The production and perception of domain-initial strengthening in Seoul, Busan, and Ulsan Korean

Kayeon Yoo

King's College

This dissertation is submitted for the degree of
Doctor of Philosophy

University of Cambridge
February 2020

Declaration

This thesis is the result of my own work and includes nothing which is the outcome of work done in collaboration except as declared in the Preface and specified in the text. It is not substantially the same as any that I have submitted, or, is being concurrently submitted for a degree or diploma or other qualification at the University of Cambridge or any other University or similar institution except as declared in the Preface and specified in the text. I further state that no substantial part of my thesis has already been submitted, or, is being concurrently submitted for any such degree, diploma or other qualification at the University of Cambridge or any other University or similar institution except as declared in the Preface and specified in the text. It does not exceed the prescribed word limit for the relevant Degree Committee.

Abstract

The production and perception of domain-initial strengthening in Seoul, Busan, and Ulsan Korean

Kayeon Yoo

Korean exhibits one of the most consistent examples of the cross-linguistic phenomenon of domain-initial strengthening (hereafter DIS; T. Cho & Keating, 2001; Keating, Cho, Fougeron, & Hsu, 2004). DIS is defined as temporal and/or spatial enhancement of segmental articulation in the initial position of prosodic domains. Broadly, this dissertation serves as a detailed case study of the production patterns and the perceptual benefits of this phenomenon.

The recent findings of denasalisation and devoicing of the initial nasals in Korean (Young Shin Kim, 2011; Yoo, 2015a) suggest that there is a striking parallelism between the lenis stops /p, t, k/ and the nasal consonants /m, n/ in their patterns of DIS. Nevertheless, we currently lack an account that captures this parallelism. In addition, there is disagreement over the categorical nature of lenis stop voicing (S.-A. Jun, 1993; Docherty, 1995) and denasalisation (Yoshida, 2008; Young Shin Kim, 2011). Despite the obvious similarities between the arguably discrete processes of lenis stop voicing and denasalisation, and the kind of gradient effects widely reported for DIS, there has been no explicit investigation of the links among them. Thus, I examined the hypothesis that DIS, operating in the phonetic component, has given rise to the categorical rules of lenis stop voicing and denasalisation in the phrase-level phonology through *rule scattering,* as predicted by the theory of the life cycle of phonological processes (Bermúdez-Otero & Trousdale, 2012; Turton, 2014).

Recordings were collected in Seoul, Busan, and Ulsan, and various auditory and acoustic analyses were conducted to examine the phonetic variation of the relevant stops. The study adopted the three-city design as these varieties were expected to be at different stages in the life cycle, particularly with regard to the stabilisation of denasalisation. In the second part of this dissertation, I conducted a perception experiment to investigate if listeners are able to use DIS patterns as a cue to a prosodic boundary.

According to the results, Seoul showed the most advanced patterns in the stabilisation of DIS. As predicted by rule scattering, speakers who showed evidence of categorical lenis stop voicing and/or denasalisation also showed an overlaid effect of a gradient phonetic process. The perception study strongly supported the hypothesis that listeners exploit DIS cues to detect the beginning of a prosodic domain. Based on these findings, this dissertation offers a unified account of lenis stop voicing, denasalisation, and DIS within a single framework, offering insights into the nature of DIS as well as its functional role in prosodic parsing.

Acknowledgements

I would like to dedicate this section to a few individuals who have supported and brightened my journey through the PhD study. Francis Nolan, my supervisor, deserves the most special thanks from me. From my undergraduate years until thus far, his mentorship and guidance have been instrumental in my development as an independent researcher. Few supervisors could be more generous than Francis at offering support for their student's academic growth as well as their mental well-being. At times when I felt challenged, his trust and encouragement have kept me going.

I am also deeply grateful to my advisor Bert Vaux who, like Francis, has been there from the beginning of my life as a Linguistics student. He has helped me build strong foundations in Phonology and challenged me to think more broadly and critically. I am always inspired by his ability to maintain both academic rigour and playful curiosity towards the subject, and I would certainly like to emulate this attitude in my own career.

I also thank my labmates and fellow PhDs for not only offering direct help and feedback on my work but also for making my time more fun and fulfilling. Li Nguyen is the heroine in my PhD life, making me laugh with her wild ideas and silly jokes, supplying me with her wonderful baked goods and all kinds of surprise presents just when I needed them. It meant so much to know that I had her to share the high's and low's of my PhD. Yang Li, who used to give me statistics advice as a labmate, is now my manager at work. Having started full-time work before the viva and corrections, it has been challenging to manage both PhD and work. Yang's role as an excellent manager has been essential to making this possible. Chris Bryant also deserves a mention for kindly proofreading my draft chapters, offering advice on Python, helping with the logistics on the day of my submission, and much more.

My research involved fieldwork in three cities and this wouldn't have been possible were it not for the generosity of Prof. Taehong Cho (Hanyang), Dr. Joongrak Kwon (Tongmyong), Dr. Joogon Kim (Tongmyong), and Prof. Jieun Kim (Ulsan), who allowed me to use the facilities and equipment at their respective universities. I am also grateful to each and every one of my participants and informants, not just those whose data I used directly in this work but also those who participated in the early pilot experiment and whom I informally consulted on various occasions. Among these, I particularly thank Dr. Hyesop Shin, Dr.

Márton Sóskuthy, Dr. Timo Roettger, and Dr. Bodo Winter for offering help on R and the statistics used in this work. Furthermore, the first three years of my PhD were fully funded by the Cambridge International Trust, which I gratefully acknowledge. I also thank King's College and the MMLL Faculty for their generosity in funding my fieldwork, conference travels, and living expenses for the term after my third year.

My friends and family – they may not have a detailed idea of what my research is about, but their love has been crucial to the completion of my PhD. Yoonseo Woo, Yeonje Jo, Jihun Kim, Heeseo Kwon, Suhyeon Lee, Saebom Kang, Pilyoon Ryu, Jihyun Kim, and Eunhee Lim. These are just a few of my friends and family members who genuinely care about my success and well-being even though many of them live miles away from me. Their occasional care parcels full of Korean snacks, products, and face masks have helped me survive (quite literally) through my PhD and the past few months of COVID-19.

Lastly but most importantly, I am deeply indebted to my Mum for being my most ardent supporter for all these years. She taught me to be positive, brave, and to achieve what I want to achieve in life. I would not have been here if it weren't for her. My Dad has passed away in the first year of my undergraduate course, but what he has given me earlier in life has continued to be a source of encouragement, inspiration, and self-confidence. I wish the work here and the past four years of my efforts have made him proud.

Contents

Speech samples, raw data, R scripts, and a Python script used in the dissertation can be found in the accompanying OSF repository (https://osf.io/8kmqz).

List of Figures

List of Tables

List of Abbreviations

DIS	Domain-initial strengthening
LSD	Lenis stop devoicing
POT	Post obstruent tensing
VOT	Voice onset time
S	Syllable
PW	Phonological Word or Prosodic Word
AP	Accentual Phrase
ip	Intermediate Phrase
IP	Intonational Phrase
U	Utterance
Si	Syllable-initial
PWi	Phonological-Word-initial
APi	Accentual-Phrase-initial
APm	Accentual-Phrase-medial
IPi	Intonational-Phrase-initial
P	Postpausal
T	Tone
H	High
L	Low
R	Rising
[c.g.]	constricted glottis
[cont]	continuant
[s.g.]	spread glottis
[son]	sonorant

Chapter 1.
Introduction

The lenis stops /p, t, k/ and the nasal consonants /m, n/ of Korean show a striking parallelism in their patterns of phonetic variation with respect to prosodic structure. Currently, there is a lack of a unified account that captures this parallel. It is the primary goal of this study to characterise the exact phonetic patterns of these segments and provide a principled account of their prosody-sensitive segmental variation.

More broadly, the dissertation is intended to serve as a case study of a cross-linguistic phenomenon known as domain-initial strengthening (hereafter DIS). DIS is defined as temporal and/or spatial enhancement of segmental articulation in domain-initial position (Fougeron & Keating, 1997; Keating et al., 2004). As will be discussed in Section 1.2.1, the Korean case bears important implications for our understanding of the nature of DIS and the workings of the prosody-phonetics interface.

In this chapter, I will first establish the view of prosody adopted in this dissertation (Section 1.1). Then, I will outline two main theoretical issues revealed by the patterns of the lenis and nasal stop variation in Korean and propose an account that addresses these issues (Sections 1.2.1 & 1.2.2). The main production study in this dissertation will be designed to test the predictions of this account. I will also provide a brief justification for the experimental design which focuses on three varieties of Korean (Section 1.2.2) as well as the motivation for a perception study (Section 1.2.3). Then, the guiding questions and the hypotheses of the production and perception studies will be summarised (Section 1.3). Finally, the overall structure of the dissertation will be provided (Section 1.4).

1.1 The theoretical framework of prosody

In this dissertation, I work under the theoretical assumption that prosody is an abstract grammatical structure that arises through the prosodic component, a subsystem within the phonological component (Nespor & Vogel, 1986, p. 6; Shattuck-Hufnagel & Turk, 1996, Section 5.3; Beckman, 1996). There is abundant evidence that prosodic structure is not isomorphic with syntactic structure, although constrained by it (for a review, see Shattuck-Hufnagel & Turk, 1996, Section 2.2). Various extra-syntactic factors, including rate of speech and symmetry in terms of the length of subconstituents of an utterance (Shattuck-Hufnagel & Turk, 1996, p. 203), appear to influence the prosodic structure of an utterance. Thus, it is postulated that there is a component of grammar, separate from syntax, which integrates and weights these factors to generate prosodic structure.

If the phonetic component is defined as that part of grammar which is responsible for the gradient and non-contrastive aspect of speech production and perception, then prosody cannot simply be viewed as a phonetic phenomenon. It is widely accepted that prosodic structure is best understood in terms of prosodic constituents and prominences (Beckman & Edwards, 1994). In this view, the placement of constituent boundaries (e.g. an Intonational Phrase boundary) and prominences (e.g. a Nuclear Pitch Accented syllable) in an utterance is categorical in the sense that a certain boundary or phrasal prominence is either present or absent. It is also contrastive as it can alter the semantic and pragmatic interpretation of the utterance. How such discrete elements are reflected in the gradient acoustic-phonetic dimensions of F0, duration, amplitude, and segment quality is a highly complex matter, and is the subject of much research (e.g. Pierrehumbert, 1980; Pierrehumbert & Beckman, 1988; S.-A. Jun, 1993; Fougeron & Keating, 1997). It is in this context of continued work on the prosody-phonetics interface that the present study investigates the effect of a preceding prosodic boundary on segmental articulation.

1.2 Issues

1.2.1 Domain-initial strengthening or non-domain-initial weakening?

The Korean lenis stops /p, t, k/ and nasal consonants /m, n/ share unmistakable similarities in their patterns of realisation across different prosodic positions. Articulatorily, both the lenis and nasal stops have been reported to show greater peak linguopalatal contact and stop seal duration in the initial position of a higher prosodic domain in the prosodic hierarchy (T. Cho & Keating, 2001). That is, these segments are realised with progressively greater linguopalatal contact and longer seal duration in Phonological Word (PW)-initial, Accentual Phrase (AP)-initial, and Intonational Phrase (IP)-initial position. According to the Strict Layer Hypothesis (Selkirk, 1986), a segment that is in IP-initial position would simultaneously be in the initial position of all the smaller prosodic domains. For concision, I will refer to initial position of the highest relevant domain (e.g. the IP), and assume that this entails initial position in the lower domains.

T. Cho and Keating (2001) also found acoustic correlates of DIS. Voice Onset Time (VOT) was progressively longer for the lenis stops in higher domain-initial positions. The percentage of voicing during the acoustic closure of the lenis stops, which measures the voicing residue from the preceding vowel, was also progressively smaller in higher positions. For the nasals, the minimum nasal energy measured at the lowest point of the RMS acoustic energy profile was lower in higher prosodic positions. This was interpreted to mean that the velum position is higher for nasals in higher prosodic positions, consistent with the findings for French (Fougeron, 2001) and Estonian (Gordon, 1997, p. 13).

More recently, Young Shin Kim (2011) found strong acoustic and aerodynamic evidence that the Korean nasals may show complete denasalisation, not just nasality[1] weakening as argued in Yoshida (2008). Yoo (2015a, 2015b) observed that the phonetic realisation of the nasals varies on a continuum from a sonorant nasal [m, n] to a plosive with a short voicing lag [p‘, t‘][2]. The existence of plosion indicates a pressure build-up in the oral cavity which is only possible when the velum is raised. Young Shin Kim and Yoo agree that nasal consonants are generally nasal AP-medially, but denasalised AP-initially.

The parallelism between the pattern of the lenis stops and that of the nasal consonants is obvious. On the one hand, the lenis stops are voiced AP-medially in intervocalic positions by a well-established rule of *lenis stop voicing* in Korean (S.-A. Jun, 1993). As the prosodic position of initial lenis stops becomes higher in the prosodic hierarchy, they become voiceless and more aspirated. On the other hand, nasal consonants are nasal and voiced AP-medially, and become denasalised, devoiced and even slightly aspirated as their prosodic position becomes higher.

Thus, it goes against the principle of parsimony (i.e. Occam's Razor)[3] that the behaviours of the lenis stops and nasals are currently accounted for by different mechanisms.

---

[1] Since this dissertation focuses on articulatory strengthening, the term "nasality" is used in an articulatory-phonetic sense. Thus, unless otherwise specified, it refers to speech production with an open velopharyngeal port as betrayed by resonance characteristics associated with acoustic coupling between the oral and nasal vocal tracts, rather than the perception of nasality or phonologically contrastive nasality. Similarly, "denasalisation" is defined as an articulatory phenomenon whereby a nasal phoneme is realised with a raised velum. The key references discussed in this work also adopt this articulatory-phonetic definition of nasality, either explicitly or implicitly, e.g. Yoshida (2008, p. 1), Kim (2011, p. 31), Styler (2015, xiii).

[2] Reversed apostrophe indicates weak aspiration as in earlier IPA usage (Pullum & Ladusaw, 1996, p. 250).

[3] The principle that among competing hypotheses, the one with the fewest assumptions should be preferred. For further discussion and examples of the usefulness of Occam's Razor in scientific research, see Standish (2004) and Domingos (1999).

The lenis stops in Korean have been assumed to be underlyingly[4] voiceless, and are accordingly represented with the voiceless phoneme symbols /p, t, k/. If so, it follows that the lenis stops become “weakened” to their voiced forms in an AP-medial position, and “strengthened” to exhibit longer VOT in the initial position of domains higher than the AP. This explanation assumes the AP-initial realisation as the default. Alternatively, the highest prosodic domain in the prosodic hierarchy could be taken as the baseline, which can be the IP or Utterance (U), depending on the exact version of the prosodic hierarchy (Nespor & Vogel, 1986; Selkirk, 1986; Pierrehumbert & Beckman, 1988). In this latter view, the forms that appear in all lower prosodic positions would be subject to different degrees of articulatory weakening.

However, it is difficult to apply the same analysis to the case of the nasal consonants. One cannot argue that the denasalised form which appears in the initial position of the AP or IP (or U) is the base form of nasal consonants. In addition to the obvious reason that this analysis is counterintuitive, it is phonetically arbitrary that phonemes that are not underlyingly nasal should acquire nasality in the vicinity of non-nasal segments, e.g. the /n/ in /t͡sin.po/ [t͡sʰin.bo] ‘progress’.

One possible solution to this problem would be to argue that the Korean lenis stops are underlyingly voiced, i.e. /b, d, g/. Then, the parallelism between the phonetic patterns of the lenis stops and nasal consonants can be captured neatly by the tendency of these underlyingly voiced segments to become strengthened in higher prosodic positions. More precisely, the segments acquire more obstruent-like features of voicelessness and aspiration, and, in the case of the nasal consonants, a lack of nasality. This would be in line with the analysis that DIS can be explained as *syntagmatic contrast enhancement*, which makes initial consonants more dissimilar to the following vowel (T. Cho & Jun, 2000). Thus, in a way, DIS can be thought of as a sort of fortition in which the effect is gradient and roughly proportional to the height of the prosodic domain. Furthermore, there are independent arguments for adopting /b, d, g/ as the underlying forms of the Korean lenis stops (M.-R. Kim, 2012; M.-R. Kim & Duanmu, 2004). This prevents Korean from being a cross-linguistic exception in having three voiceless plosives and no voiced plosive.

---

[4] This expression is in reference to the widely used notions of underlying and surface representation in generative grammar (for introduction, see McCarthy, 2007, and Cole & Hualde, 2011). An underlying representation (UR) can be defined as a basic phonological form from which contextual variants (surface representations, or SR) are derived. An underlying form of a sound category should be chosen to provide “an optimal mapping to all the observed surface forms …, maximising the function of phonological rules in specifying predictable information and in expressing regularities in the distribution of sounds in the language overall” (Cole & Hualde, 2011, p. 16).

Not only is resolving this issue crucial to the accurate description of Korean phonology and phonetics, it is fundamental to the understanding of the mechanism underlying the wider phenomenon of DIS. An account of DIS based on articulatory undershoot (Lindblom, 1963; Moon & Lindblom, 1994) has gained some support since it was first discussed by Fougeron and Keating (1997, p. 3737). According to this account, DIS is a result of longer durations of segments in higher prosodic positions, which allows articulatory targets[5] to be achieved relatively fully. On the other hand, articulatory targets in lower positions are liable to a greater degree of undershoot. For Korean, in particular, T. Cho and Keating (2001) found a strong correlation between linguopalatal contact and articulatory and acoustic duration, claiming that "lengthening and strengthening is a single effect" in Korean.

This account of DIS is problematic for two reasons. First, if true, it means that the term DIS is a misnomer, and it should instead be called "non-domain-initial weakening". The process is a result of undershot articulatory targets, closely resembling the traditional notion of lenition. Second, and more importantly, this leads to an unreasonable claim that the range of phonetic realisation of nasal consonants in Korean is a result of different degrees of articulatory undershoot of what is originally [p‘, t‘]. It also cannot account for the direction of change of domain-initial denasalisation in Seoul Korean (Yoo, 2016). In an apparent-time study, Yoo found that younger speakers were more likely to realise domain-initial nasals with obstruent-like features such as voicelessness, aspiration, and a lack of nasality. If DIS is driven primarily by articulatory undershoot leading to lenition in lower prosodic positions, the direction of sound change should be towards more sonorant and lenited forms – mirroring the synchronic process of lenition (e.g. Hansson, 2008; Bermúdez-Otero, 2015) – rather than towards more obstruent-like forms.

Thus, the first goal of this dissertation is to develop a consistent account of the prosody-sensitive variation of domain-initial lenis stops and nasal consonants in Korean. More broadly, this dissertation aims to contribute to the understanding of the nature and the mechanism underlying DIS. A production study is designed to investigate the detailed acoustic characteristics of the lenis stops and nasal consonants in a range of prosodic positions. Based on the observed patterns, we will examine different hypotheses regarding the nature and mechanism of DIS (Chapter 5).

---

[5] An articulatory target here refers to an intended goal for articulatory execution. This reflects all categorical linguistic processes but not those gradient processes that result from universal physiological/aerodynamic constraints in speech production as the latter are not "intended" effects of a speaker's grammar.

1.2.2 Lenis stop voicing, denasalisation, and domain-initial strengthening

The inconsistency of the DIS accounts discussed above means that we currently lack a clear understanding of the following: (1) the nature of the relationship between the lenis stop voicing rule and DIS of the lenis stops, and (2) that between denasalisation and DIS of the nasals in Korean. In characterising the relationships (1 & 2), researchers have adopted one of the two following approches. The first approach is to treat the three processes – lenis stop voicing, denasalisation, and DIS – to be entirely separate from each other. However, it is theoretically uneconomical to account for them in separate analyses, when they clearly describe similar processes of segmental variation as a function of prosodic position. The second approach is to take the view that these processes are in fact the same. To discuss the lenis stops first, the lenis stop voicing rule and DIS of the lenis stops can be seen to describe the same phenomenon, in which the percentage of the voiceless portion of the lenis stops becomes greater in higher prosodic positions. This percentage can range from 0% to greater than 100%, to cover the extremes of fully voiced to aspirated realisations. Then, the lenis stop voicing rule is just an incomplete description of a larger phenomenon of DIS, failing to account for the fact that the effect is not limited to just voicing of the segments.

While explicit discussion of the links between these processes is rare in the literature, researchers have tended to assume one of the two possibilities, either treating them as being essentially the same or entirely independent. In their study of DIS of Korean stops, T. Cho and Keating (2001, Section 4.4.3) used part of their data on DIS to revisit the debated issue of the grammatical status of lenis stop voicing, without discussing the relationship between the two processes. Yoshida (2008) stated that denasalisation "seems to be one instance of" DIS. On the other hand, Young Shin Kim (2011) expressed reservation about calling it a DIS process (p. 137), based on the lack of a cumulative link between the degree of denasalisation and the prosodic position of the nasal.

A similar, one-or-the-other assumption has been made in an investigation of the post obstruent tensing rule and DIS of fricatives in Korean (S. Kim, 2001, p. 6). The post obstruent tensing rule describes the process whereby the onset lenis consonants (/t, p, k, s, tʃ/) become tense when they follow an obstruent coda, e.g. /ak.sa/ → [ak.s*a] 'musician'. S.-A. Jun (1998) claimed that the domain of this rule is the AP. Relating this to DIS, S. Kim proposed two alternative hypotheses. One hypothesis was that the categoricity of the post obstruent tensing rule is a mirage, which arises due to categorical perception of a gradient pattern of DIS. That is, although DIS affects segmental articulation gradiently, listeners may perceive a categorically different sound at a certain prosodic domain. The alternative hypothesis was that post obstruent tensing is indeed a categorical process which operates independently of DIS.

However, there is a third possibility that has not been considered in these previous studies; a possibility that the rules are the "children" of DIS. In other words, the rules originate from DIS and are thus closely related, but nevertheless independent. If this is true, it would not be accurate to treat them as the same or completely different. The literature suggests a crucial distinction between DIS and the lenis stop voicing rule in terms of their grammatical status. That is, DIS is gradient, whereas the voicing rule is categorical with the AP as its domain of application[6]. This suggests that the two processes may be instances of a "scattered rule" (Bermúdez-Otero, 2010, 2015; Robinson, 1976). The notion of *rule scattering* has been proposed within the theory of the life cycle of phonological processes, under the assumption that grammar has a modular, feed-forward architecture. According to this assumption, grammar is comprised of autonomous submodules, i.e. phonetics, phonology, syntax, etc., with structures and principles of their own. These submodules of grammar together account for all language-specific aspects of a human language, categorical or gradient, leaving aside the mechanical, language-independent aspects of speech. Language processing is argued to take place in a sequence, with information flowing from one module to the next in one direction without skipping any intervening modules (see Pierrehumbert, 2002, for an overview of the modular feedforward architecture). A diagrammatic representation of the life cycle of phonological processes (Bermúdez-Otero & Trousdale, 2012) is reproduced in Figure 1.1.

The theory of the life cycle of phonological processes posits that there is a specific unidirectional pathway along which sound patterns evolve over time:

> "In the course of this life cycle, a phonetic phenomenon that is at first exhaustively determined by extragrammatical factors (physics and physiology) becomes ever more deeply embedded in the grammar of a language, first as a language-specific gradient process of phonetic implementation, later as a categorical phonological rule applying in increasingly narrow morphosyntactic domains, until it eventually escapes phonological control altogether." (Bermúdez-Otero, 2015)

Within such a theoretic framework, rule scattering occurs when a process operating at one component of the grammar enters a higher-level component without ceasing to apply in the original component. An oft-cited example is palatalisation of /s/ in English (Zsiga, 1995). /s/ is palatalised before /j/ in sequences such as *miss you*, via a gradient postlexical process of gestural overlap. Simultaneously, a categorical rule of palatalisation is found in the stem-level derivation of words such as *succession* from *success.* More recently, Turton

[6] Note that the categorical nature of the lenis stop voicing rule is debated (Docherty, 1995; S.-A. Jun, 1995c). See Section 3.4.2 for further discussion.

(2014, 2017) conducted an extensive study of the variation observed in the production of English /l/ and found evidence of an overlaid effect of categorical and gradient /l/-darkening.

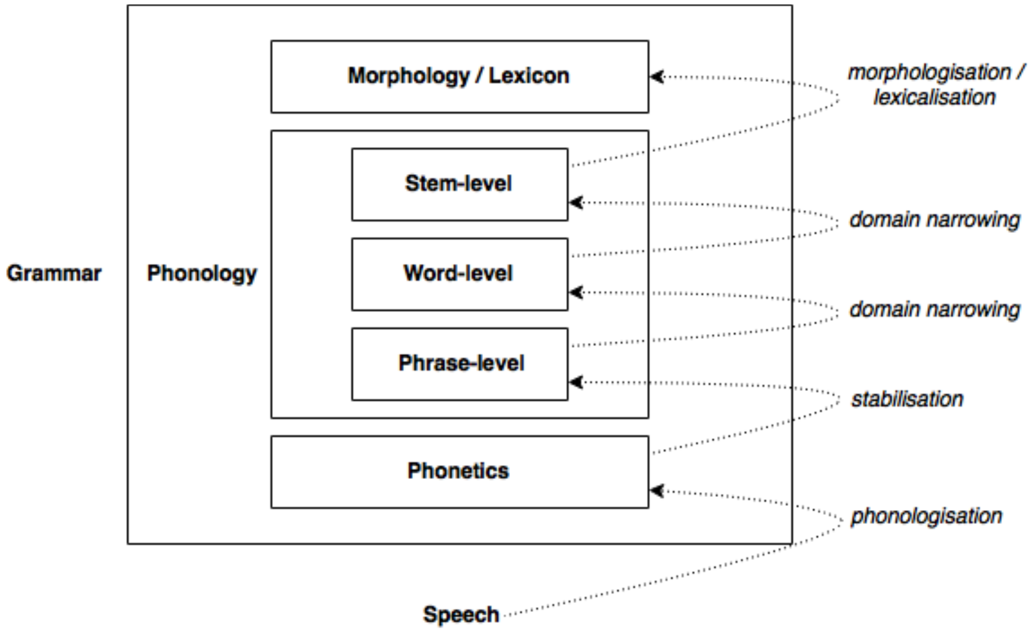


Figure 1.1: The life cycle of phonological processes (Bermúdez-Otero & Trousdale, 2012).

Applying the same explanation to the present case, one can hypothesise that the phonetic process of DIS may have undergone stabilisation and given rise to another categorical, phrase-level rule relevant to lenis stop voicing in Korean, while still operating in the phonetics component. Tentatively, the categorical rule can be formulated as shown in (1). Note that the rule is now called the lenis stop devoicing rule, following the assumption stated in Section 1.2.1 that the lenis stops are underlyingly voiced.

(1) Lenis stop devoicing: /b, d, g/ → [−voice] / $_{\mathrm{AP}}$[___

The rule states that the lenis stops become voiceless in AP-initial position. This categorical rule accounts for the apparent dichotomy of behaviours between AP-initial versus AP-medial lenis stops, as observed by S.-A. Jun (1993). At the same time, DIS continues to produce the continuum of phonetic variation on multiple parameters including the degree of voicelessness, from fully voiced to aspirated lenis stops.

This account has clear advantages over the alternatives which assume DIS as it relates to the lenis stops and the lenis stop voicing rule to be entirely separate or entirely identical. First, it parsimoniously accounts for the similarities and differences between the two processes. Secondly, it is generalisable to the case of Korean nasals, and to many similar cases observed cross-linguistically, such as the English palatalisation case above. Thirdly, it has superior predictive power. For example, over a long period of time, the lenis stop devoicing rule may be predicted to undergo domain-narrowing to apply at the Word-level, such that all PW-initial lenis stops would be categorically voiceless. Finally, it sheds some light on the controversy surrounding the categorical nature of the lenis stop voicing rule (see Section 3.4.2 for more detail), since a coexisting gradient process, namely DIS, can explain the messiness in the data. Based on the literature review in Chapter 3 and the results of the production study in Chapter 5, I will compare this account to the currently dominant view which treats the lenis stops as voiceless. Until the conclusions are drawn, however, I will continue to use the more familiar phonemic representation for the lenis stops /p, t, k/ and refer to the rule in (1) as the lenis stop (de-)voicing rule.

The relationship between denasalisation and DIS of nasal stops resembles the case of the lenis stops. There is disagreement over the existence of a cumulative link between the degree of denasalisation of initial nasals and the height of their posodic position. Young Shin Kim (2011) only found a categorical division between the denasalised AP-, IP-, and U-initial nasals and the sonorant AP-medial nasals. In contrast, other studies reported that the effect of DIS of nasals is generally proportional to the height of their prosodic position in the prosodic hierarchy (T. Cho & Keating, 2001; Yoo, 2015a). Thus, it can be hypothesised that certain aspects of DIS, such as denasalisation, have become discrete as a result of stabilisation. The resulting phrase-level rule can be formulated as shown in (2).

(2) Denasalisation: /m, n/ → [−nasal] / $_{\mathrm{AP}}$[____

This is only a preliminary rule as various aspects of the process are still uncertain, including if the discrete rule involves a complete or partial loss of nasality, if nasality is the only parameter on which DIS has stabilised, and if the AP is indeed the domain of the rule. It is one of the aims of the production study in Chapter 5 to formulate the precise rules of lenis stop devoicing and denasalisation.

Interestingly, a few existing studies have suggested that denasalisation may be an ongoing sound change which began relatively recently in Seoul and that it is in the process of spreading to other regions (Yoshida, 2008; Yoo, 2015a, 2015b, 2016). Three cities of Seoul, Busan, and Ulsan have been chosen based on the availability of previous work, accessibility, and population size. In terms of population size, Seoul (9.7 million) is roughly three times bigger than Busan (3.4 million), and, in turn, Busan is roughly three times bigger than Ulsan (1.1 million) (Korean Statistical Information Service, 2020). This set-up allows us to examine if the sound change involving denasalisation follows Trudgill's (1974) model of hierarchical diffusion, which describes how linguistic innovations jump between population centres in a cascade-like fashion. According to this model, we expect the nearest large city to adopt a sound change before intervening smaller, more rural locations. While all three cities are metropolitan centres of South Korea, based on the relative population size, we can predict that denasalisation will be more advanced in the order of Seoul, Busan, and Ulsan, even though Seoul is geographically closer to Ulsan than to Busan.

Another hypothesis can be tested thanks to the three city design. If denasalisation has indeed developed to varying degrees, the regional variation would allow us to capture the life cycle of phonological processes in action. For example, while DIS of nasals may have become a discrete rule of denasalisation with the domain of the AP in Seoul, it may still be entirely gradient in Ulsan. On the other hand, no regional variation has yet been reported for the lenis stop voicing rule. S.-A. Jun (1993) observed similar patterns of lenis stop voicing in Seoul and Jeonnam Korean, a variety spoken in the southwest of Korea.[7] Thus, it is predicted that, unlike the nasal stops, the lenis stops will show similar patterns across the three varieties.

As will be described in detail in Chapter 5, the production study in this dissertation is based on the auditory and acoustic analysis of the recordings of 10 university students from each of the three cities, Seoul, Busan and Ulsan. While previous studies on DIS and those testing the life cycle theory (e.g. T. Cho & Keating, 2001, Turton 2017) have used highly sensitive articulatory data in order to detect subtle articulatory variations, I adopt the present method for the following reasons. First, this method is one of the most favourable options for fieldwork, as it only requires a quiet room and a portable recorder. More importantly, an articulatory investigation tends to be time-consuming and expensive to conduct, imposing a practical constraint on its scale (e.g. Wrench, 2007). For example, T. Cho and Keating's study of DIS in Korean was based on three speakers including one of the authors, and Turton's study on /l/-darkening in English (2017) was based on one speaker per variety. In addition, previous works on DIS (e.g. T. Cho & Keating, 2001, Keating et al. 2004, Fougeron, 2001) have commonly found a significant amount of individual variation. Thus, the present method is intended to complement the existing literature and provide a more reliable representation of the regional varieties in question by conducting an acoustic study with a larger sample size.

The analysis of the production study will involve a separate investigation into each individual's pattern of DIS. This is in order to accurately determine the categoricity vs. gradience of a given process, e.g. lenis stop voicing, within each individual's system. In turn, this serves to answer the question regarding the status of these processes with respect to the

[7] Interestingly, S.-A. Jun (1993, p. 99) briefly mentions that, among her five Jeonnam (spelled "Chonnam" in her work) speakers and three Seoul speakers, one Seoul speaker showed a very different pattern. To quote her exact words, "[s]he produced each word so deliberately that most of lenis stop were voiceless even within a word. For this subject, it seemed that the lenis stop voicing rule does not exist at all." Although the exact nature of this speaker's pattern cannot be known from this description, it is possible that this speaker may represent an innovative stage in the life cycle of the lenis stop voicing rule, in which the rule has undergone domain-narrowing from the phrase-level stratum into a higher-level stratum in the phonological component.

grammar shown in Figure 1.1. I use statistical bimodality as evidence for categoricity in the data following previous works (Turton, 2017, 2014, Section 5.5.3.2; Bermúdez-Otero & Trousdale, 2012; Scobbie, 2005; Herd, Jongman, & Sereno, 2010, p. 508). However, when data from different individuals are pooled together, statistical bimodality in individuals' data can be masked, leading to a confusion between gradience and variance (see Section 5.3.9 for exemplification of this point and the pitfalls of the bimodality criterion). Thus, it is crucial to examine each individual's system on its own when trying to establish the grammatical status of a process.

To summarise, it is the second goal of the dissertation to test the hypothesis that the pattern of the lenis and nasal stop variation is best accounted for by rule scattering in the life cycle of phonological processes. The three-city design was adopted to potentially capture the relevant processes at different stages in the life cycle. The results of the study will have implications for the theories of speech production and perception, as the existence of DIS calls for an explicit mechanism to model how the prosodic structure of an utterance influences its phonetic realisation, and how listeners use these phonetic details to build the prosodic structure of the input utterance (T. Cho, McQueen, & Cox, 2007).

1.2.3 The perceptual role of domain-initial strengthening

The strong effects of DIS found in Korean are surprising from a functional, communicative perspective. There is a large body of evidence supporting the view that accurate transmission of meaning and effort reduction are the two driving forces of language evolution (Zipf, 1949; Piantadosi, Tily, & Gibson, 2011; Wedel, Kaplan, & Jackson, 2013; Futrell, Mahowald, & Gibson, 2015; Sóskuthy & Roettger, 2019). In the same vein, a recent study found statistical evidence that languages are significantly less likely to develop rules that modify word beginnings (Wedel, Ussishkin, King, Martin, & Geary, 2018). This is interpreted to be because phonetic cues at word-beginnings are crucial to lexical access and, therefore, to accurate conveyance of meaning. If effort must be reduced somewhere, this is likely to avoid word-beginnings.

However, when Korean nasals are strengthened domain-initially, it is not apparent how it serves either accurate transmission of meaning or effort reduction. With the crowded three-way system of oral stops, e.g. /t, t*, t^h/, it is odd that the phonetic space of the nasals should approach that of the lenis stops by losing nasality and voicing, and even gaining short aspiration (Yoo, 2015a). This increases the chance of perceptual confusion between these two types of consonants. For example, I have experienced a case in which /neilp^{h}it/ 'nail fit (a nail polish brand)' was misheard as /teipit/ 'David'. Besides, the realisation of a nasal stop phoneme as an oral plosive in IP-initial position – a position where complete denasalisation is most frequent (Yoo, 2015a) – cannot be a means of effort reduction. Since the velum must already be lowered for breathing, denasalisation of a nasal phoneme requires an extra velum raising gesture.

Assuming that Korean DIS cannot easily be understood in terms of accuracy or effort reduction, one may wonder if DIS provides other perceptual benefits. (Note, however, that the existence of perceptual benefits does not necesssarily imply that DIS has developed out of this functional need, see Section 2.4.2). While a substantial amount of literature has accumulated on the production of DIS, there is much less research on the perception side of DIS. Importantly, few studies directly test the interesting speculation that DIS aids prosodic demarcation (Fougeron & Keating, 1997, p. 3737). Previous research has focused on the perceptual roles of prosodic information in lexical segmentation and access. While articulatory strengthening can be seen as encompassing variation in duration, amplitude, and perhaps even pitch and voice quality in domain-edge segments, the main contribution of articulatory strengthening studies to existing knowledge is that prosodic structure is not only marked by what are traditionally considered prosodic resources, but also by segmental articulation.

Thus, this dissertation includes a perception experiment (Chapter 6) to investigate if listeners are able to use variation of segmental articulation as a cue to the prosodic structure of an utterance. Specifically, the experiment will test whether listeners from Seoul, Busan, and Ulsan can distinguish between a PW boundary and an AP boundary, based on the patterns of variation of segmental articulation of /t/ and /n/ found in Seoul Korean. With the three-city design, it is also possible to test whether the effect of DIS as a boundary cue is stronger for listeners whose pattern of DIS matches closely to that in the perceptual stimuli (Seoul) than for those it does not (Busan, Ulsan).

1.3 Research questions

Based on Section 1.2, research questions for the production (**RQ1 – RQ4**) and the perception study (**RQ5**) can be summarised as follows. Specific hypotheses will be given in the relevant chapters.

RQ1. What are the patterns of prosody-sensitive phonetic variation of the lenis stops and nasal stops in Seoul, Busan, and Ulsan Korean?

RQ2. What is the status with respect to the grammar (see Figure 1.1) of and the relationship among the lenis stop (de-)voicing rule, denasalisation, and DIS in the three varieties of Korean?

RQ3. What is the mechanism underlying DIS – e.g. articulatory, physiological, phonological (see Section 2.4.2 and references therein)?

RQ4. What is the best account of the prosodic variation of the lenis stops and nasal stops in Korean?

RQ5. Does DIS aid listeners in the demarcation of prosodic constituents?

1.4 Outline of dissertation

The rest of this dissertation will proceed in the following form. Chapter 2 will give a literature review of DIS, focusing on the phonetic characteristics and various accounts proposed for DIS. Chapter 3 will be about the three varieties of Korean being investigated. Any dialectal differences reported on their patterns of lenis stop devoicing, denasalisation, and DIS will be reviewed here, along with the variation in their prosodic systems. Chapter 4 will review existing research on prosodic parsing, to provide a background to the perception experiment. Then, Chapter 5 will present a production study aimed at answering **RQ1 – RQ4**. General discussion at the end of Chapter 5 will put together the findings and propose a unified account for the prosody-sensitive variation of the Korean lenis and nasal stops in domain-initial position. The implications for the description of Korean, for a general account of DIS, and for the architecture of grammar will also be discussed here. Next, Chapter 6 will report on a perception experiment conducted to answer **RQ5**. Lastly, Chapter 7 will provide key conclusions with regard to the aims of the dissertation and suggest some areas of future research.

Chapter 2.

Domain-initial strengthening

2.1 Chapter overview

In this chapter, I will provide an overview of the literature on domain-initial strengthening (DIS). First, I will introduce the phoneme inventory of Standard Korean to provide a background for understanding the Korean segments investigated in this thesis (Section 2.2). Next, I will summarise the phonetic characteristics of DIS found across different segment types, languages, and individuals (Section 2.3). Then, I will discuss the nature and the status of DIS and evaluate different accounts proposed for the phenomenon (Section 2.4). In the end, I will provide a short summary of the chapter (Section 2.5).

2.2 The phoneme inventory

Despite our focus on the lenis and the nasal stops, the complete phoneme inventory of Korean is described here for three main reasons. Firstly, when interpreting DIS patterns of a language, it is useful to consider its overall phonemic structure. For example, the unusual three-way contrast of plosives in Korean and the crowded phonetic space that results from it allow us to appreciate the bigger picture of DIS found in Korean. Secondly, establishing the entire phoneme inventory of Korean will be helpful in discussing dialectal differences in Chapter 3 and designing the experimental materials for Chapters 5 and 6, such that the materials do not accidentally create a source of bias for a particular dialect group.

Finally, this section gives me an opportunity to clarify and justify my choice of phoneme symbols as no single transcription system is uniformly adopted in the field. For example, researchers transcribe the bilabial fortis stop differently as /p*/ (T. Cho et al., 2002; Shin, 2015), /p'/ (S.-A. Jun, 1995c), or /p˭/[8] (Ho-Young Lee, 1996; Young Shin Kim, 2011). In this dissertation, I will adopt the /p*/ form for the fortis series, as it is widely used at least

[8] IPA diacritic denoting 'unaspirated'.

in the literature written in English. The symbol /p'/ may wrongly suggest that Korean fortis stops are articulated like ejectives, whereas the use of /p$^{=}$/ does not adequately capture the finding that the fortis stops are phonetically more complicated than the more usual kind of unaspirated stops found in languages such as French and Dutch.

Another set of phonemes whose transcription is controversial is the affricates. The following symbols /ɟ[9], c, t͡ʃ, t͡ɕ, t͡s/ are variably adopted to represent the lenis affricate. Similarly, /c, c*, t͡ʃ*, t͡ɕ*, t͡s*/ are used for the fortis affricate, and /c^{h}, c^{h}, t͡ʃ^{h}, t͡ɕ^{h}, t͡s^{h}/ for the aspirated affricate. While the use of the symbols based on /c/ may be taken to imply that the consonants are (palatal) stops rather than affricates, this choice seems to be largely based on the IPA guidelines to use Roman characters when possible, and to recognise the use of one-character symbols to represent affricates that constitute single phonemes for convenience (IPA, 1949, p. 15). Indeed, H.-B. Lee (1999) categorised the phonemes /ɟ, c^{h}, c/ as affricates. Nevertheless, the use of a stop symbol for an affricate may be misleading particularly for the readers who are not familiar with Korean. The alternative symbols that are also widely used are postalveolar /t͡ʃ/ (S.-A. Jun, 1995c; T. Cho et al., 2002) and alveolo-palatal /t͡ɕ/ (Ho-Young Lee, 1996; Young Shin Kim, 2011; Shin, 2015). As with /ɟ, c/, these indicate that the Korean affricates are articulated further back in the mouth compared to the alveolar stops and fricatives.

However, more recent articulatory and acoustic studies have revealed that the place of articulation for the affricates is alveolar, like the plosives (H. Kim, 1999, 2001, 2004; Ko, 2013). An MRI study by H. Kim (2004) revealed that the difference between these segments lies not so much in the location of constriction, but in the tongue shape; while the affricates are laminal, the plosives are apical or apico-laminal. There was also a tendency for the fortis and the aspirated affricates to have a more extensive contact area including the postalveolar zone than the lenis affricate. Phonological arguments for treating the affricates as alveolo-palatals rather than alveolars have also been refuted in H. Kim (1999, pp. 323-329). One of the phonological arguments comes from Hume's (1990) study of opacity effects in Korean Umlaut. Umlaut refers to a phonological phenomenon in which a back vowel becomes fronted when followed by a high front vowel or a semi vowel. Based on an informant from the Gyeongsang province, Hume (1990) observed that Umlaut is blocked across an affricate, e.g. /tʌti + ta/ [tʌ.di.da] ~ [te.di.da] 'moving slowly', but /p^{h}ʌt͡si + ta/ [p^{h}ʌ. d͡zi.da] *[p^{h}e. d͡zi.da] 'to spread out'. To explain why Umlaut is blocked by affricates but not by other coronal consonants, Hume (1990) proposed that Korean affricates are alveolo-palatals that are specified for the secondary vocalic feature [coronal], which then blocks the spreading of

[9] The use of this symbol assumes that the lenis stop is underlyingly voiced, contrary to a more widely accepted view that Korean has three voiceless stops.

[coronal] from the following front vocoid to the preceding vowel. However, H. Kim (2000) conducted a survey with 34 speakers from Gyeongsang and four other dialects and found evidence that Umlaut can in fact occur across affricates, e.g. /pʌn.t͡si.lɯ.lɯ/ [pʌn.d͡zi.lɯ.lɯ] ~ [pen.d͡zi.lɯ.lɯ], although there were dialectal, lexical, and individual differences. Thus, following Kim, I will transcribe the Korean affricates as /t͡s , t͡s*, t͡sʰ/ which reflect their alveolar place of articulation. These symbols have an additional advantage of convenience and familiarity due to the use of Roman characters.

Table 2.1 and Table 2.2 show the seven vowels and the twenty-five consonants of Standard Modern Korean, respectively. The vowel inventory follows Ho-Young Lee's (1996) decision to reflect the sound changes including the loss of length distinction and the loss of monophthongs /y/ and /ø/, which are now realised as [wi] and [we]. In addition, /w, j, ɰ/ are included in the consonant inventory as semi-vowels rather than as part of the diphthongs, as in Ho-Young Lee (1996). However, the present vowel inventory differs from Ho-Young Lee's in that it adopts the more progressive seven-vowel system (Shin, Kiaer, & Cha, 2012; Shin, 2015) as opposed to the traditional ten-vowel system (Huh, 1965; Ho-Young Lee, 1996) based on the following evidence. Kang (2013) showed that the merger of /e/ and /ɛ/ is complete among younger Seoul Koreans, with only older males retaining the distinction in word-initial position. I use the symbol /e/ to represent the merged phoneme according to Kang's finding that it is /ɛ/-raising that is driving the merger[10]. There were other sound changes reported in Kang (2016), namely /ɯ/- and /u/-fronting and /o/-raising. In particular, /ɯ/ is so fronted that the phonetic value is closer to [ɨ] than to [ɯ]. Furthermore, the phonetic value of /ʌ/ is more fronted and slightly rounded compared to its cardinal value. Nevertheless, for these three phonemes, I keep the traditional phoneme symbols as they are better recognised among the researchers of Korean Linguistics. The consonant inventory here closely follows the one in Ho-Young Lee (1996) except for treating the affricates as being alveolar in the place of articulation.

[10] Note that some researchers use /ɛ/, assuming that the original phonemes were /e/ and /æ/. Shin (2015, p. 9) compared (a) the vowels in spontaneous speech by 20 speakers of Korean, and (b) the cardinal vowels produced by Daniel Jones. The results showed that the Korean phoneme is similar to, but have slightly lower F1 than Cardinal Vowel 3. Since Cardinal Vowel 2 is not shown and the age of the informants is unclear, it leaves open the question of whether [e] or [ɛ] better reflects the phonetic value of the merged phoneme among Younger Seoul Koreans.

Table 2.1: The vowel inventory of Standard Korean. The unrounded back vowels are generally fronter than the rounded back vowels.

	Front	**Back**	
	Unrounded	*Unrounded*	*Rounded*
High	i	ɯ	u
Mid	e	ʌ	o
Low		a	

Table 2.2: The consonant inventory of Standard Korean. The arrows indicate my choice to adopt /t͡s/-based symbols for the affricates.

		Bilabial	**Dental/ Alveolar**	**Palatal**	**Velar**	**Glottal**
Plosive	*Lenis*	p	t		k	
	Fortis	p*	t*		k*	
	Aspirated	p^{h}	t^{h}		k^{h}	
Affricate	*Lenis*		t͡s ←	t͡ɕ		
	Fortis		t͡s* ←	t͡ɕ*		
	Aspirated		t͡s^{h} ←	t͡ɕ^{h}		
Fricative	*Lenis*		s			h
	Fortis		s*			
Nasal		m	n		ŋ	
Liquid			l			
Semi-vowel		w		j	ɰ	

2.3 Phonetic characteristics of domain-initial strengthening

DIS is a cross-linguistic phenomenon in which domain-initial segments show enhanced articulation on temporal and/or spatial dimensions (e.g. Fougeron & Keating, 1997; T. Cho & Keating, 2001; Keating et al., 2004). That is, domain-initial segments are articulated with larger articulatory gestures and/or longer durations than domain-medial segments. Spatial and temporal enhancements are logically separate, but T. Cho & Keating (2001, p. 179) note that they coincide in Korean DIS. It is also known that the effect of DIS is generally cumulative, such that, for example, an IP-initial segment shows “stronger” articulation than an AP-initial segment, which, in turn, shows stronger articulation than a PW-initial segment.

That said, there are few further details that can be provided about the behaviours of domain-initial segments as a whole; the precise effects of DIS vary on many levels. Below, I will summarise the variation of its phonetic characteristics by segment type, language, and individual.

2.3.1 Variation by segment type

2.3.1.1 Stops

Stops are one of the most frequently studied segments in the DIS literature and they show relatively robust effects of DIS. Different types of stops exhibit DIS effects on different phonetic parameters and to varying extents. This is clearly illustrated in the study by T. Cho and Keating (2001) who made a wide range of articulatory and acoustic measurements for the Korean alveolar stops, /n, t, t^h, t*/, in the initial position of five different prosodic domains, the U, IP, AP, PW, and S (Syllable). Due to the difficulty of designing comparable materials for all five domains, the domains from U to PW were examined in one set, and PW and S in a different set. The results are summarised below.

(1) Linguopalatal contact: This was measured through Electropalatography (EPG). Greater contact was found in a higher domain for all stops with a shift of place from the palato-alveolar to the denti-alveolar region, differentiating all above-word prosodic domains (U > IP > AP > PW). The lower levels of PW and S were only reliably distinguished for /n/, with the contact being greater PW-initially than S-initially. For two speakers, the pattern was reversed (PW < S) for /t^h, t*/. The range of variation across prosodic positions was greater for /t, n/ than /t^h, t*/, as linguopalatal contact in AP- and PW-initial positions was smaller for /t, n/.

(2) Stop seal duration: Longer duration was found in a higher domain for all four stops, differentiating all prosodic domains investigated (U > IP > AP > PW > S). The only exception to the gradient pattern was that S-initial /t^h, t*/ showed longer seal duration than PW-initial /t^h, t*/.

(3) Skewness of articulatory movement: The trajectories for linguopalatal contact were plotted against time and the skewness of the shape of this distribution was measured. If the shape of the trajectory was positively skewed or skewed to the right with a longer right tail, it suggests that the consonant closing gesture occurs more quickly than its opening gesture and vice versa. In general, the trajectory for /t, t^h, t*/ became progressively more negatively skewed in a higher domain, suggesting that the contact

peak is formed relatively later in a higher domain. Detailed statistical results were not reported and measurements for /n/ were not made with no further explanation.

(4) Acoustic closure duration: For this parameter, measurements for U-initial tokens were not taken because the onset of these tokens could not be reliably delineated from the preceding silence. All tokens of /t/ were also excluded because they were frequently realised as approximants, showing no acoustic closure. The results showed longer duration in the order of IP > AP > PW > S for /n/, and IP > AP > PW = S for /tʰ, t*/. Note that T. Cho & Keating (2001, Section 3.3) found a strong correlation between linguopalatal contact and acoustic closure duration ($R^2 = 0.66$ on average) which is unsurprising as there was also a high correlation between linguopalatal contact and articulatory closure duration (T. Cho & Keating, 2001, Section 3.1.4). This substantiates the methodology in the present study, using acoustic closure duration as a proxy for the two key articulatory measurements.

(5) % Voicing during acoustic closure: A smaller percentage of voicing was found in a higher domain (IP < AP < PW), since voicing residue at the beginning of the closure remained more or less the same, while the total closure duration became longer in higher positions. Measurements for /n/ were not taken, presumably based on the assumption that /n/ is always voiced (but see Section 3.5 of this dissertation for discussion of voiceless realisations of /n/ [t, t‘] in Korean). /t/ showed greater variation than /tʰ, t*/, as /t/ was nearly fully voiced in PW-initial position as opposed to around 50% for /tʰ, t*/.

(6) Voice Onset Time (VOT): Longer VOT was observed for /t, tʰ/ in a higher position in the order of U > IP > AP > PW and, additionally, PW > S for /tʰ/. PW- and S-initial /t/ were not compared as their VOT was always zero. /n/ was not measured because it was assumed not to show a positive VOT (again, see Section 3.5). /t*/ showed short and stable VOT across different positions so it was also excluded from the analysis.

(7) Total voiceless interval: This measure was obtained by adding voiceless closure duration and VOT, and reflects the duration of glottal opening. The measurements for /tʰ/ distinguished all prosodic levels more consistently than did VOT or closure duration alone. The other consonants were not included as they did not produce much glottal opening during the closure.

(8) RMS burst energy: This was measured for a given time window (5 ms for /t*/ and 10 ms for the rest) centred over the stop release. A general trend of less burst energy in a

higher domain was observed for /t*, t^h/ (U = IP < AP = PW), although this was not statistically supported. No clear trend was found for /t/, and /n/ was excluded with the assumption that there is no stop burst (but see Section 3.5 for exceptional findings to the contrary).

(9) Nasal energy minimum: This is a measurement of RMS acoustic energy at its lowest point during a nasal consonant. It indirectly reflects the highest point in the velum height during a nasal consonant, as the smaller the velic opening, the weaker the excitation of the nasal resonances. U-initial tokens were excluded as they were preceded by silence and thus the minimum values were always zero at the beginning. Three levels were distinguished with less nasal energy in a higher domain (IP < AP = PW < S).

(10) Duration of vowel in domain-initial CV: No statistically significant effect of position was found, although there was a trend for the vowel after /t*/ to be longer in a higher domain. However, this could arise because the aspiration of /t, t^h/ is simply devoicing the start of the vowel. In line with this view, the consonant type had a significant effect on vowel duration and a visual examination of the results in Figure 17 (T. Cho & & Keating, 2001, p. 175) reveals that /n/ behaved more similarly to /t*/ than to the other two stops.

(11) Duration of domain-final vowel: The domain-final vowel was generally longer in a higher domain in the order of U = IP > > AP > W. There was a much larger difference between the IP and the AP compared to the AP and the PW, making U- and IP-final vowels categorically longer than AP- and PW-final vowels.

While the specific patterns of DIS varied for the different stops, there was some consistency in the direction of the effects for the majority of the cases; the duration became longer, the constriction area larger, and the voicelessness greater. Nevertheless, there were two parameters, RMS burst energy and nasal energy minimum, on which the effect was perhaps not in the expected direction. RMS burst energy was observed to be smaller in a higher domain, although the effect was weaker and more inconsistent compared to the effect on the other parameters. The authors stated that this may be due to the possibility that greater contact resulted in a longer release duration, which then led to less peak burst energy (Stevens, Keyser, & Kawasaki, 1986, as cited in T. Cho & Keating, 2001, p. 180).

Alternatively, it might be that the need to form a glottal closure given the inertia of voicing from a preceding vowel (in an AP- or PW-initial position) resulted in greater vocal fold tension and supraglottal pressure leading to a stronger stop burst than when there is weak or no voicing before the segment (in a U- or IP-initial position). This may also explain the reversed pattern of a longer duration and greater contact in S-initial /t^h, t*/ than PW-initial /t^h, t*/ position. That is, the inertia of voicing in S-initial position was so strong that a longer duration may be required to achieve a complete glottal closure needed in the production of the two types of segments. As noted by the authors, the longer duration may then explain the greater contact, based on the two findings: a high correlation between the measures of seal duration and linguopalatal contact; and that the peak of the contact was formed relatively later in segments in higher domains.

The nasal energy minimum of domain-initial /n/ also showed an unexpected result. The nasal energy decreased as the prosodic domain of the initial nasal became greater. A similar pattern has been observed across different languages, including French (Fougeron, 2001), English (Fougeron & Keating, 1997), and Estonian (Gordon, 1997). This challenges the simple interpretation of DIS that it enhances segmental articulation. If the articulation of a nasal is enhanced, the velum should lower further and increase nasal energy, as velum lowering is expected to be a key articulatory gesture in the production of an underlying nasal.

One of the commonly discussed explanations is that, there are two distinct types of contrast enhancement: *syntagmatic contrast enhancement* and *paradigmatic contrast enhancement.* In this particular case, DIS serves to enhance syntagmatic contrast, which makes a segment more dissimilar to adjacent segments in terms of sonority (Fougeron, 2001, pp. 130-131; T. Cho & Keating, 2001). That is, a domain-initial consonant would become less sonorous by being realised with a longer and stronger stop closure, less voicing, and, in the case of nasals, reduced nasality. These result in an enhanced contrast between the initial consonant and the following vowel. On the other hand, some cases of DIS increase the paradigmatic contrast of segments, whereby distinctive features of a segment are enhanced such that it becomes more dissimilar to competing phonemes in the language's segmental inventory. An example from Taiwanese will be given in Section 2.3.2.

2.3.1.2 Fricatives

Across different languages, fricatives have been found to show weaker effects of DIS than stops, albeit being similar in nature. In a study of the Korean fricatives /s, s*/, S. Kim (2001) found greater linguopalatal contact and longer acoustic duration in a higher domain. For one of the two subjects, she also found a higher centroid frequency for /s*/ in a higher domain, indicating a smaller front cavity. Finally, IP- and AP-initial /s*/ showed less creakiness in the beginning of the following vowel than AP-medial /s*/, as suggested by higher H1-H2 values.

Compared to T. Cho and Keating's study of the aforementioned Korean stops, S. Kim's results showed greater degrees of between-speaker variation, smaller effect sizes, and smaller numbers of prosodic levels that are consistently distinguished, namely two levels compared to three for the stops. The finding of weaker effects of DIS in fricatives is consistent with the findings for other languages, including English (Byrd, 1994; Keating, Wright, & Zhang, 1999) and French (Fougeron, 2001). Fougeron reported less clear DIS effects in the linguopalatal contact for /s/ compared to the other consonants /t, n, k, l/ in the study. This was interpreted to be a result of a lesser degree of articulatory freedom in the production of /s/, as just the right amount of constriction is required to produce a fricative (Shadle & Scully, 1995, as cited in Fougeron, 2001, p. 125).

2.3.1.3 Vowels

Fougeron and Keating (1997) examined the realisation of the CV syllable /no/ in different prosodic positions for American English. Interestingly, they found that /n/ and /o/ behaved in opposite ways. While domain-initial /n/ showed greater linguopalatal contact compared to /n/ in domain-medial and -final position, domain-final /o/ showed less linguopalatal contact (more opening) than /o/ in domain-medial and -initial position. This was interpreted to suggest that consonants show domain-initial strengthening by means of enhancing their consonantality, whereas vowels show domain-final strengthening by means of enhancing their sonority.

However, there is an alternative explanation for this pattern. A vowel in a domain-initial CV might not have shown strengthening because it is not strictly at the left edge of the domain, i.e. #V, rather than due to some inherent difference between articulatory strengthening of consonants and vowels. Indeed, S. Kim and Cho (2012) provided evidence for the locality of DIS effects. Tongue dorsum maxima measured by Electromagnetic Articulography (EMA) showed that tongue fronting was greater IP-initially than IP-medially when it occurred in a vowel-initial word *add*, but not when it occurred in consonant-initial words, *had* and *pad*. Based on these results, the authors suggested that DIS effects may be conditioned by *phonological distance* from the boundary, as *had* behaved similarly to *pad* in which /æ/ is in the nucleus. If the effect was a function of *phonetic distance*, *had*, with the absence of phonetic interference from supralaryngeal constriction after the boundary, would have patterned with *add*.

However, the effects of DIS were observed regardless of the location of /æ/ on all the other parameters. Tongue position was lower in all three words in IP-initial position. Jaw and lip opening maxima were also greater IP-initially, but only when the target words were unaccented. On the other hand, /æ/ in accented target words exhibited a robust strengthening effect for all four parameters (i.e. lip opening, jaw opening, tongue fronting, and tongue height) regardless of prosodic position and location within the initial syllable. Two important observations can be made from these results. First, the locus of DIS is restricted to the first segment after a boundary on certain phonetic parameters, but not on others. Second, two distinct sources of articulatory strengthening exist, *boundary-induced strengthening* and *accent-induced strengthening*, which show different patterns and interact in subtle ways.

While differential effects of boundary- and accent-induced strengthening have been explored with a range of different vowels, the results are not entirely consistent among different studies. Focusing on /i, ɪ/ this time, S. Kim and Cho (2011) conducted a study similar in the design to the above, with the target words *eat-it*, *heat-hit,* and *Pete-pit*. In domain-initial position, no strengthening effect was found on any parameter except tongue height for /ɪ/ in the vowel-initial word, *it.* The authors suggested that the lack of an effect on *eat* may be due to a relatively restricted articulatory space for the tongue to become higher in /i/. However, this does not explain why /ɪ/ did not also show tongue fronting following the same logic, given that DIS was found to cause both tongue fronting and raising for /æ/ in *add*.

When accented, both /i, ɪ/ were produced with a fronter tongue body and more lip opening in all the target words. The tongue position was higher only in vowel-initial words and greater jaw opening was found only in consonant-initial words. While the details of the findings are different, there are some common findings from the two studies. First, the locality of DIS effects is confirmed again, with the *phonological* distance from a prosodic boundary determining whether a segment is affected by DIS. Second, accent-induced strengthening appears more robust as it affects a wider range of phonetic parameters.

Unfortunately, even these observations are not always supported. T. Cho (2005) examined sequences of domain-final and domain-initial syllables, bi#bi and bɑ#bɑ. Even though V2 (the second vowel in a sequence) was not immediately preceded by a prosodic boundary, DIS was still found, albeit less consistently than domain-final strengthening in V1. Unlike S. Kim and Cho's study (2011) in which no fronting or raising was found in IP-initial /i/ in *Pete,* tongue raising was found consistently in both ip (Intermediate Phrase)-initial and IP-initial *bi*. In addition, an increase in lip opening, a *decrease* in jaw opening were observed for /i/ in domain-initial syllables. The opposite direction of lip opening and jaw opening was explained as reflecting a conflict between the need to enhance sonority (syntagmatic contrast enhancement) and the need to enhance the phonological feature of /i/ (paradigmatic contrast enhancement); that is, a larger lip opening serves to enhance sonority of the vowel and a smaller jaw opening serves to enhance the feature [+high]. For /ɑ/, on the other hand, a larger lip opening, and lower and backer tongue dorsum were found but with no change in the jaw position.

In accented syllables, both /i, ɑ/ showed greater jaw and lip opening. For /ɑ/, the tongue was backer and lower, whereas for /i/, the tongue was fronter but *not necessarily higher*. Thus, the generalisation above does not hold in this case, that accent-induced strengthening is more robustly manifested on different parameters than boundary-induced strengthening. Given the patterns in this study, T. Cho suggested that articulatory strengthening in general has the two goals of sonority enhancement and phonological feature enhancement, but that *accent-induced strengthening* suppresses enhancement of features which are in conflict with sonority expansion, such as [+high]. Note, however, that this is not supported in S. Kim and Cho (2011) who found an increase in tongue height for /i, ɪ/ in vowel-initial syllables under accent.

The survey in this section did not cover the full range of segment types, leaving aside e.g. approximants, mainly due to a relative lack of research on certain segments. However, the above is sufficient to exemplify subtle and complex variation observed in the ways different segments are affected by DIS. The specific phonetic parameters that are affected, the direction and the magnitude of the effect, the number of prosodic levels distinguished, and whether the effect serves to enhance syntagmatic or paradigmatic contrast all vary by segment. And it has been mentioned that this may be partly attributed to the inherent physiological constraints in the articulation of certain types of segments, e.g. /s/ shows less prosody-sensitive variation, and /ɪ/ but not /i/ shows tongue raising.

For the purpose of conducting a production study later, the literature review here has highlighted the importance of making a range of relevant measurements, with an understanding of the particular method of measurement adopted. For example, T. Cho and Keating's (2001) study demonstrated the differences between the measurements of acoustic duration and those of articulatory seal duration obtained through EPG. Articulatory seal duration is superior at capturing the onset of a stop closure that does not produce audible acoustic noise, which is common after a silent pause. On the other hand, the sampling interval of 10 ms of EPG measurements means that, first, a closure shorter than 10ms may not be captured, and, second, the closure duration between the last frame captured by EPG and the actual release will be lost. In addition, as the aforementioned studies found much between-speaker variation among two to four subjects, conducting a larger-scale acoustic study would be a valuable complement to the existing studies.

2.3.2 Variation by language

(1) Cross-linguistic differences in DIS and their accounts

While it is clear that DIS is a phonetic process found in a wide range of different languages, the patterns of DIS show interesting cross-linguistic variation even for the same type of segment. For certain phonetic parameters, the differences are minimal. Peak linguopalatal contact and seal duration are examples of such parameters. In a range of different languages, including English, French, Korean, Taiwanese (Keating et al., 2004), and Japanese (Onaka, Palethorpe, Watson, & Harrington, 2003), the DIS patterns were similar for these parameters. The direction of the effect was consistently towards greater linguopalatal contact and longer duration in higher prosodic domains, and the pattern was cumulative in cases where more than two levels of prosodic domains were distinguished. The languages only differed in terms of the magnitude and the systematicity of the effect, such as the number of prosodic levels that were consistently distinguished through DIS. Korean showed the most consistent and clear effects of DIS, distinguishing either all five or four prosodic domains by linguopalatal contact. In comparison, Japanese and Taiwanese showed more between-speaker variation, with some speakers only distinguishing two levels, between the IP and a smaller domain.

Keating et al. (2004, p. 149) speculated that the differences in the DIS effects may be explained by the differences in the languages' prosodic systems. English has lexical stress and nuclear pitch accents, and thus the domain edges may be expected to be less marked than in languages without such prominent heads, such as Taiwanese, French, and Korean. Since Taiwanese is a lexical tone language, in which pitch has a restricted use in phrase-level prosody, it may be expected that more edge-marking is achieved through other phonetic means including DIS. On the other hand, it has been argued that the Taiwanese Tone Sandhi Group is not part of the prosodic hierarchy (Hsu & Jun, 1996), so one might expect prosodic domains to receive little phonetic marking. French and Korean have neither lexical stress nor lexical tone. They also share similar features in their prosodic structure, such as having the AP, a small phrasal domain defined by phrasal tones. However, it has been argued that Korean reinforces the beginning of the AP, while French marks the end. This is following the analysis that the end of the French AP is marked by a phrasal accent (H*) and final lengthening (Fougeron & Jun, 1998), whereas the Korean AP lack these but show initial lengthening (S.-A. Jun, 1995c, 1998). Given this tendency for phrase-initial marking, it is not surprising that Korean exhibits the most consistent and robust effects of DIS (Keating et al., 2004).

While variation found in linguopalatal contact and seal duration merely concerns the magnitude of the effect, cross-linguistic variation on VOT is more complex. In the rest of the section, I will review and critique literature on the variation of VOT for four languages: Taiwanese, English, Korean, and Dutch. The findings from these languages clearly demonstrate the complexity of the phenomenon and highlight the inconsistency in the accounts which have been proposed for each language's pattern. A summary of these findings is provided in Table 2.3 for the reader's reference.

Table 2.3: Cross-linguistic variation in VOT for domain-initial stops. Positive means that a stop shows greater positive VOT, or longer voicing lag, in a higher domain; negative means that it shows greater negative VOT, or longer voicing lead. A dash (-) means that the language does not have the relevant stop category. [±s.g.] refers to the feature 'spread glottis', and [±c.g.] to 'constricted glottis'. The final row summarises the accounts proposed for the pattern in each language.

Stop category	**Korean**	**Taiwanese**	**English**	**Dutch**
Voiceless aspirated, e.g. /t^h/ [+s.g., −c.g.]	Positive	Positive	Positive	-
Voiceless unaspirated (lenis), e.g. /t/ [−s.g., −c.g.]	Positive	No effect[11]	-	Negative
Voiceless unaspirated (fortis), e.g. /t*/ [−s.g., +c.g.]	No effect or a small effect towards negative		-	-
Voiced, e.g. /d/ [+slack vocal folds]	-	Negative	Positive	No cumulative effect (the middle category showed the longest VOT)
Account	Paradigmatic + syntagmatic contrast enhancement based on privative features	Paradigmatic contrast enhancement	Syntagmatic contrast enhancement	Phonetic feature enhancement

[11] I shall not discuss or make a claim about whether the Taiwanese voiceless unaspirated stops should be considered more like the lenis or the fortis stops in Korean.

The aforementioned study by Keating et al. (2004) showed a highly consistent DIS effect on VOT for the Korean lenis stop /t/, with progressively longer VOT in a higher domain. In contrast, no consistent effect was found for Taiwanese unaspirated /t/ and only two levels were distinguished (IP > S) for French /t/. Following these results, one may conclude that Taiwanese and French simply do not show DIS effects on the VOT parameter. However, as will be discussed below, the study by Hsu and Jun (1998) reveals the importance of considering the phonological structure and the segmental inventory (e.g. the full three-way stop contrast) of a language when analysing its patterns of DIS.

Focusing on the three-way stop contrast (e.g. /p^h, p, b/) in Taiwanese, Hsu and Jun (1998) found that these three stop categories were not affected uniformly in terms of VOT. As shown in Figure 2.1, voiceless aspirated stops, characterised by a long-lag VOT, showed an even longer VOT in a higher prosodic domain. Voiced stops, on the other hand, showed an effect in the opposite direction, showing a longer voicing lead in a higher prosodic position. Finally, voiceless unaspirated stops did not show an effect one way or the other, replicating the results by Keating et al. (2004). This was interpreted as paradigmatic contrast enhancement, as the three stops become more dispersed on the VOT continuum as a result of the strengthening.

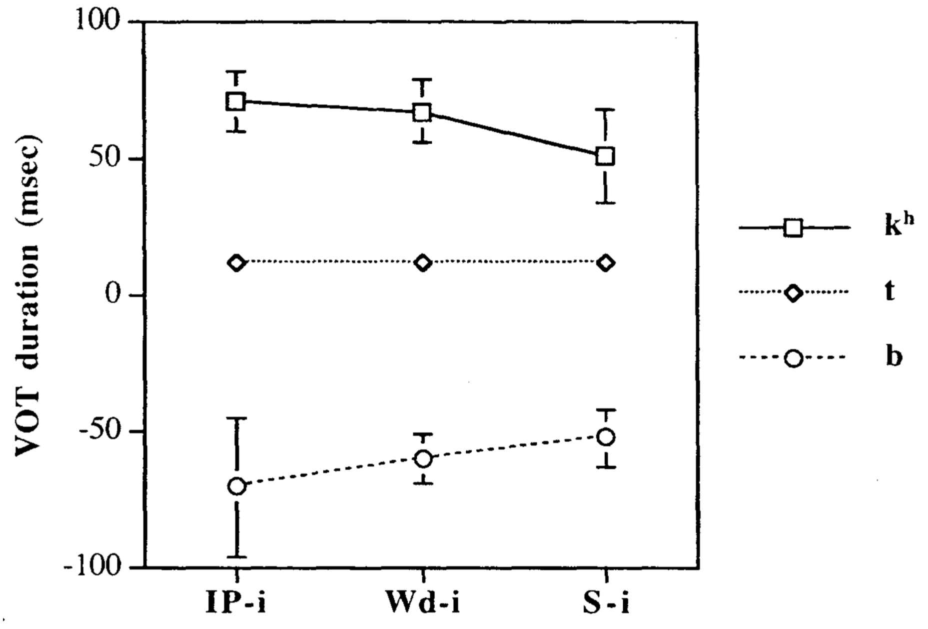


Figure 2.1: VOT variation of three stop categories in Taiwanese at three prosodic positions: IP-initial, Word-initial, and Syllable-initial positions (Hsu & Jun, 1998, p. 82). The error bars indicate one standard deviation. The reason why error bars are missing for /t/ is unclear from the original article.

This is in contrast to the pattern found in English, where VOT of the voiceless stops becomes longer phrase-initially (Pierrehumbert & Talkin, 1992) and the voiced stops are typically realised as an unaspirated voiceless stop in phrase-initial position (Davidson, 2016; Lieberman & Blumstein, 1988, p. 215). Thus, the pattern in English fits better with the explanation based on syntagmatic contrast enhancement, in which stops become more dissimilar to the neighbouring vowels in terms of sonority.

However, neither of the accounts, paradigmatic or syntagmatic contrast enhancement, can adequately capture the pattern observed for the three contrastive stops in Korean. Unlike Taiwanese, Korean is known to have three *voiceless* stops: aspirated, lenis, and fortis, e.g. /p^h, p, p*/. T. Cho and Jun (2000) investigated the aerodynamic measurements of VOT, peak airflow, and total amount of airflow for the three contrastive bilabial stops. The results for VOT are shown in Figure 2.2 (peak airflow and total airflow showed similar patterns). Both the aspirated voiceless stop /p^h/ and the lenis stop /p/ showed a large degree of strengthening on the parameters of VOT, peak airflow, and total amount of airflow. Nevertheless, VOT was consistently higher in /p^h/ than /p/ such that there was no overlap on any of the parameters. On the other hand, the fortis stop /p*/ did not show a statistically significant effect of DIS on any parameter, consistent with the results in T. Cho and Keating (2001). Thus, the pattern here is partly similar to that of Taiwanese in that the stops become more dissimilar from each other in higher prosodic domains. On the other hand, it also differs from Taiwanese in which there was polarisation of the values at the two extremes (/p^h, b/) leaving the category in the middle (/p/) unaffected.

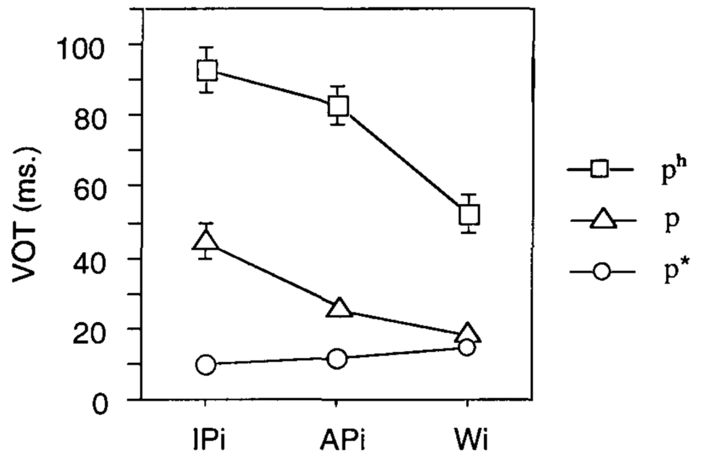


Figure 2.2: VOT variation of three stop categories in Korean at three prosodic positions (T. Cho & Jun, 2000, Section 3.1).

To account for this pattern, the authors put forward an explanation based on a combination of paradigmatic and syntagmatic contrast enhancement. Following Lombardi

(1991), the privative laryngeal features of the three Korean stops can be represented as shown in (3).

(3)	aspirated	[spread glottis]
	fortis	[constricted glottis]
	lenis	unspecified

It is proposed that DIS strengthens the relevant distinctive features paradigmatically, whereas for categories with no feature specification, syntagmatic contrast enhancement is applied. The increased VOT and airflow for the domain-initial aspirated stop can be taken to indicate enhancement of the feature [spread glottis]. While the results for the fortis stop were not entirely consistent, examination of the data from S.-A. Jun, Beckman, & Lee (1998), who investigated glottal gestures through fibrescopic technique, pointed to a similar pattern; the fortis stop showed smaller glottal opening AP-initially than AP-medially. Taken together with the VOT and airflow data from T. Cho and Jun (2000), the tendency may be seen as an enhancement of the feature [constricted glottis]. Lastly, the observation that the lenis stops still showed increased VOT and airflow was interpreted to suggest that when a stop is unspecified for a feature, it is strengthened towards enhancing syntagmatic contrast with the neighbouring sonorant segments.

To complicate the picture, Dutch showed a pattern different from all of the languages above, compatible with neither paradigmatic nor syntagmatic contrast enhancement. Voiceless stops in Dutch are characterised with relatively short VOTs (/p/: 10 ms, /t/: 15 ms, /k/: 25 ms), while voiced stops are produced with relatively long prevoicing (/b/: -85 ms; /d/: -80 ms) (Lisker & Abramson, 1964). T. Cho and McQueen (2005) examined Dutch /t, d/ along with other segments in the initial position of Big Phrase, Small Phrase, and Prosodic Word. As the prosodic structure of Dutch is still controversial, the following groups were determined by the presence or absence of a pause and a boundary tone based on de Pijper and Sanderman (1994), and Gussenhoven, Rietveld, and Terken (1999):

(a) Big Phrase: boundary tone and pause
(b) Small Phrase: boundary tone and no pause
(c) Prosodic Word: no boundary tone and no pause

The results showed that voiceless stops were produced with progressively shorter VOT in a higher prosodic domain. For voiced stops, there was no such cumulative effect; VOT was longest in Small Phrase and shortest in Big Phrase. These results cannot be accounted for by either paradigmatic or the syntagmatic contrast enhancement account. Not only does shorter VOT in voiceless stops make voiceless stops phonetically closer to voiced stops, it also reduces the sonority contrast between voiceless stops and adjacent vowels.

To account for these results, T. Cho and McQueen (2005) proposed an explanation based on phonetic feature enhancement. Phonetic features, represented with curly brackets '{}', were proposed in Keating (1984, 1990a) to capture cross-linguistic differences in the implementation of a single phonological feature. While the phonological feature [voice] is required to represent the phonological contrast between voiced and voiceless stops within a language, the phonetic feature {spread glottis} can describe the differences between languages. In languages such as Dutch, voiceless stops are unaspirated and are represented as {−spread glottis}. In contrast, English voiceless stops are produced with aspiration, and are thus represented as {+spread glottis}. According T. Cho and McQueen, this could explain why Dutch shows a different strengthening pattern compared to English; it is not the phonological feature [−voice] but the phonetic feature {±spread glottis} that is enhanced by DIS. In Dutch, {−spread glottis} is enhanced in a higher prosodic position, giving shorter VOT, whereas in English, {+spread glottis} is enhanced, which results in longer VOT.

For the voiced stops, T. Cho and McQueen (2005) suggested that the pattern of longer prevoicing in stronger positions can be seen as enhancing the phonetic feature {+slack vocal folds}[12]. However, although their results clearly showed that longer prevoicing was found in stressed and accented positions, recall that there was no consistent effect of prosody, with Big Phrase showing the shortest VOT. The authors speculated that, aerodynamically, the initiation of voicing after a pause might require more time. Thus, more prosodic levels need to be investigated in order to determine if the Big Phrase is indeed an exception to a general cumulative pattern.

The literature review in this section has highlighted the inconsistency in the accounts of DIS proposed for different languages as well as the lack of discussion on the issue. Some accounts assume privative features, while others do not. Some accounts are based on phonological or phonetic feature enhancement (i.e. paradigmatic contrast enhancement), some on sonority contrast enhancement (i.e. syntagmatic contrast enhancement), and yet others use a combination of these. While there is no doubt that there is substantial cross-linguistic variation, we should nevertheless seek theoretical consistency in accounting for this variation. In addition, unconstrained application of the two enhancement mechanisms (i.e. syntagmatic vs. paradigmatic) leads to a theory of DIS that makes neither falsifiable claims nor prediction for unseen cases. Such a theory cannot make falsifiable claims because a combination of these two mechanisms can account for DIS in any direction, e.g. increase and/or decrease in VOT. It cannot generate predictions either because there is no mechanism

[12] In Halle and Stevens (1971), the phonological features [stiff vocal folds] and [slack vocal folds] were proposed to characterise the relative degree of vocal fold tension, whereas [spread glottis] and [constricted glottis] were proposed for relative vocal fold abduction.

to predetermine whether a given language will show syntagmatic or paradigmatic enhancement, or even a mixture of both.

(2) Can a single account describe all the cross-linguistic differences?

As a potential solution to the issues discussed in (1) including theoretical inconsistency and unfalsifiability, we can examine if any one of the accounts proposed so far can adequately describe the full range of cross-linguistic variation. It is obvious that the explanation based on privative features – as proposed for Korean (T. Cho and Jun, 2000) – cannot be used to account for all the data. While the English case fits this explanation, the pattern in Dutch does not. For English, one can assume that /p, t, k/ are represented with [spread glottis] and that /b, d, g/ are unspecified. Then, the pattern that VOT is increased in a higher domain can be interpreted as enhancement of the feature [spread glottis], whereas less voicing in /b, d, g/ can be explained in terms of syntagmatic contrast enhancement. For Dutch, assuming that /b, d, g/ are represented with the privative feature [voice], it makes sense that prevoicing increases in a higher domain. However, /p, t, k/, presumably unspecified for [voice], show shorter VOT, when the privative feature-based account would predict that they show syntagmatic enhancement through longer VOT.

Next, we can examine the account based on enhancement of binary phonetic features, as initially proposed to account for the difference between Dutch and English (T. Cho and McQueen, 2005). Table 2.4 shows the specifications of four phonetic features relevant for laryngeal setting for the four languages in question (Halle & Stevens, 1971; Iverson, 1983). T. Cho and McQueen suggested that DIS enhances {+spread glottis} in the English voiceless stops but {−spread glottis} in the Dutch voiceless stops. For the voiced stops, {+slack vocal folds} is proposed as a feature that is strengthened in Dutch, resulting in longer prevoicing in higher domains. On the other hand, for English, one could interpret shorter prevoicing as a result of the enhancement of {−slack vocal folds}. For Taiwanese, the finding that VOT in the aspirated stops becomes longer in higher positions could be explained by the enhancement of {+spread glottis}, and prevoicing in the voiced stops becoming longer could be explained by the enhancement of {+slack vocal folds}. Perhaps a lack of DIS effects for the unaspirated voiceless stops could be interpreted as a result of the enhancement of both {−spread glottis} and {−slack vocal folds}, whereby the effects are cancelled out (c.f. Iverson, 1983, pp. 338-339).

So far, the phonetic-feature-based account described above can be criticised on several grounds. First, it is unsatisfying that negatively valued features can be enhanced given that, for example, {−spread glottis} is construed as a lack of a glottal spreading gesture rather than an active closing of the glottis. If {−spread glottis} was equivalent to {+constricted glottis} and {−constricted glottis} the same as {+spread glottis}, the feature combination {−spread glottis, −constricted glottis} would not be possible as spread glottis and constricted glottis cannot occur at the same time.

Table 1.4: Phonetic feature specifications for Taiwanese, English, Korean, and Dutch (c.f. Halle & Stevens, 1971; Iverson, 1983). The shaded boxes indicate the features that are proposed to be enhanced by DIS. For /t*/ in Korean, two different sets of features can be claimed to be enhanced depending on how the results from T. Cho and Jun (2000) are interpreted.

	Taiwanese			**English**		**Korean**			**Dutch**	
	/tʰ/	/t/	/d/	/t/	/d/	/tʰ/	/t/	/t*/	/t/	/d/
{spread glottis}	+	−	−	+ [13]	−	+	−	−	−	−
{constricted glottis}	−	−	−	−	−	−	−	+	−	−
{stiff vocal folds}	+	+	−	+	−	+	−	+	+	−
{slack vocal folds}	−	−	+	−	−	−	−	−	−	+

Second, the choice of features that are enhanced appears to be inconsistent. If the Taiwanese unaspirated voiceless stops are unaffected by DIS as a result of enhancement of both {−spread glottis} and {−slack vocal folds}, why is {−spread glottis} enhanced but not {−slack vocal folds} in the Dutch unaspirated voiceless stops? The same problem is encountered when we consider Korean. While longer VOT in the aspirated stops is easily accounted for by the enhancement of {+spread glottis}, one has to assume enhancement of {−constricted glottis} and/or {−slack vocal folds} for the lenis stops to explain longer VOT, ignoring the contradictory {−spread glottis}. On the other hand, yet another set of features {+constricted glottis} and/or {+stiff vocal folds} has to be enhanced for the fortis stops to show shorter VOT. Alternatively, given the lack of a statistically significant effect for the fortis stops, the same logic that was used to explain the lack of an effect in the Taiwanese unaspirated stops can be applied here; enhancement of both {−spread glottis} and {−slack vocal folds} results in no overall effect. Either way, it is evident that the choice of feature(s) across segments and languages is arbitrary. However, as will be discussed in Section 2.4.3, the arbitrariness in the choice of features may have a physiological explanation (see Footnote 15) and thus may not present a significant challenge to the account.

[13] According to Halle & Stevens (1971), the phonological feature [spread glottis] is not contrastive for English obstruents. Note that these values for the phonetic feature {spread glottis} are proposed to account for the differences in the behaviour of English and Dutch [−voice] stops.

The discussion in this section has demonstrated that the phonetic characteristics of DIS exhibit a considerable degree of cross-linguistic variation, especially for certain phonetic parameters including VOT. The accounts reviewed here work well within the particular languages they target, but often do not extend beyond those languages. In Section 2.4, I will review more general accounts for the underlying mechanism of DIS, with the ultimate goal of understanding the nature of DIS and working towards a unified theory of DIS across different languages. Before moving on to Section 2.4, I will briefly touch upon the topic of individual variation in DIS, as it is relevant for the discussion of the nature and the grammatical status of DIS.

2.3.3 Variation by individual

It is important to note that the findings summarised above for each language are either average values or patterns consistently found across all subjects for a given language. Languages show different degrees of between-speaker variation, and this also depends on the phonetic parameter and segment in question. For example, of the two Taiwanese subjects in Keating et al.'s study (2004, p. 158), one subject distinguished all five prosodic domains for /t/ and four domains for /n/ in terms of linguopalatal contact; the other subject distinguished four domains for /t/ but only two domains for /n/. In contrast, the three Korean subjects behaved similarly with each other (p. 156). In the same vein, Fougeron and Keating (1997) reported that although all three subjects (of English) made three or four pairwise cumulative distinctions by linguopalatal contact, no single pairwise distinction was made by all speakers. Even within a single individual's speech, the realisation of a segment at a given prosodic position could vary, although the details of such within-speaker variation are omitted in most studies. I informally observed in my study (Yoo, 2015b) that the phonetic form of a nasal phoneme /n/ in AP-initial position could range from a sonorant nasal [n] to a denasalised voiceless plosive [t] within an individual's speech. Thus, such extensive variation across individuals is another important characteristic of DIS to bear in mind when discussing the nature and the mechanism underlying DIS.

2.4 The nature and status of domain-initial strengthening

2.4.1 Is DIS optional?

The within-speaker variation mentioned in the previous section may be interpreted to suggest that DIS is (a) optional, in the words of Fougeron (2001, p. 126), or (b) obligatory with a wide window of articulatory targets on a given phonetic parameter, adopting Keating's window model (1990b). While the term *optional* is often used to describe a property of a categorical rule, there is no evidence or mention of categoricity of the phenomenon in

Fougeron's study. Thus, it is assumed that Fougeron's description of DIS as optional does not imply its categoricity. If so, there is little empirical difference between the seemingly contrasting interpretations in (a) and (b). For example, assuming that DIS is a gradient process which applies to the domain-initial segment /n/ and yields [n ∼ tʻ] including the canonical form [n], it is not possible to empirically test if this process is optional or obligatory.

2.4.2 Is DIS directly mediated by the prosodic structure?

Then, a more substantial question is how the variation and inconsistency observed in the DIS patterns are to be accounted for. Is DIS a direct function of the size of the prosodic domain? On the one hand, there is a cross-linguistic tendency for cumulative strengthening which is consistent with the prosodic hierarchy. Yet, on the other hand, there is a great amount of variability in how a speaker realises a phoneme in a given prosodic position, and how strictly the cumulative pattern is adhered to, e.g. AP-initial segments can be longer than IP-initial segments in French (Fougeron, 2001, p. 123).

There are at least three different ways to interpret this. One is to say that DIS is directly determined by the prosodic position of a segment but that the articulatory target of a segment in a particular prosodic position is set in terms of a range rather than a fixed point, as posited in window models (Keating, 1990b; Guenther, 1995). For example, in a given speaker's system, U-initial /n/ may be assigned a range of targets [d ∼ tʻ] which includes the maximally strengthened realisation, while AP-initial /n/ may have a range that is less extreme [n ∼ t]. Then, PW-initial /n/ may have an even more restricted window around the canonical form [n] with a varying degree of linguopalatal contact depending on the degree of lenition. In such a system, a violation of a strict cumulative pattern would be possible due to the overlap in the ranges of possible articulation. These windows would vary from one speaker to another, but it is also possible that a language as a whole may show a more or less consistent pattern depending on how wide the windows are and the phonetic distance among the windows.

Within a set window of articulation, various factors may play a role in determining the exact realisation. Rate of speech, prominence, desire for intelligibility, and style are often discussed as some of the factors which affect prosodic phrasing of a given utterance (S.-A. Jun, 1993, Section 5; Levelt, 1989, Section 10.2.3; Nespor & Vogel, 1986, Section 7.2). It is possible to imagine that, even after these factors have influenced the prosodic structure, they may still affect the degree of DIS gradiently. Let us assume that slower speech has resulted in greater DIS by dividing up a given utterance into more ip's, thereby creating more ip-initial segments which would otherwise be PW-initial. There may still be a scope of variation for these ip-initial segments to be more or less strengthened based on the exact speech rate.

A second possibility is that DIS is not a function of prosodic constituents but instead a function of various phonetic correlates of prosodic constituents. Prosodic constituents are defined in terms of the prominences and strength of the disjuncture at their edges. For example, an ip has at least one pitch accent and a phrase tone which affects pitch from the last pitch accent to the boundary. An IP is made up of one or more ip's and is distinguished from an ip by having a final boundary tone. While both an ip and an IP boundary mark a disjuncture through tonal patterns and phrase-final lengthening, an IP boundary shows these to a greater extent, representing a greater disjuncture (Break Index 4) than an ip boundary (Break Index 3) (Beckman & Ayers, 1997). Therefore, although the ip and the IP are distinct in theory, the phonetic features which distinguish them are gradient.

Then, the inconsistency of DIS patterns can be explained if DIS is not conditioned directly by prosodic structures but by one or more gradient phonetic parameters correlated with prosodic boundary strength, such as the duration, pitch, and amplitude of the segments at the edges of the boundary and/or the duration of the pause, if applicable. For expository reasons, let us say that the DIS of a segment is conditioned by the duration of the preceding syllable in addition to any silent pause. This would lead to the roughly cumulative pattern reported in the literature because, on average, IP-final segments show longer duration than ip-final segments, and so on. But the patterns will be gradient rather than categorical, and not always reliably distinguishable between all prosodic levels. However, this particular hypothesis encounters the issue of how to compute the degree of DIS in U-initial segments[14] which are preceded by a pause that may be indefinitely long. It may be that in such cases, the maximum degree of DIS applies for a segment preceded by a pause over a given durational threshold. Speculatively, one could imagine a scenario in which the threshold for applying a maximum degree of DIS is sufficiently low, i.e. the duration for accepting the presence of a pause is short enough, such that both the pause before the U and the pause before the IP exceed this threshold. Such a mechanism would explain why we sometimes observe a lack of a distinction between the U and IP, e.g. for French and Japanese (Keating et al., 2004, pp. 154-155; Onaka, 2003). Perhaps a more critical problem with this explanation is that it is not grounded in phonology or phonetics. Neither does DIS make use of the prosodic categories nor does it make obvious phonetic sense as in the physiology-based explanations discussed below.

A final possibility is that DIS can be explained by a purely physical and physiological mechanism of articulation which is only indirectly related to prosodic encoding of utterances.

[14] While the status of the U in the prosodic hierarchy is disputed, some languages clearly show a greater degree of DIS in U-initial position than in IP-initial position, e.g. Korean (Keating et al., 2004, p. 156).

The four different mechanisms suggested by Fougeron and Keating (1997, p. 3737) are summarised and discussed below.

(1) Increased duration

Given that an articulatory undershoot can be caused by shorter durations, longer durations of domain-initial segments may allow articulatory targets to be fully achieved. While this is a plausible explanation, empirical evidence for this is mixed. For English, Fougeron and Keating (1997) only found a weak correlation between DIS, measured by linguopalatal contact, and the acoustic duration of the domain-initial segment, although the latter also showed a cumulative pattern like linguopalatal contact. On the other hand, for Korean, there was a high correlation between linguopalatal contact and duration, acoustic or articulatory (T. Cho & Keating, 2001). This led the authors to propose that "strengthening and lengthening is a single effect in Korean" (p. 155). Crucially, however, this does not explain why initial nasals in Korean lose nasality in high prosodic domains. It is phonetically and phonologically arbitrary to argue that the articulatory target of a nasal phoneme is a non-nasal. Thus, there must be more to strengthening than lengthening and simply approaching an intended articulatory target.

(2) Increased distance between segments

It has been shown that, in a VCV sequence, the height of the jaw position for a high consonant is even higher when surrounded by low vowels (Keating et al., 1994). This was suggested to be a result of an overshoot due to the large displacement between the articulatory targets, which may also entail a higher velocity for articulatory movement (Imagawa, Kiritani, Masaki, & Shirai, 1985). Fougeron and Keating speculated that the same physiological mechanism may be responsible for DIS. This was based on the finding that, in the context of /...no#no…/, the difference in linguopalatal contact between /n/ and /o/ was greater when the context involved a stronger prosodic boundary, both for V1-to-C2 and for C2-to-V2. Therefore, the mechanism behind a greater linguopalatal contact for a domain-initial /n/ can be explained as follows. Because a domain-final /o/ is more open than a domain-medial /o/, there is a larger displacement between the articulatory targets for the domain-final /o/ and the following domain-initial /n/. The increased distance between the targets may cause faster articulatory gestures, which may then lead to an overshoot of the domain-initial /n/.

This explanation predicts a negative correlation between the contact of V1 and C2 and between the contact of C2 and V2. The authors found such a pattern for one of their three subjects, for whom C2 showed most contact when V1 showed least contact and, in turn, V2 showed least contact when C2 showed most contact. It also predicts that DIS will be weaker in a high vowel environment. However, the prediction has not yet been tested as most existing studies on consonant DIS do not systematically vary the vowel context. In addition, further research is needed to examine strengthening of initial consonants that are preceded by a non-vowel segment.

Importantly, however, this explanation appears problematic for the findings on reduced nasality in domain-initial nasals. If DIS is a result of a large articulatory displacement among adjacent segments, initial nasals should be more nasal around oral vowels. Of course, one could argue that the adjacent targets of the tongue and the velum positions need not be (co-)articulated in the same way. In the absence of phonologically nasal vowels in Korean, nasalisation of an oral vowel in the vicinity of a nasal consonant may be largely tolerated. On the other hand, the jaw raising gesture for the tongue tip and blade constriction in /n/ and the jaw lowering gesture for the much weaker degree of tongue dorsum constriction in /o/ are in direct opposition and coarticulation of these two gestures may be tolerated to a lesser extent perhaps due to the existence of confusable phonemes (e.g. back vowels that are less open than /o/). Thus, the large displacement between the tongue positions of the two segments may trigger extra tongue activation, leading to an overshoot in terms of the tongue gesture but not of the velum gesture (see Fowler and Saltzman, 1993, p. 174, on extra gestural activation in compensatory responses to perturbation in speech production experiments).

While it should not be assumed that a single mechanism is solely responsible for all the observed DIS effects, it is clear that this mechanism alone is insufficient. For one, it cannot explain reduced nasality in nasal segments for which lowered velum is one of the articulatory targets. In addition, an independent explanation is required for the cumulative nature of DIS. That is, the current account assumes, rather than explains, why a larger displacement is observed across a higher domain boundary. If a larger linguopalatal contact in domain-initial /n/ is due to an overshoot after a more open /o/, a further explanation is necessary as to why domain-final vowels are more open in the first place.

(3) Greater coarticulatory resistance

Following Fowler and Saltzman (1993, p. 182), Fougeron and Keating also suggested the possibility that segments at or near the edges of prosodic domains have greater coarticulatory resistance and thus show less coarticulatory blending with adjacent segments. It is posited that the components of a given articulatory gesture are assigned their own characteristic degrees of *blending strength*, which is a concept that combines *coarticulatory resistance* and *coarticulatory aggression* of phonetic gestures. Vowels are said to have the weakest blending strength and the stops, the greatest. As a result, vowels are easily influenced by neighbouring segments due to low coarticulatory resistance and are less likely to influence neighbouring segments due to low coarticulatory aggression. Stops show the opposite pattern. Fougeron and Keating (1997) speculated that greater linguopalatal contact in domain-initial /n/ and less contact in domain-final /o/ are due to increased coarticulatory resistance in the segments. While this explanation holds for some effects of DIS, it is inadequate for others. Greater VOT in initial plosives may be explained as a result of coarticulatory resistance with adjacent vowels which require the vocal folds to be held together. On the other hand, reduced nasality is again problematic, since, if anything, nasals should be more nasal and oral vowels should be less nasal if coarticulatory resistance is increased.

(4) Increased effort or energy

Finally, DIS has been proposed to result from a greater global muscular effort or energy in domain-initial positions in speech production, affecting the pulmonary, laryngeal and supralaryngeal systems (Fougeron, 2001, p. 131; Fujimura, 1990, p. 233). Alternatively, articulatory strengthening may be restricted to specific muscles involved in a given articulation. The definition of *articulatory force* suggested by Straka (1963, p. 91) is translated as the following by Fougeron (2001, p. 132): "nothing but the contraction strength of the muscles involved in a given articulation". Fougeron interpreted this to mean that the force of only the specific articulatory movements is increased, which excludes phonatory and expiratory forces.

A clear advantage of adopting the broader, global definition of articulatory effort is that it can account for the effects of DIS on source intensity and possibly F0 as a result of increased abdominal and/or thoracic activity in exhalation, as proposed for stress (Stetson, 1951; Ladefoged, 1967; Sluijter, van Heuven, & Pacilly, 1997). There are findings of increased acoustic amplitude during the stop burst and in the following vowel in initial CV syllables (T. Cho et al., 2007, p. 220). Although research on the effect of DIS on F0 is limited, it is possible that increased subglottal pressure leads to higher F0 in domain-initial segments. Modification of laryngeal articulation has also been reported, such as increased VOT and glottal constriction at the beginning of phrase-initial vowels, which may also be explained as a result of greater muscular tension in the larynx.

Fougeron (2001, p. 132) explored how a range of different DIS effects can be accounted for based on the hypothesis that DIS is simply caused by extra contraction of muscles involved in the articulation. Greater linguopalatal contact in consonants would be explained by an increased contraction of the elevator muscles of the tongue, which are genioglossus, palatoglossus and styloglossus. The decrease in nasal flow in initial nasals and the higher velum position for initial oral consonants (Krakow, Bell-Berti, & Wang, 1995) could be a result of greater contraction of the levator palatini which elevates the velum (Bell-Berti, 1993). Finally, a hard onset to phonation frequently observed in initial vowels could be caused by an increase in activation of the adductor muscles, such as the lateral cricoarytenoid, interarytenoid, and vocalis muscles (Faaborg-Andersen, 1957; Ohala & Hirano, 1967; Gay, Strome, Hirose, & Sawashima, 1972; Ludlow, Sedory, & Fujita, 1991).

While Fougeron (2001) adopts the narrow definition of articulatory effort, the definition can easily be extended to increased muscular activity in the respiratory system to account for increased amplitude, as mentioned above. Note that there are two senses in which there is greater respiratory power at the beginning of an exhalation phrase. There is naturally a higher pressure in the lungs at the start of exhalation as the diaphragm moves upward and compresses the chest cavity. This would cause the air to be pushed out with greater force and result in greater oscillation of the vocal folds compared to a later point in the exhalation, during which the air pressure gradually decreases. As T. Cho (2016, p. 9) explains, a larger prosodic domain serves as a breath group and there is augmented respiratory power at the beginning of a new phrase as the speaker resets a respiration cycle. This may be at least partially responsible for some of the DIS effects. Of course, not all beginnings of major prosodic domains coincide with the beginning of exhalation, and prosodic boundaries of the same strength can occur in different points in the course of a single exhalation. This sentence-level "declination" effect of serial position (Krakow et al., 1995; Vayra & Fowler, 1992) has not been found in the DIS literature (Pierrehumbert & Talkin, 1992; Byrd, 1994; Fougeron & Keating, 1997). In addition, the effect of DIS is known to be strictly local at least in some cases (e.g. S. Kim & Cho, 2012), in which only the very first segment after a boundary is affected. Thus, a greater subglottal pressure in domain-initial position must also be reinforced by greater contraction of intercoastal and/or abdominal muscles that is independent from the activity of the diaphragm.

2.4.3 What is the grammatical status of DIS?

The different kinds of explanations offered so far lead to the question of what the grammatical status of DIS is. The literature reviewed in Section 2.3 consistently suggests that DIS is a gradient phenomenon. On the other hand, the exact location of this phenomenon within the architecture of grammar presented in Section 1.2.2 requires further discussion. Figure 2.3 has been copied from Section 1.2.2 to remind the reader of this architecture of grammar and the theory of the life cycle of phonological processes.

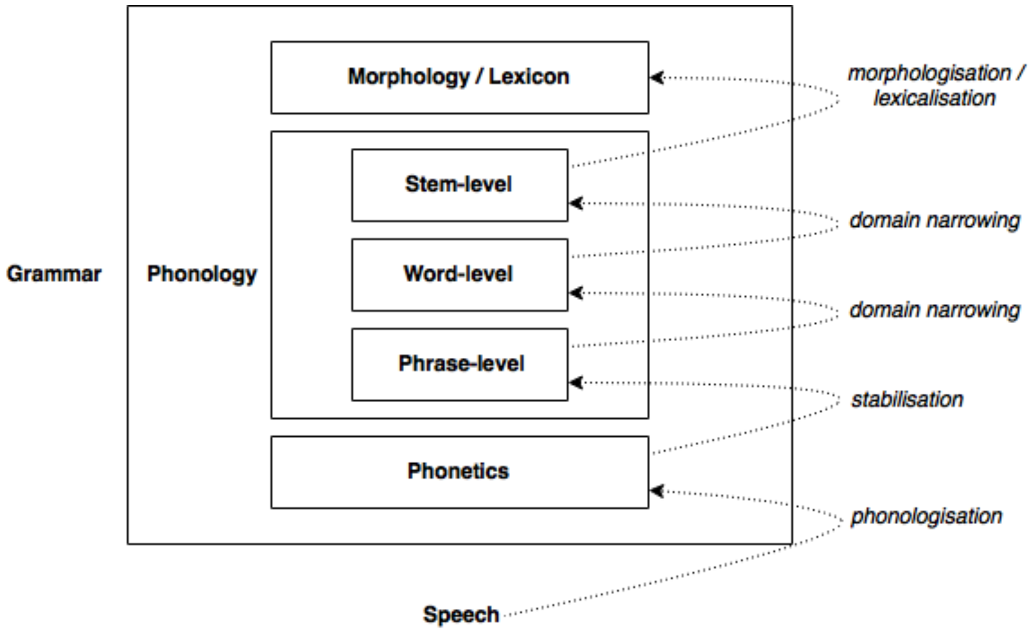


Figure 2.3: The life cycle of phonological processes (Bermúdez-Otero & Trousdale, 2012).

The physiological explanations in (1) – (4) above assume that the mechanism of DIS lies outside the language-specific grammar, indicated by "Speech" in Figure 2.3. According to these explanations, DIS is an automatic consequence of physical and physiological limitations of speech production that are universal to all languages. On the other hand, the view that DIS is directly mediated by the prosodic structure supposes that DIS resides within the language-specific phonetic component, because the prosodic structure is a language-specific linguistic construct. This component is labelled "Phonetics" in the figure. Given the postulations of the life cycle theory, these two views are not mutually exclusive. The physiological explanation may reflect the origin of DIS, whereas the explanation based on the prosodic mediation attempts to capture how DIS is represented in the grammar after phonologisation. In the rest of this section, I will argue that the combination of these two explanations provides an account of DIS superior to the abstract end-based explanations reviewed in Section 2.3, i.e. syntagmatic and paradigmatic contrast enhancement.

The strength of the physiological explanations in (1) – (4) is that they can capture the similarities in the DIS patterns of typologically distinct languages. The last explanation in (4) is particularly appealing in that various kinds of DIS effects are elegantly accounted for by a single type of process, although much empirical work is still needed to clarify the definition of articulatory force. However, one remaining challenge is to account for the cross-linguistic differences discussed in Section 2.3.2. One possible solution can be found in the diachronic course of speech processes. That is, languages may have phonologised the purely automatic consequences of speaking into the language-specific phonetic component in their own distinctive ways. Grammaticalisation of biologically determined conditions is common and well-known, such as the hypothesis of three biological codes (Gussenhoven, 2004; Ohala, 1983).

An alternative solution is to argue that different languages realise a given category of sound differently and, thus, the result of stronger muscular contraction in the articulation of a particular segment may be different cross-linguistically. Similarly, a particular phonetic realisation may be achieved with different articulatory strategies, especially given the anatomical differences among individuals and the contextual effects from neighbouring segments. Thus, not only would this account for the cross-linguistic differences but also for the between- and within-speaker variation. It should also be noted that this explanation is perfectly compatible with the first explanation based on the diachronic processes. In fact, this may partly explain why languages have phonologised the universal physiological effect of DIS in non-uniform ways.

Recall the findings regarding the cross-linguistic differences in the variation of VOT. Dutch voiceless stops show shorter VOT in a higher domain, in contrast to the pattern in English. This was interpreted to be an enhancement of the phonetic feature {−spread glottis} in Dutch, and the enhancement of {+spread glottis} in English. However, an alternative analysis can be proposed in terms of increased muscular tension involved in two different gestures in the stop articulation. Using /t/ as an example, the two relevant gestures are the constriction of the tongue against the palate and the glottal abduction. Let us assume that extra force in the lingual articulation causes a longer stop closure and leads to a delay in the timing of the stop release, in addition to a greater amount of linguopalatal contact. Extra activation of the posterior cricoarytenoid muscle in the glottal abduction gesture (Hardcastle, 1976, p. 76) can be said to cause a delay in the voice onset in the following vowel.

Then, one could explain the differences between English and Dutch as a result of extra activation of the glottal abduction gesture occurring in English but not in Dutch voiceless stops, while the lingual gesture is strengthened in both languages. Given that acoustic closure duration was lengthened in higher prosodic positions in Dutch (T. Cho & McQueen, 2005, p. 134), shortened VOT can be seen simply as a consequence of a delayed stop release caused by strengthening of an oral constriction, with more or less the same timing of the voice onset. On the other hand, English stops show longer acoustic closure as well as longer VOT domain-initially, indicating a delayed stop release and a delayed voice onset due to extra activation of both lingual and glottal gestures. Speculatively, a possible explanation for why the two languages show such differences in the glottal gesture under DIS may be as follows. Dutch voiceless stops may be articulated via relaxation of the laryngeal muscles to a resting position, i.e. a moderately open configuration. On the other hand, English voiceless stops may involve an active contraction of the posterior cricoarytenoid muscle, which would then be subject to extra activation under DIS. This could be empirically verified, for example, by measuring voicing residue in the beginning of the stop closure at different prosodic positions in the two languages.

The VOT variation in Taiwanese and Korean can also be explained this way. It was found that all three contrastive stops in Taiwanese showed increased closure duration, whereas VOT was affected in different directions: greater prevoicing in voiced stops, no modification in unaspirated stops, and longer aspiration in aspirated stops (Hsu & Jun, 1998). First, this can be interpreted to mean that all three stops showed a robust effect of stronger lingual articulation in higher positions. However, glottal articulation is affected in different ways as different muscles are involved in the voicing contrast. In voiced stops, it would be the adductor muscles such as the lateral cricoarytenoid muscle that are further activated, increasing prevoicing in higher positions. The lack of VOT variation in unaspirated stops may be because, although abduction gesture is enhanced resulting in a delayed onset of voicing, this is compensated by a delayed stop release, leaving VOT more or less unaffected. Finally, longer VOT in aspirated stops can be explained by an even greater glottal abduction gesture, which delays voicing onset longer enough that the effect on the VOT remains after some of it has been cancelled by a delay in the stop release.

As in Taiwanese, all three Korean stops showed longer duration in a higher prosodic position. VOT of the aspirated and lenis stops was increased in higher positions, whereas VOT of the fortis stops remained unchanged with a trend of slight decrease in higher positions (T. Cho & Keating, 2001). As mentioned earlier, it was also shown that the fortis stops showed smaller glottal opening AP-initially than AP-medially (S.-A. Jun et al., 1998). Given these findings, longer VOT in the aspirated and the lenis stops in higher positions can be explained by a stronger glottal abduction gesture. For the fortis stops, it is not surprising that VOT is not primarily affected by DIS, as the muscles that are known to be activated are the vocalis and the lateral criocoarytenoid muscles (Hirose, Lee, & Ushijima, 1974). The former functions to tense and stiffen the vocal folds and the latter to draw the vocal processes together.

The explanation based on increased muscular activity (with the possibility of additional phonologisation) offers a more convincing account of the universal and language-specific aspects of DIS than the abstract end-based explanations of syntagmatic and paradigmatic contrast enhancement. As discussed towards the end of Section 2.3, the contrast enhancement accounts face various criticisms. First, the effects of syntagmatic and paradigmatic contrast enhancement can be contradictory. In the Korean case, making a nasal consonant less sonorous through denasalisation and devoicing would enhance its contrast with the following vowel and possibly contribute to prosodic boundary marking; at the same time, however, it would threaten its paradigmatic contrast with the oral stops, which may have a negative effect on lexical access. In addition, such an approach fails to provide falsifiable predictions as to how a language will behave with regard to DIS; it is arbitrary to argue that Korean enhances syntagmatic contrast and sacrifices paradigmatic contrast whereas Taiwanese does exactly the opposite, unless a further explanation can be provided for their differences. Lastly, the explanations of this type have been shown to be descriptively inadequate in light of the Dutch data[15].

In contrast, the physiology-based explanation provides a more elegant account of the empirical findings, including the similarities and differences observed across different languages and segments, as well as intra-and inter-speaker variation. The similarities arise from the fact that the physical and physiological effects in speaking apply universally, while the differences are a result of different muscular activities involved in the articulation of a given segment. The mechanical, extra-grammatical phenomenon may later become or have already become phonologised and even stabilised (see Section 1.2.2). If there is indeed evidence of phonologisation and/or stabilisation, a formal account may then be proposed based on phonetic or phonological feature enhancement. However, when such an account is formulated, empirical evidence should inform which features are claimed to be enhanced, irrespective of what end-goal, e.g. syntagmatic/paradigmatic contrast enhancement, is being served.

We can contrast the two types of explanation using an example drawn from T. Cho's (2005) study of articulatory strengthening in English. In the context /bi#bi/, /i/ was found to show an increase in lip opening and a decrease in jaw opening in domain-initial positions. According to the original explanation based on contrast enhancement, this is interpreted to be a result of both syntagmatic/sonority contrast enhancement by increased lip opening, and paradigmatic/featural enhancement of [+high] by decreased jaw opening. The physiology-based explanation would account for the same finding by the single mechanism of extra muscular activation of the lip opening gesture and that of the jaw raising gesture, the gestures that are involved in the articulation of /i/ in /bi#bi/.

While the physiology-based account is favoured based on its simplicity and concreteness, much work remains to be done to define precisely which muscles are activated

---

[15] However, the alternative analysis based on the phonetic feature enhancement (T. Cho and McQueen, 2005) could provide a good framework to represent which articulatory gestures have been phonologised to undergo DIS in the language-specific phonetic component.

through DIS. In the discussion so far, it has been implicitly assumed that the articulatory muscles targeted by DIS are those that are involved in a given articulatory gesture from a neutral position. For example, when the vowel /e/ is strengthened, it is expected to become fronter and higher assuming there is greater contraction of the muscles involved in the articulation of /e/ from the neutral position /ə/. This assumption would be met in cases in which the articulation indeed starts from a neutral position such as in U-initial position, or when preceded by /ə/. However, the muscular activity required to articulate initial /e/ when it is preceded by /i/, for example, would be very different, i.e. the tongue is now required to be lowered rather than raised. In this scenario, would strengthening of the initial /e/ cause it to be realised with a lower tongue position? To answer such a question, future research should aim to test relevant hypotheses by systematically varying the segmental context of target segments and taking into account the precise muscles involved in the articulation in each context.

2.5 Chapter summary

In the first part of this chapter, I have described the detailed phonetic characteristics of DIS and reviewed various accounts proposed for the phenomenon. There are general similarities in the handful of cases reported in the DIS literature, including articulatory gestures of greater magnitude and longer durations. However, the specifics of the exact DIS pattern vary considerably depending on the segment type, the language, and even the individual. We saw that various explanations have been suggested to account for the patterns in different languages, such as syntagmatic contrast enhancement, paradigmatic contrast enhancement, a combination of both types of enhancement using privative features, and phonetic feature enhancement. In addition to other problems, I have argued that these explanations fail to provide a consistent account of DIS in light of the cross-linguistic variation.

In the second part of the chapter, I have discussed higher-level issues regarding the nature and status of DIS. Among the four alternative underlying mechanisms proposed, the one based on increased articulatory effort appears to be most descriptively adequate. Furthermore, an account that involves a reference to the physiological origin of DIS is superior to the abstract end-based explanation of contrast enhancement as it offers a simple and concrete explanation that is grounded in the mechanical properties of speech production. Following the theory of the life cycle of phonological processes (Section 1.2.2), this physiological mechanism may provide a basis for DIS to develop into a language-specific phonetic phenomenon, and subsequently become stabilised as a phonological rule.

Chapter 3.

Seoul, Busan, and Ulsan Korean

3.1 Chapter overview

Chapter 3 will be about the three varieties of Korean spoken in Seoul, Busan, and Ulsan. While there are many interesting segmental and suprasegmental differences among the three varieties of Korean, the literature review will primarily focus on those details that are relevant for the design and interpretation of the experiments in Chapters 5 and 6. The experiments will investigate the segments /p, t, k, m, n/ in a V#**C**V(C) context where the symbol # indicates a prosodic boundary. Thus, this chapter will be concerned with the phonetic properties of these lenis and nasal stops, the processes which apply to these segments in the aforementioned context, and the suprasegmental cues to prosodic boundaries of different strengths.

Having presented the phoneme inventory earlier in Section 2.2, I will summarise the prosodic system (Section 3.2) of Standard Korean. Then, I will provide a brief introduction to the three cities and major dialectal differences among the varieties spoken there (Section 3.3). I will examine the phonetic characteristics and phonological processes related to the lenis stops first (Section 3.4), followed by those related to the nasal consonants (Section 3.5), noting the dialectal differences where possible. Then, I will discuss word-initial tensification in the relevant dialects as a process that is potentially relevant to DIS (Section 3.6). Finally, I will end this chapter with a short summary (Section 3.7).

3.2 The prosodic system

This dissertation adopts a phonological model of Korean Intonation as developed by S.-A. Jun (1993, 1998). A comprehensive summary of the prosodic system of Seoul Korean can be found in S.-A. Jun (2005a; a shorter description is available in S.-A. Jun, 2005b). In her earlier proposal, she posited two prosodic units above the Phonological Word (PW): the Accentual Phrase (AP) and the Intonational Phrase (IP). As with the Autosegmental-Metrical Model of intonation proposed by Pierrehumbert et al. (Beckman & Pierrehumbert, 1986; Pierrehumbert & Beckman, 1988; Pierrehumbert, 1980), S.-A. Jun's model makes the assumption that the prosodic units in Korean are organised in a hierarchy. Following the Strict Layer Hypothesis (Nespor & Vogel, 1986; Selkirk, 1984), a given unit of the prosodic hierarchy is assumed to be composed of one or more units of the immediately lower category. Furthermore, a unit of a given level is assumed to be exhaustively contained in the superordinate unit, such that the boundaries of these units would coincide. For Korean, this means that an IP is exhaustively parsed into one or more APs, and each AP is in turn exhaustively parsed into one or more PWs, and so on. A schematic representation of the prosodic hierarchy of Seoul Korean is reproduced in Figure 3.1, based on S.-A. Jun (2005a).

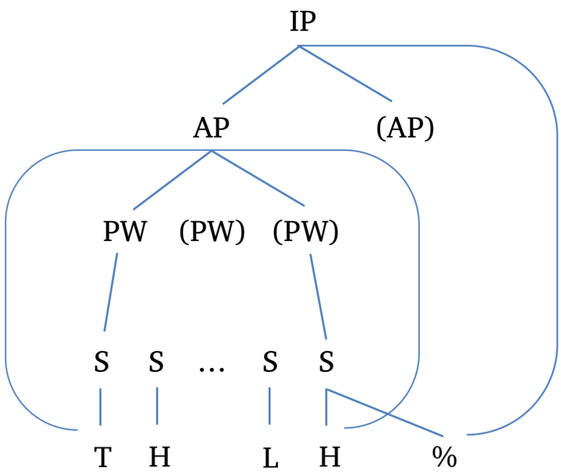


Figure 3.1: Prosodic hierarchy of Seoul Korean. The initial tone (T) is High (H) when the syllable-initial segment is aspirated, tense, /s/, or /h/; otherwise, it is Low (L), % indicates IP boundary tone (S.-A. Jun, 2005a, p. 205).

The IP in Korean is marked by phrase-final lengthening and a boundary tone (%) on the final syllable. The IP may also be followed by a pause. The AP is defined by an underlying tonal contour of either LHLH or HHLH. The initial tone is H when the AP begins with a segment that is aspirated, tense, /s/, or /h/, other it is L. The tonal contour is fully realised when there are four or more syllables in the AP. When there are fewer than four syllables, either the medial L, the medial H or both are undershot. When an AP is final to an IP, the IP boundary tone pre-empts the AP-final tone on the last syllable. Research shows that the AP-initial segment is slightly but consistently longer than the same segment in the AP-medial position (T. Cho & Keating, 2001; but see S.-A. Jun, 1995c). On the other hand, the AP-final segment is not necessarily longer than the AP-medial segment (Koo, 1986; S.-A. Jun, 1995c, 1996; T. Cho & Keating, 2001), and it is never followed by a pause unless it is also an IP-final segment. The PW in Korean often corresponds to a single word including its grammatical suffixes. It does not have any tonal specification, although it serves as one of the domains of articulatory strengthening as reviewed in Section 2.3.1.

S.-A. Jun's model of prosody was later revised to include the intermediate phrase (ip), between the IP and AP (S.-A. Jun, 2006). The ip is marked by a higher AP final boundary tone and/or an interruption of a pitch downtrend among the APs. Additionally, the ip-final syllable may also show a small degree of lengthening compared to the AP-final syllable. An ip is observed (a) when what would have otherwise been an AP receives focus or (b) at the edge of syntactically heavy constituents, e.g. a small clause or a heavy NP/VP. However, due to the difficulty of reliably demarcating ip boundaries in speech data, researchers often adopt the earlier model without the ip (e.g. Jeon & Nolan, 2017). The experiments in this dissertation will also be based on the earlier model shown in Figure 3.1.

Since our investigation is concerned with boundary-induced strengthening, the experiments will be designed to exclude focus-induced strengthening. Thus, the literature related to the effect of focus on prosody will not be reviewed here. Interested readers are directed to the following studies: Jeon & Nolan (2017); J. Jun, Kim, Kim, Lee, & Jun (2006).

3.3 The three varieties of Korean

Figure 3.2 shows the geographical locations of the three metropolitan cities, Seoul, Busan, and Ulsan, and their respective population size. Seoul is the capital and the largest city of South Korea with a population of over 9.7 million. Busan is the second largest with approximately 3.4 million inhabitants and is the economic, cultural, and educational centre of Southeastern Korea. Ulsan is the seventh largest metropolitan city with a population of approximately 1.1 million. Thus, in terms of the relative population size, Seoul is roughly three times bigger than Busan and, in turn, Busan is three times bigger than Ulsan (Korean Statistical Information Service, 2020).

The variety of Korean spoken in Seoul is recognised as the standard and prestige variety in South Korea. In "Rules of the Standard Language" (표준어 규정) issued by the South Korean Ministry of Education in 1988, the standard language is defined as "the speech widely used by people with education in present-day Seoul" (MOE, 1988, cited in I. Lee & Ramsey, 2000, p. 307). The most widely used dialect divisions to date are largely based on the extensive Korean dialect data collected by Ogura Shinpei in 1944 (I. Lee & Ramsey, 2000, p. 311; H.-M. Sohn, 1999, p. 57; H.-K. Kim, 1982; H.-K. Kim, 1972). A total of six major dialects is recognised throughout the Korean peninsula and four of them are found in South Korea: the Central, the Gyeongsang, the Jeolla, and the Jeju Dialects. Seoul Korean is part of the

Central Dialects that are further subdivided into the Gyeonggi, Chungcheong, Gangwon, and Hwanghae Dialects, which are coterminous with the political division after which they are named, like the other dialects of Korean. (The Hwanghae province is not shown on the map as it is entirely in North Korea.)

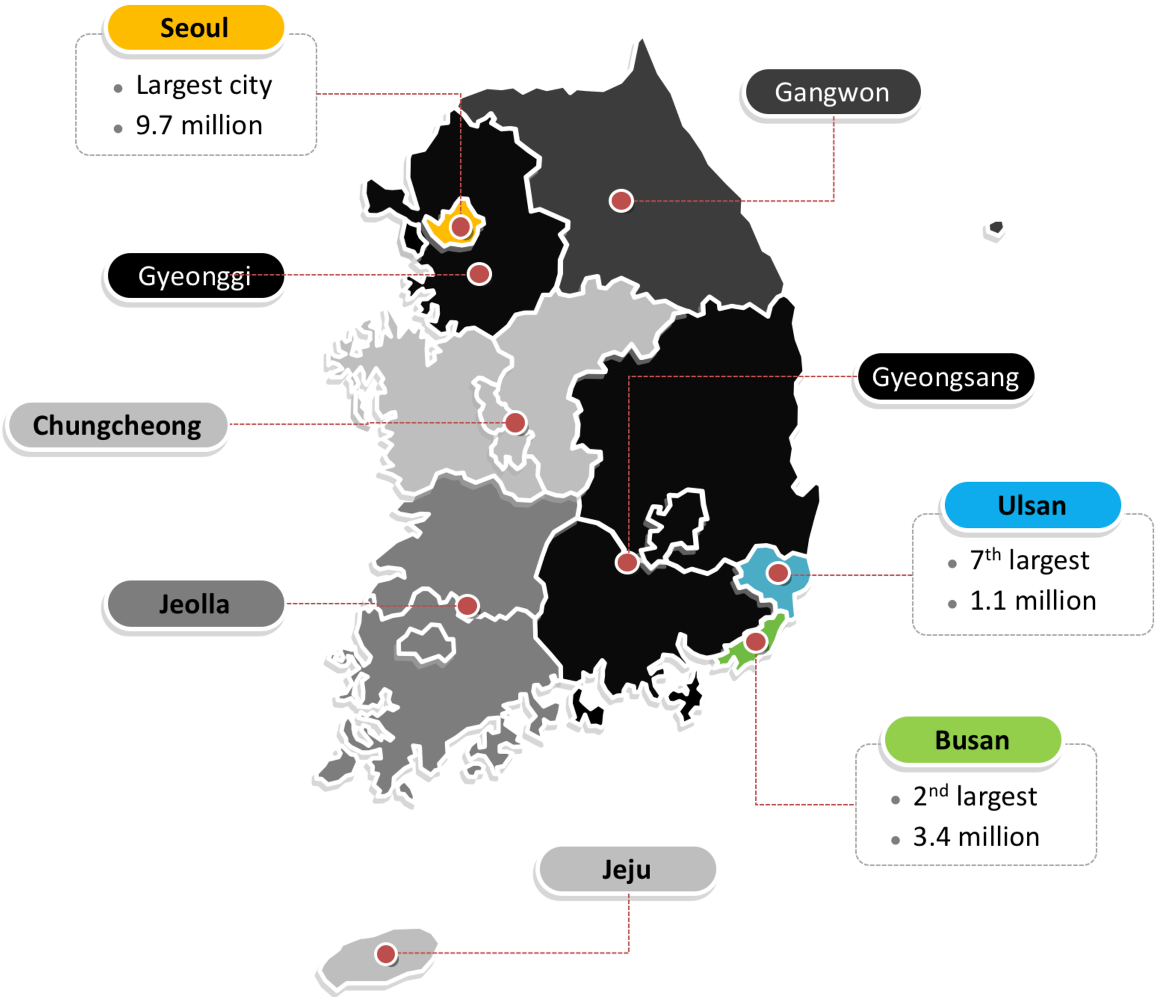


Figure 3.2: Map of South Korea. It shows the location and population of the three cities and the major provinces of South Korea which broadly correspond to the six dialectal divisions. The map has been modified using the template from Wowslides (2016) and the population data is based on Korean Statistical Information Service (2020).

Seoul Korean is grouped under Gyeonggi dialect and the dialectal differences between Seoul Korean and the rest of the area covered by Gyeonggi dialectal region are often considered negligible. One frequently discussed feature of the Gyeonggi dialect is that, in colloquial speech, the vowel /o/ is raised to /u/ in final syllables in such words as /kɯ.li.ko/ 'and', /na.to/ 'I also', /sam.t͡sʰon/ 'uncle', and /si.kol/ 'countryside'. However, it is questionable how well this serves to distinguish between the varieties of Gyeonggi and Seoul at present, as opposed to marking the differences between informal and formal speech in these areas and reflecting individual variation. Thus, it is not uncommon for even studies which investigate a phonetic variable in Seoul Korean to recruit participants from both Seoul and Gyeonggi, e.g. Silva (2006). However, as the experiments in this dissertation target very fine-grained and under-researched phonetic phenomena, I will use a precautionary principle and ensure maximum homogeneity by recruting participants who have lived in their respective cities for most, if not all, of their lives.

The other dialect of concern in this dissertation is part of Gyeongsang Dialects. The Gyeongsang Dialects can be classified into north and south following the provincial divisions. Both Busan and Ulsan belong to the South Gyeongsang province. The main phonological features of Gyeongsang dialects that are distinguished from the standard variety are summarised below (I. Lee & Ramsey, 2000, p. 324).

(1) A six-vowel system is found as opposed to a ten-vowel system traditionally proposed for Standard Korean due to the following: (a) the lack of a distinction between /ɯ/ and /ʌ/; (b) the lack of a distinction between /e/ and /ɛ/; and (c) the absence of the two front round monophthongs /y/ and /ø/. As mentioned earlier, (b) and (c) are also found in contemporary Standard Korean. Since (b) was found earlier in Gyeongsang Korean, some researchers suggest that the influx of Gyeongsang Korean speakers into Seoul may have triggered the relatively recent merger between /e/ and /ɛ/. With regard to (c), the front rounded vowels are replaced with /wi/ and /we/ in Standard Korean, but they are replaced with various other monophthongs in Gyeongsang Korean, e.g. /wi-e/ 'up-LOC' (originally /y-e/) in Seoul Korean shows up as /u-e/ in Gyeongsang Korean (Y.-Y Cho).

(2) Semivowels are lost after a consonant in words such as /p^{h}jo/ 'ticket', /sa.kwa/ 'apple', /kwi/ 'ear', and /hwak.sil.hi/ 'certainly'.

(3) There is no phonemic distinction between the fortis and lenis alveolar fricatives /s*/ and /s/ which are both pronounced [s].

(4) Palatalisation of /k/ and /h/ before /i/ or /j/, e.g. /ki.lɯm/ 'oil' → /t͡si.lɯm/[16], /kjʌt/ 'side' → /t͡sʌt/, /him/ 'strength' → /sim/.

[16] As with Standard Korean, /t͡s/ and /s/ are (more) palatalised before /i/ and /j/ than in other contexts.

In addition to these phonological differences, Gyeongsang and Seoul Korean also diverge on the morphosyntactic and lexical aspects, notably the unique set of sentence endings of Gyeongsang Korean, such as /no/ or /na/ endings for questions (for a summary, see H-.M. Sohn, 1999, pp. 70-71).

In terms of prosody, Gyeongsang Korean is a lexical pitch-accent language unlike Seoul Korean which lacks lexical pitch accent.[17] That is, tonal patterns in words are lexically contrastive in Gyeongsang Korean. A well-known example of a three-way lexical contrast is given in (4).

(4) a. ka.t͡si HL 'type' b. ka.t͡si HH 'branch' c. ka.t͡si LH 'eggplant'

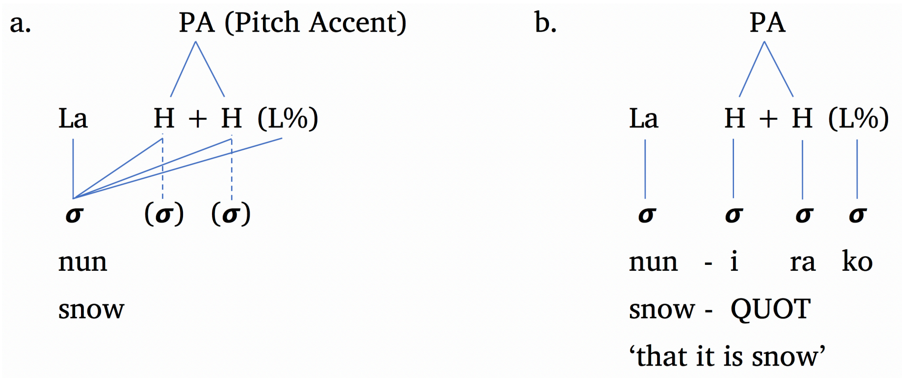
While there is some controversy regarding whether Gyeongsang Korean is a tone or a pitch-accent language, the balance of evidence seems to favour the view that it is a pitch-accent language (e.g. D. Lee & Davis, 2009). The strongest piece of evidence comes from findings of constraints on possible tonal contours over a prosodic word (PW). Unlike a tonal language which allows a free combination of tones from their tone inventory, distribution of tone is restricted in Gyeongsang Korean. For example, South Gyeongsang Korean allows neither PW-initial L sequences nor a rising pitch from L to H once there has been a falling pitch from H to L earlier in the word. According to J. Kim's model (J. Kim, 2008; J. Kim & Jun, 2009), South Gyeongsang Korean has two types of pitch accent, H+L and H+H, which can be anchored to either the first and the second syllables or the second and the third syllables. This gives four tonal contrasts, which is half of the eight logically possible contrasts if any sequence of H and L were possible. Thus, J. Kim claims that each word is lexically specified with the type of pitch accent as well as the location of the pitch accent. Apparent deviations from this simple analysis, such as a pitch rise or fall within a single syllable, are accounted for by boundary tones realised at the level of the AP or the IP; the beginning of an AP is marked by a L boundary tone (La), and the end of an IP is marked by a L boundary tone (L%) indicating a declarative sentence.

Although it is not specifically mentioned in J. Kim and Jun, what is often analysed as a Rising (R) tone in monosyllabic words such as /nun/ 'snow' R in South Gyeongsang Korean may also be analysed simply as La (AP initial boundary tone) + H (pitch accent) realised on the same syllable rather than as a third type of pitch accent. Due to the lack of a docking site for the pitch accent H+H, which is specified to be realised the second and the third syllables, the rising pattern is compressed into one syllable. Depending on whether /nun/ is realised as

[17] It is said that the presence of pitch accents/tones divide Korea into the east and the west, as pitch accents/tones are found in most of Hamgyeong in North Korea, all of Gyeongsang, and in parts of Yeongdong (eastern) area of Gangwon (I. Lee & Ramsey, 2000, p. 315).

a declarative utterance, it may also have a low IP boundary tone, giving a contour of LHL as shown in (5)a. When more syllables are added, each tone is fully realised as shown in (5)b. (For a full analysis of the pitch accent system of South Gyeongsang Korean and the interaction between pitch accent and focus, see J. Kim & Jun, 2009; for a summary of the alternative analyses of South and North Gyeongsang Korean, see H. Lee & Zhang, 2014, p. 102.)

(5)



As mentioned earlier, the varieties of Korean spoken in Busan and Ulsan are both classified as South Gyeongsang Korean. Although researchers disagree on the number and exact boundaries of subdialectal zones in South Gyeongsang, it is generally agreed that Busan and Ulsan Korean should be grouped into the same zone occupying the eastern half of South Gyeongsang. For example, J. D. Kim (2012) argues that Busan and Ulsan belong to the same subdialectal zone along with Changnyeong, Miryang, Yangsan, Changwon, and Gimhae, based on the phonological, grammatical, and lexical features.

Phonological and phonetic studies which specifically investigate Busan and Ulsan Korean are still relatively rare, although they are one of the better studied sub-varieties of Gyeongsang Korean (Kong, Kim, Yoon, & Maung, 2016, p. 248). Interestingly, however, a couple of studies note that Ulsan Korean shows pitch accent patterns which are slightly different from the rest of South Gyeongsang Korean. In an attempt to establish subdialectal areas within Gyeongsang, S.-K. Im (2014, p. 28) suggested dividing Gyeongsang into nine zones, with Ulsan Korean representing its own subdialect, separate from the variety spoken in Busan and the surrounding region in the east of South Gyeongsang. This is based on the finding that Ulsan Korean patterns with Busan Korean in terms of the pitch accents of some words beginning with an L tone, e.g. /mi.na.li/ LHL 'water parsley', but patterns with Daegu Korean in North Gyeongsang in other words beginning with an R tone, e.g. /t͡sak.eun.t͡sip/ RHL 'small house'. S.-J. Park (2012) also noted the unexpected similarity between the pitch accent patterns found in Ulsan to those found in Gyeongju, which is a city located immediately north of Ulsan and is considered to belong to a different dialectal division of North Gyeongsang dialect.

Before closing this section, it must be emphasised that the dialectal differences described thus far may not accurately reflect the state of dialectal variation in contemporary Korean, especially the variety spoken among younger speakers. For example, S.-K. Im (2014) deliberately limited the age of his informants to above 60, due to the dialect levelling widely observed across Korea. In particular, the three cities in question have seen rapid urbanisation and industrialisation, resulting in large influxes of speakers from other dialect-speaking areas. For example, Ulsan's population has proliferated from thirty thousand to one million within half a century since the 1960s (Y.-G. Jang, 2008). In addition, non-standard varieties of Korean have been stigmatised and marginalised, influenced by the ideology of Korea as a homogenous nation (Silva, 2011). Accordingly, "Standard" Seoul speech has been actively promoted on a national level through education and the media. This has caused "younger speakers of Korean to increasingly turn their backs on local language varieties and are frequently unfamiliar and/or rejecting of the dialectal lexical items and morphology used by older generations." (H.-M. Sohn, 1999, p. 460). Since the informants of the production and perception experiments in this dissertation are university students, the descriptions of traditional dialectal features here must be applied with caution. While speakers of Gyeongsang Korean are known to preserve dialectal features better than speakers of other regional dialects, some dialect levelling can still be expected among younger speakers.

3.4 The lenis stops and the three-way stop contrast of Korean

The phonetic and phonological characteristics of the Korean lenis stops /p, t, k/ have been studied extensively in the literature as part of the typologically unusual, three-way stop contrast along with the fortis /p*, t*, k*/ and aspirated stops /p^h, t^h, t^h/. Because these stops are all pulmonic egressive sounds and are viewed as underlyingly voiceless in the mainstream account, much research has focused on revealing how these stops are phonetically distinguished through acoustic, articulatory, and perceptual analyses. In earlier works such as Halle and Stevens (1971) and Lisker and Abramson (1963), the lenis stops, the partially aspirated stops, were treated as a typological anomaly under the assumption that the fortis stops are equivalent to the simple voiceless unaspirated stops in other languages. However, later studies have revealed that the fortis stops are phonetically more complex than previously known, and that the three Korean stops are signalled on multiple phonetic dimensions in a highly intricate manner. In the first half of this section, I will describe the phonetic characteristics of the three contrastive stops (Section 3.4.1). In the second half, I will focus on the phonological processes targeting the lenis stop, particularly on lenis stop voicing (Section 3.4.2).

3.4.1 The phonetic properties of the three contrastive stops

The three contrastive stops in Korean are distinguished on a number of acoustic dimensions. Their realisation is also highly sensitive to their position in the syllable as well as in larger prosodic domains such as the AP. In AP-medial position, the stop categories are signalled robustly by differences in VOT (positive or negative) and closure duration, whereas VOT (positive) and F0 are the main correlates of the stop contrast in AP-initial position, although these are by no means the only phonetic correlates of the stops (J.-I. Han, 1998; M.-S. Han & Weitzman, 1970; W.-J. Hardcastle, 1973; Hirose et al., 1974; S.-A Jun, 1993, 1995c; Kagaya, 1974; C.-W. Kim, 1965; C.-W Kim, 1970; M.-R. C. Kim, 1994; Lisker & Abramson, 1964).

Adding to the complexity of the picture, there are findings of an ongoing tonogenetic sound change in Seoul Korean (Silva, 2006). That is, more conservative speakers of Seoul Korean distinguish AP-initial stops primarily by VOT (Fortis < Lenis < Aspirated), and secondarily by F0 of the following vowel (Lenis < Fortis, Aspirated) (T. Cho et al., 2002). On the other hand, speakers with a more innovative system rely on F0[18] as a primary cue for the lenis-aspirated contrast, while showing comparable VOTs for these stops. Based on a synchronic analysis of a read-speech corpus of 120 Seoul Korean speakers, S.-H. Cho (2017) concluded that this sound change was initiated by females born in the 1950s, and has come to its completion among the speakers born in 1990s. In contrast, the study by Y. Kang and Han (2013), which is the first longitudinal instrumental phonetic study of tonogenesis, showed that the change had been ongoing in Seoul for approximately over a century. In textbook recordings made in 1935, a 41-year-old male speaker used VOT alone to signal the stop contrast, whereas a 11-year-old male speaker made use of both VOT and F0 cues. The latter speaker showed even greater reliance on F0 cues when recorded again at the age of 81 in 2005, indicating a substantial change in the direction of community-level sound change. According to S.-A. Jun (1993, 1998), such an F0 effect in Seoul Korean has been phonologised in AP-initial position, in which the underlying tone is H if the segment is either aspirated or fortis (which has the feature [+stiff vocal folds], Halle & Stevens, 1971) but it is L, otherwise. This gives the two possible underlying tonal patterns of the AP in Seoul Korean, LHLH and HHLH.

Focusing on the time course of the tonogenesis and the role of word frequency, H.-Y. Bang, Sonderegger, Kang, Clayards, & Yoon (2015, 2018) observed that both the loss of VOT distinction and the enhancement of F0 distinction were more advanced in high-frequency words, and that both occurred in parallel. This was interpreted to indicate that the loss of VOT distinction is triggered by hypo-articulation, resulting in the shortening of VOT in the

[18] Note that this system is unconnected with the lexical accent in Gyeongsang Korean.

aspirated stops (Browman & Goldstein, 1991; Pierrehumbert, 2001; Bybee, 2002), which was then compensated for by an adaptive response of F0 enhancement (see Kirby, 2013, for evidence in a similar vein from computer simulations). The explanation of the loss of VOT based on hypo-articulation is also supported by K.-H. Kang and Guion (2008) who found that the VOT difference between the lenis and aspirated stops is greater in clear speech than in conversational speech.

With this complexity in mind, I will limit the scope of the literature review below to AP-initial realisation of the three stop categories by relatively young speakers of the different varieties in question. One of the most relevant studies in this respect is by H. Lee and Jongman (2012). Focusing on the effects of tone on the three-way contrast, their study used disyllabic words with different pitch accents (HH or LH) in South Gyeongsang Korean and compared the realisation of these words by speakers[19] of South Gyeongsang Korean and speakers of non-tonal Seoul Korean. The target stop was always produced in the initial position of a disyllabic word in isolation. The study investigated the following acoustic and aerodynamic properties of the stops: VOT, F0 at following vowel onset, H1-H2 (the amplitude difference between the first and the second harmonic), airflow, and air pressure.

Consistent with the findings in the literature, VOT was shortest for the fortis, intermediate for the lenis, and longest for the aspirated stops in both dialects. However, compared to the Seoul dialect, the Gyeongsang dialect showed a shorter VOT in the lenis stops but a longer VOT in the aspirated stops. The average VOT values are given in Table 3.1. The interquartile ranges overlapped significantly between the lenis and aspirated stops for Seoul, whereas they were clearly separated for Gyeongsang. Since there was no effect of pitch accent, separate VOT values for HH and HL words are not provided here.

Table 3.1: Mean VOT value (ms) of the three contrastive stops in Seoul and South Gyeongsang Korean (H. Lee & Jongman, 2012).

	Fortis	**Lenis**	**Aspirated**
Seoul	17	65	80
South Gyeongsang	18	38	104

As shown in Table 3.2, the results for F0 also supported the previous findings, with the average value being lowest for the lenis stops, intermediate for the fortis stops, and highest for the aspirated stops in both dialects. There were also significant dialectal

[19] The speakers were aged between 21 and 48. Assuming that the recordings were made in 2012, we can estimate the birth year as 1964 - 1991.

differences. Compared to Gyeongsang Korean, Seoul Korean showed a greater F0 difference between the lenis stops and the fortis stops, and between the lenis stops and the aspirated stops. In a further analysis to understand the three-way interaction between laryngeal distinction, pitch accent, and dialect, they found no significant F0 differences between the HH lenis stops and the LH fortis stops and between the HH fortis stops and the LH aspirated stops within the South Gyeongsang data. In addition, their multiple discriminant function analysis revealed that, as a single predictor for the stop classification, VOT was found to be a more reliable cue in the South Gyeongsang dialect, whereas F0 was a more reliable cue in the Seoul dialect. From this, the authors concluded that the presence of lexical pitch for Gyeongsang speakers makes F0 an unreliable cue for the stop contrast, with lexical pitch accent apparently overriding the segmental effect.

Table 3.2: Mean F0 value (Hz) in Seoul and South Gyeongsang Korean as a function of laryngeal distinction and pitch accent (H. Lee & Jongman, 2012) (standard deviation in parentheses). The shaded cells mark the results that could have been affected by the incorrect categorisation of the pitch accents of three words in their stimuli.

	Pitch Accent[20]	**Fortis**	**Lenis**	**Aspirated**
Seoul	HH	143 (12.02)	116 (8.42)	158 (16.11)
	LH	140 (11.94)	117 (8.6)	157 (16.65)
South Gyeongsang	HH	136 (26.92)	122 (18.22)	156 (19.69)
	LH	124 (24.92)	111 (15.43)	140 (18.29)
Total mean		136 (18.92)	117 (12.90)	152 (18.18)

However, of the 18 target words used in H. Lee and Jongman (2012), the pitch accents assigned to three words did not match my intuition as a native speaker of Busan dialect: /p*aŋ-i/ (HH, 'bread-NOM'), /t^hal-i/ (LH, 'mask-NOM'), and /k^hoil/ (LH, 'coil'). In my own accent, the pitch accents of these words are LH, HH, and LH/HH (both seem possible for the last word maybe because it is a borrowed word), respectively. Quick informal recordings of four speakers (3 females in their 20s, one male in his 30s) from Busan confirmed this. All four speakers pronounced /p*aŋ-i/ as LH and /t^hal-i/ as HH. As predicted, the pitch accents of /k^hoil/ varied, with two speakers realising it as LH and the other two as HH. The authors stated that they used most of the stimuli from Kenstowicz and Park (2006) and it appears that errors were made in the process of adopting the stimuli. In Kenstowicz and Park, the /p*aŋ-i/ (HH) used was a less frequent homonym meaning 'prison' rather than 'bread', and

[20] For Seoul, this column represents words that *would* have a particular pitch accent in South Gyeongsang.

/tʰal-i/[21] was marked as having a HH, not LH pitch accent. H. Lee and Jongman reported that the data was collected from near Busan city from South Gyeongsang dialectal group, and thus it is possible that the subjects speak a slightly different variety than Busan Korean. Unfortunately, this cannot be verified as no further details are provided regarding the subjects' precise linguistic background.

This could have significantly affected their results (see the shaded cells in Table 3.2), since there were only three stimuli for each combination of laryngeal distinction and pitch accent, one per place of articulation. For example, there was only one bilabial fortis HH word in their stimuli, /p*aŋ-i/, and this was in fact a LH word. In the case of the aspirated stops, two out of three LH words could have been realised as HH. Thus, if the items had the correct pitch accents, the F0 of the HH fortis stop would be higher and the F0 of the LH aspirated stop would be lower. Recall the findings that both (a) the F0 difference between the HH lenis stop and the LH fortis stop, and (b) that the difference between the HH fortis and LH aspirated stop were not significant. While the validity of (a) is not likely to be affected, a lower mean F0 value of the LH aspirated stop may cause the difference between the HH lenis stop and the LH aspirated stop to also become insignificant. Furthermore, the result in (b) may change if the correct lexical words were chosen for the intended pitch accents. This would bring the results in H. Lee and Jongman closer to those in Kenstowicz and Park, whose main finding was that Gyeongsang Dialects (Busan and Daegu) show parallel pitch effects of the stop categories as in Seoul Korean regardless of the pitch accent type; F0 after the fortis and aspirated stops is relatively high compared to that after the lenis and nasal stops.

In addition to VOT and F0, H1-H2 was also significantly different depending on the stop category, with the pattern of Fortis < Lenis < Aspirated. This was interpreted to indicate that the vowel following the fortis stop is creakier and the vowel following the aspirated stop is breathier. The authors also found an interaction between laryngeal distinction and pitch accent, with the HH pitch showing a H1-H2 pattern of Fortis < Lenis < Aspirated and the LH pitch showing Fortis < Lenis = Aspirated, although the reliability of this is somewhat undermined by the incorrect pitch accent assignment as discussed above. However, it is worth noting that these results are in line with the recent report of a H1-H2 merger between the lenis and the aspirated stops among younger females from Seoul (M.-R. Kim, 2014). In fact, the literature is split regarding whether the lenis or the aspirated stop shows a greater H1-H2 value (Lenis < Aspirated: Ahn, 1999; H. Park, 2002; K.-H. Kang & Guion, 2008; but Aspirated < Lenis: T. Cho et al., 2002; Kenstowicz & Park, 2006).

---

[21] Interestingly, /tʰal-i/ also has a homonym which has a LH pitch accent, meaning 'trouble-NOM'. This may have caused the confusion.

