are neutralized into a single homophony. For example, the final obstruents in the morphemes /nas/ ('sickle'), /nat/ ('grain'), /natʰ/ ('a piece'), /natʃ/ ('day'), and /natʃʰ/ ('face') are all neutralized into [t].

From the viewpoint of the exemplar-based models, the categories neutralized into one homophone may have extensively overlapped phonetic distributions due to their acoustic similarities. As a category has a more number of neutralizing competitors, its exemplar space may contain a wider range of exemplars of the competing categories. This results in production targets and articulatory implementation process highly similar across the categories. This high degree of similarity will make it more likely that productions of each category will be more similar. Unless provided with contextual cues, it would be extremely difficult to make acoustic distinctions among the neutralized words, especially when the target is one member of multiple neutralizing categories.

This speculation is consistent with the current results. The speakers produce significant acoustic differences between the neutralized words with laryngeal contrasts, especially in the Q & A tasks. Within the Q & A task, the effects of incomplete neutralization are more robust in laryngeal neutralization (which involves the merger of two distinct segments in most of the cases) than in manner or palatal neutralization (which involve the merger of multiple distinct categories). In the latter case, weaker or null effects of incomplete neutralization may be due to the fact that multiple neutralizing categories may exhibit extensive overlap in their phonetic distributions, possibly with some categories overlaid on others.

6.2.4.3. Asymmetry across production and perception

Across production and perception, there is an asymmetry, in that there is perceptual neutralization of laryngeal contrasts that exhibit significant acoustic differences. This pattern is similar to near merger, the situation where speakers produce two sounds with systematic acoustic differences but cannot perceive those differences. It is similar to what Janson and Schulman (1983) find in speakers of Lycksele dialects of Swedish: they maintain the contrast between the vowels [e] and [ɛ], but fail to make a perceptual distinction between the two. Similarly, Yu (2007) show that morphologically derived and underived mid-rising tones in Cantonese are produced with significant differences in F0, but completely merged in perception. Although near merger is not conditioned by a particular phonological environment, it is very similar to incomplete neutralization in nature.

According to the exemplar-based models, perceptual neutralization or near merger occurs due to extensive overlapping in the phonetic distributions for the two categories that are kept separate (Pierrehumbert, 2002; Yu, 2007). The exemplars within a category show gradient variations that may stem from anatomical differences inherent in the vocal tract as well as from different speaking habits of individual speakers. If two categories exhibiting a given range of variations are highly similar in their phonetic properties, then their exemplar clouds overlap in their phonetic distributions. Even in this situation, each category can still remain separate, since tokens with ambiguity in some phonetic dimensions could still be correctly categorized based on other salient dimensions as well as contextual or indexical cues.

Yu (2007) suggests that as far as the two categories remain separate, the production of two merging categories exhibits distinct acoustic-phonetic differences. In this case, the magnitude of the differences is expected to be very small due to the inclusion of exemplars in the overlapping clouds. With regard to perception, however, he suggests that it is more challenging than production. Listeners would have much difficulty in identifying the correct category label for the near-merging categories, especially in a highly controlled experimental setting. Imagine a situation that listeners need to make categorical decisions solely based on very limited amount of acoustic cues in the absence of contextual information (e.g., in a forced-choice identification task). In such a setting, the identification of the phonetically ambiguous tokens that fall within the combined exemplar space is expected to be extremely difficult, thus failing to reach above chance levels.

Now, the puzzle is why the Korean language, but not languages with final devoicing, exhibits the pattern of neutralization similar to near merger. An increasing number of studies have documented incomplete neutralization, but none of the studies has reported that neutralization is complete only in perception. I suggest that distinct patterns of neutralization across languages may be related to the fact that languages differ in the dimensions of phonological contrasts (or the number of phonological features) that are neutralized as well as the number of the segments that are neutralized into a single output.

Unlike final devoicing that involves the merger of two distinct categories into a single dimension of contrast, i.e., voicing[11], coda neutralization in Korean covers phonological contrasts in two separate dimensions, laryngeal and manner[12]. Additionally, for each dimension of contrasts, two or more categories are neutralized into a single surface form in most of the cases. This is especially true for the underlyingly coronal coda obstruents (/s/, /t/, /tʰ/, /tʃ/, /tʃʰ/). All possible pairs among these obstruents are the cases of manner or palatal neutralization,

[11] It has been suggested that aspirated stops in Germans are underlyingly marked with [+spread glottis] (Anderson & Ewen, 1987; Jessen, 1996, 1998; Jessen & Ringen, 2002; Iverson & Salmons, 1995). In case of German, therefore, the phonological feature that involves neutralization is not [voice] but [spread glottis] feature.
[12] Both laryngeal and palatal neutralization involves the merger of laryngeal contrasts between the two stops or palatal affricates.

except the pair /t-t^{h}/ subject to laryngeal neutralization. Given that coda neutralization exhibits more complex patterns than final devoicing, it can impose extra processing burdens on listeners.

Imagine that in a 2AFC identification task, listeners are presented with three target words in a sequence of trials, each of which is subject to different types of neutralization (e.g., /pitʃ^{h}/ 'light' (palatal) → /tʃɪp/ 'house' (laryngeal) → /nas/ 'sickle' (manner)). Aside from the processing effort required to identify the correct underlying form, the shift in the dimension of neutralization from trial to trial (e.g., palatal → laryngeal, laryngeal → to manner) could require extra processing. The shift from the second to the third trial may impose even greater processing demands due to the presence of multiple neutralizing competitors. Specifically, when the input signal /nas/ ('sickle') is presented, the category that best matches the input signal would be activated along with its nearest neighbors (/nat/, /nat^{h}/, /natʃ/, /natʃ^{h}/, /nass/). Considering that those categories exhibit extensive overlap in their phonetic distributions, it would be extremely difficult for listeners to identify the correct underlying form solely based on acoustic cues. Consequently, listeners fail to assign the perceived token with the correct category membership, and consequently exhibit complete neutralization.

Taken together, the above speculations challenge the current exemplar-based account that explains incomplete neutralization of final devoicing, but cannot account for coda neutralization in Korean. In order to incorporate the Korean data, the account needs to be modified and extended in such a way that it explains that the degree of incomplete neutralization in production and perception can vary depending on the number of categories neutralized into a single output as well as the dimensions of phonological contrasts that are neutralized.

CHAPTER 7. CONCLUSIONS AND FUTURE DIRECTIONS

The present study investigated three types of coda neutralization in Korean in production and perception. The acoustic analysis of target words elicited from two different tasks reveals that neutralization pattern varies across the three type of coda neutralization and across the tasks. Coda pairs with laryngeal contrasts exhibit robust effects of incomplete neutralization, while coda pairs subject to manner or palatal neutralization show either weaker or null effects. For laryngeal neutralization, stronger effects of incomplete neutralization are observed in the Q & A task in which explicit orthographic cues are absent. This result suggests that incomplete neutralization occurs independently of explicit orthographic cues and is affected by a communicative context.

Unlike in production, however, all types of coda obstruents are completely neutralized in perception. The identification of the underlying form of lenis vs. non-lenis obstruents is around chance-level, regardless of the type of coda neutralization. It is also demonstrated from the regression analyses that listeners do not make an effective use of the acoustic cues available in the final obstruents to label the targets with the correct underlying forms.