Regarding the claim in Kenstowicz and Park that H1-H2 may compensate for the weaker role of F0 in Gyeongsang Korean, H. Lee and Jongman only found limited support. H1-H2 was a better predictor of the stop contrast in Gyeongsang than in Seoul Korean, but only by a small margin. Instead, VOT was a much better predictor in Gyeongsang Korean. In contrast, a more recent study with six Busan and six Suncheon[22] males found significant dialectal variation in H1-H2 (M.-R. Kim, 2017). Unlike previous findings for Seoul Korean, there was no breathiness in the vowel following a lenis stop for the Busan and the Suncheon dialects. This supports the claim that H1-H2 better distinguishes between the lenis and the aspirated stops in Gyeongsang Korean than in Seoul Korean, in line with Kenstowicz and Park (2006).

For the aerodynamic measurements in H. Lee and Jongman (2012), it was found that the intraoral airflow rate was lowest for the fortis, intermediate for the lenis, and greatest for the aspirated stops in both dialects. The dialects differed in that the airflow rates were generally greater for Seoul than Gyeongsang Korean, with the lenis stops showing a greater difference than the other stops. As for the intraoral air pressure, both dialects exhibited similar patterns that the air pressure was lower in the lenis stops than in the other stops, and it was higher in the HH pitch pattern than in the LH pattern. The general patterns confirmed the previous reports by Dart (1987) and T. Cho et al. (2002).

Of particular interest is that the interaction effect between laryngeal distinction and dialect for both VOT and airflow was carried by the lenis stops. That is, the differences in VOT and airflow were greatest in the lenis stops in the two dialects (H. Lee and Jongman, 2012, p. 162). Following the Strict Layer Hypothesis (Selkirk, 1986), if we assume that an initial segment in a word produced in isolation is also in the initial position of the Utterance, then we can infer that the VOT values reported here capture the maximum degree of DIS of the lenis stops on this parameter in the two dialects. Since the mean VOTs of U-initial lenis stops in the two dialects are significantly different, it would be interesting to examine from our production results later if the dialectal difference is maintained across different prosodic positions or if the VOT values between the two dialects only diverge in higher positions.

Focusing on the correlation between the acoustic and aerodynamic measures, the authors found a strong positive correlation between VOT and airflow (r = .613, p < .001). A correlation between VOT and airflow may not be so surprising, given previous findings of a linear correlation between VOT and glottal aperture (Kagaya, 1974), and that between glottal aperture and airflow rate (Dart, 1987). Articulatory investigations of the state of the

[22] Suncheon is located in the South of Jeolla province. See the map in Figure 3.2.

glottis in the three stops consistently showed that the glottal opening is smallest for the fortis, intermediate for the lenis and largest for the aspirated stops (C.-W. Kim, 1970; H. Kim, Honda, & Maeda, 2005). This suggests that the VOT values which will be obtained in the production study (Chapter 5) can be used, to a certain extent, as an indicator of the width of the glottal aperture. Greater VOT for the lenis stops in a higher position would suggest that DIS of the lenis stops involves a greater glottal abduction gesture.

H. Lee and Jongman also examined the correlation between F0 and air pressure, motivated by the finding that air pressure was higher for words with the HH pitch accent than for those with the LH pitch accent. The two measures showed a moderate positive correlation ($r = .398$, $p < .001$). The main driver of the higher intraoral air pressure in the fortis and aspirated stops than in the lenis stops is suggested to be the greater subglottal pressure (T. Cho et al., 2002; Dart, 1987). Thus, these results are in line with the findings of a linear correlation between subglottal pressure and F0 (Shipp & McGlone, 1971; Atkinson, 1978). T. Cho et al. (2002) further speculated that the fortis stops may be produced with a greater subglottal pressure than the aspirated stops, yet show comparable intraoral air pressure due to the smaller glottal area during the fortis stop production, preventing the intraoral pressure from rising to a higher level than in the aspirated stops.

Other potentially relevant acoustic measures that have not been discussed so far include acoustic closure duration and the percentage of voicing during acoustic closure. Using the recordings of 12 speakers – three females and three males each from Seoul and Busan born approximately between 1962 and 1974 – M.-R. C. Kim (1994) examined the VOT and closure duration of the three stops.[23] Although VOT was not different between the lenis and aspirated stops for Seoul females, closure duration remained distinct between the two stops, with the aspirated stops having a longer mean closure duration than the lenis stops. This result is in line with a more recent study by T. Cho and Keating (2001) (see Section 2.3.1 for a review focusing on the effect of prosodic position). Across a range of prosodic positions, acoustic closure duration showed a main effect of consonant type showing a pattern of /t*/ > /t^h/ > (/t/ = /n/), with a large difference in the mean value between /t^h/ and (/t/ = /n/). They also measured voicing residue interval during stop closure as an indicator that the vocal folds have not yet completely abducted. This may also be used to infer the vocal tract tension, as less vocal tract tension makes closure voicing more likely. While the raw values of voicing intervals did not vary across different consonant types, this interval relative to the entire closure duration was significantly different, with /t/ showing a greater % voicing than /t^h, t*/. This is expected as the closure duration for /t/ is significantly shorter than /t^h, t*/.

[23] The target segments were presumably in AP- or IP-initial position, considering that they were realised in the beginning of a word which was embedded in a fixed frame sentence.

For the patterns in Ulsan, phonetic studies which specifically target the variety spoken in Ulsan could not be found, even though it is one of the better studied dialects of South Gyeongsang according to the research history summarised by Kong et al. (2016). While there is an abundance of phonological research for South Gyeongsang Dialects, instrumental phonetic work is lacking in general. Another reason for the lack of research of the Ulsan dialect may be because it is expected to be very similar to the Busan dialect, which has received more attention in the literature. This is exemplified in H. Lee, Politzer-Ahles, & Jongman (2013) who investigated the difference in the use of perceptual cues for the three-way stop contrast by Seoul and South Gyeongsang Koreans. Their informants for South Gyeongsang Korean came from both Busan and Ulsan.

There are a couple of other cross-dialectal studies that are worth mentioning. Holliday and Kong (2011) examined the acoustic correlates of the stop contrast in Seoul, Daegu, and Jeju and found that speakers from Daegu, especially the males, relied more on VOT than Seoul or Jeju speakers, reflecting a conservative system with regard to the use of VOT and F0. Given that Busan and Daegu varieties behave similarly, it can be assumed that Ulsan Korean will follow a similar pattern, as Ulsan is geographically very close to Busan, while also sharing certain similarities with Daegu Korean in the pitch accent pattern (see at the end of Section 3.3).

Comparing the Seoul and Southern Jeolla (Jeonnam) dialects, H. Choi (2002) found that while Seoul Korean shows a three-way distinction in F0 and a two-way distinction in VOT, Jeonnam Korean shows the opposite pattern with a three-way distinction in VOT and a two-way distinction in F0. It was suggested that the dialectal variation can be attributed to the differences in the prosodic system. Jeonnam Korean has a salient rising intonational contour in AP-initial position, in contrast to Seoul Korean which has a rising contour realised in AP-final position. In addition, Jeonnam Korean preserves the vowel length contrast, unlike Seoul Korean. Thus, the greater use of VOT compared to F0 by Jeonnam speakers may be explained as a result of the restricted role of F0 in signalling the laryngeal contrast in initial position and their greater sensitivity to durational properties. This is similar to the argument given for Busan Korean, that F0 has a limited role as a cue because it also signals lexical contrasts by lexical pitch accent.

To summarise, research shows that the three stop categories in Korean are distinguished on the acoustic parameters of VOT, F0, and H1-H2 in AP-initial position. Reflecting the relatively recent tonogenetic sound change, younger speakers of Seoul Korean show the following pattern; VOT is longer for the lenis and aspirated stops than for the fortis stops, while F0 is higher for the aspirated and fortis stops than for the lenis stops. The previous studies also agree that H1-H2 is generally higher for the vowel following the lenis and aspirated stops than for the vowel following the fortis stops. However, the pattern is slightly different for South Gyeongsang Korean, which encompasses Busan and Ulsan Korean. Compared to Seoul Korean, VOT is shorter for the lenis stops and longer for the aspirated stops in South Gyeongsang Korean. The relative F0 differences among the three stop categories are similar to those in Seoul Korean, although each kind of stop showed significantly lower F0 values in a word with a LH pitch accent than in a word with a HH pitch accent. Some studies show that H1-H2 is significantly higher in the vowel preceded by the aspirated stops than in the vowel preceded by the lenis stops, as only the former shows breathiness. This contrasts with the pattern in Seoul Korean, where both of these stops are reported to be followed by a breathy vowel, showing comparable values of H1-H2.

3.4.2 Lenis stop voicing

Lenis stop voicing is a widely recognised postlexical phonological rule in Korean whereby a lenis stop becomes voiced between voiced segments within an AP (S.-A. Jun, 1993, Section 3.1; Lisker & Abramson, 1964; Y. Y. Cho, 1987). The formal description given in S.-A. Jun (p. 78) is shown in (6).

(6) Lenis stop voicing:

$$[-\text{cont}, -\text{c.g.}, -\text{s.g.}] \rightarrow [+\text{voice}] \ / \ _{\text{AP}}[\ldots[+\text{voice}] \_\_\_ [+\text{voice}]\ldots]$$

However, there is a controversy surrounding the categoricity of the rule. The results in S.-A. Jun showed that while AP-initial lenis stops were almost always voiceless at a normal rate of speech, there were counter examples in which AP-initial lenis stops were either entirely or partially voiced at a faster rate. Based on this finding, S.-A. Jun (pp. 107-109) suggested that lenis stop voicing is a gradient phonetic phenomenon rather than a categorical rule. She then proposed an explanation based on gestural overlap and magnitude using the gestural score model by Browman and Goldstein (1992; see also Browman & Goldstein, 1986). According to this view, the voiced realisation is not analysed as a distinct allophone of the lenis stops, but merely as a by-product of reduced magnitude of the glottal abduction (opening-and-closing) gesture in AP-medial position and greater gestural blending of CV due to a relatively short duration of C in medial position.

Challenging this view, Docherty (1995) pointed out that S.-A. Jun's data showed a largely discrete distribution of voiced and voiceless realisations. There were only a few intermediate cases of partially voiced stops, while more of these would be expected if lenis stop voicing is truly a result of gradient gestural overlap. Kagaya's (1974) fibrescopic study was cited as possible counter evidence. Although fibrescopy is perhaps less reliable than, for instance, electromyography of abductor muscles, the glottal width data suggested that intervocalic lenis stops were produced with the complete absence of a glottal abduction gesture. Such a complete lack of a gesture was interpreted to indicate that the process is categorical, rather than gradient.

Using their data on DIS, T. Cho and Keating (2001) were able to briefly examine the nature of lenis stop voicing. While they did not take a decisive position on the matter, their results, copied in Figure 3.3, evidently show a bimodal distribution of voicing with one peak around 30% voicing and the other at 100% voicing if the three graphs were to be combined into one. A bimodal distribution on a particular phonetic parameter has been used as a diagnostic for a categorical phenomenon, as this indicates two discrete targets (e.g. Bermúdez-Otero & Trousdale, 2012; Turton, 2017). T. Cho and Keating's reservation for concluding this as a categorical phenomenon was due to the great degree of variability in AP-initial stops. However, recall the model of grammar presented in Section 1.2.2 (Figure 1.1). It is common for an effect of a categorical rule to be overlaid by an independent phonetic process, which would be sensitive to factors such as speech rate. Thus, lenis stop voicing in Korean appears to be a categorical process with two discrete targets and the deviation from this could be explained as a result of a distinct lower-level process, namely, gestural blending with adjacent voiced segments. This interpretation is in line with Scobbie (1995) who pointed out that evidence of gradience is not evidence for the absence of a categorical effect.

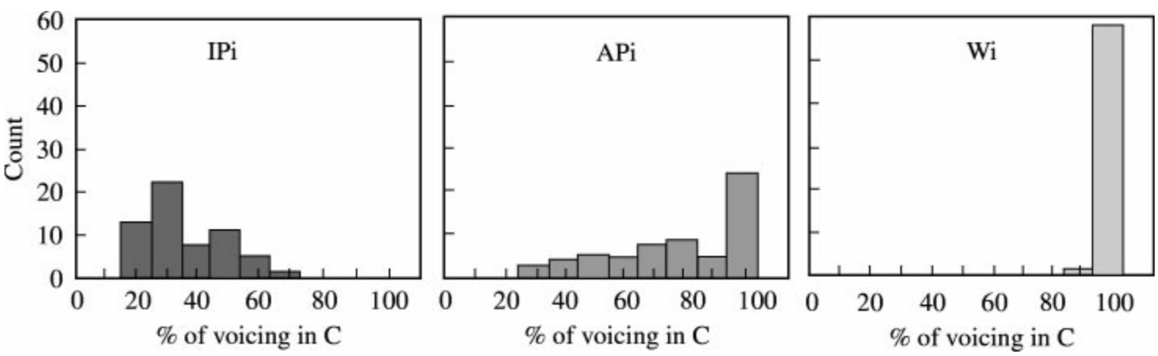


Figure 3.3: The distribution of the percentage of voicing measurements for the lenis stops in IP-initial (IPi), AP-initial (APi), and PW-initial (Wi) position (T. Cho and Keating, 2001, p. 183).

Another aspect to note from T. Cho and Keating's (2001) results (Figure 3.3) is that over 20% of the AP-initial tokens were completely voiced. This is unexpected, given S.-A. Jun's rule in (6) above, which is based on her finding that AP-initial stops were almost always entirely voiceless. There are several possible explanations. First, it may be due to the methodological differences. Unlike Cho & Keating who relied solely on acoustic evidence, Jun's study used Electroglottograph (EGG) in addition to acoustic waveforms. Based on the degree of electrical impedance between the electrodes placed on the left and right of the neck surface, EGG provides an indication of the opening and closing of the glottis. As noted in Jun (p. 91), it is possible for the vocal folds to vibrate while abducted, given an open vocal tract and a high subglottal air pressure, such as in breathy [ɦ]. Thus, in case of conflicting evidence from EGG and audio waveforms, Jun used EGG data as it reflects vocal fold vibration more directly than audio waveforms. This could explain why Jun's results indicated less voicing than Cho & Keating's.

Another possible explanation for the differences in T. Cho & Keating (2001) and S.-A. Jun (1993) is that the average speech rate in Cho and Keating's data may have been higher than in Jun's. Alternatively, it could be that the authors of the two studies made different judgments on the prosodic boundaries, as there are often cases which seem intermediate between two boundary types. A final possibility is that the exact conditioning of the lenis stop voicing rule shows a high degree of between- and within-speaker variation, such that for some speakers the devoicing targets IP-initial stops and for others it targets AP-initial stops. While it cannot be known whether any of these explanations is correct, the results from the production study in Chapter 5 may shed some light on this issue.

There is little previous work on the dialectal variation of lenis stop voicing. Although S.-A. Jun (1993) examined the production by five speakers from Jeonnam (Southern Jeolla) and three speakers from Seoul, she did not report any dialectal effect on the pattern of lenis stop voicing. As lenis stop voicing is known to be a regular feature of Korean, the working assumption for the production study will be that Busan and Ulsan Korean show similar patterns of lenis stop voicing as Seoul Korean.

3.5 The nasal consonants and denasalisation

The nasal stops in Korean have received little focus in the literature unlike the plosives which are better known as having a typologically unique three-way contrast. S.-G. Gim (1937; cited in Young Shin Kim, 2011), among others, described the Korean nasals to be phonetically no different from the English nasals. More recently, however, several studies have established that the Korean nasals /m, n/ become denasalised in onset position (Yoshida, 2008; Young Shin Kim, 2011; Yoo, 2015a; Yoo, 2016). Below, I will give a brief literature review of denasalisation, focusing on the phonetic characteristics, the grammatical status, and any known dialectal variation.

Denasalisation can be defined as the realisation of nasal phonemes with an obstruction of the velopharyngeal port. Throughout this dissertation, I use the expressions "complete" (e.g. /m/ [b]) and "partial" denasalisation (e.g. /m/ [m^b]) to refer to the temporal aspect of denasalisation. That is, if the nasality extends for the total duration of the nasal phoneme, this is described as complete denasalisation. On the other hand, partial denasalisation refers to when the nasality extends for shorter than the entire duration of the nasal phoneme. Since the release of Hunminjeongeum (King Sejong, 1443) – the first comprehensive documentation on Korean phonetics – up to 2006, there is not a single publication on Korean phonetics and Korean nasals written by Korean scholars which mentions denasalisation (Young Shin Kim, 2011, p. 16). One of the earliest descriptions of denasalisation in Korean can be found in Jones (1924). He observed that before high vowels and sometimes before /o/, the Korean onset nasals /m, n/ can either be realised with post-oralisation at the offglide, i.e. [m^b, n^d], or with complete denasalisation, i.e. [b, d]. Since then, only a handful of studies have acknowledged or examined the phenomenon and many of them are based on impressionistic judgments and/or a very small number of informants. For a more detailed review, see Young Shin Kim (2011, Section 2.13).

Even among the limited number of studies, there is disagreement on the possibility of complete denasalisation and the exact conditions in which denasalisation occurs. Martin (1951) and Umeda (1957) agree that post-oralised nasals [m^b, n^d] are free variants with sonorant nasals in Korean. S.-D. Lee and Kim (2007) stated that complete denasalisation is not likely given the presence of carryover nasalisation in the vowel after a word-initial nasal, although the proportion of vowel nasalisation was smaller than in the vowel after a word-medial nasal. In a nasometer study, Yoshida (2008) also argued that denasalisation is incomplete, although nasality does become weaker as the prosodic boundary becomes stronger. T. Cho and Keating (2001, p. 174) compared RMS acoustic energy in Korean nasals across different domain-initial positions to infer the amount of minimum nasal airflow. Although there was systematically less acoustic energy in nasals in higher domains, even IP-initial nasals showed an average of around 57dB. However, acoustic energy does not tell us how much of it is due to nasal resonance and how much is due to variations in the source intensity, i.e. voicing (for further discussion of this problem, see Section 5.3.5 (1)).

Young Shin Kim (2011) was the first relatively large-scale study on Korean denasalisation. Through acoustic, aerodynamic, and accelerometer analyses, she showed that Korean word-initial nasals can undergo complete denasalisation and concluded that denasalisation is a regular feature of Korean found across different dialects. She also reported that word-initial nasals could be realised with a release burst resembling that of a plosive. The presence of a release burst strongly suggests a closed velopharyngeal port which is necessary for the pressure build-up in the oral cavity (Yoo, 2016, p. 8). Furthermore, Yoo

(2015a) found that underlying nasals can even be realised with a loss of voicing and additionally with short aspiration, i.e. [p, t] ~ [p‘, t‘] [24].

Although earlier studies such as Jones (1924), Martin (1951), and Chen & Clumeck (1975) have suggested that the context of denasalisation is before high (back) vowels, Young Shin Kim (2011) found denasalisation regardless of the vowel context. This could be because earlier studies tend to be limited in their methodology. For example, Chen & Clumeck is based on only one informant and the authors' impressionistic judgments of 45 words spoken in isolation. Martin (1951) is also assumed to have used auditory judgment by the author as it mentions neither the methodology nor the number of informants. Another possibility is that the context for denasalisation has become more generalised over time. Evidence in line with this possibility will be discussed below.

Young Shin Kim's (2011) results also contradict the findings in Yoshida (2008) and T. Cho & Keating (2001), that the degree of denasalisation is greater in the initial position of a higher domain. Instead, Young Shin Kim (p. 27) found that denasalisation is only found in stressed word-initial nasals, which are equivalent to AP-initial nasals in the K-ToBI system (S.-A. Jun, 2000). In a relatively recent study, Yoo (2016) categorised tokens of initial nasals into five types based on the presence or absence of nasality, voicing, and aspiration: (1) Sonorant nasals [m, n]; (2) Partially denasalised nasals [m^b, n^d]; (3) (Pre-)voiced nasals [b, d]; (4) Voiceless unaspirated non-nasal [p, t]; (5) Voiceless non-nasal with aspiration longer than 10ms [p^h, t^h] (For exemplification of these categories and the detailed methodology, see Section 5.3.6. This section describes the method in Yoo & Nolan [forthcoming], which follows a very similar categorisation scheme as Yoo, 2016). Based on the frequency of tokens which belong to the five categories – rather than on the gradient degree of nasality measured by an articulatory instrument – Yoo found support for the cumulative prosodic conditioning of denasalisation. The higher the prosodic domain, the greater the frequency of tokens in a higher (more extreme) category of realisation. These results, taken together with the acoustic energy measurements in T. Cho and Keating, contradict Young Shin Kim's finding and support the view that denasalisation is phonetically gradient as a function of prosody. This in turn suggests that there may be a link between denasalisation and the general phenomenon of DIS, as discussed in Section 1.2.2. Again, the divergent findings in these studies could be due to the differences in their methods, i.e. acoustic versus articulatory, the possibility that the nature of denasalisation has changed over time from a gradient to a categorical process, and/or the high degree of individual variation.

[24] As mentioned in Chapter 1, reversed apostrophe here indicates weak aspiration as in earlier IPA usage (Pullum & Ladusaw, 1996, p. 250).

Yoo (2016) and Yoo & Nolan [forthcoming] examined the historical development of denasalisation, in an attempt to reconcile the inconsistencies in different descriptions of denasalisation in terms of the exact phonetic realisation, segmental context, and cumulative nature. The hypothesis was that the phonetic realisation of denasalisation has proceeded gradually, e.g. [n → n^d → d → t → t^h], and that the stronger cases of denasalisation with devoicing and/or aspiration are a recent development in Seoul Korean. This hypothesis was motivated by the observation that there is an overall trend in the literature to report increasingly extreme cases of denasalisation and broader contexts of application. As predicted, Yoo (2016) found evidence of a relatively rapid development of denasalisation over time, acquiring more plosive-like features of voicelessness and aspiration and applying in a wider range of vowel contexts.

There are three studies which shed light on whether there is any dialectal variation in the patterns of denasalisation. Young Shin Kim (2011) found evidence of denasalisation from speakers with a variety of geographical background: three speakers from Jeolla, two speakers from Gyeongsang, one speaker from Chungcheong, and two speakers from Seoul. However, Yoshida (2008) claimed that denasalisation had a different status in Gyeongsang and Gyeonggi Korean. His findings indicated that Gyeongsang Korean showed duration-dependent nasality weakening (partial denasalisation), as opposed to complete denasalisation. He found that the longer the nasal relative to the tautosyllabic vowel, the more denasalised the nasal. This contrasted with Gyeonggi Korean, in which nasality weakening was independent of duration. This was interpreted to mean that sub-phonemic categories have emerged in this latter variety. Yoo's (2015b) results were in the same direction. Although the study was not originally designed to compare Seoul and Busan Korean, it was observed that greater denasalisation was found in younger and less local speakers of Busan, who showed greater mobility and had more contact with speakers from Seoul. In addition, Yoo (2015b) also examined the realisation of initial nasals in Seoul Korean using news recordings made in 2015 and 1987. The findings were that more recent recordings showed more frequent denasalisation than the older recordings, and that they also showed more frequent denasalisation than the recordings of Busan Korean speakers. Taken together, it was suggested that denasalisation is an ongoing sound change which is more advanced in Seoul than in Busan.

With the hypothesis that lenis stop voicing and denasalisation are stabilised processes stemming from the phonetic phenomenon of DIS (Section 1.2.2), this dissertation may be able to capture the process of stabilisation in action. On the one hand, the lenis stop voicing rule is expected to have already entered a higher-level component in the grammar of Phrase-level Phonology in all three varieties; on the other hand, denasalisation may still be in the process of stabilisation, and in an earlier stage in Ulsan and Busan than in Seoul.

3.6 Word-initial tensification

There is little research on dialectal differences related to the patterns of DIS in general. The only previous works I have been able to find are my own (Yoo, 2015a; Yoo, 2015b; Yoo, 2016) and by Yoshida (2008), assuming that Korean denasalisation is indeed related to DIS as proposed in Section 1.2.2. Since these studies have been discussed in Section 3.5, this section focus on another phenomenon in Korean which is potentially relevant to DIS, known as *word-initial tensification (fortition)*.

Word-initial tensification refers to both the synchronic process of producing lenis consonants /k, t, p, s, t͡s/ as fortis consonants /k*, t*, p*, s*, t͡s*/ in word-initial position and the equivalent diachronic process. The earliest mention of this phenomenon is found in Choe (1937, p. 137, as cited in S. Jang, 2017), who mentioned that common words such as /k*ot͡sʰ/ 'flower' were written in historical documents as <kos>, reflecting the change in the pronunciation from the lenis to the fortis word-initially. Since then, numerous studies have mentioned word-initial tensification but extensive and systematic investigations are only found in more recent studies. Using a picture elicitation method, M. Han (2011) showed that fortis production is more frequent in males and in younger speakers. To trace the diachronic development, M. Han also made use of 'The Korean Language Digital Search Database of 2007' containing the historical forms of approximately 3300 lexical items. While views are divided regarding whether the fortis stops existed as phonemes in Korean spoken the 15^{th} century (pp. 151-152), M. Han's examination revealed that the use of tensification began to accelerate from the 17^{th} century, reaching its peak in the 19^{th} century. The historical documents showed that the lenis stops and the now-obsolete initial consonant clusters had been replaced gradually by the fortis stops.

Regarding the dialectal aspect, Oh (2011) and S. Jang (2017) both arrive at the conclusion that word-initial tensification is most frequent in and started from Gyeongsang Region, although it is now widespread across South Korea with differences only in the specific lexical items that undergo the process. Tensification of the stops and affricate /k, t, p, t͡s/ is hypothesised to have been an innovation in the south-east of North Gyeongsang and north-east of South Gyeongsang, whereas that of the fricative /s/ is believed to have started in the Southwest of South Gyeongsang. While these studies do not include the metropolitan cities in their analysis as part of the dialect speaking region of South Gyeongsang, Busan and Ulsan varieties are expected to show the same strong tendency to tensify the lenis stops in certain lexical items like the rest of Gyeongsang dialect region. Such a tendency in Busan Korean has indeed been noted in another study by Young-Seon Kim (2012, p. 88).

Word-initial tensification and DIS are similar in some ways, although the former is presumably a categorical, lexically-conditioned rule whereas the latter is a gradient phonetic process. At least in the synchronic phenomenon in which the lenis stops undergo fortition word-initially, it resembles the way initial lenis stops and nasals are "strengthened" by gaining more obstruent-like features. Thus, one may hypothesise that word-initial tensification is yet another instance of a scattered rule which originates from a gradient process of DIS. If true, the phenomenon has now reached the lexicon, the highest module in the architecture of grammar described in Section 1.2.2. However, as this question lies outside the scope of this dissertation, testing this particular hypothesis is reserved for future research.

When I designed the reading materials for the production study, I avoided the lexical items which are prone to undergo word-initial tensification. Although there was one item that could potentially be subject to word-initial tensification (/tʌpɯlpeisɯ/ 'double-base'), all speakers realise the /t/ as a lenis stop, presumably as they were reading the sentences in a relatively formal setting. This is understandable given that tensification, unlike lenis stop voicing and denasalisation, is something that speakers are aware of. It is socially considered "less polished" (Choe, 1937) and thus consciously avoided in formal situations (M. Lee, 2002, as cited in M. Han, 2011, p. 50).

3.7 Chapter summary

In this chapter, I have introduced the three varieties of Korean spoken in Seoul, Busan, and Ulsan. First, the prosodic system of Standard Korean was described, followed by a summary of the dialectal differences among the three varieties, both in terms of the segmental and suprasegmental aspect. For example, it was mentioned that Busan and Ulsan Korean lack the phonemic distinctions between /ɯ/ and /ʌ/ and between /s/ and /s*/. More importantly for our purposes, /k/ is sometimes palatalised to the affricate /t͡s/ before /i/ or /j/ in certain lexical items. This indicates that the experimental materials targeting /k/ should be chosen carefully in order to exclude these items. With regard to prosody, a crucial difference was discussed, that Gyeong Korean, unlike Seoul Korean, has a lexical pitch accent system.

As the main focus of this dissertation is on the lenis stops and nasal consonants, I examined the phonetic and phonological characteristics of these segments in greater detail. While the three-way system of plosives is signalled on a wide range of phonetic parameters, F0 and VOT are known to be primary cues. For younger speakers, studies including H. Lee & Jongman (2012) suggested that F0 is a more reliable cue for Seoul, whereas VOT is a more reliable cue for South Gyeongsang Korean. This was explained as a consequence of the dual role of F0 in South Gyeongsang, distinguishing lexical items as well as signalling the three-way stop contrast. Thus, for our production study, we may expect that Busan and Ulsan speakers will be more reluctant to vary VOT of the lenis stops under DIS compared to Seoul speakers, since the distinction between the lenis stops and the aspirated stops is still mainly maintained in terms of VOT rather than F0 in Gyeongsang Korean. Regarding the well-known rule of lenis stop voicing, little evidence could be found on its potential dialectal variation. Therefore, the working hypothesis for the later experiments will be that lenis stop voicing is a regular feature of Korean that does not show a high degree of dialectal variability.

Compared to those of the lenis stops, the phonetic characteristics of the nasals in Korean have not received much scholarly attention until relatively recently, when word-initial denasalisation was reported by Young Shin Kim (2011). Again, the literature on the potential dialectal variation is thin but there is a suggestion that Seoul represents a more innovative variety than Busan (Yoshida 2008; Yoo, 2015b) in a recent sound change involving denasalisation (Yoo & Nolan, forthcoming). Finally, a potentially relevant process of word-initial tensification was introduced, as a lexically-conditioned process whereby the lenis stops are produced as the fortis stops word-initially. Research shows that the process is more frequently found in the Gyeongsang dialects than in Standard Korean. Based on the literature review so far, a production study was conducted with the goal of understanding the characteristics and nature of DIS as applied to the lenis and nasal stops in Seoul, Busan, and Ulsan Korean. This study will be discussed in Chapter 5.

Chapter 4.

Prosodic parsing

4.1 Chapter overview

This chapter will provide a literature review relevant to the perception experiment in Chapter 6. First, I will give a brief background to the concept of prosodic parsing (Section 4.2). Next, I will present previous findings on the role of two well-known suprasegmental cues in prosodic parsing: duration and pitch (Section 4.3). Then, I will review existing research on segmental cues. I will first focus on the perceptual role of DIS in general (Section 4.4), before moving on to the perception of the lenis and nasal stops in Korean (Section 4.5). As before, there will be a short summary of the chapter at the end (Section 4.6).

4.2 Prosodic parsing

There is much existing research on the topic of speech segmentation, i.e. how listeners are able to divide the stream of speech into individual units such as words. For example, a seminal work by Davis, Marslen-Wilson, and Gaskell (2002) focuses on how listeners process temporal lexical ambiguities, e.g. between *captain* and *cap tucked.* While the first four phonemes of these two stimuli are exactly the same, i.e. /k æ p t/, there is a multitude of acoustic differences which signal the presence or absence of a word boundary between /p/ and /t/. In addition to allophonic differences such as [ʔt] vs [tʰ] (in certain varieties of English), there are additional cues such as segmental and syllable duration. Word-initial and word-final segments are longer than word-medial segments (Beckman & Edwards, 1990; Lehiste, 1960), and the syllable *sleep* is progressively shorter in the words *sleep*, *sleepy,* and *sleepless* (Lehiste, 1972).

Going beyond lexical segmentation, the rest of this section will focus on a specific problem of prosodic parsing, which can be defined as the resolving of speech into its prosodic constituents of various levels, such as the Syllable, the Word, the Accentual Phrase, and the Intonational Phrase. The concept of prosodic parsing needs to be understood with the theoretical assumption made in Section 1.1, which views prosody as an abstract grammatical structure organised in a hierarchy. The syntactic structure of an utterance constrains its prosodic structure, although there are other factors that influence the latter such as rate of speech (e.g. Shattuck-Hufnagel & Turk, 1996, Section 2.2). Therefore, given a syntactic ambiguity, parsing an utterance into its prosodic structure is key to recovering the syntactic strucutre and ultimately the intended meaning of the utterance.

There are numerous studies showing the role of prosodic phrasing on syntactic disambiguation. For example, Schafer and Jun (2002) conducted a cross-modal naming task with Korean listeners and found that a noun phrase in the form of Adjective + NP1 + NP2 (e.g. /hjʌnmjʌŋhan aki ap*a/ 'wise baby's dad') was processed faster when the accentual phrasing of the input matched its meaning (e.g. wise [baby's dad]) than when it did not (e.g. [wise baby]'s dad). This is one of many pieces of evidence showing that listeners use prosodic phrasing to infer the syntactic structure of an utterance, thereby resolving the syntactic ambiguity. The perception study in this dissertation intends to isolate the role of DIS cues from other prosodic cues in prosodic parsing during the processing of syntactically ambiguous stimuli. Thus, the following sections will review how suprasegmental (4.3) and segmental (4.4) variation provide perceptual cues to the prosodic structure of an utterance.

4.3 The role of duration and pitch

The role of duration and pitch in syntactic disambiguation and prosodic-boundary detection has been studied extensively (e.g. Lehiste, 1973; Streeter, 1978; Beach, 1991; Price et al., 1991, de Pijper & Sanderman, 1994; Schafer, 1997, Schafer et al., 2005). An early work by Lehiste (1973) found evidence of F0 and durational cues on the disambiguation of sentences with different bracketings, e.g. *Steve or Sam and Bob will come*. The relative importance of these cues seems to depend on the language and the type of prosodic boundary involved. Looking at similar bracketing ambiguity in English as Lehiste (1973), Streeter (1978) showed that duration and intonation both had significant effects in changing perceived meaning, whereas intensity contributed to the perception of meaning only in the presence of other cues.

In Korean, the IP and AP boundary have different pitch and timing properties as reviewed in Section 3.2. While the IP is marked by substantial phrase-final lengthening, an optional pause, and a boundary tone, the AP is defined by an underlying tonal contour which is in turned determined by the type of the AP-initial segment. The existence of AP-final lengthening is controversial but the balance of evidence suggests that it is at best limited and not necessarily longer than the AP-medial segment (Koo, 1986; S.-A. Jun, 1993, 1995c, 1996; T. Cho & Keating, 2001).

Reflecting these production patterns, the findings by S. Kim and Cho (2009) and Jeon and Nolan (2010) consistently indicate that pitch cues play a primary role in the detection of the AP boundary, while duration cues are either insignificant altogether (Jeon and Nolan, 2010), or significant only when there is a lack of a clear pitch cue (S. Kim & Cho, 2009). In a more recent study, however, Jeon and Nolan (2013) found that timing cues were more effective than pitch cues. It was speculated that this may be because their manipulated stimuli contained gradient temporal variations but only limited pitch variations. This suggests that the precise interplay between timing and duration cues depends not only on the language and boundary type, but also on the availability of acoustic cues in the stimuli. It is also worth pointing out that the effect of duration observed in Jeon and Nolan (2013) may be extra-linguistic rather than reflective of the patterns of prosodic parsing in Korean. As implied by the authors' analogy to musical grouping, a substantial lengthening (e.g. by 70% in Jeon & Nolan, 2013) on the AP-final syllable may signal a natural break even though Korean AP's only exhibit a small degree of final lengthening.

To summarise, the background review in this section has demonstrated that duration and pitch cues are exploited by listeners to infer the prosodic structure of an utterance and to resolve any syntactic ambiguities. This indicates that it is important to eliminate the confounding effects of pitch and durations cues in the perception experiment in Chapter 6, in order to tease out the effects of DIS in prosodic parsing.

4.4 The role of DIS

Next, we turn to the discussion on the role of DIS in prosodic parsing. Fougeron and Keating (1997, p. 3738) suggested three ways in which articulatory strengthening at the edges of prosodic domains could benefit speech comprehension. First, it could serve as a cue to prosodic demarcation. This is based on their finding that domain-initial consonants show more linguopalatal contact than domain-medial or -final consonants, in contrast with the pattern that domain-final vowels show less linguopalatal contact (i.e. a more decisively vocalic articulation) than domain-initial or -medial vowels. This implies that whether there is a prosodic boundary between two CV syllables (CV#CV as opposed to CVCV), for instance, could be marked by an enhanced articulatory contrast between the first vowel and the second consonant. While linguopalatal contact may not have a direct acoustic correlate, subsequent studies have shown that segmental articulation is rich with acoustic cues which can signal if there is a strong boundary preceding the initial segment, including closure duration, VOT, and nasality (e.g. T. Cho & Keating, 2001).

Second, more speculatively, the degree of articulatory strengthening could mark the strength of a prosodic boundary. Based on the broadly cumulative pattern of strengthening, listeners may go beyond simply perceiving the presence of a prosodic boundary and exploit the degree of DIS to as a cue to the strength of the prosodic boundary. However, Fougeron and Keating (1997, p. 3738) are careful to suggest that while IP boundaries may be distinguished from Word boundaries, it would be unlikely that naive listeners could distinguish two close boundary types – e.g. IP boundaries and ip boundaries – in their everyday linguistic interactions, given the large degree of between-speaker variation observed in their data.

Third, DIS may facilitate lexical access. Assuming that DIS enhances segment-specific articulation, it may aid the identification of a phoneme against perceptually similar phonemes in the language. This would be particularly helpful in domain-initial positions where there is less top-down syntactic and semantic information available to listeners (Gow, Melvold, & Manuel, 1996). A similar argument has been proposed for stress by de Jong (1995). He described the effect of stress on segmental articulation as local hyper-articulation, which would make lexical contrasts more distinctive. However, this explanation does not hold for all cases of DIS. Domain-initial nasals show reduced nasal energy in languages such as Korean (Young Shin Kim, 2011) and French (Fougeron, 2001), which has the opposite effect of enhancing segment-specific articulation. Similarly, Dutch voiceless stops show less aspiration in higher positions, which cannot be seen as enhancing the contrast with the voiced stops.

A different way that DIS may aid lexical access is by facilitating lexical segmentation. Because DIS is not found in a word-medial position, it could signal the start of a new word. Note that this follows naturally from the first possibility discussed above, that DIS may signal the presence of a prosodic boundary. In word monitoring and phoneme detection tasks with French participants, Christophe, Peperkamp, Pallier, Block, & Mehler (2004) tested the hypothesis that the Phonological Phrase boundary constrains lexical access. They found a delay in lexical access of monosyllabic words (e.g. *chat*) within a Phonological Phrase (e.g. [un *chat* grincheux]) in the presence of a potential competitor (e.g. *chagrin*), but not when the target word was followed by a Phonological Phrase boundary (e.g. [son grand *chat*] [grimpait…]). Based on these results, the authors claimed that listeners terminate pending lexical searches when they encounter Phonological Phrase boundaries. In a similar vein, Gout, Christophe, & Morgan (2004) used the head-turn preference paradigm to show that 13-month-old American infants exploit Phonological Phrase boundaries to constrain lexical access. They preferred listening to a familiarised disyllabic word (e.g. *paper*) when it occurred within a sentence, but not when there was a Phonological Phrase boundary between the syllables (e.g. *pay#permit*). However, since these studies did not focus on the role of DIS, their stimuli included a range of cues which could have contributed to the detection of the boundaries, including pitch and duration as well as potential DIS cues in segmental articulation.

T. Cho et al. (2007) appears to be the first study that attempted to isolate the effect of DIS on perception. They conducted cross-modal identity-priming experiments in which English listeners heard sentences containing a two-word sequence, e.g. *bus tickets,* leading to a temporary lexical ambiguity, e.g. *bust*. At the onset of the second word, e.g. *tickets*, listeners made lexical decisions to letter strings presented visually on a computer screen. These visual targets were either related to (a) pre-boundary words, e.g. *bus* (identical) vs. *mill* (unrelated), or (b) post-boundary words, e.g. *tickets* (identical) vs. *company* (unrelated). The two-word sequences either straddled a PW boundary or an IP boundary, as in *John forgot to buy BUS TICKETS for his family* (PW-initial condition) versus *When you get on the BUS, TICKETS should be shown to the driver* (IP-initial condition). The initial CV of the second word (e.g. [tɪ]) was spliced from another token of the sequence in IP- or PW-initial position. It was hypothesised that in both IP-boundary and PW-boundary sentences, splicing an IP-initial CV would lead to stronger priming effects than splicing a PW-initial CV, because greater DIS is expected to assist lexical segmentation and facilitate recognition of the words.

The results were mixed. Of the four different experimental conditions (2 host sentences x 2 groups of visual targets, either related to the pre- or post-boundary words), an expected difference in the priming effects between same-spliced and cross-spliced conditions was found only when related targets were pre-boundary words and when the host sentences involved a PW-boundary. These results were interpreted to suggest that the benefit of DIS is most clear under two conditions: (a) when there is a temporary lexical ambiguity (*bus* vs. *bust,* as opposed to *ticket* vs. *icket* [a non-word]), and (b) when other prosodic cues to prosodic boundaries are weaker. To expand on (b), when visual targets were post-boundary words, no effect was found possibly because competitors beginning with [tɪ] all received the benefit of DIS. Even when related targets were pre-boundary words, the effect of the additional cues from DIS was less noticeable when the two words straddled an IP boundary, presumably because there were stronger cues to an IP boundary, such as boundary tone and final lengthening. In contrast, splicing IP-initial CVs into PW-boundary contexts may significantly help segmentation as prosodic cues to PW-boundaries are relatively weak. This provides evidence that, in certain contexts, DIS can aid speech segmentation and facilitate lexical access.

According to T. Cho et al.'s (2007) prior examination of the stimuli, however, the acoustic cues available to the listeners were duration and amplitude differences in the IP-initial and PW-initial CV sequences. Cues that are strictly related to segmental articulation, such as spectral differences in /s/ or the amount of nasality have not been reported. Thus, it remains unknown whether listeners attend to and exploit variation of *segmental articulation* at domain-initial positions.

4.5 Perception of the lenis and nasal stops

Having reviewed the literature on the perceptual role of DIS in general, this section will now focus on the perception of the specific segments in Korean. To discuss the lenis stops first, an interesting aspect of the perception of Korean lenis stops is the crucial role played by the prosodic context. Baker (2002) conducted a phoneme identification test for the lenis, fortis, and aspirated stops by splicing a domain-initial CV syllable into the initial position of a different prosodic domain. The misrecognition rates were found to be highest for the lenis stops compared to the other stop categories. In particular, a majority of listeners perceived underlying U-initial and IP-initial lenis tokens as aspirated phonemes when spliced into a Syllable-initial position. This is unsurprising given the well-known rule of *lenis stop voicing* (reviewed in Section 3.4.2) and the DIS effects (reviewed in Section 2.3.1.1); the lenis stops are voiced AP-medially but voiceless AP-initially, with longest VOT in U- and IP-initial position. Thus, when an aspirated realisation of the lenis stop is spliced into a low prosodic position, listeners are likely to perceive an aspirated stop.

Moving on to the nasal stops, Young Shin Kim (2011) conducted a similar phoneme identification task for Korean nasal consonants with Korean and English listeners. The findings were as follows. When a denasalised segment was same-spliced into the initial position of a **C**VCV sequence, Korean listeners perceived a nasal phoneme whereas English listeners heard a plosive phoneme. On the other hand, when a denasalised segment in the same context was cross-spliced to the medial position of a V**C**V sequence, both Korean and English listeners heard a plosive. Conversely, when a lenis plosive was moved from the medial to initial position, Korean listeners identified it as a nasal while English listeners identified it as a plosive. These results indicate that Korean listeners' interpretation of the phonetic variations of a nasal phoneme is crucially dependent on the prosodic context it is heard. In this particular case, the presence of a preceding Utterance boundary (#**C**VCV) as opposed to a Syllable boundary (V**C**V) and vice versa is able to sway one's perception of a nasal and a plosive to a different phoneme.

In addition to its effect on phoneme identification, prosodic context has also been found to play a role in lexical access. Shin and Tremblay (2018) conducted cross-modal identity priming tasks in a structure similar to Cho et al. (2007) discussed in Section 4.4.

Korean listeners were visually presented with two types of target words. In Experiment 1, they saw words beginning with a nasal (e.g. *noru* 'roe deer') and in Experiment 2, they saw words beginning with a lenis stop (e.g. *toru* '(in baseball) stealing a base'). In the experimental condition, the auditory primes were words that begin with a denasalised nasal and rhyme with the target (e.g. /*noru*/). In the control condition, they were words that are phonologically and semantically unrelated to the target (e.g. /tʃote/ 'invitation'). These primes were either heard in AP-initial position or in AP-medial position (recall that Korean denasalisation is found to occur in AP-initial position).

In Experiment 1, the recognition of nasal-initial target words was facilitated by the denasalised primes in AP-initial position, but not in AP-medial position. These results suggest that Korean listeners process denasalised nasals as nasal phonemes only in licit contexts (i.e. in AP-initial position). However, in Experiment 2, the recognition of the plosive-initial words was not facilitated by the denasalised primes in both prosodic contexts. This is unexpected given Kim's (2011) results that denasalised nasals were categorised as plosives in AP-medial position. This lack of a priming effect in Shin and Tremblay (2018) was explained as a consequence of the acoustic differences between AP-initial denasalised nasals and AP-medial lenis stops. An acoustic analysis of their raw stimuli revealed the following statistically significant differences. Compared to AP-medial lenis stops, AP-initial, denasalised nasals had a higher RMS amplitude, higher RMS burst energy, and a shorter burst duration. Considering these acoustic differences, the inconsistent results between Kim (2011) and Shin and Tremblay (2018) seem to be due to the different experimental setup. That is, in a forced identification task, Kim's participants may have categorised denasalised nasals as plosives in AP-medial position because the acoustic features were *more* similar to a plosive than a canonical nasal as expected from the prosodic context. Nevertheless, this similarity may not be sufficient to trigger priming effects during a task involving lexical access.

With a different focus, J. Yun and Arai (2020) investigated the effects of different acoustic characteristics on Korean listeners' evaluation of the goodness of a given stimulus as /na/ in an Utterance-initial position. Their stimuli were generated with a Klatt synthesiser by manipulating various parameters including the amplitude of voicing, extra tilt of voicing spectrum, open quotient, F0-F3, bandwidths of F1-F5, frequency of nasal pole, and frequency of nasal zero. These parameters were adjusted to create continua representing prevoicing, nasal murmur, and vowel nasalisation. The results showed that, despite the tendency for U-initial nasals to show reduced nasality and voicing in production, longer prevoicing and longer nasal murmur generally increased the /na/-likeness rating in perception. However, there was a high degree of variation among the participants who broadly showed two different patterns. Two-thirds of the participants accepted the stimuli with no nasality in the consonant as good exemplars of /na/. These participants also preferred oral vowels to nasalised vowels, in line with the finding by Kim (2011) than vowel nasalisation is weaker after a denasalised nasal than after a sonorant nasal. In contrast, the other one-third of the participants gave higher ratings to the stimuli with nasality in both the consonant and the vowel. Such a grouping of patterns is in line with the view that Korean denasalisation is an ongoing sound change with a relative innovative group and a relatively conservative group (see Section 3.5). Another point to note from these findings is that the acoustic cues in both the onset nasals and the following vowels are important in perception, supporting the methodological decision in Chapter 6 to splice the entire CV syllable rather than just the target consonant.

The key finding from the studies discussed so far is that prosodic context determines how listeners perceive a given segment. In the perception study of the present dissertation, we will explore the exact opposite question: can the phonetic variation of a segment determine whether listeners perceive a certain prosodic boundary (e.g. AP) as opposed to another (e.g. PW)? The results of this perception experiment will contribute to our understanding of the role of prosody in speech perception and the phonetic markings of prosodic structures in speech.

4.6 Chapter summary

This chapter has summarised and discussed existing research relevant to the perception experiment in Chapter 6. The definition of prosodic parsing has been established in relation to the theoretical assumption of prosody made at the beginning of the dissertation. Then, the literature on the role of duration and pitch cues in prosodic parsing has been presented, with the conclusion that it is crucial to control their effects in the perception experiment. Next, previous studies on the role of DIS have been reviewed, identifying a gap in the literature which forms the motivation for the perception experiment in this dissertation. Finally, the review of the perception of Korean lenis and nasal stops showed that their perception is highly complex, being sensitive to the prosodic position of the segment and, in the case of the nasals, various acoustic parameters including the nasality of the following vowel.

Chapter 5.

A production study:

the phonological status and phonetic patterns of DIS

5.1 Chapter overview

This chapter will describe a production study which aims to answer the following research questions formulated in Section 1.3:

> **RQ1**. What are the patterns of prosody-sensitive phonetic variation of the lenis stops and nasal stops in Seoul, Busan, and Ulsan Korean?
> **RQ2**. What is the status with respect to the grammar (see Figure 1.1) of and the relationship among the lenis stop (de-)voicing rule, denasalisation, and DIS in the three varieties of Korean?
> **RQ3**. What is the mechanism underlying DIS – e.g. articulatory, physiological, phonological (see Section 2.4.2 and references therein)?
> **RQ4**. What is the best account of the prosodic variation of the lenis stops and nasal stops in Korean?

To answer these questions, this study compared the acoustic characteristics of the lenis stops and nasal consonants in a range of different prosodic positions, as produced by speakers from the three Korean cities, Seoul, Busan, and Ulsan. The subsequent sections will focus on the following aspects of the study: Hypotheses (5.2), Method (5.3), Statistical analyses (5.4), Results (5.5), and General Discussion (5.6). At the end, a summary of the chapter will be given (5.7).

5.2 Hypotheses

Based on the literature review in Chapters 2 and 3, hypotheses for the production study can be formulated as follows. The hypotheses in **H1** - **H7** are related to the more descriptive questions in **RQ1** and **RQ2**, and the rest of the hypotheses in **H8 – H11** are concerned with **RQ3** and **RQ4**, which are of a more explanatory nature.

H1. Both the lenis and nasal stops show the following cumulative effects of DIS as their position moves up the prosodic hierarchy:

a) Longer acoustic closure duration
b) A smaller percentage of voicing during acoustic closure (residual voicing itself does not show DIS effects)
c) Longer Voice Onset Time (VOT)
d) Longer total voiceless interval (voiceless portion of the closure + VOT)
e) For nasal consonants only, lower minimum nasal energy

H2. For the lenis stops, there is no regional variation.

H3. For the nasal stops, the DIS effects are greater in the order of Seoul, Busan, and Ulsan, following the model of hierarchical diffusion (Trudgill, 1974).

H4. The lenis stops show a categorical effect of the lenis stop (de-)voicing rule, regardless of the variety (see Section 3.4.2).

H5. The domain of the categorical process of lenis stop (de-)voicing is the AP.

H6. The nasal stops show a categorical effect of the denasalisation rule, at least in the more advanced varieties such as Seoul (see Section 3.5).

H7. The domain of the categorical process of denasalisation is the AP.

Hypotheses in **H8 - H11** test the duration-based account of DIS as reviewed in Section 2.4.2 (1). T. Cho and Keating (2001) found a robust correlation between linguopalatal contact and duration for Korean. They suggested that DIS is a consequence of the tendency for segment duration to be longer in a higher prosodic position, allowing articulatory targets to be fully achieved. According to this account, we can expect to find the following correlation between duration and the two acoustic correlates of DIS, closure voicing and nasality.

H8. The longer the closure duration, the lower the percentage of voicing during closure for the lenis stops.

H9. The longer the closure duration, the lower the level of nasality for the nasal stops.

The duration-based account is based on the assumption that DIS is not directly mediated by the prosodic structure itself, but only indirectly influenced by it through the variation in the duration of the segments in different prosodic positions. This leads to the following hypothesis.

> **H10**. DIS effects are best modelled as a function of closure duration.

On the other hand, the account of DIS based on rule scattering predicts that, for those speakers who show a categorical effect of lenis stop devoicing and/or denasalisation, there will also be an overlaid, gradient effect of DIS. The logic here will be explained further in Section 5.4.6.

> **H11**. DIS effects are best modelled as a function of the combination of prosodic position and closure duration.

Note that the two accounts are not mutually exclusive. Within the framework of the life cycle of phonological processes, the pattern in **H10** is compatible with a stage in which DIS effects are operating in the phonetic component and have not yet been stabilised into the phonological component. Then, the pattern in **H11** is compatible with the next stage in the life cycle, in which DIS has stabilised into a phrase-level phonological rule where the gradient effect of DIS continues to influence segmental articulation in a fine-grained manner.

5.3 Method

5.3.1 Recording materials

The recording materials were of the following form: Part A, consisting of 10 sets of four sentences, and Part B, consisting of seven sets of six sentences (speech samples for one set from Part B are available from https://osf.io/8kmqz). This resulted in a total of 82 sentences. All four sentences within each set in Part A were made up of the identical segmental string but had slightly different meanings due to a syntactic ambiguity. These differences in meaning can be distinguished by the prosodic structure of the utterance. See Table 5.1 below for an example.

The study took advantage of several features of Korean to create 10 sets of sentence minimal quadruplets. Korean is a pro-drop language in which subjects are frequently omitted. Possessive and nominative particles are also dropped in casual speech. As will be described shortly, subjects were instructed to drop these particles (e.g. /ɯi/ and /ka/ in Table 5.1) when reading the materials. In addition, Korean has a (subject-)object-verb order, which can be used to create an ambiguity with utterances with a vocative.

Table 5.1: Phonemic transcriptions of an example set from Part A of the recording materials, comparing four different prosodic contexts (full materials are available in Appendix A, https://osf.io/8kmqz). The bold and italicised /t/ show the target segment in each sentence.

(a) PW-initial /t/

$_{IP}\{_{AP}\{$ʌmma(-ɯi) ***t***alimi} kjʌlkuk ko͡tsaŋnasʌ pʌljʌs*ʌjo}
Mom(-POSS) iron finally broke throw away
'Did you throw away Mom's iron because it finally broke?'

(b) AP-initial /t/

$_{IP}\{$ʌmma(-ka) $_{AP}\{$***t***alimi} kjʌlkuk ko͡tsaŋnasʌ pʌljʌs*ʌjo]}
Mom(-NOM) iron finally broke throw away
'Did Mom throw away the iron because it finally broke?'

(c) IP-initial /t/ with no preceding pause

ʌmma, $_{IP}\{_{AP}\{$***t***alimi} kjʌlkuk ko͡tsaŋnasʌ pʌljʌs*ʌjo}
Mom(-VOC) iron finally broke throw away
'Mom, did you throw away the iron because it finally broke?'

(d) IP-initial /t/ preceded by a pause

ʌmma, (pause) $_{IP}\{_{AP}\{$***t***alimi} kjʌlkuk ko͡tsaŋnasʌ pʌljʌs*ʌjo}
Mom(-VOC) iron finally broke throw away
Mom, (pause) did you throw away the iron because it finally broke?

The prosodic boundaries commonly used before the crucial syllable /ta/ given the sentences in (a)-(c) are PW, AP, and IP, respectively. While a speaker can choose to read these sentences with a different prosodic phrasing, the use of these three boundaries helps distinguish the meanings of the three questions most clearly. Sentences (c) and (d) were identical except that the speakers were told to insert a short pause before /ta/ in (d). This is to examine if the patterns of domain-initial strengthening were different in the presence or absence of a pause preceding an IP-initial /t, n/.

The rationale for including the type-(d) sentences instead of sentences which test the level of the Utterance is as follows. Previous studies have indeed included the Utterance[25] as a potential prosodic domain above the IP and found that some speakers distinguished these two levels on a range of phonetic parameters (Fougeron & Keating, 1997; T. Cho & Keating, 2001; Keating et al., 2004). However, it is not always obvious from the studies whether the

[25] Note the debate surrounding the status of the Utterance (e.g. Nespor & Vogel, 1986; Selkirk, 1986; Wightman, Shattuck-Hufnagel, Ostendorf, & Price, 1992).

differences found between the Utterance and the IP are simply due to the presence or the absence of a pause. For example, T. Cho and Keating defined the Utterance to be the domain that is preceded by a pause greater than break index number 3 in K-ToBI (S.-A. Jun, 2000), triggered by an orthographic period. In turn, Break index number 3 is a strong phrasal disjuncture found at the edges of the IP, "with a strong subjective sense of pause, *whether it be an objective visible pause or only the virtual pause cued by final lengthening*". In other words, the Utterance was followed by a substantial pause, whereas the IP was followed by a short pause or final-lengthening. Thus, to tease out the role of a pause more clearly, I divided the IPs into two types, those preceded by an actual pause as well as final lengthening (d) and those preceded by final lengthening but not by an actual pause (c).

The availability of alternative prosodic structures for each of the sentences meant that there had to be a way of triggering the desired prosodic structure. In addition, as the sentences would look identical in the materials (in Korean orthography) given to subjects to read out, there had to be a method of signalling the differences in the intended meaning. To achieve this, the underlined parts shown in Table 5.1 were added to the reading materials in Korean orthography. Sentence (a) was marked with a possessive particle /ɯi/ in brackets and square brackets around the first two words to prevent the tendency of the speakers to insert an AP boundary before /talimi/. Sentence (b) was marked with the nominative particle /ka/ in brackets, (c) and (d) with a comma, and (d) with the English word *pause* in brackets. These were additional to spoken instructions to be described below, including that they should not read what is in the brackets.

Five of the sets in Part A tested /t/ and the other five tested /n/. The /t/ and /n/ sets were matched for the following vowels /a, ɛ, ʌ, o, u/, as the following vowel is known to influence the degree of denasalisation and devoicing of /n/ (Yoo, 2016). The following vowel may also affect VOT of /t/, given the cross-linguistic tendency in which VOT of a stop is longer before tense high vowels than before lax low vowels (Chao & Chen, 2008; Klatt, 1975; Port & Rotunno, 1979; Weismer, 1979).

As shown in Table 5.2, the seven sets of six sentences in Part B[26] were designed similarly to those in Part A, but included two additional positions for the target segment: morpheme-initial and morpheme-medial. In the prosodic hierarchy, both of these positions would simply be Syllable-initial. These levels were included in order to investigate the full

[26] During the design stage of this study, Jeon informed me of her previous work (2008) which made use of similar materials. With her permission, I have modified the two sentences from her study to create Sentences 15 and 16 in Part B.

spectrum of possible domains of a process in the life cycle of phonological processes (Bermúdez-Otero, 2010, 2015), as described in Section 1.2.2.

Table 5.2: Phonemic transcriptions of an example set from Part B of recording materials comparing six different contexts. The bold and italicised /k/ show the target segment in each sentence. Speech samples are available from https://osf.io/8kmqz.

(a) Morpheme-medial /k/					
$_{\text{IP}}${$_{\text{AP}}${ʌlin	a***k***a}	sanɯn	t͡sipe	tuko	was*tako}
Young	baby	live	house	leave	came
‘You left it at the house where a young baby lives?’					
(b) Morpheme-initial /k/					
$_{\text{IP}}${$_{\text{AP}}${t͡sijʌni-***k***a}	sanɯn	t͡sipe	tuko	was*tako}	
Jiyeon-SBJ	live	house	leave	came	
‘You left it at the house where Jiyeon lives?’					
(c) PW-initial /k/					
$_{\text{IP}}${$_{\text{AP}}${t͡sijʌni(-ɯi)	***k***asanɯn}	t͡sipe	tuko	was*tako}	
Jiyeon(-POSS)	lyrics.	house	leave	came	
‘You left Jiyeon’s lyrics at home?’					
(d) AP-initial /k/					
$_{\text{IP}}${t͡sijʌni(-ka)	$_{\text{AP}}${***k***asanɯn}	t͡sipe	tuko	was*tako}	
Jiyeon(-NOM)	lyrics.	house	leave	came	
‘Jiyeon left her lyrics at home?’					
(e) IP-initial /k/ with no preceding pause					
t͡sijʌni,_	$_{\text{IP}}${$_{\text{AP}}${***k***asanɯn}	t͡sipe	tuko	was*tako}	
Jiyeon(-VOC)	lyrics	house	leave	came	
‘Jiyeon, you left the lyrics at home?’					
(f) IP-initial /k/ preceded by a pause (P)					
t͡sijʌni, (pause)	$_{\text{IP}}${$_{\text{AP}}${***k***asanɯn}	t͡sipe	tuko	was*tako}	
Jiyeon(-VOC)	lyrics	house	leave	came	
‘Jiyeon, you left the lyrics at home?’					

The target consonants and the following vowels in Part B were /ka/, /to/(2), /ma/, /na/(2), and /ni/. Due to the difficulty of devising a segmental string which can be read with a given segment in six different positions, the place of articulation and the following vowel environment could not be matched perfectly between the nasal and lenis stop groups. For the same reason, not all sentences within a set had the identical segmental string. Five sets had the same segmental string from (b)-(f), and the remaining two had one sentence within (b)-(f) which had the same segmental string up to and including the target syllable but differed in the remaining part. Sentences (a) and (b) were slightly different in most sets, differing up to three syllables immediately before the target segments. When the syllables immediately before the target consonant were different, effort has been made to make them as phonetically close as possible. For example, (a) and (b) in Table 5.2 are different in the first three syllables, but the syllables immediately preceding the target consonant are phonetically similar, which are /na/ and /ni/, respectively. All sentences within each set had exactly the same syllable count, except one set which had one sentence that was longer than the rest of the sentences by one syllable. Lastly, all sentences had the same vowel following the target consonant within each set. A relatively relaxed design of Part B was possible because only Part A was used as the materials for the perception experiment in Chapter 6, for which having the same underlying segmental string was crucial.

In the rest of this chapter, the following short forms will be used to refer to the prosodic positions relevant in the study: Mm (Morpheme-medial), Mi (Morpheme-initial), PWi (Phonological-Word-initial), APi (Accentual-Phrase-initial), IPi (Intonational-Phrase-initial with no preceding pause), and P (Postpausal, Intonational-Phrase-initial with a preceding pause).

5.3.2 Speakers and procedure

A total of 30 university students (mean age: 22.3, range: 18-26, SD: 2.3) participated in the paid recording with five males and five females from each of the three cities, Seoul, Busan, and Ulsan. These students were recruited through emails and flyers within their universities, which were Hanyang University, Tongmyong University, and University of Ulsan, respectively. Twenty-four of the speakers had lived in their respective cities for their entire lives, and four speakers had spent up to two years in a nearby area. The remaining two were a 23 year-old female who lived in Pyeongtaek, a city in Gyeonggi Province which surrounds Seoul, from the age of 12 to 18, and a 22 year-old female who lived in Suwon, the capital of Gyeonggi Province, from 7 to 20. Given that even a second language acquired after puberty may influence one's L1, such as perception and production of VOT (Stölten, 2013), subjects with little or no experience of living abroad were recruited. Two speakers from Seoul and one speaker from Busan lived either in England or the US for less than 10 months and the remaining subjects have not lived abroad.

The recordings of Seoul Koreans were made in a sound-treated booth at Hanyang Phonetics and Psycholinguistics Laboratory with a Tascam DR-680 multi-channel digital recorder and a Shure VP88 condenser microphone at a bit-depth of 24 and a sampling rate of 44.1kHz. Due to the difficulty of finding a sound-proof room in Busan and Ulsan, recordings took place in a quiet room with a Nagra ARES-M II digital recorder at a bit-depth of 16 at a 44.1kHz sampling rate.

The speakers read the recording materials at two self-selected rates, *moderate* and *slightly fast*, in two repetitions, giving a total of 9840 sentences for analysis (30 speakers × 82 sentences × 2 speech rates × 2 repetitions). Speaking rate was varied in order to elicit a wider range of duration for the target segments. This would allow us to effectively test the duration-based hypotheses described in **H8 - H11** (Section 5.2). Of these sentences, 1680 sentences (30 speakers, 14 sentences × 2 speech rates × 2 repetitions) involved morpheme-medial and morpheme-initial realisations of the target segments, and the remaining 8160 sentences were for PW, AP, and IP-initial contexts.

The speakers were instructed to read the sentences naturally as if talking, and not to emphasise particular words in an utterance to prevent complication by the effect of narrow focus. In order to avoid confusion and to elicit consistent use of the intended prosodic structures, all the type (a) sentences of the 10 sets were listed first, followed by all type (b) sentences, and so on. In other words, the intended order of recording was 1a normal, 1a fast, 2a normal, 2a fast, ..., 1b normal, 1b fast, 2b normal, 2b fast, … before moving onto Part B.[27] Prior to recording, the experimenter also demonstrated two example sets which used /s/ and /w/ as target consonants after explaining the situations in which each of the sentences would be appropriate using visual aids. At this point, the subjects were also trained not to pronounce the particles in brackets. Then, the subjects were given sufficient time to familiarise themselves with the materials, and asked to read the sentence again when they made mistakes.

Despite these measures to ensure the production of the intended prosodic contours, the four to six-way contrast in the prosodic phrasing turned out to be challenging for some speakers. These speakers were interrupted a couple of times in the beginning to briefly demonstrate the desired prosodic phrasing again. In cases when this did not immediately result in the intended phrasing, no further instructions were given so as not to influence the speakers' natural productions.

Due to the minimal interruption, there were different tokens missing from different speakers. Some speakers struggled to produce PW-initial tokens, producing them in AP-initial

[27] By not randomising the order of reading, there is a possibility of bias, perhaps due to a global declination effect over the course of a recording session in which there is generally more articulatory strengthening in the beginning when a speaker may have more energy. However, this is not a significant problem as it will always work against the hypothesis that more strengthening will be found in a higher domain.

position instead, an acceptable alternative phrasing given the meaning of the sentence. Others produced the sets intended to elicit an AP boundary with an IP boundary. Quite a few speakers either produced both IP conditions with no preceding pause or produced both with a preceding pause.

There were sentences prone to a particular phrasing other than the intended. For example, most speakers produced Sentence 3a with an AP boundary instead of a PW boundary. This is presumably due to the number of syllables. A typical accentual phrase in Korean contains between two and five syllables (S.-A. Jun, 1998, p. 192). In a study of newspaper reading, only 7.8% of 165 APs had more than six syllables (Chung et al., 1996). Since the intended AP in 3a contains seven syllables, the speakers were more likely to break it up into two APs.

3a. $_{\text{IP}}\{$ $_{\text{AP}}$\{op*a **t**ʌpɯlpeisɯ\} pilljʌ osjʌs*ʌjo$\}$

brother double bass borrow come

'Did you borrow and bring you brother's double bass?'

In Sentence 16c, a few speakers produced a voiceless vowel [i̥] after the voiceless [s*], making it difficult to measure closure duration and residual voicing into closure for the target consonant /t/.

16c. $_{\text{IP}}\{$ $_{\text{AP}}$\{a͡tsʌs*i **t**ok*ika is*-ɯmjʌn t͡sot͡sʰi ankʰes*ʌjo$\}$

mister axe have-if good would not

'Wouldn't it be good if we had Mister's axe?'

Given the various problems, it was unlikely that both repetitions were consistently analysable and in the intended prosodic phrasings. Thus, only one pair of utterances, one at a normal and one at a fast pace, were analysed for each sentence. The procedure for selecting up to two tokens per sentence was as follows. First, the prosodic position of each token was determined in a post-hoc manner, following the criteria in Table 5.3. The exact criteria were the same as those from T. Cho and Keating (2001, p.161) for the Syllable (called Morpheme-initial in the present study), Phonological Word, and Accentual Phrase boundaries. As explained earlier, Intonational Phrase boundaries were further divided into two types as shown in Table 5.3. There were speakers who arguably produced Utterance boundaries, instead of IP boundaries with a preceding pause. That is, they produced the kind of long pause which would normally be triggered by a full stop. The definition of a Utterance boundary used in T. Cho and Keating was that it was associated with a disjuncture greater than break index number 3 which corresponds to "a strong phrasal disjuncture with a strong subjective sense of pause, whether it be an objective visible pause or a virtual pause cued by final lengthening" according to K-ToBI labelling conventions (S.-A. Jun, 2000). As such, it is quite vague when an IP boundary with an actual pause becomes an Utterance boundary. In addition, the existence of the Utterance as a prosodic unit separate from the IP is debated (see Shattuck-Hufnagel & Turk, 1996, pp. 209 - 210). On practical grounds, it is not possible to measure acoustic closure duration and residual voicing for tokens preceded by a pause of any length anyway. Thus, such tokens were accepted as type (f), the IP-pause condition. Since morpheme-medial and morpheme-initial tokens were invariably realised within a PW, they did not require post-hoc judgement of their prosodic position.

Table 5.3: A modified version of the criteria for prosodic coding as used in T. Cho and Keating (2001, p. 161), based on K-ToBI (S.-A. Jun, 2000).

(a) Morpheme-medial Initial consonant of the last syllable of a content word, within a PW
(b) Morpheme-initial (Syllable-initial) Initial consonant of a suffix, within a PW
(c) PW-initial No tonal specification. No perceived break (break index number 1)
(d) AP-initial LHLH or HHLH underlying phrasal tones[28] A minimal phrase disjuncture
(e) IP-initial with no preceding pause Boundary tone Subjective sense of pause due to final lengthening but with no visible pause
(f) IP-initial /k/ preceded by a pause Boundary tone Final lengthening with a visible pause (acoustic silence)

The first repetition was analysed this way, discarding the tokens that were not in the desired prosodic positions. Then, the missing tokens were analysed in the second repetition. If these tokens were also in an unintended position, I turned to the tokens that were inadvertently produced in the intended positions from different sentences. For example, if a

[28] As mentioned in Section 3.2, LHLH or HHLH are the underlying tonal contours which are only fully realised in AP's with four or more syllables. For AP's with fewer syllables, either the medial L, the medial H or both are undershot.

token in 1a intended for a PW-initial token was realised in an AP-initial position, this was used when all tokens in 1b were inappropriate. After this process, there was a total of 4137 tokens for analysis, with 1417, 1343, 1377 tokens from Seoul, Busan, and Ulsan, respectively.

5.3.3 Acoustic measurements

The acoustic measurements in the present study were largely based on the study of DIS in Seoul Korean by T. Cho and Keating (2001), which was reviewed in Section 2.3.1. However, the exact set of measurements and the method of segmentation was slightly different. First, T. Cho and Keating did not take measurements of acoustic closure duration for /t/ due to frequent lenition. For the present study, it was deemed helpful to include the cases in which a full closure was not achieved. Recording the lenited realisations would provide a complete picture of the realisation of the lenis stops and contribute to the discussion of the mechanism underlying DIS. T. Cho and Keating also did not consider VOT and the total voiceless interval for /n/. This is presumably due to the assumption that nasal consonants do not show a voicing lag or a voiceless portion during the closure. Nevertheless, these measurements were taken since more recent studies of Korean denasalisation have suggested that these measures are indeed relevant (Kim, 2011; Yoo, 2015a) (Section 3.5).

Second, RMS burst energy and vowel duration in domain-initial CV were not measured in the present study because they did not show clear effects for /n, t/ in T. Cho and Keating's study. As there were a high number of parameters to investigate in this study, the intention was to focus on those measurements that were known to show clear DIS effects.

Third, in T. Cho and Keating's study, the level of nasal energy was indirectly measured by RMS acoustic energy at the lowest point during a nasal consonant. Since there are limitations of this method, alternative methods of nasality measurement will be tested and compared in Section 5.3.8. In the current section, I will first discuss the measurement of acoustic correlates of DIS other than nasality. All tokens were segmented manually in Praat (Boersma & Weenink, 2018) following the method outlined below. Then, the duration of each segmented interval was extracted automatically using the segment duration script by Lennes (2002). The automatic extraction was always double-checked by a comparison between a few hand measured values with the extracted values and a comparison between the hand-recorded labels and the extracted labels.

(1) Acoustic closure duration (ms)

Acoustic closure duration was taken as an indirect measurement of articulatory closure duration[29], which is the period of occlusion in the oral cavity. A common method of segmenting an intervocalic oral stop is to measure from the offset of F2 in the preceding vowel up to the beginning of a stop burst (e.g. T. Cho & Keating, 2001; Penney, Cox, & Szakay, 2019). While this method has the advantage of being relatively objective and simple, it can be unreliable since F2 can survive during the initial portion of a closure when voicing is sufficiently strong. In this study, I use an alternative method involving a holistic approach, looking for points of the sharpest discontinuity in the spectrum and waveform, corresponding to the opening or the closing of the vocal tract (e.g. Jacewicz, Fox, & Lyle, 2009; Kraehenmann & Lahiri, 2008). More specifically, the beginning of a closure was determined by the point of a sudden loss of energy in the high frequency region in the spectrum and a sharp downstep in amplitude in the waveform. As described in Machač & Skarnitzl (2009, p. 58), the loss of high frequency components in the spectrum is also reflected in the absence or reduction of a "spiky" or "hairy" character associated with high frequencies in the waveform. The end of a closure was straightforwardly the beginning of a stop burst. When an oral stop was lenited to a fricative, an approximant, or even a close vowel, the case was separately recorded. It was nevertheless labelled for the ease of automatic extraction of the intervals but the closure duration was later replaced by zero in R (R Core Team, 2015) at the data pre-processing stage.

For intervocalic nasal stops, T. Cho and Keating (2001) measured nasal duration, from the beginning to the end of the nasal energy. However, as the Korean nasal stops are often realised as an oral stop without any nasality, this method was not always applicable. For denasalised nasals, I measured the period of oral closure duration, following the method described above. In cases where nasal consonants were realised with nasality throughout the closure, the following cues were used to segment the closure of a nasal consonant: reduction of amplitude, decrease in high-frequency intensity, damping of vowel formants, simpler waveform without a spiky or hairy character associated with high frequency energy, anti-formants, and spreading of formant bandwidths (Machač & Skarnitzl, 2009, Chapter 4). As with the oral stops, nasal consonants realised as an approximant or vowel were recorded as cases of lenition and assigned the value zero for their closure duration.

---

[29] As one of the aims of the present study is to propose an account of the articulatory mechanism underlying DIS, the ideal data would be articulatory closure duration obtained through such instruments as Electropalatography. However, this is more suitable for studies with a small number of subjects because a new palate must be specially made for each subject. In addition, the fieldwork component of this study made it difficult to conduct an articulatory study involving a sophisticated instrument.

Finally, for postpausal stops, regardless of whether they were oral or nasal, the measurement was not taken since the beginning of the closure could not be determined due to the preceding silence.

(2) Residual voicing (ms)

Voicing during stop closure was measured for both the lenis and nasal stops from the beginning of closure to the end of periodic glottal pulsing as seen from the waveforms and the low frequency voicing bar in the spectrogram.

(3) Voice onset time (VOT) (ms)

VOT was measured from the beginning of the release burst to the voice onset, determined by zero-crossing at the upward swing of the first glottal pulse in the vowel, following H. Lee and Jongman (2012, p. 151) and Kang and Nagy (2016, p. 255). This probably leads to an early point of voice onset compared to alternative criteria found in the literature. Fischer-Jørgensen and Hutters (1981) identified three different methods for delimiting the start of a vowel: (a) at voicing start, (b) at the start of F1, (c) at the start of higher formants. For our purposes, what is important is the point at which vocal fold vibration begins rather than where the vowel begins. This is because we are measuring VOT primarily to understand how DIS affects vocal fold tension. Thus, our method can be seen as corresponding closest to method (a).

On the other hand, Lisker and Abrahamson (1964, pp. 416 - 418) appear to have used method (b) as they did not include what they call “edge vibrations”. This is because the edge vibrations were found to be inaudible in perceptual experiments. This point does not concern us here as we are interested in the production phenomena themselves, regardless of the perceptual effects. However, according to Fischer-Jørgensen and Hutters, these edge vibrations do not always show up clearly in spectrograms, unlike in oscillograms (or waveforms). Therefore, in practice, researchers who take measurements based on spectrograms alone rarely rely on the edge vibrations, and use method (b) to delimit the voice onset.

In studies including T. Cho and Keating (2001) and T. Cho et al. (2002), VOT was taken from the stop burst to the voice onset of the second formant in the following vowel, including breathy voicing in the VOT. This method of delimiting the start of a vowel corresponds to method (c). Fischer-Jørgensen and Hutters recommend using method (c) as the beginning of a “vowel” because, among other reasons, it gives the most regular temporal relations. Nevertheless, our study measured VOT strictly up to the voicing onset, rather than the model-voiced vowel onset. This way our VOT measurement reflects the period of complete voicelessness. Figure 5.1 contrasts these two methods of measuring VOT: vot1 which is adopted in this study, and vot2 which is adopted by T. Cho and colleagues. In cases where there were more than one stop bursts, as in a double burst common to velar plosives, the first instance was taken as the start of VOT.

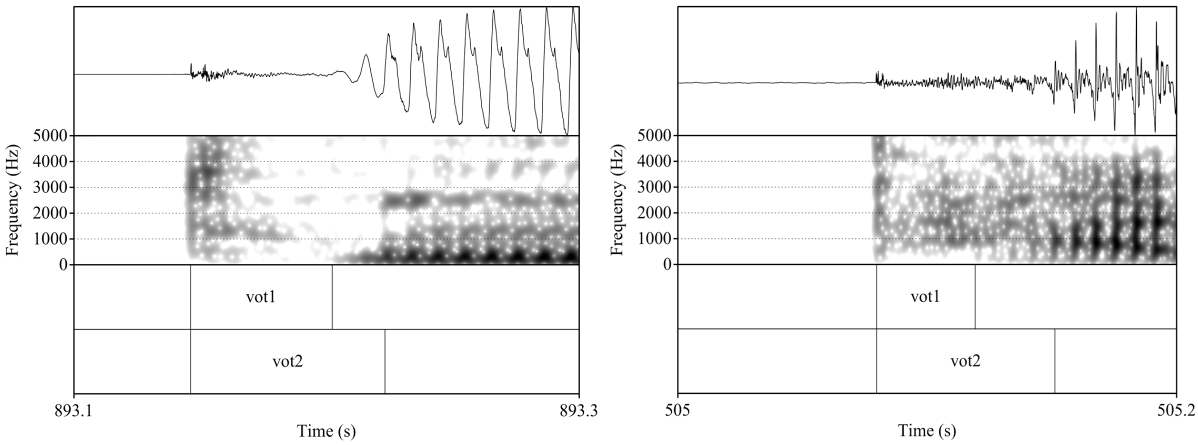


Figure 5.1: Alternative methods of measuring VOT.

(4) Prevoicing (ms)

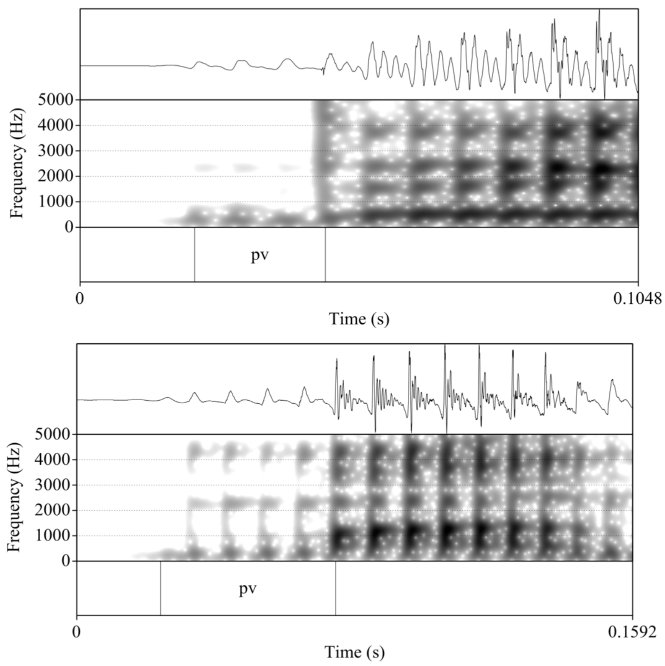


Figure 5.2: Measuring prevoicing in a denasalised nasal (top) and a sonorant nasal (bottom).

Prevoicing was measured for the tokens of the nasal phonemes in postpausal position. Figure 5.2 demonstrates the method. For denasalised nasals, as illustrated in the top figure, the prevoicing interval was measured from the upward swing of the first glottal pulse to the beginning of the release burst. For sonorant nasals, as exemplified in the bottom figure, this interval corresponded to the voiced portion of the consonant. In the absence of a release burst, the end of "prevoicing" was marked at the onset of the following vowel, as indicated by a full formant structure.

(5) Utterance duration (s)

Utterance duration was measured from the onset of the first syllable to the offset of the last syllable of each utterance, including pauses that occur naturally as part of an utterance, e.g. between two IPs. The onset and the offset were defined as the point at which acoustic energy first appears and completely disappears from the spectrogram.

Using the measurements above, six further measures were derived.

(6) Speaking rate (syllables/s) = *number of syllables* ÷ *utterance duration (s)*

Speaking rate was calculated by dividing the number of syllables in an utterance by the duration of the utterance in seconds.

(7) Rate normalised closure (ms) $= \boldsymbol{closure} \times \frac{\boldsymbol{utterance\ speaking\ rate}}{\boldsymbol{average\ speaking\ rate}}$

Even though the speakers were instructed to read the sentences in two speaking rates, normal and fast, each speaker differed greatly in how much they distinguished these two modes and how consistent they were within each mode. To control for the effect of different speaking rates, the acoustic closure duration of the target consonant was multiplied by the ratio of the speaking rate of the given utterance to the average speaking rate across the entire dataset. The average speaking rate was calculated for the entire dataset and not for each individual because, this way, the differences between a slow speaker and a fast speaker could be corrected for, in addition to the within-speaker variation for the normal and fast condition. Furthermore, it was important to normalise using speaking rate and not utterance duration, because the number of syllables in the utterances varied widely from 9 to 19.

(8) Rate normalised residual voicing (ms) $= residual\ voicing \times \frac{\boldsymbol{utterance\ speaking\ rate}}{\boldsymbol{average\ speaking\ rate}}$

As reviewed in Chapter 2, T. Cho and Keating found that for the three oral stops /t, t^h, t*/, raw residual voicing remained relatively constant across prosodic context. This contrasted with raw closure duration which became longer in a higher position. Thus, it can be expected that residual voicing would not vary as a function of speaking rate, either. However, a Pearson's correlation of the present data showed that residual voicing was negatively, albeit weakly, correlated with speaking rate (Pearson's r = -0.28, p < 0.0001)[30], as shown in Figure 5.3. A more detailed analysis by consonant and prosodic position revealed that this is mainly driven by nasal consonants, as a majority of nasal tokens were fully voiced. Thus, for nasal tokens, residual voicing was often the same as closure duration. However, the lenis tokens in the IPi position also showed a weak negative correlation (Pearson's r = -0.1, p = 0.075). This suggests that residual voicing becomes shorter when speech rate becomes faster. Therefore, residual voicing was also normalised for the effect of different speaking rates, following the method in 7.

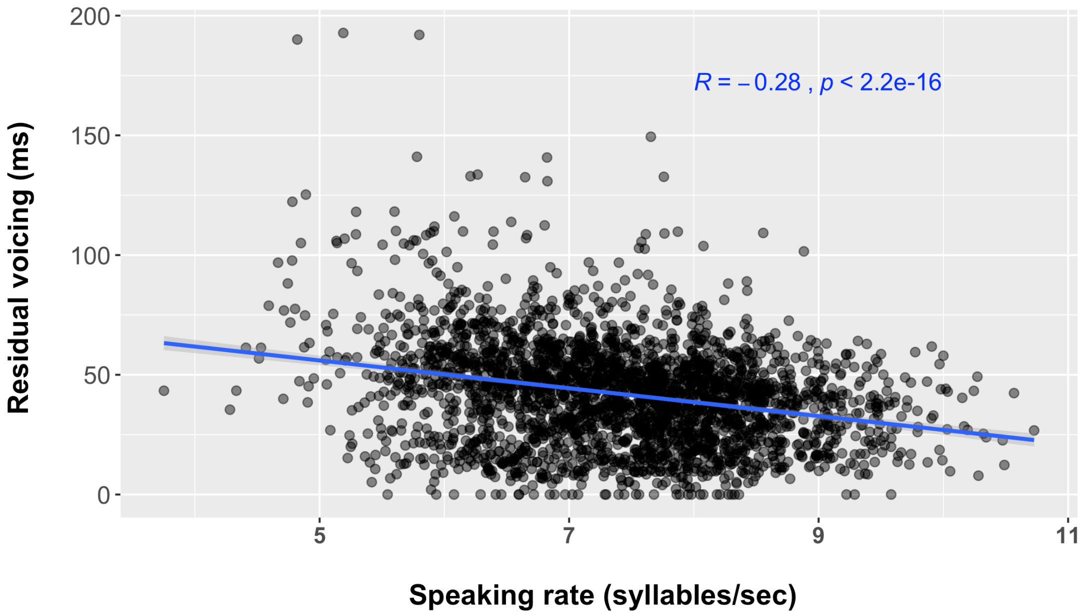


Figure 5.3: Correlation between speaking rate and residual voicing.

(9) $$\textbf{\% voicing during closure} = \frac{\textbf{residual voicing}}{\textbf{closure duration}} \times \textbf{100}$$

The percentage of residual voicing relative to closure duration was calculated as above. Lower values of this measure can be seen to indicate greater vocal tract tension. Further normalisation was not conducted as the division by closure duration cancels out the effect of speaking rate.

[30] Here, we follow the conventional cut-off points and treat a coefficient between 0.1 and 0.4 as a weak correlation (Overholser & Sowinski, 2008, p. 81; but see Schober, Boer, & Schwarte, 2018 for a recommendation again cut-offs).

(10) Rate normalised VOT (ms) $= VOT \times \frac{\textbf{utterance speaking rate}}{\textbf{average speaking rate}}$

Similar to 7 above, speaking rate differences were normalised by multiplying VOT by the ratio of the speaking rate of the given utterance to the average speaking rate of the dataset. This was based on the finding that VOT values of the present data were weakly correlated with speaking rate (Pearson's $r = -0.15$, $p < 0.0001$). This is also supported by Kessinger & Blumstein (1997)'s finding that VOT changed as a function of speaking rate for the stop categories with a long lag VOT, but not for those with a short lag VOT. Assuming that the Korean lenis stops belong to the categories with a long lag VOT, our result is consistent with Kessinger & Blumstein's.

(11) Total voiceless interval (ms) $=$ *closure duration – residual voicing* $+$ *VOT*

Total voiceless interval reflected the total duration of glottal opening, calculated by adding the voiceless portion of the stop closure and VOT. In T. Cho and Keating, this measure was found to distinguish the prosodic domains most consistently for /t^h/ than VOT or closure duration alone. In addition, this measure is helpful because the variation on closure voicing and VOT can be captured by a single parameter.

(12) Rate normalised total voiceless interval (ms)

$$= \textit{total voiceless interval} \times \frac{\textbf{utterance speaking rate}}{\textbf{average speaking rate}}$$

As in 7, total voiceless interval was multiplied by speaking rate ratio to correct for the variation in the speaking rates in each utterance.

(13) Rate normalised prevoicing (ms) $= \textit{prevoicing} \times \frac{\textbf{utterance speaking rate}}{\textbf{average speaking rate}}$

The duration of prevoicing was normalised for differences in speaking rates. Prevoicing became slightly shorter in faster speech (Pearson's $r = -0.077$, $p = 0.083$).

5.3.4 Room reverberation

In the recordings made in Busan and Ulsan, there was what appeared to be room reverberation. During stop closures and at the end of utterances, there was acoustic energy of which the spectral balance mirrored that of the preceding vowel. The auditory impression was also similar to the preceding vowel. External noise sources such as air conditioning could be ruled out as the noise was absent when there was no speech. These patterns were not found in the recordings from Seoul, which were the only recordings made in a sound-treated phonetic booth.

Two examples of suspected room reverberation during closure duration are shown in Figure 5.4. It is obvious that this poses a significant difficulty in the segmentation of residual voicing. Attempting to obtaining new, better-quality data was not considered an option, given the time and financial constraints as well as the practical difficulty of finding a sound-treated booth in Busan and Ulsan. Thus, the following method was used in order to determine the end of residual voicing as reliably as possible. In Figure 5.4, voicing shows a steady decay the early part of the closure before levelling out. This is particularly clear in the second example. The end of such steady decaying as visible in the waveform was taken to be the end of true voicing. The weak periodicity that continues after this point was seen as a consequence of room reverberation. This is because there is no reason for voicing to show a weakening trend only up to a certain point, given that it suggests that increasing pressure above the glottis is stifling the glottal flow. When it was difficult to pinpoint where voicing begins to flatten out, perhaps as in the first example, the shape of the periodic waves was used as a cue. The end of voicing was marked at the point at which the waveform starts to “lose shape” and ceases to look like a continuation of the periodicity of the vowel.

The pitch contour shown by the blue line supports the method described above. For example, the pitch in the preceding vowel in the first example is generally about 160 Hz, which drops to about 155 Hz around the point of closure, presumably as the vocal fold vibration is coming to a stop as the transglottal pressure drop is inhibited by the oral occlusion. Then there is a rising of the pitch after residual voicing, back to approximately 160 Hz. There is a similar rise of pitch in the second example. If it were not for the reverberation of the preceding vowel, there is no phonetic motivation for such a pitch step-up. This is particularly true for the first example, in which the pitch of the first vowel is higher than that of the second vowel, which rules out pitch coarticulation. However, since the exact pitch contour varies depending on the pitch range setting in Praat, the interval for residual voicing was determined primarily based on evidence in the waveforms.

While measurements of residual voicing based on the above method may not be as reliable as the measurements made under ideal conditions, these are not likely to significantly influence the hypothesis testing in the present study. This is because when cross-regional comparison is made, the measurements are compared across different prosodic positions within each region first before the differences across a given prosodic position pair are compared between regions. For example, regional comparison is made by comparing (a) the difference between residual voicing in /t/ in AP- and IP-initial position for Seoul and (b) the same difference for Busan. Thus, even if residual voicing was consistently underestimated for Busan but not for Seoul, this underestimation for Busan would be cancelled out when taking the difference between the measurements of residual voicing for AP- and IP-initial position.

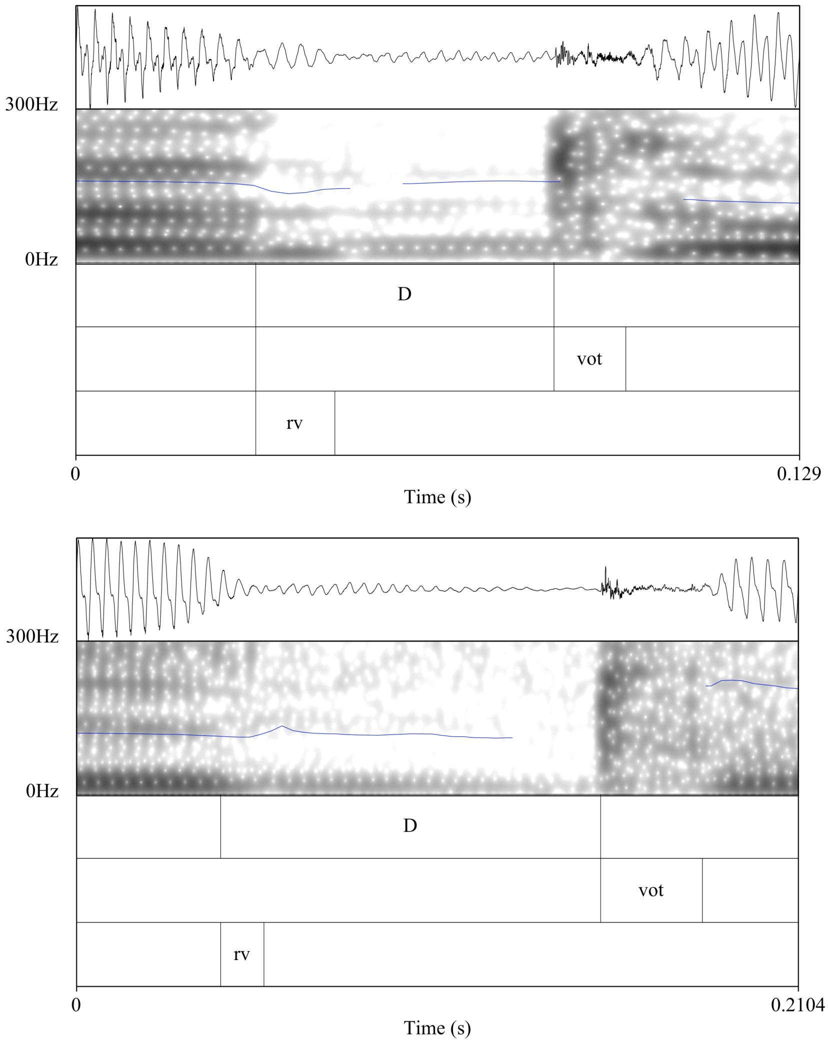


Figure 5.4: Examples of room reverberation. The three tiers are for closure duration, VOT and residual voicing, respectively. The blue line shows pitch. Speech samples to accompany these spectrograms and for most of the other spectrograms in this chapter are available from https://osf.io/8kmqz.

5.3.5 Potential measurements for nasality

The two previous sections discussed the methodology regarding the acoustic measurements other than nasality. As acoustic quantification of nasality is notoriously difficult, it requires more detailed discussion. Below, I will discuss four alternative measurements which have been used in the literature and outline the limitations of each of these measurements.

(1) Minimum intensity

T. Cho and Keating used minimum intensity, measured at the lowest point of the RMS acoustic energy profile during a nasal consonant, in order to infer its level of nasality. The energy at the valley was measured, not at the peak, because maximum energy was always found at the edges of a nasal, next to the surrounding vowels. While it is generally observed that the amplitude of a nasal consonant is higher than its oral counterpart, this measure may be problematic in the Korean case. Following the argument in Yoo and Nolan (forthcoming), the nasal consonant variation in Korean is extremely intricate and multidimensional, involving the interrelated aspects of nasality, voicing, and amplitude. For example, sonorant nasals in Korean can be realised with short and weak voicing, especially, but not exclusively, in a postpausal position. In such cases, non-nasals with strong voicing can show higher amplitude than sonorant nasals.

In T. Cho and Keating's study, Utterance-initial tokens were excluded from the measurement of nasal energy, because their minimum amplitude was always zero at the beginning, preceded by silence. While this avoids probably some of the most problematic cases in inferring nasality through amplitude for Korean nasal consonants, other limitations remain. In addition to the obvious issue of a substantial loss of data, the main problem is that amplitude is influenced by factors irrelevant to nasality, such as how loudly a subject speaks, and the distance between the mouth and the microphone.

(2) P0

As an alternative measurement, the amplitude of P0 was used in J. Jang, Kim and Cho (2018) to measure nasality in a nasal consonant. P0, also known as N1 (the first nasal formant), refers to a nasal pole or nasal peak that has been reported in studies of vowel nasalisation. In a nasalised vowel, P0 is observed between 250 and 450 Hz with an amplitude increase between 3 and 5.5 dB relative to its counterpart. P0 usually corresponds to the first or second harmonic (H1 or H2) (Chen, 1997, pp. 2360, 2362). In practical terms, this is often measured by the amplitude of the higher of the first two harmonics. The appearance of this nasal pole was attributed to new resonances introduced by the sphenoid and maxillary sinuses (Chen, p. 2362).

However, P0 is subject to interference by F1, especially in high vowels where F1 occurs in the 250 – 450 Hz range. In such cases, P0 is either affected or completely overwhelmed by F1 such that its identification and measurement become unreliable (Styler, 2015, p.24). A second nasal pole, P1, was discussed in Chen (p. 2362), occurring between 790 to 1100 Hz. While P1 has been used as an alternative for high vowels, Styler noted that P1 is also vulnerable to interference from F1 and F2. Finally, Schwartz (1968) discussed a higher nasal pole, P2, around 1250 Hz. Again, the range in which P2 is known to appear overlaps with that of F2.

The use of P0 is questionable for two main reasons. First, according to Styler's (2015) large-scale investigation of potential acoustical features of nasality in English and French, none of the amplitudes of the three nasal poles, P0, P1 and P2, were reliably correlated with nasality[31]. While P0 was significantly correlated with nasality in English, it did not reach significance in French (Styler, 2015, Section 4.5.1; see also Styler 2017b). On the other hand, P1 and P2 were correlated with nasality in French but not in English. One of the important conclusions of Styler (2015; 2017b) was that the acoustic nature of vowel nasalisation is highly language-specific. Thus, the reliability of P0 as a correlate of nasality in Korean requires further investigation. Second, all of these measurements have been tested in the context of vowel nasalisation, not to measure the level of nasality in nasal consonants. One cannot be certain whether measurements for vowel nasalisation can be extended to consonants, particularly in Korean where nasal consonants may show voicing that is very weak and/or short in duration.

(3) P0's prominence

Unlike the amplitude of P0, the prominence of P0 was a feature which was reliably correlated with nasality in both French and English (Styler, 2015). P0's prominence was measured by subtracting the average of the harmonics on either side from P0. This measure is more useful than the raw amplitude of P0 because the harmonics are enhanced *relative* to the surroundings and it is possible that P0 in the oral counterpart is still higher in absolute terms. Nevertheless, Styler recommends against using this measure as the correlation with nasality is still relatively poor compared to A1-P0, as will be explained below.

(4) A1-P0

Styler (2017b) found A1-P0 to be the single most robust correlate of nasality among the set of 22 features he investigated. A1 refers to the amplitude of the highest harmonic in the first formant of the vowel (F1). P0, as explained earlier, is the amplitude of the first nasal pole, which usually corresponds to H1 or H2. The measure exploits the phenomenon of damping, a well-known acoustic effect of nasalisation. In particular, damping of F1 is widely discussed in the literature, and is known to result in a sharp reduction of the amplitude of F1 relative to the surrounding signal (Pruthi & Espy-Wilson, 2004; Macmillan, Kingston, Thorburn, Dickey, & Bartels, 1999; Stevens, 1998; Schwartz, 1968). As A1 is expected to drop and P0 to rise with nasality, a smaller A1-P0 measure indicates greater nasality.

[31] For English, the values for each acoustical feature were compared between CVC-context vowels and CVN/NVC/NVN-context vowels. For French, the comparisons were simply made between an oral vowel and a nasal vowel.

As explained above, P0 is prone to interference from F1 in high vowels, as the lower range of F1 overlaps with P0's range. To deal with the problem, a compensation formula is devised by Chen (pp. 2363-4), using the relative position and bandwidths of nearby formants. However, Styler's (2015, p. 2477) study found A1-P0 to be the strongest cue for nasality even in high vowels. Although the compensation formula led to a small improvement for the English data set with 11 vowels, it did not lead to a better performance in the French data with three vowels. Taking simplicity into account, raw A1-P0 was concluded to be more preferable.

While A1-P0 cannot be applied directly to the present study which deals with nasal consonants, it may be possible to measure vowel nasalisation as an indirect reflection of the degree of nasality in the preceding nasal consonant. Unfortunately, the two studies which examined vowel nasalisation after a domain-initial nasal in Korean reported contradictory results (Young Shin Kim, 2011; J. Jang, Kim, & Cho, 2018). In an accelerometer study, Young Shin Kim (2011, Chapter 5) measured nasal energy in eight types of syllables, of which six are relevant for our purposes: CVC, N0VC, NVC, CVN, N0VN, and NVN, in which C indicates a lenis stop /p/ or /t/, N0 a denasalised nasal /m/ [b] or /n/ [t] and N a sonorant nasal /m/ [m] or /n/ [n]. Whether a nasal stop was denasalised or not was determined by the auditory-acoustic judgement of the author. Figure 5.5 shows a summary of the results from four subjects, taken from Kim (2011, p. 110).

The left-hand side of the figure shows the nasal energy during the three types of initial consonants. As the y-axis shows the negative of nasal energy, the results indicate that the nasal energy is clearly higher in the sonorant nasals compared to the denasalised nasals and lenis stops. While the denasalised nasals seem to show a slightly lower overall nasality, the difference from the lenis stops is neither big nor consistent. Crucially, vowel nasality in the middle of the figure suggests that, to some extent, the degree of nasality in the vowel is indeed reflected in the preceding consonant. The nasal energy of the vowel in NVC is by far the greatest of the three syllable types ending in an oral consonant. The vowel in N0VC shows slightly more nasality than the vowel in CVC, except in the last 30% of the vowel.

The vowels in NVN, N0VN and CVN also show different levels of nasal energy depending on the type of the initial consonant.[32] Nasal energy in these vowels is greater in the order of NVN, N0VN and CVN, the same as the order of the level of nasality in the preceding consonants. Towards the end of the vowels, however, the differences among the three types of syllables become smaller due to the coarticulation with the following nasal

[32] Note that these syllables are relevant for the present study because two of the segmental strings contain target syllables in this structure (/noŋ/ in Sentence 9 of Part A, and /man/ in Sentence 1 of Part B).

coda. A similar pattern can be observed in her aerodynamic study, which used an oro-nasal dual chamber mask (pp. 85-86). As in the accelerometer study, the results indicated that nasality in the beginning of the vowel generally reflected the nasality in the preceding consonant. Thus, although the correlation between vowel nasality and the nasality in the initial consonant is not perfect, Kim's results indicate that vowel nasalisation can be used as an indirect measure of nasality in the initial nasals.

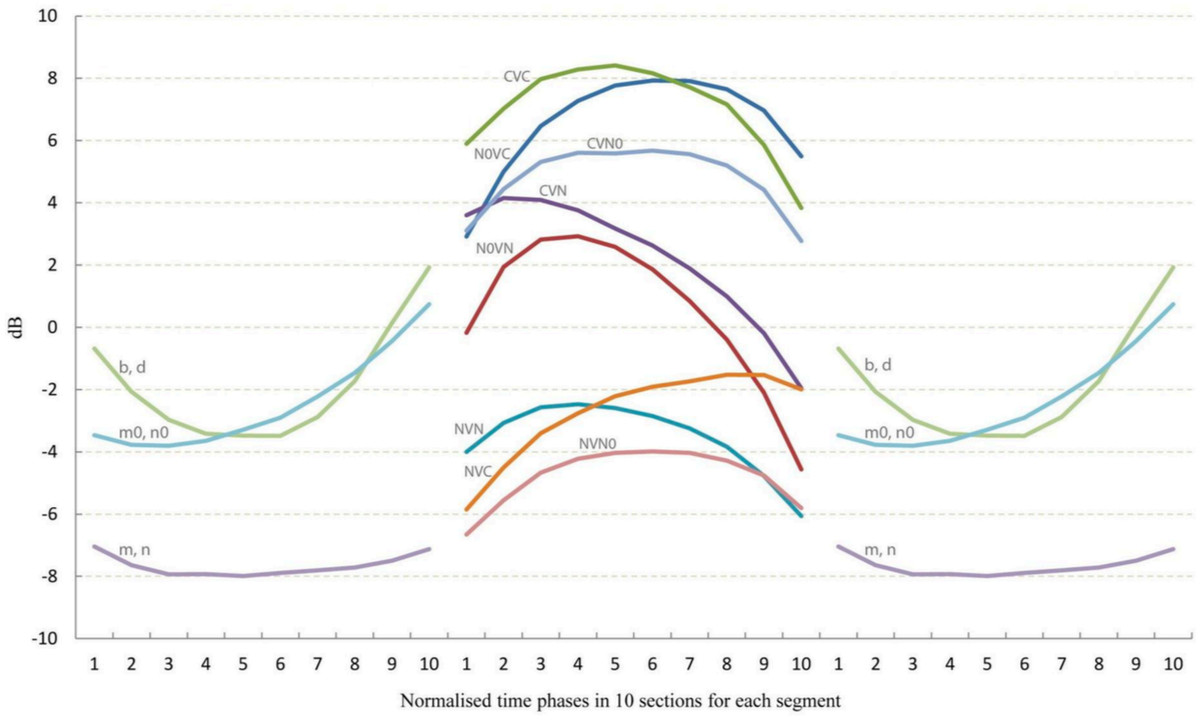


Figure 5.5: Mean nasal energy measurements in consonant-vowel-consonant sequences. The x-axis represents 10 time-normalised phases in the corresponding segment. The y-axis shows the value obtained by the formula $-\mathrm{N} = -(\text{accelerometer} - \text{microphone})$ in dB, where N represents the nasal energy. The nasal energy can be isolated by subtracting the energy picked up by the microphone from the accelerometer which detects nasal energy as well as the entire speech energy conveyed through tissue vibrations. The negative of N has been calculated for the visualisation of downward movement of the opening velum in nasalisation. Thus, a smaller value indicates greater nasal energy (Kim, 2011, p. 110).

In contrast, the study by J. Jang, Kim, and Cho (2018) suggests otherwise. Using the A1-P0 measure to quantify vowel nasalisation, they found that there was no boundary-induced modification of vowel nasalisation in #NVC syllables measured at 25%, 50%, and 75% time points. That is, even though they found significantly lower nasal energy and shorter duration for IP-initial nasals than for W-initial nasals, vowel nasalisation was not significantly different between the two groups. Given that IP-initial nasals are frequently realised with denasalisation whereas W-initial nasals are not (Kim, 2011; Yoo, 2015a), the lack of a significant difference between the vowel nasalisation contradicts the findings in Kim. Given the mixed results from the previous studies, the reliability of A1-P0 will be examined further using a subset of our data in Section 5.3.8.

5.3.6 Qualitative categorisation

From the discussion above, minimum intensity and vowel nasalisation appear to be the two best available measures of the level of nasality in initial nasal consonants. However, we have seen that even these two are not without limitations. In addition, there is a more fundamental problem in the attempt to quantify the phonetic variation of the Korean nasals: the multidimensional and dynamic nature of the phenomenon. As discussed before, the variation of the nasals in Korean cannot be captured by a nasality measure alone, as the nasals vary along the dimension of voicing as well as nasality. The finding of partial denasalisation also indicates that denasalisation is a dynamic event. For these reasons, Yoo (2015a) and Yoo and Nolan (forthcoming) adopted a qualitative method which involved qualitative categorisation of nasal tokens into five types of realisation, based on acoustic and auditory evidence:

- Category (1) Sonorant nasals [N]
- Category (2) Partially denasalised nasals [N^D]
- Category (3) (Pre-)voiced non-nasals [D]
- Category (4) Voiceless non-nasals [T], and
- Category (5) Voiceless non-nasals with aspiration (of more than 10ms) [T^H]

These categories capture an increasing level of obstruentisation of a nasal phoneme from Category (1) [m, n] to Category (5) [p^h, t^h]. The acoustic cues used in the categorisation are best explained with examples. Figure 5.6 has been taken from Yoo and Nolan (forthcoming) to demonstrate a canonical example of a token in each of the five categories. The spectrogram (1) – (5) show examples corresponding to Category (1) – (5), respectively. Category (1) displays typical features of a sonorant nasal, such as nasal formants, voicing, relatively high acoustic energy, and a smooth transition into the following vowel. On the other hand, Category (2) shows such features only in the first part of the nasal phoneme. The second part shows reduced amplitude and a sharper transition into the following vowel with a burst-like energy transient. Category (3) resembles this second part of the partially denasalised token in (2), showing regular pulses of voicing and a release burst, but without nasal formants. Category (4) is similar to (3) but without voicing in the consonant and finally, Category (5) is like Category (4) but with a voicing lag longer than 10ms.

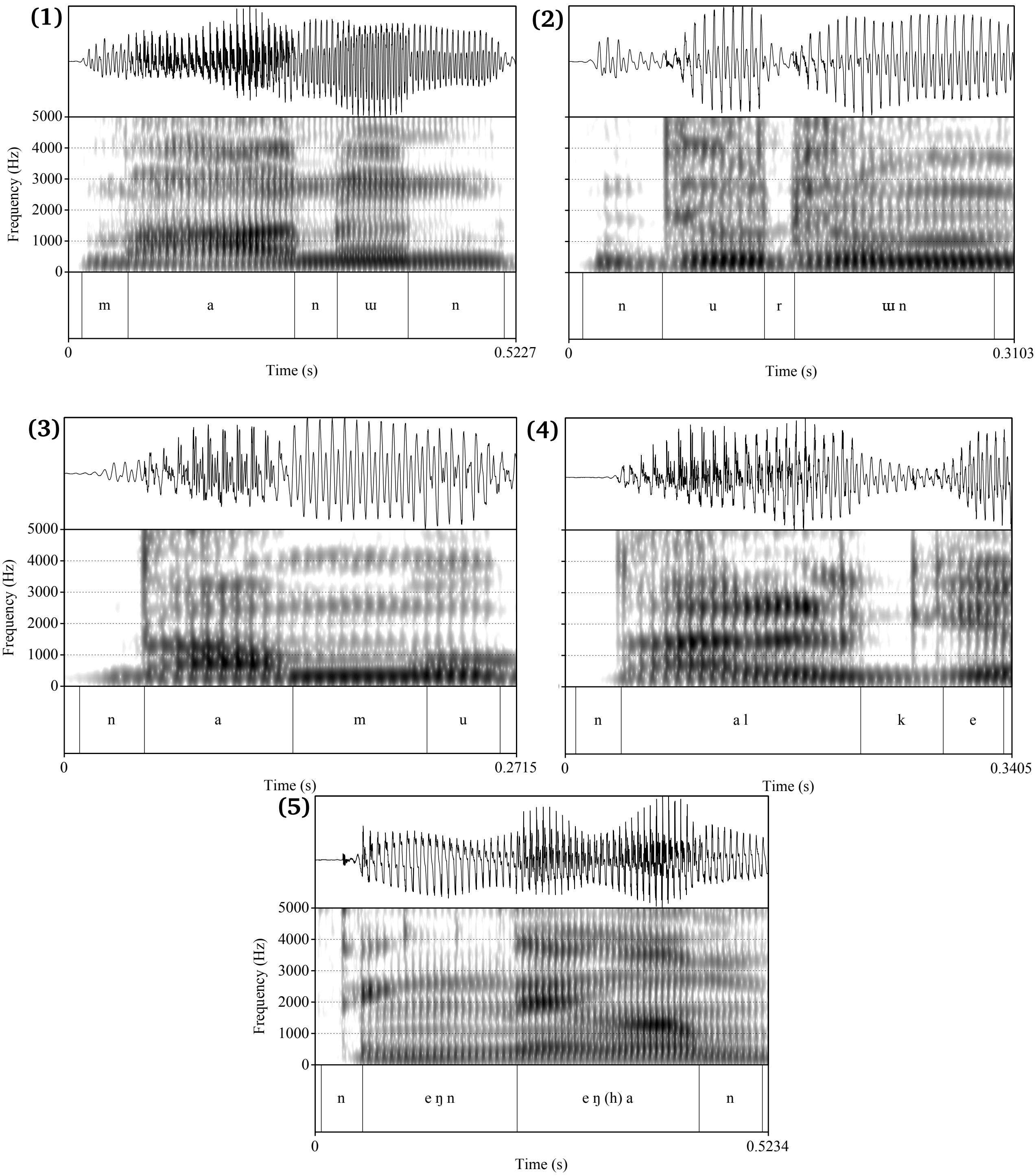


Figure 5.6: Examples of the five categories of realization of nasal phonemes in initial position: Category (1) Sonorant nasals [N]; Category (2) Partially denasalized nasals [N^D]; Category (3) (Pre-)voiced non-nasals [D]; Category (4) Voiceless non-nasals [T], and; Category (5) Voiceless non-nasals with a voicing lag of more than 10ms [T^H]. The tier below the spectrogram shows the underlying phonemes (Yoo & Nolan, forthcoming).

In addition to the auditory and acoustic evidence for the presence or absence of nasality, voicing and a voicing lag, the following additional cues were used to aid the categorisation process according to Yoo & Nolan (forthcoming). First, when it was difficult to judge whether there is nasality in a given token, nasality in the following vowel was used as a secondary cue. This was particularly helpful in post-pausal nasals which were often very brief in duration. Second, a salient acoustic burst helped to rule out Category (1), as such a burst indicates a sudden release of the pressure built up in the oral cavity. This is only possible with a velic closure (Kim, 2011, p. 56).

An important note from Yoo & Nolan (forthcoming) regarding the use of a burst in categorisation is that, while a clear burst indicates a raised velum, the absence of a burst cannot be taken as evidence of a lowered velum. For Category (3) tokens, the appearance of a burst was unreliable because, during a voiced closure, air flows into the oral cavity for only about a third of the time. This leads to lower intraoral pressure in a voiced stop than in a voiceless stop. For Category (2) tokens, the non-nasal portion of the token was often not long enough for sufficient pressure to build up. Lastly for Category (4) and (5), aspiration noise can easily mask a release burst.

5.3.7 A hybrid method of analysing nasality

While the qualitative categorisation method described above has the advantage of capturing the full range of variation by a single set of categories, there were two main shortcomings. First, it depended solely on the capacity of the two phoneticians to consistently classify thousands of tokens into the five types. Second, the boundaries between the categories which are imposed on the continua of nasality and voicing are artificial and arbitrary.

Therefore, the present study took a hybrid approach. Accepting the advantages of qualitative categorisation method, all nasal tokens were classified into one of the five categories. The method of categorisation here closely followed Yoo (2015a) and Yoo and Nolan (forthcoming). For the tokens realised as an approximant, the categorisation was based on whether there was nasality in the approximant (Category 1) or not (Category 3). Occasionally, there were tokens which had the properties of more than one category. For example, one speaker from Seoul produced three denasalised and slightly aspirated tokens of /n/ in a postpausal position with prevoicing followed by silence. Two of these examples are shown in Figure 5.7. Although these tokens showed prevoicing, the perceptual effect was different from Category (3) tokens due to the presence of silence and a voicing lag. Thus, these were classified into either Category (4) or (5) based on the length of the voicing lag.

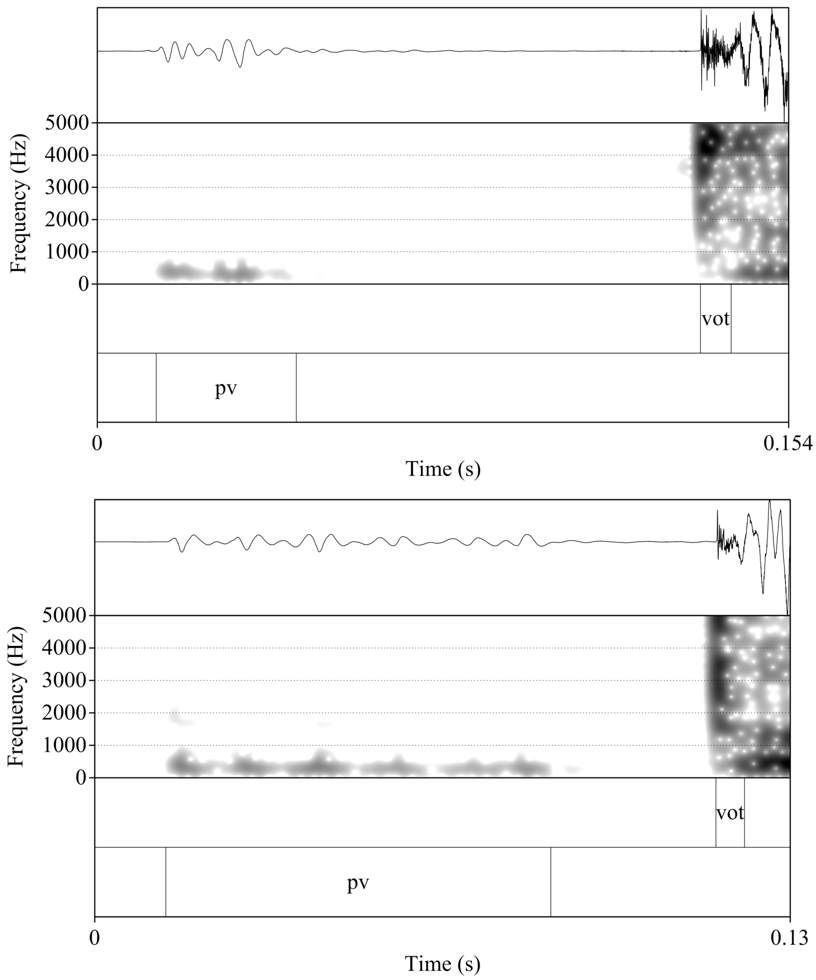


Figure 5.7: Two examples of /n/ realised as a sequence of voiced non-nasal + silence + voicing lag.

The qualitative categorisation was then complemented with quantitative measurements. In the following section, I will present my analysis of the reliability of the two alternatives for measuring nasality, (i) minimum intensity of the nasal tokens and (ii) A1-P0 of the following vowel. The better method of nasality measurement will be determined based on the outcome of this section.

For the intensity measurements, I used the Praat script by Kawahara (2010). For A1-P0 measurements, I used the Nasality Automeasure Script by Styler (Styler, 2017a). For the intensity script, the tier for closure duration was used as the input. For the nasality script, however, two additional tiers were created for the vowels and the words containing the target vowels. While the script allows the measurements to be made multiple times throughout the vowel, only the A1-P0 measurement at the first 25% point was used for each vowel. This is based on the observation in Section 5.3.5 (4) that the nasal energy in the early portion of the vowel better reflects the nasal energy in the preceding consonants. The measurements right at the start of the vowel at 0% were not used as they were occasionally “flagged” at the full-auto mode of the script with the error message, “A1 is P0”. This is presumably due to the fact that vowels were more closed in the initial position, with a low F1 which overlapped with the frequency range of P0.

5.3.8 Comparing the alternative measurements of nasality

The two alternative methods of inferring the amount of nasality, (a) minimum intensity of the nasal consonant and (b) A1-P0 of the following vowel, were tested for reliability. First, I determined whether each measurement could be used to perform a basic task which is considered effortless for a human ear: distinguishing between the level of nasality in the lenis stops and in unambiguously sonorant nasal stops. To ensure that the nasal stops are not denasalised, only the nasals in PWi position were used. PWi nasals are assumed to be phonetically nasal based on the findings in the literature and the results of my own categorisation of the present data. The existing studies agree that denasalisation only targets nasals that are in the initial position of the AP or a higher domain (Section 3.5). My categorisation of the data was in line with this finding, as only eight out of 446 PWi nasals (1.0%) were judged to be in a category other than Category (1) N. These tokens were excluded from the present analysis.

In addition, the tokens realised without a full closure (211 out of 775 tokens of PWi nasals and lenis stops, 27.2%) were also excluded. This is because a lenis stop that is lenited to a fricative or an approximant can potentially be realised with a greater intensity than a nasal stop that is produced with a full closure, rendering the intensity measurements unreliable. In addition, lenited tokens are problematic for A1-P0 measurements because it was often very difficult to determine the boundary between an underlying consonant and the following vowel. Such a scenario is more problematic for A1-P0 than for intensity because intensity was measured at the lowest point, which is likely to fall inside the consonant rather than in the following vowel, even if the segmentation with the following vowel is imprecise. However, an accurate boundary between the consonant and the vowel is more important for A1-P0 measurements as they were taken at the first 25% point of the vowel. The resulting data comprised a total of 557 tokens, consisting of 352 nasals and 205 lenis stops (more lenis stop tokens were eliminated as they were more likely to be lenited). I will first discuss the results for intensity before moving on to A1-P0.

(a) Minimum intensity

As described in Section 5.3.7, Nasality Automeasure Script (Styler, 2017a) was used to measure the mean, minimum, and maximum intensity for the lenis and nasal tokens in PWi position. The default settings in Praat were used, e.g. intensity is smoothed by the window of 42.7 ms, assuming the standard pitch floor of 75 Hz. The window itself moves on each computation by a quarter of its length (10.7 ms). The results are summarised in Table 5.4. As expected, the mean intensity was clearly lower in the lenis stops than in the nasals. However, the range of intensity of the two kinds of stops largely overlapped between 54.9 and 71.6 dB. This indicates that absolute intensity measurements are not sufficient to distinguish between nasal and lenis stops in this data.

Table 5.4: Mean, minimum, and maximum intensity for lenis and nasal tokens in PWi position in the data pooled across 30 speakers. Recall that the intensity measurements were taken at the valley of the intensity contour during the acoustic closure of a given token. Thus, the minimum and the maximum intensity indicates the lowest and the highest observation of the *minimum* intensity for all tokens of the lenis and nasal stops.

Consonant	Mean intensity (dB)	Min intensity (dB)	Max intensity (dB)
Lenis	57.2	38.3	71.6
Nasal	69.5	54.9	81.9

To test whether the measurements are more reliable *within* a given speaker, I took the same measurements for each speaker as summarised in Table 5.5. Unlike in the pooled data, there was no overlap in the range of intensity for the majority of the speakers. There were seven speakers for whom the range of intensity measurements did overlap, as shown in the shaded cells. For five of these speakers, only one lenis token was in the overlapping range and for the remaining two, two lenis stop tokens were responsible for the overlap. This gave a total of nine lenis tokens that showed higher intensity than the minimum intensity of the nasal tokens within the same speaker. This indicates that 98.4% of the tokens (548 out of 557) can be correctly classified into tokens with or without nasality by setting appropriate thresholds of minimum intensity for each speaker.

A more detailed examination of the tokens in the overlapping range revealed that five out of the nine tokens were tokens of a lenis stop from Sentence 17c (see Appendix A for the list of the recording materials). Sentence 17c is the only sentence with a preceding coda /n/ before the target /t/. In such an /nt/ sequence, the nasality disappeared relatively close to the boundary with the following vowel that the part corresponding to /t/, realised as [d], was either very short or non-existent. Thus, the high intensity during [d] appears to be due to the short duration of [d] flanked by highly sonorant segments.

Table 5.5: Mean, minimum, and maximum intensity for nasal and lenis stops in PWi position in each of the 30 speakers. The lenis stop category is missing for Speaker 10 from Busan as this speaker produced no lenis tokens in PWi position, realising them in APi position instead.

Region	Subject ID	Consonant	Mean intensity	Minimum intensity	Maximum intensity
Seoul	1	lenis	62.9	58.2	67.9
		nasal	74.1	68.9	78.0
	2	lenis	62.8	59.2	66.6
		nasal	72.4	68.6	75.0
	3	lenis	60.0	52.7	65.5
		nasal	71.2	68.5	75.2
	4	lenis	54.2	44.8	65.8
		nasal	68.9	65.9	74.2
	5	lenis	63.4	58.9	66.7
		nasal	72.6	68.4	75.8
	6	lenis	65.1	62.1	67.7
		nasal	73.1	69.8	75.5
	7	lenis	50.7	38.3	64.9
		nasal	65.2	58.0	70.0
	8	lenis	58.3	53.1	65.0
		nasal	73.2	69.8	76.6
	9	lenis	55.8	49.3	61.6
		nasal	69.0	65.0	75.1
	10	lenis	52.8	38.9	66.6
		nasal	67.5	61.3	71.5
Busan	1	lenis	49.8	43.1	55.9
		nasal	62.5	55.6	69.7
	2	lenis	62.5	62.5	62.5
		nasal	70.4	68.4	73.0
	3	lenis	58.1	55.9	63.2
		nasal	67.5	63.8	70.7
	4	lenis	60.4	57.5	63.8
		nasal	69.4	66.9	71.9
	5	lenis	50.4	46.5	52.9
		nasal	67.0	60.4	70.5
	6	lenis	57.0	53.5	59.3
		nasal	70.6	64.2	74.0

	7	lenis	56.0	54.1	57.7
		nasal	66.0	60.8	71.2
	8	lenis	55.5	49.1	63.0
		nasal	63.6	60.4	65.9
	9	lenis	47.3	43.5	51.3
		nasal	67.1	62.5	72.1
	10	nasal	59.0	54.9	61.8
Ulsan	1	lenis	56.1	50.3	59.1
		nasal	64.1	61.7	67.2
	2	lenis	56.0	47.3	65.5
		nasal	71.3	65.5	73.4
	3	lenis	60.4	57.3	65.2
		nasal	72.3	68.3	77.1
	4	lenis	54.1	51.4	57.2
		nasal	69.1	62.5	72.6
	5	lenis	52.6	46.0	59.8
		nasal	72.3	69.3	76.3
	6	lenis	66.0	62.7	71.6
		nasal	75.2	70.3	78.9
	7	lenis	54.6	48.0	64.2
		nasal	70.7	69.2	72.8
	8	lenis	63.0	54.9	70.6
		nasal	70.0	66.4	72.9
	9	lenis	59.7	54.6	67.0
		nasal	69.9	66.1	72.8
	10	lenis	61.9	52.3	69.2
		nasal	75.5	71.3	81.9

In a different case from the above, a lenis stop token with an unusually high minimum intensity of 70.6 dB turned out to show auditory and acoustic evidence of nasality. This token occurred in the context /$_{\mathrm{AP}}${ʌmma ***t***alimi} / 'Mother's Iron' from Sentence 1a, following a nasalised vowel. The spectrogram of this token is shown in Figure 5.8. The third formant is clearly visible throughout the closure, which would be unlikely in a non-nasal stop closure. However, there is a sharp onset to the following vowel, suggesting that the amount of nasality is small enough near the release to cause the apparent difference in the acoustic energy between the stop closure and the following vowel. I have noticed a couple more of such cases during the segmentation process, in which an underlying lenis stop is realised with nasality in a PWi position. For the case shown in Figure 5.8, the nasality in the lenis stop seems to be due to the carryover nasalisation from the preceding syllable /ma/.

On the whole, it can be seen that the intensity measurements are highly reliable within a given speaker for the task of distinguishing a lenis stop from a sonorant nasal in the PWi position. This suggests that intensity can indeed reflect the differences in the amount of nasality during a stop closure, regardless of the different places of articulation of the nasals and the variation in the neighbouring segments.

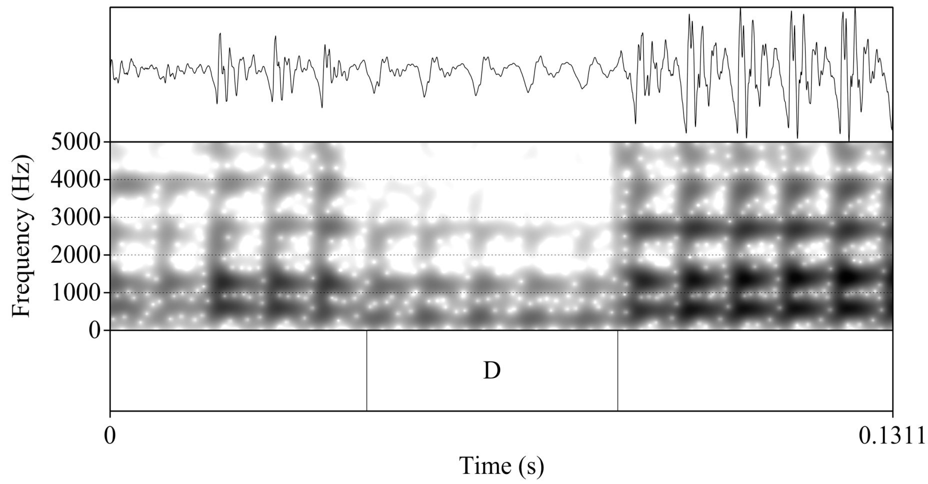


Figure 5.8: A PW-initial lenis stop phoneme realised with nasality.

Next, I investigated whether the intensity measurements reflected the differences in the tokens of the nasal phonemes in each of the five types of phonetic realisation. Of course, neither the intensity measurements nor the type of phonetic realisation as classified by myself can be taken as completely reliable. Nevertheless, if they are consistent with each other, this would add more credibility to the methodology. Regarding the differences in minimum intensity across the categories, I made the following predictions.

- The tokens in Category (1) N will show greater intensity than those in Category (3) D.
- Category (2) N^D will show an intermediate value between (1) N and (3) D.
- Category (4) T and (5) T^H will show even lower values of intensity than (3) D, due to the lack of voicing.
- The lenis stops will show values similar to Categories (3) D – (5) T^H as they were realised as, e.g., [d ~ t^h].

For the analysis here, I used a subset of my data, the data from the 10 speakers from Seoul, for the ease of analysis and presentation. Excluding the postpausal tokens in the IP-pause condition and the lenited tokens, there was a total of 916 tokens, with 546 tokens of the nasal phonemes and 370 tokens of the lenis phonemes.

The results are summarised in Table 5.6. The mean intensity indicated that, as predicted, intensity generally decreased as the category of realisation became less sonorant, from Category (1) N to Category (5) T^H. The lenis stops always showed a mean intensity lower than Category (3) D and higher than Category (5) T^H for those speakers who produced any token in Category (5) in a non-postpausal position. This indicates that there is a close match between the automatically extracted, quantitative measurements of intensity and my manual, subjective classification into the five categories.

Table 5.6: Intensity measurements for the tokens of the nasal phonemes the category of realisation for each of the 10 speakers from Seoul. The “lenis” rows show the intensity measurements of the lenis stop tokens as a point of reference. The tokens in IP-pause (postpausal) condition were excluded since intensity at the minimum was always zero due to the preceding silence.

Subject	Category	Mean intensity (dB)	Min intensity (dB)	Max intensity (dB)	Token count (total = 916)
Seoul_1	N	73.9	67.9	78.0	23
	N^D	69.4	69.4	69.4	1
	D	62.2	47.0	69.8	30
	T	46.3	43.3	49.3	2
	Lenis	56.0	26.6	67.9	38
Seoul_2	N	70.2	47.6	76.6	34
	N^D	68.3	68.3	68.3	1
	D	59.8	44.3	67.2	9
	Lenis	57.6	34.8	66.6	30
Seoul_3	N	70.0	47.3	75.5	48
	N^D	61.2	59.5	63.9	3
	D	66.6	65.3	67.8	2
	Lenis	54.8	30.3	68.4	32
Seoul_4	N	68.1	62.2	74.2	28
	N^D	63.6	63.3	63.9	2
	D	63.9	53.4	69.7	16
	T	54.0	54.0	54.0	1
	T^H	47.4	47.4	47.4	1
	Lenis	50.2	36.0	65.8	34
Seoul_5	N	71.4	60.5	75.8	34
	N^D	59.2	55.6	62.9	5
	D	60.3	56.3	65.6	21
	T	58.5	58.5	58.5	1
	Lenis	54.5	35.6	66.7	37
Seoul_6	N	73.8	69.8	77.8	32
	N^D	58.8	49.3	66.5	5
	D	64.5	59.4	72.2	23
	T	54.8	54.8	54.8	1
	Lenis	56.4	28.0	67.7	30
Seoul_7	N	62.0	32.9	70.0	29

	N^D	53.4	46.8	59.9	8
	D	47.5	18.1	59.0	12
	Lenis	45.1	28.7	64.9	39
Seoul_8	N	73.0	68.4	76.6	27
	N^D	65.7	57.2	70.5	18
	D	61.7	58.7	65.3	4
	Lenis	52.1	31.9	68.9	38
Seoul_9	N	69.1	63.8	75.1	31
	N^D	50.3	29.0	57.8	5
	D	52.1	25.0	63.1	13
	T	38.8	31.2	49.1	5
	T^H	39.6	22.1	48.6	7
	Lenis	40.7	13.9	63.0	43
Seoul_10	N	65.6	50.3	71.5	43
	N^D	56.9	38.8	64.9	13
	D	51.5	23.0	61.7	7
	T	37.2	37.2	37.2	1
	Lenis	47.4	16.8	66.6	49

However, there were a few exceptions to the pattern that intensity progressively decreased in a higher, less sonorous category. For half of the subjects, Subjects 3, 4, 5, 6 and 9, Category (2) N^D showed a lower average intensity than Category (3) D. This seems to be a result of the relatively long closure duration for tokens in Category (2), in which residual voicing became weaker. On the other hand, the voiced tokens in Category (3) tended to show short but stronger voicing. The mean closure durations for Categories (1) – (5) were 49.9 ms, 67.7 ms, 56.7 ms, 87.3 ms and 107.6 ms, respectively. The tendency for the tokens in Category (2) to be longer than those in Category (3) was particularly strong for those five subjects.

For one of the two speakers who produced Category (5) T^H in a non-postpausal position, Subject 9, the mean intensity was higher in Category (5) T^H than in Category (4) T. This can also be attributed to the differences in voicing. That is, for this particular speaker, the percentage of residual voicing during the closure tended to be higher for tokens in Category (5) than for those in Category (4), leading to a slightly higher average intensity.

In addition, for most tokens in Categories (4) & (5) in the non-postpausal positions, residual voicing ended shortly before the release. As a result, the intensity contour appeared not to capture the brief period of weak or no voicing before the high energy release burst. This is illustrated in Figure 5.9. A similar tendency was observed with the tokens in Category (2) N^D, especially for those with a relatively short duration. As exemplified in Figure 5.10, a drop in nasality was not reflected in the intensity curve. However, it must be noted that as there were relatively small numbers tokens in Categories (2), (4), and (5), the results may not be very representative.

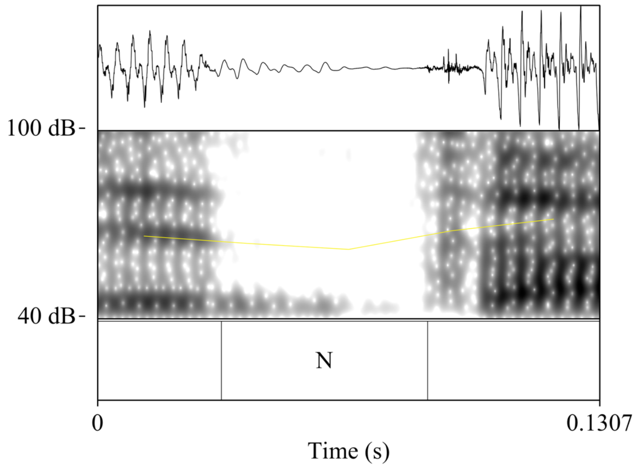


Figure 5.9: Intensity of a nasal phoneme in Category (5) T^H. Residual voicing during closure dies out before the release burst, but the intensity curve does not reflect the loss of voicing, perhaps because it is too brief[33] or because of the dominating effect of the energy of the release burst.

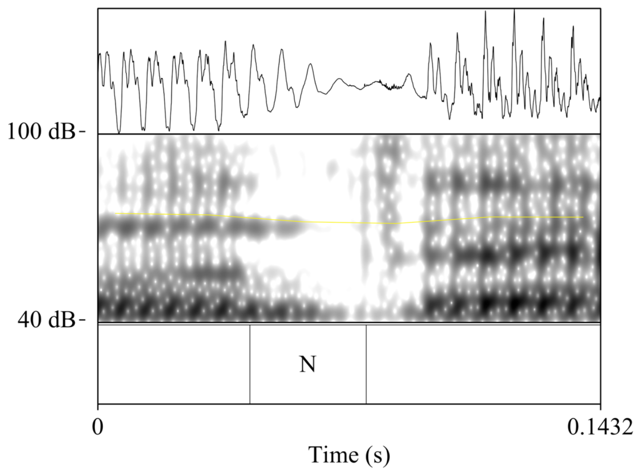


Figure 5.10: The intensity curve of a nasal phoneme in Category (2) N^D. Nasal energy seems to decrease towards the release burst, but the curve does not capture the loss of nasal energy.

A more important issue for this study is that the intensity range of Category (1) N and (3) D overlapped for every speaker, except Subjects 8 and 9. For some speakers, the overlap is fairly large, e.g. Subject 3. For this speaker, the range of intensity of Category (3) is

[33] As previously mentioned, intensity is smoothed by the window of 42.7 ms, assuming the standard pitch floor of 75 Hz. The window itself moves on each computation by a quarter of its length (10.7 ms). Thus, if the drop in intensity is shorter than this, it may not be reflected in the intensity curve.

completely subsumed under that of Category (1). Examining more closely the sonorant nasals with low minimum intensity and voiced non-nasals with high minimum intensity, the explanation again seems to lie in the variation in the strength of voicing, a concern that has been expressed earlier in Section 5.3.5 (1). Figure 5.11 shows two sonorant nasals realised with a relatively low acoustic energy, due to weak and/or intermittent voicing. On the other hand, Figure 5.12 shows an example of a denasalised nasal, realised without nasal formants but with stronger voicing, leading to higher acoustic energy compared to that of the nasals in Figure 5.11. Figure 5.13 provides strong support for this explanation, as there is a robust correlation between closure duration and minimum intensity.

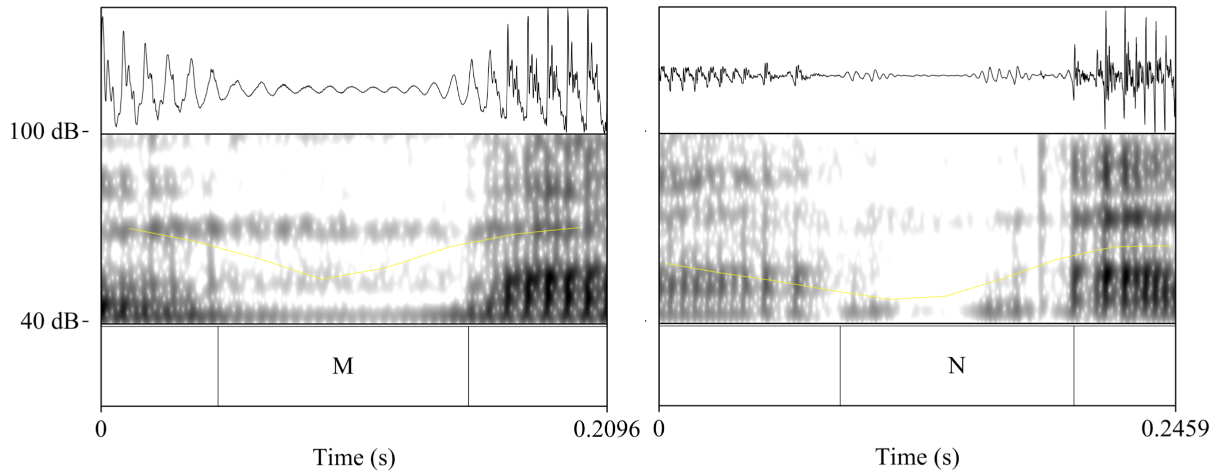


Figure 5.11: Two sonorant nasals in Category (1) N. The figure on the left shows a bilabial nasal phoneme (annotated M) realised with weak voicing, and the figure on the right shows an alveolar nasal phoneme (annotated N) realised with intermittent voicing.

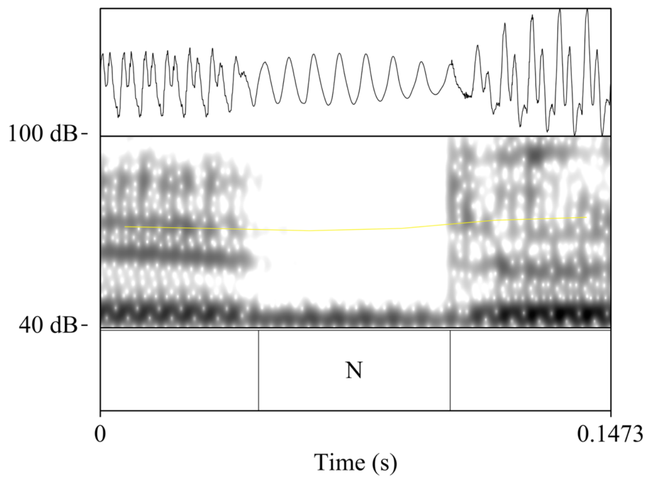


Figure 5.12: A nasal phoneme in Category (3) D, realised with strong voicing.

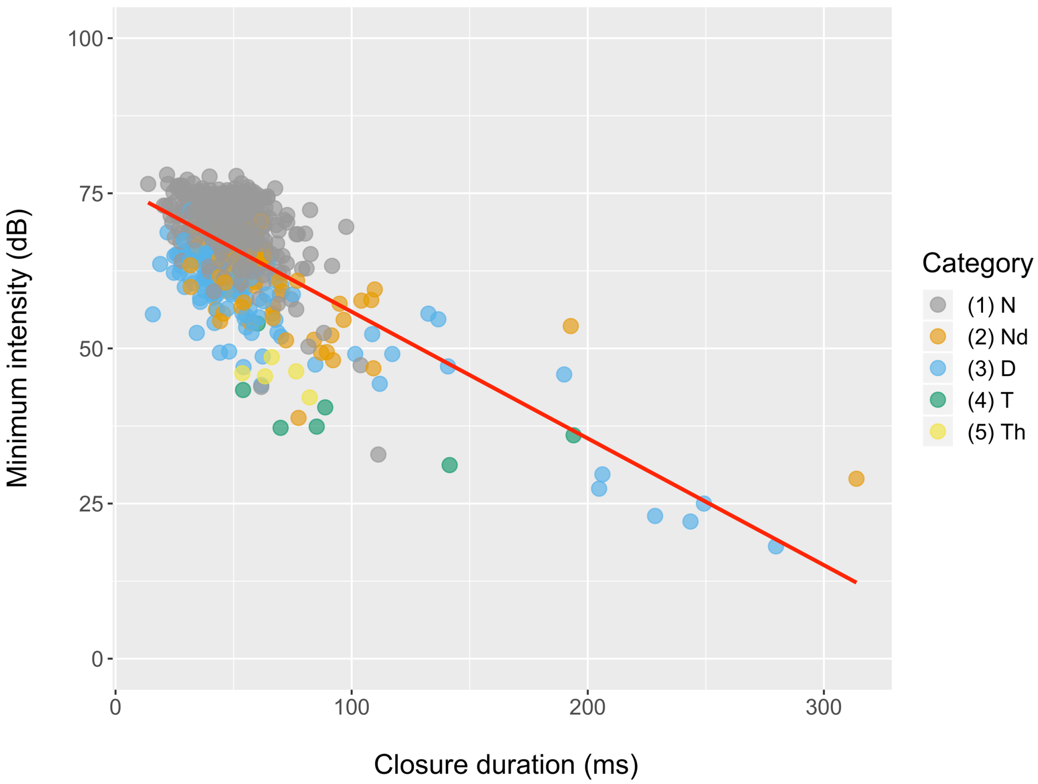


Figure 5.13: The correlation between closure duration (ms) and minimum intensity (dB) of the nasal tokens in five categories of phonetic realisation.

A final point in favour of the usefulness of measuring minimum intensity is that the intensity measurements corresponded closely to my subjective sense of the level of nasality. That is, tokens in the overlapping range of intensity between the two categories (1) N and (3) D were often the tokens that were relatively difficult to classify. For example, tokens in Category (3) with a relatively high intensity value, such as the one shown in Figure 5.12, were perceived as having an amount of nasality intermediate between Category (1) and Category (3). I could detect a hint of nasality even though the token sounded denasalised in context. The intensity measurements could also reflect my perception of whether a token in Category (2) N^D was closer to Category (1) or Category (3). Taken together with the overall result that intensity was generally consistent with the assigned categories of the tokens, it can be concluded that intensity can indeed serve as a helpful complement of the discrete categorisation scheme. However, one must be cautious when interpreting the results as there is a conflating effect of voicing.

(b) A1-P0 of the following vowel

As mentioned in Section 5.3.7, the Nasality Automeasure Script by Styler requires two tiers, the vowel tier and the word tier. This is in addition to the four existing tiers for the various duration measurements: the closure duration, VOT, residual voicing and utterance duration. Thus, before segmenting the vowels and the words for the entire dataset, I first conducted a preliminary investigation to test the reliability of A1-P0 using the data from one Seoul speaker. For all pairs of PWi lenis and nasal tokens from matching sentences, the following vowel and the word containing the token were segmented (see Section 5.3.3 (1) for the method of segmentation). Since Part B was not designed in a parallel structure between the lenis and the nasal stops, the tokens in this set were not used. This is to maximise the reliability of A1-P0 measurements which are influenced by the type of vowel, as discussed in Section 5.3.5 (4). Excluding the tokens which were lenited and the one token flagged with an error message[34], there was a total of 21 tokens, with 8 lenis and 13 nasal tokens.

Table 5.7: The mean, minimum and maximum values of A1-P0 measurements of the vowels following the PWi nasals and lenis stops. The data is from one subject from Seoul. Note that there are rows with only one token and thus contain the same values for the mean, minimum and maximum A1-P0.

Vowel	Consonant	Mean A1-P0	Min A1-P0	Max A1-P0	Token Count
e	lenis	5.995	2.82	9.17	2
	nasal	0.46	-3.77	4.69	2
ʌ	lenis	7.47	7.47	7.47	1
	nasal	2.65	2.65	2.65	1
o	lenis	-2.128	-8.34	3.7	5
	nasal	0.36	0.36	0.36	1
u	nasal	-1.27	-1.27	-1.27	1
a	nasal	5.67875	2.24	8.68	8

The results in Table 5.7 indicate that reliability of A1-P0 depends on the vowel. Recall that greater nasality is reflected in a lower value of A1-P0. For the vowels /e/ and /ʌ/, the mean A1-P0 was higher in the lenis context than in the nasal context, as expected. On the other hand, the reverse pattern was found for /o/. Considering that the /o/ phoneme in Korean is phonetically higher than the cardinal vowel 7 [o], the inaccurate result may be understood as a result of the low first formant interfering with the nasal pole. However, the

[34] The error message was, again, "A1isP0", indicating that a reliable measurement could not be obtained as the identified first formant was the same as the nasal pole which is often the case with high vowels. In this case, the vowel was /i/.

token count was very low and too unbalanced to make any conclusive statements. Lastly, due to the lack of unlenited lenis tokens followed by /a/ and /u/, the comparison could not be made for these vowels. Note that the situation of low token counts cannot be improved by pooling data across different speakers. An analysis using A1-P0 has to be made within a single speaker, because A1-P0 shows a high amount of variation in baseline and range from one speaker to another (Styler, 2017b).