As the first large-scale study that establishes acoustic and perceptual aspects of coda obstruents in Korean, this study provides evidence that the degree of neutralization varies across the three types of coda neutralization and across the production and perception. We suggest that these asymmetries may be closely related to the complex structure of coda neutralization - the neutralization of two or multiple categories across the two dimensions of phonological contrasts. The finding of incomplete neutralization extends our understanding that incomplete neutralization is not a phenomenon confined to final devoicing, but observed in a qualitatively different type of neutralization. Additionally, it suggests that phonological process of neutralization that has long been assumed to be categorical is gradient in nature. This supports the exemplar-based account that views incomplete neutralization as the reflection of fine-grained lexical representations or as the influence of morphological neighbors. At the same time, however, this account is challenged by our finding of perceptual neutralization, since it cannot explain why significant acoustic differences are not perceived by Korean listeners even though the differences are similar in their magnitudes to those observed in previous studies.

The findings of the current study raise several important questions to be addressed in future work. As discussed in section 6.2.3, the possibility exists that complete neutralization in perception is due to the difficult nature of the task - a high variability identification task that asks listeners to make a categorical judgment. Future research would be needed to further investigate whether perceptual neutralization is due to the nature of tasks or due to other factors such as small magnitude of acoustic differences or listeners' insensitivity to them.

One line of a follow-up study is to conduct a discrimination task in low-variability listening condition as in Warner et al. (2004). In such an experiment, acoustic cues could be manipulated in such a way that vowel duration is the only acoustic cue to be presented, with the stimuli blocked by speakers and words. If we observe that identification accuracy is significantly improved, it may imply that listeners are able to make use of vowel. If then, we may be able to suggest that complete neutralization in perception observed in this study is influenced by the nature of the task, and therefore that the degree of neutralization in perception may vary depending on tasks.

Secondly, this study has shown that coda obstruents with laryngeal contrasts are incompletely neutralized, with its effects stronger in the targets elicited from the Q & A task. It remains unclear whether incomplete neutralization is driven by some other factors inherent in the experimental design such as the position of targets and the order of tasks. Future work will need to control them in order to confirm that incomplete neutralization is not an artifact of these experimental compounds. For example, we can examine the effect of focus marking on incomplete neutralization by comparing (within the same experimental setting) acoustic properties of the targets that are marked with contrastive focus vs. those that are not. If the effects of incomplete neutralization are observed only in targets in focus, it may suggest that incomplete neutralization is the output of focal lengthening. If the effects are observed in both types of targets, on the other hand, it will imply that incomplete neutralization is a genuine linguistic phenomenon that occurs regardless of the focus-marking of targets.

Thirdly, the current investigation of coda obstruents is confined to monosyllabic words in certain lexical categories due to the constraint inherent in the experimental design. We used the exhaustive list of minimal pairs whose lexical category is either noun or verb that can be placed in the beginning of the sentence. Even though the present study has revealed the overall patterns of three types of coda neutralization in production and perception, further investigation using a broader range of items would be required in order to understand the more precise nature of coda neutralization. Particularly, more research on manner/palatal neutralization would be necessary, since the lack of robust incomplete neutralization effects could be due to the insufficient number of target items.

Another interesting question worthy of further analysis would be the effect of lexical properties on incomplete neutralization. In this study, we only provided the results of acoustic analyses without considering word frequency of targets. So one testable question is whether the degree of incomplete neutralization is modulated by the frequency of targets as well as the frequency of co-activated inflected forms of the targets. In the exemplar-based model, as discussed in section 2.2.2, more frequent categories have a larger number of exemplars than less frequent categories whose exemplars are less numerous. Since high-frequency items have higher activation levels, they are expected to be accessed quickly and accurately than low-frequency items in experimental settings. Based on this assumption, we can test whether high frequency targets show stronger effects of incomplete neutralization than low frequency targets, if all else being equal.

With regard to the effect of inflected forms, a singular word with highly frequent inflected forms may exhibit a larger degree of incomplete neutralization than a word that has less frequent inflected forms. However, there could be the case that many of highly frequent inflected forms of a target contain neutralized obstruents. In Korean, nouns and verbs inflected via suffixation have a wide array of inflected forms that contain either neutralized or non-neutralized stem-final obstruents. So the activation of the word 'house' (/tʃip/), for example, co-activates not only the inflected forms that contain non-neutralized obstruent (e.g., [tʃi.be], [tʃi.bi], [tʃi.bɨn], etc.) but also the forms with neutralized stem-final obstruents (e.g., [tʃip.k'wa], [tʃip.t'o], [tʃɪp.t'ɨl], etc.; See Appendix F for more examples)

Given this property, it is possible that lesser degree of incomplete neutralization is observed in targets with highly frequent inflected forms containing neutralized stem-final obstruents than those containing non-neutralized stem-final obstruents. Taken together, this line of research would suggest that the investigation of incomplete neutralization needs to consider the lexical properties of target items and their morphological neighbors within a given language under analysis.

REFERENCES

Ahn, H. (1999b). Post-release phonatory processes in English and Korean: acoustic correlates and implications for Korean phonology. Ph.D. dissertation, University of Texas at Austin.

Albright, A. (2008). Explaining universal tendencies and language particulars in analogical change. In J. Good (Ed.), *Language Universals and Language Change* (pp. 144-181). New York: Oxford University Press.

Alegre, M., & Gordon, P. (1999a). Frequency effects and the representational status of regular inflections. *Journal of Memory and Language*, *40*(1), 41-61.

Anderson, J. M., & Ewen, C. J. (1987). *Principles of dependency phonology*. Cambridge: Cambridge University Press.

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, *59*(4), 390-412.

Baayen, R. H., McQueen, J., Dijkstra, T., & Schreuder, R. (2003). Frequency effects in regular inflectional morphology: Revisiting Dutch plurals. In R. H. Baayen & R. Schreuder (Eds.), *Morphological Structure in Language Processing* (pp. 355-390). Berlin: Mouton de Gruyter.

Baayen, R. H., Dijkstra, T., & Schreuder, R. (1997). Singulars and plurals in Dutch: Evidence for a parallel dual-route model. *Journal of Memory and Language*, *37*(1), 94-117.

Bannert, R. (1979). The effect of sentence accent on quantity. In *Proceedings of the Ninth International Congress of Phonetic Sciences,* Copenhagen, pp. 253-259.

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language,* *68*(3), 255-278.

Belasco, S. (1953). The influence of force of articulation of consonants on vowel duration. *Journal of the Acoustical Society of America*, *25*(5), 1015-1016.

Bolinger, D. (1961). Contrastive accent and contrastive stress. *Language*, *37*(1), 83-96.

Braver, A. (2011). Incomplete neutralization in American English flapping: A production study. *University of Pennsylvania Working Papers in Linguistics*, *17*(1), 31-40.

Braver, A., & Kawahara, S. (2014). Incomplete vowel lengthening in Japanese: A first study. In *Proceedings of the 31st Meeting of the West Coast Conference on Formal Linguistics*, Somerville, pp. 86-95.

Bybee, J. (2001). *Phonology and language use*, Cambridge: Cambridge University Press.

Bybee, J. (2005). Restrictions on phonemes in affixes: A crosslinguistic test of a popular hypothesis. *Linguistic Typology*, *9*(2), 165-222.

Chang, C. B. (2007). Korean fricatives: Production, perception, and laryngeal typology. *UC Berkeley Phonology Lab Annual Report*, 20-70.

Charles-Luce, J. (1985). Word-final devoicing in German: effects of phonetic and sentential contexts. *Journal of Phonetics*, *13*(3), 309-324.

Charles-Luce, J. (1993). The effects of semantic context on voicing neutralization. *Phonetica*, *50*(1), 28-43.

Charles-Luce, J., & Walker, L. A. (1981). Effects of linguistic context on the durations of lexical categories. *Research on Speech Perception: Progress Report 7*. Speech Research Laboratory, Indiana University, 263-272.

Chen, M. (1970). Vowel length variation as a function of the voicing of the consonant environment. *Phonetica*, *22*(3), 129-159.

Cho, T., Jun, S. A., & Ladefoged, P. (2002). Acoustic and aerodynamic correlates of Korean stops and fricatives. *Journal of Phonetics*, *30*(2), 193-228.

Chung, K. (1980). *Neutralization in Korean: a functional view*. Ph.D. dissertation, University of Texas at Austin.

De Jong, N. H., Feldman, L. B., Schreuder, R., Pastizzo, M., & Baayen, R. H. (2002). The processing and representation of Dutch and English compounds: Peripheral morphological, and central orthographic effects. *Brain and Language*, *81*(1), 555-567.

De Jong, N. H., Schreuder, R., & Baayen, R. H. (2000). The morphological family size effect and morphology. *Language and Cognitive Processes*, *15*(4-5), 329-365.

De Jong, K., & Zawaydeh, B. (2002). Comparing stress, lexical focus, and segmental focus: Patterns of variation in Arabic vowel duration. *Journal of Phonetics*, *30*(1), 53-75.

Delattre, P. (1965). *Comparing the phonetic features of English, French, German and Spanish: An interim Report*. Heidelberg: Julius Groos Verlag.

Dell, G. S. (1986). A spreading activation theory of retrieval in sentence production. *Psychological Review*, *93*(3), 283-321.

Dinnsen, D. A., & Charles-Luce, J. (1984). Phonological neutralization, phonetic implementation and individual differences. *Journal of Phonetics*, *12*(1), 49-60.

Dinnsen, D. A. (1985). A re-examination of phonological neutralization. *Journal of Linguistics*, *21*, 265-279.

Dmitrieva, O., Jongman, A., & Sereno, J. (2010). Phonological neutralization by native and non-native speakers: The case of Russian final devoicing. *Journal of Phonetics*, *38*(3), 483-492.

Ernestus, M., & Baayen, R. H. (2003). Predicting the unpredictable: Interpreting neutralized segments in Dutch. *Language*, 79, 5-38.

Ernestus, M., & Baayen, R. H. (2006). The functionality of incomplete neutralization in Dutch: The case of past-tense formation. In L. M. Goldstein, D. H. Whalen, & C. T. Best (Eds.), *Laboratory Phonology*, 8 (pp. 27-49). Berlin: Mouton de Gruyter.

Ernestus, M., & Baayen, R. H. (2007). Intraparadigmatic effects on the perception of voice. In E. J. van der Torre & J. van de Weijer (Eds.), *Voicing in Dutch: (De)voicing - phonology, phonetics, and psycholinguistics* (pp. 153-173). Amsterdam: John Benjamins.

Fischer-Jørgensen, E. (1976). Some data on North German stops and affricates. *Annual Report of the Institute of Phonetics, University of Copenhagen*, *10*, 149-200.

Fourakis, M., & Iverson, G. K. (1984). On the 'incomplete neutralization' of German final obstruents. *Phonetica*, *41*(3), 140-149.

Fox, R., & Terbeek, D. (1977). Dental flaps, vowel duration and rule ordering in American English. *Journal of Phonetics*, *5*, 27-34.

Frick, R. W. (1995). Accepting the null hypothesis. *Memory & Cognition*, *23*(1), 132-138.

Goldinger, S. D. (1996). Words and voices: Episodic traces in spoken word identification and recognition memory. *Journal of Experimental Psychology: Learning*, *Memory*, *and Cognition*, *22*, 1166-1183.

Goldinger, S. D. (1998). Echoes of echoes? An episodic theory of lexical access. *Psychological review*, *105*(2), 251-279.

Goldinger, S. D., Pisoni, D. B., & Logan, J. S. (1991). On the nature of talker variability effects on recall of spoken word lists. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *17*(1), 152-162.

Goldrick, M., Baker, H. R., Murphy, A., & Baese-Berk, M. (2011). Interaction and representational integration: Evidence from speech errors. *Cognition*, *121*(1), 58-72.

Goldrick, M., & Blumstein, S. E. (2006). Cascading activation from phonological planning to articulatory processes: Evidence from tongue twisters. *Language and Cognitive Processes*, *21*(6), 649-683.

Gósy, M., & Ringen, C. (2009). Everything you always wanted to know about voicing in Hungarian. Paper presented at the International Conference on the Structure of Hungarian 9, Debrecen, Hungary.

Green, D. M., & Swets, J. A. (1966). *Signal detection theory and psychophysics*. New York: Wiley.

Han, J.-I. (1994). Perceptual cues for Korean tense and lax consonants. *Working Papers of the Cornell Phonetics Laboratory*, *9*, 115-135.

Harrington, J. (2006). An acoustic analysis of 'happy-tensing' in the Queen's Christmas broadcasts. *Journal of Phonetics*, *34*(4), 439-457.

Hay, J., Drager, K., & Warren, P. (2009). Careful who you talk to: An effect of experimenter identity on the production of the NEAR/SQUARE merger in New Zealand English. *Australian Journal of Linguistics*, *29*(2), 269-285.

Heldner, M., & Strangert, E. (2001). Temporal effects of focus in Swedish. *Journal of Phonetics*, *29*(3), 329-361.

House, A. S., & Fairbanks, G. (1953). The influence of consonant environment upon the secondary acoustical characteristics of vowels. *Journal of the Acoustical Society of America*, *25*(1), 105-113.

Huff, C. T. (1980). Voicing and flap neutralization in New York City English. *Research in Phonetics*, *1*, 233-256.

Ingram, J. C., & Park, S. G. (1997). Cross-language vowel perception and production by

Japanese and Korean learners of English. *Journal of Phonetics*, *25*(3), 343-370.

Iverson, G. K., & Salmons, J. C. (1995). Aspiration and laryngeal representation in Germanic. *Phonology*, *12*, 369-396.

Jaeger, T. F. (2008). Categorical data analysis: Away from ANOVAs (transformation or not) and towards logit mixed models. *Journal of Memory and Language*, *59*(4), 434-446.

Janson, T., & Schulman, R. (1983). Non-distinctive features and their use. *Journal of linguistics*, *19*(2), 321-336.

Jassem, W., & Richter, L. (1989). Neutralization of voicing in Polish obstruents. *Journal of Phonetics*, *17*(4), 317-326.

Jessen, M. (1996). The relevance of phonetic reality for underlying phonological representation: the case of tense versus lax obstruents in German. In U. Kleinhenz (Ed.), *Interfaces in Phonology* (pp. 294-328). Berlin: Akademie-Verlag.