Next, I analysed A1-P0 values by the five categories of realisation, using a larger subset of data from all 10 Seoul speakers. There was a total of 608 tokens of vowels following a nasal phoneme, after removing 39 tokens which were lenited and 113 tokens which were error flagged. The results summarised in Table 5.8 show that the mean A1-P0 values were highly inconsistent with the amount of nasality indicated by the five categories. Note that these results lump together the different types of vowels. As expected, the average A1-P0 was indeed lower following a Category (1) token than a Category (3) token in all subjects, except Subject 1, 7 and 10. However, the mean A1-P0 was often the highest after Category (2) N^D and the lowest after Category (4) T or (5) T^H within a given speaker. This is problematic, because there is no reason to expect vowel nasalisation to be greatest after a voiceless non-nasal and smallest after a partially denasalised nasal.

Table 5.8: The mean, minimum and maximum values of A1-P0 measurements by the category of realisation of the preceding nasal. The data is from 10 subjects from Seoul.

Subject	Category	Mean A1-P0	Min A1-P0	Max A1-P0	Token Count
Seoul_1	N	3.14842	-5.42	9.01	42
	N^D	10.14	10.14	10.14	19
	D	2.23738	-21.34	10.39	1
	T	-4.76	-20.82	4.83	3
	T^H	1.26	1.26	1.26	1
Seoul_2	N	2.49095	-3.77	8.68	11
	N^D	10.7	10.7	10.7	42
	D	6.44909	-2.49	12.87	1
	T	-0.65	-0.65	-0.65	1
	T^H	-4.52	-4.52	-4.52	1
Seoul_3	N	-3.9687	-9.53	1.8	3
	N^D	-2.885	-5.04	-0.14	54
	D	-2.10333	-2.57	-1.3	10
Seoul_4	N	1.54304	-10.73	11.47	17
	N^D	7.255	5.21	9.3	23
	D	7.08	-4.33	17.54	2
	T	6.866	-0.58	13.02	5
	T^H	14.16	14.16	14.16	1

Seoul_5	N	-0.79643	-16.37	7.79	21
	N^D	4.2125	-2.56	9.01	42
	D	6.25857	-11.61	12.45	8
	T	7.81	7.81	7.81	1
	T^H	6.69	6.69	6.69	1
Seoul_6	N	2.38818	-2.26	7.6	24
	N^D	4.42455	-2.06	8.38	33
	D	6.59167	-1.03	10.29	11
	T	8.59	8.59	8.59	1
	T^H	9.05	9.05	9.05	1
Seoul_7	N	3.99286	-9.26	13.99	10
	N^D	3.664	-3.82	8.32	21
	D	3.861	-0.93	8.02	10
	T^H	-2.21	-2.21	-2.21	1
Seoul_8	N	-3.74278	-23.39	5.44	5
	N^D	0.23857	-3.9	7.72	36
	D	1.568	-1.15	5.59	14
	T^H	1.37	1.37	1.37	1
Seoul_9	N	3.32955	-7.45	39.39	18
	N^D	3.44375	-3.24	11.86	22
	D	5.30167	-2.17	14.7	8
	T	6.76625	1.08	10.17	8
	T^H	5.82429	-1.99	11.89	7
Seoul_10	N	-0.64174	-15.09	13.61	9
	N^D	-0.39889	-8.25	6.08	46
	D	-0.66222	-8.64	5.78	9
	T	-4.98	-4.98	-4.98	1
	T^H	0.135	-0.09	0.36	2

Then, the same analysis was conducted while taking the vowel quality into consideration. Even within a matching vowel, A1-P0 was not consistent with the level of nasality suggested by the category of realisation. The pattern was similar to above, with a high A1-P0 after Category (2) and a low A1-P0 after Categories (4) and (5). Crucially, the mean A1-P0 was often higher in Category (1) than in Category (3) for high vowels /o u i/ even after removing the tokens which were error flagged by the script.

Perhaps, the observed unreliability of A1-P0 may be related to variation in voice quality. Voice quality is often measured by the amplitude difference between the first and the second harmonic, H1-H2 (Klatt & Klatt, 1990; Gordon & Ladefoged, 2001; Keating & Esposito, 2006; Keating, Esposito, Garellek, Khan, & Kuang, 2010). Breathy voice results in a stronger H1 relative to H2, whereas creaky voice typically shows a weaker H1 relative to H2 (Gordon & Ladefoged, 2001, Section 5.3). Given that P0, the first nasal pole, usually corresponds to either H1 or H2 (Chen, 1997, pp. 2360, 2362), one can expect complicated interference effects between vowel nasalisation and voice quality. Indeed, Zhang (2015, Section 4.2.1.2) found that in Mandarin Chinese, increased A1-P0 was observed in the presence of creaky voice, even though the articulatory measurement of nasalance indicated increased nasality. As a higher A1-P0 value is associated with lower nasality, this result was taken to indicate that creaky voice creates a bias in A1-P0 measurements such that vowel nasalisation seems weaker. In a study comparing the acoustic properties of vowel nasality in English and French, Styler (2017b) found that the measures of spectral tilt including H1-H2 were the strongest indicators of nasality in French but not in English. The results were interpreted to mean that French speakers use breathy voice to further enhance the spectral tilt which is present in nasal vowels in order to strengthen the contrast between nasal and oral vowels. The co-occurrence of nasality and breathiness has also been reported in a study of three Yi (Loloish) languages (Garellek, Ritchart, & Kuang, 2016).

While voice quality was not a primary focus of the present study, I observed voice quality variation in vowels which were preceded by a nasal token. One pattern which stood out was creakiness in the vowel(s) on one or both sides of a nasal token. This was usually found across a relatively strong domain boundary, such as the IP, and regardless of whether the nasal was realised with or without nasality. An example is provided in Figure 5.14. Given the discussed effects of voice quality on A1-P0, the voice quality variation may have introduced noise to our measurements of nasality, rendering A1-P0 less desirable as an indicator of nasality for our data. Nevertheless, it remains unclear to me why A1-P0 was often the highest in the vowels following a token in Category (2) N^D, as there is no particular reason to expect more creakiness in these vowels.

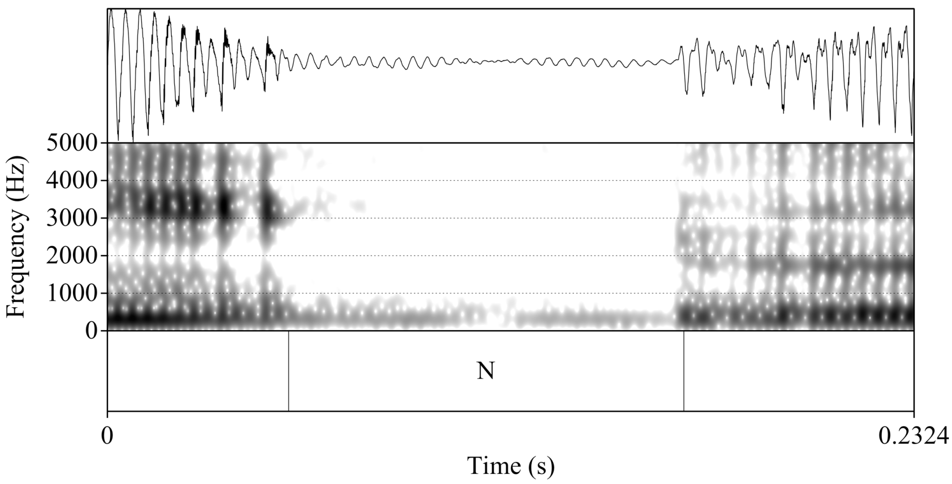


Figure 5.14: An example of creaky vowels flanking an IPi token of a nasal phoneme.

Another disadvantage of A1-P0 is that many tokens are lost due to various problems which are helpfully specified in the error warnings produced by the script. By far the most common problem was that the frequency range of the first formant overlapped with that of P0 in high vowels. Even when there were no error messages, the measurements for high vowels were highly inaccurate. Furthermore, as mentioned earlier, segmentation and labelling of two additional tiers to serve as the input of the nasality script would require a significant amount of time. In sum, while A1-P0 has the advantage of allowing the inclusion of postpausal tokens unlike intensity measurements, the disadvantages clearly outweigh the advantages. Thus, in the main analysis, I used minimum intensity as the measurement of nasality. The following method was used to normalise individual variation in the group analysis.

(14) Speaker normalised minimum intensity (dB)
= *highest minimum intensity by the speaker – minimum intensity of a given token*

For nasal tokens, minimum intensity was normalised for individual differences by subtracting the minimum intensity of a given token from the highest minimum intensity value obtained from each speaker's data. The highest minimum intensity reflected the minimum intensity of the most sonorant token in the speaker's data. Thus, the normalised minimum intensity could be seen as an indicator of the degree of loss of sonority in a given nasal token. To reflect this more straightforwardly, this speaker normalised measure will be called *nasal attenuation* from now on.

Note that the normalisation process also helps to minimise the effect of room reverberation mentioned in Section 5.3.4. Based on the discussion in the section, it is expected that minimum intensity measured for tokens for Busan and Ulsan can be overestimated due to reverberated acoustic energy from the preceding segments. On the assumption that the levels of room reverberation will be more or less similar within each individual's recording, the normalisation serves to remove, as much as possible, the added acoustic energy from room reverberation.

Another factor to consider is that room reverberation tapers off over time. This implies that the reverberation will affect the minimum intensity measurements for shorter segments more than for longer segments. Since tokens in a higher prosodic position tend to have longer closure durations (T. Cho & Keating, 2001, as reviewed in Section 2.3.1.1), this may bias our results toward overestimating the effect of prosodic position on denasalisation for Busan and Ulsan. However, as this works against hypothesis **H3**, any evidence supporting the hypothesis is considered reliable as it was found despite this bias.

H3. For the nasal stops, the DIS effects are greater in the order of Seoul, Busan, and Ulsan, following the model of hierarchical diffusion (Trudgill, 1974).

5.3.9 Determining categoricity vs. gradience

One of the main aims of this dissertation is to investigate the grammatical status of lenis stop (de-)voicing and denasalisation. We are interested in whether the phonetic patterns observed in the data are best explained as a result of a gradient, language-specific phonetic process, a categorical phonological process, or a combination of both, in the manner predicted by the life cycle of phonological processes (Section 1.2.2).

In order to find evidence of categoricity in the data, or the lack thereof, I used statistical bimodality as a diagnostic (Turton, 2017, 2014, Section 5.5.3.2; Bermúdez-Otero & Trousdale, 2012; Scobbie, 2005; Herd, Jongman, & Sereno, 2010, p. 508). The left panel in Figure 5.15 demonstrates a bimodal distribution of minimum intensity measurements for the nasal tokens from a Seoul speaker. Bermúdez-Otero and Trousdale (p. 6) stated that bimodality in high quality phonetic data from an individual constitutes compelling evidence of a categorical process, as it cannot be accounted for by a continuous process which varies as a function of a phonetic variable, e.g. segment duration. Note that the distribution must be investigated on the level of an individual in order not to confuse between gradience and variance (see Turton, 2017). That is, aggregated data from more than one individual who shows different distributional patterns may mask the presence of a bimodal distribution in separate individuals. This is exemplified in the right panel of Figure 5.15.

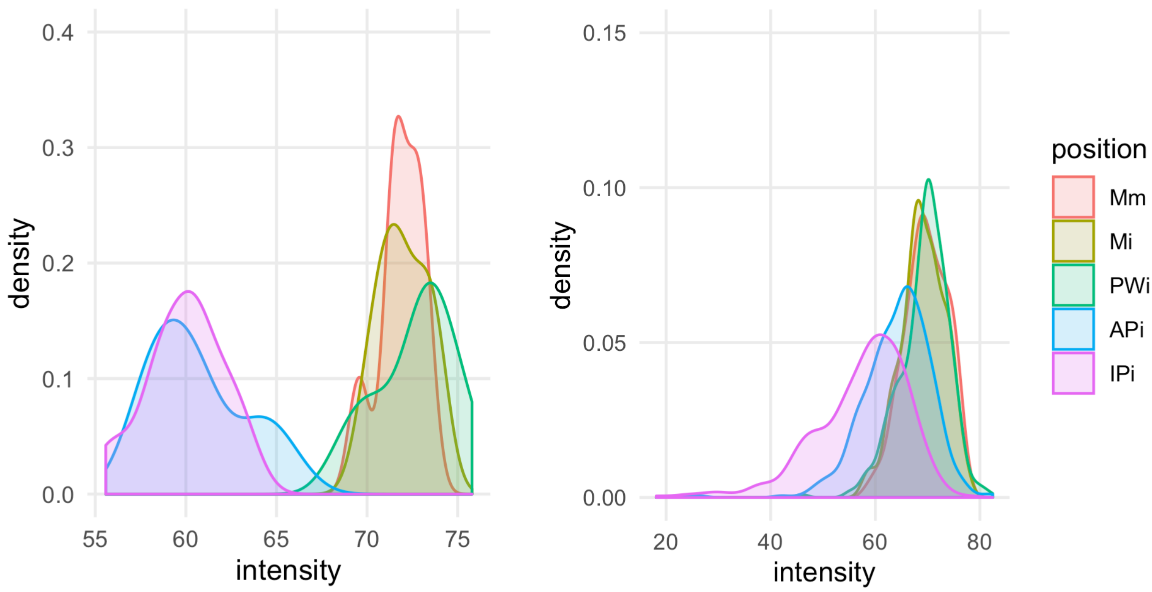


Figure 5.15: The distribution of minimum intensity for the nasal tokens across prosodic position. The bimodal distribution by the speaker on the left panel is masked in the apparent unimodal distribution of the aggregated data in the right panel.

However, this bimodality criterion has some important pitfalls and limitations, as summarised in Turton (2014, Section 4.4). A first limitation of this criterion is that an absence of bimodality does not necessarily entail an absence of categoricity (Bermúdez-Otero & Trousdale, 2012, p. 7). According to Schilling, Watkins, and Watkins (2002), a mixture of equally weighted normal distributions with the same standard deviation is bimodal only if the difference between the means of the distributions is greater than twice the standard deviation. In other words, evidence of categoricity may not be detected if the two distributions are too close to each other on a given parameter. Bermúdez-Otero and Trousdale (p. 7) noted another pitfall of the bimodality diagnostic that, at times, highly sensitive articulatory data may be necessary to reveal hidden residual gestures (see Ellis & Hardcastle, 2002; Lawson, Stuart-Smith, & Scobbie, 2008). For example, acoustic data may not be able to capture a subtle tongue-tip raising gesture. In a similar vein, Browman (1995) pointed out that articulatory gradience may result in an abrupt change in the acoustics.

The last issue is particularly relevant for the present study as minimum intensity is used ultimately to infer the velum height. It is possible that a gradual raising of the velum does not result in a gradual decrease in the acoustic intensity. For example, as a certain velum height is reached, the acoustic coupling of the oral cavity and the nasal cavity might be lost, leading to an abrupt fall in nasal resonance and hence an abrupt fall in intensity. However, the same problem exists for any articulatory method that does not measure velum height directly. While techniques such as Electromagnetic articulography (Perkell et al., 1992), Electromyography (Bell-Berti, 1976; Freitas, Teixeira, Silva, Oliveira, & Dias, 2014), and Velotrace (Horiguchi & Bell-Berti, 1987) can offer a more direct investigation of velum movement, these methods have other disadvantages, e.g. involving a direct contact between the instrument and the velum which may affect its natural movements. In the absence of a better alternative currently available, acoustic measurements of intensity will be used for the investigation of the grammatical status of denasalisation, keeping in mind the pitfalls and limitations discussed in this section.

5.4 Statistical analyses

The primary purpose of the statistical modelling in this chapter is to test the hypotheses in Section 5.2. Therefore, the structures of the models in Sections 5.4.1 and 5.4.2 have been predetermined on the basis of the literature reviewed in Chapter 2 and the hypotheses of this dissertation. No further model selection process has been applied, since alternative approaches including different forms of stepwise regression are known to lead to inflation of Type-I error, i.e. false positive findings (Mundry & Nunn, 2009; Gelman et al, 2013; Roettger, 2019). The only exception to this rule is Section 5.4.6; there, the goal is to find the best-fitting model, which is required to test the hypotheses **H10** and **H11**. Further details of the model comparison process will be given in Section 5.4.6.

5.4.1 Bayesian modelling

For the analysis involving the categories of realisation, Bayesian mixed effect logistic models were fitted using the *brms* package in R (R Core Team, 2015). The R scripts used in the present study are available in the osf repository, along with the raw data (https://osf.io/8kmqz). As input to a Bayesian model, prior distributions (priors for short) need to be set by the researcher based on their beliefs about the plausible values of model parameters. Then, data is used to update the prior distributions to arrive at a posterior distribution, which is the final output of a Bayesian model. The posterior distribution shows the probability distributions of the model parameters, given the prior belief and the data. For examples of Bayesian analyses within Linguistics, see Sóskuthy & Roettger (2019), Nalborczyk, Batailler, Lœvenbruck, Vilain, & Bürkner (2019), and Vasishth, Nicenboim, Beckman, Li, & Kong (2018).

Bayesian models offer many advantages over frequentist linear mixed models (Nalborczyk et al., 2019, pp. 147-8). One important advantage is that Bayesian models can quantify the uncertainty regarding the magnitude of an effect, instead of forcing researchers to draw a binary conclusion of whether or not an effect of a certain size is present or absent. Another strength that is particularly relevant to our case is that Bayesian models are free from convergence issues. Model convergence failures often pose severe problems for linear mixed models, especially when the random effects structure is relatively complicated, as in the present case.

For simplicity, only two binary variables were used as the outcome variable of the models: *denasalised* (no, yes) and *devoiced* (no, yes), with *no* as the reference level. The variable *denasalised* was created by assigning the value *no* to tokens in Category 1 [N], and assigning *yes* to tokens in Categories 2 [N^D] – 5 [T^H]. Similarly, the variable *devoiced* was created by assigning *no* to tokens in Categories 1 [N] – 3 [D], and *yes* to tokens in Categories 4 [T] and 5 [T^H].

While the main focus of this part of the analysis is on the nasal tokens, an analysis of the lenis tokens will also be provided in order to see if there are any parallel patterns between the two types of consonants. Four categories were created for the lenis tokens:

- Category (1) Lenited as a fricative or an approximant [L]
- Category (2) Fully voiced [D]
- Category (3) Not fully voiced [T] (with a value of % voicing smaller than 100%)

- Category (4) Not fully voiced with a voicing lag (of more than 10ms) [T^H]

For the lenis tokens, only the variable *devoicing* was used as the outcome of the model, by combining Categories 1 [L] and 2 [D] into a group without devoicing (*no*) and combining Categories 3 [T] and 4 [T^H] into another group with devoicing (*yes*).

The fixed factors were *prosodic position* (Morpheme-medial, Morpheme-initial, PW-initial, AP-initial, IP-initial, Postpausal, shortened as Mm, Mi, PWi, APi, IPi, P, respectively), *region* (Seoul, Busan, Ulsan), *repetition* (1, 2), and *gender* (female, male). Note that *speaking rate* was not included in the model as the relevant outcome variables have been normalised for differences in speaking rates. *Repetition* was included to control for the fact that some tokens were immediately preceded by a token in the same utterance, hence more susceptible to lenition. The code '1' indicated that the token was not immediately preceded by a token in the same utterance, and '2' indicated that it was. The model also included *Gender* as it is a well-established factor in conditioning sociolinguistic variation (e.g. Labov, 1990; Eckert, 1989). However, *Repetition* and *Gender* serve merely as control variables in this dissertation, rather than as foci of hypothesis testing. Thus, these two variables will be given less weight than the others in the subsequent results and discussion. The interaction between *position* and *region* was also included, to investigate if the effect of prosodic position differed by region. The models also included random intercepts for *subject* (30 speakers) and *item* (10 segmental strings), a random slope over *position* by *subject* and a random slope over *region* by *item*. This can be summarised by the following *brms* formula:

```
brm(outcome ~ position * region + repetition + gender + (position | subject)
+ (region | item))
```

Following common practice, weakly informative (regularising) priors were used for all models (Gelman, Simpson, & Betancourt, 2017). These are the priors that assume the null hypothesis by assigning the highest probability mass to parameter values which correspond to no difference between groups. In line with the recommendation in Gelman (2019), I used a student-t distribution centred around zero with df = 3 and scale = 2.5 for the intercepts, slopes, and standard deviations. A student t-distribution is similar to a normal distribution but has heavier tails. Adjusting the degree of freedom to 3 makes the tails lighter, i.e. fall off more gradually than when the degree of freedom is 1. The scale of 2.5 means that I make the prior assumption that the responses in the log odds space are distributed as a student-t distribution centred around 0, with a standard deviation of 2.5. That is, 68% of the values are expected to lie between -2.5 and 2.5 in log odds space, which corresponds to 7.6 and 92.4% in probability terms. This is considered a conservative, lenient prior, which allows a very wide range of values for the parameters.

The results of the Bayesian models were used to test the hypotheses regarding whether there is any regional variation in the patterns of the lenis and nasal stop variation (**H2** & **H3**).

5.4.2 Mixed effects linear regression

For the continuous outcome variables – that is, all outcome variables that are not about categories of realisation – standard (non-Bayesian) mixed effects linear regression models were fitted using the *lmer* function from the *lme4* package (Bates, Mächler, Bolker, & Walker, 2015) in R (R Core Team, 2015). Standard regression models were used here mainly for reasons of efficiency. Bayesian models are relatively slow to compute and interpret, and there were 80 models to run (see Table 5.9 for a summary of models). In addition, standard regression models tend to run into more convergence issues with a multi-level categorical outcome than with a continuous outcome variable, contributing to the decision to use Bayesian models in Section 5.4.1 but not here. The focus in these sets of analysis was to test the hypothesis in **H1** regarding the effects of prosodic position on the phonetic patterns of voicing and nasality in the three varieties. Hypotheses concerning the regional variation (**H2** & **H3**) were also tested by these models**.**

In these models, it was important to use normalised measurements in order to make sure the measurements are comparable across different speakers and at different speaking rates. Thus, the following variables were used as the outcome variables: 1) *rate normalised closure duration*, 2) *rate normalised residual voicing*, 3) *% voicing during normalised closure*, 4) *rate normalised VOT*, 5) *rate normalised total voiceless interval*, 6) *rate normalised prevoicing*, and 7) *speaker normalised nasal attenuation*. For convenience, these are simply referred to as *closure duration*, *residual voicing*, etc. in the results section.

For each of the outcome variables except 6) *prevoicing* and 7) *nasal attenuation*, two separate models were fitted for the lenis and the nasal tokens from the entire data. Then, a combined model was fitted to compare the two types of consonants. For *prevoicing* and *intensity*, only the nasal tokens were used, and a combined model was not necessary. Since *intensity* varied both as a function of nasality and voicing, only the voiced tokens of the nasal phonemes (Categories (1) – (3)) were used to minimise the effect of voicing. In total, there were five lenis models, seven nasal models, and five combined models.

For the lenis and nasal models, the fixed factors were the same as those in the Bayesian models: *prosodic position* (Mm, Mi, PWi, APi, IPi, P), *region* (Seoul, Busan, Ulsan), *gender* (female, male), and *repetition* (1, 2). The P position was not included in the models except with the outcome variable of *VOT* and *prevoicing*, because the preceding silence made it impossible to determine the beginning of the closure. Therefore, by extension, *residual voicing*, *total voiceless interval*, and *minimum intensity* were also not meaningful in this position. For *prevoicing*, only the tokens in P position were used, thus *prosodic position* was not used as a fixed factor. Instead, *category* (N, Nd, D, T, Th) was used to investigate if prevoicing varied according to the phonetic category of the nasal phoneme. The random intercepts for *subject* (30 speakers) and *item* (10 segmental strings) were added. While including the random slopes as in the Bayesian models would have been ideal, the random slope models did not converge with any of the available optimisers as tested using the *allFit* function from the *lme4* package. Thus, the intercept models were fitted using the following formula:

```
lmer(outcome ~ position * region + gender + repetition + (1 | subject)
+ (1 | item)
```

For the combined models, there was also an interaction term between *position* and *consonant* and a main effect term of *consonant*. The resulting formula is shown below:

```
lmer(outcome ~ position * region + position : consonant + consonant + gender
+ repetition + (1 | subject) + (1 | item)
```

One reason for not including the interaction of *position* * *region* * *consonant* in the models was because it is not central to our hypotheses to understand this high-order interaction, as long as the effects of *position* and *region* are investigated for each of the two types of consonants. In addition, the results of the three-way interaction would be difficult to interpret. If we were to conduct all pairwise comparisons for the three-way interaction term (explained further in Section 5.4.3 below), as many as 435 combinations would be required.

As a final step, a simple linear regression was run for each individual in order to investigate the grammatical status of lenis stop (de-)voicing and denasalisation (**H4**, **H6**) as well as the domain of application of the processes (**H5**, **H7**) for each individual. Recall the discussion in Section 5.3.9 that such an analysis must be performed at the level of an individual, in order not to conflate gradience and (group-level) variance. The outcome variables were % voicing and minimum intensity and the fixed variables were *position* and *repetition*[35]. Although it was not necessary, speaker normalised nasal attenuation was used because it is more intuitive in the sense that a higher value indicates greater denasalisation. The model formula is shown below:

```
lm(outcome ~ position + repetition)
```

Table 5.9 provides a summary of the statistical models used in the study.

[35] There were five speakers for whom all tokens came from the first repetition. Thus, repetition was removed for these speakers.

Table 5.9: Summary of the statistical models.

Type	**Level of analysis**	**Consonant**	**Tokens used**	**Outcome variable**	**# of models (Total = 80)**
Bayesian	group	lenis	all	*devoiced*	1
		nasal	all	*denasalised* *devoiced*	2
Non-Bayesian		lenis	all except postpausal	*closure voiceless interval*	5
			all except postpausal or lenited	*residual voicing* *% voicing*	
			all	*VOT*	
		nasal	all except postpausal	*closure voiceless interval*	7
			all except postpausal or lenited	*residual voicing* *% voicing* *intensity*	
			all	*VOT*	
			only postpausal	prevoicing	
		combined	all except postpausal	*closure voiceless interval*	5
			all except postpausal or lenited	*residual voicing* *% voicing*	
			all	*VOT*	
	individual	lenis	all except postpausal or lenited	% voicing	30
		nasal		minimum intensity	30

5.4.3 Tukey HSD Tests

Post-hoc analyses of the non-Bayesian models were conducted through Tukey's Honest Significant Difference (HSD) Tests using the *emmeans* package (Lenth, 2018). Tukey HSD uses ANOVA to compare the differences in the mean between all pairs of levels in a categorical variable. Note that Tukey HSD is similar to Bonferroni correction which is used to control the family-wise Type I error rate (i.e. false positives). Since we are interested in all pairwise comparisons rather than a subset of comparisons, Tukey is considered a better option than Bonferroni correction.

For the group-level models, two rounds of Tukey Tests were performed:

(a) First, pairwise comparisons were made across *position* within each *region*. These results showed which positions are distinguished in terms of the outcome variable for each region.

(b) Then, interaction contrasts were investigated by conducting comparisons between *regions* for each pairwise differences in *position*. These results show whether the direction or the degree of a contrast between positions are different between regions.

For the individual models, I conducted only one round of pairwise comparisons across *position* using individual data, as there was no interaction term in these models. For the Bayesian models, simpler regional comparisons were also made *within* each position in addition to (a) and (b). These results provide insights into whether the patterns of the three varieties differ in a given position.

5.4.4 Hartigan's dip test

Hartigan's dip test (Hartigan & Hartigan, 1985) was used to measure non-unimodality on two parameters, % voicing and minimum intensity. These parameters were used to determine whether lenis stop (de-)voicing and denasalisation are categorical within each individual's grammar, respectively (**H4 – H7**). Hartigan's dip test operates by assessing potential dips in a given sample and returns a *p*-value to indicate whether the dip is statistically significant. The test also produces the dip statistic, or D statistic, which reflects the degree to which the distribution is different from a unimodal distribution. A value closer to zero indicates the distribution is closer to a unimodal distribution. While there are alternative statistical measures of bimodality, the results from the recent studies favoured Hartigan's dip test as the most reliable (Freeman & Dale, 2013; Pfister, Schwartz, Janczyk, Dale, & Freeman, 2013).

5.4.5 Correlation

Pearson's product moment correlation coefficient r was calculated using the *cor.test* function in base *R* to measure the linear correlation between closure duration and % voicing and between closure duration and intensity. A separate analysis was performed separately for each individual. Raw closure duration was used instead of normalised closure duration, because we are interested in how actual closure duration affects % voicing regardless of the fact that speech rate influences actual closure duration. The correlation tests were relevant to testing the duration-based accounts of DIS (**H8 – H10**). A value closer to 1 indicates a stronger positive correlation and a value closer to -1 indicates a stronger negative correlation. The package *ggpubr* was used to add correlation coefficients and the *p*-values to scatter plots (Kassambara, 2019).

5.4.6 Model comparison

In order to test the duration-based account and the rule scattering account of DIS, the four following linear regression models were fitted for each speaker with the outcome *% voicing* and *minimum intensity:*

- Model 1: outcome ~ closure duration
- Model 2: outcome ~ position
- Model 3: outcome ~ closure duration + position
- Model 4: outcome ~ closure duration : position + closure duration + position[36]

Model 1 is consistent with the duration-based account of DIS that DIS can be explained purely as a gradient effect of closure duration (**H10**). On the other hand, Model 2 represents the account that DIS can be predicted based on the prosodic position of a given segment. Here, only three levels of position were used: AP-medial, AP-initial and IP-initial. This decision was made retrospectively based on the results of the Tukey test which showed that 29 out of 30 speakers did not distinguish the three lower positions that are AP-medial. Model 3 uses both duration and position to describe DIS. This model reflects the prediction based on rule scattering that the current pattern of lenis and nasal stop variation is best explained as a categorical effect conditioned by prosody overlaid with a gradient effect of duration (**H11**). Finally, Model 4 takes into account another scenario which is consistent with the rule-scattering account; there is an interaction between duration and position because the gradient effect of duration is different depending on the prosodic position.

The four models were compared using the *anova* function from the *car* package (Fox & Weisberg, 2019). The function returns an ANOVA testing whether the more complex model is significantly better at describing the data than the simpler model. Thus, comparisons were made between Model 1 vs Model 3, Model 2 vs. Model 3, and Model 3 vs. Model 4, to determine the best fitting model using a cut-off of $p < 0.05$. When it was necessary to compare Model 1 and Model 2, i.e. when Model 3 was neither significantly better than Model 1 nor Model 2, the model with a greater adjusted r^2 value was chosen. An adjusted r^2 closer

[36] The R syntax here can be understood as the following. Model 4 is based on the assumption that the outcome is a function of an interaction between closure duration and position, in addition to the main effects of closure duration and position.

to 1 indicates that a greater proportion of the variance is captured by the model. The method described between Section 5.4.3 – Section 5.4.6 is very similar to the approach in Turton (2014) who investigated the issue of categoricity versus gradience of /l/-darkening in English varieties.

5.5 Results

This section will be structured as follows. First, the results from the auditory-acoustic categorisation and the subsequent Bayesian modelling will be reported for the lenis tokens (Section 5.5.1) and the nasal tokens (Section 5.5.2). An interim summary of these results will be provided, and the relevant hypotheses will be evaluated (Section 5.5.3). Next, the results of the acoustic measurements will be presented, along with the results of the standard mixed linear regression models (Section 5.5.4). The order of presentation will be *closure duration* (5.5.4.1), *residual voicing* (5.5.4.2), *percentage of voicing during closure* (5.5.4.3), *VOT* (5.5.4.4), *total voiceless interval* (5.5.4.5), *prevoicing,* (5.5.4.6) and *nasal attenuation* (5.5.4.7). While the focus here is the variation by *prosodic position* and *region,* the effects of *gender* (5.5.4.8) and *repetition* (5.5.4.9) will also be discussed briefly. An interim summary for the acoustic results will be given, as well as evaluation of the relevant hypotheses (Section 5.5.5). Then, the status of lenis stop (de-)voicing (Section 5.5.6) and denasalisation (Section 5.5.7) in relation to the life cycle of phonological processes will be discussed, using the results from Hartigan's dip test and Pearson's correlation. Finally, the results from the model comparison will be reported in order to test the duration-based account and the rule-scattering account of DIS (Section 5.5.8). At the end, a summary of the last three sections will be provided and the remaining hypotheses will be evaluated (Section 5.5.9).

5.5.1 Categories of realisation for lenis stops

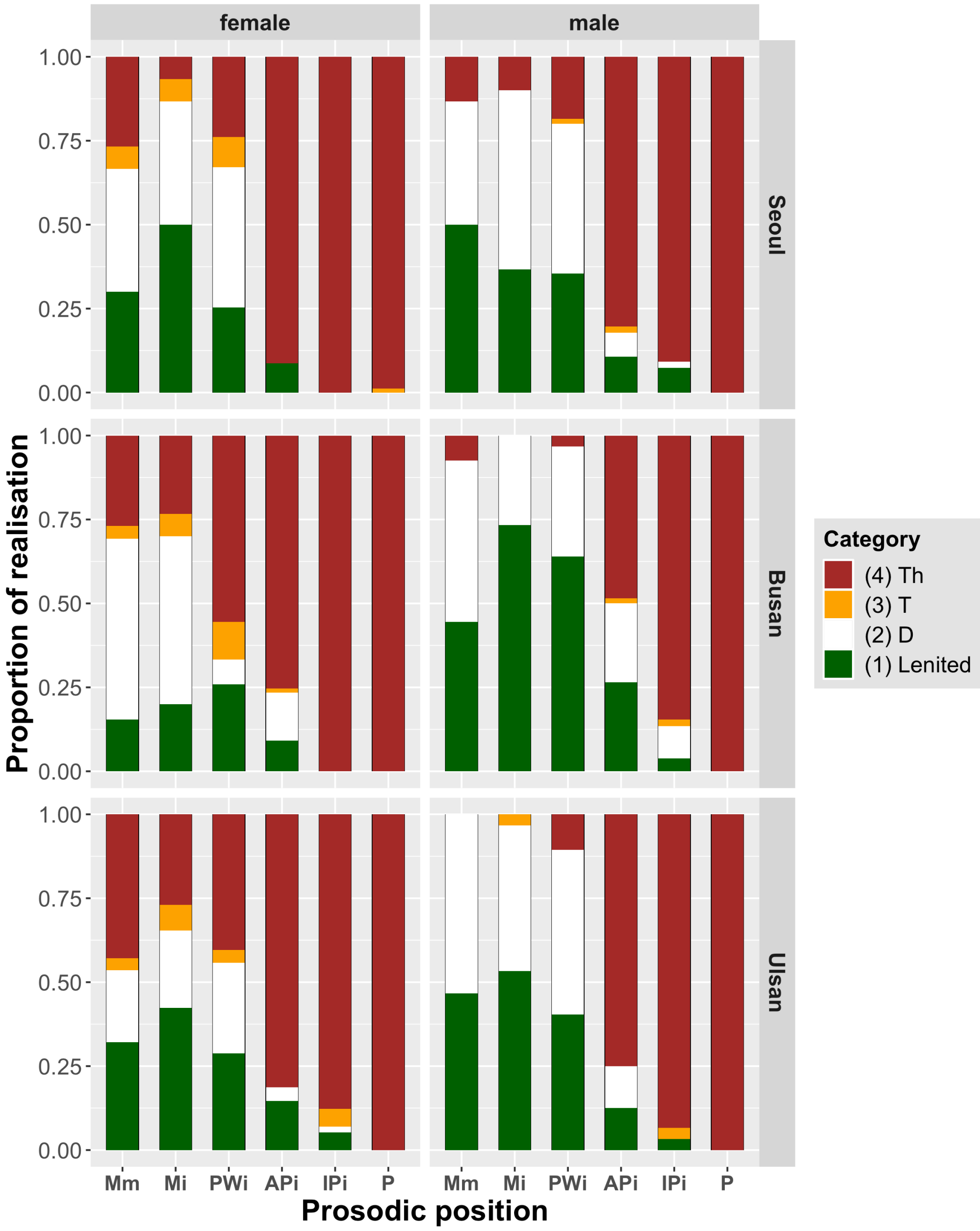


Figure 5.16: Proportions of lenis stops in four categories of realisation across prosodic position for Seoul (Top), Busan (Middle), and Ulsan (Bottom). The left and right panels are for the females and males, respectively.

Figure 5.16 presents the distribution of categories for the lenis tokens across six prosodic positions. I will first make some informal observations. Overall, there were very low proportions of Category (3) T (tokens with VOT of 0 - 10 ms), a category which was included to emulate Category (4) for the nasals. This is not surprising, given that the mean VOT value of the lenis stops is expected to be much longer than 10 ms, approximately 65 ms for Seoul and 40 ms for South Gyeongsang according to H. Lee & Jongman (2012). The mean VOT values from our data will be given in Section 5.5.4.4 below.

Perhaps more surprising is the relatively high proportions of lenited realisations. For the Busan males, for example, almost 75% of the Mi tokens were realised with lenition. Generally, there is a trend for the lenition rate to fall in a higher prosodic position. However, there is an exception to this overall trend, namely that lenition is more frequent in the Mi position than in the Mm position (Note that there were both in the Syllable-initial position). This can be explained by the fact that the Mi tokens were in the initial position of a grammatical particle. Thus, they may have been subject to less strengthening effects than the Mm tokens which were part of a content word.

Considering that all lenited tokens were realised with voicing, we can obtain the proportion of voiced realisations by combining Categories (1) and (2). Given the Lenis Stop Voicing Rule, a dichotomous pattern is expected in which most AP-medial tokens are voiced and most AP-, IP-initial and P tokens are voiceless. Indeed, there is a tendency for a split between Mm – PWi vs. APi – P. For the Seoul group and the Ulsan males, we can observe a sharp decrease in the proportion of voiced realisation between the PWi and APi positions. For the other groups, the dichotomy is less clear. Regardless of whether there is a sign of dichotomy, we can observe a gradual decline in the proportion of voiced realisation towards a higher position for all groups, with the aforementioned exception of the Mi position. The issue of categoricity versus gradience will be discussed further in Section 5.5.6.

The results of the Bayesian model with the outcome *devoiced* confirmed these trends. From the 18 pairwise comparisons across region within each prosodic position, I mainly discuss those results which constitute *strong evidence* for a difference. I take as strong evidence a posterior distribution with a 95% CI interval that does not include 0. For the full results, see Appendix B. Strong evidence for a group difference was only found for the APi position between Seoul and Busan. To interpret the results further, predicted probabilities were calculated manually using the model estimates[37]. The predicted probability that an APi token will be devoiced was 0.972 with a 95% CI of [0.956, 0.997] for Seoul, 0.864 with a CI of [0.613, 0.976] for Busan, and 0.941 with a CI of [0.796, 0.993] for Ulsan. The difference between Seoul and Busan was 0.108 with a 95% CI of [0.001, 0.341]. Given that the CI does

[37] More specifically, the calculation involved extracting the intercept and the relevant fixed effects estimates from the posterior samples matrix and performing a matrix multiplication to obtain the model predictions for the outcome variable. For the R code, see "PhD_Production_Study_02_Bayesian_models.Rmd" from https://osf.io/72k9u/.

not include 0, the model provides strong support that the lenis stops are more likely to be devoiced in APi position for Seoul than for Busan. The difference between Seoul and Ulsan was 0.031 with a 95% CI of [-0.034, 0.158]. While zero is within the CI, 80.1% of the probability mass lies above 0. This suggests that the model is 80.1% confident that APi tokens are more likely to be devoiced for Seoul than for Ulsan. Finally, between Busan and Ulsan, the difference was -0.077 with a CI of [-0.303, 0.071], with 13.4% of the probability mass above zero. Thus, we have 86.6% confidence that APi tokens are more likely to be devoiced for Ulsan than for Busan.

Despite the regional differences within each position, the cumulative trend across position was similar. For all three regions, the model provided strong evidence for the following hierarchy: Mm = Mi = PWi < APi < IPi < P. In other words, the tokens in P position were more likely to be devoiced than the tokens in all the lower positions, the tokens in IPi position were more likely to be devoiced than those in all the lower positions, and so on. For the AP-medial positions, Mm, Mi and PWi, the model did not provide evidence for a reliable difference. The pairwise comparisons between regions for the contrasts between positions did not show strong evidence for any group difference.

In addition, it is possible to observe an effect of gender in the plot. For Busan and Ulsan in particular, females produced more voiceless realisations than males in most positions, except for P position in which both females and males produced 100% of the tokens with devoicing. The model estimate indicated a strong effect of *gender* with a mean of -2.07 with a 95% CI of [-3.16, -1.07], in log odds. The predicted probability of devoiced realisation was 0.509 with a CI of [0.184, 0.814] for females but only 0.137 with a CI of [0.027, 0.345] for males. The difference was 0.373 [0.130, 0.604]. Since zero is well outside the CI, the model provides compelling evidence that females are more likely to produce the lenis stops with devoicing than males.

Finally, recall that *repetition* was included in the model to control for the fact that some utterances were immediately preceded by another instance of the same utterance. The model provided some evidence that devoicing is more likely in the first repetition, i.e. when an utterance was not immediately preceded by another instance of the utterance, than in the second repetition, i.e. when an utterance was immediately preceded by another instance of the utterance. The predicted probability of devoicing was 0.51 [0.18, 0.81] for the first repetition but 0.45 [0.16, 0.76] for the second repetition. Their estimated difference was 0.06 [-0.02, 0.14]. While zero was not included in the CI for the difference, 92.4% of the probability was above 0.

5.5.2 Categories of realisation for nasal stops

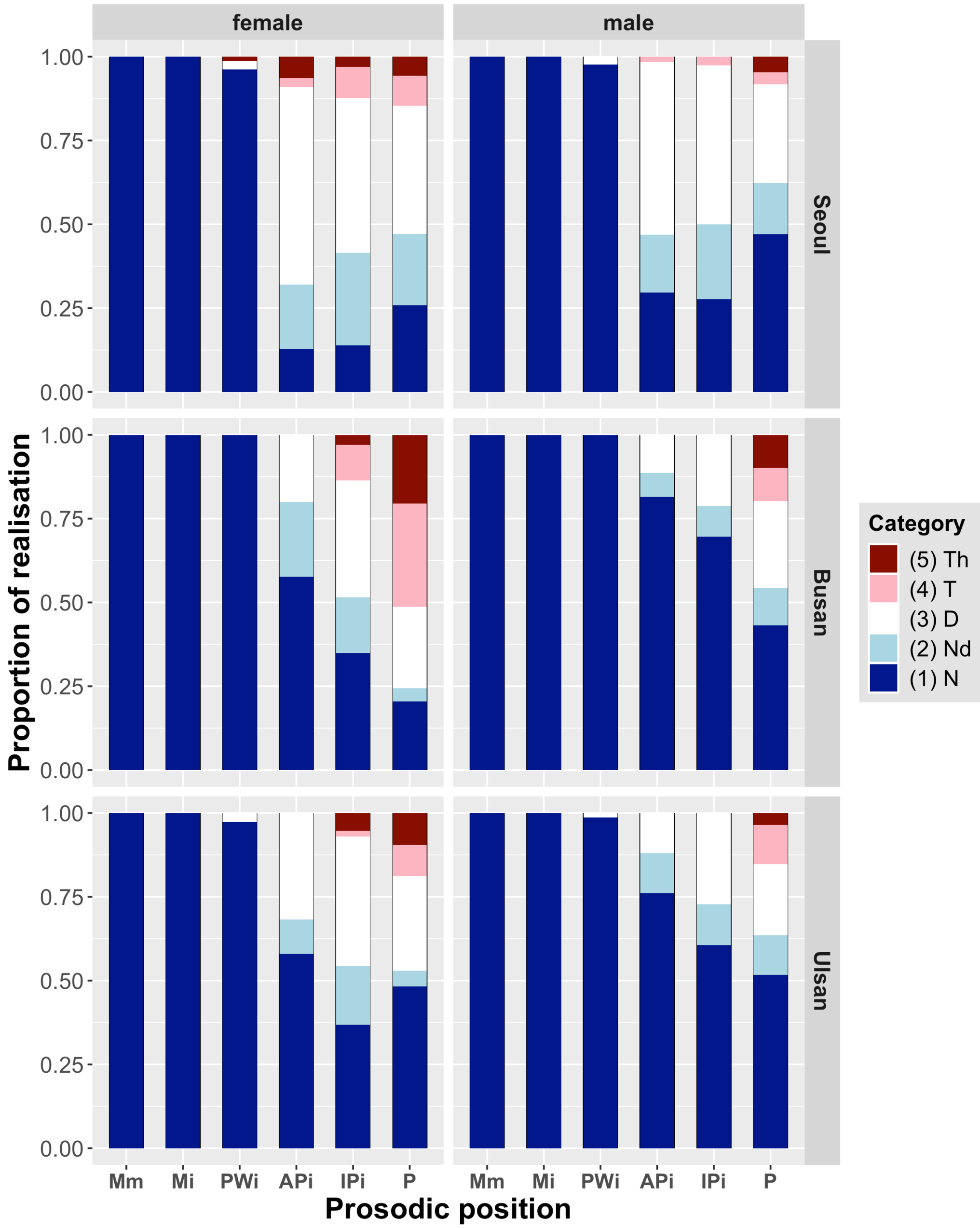


Figure 5.17: Proportions of nasal stops in five categories of realisation across prosodic position for Seoul (Top), Busan (Middle), and Ulsan (Bottom). The left and right panels are for the females and males, respectively.

Figure 5.17 presents the distribution of categories for the nasal tokens across six prosodic positions. The exact proportions can be found in Appendix C. First, I will consider the contrast between Category (1) and the other categories, to compare the proportion of tokens realised with some degree of denasalisation, partial or complete. A clear prosodic pattern can be observed in all six groups. In Mm – PWi positions, either all or a vast majority (96 ~ 100%) of the tokens in Mm – PWi positions were in Category (1). On the other hand, in APi – P positions, we can see higher proportions of other categories. In Seoul, a majority (53 – 87%) of the tokens in these positions were either partially (Category 2) or fully denasalised (Categories 3- 5). It is noteworthy that P tokens showed a higher proportion of Category (1) than APi or IPi tokens, which indicates that denasalisation is less common in this position. In contrast, both Busan females and males as well as Ulsan males showed a cumulative pattern in which denasalisation is progressively more common in a higher position. For the Ulsan female group, P tokens show an intermediate proportion of Category (1) between APi and IPi tokens.

Next, we can observe regional variation within each prosodic position. Pooling data from both genders, Seoul showed the lowest proportion of Category (1) in the PWi – IPi positions (97%, 20%, 21%, respectively). Busan and Ulsan showed similar proportions in the PWi – IPi positions but Ulsan showed a slightly lower proportion in each of these positions (Busan: 100%, 68%, 52%; Ulsan: 98%, 66%, 50%). In P position, Busan showed the lowest proportion of Category (1), with Seoul showing a slightly higher proportion and Ulsan showing a much higher proportion (Seoul: 36%; Busan: 32%; Ulsan 50%). Overall, the proportion of Category (1) was 56% for Seoul, 70% for Busan, and 73% for Ulsan, indicating that denasalisation is more common for Seoul than for Busan and Ulsan.

The results of the Bayesian model with the outcome *denasalised* were consistent with these observations. In APi and IPi positions, there was strong evidence for a difference between Seoul and Busan and between Seoul and Ulsan, according to the pairwise comparisons of estimated marginal means. In APi position, the predicted probability of denasalisation was 0.873 with a 95% CI of [0.710, 0.963] for Seoul. For Busan, the probability was only 0.377 [0.155, 0.657] and for Ulsan, the probability was slightly higher at 0.419 [0.179, 0.694]. The difference between Seoul and Busan was 0.495 [0.237, 0.721] and the difference between Seoul and Ulsan was 0.454 [0.206, 0.680]. Since zero is not within the CIs for the differences between Seoul vs. Busan and Ulsan, we have strong evidence that denasalisation is more likely for Seoul than in the two other regions in APi position. On the other hand, the difference between Busan and Ulsan was -0.041 with a CI of [-0.317, 0.243]. Given that zero is well within the CI, the model does not provide any support for a difference between Busan and Ulsan in terms of denasalisation in APi position.

The pattern was similar for the IPi position. For Seoul, the predicted probability of denasalisation was as high as 0.889 [0.727, 0.974]. For Busan, the probability was 0.615 [0.308, 0.873] and for Ulsan, it was 0.621 with a CI of [0.324, 0.864]. The estimated difference between Seoul and Busan was 0.274 [0.028, 0.570] and the difference between Seoul and Ulsan was 0.268 [0.035, 0.552], which strongly suggests that denasalisation in IPi position is more likely for Seoul than in the other regions. On the other hand, the estimated difference between Busan and Ulsan was -0.006 [-0.328, 0.335], which means that we cannot establish any difference between Busan and Ulsan for the pattern of denasalisation in IPi position.

The prosodic position within each region was also similar to our visual observations above. For Seoul, pairwise comparisons across position found the following hierarchy: Mm, Mi, PWi < APi, IPi, P. That is, the 95% CI for the contrast between an AP-medial position and an AP-initial (or a higher position) did not include 0, providing compelling evidence that AP-initial tokens are more likely to be denasalised than AP-medial tokens for Seoul. Additionally, PWi tokens were more likely to be denasalised than Mi tokens, and IPi tokens were more likely to be denasalised than P tokens, consistent with the pattern in the plot. In contrast, both Busan and Ulsan showed the pattern, Mm = Mi = PWi < APi < IPi = P. This indicates that denasalisation follows a cumulative trend more closely for these two regions, making a three-way distinction among the (a) AP-medial positions, (b) APi position, and (c) IPi and P positions.

The pairwise comparisons between regions for the contrasts between positions showed strong evidence for a group difference between Seoul and Busan for the following contrasts: between PWi vs. P (3.80 [0.14, 9.74]), APi vs. P (2.74 [1.63, 3.89]) and IPi vs. P (1.87 [0.66, 3.26]). The contrast between PWi and P was bigger for Busan than for Seoul. In other words, P was more likely to show denasalisation than PWi by a greater degree for Busan than for Seoul. For the contrasts between APi vs. P and IPi vs. P, the direction of the effect was opposite for Seoul and Busan. For Seoul, APi and IPi tokens were more likely to show denasalisation than P tokens, whereas for Busan, APi and IPi were less likely to show denasalisation than P tokens. Ulsan also showed a different pattern from Seoul, in that APi tokens were less likely to show denasalisation than P tokens (1.79 [0.63, 2.83]). To summarise, these results showed that between Seoul and the other regions, there is a difference in how much more or less likely P tokens are to be denasalised relative to the tokens in PWi, APi and IPi positions.

Comparing across gender, for all three regions, it can be confirmed visually that females showed a lower proportion of Category (1) than males in each prosodic position. The overall proportion of Category (1) was 60% for females and 73% for males. This suggests that females are more likely to denasalise a nasal phoneme in a given position. The model prediction for *gender* indicated strong evidence for the variable *gender* with a mean of -0.97 with a 95% CI of [-1.84, -0.11], in log odds. The predicted probability of denasalisation was

0.00825 [0.00027, 0.03604] for females and 0.00319 [0.00009, 0.01336] for males. Their difference was 0.00506 [0.00004, 0.02439]. Although the numbers are small due to the low overall probability of denasalisation, the CI does not include 0. Thus, the model offers strong evidence that females are more likely to show denasalisation than males.

Next, consider the contrast between Categories (1) – (3) versus Categories (4) and (5), which divides the categories into those with full voicing and those with some degree of devoicing (Note that a token that is voiceless at the release was classified as Category (4) T, even if there was residual voicing during the closure). The prosodic pattern can be described as cumulative in all six groups, as an equal or a greater proportion of Categories (4) + (5) is found in a higher position.

An interesting regional difference can be observed. While the positions in which a voiceless denasalised nasal is found are more restricted for Busan and Ulsan than for Seoul, the proportion of voiceless realisation is higher in P position in these two regions. In particular, for the Busan female group, a total of 52% of P tokens was realised as a voiceless non-nasal.

The Bayesian model with an outcome *devoiced* found the following hierarchies in agreement with the observed cumulative trend. For Seoul, there was a pattern in which IPi and P nasals were more likely to be devoiced than AP-medial nasals (Mm = Mi = PWi < IPi = P). Additionally, P nasals were more likely to be devoiced than APi nasals (APi < P). For Busan, there was a three-way distinction in which P nasals were more likely to be devoiced than IPi nasals, and in turn, IPi nasals were more likely to be devoiced than IP-medial nasals (Mm = Mi = PWi = APi < IPi < P). For Ulsan, there was only a two-way distinction in which P-nasals were more likely to be devoiced than the rest of the nasals (Mm = Mi = PWi = APi = IPi < P).

The interaction contrasts showed that Seoul and Busan were different for the contrasts between P and the lower positions of Mm (2.59 [0.14, 5.58]), APi (5.16 [0.43, 11.2]), and IPi (1.85 [0.05, 3.87]). The contrast between P and the three lower positions was greater for Busan than for Seoul. This is consistent with the modelling result that, within each prosodic position, the only strong support for a regional difference was between Seoul and Busan in P position (-1.85 [-3.13, -0.63]).

To understand the regional difference in P position better, the predicted probability of devoicing for P nasals was manually calculated for each region. The predicted probability of devoicing was 0.12 [0.03, 0.31] for Seoul, 0.44 [0.18, 0.72] for Busan, and 0.24 [0.07, 0.53] for Ulsan. The difference between Seoul and Busan was -0.32 with a 95% CI of [-0.56, -0.10]. Since the CI does not include zero, the model provides compelling evidence that Seoul is less likely to show devoicing than Busan in P position. The difference between Seoul and Ulsan was -0.12 with a CI of [-0.34, 0.04]. While the CI includes 0, 80.1% of the probability mass lies below 0. Thus, there is some evidence that devoicing in P position is less likely for Seoul than for Ulsan. Finally, the difference between Busan and Ulsan was 0.20 with a CI of [-0.04, 0.45]. Again, 86.7% of the probability mass lies below 0, providing some evidence that devoicing is more likely for Busan than for Ulsan in P position. Overall, the results suggest that devoicing in P position is most likely for Busan, intermediate for Ulsan, and least likely for Seoul.

Between the females and the males, the females always showed an equal or a higher proportion of voiceless realisation than the males in each prosodic position for all three regions. The model provided strong evidence for the effect of *gender* with a coefficient -1.07 with a CI of [-1.92, -0.26]. The predicted probability of devoicing for females was 0.00333 with a 95% CI of [0.00006, 0.01593]. for males, the estimated probability is 0.00117 with a 95% CI of [0.00002, 0.00556]. Their estimated difference was 0.00216 with a 95% CI of [0.00003, 0.01054]. Similar to the probability of denasalisation, the low estimates are explained by the fact the probability of devoicing is generally very low. What is important is that zero lies outside the CI, which constitutes strong evidence that devoicing of domain-initial nasals is more likely in females than in males.

Finally, the models for *denasalisation* and *devoicing* did not provide any support for the effect of repetition. For denasalisation, the estimated probability was 0.00825 [0.00027, 0.03604] for the first repetition and 0.00828 [0.00029, 0.03536] for the second repetition. Their estimated difference was 0.00003 [-0.00341, 0.00363]. Given that zero is almost in the centre of the distribution, the model gives no support for the difference between the two repetitions. For devoicing, the estimated probability was 0.00333 [0.00006, 0.01593] for the first repetition and 0.00314 [0.00006, 0.01457] for the second repetition. Again, their difference was not reliable at all, with a mean of 0.00019 with a 95% CI of [-0.00173, 0.00275].

5.5.3 Summary and discussion I

Sections 5.5.1 and 5.5.2 presented the analysis of the types of phonetic realisation for the lenis and the nasal tokens. In general, denasalisation and devoicing were more likely in a higher position for all three varieties, although they sometimes differed regarding which and how many positions were distinguished.

For devoicing of the lenis stops, all three varieties showed a highly regular pattern.

Seoul, Busan, Ulsan: Mm = Mi = PWi < APi < IP < P

Despite the similarity in the hierarchical pattern, a regional difference was found in APi position between Seoul and Busan. APi lenis stops were more likely to be devoiced for Seoul than Busan. There was some evidence that Ulsan was intermediate between Seoul and Busan. Therefore, the results provide evidence against **H2** in its strictest sense: for the lenis stops, there is no regional variation.

For denasalisation, the pattern can be summarised as below.

Seoul: Mm, Mi, PWi < APi, IPi, P; Mi < PWi; IPi > P

Busan & Ulsan: Mm = Mi = PWi < APi < IPi = P

Note that there was a violation of a cumulative pattern for Seoul in which IPi tokens were more likely to be denasalised than P tokens. This will be discussed further in Section 5.6.5. The regional difference was robust between Seoul vs. Busan & Ulsan for APi and IPi positions. In both positions, Seoul was more likely to show denasalisation than the two other regions. Taken together with the raw proportions of denasalisation across position, the results suggest that denasalisation is highly and similarly likely in the three high positions APi – P for Seoul, whereas it is progressively more likely in a higher position for Busan and Ulsan. The pattern that denasalisation is most likely for Seoul in APi and IPi positions partially supports **H3**: for the nasal stops, the DIS effects are greater in the order of Seoul, Busan, and Ulsan, following the model of hierarchical diffusion (Trudgill, 1974)[38]. However, there was no evidence for a difference between Busan and Ulsan.

The last model for nasal devoicing showed the following the prosodic pattern.

Seoul: Mm = Mi = PWi < IPi = P; APi < P

Busan: Mm = Mi = PWi = APi < IP < P

Ulsan: Mm = Mi = PWi = APi = IP < P

[38] Yoshida (2008, p. 2) cites Umeda (1989) who observed that the further north in the Korean peninsula, the stronger the tendency for denasalisation. This may suggest that denasalisation started from dialects further north to Seoul, against the prediction of hierarchical diffusion. However, a recent study (S. Yun & Kang, 2018) found that speakers from Seoul show more frequent denasalisation than speakers from Northern Hamkyeong in North Korea.

Unlike the results for denasalisation, none of the three varieties distinguished APi position from the AP-medial positions. However, tokens in IPi position and/or P position were more likely to be devoiced than tokens in lower positions. The dialectal differences were found in P position. There was strong evidence that nasal devoicing in P position is more likely for Busan than Seoul and some evidence that Ulsan is intermediate between Busan and Seoul. Observe that this pattern is opposite to the pattern of lenis devoicing, which was strongest for Seoul and weakest for Busan, in APi position. The results for nasal devoicing go against the hypothesis in **H3** which states that greater effects of DIS in the nasal stops are found in the order of Seoul, Busan, and Ulsan.

The last finding above might not be in the expected direction, but there is a possible explanation. While it is only speculation at this stage, greater nasal devoicing in P position for Busan and Ulsan speakers may be related to the phenomenon known as word-initial tensification, whereby lenis consonants are realised as fortis consonants in word-initial position (see Section 3.6). Previous research has shown that word-initial tensification is likely to have started in Gyeongsang Korean, where the phenomenon is still more prevalent (J. Oh, 2011; S. Jang, 2017). If we assume that word-initial tensification comes from a tendency for a glottal constriction in word-initial position, this tendency may also work against the production of (pre-)voiced obstruents in word-initial position. Recall that denasalisation is most likely in P position, causing nasals to be realised as a pre-voiced obstruent. If initial voiced obstruents are indeed less acceptable in Gyeongsang Korean, i.e. Busan and Ulsan Korean, this would explain why these segments are more likely to undergo devoicing in Gyeongsang Korean than in Seoul Korean.

Although variation by gender was not the focus of the study, the models provided compelling evidence that females are more likely to show denasalisation (for nasals) and devoicing (for both consonants) than males. Finally, the model provided some evidence that lenis devoicing was more likely the first repetition than in the second repetition, but no such evidence was found for denasalisation and nasal devoicing.

5.5.4 Results of acoustic analysis

In this section, I will report the results of the mixed effects linear regression models and the post-hoc analysis using Tukey HSD Test for the following acoustic measurements: (1) *closure duration,* (2) *residual voicing,* (3) *percentage of voicing during closure,* (4) *VOT,* (5) *total voiceless interval,* (6) *prevoicing,* and (7) *nasal attenuation* (measured through acoustic intensity). The full results from the regression models and the Tukey tests are provided in Appendices D and E, respectively. In these sections, I will only focus on the significant results for the main predictor variables of *prosodic position* and *region*. Note that the regression models also included the control variables *gender* and *repetition* and the results from the Tukey Tests were averaged over the levels of *gender* and *repetition*, such that the results for *position* and *region* would not be biased. The subsequent sections will address these factors briefly: (8) *gender* and (9) *repetition*.

5.5.4.1 Closure duration

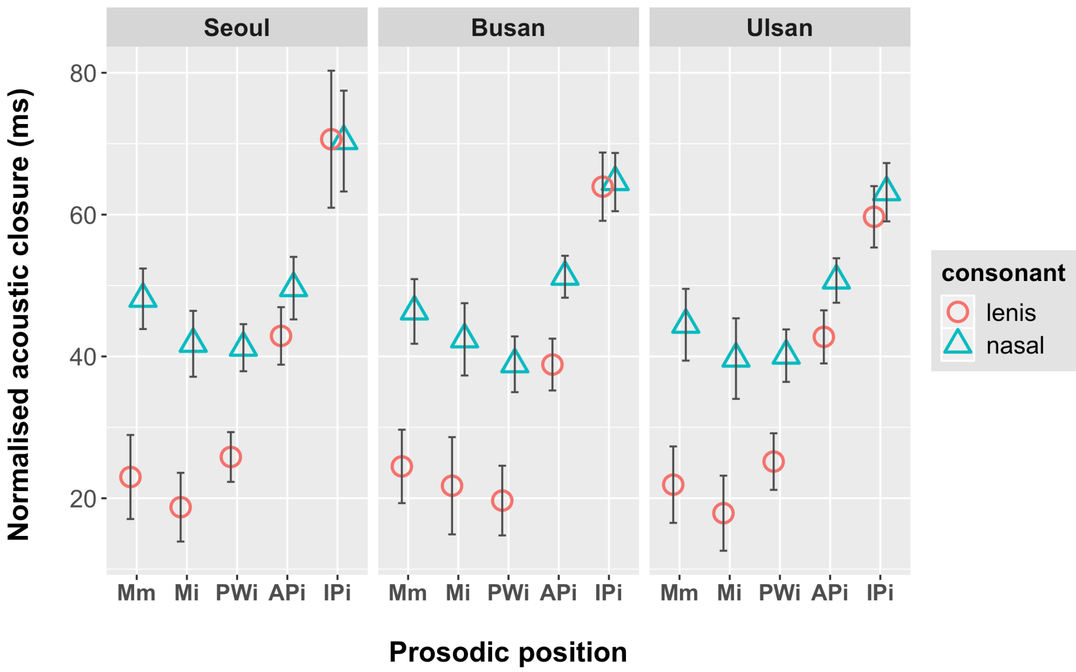


Figure 5.18: Normalised acoustic closure (ms) by position. Error bars show 95% CIs.

Figure 5.18 shows the variation of normalised acoustic closure (ms) of the lenis and nasal tokens. First, consider the prosodic variation within each region. As predicted, closure duration is generally longer in a higher domain. This cumulative pattern is clearer for the lenis than for the nasal stops. The Tukey test showed that, for all three varieties, closure duration of the lenis stops was significantly longer in IPi position than in APi position, which was in turn significantly longer than in PWi position. No pairwise comparison was significant among the lower domains of Mm, Mi and PWi positions (Mm = Mi = PWi < APi < IPi).

Despite the visual similarity across *region* in Figure 5.18, there were some statistically significant regional differences in terms the prosodic variation of the lenis stops. The differences between the IPi position and the lower positions of the APi and PWi positions were significantly larger for Seoul and Busan than for Ulsan. In addition, the difference between the IPi position and the Mi position was also significantly greater for Seoul than for Ulsan. These differences seem to be driven mainly by the relatively long closure for the IPi position for Seoul and the relatively short closure for the APi and PWi positions for Busan.

For the nasal stops, the Tukey results for prosodic variation were slightly different by region. For Seoul, closure duration was significantly longer in IPi position than in all the other positions (Mm = Mi = PWi = APi < IPi). The difference between PWi and APi positions did not reach statistical significance ($p = 0.07$). For Busan, IPi nasals were also significantly longer than nasals in all the other positions. Additionally, APi nasals were significantly longer than PWi nasals but not Mi and Mm nasals (Mm = Mi = PWi < IPi; PWi < APi < IPi). For Ulsan, the pattern was similar to that for Busan, except that APi nasals were also significantly longer than Mi nasals (Mm = Mi = PWi < IPi; Mi = PWi < APi < IPi). The tendency for PWi nasals to be shorter than Mm and/or Mi nasals, which are prosodically in PW-medial, Syllable-initial position, is in line with the finding in T. Cho and Keating (2001, p. 172) that PWi nasals are shorter than Syllable-initial nasals (which were Morpheme-medial), both in terms of acoustic closure and articulatory seal duration.

The interaction contrasts showed that the difference between the APi and IPi positions was significantly greater for Seoul than for Busan and Ulsan. This is presumably due to the relatively short duration of APi nasals and the relatively longer duration of IPi nasals for Seoul. None of the other differences reached statistical significance.

Finally, Figure 5.18 suggests a systematic interaction between *consonant* and *position*. In all three varieties, the nasal stops were generally longer than the lenis stops but the difference gradually disappeared in a higher position. This pattern was statistically validated using a combined model with both types of consonants and subsequent pairwise comparisons between the lenis and the nasal stops within each prosodic position. The estimated difference gradually decreased in a higher position from 25.25 ms in Mm position to 12.07 ms in APi position. The difference between the consonants was not statistically significant in IPi position.

5.5.4.2 Residual voicing

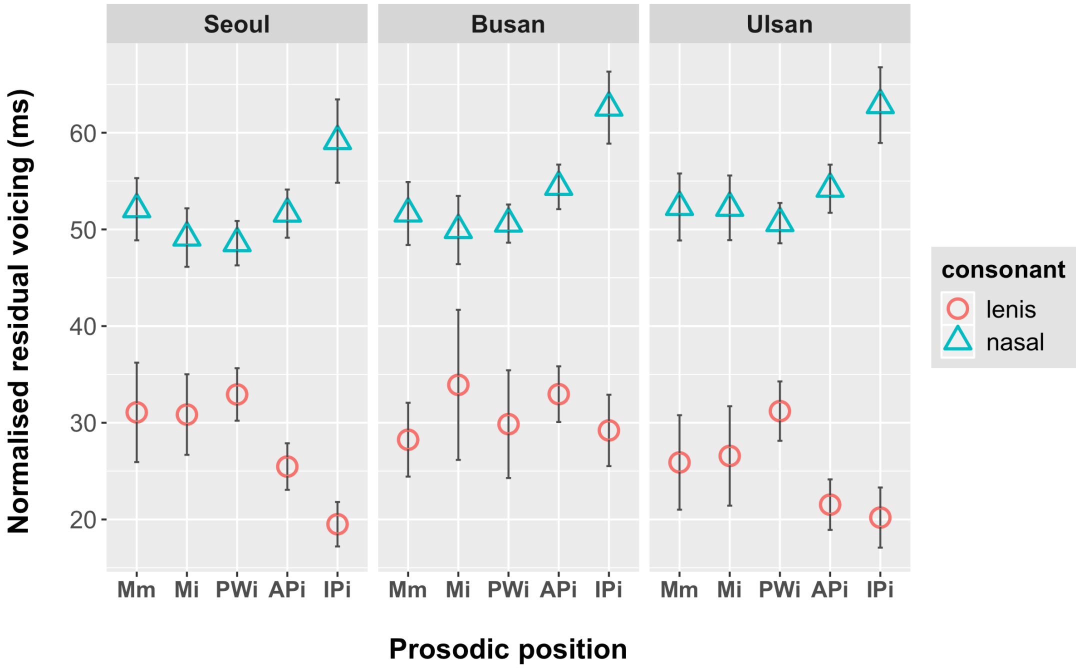


Figure 5.19: Normalised residual voicing (ms) by position. Error bars show 95% CIs.

Figure 5.19 shows the pattern of normalised residual voicing across prosodic position in the three varieties. Recall that for fully voiced consonants, residual voicing was the same as closure voicing. In contrast to the finding in T. Cho and Keating (2001, p. 172), the amount of voicing into closure for the lenis stops appears to vary across *position*. A visual inspection of Figure 5.19 suggests that, for the lenis stops, residual voicing is shorter in tokens in APi and IPi positions than tokens in the lower positions for Seoul and Ulsan, but not for Busan. It is also interesting to compare these results with the results for closure duration. While the prosodic variation of the nasals is roughly parallel for closure duration and residual voicing, the prosodic variation of the lenis stops shows diverging trends. That is, closure duration of the lenis stops tends to be longer in APi and IPi positions, whereas residual voicing tends to be shorter in the same positions.

The statistical results revealed that, for Seoul, the lenis tokens in IPi position showed significantly shorter residual voicing than in all the other positions (Mm = Mi = PWi, APi > IPi). Residual voicing was also significantly shorter in APi position than in PWi position (PWi > APi). For Busan, however, there was no significant difference between any two positions (Mm = Mi = PWi = APi = IPi). For Ulsan, residual voicing was significantly shorter for IPi tokens compared to Mm, Mi and PWi tokens. It was also significantly shorter in APi tokens than in PWi tokens. Unlike for Seoul, APi and IPi tokens were not significantly different (Mm = Mi = PWi > APi = IPi).

Consistent with the visual trend in Figure 5.19, the interaction contrasts showed that Busan is the odd-one-out in the pattern of prosodic variation. The contrast between PWi vs. APi positions and the contrast between PWi vs IPi positions were both greater for Seoul and Ulsan compared to Busan. In addition, the contrasts between Mm vs. APi, Mm vs. IPi, and Mi vs. IPi were all greater for Seoul than for Busan. On the other hand, there was no significant difference between Seoul and Ulsan.

For the nasal stops, the three varieties showed a highly similar pattern of prosodic variation, as we can observe in Figure 5.19. For all three varieties, IPi nasals showed significantly longer residual voicing than the nasals in the lower positions (Mm = Mi = PWi = APi < IPi). Interaction contrasts confirmed that there was no significant difference between all pairs of regions in terms of prosodic variation.

Finally, a combined model showed that the difference between the lenis and the nasal stops is statistically significant within each level of prosodic position. The consonantal difference was greatest in the IPi position, with the estimated difference of 41.9 ms, whereas the differences in the lower positions ranged between 21.7ms (PWi) and 29.8 ms (APi).

5.5.4.3 Percentage of voicing during closure

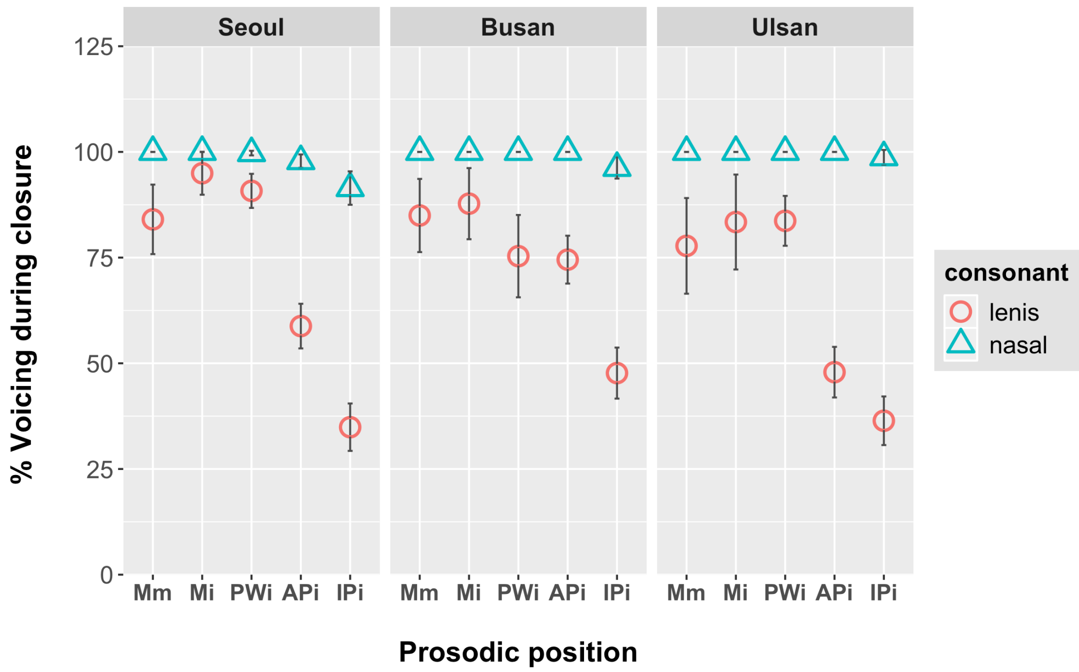


Figure 5.20: Percentage of voicing during closure by position. Error bars show 95% CIs.

As shown in Figure 5.20, the nasal stops did not vary as much across position, although a slightly lower % voicing was observed for IPi nasals. On the other hand, the lenis stops generally showed a lower % voicing during closure in a higher prosodic position. Given the Lenis Stop Voicing Rule, we would expect the AP-medial tokens of lenis stops to show a similar value near 100% and APi and IPi tokens to show a lower % of voicing. Indeed, this is roughly the pattern for Seoul and Ulsan. For Busan, however, only IPi tokens show a clearly lower mean of % voicing compared to the tokens in the lower positions.

Statistical analyses for the lenis tokens showed a three-way distinction for Seoul and Ulsan. That is, % voicing for APi position was lower than all the lower positions and % voicing for IPi position was lower than for all the lower positions (Mm = Mi = PWi > APi > IPi). On the other hand, there was only a two-way contrast for Busan between IPi position and the rest (Mm = Mi = PWi = APi > IPi).

For regional variation, the differences between APi position and the lower positions Mm – PWi and the differences between the IPi and the lower positions Mm – PWi were all significantly greater for Seoul than for Busan. The difference between APi and IPi positions was greater for Seoul and Busan than for Ulsan. Lastly, the differences between APi and the lower positions and between IPi and PWi positions were greater for Ulsan than for Busan.

Moving on to the nasals, pairwise comparisons between prosodic positions within each region showed that, for Seoul and Busan, % voicing was significantly lower for IPi nasals than for nasals in all the other positions (Mm = Mi = PWi = APi > IPi). For Ulsan, there was no significant difference between any two positions (Mm = Mi = PWi = APi = IPi).

Regional differences were greatest between Seoul and Ulsan. The differences in % closure between IPi vs. Mm, Mi, and PWi positions were significantly greater for Seoul than for Busan and Ulsan. In addition, the differences between IPi and APi positions were also significantly greater for Seoul than for Ulsan. Comparing these results to the results for closure duration and residual voicing, one can make the following observation. The regional differences in % voicing are mainly due to the differences in closure duration rather than the differences in residual voicing. This is based on the finding that there were no significant regional differences in residual voicing whereas Seoul vs. Busan and Ulsan behaved differently in terms of closure duration.

Finally, a combined model suggested that the lenis stops show significantly lower % voicing than the nasal stops in each level of prosodic position. The difference was smallest in PWi position, intermediate in APi position, and largest in IPi position, with Mm and Mi positions falling between PWi and APi position.

5.5.4.4 VOT

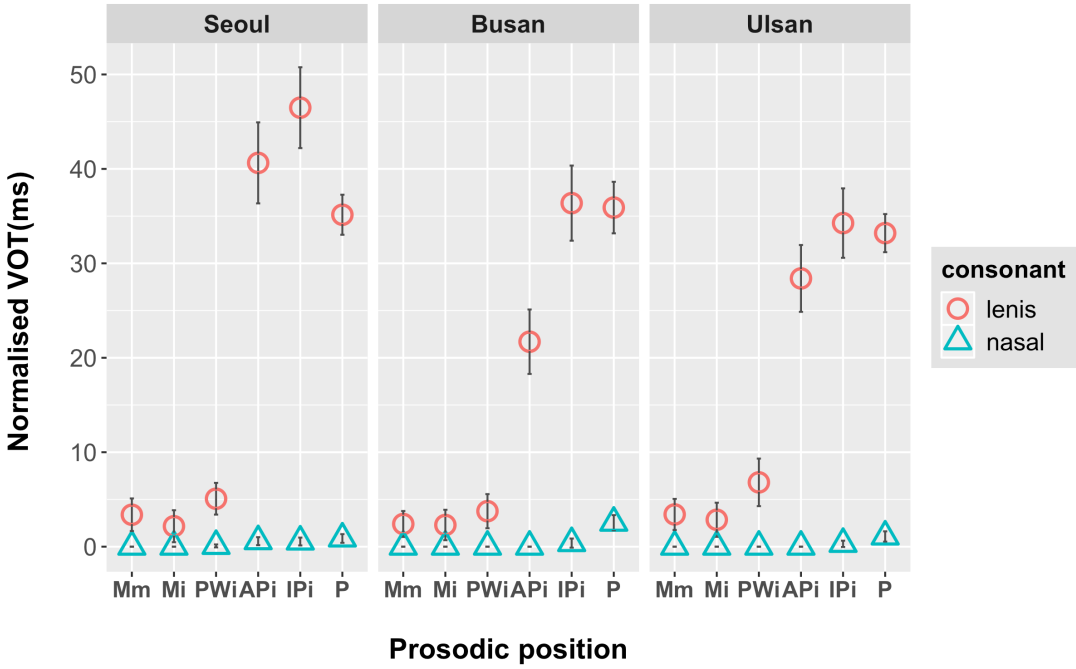


Figure 5.21: Normalised VOT (ms) by position. Error bars show 95% CIs.

The results in Figure 5.21 show that the mean normalised VOT of the lenis stops was clearly longer in APi – P positions than in the lower positions. However, a strictly cumulative pattern found in T. Cho and Keating (p. 173) was violated here as P tokens showed a similar or a lower mean VOT than IPi tokens. On the other hand, VOT of the nasal tokens did not vary greatly, remaining close to zero across different positions. However, it is worth noting that the mean VOT of P tokens was slightly higher than that of tokens in the other positions, particularly for Busan.

The statistical analyses were in line with the visually observed patterns. For the lenis stops, VOTs in Mm – PWi positions were significantly shorter than in APi – P positions for all three regions. In addition, APi tokens showed a significantly shorter VOT than IPi tokens for all groups. However, the patterns with P tokens were inconsistent. For Seoul, both APi and IPi tokens showed significantly longer VOT than P tokens (Mm = Mi = PWi < P < APi = IPi). For Busan, only APi tokens showed significantly shorter VOT than P tokens (Mm = Mi = PWi < APi, IPi, P; APi < P). Finally, for Ulsan, both APi and IPi tokens were not significantly different from P tokens (Mm = Mi = PWi < APi = IPi = P).

Many of these regional differences were statistically significant. The VOT differences between APi and IPi tokens vs. the lower positions were greater for Seoul than for Busan and

Ulsan. This is presumably driven by the high VOT values of APi and IPi tokens from the Seoul group. The difference between PWi and P tokens were also greater for Seoul and Busan than for Ulsan. For the higher positions, all pairwise contrasts across APi – P positions were significantly different between regions except APi – IPi between Seoul and Ulsan, and IPi – P between Busan and Ulsan.

For the nasal stops, only P tokens were significantly different from the lower positions. For Seoul, the VOT was significantly longer in P than PWi position ($p = 0.01$), but only by a very small difference of 0.79 ms (PWi $<$ P). For Busan, the P tokens showed significantly longer VOT than the tokens in all the other positions by 2 – 3 ms (Mm $=$ Mi $=$ PWi $=$ APi $=$ IPi $<$ P). Similarly, for Ulsan, VOT was significantly longer for the P tokens than for the tokens in all the lower positions by approximately 1 ms (Mm $=$ Mi $=$ PWi $=$ APi $=$ IPi $<$ P).

These differences across region were corroborated by the results of the interaction contrasts. The differences of VOT between the P tokens and the tokens in all the other positions were significantly greater for Busan than for Seoul and Ulsan by 1 – 2 ms. Between Seoul and Ulsan, the difference between the APi and P tokens was greater for Ulsan than for Seoul.

Finally, the combined model indicated that VOT was significantly greater for the lenis stops than for the nasal stops, except in Mm and Mi positions.

5.5.4.5 Total voiceless interval

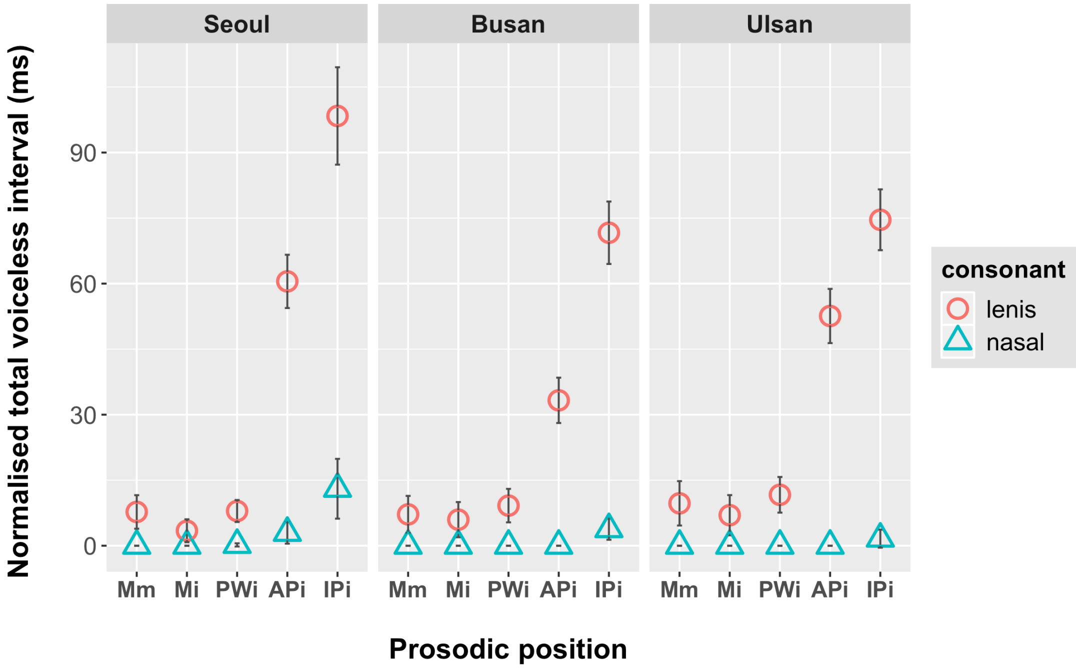


Figure 5.22: Normalised total voiceless interval (ms) by position. Error bars show 95% CIs.

As shown in Figure 5.22, the prosodic variation of normalised total voiceless interval followed the expected trend between PWi and IPi positions, in which tokens in a higher position showed a progressively longer voiceless interval than tokens in a lower position. For Seoul, the trend is visible not only for the lenis tokens but also for the nasal tokens, albeit much weaker. Among the three varieties, the cumulative trend appears to be strongest for Seoul.

For the lenis tokens, the cumulative trend across PWi, APi and IPi positions was statistically supported for all three groups. While tokens in Mm, Mi and PWi positions were not significantly different among themselves, APi tokens showed a greater voiceless interval than the tokens in all the lower positions. In turn, IPi tokens showed a greater voiceless interval than those in the lower positions (Mm = Mi = PWi < APi < IPi).

In terms of the regional differences, the Seoul group showed greater prosodic variation than the other two groups. The differences between APi position vs. the lower positions and the differences between IPi vs. the lower positions were greater for Seoul than for Busan and Ulsan, with the following exceptions. Between Busan and Ulsan, the contrast between APi vs. the lower positions was greater for Ulsan than for Busan. On the other hand, the difference between APi and IPi tokens was greater for Busan than for Ulsan.

For the nasal stops, too, Seoul showed the most robust effects of prosodic position. IPi tokens showed a significantly longer voiceless interval than the tokens in all the other positions (Mm = Mi = PWi = APi < IPi). For Busan, IPi tokens showed a longer voiceless interval only compared to the tokens in the APi position (APi < IPi). For Ulsan, none of the comparisons between two positions showed a significant difference (Mm = Mi = PWi = APi = IPi).

The regional differences were statistically significant between Seoul and the other two groups. The differences between IPi and all the other positions were significantly greater for Seoul than for Busan and Ulsan. The differences between Busan and Ulsan were not statistically significant.

Lastly, the comparison between the lenis and the nasal stops based on the combined model indicated that the two consonants were significantly different within each level of position. While the estimated differences were relatively small in Mm (8.97 ms), Mi (5.96 ms) and PWi (9.16 ms) positions, they were much greater in APi (47.6 ms) and IPi (74.8 ms) positions.

5.5.4.6 Prevoicing

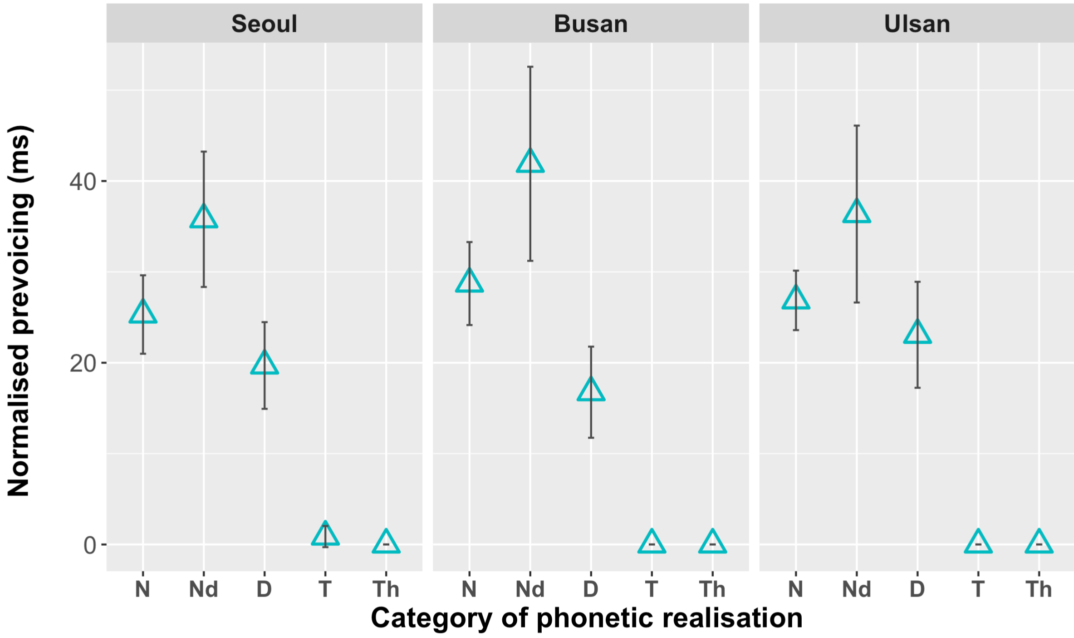


Figure 5.23: Normalised prevoicing (ms) by category for the nasal tokens in a postpausal position. Error bars show 95% CIs.

Figure 5.23 shows that the pattern of prevoicing was similar across the three varieties for the nasal tokens in a postpausal position. Note that "prevoicing" was in fact the duration of nasal voicing for the tokens realised with nasality. Unsurprisingly, the mean prevoicing for the tokens in Categories (4) T and (5) T^H was close to zero. For the tokens in Categories (1) N – (3) N^D, Category (2) N^D showed the highest value. Statistically, most of the pairwise differences across category were significant. The exceptions were the difference between T and T^H in all three groups and the difference between N and N^D for Seoul and Ulsan. Additionally for Ulsan, the difference between N and D and that between N^D and D were also not significant (Seoul: $N = N^D > D > T = T^H$; Busan: $N^D > N > D > T = T^H$; Ulsan: $N = N^D = D > T = T^H$). The interaction contrasts showed that none of the pairwise comparisons between regions was significant except for the contrast of N^D vs. D between Busan and Ulsan. That is, prevoicing was shorter for D than for N^D to a greater extent for Busan than for Ulsan.

5.5.4.7 Nasal attenuation

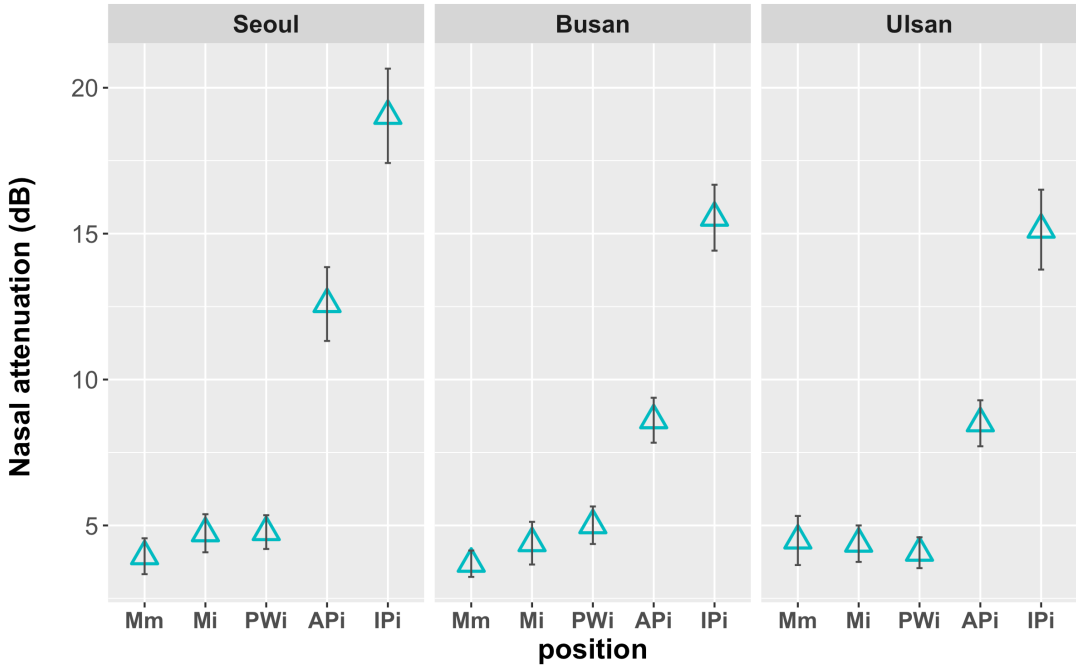


Figure 5.24: Speaker normalised nasal attenuation (dB) by position for the nasal tokens in Categories (1) - (3). Error bars show 95% CIs.

Figure 5.24 shows the results for speaker normalised nasal attenuation of the *voiced* tokens of the nasal phonemes (1534 out of 1566 tokens). Recall that speaker normalised nasal attenuation was obtained by subtracting the minimum intensity of a given token from the highest minimum intensity reading by the same speaker. As the highest minimum intensity was found in a strongly voiced, sonorant nasal, the measure reflected the degree of denasalisation and devoicing. By excluding explicitly voiceless realisations in Categories (4) & (5), I attempted to isolate the effect of denasalisation as much as possible. Overall, a cumulative trend was observed across PWi – IPi positions in all groups, in which progressively greater nasal attenuation was found in a higher position. While the values across region are similar for the three lower positions, Seoul showed the highest nasal attenuation in APi and IPi positions. The results of Tukey tests corroborated the cumulative pattern seen in Figure 5.24. All three groups made a consistent three-way distinction across the positions: Mm = Mi = PWi < APi < IPi.

There were significant regional differences in the degree of the prosodic effect. The differences in the nasal attenuation between APi position and the lower positions were significantly greater for Seoul than for Busan and Ulsan. In addition, the differences between IPi position and Mi, Mm and PWi positions were also greater for Seoul compared to the two other groups. Busan and Ulsan did not show any significant difference in the nasal attenuation measure. This indicates that the APi and the IPi nasal tokens showed a greater degree of denasalisation compared to the tokens in the lower positions for Seoul than for Busan and Ulsan.

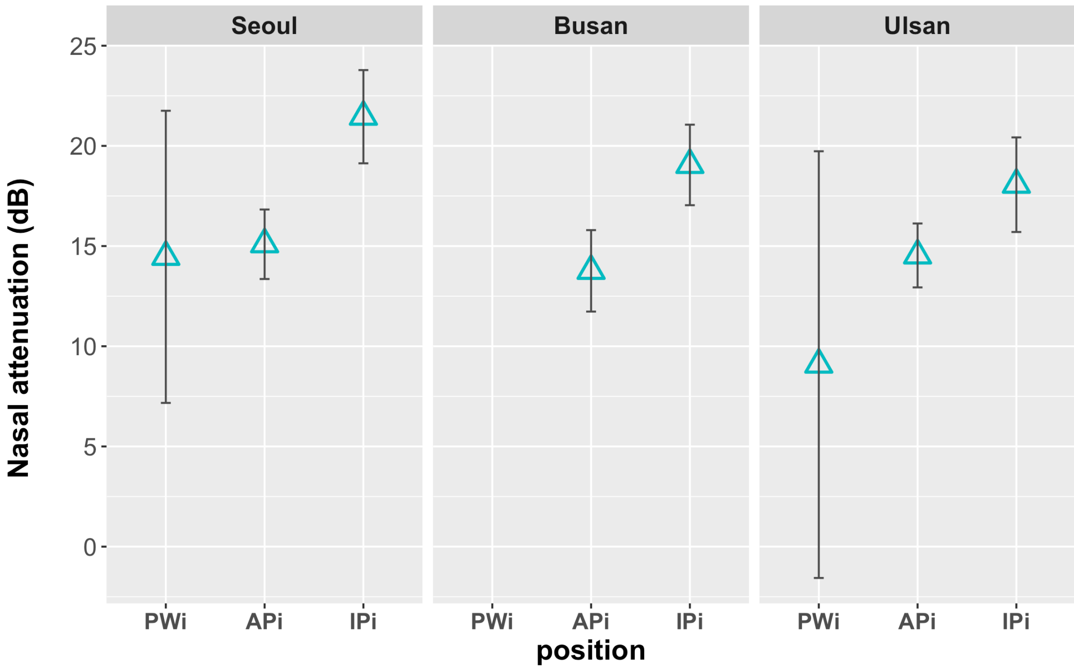


Figure 5.25: Nasal attenuation (dB) by position for the nasal tokens in Categories (3) - (5). Error bars show 95% CIs.

Figure 5.25 shows nasal attenuation for the completely denasalised tokens (Categories 3-5) of the nasal phonemes (301 out of 1566 tokens). Since no complete denasalisation was found in lower positions, they were omitted from the plot. The plot illustrates the point that nasal attenuation is sensitive to voicing. We know from Section 5.5.4.3 on % voicing that the nasal stops tended to show less voicing in a higher position. Thus, nasal attenuation in Figure 5.25 appears to reflect this variation on voicing rather than nasality. This justifies the decision to exclude Categories (4) & (5) from the analyses for nasal attenuation.

5.5.4.8 Gender

In all the statistical models in this section, *gender* showed a robust effect in a majority of the models. Of the 12 models, I will only summarise the results for the seven separate models for the lenis and nasal tokens (the full results are available in Appendix D).

The model estimates in Table 5.10 are for the male group with reference to the female group. The table indicates that the males showed shorter closure duration than the females for the lenis and the nasal tokens by 6.56 ms and 8.84 ms, respectively. For the lenis tokens, the males also produced greater residual voicing and % voicing during closure. For the nasal stops, however, the effect of gender was significant for % voicing during closure, but not for residual voicing. This suggests that, for the nasals, the significant effect of gender for % voicing is mainly driven by the differences in closure duration rather than the differences in the residual voicing. Next, VOT and total voiceless interval were shorter in male speech for both kinds of stops, although the magnitude of the effect was much larger for the lenis stops[39].

The models for prevoicing and nasal attenuation were run only for the nasal tokens. While prevoicing did not vary significantly by gender, nasal attenuation did. Since higher nasal attenuation indicates greater denasalisation of nasal tokens, the negative coefficient indicates that the males produced the nasal tokens with 1.87 dB less attenuation. Overall, the results in Table 5.10 suggest a pattern in which the males produce the lenis and the nasal stops with greater closure voicing, shorter VOT, and greater nasality and the females with less closure voicing, longer VOT, and less nasality.

Table 5.10: The estimated *B* coefficients and standard errors (in brackets) for *gender* from seven regression models. The reference level was *female* and the asterisks indicate the levels of statistical significance ($^{*}p < 0.05$; $^{**}p < 0.01$; $^{***}p < 0.001$).

	Lenis	**Nasal**
Normalised closure (ms)	-6.56^{*} (3.28)	-8.84^{**} (3.26)
Normalised residual voicing (ms)	7.60^{***} (2.01)	-2.06 (2.33)
% Voicing during closure	16.17^{***} (4.01)	2.23^{**} (0.86)
Normalised VOT (ms)	-5.02^{*} (2.34)	-0.44^{**} (0.16)
Normalised total voiceless interval (ms)	-14.78^{**} (4.68)	-3.18^{*} (1.26)
Normalised prevoicing (ms)		1.08 (3.28)
Normalised nasal attenuation (dB)		-1.87^{**} (0.63)

[39] This is in contrast to the finding in E. Oh (2011) that there was no gender effect on VOT for the lenis stops.

Recall that the interaction terms for *gender*, *region,* and *position* were not included in the models because their interaction was not the main focus of the study. Thus, we will only make visual observations for the key outcome variables *total voiceless interval* (ms) and *nasal attenuation* (dB). Figure 5.26 shows *total voiceless interval* for the lenis tokens. For Busan and Ulsan, females consistently showed a higher mean for total voiceless interval than males across all prosodic positions. On the other hand, the gender effect is only clear in the APi and IPi positions for Seoul.

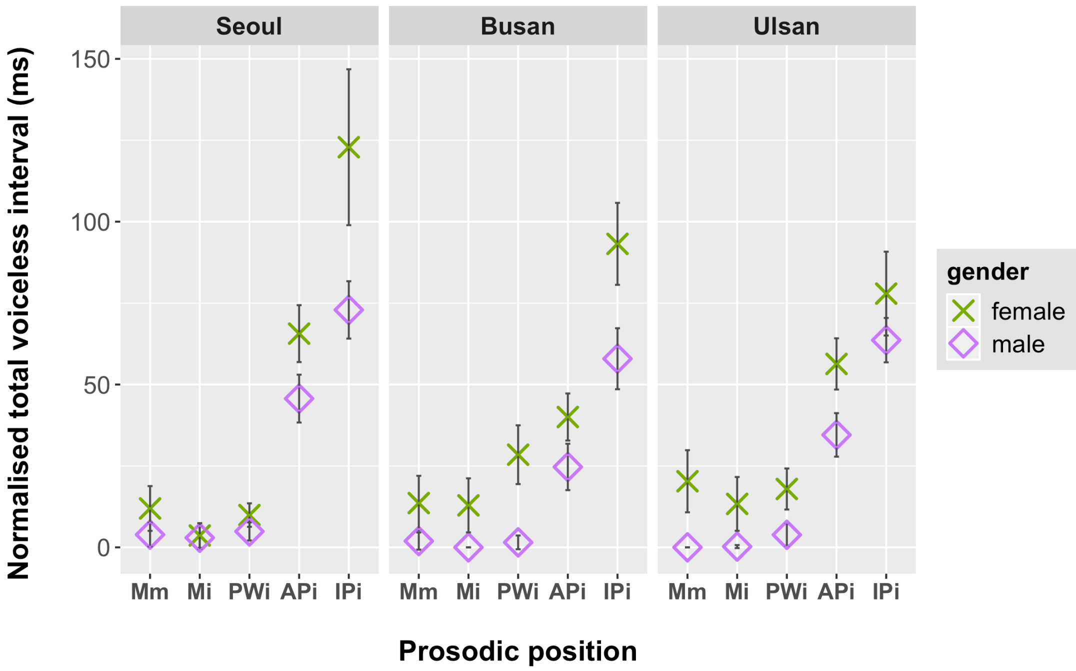


Figure 5.26: The effect of gender on normalised total voiceless interval (ms) for the lenis tokens. Error bars show 95% CIs.

Figure 5.27 shows *total voiceless interval* for the nasal tokens. Devoicing (and a possible voicing lag) for the nasal tokens is only clear for the female group. It is also only found in IPi position and, additionally for Seoul, in APi position. The gender effect appears to be stronger in the order of Seoul, Busan, and Ulsan.

Finally, Figure 5.28 shows the gender effect for nasal attenuation for the nasal tokens in Categories (1) – (3). For Seoul and Ulsan, the females showed higher mean nasal attenuation than the males across all positions. The difference by gender was greatest in IPi position. On the other hand, Busan did not show such a clear gender effect.

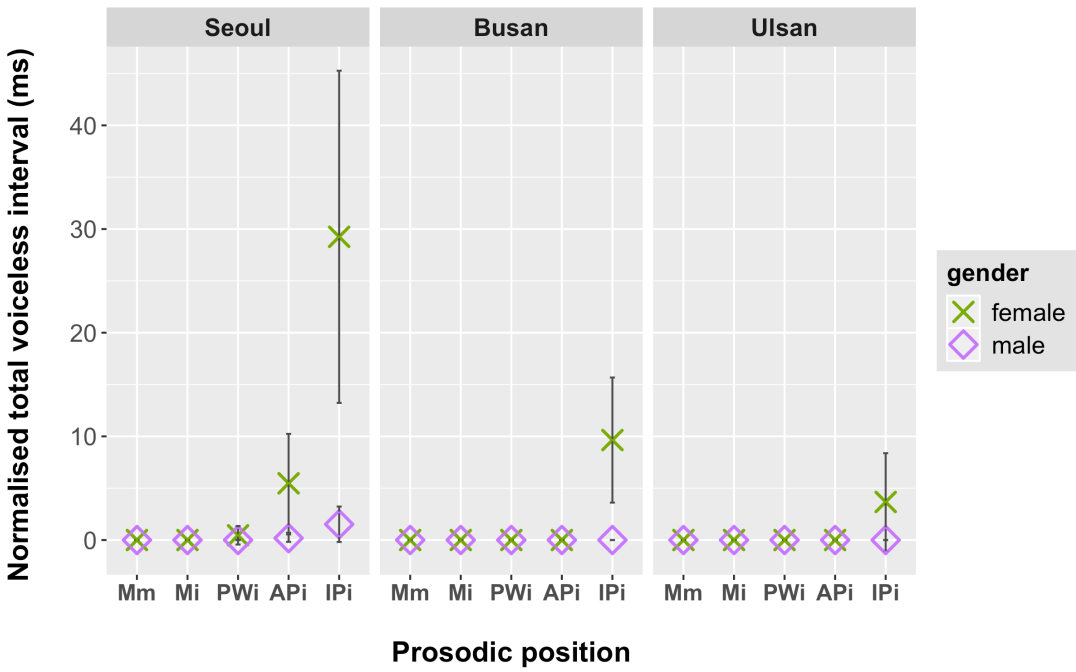


Figure 5.27: The effect of gender on normalised total voiceless interval (ms) for the nasal tokens. Error bars show 95% CIs.

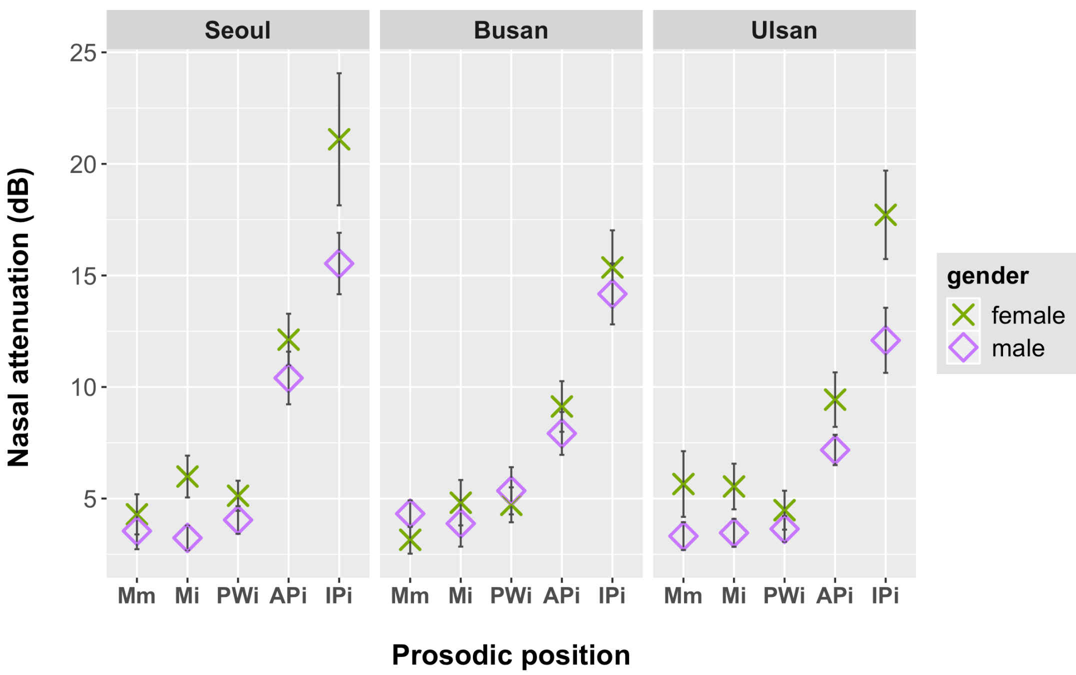


Figure 5.28: The effect of gender on nasal attenuation (dB) for the nasal tokens in Categories (1) – (3). Error bars show 95% CIs.

5.5.4.9 Repetition

Recall from Section 5.5.1 that the lenis tokens were more likely to be realised with devoicing (Categories 3 – 4) in the first repetition than in the second repetition. In most of the models for the acoustic measurements, the effect of repetition was not significant (see Appendix D). This seems to be because the outcome variables in this section were normalised for the effect of speech rate, whereas the category results did not take speech rate into account. Two additional models were run for the outcome *raw closure duration* and *raw residual voicing* to test this speculation. Although normalised closure duration did not vary by repetition, raw closure duration was significantly shorter in the second repetition by 4.19 ms and 3.31 ms for both types of consonants. On the other hand, normalised residual voicing was significantly longer in the second repetition for the lenis stops by 1.98 ms but not for the nasal stops. In contrast, raw residual voicing was significantly shorter by 1.82 ms for the nasals but not for the lenis stops[40]. These results can be interpreted in the following way. For both consonants, closure duration becomes shorter in the second repetition because, expectedly, speech is faster in the second repetition. For the nasal stops which are normally fully voiced, residual voicing also becomes shorter in the second repetition. For the lenis stops, however, residual voicing is not significantly affected by speech rate, remaining relatively constant between the two repetitions. Therefore, normalising residual voicing for the lenis stops had misleadingly produced a significant effect of repetition. By extension, we can also ignore the significant effect of repetition for the lenis stops for the models with the outcome % voicing. Indeed, a model for the outcome % voicing, computed using normalised closure but raw residual voicing, showed no effect of repetition for the lenis stops. For the rest of the models, for *total voiceless interval*, *VOT*, *prevoicing,* and *nasal attenuation*, the effect of *repetition* was not significant.

[40] Note that less than 2ms is likely to be below the threshold for what is audible.

5.5.5 Summary and discussion II

Here, I will summarise results from Section 5.5.4 and evaluate the hypotheses in **H1** - **H3**. First, I will consider **H1**.

H1. Both the lenis and nasal stops show the following cumulative effects of DIS as their position moves up the prosodic hierarchy:

(a) Longer acoustic closure duration
(b) Smaller percentage of voicing during acoustic closure (residual voicing itself does not show DIS effects)
(c) Longer Voice Onset Time (VOT)
(d) Longer total voiceless interval (voiceless portion of the closure + VOT)
(e) For nasal consonants only, lower minimum nasal energy

The results for normalised closure duration can be summarised in the following way.

(i) Lenis
- All: Mm = Mi = PWi < APi < IPi

(ii) Nasal
- Seoul: Mm = Mi = PWi = APi < IPi
- Busan: Mm = Mi = PWi < IPi; PWi < APi < IPi
- Ulsan: Mm = Mi = PWi < IPi; Mm = PWi < APi < IPi

These results support **H1** (a), although the cumulative pattern is more consistent for the lenis stops.

For normalised residual voicing, the following results were found.

(i) Lenis
- Seoul: Mm = Mi = PWi, PWi > APi, APi > IPi
- Busan: Mm = Mi = PWi = APi = IPi
- Ulsan: Mm = Mi = PWi > APi = IPi

(ii) Nasal
- All: Mm = Mi = PWi = APi < IPi

Given the impact of normalisation on the effect of repetition (Section 5.5.4.9), the models were run again with raw residual voicing as the outcome but the hierarchies remained unchanged (see Appendix E). These results go against the prediction in **H1** (b) that residual voicing will not show effects of DIS, except for Busan for the lenis stops. The results suggest that residual voicing becomes shorter in a higher position for the lenis stops and it becomes longer in a higher position for the nasal stops. The difference between the consonants is presumably due to the fact that nasal consonants are normally voiced throughout the closure, which is longer in a higher position.

Next, the results for % voicing were as follows.

(i) Lenis

- Seoul & Ulsan: Mm = Mi = PWi > APi > IPi
- Busan: Mm = Mi = PWi = APi > IPi

(ii) Nasal

- Seoul & Busan: Mm = Mi = PWi = APi > IPi
- Ulsan: Mm = Mi = PWi = APi = IPi

These results are generally consistent with the prediction in **H1** (b) that % voicing is smaller in a higher position, although more positions are distinguished this way for the lenis than for the nasal stops. However, for Ulsan nasals, there is no effect of DIS on % voicing.

The results for normalised VOT can be summarised as below.

(i) Lenis

- Seoul: Mm = Mi = PWi < P < APi = IPi
- Busan: Mm = Mi = PWi < APi, IPi, P; APi < P
- Ulsan: Mm = Mi = PWi < APi = IPi = P

(ii) Nasal

- Seoul: PW < P
- Busan & Ulsan: Mm = Mi = PWi = APi = IPi < P

The hypothesis in **H1** (c) predicts that VOT will be longer in a higher position. While all three varieties showed at least a two-way contrast in the expected direction, there was a violation of a cumulative trend for Seoul. For the lenis stops, APi and IPi tokens showed longer VOT than P tokens. Thus, there is only a partial support for this hypothesis.

The following results for normalised total voiceless interval were relevant for **H1** (d):

(i) Lenis

- All: Mm = Mi = PWi < APi < IPi

(ii) Nasal

- Seoul: Mm = Mi = PWi = APi < IPi
- Busan: APi < IPi
- Ulsan: Mm = Mi = PWi = APi = IPi

The results for the lenis stops provide support for **H1** (d) that total voiceless interval is progressively greater in a higher position. For the nasal stops, the results were mixed. Seoul and Busan showed a two-way contrast in the expected direction but Ulsan did not distinguish any position in terms of total voiceless interval.

Due to the lack of previous research, no prediction was made regarding the pattern of (normalised) prevoicing for the nasal tokens in a P position. Note that "prevoicing" was in fact the duration of nasal voicing for the tokens realised with nasality. Thus, unsurprisingly, the voiceless categories showed a value close to zero. Among the three voiced categories, the results generally showed that mean prevoicing was longest for tokens in Category (2) N^D and shortest for tokens in Category (3) D. The statistical comparisons showed the following results.

- Seoul: N = N^D > D > T = T^H
- Busan: N^D > N > D > T = T^H
- Ulsan: N = N^D = D > T = T^H

Given that closure can only be identified at the onset of voicing in a postpausal position, prevoicing *is* the acoustic closure duration for these tokens. Thus, it is not surprising that the phonetically more complex, N^D type realisation shows longest prevoicing, as both nasal and denasalised portions must be voiced to be audible. It is therefore possible that tokens in Category (3) D in fact had a voiceless nasal portion which did not show up in the spectrogram. In addition, the genuine prevoicing for D is likely to be shorter than the duration of nasal voicing for N because it is harder to sustain voicing during a non-nasal than during a nasal due to increased intraoral air pressure.

The last acoustic measurement was nasal attenuation and the results were consistent across the three varieties.

All: Mm = Mi = PWi < APi < IPi

These results can be expressed in terms of raw minimum intensity by reversing the signs: Mm = Mi = PWi > APi > IPi. This provides support for the prediction in **H1** (e) that progressively lower nasal energy will be found in a higher position.

Next, I will evaluate **H2** regarding regional variation for the lenis stops.

H2. For the lenis stops, there is no regional variation.

For the lenis stops, there were many instances of a significant regional difference for all of the relevant acoustic measurements, in contrast to the prediction in **H2**. The results are summarised below. To illustrate how the notation works, the results in (d)(i) should be interpreted as follows. The differences between IPi and the lower positions of Mm, Mi, and PWi were significantly greater for Seoul compared to Busan and Ulsan. The differences between APi and the lower positions of Mm, Mi and PWi were also significantly greater for Seoul than for Busan and Ulsan. Comparing IPi and P, Seoul showed a greater difference than the other two in the direction of longer VOT in IPi position than in P position.

(a) Normalised closure:

(i) Seoul, Busan > Ulsan:
- PWi, APi vs. IPi

(ii) Seoul > Ulsan:
- Mi vs. IPi

(b) Normalised residual voicing:
   (i) Seoul, Ulsan > Busan:
      - PWi vs. APi, IPi
   (ii) Seoul > Busan:
      - Mm vs. APi
      - Mm, Mi vs. IPi

(c) % Voicing:
   (i) Seoul > Busan:
      - Mm, Mi, PWi vs. APi, IPi
   (ii) Seoul, Busan > Ulsan:
      - APi vs. IPi
   (iii) Ulsan > Busan:
      - Mm, Mi, PWi vs. APi
      - PWi vs. IPi

(d) Normalised VOT:
   (i) Seoul > Busan, Ulsan:
      - Mm, Mi, PWi vs. APi, IPi
      - IPi vs. P (in the direction of IPi > P)
   (ii) Seoul, Busan > Ulsan:
      - PWi vs. P
   (iii) Busan > Seoul, Ulsan:
      - APi vs. IPi
   (iv) Busan > Ulsan:
      - APi vs. P

Additionally for APi vs. P, Seoul is significantly different from both Busan and Ulsan because the direction of the effect is reverse, i.e. APi > P.

(e) Normalised total voiceless interval:
   (i) Seoul > Busan, Ulsan:
      - Mm, Mi, PWi vs. APi
      - Mm, Mi, PWi, APi vs. IPi
   (ii) Ulsan > Busan:
      - Mm, Mi, PWi vs APi
   (iii) Busan > Ulsan:
      - APi vs. IPi

One broad generalisation can be drawn from these results. While the differences between Busan and Ulsan were inconsistent, Seoul always showed stronger effects of DIS than Busan and Ulsan, for all outcomes except VOT. The patterns of VOT were relatively complicated in terms of the position by region interaction. For VOT, Seoul showed a different pattern from the other groups in which P tokens showed intermediate values higher than Mm - PWi tokens but lower than APi and IPi tokens. Note how this is similar to the pattern from the category analysis for the nasal tokens (Section 5.5.2). Seoul was the only group to show a violation of the expected cumulative trend; P tokens were intermediate between Mm - PWi tokens and IPi tokens in terms of the predicted probability of denasalisation.

Independent from **H2**, I have made a speculative prediction regarding VOT variation in the three varieties in Section 3.4.1. For the distinction of the three contrastive plosives in Korean, H. Lee and Jongman (2012) found that F0 was a more reliable cue for Seoul, whereas VOT was a more reliable cue for South Gyeongsang. As this implies that the distinction between the lenis stops and the aspirated stops is still mainly maintained in terms of VOT rather than F0 in Gyeongsang Korean, I predicted that Busan and Ulsan speakers will be more reluctant to vary VOT of the lenis stops under DIS compared to Seoul speakers. The results summarised above are generally in line with this prediction, with the exception of the unexpected pattern at P position. This pattern will be discussed further in Section 5.6.5.

Finally, I will summarise the results for the nasal stops which are relevant for **H3**.

> **H3**. For the nasal stops, the DIS effects are greater in the order of Seoul, Busan and Ulsan, following the model of hierarchical diffusion (Trudgill, 1974).

(a) Normalised closure:
- (i) Seoul > Busan, Ulsan:
  - APi vs. IPi

(b) Normalised residual voicing: no interaction

(c) % Voicing:
- (i) Seoul > Busan, Ulsan:
  - Mm, Mi, PWi vs. IPi
- (ii) Seoul > Ulsan:
  - APi vs. IPi

(d) Normalised VOT:
- (i) Busan > Seoul, Ulsan:
  - Mm, Mi, PWi, APi, IPi vs. P
- (ii) Ulsan > Seoul:
  - APi vs. P

(e) Normalised total voiceless interval:
- (i) Seoul > Busan, Ulsan:
  - Mm, Mi, PWi, APi vs. IPi

(f) Normalised nasal attenuation:
- (i) Seoul > Busan, Ulsan:

- Mm, Mi, PWi vs. APi, IPi

Apart from the patterns of VOT, DIS effects were indeed stronger for Seoul compared to Busan and Ulsan as predicted in **H3**. However, the results provided no evidence for the prediction that the effects will be stronger for Busan than for Ulsan. For VOT, the DIS effects were stronger for Busan than for Seoul and Ulsan. This was driven by the fact that Busan showed the longest VOT in P position for the nasal stops. Overall, these results suggest that the effects of DIS for the nasal stops is strongest for Seoul on the parameters of closure voicing and denasalisation, but strongest for Busan on the parameter of voicing lag.

5.5.6 The life cycle of phonological processes: the status of lenis stop (de-)voicing

In this section, I will report the results of the individual models with the outcome *% voicing during closure* for the lenis stop tokens. Table 5.11 shows the results of Hartigan's dip test (in the middle two columns) and a summary of the post-hoc analyses for the linear regression models (in the rightmost column). The full results of the post-hoc analyses are available in Appendix G[41]. The dip test results in Table 5.11 indicate that the distribution of % voicing during closure of the lenis stop tokens was significantly non-unimodal for 16 speakers at $p < 0.01$ and for one additional speaker at $p < 0.1$. This suggests that, at least for these speakers, there is support for categoricity of the pattern of lenis stop (de-)voicing. While the results of the Tukey tests were different for each speaker, a common pattern was that none of the speakers showed any significant difference among the three lower positions which are AP-medial (APm), except for Busan_5. On the other hand, most speakers (23 out of 30) significantly distinguished at least two levels between an APm position and a higher position, such that % voicing was significantly lower for tokens in a higher position. Two speakers made a consistent three-way distinction among APm, APi and IPi positions, with progressively lower % voicing in that order. There was no significant difference between two positions against the expected cumulative trend of lower % voicing in a higher position. In terms of regional differences, it is interesting that none of the Busan speakers significantly distinguished between APi position and a lower position. This contrasts with the fact that six Seoul speakers and six Ulsan speakers showed a significant difference between APi position and a lower position.

Table 5.11: The results of Hartigan's dip tests and Tukey pairwise comparisons for % voicing during closure (*$p<0.1$; **$p<0.05$; ***$p<0.01$). For Tukey pairwise comparisons, the sign $>$ was used between levels that are significantly different at $p < 0.05$; the sign $=>$ was used for two levels that are different with a p-value between 0.05 and 0.1[42].

Subject	D-statistic	*p*-value	Pairwise comparison
Seoul_1	0.070	0.146	Mi = PWi > IPi, APi => IPi ($p = 0.065$)
Seoul_2	0.055	0.707	Mm = Mi = PWi > IPi, Mi => APi ($p = 0.058$), PWi > APi
Seoul_3	0.035	0.993	Mm = Mi = PWi = APi > IPi

[41] The regression estimates will not be reported as the Tukey results are more comprehensive. While regression results provide the comparison for each of the prosodic position against one fixed reference position, a Tukey test shows all pairwise comparisons between two prosodic levels.

[42] The additional use of 0.1 as a threshold is intended to make more information available to the reader by reporting weaker statistical trends as well as more robust ones. This is especially because the number of tokens were low in general, as mentioned throughout the section. For more precise understanding of the results, exact p-values are reported in Appendix G.

Seoul_4	0.099	0.006 ***	Mm = Mi = PWi = APi = IPi (but only 1 IPi token)
Seoul_5	0.120	< 0.001 ***	Mm = Mi = PWi > APi > IPi (but Mm = APi, p = 0.448)
Seoul_6	0.123	< 0.001 ***	Mm = Mi = PWi > APi = IPi
Seoul_7	0.092	0.008 ***	Mi = PWi > IPi, PWi > APi, Mm => IPi (p = 0.070)
Seoul_8	0.079	0.052 *	Mm = Mi = PWi = APi > IPi
Seoul_9	0.151	< 0.001 ***	Mm = Mi = PWi > APi > IPi
Seoul_10	0.064	0.129	Mi > APi = IPi, PWi > IPi
Busan_1	0.065	0.290	Mm = Mi = PWi = APi = IPi
Busan_2	0.060	0.615	Mm = Mi = PWi = APi = IPi
Busan_3	0.050	0.756	Mm => IPi (p = 0.0866), PWi => IPi (p = 0.077)
Busan_4	0.138	< 0.001 ***	Mi => IPi (p = 0.080)
Busan_5	0.136	< 0.001 ***	Mm = Mi > PWi = IPi
Busan_6	0.148	< 0.001 ***	Mm = Mi = PWi = APi = IPi
Busan_7	0.128	< 0.001 ***	Mm = PWi = APi > IPi (zero Mi tokens)
Busan_8	0.084	0.240	Mm = Mi = PWi = APi > IPi
Busan_9	0.109	< 0.001 ***	Mm = APi > IPi
Busan_10	0.075	0.194	Mm = Mi = APi = IPi (zero PWi tokens)
Ulsan_1	0.044	0.908	Mm = Mi = PWi = APi = IPi
Ulsan_2	0.107	< 0.001 ***	PWi => IPi (p = 0.051)
Ulsan_3	0.113	< 0.001 ***	Mm = Mi = PWi = APi > IPi (but Mm => APi, p = 0.063)
Ulsan_4	0.067	0.361	Mm = PWi = APi = IPi (zero Mi tokens)
Ulsan_5	0.068	0.195	Mi = PWi > IPi, PWi > APi, Mm => IPi (p = 0.054)
Ulsan_6	0.108	< 0.001 ***	Mm = Mi = PWi > APi > IPi
Ulsan_7	0.106	< 0.001 ***	Mm = Mi = PWi > APi = IPi (but Mm => APi, p = 0.072; Mi => APi, p = 0.084)
Ulsan_8	0.062	0.446	Mm = Mi = PWi > IPi
Ulsan_9	0.136	< 0.001 ***	Mm = Mi = PWi > IPi, PWi > APi, Mi => APi (p = 0.067)
Ulsan_10	0.103	0.004 ***	Mm => PWi (p = 0.063), PWi => APi (p = 0.096), PWi => IPi (p = 0.051)

The probability density plots for each speaker were produced to allow a more detailed examination of the patterns by prosodic position. The five positions were collapsed into three, by combining Mm, Mi and PWi tokens into one group, APm. The APi and IPi groups were kept unchanged. This decision is based on (a) the sparsity of tokens, and (b) the results of the Tukey test that no speaker showed a significant difference ($p < 0.05$) among the three lower positions, with the exception of Busan_5.

Due to the limitation of space, I will focus the discussion on a few speakers (see Appendix F for all probability density plots). To select these speakers, I first grouped the speakers by the kind of distribution pattern exhibited by their tokens (e.g. bimodal, unimodal, intermediate between the two) and selected a representative speaker from each group. We will first examine speakers with relatively straightforward patterns. The density plot for Seoul_9 in Figure 5.29 provides a clear example of a bimodal distribution, as expected from the result of the dip test. While APm tokens are clustered around 100% voicing, APi and IPi tokens together form a lower mode around 20%. The distribution of IPi tokens is narrower than that of the APi and the APm tokens which span a larger range of % voicing. The pattern in Figure 5.29 is also consistent with the result of the Tukey test (Mm = Mi = PWi > APi > IPi), as APi tokens take an intermediate value between APm and IPi tokens, although nearer to the IPi tokens.

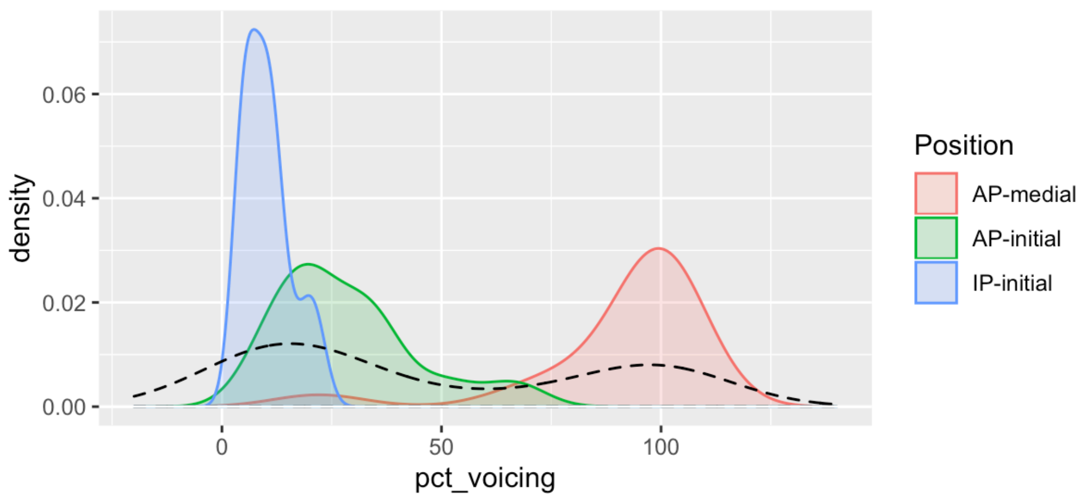


Figure 5.29: Density plot of % voicing for lenis stops produced by Seoul_9. Note that the red tail that extends beyond 100% is an artefact of modelling. There was no token which showed more than 100% voicing, which would not be a possible measurement given the definition of % voicing.

In order to understand the effect of *closure duration* on *% voicing*, correlation lines were fitted separately for the three groups of prosodic positions. Appendix H contains the results from the correlation tests and the scatter plots with fitted correlation lines for each speaker. Of the 90 groups which result from the combination of 3 prosodic positions and 30 subjects, 26 groups showed a significant correlation between duration and % voicing at $p <$

0.05 and an additional eight groups showed significance at $p < 0.1$. The correlation was always negative except for one group, meaning that % voicing became lower as duration became longer. The exception was for the APi tokens from Busan_8, but there were only four tokens and the p-value was relatively high at $p = 0.09875$. There was no striking pattern in which a particular position or region was favoured for a significant correlation. Across the entire data, 10 (9) AP-medial, 14 (9) APi and 10 (8) IPi groups showed a significant correlation, and 11 (9) Seoul, 9 (6) Busan and 14 (11) Ulsan speakers showed a significant correlation at $p < 0.1$ (the numbers in brackets are effects at $p < 0.05$).

Figure 5.30 shows a scatter plot for Seoul_9 with fitted regression lines. Consistent with the density plot above, there was a clustering of APm tokens around 100% voicing and a clustering of AP- and IPi tokens around 20% voicing. The Pearson's r and the p-values on the top right corner of the figure indicate that APi and IPi tokens showed a significant negative correlation between duration and % voicing, whereas AP-medial tokens did not show a significant correlation. Note that Pearson's r values for the APi and IPi tokens were -0.56 and -0.58, respectively, and are surprisingly similar despite the apparent differences in the slope of their regression lines. This is because Pearson's r takes into account standard error, which is greater for the APi than the IPi tokens as indicated by the thickness of the shaded area around the regression line.

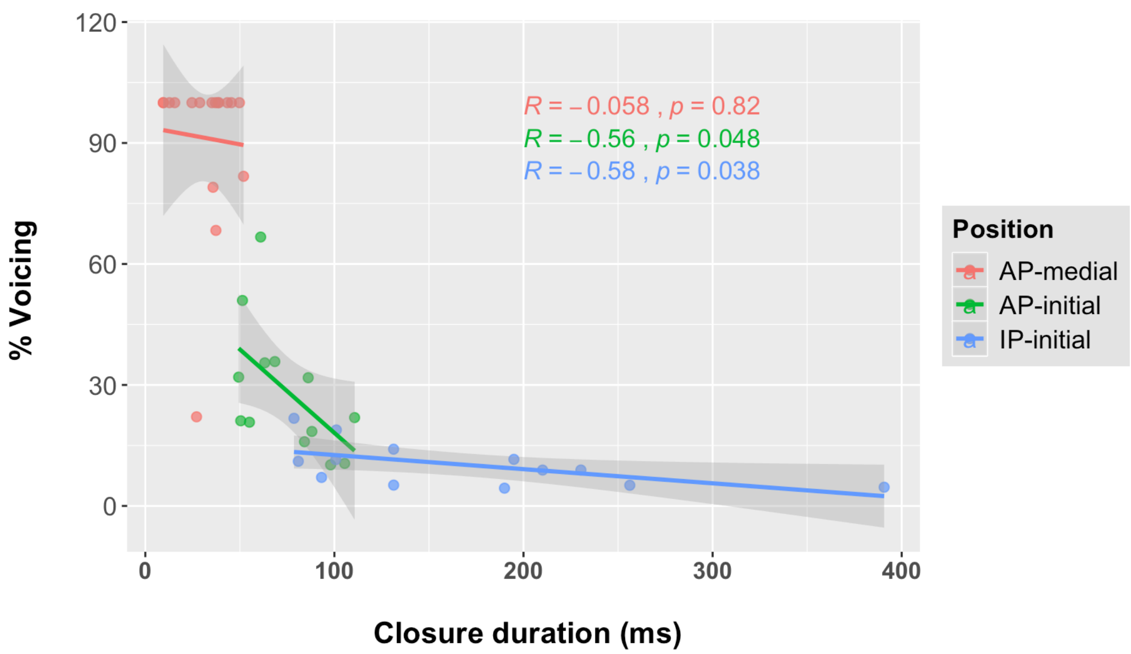


Figure 5.30: % Voicing during closure against closure duration (ms) for lenis stops produced by Seoul_9.

An important pattern revealed by Figure 5.30 is that there was relatively little overlap among the three prosodic positions in terms of closure duration. Most of the AP-medial tokens were shorter than 50 ms, most of the IPi tokens were longer than approximately 90 ms, and the APi tokens took intermediate values. Ignoring the position of the tokens for now, we can observe an s-shaped curve with a sudden decrease in % voicing at around 50 ms, which then plateaus around 90 ms.

Such a pattern is compatible with two rather different explanations. One explanation is that lenis stop (de-)voicing is a discrete phenomenon which categorically targets APm tokens but not APi and IPi tokens, leading to a bimodal distribution as seen in Figure 5.29. The negative correlation found for the APi and IPi tokens is explained by a gradient effect of phonetic voicing, as predicted by rule scattering. The lack of such correlation for the APm tokens could be understood if one assumes that these tokens were not sufficiently long enough to be subject to the gradient voicing.

An alternative explanation is that the pattern of closure voicing is solely accounted for by a curvilinear function of closure duration in which the prosodic categories play no role. This would entail that the apparent effect of prosodic position is simply a side effect of the fact that the range of closure duration for these positions does not show much overlap. However, this explanation is less convincing as it implies that it was just a coincidence that the drastic fall of % voicing at 50 ms happens to divide AP-medial tokens and AP-initial tokens on either side. In addition, without making any reference to the prosodic position, it is difficult to explain why the relationship between duration and % voicing would be curvilinear.

We can further observe the categoricity of the voicing pattern between tokens in APm position and tokens in a higher position in a plot showing residual voicing against closure duration (Figure 5.31). For APm tokens, there was a strong positive correlation between closure duration and residual voicing simply because residual voicing was usually as long as closure duration. However, for APi and IPi tokens, this correlation completely disappeared, as residual voicing stayed relatively constant around 10 - 30 ms for these tokens. Crucially, residual voicing fell drastically from approximately 50 ms to 20 ms near the closure duration of 50 ms. This suggests that there exists another source of variation for the pattern of residual voicing that is independent from closure duration. Perhaps, the explanation may lie in the different levels of vocal fold tension between the lenis stops in APm vs. the other positions. That is, the vocal fold tension could be categorically weak for APm lenis stops, allowing residual voicing to be sustained throughout the stop closure, whereas the tension is stronger for APi and IPi stops, causing residual voicing to end after around 20 ms.

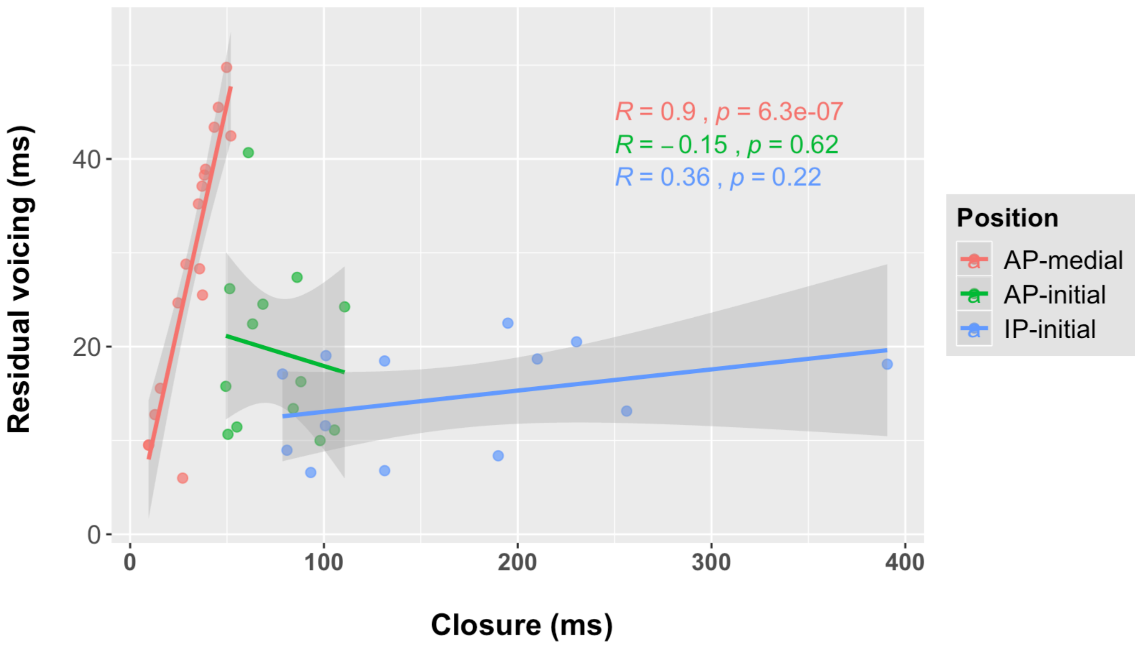


Figure 5.31: Residual voicing (ms) against closure duration (ms) for lenis stops produced by Seoul_9.

However, Seoul_9 is exceptional in that it shows such a clear bimodal distribution of % voicing by prosodic position. For some speakers, the lack of a significant effect from a correlation test and/or a dip test appeared to be due to a low token count (see Appendix H for token counts). Some other individual differences could be accounted for as a result of the distribution of the tokens on closure duration. For example, consider the difference between Seoul_9 and Seoul_3. The density plot for Seoul_3 (Figure 5.32) shows that the APm and APi tokens cluster together at a much higher value than the IPi tokens. On the other hand, the scatter plot for Seoul_3 (Figure 5.33) reveals that, unlike Seoul_9, the duration range of the APi tokens greatly overlapped with that of the APm tokens, whereas the IPi tokens showed a significantly higher range of duration. Given the negative trend for the APi tokens, one can speculate that, had there been more tokens of APi tokens which were longer than 50 ms, the closure duration of the APi tokens may have been closer to that of the IPi tokens. Note also that there were only four tokens for the IPi position, making the model less reliable. Again, it is possible that with more IPi tokens, Seoul_3 might have also shown a significant bimodal distribution, like Seoul_9.

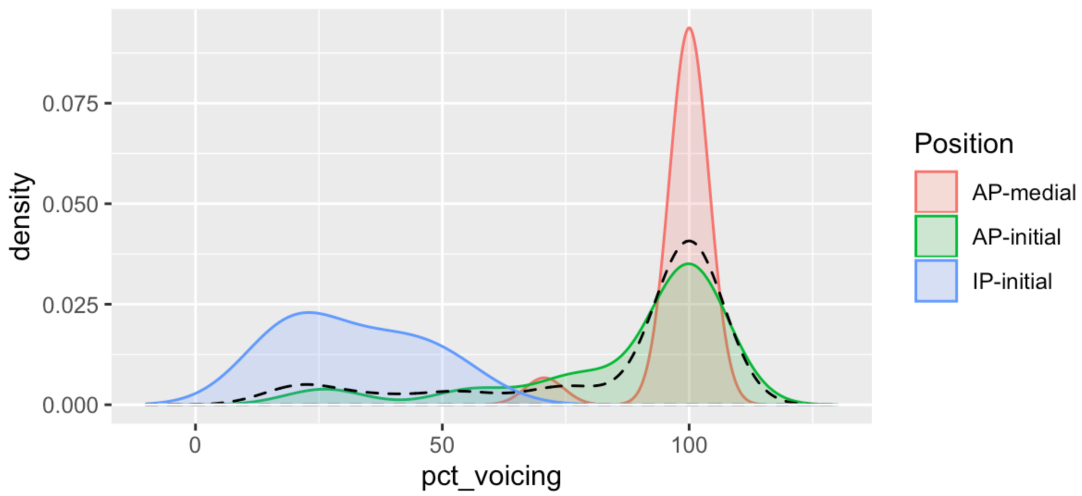


Figure 5.32: Density plot of % voicing for lenis stops produced by Seoul_3.

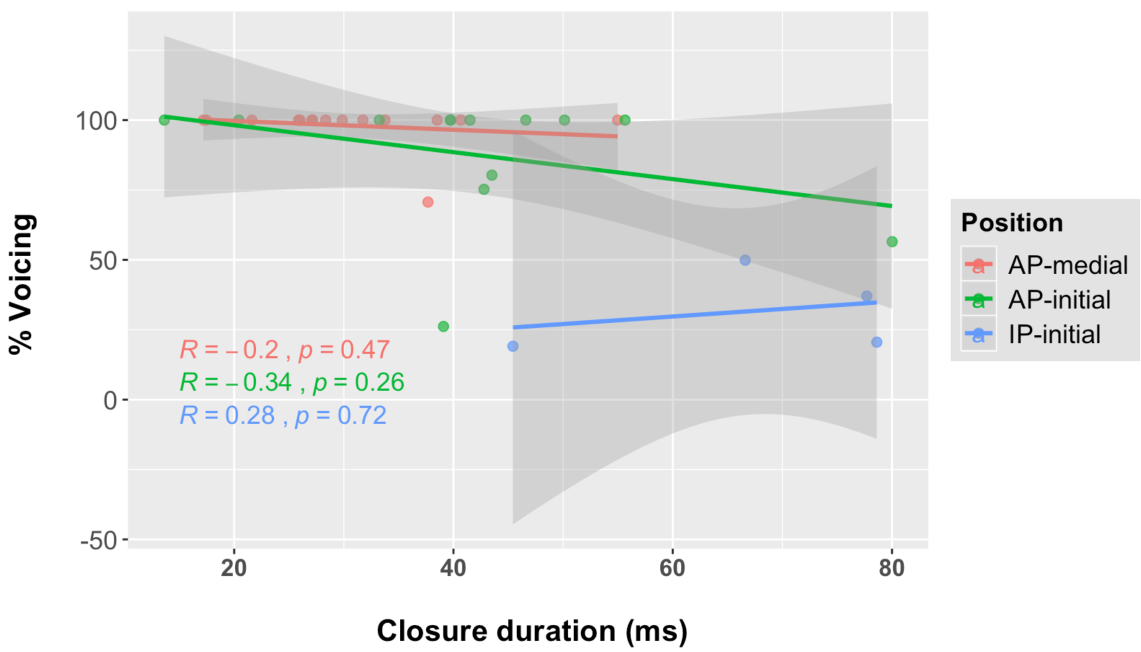


Figure 5.33: % Voicing against closure duration (ms) for lenis stops produced by Seoul_3.

Similarly, Figure 5.34 suggests that the distribution for Seoul_1 was very different from those of Seoul_9 and Seoul_3 discussed above. For this speaker, the APi tokens showed intermediate values between the AP-medial and IPi tokens. However, the apparent differences between Seoul_1 and the other two speakers may be superficial as opposed to stemming from the differences in their underlying system. As seen from Figure 5.35, the three groups of tokens for Seoul_1 showed much more overlap in closure duration compared to Seoul_9. The majority of the tokens from Seoul_1 were shorter than 50 ms regardless of prosodic position. This contrasts with the pattern for Seoul_9 in which nearly all AP-medial tokens were shorter than 50ms whereas almost all tokens in a higher position were longer than 50 ms. In addition, Seoul_1 and Seoul_9 were similar in that the slopes for the APi and IPi tokens were negative (albeit non-significantly for Seoul_1) whereas the slope for the AP-medial tokens was closer to zero. Although it is only speculation at this stage, one can suspect that the patterns of Seoul_9 and Seoul_1 might be more similar than suggested by the significant effects, if there were more tokens from Seoul_1 covering a larger range of durations.

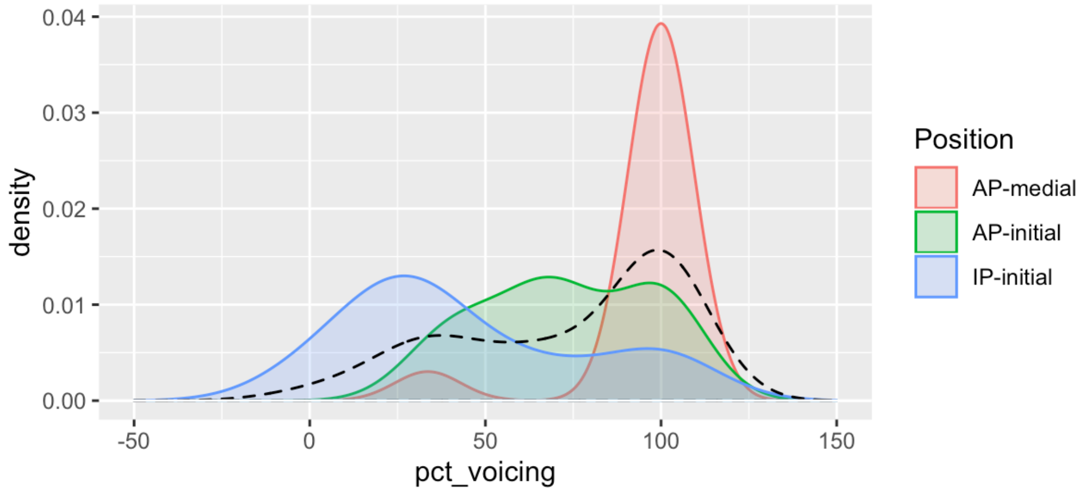


Figure 5.34: Density plot of % voicing for lenis stops produced by Seoul_1.

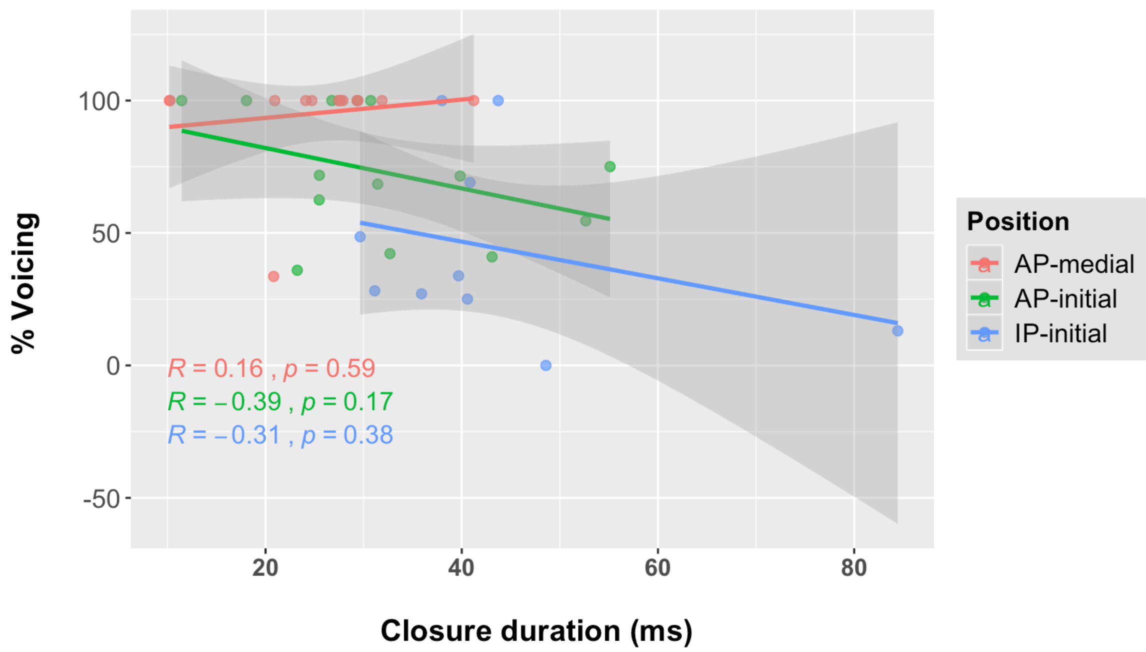


Figure 5.35: % Voicing against closure duration (ms) for lenis stops produced by Seoul_1.