Jessen, M. (1998). *Phonetics and phonology of tense and lax obstruents in German*. Amsterdam: Benjamins.

Jessen, M. (2001). Phonetic implementation of the distinctive auditory features [voice] and [tense] in stop consonants. In T. A. Hall (Ed.), *Distinctive Feature Theory* (pp. 237-294). Berlin & New York: Mouton de Gruyter.

Jessen, M., & Ringen, C. (2002). Laryngeal features in German. *Phonology*, *19*(02), 189-218.

Johnson, K. (1997). Speech perception without speaker normalization. In K. Johnson & J. Mullennix (Eds.), *Talker variability in speech processing* (pp. 145-166). San Diego: Academic Press.

Joos, M. (1942). A phonological dilemma in Canadian English. *Language*, 141-144.

Jun, J. (2007). Variation in Korean Stem-final Obstruents. Ms, Seoul National University.

Jun, J., & Lee, J. (2007). Multiple stem-final variants in Korean native nouns and loanwords. *Journal of the Linguistic Society of Korea*, *47*, 159-187.

Jun, S. A. (1994). The status of the lenis stop voicing rule in Korean. *Theoretical Issues in Korean Linguistics*, 101-114.

Kang, E., Lee, H-Y., & Kim, J. (2004). The phonetic realization of syllable codas in Korean, *Malsori*, 49, 1-30.

Kharlamov, V. (2012). *Incomplete neutralization and task effects in experimentally-elicited speech: Evidence from the production and perception of word-final devoicing in Russian.* Ph.D. dissertation, University of Ottawa.

Kharlamov, V. (2014). Incomplete neutralization of the voicing contrast in word-final obstruents in Russian: Phonological, lexical, and methodological influences. *Journal of Phonetics, 43*, 47-56.

Kharlamov, V. (2015). Perception of incompletely neutralized voicing cues in word-final obstruents: The role of differences in production context. *Laboratory Phonology*, *6*(2), 147-165.

Kim, C.-W. (1965). On the autonomy of the tensity feature in stop classification (with special reference to Korean stops). *Word*, *21*(3), 339-359.

Kim, C.-W. (1979). Neutralization in Korean revisited. *Studies in the Linguistic Sciences*, *9*, 147-155.

Kim, D. W. (2002). The vowel length as a function of the articulatory force of the following consonants in Korean. *Speech Sciences*, *9*(3), 143-153.

Kim, H., Honda, K., & Maeda, S. (2005). Stroboscopic-cine MRI study of the phasing between the tongue and the larynx in the Korean three-way phonation contrast. *Journal of*

Phonetics, 33(1), 1-26.

Kim, H., & Jongman, A. (1996). Acoustic and perceptual evidence for complete neutralization of manner of articulation in Korean. *Journal of Phonetics*, *24*(3), 295-312.

Kim, M.-R. C. (1994). *Acoustic characteristics of Korean stops and perception of English stop consonants*. Ph.D. dissertation, University of Wisconsin.

Kim, M.-R. C., Lee, J., & Lotto, A. J. (2003). An investigation of the manner contrast in inter-vocalic Korean stops. *The Korean Language in America*, *8*, 75-84.

Kim, M.-R. C., & Lotto, A. J. (2002). An investigation of acoustic characteristics of Korean stops produced by non-heritage learners. *The Korean Language in America*, *7*, 177-188.

Kim-Renaud, Y-K. (1974). *Korean Consonantal Phonology*. Ph.D. dissertation, University of Hawaii.

Kim, T. (2014). A phonetic study on two successive stop consonants with the same place in Korean. *Journal of Language Sciences*, *21*(3), 45-64.

Klatt, D. H. (1976). Linguistic uses of segmental duration in English: Acoustic and perceptual evidence. *Journal of the Acoustical Society of America*, *59*(5), 1208-1221.

Kleber, F., John, T., & Harrington, J. (2010). The implications for speech perception of incomplete neutralization of final devoicing in German. *Journal of Phonetics, 38*(2), 185-196.

Kopkalli, H. (1993). *A phonetic and phonological analysis of final devoicing in Turkish*. Ph.D. dissertation, University of Michigan.

Lahiri, A., Schriefers, H., & Kuijpers, C. (1987). Contextual neutralization of vowel length: Evidence from Dutch. *Phonetica*, *44*(2), 91-102.

Levelt, W. J. M. (1989). *Speaking: From Intention to Articulation*. Cambridge, MA: The MIT

Press.

Levelt, W. J. M., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences*, *22*(01), 1-75.

Lieberman, P. (1963). Some effects of semantic and grammatical context on the production and perception of speech. *Language and Speech*, *6*(3), 172-187.

Lindblom, B. (1990). Explaining phonetic variation: a sketch of the H&H theory. In W. J.

Hardcastle & A. Marchal (Eds.), *Speech production and speech modelling* (pp. 403-439). Dordrecht, The Netherlands: Kluwer Academic Publisher.

Macmillan, N. A. (2002). Signal detection theory. In H. Pashler (Series Ed.) & J. Wixted (Vol. Ed.), *Stevens' handbook of experimental psychology, Vol. 4: Methodology in experimental psychology* (pp. 43-90). New York: Wiley.

MacMillan, N. A., & Creelman, C. (1991). *Detection theory: A user's guide*. New York: Cambridge University Press.

Magen, H. S., & Blumstein, S. E. (1993). Effects of speaking rate on the vowel length distinction in Korean, *Journal of Phonetics*, *21*, 387-409.

Marslen-Wilson, W. D. (1987). Functional parallelism in spoken word-recognition. *Cognition*, *25*(1), 71-102.

McClelland, J., & Elman, J. (1986). The TRACE model of speech perception. *Cognitive Psychology*, *18*(1), 1-86.

Mitleb, F. M. (1982). *Segmental and non-segmental structure in phonetics: evidence from foreign accent*. Ph.D. dissertation, Indiana University.

Mulder, K., Schreuder, R., & Dijkstra, T. (2013). Morphological Family Size effects in L1 and

L2 processing: An electrophysiological study. *Language and Cognitive Processes*, *28*(7), 1004-1035.

Mullennix, J. W., Pisoni, D. B., & Martin, C. S. (1989). Some effects of talker variability on spoken word recognition. *Journal of the Acoustical Society of America*, *85*(1), 365-378.

Nadig, A., & Shaw, H. (2015). Acoustic marking of prominence: how do preadolescent speakers with and without high-functioning autism mark contrast in an interactive task? *Language, Cognition and Neuroscience*, *30*(1-2), 32-47.

Norris, D. (1994). Shortlist: a connectionist model of continuous speech recognition. *Cognition*, *52*(3), 189-234.

O'Dell, M., & Port, R. (1983). Discrimination of word-final voicing in German. *Journal of the Acoustical Society of America*, *73*(S1), S31-S31.

Oh, J., & Shin, J. (2007). A study on sound changes in noun-final consonants and collision of the phonetic Realization. *Korean Linguistics*, 34, 209-232.

Park, J.-W. (1994). Variation of vowel length in Korean. In Y. Kim-Renaud (Ed.), *Theoretical issues in Korean linguistics*, 175-187.

Park, H. (1999). The phonetic nature of the phonological contrast between the lenis and fortis fricatives in Korean. In *Proceedings of the 14th International Congress of Phonetic Sciences*, San Francisco, pp. 424-427.

Park, H.-J., Shin, H.-Y., & Yang, B. (2002). A durational study of vowels followed by voiced or voiceless consonants. *Speech Sciences*, *9*(4), 175-185.

Park, I. K. (2013). The influence of the manner and place of articulation on vowel duration. *Studies in Linguistics (언어학연구)*, *28*, 81-96.

Peterson, G. E., & Lehiste, I. (1960). Duration of syllable nuclei in English. *Journal of the Acoustical Society of America*, *32*(6), 693-703.

Pierrehumbert, J. B. (2001). Exemplar dynamics: Word frequency, lenition, and contrast. In J. Bybee & P. Hopper (Eds.), *Frequency effects and the emergence of lexical structure* (pp. 137-157). Amsterdam: John Benjamins.

Pierrehumbert, J. B. (2002). Word-specific phonetics. In C. Gussenhoven & N. Warner (Eds.), *Laboratory Phonology*, *Vol. VII* (pp. 101-139). Berlin: Mouton de Gruyter.

Pinheiro, J. C., & D. M. Bates. (2000). *Mixed-effects models in S and S-PLUS*. New York: Springer.

Piroth, H. G., & Janker, P. M. (2004). Speaker-dependent differences in voicing and devoicing of German obstruents. *Journal of Phonetics*, *32*(1), 81-109.

Pisoni, D. B. (1997). Some thoughts on normalization in speech perception. In K. Johnson & J. W. Mullennix (Eds.), *Talker variability in speech processing* (pp. 9-32). San Diego: Academic Press.

Port, R. (1976). *The influence of speaking tempo on the duration of stressed vowel and medial stop in English trochee words*. Ph.D. dissertation, University of Connecticut.

Port, R., & Crawford, P. (1989). Incomplete neutralization and pragmatics in German. *Journal of Phonetics*, *17*(4), 257-282.

Port, R., Mitleb, F., & O'Dell, M. (1981). Neutralization of obstruent voicing in German is incomplete. *Journal of the Acoustical Society of America*, *70*(S1), S13-S13.

Port, R., & O'Dell, M. (1985). Neutralization of syllable-final voicing in German. *Journal of Phonetics*, *13*(4), 455-471.

Rabbitt, P. (1966). Recognition: Memory for words correctly heard in noise. *Psychonomic*

Science, *6*(8), 383-384.

Rapp, B., & Goldrick, M. (2000). Discreteness and interactivity in spoken word production. *Psychological Review*, *107*(3), 460-499.

Repp, B. H. (1984). Closure duration and release burst amplitude cues to stop consonant manner and place of articulation. *Language and Speech*, *27*(3), 245-254.

Roettger, T. B., Winter, B., Grawunder, S., Kirby, J., & Grice, M. (2014). Assessing incomplete neutralization of final devoicing in German. *Journal of Phonetics*, *43*, 11-25.

Schreuder, R., & Baayen, R. H. (1997). How complex simplex words can be. *Journal of Memory and Language*, *37*(1), 118-139.

Seidenberg, M. S., & Tanenhaus, M. K. (1979). Orthographic effects on rhyme monitoring. *Journal of Experimental Psychology: Human Learning* & *Memory,* 5(6), 546-554.

Shrager, M. (2012). Neutralization of word-final voicing in Russian. *Journal of Slavic Linguistics*, *20*(1), 71-99.

Sin, C. Y. (1997). *Consonantal production and coarticulation in Korean.* Ph.D. dissertation, University of London.

Silva, D. J. (1992). *The phonetics and phonology of stop lenition in Korean.* Ph.D. dissertation, Cornell University.

Sin, C.-Y., Shin, J., Kiaer, J., & Cha, J. (2012). *The sounds of Korean*. Cambridge: Cambridge University Press.

Slowiaczek, L. M., & Dinnsen, D. A. (1985). On the neutralizing status of Polish word-final devoicing. *Journal of Phonetics*, *13*(3), 325-341.

Smith, B. L., Hayes-Harb, R., Bruss, M., & Harker, A. (2009). Production and perception of voicing and devoicing in similar German and English word pairs by native speakers of

German. *Journal of Phonetics*, *37*(3), 257-275.

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

Stemberger, J. P. (1985b). An interactive activation model of language production. In A. Ellis (Ed.), *Progress in the Psychology of Language* (Vol. 1, pp. 143-186). London: Erlbaum.

Stemberger, J. P., & MacWhinney, B. (1986). Frequency and the lexical storage of regularly inflected forms. *Memory & Cognition*, *14*(1), 17-26.

Stone, G. O., Vanhoy, M., & Van Orden, G. C. (1997). Perception is a two-way street: Feedforward and feedback phonology in visual word recognition. *Journal of Memory & Language*, *36* (3), 337-359.

Uchanski, R. M. (1988). Spectral and temporal contributions to speech clarity for hearing impaired listeners. Ph.D. dissertation, Massachusetts Institute of Technology.

Uchanski, R. M. (1992). Segment durations in conversational and clear speech. *Unpublished manuscript*.

Warner, N., Good, E., Jongman, A., & Sereno, J. (2006). Orthographic vs. morphological incomplete neutralization effects. *Journal of Phonetics*, *34*(2), 285-293.

Warner, N., Jongman, A., Sereno, J., & Kemps, R. (2004). Incomplete neutralization and other sub-phonemic durational differences in production and perception: evidence from Dutch. *Journal of Phonetics*, *32*(2), 251-276.

Watson, D. G. (2010). The many roads to prominence: Understanding emphasis in conversation. In B. Ross (Ed.), *Psychology of learning and motivation* (Vol. 52, pp. 163-183). San Diego: Academic Press.

Winter, B., & Röttger, T. (2011). The nature of incomplete neutralization in German: Implications for laboratory phonology. *Grazer Linguistische Studien*, *76*, 55-74.

Yoon, K. (1998). *An acoustic analysis of Korean aspirated and tense alveolar fricatives using Multi-Speech*. MA thesis, University of Kansas.

Yu, A. C. L. (2007). Understanding near mergers: The case of morphological tone in Cantonese. *Phonology*, *24*(01), 187-214.

Ziegler, J. C., & Ferrand, L. (1998). Orthography shapes the perception of speech: The consistency effect in auditory word recognition. *Psychonomic Bulletin & Review*, *5*(4), 683-689.

Ziegler, J. C., Petrova, A., & Ferrand, L. (2008). Feedback consistency effects in visual and auditory word recognition: Where do we stand after more than a decade? *Journal of Experimental Psychology: Learning Memory & Cognition*, *34*(3), 643-661.