Meanwhile, some other individual differences appeared to be more fundamental. Figure 5.36 shows a density plot for Busan_5. The shape of the distribution looks similar to that of the other speakers discussed above. There were two clear peaks, one at 100% and the other at a much lower value around 30%. A dip test for this speaker indicated that the distribution is significantly non-unimodal. Crucially, however, the bimodality for this speaker is not at all conditioned by prosodic context. All three contexts overlap greatly, showing two separate peaks within each context.

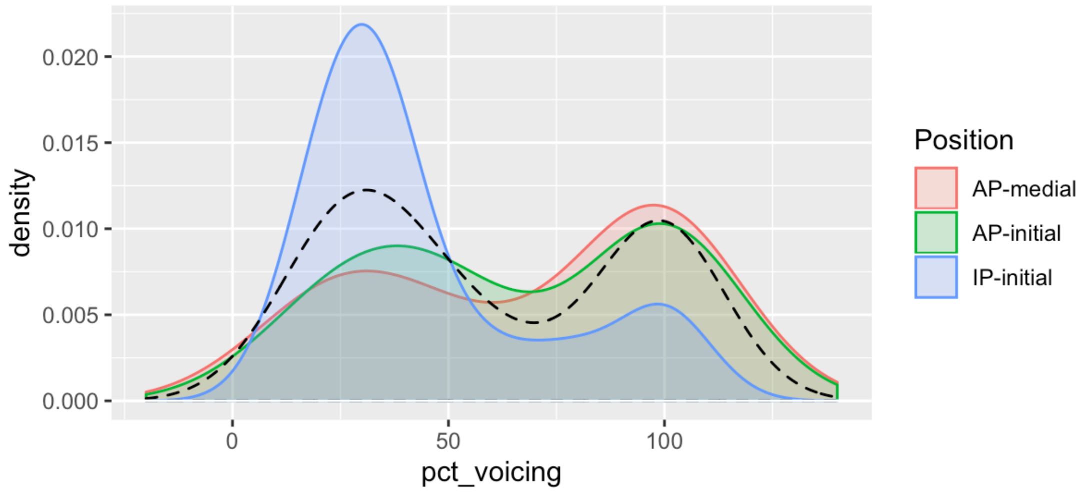


Figure 5.36: Density plot of % voicing for lenis stops produced by Busan_5.

The scatter plot in Figure 5.37 allows a more detailed investigation of the patterns. Recall that this speaker was the only exception to show a significant distinction among the three lower domains in the Tukey test (Mm = Mi > PWi = IPi). To understand the patterns within the AP-medial group, a correlation line was fitted separately for all five positions for this speaker. First, the scatter plot confirms the bimodal pattern with two separate clusters of tokens around both 100% and 30%. Although there were not many tokens, there was a clear division in the voicing pattern of the Mm and Mi tokens versus the PWi tokens. While the Mm and Mi tokens were almost always fully voiced regardless of closure duration, the PWi tokens showed a categorically lower % voicing and duration sensitivity. While the IPi tokens showed a large standard error due to the lack of tokens above the duration of 100 ms, the slope of the regression is negative. Surprisingly, the APi tokens do not show any consistent pattern with regard to duration, realised either with full voicing or with around 40% voicing. Thus, for this speaker, lenis stop (de-)voicing does seem to be a categorical phenomenon with an overlaid effect of gradient voicing, but its conditioning cannot be fully explained with reference to prosodic position and duration. Once again, more data is required to confirm the speculations here.

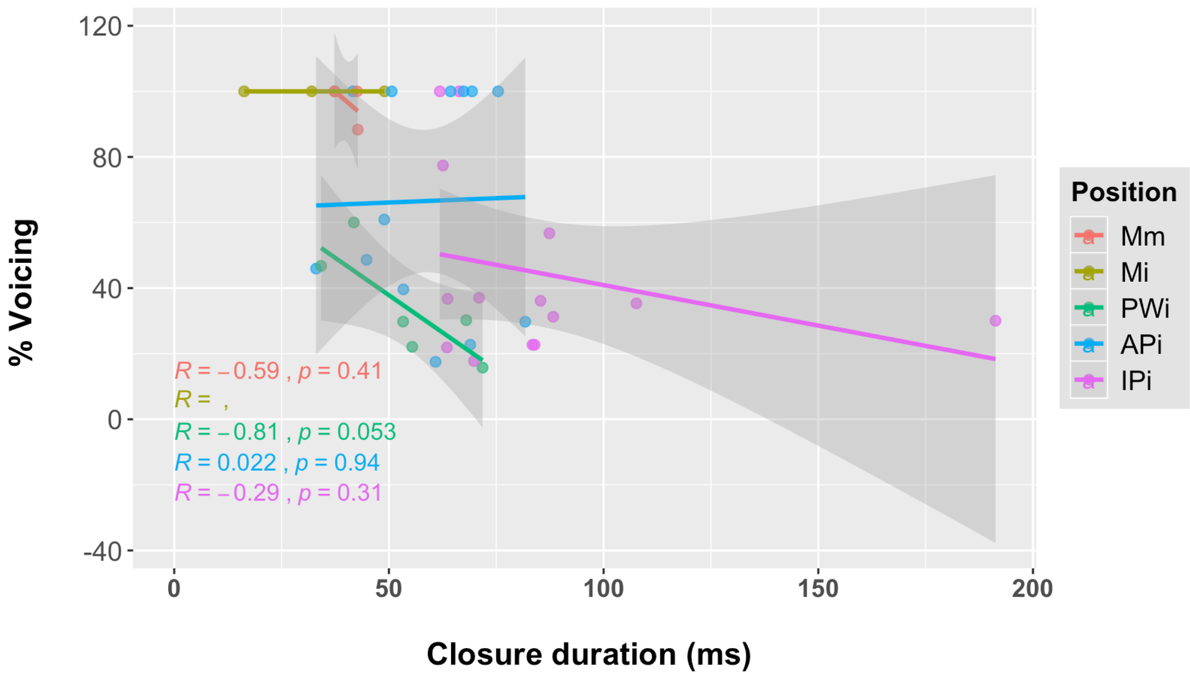


Figure 5.37: % Voicing against closure duration (ms) for lenis stops produced by Busan_5.

Lastly, there were three speakers for whom all prosodic contexts showed duration sensitivity. Take Ulsan_2 for example. Ulsan_2 showed a significant non-unimodal distribution which is clear from the density plot in Figure 5.38. Similar to Busan_5, there was a substantial overlap among the three prosodic positions. AP- and IPi tokens each show a small bump at 100% as well as a larger bump at a lower value, although this was due to one AP- and one IPi token. The substantial overlaps across the prosodic positions are consistent with the results of the pairwise comparisons for this speaker (PWi $=>$ IPi, $p = 0.051$).

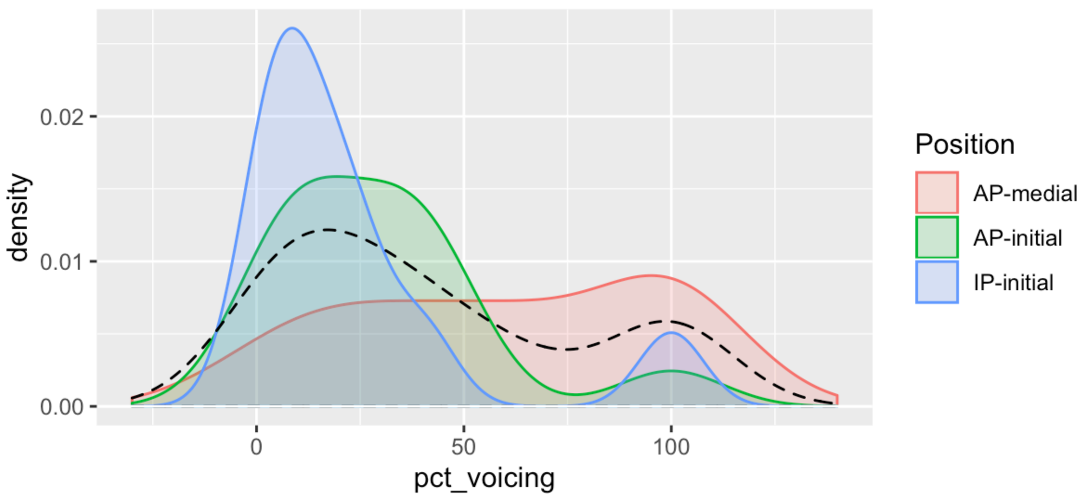


Figure 5.38: Density plot of % voicing for lenis stops produced by Ulsan_2.

Figure 5.39 shows the fitted regression lines for all five positions. All slopes were negative although the correlation was non-significant for the PWi tokens. This is a new pattern because for the speakers discussed above, a vast majority of the APm tokens were fully voiced regardless of closure duration. Although more tokens are required to interpret these results with greater certainty, lenis stop (de-)voicing appears to be categorical for this speaker although conditioned by a factor independent from prosodic position. In addition, the voicing pattern of the lenis stops is sensitive to duration regardless of prosodic position. The findings in this section will be summarised and discussed in Section 5.5.9.

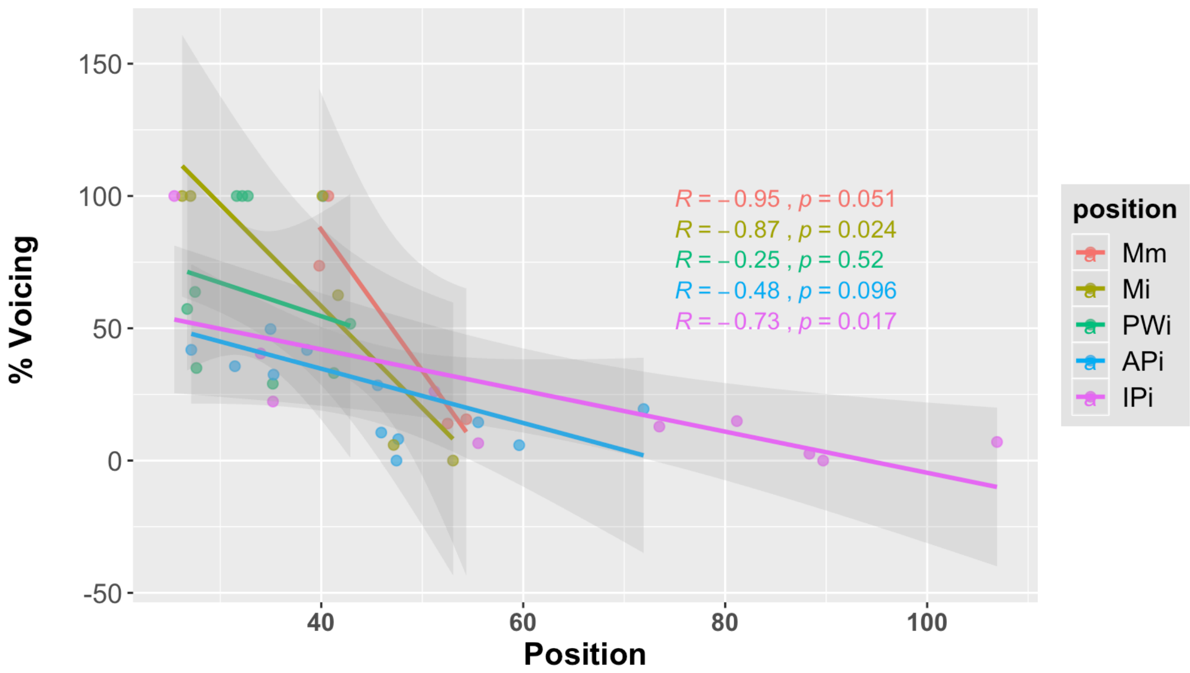


Figure 5.39: % Voicing against closure duration (ms) for lenis stops produced by Ulsan_2.

5.5.7 The life cycle of phonological processes: the status of denasalisation

This section will be structured similarly to Section 5.5.6, but instead focus on the modelling results of the outcome *nasal attenuation* for the nasal tokens in Categories (1) – (3). Table 5.12 shows the results of Hartigan's dip test (in the middle two columns) and a summary of the post-hoc analyses for the linear regression models (in the rightmost column). See Appendix G for the full results of the post-hoc analyses. First, Hartigan's dip-test indicated that only three speakers, Seoul_1, Seoul_5 and Busan_6, exhibited a distribution of nasal attenuation which is statistically non-unimodal at $p < 0.05$. Another Seoul speaker showed a significant non-unimodal distribution at $p < 0.1$. This provides evidence that, at least for these speakers, the process of nasal attenuation has stabilised into a categorical process.

On the other hand, the results of the Tukey test showed a great deal of similarity across the speakers regardless of whether they showed statistical bimodality. Including the effects with a p-value between 0.1 and 0.05, all speakers showed at least a two-way distinction in which there was greater nasal attenuation in a higher position. All but one speaker, Seoul_4, showed significantly greater nasal attenuation in IPi position than in at least two lower positions. Eleven speakers showed the pattern, Mm = Mi = PWi = APi < IPi (2 Seoul, 4 Busan & 4 Ulsan) and one speaker showed the pattern, Mm = Mi = PWi < APi = IPi (1 Seoul). Nine speakers made a consistent three-way distinction, Mm = Mi = PWi < APi < IPi (4 Seoul, 2 Busan & 3 Ulsan). Only Ulsan_10 showed a violation of a cumulative pattern in which the Mm tokens showed greater nasal attenuation than the PWi tokens. This was also the only speaker who showed a significant difference within the APm positions.

First, we will focus on interpreting the results for the four speakers who passed the bimodality criterion, before discussing the results for the remaining speakers. The Tukey test showed the hierarchy of Mm = Mi = PWi < APi = IPi for Seoul_5 and Mm = Mi = PWi < APi < IPi for the three other speakers (although note Mi <= APi, $p < 0.09$ for Busan_6). To visualise the distribution across position, density plots were produced. Unlike in the previous section, the distribution was shown separately for all five positions as there were more tokens for the nasal stops.

Table 5.12: The results of Hartigan's dip tests and Tukey pairwise comparisons for nasal attenuation (dB) ($^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$). For Tukey pairwise comparisons, the sign > was used between levels that are significantly different at $p < 0.05$; the sign => was used for two levels that are different with a p-value between 0.05 and 0.1[43].

Subject	**D-statistic**	***p*-value**	**Tukey test**
Seoul_1	0.068777	0.0423**	Mm = Mi = PWi < APi < IPi
Seoul_2	0.029006	0.9941	Mm = Mi = PWi < IPi
Seoul_3	0.035541	0.9247	Mm = Mi = PWi = APi < IPi
Seoul_4	0.035722	0.9609	Mm = PWi < APi
Seoul_5	0.092376	0.0003***	Mm = Mi = PWi < APi = IPi
Seoul_6	0.061389	0.0817*	Mm = Mi = PWi < APi < IPi
Seoul_7	0.034364	0.9655	Mm = Mi = PWi = APi < IPi
Seoul_8	0.037209	0.9174	Mm = Mi = PWi < APi < IPi
Seoul_9	0.053086	0.3685	Mm = Mi = PWi < APi < IPi
Seoul_10	0.041008	0.6229	Mm = Mi = PWi = APi < IPi
Busan_1	0.050841	0.3511	Mm = Mi = PWi < APi < IPi
Busan_2	0.05186	0.6375	Mm < IPi, PWi <= IPi ($p < 0.06$)
Busan_3	0.036373	0.9412	Mm = Mi = PWi = APi < IPi
Busan_4	0.030195	0.9818	Mm = PWi < APi < IPi, Mi < IPi
Busan_5	0.0411	0.6878	Mm = PWi < APi < IPi, Mi < IPi
Busan_6	0.067519	0.04564**	Mm = PWi < APi < IPi, Mi < IPi, Mi <= APi ($p < 0.09$)
Busan_7	0.042436	0.7914	Mm = PWi = APi < IPi, Mi <= IPi ($p < 0.09$)
Busan_8	0.03175	0.9909	Mm = Mi = PWi = APi < IPi
Busan_9	0.046454,	0.465	Mm = Mi = PWi < APi < IPi
Busan_10	0.038725	0.9682	Mm = Mi = PWi = APi < IPi
Ulsan_1	0.034911	0.9902	Mm = Mi = PWi = APi < IPi
Ulsan_2	0.028064	0.994	Mm = Mi = PWi <= APi <= IPi ($p < 0.06$)
Ulsan_3	0.044304	0.6298	Mi = PWi = APi < IPi & Mm < APi < IPi
Ulsan_4	0.033585	0.9715	Mm = Mi = PWi = APi < IPi
Ulsan_5	0.032115	0.9714	Mm = Mi = PWi = APi < IPi
Ulsan_6	0.029042	0.9919	Mm = Mi = PWi = APi < IPi
Ulsan_7	0.061407	0.1387	Mm = Mi = PWi < APi < IPi
Ulsan_8	0.040298	0.7869	Mm = Mi = PWi < APi < IPi (but Mm <= APi, $p < 0.09$)
Ulsan_9	0.054648	0.2786	Mm = Mi = PWi < IPi
Ulsan_10	0.046731	0.5657	Mm > PWi, Mi = PWi = APi < IPi

[43] As explained in Footnote 42, the additional use of 0.1 as a threshold is intended to make more information available to the reader by reporting weaker statistical trends as well as more robust ones. For more precise understanding of the results, see the exact p-values in Appendix G.

Figure 5.40 shows a density plot for Seoul_5. The x-axis is for raw minimum intensity and, thus, a lower value indicates greater denasalisation. (It does not matter whether one uses raw or speaker-normalised nasal attenuation as the models were for individual data.) The distribution clearly shows a bimodal distribution with the APi and IPi tokens forming one category with a mode around 60 dB and the other tokens forming another category with a mode around 72.5 dB. This pattern mirrors the results of the Tukey test. Recall the discussion at the end of Section 5.3.9 that a bimodal distribution may result not only from a truly categorical process but also from a potential non-linear relationship between the velum height and acoustic intensity. In light of the current pattern, however, it is difficult to imagine that a gradient change in the velum height could have led to such a strong bimodal distribution of acoustic intensity, with the two modes that are more than 10 dB apart from each other.

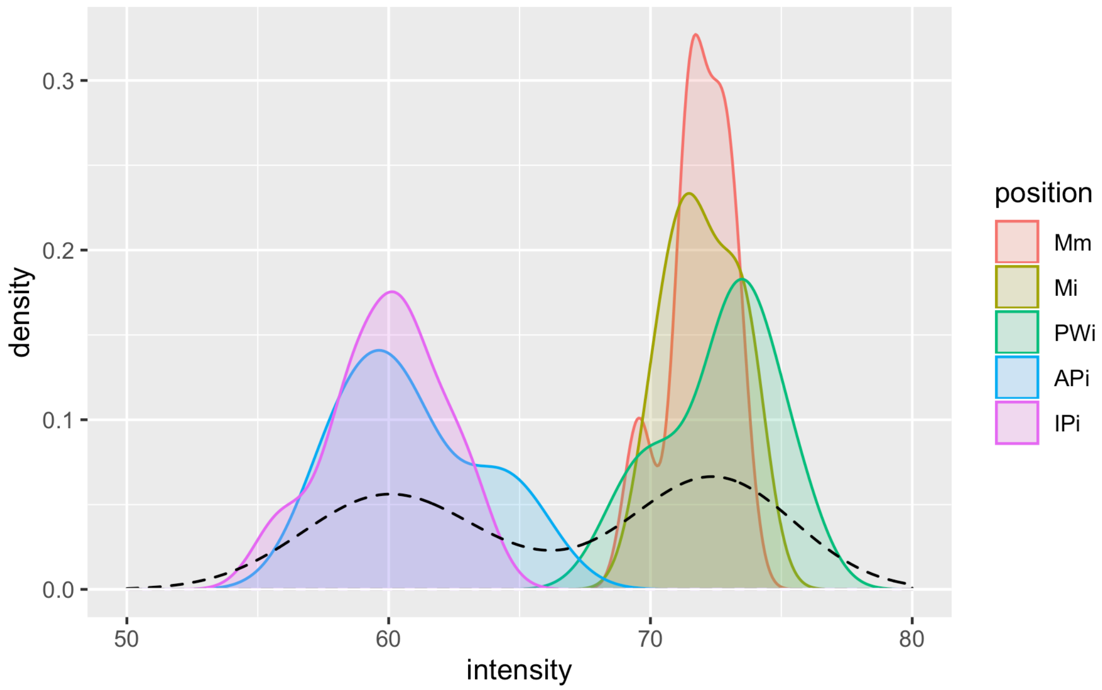


Figure 5.40: Density plot of minimum intensity (dB) for nasal stops produced by Seoul_5.

As shown in Figure 5.41, the linear correlation between closure duration and minimum intensity was not statistically significant for any of the three prosodic positions for Seoul_5 (The correlation results for all 30 speakers are available in Appendix H). As can be predicted from Figure 5.40, the APm tokens clustered around the intensity of 72.5 dB, whereas the APi and IPi tokens clustered around 60 dB, regardless of closure duration. The pattern can be interpreted to suggest that, for this speaker, denasalisation is a categorical phenomenon which targets the nasal stops in AP- and IPi positions.

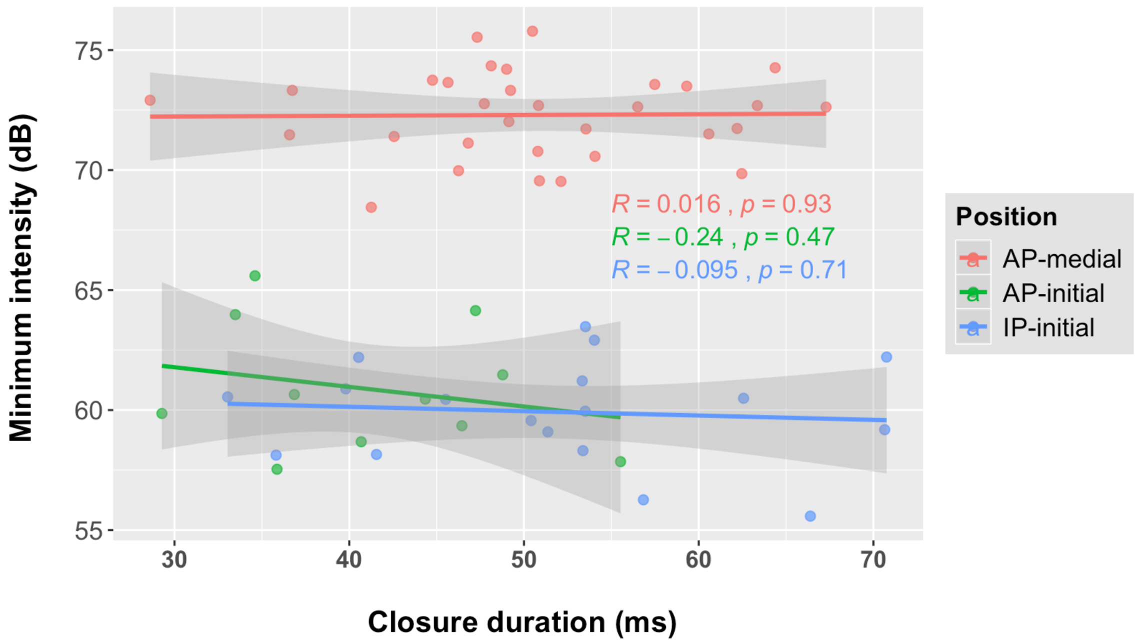


Figure 5.41: Minimum intensity (dB) against closure duration (ms) for nasal stops produced by Seoul_5.

Next, Figure 5.42 shows a density plot by Seoul_1, another speaker for whom Hartigan's dip test indicated a significant non-unimodal distribution. The pattern for this speaker is slightly different from the pattern for Seoul_5 above in that a small number of IPi tokens formed a bump with a mode between 45 and 50 dB. While the dip-test doesn't tell us whether this particular bump is due to chance and, thus, whether the distribution is bimodal or trimodal, it is consistent with the pairwise comparisons that IPi tokens showed a lower minimum intensity than APi tokens. As shown in Appendix F, a similar pattern was found for many other speakers in which there was a small bump with a substantially low intensity value, most often formed by IPi tokens, but occasionally by APi tokens. The pattern of correlation for Seoul_1 (Figure 5.42) differed from that of Seoul_5 (Figure 5.41). Although the two were similar in that the AP and IPi tokens showed lower minimum intensity than the APm tokens, the correlation was significant for the IPi tokens of Seoul_1 with Pearson's r = - 0.66 (p = 0.0074). This suggests that in addition to categorical denasalisation which targets the AP- and IPi tokens, there is a gradient effect of duration affecting IPi tokens.

The scatter plot in Figure 5.43 also reveals that the small bump observed in the density plot was caused by two IPi tokens with intensity below 50 dB. The token with a duration value around 100 ms is close to the regression line whereas the token around 55 ms shows a much lower intensity than would be predicted by the regression. Further examination of the latter token indicated that it was one of the unusual tokens in Category (3) D in which there was a brief period of voicelessness in the middle of the closure. Thus, the unexpectedly low intensity was due to weaker voicing in addition to weaker nasality. Setting the unusual case aside, only one token was causing the bump in the distribution and was consistent with the pattern of regression. This suggests that the bump in the distribution is most likely a result of the sparsity of tokens with long duration. Therefore, the bump cannot be seen as constituting a category of its own.

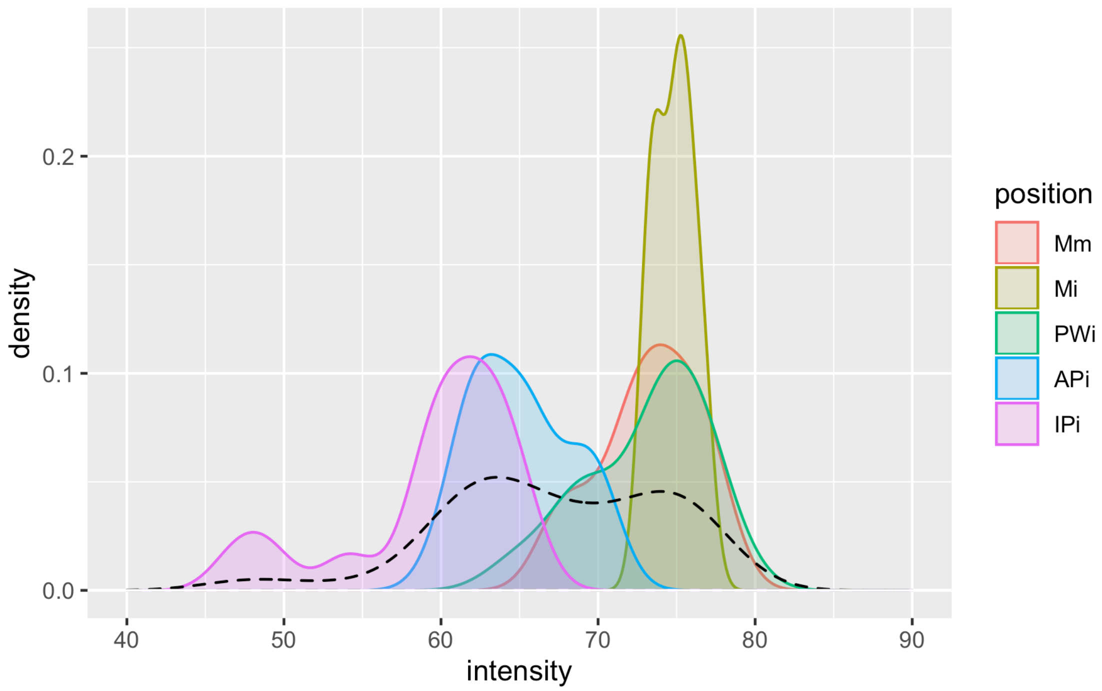


Figure 5.42: Density plot of minimum intensity for nasal stops produced by Seoul_1.

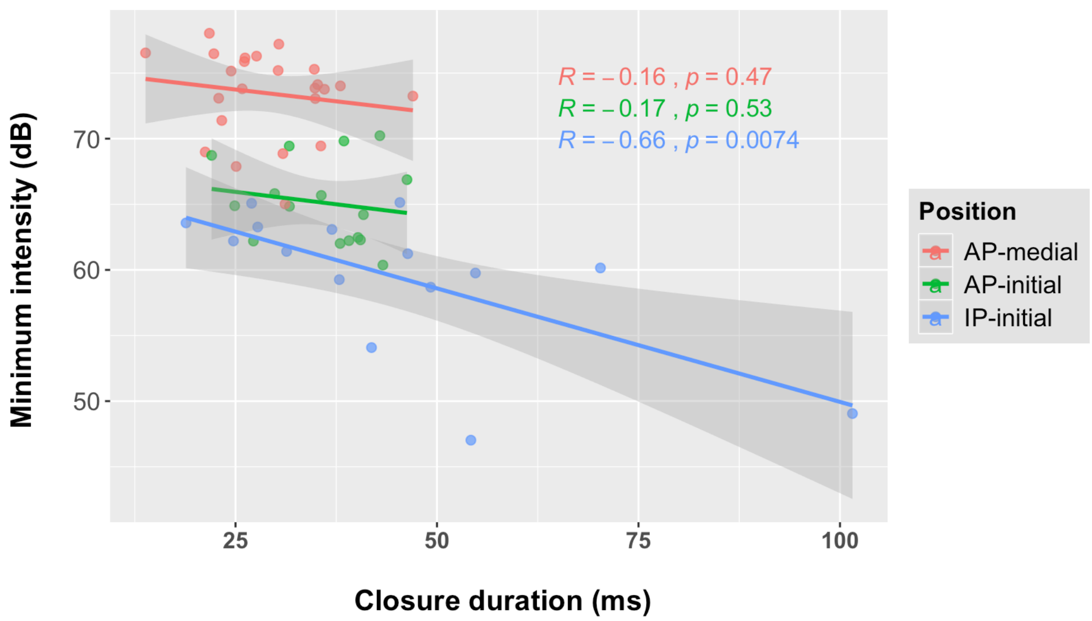


Figure 5.43: Minimum intensity against closure duration for nasal stops produced by Seoul_1.

Another point to mention with regard to the sparsity of tokens is that not only were the tokens sparse above the duration of 50ms, all of these tokens were in IPi position. Given that Pearson's *r* is highly sensitive to extreme values, the differences in the slopes and the *p*-values of correlation for the three prosodic positions may simply be due to the fact that only the range of closure duration was wide enough to reveal a significant correlation with intensity. Figure 5.44 shows the correlation for the three positions after removing the IPi tokens longer than 50 ms. As expected, the significant correlation for the IPi tokens disappeared. Similar to the speculation in the previous section, it is possible that the seemingly different effects of duration on intensity across prosodic position are only superficial, stemming from the differences in the range of duration found in the three prosodic positions.

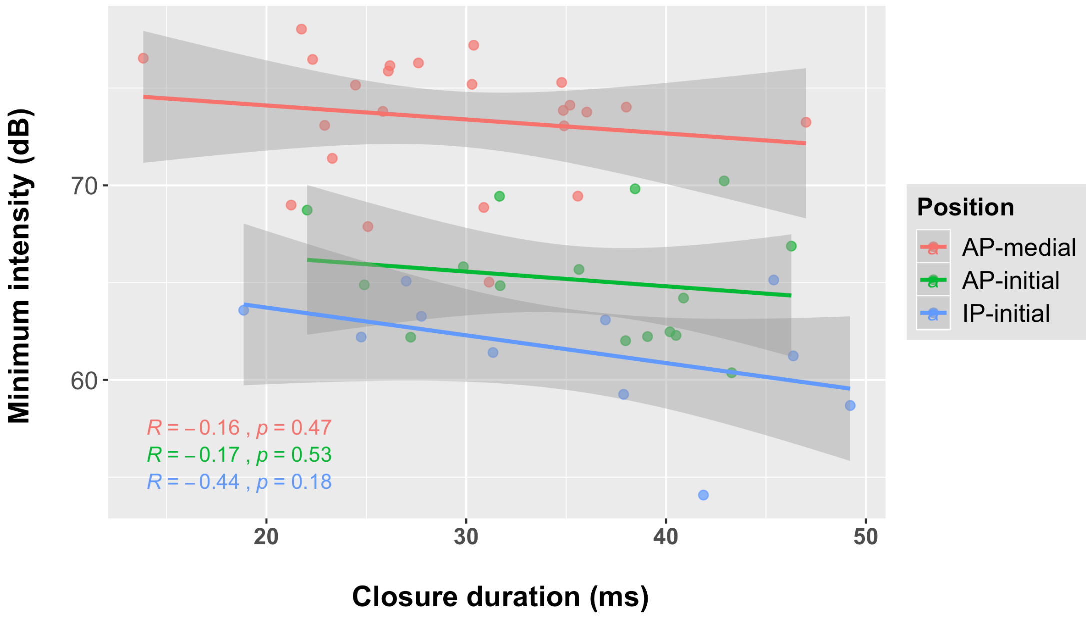


Figure 5.44: Minimum intensity against closure duration for nasal stops produced by Seoul_1, excluding tokens longer than 50 ms.

In support of this speculation, notice that the correlation lines now look similar between Seoul_1 and Seoul_5. Unlike Seoul_5, Seoul_1 did not produce very long IPi tokens. This suggests the initial difference observed between the two speakers was really due to the four IPi tokens in the data by Seoul_1 which were longer than 50 ms. Had more data been collected from Seoul_1, we might have seen IPi tokens with a long duration and a low minimum intensity, giving a significant correlation between duration and intensity for IPi tokens.

The scatter plots for the rest of the speakers are also in line with this speculation; the speakers for whom the IPi tokens did not show a significant correlation between duration and intensity were also the speakers for whom the IPi tokens showed a relatively short range of duration; see Figure 5.45, for example. In contrast, those speakers with a very robust correlation turned out to be those with IPi tokens which were well-spread-out over a longer range of duration. Figure 5.46 and Figure 5.47 illustrate this point. The exceptions to this generalisation were Ulsan_4 and Ulsan_5, as their IP-tokens spanned a relatively long range of duration yet did not show a significant correlation with intensity. However, the speakers also had a relatively few tokens for IPi position (less than 10). This is partly because, as mentioned above, IPi tokens which were in Categories (4) and (5) were removed from the analysis to minimise the effect of voicing on intensity.

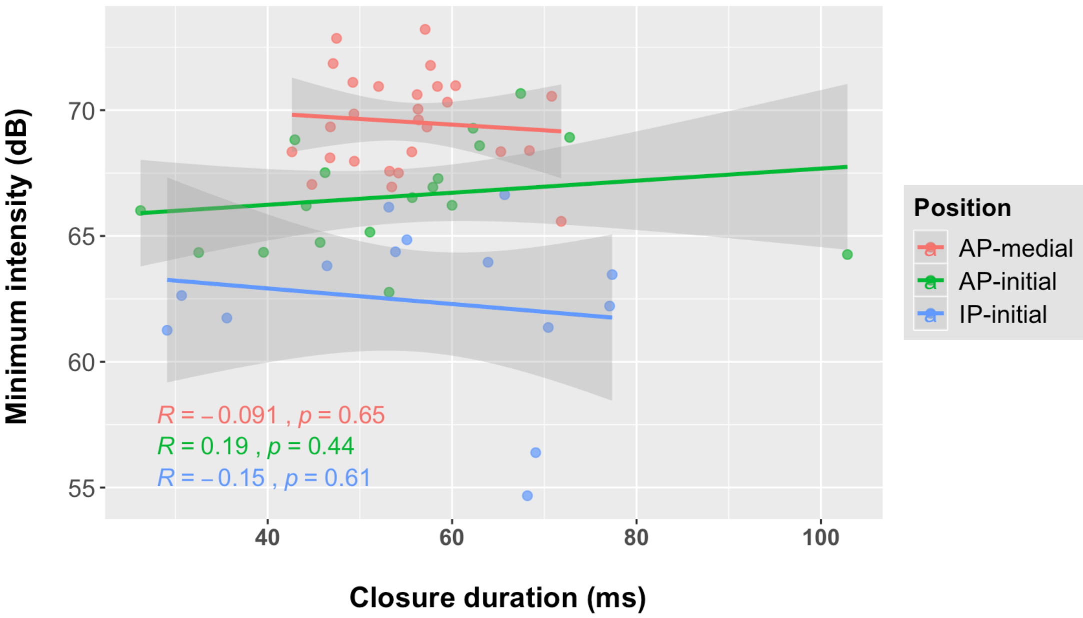


Figure 5.45: Minimum intensity against closure duration for nasal stops produced by Busan_4.

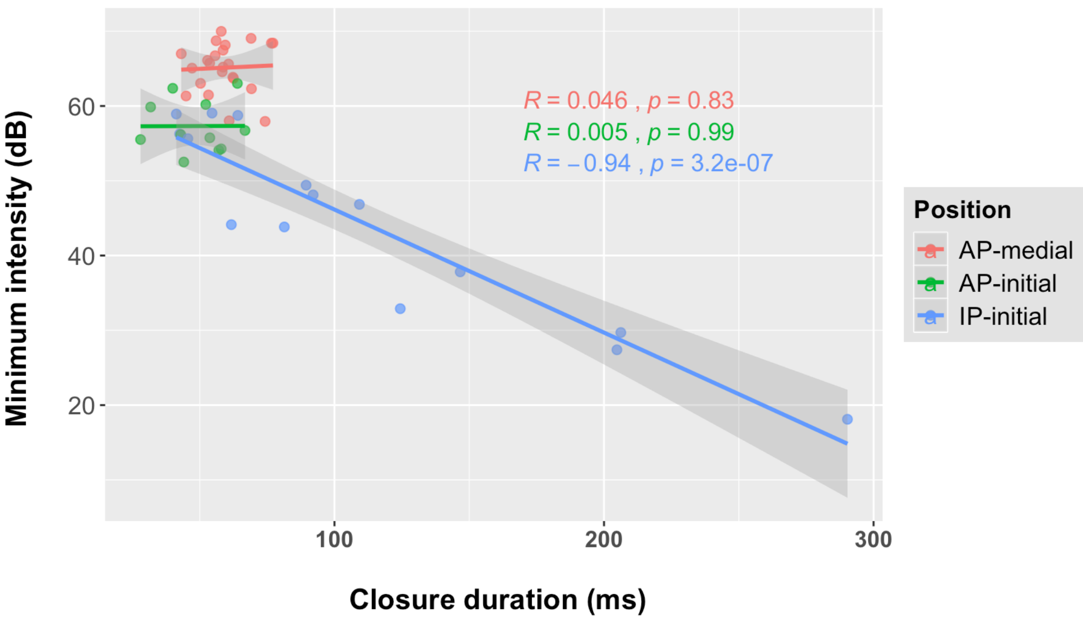


Figure 5.46: Minimum intensity against closure duration for nasal stops produced by Seoul_7.

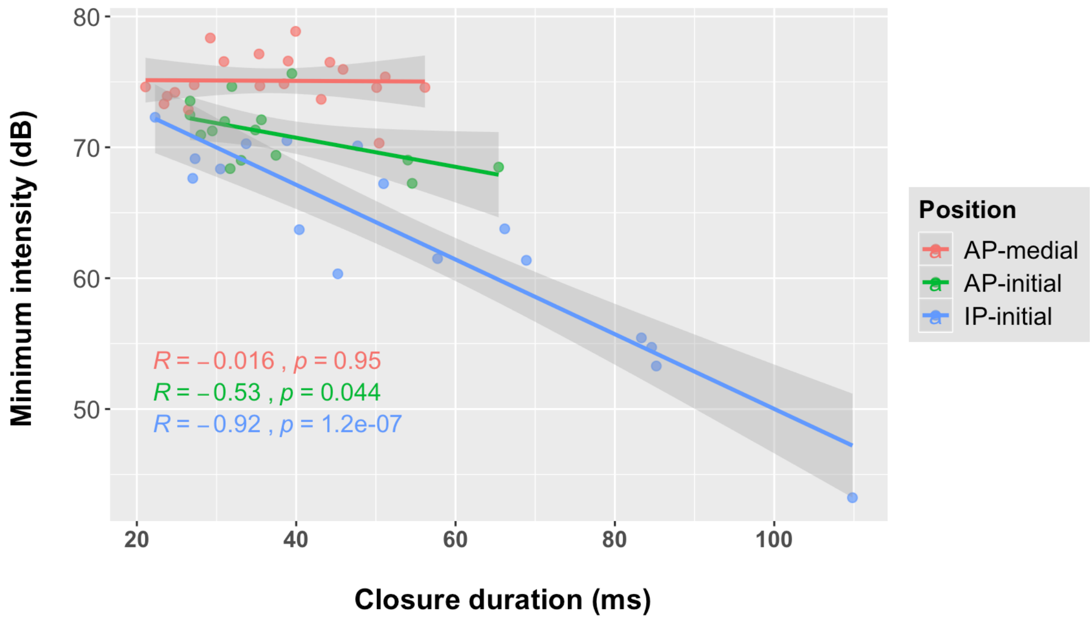


Figure 5.47: Minimum intensity against closure duration for nasal stops produced by Ulsan_6.

Of the two remaining speakers with a non-unimodal distribution, Seoul_6 showed a pattern very similar to Seoul_1. This leaves us to discuss the results for Busan_6. The distribution shown in Figure 5.48 is clearly bimodal. However, the pattern differs from the other speakers in that the distinction between the APm vs. APi & IPi tokens was not as clear-cut.

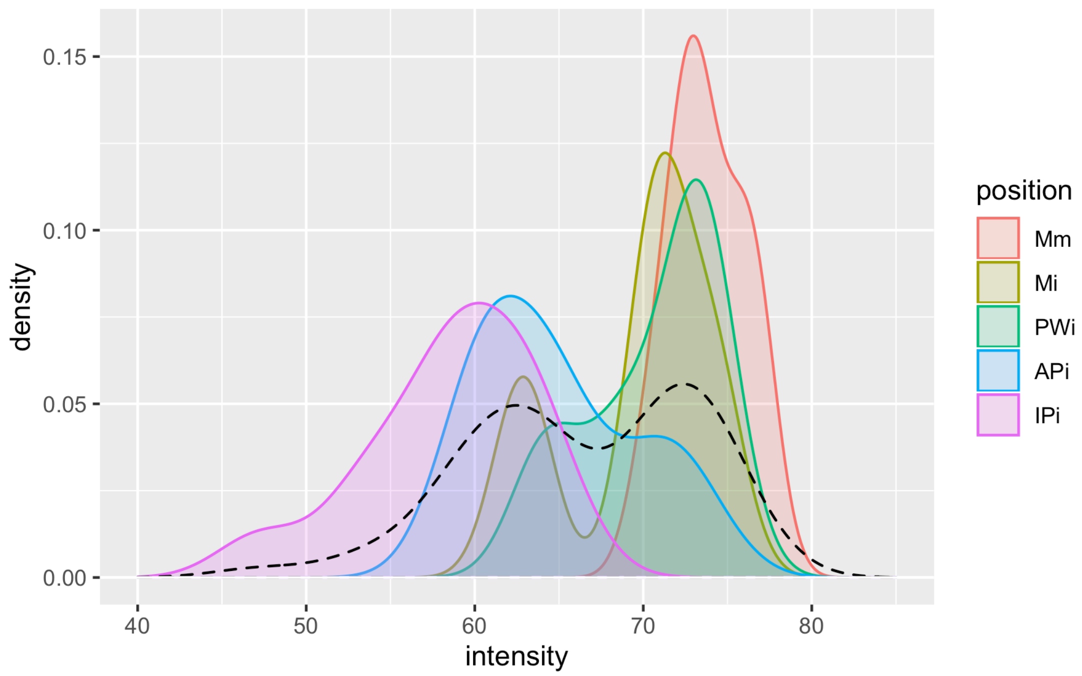


Figure 5.48: Density plot of minimum intensity for nasal stops produced by Busan_6.

The scatter plot in Figure 5.49 shows that there were four AP-medial tokens with an intensity value below 65 dB. Auditory-acoustic characteristics of these four tokens were similar. Figure 5.50 shows an example. While these tokens were not completely denasalised, the first half of the tokens showed less nasality than the second half. This is an unusual pattern as partial denasalisation often takes the form [m^{b} n^{d}].

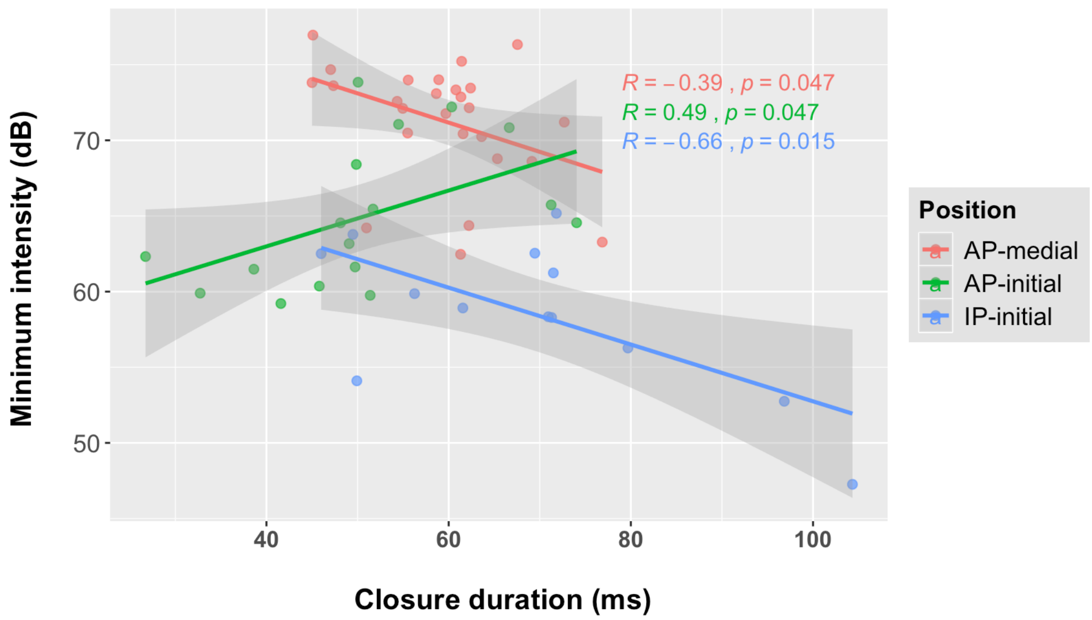


Figure 5.49: Minimum intensity (dB) against closure duration (ms) for nasal stops produced by Busan_6.

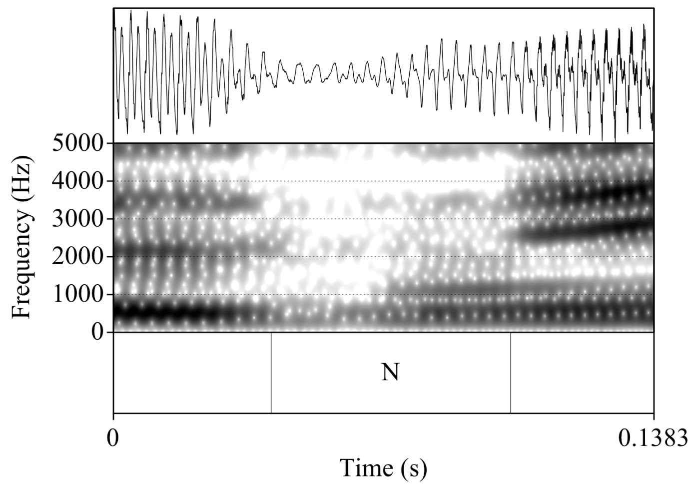


Figure 5.50: Spectrogram showing a pre-denasalised Morpheme-initial nasal by Busan_6.

Another unusual pattern in Busan_6's data was that the APi tokens showed a positive correlation between duration and intensity (p = 0.04694), whereas the APm and IPi tokens showed a negative correlation. This is visually clear from Figure 5.49. The APi tokens from Busan_6 were one of the six cases which showed a significant positive correlation out of the 90 combinations of positions (3) $\times$ speaker (30). While I do not have an explanation for these cases, note that the p-value is only just below the threshold of $p = 0.05$. The remaining five cases also showed a relatively high p-value between 0.047 and 0.097. The negative correlation for the APm tokens was also barely significant.

A more detailed breakdown of the APm tokens can be seen in Figure 5.51. The separate correlation lines reveal that it is mainly the Mi tokens which contributed to the previously positive correlation of the APm tokens, although none of the APm positions shows statistically significant correlation when considered separately. The Mi tokens with a relatively long duration are produced with a relatively low minimum intensity. This explains the unusual Tukey result from this speaker that the Mi tokens were not significantly different from the APi tokens at $p < 0.05$. The pattern in which the Mi tokens were not significantly different from APi and/or IPi tokens was also found with a few other speakers (see Table 5.12 above).

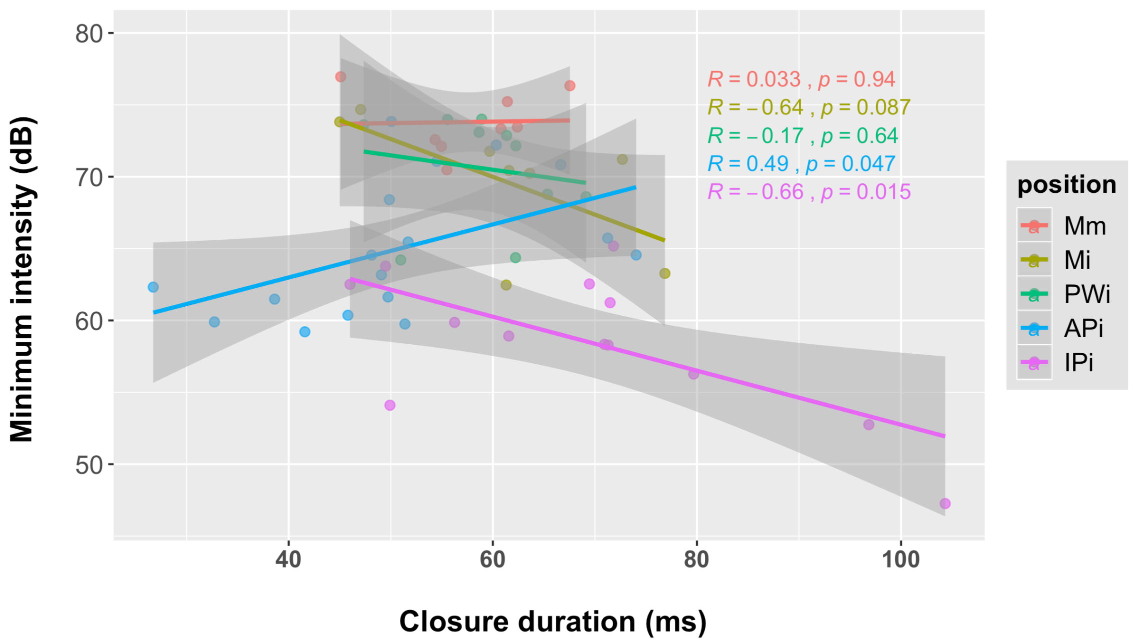


Figure 5.51: Minimum intensity (dB) against closure duration (ms) for nasal stops by Busan_6 for individual prosodic positions.

In line with the pattern, the reanalysis of the data for Seoul_1 with five prosodic positions (Figure 5.52) revealed that only the Mi tokens out of the three types of AP-medial tokens showed a significant negative correlation between duration and intensity with Pearson's r = - 0.92 (p = 0.026). On the other hand, the same reanalysis for Seoul_5 and Seoul_6 showed that none of the AP-medial tokens showed a significant correlation (see

Figure 5.53, for example). A possible explanation may lie in the fact that the Mi tokens spanned a longer range of duration than the Mm and PWi tokens for Busan_6 and Seoul_1, but not for Seoul_5 and Seoul_6. As mentioned above, a sufficiently long duration span seems to be required for a significant effect of duration to be found. Once again, this suggests that the differences across speakers and positions regarding duration sensitivity might be superficial and may disappear with more tokens showing a wider range of closure duration.

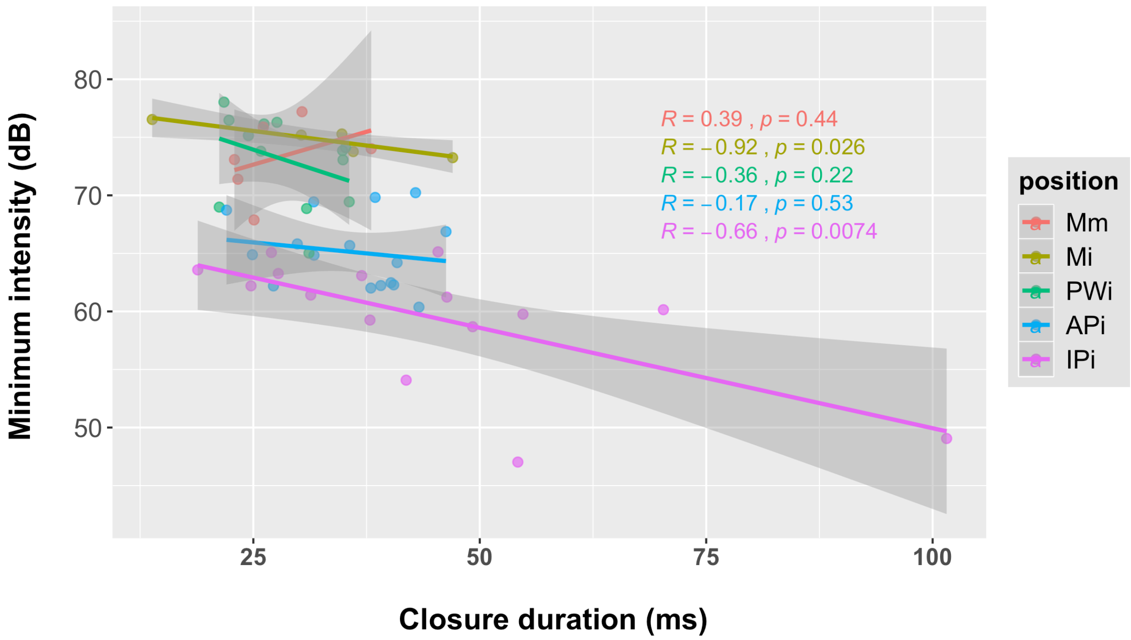


Figure 5.52: Minimum intensity (dB) against closure duration (ms) for nasal stops by Seoul_1 for individual prosodic positions.

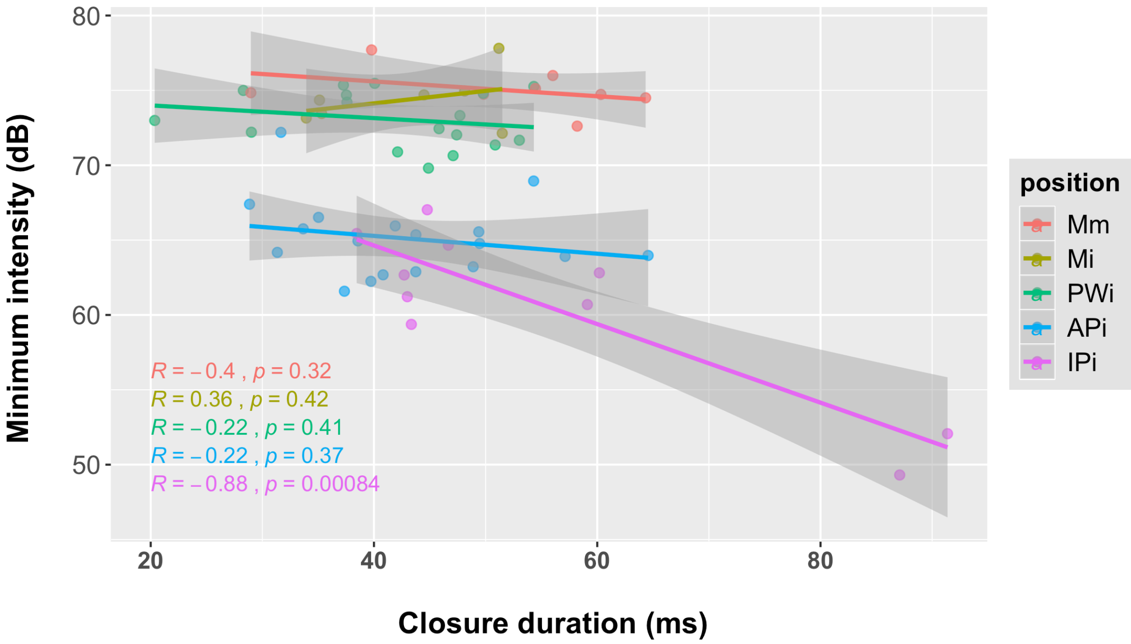


Figure 5.53: Minimum intensity (dB) against closure duration (ms) for nasal stops by Seoul_6 for individual prosodic positions.

Before closing this section, I would like to briefly discuss speakers who did not show a significant non-unimodal distribution. The distribution of intensity measurements from Seoul_9, plotted in Figure 5.54, looks similar to the distribution from Seoul_1, yet Hartigan's dip test was not statistically significant for Seoul_9 (D $= 0.053086, p = 0.3685$). On the other hand, compare Seoul_9 with Ulsan_10, another speaker whose dip test result was also not significant (D $= 0.046731, p = 0.5657$). The density plot for Ulsan_10 in Figure 5.55 is much closer to a true unimodal distribution than is the plot for Seoul_9. Thus, even though the dip test for the majority of the speakers did not produce significant *p*-values, there were some clear differences among these speakers. It is possible that with more tokens, speakers such as Seoul_9 might have shown a statistically significant bimodal distribution.

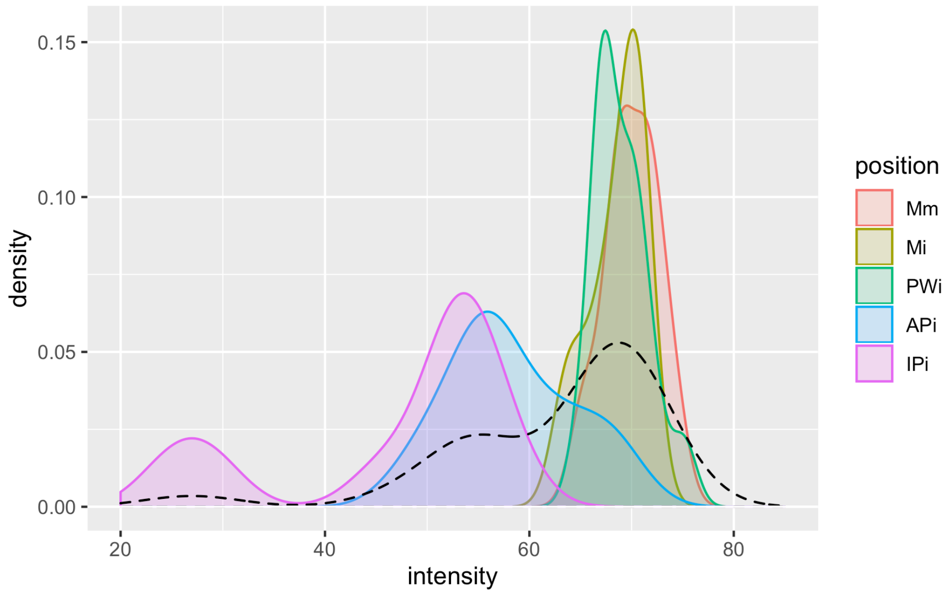


Figure 5.54: Density plot of minimum intensity for nasal stops produced by Seoul_9.

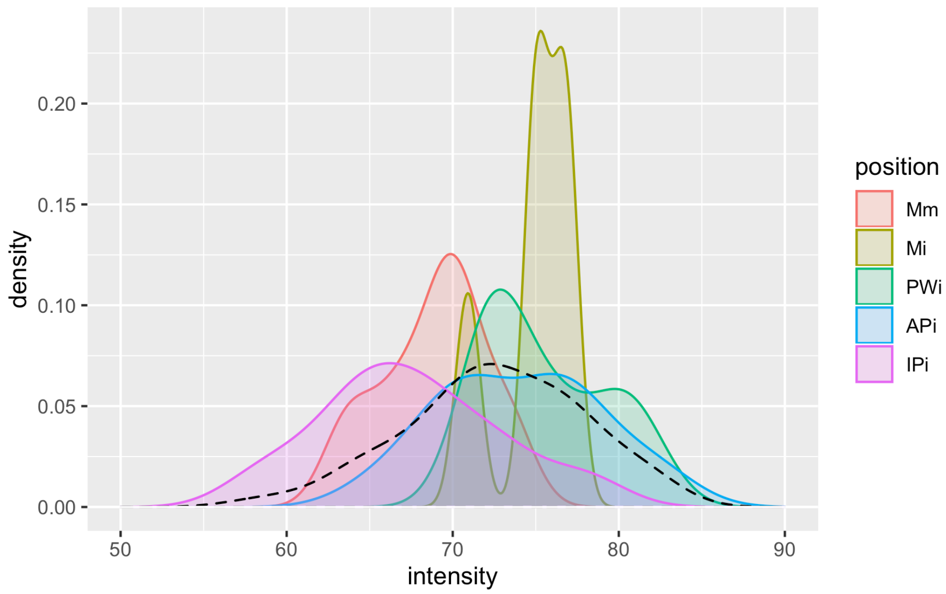


Figure 5.55: Density plot of minimum intensity for nasal stops produced by Ulsan_10.

So far, the speakers with a significant non-unimodal distribution showed a pattern in which APi and IPi tokens formed a group against APm tokens. This was then interpreted to suggest that denasalisation categorically targets AP- and IPi tokens for these speakers. However, according to the Tukey tests in Table 5.12, there were 11 speakers who showed a two-way distinction between IPi position and one or more lower domains at $p < 0.1$, but not between APi position and a lower domain. Indeed, examination of the density plots by these speakers showed a pattern which resembled bimodality between IPi tokens and the other tokens; see Figure 5.56 for Busan_3, for example. While the distribution was not significantly non-unimodal according to the dip test, it certainly appears as if more IPi tokens could have led to bimodality with one mode around 50 dB and another mode around 67 dB. A summary and discussion of the results in this section will be provided in Section 5.5.9.

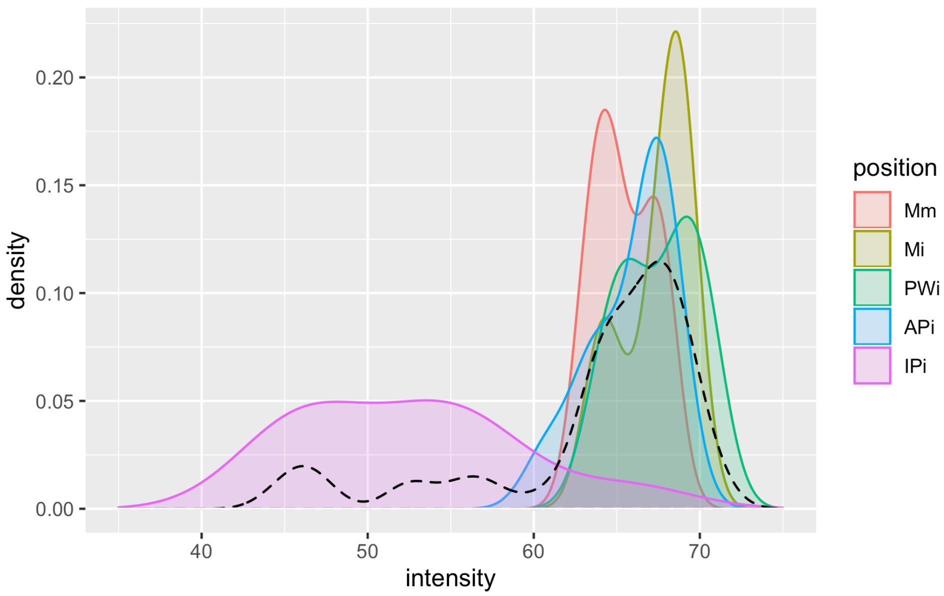


Figure 5.56: Density plot of minimum intensity for nasal stops produced by Busan_3.

5.5.8 Comparing alternative models of DIS

In this section, I will report the results from the ANOVA comparison of the four models with the outcome *% voicing* and *minimum intensity*:

- Model 1: outcome ~ closure duration
- Model 2: outcome ~ position
- Model 3: outcome ~ closure duration + position
- Model 4: outcome ~ closure duration : position + closure duration + position

Table 5.13 shows a summary of the best fitting model for each speaker (see Appendix I for the full results). Overall, there is no strong tendency towards a single type of model. Different models provided the best fit for different speakers. However, for % voicing, an interesting regional pattern can be observed. On the one hand, Model 1 was never the best fitting model for any speaker from Seoul. This indicates that, for Seoul speakers, % voicing is best explained with reference to prosodic position, whether or not it is in conjunction with the gradient effect of duration. On the other hand, Model 1 provided the best fit for six Busan speakers and four Ulsan speakers. This suggests that, for these speakers, % voicing is better described as a function of closure duration alone than with reference to the three prosodic levels, APm, APi and IPi. For minimum intensity, Model 1 was not the best model for any speaker. This suggests that denasalisation is not simply an effect of duration but is conditioned by prosodic structure, either with (Model 3 & 4) or without (Model 2) an overlaid effect of duration.

According to the results from the Tukey pairwise comparisons (Sections 5.5.6 and 5.5.7), there were seven speakers who did not show any significant difference across the five prosodic positions for the outcome of % voicing: Seoul_4, Busan_1, Busan_2, Busan_6, Busan_10, Ulsan_1, and Ulsan_4. These speakers are shown by shading in Table 5.13. (All speakers showed at least some distinction across the prosodic positions for minimum intensity.) Of these seven speakers, Model 1 provided the best fit for only four of them. Model 2 was chosen as the best fitting model for Busan_10 and Ulsan_4. This suggests that, even though the difference across the prosodic positions did not reach statistical significance, the model based on prosodic position still explained more variance in the data than the model based on duration. For Seoul_4, Model 3 turned out to be best. Adjusted r^2 was 0.26 for Model 1 and 0.13 for Model 2, but significantly higher at 0.42 for Model 3. ANOVA comparisons showed that Model 3 was significantly better at capturing the data than both Model 1 and Model 2. This suggests that both closure and position play a role in the variation of lenis stop (de-)voicing for this speaker, and that the effect of prosodic position was not apparent from the pairwise comparisons due to the overlaid effect of duration.

Lastly, the speakers for whom Model 4 was the best fitting model can be analysed further in order to understand how exactly duration and position interacted. For % voicing, there were five such speakers: Seoul_2, Seoul_5, Busan_7, Ulsan_5 and Ulsan_9. Examining the correlation results for these speakers (Appendix H) revealed that, for four speakers, the significant interaction was due to the fact that there was no effect of duration for the APm tokens, whereas duration negatively affected % voicing for the APi tokens (Busan_7) or both the APi and IPi tokens (Seoul_2, Seoul_5, & Ulsan_9). Ulsan_5 showed a unique pattern in which only the APm tokens were negatively correlated with duration.

For minimum intensity, there were 14 speakers for whom the best fitting model was Model 4. For nine of these speakers, only the IPi tokens showed a significant negative correlation between duration and minimum intensity (Seoul_2, 3, 6, 7, 10, Busan_3, Ulsan_1, 4[44] & 9). For three speakers, there was a significant negative correlation for both the IPi and APi tokens but not for APm tokens (Seoul_8, Ulsan_3 & 6). For one speaker, Busan_9, only the APi tokens showed a significant negative correlation. Finally, Busan_6 showed a significant correlation for all three positions, but in different directions; the APm tokens and IPi tokens showed a negative correlation while the APi tokens showed a positive correlation. Overall, the pattern of interaction observed here is one in which duration is negatively correlated with the outcome variable only in high prosodic positions.

Table 5.13: The best fitting model out of Model 1 - Model 4.

	% Voicing			**Minimum intensity**		
	Seoul	Busan	Ulsan	Seoul	Busan	Ulsan
Speaker 1	2	1	1	3	2	4
Speaker 2	4	1	1	4	2	2
Speaker 3	2	2	1	4	4	4
Speaker 4	3	1	2	2	2	4
Speaker 5	4	1	4	2	3	2
Speaker 6	2	1	2	4	4	4
Speaker 7	3	4	3	4	3	2
Speaker 8	2	2	3	4	3	2
Speaker 9	2	1	4	3	4	4
Speaker 10	3	2	1	4	2	3

[44] According to the correlation test, the apparent negative correlation for IPi position did not reach statistical significance for Ulsan_4. However, the regression results indicated that the slope for duration was significantly more negative for IP position compared to APm position.

5.5.9 Summary and discussion III

Sections 5.5.6 and 5.5.7 reported the individual patterns with regard to the prosody-sensitive variation of the lenis stops and nasal stops. First, I will provide a summary of the results relevant to the lenis stop variation and evaluate **H4** and **H5** based on the results.

> **H4**. The lenis stops show a categorical effect of the lenis stop (de-)voicing rule, regardless of the variety.
>
> **H5**. The domain of the categorical process of lenis stop (de-)voicing is the AP.

Regarding the categoricity of the pattern of lenis stop (de-)voicing, 16 out of 30 speakers showed a significantly non-unimodal distribution according to Hartigan's dip test at $p < 0.01$. Of the 16 speakers, five were from Seoul, five were from Busan and the remaining six were from Ulsan. One additional speaker from Seoul showed a significant non-unimodal distribution at $p < 0.1$. Thus, there is a support for the hypothesis in **H4** that lenis stop (de-)voicing is a categorical phenomenon and this is true of any one of the three varieties, at least for the 16 speakers.

Surprisingly, however, the effect of prosody on the pattern of lenis stop (de-)voicing was not always consistent with what we would expect from the lenis stop (de-)voicing rule. According to the lenis stop (de-)voicing rule, the percentage of voicing during closure in APi and IPi position should be significantly smaller than that in APm position. This prediction can be expressed as the following hierarchy: Mm, Mi, PWi > APi, IPi. Including effects with a *p*-value between 0.1 and 0.05, only four of the 17 speakers who conformed to this pattern showed significant non-unimodality: Seoul_6, Seoul_9, Ulsan_6, and Ulsan_7. Therefore, we only have evidence for **H5** from these four speakers.

Of the 13 speakers who showed a prosodic pattern that deviated from the expected hierarchy of Mm, Mi, PWi > APi, IPi, three speakers were not very far from it. For two speakers, the hierarchy differed only by the fact that Mm position was not significantly different from APi position (Seoul_5, Ulsan_9), and for the remaining was Mi was also not significantly different from APi (Seoul_7). The pattern for three additional speakers was consistent with the hierarchy of Mm, Mi, PWi, APi > IPi (Seoul_4, Busan_7, Ulsan_3). Of the remaining seven, five speakers showed a significant distinction between IPi position and at least one APm position.

It is also worth mentioning that there were two speakers, Seoul_4 and Busan_6, who showed categoricity despite showing no significant difference at all across the five prosodic positions. Of these two, Seoul_4 had only one token for the IPi position, making the result less reliable. However, this was not the case for Busan_6, and it suggests that for this speaker, prosodic position does not condition the discrete application of lenis stop (de-)voicing.

Next, I will summarise the results for the nasal stop variation and evaluate **H6** and **H7**.

H6. The nasal stops show a categorical effect of the denasalisation rule, at least in the more advanced varieties such as Seoul.

H7. The domain of the categorical process of denasalisation is the AP.

According to Hartigan's dip test, there was evidence that denasalisation is categorical for three speakers, Seoul_1, Seoul_5 and Busan_6 at $p < 0.05$ and for another speaker, Seoul_6 at $p < 0.1$. These results are in line with **H6**, albeit for a small number of speakers. All four of these speakers showed the hierarchy Mm, Mi, PWi $<$ APi, IPi for nasal attenuation, suggesting that there is significantly greater denasalisation in APi and IPi positions than in APm positions. This clearly supports the prediction in **H7**.

In Section 5.5.6, Pearson's *r* was calculated to measure the correlation between closure duration and % voicing during closure. This can be used to test the following hypothesis.

H8. The longer the closure duration, the lower the percentage of voicing during closure for the lenis stops.

Twenty-six out of ninety groups (3 positions $\times$ 30 subjects) showed a significant correlation at $p < 0.05$ and all showed a negative correlation. These results support the prediction in **H8**.

Then in Section 5.5.7, Pearson' *r* was calculated between closure duration and minimum intensity to investigate whether denasalisation is correlated with closure duration for the nasal stops. This is relevant to testing **H9** as shown below.

H9. The longer the closure duration, the lower the level of nasality for the nasal stops.

Twenty-nine out of ninety groups (3 positions $\times$ 30 subjects) showed a significant correlation at $p < 0.05$, and all but two showed a negative correlation. The two groups with a positive correlation had a relatively high *p*-value, above 0.046. Thus, our results provide evidence for **H9**.

The results of model comparison reported in Section 5.5.8 can be used to test the predictions of the duration-based account (**H10**) and the rule scattering account (**H11**) of DIS.

H10. DIS effects are best modelled as a function of closure duration

H11. DIS effects are best modelled as a function of the combination of closure duration and prosodic position.

For % voicing there was support for **H10** from six Busan speakers and four Ulsan speakers as Model 1 provided the best fit for their data. However, for the remaining 20 speakers, duration alone was insufficient to capture the variation in % voicing. Importantly, for five Seoul speakers, one Busan speaker and four Ulsan speakers, the variation was best captured by a combination of both closure duration and prosodic position, either with or without interaction. These results corroborate **H11**.

For minimum intensity, on the other hand, the duration-only model was not sufficient for any of the speakers, providing no evidence for **H10**. For 20 out of 30 speakers (8 Seoul, 6 Busan & 6 Ulsan), the best fitting models were ones which were based on both closure duration and prosodic position (Models 3 &4). This is consistent with the prediction in **H11**.

5.6 General discussion

5.6.1 Towards a unified account of DIS for the lenis stops and nasal stops

The present results have overwhelmingly shown that the lenis stops and nasal stops pattern similarly in their response to DIS, in line with the literature. The results provided a strong support for our first hypothesis that both the lenis stops and nasal stops show the following acoustic characteristics in the initial position of a higher domain: longer closure, lower % voicing, longer VOT, and longer total voiceless interval. For the nasal stops, nasal attenuation also became progressively greater in a higher position, indicating greater denasalisation. The acoustic measurements for a given position consistently reflected either equivalent or greater strengthening than for a lower position.[45] This corroborates the previous findings that DIS tends to be cumulative across prosodic positions.

On the other hand, the pattern with residual voicing appears to work against our claim that the lenis stop and nasal stop variation show a close parallel. Contrary to the finding in T. Cho and Keating (2001), residual voicing, normalised or raw, also varied as a function of prosodic position except for Busan. The direction of variation was different for the lenis stops and for the nasal stops. For the lenis stops, residual voicing became shorter in APi and IPi positions compared to the lower positions for Seoul and Ulsan. This suggests that, along with closure duration, residual voicing itself is also influenced by the prosodic structure, contributing to the pattern of lower % voicing in a higher position. However, for the nasal stops, residual voicing became longer in IPi position for all three varieties. This is because of the way we defined residual voicing in this study, which refers to voicing during the acoustic closure of an oral or a nasal stop whether or not they are fully voiced. Our results showed that, unsurprisingly, closure duration of the nasal tokens became longer in a higher position and that a vast majority of the nasal stops were realised with full voicing. Nevertheless, mean % voicing during a nasal stop was significantly smaller in IPi position than in all the lower domains for Seoul and Busan. This suggests that although residual voicing itself does not vary as a function of prosody for the nasal stops, the overall effect of DIS for the lenis and the nasal stops is in the same direction towards a lower % voicing during closure.

Such a parallelism observed between the lenis stops and nasal stops calls for a unified account of DIS. As briefly outlined in Section 1.2.1, the existing interpretation of the lenis stop variation is inconsistent with that of the nasal stop variation. In the literature, the lenis stop phonemes in Korean are widely accepted to be underlyingly voiceless, represented by the symbols /p, t, k/. They are then known to become voiced in an intervocalic position

[45] Note that there was one exception to this cumulative tendency; the APi and IPi lenis tokens showed longer VOT than the P tokens for Seoul. This will be addressed shortly in Section 5.6.5.

within the AP via the lenis stop voicing rule. Under DIS, they also exhibit progressively smaller % voicing and longer VOT in a higher prosodic position. Assuming that the lenis stops are underlyingly voiceless, it follows that there is a process of weakening from their voiceless forms in AP-initial position to their voiced forms in AP-medial position, and a process of strengthening to more obstruent-like realisations, i.e. shorter % voicing and longer VOT, from AP-initial position to a higher domain-initial position.

However, this explanation cannot be extended to the patterns of DIS of the nasal stops. The findings here as well as in previous studies unequivocally show that the initial nasals become progressively denasalised and devoiced in a higher domain, with the AP-initial position being the lowest possible position for denasalisation. In this case, we must assume that the nasal phonemes are underlyingly nasal, and that the default realisation is nasal realisation found in the AP-medial position, which then becomes progressively strengthened in a higher position. It is uneconomical from a theoretical perspective that, despite their similarities, the lenis stop variation is explained by invoking both weakening and strengthening processes when strengthening alone is used to explain the pattern with the nasals.

As proposed in Section 1.2.1, a more parsimonious account of the Korean sound system can be achieved if we assume the lenis stops to be underlyingly voiced, i.e. /b, d, g/. This neatly captures the parallelism between the phonetic patterns of the lenis stops and nasal consonants by the tendency of these underlyingly voiced segments to become progressively strengthened in a higher prosodic position. By this strengthening, the segments acquire more obstruent-like features of voicelessness and aspiration, and, in the case of nasal consonants, a lack of nasality. Recall from Section 1.2.1 that another advantage of this account is that Korean ceases to be a cross-linguistic exception in having three voiceless plosives and no voiced plosive (see M.-R. Kim, 2012, and M.-R. Kim & Duanmu, 2004, for additional arguments for treating the lenis stop phonemes as voiced).

In a non-AP-initial position, there are several contexts in which the lenis stops (and the lenis affricate /d͡z/) are realised without voicing.

(a) The lenis obstruents are voiceless in AP-final position.
(b) The lenis obstruents are voiceless before the fortis and aspirated obstruents.
(c) A consonant cluster consisting of a lenis obstruent and /h/ is realised as an aspirated obstruent. This is often described as a process of consonant coalescence (see, for example, L. Kim & Alderete, 2008).
(d) The lenis obstruents are realised as the homorganic fortis obstruent after another obstruent. This is known as the post obstruent tensing (POT) rule, as briefly mentioned

in Section 1.2.2. According to Ahn (1998, p. 50), the POT rule can be represented as shown in (7).

(7) Post obstruent tensing (POT): $[-\text{son}] \rightarrow [+\text{c.g.}] \,/\, [-\text{son}]\ \_\_\_$

To account for (a) and (b) as well as devoicing in AP-initial position, we need to expand our initial lenis stop devoicing rule as shown in (8).

(8) Lenis stop devoicing (LSD): $/\text{b, d, g}/ \rightarrow [-\text{voice}] \,/ \left\{ \begin{array}{l} {}_{\text{AP}}[\ \_\_\_ \\ \_\_\_\ ]_{\text{AP}} \\ \_\_\_\ [-\text{voice}] \end{array} \right\}$

To account for (c), the coalescence rule can work based on the feature [spread glottis] regardless of whether or not the lenis stops are assumed to be underlyingly voiced or voiceless. Finally, (d) can be explained if POT is assumed to be ordered before LSD. For illustration, the derivation of the surface form [ak.p*o] from the underlying /ag.bo/ is shown in Table 5.14.

Table 5.14: Rule ordering between Post Obstruent Tensing and Lenis Stop Devoicing.

Underlying Representation	**/ag.bo/ 'musical score'**
Rule 1: POT	ag.p*o
Rule 2: LSD	ak.p*o
Surface Representation	[ak.p*o]

Note that the proposed lenis stop devoicing rule is equally phonetically natural as the original lenis stop voicing rule. The devoicing of a lenis stop before a voiceless stop can be explained by a process of regressive assimilation. Furthermore, domain-final obstruent devoicing is observed cross-linguistically, for which there are several motivations including the difficulty of maintaining voicing in final obstruents and the difficulty of perceiving a voice contrast in a final position (Blevins, 2006).

Given the discussion above, one may question why the lenis stops have been widely regarded as voiceless in the literature. Some of the earlier works to discuss the Korean stop contrast include Lisker & Abramson (1964, pp. 396–397) and C.-W. Kim (1965). However, it is difficult to find detailed discussion of the rationale for treating the lenis stops as underlyingly voiceless. It may be the case that the phonetic realisation of the lenis stops in the initial position was accepted as the underlying form without further considerations. The data from Lisker and Abrahamson (1964, p. 391) is based on the "initial stops of isolated words" which are essentially tokens in U-initial position. In addition, they specifically clarify than their choice of symbols is based on their broad transcription rather than on the practices of linguistics specialising in Korean (p. 396). C.-W. Kim (1965, p. 346) also simply mentions word-medial voicing of (voiceless) lenis consonants without further citation or debate.

A possible reason why the view that the lenis stops are voiceless has been dominant in the literature may be based on the following. Intervocalic voicing of the lenis stops has been considered a result of gradient gestural overlap with the neighbouring vowels (e.g. Jun, 1993). For example, S.-D. Cho & Whitman (2019, p. 69) criticise the treatment of the lenis stops as voiced, based on the finding that even in intervocalic positions, the lenis stops may not be completely voiced. However, this may now be understood as a result of the gradient process of domain-initial strengthening applying on underlyingly voiced lenis stops, some of which may be in intervocalic positions.

While the account proposed in this section best explains the results of this dissertation, more research is needed to validate whether this account can be extended beyond the case of the lenis and nasal stops and to consider the implications on the entire feature system of Korean. For example, Korean /h/ is known to show intervocalic voicing or deletion similarly to the lenis stops (Um, 2014; H. Kang & Hyo-Young. Lee, 2019). If so, should we treat /h/ to be underlyingly voiced, i.e. /ɦ/, to maintain a parallel with the lenis and nasal stops? Given that /h/ was found to be deleted more often in high frequency words (T.-H. Choi, N.-S. Lim, & J.-I. Han, 2006), it seems more convincing to regard it as a case of lenition with the voiceless /h/ as the underlying representation. Moreover, /h/ can undergo complete deletion but it would not make sense to claim that the deleted form is the underlying form of /h/. In a similar vein, further investigation is required to understand how the forces of lenition and fortition are to be reconciled in a linguistic theory given that these forces work toward opposite directions. Are we to assume that there is a single underlying form that is subject to fortition in certain contexts and subject to lenition in others? If so, how can we pinpoint the true underlying form?

Before closing this section, we can revisit the result that a Busan speaker showed a significant distinction between the IPi and PWi positions from the Mi and Mm positions in terms of % voicing for the lenis stops (see the end of Section 5.5.9). This contrasts with the rest of the speakers who only made a distinction between the APi and the lower positions or the IPi and the lower positions. Assuming that the lenis stops are voiceless and there is a rule of lenis stop voicing, this Busan speaker can be seen as representing the most conservative pattern according to the prediction of domain-generalisation. This is because voicing within the domain of the PW corresponds to a narrower range of contexts than voicing within the domain of the AP or the IP. However, if we assume that the lenis stops are voiced and there is lenis stop devoicing instead of lenis stop voicing, this speaker can be seen as the most innovative speaker. Devoicing the lenis stops in the PWi position as well as in the higher positions can be described as a generalisation of the lenis stop devoicing rule to include more contexts.

5.6.2 The mechanism underlying DIS

The present study has investigated the view in T. Cho and Keating (2001) that DIS arises as a result of variation in closure duration across different prosodic contexts. Having found a strong correlation between linguopalatal contact and duration, they claimed that "strengthening and lengthening is a single effect in Korean." In line with this claim, the results of the present study showed that as the closure duration became longer, % voicing of the lenis stops became lower, and nasal attenuation of the nasal stops became higher.

Based on the strong correlation between duration and strengthening, T. Cho and Keating suggested that the mechanism underlying DIS is one of articulatory undershoot. That is, the articulatory target of a consonant is more likely to be undershot in a lower position where the segment duration is shorter, whereas the articulatory target is more likely to be achieved fully in a higher position due to a longer duration. As pointed out in Section 1.2.1, this account does not hold in light of the pattern of the nasal stops. The nasal consonants realised with a lack of nasality and voicing cannot merely be seen as a result of a full achievement of the intended articulatory target. It has also been pointed out that, strictly speaking, this interpretation makes "domain-initial strengthening" a misnomer as it assumes that the process in fact stems from "non-domain-initial weakening".

As discussed in Section 2.4.2, there is an alternative explanation of the origin of DIS based on increased articulatory effort or energy (Fougeron, 2001, pp. 131-132; Fougeron & Keating, 1997, p. 3737). One interpretation is that there is extra contraction of the muscles involved in the articulation of domain-initial segments. This can potentially explain a variety of the observed temporal and spatial effects of DIS, such as longer segmental duration, greater linguopalatal contact, less voicing, longer VOT, and higher velum position (Section 2.4.2). This account also offers an explanation for why there was a strong correlation between closure duration and linguopalatal contact, if we assume that both of these parameters are influenced by the level of articulatory effort.

The findings of this dissertation were in line with the view in Fougeron (2001, pp. 131-132) and Fougeron & Keating (1997, p. 3737). The purely gradient, phonetic account of DIS predicts that DIS effects should be best modelled as a function of closure duration, which serves as a proxy for the degree of increased articulatory effort or any other phonetic source of DIS. This was indeed the case for some of the speakers from Busan and Ulsan for % voicing of the lenis stops. However, for the rest of the speakers, the duration-only model did not provide the best fit for the pattern of lenis and nasal stop variation. Instead, it was the models which used a combination of the gradient effect of duration and the categorical effect of prosodic position that captured their data best. These results lead us to the final proposal of this dissertation in the section below.

5.6.3 Lenis stop devoicing and denasalisation as scattered rules of DIS

In Section 1.2.2, I proposed an overarching account of the variation of the lenis stops and nasal stops in Korean based on the framework of the life cycle of phonological processes (Bermúdez-Otero & Trousdale, 2012; Turton 2014). The life cycle provides a diachronic explanation for patterns of synchronic variation. It posits that, with the assumption of a modular, feed-forward architecture of grammar, a sound pattern travels chronologically along a specific unidirectional pathway. Starting as an extra-grammatical phenomenon which originates from a physiological or aerodynamic effect of speech production and perception, a sound pattern first becomes *phonologised* to a language-specific phonetic phenomenon. As a next step, it may then get *stabilised* into the phonological component of the grammar, initially as a phrase-level process, before undergoing successive *domain-narrowing* into the word-level and the stem-level stratum. The last step in the life cycle is *morphologisation* and *lexicalisation*, at which point the sound ceases to be a synchronically active phonological rule and becomes a lexical phenomenon. Within the theory, the notion of rule scattering was introduced, which refers to the phenomenon in which a process operating at one component of the grammar (e.g. phonetics) enters a higher-level component (e.g. phonology) without ceasing to apply in the original component (Robinson, 1976; Bermúdez-Otero, 2010, 2015).

The findings of the present study are consistent with the prediction of the life cycle. As mentioned in the previous section, some speakers' patterns were best explained as a function of duration alone. This supports the view that, for these speakers, DIS is a gradient phonetic effect originally motivated by some physical or physiological consequence of speech production. As discussed in Section 2.4.2, a possible source of this effect is increased articulatory effort, which may be conceived as a result of extra contraction of muscles involved in segmental articulation in domain-initial positions. However, for other speakers, the best-fitting model was the one that made reference to both duration and prosodic positions. For a subset of these speakers, there was evidence from Hartigan's dip test that % voicing and/or the distribution of minimum intensity showed statistical bimodality. This can be interpreted to suggest that there are at least some speakers for whom the prosodic position directly mediates the pattern of DIS in a categorical manner. For these speakers, the gradient process of DIS of the lenis stops and nasal stops can be seen to have stabilised into a categorical phonological process of lenis stop devoicing and denasalisation, respectively. Crucially, for some of the speakers, the gradient effect of duration was still significant alongside the effect of prosodic position. This provides support for the account that lenis stop devoicing and denasalisation are instances of scattered rules which originate from DIS, which continues to affect the articulation of the lenis and the nasal stops gradiently. At the same time, the coexistence of a categorical effect and a gradient effect explains why there has been a debate in the literature concerning the grammatical status of lenis stop devoicing (T. Cho & Keating, 2001, Section 4.4.3; Docherty, 1995; S.-A. Jun, 1995c) and denasalisation (Young Shin Kim, 2011, Section 4.6; Yoshida, 2008).

Before closing this section, it must be pointed out that some of the results were not compatible with the two aforementioned scenarios within the life cycle: the rule scattering scenario and the scenario in which DIS only exists in the phonetic component. For some speakers, the model with the prosodic position as the sole variable provided the best fit for the pattern of % voicing and/or minimum intensity. This could be interpreted to suggest that DIS has stopped being active in the phonetic process as it became stabilised into a phonological process. An alternative explanation is that for these speakers, adding the effect of duration failed to improve the model fit because, as discussed in Sections 5.5.6 and 5.5.7, the range of duration exhibited by the tokens in the data was too narrow for its subtle effect to emerge and amount to statistical significance.

5.6.4 Seoul as the leader of the sound change

The patterns of regional variation suggest that the three varieties may be at different stages in the life cycle. The Bayesian models provided strong evidence that the lenis stops are more likely to be devoiced in APi position for the Seoul speakers than those in Busan. There was also some evidence that Ulsan was intermediate between Seoul and Busan in terms of the probability of devoicing in APi position. For the nasals, Seoul was more likely to show denasalisation than the two other regions in APi and IPi positions. In addition, the results from the acoustic measurements showed that the differences across the prosodic positions were generally greater for Seoul than for Busan and Ulsan. This was true of both the lenis stops and nasal stops. The only exception was that, for the nasal stops, Busan showed a greater VOT difference between postpausal position and the lower positions than Seoul and Ulsan. Then, the individual analyses showed that for % voicing of the lenis stops, none of the Busan speakers showed a significant difference between the APi position and a lower position, while six speakers from Seoul and six speakers from Ulsan showed a significant distinction between the APi position and a lower position.

These results suggest that the domain of lenis stop devoicing is the IP for Busan (except for Busan_5 with the PW as the domain), whereas it is the AP, at least for some speakers from Seoul and Ulsan. Under the assumption that rule generalisation proceeds from a higher domain to a lower domain to include a wider range of environments (Bermúdez-Otero, 2007; Vennemann, 1972, pp. 186-187; Kiparsky, 1988, pp. 393-394), the results can be interpreted such that Busan Korean is at a more advanced stage in the life cycle regarding lenis stop devoicing. Similarly, there were three speakers from Seoul and one speaker from Busan who showed categorical denasalisation at the domain of the AP. Although we must be careful given that the study only sampled 10 speakers from each city, the results suggest that Seoul may be at a more advanced stage with the stabilisation of DIS into AP-initial denasalisation. This confirms the preliminary rule as formulated earlier in Section 1.2.1.

(2) Denasalisation: /m, n/ → $[-\text{nasal}]$ / $_{\text{AP}}[$ ___

5.6.5 Presence of pause as a factor

In this last section of discussion, I will address the result that there was an exception to the cumulative trend that DIS is progressively stronger in a higher position. In the Seoul group, VOT of the lenis stops was significantly longer in APi and IPi position than in P (Postpausal) position. Note that similar results were obtained in a recent study by Yoo and Nolan (forthcoming). One possible explanation for this is that the role of VOT in signalling a prosodic boundary is less important in P position than in APi and IPi position, as there is a preceding pause in P position which serves as a strong boundary cue. (Recall that the IPi condition in this study was strictly those IPi segments which were preceded by final lengthening but not a pause.) A similar reasoning is found in Cho et al. (2007) who found a perceptual benefit of DIS across a Word boundary but not across an IP boundary (Section 4.4). In a listener-based theory of sound change (e.g. Ohala, 1992), we may speculate that listeners register DIS as a marker of a prosodic boundary only in the absence of other more salient cues such as a pause. Over time, this bias in perception may become reflected in the pattern of production such that DIS in U-initial position becomes weaker. However, this explanation is somewhat weakened by the fact that IP-final lengthening, often considered a strong boundary marker, does not also lead to weaker DIS effects.

Alternatively, a production-based explanation may be given, that complete denasalisation is less likely after a pause because the velum is typically lowered at resting position. In the middle of an utterance, however, the velum is not likely to be fully lowered except for the production of a sonorant nasal segment and any coarticulated segments. Thus, if we assume that DIS results in a given amount of velum raising, complete denasalisation would be less likely after a pause where the velum movement starts from a lower position.

This also implies that the preceding segment would be an important factor that conditions denasalisation, e.g. greater denasalisation would be observed in the context of i#n than a#n as the velum is expected to be higher in /i/ than in /a/. Further research is required to empirically test this speculation.

5.7 Chapter summary

This chapter has investigated the production patterns of DIS using the recordings from 30 speakers of the varieties of Korean spoken in the cities of Seoul, Busan, and Ulsan. This is the first study in the literature to survey the patterns of DIS across different regional varieties. Through a detailed case study of DIS, the chapter has shed some light on the possible mechanism underlying DIS. Based on the observed patterns and the literature review, the claim that DIS is driven by articulatory undershoot was refuted. Instead, the results were interpreted with an alternative mechanism that the origin of DIS lies in increased contraction of the muscles involved in speech production.

Furthermore, an overarching account was proposed by making a novel connection among the three phenomena of lenis stop (de-)voicing, denasalisation, and DIS, within the framework of the life cycle of phonological processes. Arguing that the lenis stops are underlyingly voiced, the account states that lenis stop devoicing and denasalisation are instances of scattered rules which originate from a more general, phonetically-motivated phenomenon of DIS. The results provided evidence that, for a subset of the speakers, there is an overlaying of the categorical effects of lenis stop devoicing and denasalisation, and the gradient effects of DIS. This is consistent with the notion of rule scattering. The results also provided support for the view that Seoul is the most advanced variety in the diachronic development modelled by the life cycle of phonological processes. In the next chapter, we will examine the one remaining research question: the perceptual role of DIS in the detection of a prosodic boundary.

Chapter 6.

A perception experiment:

the role of DIS in prosodic parsing

6.1 Chapter overview

This chapter is based on a perception experiment guided by the broad research question introduced in Section 1.3:

RQ5. Does DIS aid listeners in the demarcation of prosodic constituents?

The experiment was designed to investigate whether listeners can exploit DIS cues in the prosodic parsing of an utterance, specifically in distinguishing a stronger boundary (AP) from a weaker boundary (PW). Listeners from three different cities, Seoul, Busan, and Ulsan, responded to manipulated stimuli based on recordings by Seoul Koreans. The sections in this chapter will concern the motivation (5.2), methodology (5.3), hypotheses (5.4), results (5.5), discussion (5.6), and conclusions (5.7) of the perception experiment. In the end, a chapter summary will be provided (5.8).

6.2 Motivation

Based on the literature review in Chapter 4, the perception study here is designed to test the speculation by Fougeron and Keating (1997, p. 3738) that listeners can infer the strength of a prosodic boundary by the degree of articulatory strengthening. Despite the authors' scepticism about the possibility of listeners distinguishing two close boundary types, i.e. those belonging to domains adjacent in the prosodic hierarchy, the current study is designed to explore strengthening in PW-initial and AP-initial positions. This decision is based on three reasons. First, the AP seems to have a special status in Korean DIS. S.-A. Jun (1993) claims that the domain of the lenis stop voicing rule is the AP, and Young Shin Kim (2011) and Yoo (2015a) agree that only nasals in initial position of AP or higher domains are subject to denasalisation[46]. The results of our production study (Chapter 5) suggest that a subset of our speakers have stabilised lenis stop devoicing and denasalisation into categorical phonological processes, which apply at the level of the AP. Second, our results are also consistent with the previous reports that the acoustic range of Korean /t, n/ across different domain-initial positions is relatively wide. Thus, their variation is more likely to provide salient perceptual cues than in cases where the segmental variation under DIS is minimal. Finally, Korean shows most consistency in terms of the cumulative link between the degree of strengthening of a segment and the height of the prosodic position of the segment, compared to English, French, or Taiwanese (Keating et al., 2004). This is corroborated by our results particularly for the parameters of total voiceless interval for the lenis stops and nasal attenuation for the nasal stops. Taken together, these findings justify the prediction that Korean listeners are able to distinguish the AP boundary from the PW boundary based on DIS patterns.

Unlike T. Cho et al.'s (2005) study reviewed in Section 4.4, the present study attempts to isolate the effect of segmental articulation from that of other prosodic cues in prosodic parsing. While articulatory strengthening may be regarded as encompassing changes in duration, amplitude, and perhaps even pitch and voice quality in domain-initial segments, the main contribution of articulatory strengthening studies to the existing literature is that prosodic structure is not only marked by what are traditionally considered as prosodic resources, but also by the variation in segmental articulation. Thus, this experiment will investigate the specific question of whether listeners are sensitive to acoustic cues in the segmental variation when determining the presence or absence of a given prosodic boundary.

6.3 Hypotheses

Based on the literature review in Chapter 4, hypotheses of the present study are formulated as follows.

H1. In the absence of key prosodic cues, listeners can use different DIS patterns at AP and PW boundaries to parse ambiguous utterances.

If **H1** is supported, this can be seen as evidence for the following speculations in Fougeron and Keating (1997), the second speculation being a stronger version of the first:

(1) DIS can serve as a cue to a prosodic boundary

(2) The degree of DIS can mark the strength of a prosodic boundary

[46] To be precise, Young Shin Kim (2011) is based on Ho-Young Lee's (1996) model of Korean Prosody which views Korean as a fixed stress language. Thus, her dissertation concludes that it is stressed word-initial nasals that are subject to denasalisation. However, it is explicitly acknowledged in Kim (p. 26) that this position would correspond to the AP-initial position in the more widely used framework of Korean Prosody by S.-A. Jun (1993).

There are several ways in which listeners can achieve the task described in **H1**. The first way is by correctly associating a certain degree of DIS with an AP boundary and a different, weaker degree of DIS with a PW boundary. The second way is through a relative comparison between the two sets of DIS found in the perceptual input, thereby distinguishing a stronger boundary from a weaker boundary. The final way is by associating a certain degree of DIS with only one of the two boundary types, then inferring weaker or stronger DIS to indicate a weaker or a stronger boundary. Strictly speaking, this last method would support (1) but not (2), because the listener only has to be able to tell whether there is an AP boundary and when it is absent, assume the segment to be in a lower position in the prosodic hierarchy. The present study does not distinguish among the three possible mechanisms.

> **H2**. For the lenis stops, cross-splicing a PW-initial CV to an AP-initial position has a stronger adverse effect for the perception of the "correct" (AP-initial) boundary than vice versa. In contrast, for the nasal stops, cross-splicing an AP-initial CV to a PW-initial position has a stronger adverse effect for the perception of the "correct" (PW-initial) boundary than vice versa.

For ease of reference, "correct" answers that are arbitrarily defined as those answers that are based on the host (recipient) sentence, rather than on the sentence from which the target CV is spliced. This hypothesis is based on the production results in Section 5.5.1 and 5.5.2. While the majority of the lenis stops were realised with a voicing lag longer than 10ms in AP-initial position, their realisations in PW-initial position were more varied from a fully voiced and lenited form to a stop with a voicing lag. Thus, cross-splicing a fully voiced realisation into an AP-initial position is expected to be more misleading than cross-splicing an aspirated realisation into a PW-initial position. On the other hand, for the nasals, the realisation in PW-initial position is much more consistent as a sonorant nasal, whereas their realisation in the AP-initial varies more widely between a sonorant nasal and an oral stop with a voicing lag. We can therefore predict that cross-splicing a denasalised and devoiced nasal into a PW-initial position will be less acceptable than cross-splicing a sonorant nasal into an AP-initial position.

> **H3**. The effectiveness of DIS as a cue to a prosodic boundary depends on the goodness of match between the pattern in the perceptual input and the pattern in the listener's regional variety.

Since the perceptual input was generated based on two Seoul speakers, the DIS cues are predicted to be more effective for the listeners from Seoul than for the listeners from Busan and Ulsan. The results from Chapter 5 are relevant here, in particular the regional variation in the PWi and APi positions. For the lenis stops, the production results generally suggest that the DIS patterns between PWi and APi positions are more strongly distinguished for Seoul than for Busan. Ulsan patterned either with Seoul or Busan, or lay between the two, depending on the exact phonetic parameter. Thus, we can predict that, for the lenis stops, the effectiveness of DIS as a perceptual cue will be stronger for Seoul than for Busan. Less certainly, we can also hypothesise that Ulsan may take an intermediate position between Seoul and Busan.

For the nasal stops, the only significant regional difference was observed in terms of nasal attenuation, a measure which reflects the degree of denasalisation. Seoul showed a greater difference between the PWi and APi positions compared to Busan and Ulsan. This leads to the hypothesis that DIS cues based on Seoul will be more effective for Seoul listeners than for Busan or Ulsan listeners. We would not expect to see a difference between Busan and Ulsan listeners.

6.4 Method

6.4.1 Stimulus manipulation

Of the recording materials in Chapter 5, 10 pairs of sentences (a) and (b) in Part A were used in the perception experiments (see Appendix A for full materials). As previously explained, the minimal pair sentences were identical except for the presence or absence of the AP boundary before the target syllable. An example pair is shown in the leftmost column of Table 6.1. The column is headed *host sentence* because they receive a CV spliced from another sentence.

Stimulus manipulation consisted of three stages: *Stage I - CV splicing, Stage II - duration restoration,* and *Stage III - pitch flattening*. The three stages ensured that the specific effect of segmental articulation could be tested, by eliminating the influence of the two most important prosodic cues to boundaries, duration and pitch (see Section 4.3 for a review).

In *Stage I*, two recordings of each of (a) and (b) were used to create two conditions, same-spliced and cross-spliced, resulting in four versions of the segmental string per each set. For example, to create a same-spliced version of (a), the target syllable /ta/ was spliced from another utterance of (a). On the other hand, for the cross-spliced condition, the target syllable was spliced from an utterance of (b) into (a). As mentioned earlier, answers that are based on the host (recipient) sentence are arbitrarily labelled “correct”, and the answers based on the sentence from which the target CV is spliced, “incorrect”. If a listener’s perception is swayed by DIS cues, then we expect to find more incorrect answers in the cross-spliced condition than in the same-spliced condition.

Table 6.1: An example of minimal pair sentences used as stimuli in the perception experiment. Two versions of each of (a) and (b) were created by same-splicing and cross-splicing the target syllable in bold and italics. Answers based on the original, pre-splicing structure of the host (recipient) sentence are arbitrarily referred to as “correct”.

<table>
<tr><th>Host sentence</th><th>Same-spliced CV</th><th>Cross-spliced CV</th></tr>
<tr><td rowspan="2">(a) PW-initial /ta/
$_{IP}${$_{AP}${ʌmma talimi} $_{AP}${kjʌlkuk} $_{AP}${ko͡tsaŋnasʌ} $_{AP}${pʌljʌs*ʌjo}}
Mom iron finally broke throw away
‘Did you throw away Mom's iron because it finally broke?’</td><td>Answer: (a)
Correct</td><td>Answer: (a)
Correct</td></tr>
<tr><td>Answer: (b)
Incorrect</td><td>Answer: (b)
Incorrect</td></tr>
<tr><td rowspan="2">(b) AP-initial /ta/
$_{IP}${$_{AP}${ʌmma} $_{AP}${talimi} $_{AP}${kjʌlkuk} $_{AP}${ko͡tsaŋnasʌ} $_{AP}${pʌljʌs*ʌjo}}
Mom iron finally broke throw away
‘Did Mom throw away the iron because it finally broke?’</td><td>Answer: (b)
Correct</td><td>Answer: (b)
Correct</td></tr>
<tr><td>Answer: (a)
Incorrect</td><td>Answer: (a)
Incorrect</td></tr>
</table>

The left edge of splicing was made at the zero-crossing nearest the onset of a stop closure. The closure onset was determined by a sudden loss of energy in the spectrum, particularly in the high frequency region, and a sharp downstep in amplitude in the waveform. The right edge of splicing was made at the zero-crossing nearest to the end of the following vowel. For the majority of the sentences, the target CV syllable was followed by one of the following types of onset consonants: a liquid phoneme realised as an alveolar tap /l/[ɾ], or a lenis stop often realised as a voiced allophone /p/[b], /k/[g], /t/[d] or a tensified allophone /k/[k*]. In such cases, the right edge of splicing was determined using the same method as the left edge of splicing, by delimiting the boundary between the vowel and the consonant.

However, there were two pairs of sentences in which the right edge of splicing could not be determined in this way: (1) Sentence 4a and 4b in which the target consonant was followed by a coda nasal and an onset nasal, /***t***oŋne/, and (2) Sentences 9a & 9b in which the target consonant was followed by a coda nasal and a lenis stop, /***n***oŋtam/. Due to coarticulation, it was difficult to reliably segment a nasalised vowel and the following coda nasal. Thus, for /***t***oŋne/, the right edge of splicing was determined at the zero-crossing nearest to the midpoint of the portion corresponding to /oŋn/. This resulted in a splicing point that was in the steady state of the coda nasal, especially since coda nasal tended to be significantly longer than the following onset nasal. Similarly for /***n***oŋtam/, segmentation was tricky due to the gradual increase and the subsequent decrease of acoustic nasality in /oŋt/. The midpoint of /oŋt/ was thus used as the right boundary of splicing.

Unlike the production study, it is imperative that sentence pairs be completely identical – that is, except the prosodic position of the target syllable – including the prosodic structure of the later part of the sentences. For example, one of the speakers in the production study chose an alternative prosodic structure to the one presented in (b) in Table 6.1, not producing an AP boundary before /kjʌlkuk/. However, the lack of an AP boundary could serve as an additional cue to the AP boundary before the target segment /t/, given the finding in S. Kim (2004) that 99% of the Korean APs analysed in her study were made up of one or two content words. It is highly unlikely that the first three words all belong to one AP. This was less of a concern for the production study, as there is no suggestion that the segmental articulation of an AP-initial segment is influenced by the number of syllables in the AP[47], or the prosodic structure of the rest of the sentence. Therefore, in deciding which speaker's recording to use for the perception experiments, I eliminated those who did not produce the prosodic structures exactly as shown in Table 6.1 and in Appendix A for the rest of the sentence pairs.

Unfortunately, no speaker produced all of the test sentences in the desired prosodic structure. This is mainly because, as mentioned in Section 5.3.1, many speakers were inclined to insert an AP boundary before the target syllable in type (a) sentences. This was presumably due to the tendency of Korean APs to be short, with 3.39 syllables per AP on average (S. Kim, 2004). Seoul_10 was the speaker who produced most sentences in the desired structures. Thus, Seoul_10 contributed seven pairs, which was all of the pairs that this speaker produced as intended. For the remaining three pairs (Sets 6, 9, 10 in Appendix A), the recording from Seoul_9 was used[48]. The two speakers were both females and only the sentences produced at *moderate* speed were used.

The rationale behind the choice of splicing the entire syllable, as opposed to the initial segment alone, was as follows. The phonetic realisation of an initial segment strongly influences how the following vowel is realised. For example, T. Cho et al. (2002) observed that one of the systematic acoustic correlates of Korean lenis stops in initial position of words

---

[47] There is a finding that the amplitude of the initial syllable of an AP is influenced by the number of syllables in the AP (S.-A. Jun, 1995b), such that the amplitude of the first syllable is greater than the second syllable if the AP has three syllables, but the reverse is true if the AP has four or more syllables. However, this was not an issue as syllable amplitude was not a parameter of interest for the production study in Chapter 5.

[48] The individual analyses of the production patterns in Sections 5.5.6 and 5.5.7 show that the patterns for Seoul_9 and Seoul_10 do not deviate from the general patterns observed for the other speakers from Seoul (Tables 5.11 & 5.12). For lenis stop voicing, five out of 10 speakers from Seoul showed evidence of stabilisation (statistical bimodality at $p < 0.05$). Seoul_9 was one of the five who showed evidence of stabilisation, whereas Seoul_10 was not. For denasalisation, only two out of 10 Seoul speakers showed evidence of stabilisation at $p < 0.05$. Seoul_9 and Seoul_10 were part of the majority who did not show evidence of stabilisation.

pronounced in isolation is the breathy phonation in the following vowel, in addition to the well-known characteristics of aspiration. While their study does not consider other prosodic positions, such breathiness in the vowel is expected to disappear following an intervocalic lenis stop within an AP, as the lenis stops in this environment are voiced.[49] Similarly, a vowel following a sonorant nasal has been shown to significantly differ in the degree of coarticulatory nasalisation than a vowel following a denasalised nasal (Section 5.3.5 (4)). Vowel nasalisation has also been shown to affect the perception of the syllable /na/ in Yun & Arai (2020) (Section 4.5). In addition, the effect of DIS on the lingual articulation may be reflected in the following vowel through formant transitions. First of all, such coarticulation between C and V implies a practical difficulty of cross-splicing without creating an undesirable perceptual juncture at splicing points. Second, the following vowel may contain crucial acoustic cues that signal the phonetic quality of the initial segment, and hence allow the experimental stimuli to test the full effects of DIS.

On the other hand, the pre-boundary segment(s) was not spliced together with the post-boundary CV, although there may possibly be domain-final strengthening cues which listeners could exploit. This is so that the current study can specifically investigate the role of DIS in marking of a prosodic boundary. In addition, T. Cho & Keating (2001) showed that AP-final lengthening in Korean is very small compared to IP-final lengthening, which explains why earlier studies concluded that there is no significant AP-final lengthening (S.-A. Jun, 1993, 1995c). In addition, no greater amplitude was found at the end of the AP (S.-A. Jun, 1995a). Nevertheless, there may be spatial cues in domain-final vowels such as less linguopalatal contact reflected in the vowel formants (Fougeron & Keating, 1997). The fact that potential domain-final strengthening cues were left unspliced implies that, if the results turn out to support our hypothesis (**H1**), this would be despite such cues working against the hypothesis. Furthermore, some studies showed that prosodic boundary effects can be extended to even three syllables from the boundary (Byrd & Riggs, 2008). However, no such extended effects have been reported for Korean. Even if there were any cues in the syllables following the domain-initial syllable, these would also work against our hypothesis as they were not spliced with the domain-initial syllable.

Thus, in the second stage – duration restoration – only the durations of the C and V which had been spliced in were modified to match the durations of the original C and V slots in the host utterance. As C to V ratio increased after a stronger boundary as well as absolute

[49] Recall the finding reviewed in Section 3.4.1 that the vowel after the lenis stops is breathy in Seoul Korean but not in Gyeongsang Korean (M.-R. Kim, 2017). This could potentially bias our results because vowel breathiness may serve as a signal to the prosodic position of the lenis stops more strongly for the participants from Seoul.

CV duration, C and V durations were individually restored to original values. All duration modification was conducted in Praat's Manipulation Editor, as exemplified in Figure 6.1.

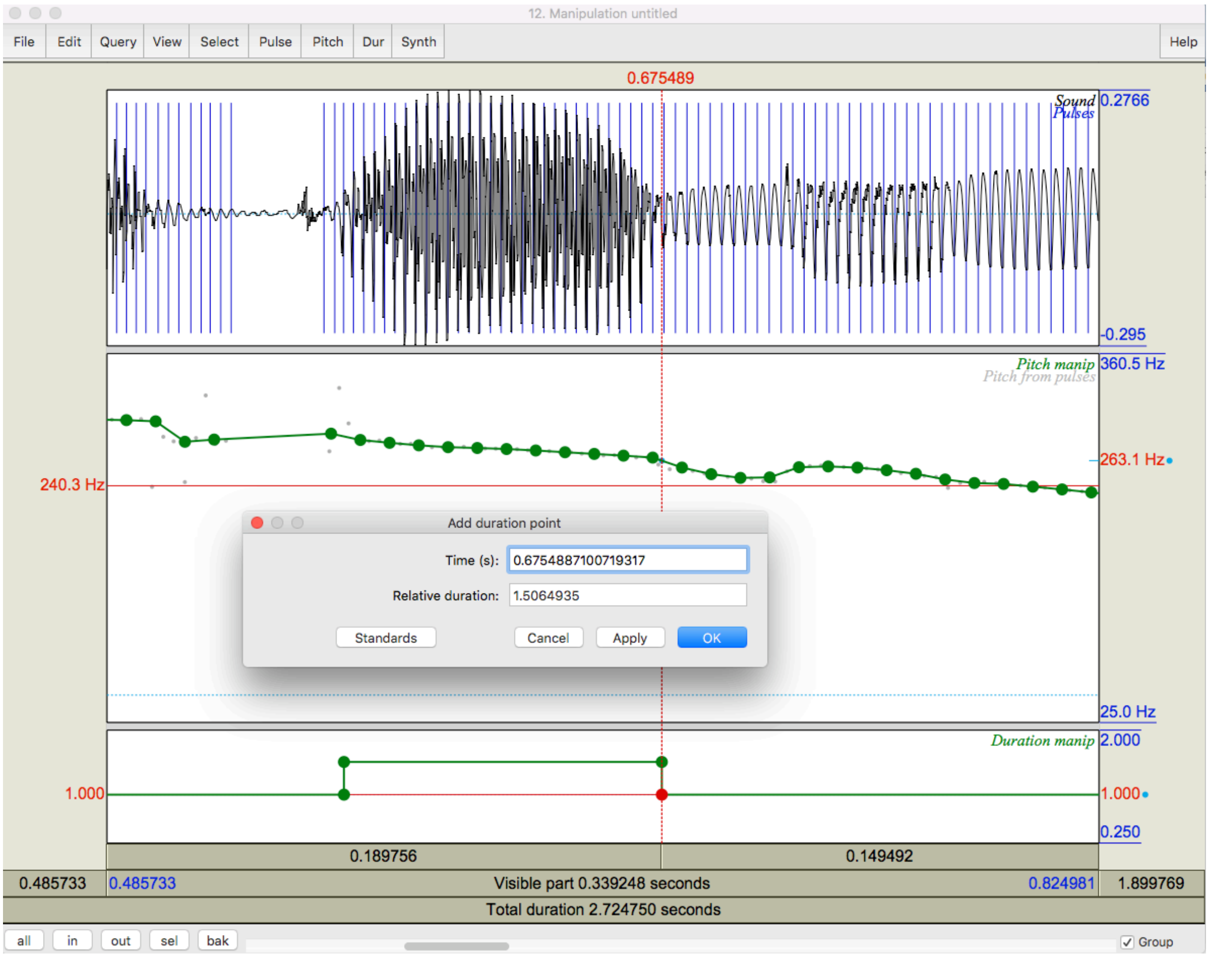


Figure 6.1: Example of duration manipulation of a spliced vowel.

Using the duration measurements of the spliced segment and of its counterpart in the host utterance, the relative duration factor was calculated. For example, if the spliced segment is 77ms and its counterpart in the host sentence is 116ms, the relative duration of the latter is 1.5064935. In the Manipulation Editor, a duration point was created at the onset and another at the offset of the segment. Then, another pair of duration points were created with the calculated relative duration, immediately inside of the two duration points by changing the very last digit of the value displayed after "Time (s)" by one, e.g. 8 to 7 in the example. This way, the spliced segment was lengthened or shortened to match the duration of the original segment in the recipient utterance.

In the final stage, the pitch contours of the utterances were flattened to the average F0 of the speaker of the utterance. In the Manipulation Editor, the existing pitch points were removed and two pitch points were added at the average frequency, one at the beginning and the other at the end of the utterance. While monotonising stimuli can be seen to undermine the ecological validity of the study, this was judged to be the best method of creating stimuli that are neutral between (a) and (b) (see Table 6.1) for the three varieties of Korean. This is particularly true in the light of the literature reviewed in Section 4.3 showing the important role played by pitch in signalling the AP boundary. The full experimental stimuli for the same-spliced and cross-spliced conditions are available from https://osf.io/8kmqz.

6.4.2 Participants

A total of 63 local university students from Seoul (n = 22), Busan (n = 19), and Ulsan (n = 22) were recruited in their respective universities – Hanyang University, Tongmyong University, and University of Ulsan – using emails and flyers. All participants were required to be aged between 20 and 28 and those who took part in the production study were not allowed to participate in the perception experiment. Ideally, participants should be as local to their cities as possible, because we do not know how subtle the regional variation may be, and how quickly and easily one's DIS pattern can be influenced through language and dialectal contact.

Before the experiment, the participants were given a short questionnaire to make sure they were local to their cities. The data from one participant from Seoul and two participants from Busan were excluded from the result analysis, as their answers to the questionnaire revealed that they grew up in another part of Korea. Another Ulsan speaker's data was lost due to a technical problem. Thus, the analysis reported in this chapter is based on 59 students in total whose age ranged from 18 to 27. The details of the participants used in the analysis are provided in Table 6.2.

Table 6.2: Participants for the perception experiment.

	Seoul	**Busan**	**Ulsan**	**Total**
No of participants	21	17	21	59
No of females	10	8	11	29
No of males	11	9	10	30
Mean age	23.5*	21.7*	20.6*	22.0**

* SD = 2.1 ** SD = 2.4

The participants spent all or a vast majority of their lives in the city of recording, except one speaker in the Seoul group who spent 13 years in Uijeongbu, a city just north of Seoul. Ten participants had lived abroad from four months to two and a half years, with only two of them spending longer than a year in a foreign country. Of these 10 participants, only one of them reported not to have spoken Korean every day or almost every day. The participants had no known hearing problem.

6.4.3 Procedure

The perception experiment was conducted in a quiet room in each of the local universities. The participants listened to the perceptual stimuli using Sennheiser HD 520 Headphones and my personal laptop. The experiment was in the form of two alternative forced-choice questions. A screenshot of the interface of the experiment is shown in Figure 6.2. Upon listening to either a same-spliced or cross-spliced stimulus based on type (a) or (b) sentence (see Table 6.1), the subject chose one of two answer choices, (a) or (b), presented visually in Korean orthography. As in the production study, the answers were marked with a possessive particle for (a) and a nominative particle for (b) in brackets to make the differences in meaning clear. In other words, these questions were designed to investigate whether the listeners' interpretation of the syntactically ambiguous utterance is influenced by the different DIS patterns at the AP and PW boundary. The stimulus for each question was played twice and all the stimuli were concatenated in a long file with a two second-silence after each stimulus to control the pace of the experiment. The concatenation procedure was automated with a script written in Python 3 using the Parselmouth library (Jadoul, Thompson, & de Boer, 2018) which provides an interface with Praat. This script is available from https://osf.io/8kmqz.

Prior to the perception experiment, I used two extra pairs of sentences which contained /s/ and /w/ as target consonants for the purpose of demonstration, as in the production study. With the help of simple hand-drawn visual aids, the participants were shown the situations in which each of the sentences would be appropriate in order to make sure they understand the differences in the two versions of the sentences. The participants were then given two practice questions based on those examples (see Figure 6.2) and another practice question based on an unseen sentence pair using the lenis stop /k/ taken from Part B of the recording materials (Appendix A). These questions used unmanipulated recordings of myself reading one of the sentence pairs. This ensured that the speakers understood the task. After the practice questions, the participants were told that the subsequent questions would involve "robotic voices" in order to inform them of the monotonicity of the stimuli without explicitly mentioning of pitch manipulation, as this might bias their responses. They were instructed to simply respond according to what came to mind as with a normal voice. They then listened to monotonised versions of the practice questions as illustration before proceeding to the main experiment.

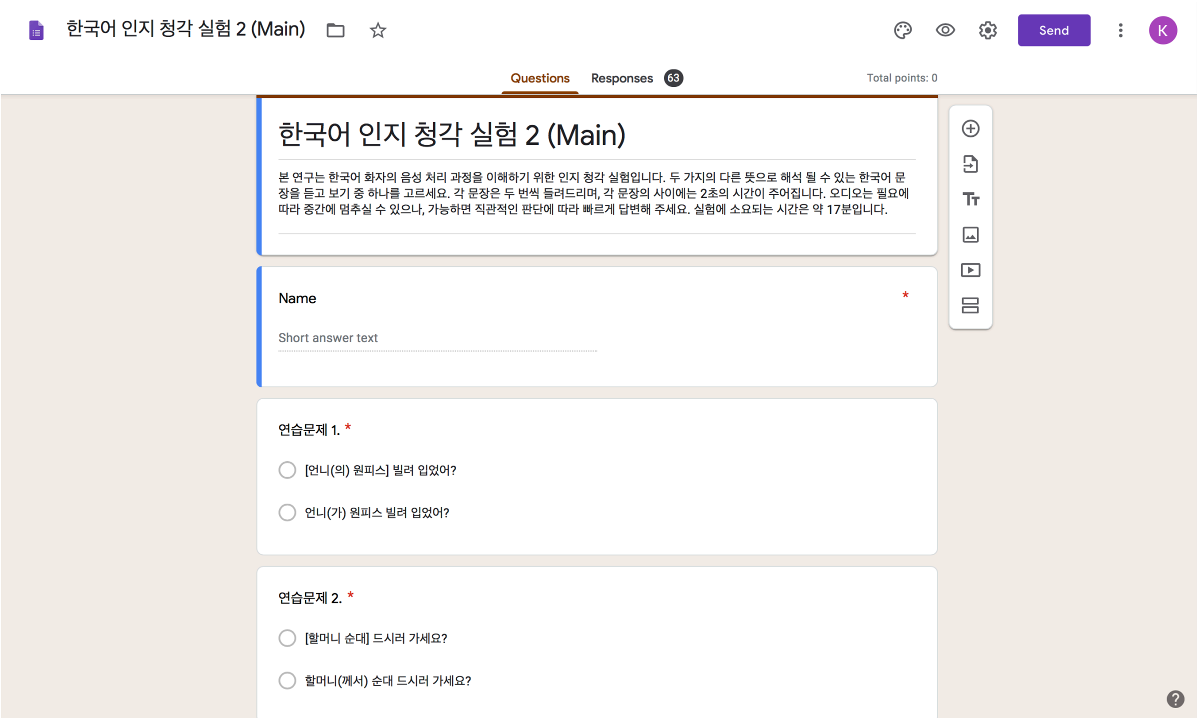


Figure 6.2: Screenshot of the experiment showing the title of the experiment (Korean auditory perception experiment 2), instructions, and the first two (practice) questions. The instructions are translated as follows: “This study is an auditory perception experiment designed to understand how speech signals are processed by Korean speakers. Listen to the following Korean sentences which can have two different interpretations, and select one of the two interpretations. Each sentence will be played twice with a two-second interval. While you may pause the audio file as needed, try to respond immediately following your intuitive judgement. This experiment will require approximately 17 minutes.”

There was a total of 120 questions (10 segmental strings $\times$ 4 versions – same & cross-spliced versions of (a) & (b) $\times$ 3 repetitions). As mentioned before, the stimulus for each question was played twice with a two-second interval. The three repetitions were pseudo-randomised in three blocks such that no two adjacent questions were based on the same segmental string. The entire duration of the experiment was approximately 17 minutes and there were 7080 observations in total (120 questions $\times$ 59 participants).

6.4.4 Statistical analysis

Similar to the method in Section 5.4.1, Bayesian mixed effect logistic models were fitted using the *brms* package in R. The R script and the raw data are available from https://osf.io/8kmqz. As before, a student-t distribution centred around zero with df = 3 and scale = 2.5 was used as the prior distribution for the intercepts, slopes, and standard deviations. These are weakly informative priors which assume the null hypothesis.

As the outcome variable, the binary variable *correct* was used, which indicated whether the listener's response was based on the interpretation of the host (recipient) sentence (correct) or not (incorrect). The predictor variables were *condition* (same-spliced, cross-spliced), *consonant* (t, n), *region* (Seoul, Busan, Ulsan), and *host sentence* (type (a) with a PW boundary, type (b) with an AP boundary). For the fixed effects, a four-way interaction term and all lower terms were included in order to test **H2** and **H3** which make predictions regarding high-order interaction effects. For the random effects, the maximal structure was used, including the random intercepts and all the logically possible random slopes by *subject* and *item*. The resulting model structure can be summarised with the following *brms* formula:

```
correct ~ condition * consonant * region * host_sentence + (condition
* consonant * host | subject) + (condition * region * host_sentence | item)
```

To interpret the results of the model with a complicated four-way interaction, the mean predicted probabilities for each combination of factor levels and their credible intervals (CI) were calculated manually using the posterior samples and the model estimates. There were 24 levels in total (2 conditions $\times$ 2 consonants $\times$ 3 regions $\times$ 2 host sentences).

6.5 Results

Figure 6.3 shows the predicted probabilities of correct answers in the same-spliced and cross-spliced conditions for listeners from Seoul, Busan, and Ulsan. As predicted, there is a clear visual pattern for a higher probability of correct answers in the same-spliced condition than in the cross-spliced condition, regardless of the variety, consonant type, and host sentence. Table 6.3 provides the differences between the same- and cross-spliced condition for the predicted probabilities of a correct answer and the CI intervals (for full results, see Appendix J). A 95% CI interval that does not include zero was regarded as "strong evidence" for a true effect of condition.

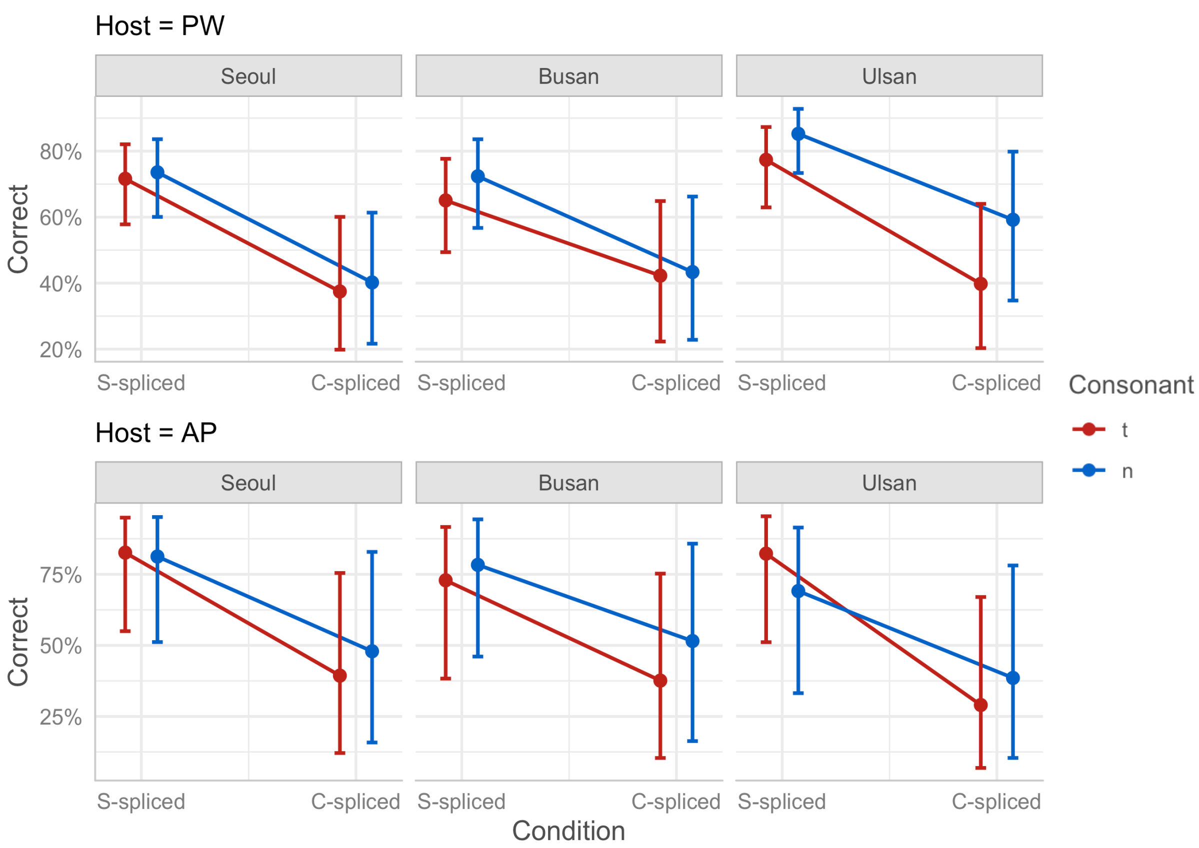


Figure 6.3: Predicted probabilities and 95% CIs of correct answers by experimental condition for the three varieties. The top panels are for the host sentences with a PW boundary (hereafter PW condition) and the bottom panels are for the host sentences with an AP boundary (hereafter AP condition). Red circles are for /t/ and blue triangles are for /n/.

As shown in Table 6.3, for 10 of the 12 possible combinations of the three factors – region, consonant, and host sentence – there was strong evidence for a difference between the same-spliced and cross-spliced condition, consistent with our prediction in **H1**. The two exceptions were (1) Busan subjects listening to /n/'s spliced to an AP-initial position and similarly (2) Ulsan subjects listening to /n/s spliced to an AP-initial position, highlighted with grey shading. The predicted probability that Busan listeners will answer correctly when /n/'s were *same*-spliced into AP-initial position was 0.713 with a 95% CI of [0.325, 0.944]. On the other hand, the probability that Busan listeners will answer correctly when /n/ was *cross*-spliced into AP-initial position was only 0.516 with a 95% CI of [0.163, 0.858]. Thus, the difference between the same- and cross-spliced condition was 0.197 [-0.149, 0.544]. While zero was included in the CI, 87.6% of the probability mass lay above zero. This is interpreted to mean that, with 87.6% probability, a correct answer is more likely in the same-spliced than in the cross-spliced condition.

Table 6.3: Differences between the same- and cross-spliced condition for the predicted probabilities of a correct answer. Grey shading indicates the two exceptions where strong evidence for a difference between the same- and cross-spliced condition was not found.

Consonant	**Region**	**Host-sentence**	**Mean**	**Lower Quantile (0.025)**	**Upper Quantile (0.975)**
t	Seoul	PW	0.33119	0.13278	0.50581
t	Seoul	AP	0.39733	0.11330	0.63915
t	Busan	PW	0.22070	0.00570	0.41830
t	Busan	AP	0.31614	0.00891	0.58374
t	Ulsan	PW	0.36388	0.14947	0.55413
t	Ulsan	AP	0.48516	0.19078	0.71142
n	Seoul	PW	0.32628	0.12283	0.50576
n	Seoul	AP	0.30923	0.01957	0.60230
n	Busan	PW	0.28056	0.06549	0.47626
n	Busan	AP	0.19736	-0.14884	0.54381
n	Ulsan	PW	0.26010	0.06903	0.46404
n	Ulsan	AP	0.27459	-0.04429	0.57419

Similarly, the predicted probability that Ulsan listeners will answer correctly when listening to nasals same-spliced into an AP-initial position was 0.675 [0.332, 0.915], whereas the probability that they will answer correctly in the cross-spliced condition was 0.400 [0.104, 0.781]. The difference between the two conditions was 0.275 [-0.044, 0.574]. In this case, 95.7% of the probability mass lay above zero, providing a greater level of certainty than in the Busan case above, that a correct answer is more likely in the same-spliced than in the cross-sliced condition.

Contrary to our prediction in **H2**, there was no strong statistical support for the interaction between *condition* and *host sentence*. For the lenis stops, the mean same-cross difference in the probability of a correct answer was greater in the AP condition than in the PW condition (Seoul: -0.066 [-0.317, 0.218], Busan: -0.095 [-0.363, 0.202], and Ulsan: -0.121 [-0.376, 0.175]). However, although the mean was consistently higher for the AP condition than for the PW condition, zero was well within the CI. Specifically, the results indicated 69.3%, 75.1%, and 81.8% probability that splicing a PW-initial /t/ into an AP-initial position has a stronger adverse effect for the perception of the host-sentence boundary than vice versa, for listeners from Seoul, Busan, and Ulsan, respectively.

For the nasal stops, the mean same-cross difference in the probability of a correct answer was greater in the PW than in the AP condition for Seoul and Busan, but not for Ulsan (Seoul: 0.017 [-0.276, 0.318], Busan: 0.083 [-0.341, 0.509], and Ulsan: -0.014 [-0.330, 0.309]). Again, zero was included in the CI with 54.6% and 65.4% of the probability mass lying above zero for Seoul and Busan, and 54.4% of the probability mass lying below zero for Ulsan. On the whole, there is no compelling evidence that the listeners perform differently depending on the host sentence. In addition, there was no strong support for the notion that consonant type may underlie the *condition* and *host-sentence* interaction, as zero was well within the CI for all three varieties.

Lastly, we can examine the interaction among *condition, consonant,* and *region* in relation to **H3**. Visually, for the lenis stops, Seoul and Ulsan appear to show stronger effects of *condition* compared to Busan, as the slopes for Seoul and Ulsan are steeper than that of Busan within the same host sentence. For the nasals, Seoul shows slightly steeper slopes compared to Busan and Ulsan.

To statistically validate these patterns, pairwise regional differences in the predicted probabilities were calculated for each of the four combinations of *consonant* type and *host sentence*. The results are presented in Table 6.4. I will focus the discussion on the regional differences for which the probability mass lying on either side of zero was greater than 90% which can be seen to provide some degree of confidence. For the lenis stops in the PW condition, the same-cross difference was greater for Seoul than for Busan (Seoul - Busan: 0.110 [-0.033, 0.249], with 93.9% of the probability mass lying above zero). This provides some support that cross-splicing a lenis stop is more misleading for the perception for Seoul listeners than for Busan listeners in the PW condition. Between Busan and Ulsan, the same-cross difference was smaller for Busan than for Ulsan (Busan - Ulsan: -0.143 [-0.306, 0.015], with 96.1% of the probability lying below zero). For the lenis stops in the AP condition, there is a 96.7% probability that the same-cross difference is greater for Ulsan than for Busan

(Busan - Ulsan: -0.169 [-0.352, 0.011). These results indicate that there is a 96.1% and a 96.7% probability that cross-splicing a lenis stop is more misleading for the perception for Ulsan listeners than for Busan listeners, in the PW and AP condition, respectively. For the nasal stops, there was no regional effect for which the probability mass lying on either side of zero was above 90%.

Table 6.4: Pairwise comparison of region for the same-cross condition differences in the predicted probabilities of a correct answer.

Consonant	Region	Host	Mean	Lower Quantile	Upper Quantile	Probability < 0	Probability > 0
t	Seoul - Busan	PW	0.11049	-0.03253	0.24894	0.0615	0.9385
t	Seoul - Ulsan	PW	-0.03268	-0.16397	0.10689	0.686	0.314
t	Busan - Ulsan	PW	-0.14318	-0.3058	0.01526	0.96125	0.03875
t	Seoul - Busan	AP	0.08119	-0.08198	0.24889	0.16575	0.83425
t	Seoul - Ulsan	AP	-0.08783	-0.24277	0.06394	0.87575	0.12425
t	Busan - Ulsan	AP	-0.16902	-0.3524	0.01077	0.9665	0.0335
n	Seoul - Busan	PW	0.04573	-0.09299	0.18615	0.2595	0.7405
n	Seoul - Ulsan	PW	0.06618	-0.07712	0.20221	0.17725	0.82275
n	Busan - Ulsan	PW	0.02045	-0.14173	0.17459	0.39675	0.60325
n	Seoul - Busan	AP	0.11187	-0.13498	0.41329	0.20625	0.79375
n	Seoul - Ulsan	AP	0.03465	-0.14039	0.2091	0.346	0.654
n	Busan - Ulsan	AP	-0.07723	-0.38483	0.21363	0.70425	0.29575

6.6 Discussion and conclusions

The perceptual experiment in this chapter was designed to answer the following research question: does DIS aid listeners in the demarcation of prosodic constituents? There were three hypotheses for the experiment. The first hypothesis (**H1**) was that, in the absence of key prosodic cues, listeners can use different DIS patterns at AP and PW boundaries to parse ambiguous utterances. Our results provide strong support for this hypothesis. For 10 out of the 12 combinations of the factor levels (*region* $\times$ *consonant* $\times$ *host sentence*), there was strong evidence that a correct answer was more likely in the same-spliced than in the cross-spliced condition. The two remaining groups were (1) Busan and (2) Ulsan subjects, listening to /n/'s spliced to an AP-initial position. For these groups, the posterior distribution indicated that there was 87.6% and 95.7% probabilities, respectively, that the listeners perform better in the same- rather than cross-spliced condition. These group differences will be discussed further with the third hypothesis (**H3**). Group differences aside, these results are interpreted to mean that having the correct DIS cues aids the detection of the correct prosodic boundary in general. This implies that prosodic boundaries are not only demarcated by prosodic cues including pitch and duration, but also by fine-grained phonetic variation in the segmental articulation.

However, there remains a possibility that other prosodic cues may have been spliced together with the target syllable, such as the amplitude- and the voice quality-related variation. This means that the differences in the same- and cross-spliced conditions may be due to these remaining prosodic cues rather than the strictly segmental variation at different prosodic positions. Although pitch and duration cues are known to be primary cues to a prosodic boundary, acoustic cues related to the voice source, such as harmonic structure and spectral tilt, as well as amplitude are indeed found to be relevant for prosodic boundary detection (J.-Y. Choi, Hasegawa-Johnson, & Cole, 2005). In addition, as the following vowel was spliced with the initial consonant, the potential cues in the vowel such as breathiness after a lenis stop in AP-initial position could also have benefited listeners in the boundary detection, and possibly more strongly for Seoul listeners (M.-R. Kim, 2017). Thus, further research is required to understand (a) the effect of these cues in boundary detection in Korean, and (b) whether DIS cues are indeed useful in boundary detection, when these prosodic cues have been isolated from strictly segmental cues such as voicing and nasality.

Next, our second hypothesis (**H2**) was as follows; for the lenis stops, cross-splicing a PW-initial CV to an AP-initial position is more misleading for the perception of the "correct" (PW-initial) boundary than vice versa. In contrast, for the nasal stops, cross-splicing an AP-initial CV to a PW-initial position is more misleading for the perception of the "correct" (PW-initial) boundary than vice versa. While the mean predicted probabilities were in the expected direction, there was no compelling statistical support for this hypothesis. For the lenis stops, the results did not provide reliable evidence. The posterior distribution showed 69.3%, 75.1%, and 81.8% probabilities that splicing a PW-initial /t/ into an AP-initial position is more misleading for the perception of the host sentence boundary than vice versa, for listeners from Seoul, Busan, and Ulsan, respectively. For the nasal stops, there was no support for an effect one way or another as the mean probabilities were close to zero, with 54.6% and 65.4% of the probability mass lying above zero for Seoul and Busan, and 54.4% of the probability mass lying below zero for Ulsan. We can therefore conclude that, in this study, the direction of splicing does not influence the listeners' ability to detect the boundary in the host sentence.

Note that these results differ from the finding by T. Cho et al. (2007), as reviewed in Section 4.4. In cross-modal identity-priming experiments with English listeners, they found an effect of the direction of splicing, such that DIS cues facilitated lexical access only when the host sentence of splicing involved a PW boundary and not when it involved an IP boundary. This was interpreted to mean that the benefit of DIS is most clear when other prosodic cues are weaker. Since there are more powerful cues to an IP boundary including a boundary tone and final lengthening, additional cues from DIS may not be so strong. On the other hand, splicing IP-initial CVs into PW-boundary contexts may significantly help segmentation as prosodic cues to PW-boundaries are relatively weak. In our case, however, all of the major prosodic cues from pitch and duration were eliminated. Therefore, our results in fact corroborate T. Cho et al. (2007)'s interpretation that the asymmetrical effect found in their study is due to other prosodic cues present in their stimuli, as such an effect was not found in this study where pitch and duration cues were removed from the stimuli. It is also worth pointing out that, while T. Cho et al. (2007) measured lexical decision reaction times to investigate the use of DIS patterns in lexical segmentation, this study directly tested the effect of DIS patterns on the comprehension of globally ambiguous sentences. Thus, while this study provides evidence that DIS patterns can be so powerful as to affect the semantic interpretation of an utterance, it may be more difficult for subtler effects to manifest in this kind of experimental set-up.

Finally, our last hypothesis (**H3**) was stated as follows: the effectiveness of DIS as a cue to a prosodic boundary depends on the goodness of match between the pattern in the perceptual input and the pattern in the listener's regional variety. We will first discuss the results for the lenis stops. Based on our production results in Chapter 5, we predicted that, for the lenis stops, the effectiveness of DIS as a perceptual cue will be stronger for Seoul than for Busan. With a lower degree of certainty, we also hypothesised that Ulsan may show an intermediate value between Seoul and Busan. While there was no strong indication, the results gave some supporting evidence for these hypotheses. In the PW condition, we can conclude that, with 93.9% certainty, the effect of DIS cues was stronger for Seoul than for Busan. Furthermore, the results showed a 96.1% probability that the effect was stronger for Ulsan than for Busan. In the AP condition, there was limited evidence that the effect was stronger for Seoul than for Busan (83.4% probability) and for Ulsan than for Busan (96.7% probability). Surprisingly, there was also limited indication that the effect of DIS cues was greater for Ulsan than for Seoul (87.6% probability), despite the fact that the perceptual stimuli were based on the recordings of two speakers from Seoul.

For the nasal stops, our prediction was that DIS cues would be more effective for Seoul listeners than for Busan or Ulsan listeners. No difference was predicted between Busan and Ulsan listeners. There was only limited indication towards these predictions. In the PW condition, the results indicated, with 82.3% certainty, that we can conclude that correct DIS cues were more likely to lead to a correct answer for Seoul listeners than for Ulsan listeners. In the AP condition, there was a 79.4% probability that the correct DIS cues were more helpful for Seoul than for Busan listeners. Overall, there is only limited evidence that DIS cues are more effective if there is a close match between the pattern in the perceptual input and the pattern in the listener's regional variety.

Although there was no strong evidence for regional differences, recall the results for the effect of condition reported at the beginning of this section. There were two out of 12 groups for which we failed to find a strong support for a reliable effect of condition: (1) Busan and (2) Ulsan subjects, listening to /n/'s spliced to an AP-initial position. This suggests that, for Busan and Ulsan speakers, it is not as helpful to hear a denasalised and devoiced nasal as opposed to a sonorant nasal in an AP-initial position, for the detection of an AP boundary. A possible explanation for this finding can be found in the results of our production study. In Section 5.6.4, we discussed the possibility that Seoul serves as the leader of the sound change, showing more denasalisation in AP-initial and IP-initial positions compared to Busan and Ulsan. Therefore, it is not surprising that Seoul listeners are better able to exploit denasalisation and devoicing of nasals as a cue to an AP boundary than Busan and Ulsan listeners.

Then, this raises the question as to why we did not observe the same pattern when /n/s were spliced into a PW-initial position. This may be because, for Busan and Ulsan listeners, a weaker degree of DIS in a higher prosodic position is more acceptable than a stronger degree of DIS in a lower prosodic position. This is based on the finding that DIS is cumulative and that the phonetic realisation of nasals is more limited in a PW-initial position, e.g. [n], whereas it is much more varied in an AP-initial position, e.g. [n $\sim$ t^h]. Thus, while the presence of a cross-spliced [t^h] in a PW-initial position strongly sways the perception towards an AP boundary, the presence of a cross-spliced [n] in an AP-initial position does not have the same effect. For Seoul listeners, however, the presence of a cross-spliced [n] in an AP-initial position also seems to weight the perception towards a PW boundary. This suggests that Seoul listeners treat a higher degree of DIS to be a crucial cue in perceiving an AP boundary.

While there were regional differences for the realisation of the lenis stops, too, the regional differences may perhaps have been better reflected in the perception of the nasal stops because denasalisation is a more recent phenomenon (Yoo, 2016). Our results in Sections 5.5.6 and 5.5.7 are also consistent with this view. For lenis stop (de-)voicing, a total of 17 out of 30 speakers showed evidence of stabilisation, with six from Seoul, five from Busan and six from Ulsan. For denasalisation, however, only four speakers showed evidence of stabilisation, with three from Seoul, one from Busan and none from Ulsan. However, more research is needed to establish the regional effect in the perception of DIS, as there was no compelling statistical support from this study.

Before closing the section, it is useful to acknowledge a limitation of this study in terms of its ecological validity. Because the perceptual stimuli which were monotonised and consisted of semantically ambiguous utterances, it remains unknown whether these conclusions can be extended to everyday listening situations. For example, it is possible that listeners hardly rely on DIS cues in their daily lives because of the availability of more powerful cues such as pitch and duration. Thus, further research is needed to investigate whether listeners use DIS cues for prosodic demarcation when they listen to unmodified speech.

6.7 Chapter summary

This chapter has presented a perception experiment conducted with a total of 59 students from Seoul, Busan, and Ulsan. The goal of this experiment was to test the speculation in the literature that DIS provides perceptual benefits to listeners. The study focused on the following question.

RQ5. Does DIS aid listeners in the demarcation of prosodic constituents?

While there is a previous finding that DIS plays a role in facilitating lexical access (T. Cho et al., 2007), little is known about the function of DIS in signalling a prosodic boundary in the prosodic parsing of syntactically ambiguous utterances.

The results of this experiment provided strong support for the main hypothesis that listeners exploit DIS cues to identify the left-edge of a prosodic domain. A novel finding from this study is that prosodic boundaries are not only signalled by the well-known prosodic cues including pitch and duration, but also by fine-grained phonetic variation in the segmental articulation. There was some evidence that the perceptual exploitation of these subtle phonetic cues is a highly sensitive process, such that listeners from different regions performed differently. As the perceptual stimuli were based on the recordings of Seoul speakers, there was a tendency for the effect to be stronger for Seoul listeners than for Busan and Ulsan listeners. However, as there was a lack of strong statistical support for this observation, further work is required to establish any effect of regional variety in the perception of DIS cues.

Chapter 7.