APPENDICES

Appendix A. Stimuli for the production experiment

Item	Class	Contrast	UR → SR	KO Ortho-graphy	Meaning	Underlying form	Following suffix: N	Following suffix: NN
Targets	Laryngeal (12)	Lenis - Nonlenis	/p-pʰ/ → [p]	입 - 잎	mouth (n) - leaf (n)	/ip - ipʰ/	to	ɨl
				갑 - 갚	group (n) - pay back (v)	/kap - kapʰ/	to	
				집 - 짚	house (n) - straw (n)	/tʃip - tʃipʰ/	to	
				업 - 엎	carry on the back (v) - overturn (v)	/up - upʰ/	tən	
			/t-tʰ/ → [t]	걷 - 겉	roll (v) - exterior (n)	/kət - kətʰ/	tən - to	
				믿 - 밑	believe (v) - bottom (n)	/mit - mitʰ/	tən - to	
				받 - 밭	receive (v) - field (n)	/pat - patʰ/	tən - to	
			/k-k'/ → [k]	묵 - 묶	stay (v) - tie (v)	/muk - muk'/	ko	
				낙 - 낚	joy (n) - catch (v)	/nak - nak'/	kwa - ko	
				박 - 밖	gourd (n) - outside (n)	/pak - pak'/	kwa	
				복 - 볶	fortune (n) - fry (v)	/pok - pok'/	kwa - ko	
				역 - 엮	station (n) - weave (v)	/yək - yək'/	kwa - ko	
	Manner (14)	Fricative-Stop (4)	/s-t/ → [t]	굿 - 굳	exorcism (n) - harden (v)	/kus - kut/	to - tən	
			/s-tʰ/ → [t]	맛 - 맡	taste (n) - smell (v)	/mas - matʰ/	to - tən	
				붓 - 붙	brush (n) - pass (v)	/pus - putʰ/	to - tən	
				솟 - 솥	soar (v) - cooker (n)	/sos - sotʰ/	tən - to	
		Affricate-Stop (4)	/tʃ-tʰ/ → [t]	맞 - 맡	be hit (v) - smell (v)	/matʃ - matʰ/	tən	
			/tʃʰ-t/ → [t]	닻 - 닫	anchor(n) - close (v)	/tatʃʰ - tat/	tən - to	
				돛 - 돋	sail (n) - break out (v)	/totʃʰ - tot/	tən - to	
			/tʃʰ-tʰ/ → [t]	숯 - 숱	charcoal (n) - thickness (n)	/sutʃʰ - sutʰ/	to	
		Fricative-Affricate (6)	/s-tʃ/ → [t]	맛 - 맞	taste (n) - be hit (v)	/mas - matʃ/	to - tən	
				멋 - 멎	style (n) - stop (v)	/məs - mətʃ/	to - tən	
				낫 - 낮	sickle (n) - day (n)	/nas - natʃ/	to	
				빗 - 빚	comb (n) - debt (n)	/pis - pitʃ/	to	
			/s-tʃʰ/ → [t]	낫 - 낯	sickle (n) - face (n)	/nas - natʃʰ/	to	
				빗 - 빛	comb (n) - light (n)	/pis - pitʃʰ/	to	
	Palatal (2)	Lenis-Nonlenis	/tʃ-tʃʰ/ → [t]	낮 - 낯	day (n) - face (n)	/natʃ - natʃʰ/	to	
				빚 - 빛	debt (n) - light (n)	/pitʃ - pitʃʰ/	to	
Fillers				간 - 감	liver (n) - wash (v)	/kan - kam/	kwa - ko	ɨl
				돈 - 돌	money (n) - stone (n)	/ton - tol/	kwa	
				목 - 몸	neck (n) - body (n)	/mok - mom/	to	
				문 - 물	door (n) - water (n)	/mun - mul/	to	
				발 - 밤	foot (n) - night (n)	/pal - pam/	kwa	
				술 - 숨	alcohol (n) - hide (v)	/sul - sum/	to - tən	
				참 - 찾	endure (v) - find (v)	/tʃʰam - tʃʰatʃ/	tən	

Appendix B. Examples of the trials for the Q & A task

1	NN*	Q: This is a verb meaning “be given or presented with something from others” and the opposite of the word ‘give’. What is the opposite expression of “the person who will give a present”? Please say “the person who will do this”. A: The person who will receive.
	N	Q: This is a verb that means “be given or presented with something from others” and is the opposite of the word ‘give’. What is the opposite expression of “people who used to give some help”? Please say “people who used to do this”. A: People who used to receive.
2	NN	Q: This is a verb meaning “to move something such as an open door back to the original position so that things cannot pass through an opening” and is the opposite of the word ‘open’. What is the opposite expression of “(Somebody) will open the window”? Please say “(Somebody) will do this”. A: (Somebody) will close.
	N	Q: This is a verb meaning “to move something such as an open door back to the original position so that things cannot pass through an opening” and is the opposite of the word ‘open’. What is the opposite expression of “people who used to open the window”? Please say “people who used to do this”. A: People who used to close.
3	NN	Q: This is a device used for making hair neat. What do people use when they manage tangled hair? Please say “(People) use this”. A: (People) use a comb.
	N	Q: This is a device used for making hair neat. What does also vary if hair style changes? Please say “This also varies”. A: A comb also varies.
4	NN	Q: This is a place where a bus or a train stops. Which do you say you passed this while you were dozing off in a train? Please say “(I) passed this”. A: (I) passed a station.

	N	Q: This is a place where a bus or a train stops. What is the opposite expression of the following? "(It is) near this". A: (It is) far from a station.
5	NN	Q: This is a word meaning glittering radiance. It is said there is no lay of this in the darkness. What is lack of? Please say "there is no (lay of) this". A: There is no (lay of) light.
	N	Q: This is a word meaning glittering radiance. What does a firefly give out from its body? Please say "(It) gives out this" A: (It) gives out light.

* NN: non-neutralizing condition where targets are followed by a vowel
  N: neutralizing condition where targets are followed by a consonant

Appendix C. Individual talker and item means of differences in voicing into closure duration (Laryngeal neutralization, Q & A task)

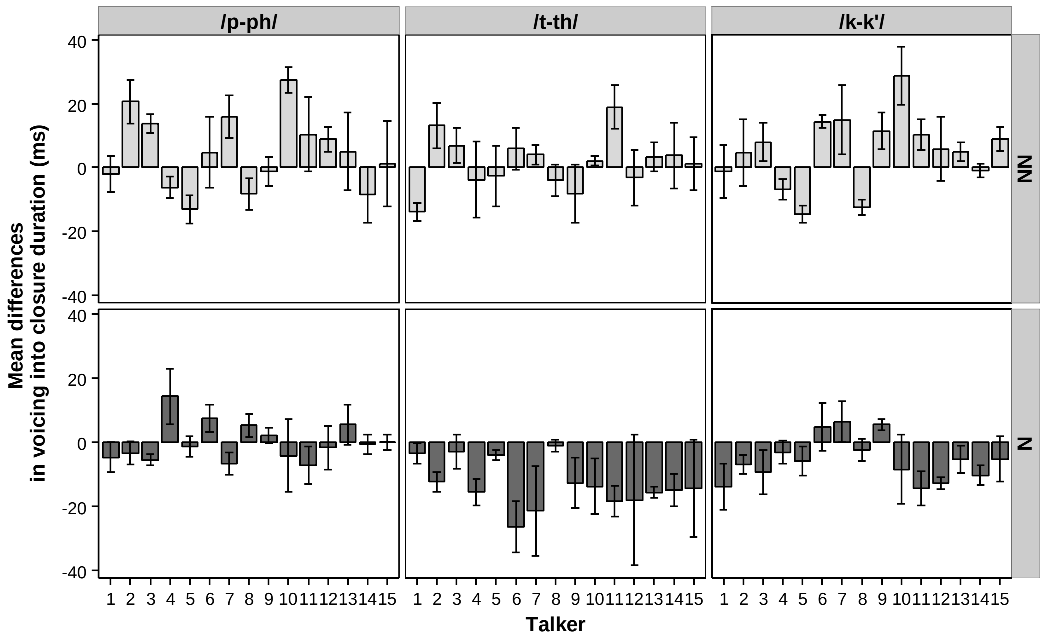

Figure 24. Mean differences in voicing into closure duration in pairs subject to laryngeal neutralization per talker (Q & A task; NN = non-neutralizing; N = neutralizing)

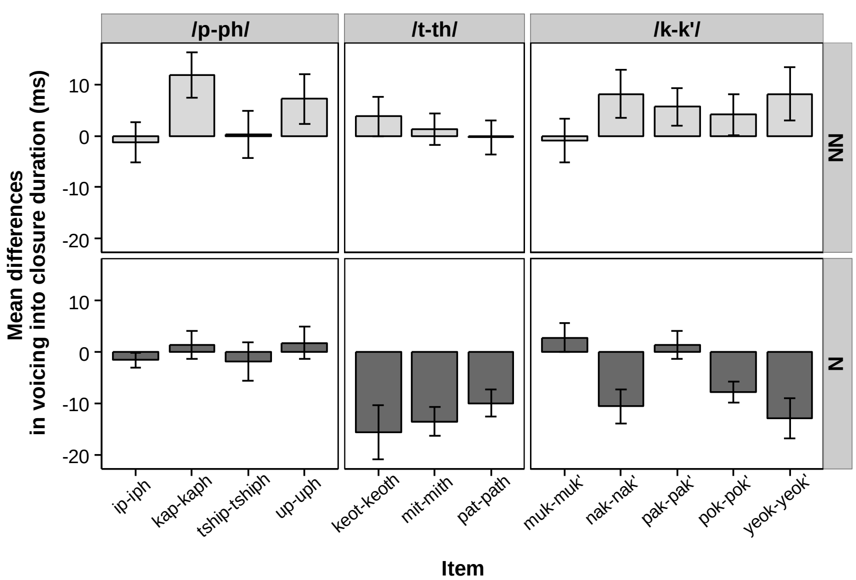

Figure 25. Mean differences in voicing into closure duration in pairs subject to laryngeal neutralization per item (Q & A task; NN = non-neutralizing; N = neutralizing)

Appendix D. Comparison of the original production stimuli with the subset stimuli for the perception experiment

D1. Laryngeal neutralization

Original	Vowel duration			Voicing into closure duration			Closure duration		
	Lenis	Non-lenis	Difference	Lenis	Non-lenis	Difference	Lenis	Non-lenis	Difference
Bilabial	47.9 (2.0)	43.0 (1.9)	4.9 (1.1)	49.9 (2.4)	50.0 (2.7)	0.1 (1.4)	140.4 (4.0)	144.2 (4.4)	-3.8 (3.6)
Alveolar	50.3 (1.5)	48.9 (1.6)	1.4 (1.0)	40.3 (2.8)	53.4 (2.1)	-13.1 (2.2)	164.8 (3.7)	172.7 (4.0)	-7.9 (2.4)
Velar	64.7 (2.1)	56.4 (2.0)	8.3 (1.0)	41.7 (1.9)	49.9 (3.3)	-8.2 (3.0)	180.3 (3.3)	180.7 (3.1)	-0.4 (2.7)
Mean (SE)	55.5 (1.3)	50.0 (1.2)	5.5 (0.7)	48.4 (1.0)	42.0 (1.3)	-5.6 (1.0)	163.1 (2.5)	166.5 (2.5)	-3.4 (1.8)
Regression results	$\beta = 4.827$, SE = 0.829, $\chi^2(1) = 18.20$, $p < 0.001$			$\beta = -6.021$, SE = 1.439, $\chi^2(1) = 11.47$, $p < 0.001$			$\beta = 3.820$, SE = 3.187, $\chi^2(1) = 1.36$, $p = 0.24$		

Subset	Vowel duration			Voicing into closure duration			Closure duration		
	Lenis	Non-lenis	Difference	Lenis	Non-lenis	Difference	Lenis	Non-lenis	Difference
Bilabial	48.9 (2.1)	42.7 (1.9)	6.2 (1.5)	50.0 (2.6)	50.2 (2.8)	-0.1 (2.1)	139.6 (4.5)	144.9 (4.7)	-5.3 (4.4)
Alveolar	51.2 (1.6)	49.0 (1.7)	2.1 (1.1)	40.4 (3.0)	54.9 (2.3)	-14.5 (2.3)	166.5 (4.2)	171.7 (4.8)	-5.2 (3.1)
Velar	64.3 (2.2)	56.3 (2.0)	7.9 (1.2)	40.7 (2.0)	45.8 (2.3)	-5.1 (2.1)	179.4 (3.6)	181.0 (3.9)	-1.6 (3.6)
Mean (SE)	55.9 (1.3)	50.0 (1.2)	5.9 (0.8)	43.7 (1.4)	49.5 (1.5)	-5.8 (1.3)	162.9 (2.7)	166.6 (2.8)	-3.7 (2.2)
Regression results	$\beta = 5.916$, SE = 1.139, $\chi^2(1) = 15.37$, $p < 0.001$			$\beta = -5.792$, SE = 2.190, $\chi^2(1) = 5.54$, $p < 0.05$			$\beta = -3.713$, SE = 3.652, $\chi^2(1) = 0.99$, $p = 0.31$		

D2. Manner neutralization

Original	Vowel duration			Voicing into closure duration			Closure duration		
	Lenis	Non-lenis	Difference	Lenis	Non-lenis	Difference	Lenis	Non-lenis	Difference
Fricative /s/ vs. Aspirated Stop /tʰ/	52.5 (2.8)	48.0 (2.6)	4.5 (1.0)	45.3 (2.5)	44.5 (2.3)	0.7 (1.8)	181.4 (3.6)	186.9 (3.4)	-5.5 (2.5)
Fricative /s/ vs. Aspirated Affricate /tʃʰ/	59.1 (4.0)	57.7 (4.2)	1.4 (1.5)	44.7 (3.0)	46.5 (3.1)	-1.8 (2.4)	189.1 (4.6)	184.7 (4.4)	4.4 (3.1)
Stop /t/ vs. Aspirated Affricate /tʃʰ/	47.8 (2.0)	48.5 (2.3)	-0.6 (1.5)	44.1 (2.9)	47.6 (3.3)	-3.5 (2.5)	183.6 (4.2)	192.7 (5.7)	-9.1 (4.5)
Affricate /tʃ/ vs. Aspirated Stop /tʰ/	68.5 (3.1)	67.2 (2.4)	1.3 (1.8)	32.3 (5.4)	34.3 (3.7)	-2.0 (4.2)	175.4 (5.1)	187.9 (5.9)	-12.6 (3.7)
Mean (SE)	55.0 (1.7)	53.0 (1.7)	2.0 (0.7)	43.2 (1.6)	44.5 (1.5)	-1.3 (1.2)	183.1 (2.2)	187.9 (2.3)	-4.8 (1.8)
Regression results	β = 2.017, SE = 1.008, χ^2(1) = 3.28, p = 0.69			β = -1.280, SE = 1.309, χ^2(1) = 0.94, p = 0.32			β = -4.828, SE = 3.511, χ^2(1) = 1.69, p = 0.19		