Conclusions and future directions

7.1 Conclusions

The dissertation has set out to achieve the three following research goals (Section 1.2) through a production study and a perception experiment. The first goal was to develop a consistent account of the prosody-sensitive variation of the lenis and nasal stops in domain-initial position. Closely related to this, the second goal was to investigate whether the lenis and nasal stop variation in Korean can be accounted for by the notion of rule scattering within the theory of the life cycle of phonological processes (Bermúdez-Otero, 2015). The three varieties spoken in Seoul, Busan, and Ulsan were included because previous research suggested that there was dialectal variation in the patterns of DIS. This potentially provided an opportunity to capture the relevant processes at different stages in the life cycle. Through such a case study, I ultimately sought to contribute to the understanding of the nature and the mechanism of the wider phenomenon of domain-initial strengthening (DIS). Finally, the third goal was to examine if DIS provides a perceptual benefit to listeners in terms of the demarcation of prosodic boundaries. Exploiting the three-city design, I also asked whether the effectiveness of DIS as a perceptual cue would depend on how well a listener's pattern of DIS matches that in the perceptual stimuli.

Following a detailed literature review and a production study, the dissertation proposed an account of the lenis and nasal stop variation in relation to the first goal. Despite the widespread view that the Korean lenis stops are underlyingly voiceless, it was concluded that the DIS patterns in Korean can be best captured if we assume the lenis stops to be underlying voiced, i.e. /b, d, g/. The similarities between the behaviours of the lenis and nasal stops under DIS were obvious from the investigation of the acoustic parameters such as closure duration and total voiceless interval during closure. With the assumption that the lenis stops are voiced, we can neatly account for this parallelism as follows. These stops share a tendency to show progressively greater devoicing, aspiration, and – additionally for the nasals – denasalisation, in a higher prosodic position. It has been pointed out that, treating the lenis stops to be underlyingly voiced also prevents Korean from being a cross-linguistic exception in having three voiceless plosives, e.g. /p, p*, p^h/.

Regarding the second goal, our production study has arrived at the conclusion that the three independently observed phenomena of lenis stop devoicing, denasalisation, and DIS can indeed be accounted for within the framework of the life cycle of phonological processes. Specifically, I put forward the view that lenis stop devoicing and denasalisation were categorical rules that had originated from DIS, a gradient phonetic process. Our results provided evidence that, for some speakers, lenis stop devoicing had stabilised into a phrase-level phonological rule operating at the level of the AP or the IP. For a smaller number of speakers, denasalisation was also found to be operating at the phrase-level phonological component, targeting AP-initial nasals. For other speakers, the lenis and nasal stop variation were most effectively modelled as a function of closure duration rather than as a function of discrete prosodic positions, suggesting that their pattern is a consequence of a gradient phonetic process. Crucially, for those speakers who showed a categorical pattern regarding lenis stop devoicing and/or denasalisation, there was also evidence of a simultaneous gradient effect. This is exactly what we had predicted based on the rule scattering hypothesis, which posits that a process operating at one component of the grammar (e.g. language-specific phonetics) may enter a higher-level component (e.g. phrase-level phonology) without ceasing to apply in the original component (e.g. Zsiga, 1995). One important advantage of this explanation is that the categorical vs. gradience debate in the literature for lenis stop voicing (T. Cho & Keating, 2001, Section 4.4.3; Docherty, 1995; S.-A. Jun, 1995c) and denasalisation (Young Shin Kim, 2011, Section 4.6; Yoshida, 2008) can be understood as stemming from the simultaneous existence of a categorical and a gradient effect.

These results also suggest that the nature of DIS is most convincingly understood as a gradient phonetic phenomenon which has its origins in the inherent properties of the workings of the speech apparatus. Among several possible mechanisms proposed for DIS (Fougeron & Keating, 1997, p. 3737), I have favoured the explanation based on increased articulatory energy in domain-initial position. An alternative view that DIS in Korean is a result of articulatory undershoot in a lower position where segments have shorter duration (T. Cho & Keating, 2001) has been shown to be inadequate in light of the pattern with the nasal stops. Denasalisation and devoicing of the nasal stops cannot be seen as a result of a mere full achievement of an intended articulatory target. An explanation grounded in phonetics has also been argued to be more desirable than the abstract end-based explanations of syntagmatic or paradigmatic contrast enhancement (e.g. T. Cho & Jun, 2000; Hsu & Jun, 1998). Not only does unconstrained application of the two enhancement mechanisms result in an unfalsifiable theory, these contrast-enhancement-based explanations fail to provide a consistent account of the full range of cross-linguistic phenomena reported in the literature.

While the origin of DIS can be seen as a consequence of some physical or physiological effect of speech production, it is possible that such an extra-grammatical effect is re-interpreted as a language-specific phonetic process, as predicted by the life cycle. This can potentially explain why languages exhibit different patterns of DIS for a given type of segment (e.g. T. Cho & McQueen, 2005). As a gradient phonetic process, DIS may be mediated directly or indirectly by the prosodic structure. We have discussed in Section 2.4.3 that these two scenarios are not mutually exclusive. Given the view that DIS is directly governed by the prosodic structure, one can explain the inconsistencies in the cumulative pattern of DIS as follows. While DIS is determined by the prosodic position of a segment, the articulatory target of a segment in a given position is specified as a range rather than a fixed point, following window models (Keating, 1990b; Guenther, 1995). This could explain why, for example, an AP-initial nasal can be realised with a greater degree of denasalisation than an IP-initial nasal.

Further along the life cycle, these language-specific processes can become stabilised as a categorical phonological rule, as we have suggested for lenis stop devoicing and denasalisation. An important implication of this is that, across different languages, we may find other phonological rules targeting domain-initial segments which have their origin in DIS. In addition, these rules as well as DIS itself are predicted to continue to travel along the path as laid out by the theory of the life cycle. If phrase-level phonological rules are found to operate at a specific prosodic domain, they are expected to show rule generalisation , applying in a wider range of prosodic environments (Bermúdez-Otero, 2007; Vennemann, 1972, pp. 186-187; Kiparsky, 1988, pp. 393-394).

In line with our expectations, the three varieties of Korean showed interesting variation with regard to their patterns of DIS, lenis stop devoicing, and denasalisation. The results of the various statistical analyses in our production study pointed to the pattern that speakers from Seoul and Ulsan represented a more advanced stage in the life cycle than those from Busan with regard to the stabilisation of lenis stop devoicing. This is assuming that rule generalisation proceeds from a narrower to a wider range of prosodic contexts. The domain of the lenis stop devoicing rule was shown to be the AP for Seoul and Ulsan Korean, whereas it was the IP for Busan Korean, at least for those who showed evidence of categoricity in Hartigan's dip test. The acoustic analyses also revealed that Seoul Koreans showed a greater magnitude of lenis stop variation under DIS than Busan Koreans, with some evidence that Ulsan Koreans showed an intermediate pattern between Seoul and Ulsan. For denasalisation of the nasal stops, only three Seoul speakers and one Busan speaker showed evidence of stabilisation, consistent with the view that denasalisation has emerged more recently than lenis stop devoicing. For all four speakers, the domain of categorical denasalisation was found to be the AP. Together with the acoustic analyses, our results corroborate the hypothesis that

Seoul is leading the sound change whereby DIS of the nasal consonants is stabilised into a categorical AP-initial denasalisation rule.

The overall pattern above only partially supports our prediction that the pattern of dialectal variation will follow Trudgill's (1974) model of hierarchical diffusion, which posits that the nearest large city will adopt a sound change before intervening smaller, more rural locations. While all three cities are relatively large metropolitan centres, Seoul has a population size that is approximately three times bigger than Busan, and, in turn, Busan has a population roughly three times bigger than Ulsan. Thus, even though Ulsan is geographically slightly closer to Seoul than Busan is to Seoul, denasalisation was expected to be more advanced in the order of Seoul, Busan, and Ulsan. Our results only provided evidence that Seoul shows a more advanced pattern than Busan and Ulsan, while no significant difference was found between Busan and Ulsan. In certain aspects, such as lenis stop devoicing, the pattern in Ulsan seemed to be intermediate between Seoul and Busan, going against the prediction based on the hierarchical diffusion. This is perhaps due to Busan maintaining a relatively strong dialectal identity in the midst of a general trend for dialect levelling across South Korea, following marginalisation of non-standard varieties of Korean (Silva, 2011; H.-M. Sohn, 1999).

For the final research goal regarding the perceptual role of DIS, our perception experiment provided strong evidence that listeners exploit DIS cues to detect the beginning of a prosodic domain. Our results also gave a limited indication that, as the experiment stimuli were based on speakers from Seoul, the DIS cues in these stimuli were more effective for Seoul listeners than for Busan and Ulsan listeners. Our findings have important implications for speech perception theories as they must provide a mechanism for how listeners are able to incorporate not only the widely-known prosodic cues such as duration and pitch but also fine-grained segmental variation in recovering the prosodic structure of an utterance (T. Cho et al., 2007).

Furthermore, we have mentioned earlier that denasalisation and/or devoicing of AP-initial nasals can neither be explained by accurate transmission of meaning nor effort reduction, which are often viewed as primary driving forces of language evolution (Labov, 2002; Xu & Prom-on, 2019). In terms of accurate transmission and phonological dispersion (Martinet, 1962), one could claim that the process has an opposite effect as Korean already has a crowded three-way system of contrastive plosives, all of which are voiceless in AP-initial position. The principle of least effort also predicts against denasalisation. In a higher prosodic position where denasalisation is most frequently found, there is often a preceding silence where the velum is lowered for breathing. Thus, denasalisation involves an extra velum raising gesture which cannot be considered a result of effort reduction. However, the results of our perception study have suggested that listeners potentially benefit from denasalisation and other DIS-related processes in a previously unseen way, namely the demarcation of a prosodic boundary in the perception of an ambiguous utterance. This implies that not only are listeners able to distinguish four different phonemes that are realised as a voiceless non-nasal in AP-initial position, they go even further and utilise fine-grained segmental variation to infer the prosodic position of a given phoneme.

7.2 Limitations and future directions

As summarised above, the dissertation has been able to explore some key research questions related to the production of DIS through a detailed auditory and acoustic investigation. What is not answered in our production study, however, is how far these acoustic analyses can be used to infer the subtle articulatory modifications at domain-initial position. It has been noted earlier that only articulatory data may be able to reveal certain hidden articulatory gestures, e.g. a tongue-tip gesture. Furthermore, what appears to be an abrupt change in the acoustics may result from a continuous articulatory movement. Crucially for our study, a gradual raising of the velum may not necessarily translate into a gradual decrease in the acoustic intensity. I suggested the possibility that at a certain point during a velum raising gesture, acoustic coupling between the oral tract and the nasal tract may be lost and show up on the spectrogram as an abrupt fall in intensity. This can lead to a false indication of a categorical behaviour when in reality the underlying articulatory pattern is gradient. In addition, I have inferred the degree of denasalisation through the decrease in the minimum acoustic intensity during the closure of a nasal phoneme. Strictly speaking, this measure reflects the combined degree of denasalisation and weakening of voicing. Thus, we cannot exclude the possibility that the speakers whom we have concluded show categorical denasalisation in fact only exhibit categorical weakening of voicing without any denasalisation.

In order to confirm or disconfirm the conclusions here, an articulatory study which tracks the velum movement more directly, such as Electromagnetic articulography (Perkell et al., 1992), Electromyography (Bell-Berti, 1976; Freitas et al., 2014), and Velotrace (Horiguchi & Bell-Berti, 1987) would be useful. Of course, an articulatory study is not without its own shortcomings, such as being invasive, time-consuming and/or expensive to conduct, which often leads researchers to compromise the scale of a study. Unfortunately, scale is important for any study that attempts to understand the general pattern in a given dialect, including the present study. Thus, it would be ideal if a comparable articulatory investigation could be conducted as a complement to our study.

Another limitation can be found in the ecological validity of our perception experiment. In order to eliminate the effect of important prosodic cues and isolate the effect of cues from segmental variation, I have used monotonised speech as the experimental material. The duration of the key segments at domain-initial position has also been manipulated to remove any durational cues to the prosodic boundary. In addition, the experiment was in the form of forced-choice questions which asked listeners to parse a semantically ambiguous sentence in the absence of any context. It is unlikely that listeners are faced with such a situation in their everyday lives. Thus, one might question whether our findings are generalisable to how listeners use DIS cues in everyday listening. Therefore, it would be interesting to test if the findings in our study could be replicated using unmodified speech as perceptual stimuli.

On the other hand, future research could be led in an opposite direction, towards controlling the experimental manipulation more precisely. While pitch and duration cues have been eliminated from the stimuli in our perception experiment, there may have been cues that were spliced along with the CV syllable. These cues include harmonic structure, spectral tilt, amplitude, and vowel breathiness, which may show systematic variation with regard to prosodic position. If these factors could be controlled more strictly, it would strongly corroborate our conclusion that segmental variation alone could signal a prosodic boundary.

Having discussed some limitations of the present study and how future studies could improve upon them, I will now turn to those areas that have not been addressed given the scope of our study but would greatly benefit from further research. First, there are various aspects of DIS which are poorly understood. For example, DIS of consonants is most often studied in the context of CV#CV. However, in order to distinguish among different proposals for the mechanism underlying DIS, it is important to understand the pattern of DIS in different contexts. For example, one of the proposed mechanisms is that DIS is a consequence of an overshoot caused by an increased articulatory distance between the domain-final vowel and the domain-initial consonant (Fougeron & Keating, 1997). This is based on the finding that domain-final vowels are usually more open and domain-initial consonants tend to be more closed. To test this hypothesis, it is crucial to study the behaviour of domain-initial consonants when preceded by a coda consonant and when the height of the adjacent vowels is varied systematically.

As the most plausible mechanism underlying DIS, we have discussed the view that there is increased articulatory effort at a higher domain-initial position which leads to stronger activation of the muscles used in a given segmental articulation (Fougeron & Keating, 1997). This suggestion can be empirically tested by varying the segmental context of a target segment, making specific predictions about the precise muscles that would be involved in articulating the target segment. For example, if /e/ is the target domain-initial segment, it is expected to be realised with a higher tongue position when the preceding domain-final segment is /a/ than when it is /i/. This is based on the assumption that a tongue raising gesture would be activated in the articulation of /a#e/ where a tongue lowering gesture would be involved in the articulation of /i#e/.

Another potentially fruitful area for future work is the factors that affect DIS other than those that have been studied here. While we have explored how factors such as prosodic position, consonant type, and dialectal variety influence the pattern of DIS, we are far from understanding the effect of style and rate of speech. In addition, we have mentioned earlier the possibility that speakers vary their pattern of DIS as part of social accommodation, a process by which a speaker adjust aspects of their speech according to the style of their interlocuter. To my knowledge, there is no previous research on the effect of social accommodation on DIS patterns.

Throughout this dissertation, I have worked under the assumption that, over time, a phonological rule operating in the phrase-level stratum may undergo rule generalisation within the stratum such that its domain of application would become more general, including a wider range of prosodic positions (Bermúdez-Otero, 2007; Vennemann, 1972, pp. 186-187). While our study could not directly test this assumption, a longitudinal or an apparent-time study of denasalisation in Korean may be able to capture an ongoing rule generalisation in Korean, assuming that denasalisation is still a sound change in progress (c.f. Yoo & Nolan, forthcoming).

Moving onto the perception of DIS, an interesting area for future research is understanding exactly how a listener is able to distinguish two types of prosodic boundary through the patterns of segmental variation. While the present study showed that listeners can distinguish between the AP and the PW boundary based on DIS cues, there is more than one mechanism that could allow this. As briefly discussed earlier, a listener may be able to achieve this relatively by associating a certain degree of DIS with a specific boundary type and inferring that a weaker DIS pattern indicates a lower boundary. In contrast, there may be a more precise mechanism in place by which a listener may recognise a certain DIS pattern as indicating one boundary type and another pattern, a different boundary type. Finally, it might be that a listener simply learns the range of DIS pattern present in the perceptual stimuli and distinguishes a stronger pattern from a weaker pattern purely by comparison. An additional topic to investigate is whether listeners are able to achieve the same task when listening to a speaker whose pattern of DIS has yet to be stabilised into a categorical phenomenon. Our study used perceptual stimuli based on two Seoul speakers' recordings, and both of them showed a categorical pattern for lenis stop devoicing. For denasalisation, only one of them showed evidence of categorical denasalisation. Thus, it is unknown whether we would have obtained the same results if our listening materials had been based only on those speakers who showed no evidence of stabilisation.

As a final suggestion for further work, one could investigate how a given individual's production pattern is reflected in their ability to exploit a particular DIS cue in perception. It is conceivable that listeners may be capable of exploiting certain patterns they are exposed in speech perception, even when they produce slightly different patterns or do not produce those patterns themselves. Conversely, one may produce certain patterns but fail to exploit them in perception. Our study has attempted to get closer to this understanding by examining whether a Seoul listener can use DIS cues more effectively than a Busan or Ulsan listener if the listening materials were drawn from another Seoul listener. However, we have discovered a wide range of individual variation even within a given region. Thus, a study that examines this on an individual level would be able to contribute further to our understanding of the nature of the link between speaker-listener's articulatory and perceptual repertoires, a relatively under-researched topic in Linguistics (Beddor, 2015).

In conclusion, this dissertation has explored DIS phenomena in Korean, a language which shows particularly interesting patterns in the prosody-sensitive variation of the lenis stops and nasal stops. Based the results of a production study and the theory of the life cycle of phonological processes (Bermúdez-Otero, 2015), this dissertation has proposed an account that captures both the striking similarity between the two types of stops and the previously unseen connection among the three phenomena of lenis stop voicing, denasalisation, and DIS. As the first dialectal survey of DIS, the production study also revealed fine-grained dialectal variation among Seoul, Busan, and Ulsan. The results suggested that Seoul is a more advanced variety than the other two in the sound change towards more extreme DIS patterns and phonological stabilisation of these patterns. Finally, the perception experiment in the dissertation provided evidence that listeners exploit DIS cues in the demarcation of prosodic boundaries. These findings have implications for the phonetic and phonological description of Korean, understanding of the nature of DIS, theories of linguistic change and diffusion, and theories of speech production and perception.

References

Ahn, H. (1999). *Post-release phonatory processes in English and Korean: Acoustic correlates and implications for Korean phonology* (Doctoral dissertation, University of Texas at Austin). Retrieved from http://www.riss.kr/link?id=T8445451

Atkinson, J. E. (1978). Correlation analysis of the physiological factors controlling fundamental voice frequency. *The Journal of the Acoustical Society of America*, *63*(1), 211–222. https://doi.org/10.1121/1.381716

Bang, H.-Y., Sonderegger, M., Kang, Y., Clayards, M., & Yoon, T.-J. (2015). The effect of word frequency on the time course of tonogenesis in Seoul Korean. *Proceedings of the 18th International Congress of Phonetic Sciences.*

Bang, H.-Y., Sonderegger, M., Kang, Y., Clayards, M., & Yoon, T.-J. (2018). The emergence, progress, and impact of sound change in progress in Seoul Korean: Implications for mechanisms of tonogenesis. *Journal of Phonetics*, *66*, 120–144. https://doi.org/10.1016/j.wocn.2017.09.005

Bates, D., Mächler, M., Bolker, B. M., & Walker, S. C. (2015). *lme4: Linear mixed-effects models using Eigen and S4.* https://cran.r-project.org/web/packages/lme4, R package version 1.1-8.

Beach, C. M. (1991). The interpretation of prosodic patterns at points of syntactic structure ambiguity: Evidence for cue trading relations. *Journal of Memory and Language, 30,* 644-663. https://doi.org/10.1016/0749-596X(91)90030-N

Beckman, M. E. (1996). The parsing of prosody. *Language and Cognitive Processes*, *11*(1–2), 17–67. https://doi.org/10.1080/016909696387213

Beckman, M. E., & Ayers, G. M. (1997). Guidelines for ToBI labelling. *The OSU Research Foundation, 3*, 30. Retrieved from http://www.cs.columbia.edu/~agus/tobi/labelling_guide_v3.pdf

Beckman, M. E., & Edwards, J. (1990). Lengthenings and shortenings and the nature of prosodic constituency. *Between the grammar and physics of speech: Papers in laboratory phonology I,* 152-178.

Beckman, M. E., & Edwards, J. (1994). Articulatory evidence for differentiating stress categories. In J. Kingston & M. E. Beckman (Eds.), *Papers in laboratory phonology I: Between the grammar and physics of speech* (pp. 152-193). Cambridge, UK: Cambridge University Press.

Beckman, M. E., Edwards, J., & Fletcher, J. (1992). Prosodic structure and tempo in a sonority model of articulatory dynamics. *Papers in Laboratory Phonology II*, 68–86. https://doi.org/10.1017/CBO9780511519918.004

Beckman, M. E., & Pierrehumbert, J. B. (1986). Intonational structure in Japanese and English. *Phonology Yearbook*, *3*, 255–309. https://doi.org/10.1017/S095267570000066X

Beddor, P. S. (2015). The relation between language users' perception and production repertoires. *Proceedings of the 18th International Congress of Phonetic Sciences*, 171–204.

Bell-Berti, F. (1976). An electromyographic study of velopharyngeal function in speech. *Journal of Speech and Hearing Research, 19*(2), 225–240. https://doi.org/10.1044/jshr.1902.225

Bell-Berti, F. (1993). Understanding velic motor control: Studies of segmental context. In M. K. Huffman & R. A. Krakow (Eds.), *Nasals, nasalization, and the velum* (Vol. 5, pp. 63–85). Academic Press. https://doi.org/10.1016/C2009-0-21373-0

Bermúdez-Otero, R. (2007). Word-final prevocalic consonants in English: Representation vs derivation. Paper presented at The Old World Conference in Phonology 4, Rhodes, Greece. Handout retrieved from http://www.bermudez-otero.com/OCP4.pdf

Bermúdez-Otero, R. (2010). *Morphologically conditioned phonetics? Not proven*. Paper presented at On Linguistic Interfaces II, Belfast, UK. Handout retrieved from http://www.bermudez-otero.com/Belfast_handout.pdf

Bermúdez-Otero, R. (2015). Amphichronic explanation and the life cycle of phonological processes. In P. Honeybone & J. C. Salmons (Eds.), *The Oxford handbook of historical phonology* (pp. 374–399). Oxford, UK: Oxford University Press. https://www.oxfordhandbooks.com/view/10.1093/oxfordhb/9780199232819.001.0001/oxfordhb-9780199232819-e-014

Bermúdez-Otero, R., & Trousdale, G. (2012). Cycles and continua: On unidirectionality and gradualness in language change. In T. Nevalainen & E. C. Traugott (Eds.), *The Oxford Handbook of the History of English* (pp. 691–720). Oxford, UK: Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199922765.013.0059

Blevins, J. (2006). A theoretical synopsis of Evolutionary Phonology. *Theoretical Linguistics*, *32*(2), 117–166. https://doi.org/10.1515/TL.2006.009

Boersma, P., & Weenink, D. (2018). *Praat, a system for doing phonetics by computer [Computer software].*

Browman, C. P. (1995). Assimilation as gestural overlap: Comments on Hoist and Nolan. In B. Connell & A. Arvaniti (Eds.), *Phonology and phonetic evidence: Papers in Laboratory Phonology IV* (pp. 334–342). Cambridge, UK: Cambridge University Press. https://doi.org/10.1017/CBO9780511554315.023

Browman, C. P., & Goldstein, L. M. (1986). Towards an articulatory phonology. *Phonology*, *3*, 219–252. https://doi.org/10.1017/S0952675700000658

Browman, C. P., & Goldstein, L. (1991). Gestural structures: Distinctiveness, phonological processes, and historical change. *Modularity and the Motor Theory of Speech Perception: Proceedings of a Conference to Honor Alvin M. Liberman*, 313–338.

Browman, C. P., & Goldstein, L. (1992). Articulatory phonology: An overview. *Phonetica*, *49*(3–4), 155–180. https://www.karger.com/Article/Abstract/261913

Bybee, J. L. (2002). Word frequency and context of use in the lexical diffusion of phonetically conditioned sound change. *Language Variation and Change*, *14*(3), 261–290. https://doi.org/10.1017/S0954394502143018

Byrd, D. (1994). *Articulatory timing in English consonant sequences* (Doctoral dissertation, University of California, Los Angeles, distributed as *UCLA Working Papers in Phonetics, 86*). Retrieved from http://escholarship.org/uc/item/4sr0s3jw

Byrd, D., & Riggs, D. (2008). Locality interactions with prominence in determining the scope of phrasal lengthening. *Journal of the International Phonetic Association*, *38*(2), 187–202. https://doi.org/10.1017/S0025100308003460

Chao, K.-Y., & Chen, L. (2008). A cross-linguistic study of voice onset time in stop consonant productions. *International Journal of Computational Linguistics & Chinese Language Processing, 13*(2), 215–232.

Chen, M. Y. (1997). Acoustic correlates of English and French nasalized vowels. *The Journal of the Acoustical Society of America*, *102*(4), 2360–2370. https://doi.org/10.1121/1.419620

Chen, M., & Clumeck, H. (1975). Denasalization in Korean: A search for universals. In C. A. Ferguson, L. M. Hyman, & J. J. Ohala (Eds.), *Nasálfest: Papers from a symposium on nasals and nasalization*, 125-131. Standford, CA: Stanford University Linguistics Dept.

Cho, S.-D., Whitman, J. (2019). Phonology and Phonetics. In *Korean: A Linguistic Introduction* pp. 63-95). Cambridge, UK: Cambridge University Press. doi:10.1017/9781139048842.005

Cho, S.-H. (2017). *Development of pitch contrast and Seoul Korean intonation* (Doctoral dissertation, University of Pennsylvania). Retrieved from https://repository.upenn.edu/edissertations/2226/

Cho, T. (2005). Prosodic strengthening and featural enhancement: Evidence from acoustic and articulatory realizations of/ɑ, i/in English. *The Journal of the Acoustical Society of America*, *117*(6), 3867–3878. https://doi.org/10.1121/1.1861893

Cho, T. (2016). Prosodic boundary strengthening in the phonetics-prosody interface. *Linguistics and Language Compass*, *10*(3), 120–141. https://doi.org/10.1111/lnc3.12178

Cho, T., & Jun, S.-A. (2000). Domain-initial strengthening as enhancement of laryngeal features: Aerodynamic evidence from Korean. *UCLA Working Papers in Phonetics*, 57–70.

Cho, T., Jun, S.-A., & Ladefoged, P. (2002). Acoustic and aerodynamic correlates of Korean stops and fricatives. *Journal of Phonetics*, *30*(2), 193–228. https://doi.org/10.1006/jpho.2001.0153

Cho, T., & Keating, P. A. (2001). Articulatory and acoustic studies on domain-initial strengthening in Korean. *Journal of Phonetics*, *29*(2), 155–190. https://doi.org/10.1006/jpho.2001.0131

Cho, T., & McQueen, J. M. (2005). Prosodic influences on consonant production in Dutch: Effects of prosodic boundaries, phrasal accent and lexical stress. *Journal of Phonetics*, *33*(2), 121–157. https://doi.org/10.1016/j.wocn.2005.01.001

Cho, T., McQueen, J. M., & Cox, E. A. (2007). Prosodically driven phonetic detail in speech processing: The case of domain-initial strengthening in English. *Journal of Phonetics*, *35*(2), 210–243. https://doi.org/10.1016/j.wocn.2006.03.003

Cho, Y. Y. (1987). Phrasal phonology of Korean. *Harvard Studies in Korean Linguistics*, *2*, 328–340.

Choe, H. (1937). *우리말본 [Urimalbon]*. Seoul, South Korea: 정음사 [Cengumsa].

Choi, H. (2002). Acoustic cues for the Korean stop contrast: Dialectal variation. *ZAP Papers in Linguistics, 28,* 1-12.

Choi, J.-Y., Hasegawa-Johnson, M., & Cole, J. (2005). Finding intonational boundaries using acoustic cues related to the voice source. *The Journal of the Acoustical Society of America, 118*(4), 2579-2587. https://doi.org/10.1121/1.2010288

Choi, T.-H., Lim, N.-S., & Han, J.-I. (2006). Effects of word frequency on a lenition process: Evidence from stop voicing and /h/ reduction in Korean. *음성과학 [The Korean Journal of Speech Sciences]*, 13(3), 35-48.

Christophe, A., Peperkamp, S., Pallier, C., Block, E., & Mehler, J. (2004). Phonological phrase boundaries constrain lexical access I: Adult data. *Journal of Memory and Language, 51*(4), 523–547. https://doi.org/10.1016/j.jml.2004.07.001

Chung, K., Chang, S., Choi, J., Nam, S., Lee, M., Chung, S., Koo, H., Kim, K., Kim, J., Lee, C., & Han, S. (1996). A study of Korean prosody and discourse for the development of speech synthesis/recognition system. *Korea Telecom Research and Development Group*.

Dart, S. N. (1987). An aerodynamic study of Korean stop consonants: Measurements and modeling. *The Journal of the Acoustical Society of America*, *81*(1), 138–147. https://doi.org/10.1121/1.395023

Davidson, L. (2016). Variability in the implementation of voicing in American English obstruents. *Journal of Phonetics, 54*, 35–50. https://doi.org/10.1016/j.wocn.2015.09.003

Davis, M. H., Marslen-Wilson, W. D., & Gaskell, M. G. (2002). Leading up the lexical garden path: Segmentation and ambiguity in spoken word recognition. *Journal of Experimental Psychology: Human Perception and Performance, 28*(1), 218-244. https://doi.org/10.1037/0096-1523.28.1.218

de Jong, K. J. (1995). The supraglottal articulation of prominence in English: Linguistic stress as localized hyperarticulation. *The Journal of the Acoustical Society of America, 97*(1), 491–504. https://doi.org/10.1121/1.412275

de Pijper, J. R., & Sanderman, A. A. (1994). On the perceptual strength of prosodic boundaries and its relation to suprasegmental cues. *The Journal of the Acoustical Society of America*, *96*(4), 2037–2047. https://doi.org/10.1121/1.410145

Docherty, G. J. (1995). On a gestural account of lenis stop voicing in Korean: Comments on Jun. *Phonology and phonetic evidence: Papers in Laboratory Phonology IV* (pp. 254-264). https://doi.org/10.1017/CBO9780511554315.018

Domingos, P. (1999). The role of Occam's Razor in knowledge discovery. *Data Mining and Knowledge Discovery, 3,* 409-425. https://doi.org/10.1023/A:1009868929893

Eckert, P. (1989). The whole woman: Sex and gender differences in variation. *Language variation and change, 1*(3), 245-267. https://doi.org/10.1017/S095439450000017X

Edwards, J., & Beckman, M. E. (1988). Articulatory timing and the prosodic interpretation of syllable duration. *Phonetica*, *45*(2–4), 156–174.

Ellis, L., & Hardcastle, W. J. (2002). *Categorical and gradient properties of assimilation in alveolar to velar sequences: Evidence from EPG and EMA data.* http://doi:10.1006/jpho.2001.0162

Faaborg-Andersen, K. K. (1957). *Electromyographic investigation on intrinsic laryngeal muscles in humans; an investigation of subjects with normally movable vocal cords and patients with vocal cord paresis* (Doctoral dissertation). University of Copenhagen, Copenhagen, Denmark.

Fischer-Jorgensen, E., & Hutters, B. (1981). Aspirated stop consonants before low vowels, a problem of delimitation: Its causes and consequences. *Annual Reports of the Institute of Phonetics, University of Copenhagen (ARIPUC)*, *15*, 77–102.

Fougeron, C. (1998). *Variations articulatoires en début de constituants prosodiques de différents niveaux en français [Articulatory variations at the start of prosodic constituents of different levels]* (Doctoral dissertation, University of Paris III, Paris, France). Retrieved from http://lpp.in2p3.fr/IMG/pdf/thesecfougeron-nonve_rifie_e.pdf

Fougeron, C. (2001). Articulatory properties of initial segments in several prosodic constituents in French. *Journal of Phonetics*, *29*(2), 109–135. https://doi.org/10.1006/jpho.2000.0114

Fougeron, C., & Jun, S.-A. (1998). Rate effects on French intonation: Prosodic organization and phonetic realization. *Journal of Phonetics, 26*: 45.69. https://doi.org/10.1006/jpho.1997.0062

Fougeron, C., & Keating, P. A. (1997). Articulatory strengthening at edges of prosodic domains. *Journal of the Acoustical Society of America*, *101*(6), 3728–3740. https://doi.org/10.1121/1.418332

Fowler, C. A., & Saltzman, E. (1993). Coordination and coarticulation in speech production. *Language and Speech*, *36*(2–3), 171–195. https://doi.org/10.1177/002383099303600304

Fox, J., & Weisberg, S. (2019). *An R companion to applied regression* (3rd ed.). Thousand Oaks, CA: Sage. Available from https://socialsciences.mcmaster.ca/jfox/Books/Companion/ https://doi.org/10.1080/10543406.2012.635980

Freeman, J. B., & Dale, R. (2013). Assessing bimodality to detect the presence of a dual cognitive process. *Behavior Research Methods*, *45*(1), 83–97. https://doi.org/10.3758/s13428-012-0225-x

Freitas, J., Teixeira, A., Silva, S., Oliveira, C., & Dias, M. S. (2014). Velum movement detection based on surface Electromyography for speech interface. *Proceedings of the International Conference on Bio-inspired Systems and Signal Processing*, 13–20.

Fujimura, O. (1990). Methods and goals of speech production research. *Language and Speech*, *33*(3), 195–258. https://doi.org/10.1177/002383099003300301

Futrell, R., Mahowald, K., & Gibson, E. (2015). Large-scale evidence of dependency length minimization in 37 languages. *Proceedings of the National Academy of Sciences*, *112*(33), 10336–10341.

Garellek, M., Ritchart, A., & Kuang, J. (2016). Breathy voice during nasality: A cross-linguistic study. *Journal of Phonetics*, *59*, 110–121. https://doi.org/10.1016/j.wocn.2016.09.001

Gay, T., Strome, M., Hirose, H., & Sawashima, M. (1972). Electromyography of the intrinsic laryngeal muscles during phonation. *Annals of Otology, Rhinology & Laryngology*, *81*(3), 401–409. https://doi.org/10.1177/000348947208100311

Gelman, A. (2019, Dec 30). *Prior choice recommendations*. Retrieved from https://github.com/stan-dev/stan/wiki/Prior-Choice-Recommendations

Gelman, A., & Loken, E. (2013). The garden of forking paths: Why multiple comparisons can be a problem, even when there is no "fishing expedition" or "p-hacking" and the research hypothesis was posited ahead of time. *Pyshological Bulletin, 140,* 1272-1280. https://doi.org/10.1037/a0037714

Gelman, A., Simpson, D., & Betancourt, M. (2017). The prior can often only be understood in the context of the likelihood. *Entropy*, *19*(10), 555. https://doi.org/10.3390/e19100555

Georgeton, L., & Fougeron, C. (2014). Domain-initial strengthening on French vowels and phonological contrasts: Evidence from lip articulation and spectral variation. *Journal of Phonetics*, *44*, 83–95. https://doi.org/10.1016/j.wocn.2014.02.006

Gim, S.-G. (1937). *The phonetics of the Korean Language*. Master's thesis in comparative philology probably supervised by Daniel Jones. One of the two photocopies prepared in March 2009 from the original held at the Senate House Library in London with permission from the author's next of kin, Bohi Gim. London, UK: UCL.

Gordon, M. (1997). Phonetic correlates of stress and the prosodic hierarchy in Estonian. In I. Lehiste & J. Ross (Eds.) *Estonian prosody: Papers from a symposium*, 100–124. Retrieved from http://gordon.faculty.linguistics.ucsb.edu/estonian.pdf (pp. 1-25)

Gordon, M., & Ladefoged, P. (2001). Phonation types: A cross-linguistic overview. *Journal of Phonetics*, *29*(4), 383–406. https://doi.org/10.1006/jpho.2001.0147

Gout, A., Christophe, A., & Morgan, J. L. (2004). Phonological phrase boundaries constrain lexical access II: Infant data. *Journal of Memory and Language*, *51*(4), 548–567. https://doi.org/10.1016/j.jml.2004.07.002

Gow, D. W., Melvold, J., & Manuel, S. (1996). How word onsets drive lexical access and segmentation: Evidence from acoustics, phonology and processing. *Proceedings of the international conference on spoken language processing,* 66–69. https://doi.org/10.1109/ICSLP.1996.607031

Guenther, F. H. (1995). Speech sound acquisition, coarticulation, and rate effects in a neural network model of speech production. *Psychological Review*, *102*(3), 594–621. Retrieved from https://hdl.handle.net/2144/2150

Hansson, G. Ó. (2008). Diachronic explanations of sound patterns. *Language and Linguistics Compass, 2*(5), 859-893. https://doi.org/10.1111/j.1749-818x.2008.00077.x

Gussenhoven, C. (2004). Paralinguistics: Three biological codes. *The phonology of tone and intonation* (pp. 71-96). Cambridge, UK: Cambridge University Press. https://doi.org/10.1017/CBO9780511616983.006

Gussenhoven, C., Rietveld, T., Kerkhoff, J., & Terken, J. (1999). ToDI: Transcription of Dutch intonation. Retrieved from http://todi.science.ru.nl/

Halle, M., & Stevens, K. (1971). A note on laryngeal features. *MIT-RLE Quarterly Progress Report*, *101*, 198–213.

Han, J.-I. (1996). *The phonetics and phonology of "tense" and "plain" consonants in Korean* (Doctoral dissertation, Cornell University). Retrieved from http://www.riss.kr/search/detail/DetailView.do?p_mat_type=be54d9b8bc7cdb09&control_no=d9ec3df8b0ecb9b6

Han, M. (2011). *한국어의 어두 경음화 현상 연구 [A study on the word-initial glottalization in Korean]* (Doctoral dissertation, Konkuk University, Seoul, South Korea). Retrieved from http://www.riss.kr/search/detail/DetailView.do?p_mat_type=be54d9b8bc7cdb09&control_no=e6bf65598b433d12ffe0bdc3ef48d419

Han, M. S., & Weitzman, R. S. (1970). Acoustic features of Korean /P, T, K/, /p, t, k/ and /p^h, t^h, k^h/. *Phonetica*, *22*(2), 112–128. https://doi.org/10.1159/000259311

Hardcastle, W. J. (1973). Some observations on the *tense-lax* distinction in initial stops in Korean. *Journal of Phonetics*, *1*(3), 263–272. https://doi.org/10.1016/S0095-4470(19)31390-7

Hardcastle, W. J. (1976). *Physiology of speech production: An introduction for speech scientists*. Cambridge, MA: Academic Press.

Hartigan, J. A., & Hartigan, P. M. (1985). The dip test of unimodality. *The Annals of Statistics*, *13*(1), 70–84. https://doi.org/10.1214/aos/1176346577

Herd, W., Jongman, A., & Sereno, J. (2010). An acoustic and perceptual analysis of /t/ and /d/ flaps in American English. *Journal of Phonetics*, *38*(4), 504–516. https://doi.org/10.1016/j.wocn.2010.06.003

Hirose, H., Lee, C. Y., & Ushijima, T. (1974). Laryngeal control in Korean stop production. *Journal of Phonetics*, *2*(2), 145–152. https://doi.org/10.1016/S0095-4470(19)31189-1

Hlavac, M. (2018). *stargazer: Well-formatted regression and summary statistics tables*. https://CRAN.R-project.org/package=stargazer, R package version 5.2.1.

Holliday, J. J., & Kong, E. J. (2011). Dialectal variation in the acoustic correlates of Korean stops. *Proceedings of the 17th International Congress of Phonetic Sciences,* 878–881.

Horiguchi, S., & Bell-Berti, F. (1987). The Velotrace: A device for monitoring velar position. *The Cleft Palate Journal*, *24*(2), 104–111.

Hsu, C.-S., & Jun, S.-A. (1996). Is Tone Sandhi Group part of the prosodic hierarchy in Taiwanese? *The Journal of the Acoustical Society of America*, *100*(4), 2824. https://doi.org/10.1121/1.416637

Hsu, C.-S., & Jun, S.-A. (1998). Prosodic strengthening in Taiwanese: Syntagmatic or paradigmatic? *UCLA Working Papers in Phonetics*, 69–89. Retrieved from https://escholarship.org/uc/item/2m2416qp

Huh, W. (1965). *국어음운론 [Korean Phonology]*. Seoul, South Korea: 정음사 [Cengumsa].

Hume, E. (1990). Front vowels, palatal consonants and the rule of umlaut in Korean. *Proceedings of North Eastern Linguistics Society 20,* 230-243.

Im, S. K. (2014). 성조 현상을 토대로 한 방언구획론을 위하여 [A study on the dialectal division based on tonal processes of Kyeongsang-dialect]. *방언학 [The Journal of Korean Dialectology]*, *19,* 5–40. Retrieved from https://www.kci.go.kr/kciportal/ci/sereArticleSearch/ciSereArtiView.kci?sereArticleSearchBean.artiId=ART001891513

Imagawa, H., Kiritani, S., Masaki, S., & Shirai, K. (1985). Contextual variation in the jaw position for the vowels in /CVC/ utterances. *Annual Bulletin of the Research Institute of Logopedics and Phoniatrics*, *19*, 7–19.

International Phonetic Association. (1949). The principles of the International Phonetic Association. *Journal of the International Phonetic Association, 40*(3). https://doi.org/10.1017/S0025100311000089

Iverson, G. K. (1983). On glottal width features. *Lingua*, *60*(4), 331–339. https://doi.org/10.1016/0024-3841(83)90052-9

Jacewicz, E., Fox, R. A., & Lyle, S. (2009). Variation in stop consonant voicing in two regional varieties of American English. *Journal of the International Phonetic Association, 39*(3), 313–334. https://doi.org/10.1017/S0025100309990156

Jadoul, Y., Thompson, B., & de Boer, B. (2018). Introducing Parselmouth: A Python interface to Praat. *Journal of Phonetics*, *71*, 1–15. https://doi.org/10.1016/j.wocn.2018.07.001

Jang, J., Kim, S., & Cho, T. (2018). Focus and boundary effects on coarticulatory vowel nasalization in Korean with implications for cross-linguistic similarities and differences. *The Journal of the Acoustical Society of America*, *144*(1), EL33–EL39. https://doi.org/10.1121/1.5044641

Jang, S. (2017). *현대 한국어의 어두 경음화 연구 [Word-initial tensification in Modern Korean]* (Master's thesis, Seoul National University, Seoul, South Korea). Retrieved from http://s-space.snu.ac.kr/handle/10371/132177

Jang, Y. (2008). A phonological study on the place-names of Ulsan region. *신라문화 [The Journal of the Research Institute for Silla Culture]*, *32,* 359–388.

Jeon, H.-S. (2010). A phonetic study on phrasing in Seoul Korean. In J. Yeon & J. Kiaer (Eds.)*, Selected Papers from 2nd European Conference on Korean Linguistic,* 88-101. Munich: LINCOM Europa.

Jeon, H.-S., & Nolan, F. (2010). Segmentation of the Accentual Phrase in Seoul Korean. *Proceedings of Speech Prosody*, *100023,* 1-4.

Jeon, H.-S., & Nolan, F. (2013). The role of pitch and timing cues in the perception of phrasal grouping in Seoul Korean. *Journal of the Acoustical Society of America, 133,* 3039-3049. https://doi.org/10.1121/1.4798663

Jeon, H.-S., & Nolan, F. (2017). Prosodic marking of narrow focus in Seoul Korean. *Laboratory Phonology: Journal of the Association for Laboratory Phonology, 8*(1), 2. https://doi.org/10.5334/labphon.48

Jones, D. (1924). spesimɛn kəriən [The specimen of Korean]. *lə mɛːtrə fɔnetik ɔrgan də l asɔsjaːsjɔ̃ fɔnetik ɛ̃ːtɛrnasjɔnal [Journal of the International Phonetic Association]*, 14–15.

Jun, J., Kim, J., Kim, J., Lee, H., & Jun, S.-A. (2006). The prosodic structure and pitch accent of Northern Kyungsang Korean. *Journal of East Asian Linguistics, 15*(4), 289–317. https://doi.org/10.1007/s10831-006-9000-2

Jun, S.-A. (1993). *The phonetics and phonology of Korean prosody* (Doctoral dissertation, Ohio State University). Retrieved from https://etd.ohiolink.edu/!etd.send_file?accession=osu1220465077&disposition=inline

Jun, S.-A. (1995a). *An acoustic study of Korean stress*. Poster presented at the 130th meeting of Acoustical Society of America, St. Louis, MO.

Jun, S.-A. (1995b). A phonetic study of stress in Korean. *The Journal of the Acoustical Society of America*, *98*(5), 2893–2893. https://doi.org/10.1121/1.414317

Jun, S.-A. (1995c). Asymmetrical prosodic effects on the laryngeal gesture in Korean. In B. Connell (Ed.), *Phonology and phonetic evidence: Papers in Laboratory Phonology IV* (pp. 235–253). https://doi.org/10.1017/CBO9780511554315.017

Jun, S.-A. (1996). *The phonetics and phonology of Korean prosody* (1st ed.). London, UK: Routledge. https://doi.org/10.4324/9780429454943

Jun, S.-A. (1998). The Accentual Phrase in the Korean prosodic hierarchy. *Phonology*, *15*(2), 189–226. https://doi.org/10.1017/s0952675798003571

Jun, S.-A. (2000). K-ToBI (Korean ToBI) labeling conventions. *Korean Journal of Speech Science, 7, 143-169.* Retrieved from https://linguistics.ucla.edu/people/jun/ktobi/k-tobi.html

Jun, S.-A. (2005a). Korean intonational phonology and prosodic transcription. In S.-A. Jun (Ed.), *Prosodic typology: The phonology of intonation and phrasing* (pp. 201–229).

Oxford, UK: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199249633.001.0001

Jun, S.-A. (2005b). Prosody in sentence processing: Korean vs. English. *UCLA Working Papers in Phonetics*, *104*, 26–45.

Jun, S.-A. (2006). Intonational phonology of Seoul Korean revisited. *UCLA Working Papers in Phonetics*, *104*, 14–25.

Jun, S.-A., Beckman, M. E., & Lee, H.-J. (1998). Fiberscopic evidence for the influence on vowel devoicing of the glottal configurations for Korean obstruents. *UCLA Working Papers in Phonetics*, 43–68.

Kagaya, R. (1974). A fiberscopic and acoustic study of the Korean stops, affricates and fricatives. *Journal of Phonetics*, *2*, 161–180. https://doi.org/10.1016/S0095-4470(19)31191-X

Kang, H., & Hyo-Young Lee. (2019). The Realizations of /h/ in Seoul and Gwangju Koreans. *언어학 [The Linguistic Association of Korea Journal]*, 84, 175-198.

Kang, K.-H., & Guion, S. G. (2008). Clear speech production of Korean stops: Changing phonetic targets and enhancement strategies. *The Journal of the Acoustical Society of America*, *124*(6), 3909–3917. https://doi.org/10.1121/1.2988292

Kang, Y. (2016). A corpus-based study of positional variation in Seoul Korean vowels. In M. Kenstowicz, T. Levin, & R. Masuda (Eds.), *Japanese/Korean Linguistics* (Vol. 23, pp. 3–22). Stanford, CA: CSLI.

Kang, Y., & Han, S. (2013). Tonogenesis in early Contemporary Seoul Korean: A longitudinal case study. *Lingua*, *134*, 62–74. https://doi.org/10.1016/j.lingua.2013.06.002

Kassambara, A. (2019). *ggpubr: “ggplot2” Based Publication Ready Plots*. https://CRAN.R-project.org/package=ggpubr, R package version 0.2.5.

Kawahara, S. (2010). *Get intensity min max [Praat script]*. Retrieved from http://user.keio.ac.jp/~kawahara/scripts/get_intensity_minmax.praat

Keating, P. A. (1984). Phonetic and phonological representation of stop consonant voicing. *Language, 60*(2), 286–319. https://doi.org/10.2307/413642

Keating, P. A. (1990a). Phonetic representations in a generative grammar. *Journal of Phonetics*, *18*(3), 321–334. https://doi.org/10.1016/S0095-4470(19)30377-8

Keating, P. A. (1990b). *The window model of coarticulation: Articulatory evidence*. In J. Kingston & M. E. Beckman (Eds.), *Papers in Laboratory Phonology* (pp. 451-470). Cambridge, UK: Cambridge University Press. https://doi.org/10.1017/CBO9780511627736.026

Keating, P. A., Cho, T., Fougeron, C., & Hsu, C.-S. (2004). Domain-initial strengthening in four languages. In J. Local, R. Ogden, & R. Temple (Eds.), *Phonetic Interpretation: Papers in Laboratory Phonology VI* (pp. 145–163). Cambridge, UK: Cambridge University Press. https://doi.org/10.1017/CBO9780511486425.009

Keating, P. A., & Esposito, C. (2006). Linguistic voice quality. *UCLA Working Papers in Phonetics*, *105*, 85–91.

Keating, P. A., Esposito, C., Garellek, M., Khan, S., & Kuang, J. (2010). Phonation contrasts across languages. *UCLA Working Papers in Phonetics, 108,* 188-202.

Keating, P. A., Lindblom, B., Lubker, J., & Kreiman, J. (1994). Variability in jaw height for segments in English and Swedish VCVs. *Journal of Phonetics*, *22*(4), 407–422. https://doi.org/10.1016/S0095-4470(19)30293-1

Keating, P. A., Wright, R., & Zhang, J. (1999). Word-level asymmetries in consonant articulation. *UCLA Working Papers in Phonetics*, *97*, 157–173.

Kenstowicz, M., & Park, C. (2006). Laryngeal features and tone in Kyungsang Korean: A phonetic study. *음성음운형태론 연구 [Studies in Phonetics, Phonology, and Morphology]*, *12*(2), 247–264.

Kessinger, R. H., & Blumstein, S. E. (1997). Effects of speaking rate on voice-onset time in Thai, French, and English. *Journal of Phonetics*, *25*(2), 143–168. https://doi.org/10.1006/jpho.1996.0039

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. https://doi.org/10.1080/00437956.1965.11435434

Kim, C.-W. (1970). A theory of aspiration. *Phonetica*, *21*(2), 107–116. https://doi.org/10.1159/000259293

Kim, H.-K. (1972). *국어학개론 [An introduction to Korean linguistics]*. Seoul, South Korea: 일조각 [Ilchokak].

Kim, H.-K. (1982). *한국 방언 연구 [A study of Korean dialects]* (Vols. 1-2). Seoul, South Korea: Seoul National University Press.

Kim, H. (1999). The place of articulation of Korean affricates revisited. *Journal of East Asian Linguistics*, *8*(4), 313–347. https://doi.org/10.1023/A:1008331010694

Kim, H. (2000). Korean Umlaut revisited. *언어연구[Language Research], 36*(2)*, 381-416.* Seoul: Seoul National University.

Kim, H. (2001). The place of articulation of the Korean plain affricate in intervocalic position: An articulatory and acoustic study. *Journal of the International Phonetic Association, 31*(2), 229–257. https://doi.org/10.1017/S0025100301002055

Kim, H. (2004). Stroboscopic-cine MRI data on Korean coronal plosives and affricates: Implications for their place of articulation as alveolar. *Phonetica*, *61*(4), 234–251. https://doi.org/10.1159/000084160

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. https://doi.org/10.1016/j.wocn.2004.06.001

Kim, J. (2008). *Focus realization: Not by focus-to-accent but prosodic structure.* In A. Schardl, M. Walkow, & M. Abdurrahman (Eds.), *Proceedings of the 38th Annual Meeting of the North East Linguistics Society, 2,* 1-14. Amherst, MA: GLSA.

Kim, J., & Jun, S.-A. (2009). Prosodic structure and focus prosody of South Kyungsang Korean. *언어연구 [Language Research]*, *45*(1), 43–66.

Kim, J. D. (2012). 경남방언 구획 문제를 다시 생각한다 [Demarcation of Gyeongnam dialect area revisited]. *배달말 [Korean Language]*, *51*, 45–90.

Kim, L., & Alderete, J. (2008). Why [h] and aspirated consonants are different in Korean. *Toronto Working Papers in Linguistics*, *28*. Retrieved from https://twpl.library.utoronto.ca/index.php/twpl/article/view/6140

Kim, M.-R. (2012). The [voice] system of Korean stops revisited with special reference to the aspirated-lax merger. *음성음운형태론 연구 [Studies in Phonetics, Phonology, and Morphology], 18*(2), 211–243. https://doi.org/10.17959/sppm.2012.18.2.211

Kim, M.-R. (2014). Ongoing sound change in the stop system of Korean: A three- to two-way categorization. *음성음운형태론 연구 [Studies in Phonetics, Phonology, and Morphology], 20*(1), 51–82.

Kim, M.-R. (2017). Dialectal variability on sound change in Contemporary Korean. *언어 [Korean Journal of Linguistics]*, *42*(3), 323–341. https://doi.org/10.18855/lisoko.2017.42.3.005

Kim, M.-R., & Duanmu, S. (2004). "Tense" and "lax" stops in Korean. *Journal of East Asian Linguistics*, *13*(1), 59–104. https://doi.org/10.1023/B:JEAL.0000007344.43938.4e

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

Kim, S. (2001). *The interaction between prosodic domain and segmental properties: Domain-initial strengthening of fricatives and Post Obstruent Tensing rule in Korean.* (Master's thesis, University of California, Los Angeles, CA). Retrieved from http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.5.6932&rep=rep1&type=pdf

Kim, S. (2004). *The role of prosodic phrasing in Korean word segmentation* (Doctoral dissertation, University of California, Los Angeles). Retrieved from http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.658.5471&rep=rep1&type=pdf

Kim, S., & Cho, T. (2009). The use of phrase-level prosodic information in lexical segmentation: Evidence from word-spotting experiments in Korean. *The Journal of* *the Acoustical Society of America, 125,* 3373-3386. https://doi.org/10.1121/1.3097777

Kim, S., & Cho, T. (2011). Articulatory manifestation of prosodic strengthening in English /i/ and /ɪ/. *말소리와 음성과학 [Phonetics and Speech Sciences]*, *3*(4), 13–21.

Kim, S., & Cho, T. (2012). Prosodic strengthening in the articulation of English /æ/. *음성음운형태론 연구 [Studies in Phonetics, Phonology, and Morphology]*, *18*(2), 321–337. https://doi.org/10.17959/sppm.2012.18.2.321

Kim, Young Seon. (2012). 국어학: 부산 방언 연구와 관련된 몇 문제 [Some problems in the study of Busan Dialect]. *배달말 [Korean Language]*, *50*, 61–91.

Kim, Young Shin. (2011). *An acoustic, aerodynamic and perceptual investigation of word-initial denasalization in Korean* (Doctoral dissertation, University College London, London, UK). Retrieved from https://discovery.ucl.ac.uk/id/eprint/1324541/1/1324541.pdf

Kirby, J. (2013). The role of probabilistic enhancement in phonologization. In A. Yu (Ed.), *Origins of Sound Change: Approaches to Phonologization* (pp. 228-246). Oxford, UK: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199573745.003.0011

Klatt, D. H. (1975). Voice onset time, frication, and aspiration in word-initial consonant clusters. *Journal of Speech and Hearing Research*, *18*(4), 686–706. https://doi.org/10.1044/jshr.1804.686

Klatt, D. H., & Klatt, L. C. (1990). Analysis, synthesis, and perception of voice quality variations among female and male talkers. *The Journal of the Acoustical Society of America*, *87*(2), 820–857. https://doi.org/10.1121/1.398894

Ko, I. (2013). *The articulation of Korean coronal obstruents: Data from heritage speakers and second language learners* (Doctoral dissertation, University of Hawaii). Available from https://search.proquest.com/docview/1430292739

Kong, T., Kim, Y., Yoon, J., & Maung, H. (2016). 경남 방언의 음운론 연구사 [The phonological research history of South Gyeongsang dialects]. *Urimal [우리말연구]*, *46*, 1–25.

Koo, H. S. (1986). *An experimental acoustic study of the phonetics of intonation in Standard Korean* (Doctoral dissertation, University of Texas at Austin). Retrieved from http://www.riss.kr/search/detail/DetailView.do?p_mat_type=be54d9b8bc7cdb09&control_no=08800726062a4710

Kraehenmann, A., & Lahiri, A. (2008). Duration differences in the articulation and acoustics of Swiss German word-initial geminate and singleton stops. *The Journal of the Acoustical Society of America, 123*(6), 4446–4455. https://doi.org/10.1121/1.2916699

Krakow, R. A., Bell-Berti, F., & Wang, Q. E. (1995). Supralaryngeal declination: Evidence from the velum. In F. Bell-Berti & L. J. Raphael (Eds.), *Producing speech: Contemporary issues for Katherine Safford Harris* (pp. 333–353). Woodbury, NY: AIP Press.

Labov, W. (1990). The intersection of sex and social class in the course of linguistic change. *Language variation and change, 2*(2), 205-254. https://doi.org/10.1017/S0954394500000338

Labov, W. (2002). Driving forces of linguistic change. *Proceedings of the International Conference on Korean Linguistics, Seoul, South Korea, 2 August 2002*. Retrieved from https://www.ling.upenn.edu/~wlabov/Papers/DFLC.htm

Ladefoged, P. (1967). *Three areas of experimental phonetics.* London, UK: Oxford University Press.

Lawson, E., Stuart-Smith, J., & Scobbie, J. (2008). Articulatory insights into language variation and change: Preliminary findings from an ultrasound study of derhoticization in Scottish English. *University of Pennsylvania Working Papers in Linguistics, 14*(2). Retrieved from https://repository.upenn.edu/pwpl/vol14/iss2/13

Lee, D., & Davis, S. (2009). *On the pitch-accent system of South Kyungsang Korean: A phonological perspective*. *언어연구 [Language Research]*, *45*(1), 3-22.

Lee, H. B. (1999). Korean. In International Phonetic Association*, Handbook of the International Phonetic Association: A guide to the use of the International Phonetic Alphabet* (pp. 120-123). Cambridge, UK: Cambridge University Press. Retrieved from http://archive.org/details/rosettaproject_kor_phon-2

Lee, H., & Jongman, A. (2012). Effects of tone on the three-way laryngeal distinction in Korean: An acoustic and aerodynamic comparison of the Seoul and South Kyungsang dialects. *Journal of the International Phonetic Association*, *42*(2), 145–169. https://doi.org/10.1017/S0025100312000035

Lee, H., Politzer-Ahles, S., & Jongman, A. (2013). Speakers of tonal and non-tonal Korean dialects use different cue weightings in the perception of the three-way laryngeal stop contrast. *Journal of Phonetics*, *41*(2), 117–132. https://doi.org/10.1016/j.wocn.2012.12.002

Lee, H., & Zhang, J. (2014). The nominal pitch accent system of South Kyungsang Korean. *Journal of East Asian Linguistics*, *23*(1), 71–111. https://doi.org/10.1007/s10831-013-9119-x

Lee, Ho-Young. (1996). *국어 음성학 [Korean phonetics]*. Paju, South Korea: 태학사 [Thaehaksa].

Lee, I., & Ramsey, S. R. (2000). *The Korean language*. Albany, NY: SUNY Press.

Lee, M. (2002). 어두 경음화에 관한 사회 언어학적 고찰 -언어 변화의 측면에서- [Sociolinguistic study of word-initial tensification: From the perspective of language change]. *말소리 (대한음성학회) [Malsori]*, 167–178.

Lee, S.-D., & Kim, S.-J. (2007). Aerodynamic approach to nasals in Korean and English. 새한영어영문학 *[The New Korean Journal of English Language and Literature]*, *49*(3), 85–100. https://doi.org/10.25151/nkje.2007.49.3.005

Lehiste, I. (1960). An acoustic-phonetic study of internal open juncture. *Phonetica, 5*(Suppl. 1), 5-54. https://doi.org/10.1159/000258062

Lehiste, I. (1972). Timing of utterances and linguistic boundaries. *The Journal of the Acoustical Society of America, 51,* 2018-2024. https://doi.org/10.1121/1.1913062

Lehiste, I. (1973). Phonetic disambiguation of syntactic ambiguity. *Glossa, 7,* 107-122.

Lennes, M. (2002). *Calculate_segment_durations.praat [Praat script]*. Retrieved from https://github.com/lennes/spect/blob/master/scripts/calculate_segment_durations.praat

Lenth, R. (2018). *Emmeans: Estimated marginal means, aka least-squares means.* https://CRAN.R-project.org/package=emmeans, R package version, 1.1.

Levelt, W. J. M. (1989). *Speaking: From intention to articulation*. Cambridge, MA: MIT Press.

Lieberman, P., & Blumstein, S. E. (1988). *Speech physiology, speech perception, and acoustic phonetics*. Cambridge, UK: Cambridge University Press. https://doi.org/10.1017/CBO9781139165952

Lindblom, B. (1963). Spectrographic study of vowel reduction. *Journal of the Acoustical Society of America*, *35*(11), 1773–1781. https://doi.org/10.1121/1.1918816

Lisker, L., & Abramson, A. S. (1964). Crosslanguage study of voicing in initial stops. *The Journal of the Acoustical Society of America*, *35*(11), 1889–1890. https://doi.org/10.1080/00437956.1964.11659830

Lombardi, L. (1991). *Laryngeal features and laryngeal neutralization* (Doctoral dissertation). University of Massachusetts. Published version available from https://www.taylorfrancis.com/books/9780429454929

Ludlow, C. L., Sedory, S. E., & Fujita, M. (1991). Neurophysiological control of vocal fold adduction and abduction for phonation onset and offset during speech. In J. Gauffin & B. Hammarberg (Eds.), *Vocal fold physiology: Acoustic, perceptual, and physiological aspects of voice mechanism* (pp. 197–205). San Diego, CA: Singular.

Machač, P., & Skarnitzl, R. (2009). *Principles of phonetic segmentation*. Prague, Czech Republic: Epocha Publishing House. https://doi.org/10.1159/000331902

Macmillan, N. A., Kingston, J., Thorburn, R., Dickey, L. W., & Bartels, C. (1999). Integrality of nasalization and F1. II. Basic sensitivity and phonetic labeling measure distinct sensory and decision-rule interactions. *The Journal of the Acoustical Society of America*, *106*(5), 2913–2932. https://doi.org/10.1121/1.428113

Martin, S. E. (1951). Korean phonemics. *Language*, *27*(4), 519–533. https://doi.org/10.2307/410039

Martinet, A. (1962). *A functional view of language*. Oxford, UK: Clarendon Press.

Moon, S.-J., & Lindblom, B. (1994). Interaction between duration, context, and speaking style in English stressed vowels. *The Journal of the Acoustical Society of America*, *96*(40), 40–55. https://doi.org/10.1121/1.410492

Mundry, R., & Nunn, C. (2009). Stepwise model fitting and statistical inference: Turning noise into signal pollution. *American Naturalist, 173*(1), 119-123. https://doi.org/10.1086/593303

Nalborczyk, L., Batailler, C., Lœvenbruck, H., Vilain, A., & Bürkner, P.-C. (2019). An introduction to Bayesian multilevel models using brms: A case study of gender effects on vowel variability in Standard Indonesian. *Journal of Speech, Language, and Hearing Research*, *62*(5), 1225–1242. https://doi.org/10.1044/2018_JSLHR-S-18-0006

Nespor, M., & Vogel, I. (1986). *Prosodic phonology*. Dordrecht, the Netherlands: Foris.

Oh, E. (2011). Effects of speaker gender on voice onset set in Korean stops, *Journal of Phonetics, 39*(1), 59–67. https://doi.org/10.1016/j.wocn.2010.11.002

Oh, J. (2011). 국어 방언에 반영된 어두경음화 [Word-initial tensification reflected in Korean dialects]. *한민족어문학 [Hanminjok Emunhak]*, *58*, 239–271.

Ohala, J. J., & Hirano, M. (1967). An experimental investigation of pitch change in speech. *The Journal of the Acoustical Society of America*, *42*(5), 1208–1209. https://doi.org/10.1121/1.2144216

Ohala, J. J. (1983). Cross-language use of pitch: An ethological view. *Phonetica*, *40*(1), 1–18. https://doi.org/10.1159/000261678

Ohala, J. J. (1992). What’s cognitive, what’s not, in sound change. In G. Kellerman & M. D. Morrissey (Eds.), *Diachrony within synchrony - language history and cognition: papers from the international symposium at the University of Duisburg, 26-28 March 1990* (pp. 309–355). Frankfurt, Germany: Peter Lang.

Onaka, A. (2003). Domain-initial strengthening in Japanese: An acoustic and articulatory study. *Proceedings of the 15th International Congress of Phonetic Sciences*, 2091–2094.

Onaka, A., Palethorpe, S., Watson, C., & Harrington, J. (2003). An acoustic analysis of domain-initial strengthening effect in Japanese. *Proceedings of the Sixth International Seminar on Speech Production*, 201–206.

Overholser, B. R., & Sowinski, K. M. (2008). Biostatistics primer: Part 2. *Nutrition in clinical practice*, *23*(1), 76–84. https://doi.org/10.1177/011542650802300176

Park, H. (2002). *Temporal and spectral characteristics of Korean phonation types* (Doctoral dissertation, University of Texas at Austin). Retrieved from https://repositories.lib.utexas.edu/bitstream/handle/2152/837/parkh026.pdf

Park, S. (2012). *경주방언과 울산방언의 성조 비교 연구 [A Contrastive Study on the tone between Gyeongju and Ulsan Dialect]* (Master's thesis, Kyungpook National University, Daegu, South Korea). Retrieved from http://dcollection.knu.ac.kr/public_resource/pdf/000000052118_20200220195840.pdf

Penney, J., Cox, F., & Szakay, A. (2019). Glottalisation of word-final stops in Australian English unstressed syllables. *Journal of the International Phonetic Association*, 1–32. https://doi.org/10.1017/S0025100319000045

Perkell, J. S., Cohen, M. H., Svirsky, M. A., Matthies, M. L., Garabieta, I., & Jackson, M. T. T. (1992). Electromagnetic midsagittal articulometer systems for transducing speech articulatory movements. *The Journal of the Acoustical Society of America*, *92*(6), 3078–3096. https://doi.org/10.1121/1.404204

Pfister, R., Schwarz, K. A., Janczyk, M., Dale, R., & Freeman, J. B. (2013). Good things peak in pairs: A note on the bimodality coefficient. *Frontiers in Psychology, 4*, 700. https://doi.org/10.3389/fpsyg.2013.00700

Piantadosi, S. T., Tily, H., & Gibson, E. (2011). Word lengths are optimized for efficient communication. *Proceedings of the National Academy of Sciences*, *108*(9), 3526–3529. https://doi.org/10.1073/pnas.1012551108

Pierrehumbert, J. B. (1980). *The phonology and phonetics of English intonation* (Doctoral dissertation, Massachusetts Institute of Technology). Retrieved from http://dspace.mit.edu/handle/1721.1/16065

Pierrehumbert, J. B. (2001). Exemplar dynamics: Word frequency, lenition and contrast. In J. L. Bybee & P. J. Hopper (Eds.), *Frequency and the emergency of linguistic structure* (Typological Studies in Language, Vol. 45, p. 137). Amsterdam, the Netherlands, & Philadelphia, PA: John Benjamins. https://doi.org/10.1075/tsl.45.08pie

Pierrehumbert, J. B. (2002). Word-specific phonetics. In C. Gussenhoven & N. Warner (Eds.), *Laboratory Phonology* (Vol. 7, pp. 101–139). Berlin, Germany: Mouton de Gruyter. https://doi.org/10.1515/9783110197105

Pierrehumbert, J. B., & Beckman, M. E. (1988). Japanese tone structure. *Linguistic Inquiry Monographs*, *15*, 1–282.

Pierrehumbert, J. B., & Talkin, D. (1992). Lenition of /h/ and glottal stop. In G. Docherty & D. Ladd (Eds.), *Papers in Laboratory Phonology II: Gesture, Segment, Prosody* (pp. 90–127). Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511519918.005

Korean Statistical Information Service. (2020). *Population statistics based on resident registration*. Retrieved from https://kosis.kr/eng/statisticsList/statisticsListIndex.do?menuId=M_01_01&vwcd=MT_ETITLE&parmTabId=M_01_01&statId=1994042&themaId=#SelectStatsBoxDiv

Port, R. F., & Rotunno, R. (1979). Relation between voice-onset time and vowel duration. *The Journal of the Acoustical Society of America*, *66*(3), 654–662. https://doi.org/10.1121/1.383692

Price, P. J., Ostendorf, S., Shattuck-Hufnagel, S., & Fong, C. (1991). The use of prosody in syntactic disambiguation. *The Journal of the Acoustic Society of America, 9,* 2956-2970. https://doi.org/10.1121/1.401770

Pruthi, T., & Espy-Wilson, C. Y. (2004). Acoustic parameters for automatic detection of nasal manner. *Speech Communication*, *43*(3), 225–239. https://doi.org/10.1016/j.specom.2004.06.001

Pullum, G. K., & Ladusaw, W. A. (1996). *Phonetic symbol guide* (2nd ed.). Chicago, IL: University of Chicago Press.

Ministry of Education. (1988). *표준어 규정 [Rules of the Standard Language].*

R Core Team. (2015). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. Available from https://www.r-project.org

Repp, B. H. (1982). Phonetic trading relations and context effects: New experimental evidence for a speech mode of perception. *Psychological Bulletin*, *92*(1), 81–110. https://doi.org/10.1037//0033-2909.92.1.81

Robinson, O. W. (1976). A “scattered” rule in Swiss German. *Language*, *52*(1), 148–162. https://doi.org/10.2307/413214

Roettger, T. B. (2019). Researcher degrees of freedom in phonetic research. *Laboratory Phonology: Journal of the Association for Laboratory Phonology, 10*(1), 1. https://doi.org/10.5334/labphon.147

Schilling, M. F., Watkins, A. E., & Watkins, W. (2002). Is human height bimodal? *The American Statistician*, *56*(3), 223–229. https://doi.org/10.1198/00031300265

Schober, P., Boer, C., & Schwarte, L. A. (2018). Correlation coefficients: Appropriate use and interpretation. *Anesthesia & Analgesia*, *126*(5), 1763–1768. https://doi.org/10.1213/ANE.0000000000002864

Schwartz, M. F. (1968). The acoustics of normal and nasal vowel production. *The Cleft Palate Journal*, *5*(2), 125–140.

Scobbie, J. M. (1995). What do we do when phonology is powerful enough to imitate phonetics? Comments on Zsiga. In B. Connell & A. Arvaniti (Eds.), *Phonology and phonetic evidence: Papers in Laboratory Phonology IV* (pp. 303-314). Cambridge, UK: Cambridge University Press. http://doi.org/10.1017/CBO9780511554315.021

Scobbie, J. M. (2005). The phonetics phonology overlap. *QMU Speech Science Research Centre Working Papers, WP-1*. Retrieved from https://eresearch.qmu.ac.uk/handle/20.500.12289/138

Sejong, King of Chosun. (1443). *훈민정음 [Hunmin Chongum]*. Seoul, South Korea: 집현전 [Jiphyeonjeon]. Photographic reproduction of the "Haeryebon" edition of the Hunmin Chongum by King Sejong, Chong In-ji and others, with a translation into modern Korean. (1967).

Selkirk, E. (1984). *Phonology and syntax: The relation between sound and structure*. Cambridge, MA: MIT Press.

Selkirk, E. (1986). On derived domains in sentence phonology. *Phonology*, *3*, 371–405. https://doi.org/10.1017/S0952675700000695

Shadle, C. H., & Scully, C. (1995). An articulatory-acoustic-aerodynamic analysis of [s] in VCV sequences. *Journal of Phonetics*, *23*(1), 53–66. https://doi.org/10.1016/S0095-4470(95)80032-8

Schafer, A. J. (1997). *Prosodic parsing: The role of prosody in sentence comprehension.* (Doctoral dissertaion, University of Massachusetts). Retrieved from http://www2.hawaii.edu/~aschafer/Schafer1997.pdf

Schafer, A. J., & Jun, S.-A. (2002). Effects of accentual phrasing on adjective interpretation in Korean. In M. Nakayama (Ed.), *East Asian Language Processing,* 223-255. Stanford, CA: CSLI.

Schafer, A. J., Speer, S. R., & Warren, P. (2005). Prosodic influences on the production and comprehension of syntactic ambiguity in a game-based conversation task. In M. Tanenhaus & J. Trueswell (Eds.), *Approaches to Studying World Situated Language Use: Psycholinguistic, Linguistic and Computational Perspectives on Bridging the Product and Action Tradition* (pp. 209-225). Cambridge, MA: MIT Press.

Shattuck-Hufnagel, S., & Turk, A. E. (1996). A prosody tutorial for investigators of auditory sentence processing. *Journal of Psycholinguistic Research*, *25*(2), 193–247. https://doi.org/10.1007/BF01708572

Shin, J. (2015). Vowels and consonants. In L. Brown & J. Yeon (Eds.), *The Handbook of Korean Linguistics* (pp. 1–21). New York, NY: John Wiley & Sons. https://doi.org/10.1002/9781118371008.ch1

Shin, J., Kiaer, J., & Cha, J. (2012). *The Sounds of Korean*. Cambridge, UK Cambridge University Press. https://doi.org/10.1017/CBO9781139342858

Shin, S., & Tremblay, A. (2018). Effect of prosodic context on lexical access: An investigation of Korean denasalization. *Proceedings of the 9th International Conference on Speech Prosody*, 408-412.

Shipp, T., & McGlone, R. E. (1971). Laryngeal dynamics associated with voice frequency change. *Journal of Speech, Language, and Hearing Research*, *14*(4), 761–768. https://doi.org/ 10.1044/jshr.1404.761

Silva, D. (2006). Acoustic evidence for the emergence of tonal contrast in Contemporary Korean. *Phonology*, *23*(02), 287–308. https://doi.org/10.1017/S0952675706000911

Silva, D. (2011). Out of one, many: The emergence of world Korean(s). In J. H.-S. Yoon, Y. Kang, J.-Y. Han, & C.-K. Shi (Eds.), *Inquiries into Korean Linguistics IV*, (pp. 9–37). Paju, South Korea: 태학사 [Thaehaksa].

Sluijter, A. M. C., van Heuven, V. J., & Pacilly, J. J. A. (1997). Spectral balance as a cue in the perception of linguistic stress. *The Journal of the Acoustical Society of America, 101*(1), 503–513. https://doi.org/10.1121/1.417994

Sohn, H.-M. (1999). *The Korean Language* (Cambridge Language Surveys, Vol. 12). Cambridge, UK: Cambridge University Press. https://doi.org/10.1075/sl.32.2.15hor

Sóskuthy, M., & Roettger, T. B. (2019). *When the tune shapes morphology: The origins of vocatives*. https://doi.org/10.31234/osf.io/ef52k

Standish, R. K. (2004). Why Occam's Razor. *Foundations of Physics Letters, 17*(3), 255-266. https://doi.org/10.1023/B:FOPL.0000032475.18334.0e

Stetson, R. H. (1951). *Motor phonetics: A study of speech movements in action* (2nd ed.). Amsterdam, the Netherlands: North Holland Publishing.

Stevens, K. N. (1998). *Acoustic phonetics*. Cambridge, MA: MIT Press.

Stevens, K. N., Keyser, S. J., & Kawasaki, H. (1986). Toward a phonetic and phonological theory of redundant features. In J. S. Perkell & D. H. Klatt (Eds.), *Invariance and variability in speech processes* (pp. 426–463). Hillsdale, NJ: Lawrence Erlbaum.

Stölten, K. (2013). *The effects of age of onset on VOT in L2 acquisition and L1 attrition: A study of the speech production and perception of advanced Spanish-Swedish bilinguals* (Doctoral dissertation, Stockholm University, Stockholm, Sweden). Retrieved from http://urn.kb.se/resolve?urn=urn:nbn:se:su:diva-94887

Straka, G. (1963). La Division des sons du langage en voyelles et consonnes peut-elle être justifiée? *Travaux de Linguistique et de Litt*érature, *1,* 17-99.

Streeter, L. A. (1978). Acoustic determinants of phrase boundary perception. *The Journal of the Acoustical Society of America, 64,* 1582-1592. https://doi.org/10.1121/1.382142

Styler, W. (2015). *On the acoustical and perceptual features of vowel nasality* (Doctoral dissertation, University of Colorado). Retrieved from https://scholar.colorado.edu/ling_gradetds/56

Styler, W. (2017a). *Nasality automeasure script package* (Version 5.9). Retrieved from https://github.com/stylerw/styler_praat_scripts/tree/master/nasality_automeasure

Styler, W. (2017b). On the acoustical features of vowel nasality in English and French. *The Journal of the Acoustical Society of America*, *142*(4), 2469–2482. https://doi.org/10.1121/1.5008854

Trudgill, P. (1974). *The social differentiation of English in Norwich* (Vol. 13). Cambridge, UK: Cambridge University Press.

Turton, D. (2014). *Variation in English /l/: Synchronic reflections of the life cycle of phonological processes* (Doctoral dissertation, University of Manchester, Manchester, UK). Retrieved from https://pdfs.semanticscholar.org/7754/cd45c3a4d107efabf677795a75351b8dd113.pdf

Turton, D. (2017). Categorical or gradient? An ultrasound investigation of/l/-darkening and vocalization in varieties of English. *Laboratory Phonology: Journal of the Association for Laboratory Phonology*, *8*(1), 13. https://doi.org/10.5334/labphon.35

Um, H.-Y. (2014). Phonetic realizations of /h/ after a sonorant in Korean. *음성음운형태론연구[Studies in Phonetics, Phonology, and Morphology]*, 20(3), 357–377.

Umeda, H. (1957). The phonemic system of Modern Korean. *言語研究 [Gengo Kenkyu]*, *32*, 60–82.

Vasishth, S., Nicenboim, B., Beckman, M. E., Li, F., & Kong, E. J. (2018). Bayesian data analysis in the phonetic sciences: A tutorial introduction. *Journal of Phonetics, 71*, 147–161. https://doi.org/10.1016/j.wocn.2018.07.008

Vayra, M., & Fowler, C. A. (1992). Declination of supralaryngeal gestures in spoken Italian. *Phonetica*, *49*(1), 48–60. https://doi.org/10.1159/000261902

Vennemann, T. (1972). Phonetic analogy and conceptual analogy. In T. Vennemann & T. H. Wilbur (Eds.), *Schuchardt, the neogrammarians, and the transformational theory of phonological change: Four essays* (pp. 181–204). Königstein im Taunus, Germany: Athenäum Verlag.

Wedel, A., Kaplan, A., & Jackson, S. (2013). High functional load inhibits phonological contrast loss: A corpus study. *Cognition*, *128*(2), 179–186. https://doi.org/10.1016/j.cognition.2013.03.002

Wedel, A., Ussishkin, A., King, A., Martin, J., & Geary, J. (2018). Phonological grammars evolve to preserve information at word beginnings. *Proceedings of the 12th International Conference on the Evolution of Language*.

Weismer, G. (1979). Sensitivity of voice-onset time (VOT) measures to certain segmental features in speech production. *Journal of Phonetics*, *7*(2), 197–204. https://doi.org/10.1016/S0095-4470(19)31041-1

Wightman, C. W., Shattuck-Hufnagel, S., Ostendorf, M., & Price, P. J. (1992). Segmental durations in the vicinity of prosodic phrase boundaries. *The Journal of the Acoustical Society of America*, *91*(3), 1707–1717. https://doi.org/10.1121/1.402450

Wowslides. (2016). *Free South Korea editable map*. Retrieved from https://yourfreetemplates.com/free-south-korea-editable-map/

Xu, Y., & Prom-on, S. (2019). Economy of effort or maximum rate of information? Exploring basic principles of articulatory dynamics. *Frontiers in Psychology*, *10*. https://doi.org/10.3389/fpsyg.2019.02469

Yoo, K. (2015a). Domain-initial denasalisation in Busan Korean: A cross-generational case study. *Proceedings of the 18th International Congress of Phonetic Sciences.*

Yoo, K. (2015b). *The effect of prosodic position on the denasalisation of Korean word-initial nasals* (Undergraduate dissertation). University of Cambridge, Cambridge, UK.

Yoo, K. (2016). *The historical development of domain-initial denasalisation in Seoul Korean* (Master's thesis). University of Cambridge, Cambridge, UK.

Yoo, K., & Nolan, F. (forthcoming). Sampling the progression of domain-initial denasalization in Seoul Korean. *Laboratory Phonology: Journal of the Association for Laboratory Phonoloy.*

Yoshida, K. (2008). Phonetic implementation of Korean denasalization and its variation related to prosody. *IULC Working Papers*, *8*(1). Retrieved from https://pdfs.semanticscholar.org/ed7a/e9a10afc1d6ed7e7a99f31693fb96fa0ce0e.pdf

Yun, J., & Arai, T. (2020). Perception of synthesised /na/ by Korean listeners. *Acoustical Science and Technology, 41*(2), 501-512. https://doi.org/10.1250/ast.41.501

Yun, S., & Kang, Y. (2018). *Dialectal variation and sound change in Korean nasals.* Poster presented at LabPhon16, Lisbon, Portugal. Retrieved from https://www.yoonjungkang.com/uploads/1/1/6/2/11625099/nasal_labphon_final.pdf

Zhang, H. (2015). *Production and acoustics of creaky nasal vowels* (Doctoral dissertation, University of Colorado). Retrieved from https://scholar.colorado.edu/ling_gradetds/49

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

Zsiga, E. C. (1995). An acoustic and electropalatographic study of lexical and postlexical palatalization in American English. In B. Connell & A. Arvaniti (Eds.), *Phonology and phonetic evidence: Papers in Laboratory Phonology IV* (pp. 282-302). https://doi.org/10.1017/CBO9780511554315.020

Appendix A.

Experimental stimuli

In (I) below, the phonemic transcriptions, interlinear glosses, and English translations are provided for the experimental stimuli of the production and perception experiments in this dissertation. Note that the materials inside the round brackets were included for the clarity of the intended meaning, and the participants were instructed not to read them. These materials are only shown in the practice sets. The original materials in Korean, including the instructions and the English translations of the instructions, can be found in (II).

I. Phonemic transcriptions

Practice Set 1. /w/

a. PW-initial /w/

$_{\text{IP}}${ $_{\text{AP}}${ʌnni(-ɯi) ***w***ʌnp^{h}isɯ} pilljʌ ipʌs*ʌ}
sister(-POSS) dress borrow wear
'Did you borrow and wear your sister's dress?'

b. AP-initial /w/

$_{\text{IP}}${ ʌnni(-ka) $_{\text{AP}}${***w***ʌnp^{h}isɯ} pilljʌ ipʌs*ʌ}
sister(-NOM) dress borrow wear
'Did your sister borrow and wear a dress?'

c. IP-initial /w/ with no preceding pause

ʌnni(,) $_{\text{IP}}${***w***ʌnp^{h}isɯ pilljʌ ipʌs*ʌ}
sister(implied VOC) dress borrow wear
'Sister, did you borrow and wear a dress?'

d. IP-initial /w/ preceded by a pause

ʌnni (pause) $_{\text{IP}}${***w***ʌnp^{h}isɯ pilljʌ ipʌs*ʌ}
sister(implied VOC) dress borrow wear
'Sister, (pause) did you borrow and wear a dress?'

Practice Set 2. /s/

a. PW-initial /s/

$_{\text{IP}}${ $_{\text{AP}}${halmʌni(-ne)	**s**unte}	tɯsilʌ	kasejo}
grandma(-POSS)	*sundae* (Korean street food)	eat	go

‘Are you going to go eat *Grandma’s Sundae* (restaurant name)?’

b. AP-initial /s/

$_{\text{IP}}${ halmʌni(-ka)	$_{\text{AP}}${***s***unte}	tɯsilʌ	kasejo}
grandma(-NOM)	*sundae* (Korean street food)	eat	go

‘Is Grandma going to go eat *sundae*?’

c. IP-initial /s/ with no preceding pause

halmʌni(,)	$_{\text{IP}}${***s***unte	tɯsilʌ	kasejo}
grandma(implied VOC)	*sundae* (Korean street food)	eat	go

‘Grandma, are you going to go eat *sundae*?’

d. IP-initial /s/ preceded by a pause

halmʌni	(pause)	$_{\text{IP}}${***s***unte	tɯsilʌ	kasejo}
grandma(implied VOC)		*sundae* (Korean street food)	eat	go

‘Grandma, (pause) are you going to go eat *sundae*?’

<PART A>

(a) PW-initial /t, n/

1a. $_{\text{IP}}\{$ $_{\text{AP}}\{$ʌmma	***t***alimi}	kjʌlkuk	kot͡saŋnasʌ	pʌljʌs*ʌjo}
mom	iron	finally	broke	throw away

‘Did you throw away Mom's iron because it finally broke?’

2a. $_{\text{IP}}\{$ $_{\text{AP}}\{$ap*a	***t***eli}	t͡sat͡su	ijoŋhasintakojo}
dad	chauffeur service	often	use

‘You use *Dad’s Chaffeur Service* (company name) often?’

3a. $_{\text{IP}}\{$ $_{\text{AP}}\{$op*a	***t***ʌpɯlpeisɯ}	pilljʌ	osjʌs*ʌjo}
brother	double bass	borrow	come

‘Did you borrow and bring you brother’s double bass?’

4a. $_{\text{IP}}\{$ $_{\text{AP}}\{$at͡sʌs*i	***t***oŋneesʌ}	ilhasinɯn	kʌjejo}
mister	town	working	that

‘You are working in Mister’s town?’

5a. $_{\text{IP}}\{$ $_{\text{AP}}\{$halmʌni	***t***ulut͡s^{h}iki}	tɯsilʌ	kasejo}
grandma	*duruchigi* (Korean dish)	eat	go

‘Are you going to go eat *Grandma’s Duruchigi* (restaurant name)?’

6a. $_{\text{IP}}\{$ $_{\text{AP}}\{$imo	***n***ek*wa}	an	tanisejo}
aunt	internal medicine clinic	not	go

‘Don’t you go to *Auntie’s Internal Medicine Clinic* (name of a clinic)?’

7a. $_{\text{IP}}\{$ $_{\text{AP}}\{$halapʌt͡si	***n***amu}	kjʌlkuk	p^{h}alkilo	hasjʌs*ʌjo}
grandpa	tree	finally	sell	decide

‘Did you finally decide to sell Grandpa’s tree?’

8a. $_{\text{IP}}\{$ $_{\text{AP}}\{$appa	***n***ʌkuli}	t͡sapɯlʌ	kasejo}
dad	racoon	catch	go

‘Are you going to go catch racoon dad?’

9a. $_{\mathrm{IP}}${ $_{\mathrm{AP}}${at͡sʌs*i ***n***oŋtam} kat^{h}ɯn kʌ t͡soahasejo}
mister jokes like things like
'Do you like old-fashioned jokes?'

10a. $_{\mathrm{IP}}${ $_{\mathrm{AP}}${halmʌni ***n***uluŋt͡si} tɯsilʌ kasejo}
grandma rice crusts eat go
'Are you going to go eat *Grandma's Rice Crusts* (restaurant name)?'

(b) AP-initial /t, n/

1b. $_{\mathrm{IP}}${ʌmma $_{\mathrm{AP}}${***t***alimi} kjʌlkuk kot͡saŋnasʌ pʌljʌs*ʌjo}
mom iron finally broke throw away
'Did Mom throw away the iron because it finally broke?'

2b. $_{\mathrm{IP}}${ap*a $_{\mathrm{AP}}${***t***eli} t͡sat͡su ijoŋhasintakojo}
dad chauffeur service often use
'Dad uses chauffeur service often?'

3b. $_{\mathrm{IP}}${op*a $_{\mathrm{AP}}${***t***ʌpɯlpeisɯ} pilljʌ osjʌs*ʌjo}
brother double bass borrow come
'Did Brother borrow and bring a double bass?'

4b. $_{\mathrm{IP}}${at͡sʌs*i $_{\mathrm{AP}}${***t***oŋneesʌ} ilhasinɯn kʌjejo}
mister town working that
'Is Mister working in town?'

5b. $_{\mathrm{IP}}${halmʌni $_{\mathrm{AP}}${***t***ulut͡s^{h}iki} tɯsilʌ kasejo}
grandma *duruchigi* (Korean dish) eat go
'Is Grandma going to go eat *duruchigi*?'

6b. $_{\mathrm{IP}}${imo $_{\mathrm{AP}}${***n***ek*wa} an tanisejo}
aunt internal medicine clinic not go
'Doesn't Auntie go to Internal Medicine Clinic?'

7b. $_{\mathrm{IP}}${halapʌt͡si $_{\mathrm{AP}}${***n***amu} kjʌlkuk p^{h}alkilo hasjʌs*ʌjo}
grandpa tree finally sell decide
'Did Grandpa finally decide to sell his tree?'

8b.	$_{\text{IP}}${appa	$_{\text{AP}}${***n***ʌkuli}	t͡sapɯlʌ	kasejo}	
	dad	racoon	catch	go	

‘Is Dad going to go catch a racoon?’

9b.	$_{\text{IP}}${at͡sʌs*i	$_{\text{AP}}${***n***oŋtam}	kat^{h}ɯn	kʌ	t͡soahasejo}
	mister	jokes	like	things	like

‘Does Mister like jokes?’

10b.	$_{\text{IP}}${halmʌni	$_{\text{AP}}${***n***uluŋt͡si}	tɯsilʌ	kasejo}	
	grandma	rice crusts	eat	go	

‘Is Grandma going to go eat rice crusts?’

(c) IP-initial /t, n/ with no preceding pause

1c.	ʌmma	$_{\text{IP}}${***t***alimi	kjʌlkuk	kot͡saŋnasʌ	pʌljʌs*ʌjo}
	mom	iron	finally	broke	throw away

‘Mom, did you throw away the iron because it finally broke?’

2c.	ap*a	$_{\text{IP}}${***t***eli	t͡sat͡su	ijoŋhasintakojo}
	dad	chauffeur service	often	use

‘Dad, you use chauffeur service often?’

3c.	op*a	$_{\text{IP}}${***t***ʌpɯlpeisɯ	pilljʌ	osjʌs*ʌjo}
	brother	double bass	borrow	come

‘Brother, did you borrow and bring a double bass?’

4c.	at͡sʌs*i	$_{\text{IP}}${***t***oŋneesʌ	ilhasinɯn	kʌjejo}
	mister	town	working	that

‘Mister, you are working in town?’

5c.	halmʌni	$_{\text{IP}}${***t***ulut͡s^{h}iki	tɯsilʌ	kasejo}
	grandma	*duruchigi* (Korean dish)	eat	go

‘Grandma, are you going to go eat *duruchigi*?’

6c.	imo	$_{\text{IP}}${***n***ek*wa	an	tanisejo}
	aunt	internal medicine clinic	not	go

‘Auntie, do you go to Internal Medicine Clinic?’

7c. halapʌt͡si $_{IP}${***n***amu   kjʌlkuk   p^{h}alkilo   hasjʌs*ʌjo}
grandpa   tree   finally   sell   decide
‘Grandpa, did you finally decide to sell the tree?’

8c. appa $_{IP}${***n***ʌkuli   t͡sapɯlʌ   kasejo}
dad   racoon   catch   go
‘Dad, are you going to go catch a racoon?’

9c. at͡sʌs*i $_{IP}${***n***oŋtam   kat^{h}ɯn   kʌ   t͡soahasejo}
mister   jokes   like   things   like
‘Mister, do you like jokes?’

10c. halmʌni $_{IP}${***n***uluŋt͡si   tɯsilʌ   kasejo}
grandma   rice crusts   eat   go
‘Grandma, are you going to go eat rice crusts?’

(d) IP-initial /t, n/ preceded by a pause

1d. ʌmma (pause) $_{IP}${***t***alimi   kjʌlkuk   kot͡saŋnasʌ   pʌljʌs*ʌjo}
mom   iron   finally   broke   throw away
‘Mom, (pause) did you throw away the iron because it finally broke?’

2d. ap*a (pause) $_{IP}${***t***eli   t͡sat͡su   ijoŋhasintakojo}
dad   chauffeur service   often   use
‘Dad, (pause) you use chauffeur service often?’

3d. op*a (pause) $_{IP}${***t***ʌpɯlpeisɯ   pilljʌ   osjʌs*ʌjo}
brother   double bass   borrow   come
‘Brother, (pause) did you borrow and bring a double bass?’

4d. at͡sʌs*i (pause) $_{IP}${***t***oŋneesʌ   ilhasinɯn   kʌjejo}
mister   town   working   that
‘Mister, (pause) you are working in town?’

5d. halmʌni (pause) $_{IP}${***t***ulut͡s^{h}iki   tɯsilʌ   kasejo}
grandma   *duruchigi* (Korean dish)   eat   go
‘Grandma, (pause) are you going to go eat *duruchigi*?’

6d. imo (pause) $_{\text{IP}}${***n***ek*wa an tanisejo}
aunt internal medicine clinic not go
‘Auntie, (pause) do you go to Internal Medicine Clinic?’

7d. halapʌt͡si (pause) $_{\text{IP}}${***n***amu kjʌlkuk p^{h}alkilo hasjʌs*ʌjo}
grandpa tree finally sell decide
‘Grandpa, (pause) did you finally decide to sell the tree?’

8d. appa (pause) $_{\text{IP}}${***n***ʌkuli t͡sapɯlʌ kasejo}
dad racoon catch go
‘Dad, (pause) are you going to go catch a racoon?’

9d. at͡sʌs*i (pause) $_{\text{IP}}${***n***oŋtam kat^{h}ɯn kʌ t͡soahasejo}
mister jokes like things like
‘Mister, (pause) do you like jokes?’

10d. halmʌni (pause) $_{\text{IP}}${***n***uluŋt͡si tɯsilʌ kasejo}
grandma rice crusts eat go
‘Grandma, (pause) are you going to go eat rice crusts?’

<PART B>

(a) Morpheme-internal /m, n, k, t/

11a. $_{\mathrm{IP}}${ $_{\mathrm{AP}}${kaŋt͡su***m***an} hwakalɯl mannakilo hes*tako}
Jooman Kang Artist meet decided
'Did you decide to meet Artist Jooman Kang?'

12a. $_{\mathrm{IP}}${ $_{\mathrm{AP}}${t͡s^{h}wekju***n***a} kɯnelɯl t͡soahantako}
Kyuna Choi swing like
'Kyuna Choi likes swings?'

13a. $_{\mathrm{IP}}${ $_{\mathrm{AP}}${pallent^{h}i***n***a} sʌnseŋnim-hako jekinɯn heponkʌja}
Valentina teacher-with talk try
'Has Valentina tried talking to your teacher?'

14a. $_{\mathrm{IP}}${ $_{\mathrm{AP}}${t͡se***n***i} sɯk^{h}i t^{h}aponkʌja}
Jenny ski try
'Has Jenny tried skiing before?'

15a. $_{\mathrm{IP}}${ $_{\mathrm{AP}}${ʌlin a***k***a} sanɯn t͡sipe tuko was*tako}
Young baby live house leave came
'You left it at the house where a young baby lives?'

16a. $_{\mathrm{IP}}${ $_{\mathrm{AP}}${alp^{h}ɯle***t***o} k*ika is*-ɯmjʌn t͡sot͡s^{h}i ank^{h}es*ʌjo}
Alfredo talent have-if good would not
'Wouldn't it be good if Alfredo had a talent?'

17a. $_{\mathrm{IP}}${ $_{\mathrm{AP}}${p^{h}elɯnan***t***o} t͡saki-man t͡suŋjohake seŋkakhanɯn kʌs kat^{h}tako}
Fernando self-only important think that as if
'You mean Fernando seems to think only himself is important?'

(b) Morpheme-initial /m, n, k, t/

11b. $_{\mathrm{IP}}${ $_{\mathrm{AP}}${kaŋt͡su-***m***an} hwakalɯl mannakilo hes*tako}
Kangjoo-only Artist meet decided
'Only Kangjoo decided to meet an artist?'

12b. ${}_{IP}${ ${}_{AP}${hjʌnkju-***n***a} kɯnelɯl t͡soahantako}
Hyeongyu-only swing like
‘Only people like Hyeongyu likes swings?’

13b. ${}_{IP}${ ${}_{AP}${kanhosa-***n***a} sʌnseŋnim-hako jekinɯn heponkʌja}
Nurse-or doctor-with talk try
‘Have you tried talking to a nurse or a doctor?’

14b. ${}_{IP}${ ${}_{AP}${mint͡se-***n***i} sɯk^{h}i t^{h}aponkʌja}
Minje-QUES ski try
‘Is it Minje who has skied before?’

15b. ${}_{IP}${ ${}_{AP}${t͡sijʌni-***k***a} sanɯn t͡sipe tuko was*tako}
Jiyeon-SBJ live house leave came
‘You left it at the house where Jiyeon lives?’

16b. ${}_{IP}${ ${}_{AP}${at͡sʌs*i-***t***o} k*ika is*-ɯmjʌn t͡sot͡s^{h}i ank^{h}es*ʌjo}
mister-too talent have-if good would not
‘Wouldn’t it be good if the mister had a talent, too?’

17b. ${}_{IP}${ ${}_{AP}${t͡s^{h}wekɯmnan-***t***o} t͡saki-man t͡suŋjohake seŋkakhanɯn kʌs kat^{h}tako}
Geumnan Choi-too self-only important think that as if
‘You mean Geumnan Choi seems to think only herself is important?’

(c) PW-initial /m, n, k, t/

11c. ${}_{IP}${ ${}_{AP}${kaŋt͡su ***m***anhwakalɯl} mannakilo hes*tako}
Kangjoo cartoon artist meet decided
‘You decided to meet Cartoon Artist Kangjoo?’

12c. ${}_{IP}${ ${}_{AP}${hjʌnkju ***n***akɯnelɯl} t͡soahantako}
Hyeongyu drifter like
‘You like Hyeongyu the Drifter?’

13c. ${}_{IP}${ ${}_{AP}${kanhosa ***n***alae tehesʌ} jekinɯn heponkʌja}
nurse world about talk try
‘Have you tried talking about the world of nurses?’

14c. $_{\text{IP}}${ $_{\text{AP}}${mint͡se	***n***isɯ}	s*ʌpon	t͡sʌk	is*ʌ}	
	Minje	varnish	try using	ever	has

‘Have you tried using *Minje Varnish* (product name) before?’

15c. $_{\text{IP}}${ $_{\text{AP}}${t͡sijʌni	***k***asanɯn}	t͡sipe	tuko	was*tako}	
	Jiyeon	lyrics	home	leave	came

‘You left Jiyeon’s lyrics at home?’

16c. $_{\text{IP}}${ $_{\text{AP}}${at͡sʌs*i	***t***ok*ika	is*-ɯmjʌn	t͡sot͡sʰi	ankʰes*ʌjo}	
	mister	axe	have-if	good	would not

‘Wouldn’t it be good if we had Mister’s axe?’

17c. $_{\text{IP}}${ $_{\text{AP}}${t͡sʰwekɯmnan	***t***ot͡saki-man}	t͡suŋjohake	seŋkakhanɯn	kʌs	katʰtako}	
	Geumnan Choi	pottery-only	important	think	that	as if

‘You mean they seem to think only Geumnam Choi’s pottery is important?’

(d) AP-initial /m, n, k, t/

11d. $_{\text{IP}}${kaŋt͡su	$_{\text{AP}}${***m***anhwakalɯl}	mannakilo	hes*tako}
Kangjoo	cartoon artist	meet	decided

‘Kangjoo decided to meet a cartoon artist?’

12d. $_{\text{IP}}${hjʌnkju	$_{\text{AP}}${***n***akɯnelɯl}	t͡soahantako}
Hyeongyu	drifter	like

‘Hyeongyu likes drifters?’

13d. $_{\text{IP}}${kanhosa	$_{\text{AP}}${***n***a-sʌnseŋnim-hako}	jekinɯn	heponkʌja}
nurse	Mr./Ms. Na-with	talk	try

‘Has the nurse tried talking to Mr./Ms. Na?’

14d. $_{\text{IP}}${mint͡se	$_{\text{AP}}${***n***isɯ}	s*ʌpon	t͡sʌk	is*ʌ}
Minje	varnish	try using	ever	has

‘Has Minje ever tried using varnish before?’

15d. $_{\text{IP}}${t͡sijʌni	$_{\text{AP}}${***k***asanɯn}	t͡sipe	tuko	was*tako}
Jiyeon	lyrics	home	leave	came

‘Jiyeon left the lyrics at home?’

16d. $_{IP}${at͡sʌs*i $_{AP}${***t***ok*ika is*-ɯmjʌn t͡sot͡s^{h}i ank^{h}es*ʌjo}
mister axe have-if good would not
'Wouldn't it be good if the mister had an axe?'

17d. $_{IP}${t͡s^{h}wekɯmnan $_{AP}${***t***ot͡saki-man} t͡suŋjohake seŋkakhanɯn kʌs kat^{h}tako}
Geumnan Choi pottery-only important think that as if
'You mean Geumnan Choi seems to think only pottery is important?'

(e) IP-initial /m, n, k, t/ with no preceding pause

11e. kaŋt͡su $_{IP}${ $_{AP}${***m***anhwakalɯl} mannakilo hes*tako}
Kangjoo cartoon artist meet decided
'Kangjoo, you decided to meet a cartoon artist?'

12e. hjʌnkju $_{IP}${ $_{AP}${***n***akɯnelɯl} t͡soahantako}
Hyeongyu drifter like
'Hyeongyu, you like drifters?'

13e. kanhosa $_{IP}${ $_{AP}${***n***a-sʌnseŋnim-hako} jekinɯn heponkʌja}
nurse Mr./Ms. Na-with talk try
'Nurse, have you tried talking to Mr./Ms. Na?'

14e. mint͡se $_{IP}${ $_{AP}${***n***isɯ} s*ʌpon t͡sʌk is*ʌ}
Minje varnish try using ever has
'Minje, have you ever tried using varnish before?'

15e. t͡sijʌni $_{IP}${ $_{AP}${***k***asanɯn} t͡sipe tuko was*tako}
Jiyeon lyrics home leave came
'Jiyeon, you left the lyrics at home?'

16e. at͡sʌs*i $_{IP}${ $_{AP}${***t***ok*ika is*-ɯmjʌn t͡sot͡s^{h}i ank^{h}es*ʌjo}
mister axe have-if good would not
'Mister, wouldn't it be good if we had an axe?'

17e. t͡s^{h}wekɯmnan $_{IP}${ $_{AP}${***t***ot͡saki-man} t͡suŋjohake seŋkakhanɯn kʌs kat^{h}tako}
Geumnan Choi pottery important think that as if
'Geumnan Choi, you mean they think only pottery is important?'

(f) IP-initial /m, n, k, t/ preceded by a pause

11f. kaŋt͡su (pause) $_{\text{IP}}${ $_{\text{AP}}${***m***anhwakalɯl} mannakilo hes*tako}
Kangjoo cartoon artist meet decided
‘Kangjoo, (pause) you decided to meet a cartoon artist?’

12f. hjʌnkju (pause) $_{\text{IP}}${ $_{\text{AP}}${***n***akɯnelɯl} t͡soahantako}
Hyeongyu drifter like
‘Hyeongyu, (pause) you like drifters?’

13f. kanhosa (pause) $_{\text{IP}}${ $_{\text{AP}}${***n***a-sʌnseŋnim-hako} jekinɯn heponkʌja}
nurse Mr./Ms. Na-with talk try
‘Nurse, (pause) have you tried talking to Mr./Ms. Na?’

14f. mint͡se (pause) $_{\text{IP}}${ $_{\text{AP}}${***n***isɯ} s*ʌpon t͡sʌk is*ʌ}
Minje varnish try using ever has
‘Minje, (pause) have you ever tried using varnish before?’

15f. t͡sijʌni (pause) $_{\text{IP}}${ $_{\text{AP}}${***k***asanɯn} t͡sipe tuko was*tako}
Jiyeon lyrics home leave came
‘Jiyeon, (pause) you left the lyrics at home?’

16f. at͡sʌs*i (pause) $_{\text{IP}}${ $_{\text{AP}}${tok*ika is*-ɯmjʌn t͡sot͡s^{h}i ank^{h}es*ʌjo}
mister axe have-if good would not
‘Mister, (pause) wouldn’t it be good if we had an axe?’

17f. t͡s^{h}wekɯmnan (pause) $_{\text{IP}}${ $_{\text{AP}}${tot͡saki-man} t͡suŋjohake seŋkakhanɯn kʌs kat^{h}tako}
Geumnan Choi pottery important think that as if
‘Geumnan Choi, (pause) you mean they think only pottery is important?’

II. Original materials

<u>녹음 방법 및 주의 사항</u>

1. 속도를 다르게 해서 두 번 읽어주세요. 처음에는 보통으로, 그 다음에는 조금 빠르게 읽어주세요. 한 세트를 읽고 나면 한 번 더 전과 같이 읽어주세요.
2. 말하듯이 자연스럽게 읽어주세요.
3. 소괄호 안의 단어는 읽지 마세요.
4. 문장의 어떤 단어를 강조하지 않도록 주의해 주세요.
5. 충분히 연습을 해보시고 문장의 뜻과 문장을 말하는 상황을 이해하고 말해주세요.

<u>Recording instructions</u>

1. Read the materials in two different speech rates, first at a moderate rate, then at a slightly fast rate. Once you finish reading each set, repeat one more time.
2. Read as if you are talking to someone.
3. Do not read the words in the round brackets.
4. Be careful not to emphasise certain word(s) in a sentence.
5. Take as much time as you need to practice. As you record, make sure you understand the meaning of the sentence you read.

Practice Set 1.

a. (너 혹시) [언니 원피스] 빌려 입었어?
b. (엄마,) 언니(가) 원피스 빌려 입었어?
c. 언니, 원피스 빌려 입었어?
d. 언니, (pause) 원피스 빌려 입었어?

Practice Set 2.

a. [할머니 순대] 드시러 가세요?
b. (엄마,) 할머니(께서 지금) 순대 드시러 가세요?
c. 할머니, 순대 드시러 가세요?
d. 할머니, (pause) 순대 드시러 가세요?

<SET A>

(a)
1a. [엄마 다리미] 결국 고장 나서 버렸어요?
2a. [아빠 대리] 자주 이용하신다고요?
3a. [오빠 더블베이스] 빌려 오셨어요?
4a. [아저씨 동네]에서 일하시는 거에요?
5a. [할머니 두루치기] 드시러 가세요?
6a. [이모 내과] 안 다니세요?
7a. [할아버지 나무] 결국 팔기로 하셨어요?
8a. [아빠 너구리] 잡으러 가세요?
9a. [아저씨 농담] 같은 거 좋아하세요?
10a. [할머니 누룽지] 드시러 가세요?

(b)
1b. (할머니,) 엄마(가) 다리미 결국 고장 나서 버렸어요?
2b. (엄마,) 아빠(가) 대리 자주 이용하신다고요?
3b. (언니,) 오빠(가) 더블베이스 빌려 오셨어요?
4b. (아줌마,) 아저씨가 동네에서 일하시는 거에요?
5b. (아빠,) 할머니(께서) 두루치기 드시러 가세요?
6b. (엄마,) 이모(는) 내과 안 다니세요?
7b. (엄마,) 할아버지(께서) 나무 결국 팔기로 하셨어요?
8b. (엄마,) 아빠(가) 너구리 잡으러 가세요?
9b. (아줌마,) 아저씨(가) 농담 같은 거 좋아하세요?
10b. (아빠,) 할머니(께서) 누룽지 드시러 가세요?

(c)
1c. 엄마, 다리미 결국 고장 나서 버렸어요?
2c. 아빠, 대리 자주 이용하신다고요?
3c. 오빠, 더블베이스 빌려 오셨어요?
4c. 아저씨, 동네에서 일하시는 거에요?
5c. 할머니, 두루치기 드시러 가세요?
6c. 이모, 내과 안 다니세요?
7c. 할아버지, 나무 결국 팔기로 하셨어요?
8c. 아빠, 너구리 잡으러 가세요?
9c. 아저씨, 농담 같은 거 좋아하세요?
10c. 할머니, 누룽지 드시러 가세요?

(d)
1d. 엄마, (pause) 다리미 결국 고장 나서 버렸어요?
2d. 아빠, (pause) 대리 자주 이용하신다고요?
3d. 오빠, (pause) 더블베이스 빌려 오셨어요?
4d. 아저씨, (pause) 동네에서 일하시는 거에요?
5d. 할머니, (pause) 두루치기 드시러 가세요?
6d. 이모, (pause) 내과 안 다니세요?
7d. 할아버지, (pause) 나무 결국 팔기로 하셨어요?
8d. 아빠, (pause) 너구리 잡으러 가세요?
9d. 아저씨, (pause) 농담 같은 거 좋아하세요?
10d. 할머니, (pause) 누룽지 드시러 가세요?

<SET B>

(a)
11a. [강주만] 화가를 만나기로 했다고?
12a. 최규나(가) 그네를 좋아한다고?
13a. (데이빗,) 발렌티나(가) (사라) 선생님하고 얘기는 해 본거야?
14a. 제니(가) 스키 타 본 적 있어?
15a. 어린 아가 사는 집에 두고 왔다고?
16a. 알프레도(가) 끼가 있으면 좋지 않겠어요?
17a. 페르난도(가) 자기만 중요하게 생각하는 것 같다고?

(b)
11b. [강주]만 화가를 만나기로 했다고?
12b. [현규]나 그네를 좋아한다고?
13b. [간호사]나 선생님하고 얘기는 해 본거야?
14b. [민제]니 스키 타봤던 애가?
15b. 지연이가 사는 집에 두고 왔다고?
16b. 아저씨도 끼가 있으면 좋지 않겠어요?
17b. 최금난도 자기만 중요하게 생각하는 것 같다고?

(c)
11c. [강주 만화가]를 만나기로 했다고?
12c. [현규 나그네]를 좋아한다고?
13c. [간호사 나라]에 대해서 얘기는 해 본거야?
14c. [민제 니스] 써본 적 있어?
15c. [지연이 가사]는 집에 두고 왔다고?
16c. [아저씨 도끼]가 있으면 좋지 않겠어요?
17c. [최금난 도자기]만 중요하게 생각하는 것 같다고?

(d)
11d. (지영아,) 강주(가) 만화가를 만나기로 했다고?
12d. 현규(가) 나그네를 좋아한다고?
13d. 간호사(가) 나 선생님하고 얘기는 해 본거야?
14d. 민제(가) 니스 써본 적 있어?
15d. (지수야,) 지연이(가) 가사는 집에 두고 왔다고?
16d. (아주머니,) 아저씨(가) 도끼가 있으면 좋지 않겠어요?
17d. 최금난(이) 도자기만 중요하게 생각하는 것 같다고?

(e)
11e. 강주, 만화가를 만나기로 했다고?
12e. 현규, 나그네를 좋아한다고?
13e. 간호사, 나 선생님하고 얘기는 해 본거야?
14e. 민제, 니스 써본 적 있어?
15e. 지연이, 가사는 집에 두고 왔다고?
16e. 아저씨, 도끼가 있으면 좋지 않겠어요?
17e. 최금난, (내가) 도자기만 중요하게 생각하는 것 같다고?

(f)
11f. 강주, (pause) 만화가를 만나기로 했다고?
12f. 현규, (pause) 나그네를 좋아한다고?
13f. 간호사, (pause) 나 선생님하고 얘기는 해 본거야?
14f. 민제, (pause) 니스 써본 적 있어?
15f. 지연이, (pause) 가사는 집에 두고 왔다고?
16f. 아저씨, (pause) 도끼가 있으면 좋지 않겠어요?
17f. 최금난, (pause) 도자기만 중요하게 생각하는 것 같다고?

Appendix B.

Bayesian models for category analyses

Table B.1: Output of the lenis devoicing model.

```
Family: bernoulli
  Links: mu = logit
Formula: devoiced ~ position * region + repetition + gender + (1 + position |
subject) + (1 + region | item_no)
   Data: lenis_data (Number of observations: 1874)
Samples: 4 chains, each with iter = 2000; warmup = 1000; thin = 1;
         total post-warmup samples = 4000

Group-Level Effects:
~item_no (Number of levels: 11)
                              Estimate Est.Error l-95% CI u-95% CI Eff.Sample Rhat
sd(Intercept)                     1.07      0.40     0.51     2.05       1739 1.00
sd(regionBusan)                   0.60      0.38     0.03     1.48       1510 1.00
sd(regionUlsan)                   0.69      0.43     0.06     1.66       1937 1.00
cor(Intercept,regionBusan)       -0.44      0.40    -0.96     0.53       3395 1.00
cor(Intercept,regionUlsan)        0.17      0.42    -0.65     0.87       3921 1.00
cor(regionBusan,regionUlsan)     -0.17      0.46    -0.91     0.74       2112 1.00

~subject (Number of levels: 30)
                              Estimate Est.Error l-95% CI u-95% CI Eff.Sample Rhat
sd(Intercept)                     1.23      0.30     0.71     1.89       1781 1.00
sd(positionMi)                    1.12      0.68     0.06     2.64        837 1.00
sd(positionPWi)                   0.51      0.36     0.02     1.38       1249 1.00
sd(positionAPi)                   1.66      0.41     0.95     2.55       1634 1.00
sd(positionIPi)                   2.09      0.70     0.90     3.68       1972 1.00
sd(positionP)                     1.44      1.30     0.05     4.76       3730 1.00
cor(Intercept,positionMi)         0.09      0.32    -0.53     0.68       4188 1.00
cor(Intercept,positionPWi)        0.16      0.35    -0.55     0.78       4955 1.00
cor(positionMi,positionPWi)       0.04      0.37    -0.66     0.71       2895 1.00
cor(Intercept,positionAPi)       -0.19      0.25    -0.63     0.37       1633 1.00
cor(positionMi,positionAPi)      -0.18      0.31    -0.75     0.45        845 1.01
cor(positionPWi,positionAPi)      0.04      0.35    -0.65     0.70        700 1.00
cor(Intercept,positionIPi)       -0.31      0.27    -0.75     0.28       2378 1.00
cor(positionMi,positionIPi)      -0.23      0.34    -0.81     0.51       1201 1.00
cor(positionPWi,positionIPi)      0.04      0.35    -0.64     0.70       1357 1.00
cor(positionAPi,positionIPi)      0.38      0.25    -0.17     0.78       2742 1.00
cor(Intercept,positionP)         -0.05      0.38    -0.74     0.67       6541 1.00
cor(positionMi,positionP)        -0.01      0.38    -0.71     0.70       5360 1.00
```

```
cor(positionPWi,positionP)      -0.01      0.38     -0.71      0.70        4640 1.00
cor(positionAPi,positionP)       0.03      0.38     -0.68      0.74        5747 1.00
cor(positionIPi,positionP)       0.04      0.38     -0.69      0.72        5463 1.00

Population-Level Effects:
                         Estimate Est.Error l-95% CI u-95% CI Eff.Sample Rhat
Intercept                    0.04      0.75    -1.49     1.48       2163 1.00
positionMi                  -1.23      0.68    -2.67     0.00       3136 1.00
positionPWi                 -0.25      0.47    -1.22     0.64       3232 1.00
positionAPi                  3.85      0.69     2.52     5.22       2293 1.00
positionIPi                  5.88      1.04     4.10     8.08       2327 1.00
positionP                   13.54      5.28     7.66    25.57        426 1.01
regionBusan                 -0.64      0.79    -2.20     0.86       2272 1.00
regionUlsan                 -0.56      0.78    -2.10     0.88       1900 1.00
repetition                  -0.27      0.19    -0.64     0.09       8261 1.00
gendermale                  -2.07      0.52    -3.16    -1.07       1905 1.00
positionMi:regionBusan       0.42      0.98    -1.62     2.28       3292 1.00
positionPWi:regionBusan      0.49      0.73    -0.96     1.92       3710 1.00
positionAPi:regionBusan     -1.17      0.91    -2.95     0.66       2843 1.00
positionIPi:regionBusan      0.44      1.23    -2.03     2.89       2942 1.00
positionP:regionBusan        1.05      3.08    -3.94     8.13       2703 1.00
positionMi:regionUlsan       0.40      0.95    -1.48     2.26       3109 1.00
positionPWi:regionUlsan      0.35      0.70    -1.02     1.72       2898 1.00
positionAPi:regionUlsan     -0.22      0.93    -2.01     1.68       2566 1.00
positionIPi:regionUlsan      0.08      1.23    -2.38     2.42       2557 1.00
positionP:regionUlsan        1.08      3.05    -4.20     7.99       2624 1.00
```

Table B.2: Pairwise comparisons of position by region for the lenis devoicing model.

```
$emmeans
region = Seoul:
 position emmean lower.HPD upper.HPD
 Mm        -1.37    -2.670    -0.112
 Mi        -2.55    -4.271    -0.897
 PWi       -1.60    -2.890    -0.306
 APi        2.47     0.945     4.125
 IPi        4.44     2.551     6.664
 P         11.06     5.768    21.744

region = Busan:
 position emmean lower.HPD upper.HPD
 Mm        -1.99    -3.297    -0.483
 Mi        -2.75    -4.819    -1.012
 PWi       -1.75    -3.225    -0.375
 APi        0.68    -0.756     2.109
 IPi        4.25     2.310     6.453
 P         11.43     5.665    22.356
```

```
region = Ulsan:
 position emmean lower.HPD upper.HPD
 Mm        -1.91    -3.382    -0.434
 Mi        -2.68    -4.744    -0.695
 PWi       -1.81    -3.403    -0.395
 APi        1.70     0.152     3.343
 IPi        3.99     2.063     6.156
 P         11.56     5.314    21.879

$contrasts
region = Seoul:
 contrast  estimate lower.HPD upper.HPD
 Mm - Mi      1.200   -0.0315    2.6035
 Mm - PWi     0.239   -0.6665    1.1899
 Mm - APi    -3.842   -5.2129   -2.5156
 Mm - IPi    -5.799   -7.9511   -4.0320
 Mm - P     -12.366  -22.5889   -6.8131
 Mi - PWi    -0.951   -2.3976    0.4102
 Mi - APi    -5.053   -6.7935   -3.1878
 Mi - IPi    -7.049   -9.4094   -4.8401
 Mi - P     -13.654  -24.1134   -7.9546
 PWi - APi   -4.085   -5.4780   -2.7384
 PWi - IPi   -6.073   -8.1619   -4.1514
 PWi - P    -12.615  -22.8110   -7.0548
 APi - IPi   -1.963   -4.2354   -0.0348
 APi - P     -8.614  -18.8280   -2.8037
 IPi - P     -6.574  -17.4207   -0.8095

region = Busan:
 contrast  estimate lower.HPD upper.HPD
 Mm - Mi      0.745   -0.8287    2.5325
 Mm - PWi    -0.242   -1.5409    0.9783
 Mm - APi    -2.668   -4.0347   -1.1908
 Mm - IPi    -6.241   -8.4446   -4.1781
 Mm - P     -13.420  -24.0745   -7.5091
 Mi - PWi    -0.993   -2.7897    0.8272
 Mi - APi    -3.405   -5.5999   -1.6525
 Mi - IPi    -7.030   -9.8432   -4.6944
 Mi - P     -14.239  -25.4168   -8.3066
 PWi - APi   -2.431   -3.9167   -0.9739
 PWi - IPi   -6.011   -8.2546   -3.9447
 PWi - P    -13.147  -24.1095   -7.4088
 APi - IPi   -3.554   -5.8523   -1.6851
 APi - P    -10.778  -21.7460   -5.0790
 IPi - P     -7.104  -18.5629   -1.1463
```

```
region = Ulsan:
 contrast  estimate lower.HPD upper.HPD
 Mm - Mi      0.774   -0.7833    2.4004
 Mm - PWi    -0.100   -1.3577    1.0049
 Mm - APi    -3.599   -5.1291   -2.1784
 Mm - IPi    -5.889   -8.0908   -4.0667
 Mm - P     -13.473  -24.1546   -7.9715
 Mi - PWi    -0.878   -2.6919    0.6583
 Mi - APi    -4.389   -6.4801   -2.4148
 Mi - IPi    -6.674   -9.3794   -4.4832
 Mi - P     -14.339  -24.7892   -8.2903
 PWi - APi   -3.520   -4.9141   -2.1746
 PWi - IPi   -5.799   -7.8602   -3.9775
 PWi - P    -13.370  -24.2091   -7.8558
 APi - IPi   -2.282   -4.2248   -0.5500
 APi - P     -9.856  -20.0711   -3.7652
 IPi - P     -7.512  -18.0733   -1.1962
```

Table B.3: Interaction contrasts for the lenis devoicing model.

position_pairwise	region_pairwise	estimate	lower.HPD	upper.HPD
Mm - Mi	Seoul - Busan	0.4526	-1.543	2.346
Mm - PWi	Seoul - Busan	0.4994	-0.900	1.969
Mm - APi	Seoul - Busan	-1.1701	-2.956	0.660
Mm - IPi	Seoul - Busan	0.4215	-2.094	2.806
Mm - P	Seoul - Busan	0.7869	-4.449	7.341
Mi - PWi	Seoul - Busan	0.0409	-1.991	2.115
Mi - APi	Seoul - Busan	-1.6183	-4.087	0.922
Mi - IPi	Seoul - Busan	-0.0273	-2.966	3.025
Mi - P	Seoul - Busan	0.3466	-5.376	6.875
PWi - APi	Seoul - Busan	-1.6598	-3.503	0.242
PWi - IPi	Seoul - Busan	-0.0610	-2.634	2.406
PWi - P	Seoul - Busan	0.2596	-5.267	6.894
APi - IPi	Seoul - Busan	1.6205	-0.889	4.132
APi - P	Seoul - Busan	1.9523	-3.364	8.841
IPi - P	Seoul - Busan	0.3289	-5.324	7.212
Mm - Mi	Seoul - Ulsan	0.4100	-1.492	2.239
Mm - PWi	Seoul - Ulsan	0.3465	-1.022	1.717
Mm - APi	Seoul - Ulsan	-0.2507	-2.022	1.635
Mm - IPi	Seoul - Ulsan	0.0977	-2.327	2.460
Mm - P	Seoul - Ulsan	0.8129	-4.855	7.062
Mi - PWi	Seoul - Ulsan	-0.0737	-1.968	1.892
Mi - APi	Seoul - Ulsan	-0.6309	-2.934	2.007
Mi - IPi	Seoul - Ulsan	-0.3223	-3.232	2.596
Mi - P	Seoul - Ulsan	0.4743	-5.219	7.199
PWi - APi	Seoul - Ulsan	-0.5748	-2.416	1.230
PWi - IPi	Seoul - Ulsan	-0.2806	-2.730	2.217
PWi - P	Seoul - Ulsan	0.5011	-5.341	6.732
APi - IPi	Seoul - Ulsan	0.3484	-2.140	2.950
APi - P	Seoul - Ulsan	1.0344	-4.943	7.363
IPi - P	Seoul - Ulsan	0.7693	-5.088	7.599
Mm - Mi	Busan - Ulsan	-0.0365	-2.069	2.199
Mm - PWi	Busan - Ulsan	-0.1397	-1.818	1.582
Mm - APi	Busan - Ulsan	0.9322	-1.030	2.990
Mm - IPi	Busan - Ulsan	-0.3453	-3.081	2.215
Mm - P	Busan - Ulsan	0.0582	-7.828	8.438
Mi - PWi	Busan - Ulsan	-0.0970	-2.261	1.989
Mi - APi	Busan - Ulsan	1.0066	-1.561	3.812
Mi - IPi	Busan - Ulsan	-0.3495	-3.582	2.944
Mi - P	Busan - Ulsan	0.0846	-7.960	8.714
PWi - APi	Busan - Ulsan	1.0852	-0.894	3.034
PWi - IPi	Busan - Ulsan	-0.2232	-3.017	2.302
PWi - P	Busan - Ulsan	0.1742	-7.388	9.205
APi - IPi	Busan - Ulsan	-1.2901	-3.809	1.333
APi - P	Busan - Ulsan	-0.8926	-9.182	7.436
IPi - P	Busan - Ulsan	0.4110	-7.311	9.649

Table B.4: Pairwise comparisons of region by position for the lenis devoicing model.

```
$emmeans
position = Mm:
 region emmean lower.HPD upper.HPD
 Seoul   -1.37    -2.670    -0.112
 Busan   -1.99    -3.297    -0.483
 Ulsan   -1.91    -3.382    -0.434

position = Mi:
 region emmean lower.HPD upper.HPD
 Seoul   -2.55    -4.271    -0.897
 Busan   -2.75    -4.819    -1.012
 Ulsan   -2.68    -4.744    -0.695

position = PWi:
 region emmean lower.HPD upper.HPD
 Seoul   -1.60    -2.890    -0.306
 Busan   -1.75    -3.225    -0.375
 Ulsan   -1.81    -3.403    -0.395

position = APi:
 region emmean lower.HPD upper.HPD
 Seoul    2.47     0.945     4.125
 Busan    0.68    -0.756     2.109
 Ulsan    1.70     0.152     3.343

position = IPi:
 region emmean lower.HPD upper.HPD
 Seoul    4.44     2.551     6.664
 Busan    4.25     2.310     6.453
 Ulsan    3.99     2.063     6.156

position = P:
 region emmean lower.HPD upper.HPD
 Seoul   11.06     5.768    21.744
 Busan   11.43     5.665    22.356
 Ulsan   11.56     5.314    21.879
```

```
$contrasts
position = Mm:
 contrast      estimate lower.HPD upper.HPD
 Seoul - Busan   0.6491    -0.852     2.209
 Seoul - Ulsan   0.5299    -0.876     2.108
 Busan - Ulsan  -0.1110    -1.856     1.635

position = Mi:
 contrast      estimate lower.HPD upper.HPD
 Seoul - Busan   0.2121    -1.983     2.379
 Seoul - Ulsan   0.1203    -1.814     2.391
 Busan - Ulsan  -0.0444    -2.397     2.440

position = PWi:
 contrast      estimate lower.HPD upper.HPD
 Seoul - Busan   0.1447    -1.420     1.784
 Seoul - Ulsan   0.2011    -1.248     1.906
 Busan - Ulsan   0.0475    -1.769     1.832

position = APi:
 contrast      estimate lower.HPD upper.HPD
 Seoul - Busan   1.7967     0.108     3.720
 Seoul - Ulsan   0.7862    -1.037     2.585
 Busan - Ulsan  -1.0178    -2.865     0.998

position = IPi:
 contrast      estimate lower.HPD upper.HPD
 Seoul - Busan   0.2076    -2.100     2.809
 Seoul - Ulsan   0.4705    -2.097     2.688
 Busan - Ulsan   0.2701    -2.195     2.878
position = P:
 contrast      estimate lower.HPD upper.HPD
 Seoul - Busan  -0.1492    -6.408     5.592
 Seoul - Ulsan  -0.3277    -6.366     5.930
 Busan - Ulsan  -0.1461    -8.170     8.407
```

Table B.5: Output of the denasalisation model.

```
Family: bernoulli
  Links: mu = logit
Formula: denasalised ~ position * region + repetition + gender + (1 + position |
subject) + (1 + region | item_no)
   Data: nasal_data (Number of observations: 2263)
Samples: 4 chains, each with iter = 2000; warmup = 1000; thin = 1;
         total post-warmup samples = 4000

Group-Level Effects:
~item_no (Number of levels: 9)
                             Estimate Est.Error l-95% CI u-95% CI Eff.Sample Rhat
sd(Intercept)                    0.65      0.25     0.30     1.29       1201 1.00
sd(regionBusan)                  0.45      0.29     0.02     1.13       1283 1.01
sd(regionUlsan)                  0.33      0.24     0.02     0.92       1790 1.00
cor(Intercept,regionBusan)      -0.19      0.42    -0.86     0.73       3826 1.00
cor(Intercept,regionUlsan)       0.07      0.46    -0.78     0.88       4531 1.00
cor(regionBusan,regionUlsan)     0.16      0.49    -0.81     0.92       2290 1.00

~subject (Number of levels: 30)
                             Estimate Est.Error l-95% CI u-95% CI Eff.Sample Rhat
sd(Intercept)                    1.09      0.24     0.63     1.60       1743 1.00
sd(positionMi)                   1.35      1.10     0.07     4.14       3774 1.00
sd(positionPWi)                  0.72      0.60     0.03     2.29       2113 1.00
sd(positionAPi)                  0.48      0.31     0.02     1.18        967 1.00
sd(positionIPi)                  0.79      0.36     0.11     1.52        776 1.00
sd(positionP)                    0.59      0.31     0.05     1.21        788 1.00
cor(Intercept,positionMi)       -0.05      0.37    -0.72     0.66       7166 1.00
cor(Intercept,positionPWi)       0.01      0.37    -0.69     0.68       5155 1.00
cor(positionMi,positionPWi)      0.01      0.38    -0.71     0.71       4207 1.00
cor(Intercept,positionAPi)      -0.02      0.34    -0.64     0.67       3930 1.00
cor(positionMi,positionAPi)     -0.00      0.38    -0.68     0.70       2003 1.00
cor(positionPWi,positionAPi)     0.05      0.38    -0.69     0.73       2003 1.00
cor(Intercept,positionIPi)       0.18      0.30    -0.42     0.73       3212 1.00
cor(positionMi,positionIPi)     -0.04      0.37    -0.71     0.68       1501 1.00
cor(positionPWi,positionIPi)     0.02      0.37    -0.70     0.72       1864 1.00
cor(positionAPi,positionIPi)     0.11      0.36    -0.63     0.75       1447 1.00
cor(Intercept,positionP)        -0.08      0.33    -0.66     0.58       2227 1.00
cor(positionMi,positionP)       -0.00      0.38    -0.70     0.72       1509 1.00
cor(positionPWi,positionP)       0.02      0.37    -0.69     0.69       1705 1.00
cor(positionAPi,positionP)       0.02      0.36    -0.69     0.69       1980 1.00
cor(positionIPi,positionP)       0.03      0.35    -0.66     0.67       1895 1.00

Population-Level Effects:
                        Estimate Est.Error l-95% CI u-95% CI Eff.Sample Rhat
Intercept                  -5.41      1.29    -8.21    -3.29        970 1.00
positionMi                 -3.31      3.12   -11.38     0.89       1747 1.00
```

```
positionPWi                  1.32       1.31     -0.89     4.19         987 1.00
positionAPi                  7.46       1.26      5.53    10.32         895 1.00
positionIPi                  7.66       1.28      5.66    10.56         919 1.00
positionP                    6.68       1.25      4.77     9.50         851 1.00
regionBusan                 -1.77       1.25     -4.49     0.55        1384 1.00
regionUlsan                 -1.34       1.13     -3.58     0.84        1758 1.00
repetition                   0.02       0.15     -0.27     0.31        5703 1.00
gendermale                  -0.97       0.43     -1.84    -0.11        1860 1.00
positionMi:regionBusan      -0.74       3.13     -7.89     4.38        2768 1.00
positionPWi:regionBusan     -2.37       2.79     -9.36     1.57        1454 1.00
positionAPi:regionBusan     -0.83       1.22     -3.14     1.65        1567 1.00
positionIPi:regionBusan      0.04       1.25     -2.32     2.63        1481 1.00
positionP:regionBusan        1.91       1.23     -0.42     4.50        1461 1.00
positionMi:regionUlsan      -0.95       3.40     -8.46     4.19        1243 1.00
positionPWi:regionUlsan      0.87       1.27     -1.57     3.40        2718 1.00
positionAPi:regionUlsan     -1.06       1.08     -3.18     1.08        1949 1.00
positionIPi:regionUlsan     -0.36       1.11     -2.52     1.84        1982 1.00
positionP:regionUlsan        0.74       1.08     -1.38     2.87        1841 1.00
```

Table B.6: Pairwise comparisons of position by region for the denasalisation model.

```
$emmeans
region = Seoul:
 position   emmean lower.HPD upper.HPD
 Mm        -5.6992   -8.1704   -3.5662
 Mi        -8.5397  -15.3197   -4.4659
 PWi       -4.4824   -6.1495   -3.0614
 APi        1.5872    0.5960    2.5983
 IPi        1.7777    0.6758    3.0311
 P          0.8005   -0.1304    1.8457

region = Busan:
 position   emmean lower.HPD upper.HPD
 Mm        -7.5107  -10.9845   -4.8432
 Mi       -10.9845  -19.8181   -4.7507
 PWi       -8.1664  -14.0756   -4.5116
 APi       -1.0057   -1.9735    0.0744
 IPi        0.0556   -1.1286    1.3026
 P          0.9473   -0.0673    1.9846

region = Ulsan:
 position   emmean lower.HPD upper.HPD
 Mm        -7.0732  -10.1149   -4.4702
 Mi       -10.7765  -19.7116   -4.9606
 PWi       -4.8989   -6.8954   -3.0777
 APi       -0.8093   -1.7815    0.2039
 IPi        0.0702   -1.0710    1.2093
 P          0.2039   -0.8102    1.1579

$contrasts
region = Seoul:
 contrast  estimate lower.HPD upper.HPD
 Mm - Mi      2.672   -1.7126   9.37574
 Mm - PWi    -1.186   -3.9292   1.05568
 Mm - APi    -7.300  -10.0309  -5.38529
 Mm - IPi    -7.506  -10.2576  -5.52063
 Mm - P      -6.514   -9.0771  -4.59670
 Mi - PWi    -4.018  -11.0338  -0.06323
 Mi - APi   -10.145  -16.5857  -5.80171
 Mi - IPi   -10.362  -16.8801  -6.13204
 Mi - P      -9.334  -16.0824  -5.26636
 PWi - APi   -6.069   -7.6816  -4.74728
 PWi - IPi   -6.269   -7.8977  -4.80843
 PWi - P     -5.289   -6.8586  -4.00174
 APi - IPi   -0.184   -1.1528   0.74813
 APi - P      0.782   -0.0917   1.57048
 IPi - P      0.961    0.0806   1.91766
```

```
region = Busan:
 contrast  estimate lower.HPD upper.HPD
 Mm - Mi      3.395   -2.8307  12.93570
 Mm - PWi     0.669   -4.5064   7.02051
 Mm - APi    -6.496   -9.6459  -3.69760
 Mm - IPi    -7.571  -10.8077  -4.79874
 Mm - P      -8.474  -11.7497  -5.67378
 Mi - PWi    -2.696  -13.0302   7.10448
 Mi - APi    -9.958  -18.7778  -3.96768
 Mi - IPi   -11.026  -19.9455  -4.96347
 Mi - P     -11.944  -20.7470  -5.90626
 PWi - APi   -7.110  -13.0477  -3.67822
 PWi - IPi   -8.216  -14.0093  -4.47279
 PWi - P     -9.103  -14.8314  -5.39243
 APi - IPi   -1.049   -1.9520  -0.19025
 APi - P     -1.952   -2.7473  -1.16642
 IPi - P     -0.897   -1.7688   0.00599

region = Ulsan:
 contrast  estimate lower.HPD upper.HPD
 Mm - Mi      3.556   -2.6897  12.56579
 Mm - PWi    -2.095   -5.4286   0.80944
 Mm - APi    -6.245   -9.2632  -3.85345
 Mm - IPi    -7.152  -10.1717  -4.68078
 Mm - P      -7.263  -10.3505  -4.85889
 Mi - PWi    -5.761  -14.4611   0.68697
 Mi - APi    -9.959  -18.4598  -4.06363
 Mi - IPi   -10.868  -19.3691  -4.64656
 Mi - P     -10.986  -19.6063  -4.92232
 PWi - APi   -4.107   -6.0528  -2.52329
 PWi - IPi   -5.031   -6.8786  -3.22004
 PWi - P     -5.116   -7.0847  -3.48654
 APi - IPi   -0.885   -1.7229  -0.06109
 APi - P     -1.013   -1.7880  -0.32724
 IPi - P     -0.130   -0.9830   0.73157
```

Table B.7: Interaction contrasts for the denasalisation model.

position_pairwise	region_pairwise	estimate	lower.HPD	upper.HPD
Mm - Mi	Seoul - Busan	-0.3574	-6.637	5.0918
Mm - PWi	Seoul - Busan	-1.8774	-7.731	2.3057
Mm - APi	Seoul - Busan	-0.8652	-3.175	1.5864
Mm - IPi	Seoul - Busan	0.0139	-2.372	2.5542
Mm - P	Seoul - Busan	1.8751	-0.357	4.5067
Mi - PWi	Seoul - Busan	-1.5998	-10.408	6.2001
Mi - APi	Seoul - Busan	-0.4308	-6.352	5.8907
Mi - IPi	Seoul - Busan	0.4551	-5.125	6.9178
Mi - P	Seoul - Busan	2.3070	-3.353	8.9809
PWi - APi	Seoul - Busan	1.0401	-2.921	6.6361
PWi - IPi	Seoul - Busan	1.9274	-1.897	7.8183
PWi - P	Seoul - Busan	3.7994	0.136	9.7357
APi - IPi	Seoul - Busan	0.8708	-0.361	2.1696
APi - P	Seoul - Busan	2.7358	1.629	3.8887
IPi - P	Seoul - Busan	1.8739	0.656	3.2572
Mm - Mi	Seoul - Ulsan	-0.5549	-7.006	5.2988
Mm - PWi	Seoul - Ulsan	0.8736	-1.493	3.4365
Mm - APi	Seoul - Ulsan	-1.0612	-3.286	0.9599
Mm - IPi	Seoul - Ulsan	-0.3770	-2.466	1.8787
Mm - P	Seoul - Ulsan	0.7106	-1.448	2.7760
Mi - PWi	Seoul - Ulsan	1.4328	-4.530	8.8765
Mi - APi	Seoul - Ulsan	-0.4731	-6.025	6.5851
Mi - IPi	Seoul - Ulsan	0.1746	-5.512	7.3806
Mi - P	Seoul - Ulsan	1.3385	-4.383	8.4240
PWi - APi	Seoul - Ulsan	-1.9587	-4.023	0.0122
PWi - IPi	Seoul - Ulsan	-1.2578	-3.338	0.9446
PWi - P	Seoul - Ulsan	-0.1743	-2.186	1.8844
APi - IPi	Seoul - Ulsan	0.7058	-0.471	1.9716
APi - P	Seoul - Ulsan	1.7973	0.625	2.8314
IPi - P	Seoul - Ulsan	1.0929	-0.102	2.3110
Mm - Mi	Busan - Ulsan	-0.1291	-8.925	8.6168
Mm - PWi	Busan - Ulsan	2.7853	-2.020	9.1206
Mm - APi	Busan - Ulsan	-0.2294	-3.446	2.5568
Mm - IPi	Busan - Ulsan	-0.4109	-3.451	2.6488
Mm - P	Busan - Ulsan	-1.1449	-4.320	1.7908
Mi - PWi	Busan - Ulsan	3.1474	-6.451	14.7259
Mi - APi	Busan - Ulsan	-0.1055	-9.161	8.8963
Mi - IPi	Busan - Ulsan	-0.2628	-9.015	9.0615
Mi - P	Busan - Ulsan	-1.0843	-9.973	8.0336
PWi - APi	Busan - Ulsan	-2.9869	-9.369	0.8043
PWi - IPi	Busan - Ulsan	-3.1720	-9.045	1.1408
PWi - P	Busan - Ulsan	-3.9439	-10.206	0.2140
APi - IPi	Busan - Ulsan	-0.1526	-1.324	1.0476
APi - P	Busan - Ulsan	-0.9376	-1.991	0.2099
IPi - P	Busan - Ulsan	-0.7790	-1.950	0.4729

Table B.8: Pairwise comparisons of region by position for the denasalisation model.

```
$emmeans
region = Seoul:
 position   emmean lower.HPD upper.HPD
 Mm        -5.6992   -8.1704   -3.5662
 Mi        -8.5397  -15.3197   -4.4659
 PWi       -4.4824   -6.1495   -3.0614
 APi        1.5872    0.5960    2.5983
 IPi        1.7777    0.6758    3.0311
 P          0.8005   -0.1304    1.8457

region = Busan:
 position   emmean lower.HPD upper.HPD
 Mm        -7.5107  -10.9845   -4.8432
 Mi       -10.9845  -19.8181   -4.7507
 PWi       -8.1664  -14.0756   -4.5116
 APi       -1.0057   -1.9735    0.0744
 IPi        0.0556   -1.1286    1.3026
 P          0.9473   -0.0673    1.9846

region = Ulsan:
 position   emmean lower.HPD upper.HPD
 Mm        -7.0732  -10.1149   -4.4702
 Mi       -10.7765  -19.7116   -4.9606
 PWi       -4.8989   -6.8954   -3.0777
 APi       -0.8093   -1.7815    0.2039
 IPi        0.0702   -1.0710    1.2093
 P          0.2039   -0.8102    1.1579

$contrasts
region = Seoul:
 contrast  estimate lower.HPD upper.HPD
 Mm - Mi      2.672   -1.7126   9.37574
 Mm - PWi    -1.186   -3.9292   1.05568
 Mm - APi    -7.300  -10.0309  -5.38529
 Mm - IPi    -7.506  -10.2576  -5.52063
 Mm - P      -6.514   -9.0771  -4.59670
 Mi - PWi    -4.018  -11.0338  -0.06323
 Mi - APi   -10.145  -16.5857  -5.80171
 Mi - IPi   -10.362  -16.8801  -6.13204
 Mi - P      -9.334  -16.0824  -5.26636
 PWi - APi   -6.069   -7.6816  -4.74728
 PWi - IPi   -6.269   -7.8977  -4.80843
 PWi - P     -5.289   -6.8586  -4.00174
 APi - IPi   -0.184   -1.1528   0.74813
 APi - P      0.782   -0.0917   1.57048
 IPi - P      0.961    0.0806   1.91766
```

```
region = Busan:
 contrast  estimate lower.HPD upper.HPD
 Mm - Mi      3.395   -2.8307  12.93570
 Mm - PWi     0.669   -4.5064   7.02051
 Mm - APi    -6.496   -9.6459  -3.69760
 Mm - IPi    -7.571  -10.8077  -4.79874
 Mm - P      -8.474  -11.7497  -5.67378
 Mi - PWi    -2.696  -13.0302   7.10448
 Mi - APi    -9.958  -18.7778  -3.96768
 Mi - IPi   -11.026  -19.9455  -4.96347
 Mi - P     -11.944  -20.7470  -5.90626
 PWi - APi   -7.110  -13.0477  -3.67822
 PWi - IPi   -8.216  -14.0093  -4.47279
 PWi - P     -9.103  -14.8314  -5.39243
 APi - IPi   -1.049   -1.9520  -0.19025
 APi - P     -1.952   -2.7473  -1.16642
 IPi - P     -0.897   -1.7688   0.00599

region = Ulsan:
 contrast  estimate lower.HPD upper.HPD
 Mm - Mi      3.556   -2.6897  12.56579
 Mm - PWi    -2.095   -5.4286   0.80944
 Mm - APi    -6.245   -9.2632  -3.85345
 Mm - IPi    -7.152  -10.1717  -4.68078
 Mm - P      -7.263  -10.3505  -4.85889
 Mi - PWi    -5.761  -14.4611   0.68697
 Mi - APi    -9.959  -18.4598  -4.06363
 Mi - IPi   -10.868  -19.3691  -4.64656
 Mi - P     -10.986  -19.6063  -4.92232
 PWi - APi   -4.107   -6.0528  -2.52329
 PWi - IPi   -5.031   -6.8786  -3.22004
 PWi - P     -5.116   -7.0847  -3.48654
 APi - IPi   -0.885   -1.7229  -0.06109
 APi - P     -1.013   -1.7880  -0.32724
 IPi - P     -0.130   -0.9830   0.73157
```

Table B.9: Output of the nasal devoicing model.

```
Family: bernoulli
  Links: mu = logit
Formula: devoiced ~ position * region + repetition + gender + (1 + position |
subject) + (1 + region | item_no)
   Data: nasal_data (Number of observations: 2263)
Samples: 4 chains, each with iter = 2000; warmup = 1000; thin = 1;
         total post-warmup samples = 4000

Group-Level Effects:
~item_no (Number of levels: 9)
                              Estimate Est.Error l-95% CI u-95% CI Eff.Sample Rhat
sd(Intercept)                     1.27      0.45     0.61     2.36       2147 1.00
sd(regionBusan)                   1.00      0.43     0.28     1.98       1889 1.00
sd(regionUlsan)                   0.82      0.47     0.08     1.89       1641 1.00
cor(Intercept,regionBusan)       -0.61      0.31    -0.96     0.15       3040 1.00
cor(Intercept,regionUlsan)       -0.37      0.41    -0.94     0.58       4243 1.00
cor(regionBusan,regionUlsan)      0.52      0.38    -0.43     0.97       2835 1.00


~subject (Number of levels: 30)
                              Estimate Est.Error l-95% CI u-95% CI Eff.Sample Rhat
sd(Intercept)                     0.77      0.36     0.08     1.54        800 1.00
sd(positionMi)                    1.40      1.17     0.05     4.21       3668 1.00
sd(positionPWi)                   1.34      1.10     0.05     4.06       2309 1.00
sd(positionAPi)                   2.17      1.16     0.34     4.89       1965 1.00
sd(positionIPi)                   1.27      0.74     0.09     2.91       1154 1.00
sd(positionP)                     0.46      0.32     0.02     1.15       1100 1.00
cor(Intercept,positionMi)        -0.04      0.38    -0.74     0.68       9557 1.00
cor(Intercept,positionPWi)        0.04      0.37    -0.67     0.74       8108 1.00
cor(positionMi,positionPWi)      -0.00      0.38    -0.71     0.70       4783 1.00
cor(Intercept,positionAPi)        0.14      0.34    -0.56     0.76       3114 1.00
cor(positionMi,positionAPi)      -0.02      0.38    -0.73     0.70       3229 1.00
cor(positionPWi,positionAPi)      0.16      0.36    -0.60     0.77       2575 1.00
cor(Intercept,positionIPi)        0.25      0.34    -0.47     0.81       2009 1.00
cor(positionMi,positionIPi)      -0.02      0.38    -0.72     0.71       2643 1.00
cor(positionPWi,positionIPi)      0.08      0.37    -0.65     0.73       2160 1.00
cor(positionAPi,positionIPi)      0.16      0.35    -0.55     0.77       2825 1.00
cor(Intercept,positionP)         -0.18      0.40    -0.83     0.62       3667 1.00
cor(positionMi,positionP)         0.03      0.38    -0.69     0.74       2706 1.00
cor(positionPWi,positionP)       -0.02      0.38    -0.71     0.70       2548 1.00
cor(positionAPi,positionP)       -0.02      0.37    -0.72     0.69       2395 1.00
cor(positionIPi,positionP)        0.14      0.38    -0.63     0.80       1554 1.00
```

```
Population-Level Effects:
                        Estimate Est.Error l-95% CI u-95% CI Eff.Sample Rhat
Intercept                  -6.56      1.42    -9.72    -4.12       2082 1.00
positionMi                 -3.06      3.86   -11.77     1.43        924 1.01
positionPWi                -0.67      1.81    -4.50     2.51       2819 1.00
positionAPi                 1.07      1.63    -2.30     4.24       2324 1.00
positionIPi                 3.11      1.42     0.48     6.18       2167 1.00
positionP                   4.42      1.30     2.32     7.40       2219 1.00
regionBusan                -0.82      1.43    -3.83     1.82       2586 1.00
regionUlsan                -1.31      1.46    -4.49     1.33       2814 1.00
repetition                 -0.02      0.24    -0.48     0.46      11386 1.00
gendermale                 -1.07      0.42    -1.92    -0.26       3652 1.00
positionMi:regionBusan     -0.90      3.30    -8.60     4.61       3639 1.00
positionPWi:regionBusan    -1.50      2.96    -8.39     2.87       2268 1.00
positionAPi:regionBusan    -2.86      2.90    -9.74     1.37       2197 1.00
positionIPi:regionBusan     0.82      1.52    -1.97     4.01       2834 1.00
positionP:regionBusan       2.70      1.42     0.21     5.74       2548 1.00
positionMi:regionUlsan     -0.77      3.18    -7.79     4.42       3049 1.00
positionPWi:regionUlsan    -1.40      3.03    -8.91     3.19       2380 1.00
positionAPi:regionUlsan    -2.62      2.87    -9.59     1.73       3013 1.00
positionIPi:regionUlsan     0.22      1.55    -2.72     3.47       3043 1.00
positionP:regionUlsan       2.19      1.45    -0.39     5.42       2704 1.00
```

Table B.10: Pairwise comparisons of position by region for the nasal devoicing model.

```
$emmeans
region = Seoul:
 position  emmean lower.HPD upper.HPD
 Mm        -6.993     -9.98    -4.605
 Mi        -9.462    -16.75    -5.097
 PWi       -7.549    -11.53    -4.645
 APi       -5.843     -9.53    -3.320
 IPi       -3.952     -5.92    -2.243
 P         -2.673     -3.92    -1.497

region = Busan:
 position  emmean lower.HPD upper.HPD
 Mm        -7.790    -11.33    -4.750
 Mi       -11.167    -19.96    -4.180
 PWi       -9.540    -16.20    -4.652
 APi       -9.218    -15.75    -4.811
 IPi       -3.905     -6.17    -2.230
 P         -0.816     -1.75     0.157

region = Ulsan:
 position  emmean lower.HPD upper.HPD
 Mm        -8.285    -12.19    -5.229
 Mi       -11.499    -21.54    -5.286
 PWi       -9.941    -17.37    -5.144
 APi       -9.484    -15.94    -4.988
 IPi       -4.987     -7.50    -2.620
 P         -1.818     -3.03    -0.730

$contrasts
region = Seoul:
 contrast  estimate lower.HPD upper.HPD
 Mm - Mi      2.331     -2.49     9.668
 Mm - PWi     0.529     -2.83     4.083
 Mm - APi    -1.115     -4.07     2.429
 Mm - IPi    -3.018     -5.90    -0.361
 Mm - P      -4.280     -7.07    -2.106
 Mi - PWi    -1.825     -9.49     4.180
 Mi - APi    -3.524    -11.13     2.319
 Mi - IPi    -5.460    -12.61    -0.758
 Mi - P      -6.722    -14.14    -2.605
 PWi - APi   -1.655     -6.25     2.434
 PWi - IPi   -3.591     -7.60    -0.380
 PWi - P     -4.836     -8.76    -2.121
 APi - IPi   -1.864     -5.43     1.026
 APi - P     -3.127     -6.61    -0.812
 IPi - P     -1.243     -2.83     0.266
```

```
region = Busan:
 contrast  estimate lower.HPD upper.HPD
 Mm - Mi      3.189     -3.82    12.644
 Mm - PWi     1.747     -3.58     8.346
 Mm - APi     1.420     -3.37     8.790
 Mm - IPi    -3.844     -7.47    -0.291
 Mm - P      -6.976    -10.34    -3.827
 Mi - PWi    -1.384    -12.42     9.053
 Mi - APi    -1.711    -12.58     8.445
 Mi - IPi    -7.189    -16.93    -0.532
 Mi - P     -10.350    -19.52    -3.826
 PWi - APi   -0.359     -9.06     7.838
 PWi - IPi   -5.585    -12.67    -0.444
 PWi - P     -8.722    -15.49    -4.041
 APi - IPi   -5.250    -12.23    -0.710
 APi - P     -8.386    -14.91    -4.143
 IPi - P     -3.062     -5.08    -1.570

region = Ulsan:
 contrast  estimate lower.HPD upper.HPD
 Mm - Mi      3.078     -3.77    12.539
 Mm - PWi     1.580     -3.96     9.002
 Mm - APi     1.161     -4.12     8.035
 Mm - IPi    -3.250     -6.97     0.636
 Mm - P      -6.429    -10.12    -3.440
 Mi - PWi    -1.411    -13.23     8.544
 Mi - APi    -1.774    -12.99     7.874
 Mi - IPi    -6.425    -15.88     0.520
 Mi - P      -9.631    -19.31    -3.431
 PWi - APi   -0.442     -9.02     8.040
 PWi - IPi   -4.853    -11.99     0.843
 PWi - P     -8.148    -15.35    -3.315
 APi - IPi   -4.444    -10.92     0.478
 APi - P     -7.649    -13.85    -3.285
 IPi - P     -3.141     -5.51    -1.216
```

Table B.11: Interaction contrasts for the nasal devoicing model.

position_pairwise	region_pairwise	estimate	lower.HPD	upper.HPD
Mm - Mi	Seoul - Busan	-0.63114	-7.223	5.34
Mm - PWi	Seoul - Busan	-1.08098	-7.874	3.05
Mm - APi	Seoul - Busan	-2.42346	-9.134	1.68
Mm - IPi	Seoul - Busan	0.76490	-2.153	3.77
Mm - P	Seoul - Busan	2.59108	0.143	5.58
Mi - PWi	Seoul - Busan	-0.55182	-9.773	7.77
Mi - APi	Seoul - Busan	-1.89934	-10.874	6.57
Mi - IPi	Seoul - Busan	1.43401	-4.932	8.48
Mi - P	Seoul - Busan	3.29185	-2.744	10.47
PWi - APi	Seoul - Busan	-1.31319	-9.914	6.26
PWi - IPi	Seoul - Busan	1.92461	-3.074	8.72
PWi - P	Seoul - Busan	3.73909	-0.821	10.75
APi - IPi	Seoul - Busan	3.27051	-1.240	9.98
APi - P	Seoul - Busan	5.15823	0.432	11.19
IPi - P	Seoul - Busan	1.84783	0.048	3.87
Mm - Mi	Seoul - Ulsan	-0.47346	-6.838	5.03
Mm - PWi	Seoul - Ulsan	-0.97354	-7.964	3.64
Mm - APi	Seoul - Ulsan	-2.19376	-8.555	2.55
Mm - IPi	Seoul - Ulsan	0.19983	-2.856	3.26
Mm - P	Seoul - Ulsan	2.06323	-0.545	5.15
Mi - PWi	Seoul - Ulsan	-0.53978	-9.523	7.70
Mi - APi	Seoul - Ulsan	-1.81611	-9.880	6.38
Mi - IPi	Seoul - Ulsan	0.69404	-5.161	7.93
Mi - P	Seoul - Ulsan	2.65648	-3.592	9.25
PWi - APi	Seoul - Ulsan	-1.24247	-9.309	7.04
PWi - IPi	Seoul - Ulsan	1.25162	-4.116	8.12
PWi - P	Seoul - Ulsan	3.15556	-1.828	9.94
APi - IPi	Seoul - Ulsan	2.45627	-2.596	9.13
APi - P	Seoul - Ulsan	4.42330	-0.181	10.86
IPi - P	Seoul - Ulsan	1.93810	-0.254	4.17
Mm - Mi	Busan - Ulsan	0.05838	-9.167	8.91
Mm - PWi	Busan - Ulsan	0.11067	-8.344	8.45
Mm - APi	Busan - Ulsan	0.19542	-7.530	8.64
Mm - IPi	Busan - Ulsan	-0.62400	-4.530	3.35
Mm - P	Busan - Ulsan	-0.53727	-4.110	3.51
Mi - PWi	Busan - Ulsan	-0.00233	-12.756	11.74
Mi - APi	Busan - Ulsan	0.06353	-12.600	11.03
Mi - IPi	Busan - Ulsan	-0.73810	-10.511	8.20
Mi - P	Busan - Ulsan	-0.61608	-10.605	8.09
PWi - APi	Busan - Ulsan	0.19292	-10.128	12.15
PWi - IPi	Busan - Ulsan	-0.70336	-8.755	8.35
PWi - P	Busan - Ulsan	-0.65396	-9.831	7.39
APi - IPi	Busan - Ulsan	-0.87792	-8.871	7.40
APi - P	Busan - Ulsan	-0.78875	-8.568	7.18
IPi - P	Busan - Ulsan	0.06463	-2.258	2.49

Table B.12: Pairwise comparisons of region by position for the nasal devoicing model.

```
$emmeans
position = Mm:
 region  emmean lower.HPD upper.HPD
 Seoul   -6.993     -9.98    -4.605
 Busan   -7.790    -11.33    -4.750
 Ulsan   -8.285    -12.19    -5.229

position = Mi:
 region  emmean lower.HPD upper.HPD
 Seoul   -9.462    -16.75    -5.097
 Busan  -11.167    -19.96    -4.180
 Ulsan  -11.499    -21.54    -5.286

position = PWi:
 region  emmean lower.HPD upper.HPD
 Seoul   -7.549    -11.53    -4.645
 Busan   -9.540    -16.20    -4.652
 Ulsan   -9.941    -17.37    -5.144

position = APi:
 region  emmean lower.HPD upper.HPD
 Seoul   -5.843     -9.53    -3.320
 Busan   -9.218    -15.75    -4.811
 Ulsan   -9.484    -15.94    -4.988

position = IPi:
 region  emmean lower.HPD upper.HPD
 Seoul   -3.952     -5.92    -2.243
 Busan   -3.905     -6.17    -2.230
 Ulsan   -4.987     -7.50    -2.620

position = P:
 region  emmean lower.HPD upper.HPD
 Seoul   -2.673     -3.92    -1.497
 Busan   -0.816     -1.75     0.157
 Ulsan   -1.818     -3.03    -0.730
```

```
$contrasts
position = Mm:
 contrast      estimate lower.HPD upper.HPD
 Seoul - Busan   0.7452    -1.958     3.633
 Seoul - Ulsan   1.2389    -1.345     4.455
 Busan - Ulsan   0.5219    -3.405     4.267

position = Mi:
 contrast      estimate lower.HPD upper.HPD
 Seoul - Busan   1.4307    -5.063     8.285
 Seoul - Ulsan   1.7756    -3.988     8.888
 Busan - Ulsan   0.4197    -9.195     9.405

position = PWi:
 contrast      estimate lower.HPD upper.HPD
 Seoul - Busan   1.8872    -2.809     8.672
 Seoul - Ulsan   2.2935    -2.715     8.915
 Busan - Ulsan   0.3756    -7.877     9.254

position = APi:
 contrast      estimate lower.HPD upper.HPD
 Seoul - Busan   3.2498    -1.363     9.505
 Seoul - Ulsan   3.5558    -0.927    10.038
 Busan - Ulsan   0.2301    -7.632     8.293

position = IPi:
 contrast      estimate lower.HPD upper.HPD
 Seoul - Busan  -0.0159    -2.087     2.212
 Seoul - Ulsan   1.0553    -1.175     3.458
 Busan - Ulsan   1.0898    -1.553     3.695

position = P:
 contrast      estimate lower.HPD upper.HPD
 Seoul - Busan  -1.8574    -3.134    -0.627
 Seoul - Ulsan  -0.8640    -2.160     0.334
 Busan - Ulsan   1.0053    -0.195     2.222
```

Appendix C.