Subset	Vowel duration			Voicing into closure duration			Closure duration		
	Lenis	Non-lenis	Difference	Lenis	Non-lenis	Difference	Lenis	Non-lenis	Difference
Fricative /s/ vs. Aspirated Stop /tʰ/	53.3 (2.9)	48.1 (2.7)	5.3 (1.2)	45.2 (2.8)	45.8 (2.6)	-0.6 (2.6)	181.4 (4.3)	188.6 (4.0)	-7.2 (3.3)
Fricative /s/ vs. Aspirated Affricate /tʃʰ/	58.9 (4.0)	59.9 (4.5)	-1.0 (1.6)	45.4 (3.4)	44.7 (3.1)	0.7 (2.9)	188.8 (4.9)	185.3 (4.4)	3.5 (4.4)
Stop /t/ vs. Aspirated Affricate /tʃʰ/	47.5 (2.2)	48.1 (2.3)	-0.6 (2.1)	45.3 (3.0)	49.7 (3.6)	-4.4 (2.8)	184.2 (5.2)	195.4 (6.1)	-11.1 (5.9)
Affricate /tʃ/ vs. Aspirated Stop /tʰ/	67.0 (2.7)	70.0 (3.4)	3.0 (2.5)	38.3 (4.6)	31.5 (5.4)	-6.8 (4.9)	190.4 (7.2)	178.4 (6.2)	-12.0 (6.5)
Mean (SE)	55.4 (1.7)	53.4 (1.8)	2.0 (0.9)	43.6 (1.7)	45.6 (1.6)	-2.0 (1.5)	183.6 (2.5)	189.7 (2.6)	-6.1 (2.4)
Regression results	β = -1.854, SE = 1.310, χ^2(1) = 1.78, p = 0.18			β = 1.793, SE = 1.782, χ^2(1) = 0.95, p = 0.32			β = 6.457, SE = 3.469, χ^2(1) = 3.10, p = 0.38		

D3. Palatal neutralization

Original	Vowel duration			Voicing into closure duration			Closure duration		
	Lenis	Non-lenis	Difference	Lenis	Non-lenis	Difference	Lenis	Non-lenis	Difference
Affricate /tʃ/ vs. Aspirated affricate /tʃʰ/	58.2 (3.8)	57.7 (4.2)	0.5 (1.6)	44.8 (3.2)	46.5 (3.1)	-1.7 (2.4)	188.2 (4.8)	184.7 (4.4)	3.5 (3.3)
Regression results	$\beta = 0.503$, $SE = 3.959$, $\chi^2(1) = 0.01$, $p = 0.89$			$\beta = -1.708$, $SE = 3.124$, $\chi^2(1) = 0.29$, $p = 0.58$			$\beta = 3.447$, $SE = 3.630$, $\chi^2(1) = 0.89$, $p = 0.34$		

Subset	Vowel duration			Voicing into closure duration			Closure duration		
	Lenis	Non-lenis	Difference	Lenis	Non-lenis	Difference	Lenis	Non-lenis	Difference
Affricate /tʃ/ vs. Aspirated affricate /tʃʰ/	58.2 (4.1)	59.8 (4.5)	1.6 (1.9)	44.5 (3.5)	44.7 (3.1)	-0.2 (2.6)	188.8 (5.1)	185.3 (4.4)	3.5 (4.8)
Regression results	$\beta = -1.630$, $SE = 6.047$, $\chi^2(1) = 0.07$, $p = 0.78$			$\beta = -0.173$, $SE = 4.591$, $\chi^2(1) < 0.01$, $p = 0.96$			$\beta = 3.490$, $SE= 5.230$, $\chi^2(1) = 0.44$, $p = 0.50$		

Appendix E. Target pairs selected from the production stimuli for the perception experiment

Type of Neutralization (# of minimal pair)	Place/Manner	Target Pairs
Laryngeal (12)	Bilabial	/ip - ip^{h}/ /kap - kap^{h}/ /tʃip - tʃip^{h}/ /up - up^{h}/
	Alveolar	/kət - kət^{h}/ /mit - mit^{h}/ /pat - pat^{h}/
	Velar	/muk - muk'/ /nak - nak'/ /pak - pak'/ /pok - pok'/ /yək - yək'/
Manner (8)	Fricative - Aspirated Stop	/mas - mat^{h}/ /pus - put^{h}/ /sos - sot^{h}/
	Fricative - Aspirated Affricate	/nas - natʃ^{h}/ /pis - pitʃ^{h}/
	Stop - Aspirated Affricate	/tat - tatʃ^{h}/ /tot - totʃ^{h}/
	Affricate - Aspirated Stop	/matʃ - mat^{h}/
Palatal (2)	Affricate - Aspirated Affricate	/natʃ - natʃ^{h}/ /pitʃ - pitʃ^{h}/

Appendix F. Examples of suffixes attached to the base of the noun 'house' (/tʃɪp/) and the verb 'receive' (/pat/)

Stem	Suffix		Inflected form	Role / Meaning
/tʃɪp/ 'house'	V-initial	/e/	[tʃi.be]	(locative, directive, 'at', 'to')
		/i/	[tʃi.bi]	(nominative)
		/ina/	[tʃi.bina]	(disjunctive, 'or')
		/ɨj/	[tʃi.bɨj]	(genitive, 'of')
		/ɨl/	[tʃi.bɨl]	(accusative)
		/ɨn/	[tʃi.bɨn]	(topic)
		/ɨro/	[tʃi.bɨro]	(directive, 'to')
	C-initial	/kwa/	[tʃɪp.k'wa]	(comitative 'with', conjunctive 'and')
		/man/	[tʃɪp.man]	(delimiter, 'only')
		/mankɨm/	[tʃɪp.mankɨm]	(auxiliary, 'as much as')
		/poda/	[tʃɪp.p'oda]	(comparative, 'than')
		/tɨl/	[tʃɪp.t'ɨl]	(plural)
		/to/	[tʃɪp.t'o]	(delimiter, 'also')
		/tʃələm/	[tʃɪp.tʃ'ələm]	(delimiter, 'like')

Stem	Suffix		Inflected form	Role / Meaning
/pat/ 'receive'	V-initial	/a/	[pa.da]	(complementizer)
		/at/	[pa.dat]	(past tense)
		/ɨl/	[pa.dɨl]	(future tense)
		/ɨm/	[pa.dɨm]	(norminalizer)
		/ɨmyə/	[pa.dɨmyə]	(conjunctive)
		/ɨmyən/	[pa.dɨmyən]	(conditional)
		/ɨn/	[pa.dɨn]	(pre-nominal modifier)
	C-initial	/ket/	[pat.k'et]	(presumptive modal)
		/ki/	[pat.k'i]	(nominalizer)
		/ko/	[pat.k'o]	(conjunctive, 'and')
		/nɨn/	[pat.nɨn]	(present tense)
		/tən/	[pat.t'ən]	(imperfective, 'used to')
		/sɨp/	[pat.s'ɨp]	(addressee honorific)
		/tʃi/	[pat.tʃ'i]	(negative)