Raw proportions of category

Table C.1: Overall proportions and token counts of nasal category.

Region	**Seoul**		**Busan**		**Ulsan**	
Category	Proportion	Tokens	Proportion	Tokens	Proportion	Tokens
(1) N	0.56	439	0.70	511	0.73	549
(2) N^D	0.12	93	0.07	53	0.07	49
(3) D	0.27	208	0.14	102	0.16	121
(4) T	0.03	22	0.05	39	0.03	19
(5) T^H	0.02	17	0.04	26	0.02	14

Table C.2: Proportions of nasal category by position for Seoul.

	Mm	**Mi**	**PWi**	**APi**	**IPi**	**P**
(1) N	1.00	1.00	0.969	0.204	0.213	0.362
(2) N^D	-	-	-	0.183	0.248	0.184
(3) D	-	-	0.025	0.556	0.468	0.339
(4) T	-	-	-	0.021	0.057	0.063
(5) T^H	-	-	0.006	0.035	0.014	0.052

Table C.3: Proportions of nasal category by position for Busan.

	Mm	**Mi**	**PWi**	**APi**	**IPi**	**P**
(1) N	1.00	1.00	1	0.684	0.523	0.321
(2) N^D	-	-	-	0.155	0.129	0.075
(3) D	-	-	-	0.161	0.28	0.252
(4) T	-	-	-	-	0.053	0.201
(5) T^H	-	-	-	-	0.015	0.151

Table C.4: Proportions of nasal category by position for Ulsan.

	Mm	**Mi**	**PWi**	**APi**	**IPi**	**P**
(1) N	1.00	1.00	0.98	0.658	0.496	0.5
(2) N^D	-	-	-	0.11	0.146	0.082
(3) D	-	-	0.02	0.232	0.325	0.247
(4) T	-	-	-	-	0.008	0.106
(5) T^H	-	-	-	-	0.024	0.065

Table C.5: Proportions of nasal category by position for Seoul females.

	Mm	Mi	PWi	APi	IPi	P
(1) N	1.00	1.00	0.96	0.13	0.14	0.26
(2) N^D	-	-	-	0.19	0.28	0.21
(3) D	-	-	0.03	0.59	0.46	0.38
(4) T	-	-	-	0.03	0.09	0.09
(5) T^H	-	-	0.01	0.06	0.03	0.06

Table C.6: Proportions of nasal category by position for Seoul males.

	Mm	Mi	PWi	APi	IPi	P
(1) N	1.00	1.00	0.98	0.30	0.28	0.47
(2) N^D	-	-	-	0.17	0.22	0.15
(3) D	-	-	0.02	0.52	0.47	0.29
(4) T	-	-	-	0.02	0.03	0.04
(5) T^H	-	-	-	-	-	0.05

Table C.7: Proportions of nasal category by position for Busan females.

	Mm	Mi	PWi	APi	IPi	P
(1) N	1.00	1.00	1.00	0.58	0.35	0.21
(2) N^D	-	-	-	0.22	0.17	0.04
(3) D	-	-	-	0.20	0.35	0.24
(4) T	-	-	-	-	0.11	0.31
(5) T^H	-	-	-	-	0.03	0.21

Table C.8: Proportions of nasal category by position for Busan males.

	Mm	Mi	PWi	APi	IPi	P
(1) N	1.00	1.00	1.00	0.81	0.70	0.43
(2) N^D	-	-	-	0.07	0.09	0.11
(3) D	-	-	-	0.11	0.21	0.26
(4) T	-	-	-	-	-	0.10
(5) T^H	-	-	-	-	-	0.10

Table C.9: Proportions of nasal category by position for Ulsan females.

	Mm	Mi	PWi	APi	IPi	P
(1) N	1.00	1.00	0.97	0.58	0.37	0.48
(2) N^D	-	-	-	0.10	0.18	0.05
(3) D	-	-	0.03	0.32	0.39	0.28
(4) T	-	-	-	-	0.02	0.09
(5) T^H	-	-	-	-	0.05	0.09

Table C.10: Proportions of nasal category by position for Ulsan males.

	Mm	Mi	PWi	APi	IPi	P
(1) N	1.00	1.00	0.99	0.76	0.61	0.52
(2) N^D	-	-	-	0.12	0.12	0.12
(3) D	-	-	0.01	0.12	0.27	0.21
(4) T	-	-	-	-	-	0.12
(5) T^H	-	-	-	-	-	0.04

Appendix D.

Non-Bayesian models for acoustic analyses

The summary tables in this appendix have been produced using the *stargazer* package (Hlavak, 2018) in R (R Core Team, 2015).

Table D.1: Regression results for normalised closure (ms).

	Normalised closure (ms)	
	Lenis	**Nasal**
Constant	29.73^{***} (5.32)	54.36^{***} (4.39)
positionMi	-4.25 (4.01)	-6.48 (3.61)
positionPWi	-0.37 (3.52)	-7.76^{*} (3.19)
positionAPi	17.06^{***} (3.57)	-0.94 (3.28)
positionIPi	44.24^{***} (3.70)	22.12^{***} (3.28)
regionBusan	0.21 (5.58)	-2.10 (5.23)
regionUlsan	-1.01 (5.50)	-3.74 (5.21)
gendermale	-6.56^{*} (3.28)	-8.84^{**} (3.26)
repetition2	-2.41 (1.29)	-0.76 (1.21)
positionMi:regionBusan	1.11 (5.78)	4.30 (5.17)
positionPWi:regionBusan	-4.58 (5.16)	-0.09 (4.51)
positionAPi:regionBusan	-5.43 (4.94)	4.39 (4.50)
positionIPi:regionBusan	-7.05 (5.17)	-3.91 (4.59)
positionMi:regionUlsan	0.61 (5.75)	1.48 (5.14)
positionPWi:regionUlsan	1.31 (4.96)	3.34 (4.44)
positionAPi:regionUlsan	0.37 (4.90)	5.52 (4.49)
positionIPi:regionUlsan	-9.85 (5.05)	-3.32 (4.59)
Observations	1,402	1,762
Log Likelihood	-6,320.97	-8,008.08
Akaike Inf. Crit.	12,681.95	16,056.17
Bayesian Inf. Crit.	12,786.86	16,165.65
Note:	$^{*}p<0.05$; $^{**}p<0.01$; $^{***}p<0.001$	

Table D.1’: Regression results for raw closure (ms). Compare the effect of *repetition* with Table D.1.

	Raw closure (ms)	
	Lenis	**Nasal**
Constant	35.89^{***} (5.71)	59.53^{***} (4.89)
positionMi	-4.70 (4.19)	-5.47 (3.74)
positionPWi	-0.82 (3.68)	-8.79^{**} (3.30)
positionAPi	15.56^{***} (3.74)	-1.55 (3.38)
positionIPi	44.45^{***} (3.87)	24.90^{***} (3.38)
regionBusan	2.59 (6.02)	2.66 (5.64)
regionUlsan	-0.39 (5.95)	-3.21 (5.62)
gendermale	-12.20^{***} (3.64)	-15.10^{***} (3.62)
repetition2	-4.19^{**} (1.35)	-3.31^{**} (1.25)
positionMi:regionBusan	2.45 (6.03)	1.10 (5.36)
positionPWi:regionBusan	-3.63 (5.40)	-1.34 (4.67)
positionAPi:regionBusan	-4.94 (5.17)	2.17 (4.66)
positionIPi:regionBusan	-4.48 (5.41)	-4.10 (4.75)
positionMi:regionUlsan	-0.24 (6.01)	0.74 (5.33)
positionPWi:regionUlsan	0.52 (5.19)	2.69 (4.60)
positionAPi:regionUlsan	-1.80 (5.13)	3.38 (4.65)
positionIPi:regionUlsan	-12.31^{*} (5.28)	-6.38 (4.75)
Observations	1,405	1,759
Log Likelihood	-6,397.70	-8,054.28
Akaike Inf. Crit.	12,835.40	16,148.57
Bayesian Inf. Crit.	12,940.35	16,258.02
Note:	$^{*}p<0.05$; $^{**}p<0.01$; $^{***}p<0.001$	

Table D.2: Regression results for normalised closure (ms) using both the lenis and nasal tokens.

	Normalised closure (ms)
	Combined
Constant	31.02^{***} (4.51)
positionMi	-4.36 (3.27)
positionPWi	-1.06 (2.95)
positionAPi	14.61^{***} (2.97)
positionIPi	43.23^{***} (3.05)
regionBusan	-1.07 (4.46)
regionUlsan	-2.51 (4.44)
consonantnasal	22.21^{***} (4.13)
gendermale	-7.92^{**} (2.98)
repetition2	-1.56 (0.89)
positionMi:regionBusan	3.01 (3.89)
positionPWi:regionBusan	-1.79 (3.42)
positionAPi:regionBusan	-0.05 (3.35)
positionIPi:regionBusan	-5.35 (3.46)
positionMi:regionUlsan	1.21 (3.87)
positionPWi:regionUlsan	2.51 (3.34)
positionAPi:regionUlsan	3.12 (3.34)
positionIPi:regionUlsan	-6.52 (3.42)
positionMi:consonantnasal	-2.03 (3.21)
positionPWi:consonantnasal	-5.87 (3.04)
positionAPi:consonantnasal	-13.18^{***} (2.99)
positionIPi:consonantnasal	-19.99^{***} (3.07)
Observations	3,164
Log Likelihood	-14,370.68
Akaike Inf. Crit.	28,791.35
Bayesian Inf. Crit.	28,942.84
Note:	$^{*}p<0.05$; $^{**}p<0.01$; $^{***}p<0.001$

Table D.3: Regression results for normalised residual voicing (ms).

	Normalised residual voicing (ms)	
	Lenis	**Nasal**
Constant	27.85*** (3.43)	59.41*** (6.18)
positionMi	-0.42 (3.09)	-3.53 (2.49)
positionPWi	0.32 (2.60)	-4.65* (2.21)
positionAPi	-6.57** (2.55)	-2.36 (2.23)
positionIPi	-12.71*** (2.59)	6.13** (2.20)
regionBusan	-2.56 (3.76)	-0.91 (3.66)
regionUlsan	-5.80 (3.80)	0.09 (3.67)
gendermale	7.60*** (2.01)	-2.06 (2.33)
repetition2	1.98* (0.87)	0.84 (0.83)
positionMi:regionBusan	2.52 (4.41)	3.17 (3.59)
positionPWi:regionBusan	-0.95 (3.88)	2.32 (3.15)
positionAPi:regionBusan	8.48* (3.46)	3.52 (3.06)
positionIPi:regionBusan	11.67*** (3.54)	4.23 (3.09)
positionMi:regionUlsan	1.48 (4.49)	2.91 (3.65)
positionPWi:regionUlsan	3.59 (3.69)	2.44 (3.11)
positionAPi:regionUlsan	2.82 (3.51)	2.52 (3.08)
positionIPi:regionUlsan	5.52 (3.56)	3.50 (3.12)
Observations	1,068	1,569
Log Likelihood	-4,244.17	-6,474.93
Akaike Inf. Crit.	8,528.34	12,989.86
Bayesian Inf. Crit.	8,627.81	13,097.02
Note:	*p<0.05; **p<0.01; ***p<0.001	

Table D.3': Regression results for raw residual voicing (ms).
Compare the effect of *repetition* with Table D.3.

	Raw residual voicing	
	Lenis	**Nasal**
Constant	27.76*** (3.52)	57.98*** (3.60)
positionMi	-0.69 (3.08)	-2.24 (2.66)
positionPWi	-0.25 (2.59)	-5.89* (2.35)
positionAPi	-6.92** (2.53)	-2.76 (2.38)
positionIPi	-11.93*** (2.58)	8.37*** (2.35)
regionBusan	-0.17 (3.81)	4.63 (3.96)
regionUlsan	-5.13 (3.84)	0.34 (3.97)
gendermale	3.28 (2.08)	-8.63*** (2.54)
repetition2	0.63 (0.86)	-1.82* (0.89)
positionMi:regionBusan	4.63 (4.38)	-0.24 (3.84)
positionPWi:regionBusan	-2.44 (3.86)	0.31 (3.36)
positionAPi:regionBusan	7.89* (3.44)	0.52 (3.26)
positionIPi:regionBusan	11.45** (3.52)	3.53 (3.30)
positionMi:regionUlsan	0.18 (4.47)	1.99 (3.89)
positionPWi:regionUlsan	2.58 (3.67)	1.67 (3.32)
positionAPi:regionUlsan	2.26 (3.50)	0.58 (3.29)
positionIPi:regionUlsan	4.94 (3.55)	1.78 (3.33)
Observations	1,071	1,566
Log Likelihood	-4,251.38	-6,551.33
Akaike Inf. Crit.	8,542.76	13,142.66
Bayesian Inf. Crit.	8,642.29	13,249.79
Note:	*p < 0.05; **p < 0.01; ***p < 0.001	

Table D.4: Regression results for normalised residual voicing (ms) using both the lenis and nasal tokens.

	Normalised residual voicing (ms)
	Combined
Constant	26.94^{***} (3.22)
positionMi	0.03 (2.65)
positionPWi	1.10 (2.32)
positionAPi	-4.84^{*} (2.26)
positionIPi	-9.77^{***} (2.29)
regionBusan	-1.54 (2.92)
regionUlsan	-1.72 (2.93)
consonantnasal	27.48^{***} (3.14)
gendermale	1.87 (1.83)
repetition2	1.11 (0.63)
positionMi:regionBusan	2.90 (2.91)
positionPWi:regionBusan	1.53 (2.55)
positionAPi:regionBusan	5.61^{*} (2.39)
positionIPi:regionBusan	7.90^{**} (2.43)
positionMi:regionUlsan	1.82 (2.95)
positionPWi:regionUlsan	2.46 (2.49)
positionAPi:regionUlsan	1.97 (2.41)
positionIPi:regionUlsan	3.78 (2.45)
positionMi:consonantnasal	-3.10 (2.54)
positionPWi:consonantnasal	-5.82^{*} (2.33)
positionAPi:consonantnasal	2.37 (2.22)
positionIPi:consonantnasal	14.39^{***} (2.25)
Observations	2,637
Log Likelihood	-10,843.03
Akaike Inf. Crit.	21,736.05
Bayesian Inf. Crit.	21,882.99
Note:	$^{*}p<0.05$; $^{**}p<0.01$; $^{***}p<0.001$

Table D.5: Regression results for % voicing during closure.

	% Voicing during closure	
	Lenis	**Nasal**
Constant	71.43*** (6.66)	99.02*** (1.34)
positionMi	6.88 (6.00)	-0.05 (1.48)
positionPWi	10.04* (5.05)	-0.58 (1.29)
positionAPi	-20.53*** (4.94)	-2.08 (1.31)
positionIPi	-45.82*** (5.03)	-8.80*** (1.29)
regionBusan	-0.75 (7.39)	0.01 (1.72)
regionUlsan	-8.87 (7.46)	-0.06 (1.73)
gendermale	16.17*** (4.01)	2.23** (0.86)
repetition2	4.29* (1.68)	0.33 (0.49)
positionMi:regionBusan	-3.19 (8.56)	0.15 (2.13)
positionPWi:regionBusan	-11.68 (7.53)	0.56 (1.87)
positionAPi:regionBusan	14.61* (6.71)	1.91 (1.81)
positionIPi:regionBusan	14.91* (6.86)	4.75** (1.83)
positionMi:regionUlsan	1.31 (8.71)	-0.04 (2.16)
positionPWi:regionUlsan	1.49 (7.15)	0.43 (1.85)
positionAPi:regionUlsan	0.56 (6.81)	2.01 (1.82)
positionIPi:regionUlsan	11.04 (6.91)	7.18*** (1.85)
Observations	1,068	1,569
Log Likelihood	-4,940.70	-5,632.78
Akaike Inf. Crit.	9,921.41	11,305.56
Bayesian Inf. Crit.	10,020.88	11,412.73
Note:	*p < 0.05; **p < 0.01; ***p < 0.001	

Table D.5': Regression results for % voicing with normalised closure for the lenis tokens. Compare the effect of *repetition* with Table D.5.

	% Voicing with normalised closure
	Lenis
Constant	72.36*** (6.98)
positionMi	4.87 (6.03)
positionPWi	5.73 (5.07)
positionAPi	-23.07*** (4.96)
positionIPi	-44.89*** (5.06)
regionBusan	3.39 (7.65)
regionUlsan	-9.05 (7.72)
gendermale	7.19 (4.30)
repetition	1.55 (1.69)
positionMi:regionBusan	-4.15 (8.58)
positionPWi:regionBusan	-13.58 (7.57)
positionAPi:regionBusan	14.07* (6.75)
positionIPi:regionBusan	13.12 (6.90)
positionMi:regionUlsan	-0.56 (8.77)
positionPWi:regionUlsan	0.75 (7.20)
positionAPi:regionUlsan	1.19 (6.85)
positionIPi:regionUlsan	10.86 (6.95)
Observations	1,071
Log Likelihood	-4,961.71
Akaike Inf. Crit.	9,963.41
Bayesian Inf. Crit.	10,062.94
Note:	$^{*}p<0.05$; $^{**}p<0.01$; $^{***}p<0.001$

Table D.6: Regression results for % voicing during closure using both the lenis and nasal tokens.

	% Voicing during closure
	Combined
Constant	73.79*** (3.70)
positionMi	6.98* (3.32)
positionPWi	9.05** (2.91)
positionAPi	-18.26*** (2.83)
positionIPi	-42.65*** (2.87)
regionBusan	-0.35 (3.21)
regionUlsan	-3.01 (3.23)
consonantnasal	22.43*** (3.80)
gendermale	7.96*** (1.80)
repetition2	1.79* (0.79)
positionMi:regionBusan	-1.47 (3.64)
positionPWi:regionBusan	-3.60 (3.20)
positionAPi:regionBusan	7.63* (3.00)
positionIPi:regionBusan	9.79** (3.05)
positionMi:regionUlsan	-0.70 (3.71)
positionPWi:regionUlsan	0.21 (3.12)
positionAPi:regionUlsan	0.40 (3.03)
positionIPi:regionUlsan	8.09** (3.07)
positionMi:consonantnasal	-6.74* (3.19)
positionPWi:consonantnasal	-8.70** (2.92)
positionAPi:consonantnasal	15.34*** (2.79)
positionIPi:consonantnasal	31.90*** (2.83)
Observations	2,637
Log Likelihood	-11,428.71
Akaike Inf. Crit.	22,907.42
Bayesian Inf. Crit.	23,054.36
Note:	*p < 0.05; **p < 0.01; ***p < 0.001

Table D.7: Regression results for normalised VOT (ms).

	Normalised VOT (ms)	
	Lenis	**Nasal**
Constant	2.96 (3.46)	0.26 (0.30)
positionMi	-1.22 (2.69)	0.001 (0.35)
positionPWi	2.85 (2.34)	0.04 (0.31)
positionAPi	38.24*** (2.37)	0.52 (0.32)
positionIPi	44.17*** (2.46)	0.51 (0.32)
positionP	32.46*** (2.28)	0.84** (0.31)
regionBusan	-0.80 (3.90)	-0.0000 (0.39)
regionUlsan	0.41 (3.85)	0.003 (0.39)
gendermale	-5.02* (2.34)	-0.44** (0.16)
repetition2	0.98 (0.76)	-0.03 (0.10)
positionMi:regionBusan	0.79 (3.88)	-0.01 (0.51)
positionPWi:regionBusan	0.22 (3.46)	-0.05 (0.44)
positionAPi:regionBusan	-17.68*** (3.32)	-0.56 (0.44)
positionIPi:regionBusan	-9.50** (3.46)	-0.17 (0.45)
positionP:regionBusan	1.63 (3.24)	1.67*** (0.43)
positionMi:regionUlsan	0.20 (3.86)	0.004 (0.50)
positionPWi:regionUlsan	2.28 (3.32)	-0.09 (0.43)
positionAPi:regionUlsan	-12.39*** (3.29)	-0.56 (0.44)
positionIPi:regionUlsan	-12.10*** (3.38)	-0.25 (0.45)
positionP:regionUlsan	-3.80 (3.19)	0.22 (0.43)
Observations	1,872	2,265
Log Likelihood	-7,700.78	-5,051.90
Akaike Inf. Crit.	15,447.56	10,149.80
Bayesian Inf. Crit.	15,574.86	10,281.48
Note:	$^{*}p<0.05$; $^{**}p<0.01$; $^{***}p<0.001$	

Table D.8: Regression results for normalised VOT (ms) using both the lenis and the nasal tokens.

	Normalised VOT (ms)
	Combined
Constant	2.00 (1.90)
positionMi	-1.06 (1.54)
positionPWi	3.32* (1.37)
positionAPi	33.35*** (1.38)
positionIPi	40.27*** (1.42)
positionP	31.93*** (1.33)
regionBusan	-0.38 (1.82)
regionUlsan	0.10 (1.80)
consonantnasal	0.06 (1.84)
gendermale	-2.50* (1.09)
repetition2	0.43 (0.37)
positionMi:regionBusan	0.42 (1.84)
positionPWi:regionBusan	-0.05 (1.61)
positionAPi:regionBusan	-8.79*** (1.58)
positionIPi:regionBusan	-4.05* (1.63)
positionP:regionBusan	1.80 (1.55)
positionMi:regionUlsan	0.20 (1.83)
positionPWi:regionUlsan	1.01 (1.58)
positionAPi:regionUlsan	-6.17*** (1.58)
positionIPi:regionUlsan	-5.56*** (1.62)
positionP:regionUlsan	-1.67 (1.53)
positionMi:consonantnasal	0.94 (1.52)
positionPWi:consonantnasal	-3.64** (1.40)
positionAPi:consonantnasal	-28.45*** (1.38)
positionIPi:consonantnasal	-36.84*** (1.42)
positionP:consonantnasal	-30.54*** (1.35)
Observations	4,137
Log Likelihood	-15,693.51
Akaike Inf. Crit.	31,445.01
Bayesian Inf. Crit.	31,628.52
Note:	*p<0.05; **p<0.01; ***p<0.001

Table D.9: Regression results for normalised total voiceless interval (ms).

	Normalised total voiceless interval (ms)	
	Lenis	**Nasal**
Constant	16.81** (6.10)	1.71 (1.91)
positionMi	-4.31 (5.00)	0.02 (2.07)
positionPWi	-0.89 (4.38)	0.31 (1.81)
positionAPi	50.82*** (4.44)	2.52 (1.85)
positionIPi	89.53*** (4.60)	13.04*** (1.85)
regionBusan	-0.42 (7.49)	0.06 (2.47)
regionUlsan	3.12 (7.41)	0.05 (2.46)
gendermale	-14.78** (4.68)	-3.18* (1.26)
repetition2	-2.13 (1.61)	-0.51 (0.69)
positionMi:regionBusan	3.30 (7.21)	-0.05 (2.96)
positionPWi:regionBusan	3.02 (6.44)	-0.21 (2.58)
positionAPi:regionBusan	-25.58*** (6.17)	-2.55 (2.58)
positionIPi:regionBusan	-27.04*** (6.45)	-9.22*** (2.63)
positionMi:regionUlsan	1.01 (7.17)	0.05 (2.95)
positionPWi:regionUlsan	2.62 (6.18)	-0.24 (2.55)
positionAPi:regionUlsan	-11.46 (6.12)	-2.61 (2.57)
positionIPi:regionUlsan	-26.40*** (6.30)	-11.25*** (2.63)
Observations	1,402	1,762
Log Likelihood	-6,625.79	-7,020.19
Akaike Inf. Crit.	13,291.57	14,080.38
Bayesian Inf. Crit.	13,396.48	14,189.87
Note:		$^{*}p<0.05$; $^{**}p<0.01$; $^{***}p<0.001$

Table D.10: Regression results for normalised total voiceless interval (ms) using both the lenis and the nasal tokens.

	Normalised total voiceless interval (ms)
	Combined
Constant	13.38*** (3.56)
positionMi	-3.58 (3.13)
positionPWi	-0.05 (2.80)
positionAPi	45.82*** (2.82)
positionIPi	83.46*** (2.90)
regionBusan	-0.14 (3.90)
regionUlsan	1.10 (3.87)
consonantnasal	-8.97** (2.88)
gendermale	-8.25*** (2.48)
repetition2	-1.32 (0.85)
positionMi:regionBusan	1.52 (3.72)
positionPWi:regionBusan	0.98 (3.28)
positionAPi:regionBusan	-13.66*** (3.21)
positionIPi:regionBusan	-16.92*** (3.31)
positionMi:regionUlsan	0.74 (3.70)
positionPWi:regionUlsan	1.30 (3.20)
positionAPi:regionUlsan	-6.72* (3.19)
positionIPi:regionUlsan	-17.79*** (3.28)
positionMi:consonantnasal	3.01 (3.07)
positionPWi:consonantnasal	-0.19 (2.87)
positionAPi:consonantnasal	-38.64*** (2.83)
positionIPi:consonantnasal	-65.78*** (2.90)
Observations	3,164
Log Likelihood	-14,220.79
Akaike Inf. Crit.	28,491.59
Bayesian Inf. Crit.	28,643.08
Note:	$^{*}p<0.05$; $^{**}p<0.01$; $^{***}p<0.001$

Table D.11: Regression results for normalised prevoicing (ms).

	Normalised prevoicing (ms)
	Nasal
Constant	28.28*** (3.81)
regionBusan	-2.04 (4.64)
regionUlsan	-0.89 (4.45)
categorynd	4.61 (3.34)
categoryd	-10.92*** (2.95)
categoryt	-28.78*** (4.86)
categoryth	-26.16*** (5.10)
gendermale	1.08 (3.28)
repetition2	-0.25 (1.42)
regionBusan: categorynd	10.21 (5.63)
regionUlsan: categorynd	-0.69 (5.34)
regionBusan: categoryd	1.35 (4.27)
regionUlsan: categoryd	5.87 (4.03)
regionBusan: categoryt	4.78 (5.96)
regionUlsan: categoryt	0.82 (6.15)
regionBusan: categoryth	1.51 (6.31)
regionUlsan: categoryth	1.16 (6.90)
Observations	503
Log Likelihood	-2,020.51
Akaike Inf. Crit.	4,081.02
Bayesian Inf. Crit.	4,165.43
Note:	*p < 0.05; **p < 0.01; ***p < 0.001

Table D.12: Regression results for normalised intensity (dB).

	Normalised intensity (dB)
	Nasal
Constant	5.32^{***} (0.90)
positionMi	0.77 (0.75)
positionPWi	0.03 (0.66)
positionAPi	6.71^{***} (0.68)
positionIPi	13.64^{***} (0.67)
regionBusan	-0.39 (1.03)
regionUlsan	0.64 (1.03)
gendermale	-1.87^{**} (0.63)
repetition2	0.47 (0.25)
positionMi:regionBusan	-0.20 (1.07)
positionPWi:regionBusan	0.44 (0.94)
positionAPi:regionBusan	-2.49^{**} (0.92)
positionIPi:regionBusan	-2.94^{**} (0.94)
positionMi:regionUlsan	-0.68 (1.09)
positionPWi:regionUlsan	-1.05 (0.93)
positionAPi:regionUlsan	-3.79^{***} (0.93)
positionIPi:regionUlsan	-4.17^{***} (0.94)
Observations	1,534
Log Likelihood	-4,484.43
Akaike Inf. Crit.	9,008.87
Bayesian Inf. Crit.	9,115.58
Note:	$^{*}p<0.05$; $^{**}p<0.01$; $^{***}p<0.001$

Appendix E.

Tukey HSD tests for Non-Bayesian models

Table E.1: Pairwise comparisons of prosodic position by region (outcome: normalised closure duration; data: lenis, excluding post-pausal tokens).

```
region = Seoul:
 contrast  estimate   SE   df t.ratio p.value
 Mm - Mi      4.255 4.01 1351   1.061 0.8262
 Mm - PWi     0.367 3.53 1358   0.104 1.0000
 Mm - APi   -17.057 3.58 1359  -4.770 <.0001
 Mm - IPi   -44.244 3.70 1363 -11.960 <.0001
 Mi - PWi    -3.887 3.53 1358  -1.102 0.8054
 Mi - APi   -21.312 3.58 1359  -5.960 <.0001
 Mi - IPi   -48.499 3.70 1363 -13.110 <.0001
 PWi - APi  -17.425 2.77 1359  -6.293 <.0001
 PWi - IPi  -44.611 2.94 1360 -15.199 <.0001
 APi - IPi  -27.186 3.00 1365  -9.076 <.0001

region = Busan:
 contrast  estimate   SE   df t.ratio p.value
 Mm - Mi      3.144 4.16 1352   0.755 0.9431
 Mm - PWi     4.946 3.99 1365   1.239 0.7288
 Mm - APi   -11.627 3.66 1359  -3.180 0.0130
 Mm - IPi   -37.199 3.84 1361  -9.695 <.0001
 Mi - PWi     1.802 3.87 1364   0.466 0.9903
 Mi - APi   -14.771 3.53 1359  -4.190 0.0003
 Mi - IPi   -40.343 3.70 1360 -10.891 <.0001
 PWi - APi  -16.573 3.05 1366  -5.439 <.0001
 PWi - IPi  -42.145 3.20 1361 -13.160 <.0001
 APi - IPi  -25.572 2.81 1357  -9.088 <.0001
```

```
region = Ulsan:
 contrast  estimate   SE   df t.ratio p.value
 Mm - Mi      3.644 4.12 1352   0.885 0.9024
 Mm - PWi    -0.939 3.71 1361  -0.253 0.9991
 Mm - APi   -17.430 3.61 1360  -4.834 <.0001
 Mm - IPi   -34.399 3.67 1360  -9.379 <.0001
 Mi - PWi    -4.583 3.75 1361  -1.223 0.7380
 Mi - APi   -21.074 3.65 1361  -5.777 <.0001
 Mi - IPi   -38.043 3.70 1359 -10.272 <.0001
 PWi - APi  -16.491 2.90 1364  -5.690 <.0001
 PWi - IPi  -33.460 2.95 1359 -11.332 <.0001
 APi - IPi  -16.969 2.81 1357  -6.045 <.0001
```

Table E.2: Interaction contrasts (outcome: normalised closure duration; data: lenis, excluding post-pausal tokens).

position_pairwise	region_pairwise	estimate	SE	df	t.ratio	p.value
Mm - Mi	Seoul - Busan	1.1105	5.78	1352	0.192	0.8476
Mm - PWi	Seoul - Busan	-4.5789	5.16	1356	-0.887	0.3753
Mm - APi	Seoul - Busan	-5.4308	4.95	1353	-1.098	0.2723
Mm - IPi	Seoul - Busan	-7.0453	5.17	1356	-1.363	0.1730
Mi - PWi	Seoul - Busan	-5.6895	5.07	1356	-1.123	0.2616
Mi - APi	Seoul - Busan	-6.5413	4.85	1353	-1.348	0.1779
Mi - IPi	Seoul - Busan	-8.1559	5.07	1356	-1.608	0.1080
PWi - APi	Seoul - Busan	-0.8519	4.12	1364	-0.207	0.8362
PWi - IPi	Seoul - Busan	-2.4664	4.34	1360	-0.568	0.5700
APi - IPi	Seoul - Busan	-1.6146	4.12	1362	-0.392	0.6949
Mm - Mi	Seoul - Ulsan	0.6107	5.75	1352	0.106	0.9154
Mm - PWi	Seoul - Ulsan	1.3064	4.96	1354	0.264	0.7922
Mm - APi	Seoul - Ulsan	0.3724	4.90	1354	0.076	0.9395
Mm - IPi	Seoul - Ulsan	-9.8453	5.05	1356	-1.950	0.0514
Mi - PWi	Seoul - Ulsan	0.6957	4.99	1354	0.139	0.8891
Mi - APi	Seoul - Ulsan	-0.2383	4.94	1354	-0.048	0.9615
Mi - IPi	Seoul - Ulsan	-10.4560	5.08	1355	-2.060	0.0396
PWi - APi	Seoul - Ulsan	-0.9340	4.00	1361	-0.233	0.8156
PWi - IPi	Seoul - Ulsan	-11.1517	4.16	1359	-2.682	0.0074
APi - IPi	Seoul - Ulsan	-10.2177	4.11	1362	-2.489	0.0129
Mm - Mi	Busan - Ulsan	-0.4998	5.85	1352	-0.085	0.9319
Mm - PWi	Busan - Ulsan	5.8853	5.27	1357	1.118	0.2639
Mm - APi	Busan - Ulsan	5.8032	4.96	1354	1.171	0.2419
Mm - IPi	Busan - Ulsan	-2.8000	5.10	1353	-0.549	0.5833
Mi - PWi	Busan - Ulsan	6.3851	5.20	1356	1.227	0.2200
Mi - APi	Busan - Ulsan	6.3030	4.90	1354	1.287	0.1982
Mi - IPi	Busan - Ulsan	-2.3001	5.04	1353	-0.457	0.6480
PWi - APi	Busan - Ulsan	-0.0821	4.20	1366	-0.020	0.9844
PWi - IPi	Busan - Ulsan	-8.6853	4.35	1359	-1.998	0.0459
APi - IPi	Busan - Ulsan	-8.6031	3.97	1357	-2.166	0.0305

Table E.3: Pairwise comparisons of position by region (outcome: normalised closure duration; data: nasal, excluding post-pausal tokens).

region = Seoul:

contrast	estimate	SE	df	t.ratio	p.value
Mm - Mi	6.478	3.61	1711	1.794	0.3775
Mm - PWi	7.761	3.20	1718	2.426	0.1090
Mm - APi	0.938	3.29	1717	0.285	0.9986
Mm - IPi	-22.116	3.28	1721	-6.738	<.0001
Mi - PWi	1.283	3.19	1717	0.403	0.9945
Mi - APi	-5.541	3.27	1717	-1.693	0.4384
Mi - IPi	-28.594	3.27	1721	-8.744	<.0001
PWi - APi	-6.823	2.63	1714	-2.599	0.0710
PWi - IPi	-29.877	2.63	1715	-11.353	<.0001
APi - IPi	-23.053	2.74	1719	-8.425	<.0001

region = Busan:

contrast	estimate	SE	df	t.ratio	p.value
Mm - Mi	2.178	3.70	1713	0.588	0.9768
Mm - PWi	7.847	3.34	1719	2.352	0.1294
Mm - APi	-3.450	3.24	1716	-1.065	0.8242
Mm - IPi	-18.206	3.35	1724	-5.442	<.0001
Mi - PWi	5.669	3.40	1721	1.669	0.4537
Mi - APi	-5.628	3.30	1718	-1.705	0.4313
Mi - IPi	-20.384	3.41	1727	-5.973	<.0001
PWi - APi	-11.296	2.70	1719	-4.178	0.0003
PWi - IPi	-26.053	2.82	1719	-9.238	<.0001
APi - IPi	-14.757	2.73	1724	-5.413	<.0001

region = Ulsan:

contrast	estimate	SE	df	t.ratio	p.value
Mm - Mi	5.002	3.66	1711	1.366	0.6497
Mm - PWi	4.421	3.24	1718	1.365	0.6502
Mm - APi	-4.581	3.23	1717	-1.420	0.6147
Mm - IPi	-18.796	3.36	1719	-5.599	<.0001
Mi - PWi	-0.580	3.31	1719	-0.175	0.9998
Mi - APi	-9.583	3.30	1718	-2.908	0.0302
Mi - IPi	-23.798	3.42	1720	-6.955	<.0001
PWi - APi	-9.002	2.64	1719	-3.410	0.0060
PWi - IPi	-23.218	2.79	1717	-8.309	<.0001
APi - IPi	-14.215	2.77	1715	-5.138	<.0001

Table E.4: Interaction contrasts (outcome: normalised closure duration; data: nasal, excluding post-pausal tokens).

position_pairwise	region_pairwise	estimate	SE	df	t.ratio	p.value
Mm - Mi	Seoul - Busan	4.3005	5.17	1712	0.832	0.4057
Mm - PWi	Seoul - Busan	-0.0856	4.51	1713	-0.019	0.9849
Mm - APi	Seoul - Busan	4.3873	4.50	1713	0.976	0.3294
Mm - IPi	Seoul - Busan	-3.9096	4.59	1716	-0.852	0.3942
Mi - PWi	Seoul - Busan	-4.3861	4.55	1714	-0.965	0.3349
Mi - APi	Seoul - Busan	0.0869	4.53	1714	0.019	0.9847
Mi - IPi	Seoul - Busan	-8.2101	4.63	1718	-1.774	0.0763
PWi - APi	Seoul - Busan	4.4730	3.77	1717	1.187	0.2353
PWi - IPi	Seoul - Busan	-3.8240	3.85	1716	-0.992	0.3213
APi - IPi	Seoul - Busan	-8.2969	3.86	1722	-2.149	0.0318
Mm - Mi	Seoul - Ulsan	1.4763	5.14	1711	0.287	0.7741
Mm - PWi	Seoul - Ulsan	3.3394	4.44	1712	0.751	0.4526
Mm - APi	Seoul - Ulsan	5.5183	4.49	1713	1.230	0.2189
Mm - IPi	Seoul - Ulsan	-3.3197	4.59	1714	-0.723	0.4697
Mi - PWi	Seoul - Ulsan	1.8631	4.49	1712	0.415	0.6780
Mi - APi	Seoul - Ulsan	4.0420	4.53	1713	0.893	0.3722
Mi - IPi	Seoul - Ulsan	-4.7959	4.63	1714	-1.036	0.3004
PWi - APi	Seoul - Ulsan	2.1789	3.72	1717	0.585	0.5583
PWi - IPi	Seoul - Ulsan	-6.6591	3.84	1716	-1.735	0.0830
APi - IPi	Seoul - Ulsan	-8.8380	3.89	1717	-2.273	0.0232
Mm - Mi	Busan - Ulsan	-2.8242	5.21	1712	-0.542	0.5877
Mm - PWi	Busan - Ulsan	3.4250	4.54	1713	0.755	0.4505
Mm - APi	Busan - Ulsan	1.1310	4.46	1713	0.254	0.7996
Mm - IPi	Busan - Ulsan	0.5899	4.63	1715	0.127	0.8987
Mi - PWi	Busan - Ulsan	6.2492	4.63	1714	1.349	0.1776
Mi - APi	Busan - Ulsan	3.9552	4.55	1715	0.869	0.3851
Mi - IPi	Busan - Ulsan	3.4141	4.73	1718	0.722	0.4706
PWi - APi	Busan - Ulsan	-2.2940	3.78	1719	-0.607	0.5440
PWi - IPi	Busan - Ulsan	-2.8351	3.97	1717	-0.715	0.4749
APi - IPi	Busan - Ulsan	-0.5411	3.88	1720	-0.139	0.8892

Table E.5: Pairwise comparisons of consonant by position. (outcome: normalised closure duration; data: combined, excluding post-pausal tokens).

```
position = Mm:
 contrast      estimate   SE   df t.ratio p.value
 lenis - nasal   -25.25 4.15 41.8 -6.083  <.0001

position = Mi:
 contrast      estimate   SE   df t.ratio p.value
 lenis - nasal   -24.31 4.14 42.0 -5.877  <.0001

position = PWi:
 contrast      estimate   SE   df t.ratio p.value
 lenis - nasal   -19.18 3.69 26.1 -5.202  <.0001

position = APi:
 contrast      estimate   SE   df t.ratio p.value
 lenis - nasal   -12.07 3.63 24.9 -3.322  0.0028

position = IPi:
 contrast      estimate   SE   df t.ratio p.value
 lenis - nasal    -5.06 3.72 26.6 -1.361  0.1849
```

Table E.6: Pairwise comparisons of position by region (outcome: normalised residual voicing; data: lenis, excluding post-pausal tokens and lenited tokens).

```
region = Seoul:
 contrast  estimate   SE   df t.ratio p.value
 Mm - Mi      0.419 3.09 1022  0.135  0.9999
 Mm - PWi    -0.323 2.60 1026 -0.124  0.9999
 Mm - APi     6.572 2.55 1026  2.578  0.0751
 Mm - IPi    12.709 2.60 1030  4.895  <.0001
 Mi - PWi    -0.742 2.66 1025 -0.279  0.9987
 Mi - APi     6.154 2.62 1027  2.345  0.1317
 Mi - IPi    12.291 2.66 1030  4.614  <.0001
 PWi - APi    6.896 1.83 1024  3.777  0.0016
 PWi - IPi   13.033 1.89 1026  6.880  <.0001
 APi - IPi    6.137 1.83 1034  3.362  0.0072

region = Busan:
 contrast  estimate   SE   df t.ratio p.value
 Mm - Mi     -2.098 3.14 1022 -0.668  0.9633
 Mm - PWi     0.631 3.01 1033  0.209  0.9996
 Mm - APi    -1.904 2.51 1026 -0.759  0.9420
 Mm - IPi     1.042 2.58 1029  0.404  0.9944
 Mi - PWi     2.730 3.18 1035  0.860  0.9114
 Mi - APi     0.194 2.67 1026  0.073  1.0000
 Mi - IPi     3.141 2.75 1033  1.142  0.7841
 PWi - APi   -2.535 2.35 1030 -1.079  0.8175
 PWi - IPi    0.411 2.39 1030  0.172  0.9998
 APi - IPi    2.946 1.74 1029  1.692  0.4392

region = Ulsan:
 contrast  estimate   SE   df t.ratio p.value
 Mm - Mi     -1.060 3.25 1022 -0.326  0.9975
 Mm - PWi    -3.916 2.76 1027 -1.420  0.6151
 Mm - APi     3.754 2.60 1028  1.446  0.5978
 Mm - IPi     7.190 2.62 1029  2.748  0.0479
 Mi - PWi    -2.856 2.93 1028 -0.976  0.8662
 Mi - APi     4.815 2.78 1030  1.734  0.4133
 Mi - IPi     8.251 2.78 1029  2.967  0.0255
 PWi - APi    7.671 1.98 1029  3.879  0.0011
 PWi - IPi   11.107 1.98 1027  5.616  <.0001
 APi - IPi    3.436 1.73 1025  1.991  0.2709
```

Table E.6’: Pairwise comparisons of position by region (outcome: raw residual voicing; data: lenis, excluding post-pausal tokens and lenited tokens). Compare with Table E.6.

```
region = Seoul:
 contrast  estimate   SE   df t.ratio p.value
 Mm - Mi      0.688 3.08 1025  0.224  0.9994
 Mm - PWi     0.249 2.59 1030  0.096  1.0000
 Mm - APi     6.920 2.54 1030  2.728  0.0507
 Mm - IPi    11.930 2.58 1034  4.616  <.0001
 Mi - PWi    -0.439 2.65 1029 -0.166  0.9998
 Mi - APi     6.232 2.61 1031  2.386  0.1198
 Mi - IPi    11.242 2.65 1033  4.240  0.0002
 PWi - APi    6.671 1.82 1027  3.671  0.0024
 PWi - IPi   11.681 1.89 1029  6.194  <.0001
 APi - IPi    5.010 1.82 1036  2.756  0.0469

region = Busan:
 contrast  estimate   SE   df t.ratio p.value
 Mm - Mi     -3.941 3.10 1025 -1.269  0.7102
 Mm - PWi     2.687 3.00 1036  0.896  0.8985
 Mm - APi    -0.970 2.49 1030 -0.390  0.9951
 Mm - IPi     0.483 2.56 1032  0.188  0.9997
 Mi - PWi     6.628 3.14 1040  2.109  0.2168
 Mi - APi     2.971 2.63 1031  1.130  0.7907
 Mi - IPi     4.423 2.71 1038  1.630  0.4787
 PWi - APi   -3.658 2.33 1033 -1.566  0.5194
 PWi - IPi   -2.205 2.38 1033 -0.928  0.8861
 APi - IPi    1.453 1.72 1031  0.842  0.9174

region = Ulsan:
 contrast  estimate   SE   df t.ratio p.value
 Mm - Mi      0.507 3.23 1025  0.157  0.9999
 Mm - PWi    -2.335 2.75 1031 -0.850  0.9146
 Mm - APi     4.656 2.58 1031  1.802  0.3729
 Mm - IPi     6.994 2.60 1032  2.686  0.0567
 Mi - PWi    -2.842 2.91 1032 -0.976  0.8663
 Mi - APi     4.149 2.76 1034  1.501  0.5619
 Mi - IPi     6.487 2.77 1033  2.344  0.1321
 PWi - APi    6.991 1.97 1031  3.551  0.0037
 PWi - IPi    9.330 1.97 1030  4.738  <.0001
 APi - IPi    2.339 1.72 1027  1.361  0.6526
```

Table E.7: Interaction contrasts (outcome: normalised residual voicing; data: lenis, excluding post-pausal tokens and lenited tokens).

position_pairwise	region_pairwise	estimate	SE	df	t.ratio	p.value
Mm - Mi	Seoul - Busan	2.517	4.41	1021	0.570	0.5686
Mm - PWi	Seoul - Busan	-0.955	3.88	1025	-0.246	0.8058
Mm - APi	Seoul - Busan	8.476	3.46	1021	2.449	0.0145
Mm - IPi	Seoul - Busan	11.667	3.54	1024	3.297	0.0010
Mi - PWi	Seoul - Busan	-3.472	4.04	1026	-0.858	0.3909
Mi - APi	Seoul - Busan	5.959	3.64	1021	1.639	0.1016
Mi - IPi	Seoul - Busan	9.150	3.71	1026	2.466	0.0138
PWi - APi	Seoul - Busan	9.431	2.97	1028	3.171	0.0016
PWi - IPi	Seoul - Busan	12.622	3.04	1028	4.153	<.0001
APi - IPi	Seoul - Busan	3.191	2.53	1032	1.262	0.2071
Mm - Mi	Seoul - Ulsan	1.479	4.49	1023	0.329	0.7422
Mm - PWi	Seoul - Ulsan	3.593	3.69	1023	0.974	0.3302
Mm - APi	Seoul - Ulsan	2.818	3.51	1021	0.802	0.4225
Mm - IPi	Seoul - Ulsan	5.519	3.56	1024	1.549	0.1218
Mi - PWi	Seoul - Ulsan	2.114	3.86	1022	0.548	0.5838
Mi - APi	Seoul - Ulsan	1.339	3.71	1024	0.361	0.7182
Mi - IPi	Seoul - Ulsan	4.040	3.74	1024	1.081	0.2801
PWi - APi	Seoul - Ulsan	-0.775	2.68	1026	-0.289	0.7727
PWi - IPi	Seoul - Ulsan	1.926	2.73	1026	0.706	0.4806
APi - IPi	Seoul - Ulsan	2.701	2.51	1030	1.075	0.2827
Mm - Mi	Busan - Ulsan	-1.038	4.51	1022	-0.230	0.8182
Mm - PWi	Busan - Ulsan	4.548	3.97	1025	1.146	0.2519
Mm - APi	Busan - Ulsan	-5.658	3.48	1021	-1.626	0.1043
Mm - IPi	Busan - Ulsan	-6.148	3.51	1022	-1.754	0.0797
Mi - PWi	Busan - Ulsan	5.586	4.22	1028	1.325	0.1856
Mi - APi	Busan - Ulsan	-4.620	3.75	1023	-1.231	0.2188
Mi - IPi	Busan - Ulsan	-5.110	3.78	1025	-1.351	0.1771
PWi - APi	Busan - Ulsan	-10.206	3.07	1029	-3.329	0.0009
PWi - IPi	Busan - Ulsan	-10.696	3.09	1028	-3.466	0.0006
APi - IPi	Busan - Ulsan	-0.490	2.45	1027	-0.200	0.8414

Table E.7’: Interaction contrasts (outcome: raw residual voicing; data: lenis, excluding post-pausal tokens and lenited tokens). Compare with Table E.7.

position_pairwise	region_pairwise	estimate	SE	df	t.ratio	p.value
Mm - Mi	Seoul - Busan	4.629	4.38	1025	1.058	0.2904
Mm - PWi	Seoul - Busan	-2.438	3.86	1028	-0.631	0.5283
Mm - APi	Seoul - Busan	7.891	3.44	1024	2.292	0.0221
Mm - IPi	Seoul - Busan	11.448	3.52	1027	3.249	0.0012
Mi - PWi	Seoul - Busan	-7.067	4.01	1030	-1.762	0.0784
Mi - APi	Seoul - Busan	3.262	3.59	1025	0.907	0.3644
Mi - IPi	Seoul - Busan	6.819	3.68	1029	1.855	0.0639
PWi - APi	Seoul - Busan	10.329	2.96	1031	3.493	0.0005
PWi - IPi	Seoul - Busan	13.886	3.03	1030	4.589	<.0001
APi - IPi	Seoul - Busan	3.557	2.51	1034	1.416	0.1570
Mm - Mi	Seoul - Ulsan	0.181	4.47	1026	0.041	0.9677
Mm - PWi	Seoul - Ulsan	2.585	3.67	1026	0.704	0.4817
Mm - APi	Seoul - Ulsan	2.265	3.50	1024	0.648	0.5173
Mm - IPi	Seoul - Ulsan	4.936	3.55	1027	1.391	0.1645
Mi - PWi	Seoul - Ulsan	2.403	3.84	1025	0.626	0.5316
Mi - APi	Seoul - Ulsan	2.084	3.69	1027	0.564	0.5727
Mi - IPi	Seoul - Ulsan	4.755	3.72	1027	1.277	0.2018
PWi - APi	Seoul - Ulsan	-0.320	2.67	1028	-0.120	0.9047
PWi - IPi	Seoul - Ulsan	2.351	2.72	1028	0.865	0.3871
APi - IPi	Seoul - Ulsan	2.671	2.50	1032	1.067	0.2860
Mm - Mi	Busan - Ulsan	-4.448	4.48	1025	-0.993	0.3209
Mm - PWi	Busan - Ulsan	5.023	3.95	1028	1.272	0.2038
Mm - APi	Busan - Ulsan	-5.626	3.46	1024	-1.626	0.1043
Mm - IPi	Busan - Ulsan	-6.512	3.49	1024	-1.866	0.0623
Mi - PWi	Busan - Ulsan	9.471	4.19	1031	2.263	0.0239
Mi - APi	Busan - Ulsan	-1.178	3.72	1027	-0.317	0.7513
Mi - IPi	Busan - Ulsan	-2.064	3.75	1029	-0.550	0.5821
PWi - APi	Busan - Ulsan	-10.649	3.05	1032	-3.493	0.0005
PWi - IPi	Busan - Ulsan	-11.535	3.07	1030	-3.753	0.0002
APi - IPi	Busan - Ulsan	-0.886	2.43	1029	-0.365	0.7155

Table E.8: Pairwise comparisons of position by region (outcome: normalised residual voicing; data: nasal, excluding post-pausal tokens and lenited tokens).

```
region = Seoul:
 contrast  estimate   SE   df t.ratio p.value
 Mm - Mi      3.530 2.49 1517  1.415  0.6181
 Mm - PWi     4.646 2.21 1519  2.104  0.2186
 Mm - APi     2.362 2.23 1519  1.058  0.8281
 Mm - IPi    -6.134 2.20 1520 -2.783  0.0433
 Mi - PWi     1.116 2.26 1519  0.494  0.9879
 Mi - APi    -1.169 2.28 1520 -0.511  0.9863
 Mi - IPi    -9.664 2.26 1521 -4.285  0.0002
 PWi - APi   -2.284 1.82 1521 -1.256  0.7185
 PWi - IPi  -10.780 1.80 1522 -5.996  <.0001
 APi - IPi   -8.496 1.83 1524 -4.633  <.0001

region = Busan:
 contrast  estimate   SE   df t.ratio p.value
 Mm - Mi      0.357 2.58 1524  0.138  0.9999
 Mm - PWi     2.322 2.37 1521  0.981  0.8638
 Mm - APi    -1.162 2.21 1520 -0.526  0.9847
 Mm - IPi   -10.361 2.27 1523 -4.574  0.0001
 Mi - PWi     1.965 2.45 1526  0.803  0.9296
 Mi - APi    -1.519 2.31 1529 -0.659  0.9650
 Mi - IPi   -10.718 2.36 1532 -4.538  0.0001
 PWi - APi   -3.485 1.92 1524 -1.814  0.3655
 PWi - IPi  -12.684 1.99 1526 -6.381  <.0001
 APi - IPi   -9.199 1.81 1531 -5.069  <.0001

region = Ulsan:
 contrast  estimate   SE   df t.ratio p.value
 Mm - Mi      0.616 2.66 1517  0.232  0.9994
 Mm - PWi     2.203 2.31 1518  0.956  0.8748
 Mm - APi    -0.154 2.24 1520 -0.069  1.0000
 Mm - IPi    -9.638 2.31 1521 -4.168  0.0003
 Mi - PWi     1.587 2.44 1519  0.650  0.9666
 Mi - APi    -0.771 2.39 1523 -0.323  0.9977
 Mi - IPi   -10.254 2.45 1523 -4.183  0.0003
 PWi - APi   -2.358 1.85 1523 -1.272  0.7087
 PWi - IPi  -11.841 1.93 1522 -6.127  <.0001
 APi - IPi   -9.484 1.84 1521 -5.163  <.0001
```

Table E.8’: Pairwise comparisons of position by region (outcome: raw residual voicing; data: nasal, excluding post-pausal tokens and lenited tokens). Compare with Table E.8.

```
region = Seoul:
 contrast  estimate   SE   df t.ratio p.value
 Mm - Mi      2.235 2.66 1516  0.839  0.9184
 Mm - PWi     5.891 2.35 1525  2.503  0.0904
 Mm - APi     2.764 2.38 1525  1.161  0.7735
 Mm - IPi    -8.365 2.35 1526 -3.560  0.0035
 Mi - PWi     3.656 2.41 1525  1.519  0.5500
 Mi - APi     0.529 2.44 1526  0.217  0.9995
 Mi - IPi   -10.601 2.40 1527 -4.408  0.0001
 PWi - APi   -3.127 1.94 1519 -1.610  0.4913
 PWi - IPi  -14.257 1.92 1521 -7.427  <.0001
 APi - IPi  -11.130 1.96 1523 -5.685  <.0001

region = Busan:
 contrast  estimate   SE   df t.ratio p.value
 Mm - Mi      2.473 2.76 1517  0.896  0.8986
 Mm - PWi     5.578 2.52 1526  2.211  0.1761
 Mm - APi     2.240 2.36 1524  0.950  0.8770
 Mm - IPi   -11.891 2.42 1529 -4.923  <.0001
 Mi - PWi     3.106 2.61 1526  1.189  0.7578
 Mi - APi    -0.232 2.47 1528 -0.094  1.0000
 Mi - IPi   -14.364 2.52 1532 -5.694  <.0001
 PWi - APi   -3.338 2.05 1521 -1.627  0.4801
 PWi - IPi  -17.470 2.12 1526 -8.233  <.0001
 APi - IPi  -14.132 1.94 1526 -7.288  <.0001

region = Ulsan:
 contrast  estimate   SE   df t.ratio p.value
 Mm - Mi      0.240 2.84 1516  0.085  1.0000
 Mm - PWi     4.220 2.46 1524  1.716  0.4240
 Mm - APi     2.181 2.39 1526  0.912  0.8922
 Mm - IPi   -10.145 2.47 1527 -4.115  0.0004
 Mi - PWi     3.980 2.60 1525  1.529  0.5435
 Mi - APi     1.941 2.55 1528  0.762  0.9413
 Mi - IPi   -10.386 2.61 1528 -3.973  0.0007
 PWi - APi   -2.039 1.98 1523 -1.030  0.8415
 PWi - IPi  -14.365 2.06 1522 -6.963  <.0001
 APi - IPi  -12.327 1.96 1520 -6.285  <.0001
```

Table E.9: Interaction contrasts (outcome: normalised residual voicing; data: nasal, excluding post-pausal tokens and lenited tokens).

position_pairwise	region_pairwise	estimate	SE	df	t.ratio	p.value
Mm - Mi	Seoul - Busan	3.1732	3.59	1521	0.884	0.3769
Mm - PWi	Seoul - Busan	2.3237	3.15	1519	0.738	0.4609
Mm - APi	Seoul - Busan	3.5240	3.06	1519	1.153	0.2491
Mm - IPi	Seoul - Busan	4.2274	3.09	1521	1.369	0.1711
Mi - PWi	Seoul - Busan	-0.8495	3.24	1522	-0.262	0.7935
Mi - APi	Seoul - Busan	0.3508	3.16	1525	0.111	0.9117
Mi - IPi	Seoul - Busan	1.0542	3.20	1527	0.330	0.7415
PWi - APi	Seoul - Busan	1.2003	2.64	1523	0.454	0.6500
PWi - IPi	Seoul - Busan	1.9037	2.68	1524	0.711	0.4774
APi - IPi	Seoul - Busan	0.7034	2.58	1528	0.273	0.7851
Mm - Mi	Seoul - Ulsan	2.9140	3.65	1517	0.799	0.4243
Mm - PWi	Seoul - Ulsan	2.4429	3.11	1518	0.785	0.4329
Mm - APi	Seoul - Ulsan	2.5161	3.08	1519	0.818	0.4137
Mm - IPi	Seoul - Ulsan	3.5041	3.12	1520	1.125	0.2609
Mi - PWi	Seoul - Ulsan	-0.4711	3.25	1518	-0.145	0.8847
Mi - APi	Seoul - Ulsan	-0.3979	3.22	1521	-0.123	0.9018
Mi - IPi	Seoul - Ulsan	0.5900	3.26	1521	0.181	0.8562
PWi - APi	Seoul - Ulsan	0.0733	2.60	1522	0.028	0.9775
PWi - IPi	Seoul - Ulsan	1.0612	2.64	1522	0.402	0.6878
APi - IPi	Seoul - Ulsan	0.9879	2.59	1523	0.381	0.7034
Mm - Mi	Busan - Ulsan	-0.2592	3.71	1520	-0.070	0.9443
Mm - PWi	Busan - Ulsan	0.1192	3.23	1519	0.037	0.9705
Mm - APi	Busan - Ulsan	-1.0078	3.07	1520	-0.329	0.7424
Mm - IPi	Busan - Ulsan	-0.7233	3.16	1521	-0.229	0.8189
Mi - PWi	Busan - Ulsan	0.3784	3.38	1522	0.112	0.9109
Mi - APi	Busan - Ulsan	-0.7486	3.24	1526	-0.231	0.8175
Mi - IPi	Busan - Ulsan	-0.4641	3.33	1527	-0.139	0.8892
PWi - APi	Busan - Ulsan	-1.1270	2.67	1524	-0.422	0.6728
PWi - IPi	Busan - Ulsan	-0.8425	2.77	1524	-0.304	0.7609
APi - IPi	Busan - Ulsan	0.2845	2.58	1527	0.110	0.9122

Table E.9': Interaction contrasts (outcome: raw residual voicing; data: nasal, excluding post-pausal tokens and lenited tokens). Compare with Table E.9.

position_pairwise	region_pairwise	estimate	SE	df	t.ratio	p.value
Mm - Mi	Seoul - Busan	-0.237	3.84	1516	-0.062	0.9507
Mm - PWi	Seoul - Busan	0.313	3.36	1518	0.093	0.9259
Mm - APi	Seoul - Busan	0.524	3.26	1517	0.161	0.8725
Mm - IPi	Seoul - Busan	3.526	3.30	1520	1.070	0.2850
Mi - PWi	Seoul - Busan	0.550	3.47	1518	0.159	0.8739
Mi - APi	Seoul - Busan	0.761	3.38	1519	0.225	0.8219
Mi - IPi	Seoul - Busan	3.763	3.41	1522	1.102	0.2705
PWi - APi	Seoul - Busan	0.211	2.82	1520	0.075	0.9404
PWi - IPi	Seoul - Busan	3.213	2.86	1524	1.123	0.2614
APi - IPi	Seoul - Busan	3.002	2.75	1525	1.090	0.2761
Mm - Mi	Seoul - Ulsan	1.995	3.89	1516	0.512	0.6084
Mm - PWi	Seoul - Ulsan	1.671	3.32	1517	0.503	0.6152
Mm - APi	Seoul - Ulsan	0.583	3.29	1518	0.177	0.8592
Mm - IPi	Seoul - Ulsan	1.780	3.33	1519	0.535	0.5927
Mi - PWi	Seoul - Ulsan	-0.324	3.47	1517	-0.093	0.9257
Mi - APi	Seoul - Ulsan	-1.412	3.44	1520	-0.410	0.6816
Mi - IPi	Seoul - Ulsan	-0.215	3.48	1520	-0.062	0.9507
PWi - APi	Seoul - Ulsan	-1.088	2.77	1521	-0.393	0.6947
PWi - IPi	Seoul - Ulsan	0.109	2.82	1522	0.039	0.9693
APi - IPi	Seoul - Ulsan	1.197	2.77	1522	0.432	0.6657
Mm - Mi	Busan - Ulsan	2.232	3.96	1517	0.563	0.5733
Mm - PWi	Busan - Ulsan	1.358	3.44	1518	0.394	0.6933
Mm - APi	Busan - Ulsan	0.059	3.27	1517	0.018	0.9856
Mm - IPi	Busan - Ulsan	-1.746	3.37	1520	-0.518	0.6048
Mi - PWi	Busan - Ulsan	-0.874	3.61	1518	-0.242	0.8089
Mi - APi	Busan - Ulsan	-2.173	3.47	1520	-0.627	0.5308
Mi - IPi	Busan - Ulsan	-3.978	3.56	1522	-1.118	0.2637
PWi - APi	Busan - Ulsan	-1.300	2.85	1522	-0.456	0.6485
PWi - IPi	Busan - Ulsan	-3.104	2.96	1523	-1.050	0.2937
APi - IPi	Busan - Ulsan	-1.805	2.76	1524	-0.655	0.5127

Table E.10: Pairwise comparisons of consonant by position. (outcome: normalised residual voicing; data: combined, excluding post-pausal tokens and lenited tokens).

```
position = Mm:
 contrast      estimate   SE   df t.ratio p.value
 lenis - nasal    -27.5 3.17 43.9  -8.663 <.0001

position = Mi:
 contrast      estimate   SE   df t.ratio p.value
 lenis - nasal    -24.4 3.22 48.0  -7.564 <.0001

position = PWi:
 contrast      estimate   SE   df t.ratio p.value
 lenis - nasal    -21.7 2.84 27.8  -7.622 <.0001

position = APi:
 contrast      estimate   SE   df t.ratio p.value
 lenis - nasal    -29.8 2.72 24.1 -10.973 <.0001

position = IPi:
 contrast      estimate   SE   df t.ratio p.value
 lenis - nasal    -41.9 2.75 24.7 -15.200 <.0001
```

Table E.11: Pairwise comparisons of position by region (outcome: % voicing during closure; data: lenis, excluding post-pausal tokens and lenited tokens).

region = Seoul:

contrast	estimate	SE	df	t.ratio	p.value
Mm - Mi	-6.88	6.00	1022	-1.147	0.7815
Mm - PWi	-10.04	5.05	1026	-1.988	0.2728
Mm - APi	20.53	4.94	1027	4.153	0.0003
Mm - IPi	45.82	5.03	1030	9.101	<.0001
Mi - PWi	-3.16	5.16	1025	-0.612	0.9732
Mi - APi	27.41	5.09	1027	5.387	<.0001
Mi - IPi	52.70	5.17	1029	10.201	<.0001
PWi - APi	30.57	3.54	1024	8.635	<.0001
PWi - IPi	55.86	3.67	1025	15.205	<.0001
APi - IPi	25.29	3.54	1033	7.143	<.0001

region = Busan:

contrast	estimate	SE	df	t.ratio	p.value
Mm - Mi	-3.69	6.09	1022	-0.606	0.9742
Mm - PWi	1.64	5.85	1032	0.281	0.9986
Mm - APi	5.92	4.86	1026	1.218	0.7411
Mm - IPi	30.91	5.00	1029	6.187	<.0001
Mi - PWi	5.33	6.16	1035	0.866	0.9093
Mi - APi	9.61	5.18	1026	1.854	0.3432
Mi - IPi	34.60	5.33	1032	6.488	<.0001
PWi - APi	4.28	4.56	1030	0.939	0.8817
PWi - IPi	29.27	4.63	1030	6.324	<.0001
APi - IPi	24.99	3.38	1028	7.401	<.0001

region = Ulsan:

contrast	estimate	SE	df	t.ratio	p.value
Mm - Mi	-8.19	6.30	1022	-1.300	0.6916
Mm - PWi	-11.53	5.35	1027	-2.155	0.1980
Mm - APi	19.97	5.03	1028	3.968	0.0007
Mm - IPi	34.78	5.07	1029	6.857	<.0001
Mi - PWi	-3.34	5.68	1028	-0.588	0.9768
Mi - APi	28.16	5.38	1030	5.231	<.0001
Mi - IPi	42.97	5.39	1029	7.969	<.0001
PWi - APi	31.50	3.83	1028	8.214	<.0001
PWi - IPi	46.31	3.84	1027	12.074	<.0001
APi - IPi	14.81	3.35	1024	4.426	0.0001

Table E.12: Interaction contrasts (outcome: % voicing during closure; data: lenis, excluding post-pausal tokens and lenited tokens).

position_pairwise	region_pairwise	estimate	SE	df	t.ratio	p.value
Mm - Mi	Seoul - Busan	-3.186	8.56	1021	-0.372	0.7098
Mm - PWi	Seoul - Busan	-11.679	7.53	1025	-1.551	0.1211
Mm - APi	Seoul - Busan	14.611	6.71	1020	2.177	0.0297
Mm - IPi	Seoul - Busan	14.911	6.86	1024	2.173	0.0300
Mi - PWi	Seoul - Busan	-8.493	7.84	1026	-1.083	0.2792
Mi - APi	Seoul - Busan	17.797	7.05	1021	2.524	0.0118
Mi - IPi	Seoul - Busan	18.097	7.20	1026	2.515	0.0120
PWi - APi	Seoul - Busan	26.290	5.77	1028	4.557	<.0001
PWi - IPi	Seoul - Busan	26.590	5.89	1027	4.512	<.0001
APi - IPi	Seoul - Busan	0.300	4.90	1031	0.061	0.9512
Mm - Mi	Seoul - Ulsan	1.309	8.72	1023	0.150	0.8806
Mm - PWi	Seoul - Ulsan	1.489	7.15	1022	0.208	0.8351
Mm - APi	Seoul - Ulsan	0.558	6.81	1021	0.082	0.9348
Mm - IPi	Seoul - Ulsan	11.037	6.91	1024	1.597	0.1106
Mi - PWi	Seoul - Ulsan	0.180	7.48	1022	0.024	0.9808
Mi - APi	Seoul - Ulsan	-0.752	7.19	1024	-0.105	0.9168
Mi - IPi	Seoul - Ulsan	9.728	7.25	1024	1.342	0.1800
PWi - APi	Seoul - Ulsan	-0.931	5.20	1025	-0.179	0.8580
PWi - IPi	Seoul - Ulsan	9.548	5.29	1025	1.804	0.0716
APi - IPi	Seoul - Ulsan	10.479	4.87	1030	2.150	0.0318
Mm - Mi	Busan - Ulsan	4.496	8.75	1022	0.514	0.6077
Mm - PWi	Busan - Ulsan	13.168	7.69	1025	1.712	0.0873
Mm - APi	Busan - Ulsan	-14.053	6.75	1021	-2.082	0.0376
Mm - IPi	Busan - Ulsan	-3.874	6.80	1021	-0.570	0.5688
Mi - PWi	Busan - Ulsan	8.672	8.18	1027	1.060	0.2892
Mi - APi	Busan - Ulsan	-18.549	7.28	1023	-2.547	0.0110
Mi - IPi	Busan - Ulsan	-8.370	7.34	1025	-1.141	0.2543
PWi - APi	Busan - Ulsan	-27.221	5.95	1029	-4.578	<.0001
PWi - IPi	Busan - Ulsan	-17.042	5.99	1027	-2.847	0.0045
APi - IPi	Busan - Ulsan	10.179	4.74	1026	2.146	0.0321

Table E.13: Pairwise comparisons of position by region (outcome: % voicing during closure; data: nasal, excluding post-pausal tokens and lenited tokens).

```
region = Seoul:
 contrast  estimate   SE   df t.ratio p.value
 Mm - Mi     0.0504 1.48 1520  0.034  1.0000
 Mm - PWi    0.5802 1.30 1445  0.447  0.9918
 Mm - APi    2.0837 1.31 1438  1.586  0.5067
 Mm - IPi    8.8040 1.30 1473  6.782  <.0001
 Mi - PWi    0.5298 1.33 1455  0.399  0.9947
 Mi - APi    2.0334 1.34 1453  1.512  0.5547
 Mi - IPi    8.7536 1.33 1484  6.589  <.0001
 PWi - APi   1.5035 1.08 1529  1.395  0.6310
 PWi - IPi   8.2238 1.06 1534  7.725  <.0001
 APi - IPi   6.7202 1.09 1538  6.193  <.0001

region = Busan:
 contrast  estimate   SE   df t.ratio p.value
 Mm - Mi    -0.1003 1.53 1524 -0.066  1.0000
 Mm - PWi    0.0213 1.39 1454  0.015  1.0000
 Mm - APi    0.1741 1.30 1454  0.134  0.9999
 Mm - IPi    4.0589 1.33 1485  3.044  0.0201
 Mi - PWi    0.1216 1.44 1472  0.085  1.0000
 Mi - APi    0.2744 1.35 1480  0.203  0.9996
 Mi - IPi    4.1591 1.39 1506  3.001  0.0229
 PWi - APi   0.1528 1.13 1534  0.135  0.9999
 PWi - IPi   4.0375 1.18 1544  3.436  0.0055
 APi - IPi   3.8848 1.07 1542  3.630  0.0027

region = Ulsan:
 contrast  estimate   SE   df t.ratio p.value
 Mm - Mi     0.0877 1.58 1522  0.056  1.0000
 Mm - PWi    0.1536 1.36 1484  0.113  1.0000
 Mm - APi    0.0688 1.32 1460  0.052  1.0000
 Mm - IPi    1.6194 1.36 1466  1.190  0.7573
 Mi - PWi    0.0659 1.44 1498  0.046  1.0000
 Mi - APi   -0.0189 1.41 1482 -0.013  1.0000
 Mi - IPi    1.5316 1.44 1487  1.061  0.8264
 PWi - APi  -0.0848 1.10 1537 -0.077  1.0000
 PWi - IPi   1.4658 1.14 1535  1.281  0.7029
 APi - IPi   1.5506 1.09 1531  1.425  0.6115
```

Table E.14: Interaction contrasts (outcome: % voicing during closure; data: nasal, excluding post-pausal tokens and lenited tokens).

position_pairwise	region_pairwise	estimate	SE	df	t.ratio	p.value
Mm - Mi	Seoul - Busan	0.1507	2.13	1522	0.071	0.9435
Mm - PWi	Seoul - Busan	0.5589	1.87	1526	0.299	0.7647
Mm - APi	Seoul - Busan	1.9096	1.81	1524	1.055	0.2917
Mm - IPi	Seoul - Busan	4.7451	1.83	1530	2.595	0.0096
Mi - PWi	Seoul - Busan	0.4082	1.92	1526	0.213	0.8316
Mi - APi	Seoul - Busan	1.7590	1.87	1529	0.941	0.3471
Mi - IPi	Seoul - Busan	4.5945	1.89	1535	2.432	0.0151
PWi - APi	Seoul - Busan	1.3508	1.56	1531	0.863	0.3882
PWi - IPi	Seoul - Busan	4.1862	1.58	1540	2.641	0.0083
APi - IPi	Seoul - Busan	2.8355	1.52	1540	1.861	0.0629
Mm - Mi	Seoul - Ulsan	-0.0373	2.16	1520	-0.017	0.9862
Mm - PWi	Seoul - Ulsan	0.4266	1.85	1523	0.231	0.8173
Mm - APi	Seoul - Ulsan	2.0149	1.82	1525	1.105	0.2695
Mm - IPi	Seoul - Ulsan	7.1846	1.85	1527	3.892	0.0001
Mi - PWi	Seoul - Ulsan	0.4640	1.93	1524	0.241	0.8097
Mi - APi	Seoul - Ulsan	2.0523	1.91	1530	1.075	0.2826
Mi - IPi	Seoul - Ulsan	7.2220	1.93	1531	3.744	0.0002
PWi - APi	Seoul - Ulsan	1.5883	1.54	1533	1.033	0.3018
PWi - IPi	Seoul - Ulsan	6.7580	1.56	1536	4.323	<.0001
APi - IPi	Seoul - Ulsan	5.1697	1.54	1536	3.366	0.0008
Mm - Mi	Busan - Ulsan	-0.1880	2.20	1523	-0.086	0.9318
Mm - PWi	Busan - Ulsan	-0.1323	1.91	1526	-0.069	0.9449
Mm - APi	Busan - Ulsan	0.1053	1.82	1525	0.058	0.9538
Mm - IPi	Busan - Ulsan	2.4395	1.87	1530	1.304	0.1925
Mi - PWi	Busan - Ulsan	0.0557	2.00	1525	0.028	0.9778
Mi - APi	Busan - Ulsan	0.2933	1.92	1532	0.153	0.8784
Mi - IPi	Busan - Ulsan	2.6275	1.97	1536	1.335	0.1822
PWi - APi	Busan - Ulsan	0.2376	1.58	1534	0.151	0.8804
PWi - IPi	Busan - Ulsan	2.5718	1.64	1537	1.570	0.1167
APi - IPi	Busan - Ulsan	2.3342	1.53	1537	1.530	0.1261

Table E.15: Pairwise comparisons of consonant by position. (outcome: % voicing during closure; data: combined, excluding post-pausal tokens and lenited tokens).

```
position = Mm:
 contrast      estimate   SE   df t.ratio p.value
 lenis - nasal    -22.4 3.83 46.1  -5.854 <.0001

position = Mi:
 contrast      estimate   SE   df t.ratio p.value
 lenis - nasal    -15.7 3.90 50.7  -4.024 0.0002

position = PWi:
 contrast      estimate   SE   df t.ratio p.value
 lenis - nasal    -13.7 3.40 28.1  -4.042 0.0004

position = APi:
 contrast      estimate   SE   df t.ratio p.value
 lenis - nasal    -37.8 3.24 24.0 -11.654 <.0001

position = IPi:
 contrast      estimate   SE   df t.ratio p.value
 lenis - nasal    -54.3 3.28 24.7 -16.547 <.0001
```

Table E.16: Pairwise comparisons of position by region (outcome: normalised VOT; data: lenis).

region = Seoul:

contrast	estimate	SE	df	t.ratio	p.value
Mm - Mi	1.217	2.69	1818	0.452	0.9976
Mm - PWi	-2.849	2.34	1824	-1.216	0.8294
Mm - APi	-38.236	2.37	1825	-16.111	<.0001
Mm - IPi	-44.169	2.46	1827	-17.967	<.0001
Mm - P	-32.464	2.28	1824	-14.215	<.0001
Mi - PWi	-4.066	2.34	1824	-1.735	0.5088
Mi - APi	-39.453	2.37	1825	-16.624	<.0001
Mi - IPi	-45.386	2.46	1827	-18.462	<.0001
Mi - P	-33.681	2.28	1824	-14.747	<.0001
PWi - APi	-35.387	1.85	1823	-19.095	<.0001
PWi - IPi	-41.320	1.96	1824	-21.035	<.0001
PWi - P	-29.615	1.74	1819	-17.061	<.0001
APi - IPi	-5.933	2.00	1827	-2.966	0.0361
APi - P	5.772	1.77	1821	3.269	0.0140
IPi - P	11.705	1.89	1825	6.183	<.0001

region = Busan:

contrast	estimate	SE	df	t.ratio	p.value
Mm - Mi	0.431	2.79	1819	0.155	1.0000
Mm - PWi	-3.067	2.64	1828	-1.160	0.8557
Mm - APi	-20.554	2.43	1825	-8.469	<.0001
Mm - IPi	-34.668	2.54	1826	-13.640	<.0001
Mm - P	-34.091	2.39	1825	-14.245	<.0001
Mi - PWi	-3.499	2.56	1828	-1.367	0.7467
Mi - APi	-20.985	2.34	1825	-8.972	<.0001
Mi - IPi	-35.099	2.45	1826	-14.308	<.0001
Mi - P	-34.522	2.30	1824	-15.020	<.0001
PWi - APi	-17.487	2.03	1828	-8.607	<.0001
PWi - IPi	-31.601	2.14	1825	-14.767	<.0001
PWi - P	-31.023	1.99	1827	-15.575	<.0001
APi - IPi	-14.114	1.88	1822	-7.491	<.0001
APi - P	-13.537	1.71	1821	-7.923	<.0001
IPi - P	0.577	1.85	1821	0.312	0.9996

```
region = Ulsan:
 contrast  estimate   SE   df t.ratio p.value
 Mm - Mi      1.014 2.76 1818   0.367 0.9991
 Mm - PWi    -5.130 2.46 1826  -2.086 0.2951
 Mm - APi   -25.846 2.39 1826 -10.814 <.0001
 Mm - IPi   -32.064 2.43 1826 -13.194 <.0001
 Mm - P     -28.662 2.33 1825 -12.290 <.0001
 Mi - PWi    -6.144 2.49 1826  -2.471 0.1333
 Mi - APi   -26.860 2.42 1826 -11.105 <.0001
 Mi - IPi   -33.078 2.46 1825 -13.471 <.0001
 Mi - P     -29.676 2.36 1825 -12.562 <.0001
 PWi - APi  -20.716 1.94 1827 -10.697 <.0001
 PWi - IPi  -26.934 1.98 1823 -13.634 <.0001
 PWi - P    -23.532 1.87 1825 -12.603 <.0001
 APi - IPi   -6.219 1.88 1822  -3.307 0.0123
 APi - P     -2.816 1.77 1824  -1.593 0.6030
 IPi - P      3.403 1.82 1822   1.870 0.4209
```

Table E.17: Interaction contrasts (outcome: normalised VOT; data: lenis).

position_pairwise	region_pairwise	estimate	SE	df	t.ratio	p.value
Mm - Mi	Seoul - Busan	0.785	3.88	1818	0.203	0.8395
Mm - PWi	Seoul - Busan	0.218	3.46	1821	0.063	0.9497
Mm - APi	Seoul - Busan	-17.682	3.32	1819	-5.331	<.0001
Mm - IPi	Seoul - Busan	-9.501	3.46	1821	-2.745	0.0061
Mm - P	Seoul - Busan	1.627	3.24	1819	0.502	0.6155
Mi - PWi	Seoul - Busan	-0.567	3.39	1821	-0.167	0.8673
Mi - APi	Seoul - Busan	-18.467	3.25	1819	-5.675	<.0001
Mi - IPi	Seoul - Busan	-10.287	3.40	1821	-3.028	0.0025
Mi - P	Seoul - Busan	0.841	3.17	1819	0.265	0.7907
PWi - APi	Seoul - Busan	-17.900	2.75	1826	-6.508	<.0001
PWi - IPi	Seoul - Busan	-9.719	2.90	1824	-3.348	0.0008
PWi - P	Seoul - Busan	1.409	2.64	1824	0.533	0.5941
APi - IPi	Seoul - Busan	8.181	2.75	1825	2.974	0.0030
APi - P	Seoul - Busan	19.309	2.46	1821	7.860	<.0001
IPi - P	Seoul - Busan	11.128	2.65	1824	4.205	<.0001
Mm - Mi	Seoul - Ulsan	0.203	3.86	1818	0.053	0.9580
Mm - PWi	Seoul - Ulsan	2.281	3.32	1820	0.686	0.4927
Mm - APi	Seoul - Ulsan	-12.390	3.29	1820	-3.769	0.0002
Mm - IPi	Seoul - Ulsan	-12.105	3.38	1821	-3.577	0.0004
Mm - P	Seoul - Ulsan	-3.802	3.19	1819	-1.193	0.2331
Mi - PWi	Seoul - Ulsan	2.078	3.34	1820	0.621	0.5345
Mi - APi	Seoul - Ulsan	-12.593	3.31	1820	-3.806	0.0001
Mi - IPi	Seoul - Ulsan	-12.308	3.40	1821	-3.618	0.0003
Mi - P	Seoul - Ulsan	-4.005	3.21	1819	-1.248	0.2123
PWi - APi	Seoul - Ulsan	-14.671	2.68	1825	-5.478	<.0001
PWi - IPi	Seoul - Ulsan	-14.386	2.78	1823	-5.168	<.0001
PWi - P	Seoul - Ulsan	-6.083	2.55	1822	-2.388	0.0170
APi - IPi	Seoul - Ulsan	0.286	2.75	1825	0.104	0.9172
APi - P	Seoul - Ulsan	8.588	2.50	1822	3.440	0.0006
IPi - P	Seoul - Ulsan	8.302	2.63	1824	3.162	0.0016
Mm - Mi	Busan - Ulsan	-0.583	3.93	1818	-0.148	0.8821
Mm - PWi	Busan - Ulsan	2.063	3.53	1821	0.585	0.5587
Mm - APi	Busan - Ulsan	5.292	3.32	1820	1.592	0.1116
Mm - IPi	Busan - Ulsan	-2.604	3.42	1819	-0.761	0.4469
Mm - P	Busan - Ulsan	-5.429	3.27	1819	-1.661	0.0970
Mi - PWi	Busan - Ulsan	2.645	3.49	1821	0.759	0.4480
Mi - APi	Busan - Ulsan	5.874	3.28	1820	1.789	0.0737
Mi - IPi	Busan - Ulsan	-2.021	3.38	1819	-0.598	0.5498
Mi - P	Busan - Ulsan	-4.846	3.22	1819	-1.503	0.1329
PWi - APi	Busan - Ulsan	3.229	2.80	1828	1.152	0.2496
PWi - IPi	Busan - Ulsan	-4.666	2.91	1823	-1.605	0.1087
PWi - P	Busan - Ulsan	-7.492	2.73	1825	-2.748	0.0061
APi - IPi	Busan - Ulsan	-7.895	2.66	1822	-2.968	0.0030
APi - P	Busan - Ulsan	-10.721	2.46	1821	-4.364	<.0001
IPi - P	Busan - Ulsan	-2.825	2.59	1821	-1.092	0.2750

Table E.18: Pairwise comparisons of position by region (outcome: normalised VOT; data: nasal).

region = Seoul:

contrast	estimate	SE	df	t.ratio	p.value
Mm - Mi	-0.00087	0.354	2211	-0.002	1.0000
Mm - PWi	-0.04350	0.310	2162	-0.140	1.0000
Mm - APi	-0.51965	0.318	2153	-1.633	0.5764
Mm - IPi	-0.51384	0.318	2181	-1.615	0.5884
Mm - P	-0.83538	0.307	2162	-2.723	0.0712
Mi - PWi	-0.04263	0.309	2161	-0.138	1.0000
Mi - APi	-0.51878	0.317	2151	-1.637	0.5740
Mi - IPi	-0.51297	0.317	2181	-1.619	0.5861
Mi - P	-0.83451	0.306	2160	-2.731	0.0696
PWi - APi	-0.47615	0.257	2221	-1.856	0.4300
PWi - IPi	-0.47035	0.257	2222	-1.829	0.4469
PWi - P	-0.79188	0.243	2215	-3.256	0.0146
APi - IPi	0.00581	0.267	2231	0.022	1.0000
APi - P	-0.31573	0.253	2219	-1.250	0.8120
IPi - P	-0.32153	0.254	2225	-1.268	0.8024

region = Busan:

contrast	estimate	SE	df	t.ratio	p.value
Mm - Mi	0.01303	0.362	2216	0.036	1.0000
Mm - PWi	0.00219	0.323	2157	0.007	1.0000
Mm - APi	0.04292	0.314	2153	0.137	1.0000
Mm - IPi	-0.34633	0.324	2188	-1.069	0.8935
Mm - P	-2.50869	0.312	2184	-8.033	<.0001
Mi - PWi	-0.01083	0.329	2170	-0.033	1.0000
Mi - APi	0.02989	0.320	2166	0.094	1.0000
Mi - IPi	-0.35936	0.330	2199	-1.089	0.8858
Mi - P	-2.52172	0.318	2190	-7.934	<.0001
PWi - APi	0.04072	0.264	2228	0.154	1.0000
PWi - IPi	-0.34853	0.275	2231	-1.267	0.8029
PWi - P	-2.51088	0.263	2228	-9.557	<.0001
APi - IPi	-0.38925	0.265	2236	-1.467	0.6851
APi - P	-2.55161	0.252	2224	-10.140	<.0001
IPi - P	-2.16236	0.265	2238	-8.156	<.0001

```
region = Ulsan:
 contrast  estimate    SE   df t.ratio p.value
 Mm - Mi   -0.00490 0.358 2212  -0.014 1.0000
 Mm - PWi   0.04676 0.314 2163   0.149 1.0000
 Mm - APi   0.04400 0.312 2151   0.141 1.0000
 Mm - IPi  -0.26042 0.325 2168  -0.801 0.9675
 Mm - P    -1.05082 0.307 2151  -3.419 0.0084
 Mi - PWi   0.05166 0.321 2171   0.161 1.0000
 Mi - APi   0.04891 0.319 2159   0.153 1.0000
 Mi - IPi  -0.25552 0.332 2174  -0.770 0.9725
 Mi - P    -1.04591 0.314 2161  -3.328 0.0115
 PWi - APi -0.00276 0.257 2231  -0.011 1.0000
 PWi - IPi -0.30718 0.273 2228  -1.126 0.8705
 PWi - P   -1.09758 0.251 2226  -4.370 0.0002
 APi - IPi -0.30442 0.270 2223  -1.127 0.8704
 APi - P   -1.09482 0.249 2226  -4.402 0.0002
 IPi - P   -0.79040 0.266 2233  -2.975 0.0350
```

Table E.19: Interaction contrasts (outcome: normalised VOT; data: nasal).

position_pairwise	region_pairwise	estimate	SE	df	t.ratio	p.value
Mm - Mi	Seoul - Busan	-0.01389	0.506	2214	-0.027	0.9781
Mm - PWi	Seoul - Busan	-0.04569	0.441	2216	-0.104	0.9175
Mm - APi	Seoul - Busan	-0.56257	0.440	2215	-1.279	0.2011
Mm - IPi	Seoul - Busan	-0.16751	0.448	2223	-0.374	0.7085
Mm - P	Seoul - Busan	1.67331	0.432	2213	3.874	0.0001
Mi - PWi	Seoul - Busan	-0.03180	0.444	2218	-0.072	0.9430
Mi - APi	Seoul - Busan	-0.54867	0.443	2219	-1.238	0.2160
Mi - IPi	Seoul - Busan	-0.15362	0.452	2228	-0.340	0.7338
Mi - P	Seoul - Busan	1.68721	0.435	2215	3.878	0.0001
PWi - APi	Seoul - Busan	-0.51688	0.368	2224	-1.405	0.1600
PWi - IPi	Seoul - Busan	-0.12182	0.376	2226	-0.324	0.7462
PWi - P	Seoul - Busan	1.71900	0.358	2222	4.803	<.0001
APi - IPi	Seoul - Busan	0.39506	0.376	2233	1.050	0.2936
APi - P	Seoul - Busan	2.23588	0.356	2221	6.273	<.0001
IPi - P	Seoul - Busan	1.84082	0.367	2234	5.018	<.0001
Mm - Mi	Seoul - Ulsan	0.00403	0.503	2212	0.008	0.9936
Mm - PWi	Seoul - Ulsan	-0.09026	0.435	2215	-0.208	0.8356
Mm - APi	Seoul - Ulsan	-0.56365	0.439	2217	-1.285	0.1991
Mm - IPi	Seoul - Ulsan	-0.25343	0.449	2219	-0.565	0.5724
Mm - P	Seoul - Ulsan	0.21544	0.428	2213	0.504	0.6145
Mi - PWi	Seoul - Ulsan	-0.09429	0.439	2216	-0.215	0.8299
Mi - APi	Seoul - Ulsan	-0.56769	0.443	2218	-1.282	0.2000
Mi - IPi	Seoul - Ulsan	-0.25746	0.453	2220	-0.569	0.5695
Mi - P	Seoul - Ulsan	0.21141	0.432	2215	0.489	0.6246
PWi - APi	Seoul - Ulsan	-0.47340	0.363	2227	-1.303	0.1926
PWi - IPi	Seoul - Ulsan	-0.16317	0.375	2226	-0.435	0.6634
PWi - P	Seoul - Ulsan	0.30570	0.350	2221	0.874	0.3820
APi - IPi	Seoul - Ulsan	0.31023	0.379	2228	0.818	0.4137
APi - P	Seoul - Ulsan	0.77909	0.354	2223	2.198	0.0280
IPi - P	Seoul - Ulsan	0.46886	0.367	2229	1.277	0.2018
Mm - Mi	Busan - Ulsan	0.01793	0.510	2214	0.035	0.9719
Mm - PWi	Busan - Ulsan	-0.04457	0.444	2217	-0.100	0.9200
Mm - APi	Busan - Ulsan	-0.00109	0.436	2216	-0.002	0.9980
Mm - IPi	Busan - Ulsan	-0.08591	0.453	2223	-0.190	0.8495
Mm - P	Busan - Ulsan	-1.45787	0.432	2216	-3.374	0.0008
Mi - PWi	Busan - Ulsan	-0.06249	0.453	2219	-0.138	0.8902
Mi - APi	Busan - Ulsan	-0.01901	0.445	2221	-0.043	0.9659
Mi - IPi	Busan - Ulsan	-0.10384	0.462	2227	-0.225	0.8221
Mi - P	Busan - Ulsan	-1.47580	0.441	2218	-3.345	0.0008
PWi - APi	Busan - Ulsan	0.04348	0.368	2230	0.118	0.9061
PWi - IPi	Busan - Ulsan	-0.04135	0.387	2228	-0.107	0.9150
PWi - P	Busan - Ulsan	-1.41331	0.363	2227	-3.890	0.0001
APi - IPi	Busan - Ulsan	-0.08483	0.378	2230	-0.224	0.8227
APi - P	Busan - Ulsan	-1.45679	0.354	2225	-4.118	<.0001
IPi - P	Busan - Ulsan	-1.37196	0.375	2236	-3.656	0.0003

Table E.20: Pairwise comparisons of consonant by position. (outcome: normalised VOT; data: combined).

```
position = Mm:
 contrast      estimate   SE   df t.ratio p.value
 lenis - nasal  -0.0551 1.86 50.2 -0.030  0.9765

position = Mi:
 contrast      estimate   SE   df t.ratio p.value
 lenis - nasal  -0.9956 1.85 50.6 -0.537  0.5936

position = PWi:
 contrast      estimate   SE   df t.ratio p.value
 lenis - nasal   3.5846 1.65 31.1  2.171  0.0377

position = APi:
 contrast      estimate   SE   df t.ratio p.value
 lenis - nasal  28.3951 1.62 29.6 17.480  <.0001

position = IPi:
 contrast      estimate   SE   df t.ratio p.value
 lenis - nasal  36.7810 1.66 31.8 22.104  <.0001

position = P:
 contrast      estimate   SE   df t.ratio p.value
 lenis - nasal  30.4884 1.60 28.0 19.015  <.0001
```

Table E.21: Pairwise comparisons of position by region (outcome: normalised total voiceless interval; data: lenis, excluding post-pausal tokens).

```
region = Seoul:
 contrast  estimate   SE   df t.ratio p.value
 Mm - Mi       4.31 5.00 1352   0.862 0.9108
 Mm - PWi      0.89 4.39 1350   0.203 0.9996
 Mm - APi    -50.82 4.45 1347 -11.415 <.0001
 Mm - IPi    -89.53 4.61 1359 -19.426 <.0001
 Mi - PWi     -3.42 4.39 1350  -0.779 0.9368
 Mi - APi    -55.12 4.45 1347 -12.383 <.0001
 Mi - IPi    -93.84 4.61 1359 -20.361 <.0001
 PWi - APi   -51.71 3.45 1359 -14.968 <.0001
 PWi - IPi   -90.42 3.66 1359 -24.690 <.0001
 APi - IPi   -38.72 3.74 1363 -10.354 <.0001

region = Busan:
 contrast  estimate   SE   df t.ratio p.value
 Mm - Mi       1.01 5.19 1353   0.195 0.9997
 Mm - PWi     -2.13 4.97 1349  -0.428 0.9930
 Mm - APi    -25.24 4.55 1346  -5.546 <.0001
 Mm - IPi    -62.49 4.77 1339 -13.087 <.0001
 Mi - PWi     -3.14 4.81 1347  -0.653 0.9662
 Mi - APi    -26.25 4.39 1344  -5.982 <.0001
 Mi - IPi    -63.50 4.61 1335 -13.777 <.0001
 PWi - APi   -23.11 3.80 1365  -6.076 <.0001
 PWi - IPi   -60.36 4.00 1361 -15.101 <.0001
 APi - IPi   -37.25 3.51 1358 -10.608 <.0001

region = Ulsan:
 contrast  estimate   SE   df t.ratio p.value
 Mm - Mi       3.30 5.14 1352   0.643 0.9679
 Mm - PWi     -1.73 4.62 1347  -0.374 0.9958
 Mm - APi    -39.35 4.49 1342  -8.770 <.0001
 Mm - IPi    -63.13 4.56 1338 -13.832 <.0001
 Mi - PWi     -5.03 4.67 1349  -1.078 0.8180
 Mi - APi    -42.66 4.54 1344  -9.395 <.0001
 Mi - IPi    -66.43 4.61 1341 -14.413 <.0001
 PWi - APi   -37.63 3.62 1363 -10.404 <.0001
 PWi - IPi   -61.40 3.68 1360 -16.667 <.0001
 APi - IPi   -23.78 3.50 1356  -6.790 <.0001
```

Table E.22: Interaction contrasts (outcome: normalised total voiceless interval; data: lenis, excluding post-pausal tokens).

position_pairwise	region_pairwise	estimate	SE	df	t.ratio	p.value
Mm - Mi	Seoul - Busan	3.296	7.21	1352	0.457	0.6476
Mm - PWi	Seoul - Busan	3.018	6.44	1357	0.468	0.6395
Mm - APi	Seoul - Busan	-25.578	6.17	1353	-4.145	<.0001
Mm - IPi	Seoul - Busan	-27.045	6.45	1357	-4.194	<.0001
Mi - PWi	Seoul - Busan	-0.278	6.32	1356	-0.044	0.9649
Mi - APi	Seoul - Busan	-28.874	6.05	1353	-4.769	<.0001
Mi - IPi	Seoul - Busan	-30.341	6.33	1356	-4.796	<.0001
PWi - APi	Seoul - Busan	-28.595	5.14	1363	-5.563	<.0001
PWi - IPi	Seoul - Busan	-30.062	5.42	1359	-5.549	<.0001
APi - IPi	Seoul - Busan	-1.467	5.14	1362	-0.286	0.7752
Mm - Mi	Seoul - Ulsan	1.006	7.17	1352	0.140	0.8885
Mm - PWi	Seoul - Ulsan	2.616	6.18	1354	0.423	0.6724
Mm - APi	Seoul - Ulsan	-11.463	6.12	1354	-1.874	0.0611
Mm - IPi	Seoul - Ulsan	-26.402	6.30	1356	-4.190	<.0001
Mi - PWi	Seoul - Ulsan	1.610	6.22	1354	0.259	0.7959
Mi - APi	Seoul - Ulsan	-12.468	6.16	1354	-2.025	0.0431
Mi - IPi	Seoul - Ulsan	-27.407	6.33	1355	-4.327	<.0001
PWi - APi	Seoul - Ulsan	-14.078	5.00	1360	-2.818	0.0049
PWi - IPi	Seoul - Ulsan	-29.018	5.19	1358	-5.592	<.0001
APi - IPi	Seoul - Ulsan	-14.939	5.12	1361	-2.916	0.0036
Mm - Mi	Busan - Ulsan	-2.290	7.30	1352	-0.314	0.7538
Mm - PWi	Busan - Ulsan	-0.402	6.57	1356	-0.061	0.9512
Mm - APi	Busan - Ulsan	14.115	6.18	1354	2.282	0.0226
Mm - IPi	Busan - Ulsan	0.643	6.37	1353	0.101	0.9196
Mi - PWi	Busan - Ulsan	1.888	6.49	1356	0.291	0.7713
Mi - APi	Busan - Ulsan	16.405	6.11	1354	2.686	0.0073
Mi - IPi	Busan - Ulsan	2.933	6.28	1353	0.467	0.6408
PWi - APi	Busan - Ulsan	14.517	5.24	1364	2.769	0.0057
PWi - IPi	Busan - Ulsan	1.045	5.42	1359	0.193	0.8473
APi - IPi	Busan - Ulsan	-13.472	4.96	1356	-2.719	0.0066

Table E.23: Pairwise comparisons of position by region (outcome: normalised total voiceless interval; data: nasal, excluding post-pausal tokens).

```
region = Seoul:
 contrast   estimate   SE   df t.ratio p.value
 Mm - Mi    -0.01659 2.07 1711 -0.008  1.0000
 Mm - PWi   -0.30770 1.82 1519 -0.169  0.9998
 Mm - APi   -2.52358 1.86 1493 -1.353  0.6576
 Mm - IPi  -13.04092 1.87 1559 -6.992  <.0001
 Mi - PWi   -0.29112 1.81 1515 -0.161  0.9998
 Mi - APi   -2.50699 1.86 1488 -1.350  0.6600
 Mi - IPi  -13.02433 1.86 1557 -7.009  <.0001
 PWi - APi  -2.21588 1.50 1720 -1.473  0.5802
 PWi - IPi -12.73321 1.51 1720 -8.446  <.0001
 APi - IPi -10.51734 1.57 1728 -6.715  <.0001

region = Busan:
 contrast   estimate   SE   df t.ratio p.value
 Mm - Mi     0.03792 2.12 1715  0.018  1.0000
 Mm - PWi   -0.09359 1.89 1499 -0.049  1.0000
 Mm - APi    0.02402 1.84 1497  0.013  1.0000
 Mm - IPi   -3.82049 1.90 1571 -2.010  0.2614
 Mi - PWi   -0.13151 1.93 1525 -0.068  1.0000
 Mi - APi   -0.01390 1.87 1520 -0.007  1.0000
 Mi - IPi   -3.85841 1.94 1590 -1.991  0.2709
 PWi - APi   0.11761 1.55 1727  0.076  1.0000
 PWi - IPi  -3.72691 1.61 1730 -2.309  0.1425
 APi - IPi  -3.84451 1.56 1734 -2.466  0.0990

region = Ulsan:
 contrast   estimate   SE   df t.ratio p.value
 Mm - Mi    -0.06600 2.10 1712 -0.031  1.0000
 Mm - PWi   -0.06348 1.84 1516 -0.035  1.0000
 Mm - APi    0.08858 1.83 1486  0.048  1.0000
 Mm - IPi   -1.78843 1.91 1529 -0.938  0.8821
 Mi - PWi    0.00252 1.88 1532  0.001  1.0000
 Mi - APi    0.15458 1.87 1501  0.083  1.0000
 Mi - IPi   -1.72243 1.94 1541 -0.886  0.9021
 PWi - APi   0.15206 1.51 1729  0.101  1.0000
 PWi - IPi  -1.72495 1.60 1726 -1.078  0.8178
 APi - IPi  -1.87701 1.58 1722 -1.185  0.7603
```

Table E.24: Interaction contrasts (outcome: normalised total voiceless interval; data: nasal, excluding post-pausal tokens).

position_pairwise	region_pairwise	estimate	SE	df	t.ratio	p.value
Mm - Mi	Seoul - Busan	-0.0545	2.96	1713	-0.018	0.9853
Mm - PWi	Seoul - Busan	-0.2141	2.59	1715	-0.083	0.9340
Mm - APi	Seoul - Busan	-2.5476	2.58	1714	-0.989	0.3230
Mm - IPi	Seoul - Busan	-9.2204	2.63	1721	-3.509	0.0005
Mi - PWi	Seoul - Busan	-0.1596	2.60	1717	-0.061	0.9512
Mi - APi	Seoul - Busan	-2.4931	2.60	1717	-0.960	0.3373
Mi - IPi	Seoul - Busan	-9.1659	2.65	1725	-3.459	0.0006
PWi - APi	Seoul - Busan	-2.3335	2.16	1722	-1.082	0.2796
PWi - IPi	Seoul - Busan	-9.0063	2.21	1724	-4.080	<.0001
APi - IPi	Seoul - Busan	-6.6728	2.21	1730	-3.020	0.0026
Mm - Mi	Seoul - Ulsan	0.0494	2.95	1712	0.017	0.9866
Mm - PWi	Seoul - Ulsan	-0.2442	2.55	1714	-0.096	0.9236
Mm - APi	Seoul - Ulsan	-2.6122	2.57	1716	-1.016	0.3098
Mm - IPi	Seoul - Ulsan	-11.2525	2.63	1718	-4.278	<.0001
Mi - PWi	Seoul - Ulsan	-0.2936	2.57	1715	-0.114	0.9091
Mi - APi	Seoul - Ulsan	-2.6616	2.60	1717	-1.026	0.3052
Mi - IPi	Seoul - Ulsan	-11.3019	2.65	1719	-4.261	<.0001
PWi - APi	Seoul - Ulsan	-2.3679	2.13	1724	-1.111	0.2667
PWi - IPi	Seoul - Ulsan	-11.0083	2.20	1724	-5.008	<.0001
APi - IPi	Seoul - Ulsan	-8.6403	2.23	1725	-3.880	0.0001
Mm - Mi	Busan - Ulsan	0.1039	2.98	1714	0.035	0.9722
Mm - PWi	Busan - Ulsan	-0.0301	2.60	1717	-0.012	0.9908
Mm - APi	Busan - Ulsan	-0.0646	2.55	1715	-0.025	0.9798
Mm - IPi	Busan - Ulsan	-2.0321	2.65	1721	-0.765	0.4441
Mi - PWi	Busan - Ulsan	-0.1340	2.65	1718	-0.050	0.9597
Mi - APi	Busan - Ulsan	-0.1685	2.61	1719	-0.065	0.9485
Mi - IPi	Busan - Ulsan	-2.1360	2.71	1725	-0.788	0.4306
PWi - APi	Busan - Ulsan	-0.0345	2.16	1728	-0.016	0.9873
PWi - IPi	Busan - Ulsan	-2.0020	2.27	1725	-0.881	0.3783
APi - IPi	Busan - Ulsan	-1.9675	2.22	1727	-0.885	0.3761

Table E.25: Pairwise comparisons of consonant by position. (outcome: normalised total voiceless interval; data: combined, excluding post-pausal tokens).

```
position = Mm:
 contrast      estimate   SE   df t.ratio p.value
 lenis - nasal     8.97 2.90 86.0  3.096  0.0026

position = Mi:
 contrast      estimate   SE   df t.ratio p.value
 lenis - nasal     5.96 2.89 87.1  2.060  0.0424

position = PWi:
 contrast      estimate   SE   df t.ratio p.value
 lenis - nasal     9.16 2.27 38.8  4.030  0.0003

position = APi:
 contrast      estimate   SE   df t.ratio p.value
 lenis - nasal    47.61 2.20 35.0 21.596  <.0001

position = IPi:
 contrast      estimate   SE   df t.ratio p.value
 lenis - nasal    74.75 2.31 40.7 32.415  <.0001
```

Table E.26: Pairwise comparisons of category by region (outcome: normalised prevoicing; data: nasal, only post-pausal tokens).

```
region = Seoul:
 contrast estimate   SE  df t.ratio p.value
 n - nd     -4.610 3.36 479 -1.373  0.6450
 n - d      10.921 2.97 484  3.678  0.0024
 n - t      28.779 4.88 477  5.901  <.0001
 n - th     26.163 5.14 466  5.094  <.0001
 nd - d     15.532 3.37 482  4.615  <.0001
 nd - t     33.390 5.22 476  6.401  <.0001
 nd - th    30.774 5.46 469  5.639  <.0001
 d - t      17.858 4.72 467  3.784  0.0016
 d - th     15.242 5.20 469  2.933  0.0289
 t - th     -2.616 6.60 472 -0.396  0.9948

region = Busan:
 contrast estimate   SE  df t.ratio p.value
 n - nd    -14.817 4.56 465 -3.248  0.0109
 n - d       9.571 3.12 471  3.067  0.0193
 n - t      23.998 3.50 482  6.858  <.0001
 n - th     24.654 3.73 478  6.609  <.0001
 nd - d     24.389 4.66 464  5.229  <.0001
 nd - t     38.816 4.96 472  7.833  <.0001
 nd - th    39.471 5.06 466  7.800  <.0001
 d - t      14.427 3.44 470  4.190  0.0003
 d - th     15.083 3.76 472  4.013  0.0007
 t - th      0.656 3.91 471  0.168  0.9998

region = Ulsan:
 contrast estimate   SE  df t.ratio p.value
 n - nd     -3.924 4.19 470 -0.937  0.8825
 n - d       5.050 2.76 474  1.827  0.3590
 n - t      27.960 3.80 475  7.352  <.0001
 n - th     25.007 4.67 473  5.358  <.0001
 nd - d      8.975 4.48 467  2.004  0.2654
 nd - t     31.884 5.14 468  6.205  <.0001
 nd - th    28.931 5.93 473  4.881  <.0001
 d - t      22.910 4.04 466  5.669  <.0001
 d - th     19.957 4.89 471  4.085  0.0005
 t - th     -2.953 5.42 462 -0.545  0.9825
```

Table E.27: Interaction contrasts (outcome: normalised prevoicing; data: nasal, only post-pausal tokens).

cat_pairwise	region_pairwise	estimate	SE	df	t.ratio	p.value
n - nd	Seoul - Busan	10.207	5.64	469	1.810	0.0709
n - d	Seoul - Busan	1.350	4.29	480	0.315	0.7530
n - t	Seoul - Busan	4.781	5.99	480	0.798	0.4255
n - th	Seoul - Busan	1.510	6.35	473	0.238	0.8122
nd - d	Seoul - Busan	-8.857	5.72	470	-1.548	0.1223
nd - t	Seoul - Busan	-5.426	7.20	475	-0.754	0.4514
nd - th	Seoul - Busan	-8.697	7.43	468	-1.170	0.2426
d - t	Seoul - Busan	3.431	5.87	470	0.584	0.5592
d - th	Seoul - Busan	0.159	6.42	471	0.025	0.9802
t - th	Seoul - Busan	-3.272	7.71	472	-0.424	0.6716
n - nd	Seoul - Ulsan	-0.686	5.36	473	-0.128	0.8982
n - d	Seoul - Ulsan	5.871	4.05	483	1.449	0.1481
n - t	Seoul - Ulsan	0.819	6.18	476	0.133	0.8946
n - th	Seoul - Ulsan	1.156	6.93	472	0.167	0.8676
nd - d	Seoul - Ulsan	6.557	5.60	474	1.171	0.2422
nd - t	Seoul - Ulsan	1.505	7.31	473	0.206	0.8368
nd - th	Seoul - Ulsan	1.842	8.09	473	0.228	0.8199
d - t	Seoul - Ulsan	-5.052	6.20	467	-0.815	0.4156
d - th	Seoul - Ulsan	-4.715	7.18	472	-0.656	0.5120
t - th	Seoul - Ulsan	0.337	8.59	470	0.039	0.9687
n - nd	Busan - Ulsan	-10.893	6.17	466	-1.765	0.0783
n - d	Busan - Ulsan	4.521	4.17	474	1.085	0.2786
n - t	Busan - Ulsan	-3.962	5.13	476	-0.772	0.4404
n - th	Busan - Ulsan	-0.353	5.96	475	-0.059	0.9528
nd - d	Busan - Ulsan	15.414	6.45	466	2.390	0.0172
nd - t	Busan - Ulsan	6.931	7.10	469	0.976	0.3295
nd - th	Busan - Ulsan	10.540	7.75	469	1.360	0.1745
d - t	Busan - Ulsan	-8.483	5.29	466	-1.603	0.1097
d - th	Busan - Ulsan	-4.874	6.13	470	-0.796	0.4267
t - th	Busan - Ulsan	3.609	6.67	464	0.541	0.5889

Table E.28: Pairwise comparisons of position by region (outcome: normalised minimum intensity; data: nasals in Categories (1) – (3), excluding post-pausal tokens and lenited tokens).

```
region = Seoul:
 contrast  estimate    SE   df t.ratio p.value
 Mm - Mi    -0.7710 0.746 1484  -1.033 0.8401
 Mm - PWi   -0.0284 0.661 1494  -0.043 1.0000
 Mm - APi   -6.7140 0.676 1494  -9.931 <.0001
 Mm - IPi  -13.6364 0.669 1496 -20.374 <.0001
 Mi - PWi    0.7426 0.676 1494   1.099 0.8070
 Mi - APi   -5.9430 0.691 1494  -8.599 <.0001
 Mi - IPi  -12.8654 0.684 1496 -18.809 <.0001
 PWi - APi  -6.6856 0.555 1489 -12.056 <.0001
 PWi - IPi -13.6080 0.550 1492 -24.726 <.0001
 APi - IPi  -6.9224 0.569 1494 -12.175 <.0001

region = Busan:
 contrast  estimate    SE   df t.ratio p.value
 Mm - Mi    -0.5692 0.774 1486  -0.736 0.9481
 Mm - PWi   -0.4695 0.708 1495  -0.663 0.9641
 Mm - APi   -4.2257 0.661 1493  -6.390 <.0001
 Mm - IPi  -10.6998 0.687 1499 -15.571 <.0001
 Mi - PWi    0.0997 0.733 1495   0.136 0.9999
 Mi - APi   -3.6566 0.691 1497  -5.288 <.0001
 Mi - IPi  -10.1307 0.717 1502 -14.127 <.0001
 PWi - APi  -3.7563 0.575 1490  -6.534 <.0001
 PWi - IPi -10.2304 0.605 1497 -16.901 <.0001
 APi - IPi  -6.4741 0.554 1497 -11.685 <.0001

region = Ulsan:
 contrast  estimate    SE   df t.ratio p.value
 Mm - Mi    -0.0904 0.796 1485  -0.114 1.0000
 Mm - PWi    1.0197 0.689 1493   1.479 0.5763
 Mm - APi   -2.9252 0.670 1495  -4.363 0.0001
 Mm - IPi   -9.4618 0.694 1495 -13.629 <.0001
 Mi - PWi    1.1101 0.730 1493   1.521 0.5487
 Mi - APi   -2.8348 0.714 1498  -3.969 0.0007
 Mi - IPi   -9.3714 0.735 1497 -12.746 <.0001
 PWi - APi  -3.9449 0.555 1492  -7.114 <.0001
 PWi - IPi -10.4815 0.582 1490 -17.997 <.0001
 APi - IPi  -6.5366 0.555 1489 -11.784 <.0001
```

Table E.29: Interaction contrasts (outcome: normalised minimum intensity; data: nasals in Categories (1) – (3), excluding post-pausal tokens and lenited tokens).

position_pairwise	region_pairwise	estimate	SE	df	t.ratio	p.value
Mm - Mi	Seoul - Busan	-0.2019	1.075	1485	-0.188	0.8511
Mm - PWi	Seoul - Busan	0.4410	0.943	1487	0.468	0.6401
Mm - APi	Seoul - Busan	-2.4882	0.920	1485	-2.704	0.0069
Mm - IPi	Seoul - Busan	-2.9365	0.937	1490	-3.135	0.0018
Mi - PWi	Seoul - Busan	0.6429	0.972	1486	0.662	0.5084
Mi - APi	Seoul - Busan	-2.2864	0.953	1488	-2.399	0.0165
Mi - IPi	Seoul - Busan	-2.7347	0.969	1492	-2.821	0.0048
PWi - APi	Seoul - Busan	-2.9293	0.799	1489	-3.668	0.0003
PWi - IPi	Seoul - Busan	-3.3776	0.818	1495	-4.130	<.0001
APi - IPi	Seoul - Busan	-0.4483	0.794	1496	-0.565	0.5722
Mm - Mi	Seoul - Ulsan	-0.6806	1.091	1484	-0.624	0.5328
Mm - PWi	Seoul - Ulsan	-1.0481	0.932	1485	-1.124	0.2611
Mm - APi	Seoul - Ulsan	-3.7888	0.926	1487	-4.090	<.0001
Mm - IPi	Seoul - Ulsan	-4.1746	0.941	1488	-4.435	<.0001
Mi - PWi	Seoul - Ulsan	-0.3675	0.973	1486	-0.378	0.7056
Mi - APi	Seoul - Ulsan	-3.1082	0.969	1488	-3.207	0.0014
Mi - IPi	Seoul - Ulsan	-3.4940	0.982	1489	-3.557	0.0004
PWi - APi	Seoul - Ulsan	-2.7407	0.784	1491	-3.496	0.0005
PWi - IPi	Seoul - Ulsan	-3.1265	0.801	1491	-3.902	0.0001
APi - IPi	Seoul - Ulsan	-0.3858	0.794	1493	-0.486	0.6270
Mm - Mi	Busan - Ulsan	-0.4788	1.110	1485	-0.431	0.6664
Mm - PWi	Busan - Ulsan	-1.4892	0.965	1486	-1.543	0.1230
Mm - APi	Busan - Ulsan	-1.3006	0.917	1486	-1.418	0.1565
Mm - IPi	Busan - Ulsan	-1.2381	0.954	1489	-1.298	0.1945
Mi - PWi	Busan - Ulsan	-1.0104	1.012	1486	-0.998	0.3183
Mi - APi	Busan - Ulsan	-0.8218	0.971	1490	-0.846	0.3976
Mi - IPi	Busan - Ulsan	-0.7593	1.005	1492	-0.755	0.4502
PWi - APi	Busan - Ulsan	0.1886	0.799	1491	0.236	0.8133
PWi - IPi	Busan - Ulsan	0.2511	0.839	1493	0.299	0.7646
APi - IPi	Busan - Ulsan	0.0625	0.783	1494	0.080	0.9364

Appendix F.

Probability density plots

I. % Voicing during closure for lenis stops

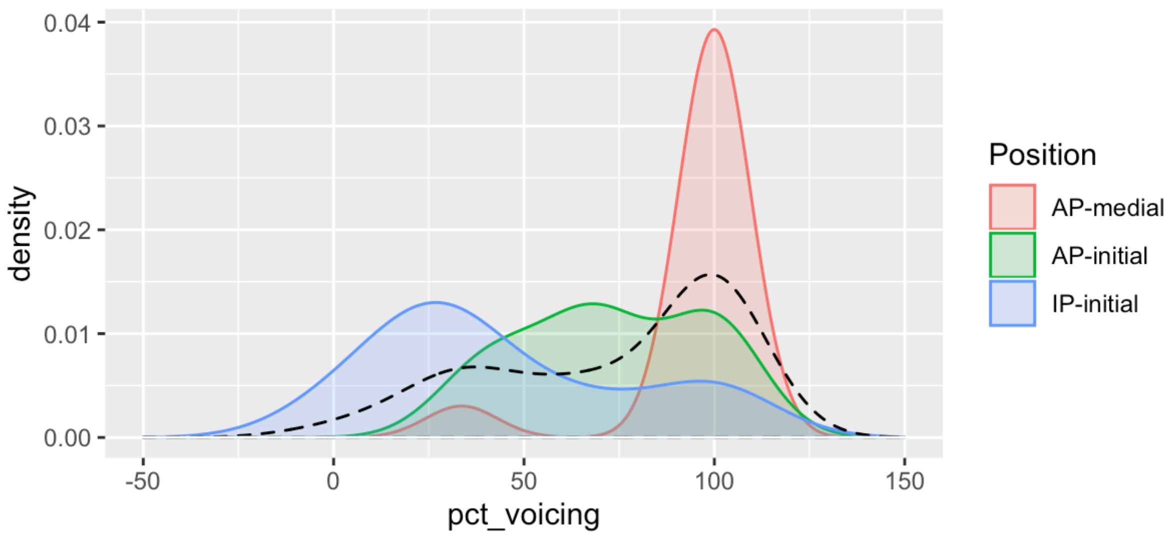


Figure F.1: Density plot of % voicing for lenis stops produced by Seoul_1.

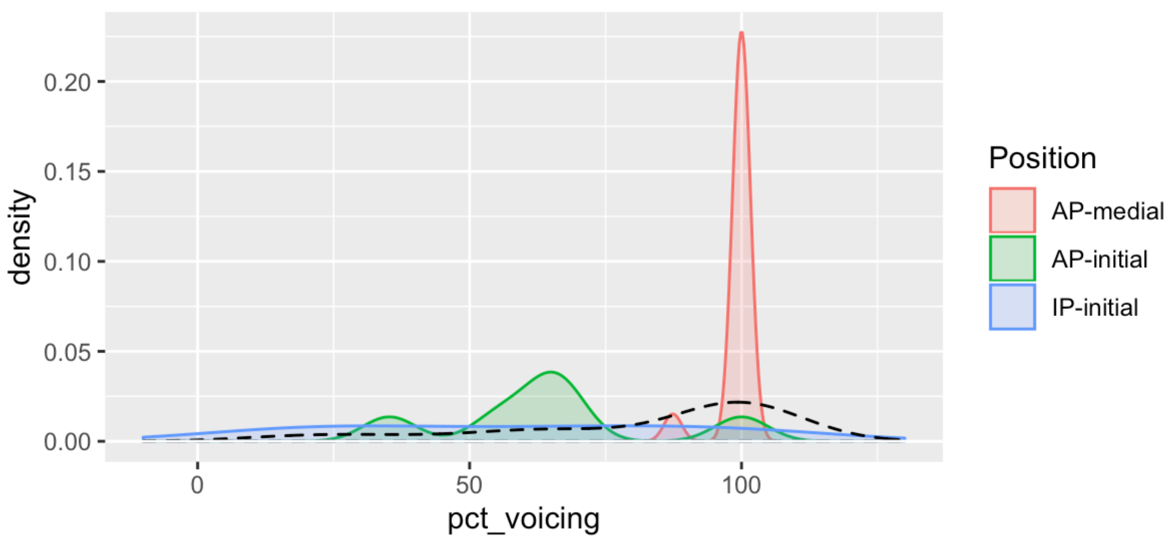


Figure F.2: Density plot of % voicing for lenis stops produced by Seoul_2.

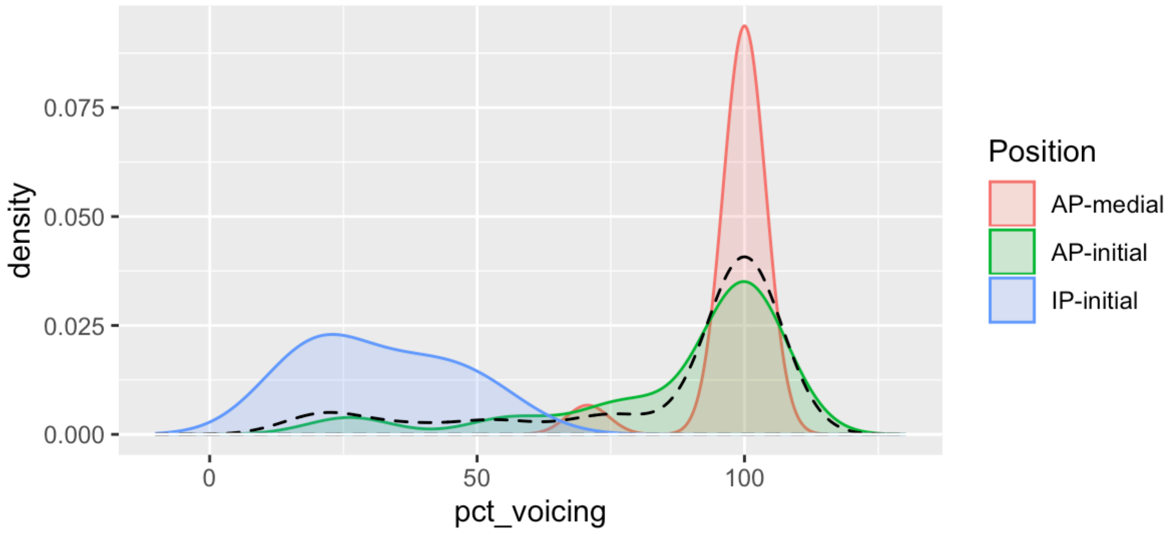


Figure F.3: Density plot of % voicing for lenis stops produced by Seoul_3.

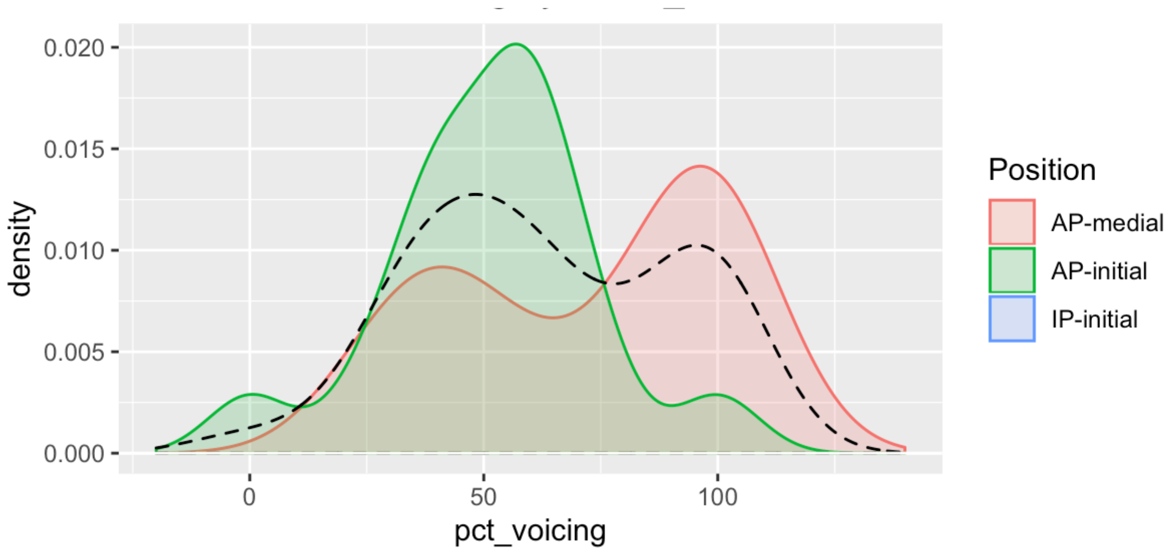


Figure F.4: Density plot of % voicing for lenis stops produced by Seoul_4.

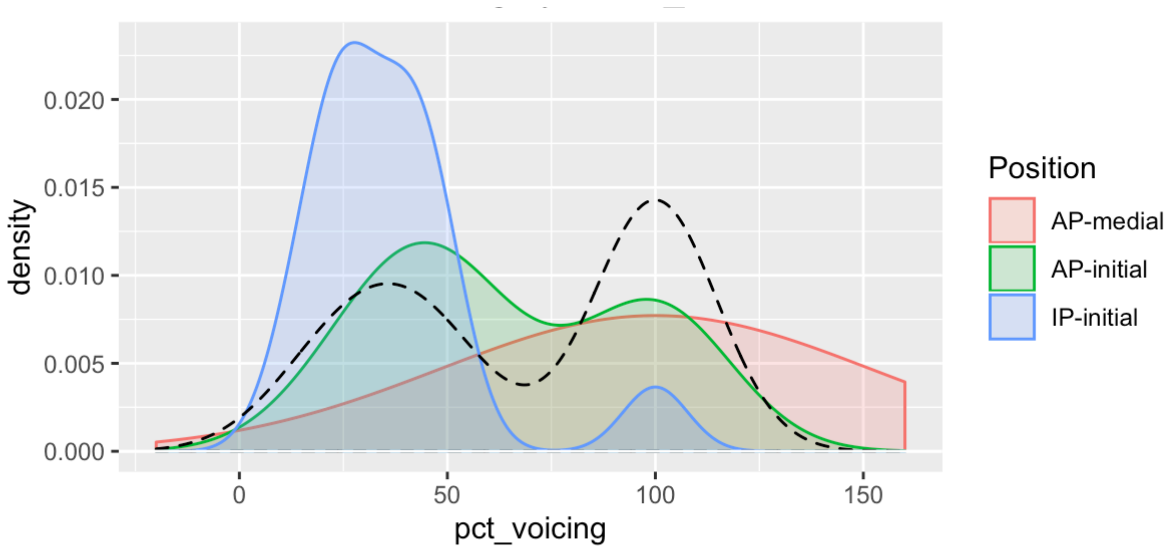


Figure F.5: Density plot of % voicing for lenis stops produced by Seoul_5.

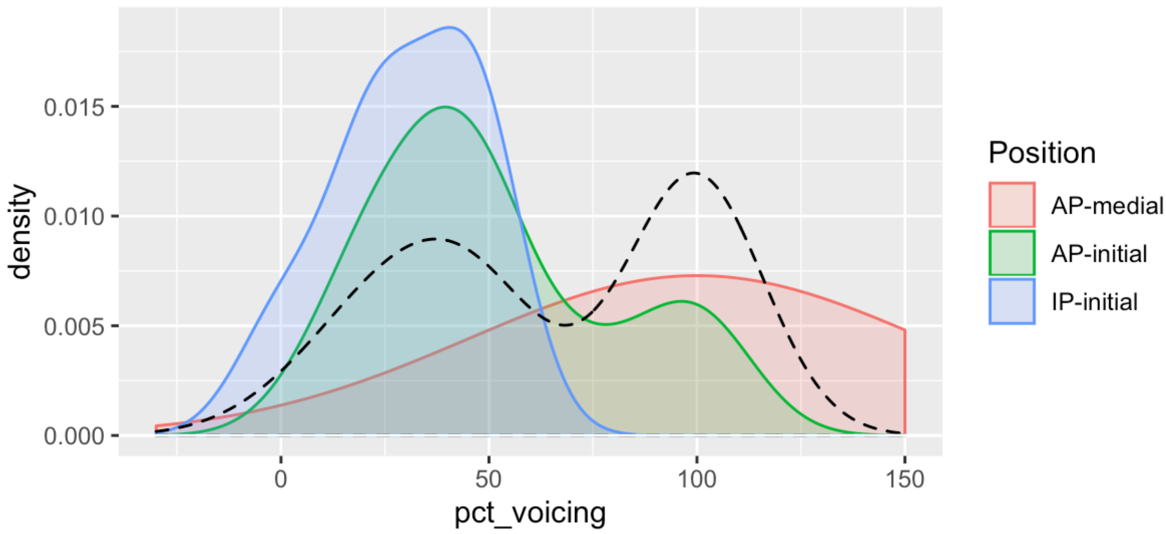


Figure F.6: Density plot of % voicing for lenis stops produced by Seoul_6.

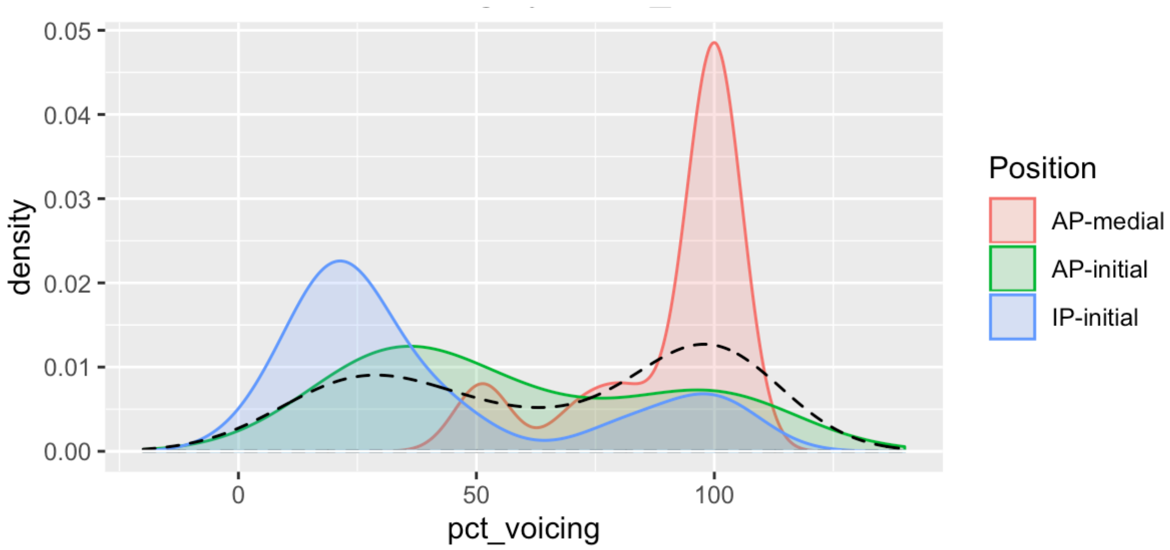


Figure F.7: Density plot of % voicing for lenis stops produced by Seoul_7.

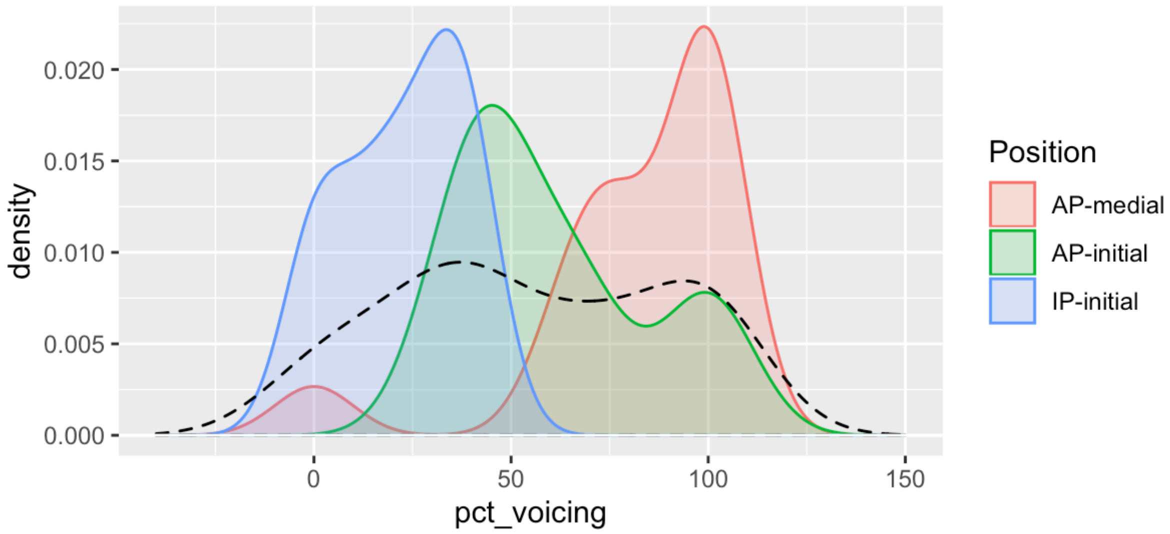


Figure F.8: Density plot of % voicing for lenis stops produced by Seoul_8.

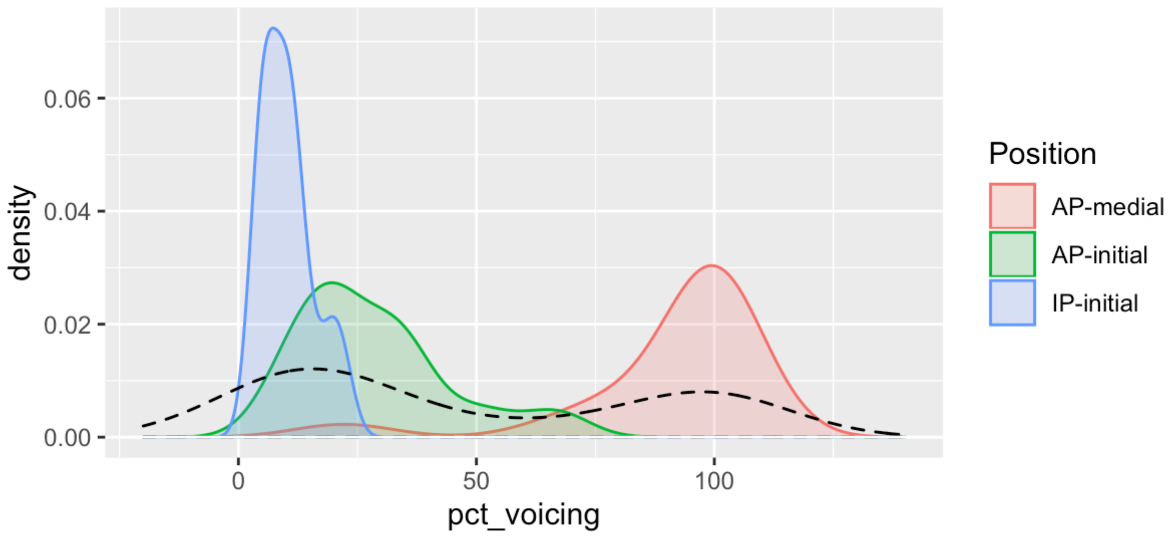


Figure F.9: Density plot of % voicing for lenis stops produced by Seoul_9.

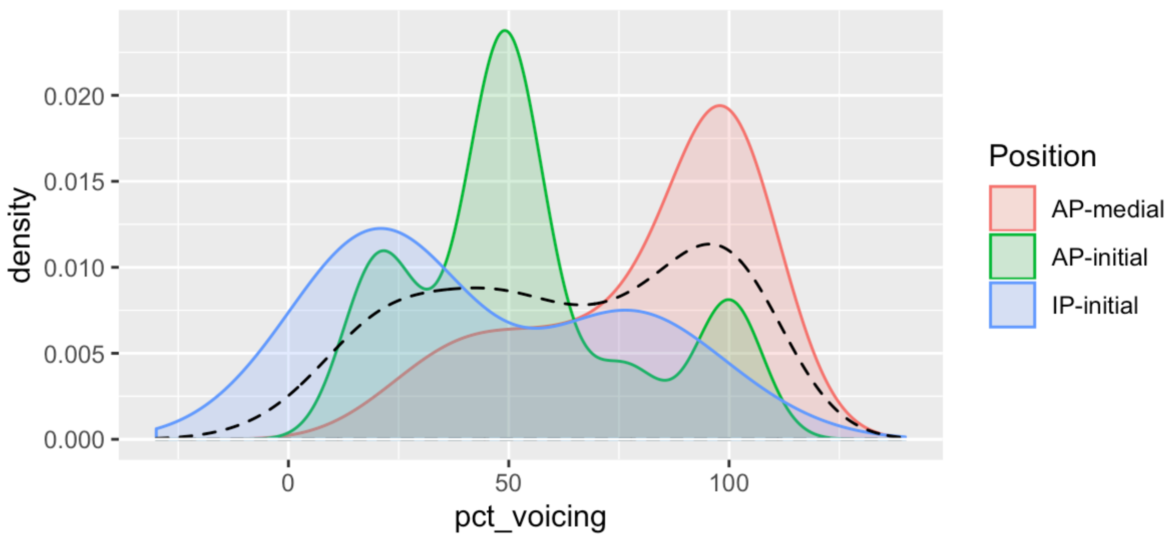


Figure F.10: Density plot of % voicing for lenis stops produced by Seoul_10.

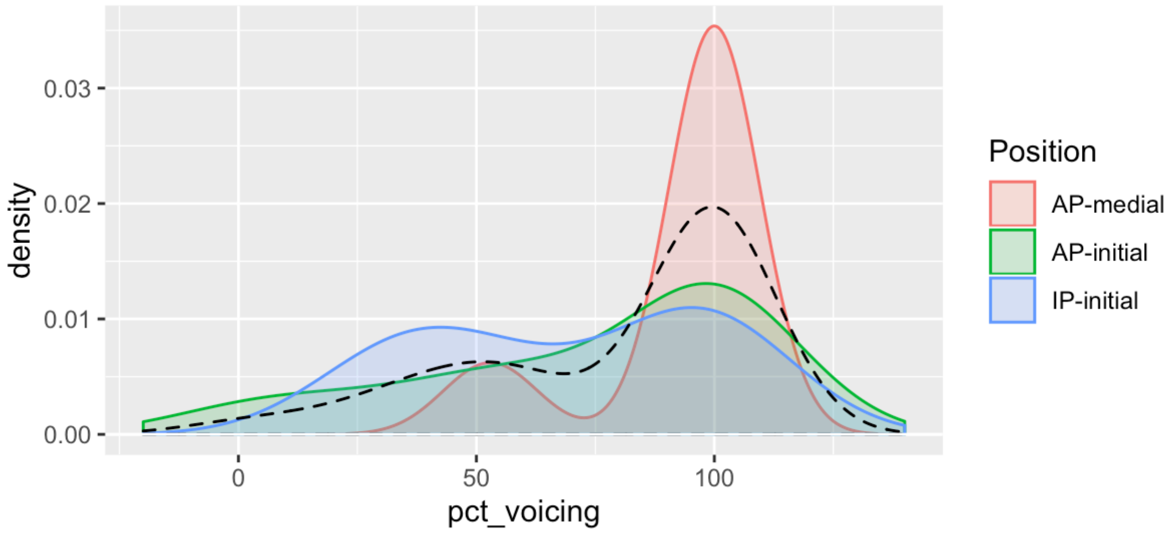


Figure F.11: Density plot of % voicing for lenis stops produced by Busan_1.

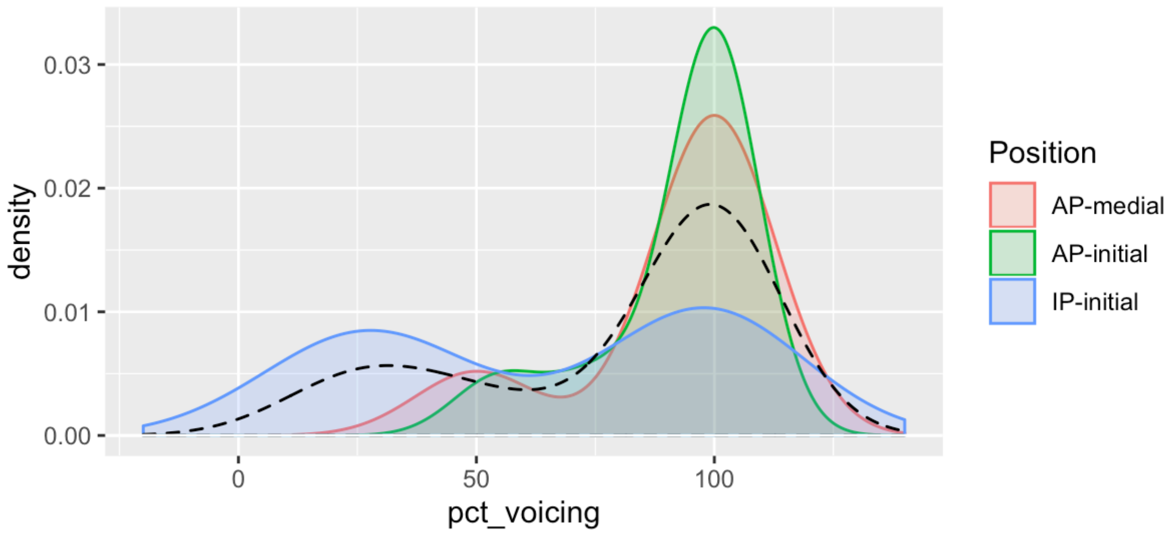


Figure F.12: Density plot of % voicing for lenis stops produced by Busan_2.

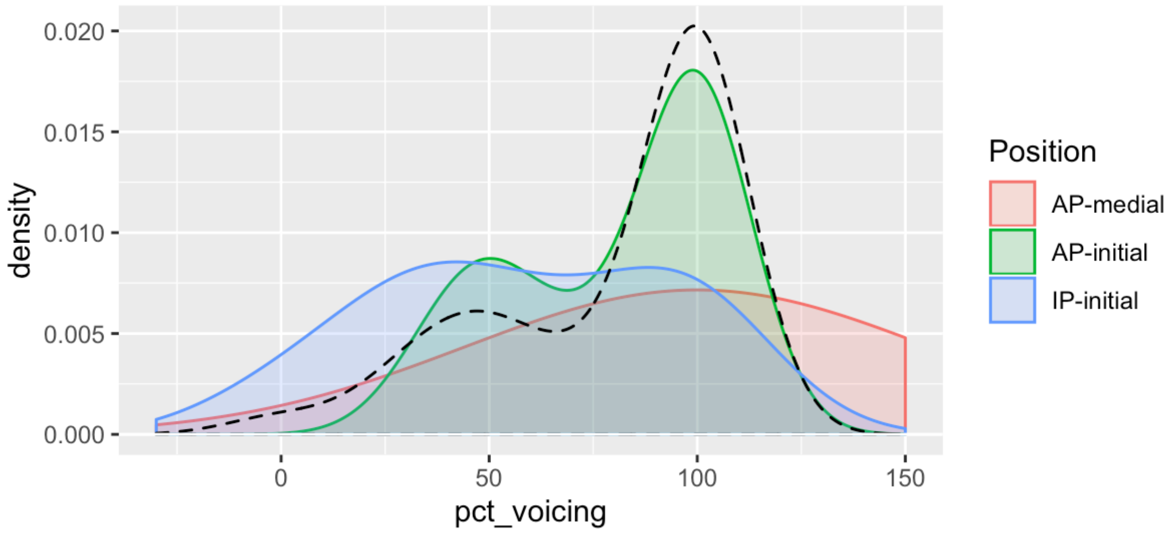


Figure F.13: Density plot of % voicing for lenis stops produced by Busan_3.

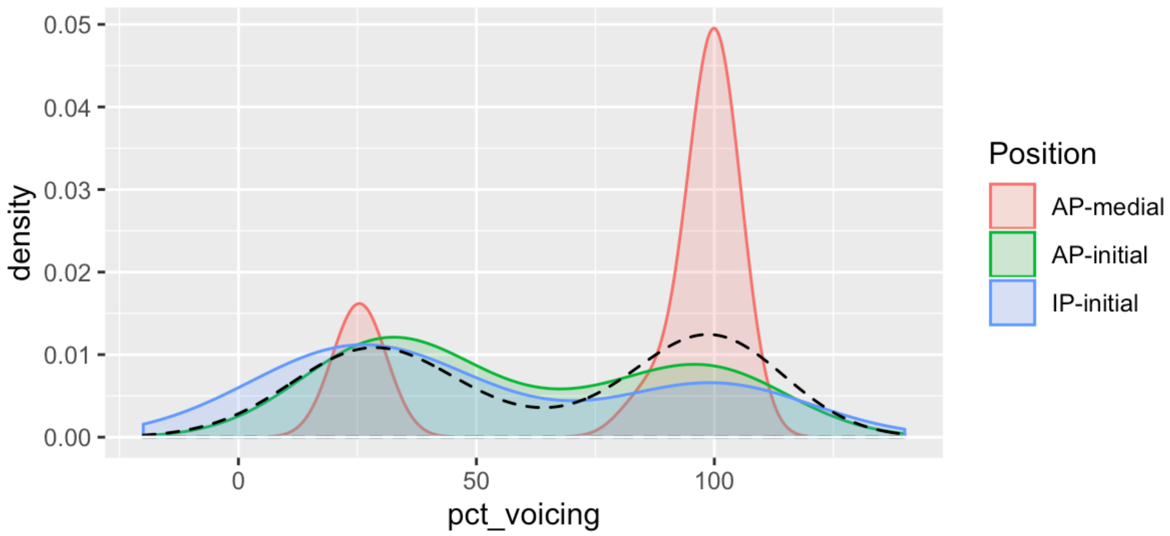


Figure F.14: Density plot of % voicing for lenis stops produced by Busan_4.

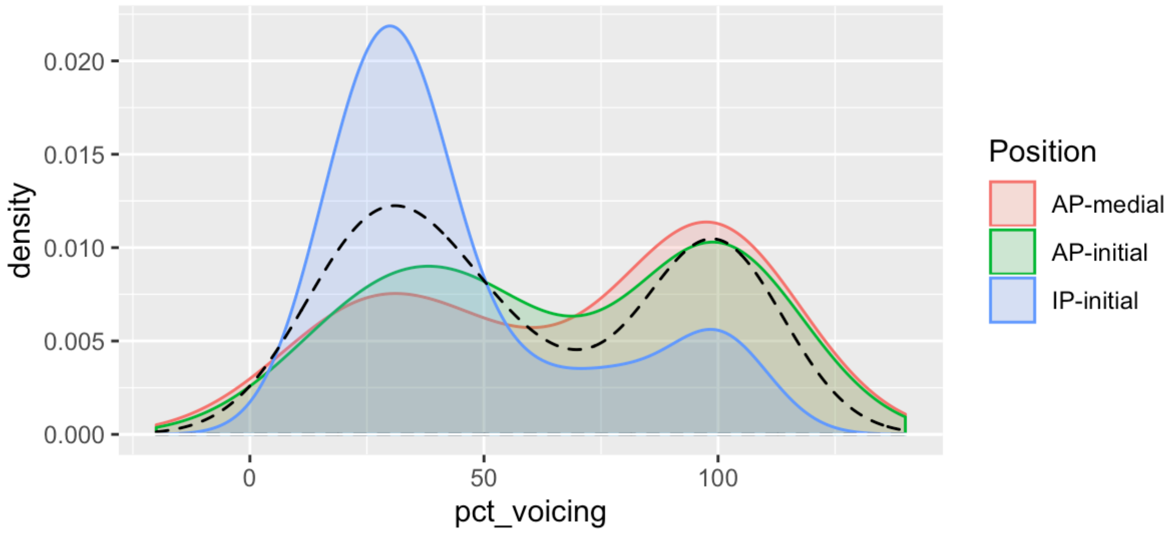


Figure F.15: Density plot of % voicing for lenis stops produced by Busan_5.

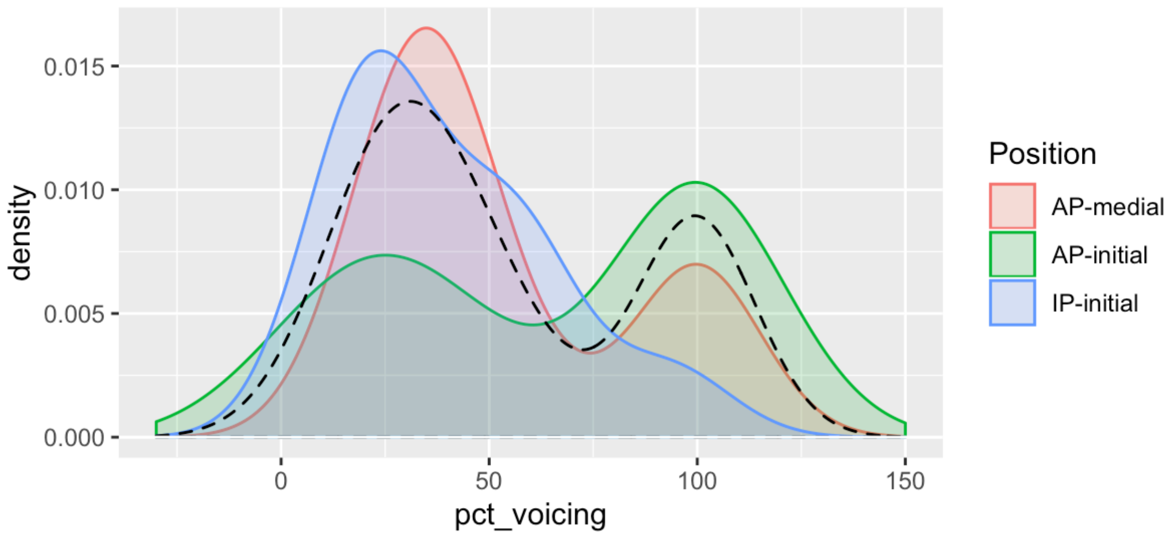


Figure F.16: Density plot of % voicing for lenis stops produced by Busan_6.

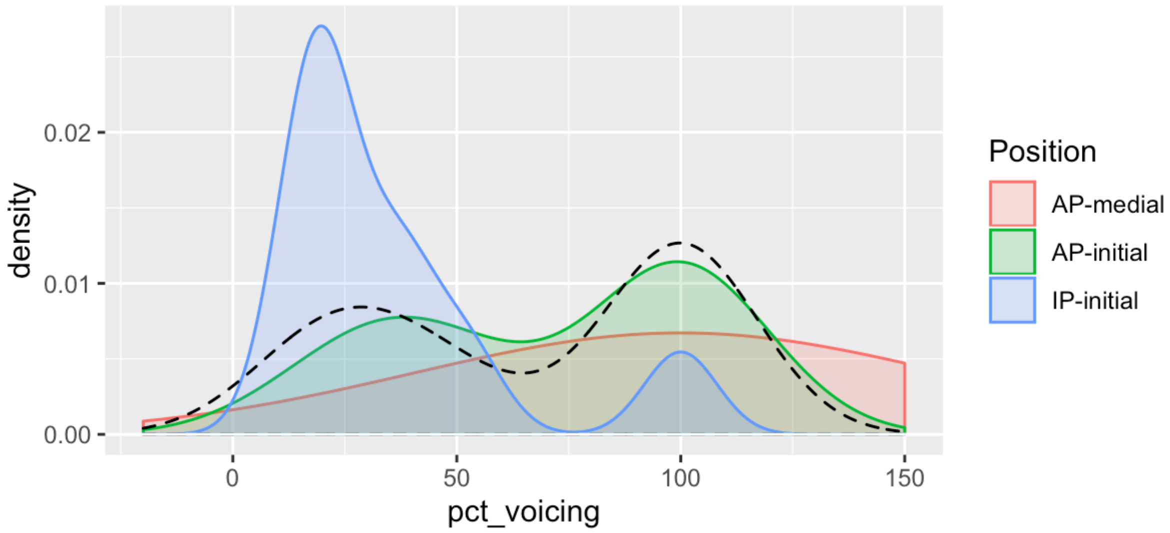


Figure F.17: Density plot of % voicing for lenis stops produced by Busan_7.

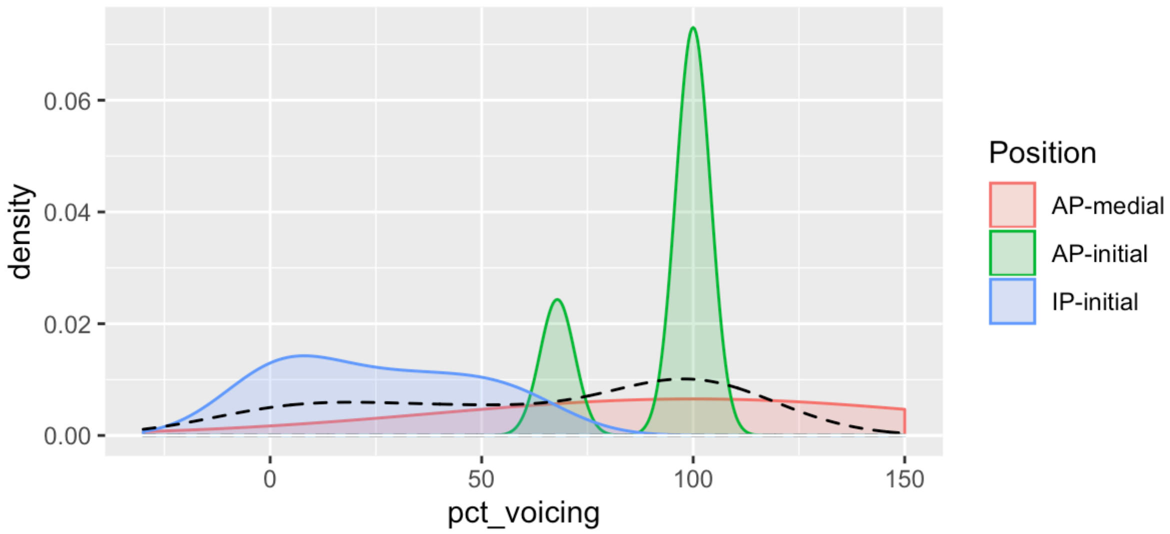


Figure F.18: Density plot of % voicing for lenis stops produced by Busan_8.

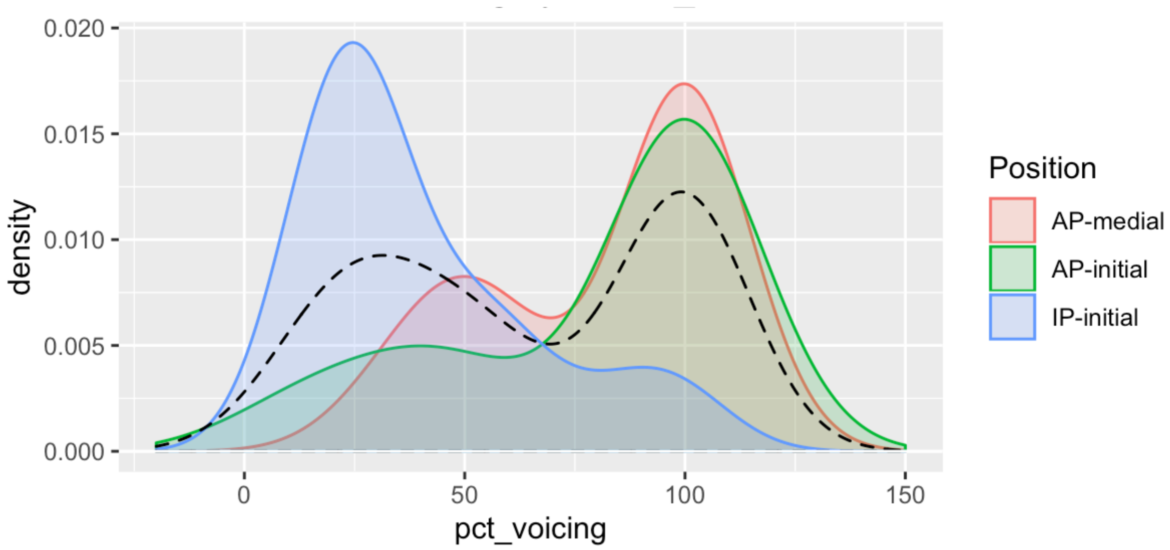


Figure F.19: Density plot of % voicing for lenis stops produced by Busan_9.

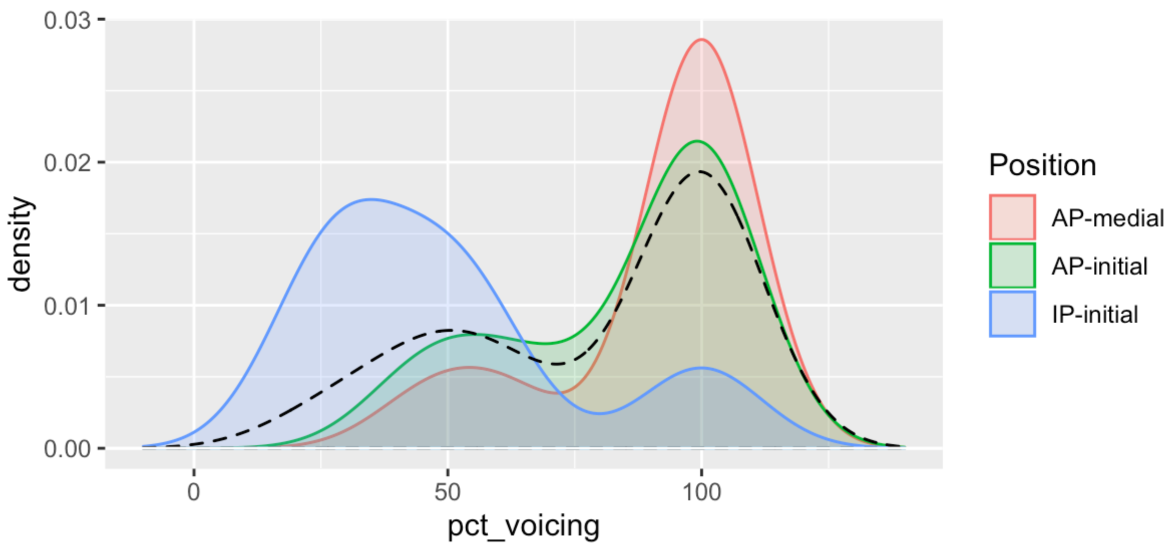


Figure F.20: Density plot of % voicing for lenis stops produced by Busan_10.

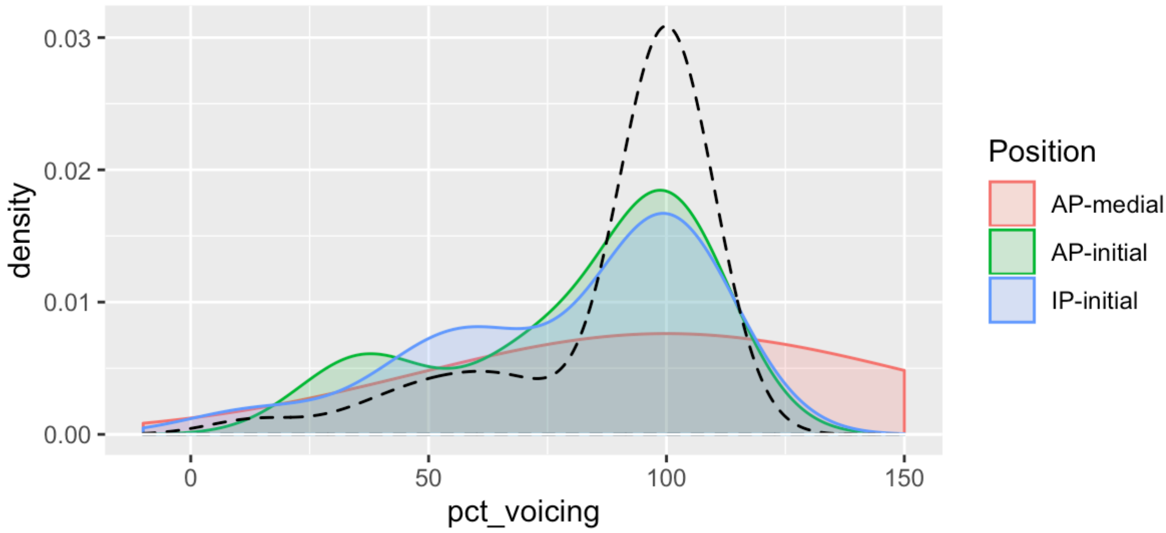


Figure F.21: Density plot of % voicing for lenis stops produced by Ulsan_1.

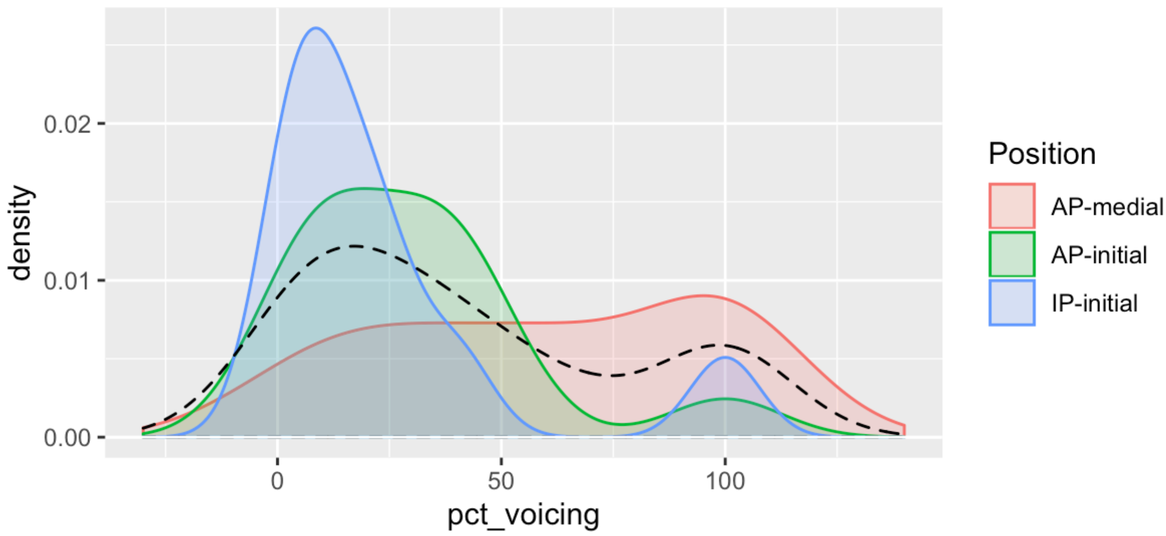


Figure F.22: Density plot of % voicing for lenis stops produced by Ulsan_2.

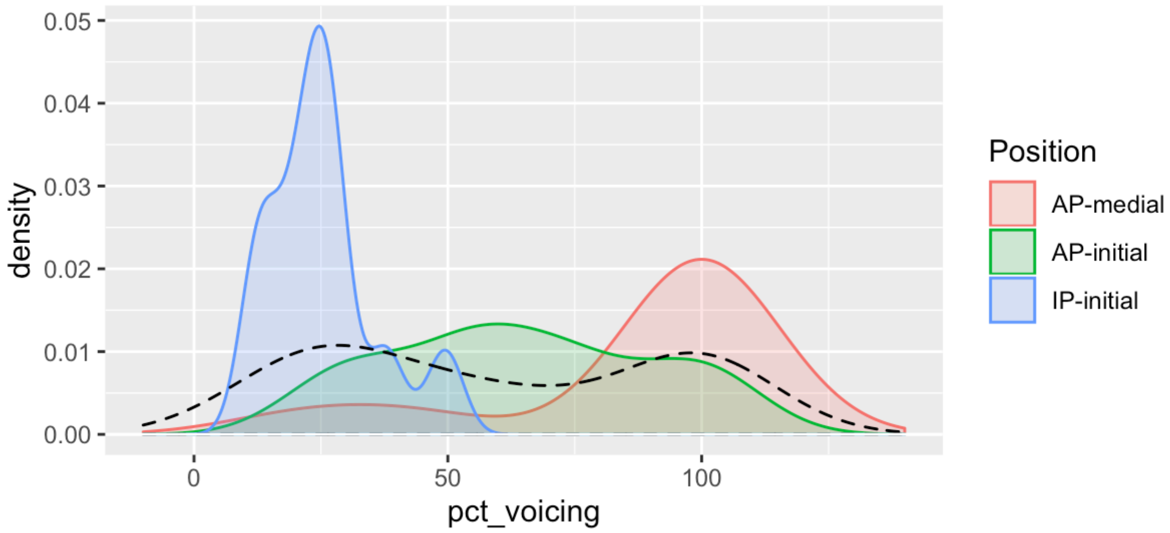


Figure F.23: Density plot of % voicing for lenis stops produced by Ulsan_3.

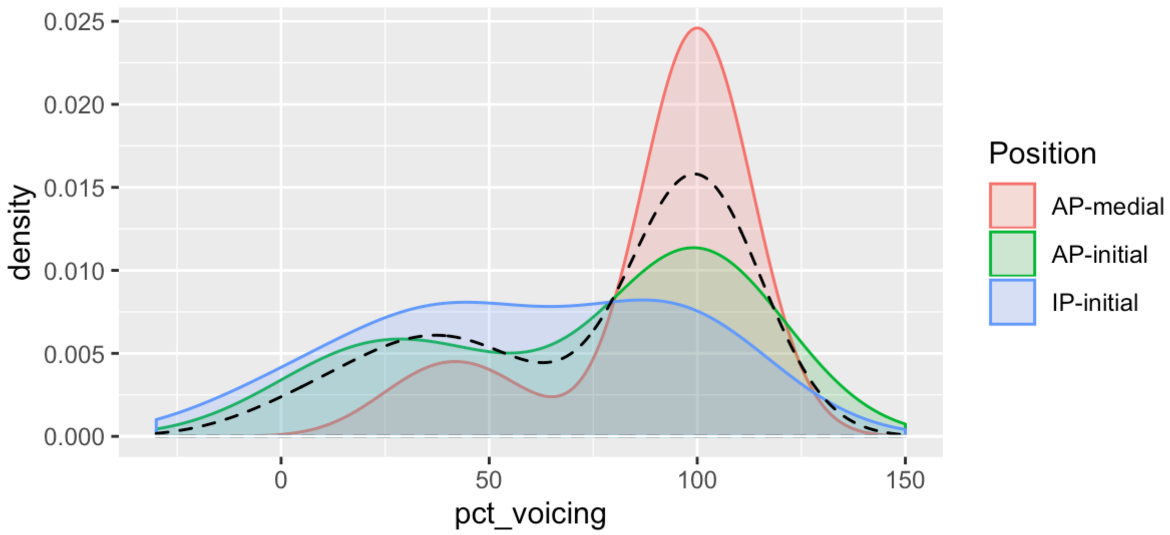


Figure F.24: Density plot of % voicing for lenis stops produced by Ulsan_4.

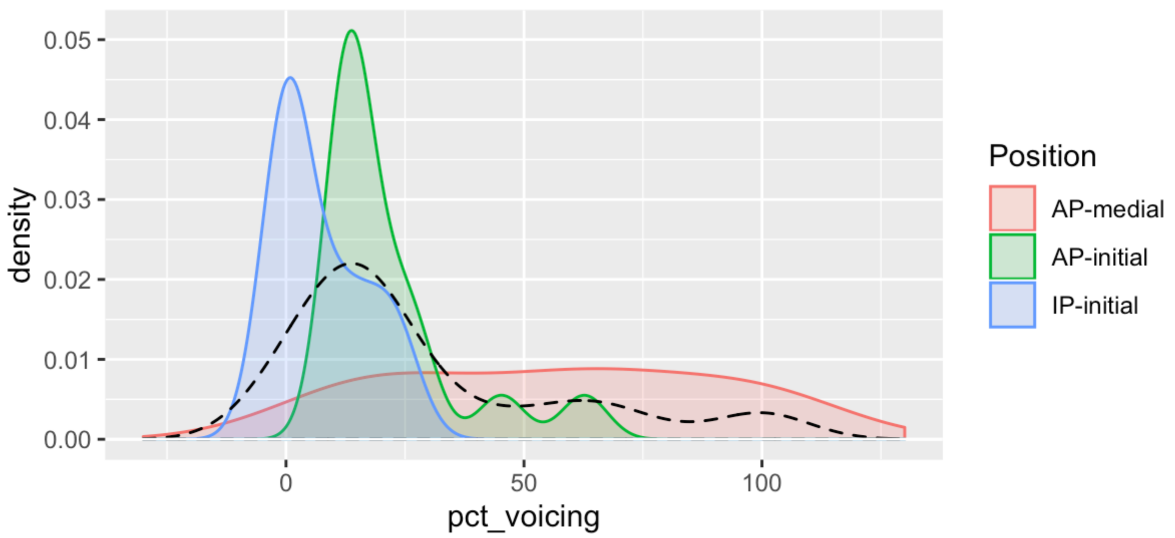


Figure F.25: Density plot of % voicing for lenis stops produced by Ulsan_5.

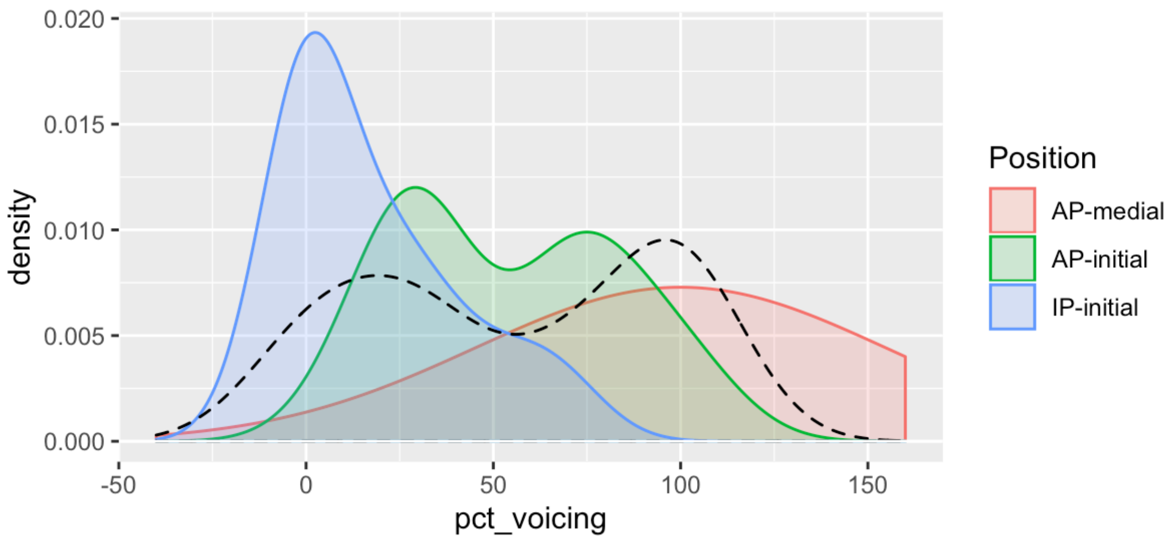


Figure F.26: Density plot of % voicing for lenis stops produced by Ulsan_6.

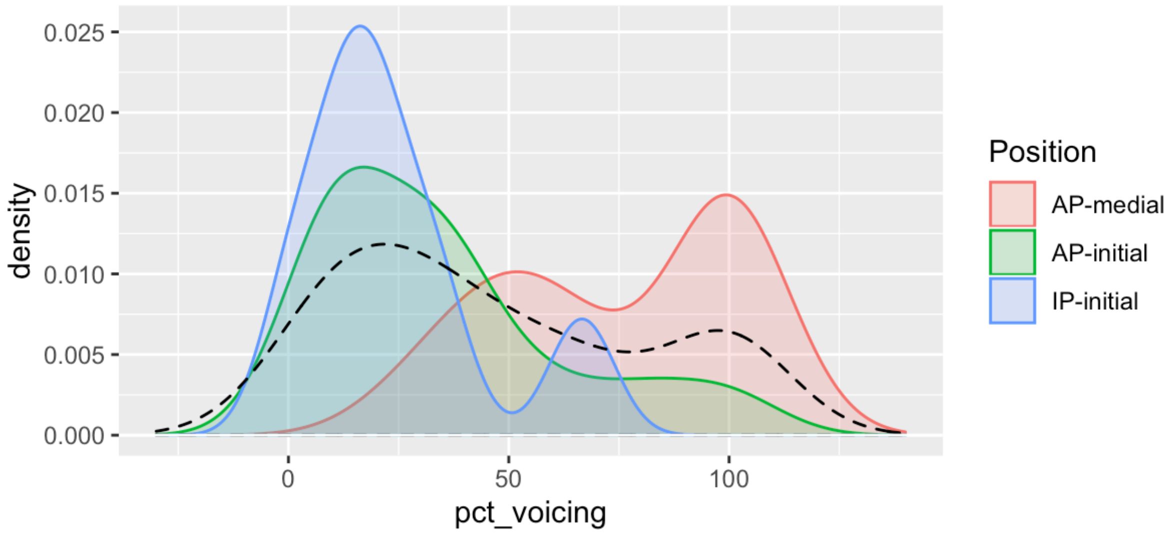


Figure F.27: Density plot of % voicing for lenis stops produced by Ulsan_7.

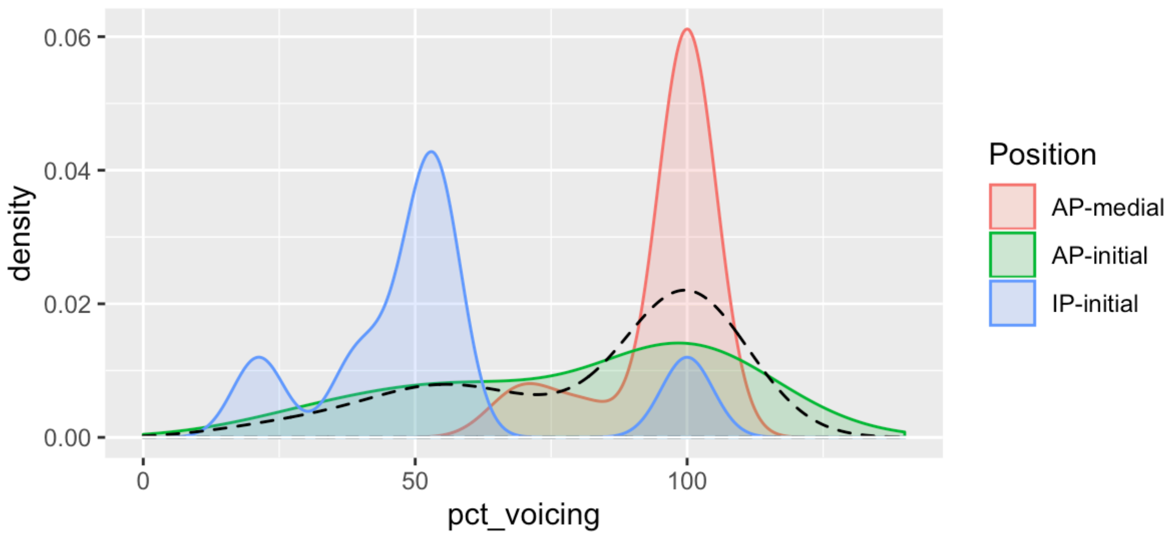


Figure F.28: Density plot of % voicing for lenis stops produced by Ulsan_8.

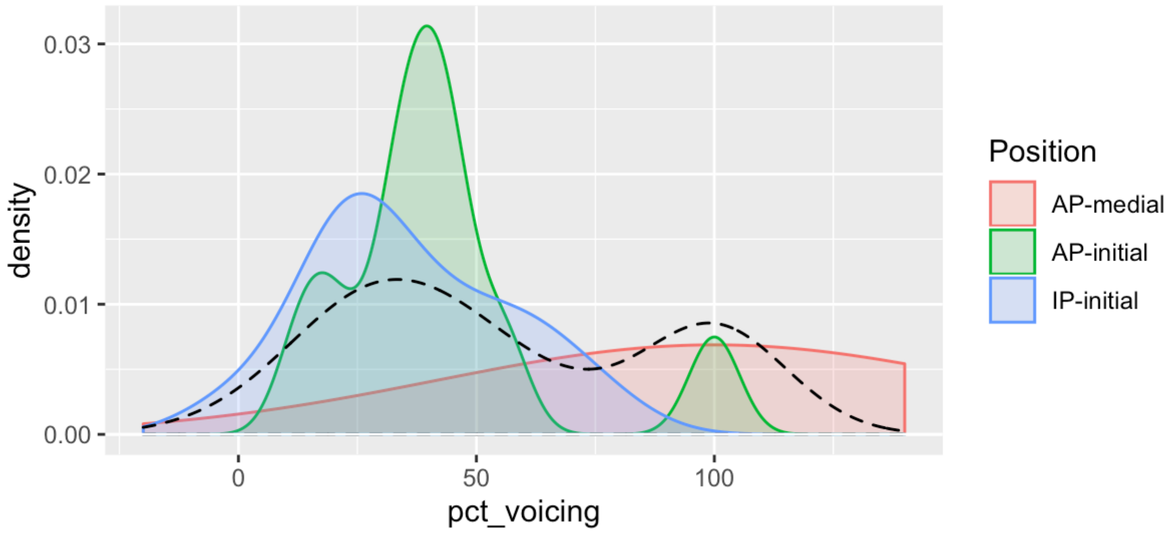


Figure F.29: Density plot of % voicing for lenis stops produced by Ulsan_9.

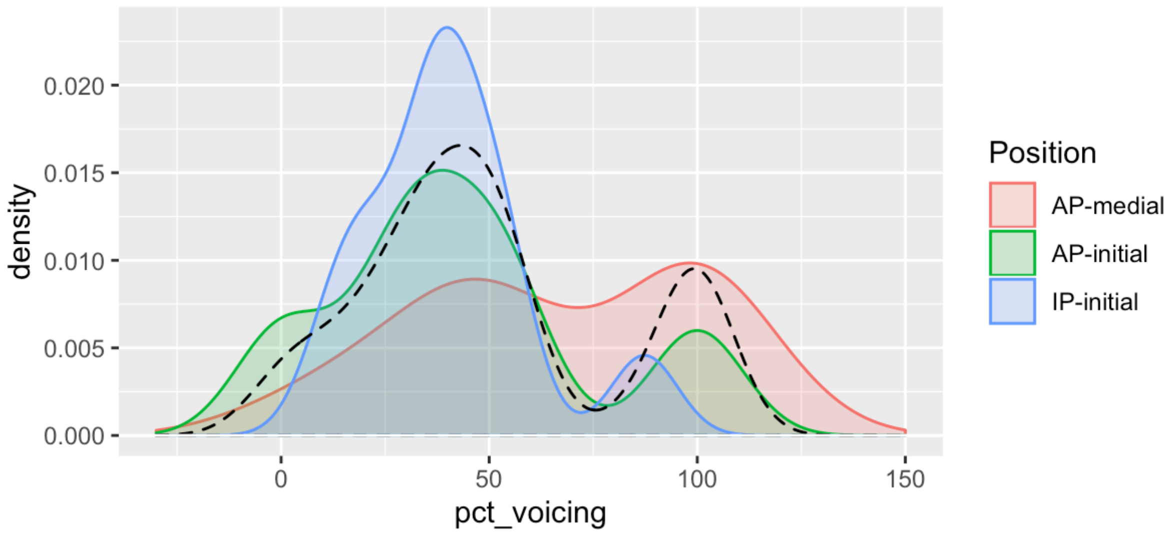


Figure F.30: Density plot of % voicing for lenis stops produced by Ulsan_10.

II. Minimum intensity for nasal stops

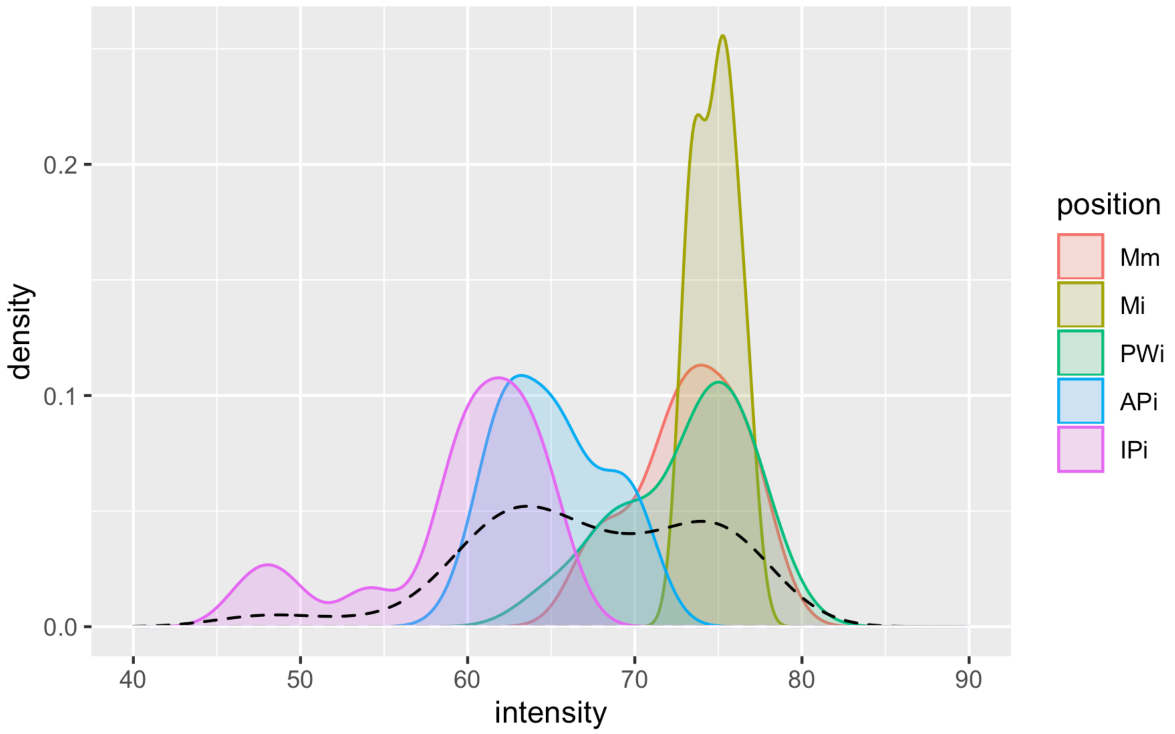


Figure F.31: Density plot of minimum intensity for nasal stops produced by Seoul_1.

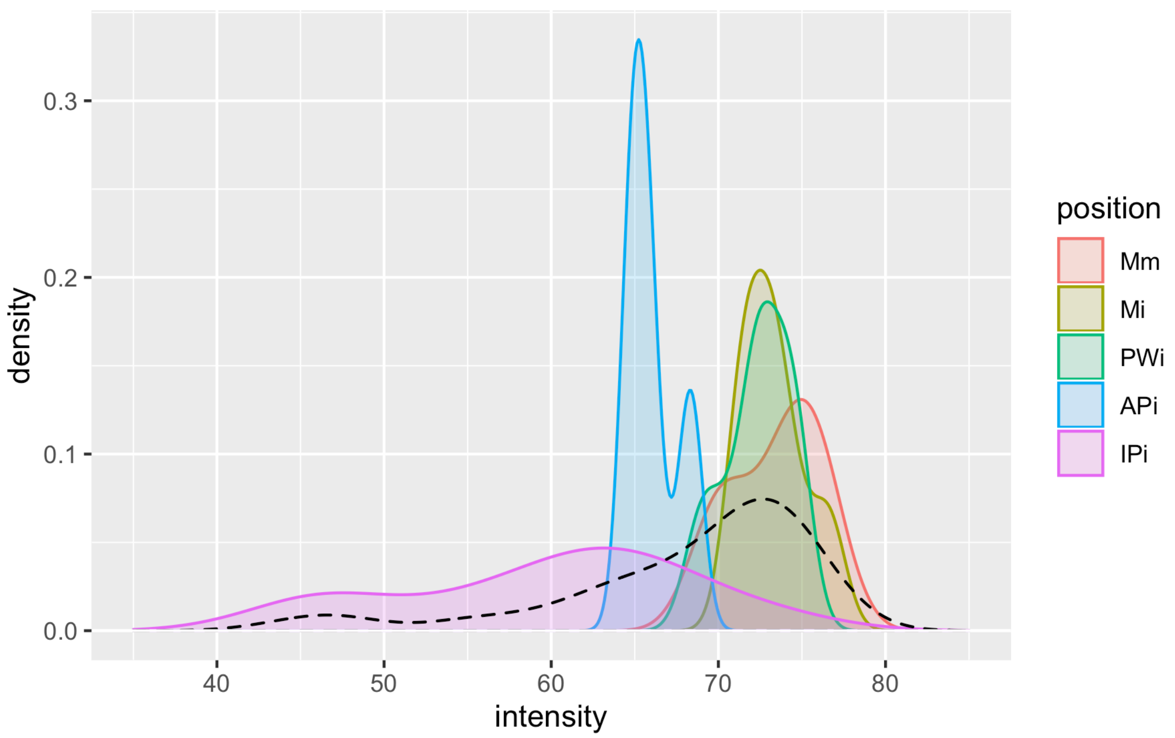


Figure F.32: Density plot of minimum intensity for nasal stops produced by Seoul_2.

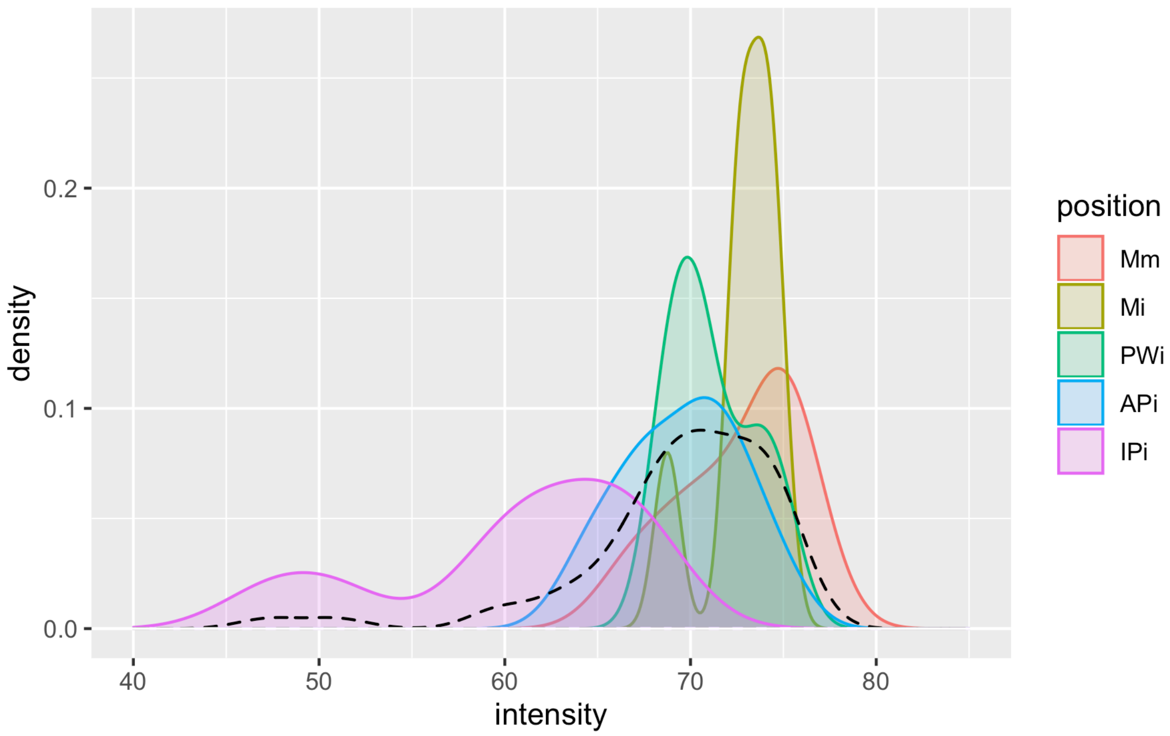


Figure F.33: Density plot of minimum intensity for nasal stops produced by Seoul_3.

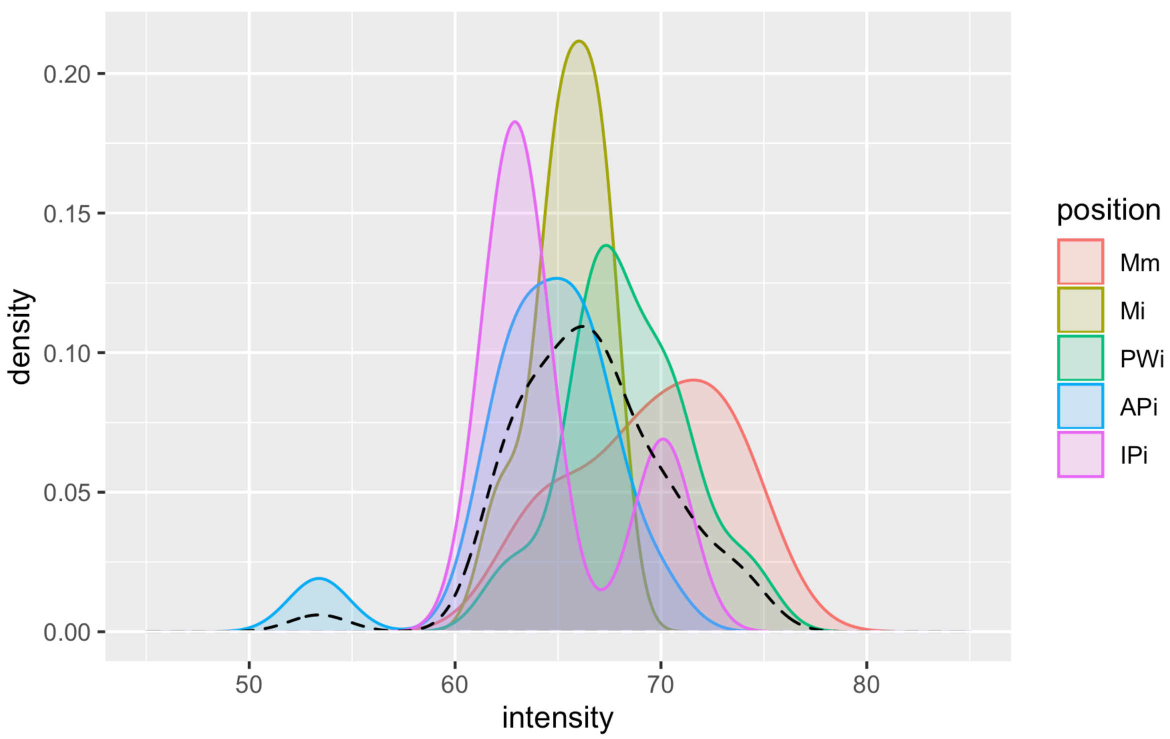


Figure F.34: Density plot of minimum intensity for nasal stops produced by Seoul_4.

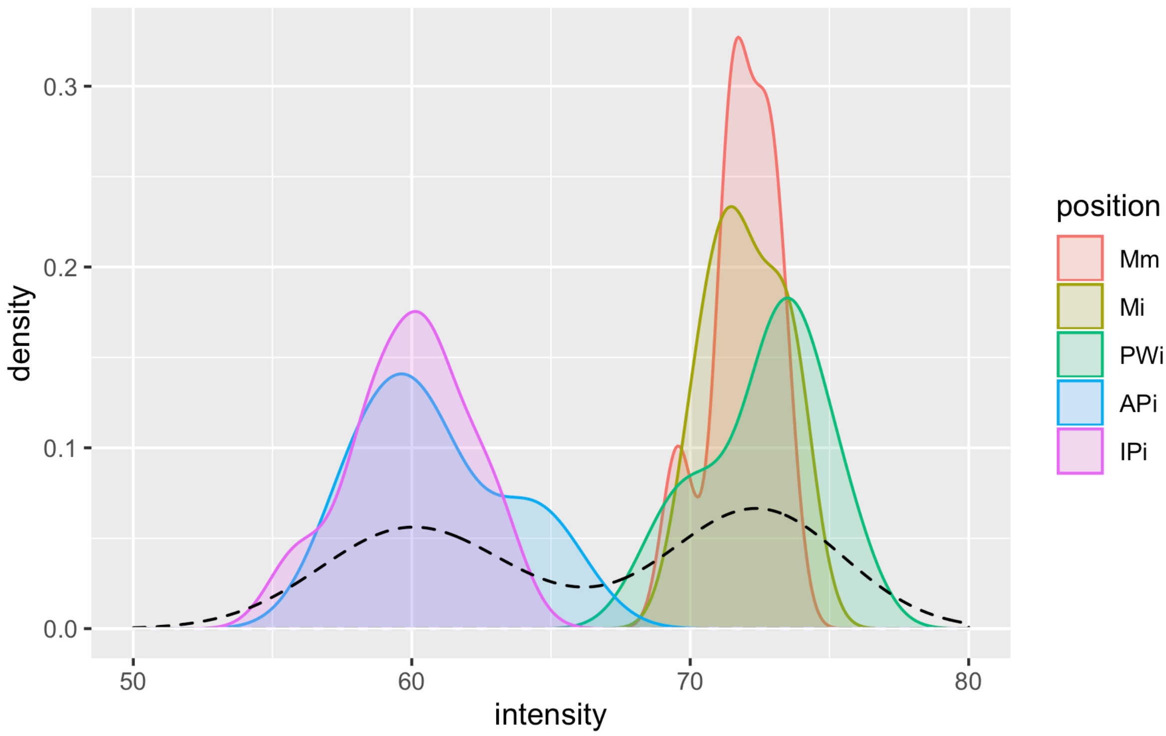


Figure F.35: Density plot of minimum intensity for nasal stops produced by Seoul_5.

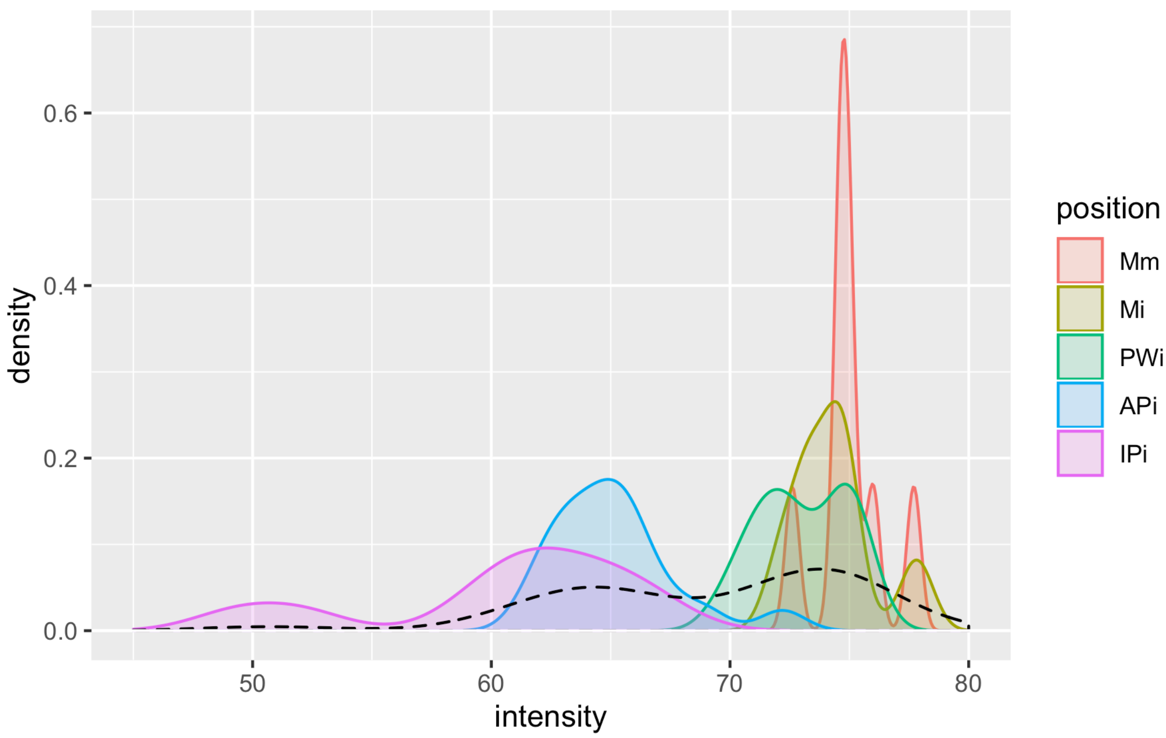


Figure F.36: Density plot of minimum intensity for nasal stops produced by Seoul_6.

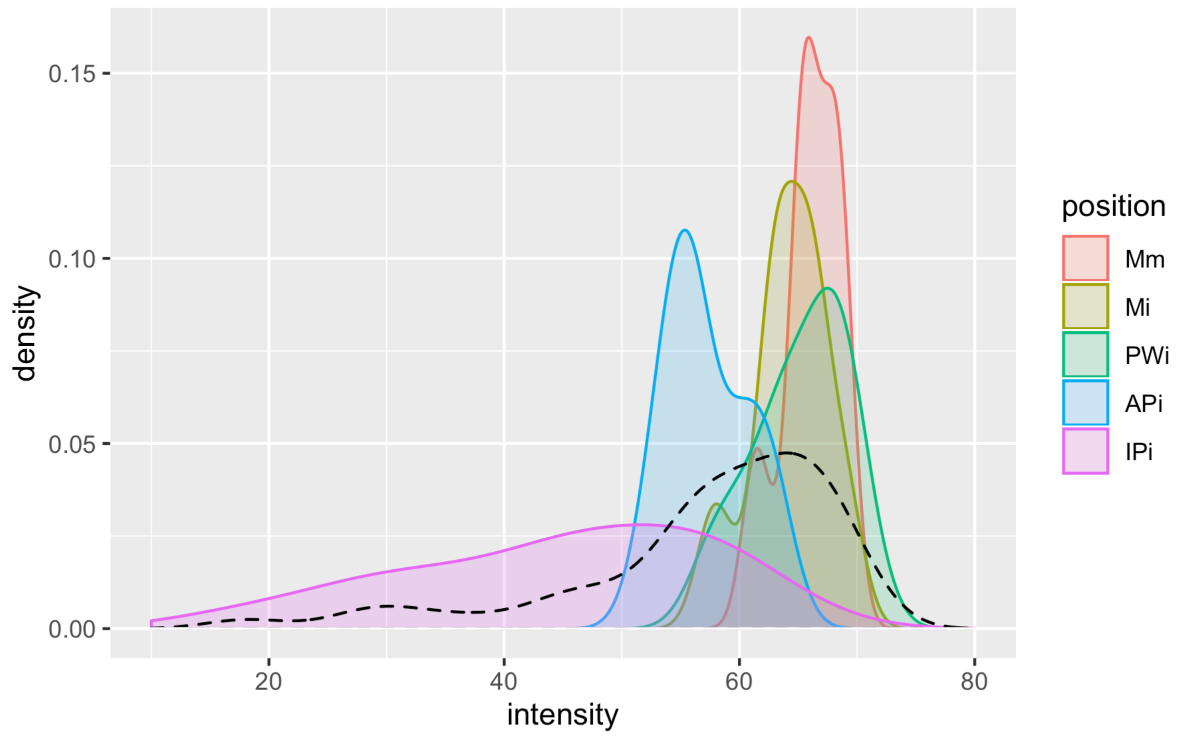


Figure F.37: Density plot of minimum intensity for nasal stops produced by Seoul_7.

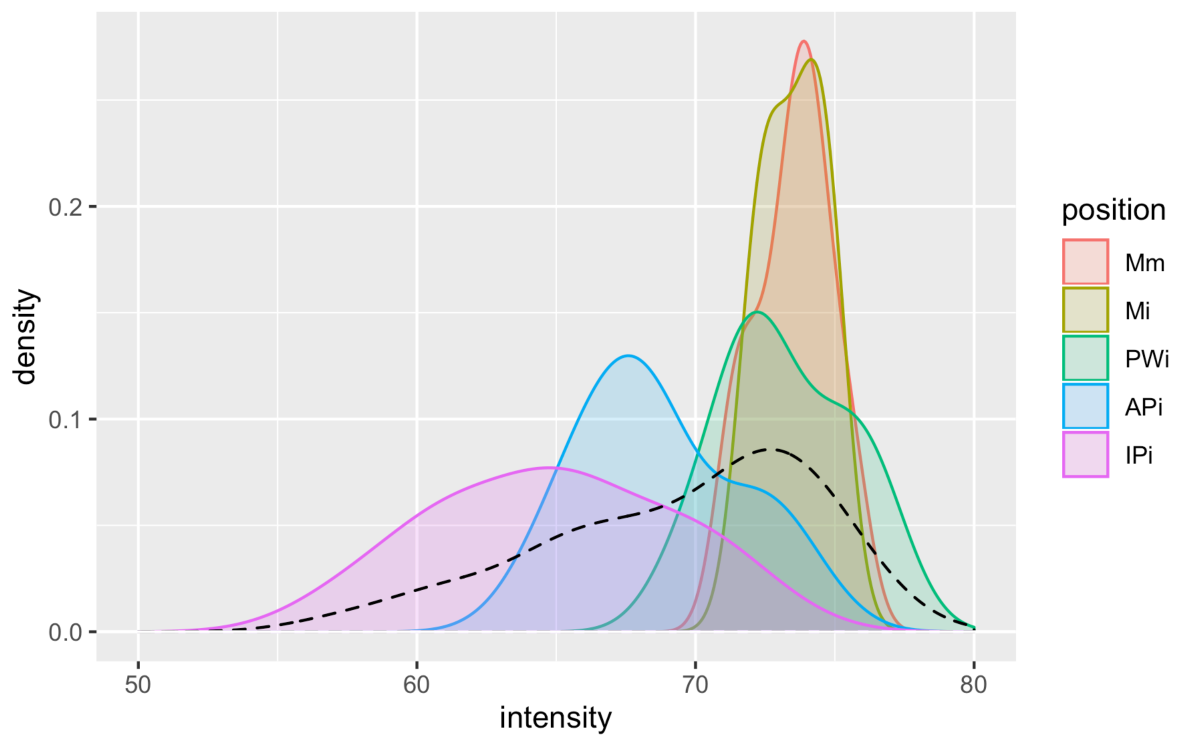


Figure F.38: Density plot of minimum intensity for nasal stops produced by Seoul_8.

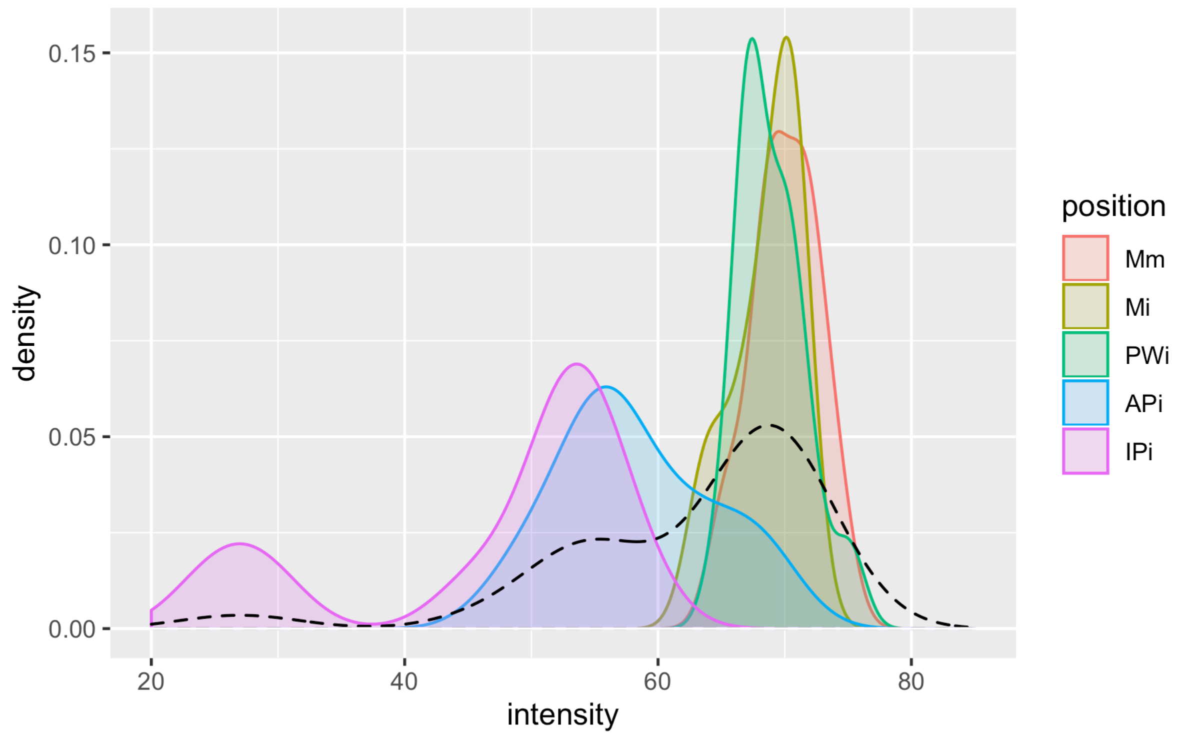


Figure F.39: Density plot of minimum intensity for nasal stops produced by Seoul_9.

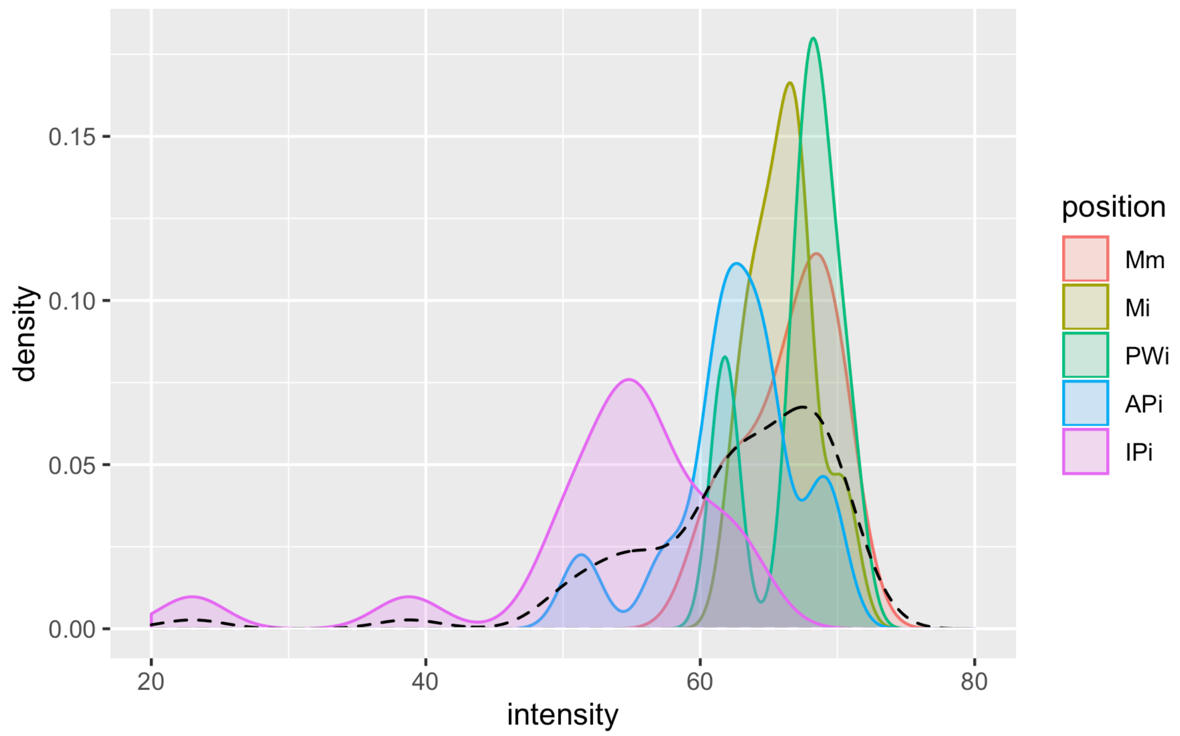


Figure F.40: Density plot of minimum intensity for nasal stops produced by Seoul_10.

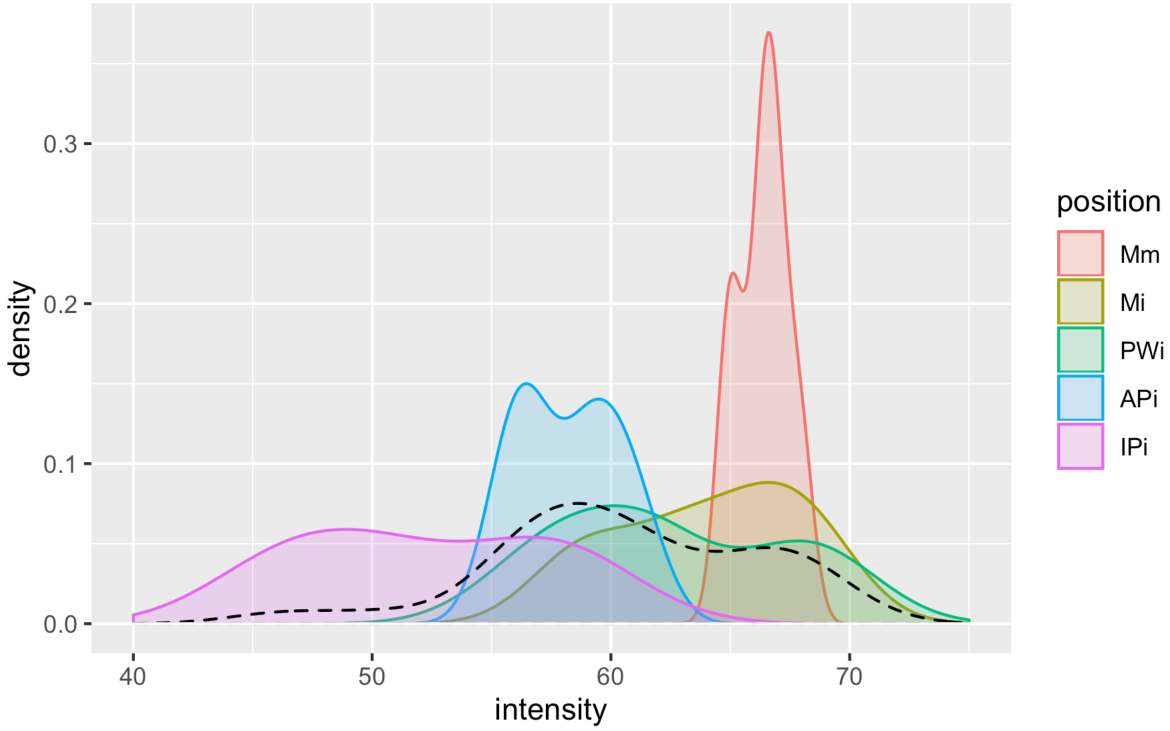


Figure F.41: Density plot of minimum intensity for nasal stops produced by Busan_1.

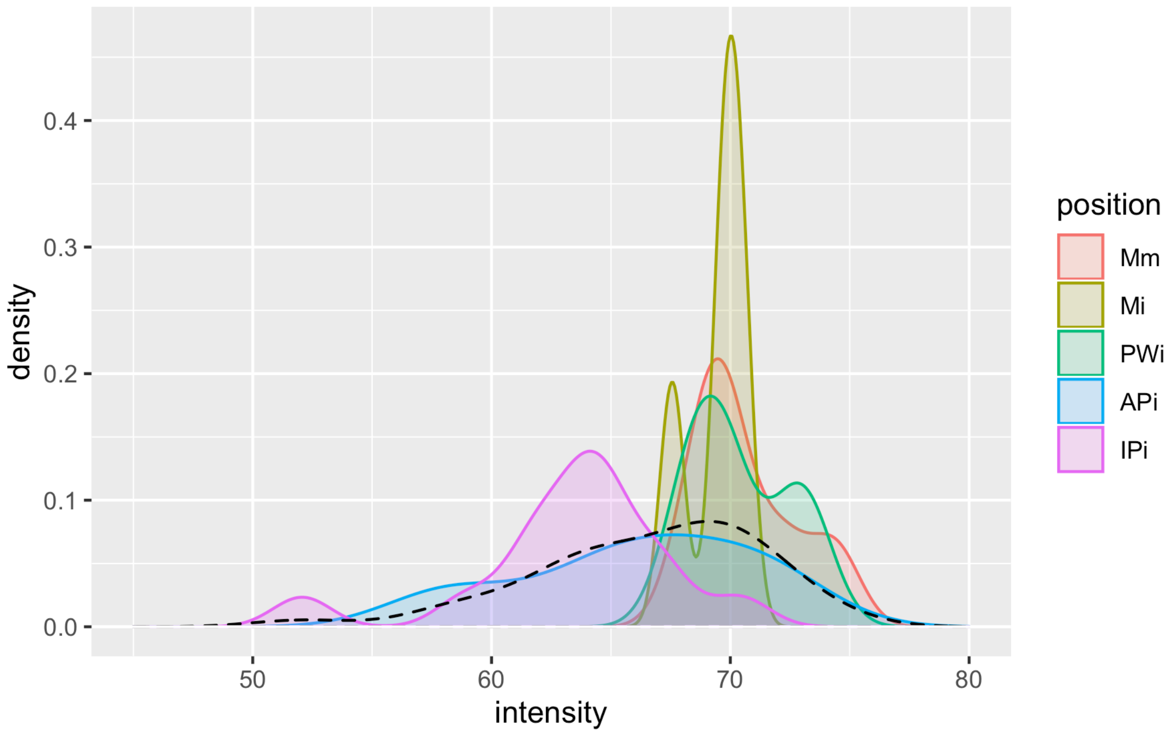


Figure F.42: Density plot of minimum intensity for nasal stops produced by Busan_2.

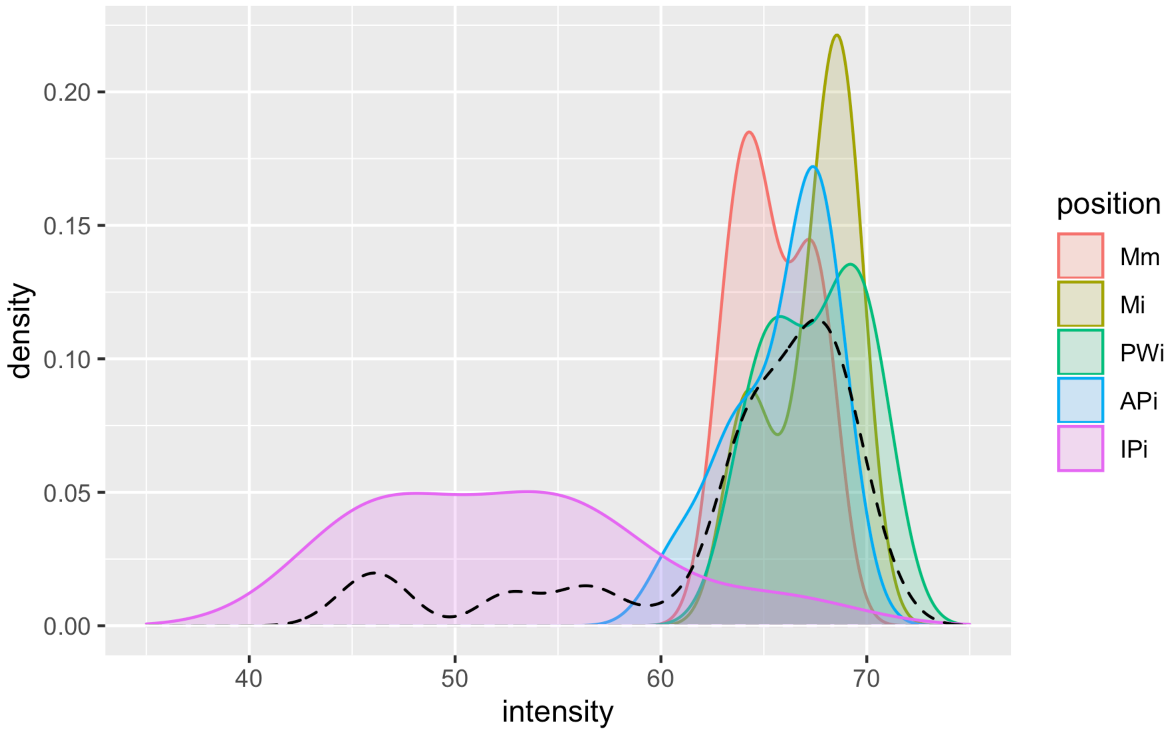


Figure F.43: Density plot of minimum intensity for nasal stops produced by Busan_3.

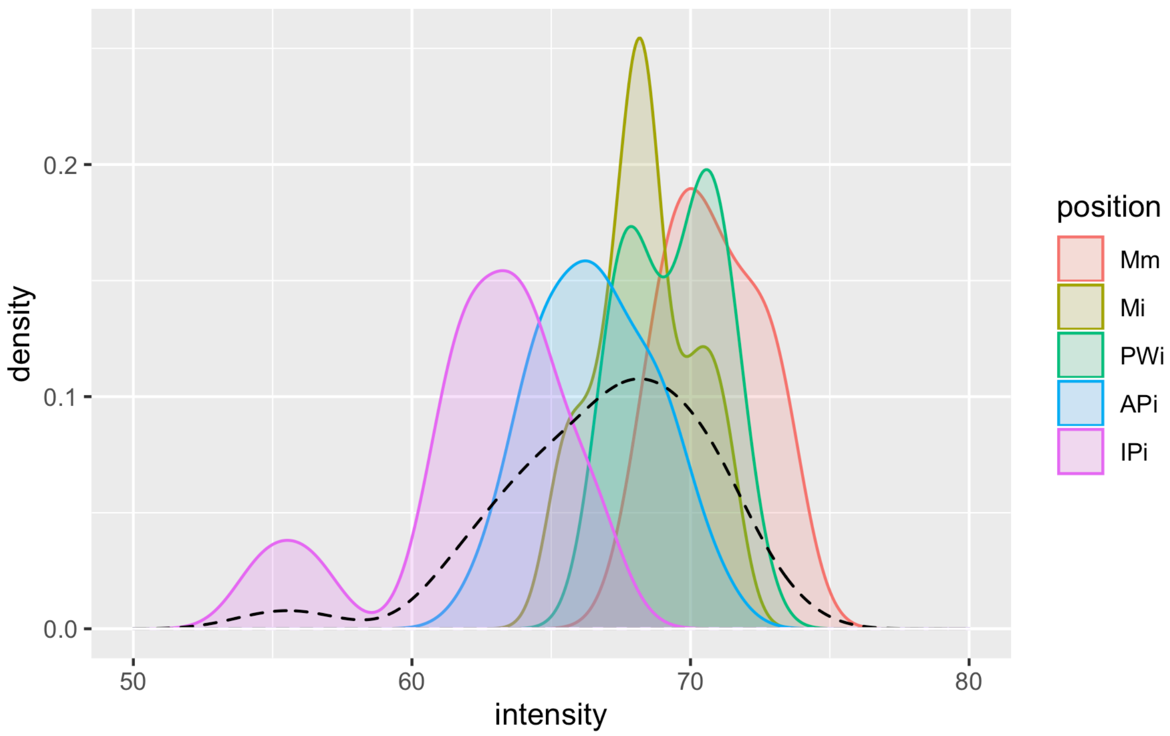


Figure F.44: Density plot of minimum intensity for nasal stops produced by Busan_4.

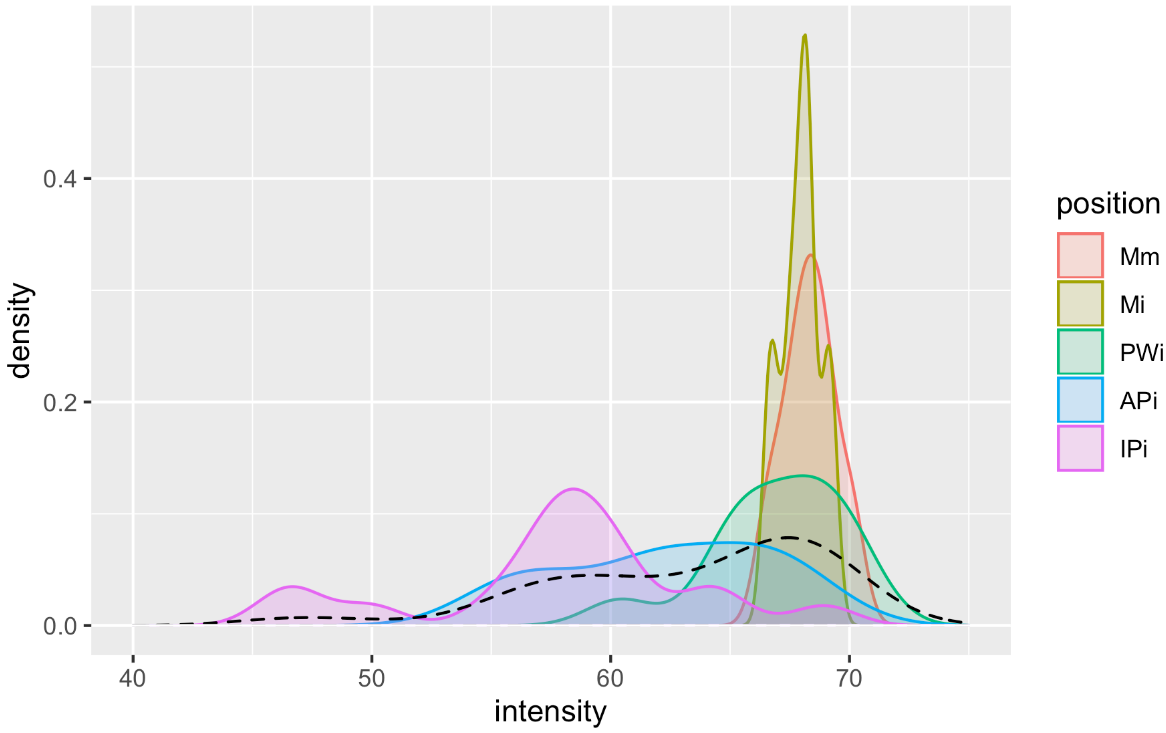


Figure F.45: Density plot of minimum intensity for nasal stops produced by Busan_5.

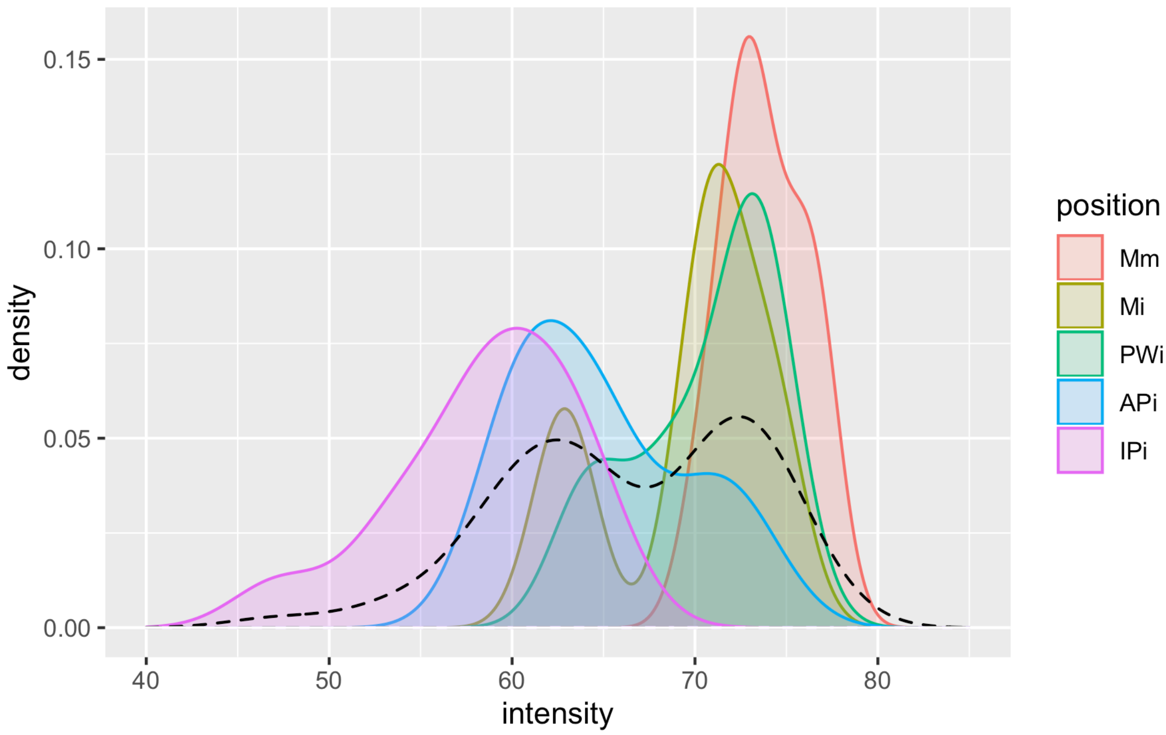


Figure F.46: Density plot of minimum intensity for nasal stops produced by Busan_6.

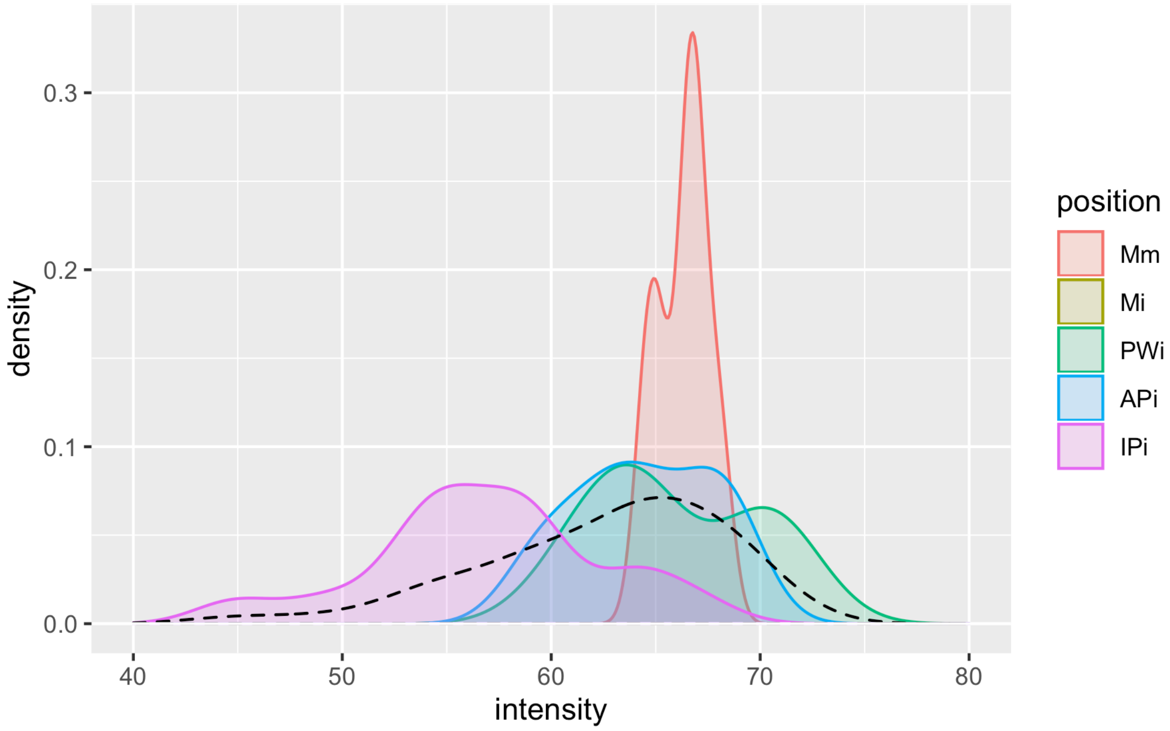


Figure F.47: Density plot of minimum intensity for nasal stops produced by Busan_7.

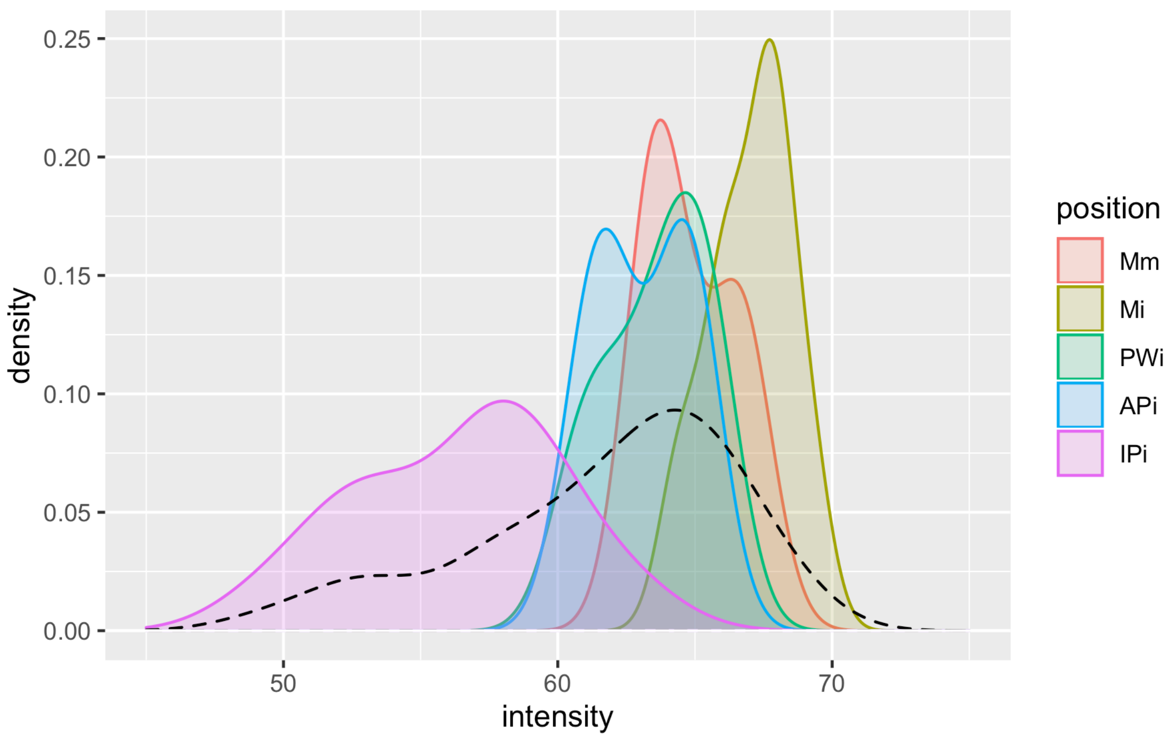


Figure F.48: Density plot of minimum intensity for nasal stops produced by Busan_8.

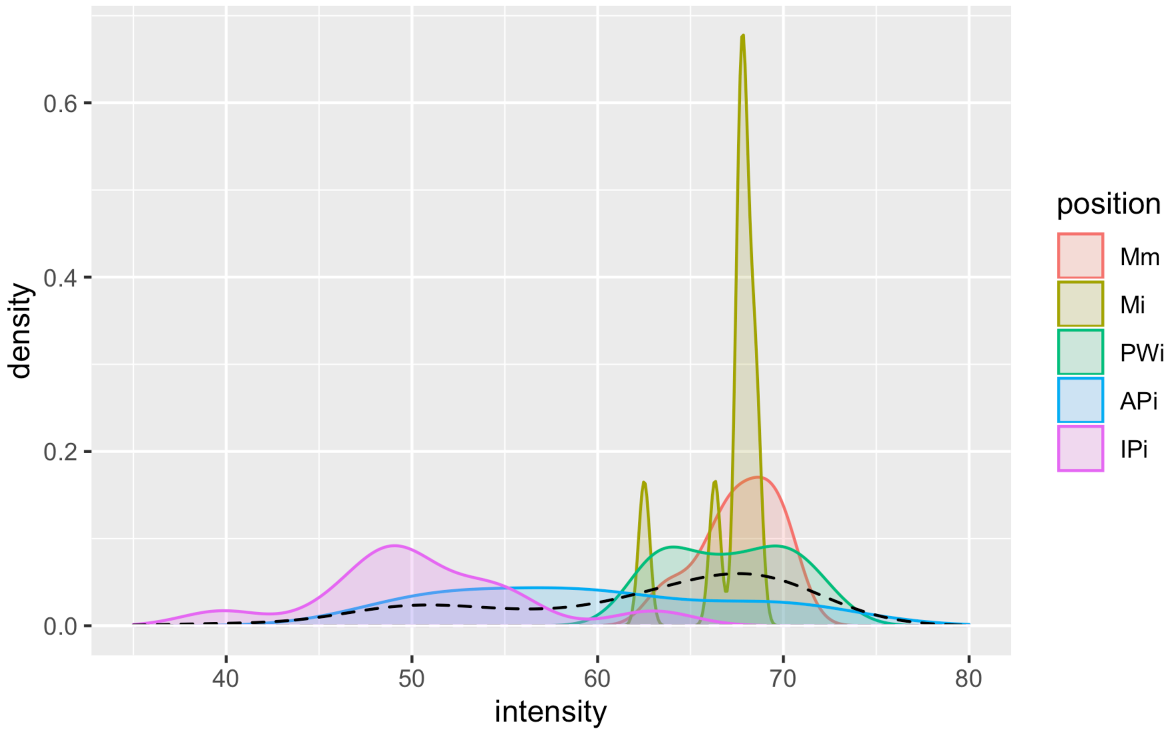


Figure F.49: Density plot of minimum intensity for nasal stops produced by Busan_9.

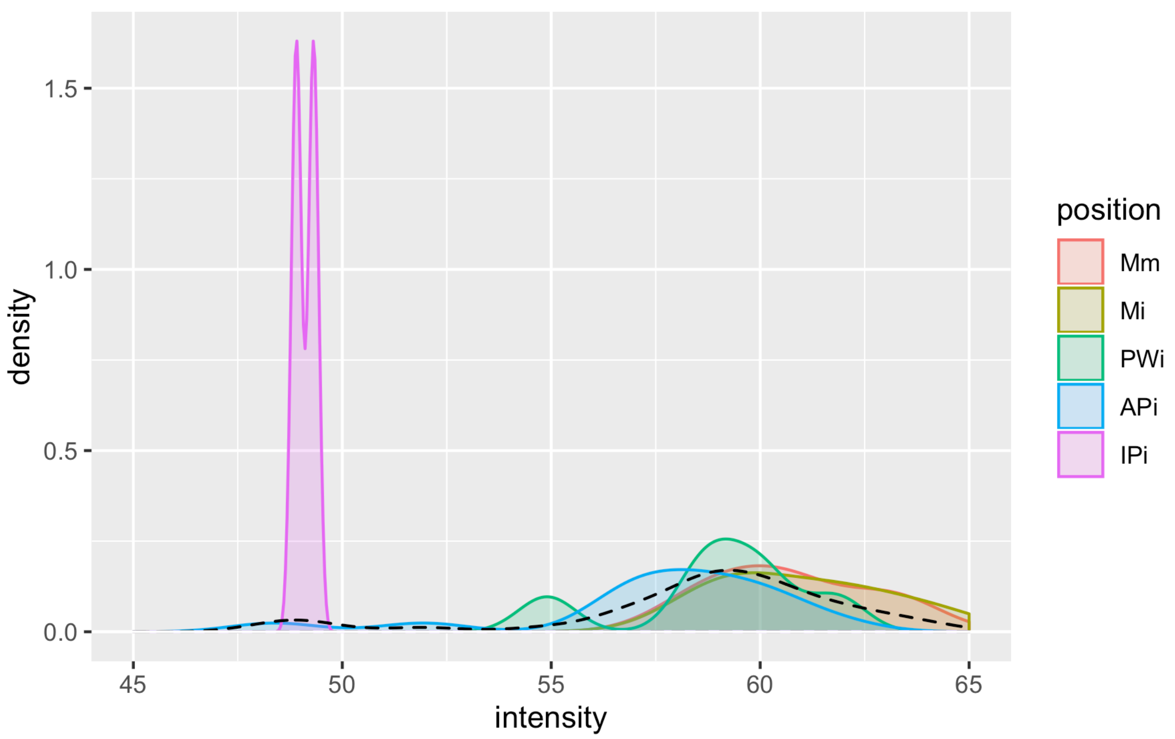


Figure F.50: Density plot of minimum intensity for nasal stops produced by Busan_10.

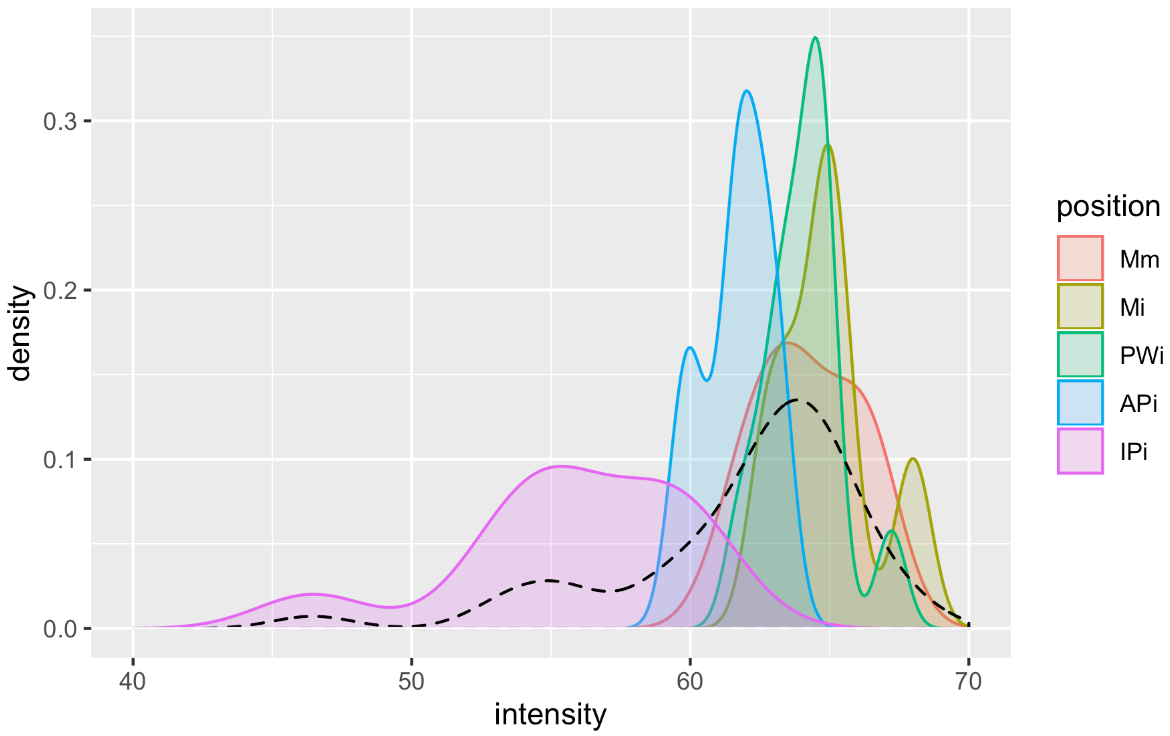


Figure F.51: Density plot of minimum intensity for nasal stops produced by Ulsan_1.

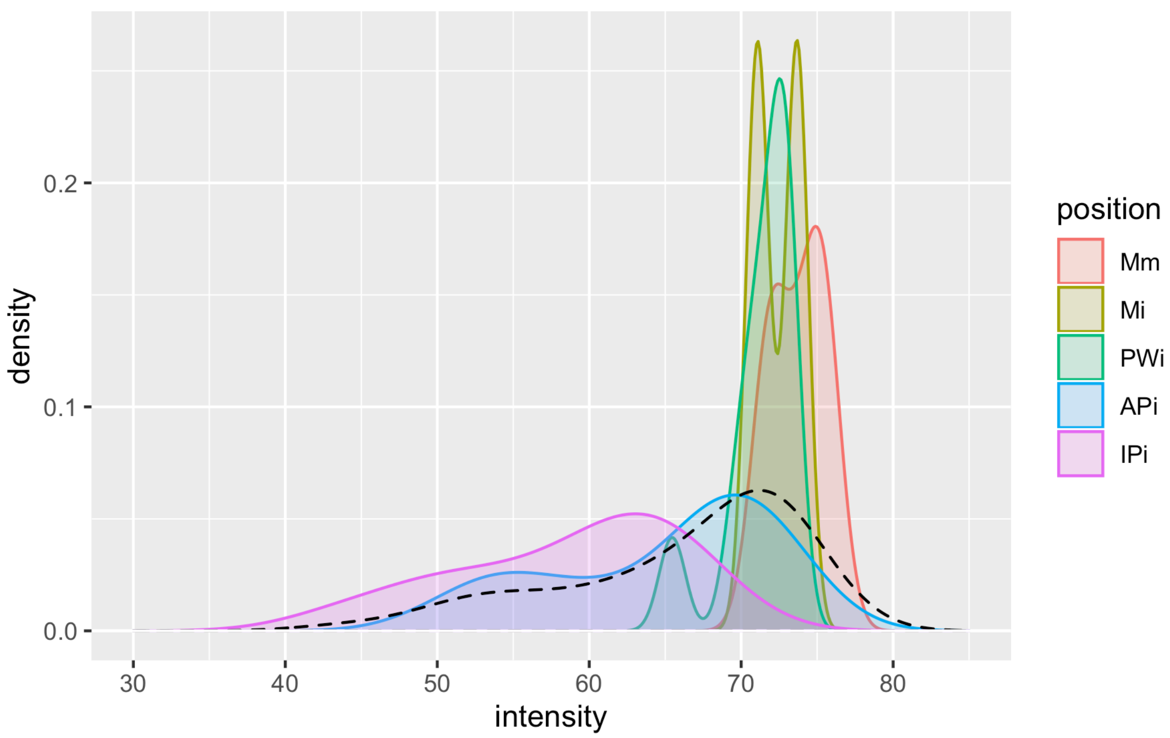


Figure F.52: Density plot of minimum intensity for nasal stops produced by Ulsan_2.

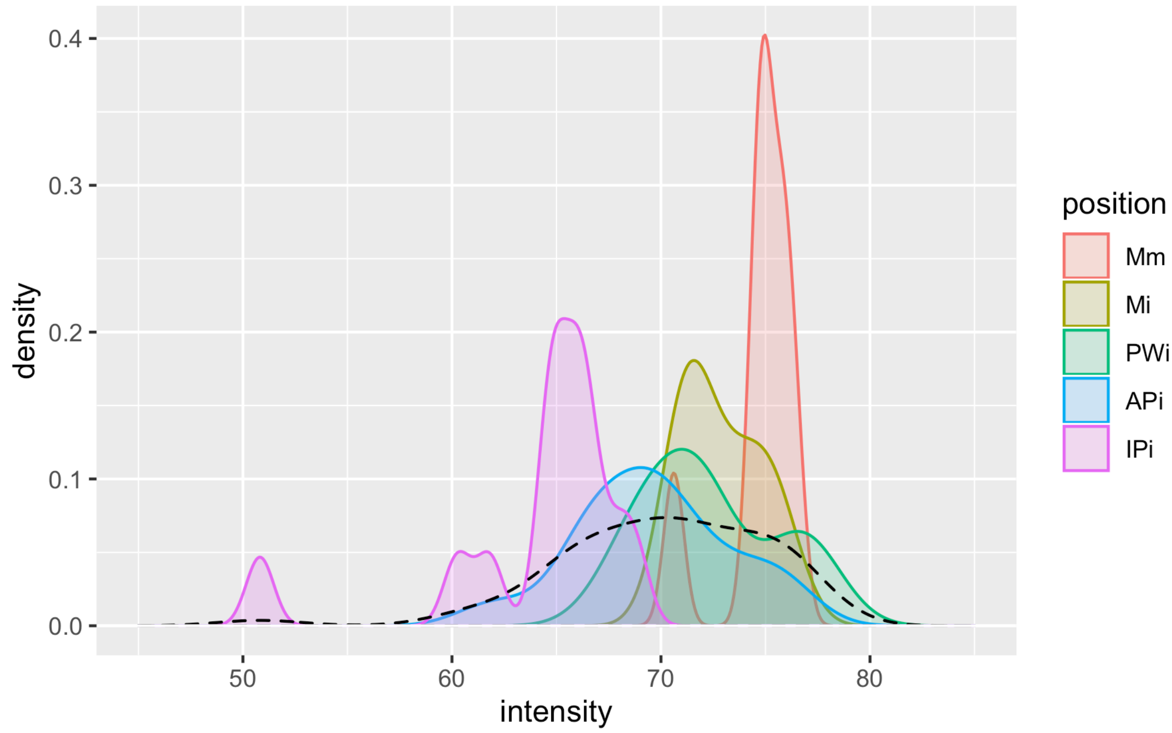


Figure F.53: Density plot of minimum intensity for nasal stops produced by Ulsan_3.

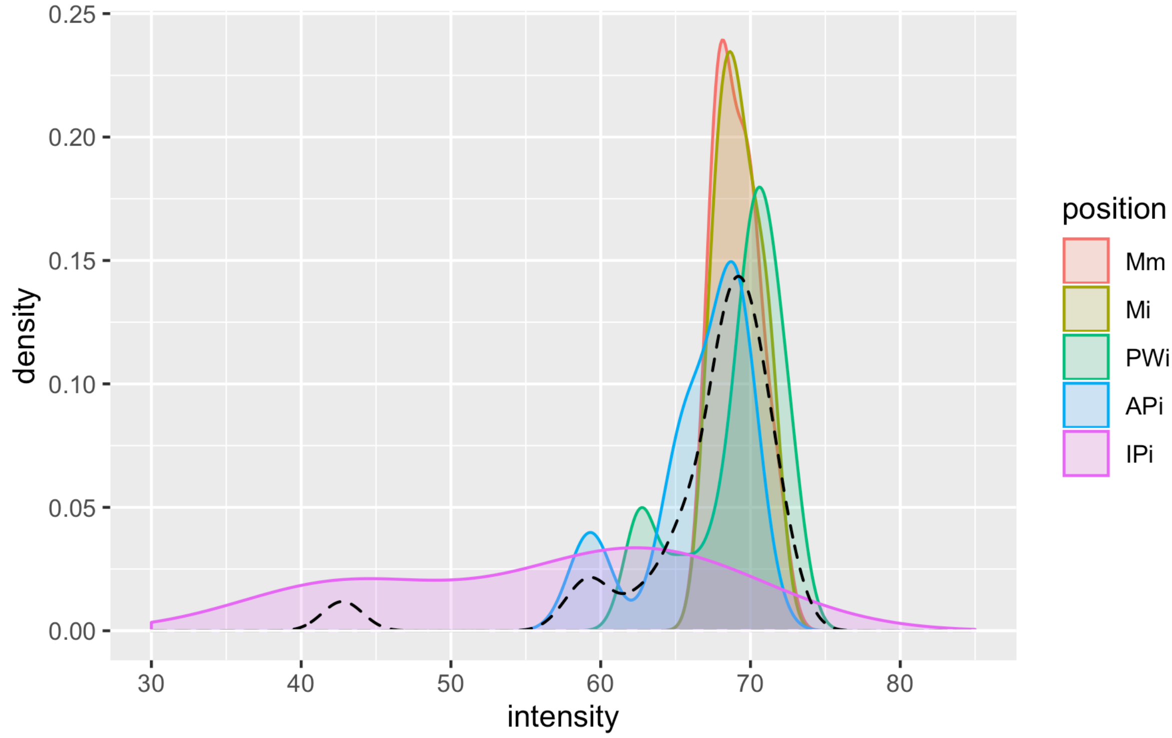


Figure F.54: Density plot of minimum intensity for nasal stops produced by Ulsan_4.

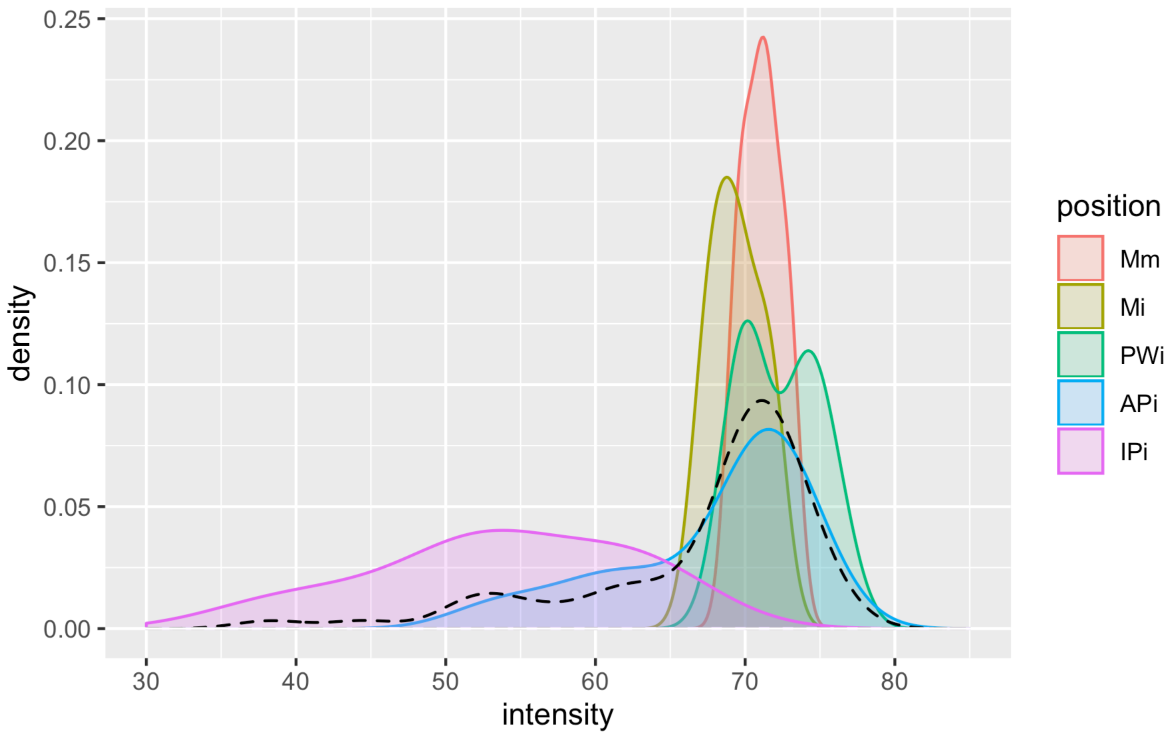


Figure F.55: Density plot of minimum intensity for nasal stops produced by Ulsan_5.

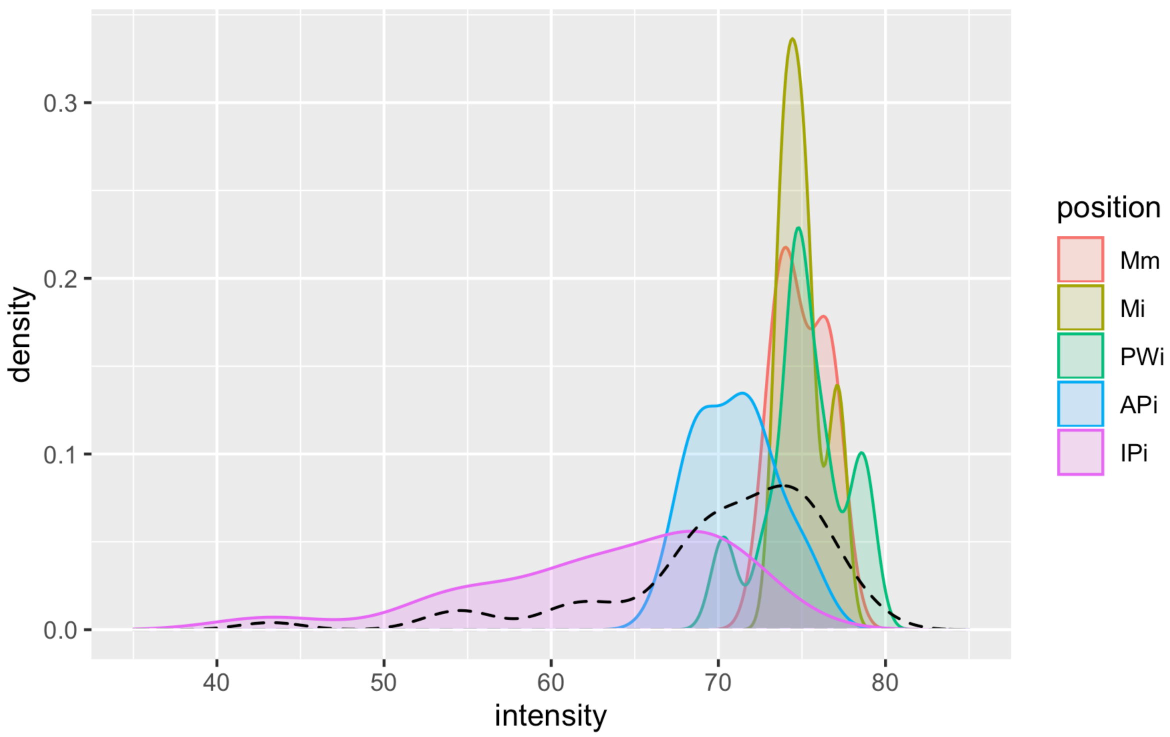


Figure F.56: Density plot of minimum intensity for nasal stops produced by Ulsan_6.

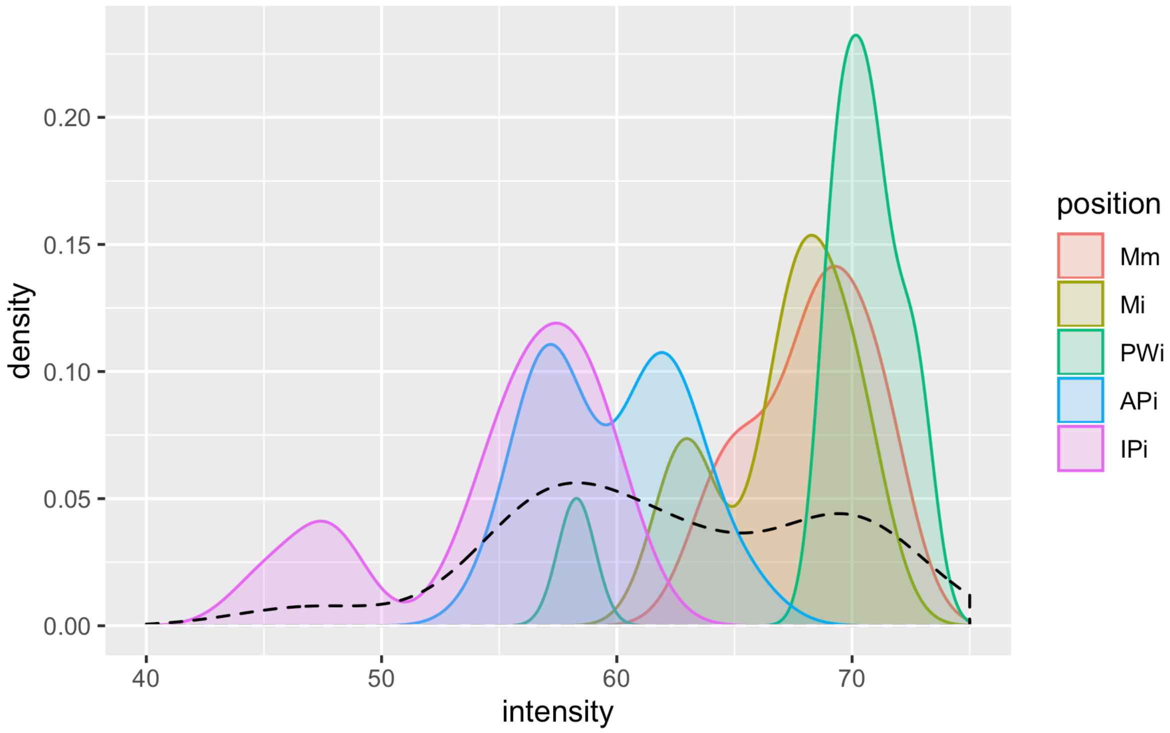


Figure F.57: Density plot of minimum intensity for nasal stops produced by Ulsan_7.

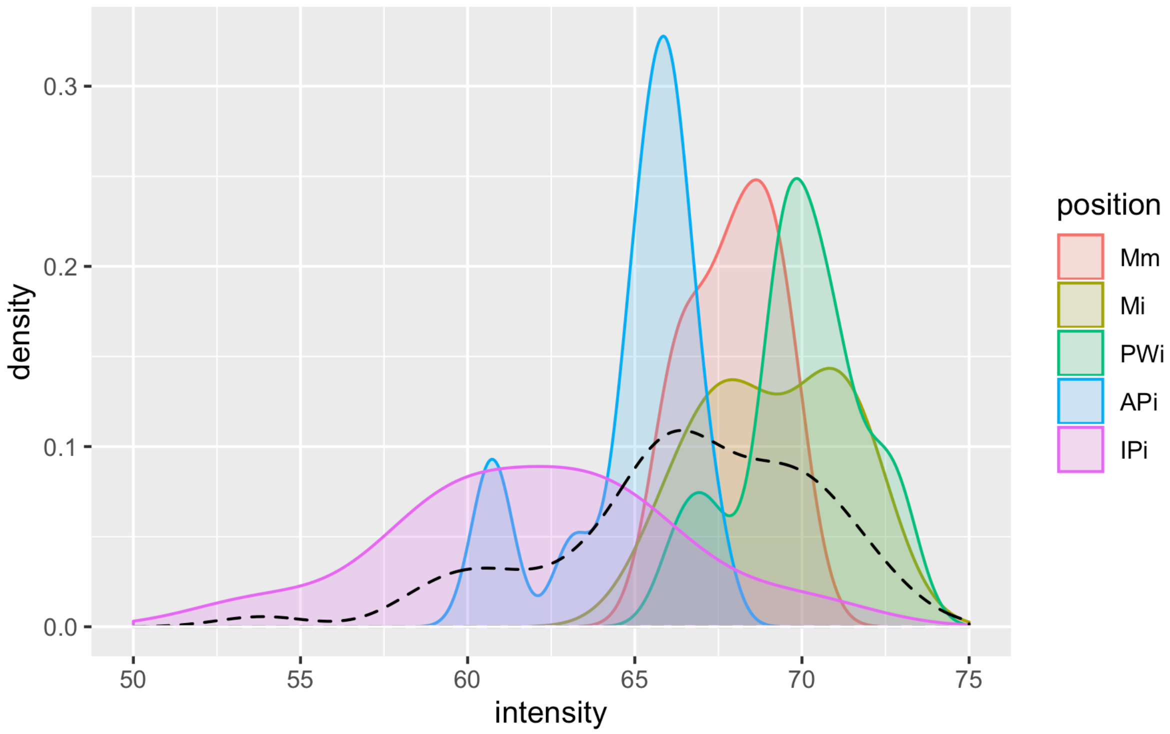


Figure F.58: Density plot of minimum intensity for nasal stops produced by Ulsan_8.

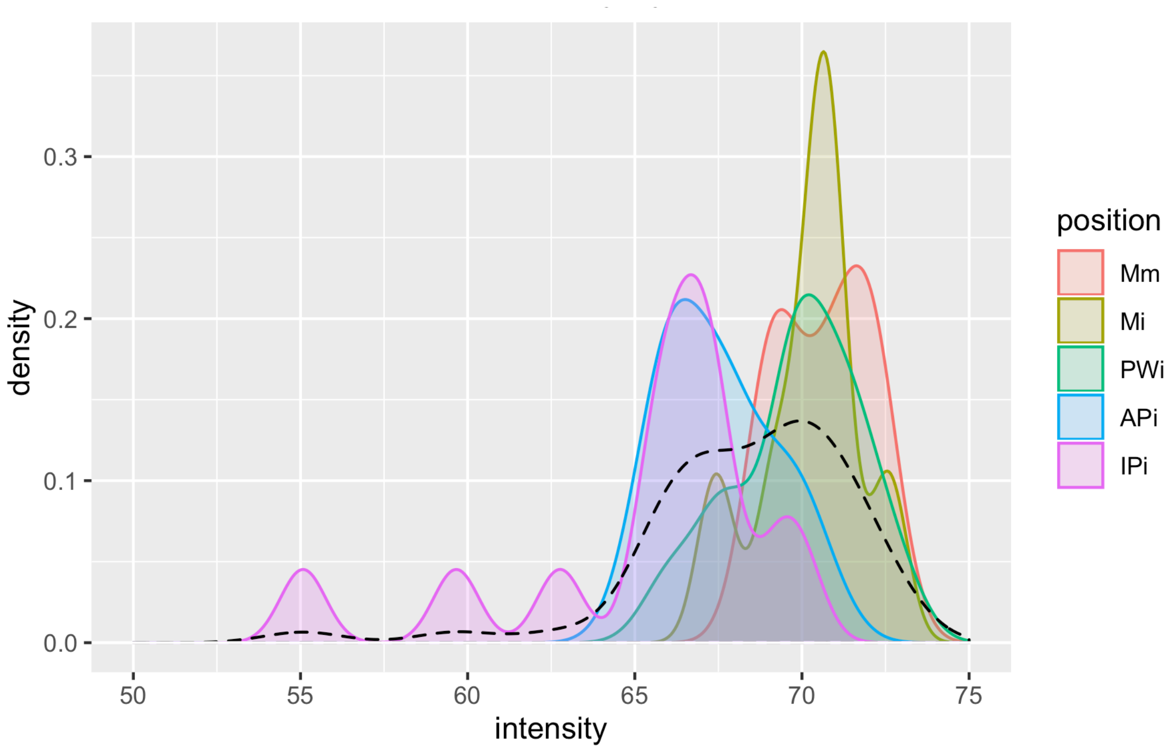


Figure F.59: Density plot of minimum intensity for nasal stops produced by Ulsan_9.

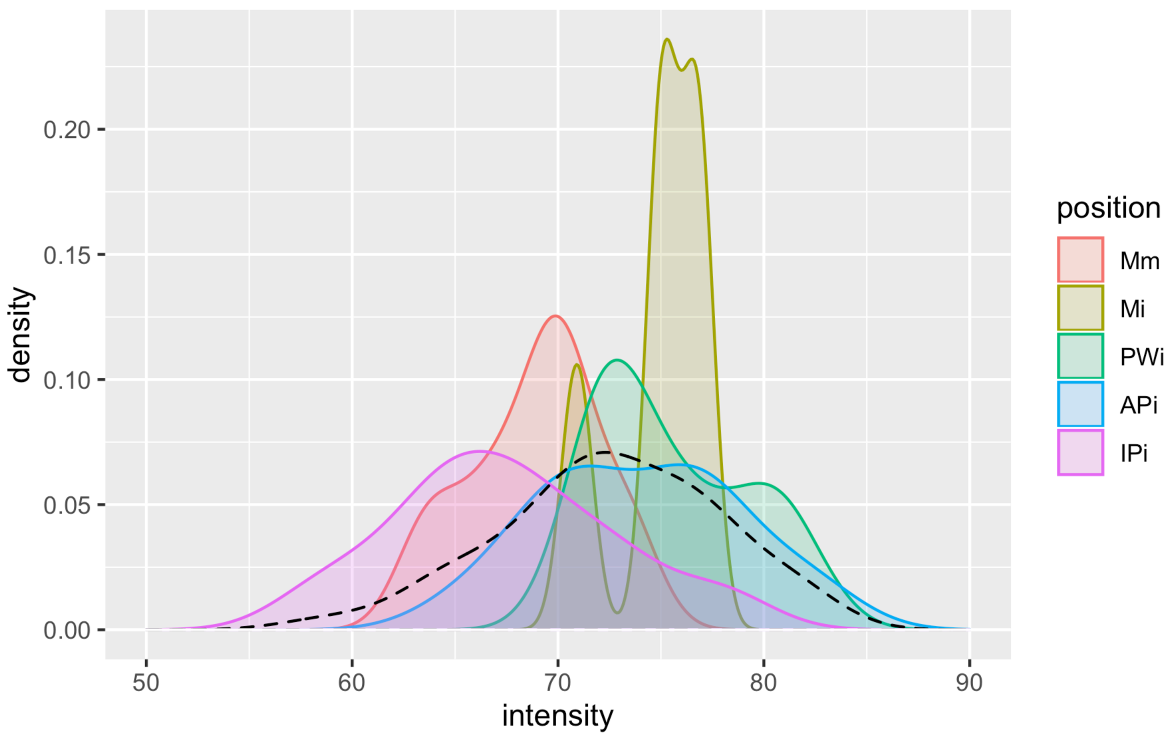


Figure F.60: Density plot of minimum intensity for nasal stops produced by Ulsan_10.

Appendix G.

Tukey HSD tests for individual analyses

I. % Voicing during closure for lenis stops

Table G.1: % Voicing during closure for lenis stops produced by Seoul_1.

contrast	estimate	SE	df	t.ratio	p.value
Mm - Mi	-33.20	21.6	33	-1.539	0.5458
Mm - PWi	-33.20	19.7	33	-1.685	0.4564
Mm - APi	-6.27	18.8	33	-0.333	0.9972
Mm - IPi	22.31	19.3	33	1.156	0.7756
Mi - PWi	0.00	15.3	33	0.000	1.0000
Mi - APi	26.93	14.1	33	1.906	0.3341
Mi - IPi	55.51	14.7	33	3.766	0.0055
PWi - APi	26.93	11.0	33	2.438	0.1304
PWi - IPi	55.51	11.8	33	4.697	0.0004
APi - IPi	28.59	10.3	33	2.771	0.0646

Table G.2: % Voicing during closure for lenis stops produced by Seoul_2.

contrast	estimate	SE	df	t.ratio	p.value
Mm - Mi	0.00	14.9	25	0.000	1.0000
Mm - PWi	1.58	13.8	25	0.114	1.0000
Mm - APi	35.50	14.5	25	2.455	0.1339
Mm - IPi	41.95	13.8	25	3.030	0.0408
Mi - PWi	1.58	11.7	25	0.135	0.9999
Mi - APi	35.50	12.4	25	2.866	0.0582
Mi - IPi	41.95	11.7	25	3.598	0.0110
PWi - APi	33.92	11.0	25	3.071	0.0373
PWi - IPi	40.37	10.2	25	3.948	0.0047
APi - IPi	6.45	11.0	25	0.584	0.9763

Table G.3: % Voicing during closure for lenis stops produced by Seoul_3.

contrast	estimate	SE	df	t.ratio	p.value
Mm - Mi	-7.33	12.83	27	-0.572	0.9781
Mm - PWi	-7.33	10.29	27	-0.713	0.9517
Mm - APi	5.11	9.60	27	0.532	0.9832
Mm - IPi	61.02	11.88	27	5.138	0.0002
Mi - PWi	0.00	11.37	27	0.000	1.0000
Mi - APi	12.45	10.76	27	1.157	0.7751
Mi - IPi	68.35	12.83	27	5.328	0.0001
PWi - APi	12.45	7.55	27	1.649	0.4808
PWi - IPi	68.35	10.29	27	6.645	<.0001
APi - IPi	55.91	9.60	27	5.821	<.0001

Table G.4: % Voicing during closure for lenis stops produced by Seoul_4.

contrast	estimate	SE	df	t.ratio	p.value
Mm - Mi	2.36	19.1	28	0.123	0.9999
Mm - PWi	-11.84	16.2	28	-0.731	0.9473
Mm - APi	12.95	16.1	28	0.806	0.9266
Mm - IPi	-29.78	30.9	28	-0.963	0.8691
Mi - PWi	-14.20	16.1	28	-0.883	0.9006
Mi - APi	10.60	15.3	28	0.693	0.9564
Mi - IPi	-32.13	30.3	28	-1.059	0.8254
PWi - APi	24.79	11.8	28	2.103	0.2471
PWi - IPi	-17.93	28.8	28	-0.623	0.9701
APi - IPi	-42.73	27.6	28	-1.547	0.5418

Table G.5: % Voicing during closure for lenis stops produced by Seoul_5.

contrast	estimate	SE	df	t.ratio	p.value
Mm - Mi	0.0	21.76	33	0.000	1.0000
Mm - PWi	0.0	20.33	33	0.000	1.0000
Mm - APi	35.1	20.65	33	1.700	0.4480
Mm - IPi	63.1	20.15	33	3.130	0.0280
Mi - PWi	0.0	11.36	33	0.000	1.0000
Mi - APi	35.1	11.92	33	2.944	0.0436
Mi - IPi	63.1	11.04	33	5.715	<.0001
PWi - APi	35.1	9.04	33	3.880	0.0040
PWi - IPi	63.1	7.84	33	8.041	<.0001
APi - IPi	28.0	8.63	33	3.242	0.0212

Table G.6: % Voicing during closure for lenis stops produced by Seoul_6.

contrast	estimate	SE	df	t.ratio	p.value
Mm - Mi	0.00	17.9	24	0.000	1.0000
Mm - PWi	-1.08	13.4	24	-0.080	1.0000
Mm - APi	47.88	12.1	24	3.955	0.0049
Mm - IPi	70.12	13.0	24	5.411	0.0001
Mi - PWi	-1.08	16.9	24	-0.064	1.0000
Mi - APi	47.88	15.9	24	3.009	0.0436
Mi - IPi	70.12	16.6	24	4.231	0.0025
PWi - APi	48.96	10.5	24	4.669	0.0008
PWi - IPi	71.19	11.6	24	6.117	<.0001
APi - IPi	22.23	10.1	24	2.198	0.2146

Table G.7: % Voicing during closure for lenis stops produced by Seoul_7.

contrast	estimate	SE	df	t.ratio	p.value
Mm - Mi	-23.9	23.9	34	-1.001	0.8528
Mm - PWi	-11.9	15.8	34	-0.752	0.9423
Mm - APi	23.6	16.0	34	1.475	0.5848
Mm - IPi	42.5	15.6	34	2.732	0.0698
Mi - PWi	12.0	21.2	34	0.567	0.9790
Mi - APi	47.5	21.6	34	2.196	0.2055
Mi - IPi	66.4	21.5	34	3.091	0.0303
PWi - APi	35.5	11.9	34	2.988	0.0388
PWi - IPi	54.4	11.3	34	4.819	0.0003
APi - IPi	18.9	11.4	34	1.654	0.4748

Table G.8: % Voicing during closure for lenis stops produced by Seoul_8.

contrast	estimate	SE	df	t.ratio	p.value
Mm - Mi	-6.19	16.70	32	-0.371	0.9958
Mm - PWi	15.88	13.45	32	1.181	0.7619
Mm - APi	27.97	12.80	32	2.184	0.2114
Mm - IPi	65.85	11.91	32	5.529	<.0001
Mi - PWi	22.07	15.87	32	1.390	0.6381
Mi - APi	34.16	15.33	32	2.228	0.1955
Mi - IPi	72.04	14.56	32	4.948	0.0002
PWi - APi	12.09	11.52	32	1.049	0.8305
PWi - IPi	49.97	10.64	32	4.699	0.0004
APi - IPi	37.89	9.81	32	3.861	0.0044

Table G.9: % Voicing during closure for lenis stops produced by Seoul_9.

contrast	estimate	SE	df	t.ratio	p.value
Mm - Mi	-19.67	12.15	37	-1.618	0.4956
Mm - PWi	-12.50	9.41	37	-1.329	0.6753
Mm - APi	51.90	9.08	37	5.715	<.0001
Mm - IPi	70.43	9.10	37	7.743	<.0001
Mi - PWi	7.17	10.46	37	0.685	0.9584
Mi - APi	71.57	10.19	37	7.024	<.0001
Mi - IPi	90.10	10.26	37	8.778	<.0001
PWi - APi	64.40	6.69	37	9.633	<.0001
PWi - IPi	82.93	6.76	37	12.264	<.0001
APi - IPi	18.53	6.27	37	2.954	0.0407

Table G.10: % Voicing during closure for lenis stops produced by Seoul_10.

contrast	estimate	SE	df	t.ratio	p.value
Mm - Mi	-25.68	18.5	43	-1.386	0.6398
Mm - PWi	-6.84	14.4	43	-0.475	0.9892
Mm - APi	19.70	14.5	43	1.356	0.6582
Mm - IPi	28.54	14.6	43	1.960	0.3027
Mi - PWi	18.84	15.6	43	1.204	0.7491
Mi - APi	45.38	15.8	43	2.871	0.0472
Mi - IPi	54.22	15.8	43	3.436	0.0110
PWi - APi	26.54	10.7	43	2.487	0.1125
PWi - IPi	35.39	10.6	43	3.328	0.0147
APi - IPi	8.85	10.9	43	0.812	0.9255

Table G.11: % Voicing during closure for lenis stops produced by Busan_1.

contrast	estimate	SE	df	t.ratio	p.value
Mm - Mi	-24.79	29.7	29	-0.834	0.9177
Mm - PWi	-20.11	23.2	29	-0.866	0.9071
Mm - APi	1.37	22.5	29	0.061	1.0000
Mm - IPi	6.07	23.5	29	0.258	0.9990
Mi - PWi	4.68	23.2	29	0.201	0.9996
Mi - APi	26.15	22.5	29	1.164	0.7710
Mi - IPi	30.85	23.5	29	1.312	0.6863
PWi - APi	21.48	12.7	29	1.690	0.4553
PWi - IPi	26.17	14.6	29	1.799	0.3935
APi - IPi	4.70	13.2	29	0.355	0.9964

Table G.12: % Voicing during closure for lenis stops produced by Busan_2.

contrast	estimate	SE	df	t.ratio	p.value
Mm - Mi	-17.01	27.8	22	-0.612	0.9717
Mm - PWi	-15.99	35.3	22	-0.453	0.9907
Mm - APi	-9.35	20.4	22	-0.458	0.9903
Mm - IPi	17.75	19.6	22	0.905	0.8918
Mi - PWi	1.02	37.7	22	0.027	1.0000
Mi - APi	7.66	23.7	22	0.323	0.9975
Mi - IPi	34.76	23.1	22	1.507	0.5692
PWi - APi	6.65	32.7	22	0.203	0.9996
PWi - IPi	33.74	32.1	22	1.050	0.8296
APi - IPi	27.09	13.2	22	2.057	0.2733

Table G.13: % Voicing during closure for lenis stops produced by Busan_3.

contrast	estimate	SE	df	t.ratio	p.value
Mm - Mi	0.652	19.8	29	0.033	1.0000
Mm - PWi	-0.978	18.4	29	-0.053	1.0000
Mm - APi	19.450	14.7	29	1.325	0.6784
Mm - IPi	40.679	15.3	29	2.653	0.0866
Mi - PWi	-1.631	20.1	29	-0.081	1.0000
Mi - APi	18.797	16.5	29	1.141	0.7839
Mi - IPi	40.026	17.2	29	2.327	0.1652
PWi - APi	20.428	14.9	29	1.366	0.6531
PWi - IPi	41.657	15.4	29	2.711	0.0766
APi - IPi	21.229	10.8	29	1.962	0.3093

Table G.14: % Voicing during closure for lenis stops produced by Busan_4.

contrast	estimate	SE	df	t.ratio	p.value
Mm - Mi	-36.24	22.5	35	-1.609	0.5019
Mm - PWi	-15.23	23.7	35	-0.642	0.9670
Mm - APi	3.33	18.8	35	0.178	0.9998
Mm - IPi	11.41	19.4	35	0.589	0.9758
Mi - PWi	21.01	22.5	35	0.933	0.8822
Mi - APi	39.57	17.2	35	2.299	0.1695
Mi - IPi	47.65	17.9	35	2.664	0.0803
PWi - APi	18.56	18.8	35	0.990	0.8580
PWi - IPi	26.64	19.4	35	1.375	0.6469
APi - IPi	8.07	12.8	35	0.630	0.9691

Table G.15: % Voicing during closure for lenis stops produced by Busan_5.

contrast	estimate	SE	df	t.ratio	p.value
Mm - Mi	-3.19	21.0	34	-0.152	0.9999
Mm - PWi	62.70	17.8	34	3.532	0.0099
Mm - APi	30.60	15.7	34	1.952	0.3105
Mm - IPi	52.50	15.5	34	3.376	0.0149
Mi - PWi	65.89	19.4	34	3.400	0.0140
Mi - APi	33.79	17.6	34	1.915	0.3292
Mi - IPi	55.69	17.6	34	3.171	0.0249
PWi - APi	-32.10	13.6	34	-2.352	0.1535
PWi - IPi	-10.20	13.5	34	-0.753	0.9420
APi - IPi	21.90	10.6	34	2.074	0.2544

Table G.16: % Voicing during closure for lenis stops produced by Busan_6.

contrast	estimate	SE	df	t.ratio	p.value
Mm - Mi	-16.497	18.7	38	-0.882	0.9016
Mm - PWi	9.822	18.0	38	0.545	0.9819
Mm - APi	-14.995	15.6	38	-0.958	0.8718
Mm - IPi	9.156	16.8	38	0.547	0.9817
Mi - PWi	26.319	18.0	38	1.460	0.5940
Mi - APi	1.503	15.6	38	0.096	1.0000
Mi - IPi	25.654	16.8	38	1.531	0.5491
PWi - APi	-24.816	14.9	38	-1.670	0.4638
PWi - IPi	-0.665	16.0	38	-0.041	1.0000
APi - IPi	24.151	13.2	38	1.825	0.3749

Table G.17: % Voicing during closure for lenis stops produced by Busan_7.

contrast	estimate	SE	df	t.ratio	p.value
Mm - PWi	0.0	18.7	21	0.000	1.0000
Mm - APi	26.5	16.0	21	1.657	0.3700
Mm - IPi	64.8	15.9	21	4.069	0.0029
PWi - APi	26.5	16.0	21	1.657	0.3700
PWi - IPi	64.8	15.9	21	4.069	0.0029
APi - IPi	38.2	12.7	21	3.011	0.0312

Table G.18: % Voicing during closure for lenis stops produced by Busan_8.

contrast	estimate	SE	df	t.ratio	p.value
Mm - Mi	2.41	23.8	15	0.101	1.0000
Mm - PWi	-1.21	16.7	15	-0.072	1.0000
Mm - APi	6.82	16.7	15	0.409	0.9935
Mm - IPi	74.84	14.8	15	5.066	0.0011
Mi - PWi	-3.62	22.4	15	-0.162	0.9998
Mi - APi	4.41	22.4	15	0.197	0.9996
Mi - IPi	72.42	20.5	15	3.533	0.0215
PWi - APi	8.03	13.5	15	0.595	0.9737
PWi - IPi	76.04	11.5	15	6.610	0.0001
APi - IPi	68.01	11.5	15	5.912	0.0002

Table G.19: % Voicing during closure for lenis stops produced by Busan_9.

contrast	estimate	SE	df	t.ratio	p.value
Mm - Mi	26.27	20.5	33	1.284	0.7024
Mm - PWi	33.82	22.0	33	1.540	0.5450
Mm - APi	20.63	16.5	33	1.247	0.7247
Mm - IPi	60.74	16.4	33	3.706	0.0064
Mi - PWi	7.55	22.0	33	0.343	0.9969
Mi - APi	-5.64	16.4	33	-0.343	0.9969
Mi - IPi	34.47	16.2	33	2.131	0.2314
PWi - APi	-13.19	18.4	33	-0.715	0.9516
PWi - IPi	26.92	18.3	33	1.474	0.5862
APi - IPi	40.11	10.9	33	3.676	0.0070

Table G.20: % Voicing during closure for lenis stops produced by Busan_10.

contrast	estimate	SE	df	t.ratio	p.value
Mm - Mi	15.29	14.7	24	1.043	0.7264
Mm - APi	12.70	13.4	24	0.945	0.7811
Mm - IPi	38.82	20.9	24	1.855	0.2737
Mi - APi	-2.59	11.7	24	-0.221	0.9961
Mi - IPi	23.54	19.9	24	1.184	0.6425
APi - IPi	26.13	16.1	24	1.627	0.3829

Table G.21: % Voicing during closure for lenis stops produced by Ulsan_1.

contrast	estimate	SE	df	t.ratio	p.value
Mm - Mi	-2.68	15.2	29	-0.176	0.9998
Mm - PWi	-6.51	14.2	29	-0.458	0.9905
Mm - APi	12.38	15.0	29	0.827	0.9202
Mm - IPi	16.91	12.2	29	1.383	0.6428
Mi - PWi	-3.83	13.7	29	-0.280	0.9986
Mi - APi	15.06	14.5	29	1.038	0.8358
Mi - IPi	19.59	12.0	29	1.631	0.4907
PWi - APi	18.89	12.5	29	1.511	0.5639
PWi - IPi	23.42	10.1	29	2.318	0.1682
APi - IPi	4.53	11.2	29	0.403	0.9941

Table G.22: % Voicing during closure for lenis stops produced by Ulsan_2.

contrast	estimate	SE	df	t.ratio	p.value
Mm - Mi	-10.60	20.8	36	-0.510	0.9858
Mm - PWi	-13.42	19.4	36	-0.693	0.9567
Mm - APi	20.30	18.4	36	1.102	0.8045
Mm - IPi	29.16	19.1	36	1.528	0.5516
Mi - PWi	-2.82	17.0	36	-0.166	0.9998
Mi - APi	30.90	15.9	36	1.943	0.3140
Mi - IPi	39.76	16.7	36	2.386	0.1425
PWi - APi	33.72	14.0	36	2.414	0.1348
PWi - IPi	42.58	14.9	36	2.862	0.0511
APi - IPi	8.86	13.6	36	0.651	0.9654

Table G.23: % Voicing during closure for lenis stops produced by Ulsan_3.

contrast	estimate	SE	df	t.ratio	p.value
Mm - Mi	0.0	19.32	30	0.000	1.0000
Mm - PWi	26.9	14.96	30	1.800	0.3920
Mm - APi	35.6	12.75	30	2.795	0.0633
Mm - IPi	74.7	13.02	30	5.738	<.0001
Mi - PWi	26.9	18.66	30	1.444	0.6055
Mi - APi	35.6	16.94	30	2.104	0.2449
Mi - IPi	74.7	17.14	30	4.358	0.0012
PWi - APi	8.7	11.74	30	0.741	0.9449
PWi - IPi	47.8	12.03	30	3.972	0.0035
APi - IPi	39.1	9.14	30	4.277	0.0015

Table G.24: % Voicing during closure for lenis stops produced by Ulsan_4.

contrast	estimate	SE	df	t.ratio	p.value
Mm - PWi	-6.4	21.8	24	-0.294	0.9910
Mm - APi	17.5	20.2	24	0.865	0.8227
Mm - IPi	28.9	20.8	24	1.391	0.5170
PWi - APi	23.9	18.4	24	1.299	0.5723
PWi - IPi	35.3	18.2	24	1.935	0.2405
APi - IPi	11.4	16.3	24	0.696	0.8975

Table G.25: % Voicing during closure for lenis stops produced by Ulsan_5.

contrast	estimate	SE	df	t.ratio	p.value
Mm - Mi	-17.09	19.61	31	-0.872	0.9051
Mm - PWi	-13.37	14.67	31	-0.911	0.8904
Mm - APi	25.47	12.79	31	1.991	0.2938
Mm - IPi	40.09	14.02	31	2.860	0.0541
Mi - PWi	3.72	18.39	31	0.202	0.9996
Mi - APi	42.55	16.96	31	2.510	0.1144
Mi - IPi	57.18	17.61	31	3.248	0.0218
PWi - APi	38.84	10.88	31	3.569	0.0097
PWi - IPi	53.46	11.87	31	4.504	0.0008
APi - IPi	14.63	9.57	31	1.529	0.5522

Table G.26: % Voicing during closure for lenis stops produced by Ulsan_6.

contrast	estimate	SE	df	t.ratio	p.value
Mm - Mi	1.85	16.36	33	0.113	1.0000
Mm - PWi	1.85	16.36	33	0.113	1.0000
Mm - APi	47.87	13.22	33	3.621	0.0081
Mm - IPi	86.04	13.67	33	6.294	<.0001
Mi - PWi	0.00	16.22	33	0.000	1.0000
Mi - APi	46.02	13.01	33	3.537	0.0100
Mi - IPi	84.18	13.24	33	6.360	<.0001
PWi - APi	46.02	13.01	33	3.537	0.0100
PWi - IPi	84.18	13.24	33	6.360	<.0001
APi - IPi	38.16	8.96	33	4.260	0.0014

Table G.27: % Voicing during closure for lenis stops produced by Ulsan_7.

contrast	estimate	SE	df	t.ratio	p.value
Mm - Mi	0.905	18.60	41	0.049	1.0000
Mm - PWi	-2.434	15.58	41	-0.156	0.9999
Mm - APi	40.175	14.91	41	2.694	0.0723
Mm - IPi	51.095	14.81	41	3.451	0.0108
Mi - PWi	-3.339	15.58	41	-0.214	0.9995
Mi - APi	39.270	14.91	41	2.633	0.0828
Mi - IPi	50.190	14.81	41	3.390	0.0128
PWi - APi	42.609	10.92	41	3.902	0.0030
PWi - IPi	53.529	10.79	41	4.961	0.0001
APi - IPi	10.920	9.78	41	1.117	0.7969

Table G.28: % Voicing during closure for lenis stops produced by Ulsan_8.

contrast	estimate	SE	df	t.ratio	p.value
Mm - Mi	4.793	16.16	25	0.297	0.9982
Mm - PWi	5.028	12.78	25	0.393	0.9946
Mm - APi	20.482	13.53	25	1.514	0.5633
Mm - IPi	46.632	13.53	25	3.447	0.0157
Mi - PWi	0.235	12.84	25	0.018	1.0000
Mi - APi	15.688	13.62	25	1.152	0.7776
Mi - IPi	41.839	13.62	25	3.073	0.0371
PWi - APi	15.453	9.47	25	1.632	0.4917
PWi - IPi	41.604	9.47	25	4.395	0.0015
APi - IPi	26.151	10.46	25	2.499	0.1229

Table G.29: % Voicing during closure for lenis stops produced by Ulsan_9.

contrast	estimate	SE	df	t.ratio	p.value
Mm - Mi	-3.346	29.35	26	-0.114	1.0000
Mm - PWi	-0.956	21.38	26	-0.045	1.0000
Mm - APi	55.655	21.45	26	2.594	0.1009
Mm - IPi	63.770	20.67	26	3.086	0.0352
Mi - PWi	2.390	22.09	26	0.108	1.0000
Mi - APi	59.001	21.12	26	2.794	0.0668
Mi - IPi	67.117	21.83	26	3.074	0.0362
PWi - APi	56.611	10.15	26	5.578	0.0001
PWi - IPi	64.726	9.38	26	6.898	<.0001
APi - IPi	8.116	9.17	26	0.885	0.8999

Table G.30: % Voicing during closure for lenis stops produced by Ulsan_10.

contrast	estimate	SE	df	t.ratio	p.value
Mm - Mi	-53.949	21.6	28	-2.498	0.1198
Mm - PWi	-54.302	19.3	28	-2.810	0.0628
Mm - APi	-17.514	17.1	28	-1.022	0.8433
Mm - IPi	-12.873	17.2	28	-0.747	0.9433
Mi - PWi	-0.354	19.3	28	-0.018	1.0000
Mi - APi	36.435	17.1	28	2.125	0.2379
Mi - IPi	41.076	17.2	28	2.384	0.1494
PWi - APi	36.789	14.1	28	2.607	0.0962
PWi - IPi	41.430	14.3	28	2.903	0.0511
APi - IPi	4.641	11.1	28	0.418	0.9933

II. Minimum intensity for nasals stops

Table G.31: Minimum intensity for nasal stops produced by Seoul_1.

contrast	estimate	SE	df	t.ratio	p.value
Mm - Mi	-1.564	2.43	50	-0.645	0.9668
Mm - PWi	0.223	1.98	50	0.113	1.0000
Mm - APi	8.112	1.92	50	4.231	0.0009
Mm - IPi	13.709	1.93	50	7.086	<.0001
Mi - PWi	1.787	2.11	50	0.848	0.9142
Mi - APi	9.676	2.05	50	4.715	0.0002
Mi - IPi	15.273	2.07	50	7.385	<.0001
PWi - APi	7.889	1.50	50	5.276	<.0001
PWi - IPi	13.486	1.52	50	8.887	<.0001
APi - IPi	5.597	1.44	50	3.889	0.0026

Table G.32: Minimum intensity for nasal stops produced by Seoul_2.

contrast	estimate	SE	df	t.ratio	p.value
Mm - Mi	0.102	3.07	39	0.033	1.0000
Mm - PWi	1.042	2.60	39	0.401	0.9943
Mm - APi	7.361	3.44	39	2.142	0.2234
Mm - IPi	13.712	2.60	39	5.279	<.0001
Mi - PWi	0.940	2.60	39	0.362	0.9962
Mi - APi	7.259	3.44	39	2.112	0.2354
Mi - IPi	13.610	2.60	39	5.240	0.0001
PWi - APi	6.319	3.02	39	2.094	0.2432
PWi - IPi	12.671	2.01	39	6.297	<.0001
APi - IPi	6.351	3.02	39	2.104	0.2388

Table G.33: Minimum intensity for nasal stops produced by Seoul_3.

Mm - Mi	-0.241	2.05	48	-0.118	1.0000
Mm - PWi	1.412	1.74	48	0.814	0.9252
Mm - APi	3.024	1.77	48	1.705	0.4404
Mm - IPi	12.095	1.93	48	6.266	<.0001
Mi - PWi	1.653	1.74	48	0.953	0.8746
Mi - APi	3.265	1.77	48	1.841	0.3626
Mi - IPi	12.336	1.93	48	6.390	<.0001
PWi - APi	1.611	1.40	48	1.150	0.7794
PWi - IPi	10.682	1.60	48	6.693	<.0001
APi - IPi	9.071	1.64	48	5.543	<.0001

Table G.34: Minimum intensity for nasal stops produced by Seoul_4.

contrast	estimate	SE	df	t.ratio	p.value
Mm - Mi	4.227	1.68	40	2.515	0.1075
Mm - PWi	1.231	1.52	40	0.811	0.9258
Mm - APi	5.400	1.47	40	3.668	0.0061
Mm - IPi	5.002	2.05	40	2.440	0.1256
Mi - PWi	-2.996	1.53	40	-1.964	0.3018
Mi - APi	1.173	1.49	40	0.788	0.9326
Mi - IPi	0.775	2.03	40	0.381	0.9953
PWi - APi	4.169	1.31	40	3.190	0.0219
PWi - IPi	3.772	1.92	40	1.959	0.3041
APi - IPi	-0.398	1.90	40	-0.210	0.9995

Table G.35: Minimum intensity for nasal stops produced by Seoul_5.

contrast	estimate	SE	df	t.ratio	p.value
Mm - Mi	-0.145	1.084	55	-0.134	0.9999
Mm - PWi	-0.805	0.949	55	-0.848	0.9142
Mm - APi	10.976	1.012	55	10.841	<.0001
Mm - IPi	11.923	0.933	55	12.783	<.0001
Mi - PWi	-0.660	0.907	55	-0.728	0.9491
Mi - APi	11.121	0.973	55	11.430	<.0001
Mi - IPi	12.068	0.890	55	13.563	<.0001
PWi - APi	11.780	0.820	55	14.364	<.0001
PWi - IPi	12.728	0.719	55	17.691	<.0001
APi - IPi	0.947	0.801	55	1.182	0.7615

Table G.36: Minimum intensity for nasal stops produced by Seoul_6.

contrast	estimate	SE	df	t.ratio	p.value
Mm - Mi	0.605	1.55	54	0.390	0.9949
Mm - PWi	1.940	1.28	54	1.511	0.5601
Mm - APi	9.916	1.27	54	7.794	<.0001
Mm - IPi	14.256	1.44	54	9.898	<.0001
Mi - PWi	1.334	1.35	54	0.992	0.8578
Mi - APi	9.310	1.33	54	6.974	<.0001
Mi - IPi	13.651	1.49	54	9.181	<.0001
PWi - APi	7.976	1.01	54	7.874	<.0001
PWi - IPi	12.317	1.21	54	10.157	<.0001
APi - IPi	4.341	1.20	54	3.602	0.0059

Table G.37: Minimum intensity for nasal stops produced by Seoul_7.

contrast	estimate	SE	df	t.ratio	p.value
Mm - Mi	1.853	4.08	43	0.455	0.9909
Mm - PWi	0.947	4.08	43	0.232	0.9993
Mm - APi	8.766	3.81	43	2.303	0.1636
Mm - IPi	21.710	3.62	43	5.992	<.0001
Mi - PWi	-0.907	3.94	43	-0.230	0.9993
Mi - APi	6.913	3.66	43	1.890	0.3383
Mi - IPi	19.856	3.45	43	5.751	<.0001
PWi - APi	7.819	3.66	43	2.137	0.2234
PWi - IPi	20.763	3.45	43	6.013	<.0001
APi - IPi	12.944	3.14	43	4.120	0.0015

Table G.38: Minimum intensity for nasal stops produced by Seoul_8.

contrast	estimate	SE	df	t.ratio	p.value
Mm - Mi	-0.151	1.78	43	-0.085	1.0000
Mm - PWi	0.450	1.56	43	0.288	0.9984
Mm - APi	4.840	1.45	43	3.342	0.0142
Mm - IPi	8.603	1.34	43	6.438	<.0001
Mi - PWi	0.601	1.79	43	0.336	0.9971
Mi - APi	4.992	1.69	43	2.961	0.0379
Mi - IPi	8.754	1.59	43	5.518	<.0001
PWi - APi	4.391	1.45	43	3.029	0.0320
PWi - IPi	8.153	1.34	43	6.070	<.0001
APi - IPi	3.763	1.21	43	3.115	0.0257

Table G.39: Minimum intensity for nasal stops produced by Seoul_9.

contrast	estimate	SE	df	t.ratio	p.value
Mm - Mi	1.490	3.31	43	0.451	0.9912
Mm - PWi	0.983	2.68	43	0.366	0.9960
Mm - APi	11.888	2.85	43	4.176	0.0013
Mm - IPi	22.769	2.98	43	7.648	<.0001
Mi - PWi	-0.507	2.96	43	-0.171	0.9998
Mi - APi	10.399	3.11	43	3.345	0.0141
Mi - IPi	21.279	3.23	43	6.590	<.0001
PWi - APi	10.906	2.43	43	4.479	0.0005
PWi - IPi	21.786	2.58	43	8.437	<.0001
APi - IPi	10.880	2.76	43	3.945	0.0026

Table G.40: Minimum intensity for nasal stops produced by Seoul_10.

contrast	estimate	SE	df	t.ratio	p.value
Mm - Mi	0.371	2.98	57	0.124	0.9999
Mm - PWi	-0.658	2.54	57	-0.259	0.9990
Mm - APi	3.899	2.72	57	1.431	0.6103
Mm - IPi	13.557	2.56	57	5.297	<.0001
Mi - PWi	-1.028	2.54	57	-0.405	0.9942
Mi - APi	3.529	2.72	57	1.295	0.6951
Mi - IPi	13.187	2.56	57	5.152	<.0001
PWi - APi	4.557	2.23	57	2.047	0.2573
PWi - IPi	14.215	2.02	57	7.041	<.0001
APi - IPi	9.658	2.25	57	4.291	0.0006

Table G.41: Minimum intensity for nasal stops produced by Busan_1.

contrast	estimate	SE	df	t.ratio	p.value
Mm - Mi	-2.27	1.90	49	-1.192	0.7556
Mm - PWi	-3.81	1.69	49	-2.258	0.1762
Mm - APi	-8.13	1.64	49	-4.960	0.0001
Mm - IPi	-14.13	1.97	49	-7.187	<.0001
Mi - PWi	-1.54	1.61	49	-0.955	0.8738
Mi - APi	-5.86	1.56	49	-3.751	0.0041
Mi - IPi	-11.86	1.90	49	-6.225	<.0001
PWi - APi	-4.32	1.29	49	-3.362	0.0125
PWi - IPi	-10.32	1.69	49	-6.118	<.0001
APi - IPi	-6.00	1.64	49	-3.656	0.0054

Table G.42: Minimum intensity for nasal stops produced by Busan_2.

contrast	estimate	SE	df	t.ratio	p.value
Mm - Mi	-1.22	2.56	31	-0.479	0.9887
Mm - PWi	-0.01	2.81	31	-0.004	1.0000
Mm - APi	-4.59	2.05	31	-2.243	0.1910
Mm - IPi	-7.28	1.93	31	-3.769	0.0058
Mi - PWi	1.21	3.03	31	0.400	0.9943
Mi - APi	-3.36	2.34	31	-1.435	0.6103
Mi - IPi	-6.06	2.25	31	-2.699	0.0771
PWi - APi	-4.58	2.62	31	-1.750	0.4199
PWi - IPi	-7.27	2.53	31	-2.876	0.0522
APi - IPi	-2.70	1.64	31	-1.644	0.4819

Table G.43: Minimum intensity for nasal stops produced by Busan_3.

contrast	estimate	SE	df	t.ratio	p.value
Mm - Mi	1.901	1.81	42	1.053	0.8291
Mm - PWi	2.029	1.68	42	1.208	0.7466
Mm - APi	0.258	1.68	42	0.153	0.9999
Mm - IPi	-13.142	1.72	42	-7.657	<.0001
Mi - PWi	0.127	1.68	42	0.076	1.0000
Mi - APi	-1.643	1.68	42	-0.975	0.8647
Mi - IPi	-15.043	1.72	42	-8.765	<.0001
PWi - APi	-1.770	1.54	42	-1.147	0.7805
PWi - IPi	-15.171	1.58	42	-9.607	<.0001
APi - IPi	-13.401	1.58	42	-8.490	<.0001

Table G.44: Minimum intensity for nasal stops produced by Busan_4.

contrast	estimate	SE	df	t.ratio	p.value
Mm - Mi	-2.46	1.214	53	-2.027	0.2676
Mm - PWi	-1.41	1.071	53	-1.314	0.6838
Mm - APi	-4.22	0.996	53	-4.235	0.0008
Mm - IPi	-8.39	1.040	53	-8.070	<.0001
Mi - PWi	1.05	1.119	53	0.941	0.8794
Mi - APi	-1.76	1.045	53	-1.682	0.4534
Mi - IPi	-5.93	1.088	53	-5.448	<.0001
PWi - APi	-2.81	0.875	53	-3.212	0.0182
PWi - IPi	-6.98	0.922	53	-7.573	<.0001
APi - IPi	-4.17	0.836	53	-4.989	0.0001

Table G.45: Minimum intensity for nasal stops produced by Busan_5.

contrast	estimate	SE	df	t.ratio	p.value
Mm - Mi	-0.767	2.42	52	-0.316	0.9978
Mm - PWi	-1.338	1.96	52	-0.682	0.9596
Mm - APi	-6.114	1.85	52	-3.297	0.0146
Mm - IPi	-10.783	1.86	52	-5.806	<.0001
Mi - PWi	-0.571	2.21	52	-0.259	0.9990
Mi - APi	-5.347	2.10	52	-2.540	0.0973
Mi - IPi	-10.016	2.10	52	-4.766	0.0001
PWi - APi	-4.776	1.56	52	-3.066	0.0272
PWi - IPi	-9.445	1.56	52	-6.050	<.0001
APi - IPi	-4.669	1.42	52	-3.294	0.0147

Table G.46: Minimum intensity for nasal stops produced by Busan_6.

contrast	estimate	SE	df	t.ratio	p.value
Mm - Mi	-4.07	2.17	50	-1.880	0.3410
Mm - PWi	-2.97	2.07	50	-1.434	0.6090
Mm - APi	-8.88	1.86	50	-4.779	0.0001
Mm - IPi	-15.32	1.95	50	-7.866	<.0001
Mi - PWi	1.11	2.07	50	0.535	0.9833
Mi - APi	-4.81	1.86	50	-2.587	0.0883
Mi - IPi	-11.25	1.95	50	-5.774	<.0001
PWi - APi	-5.91	1.75	50	-3.383	0.0117
PWi - IPi	-12.35	1.84	50	-6.699	<.0001
APi - IPi	-6.44	1.60	50	-4.034	0.0017

Table G.47: Minimum intensity for nasal stops produced by Busan_7.

contrast	estimate	SE	df	t.ratio	p.value
Mm - Mi	1.906	4.56	41	0.418	0.9934
Mm - PWi	-0.353	2.26	41	-0.156	0.9999
Mm - APi	-2.085	2.01	41	-1.038	0.8362
Mm - IPi	-9.591	2.00	41	-4.787	0.0002
Mi - PWi	-2.259	4.48	41	-0.504	0.9865
Mi - APi	-3.991	4.34	41	-0.920	0.8874
Mi - IPi	-11.497	4.35	41	-2.646	0.0805
PWi - APi	-1.732	1.82	41	-0.953	0.8743
PWi - IPi	-9.238	1.81	41	-5.096	0.0001
APi - IPi	-7.506	1.48	41	-5.069	0.0001

Table G.48: Minimum intensity for nasal stops produced by Busan_8.

contrast	estimate	SE	df	t.ratio	p.value
Mm - Mi	2.185	1.58	40	1.380	0.6437
Mm - PWi	-1.177	1.49	40	-0.790	0.9320
Mm - APi	-1.691	1.54	40	-1.096	0.8076
Mm - IPi	-8.482	1.39	40	-6.122	<.0001
Mi - PWi	-3.362	1.34	40	-2.510	0.1087
Mi - APi	-3.876	1.40	40	-2.768	0.0614
Mi - IPi	-10.666	1.23	40	-8.706	<.0001
PWi - APi	-0.514	1.28	40	-0.400	0.9944
PWi - IPi	-7.305	1.10	40	-6.655	<.0001
APi - IPi	-6.790	1.17	40	-5.795	<.0001

Table G.49: Minimum intensity for nasal stops produced by Busan_9.

contrast	estimate	SE	df	t.ratio	p.value
Mm - Mi	-0.55621	2.56	52	-0.218	0.9995
Mm - PWi	-0.55025	2.21	52	-0.249	0.9991
Mm - APi	-8.21752	2.24	52	-3.671	0.0050
Mm - IPi	-17.15965	2.38	52	-7.221	<.0001
Mi - PWi	0.00596	2.21	52	0.003	1.0000
Mi - APi	-7.66131	2.24	52	-3.422	0.0103
Mi - IPi	-16.60344	2.38	52	-6.987	<.0001
PWi - APi	-7.66727	1.84	52	-4.172	0.0011
PWi - IPi	-16.60940	2.00	52	-8.291	<.0001
APi - IPi	-8.94213	2.03	52	-4.406	0.0005

Table G.50: Minimum intensity for nasal stops produced by Busan_10.

contrast	estimate	SE	df	t.ratio	p.value
Mm - Mi	0.300	1.51	31	0.199	0.9996
Mm - PWi	-0.497	1.83	31	-0.272	0.9987
Mm - APi	-2.651	1.29	31	-2.061	0.2622
Mm - IPi	-10.882	2.26	31	-4.823	0.0003
Mi - PWi	-0.797	1.87	31	-0.426	0.9928
Mi - APi	-2.951	1.35	31	-2.182	0.2130
Mi - IPi	-11.182	2.29	31	-4.873	0.0003
PWi - APi	-2.153	1.45	31	-1.488	0.5777
PWi - IPi	-10.384	2.25	31	-4.620	0.0006
APi - IPi	-8.231	2.04	31	-4.029	0.0029

Table G.51: Minimum intensity for nasal stops produced by Ulsan_1.

contrast	estimate	SE	df	t.ratio	p.value
Mm - Mi	0.358	1.373	33	0.261	0.9989
Mm - PWi	-0.219	1.168	33	-0.188	0.9997
Mm - APi	-3.044	1.540	33	-1.977	0.2992
Mm - IPi	-8.461	1.222	33	-6.923	<.0001
Mi - PWi	-0.578	1.172	33	-0.493	0.9875
Mi - APi	-3.403	1.529	33	-2.225	0.1957
Mi - IPi	-8.820	1.229	33	-7.176	<.0001
PWi - APi	-2.825	1.363	33	-2.073	0.2558
PWi - IPi	-8.242	0.996	33	-8.276	<.0001
APi - IPi	-5.417	1.414	33	-3.832	0.0046

Table G.52: Minimum intensity for nasal stops produced by Ulsan_2.

contrast	estimate	SE	df	t.ratio	p.value
Mm - Mi	1.858	5.13	42	0.362	0.9962
Mm - PWi	2.212	3.28	42	0.674	0.9610
Mm - APi	8.683	3.07	42	2.825	0.0531
Mm - IPi	15.030	3.24	42	4.644	0.0003
Mi - PWi	0.355	4.77	42	0.074	1.0000
Mi - APi	6.825	4.61	42	1.480	0.5810
Mi - IPi	13.172	4.72	42	2.789	0.0577
PWi - APi	6.471	2.32	42	2.784	0.0585
PWi - IPi	12.818	2.54	42	5.056	0.0001
APi - IPi	6.347	2.26	42	2.806	0.0555

Table G.53: Minimum intensity for nasal stops produced by Ulsan_3.

contrast	estimate	SE	df	t.ratio	p.value
Mm - Mi	1.845	1.89	48	0.976	0.8645
Mm - PWi	2.397	1.63	48	1.474	0.5837
Mm - APi	5.159	1.53	48	3.368	0.0124
Mm - IPi	10.566	1.57	48	6.721	<.0001
Mi - PWi	0.552	1.78	48	0.311	0.9979
Mi - APi	3.314	1.69	48	1.961	0.3004
Mi - IPi	8.721	1.73	48	5.051	0.0001
PWi - APi	2.762	1.39	48	1.989	0.2869
PWi - IPi	8.169	1.43	48	5.699	<.0001
APi - IPi	5.407	1.33	48	4.078	0.0015

Table G.54: Minimum intensity for nasal stops produced by Ulsan_4.

contrast	estimate	SE	df	t.ratio	p.value
Mm - Mi	-0.0703	2.35	44	-0.030	1.0000
Mm - PWi	0.0747	1.99	44	0.037	1.0000
Mm - APi	2.6240	2.02	44	1.301	0.6919
Mm - IPi	12.7234	2.49	44	5.100	0.0001
Mi - PWi	0.1450	2.08	44	0.070	1.0000
Mi - APi	2.6942	2.10	44	1.281	0.7043
Mi - IPi	12.7936	2.59	44	4.948	0.0001
PWi - APi	2.5492	1.69	44	1.509	0.5624
PWi - IPi	12.6486	2.27	44	5.571	<.0001
APi - IPi	10.0994	2.28	44	4.424	0.0006

Table G.55: Minimum intensity for nasal stops produced by Ulsan_5.

contrast	estimate	SE	df	t.ratio	p.value
Mm - Mi	4.016	3.22	49	1.247	0.7239
Mm - PWi	-0.960	2.58	49	-0.372	0.9958
Mm - APi	4.310	2.36	49	1.826	0.3705
Mm - IPi	19.269	2.91	49	6.629	<.0001
Mi - PWi	-4.976	3.01	49	-1.651	0.4731
Mi - APi	0.295	2.52	49	0.117	1.0000
Mi - IPi	15.253	2.76	49	5.535	<.0001
PWi - APi	5.271	2.15	49	2.455	0.1182
PWi - IPi	20.229	2.70	49	7.489	<.0001
APi - IPi	14.959	2.22	49	6.726	<.0001

Table G.56: Minimum intensity for nasal stops produced by Ulsan_6.

contrast	estimate	SE	df	t.ratio	p.value
Mm - Mi	-0.0406	3.14	46	-0.013	1.0000
Mm - PWi	-0.1558	2.74	46	-0.057	1.0000
Mm - APi	3.9754	2.59	46	1.535	0.5458
Mm - IPi	11.8919	2.55	46	4.671	0.0002
Mi - PWi	-0.1152	2.74	46	-0.042	1.0000
Mi - APi	4.0159	2.59	46	1.550	0.5360
Mi - IPi	11.9324	2.55	46	4.686	0.0002
PWi - APi	4.1311	2.02	46	2.041	0.2635
PWi - IPi	12.0477	1.98	46	6.095	<.0001
APi - IPi	7.9165	1.76	46	4.505	0.0004

Table G.57: Minimum intensity for nasal stops produced by Ulsan_7.

contrast	estimate	SE	df	t.ratio	p.value
Mm - Mi	-0.716	2.28	46	-0.314	0.9978
Mm - PWi	1.256	1.79	46	0.702	0.9550
Mm - APi	-8.675	1.63	46	-5.313	<.0001
Mm - IPi	-13.578	1.69	46	-8.052	<.0001
Mi - PWi	1.972	2.15	46	0.916	0.8892
Mi - APi	-7.959	2.04	46	-3.907	0.0027
Mi - IPi	-12.862	2.08	46	-6.173	<.0001
PWi - APi	-9.931	1.45	46	-6.838	<.0001
PWi - IPi	-14.834	1.51	46	-9.806	<.0001
APi - IPi	-4.903	1.31	46	-3.741	0.0044

Table G.58: Minimum intensity for nasal stops produced by Ulsan_8.

contrast	estimate	SE	df	t.ratio	p.value
Mm - Mi	1.136	1.33	47	0.853	0.9124
Mm - PWi	1.944	1.22	47	1.588	0.5125
Mm - APi	-3.084	1.19	47	-2.587	0.0892
Mm - IPi	-6.194	1.24	47	-5.013	0.0001
Mi - PWi	0.808	1.17	47	0.688	0.9581
Mi - APi	-4.221	1.14	47	-3.701	0.0049
Mi - IPi	-7.331	1.18	47	-6.210	<.0001
PWi - APi	-5.029	1.01	47	-4.968	0.0001
PWi - IPi	-8.139	1.06	47	-7.698	<.0001
APi - IPi	-3.110	1.02	47	-3.051	0.0293

Table G.59: Minimum intensity for nasal stops produced by Ulsan_9.

contrast	estimate	SE	df	t.ratio	p.value
Mm - Mi	-0.463	1.30	46	-0.356	0.9964
Mm - PWi	-1.030	1.16	46	-0.886	0.9007
Mm - APi	-2.985	1.15	46	-2.589	0.0891
Mm - IPi	-5.708	1.18	46	-4.845	0.0001
Mi - PWi	-0.567	1.20	46	-0.472	0.9895
Mi - APi	-2.522	1.21	46	-2.087	0.2431
Mi - IPi	-5.245	1.21	46	-4.335	0.0007
PWi - APi	-1.955	1.08	46	-1.813	0.3786
PWi - IPi	-4.679	1.01	46	-4.615	0.0003
APi - IPi	-2.723	1.11	46	-2.449	0.1208

Table G.60: Minimum intensity for nasal stops produced by Ulsan_10.

contrast	estimate	SE	df	t.ratio	p.value
Mm - Mi	5.983	2.67	45	2.240	0.1838
Mm - PWi	6.570	2.15	45	3.050	0.0298
Mm - APi	4.951	2.14	45	2.309	0.1609
Mm - IPi	-1.619	2.25	45	-0.719	0.9511
Mi - PWi	0.587	2.30	45	0.256	0.9990
Mi - APi	-1.032	2.28	45	-0.452	0.9911
Mi - IPi	-7.602	2.38	45	-3.188	0.0209
PWi - APi	-1.620	1.64	45	-0.986	0.8604
PWi - IPi	-8.190	1.78	45	-4.596	0.0003
APi - IPi	-6.570	1.75	45	-3.755	0.0043

Appendix H.

Correlations

I. % Voicing during closure for lenis stops

Table H.1: % Voicing during closure for lenis stops produced by Seoul speakers.

Subject	Position	Pearson's *r*	*p*-value	Dur. range (ms)	Token no.
Seoul_1	AP-medial	0.157402	0.591	31.0276	14
	AP-initial	-0.39056	0.1674	43.6589	14
	IP-initial	-0.3096	0.384	54.8023	10
Seoul_2	AP-medial	0.089221	0.7425	37.315	16
	AP-initial	-0.686	0.1324	43.259	6
	IP-initial	-0.77368	0.02428 **	54.2837	8
Seoul_3	AP-medial	-0.20321	0.4676	37.7757	15
	AP-initial	-0.33648	0.261	66.3798	13
	IP-initial	0.284384	0.7156	33.1855	4
Seoul_4	AP-medial	-0.7317	0.000842 ***	37.4696	17
	AP-initial	-0.4021	0.1226	33.7749	16
	IP-initial	*not enough finite observations*		0	1
Seoul_5	AP-medial	NA (sd = 0)	NA	29.9712	16
	AP-initial	-0.91813	0.001289 ***	46.1656	8
	IP-initial	-0.63394	0.01491 **	87.2931	14
Seoul_6	AP-medial	NA (sd = 0)	NA	54.2605	12
	AP-initial	-0.20981	0.5358	32.0254	11
	IP-initial	-0.01757	0.9702	76.9124	7
Seoul_7	AP-medial	-0.68451	0.002436 ***	57.6961	17
	AP-initial	-0.68784	0.02791 **	61.8665	10
	IP-initial	-0.49038	0.08888 *	156.483	13
Seoul_8	AP-medial	-0.00413	0.9884	44.619	15
	AP-initial	-0.55824	0.1183	39.147	9
	IP-initial	0.20572	0.4805	45.1496	14
Seoul_9	AP-medial	-0.05833	0.824	42.4394	17
	AP-initial	-0.55607	0.04845 **	61.3072	13
	IP-initial	-0.57979	0.03781 **	312.094	13
Seoul_10	AP-medial	-0.57852	0.003829 ***	50.2525	23
	AP-initial	-0.54015	0.0567 *	47.9989	13
	IP-initial	-0.41552	0.1579	228.716	13

Table H.2: % Voicing during closure for lenis stops produced by Busan speakers.

Subject	Position	Pearson's *r*	*p*-value	Dur. range (ms)	Token no.
Busan_1	AP-medial	0.046811	0.8793	65.8737	13
	AP-initial	-0.49748	0.07029 *	61.4074	14
	IP-initial	-0.17977	0.6701	57.0208	8
Busan_2	AP-medial	-0.1637	0.7566	41.5567	6
	AP-initial	-0.07496	0.848	26.1036	9
	IP-initial	-0.71523	0.005991 ***	48.3054	13
Busan_3	AP-medial	NA (sd = 0)	NA	36.6446	11
	AP-initial	-0.32889	0.2509	30.5509	14
	IP-initial	0.448379	0.1937	53.5246	10
Busan_4	AP-medial	-0.79258	0.001233 ***	45.4815	13
	AP-initial	-0.22296	0.4065	46.5152	16
	IP-initial	-0.62846	0.02862 **	73.6787	12
Busan_5	AP-medial	-0.73412	0.004275 ***	55.5371	13
	AP-initial	0.022323	0.9423	48.7052	13
	IP-initial	-0.29484	0.3062	129.421	14
Busan_6	AP-medial	-0.19426	0.4255	81.0492	19
	AP-initial	-0.46767	0.07877 *	51.1697	15
	IP-initial	-0.36592	0.2984	76.1964	10
Busan_7	AP-medial	NA (sd = 0)	NA	32.3085	8
	AP-initial	-0.94658	0.00011 ***	54.021	9
	IP-initial	-0.40032	0.2857	64.8695	9
Busan_8	AP-medial	NA (sd = 0)	NA	38.1634	7
	AP-initial	0.901254	0.09875 *	12.2024	4
	IP-initial	0.047218	0.8969	53.4148	10
Busan_9	AP-medial	-0.12613	0.7117	20.5371	11
	AP-initial	-0.6705	0.01214 **	73.3146	13
	IP-initial	-0.23906	0.3908	154.913	15
Busan_10	AP-medial	0.057417	0.8748	110.084	10
	AP-initial	-0.26519	0.3812	50.4968	13
	IP-initial	-0.42485	0.4011	28.8583	6

Table H.3: % Voicing during closure for lenis stops produced by Ulsan speakers.

Subject	Position	Pearson's *r*	*p*-value	Dur. range (ms)	Token no.
Ulsan_1	AP-medial	NA (sd = 0)	NA	46.9379	15
	AP-initial	-0.78247	0.1177	25.636	5
	IP-initial	-0.26596	0.338	71.1152	15
Ulsan_2	AP-medial	-0.64159	0.003067 ***	28.1331	19
	AP-initial	-0.48059	0.09644 *	44.7888	13
	IP-initial	-0.72837	0.0169 **	81.4674	10
Ulsan_3	AP-medial	-0.62213	0.04096 **	48.1847	11
	AP-initial	-0.32338	0.2811	46.1374	13
	IP-initial	-0.56812	0.06825 *	33.0982	11
Ulsan_4	AP-medial	-0.21285	0.5298	30.0666	11
	AP-initial	-0.25652	0.4744	42.079	10
	IP-initial	0.067271	0.8742	21.0503	8
Ulsan_5	AP-medial	-0.67437	0.01616 **	54.3828	12
	AP-initial	-0.23389	0.4015	28.6785	15
	IP-initial	-0.25686	0.4737	75.5827	10
Ulsan_6	AP-medial	NA (sd = 0)	NA	38.0031	12
	AP-initial	-0.38305	0.1764	25.5626	14
	IP-initial	-0.33143	0.2686	89.7105	13
Ulsan_7	AP-medial	-0.53653	0.0217 **	54.42	18
	AP-initial	-0.74707	0.002136 ***	54.2085	14
	IP-initial	-0.54271	0.03658 **	30.2776	15
Ulsan_8	AP-medial	-0.43255	0.0829 *	33.5867	17
	AP-initial	-0.86864	0.01118 **	42.8236	7
	IP-initial	-0.35984	0.4279	42.9661	7
Ulsan_9	AP-medial	NA (sd = 0)	NA	35.0073	9
	AP-initial	-0.81916	0.003738 ***	38.3649	10
	IP-initial	-0.56357	0.04488 **	62.7097	13
Ulsan_10	AP-medial	-0.37368	0.2576	19.4582	11
	AP-initial	-0.81382	0.001277 ***	39.1209	12
	IP-initial	0.375189	0.2555	40.7535	11

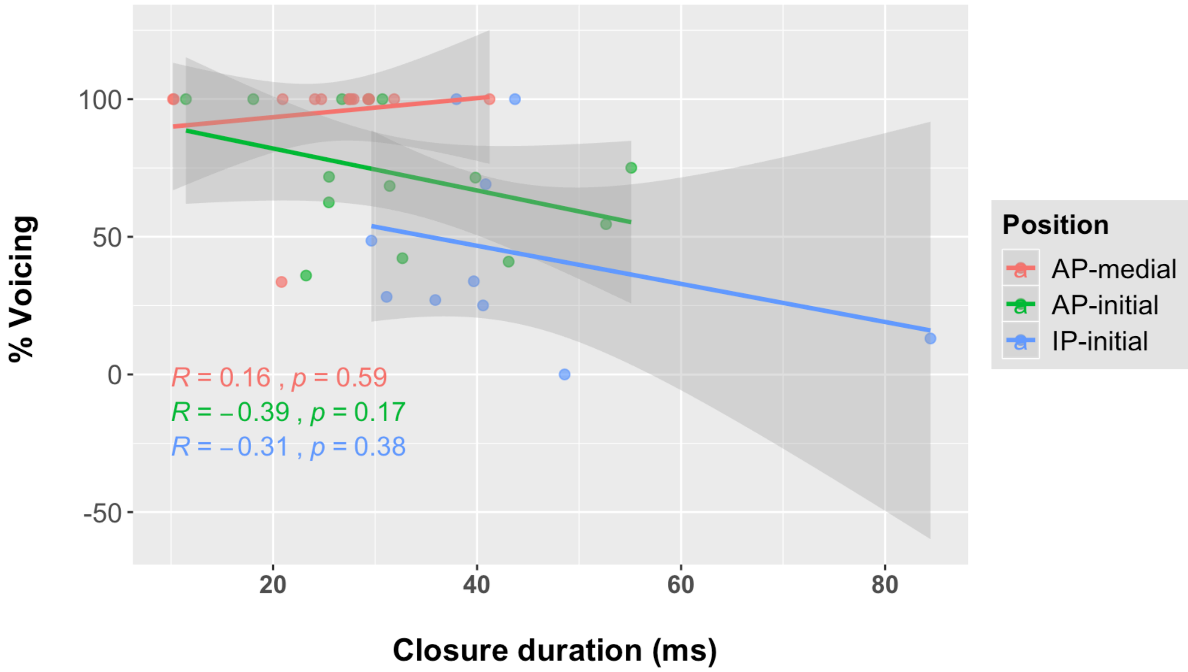


Figure H.1: % Voicing against closure duration (ms) for lenis stops produced by Seoul_1.

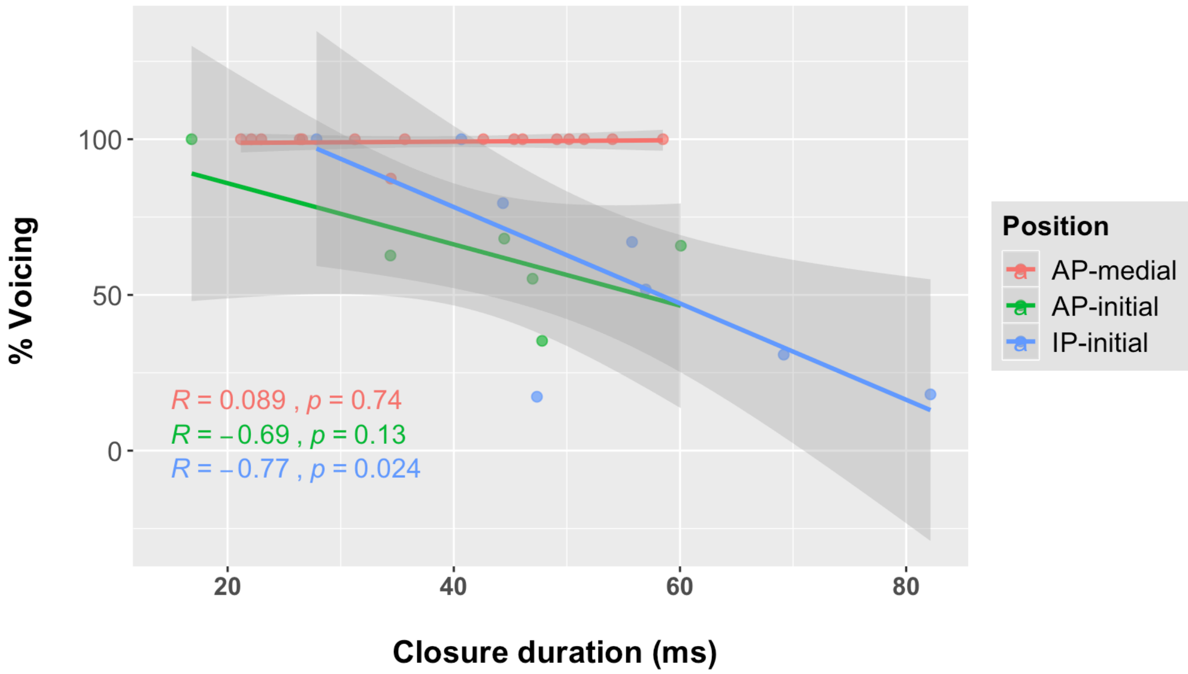


Figure H.2: % Voicing against closure duration (ms) for lenis stops produced by Seoul_2.

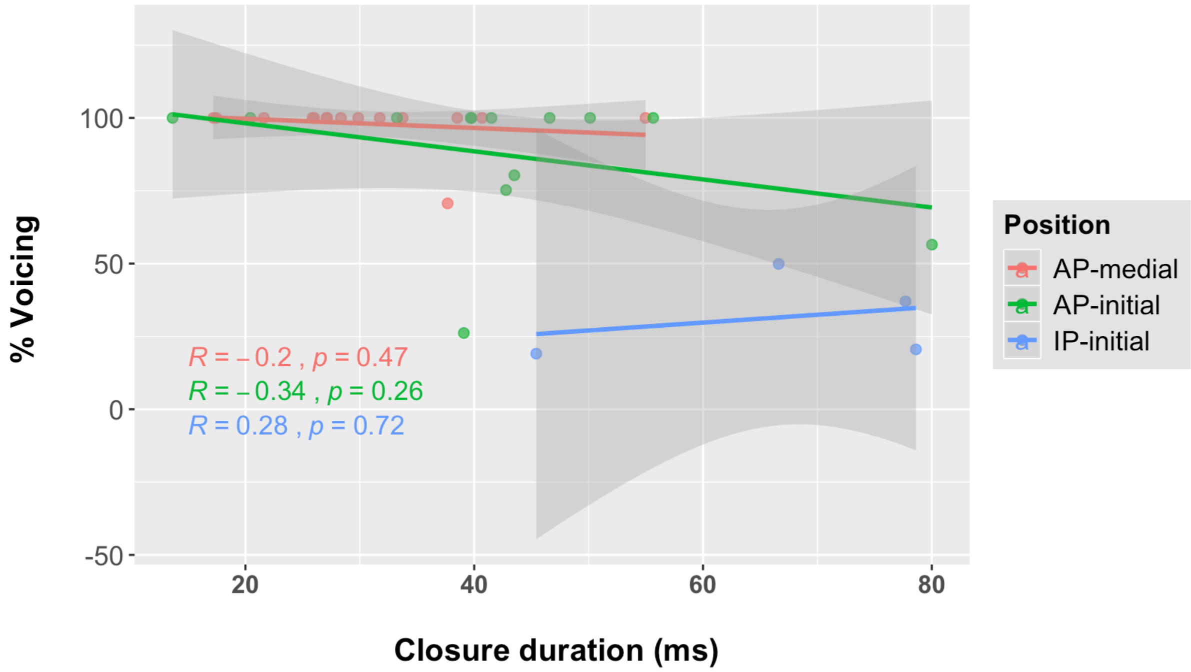


Figure H.3: % Voicing against closure duration (ms) for lenis stops produced by Seoul_3.

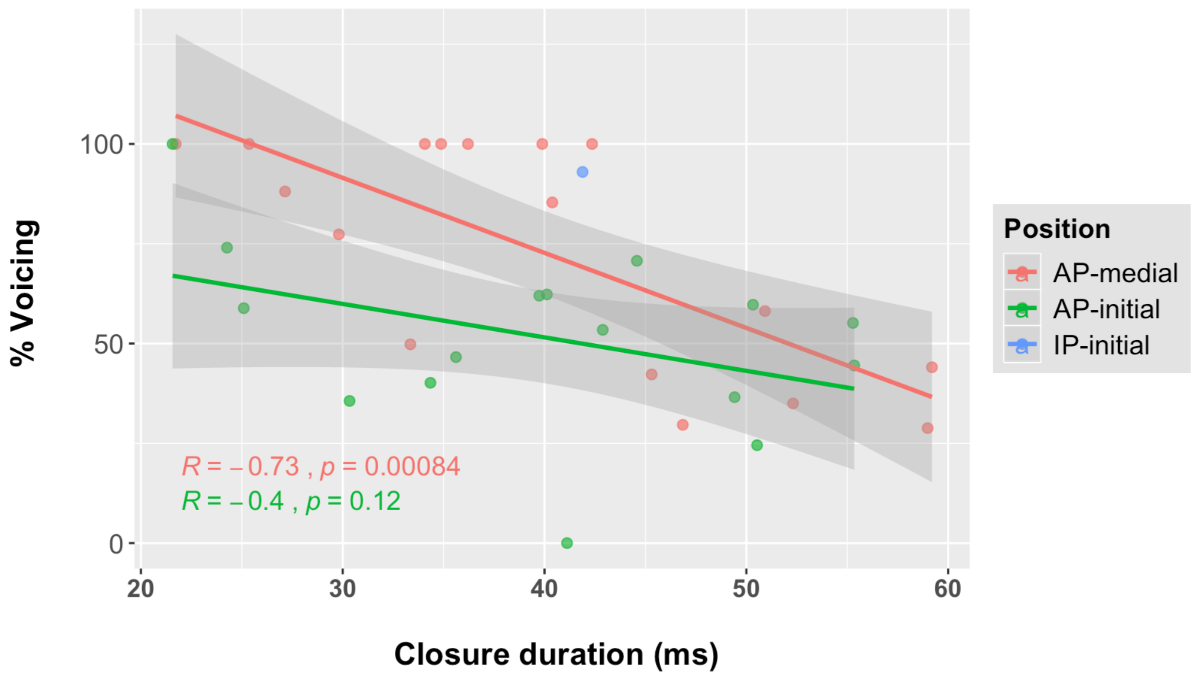


Figure H.4: % Voicing against closure duration (ms) for lenis stops produced by Seoul_4.

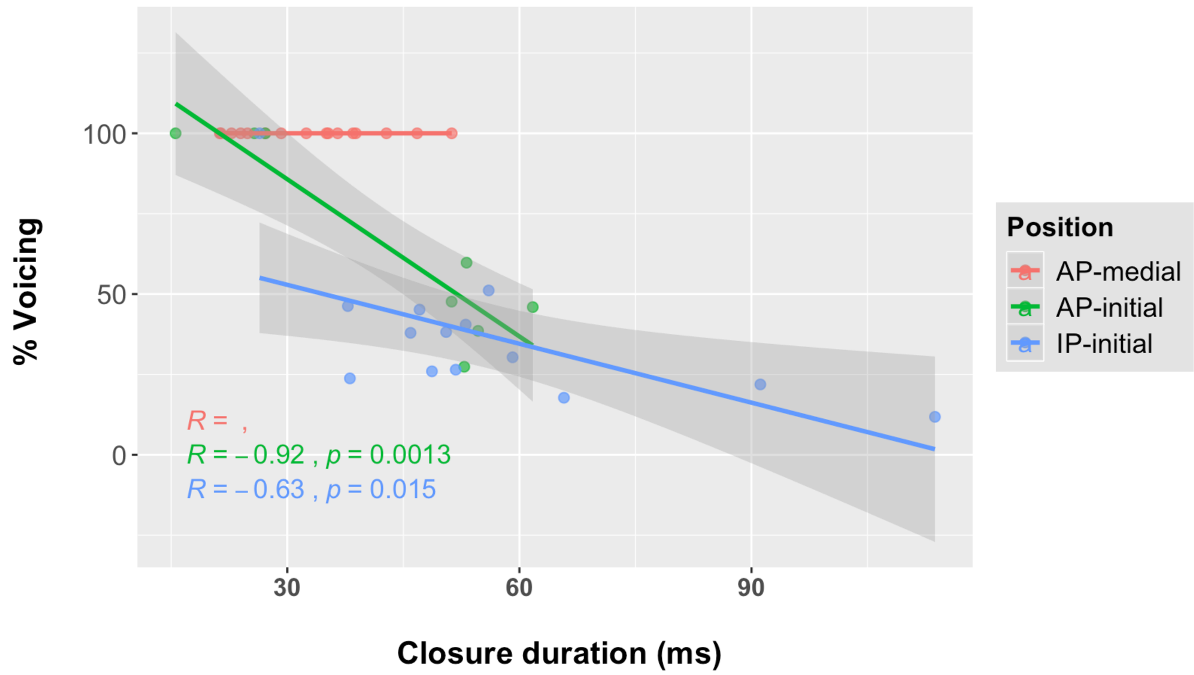


Figure H.5: % Voicing against closure duration (ms) for lenis stops produced by Seoul_5.

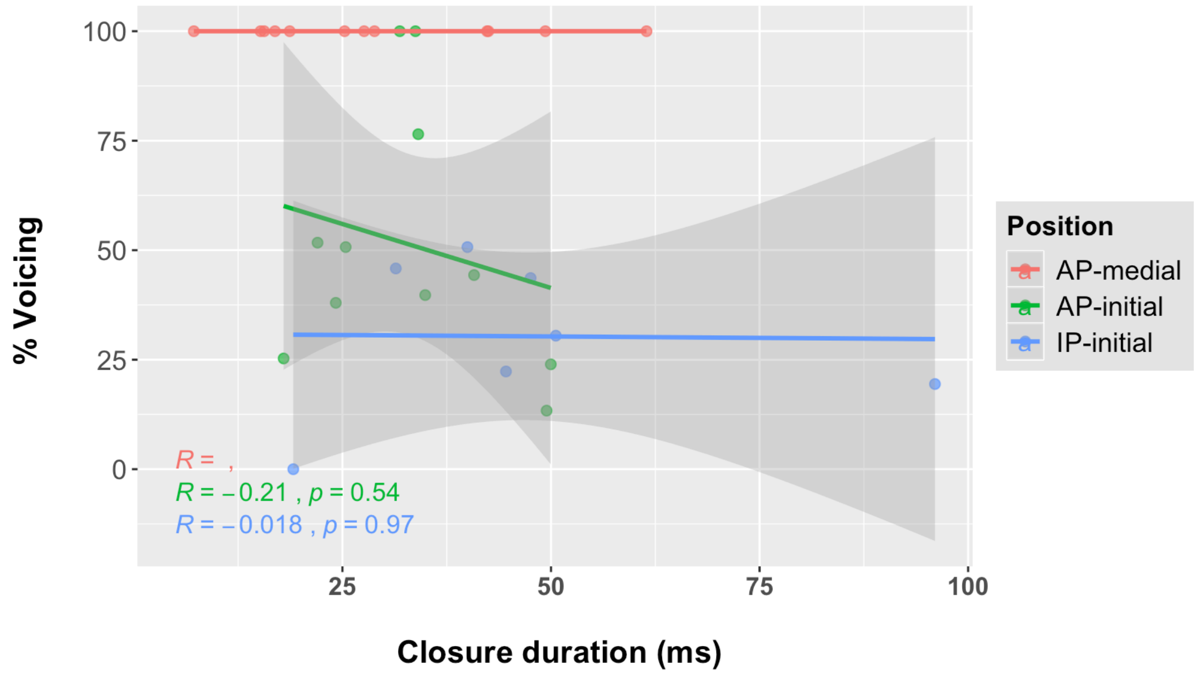


Figure H.6: % Voicing against closure duration (ms) for lenis stops produced by Seoul_6.

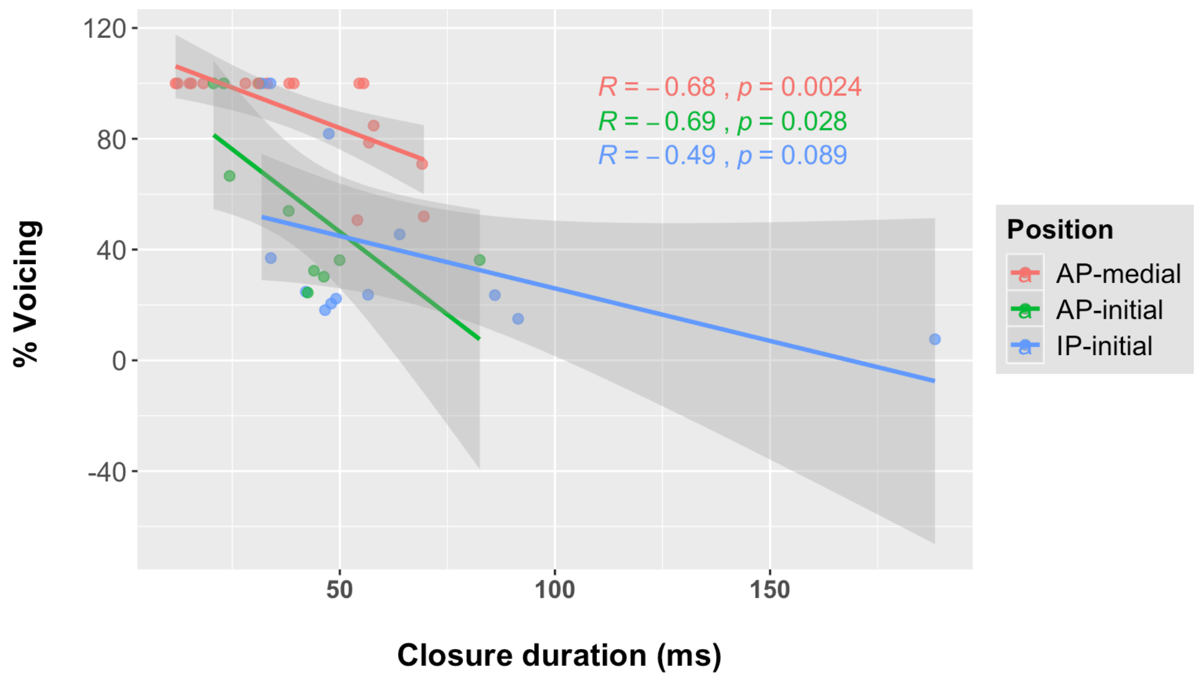


Figure H.7: % Voicing against closure duration (ms) for lenis stops produced by Seoul_7.

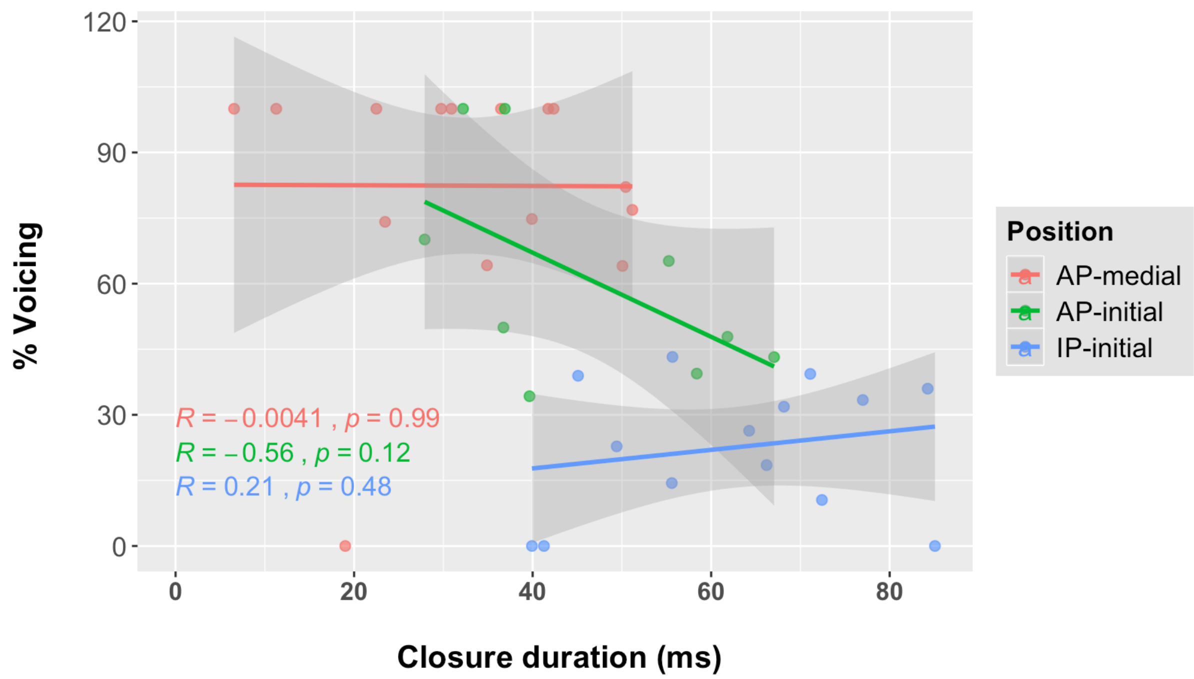


Figure H.8: % Voicing against closure duration (ms) for lenis stops produced by Seoul_8.

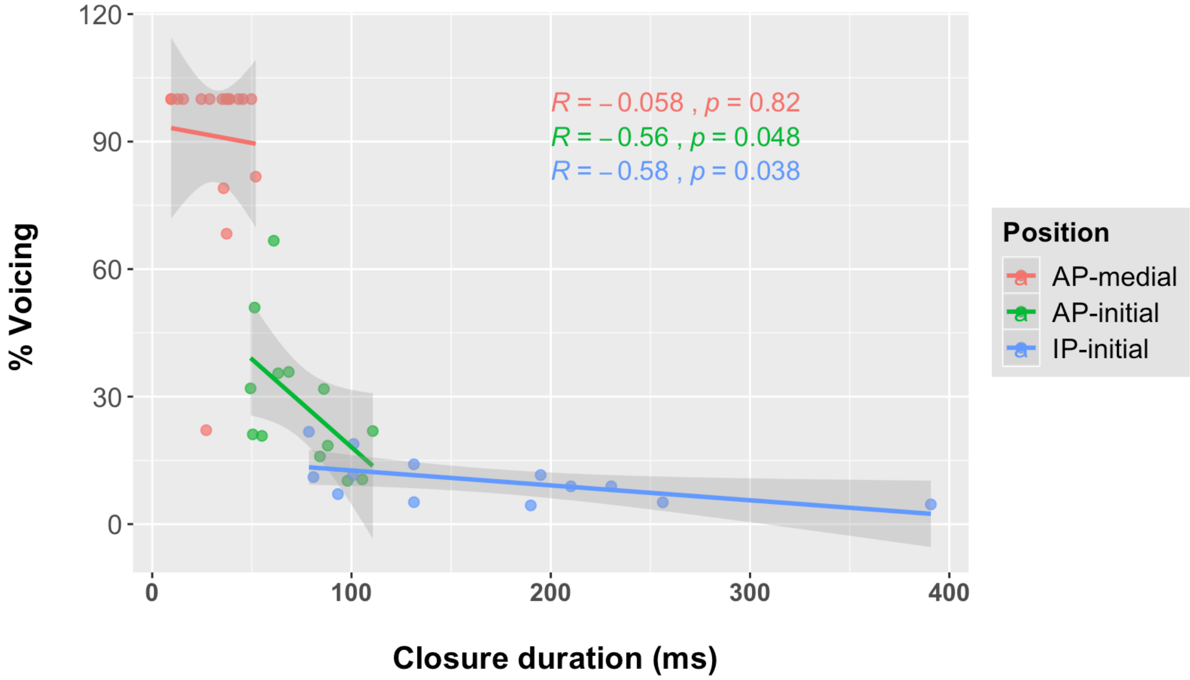


Figure H.9: % Voicing against closure duration (ms) for lenis stops produced by Seoul_9.

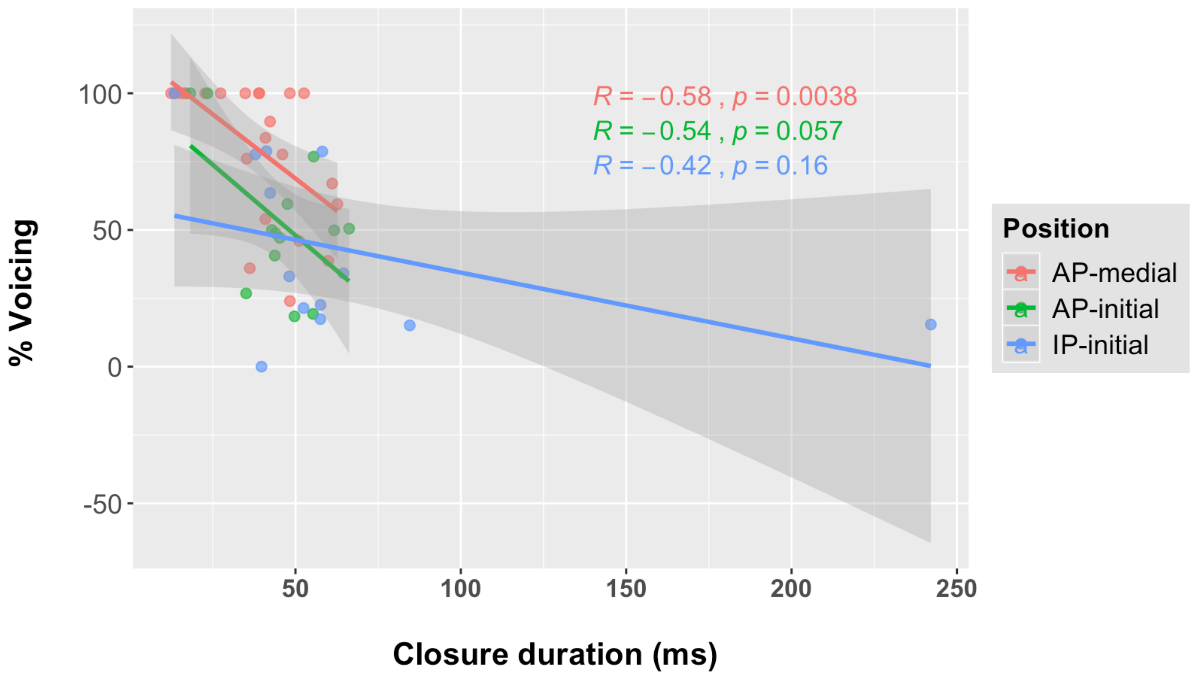


Figure H.10: % Voicing against closure duration (ms) for lenis stops produced by Seoul_10.

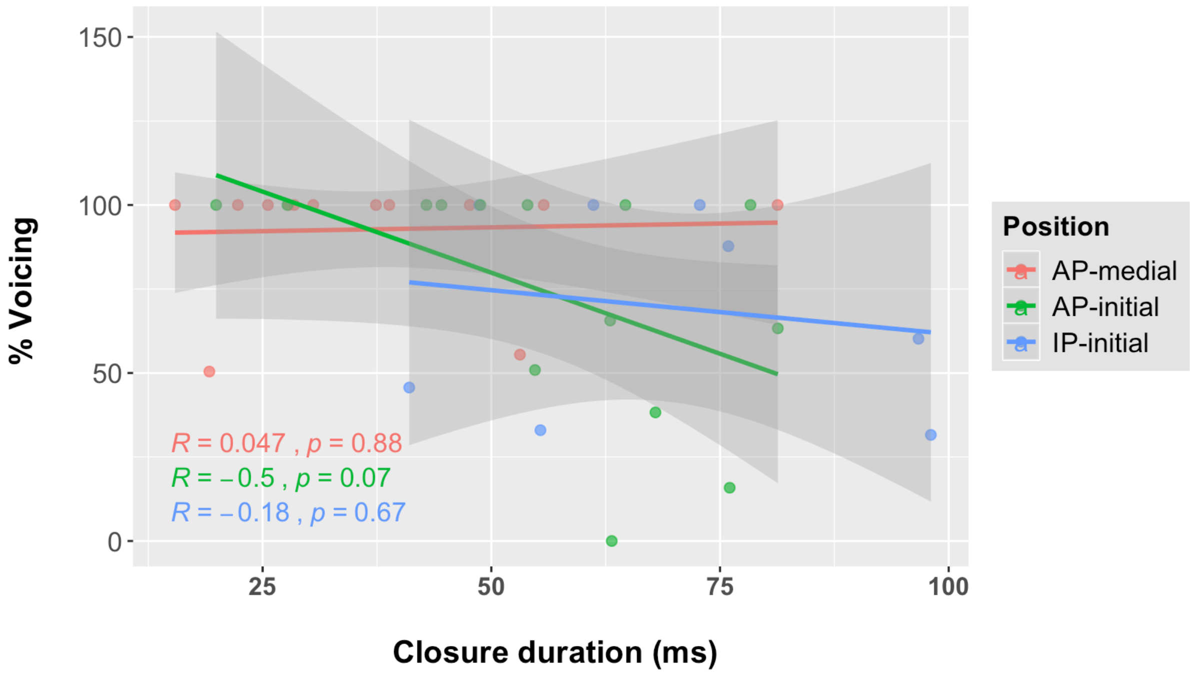


Figure H.11: % Voicing against closure duration (ms) for lenis stops produced by Busan_1.

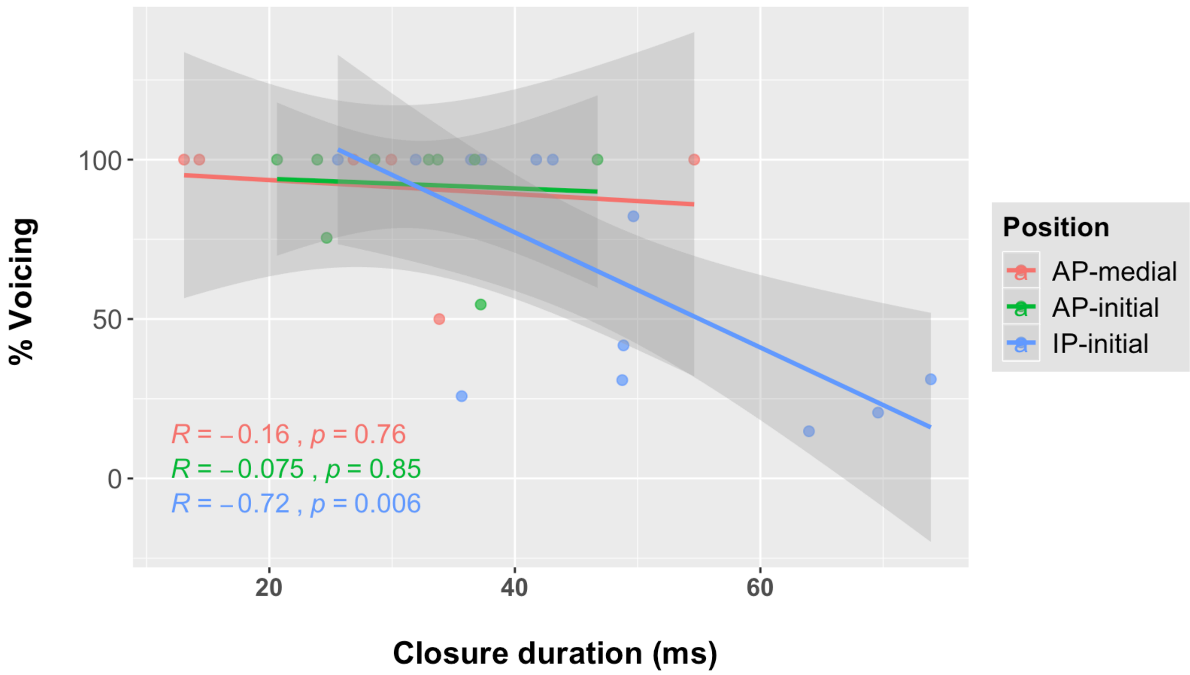


Figure H.12: % Voicing against closure duration (ms) for lenis stops produced by Busan_2.

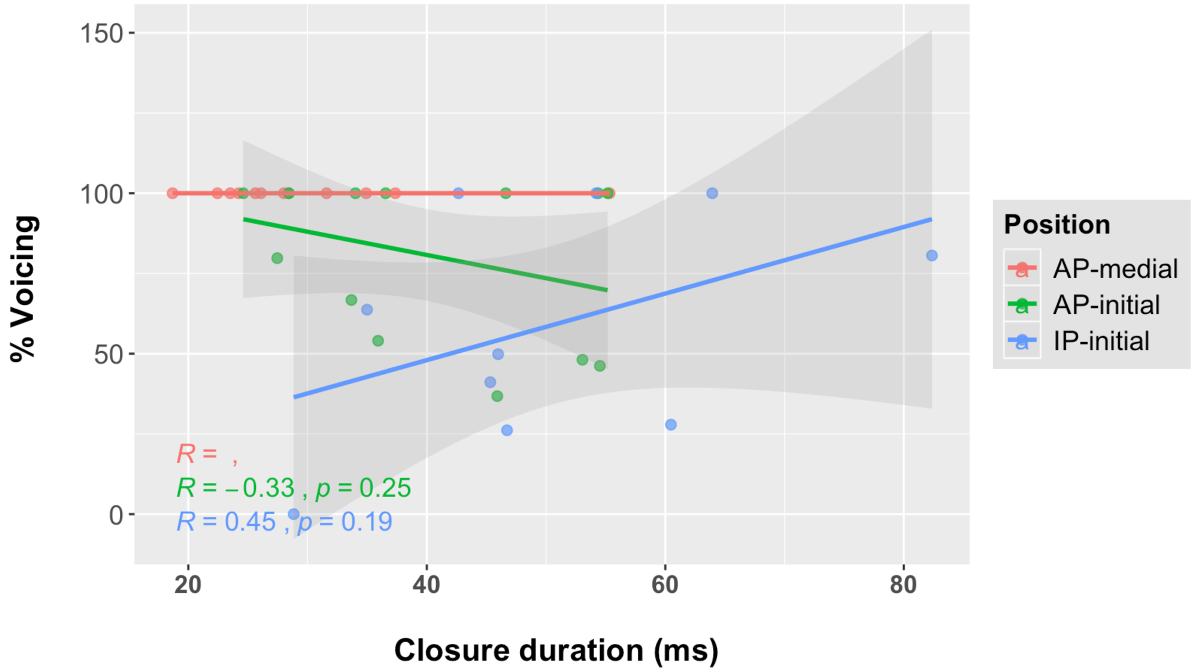


Figure H.13: % Voicing against closure duration (ms) for lenis stops produced by Busan_3.

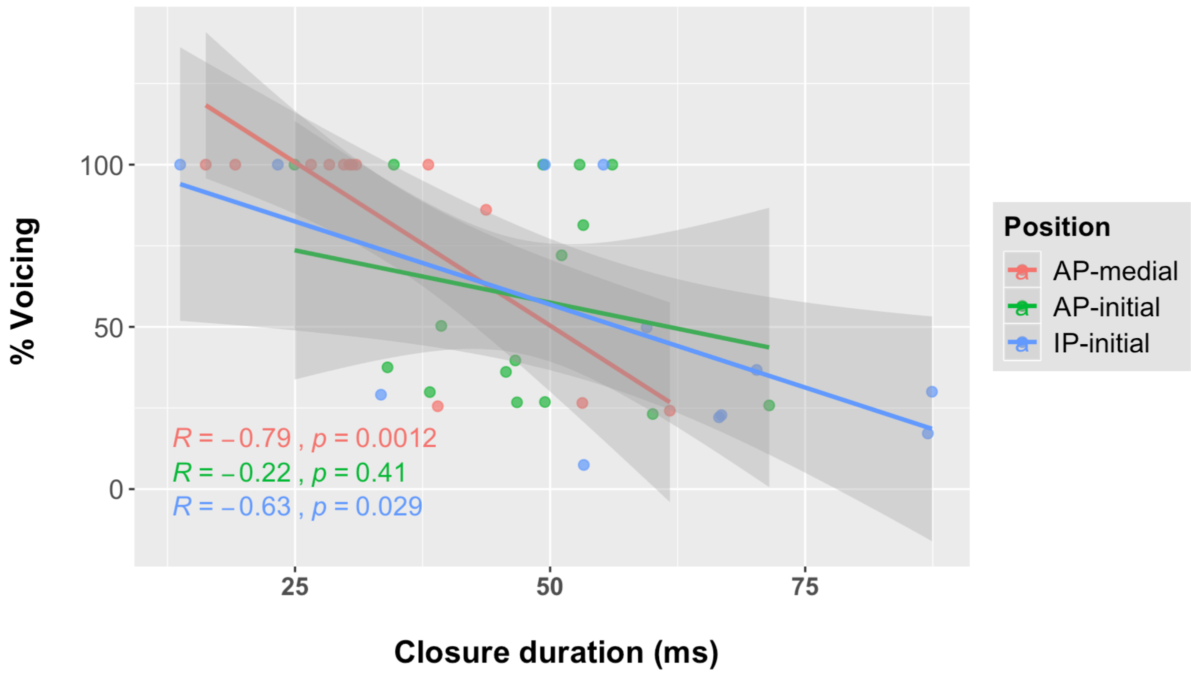


Figure H.14: % Voicing against closure duration (ms) for lenis stops produced by Busan_4.

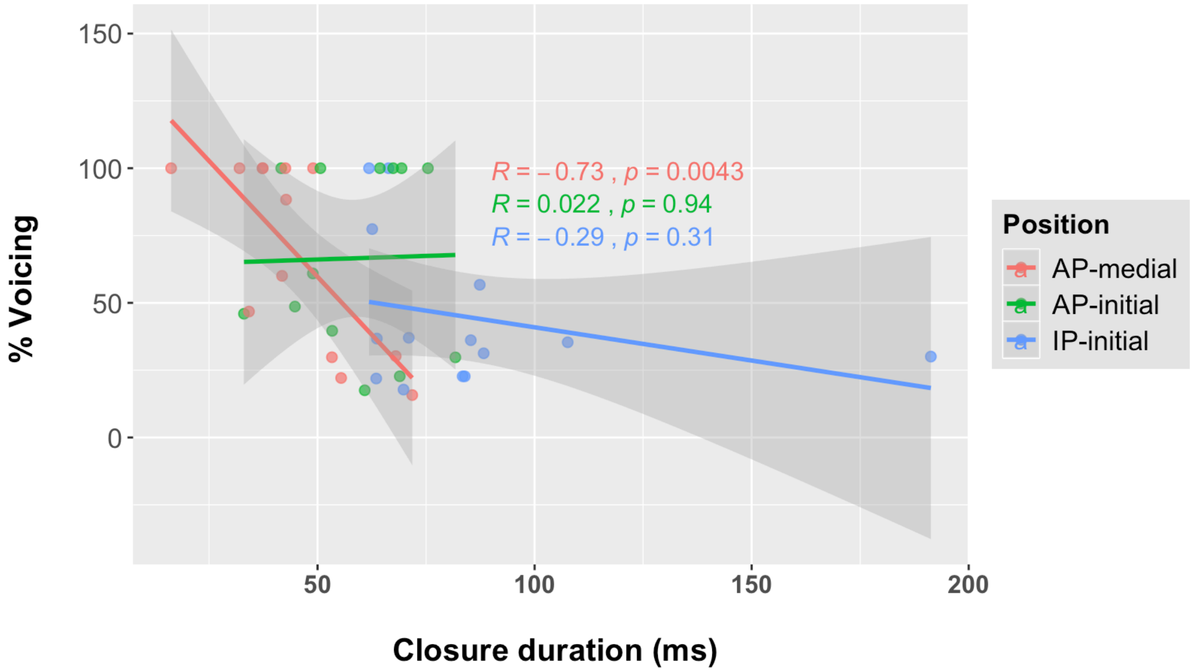


Figure H.15: % Voicing against closure duration (ms) for lenis stops produced by Busan_5.

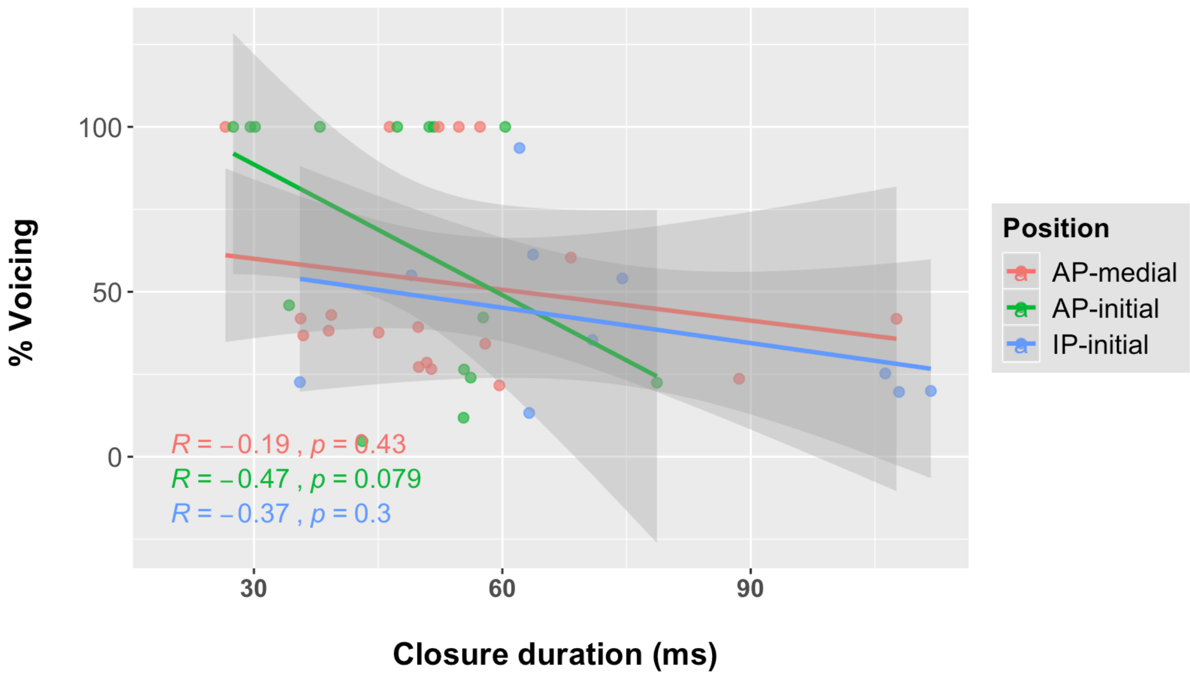


Figure H.16: % Voicing against closure duration (ms) for lenis stops produced by Busan_6.

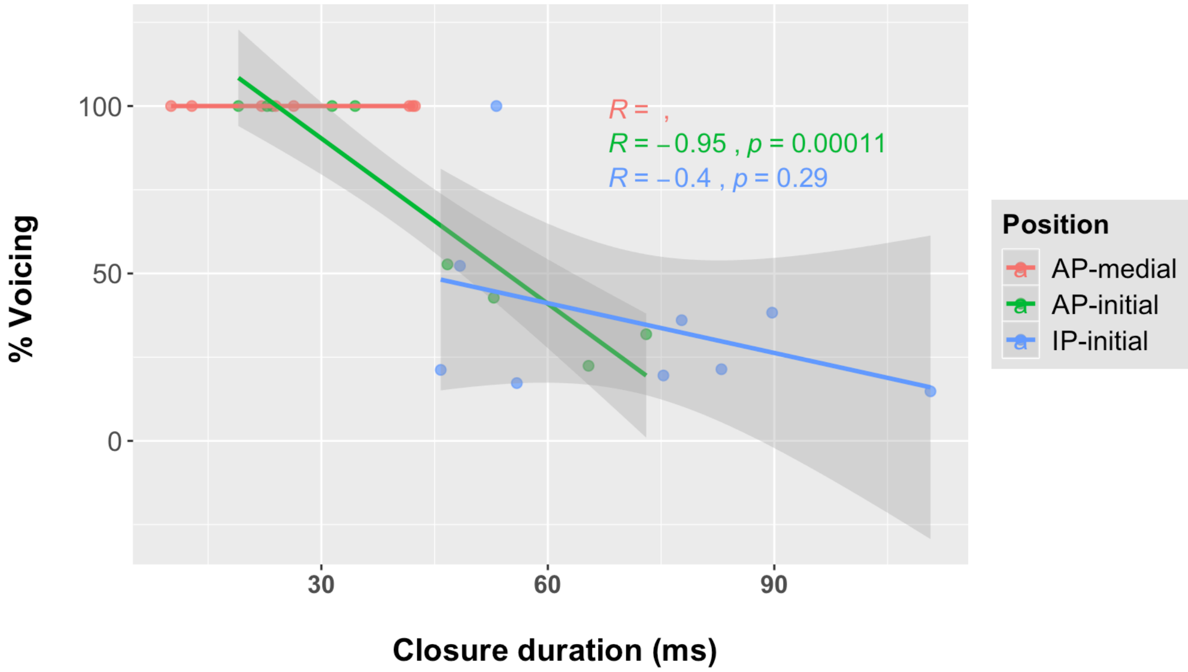


Figure H.17: % Voicing against closure duration (ms) for lenis stops produced by Busan_7.

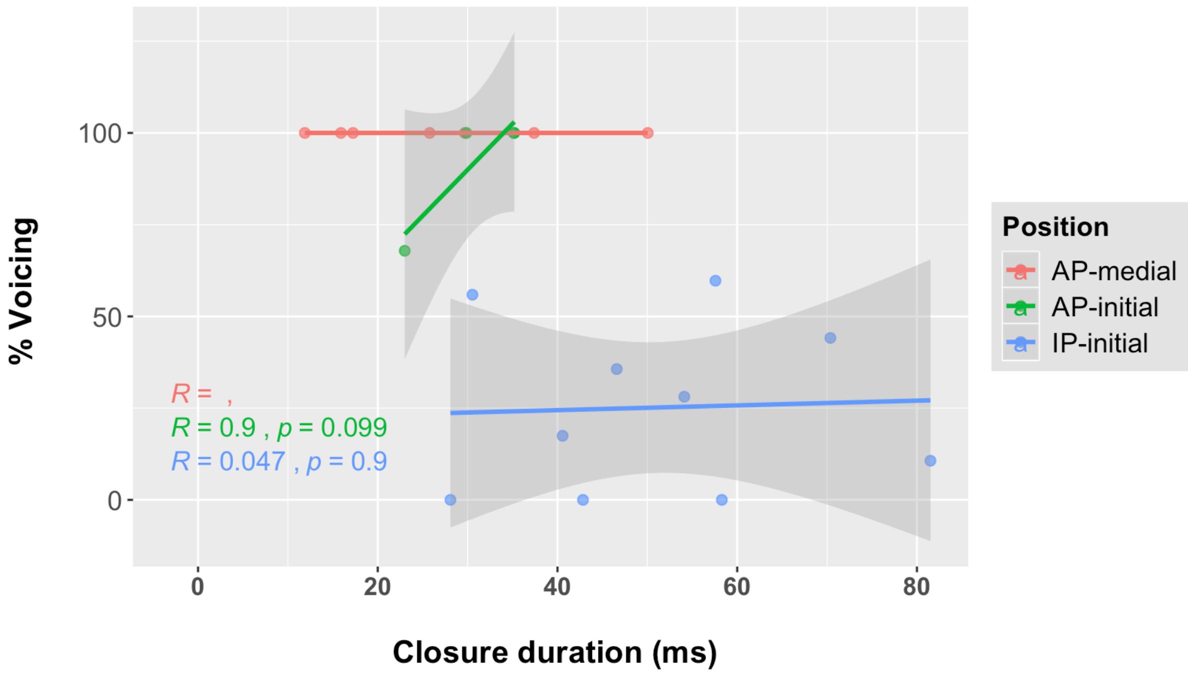


Figure H.18: % Voicing against closure duration (ms) for lenis stops produced by Busan_8.

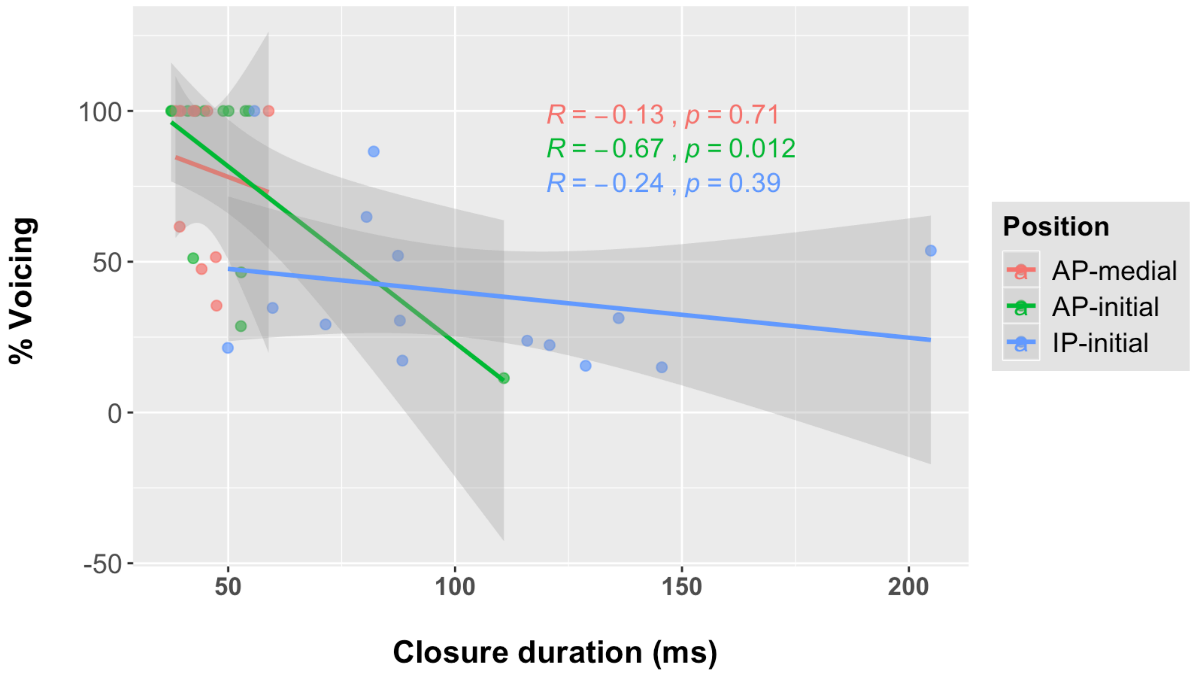


Figure H.19: % Voicing against closure duration (ms) for lenis stops produced by Busan_9.

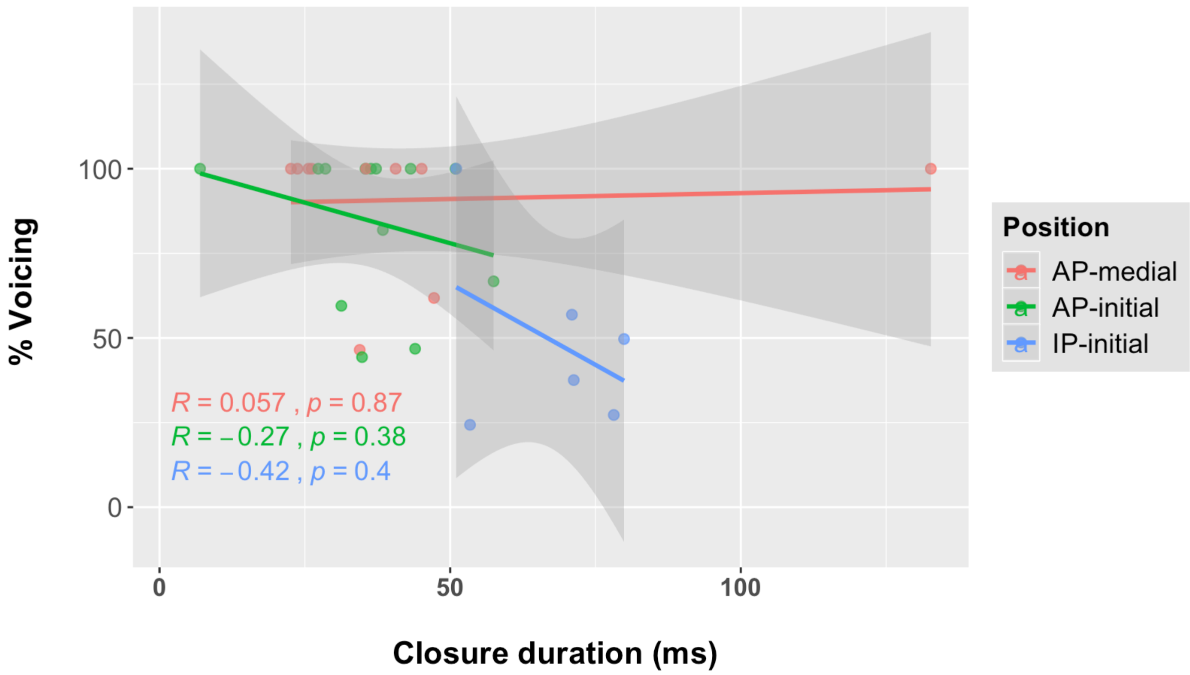


Figure H.20: % Voicing against closure duration (ms) for lenis stops produced by Busan_10.

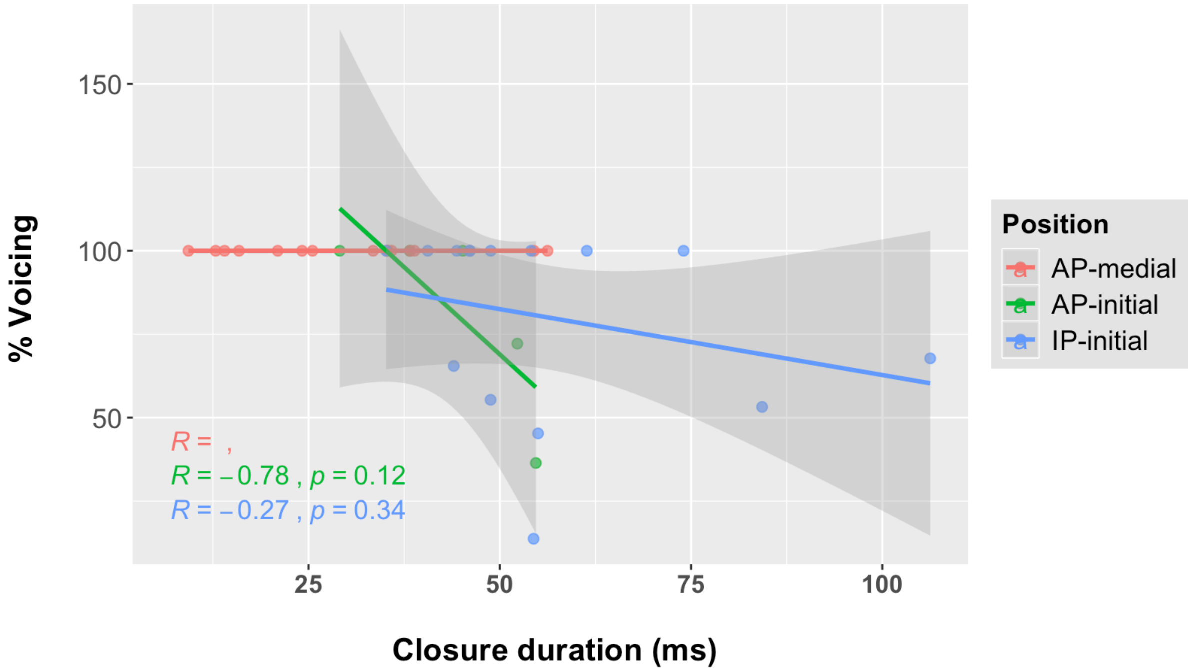


Figure H.21: % Voicing against closure duration (ms) for lenis stops produced by Ulsan_1.

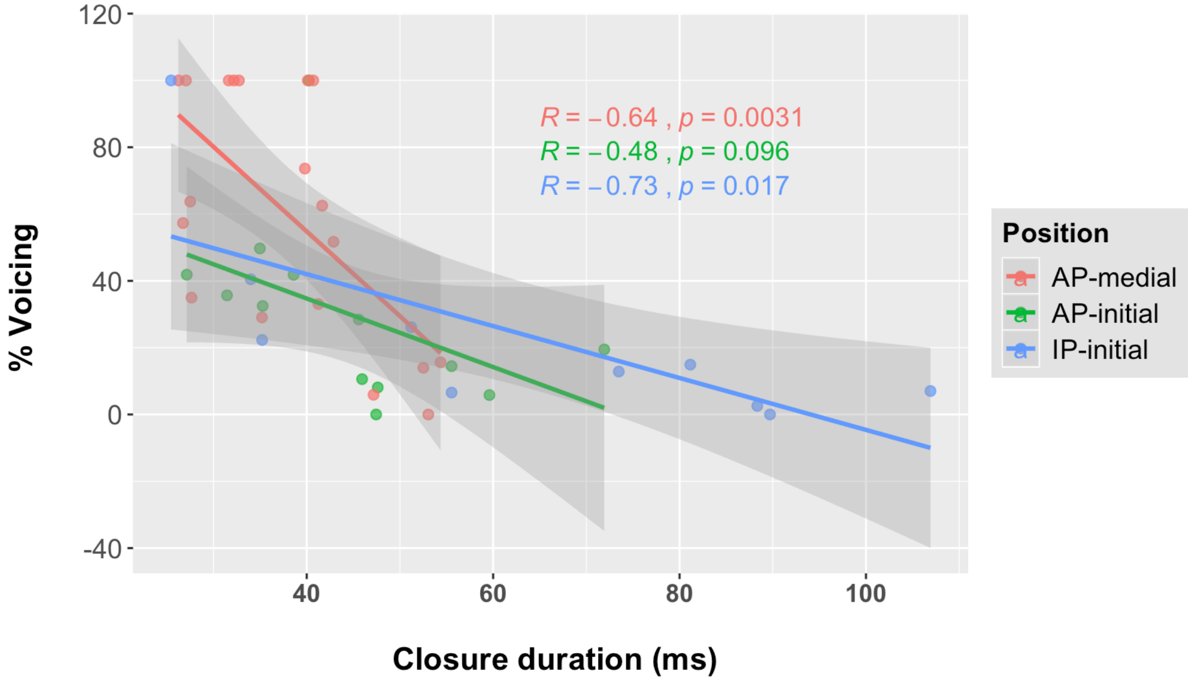


Figure H.22: % Voicing against closure duration (ms) for lenis stops produced by Ulsan_2.

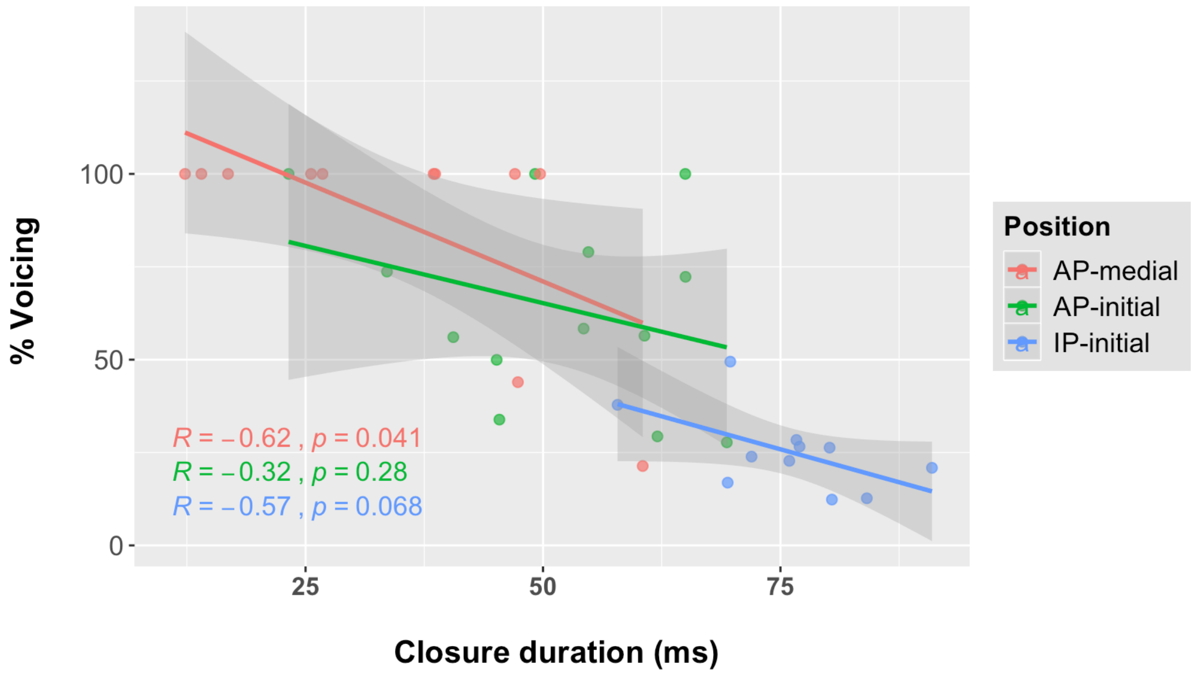


Figure H.23: % Voicing against closure duration (ms) for lenis stops produced by Ulsan_3.

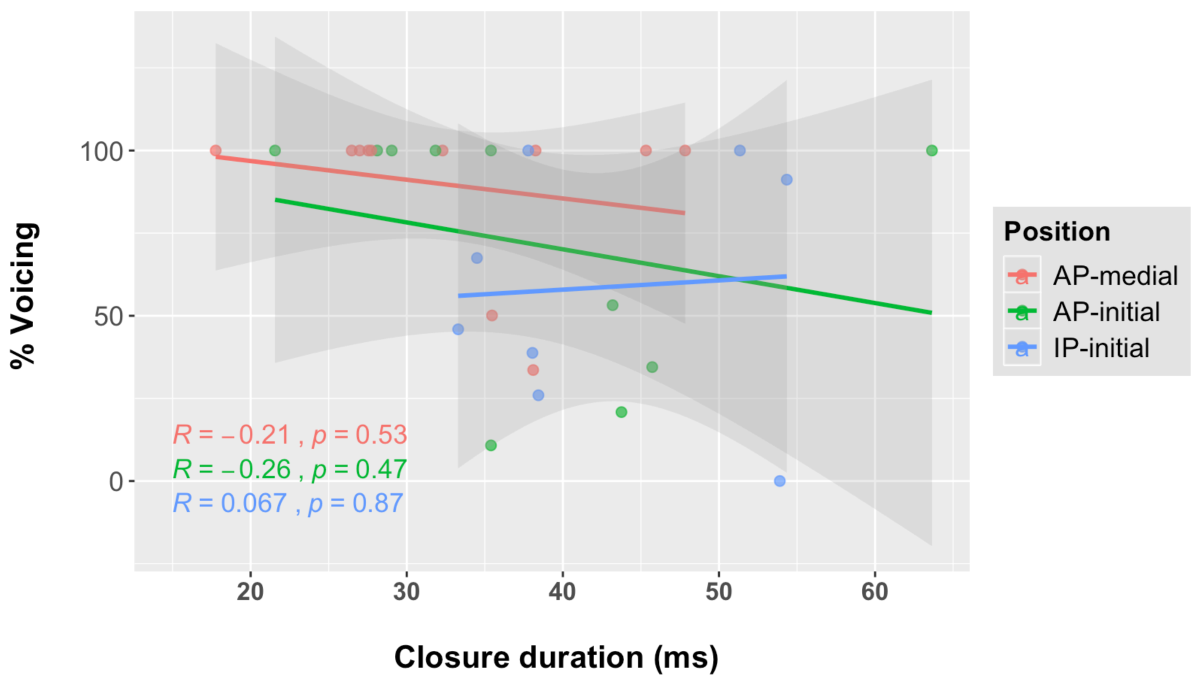


Figure H.24: % Voicing against closure duration (ms) for lenis stops produced by Ulsan_4.

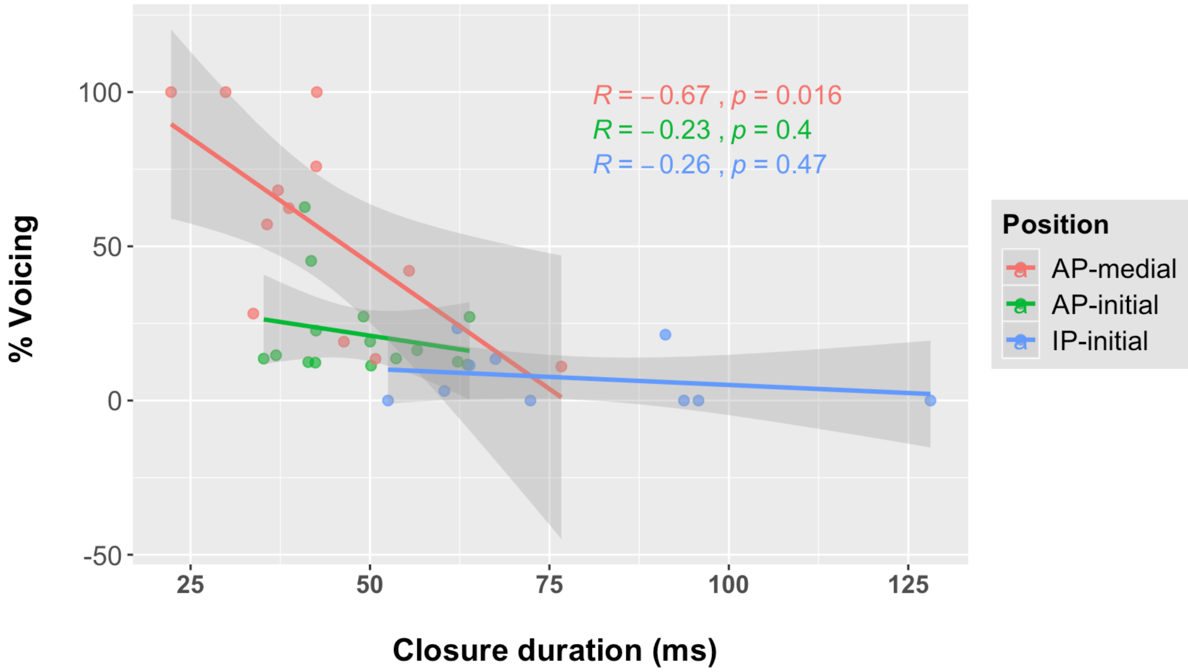


Figure H.25: % Voicing against closure duration (ms) for lenis stops produced by Ulsan_5.

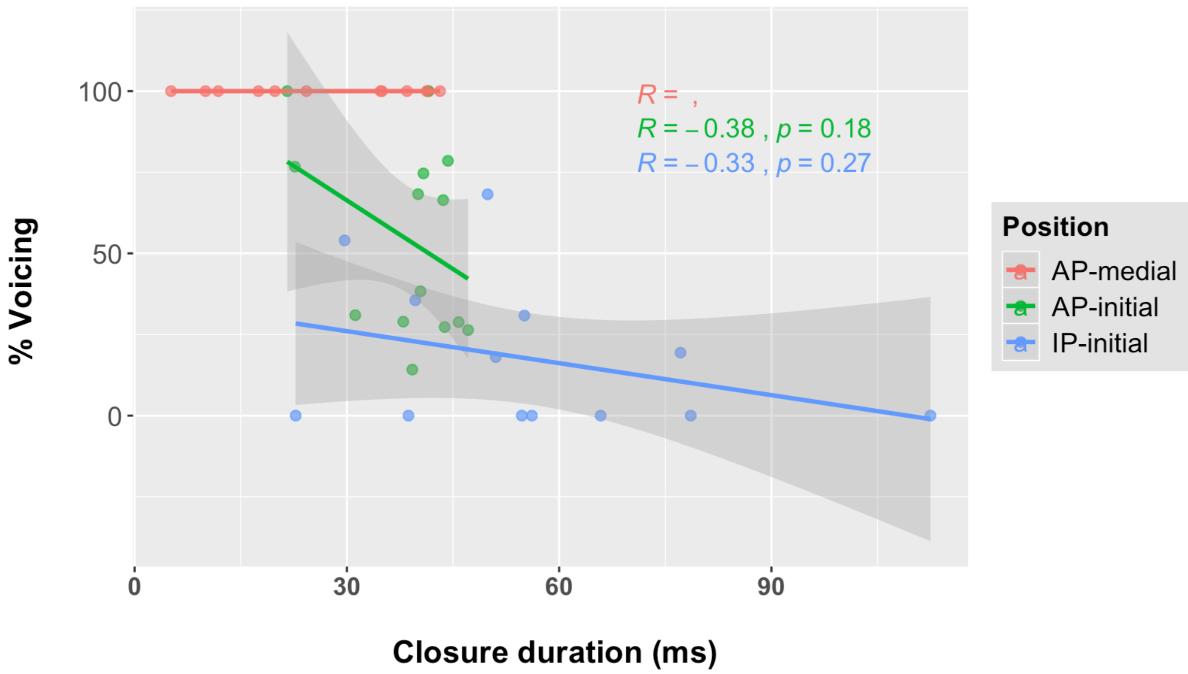


Figure H.26: % Voicing against closure duration (ms) for lenis stops produced by Ulsan_6.

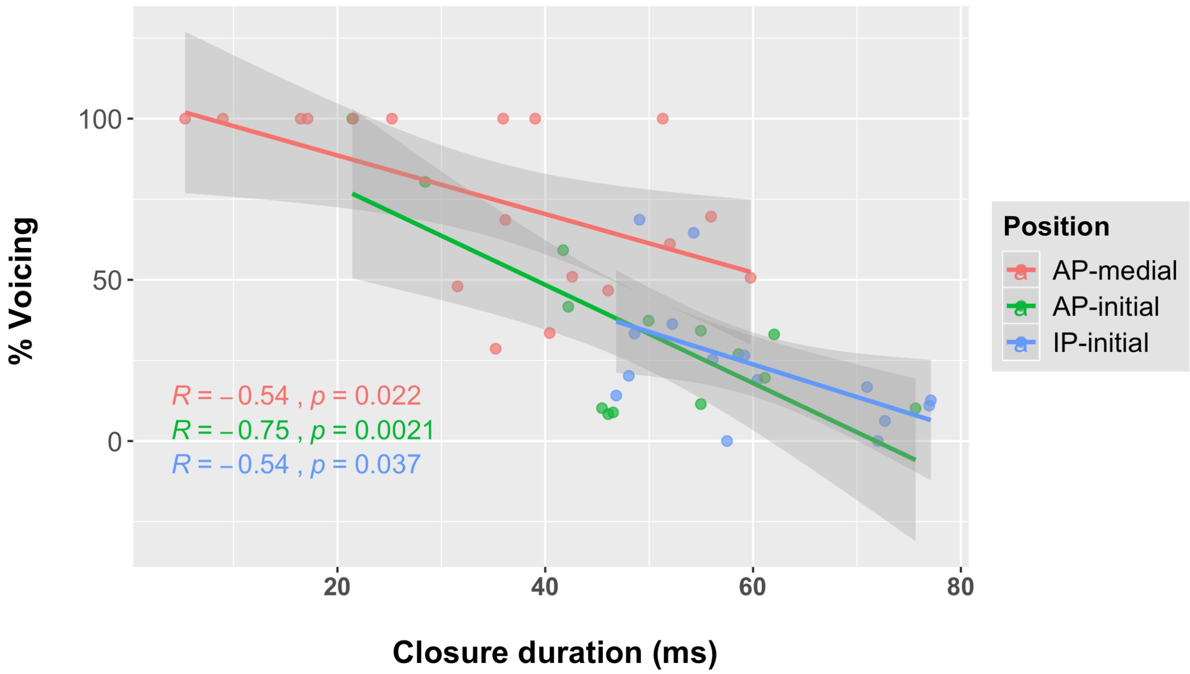


Figure H.27: % Voicing against closure duration (ms) for lenis stops produced by Ulsan_7.

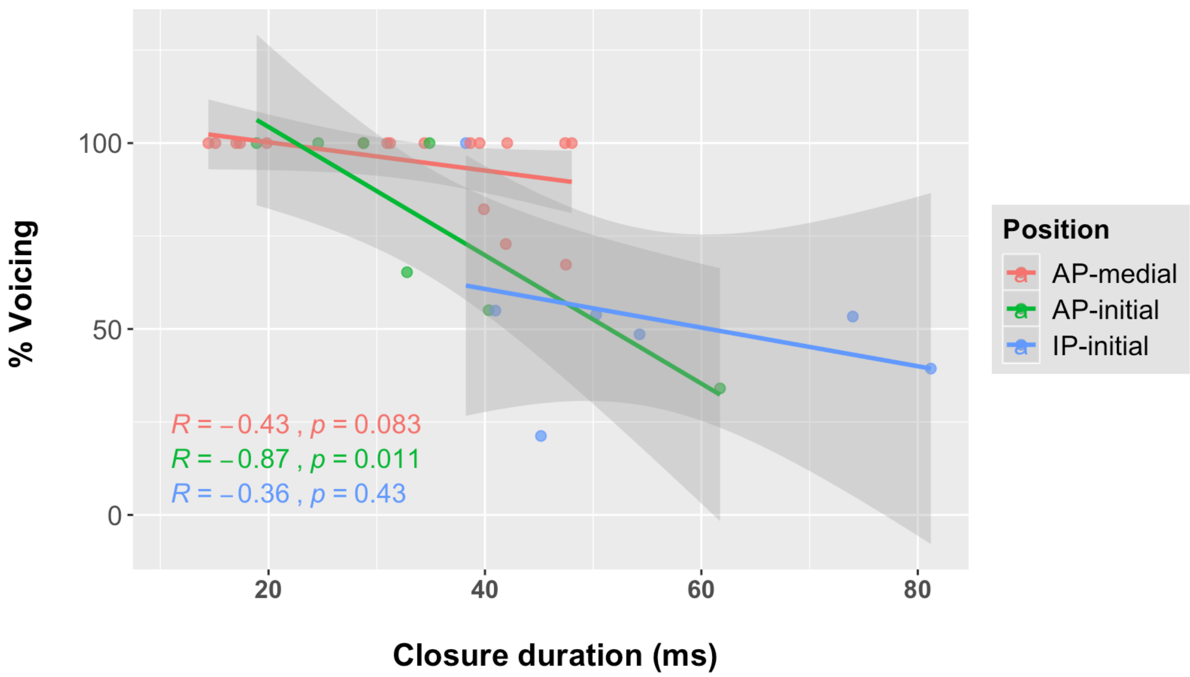


Figure H.28: % Voicing against closure duration (ms) for lenis stops produced by Ulsan_8.

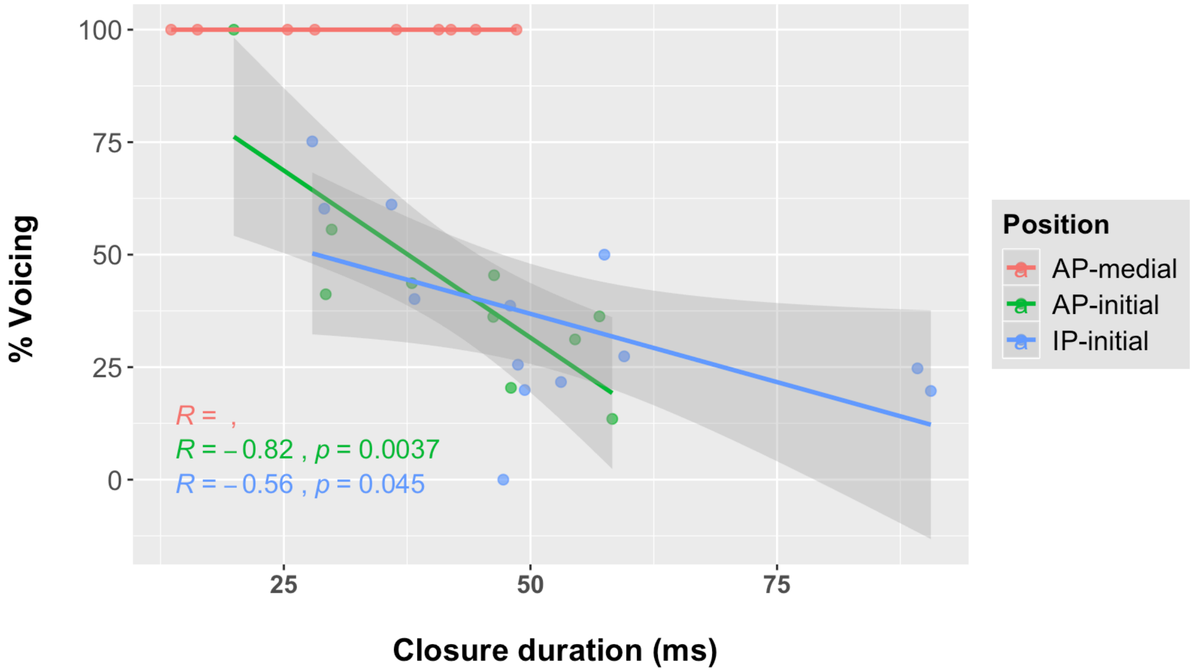


Figure H.29: % Voicing against closure duration (ms) for lenis stops produced by Ulsan_9.

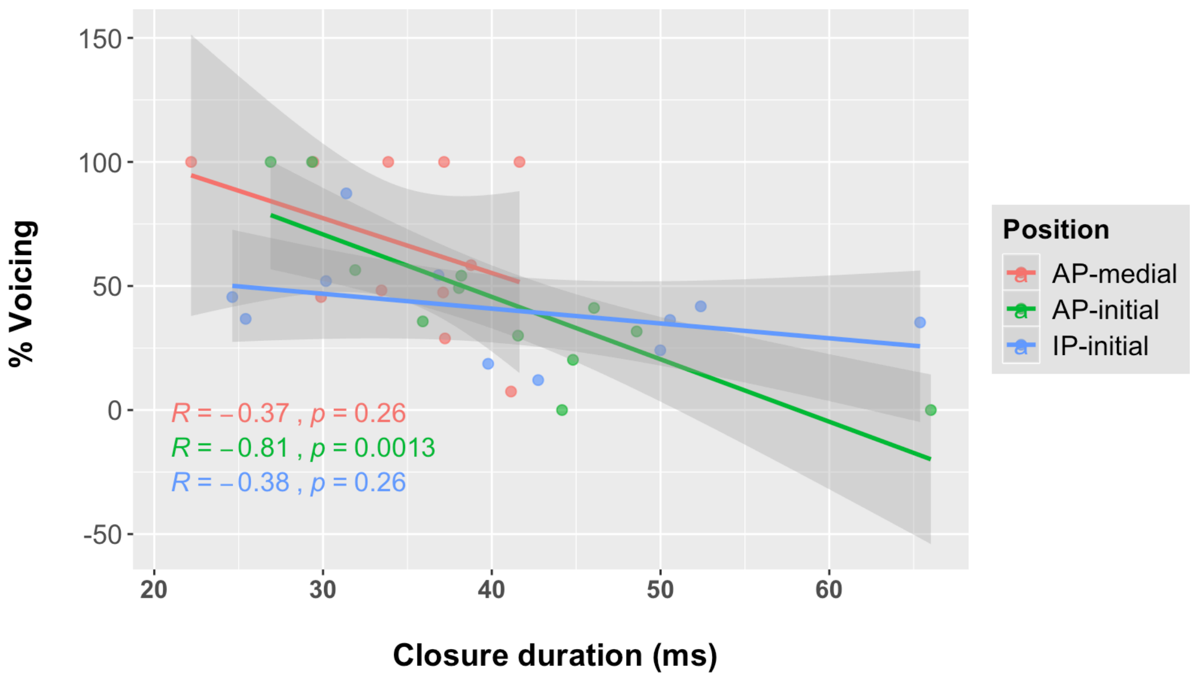


Figure H.30: % Voicing against closure duration (ms) for lenis stops produced by Ulsan_10.

II. Minimum intensity for nasal stops

Table H.4: Minimum intensity for nasal stops produced by Seoul speakers.

Subject	Position	Pearson's *r*	*p*-value	Dur. range (ms)	Token no.
Seoul_1	AP-medial	-0.15608	0.4664	33.1687	24
	AP-initial	-0.17028	0.5284	24.2337	16
	IP-initial	-0.66031	0.007379 ***	82.6836	17
Seoul_2	AP-medial	0.246002	0.2257	46.2899	26
	AP-initial	-0.10827	0.8917	19.8423	4
	IP-initial	-0.67176	0.00851 ***	109.817	14
Seoul_3	AP-medial	-0.03444	0.8566	41.0714	30
	AP-initial	-0.20794	0.4756	32.6313	14
	IP-initial	-0.75422	0.01887 **	72.2642	9
Seoul_4	AP-medial	-0.18816	0.3376	58.4606	28
	AP-initial	-0.32491	0.257	31.7704	15
	IP-initial	-0.72437	0.2756	39.8691	5
Seoul_5	AP-medial	0.01553	0.9339	38.6854	31
	AP-initial	-0.2423	0.4728	26.2483	12
	IP-initial	-0.09458	0.7089	37.7153	18
Seoul_6	AP-medial	0.028457	0.8771	43.9736	32
	AP-initial	-0.22399	0.3716	35.7065	18
	IP-initial	-0.87763	0.000844 ***	52.9218	11
Seoul_7	AP-medial	0.046399	0.8335	34.0264	23
	AP-initial	0.005049	0.9882	38.7427	11
	IP-initial	-0.93522	3.16E-07 ***	248.98	15
Seoul_8	AP-medial	0.060245	0.7953	29.9051	21
	AP-initial	-0.84098	0.00118 ***	21.4503	11
	IP-initial	-0.66968	0.003274 ***	60.7921	17
Seoul_9	AP-medial	-0.36195	0.05368 *	31.588	30
	AP-initial	-0.40072	0.222	196.133	17
	IP-initial	-0.83732	0.004857 ***	256.209	14
Seoul_10	AP-medial	0.155429	0.3801	57.9647	34
	AP-initial	-0.44073	0.1516	32.6434	12
	IP-initial	-0.81034	8.01E-05 ***	196.753	18

Table H.5: Minimum intensity for nasal stops produced by Busan speakers.

Subject	Position	Pearson's *r*	*p*-value	Dur. range (ms)	Token no.
Busan_1	AP-medial	0.331987	0.07308 *	46.1827	30
	AP-initial	-0.01702	0.9466	50.2059	18
	IP-initial	-0.31704	0.4884	53.8463	7
Busan_2	AP-medial	0.106486	0.7292	41.2534	13
	AP-initial	0.063243	0.8622	40.0949	10
	IP-initial	-0.48838	0.07641 *	38.1357	14
Busan_3	AP-internal	0.089405	0.6574	49.0716	27
	AP-initial	0.509387	0.1095	48.4634	11
	IP-initial	-0.6037	0.06459 *	56.9983	10
Busan_4	AP-internal	-0.0911	0.6513	29.1988	27
	AP-initial	0.193387	0.442	76.6595	18
	IP-initial	-0.15098	0.6064	48.2568	14
Busan_5	AP-internal	-0.08136	0.7055	40.9371	24
	AP-initial	-0.75036	0.000521 ***	61.8893	17
	IP-initial	-0.67313	0.003061 ***	77.0156	18
Busan_6	AP-internal	-0.39343	0.04676 **	31.8596	26
	AP-initial	0.487899	0.04694 **	47.3196	17
	IP-initial	-0.65761	0.01457 **	61.9174	15
Busan_7	AP-internal	-0.56431	0.02842 **	34.2108	15
	AP-initial	-0.32833	0.2144	35.976	16
	IP-initial	-0.66379	0.005049 ***	140.634	16
Busan_8	AP-internal	-0.50529	0.01645 **	36.0937	22
	AP-initial	-0.61151	0.1072	10.23	8
	IP-initial	-0.54187	0.03015 **	79.0713	16
Busan_9	AP-internal	0.348726	0.05045 *	73.0676	32
	AP-initial	-0.53638	0.03927 **	69.6611	15
	IP-initial	0.525262	0.09707 *	118.966	17
Busan_10	AP-internal	0.397331	0.09208 *	72.6301	19
	AP-initial	-0.33094	0.2106	44.4032	16
	IP-initial	*not enough finite observations*		18.0184	2

Table H.6: Minimum intensity for nasal stops produced by Ulsan speakers.

Subject	**Position**	**Pearson's *r***	***p*-value**	**Dur. range (ms)**	**Token no.**
Ulsan_1	AP-internal	0.255587	0.2175	35.3379	25
	AP-initial	-0.81203	0.188	16.4416	4
	IP-initial	-0.56944	0.08575 *	44.3942	10
Ulsan_2	AP-internal	0.167243	0.5071	51.2536	18
	AP-initial	-0.03633	0.8862	27.3817	18
	IP-initial	-0.59264	0.04228 **	57.2151	13
Ulsan_3	AP-internal	-0.02973	0.8878	56.3995	25
	AP-initial	-0.62575	0.01259 **	38.1308	15
	IP-initial	-0.78658	0.001425 ***	101.011	13
Ulsan_4	AP-internal	0.089563	0.6379	50.2232	30
	AP-initial	0.218066	0.4539	34.4539	14
	IP-initial	-0.60026	0.2078	81.1212	6
Ulsan_5	AP-internal	-0.12091	0.5648	61.4827	25
	AP-initial	0.206415	0.3693	69.419	21
	IP-initial	0.045461	0.9075	80.719	9
Ulsan_6	AP-internal	-0.01601	0.9466	35.0665	20
	AP-initial	-0.52627	0.04388 **	38.7404	15
	IP-initial	-0.92368	1.22E-07 ***	87.5185	17
Ulsan_7	AP-internal	-0.16077	0.4863	37.48	21
	AP-initial	0.082507	0.7529	44.4587	17
	IP-initial	-0.47216	0.08825 *	56.7486	17
Ulsan_8	AP-internal	0.385063	0.04732 **	45.9997	27
	AP-initial	0.11275	0.7012	33.9735	14
	IP-initial	-0.01646	0.9595	53.9537	12
Ulsan_9	AP-internal	0.137323	0.4946	33.9068	27
	AP-initial	0.088525	0.7844	35.636	12
	IP-initial	-0.72197	0.005327 ***	75.5016	13
Ulsan_10	AP-internal	-0.42388	0.03472 **	29.2442	25
	AP-initial	-0.54745	0.03466 **	55.7611	15
	IP-initial	-0.64401	0.03247 **	44.7013	11

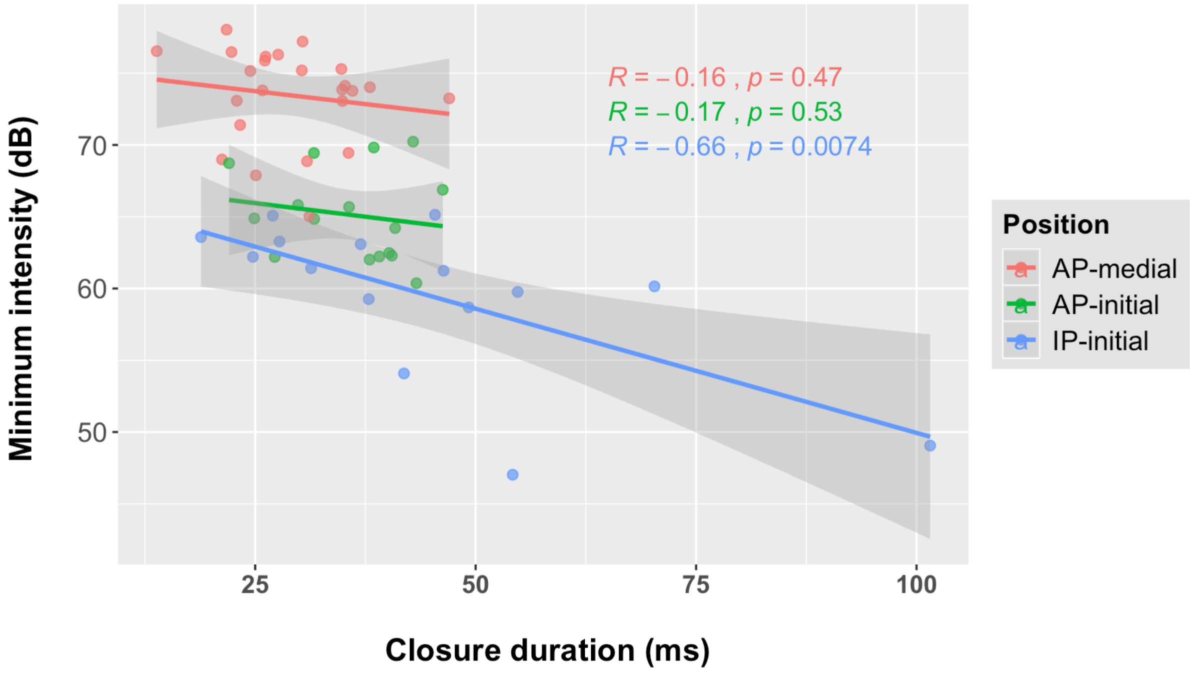


Figure H.31: Minimum intensity (dB) against closure duration (ms) for nasal stops by Seoul_1.

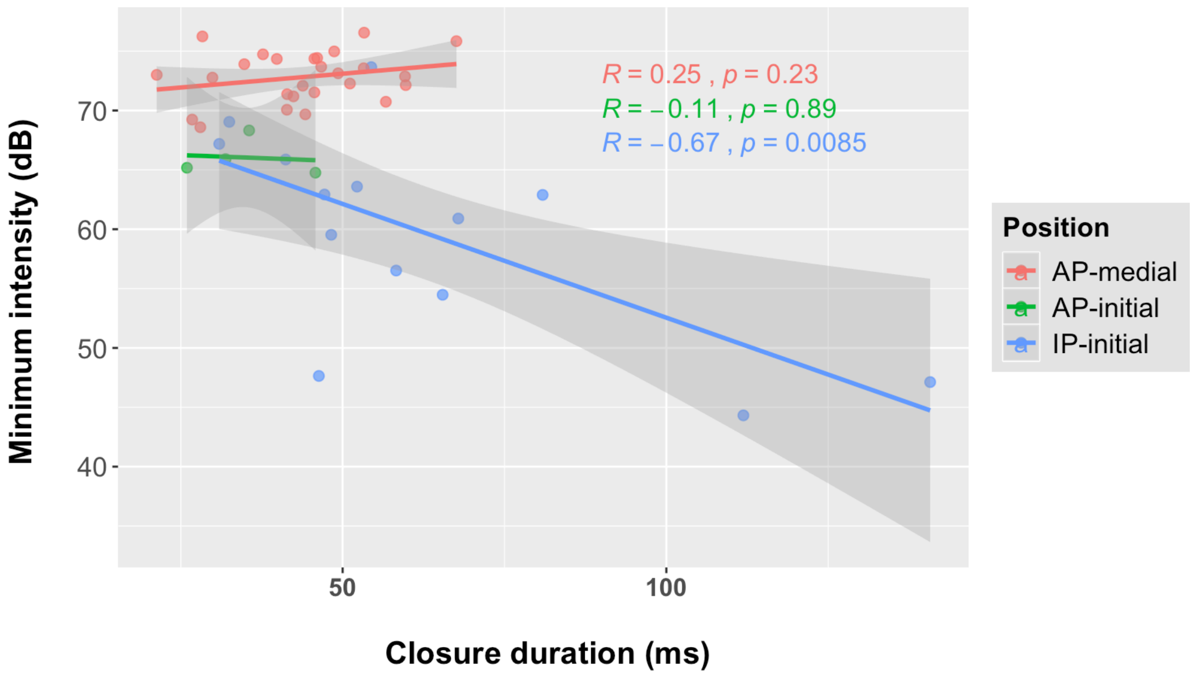


Figure H.32: Minimum intensity (dB) against closure duration (ms) for nasal stops by Seoul_2.

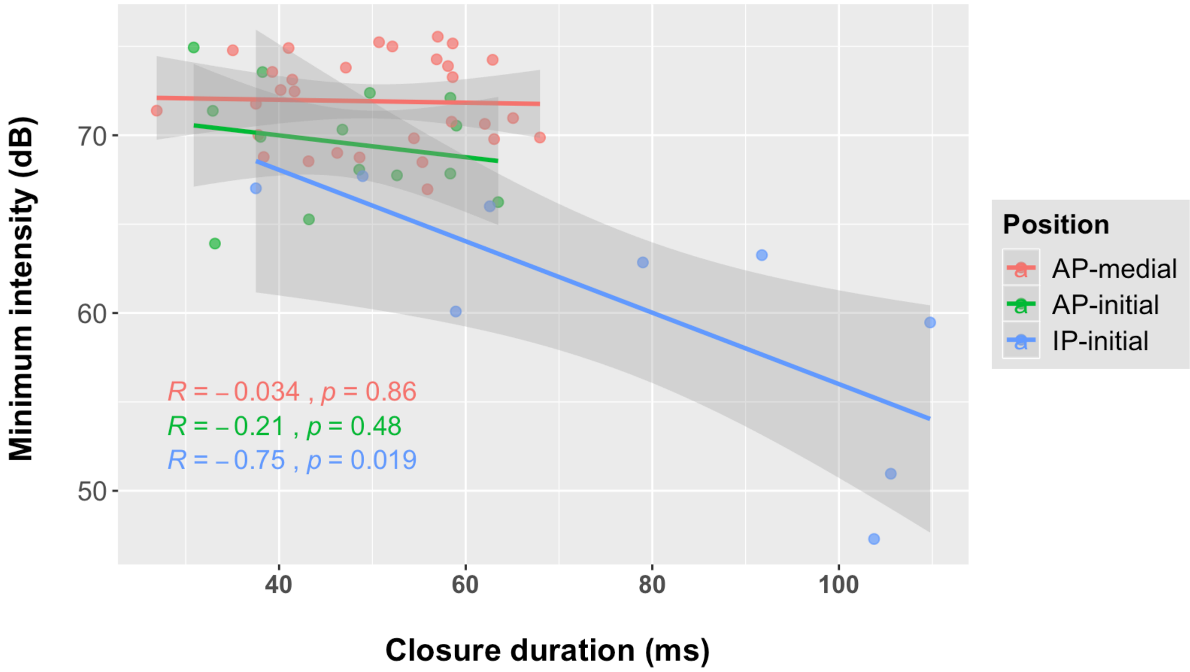


Figure H.33: Minimum intensity (dB) against closure duration (ms) for nasal stops by Seoul_3.

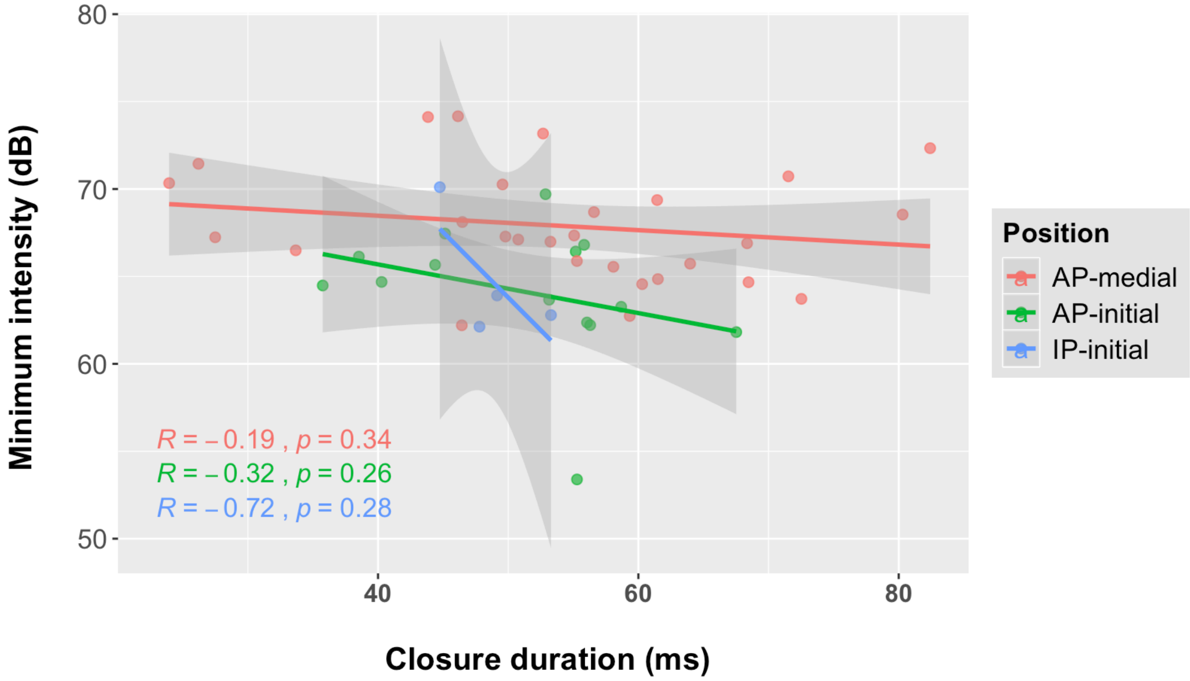


Figure H.34: Minimum intensity (dB) against closure duration (ms) for nasal stops by Seoul_4.

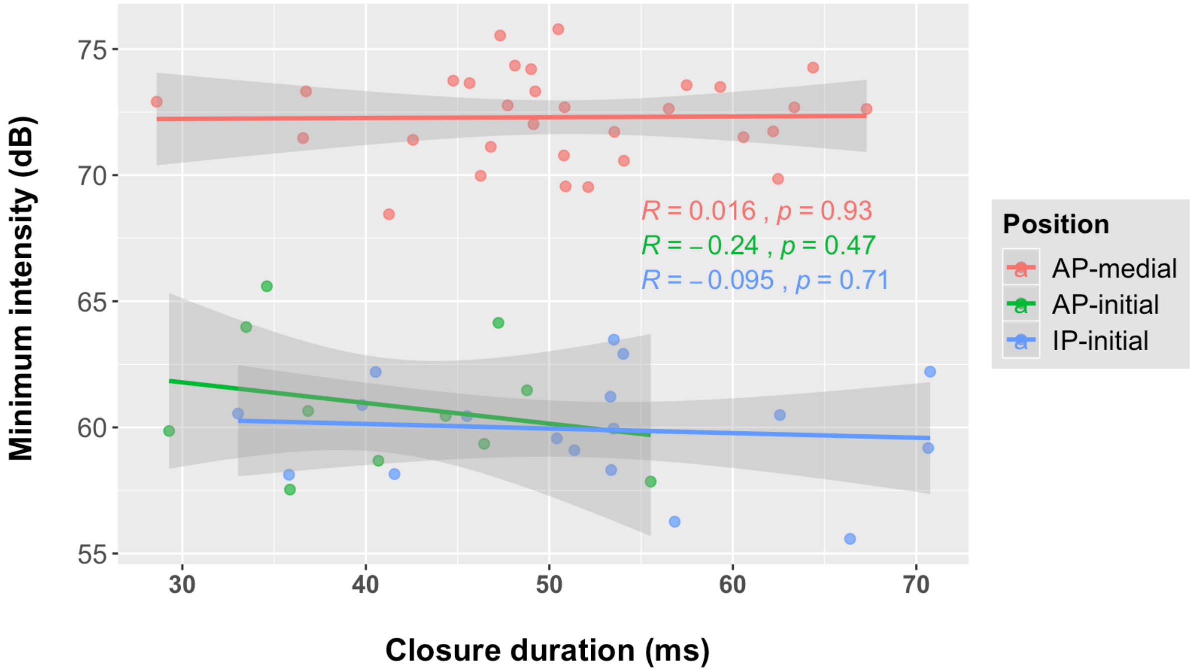


Figure H.35: Minimum intensity (dB) against closure duration (ms) for nasal stops by Seoul_5.

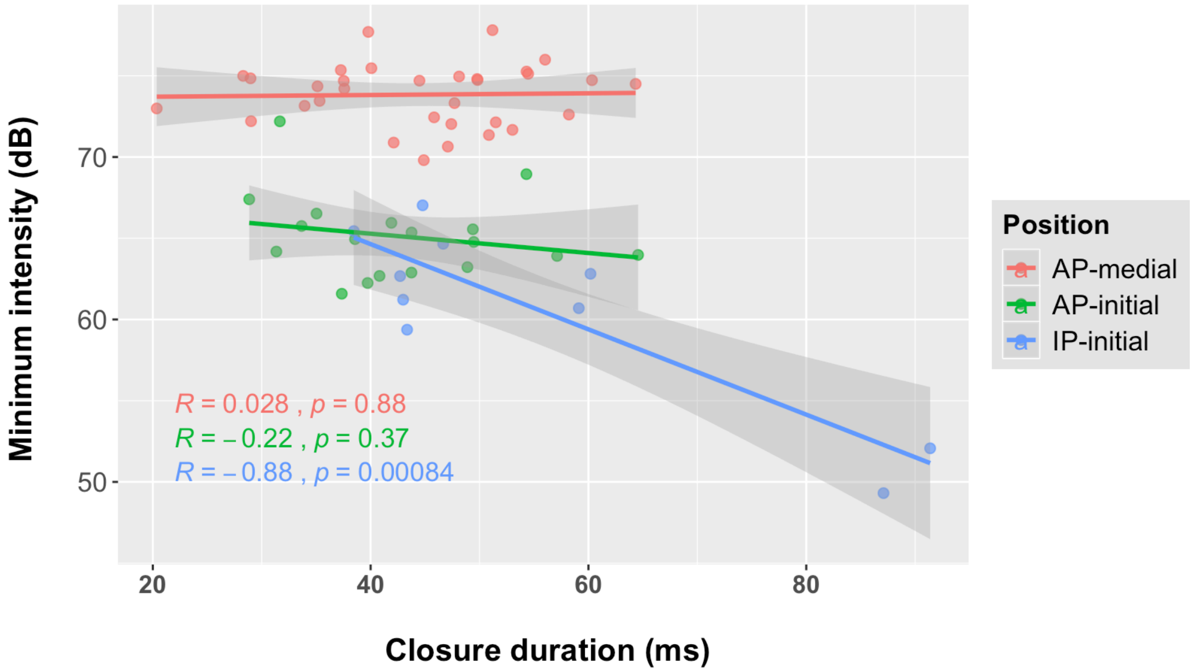


Figure H.36: Minimum intensity (dB) against closure duration (ms) for nasal stops by Seoul_6.

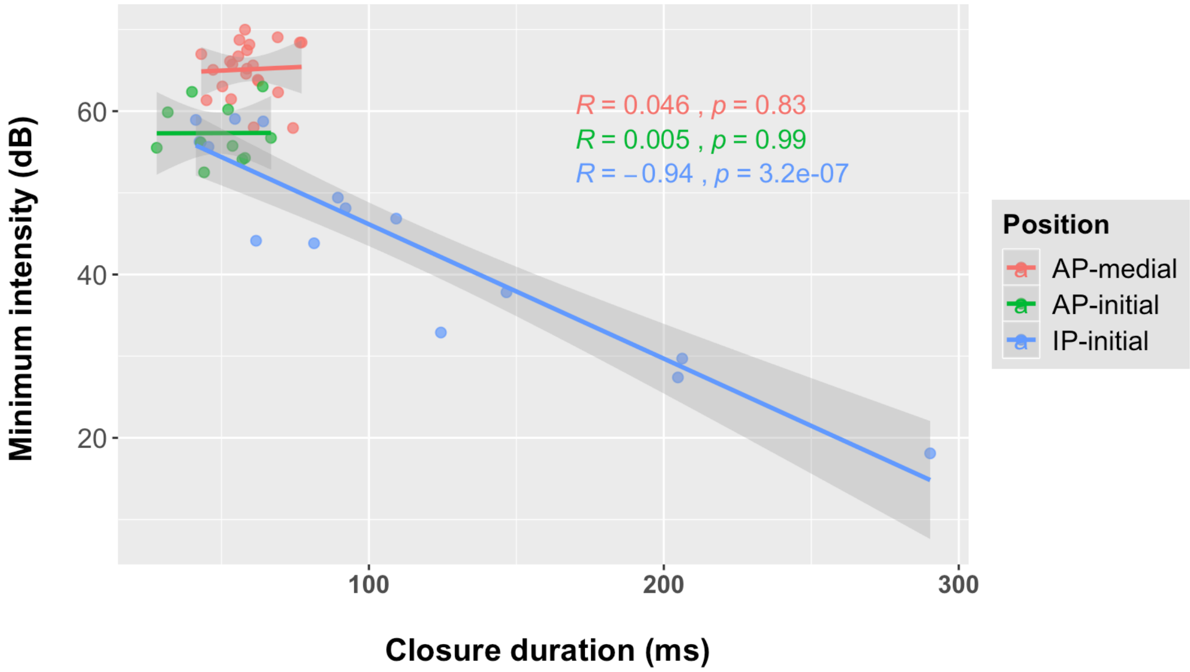


Figure H.37: Minimum intensity (dB) against closure duration (ms) for nasal stops by Seoul_7.

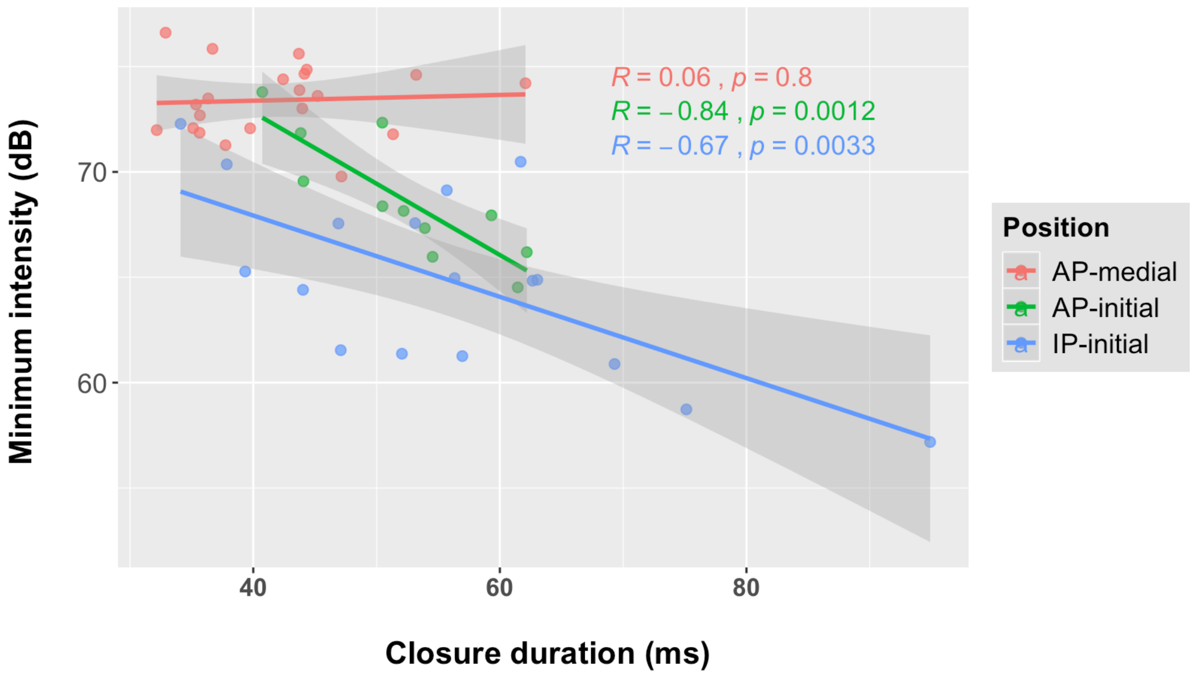


Figure H.38: Minimum intensity (dB) against closure duration (ms) for nasal stops by Seoul_8.

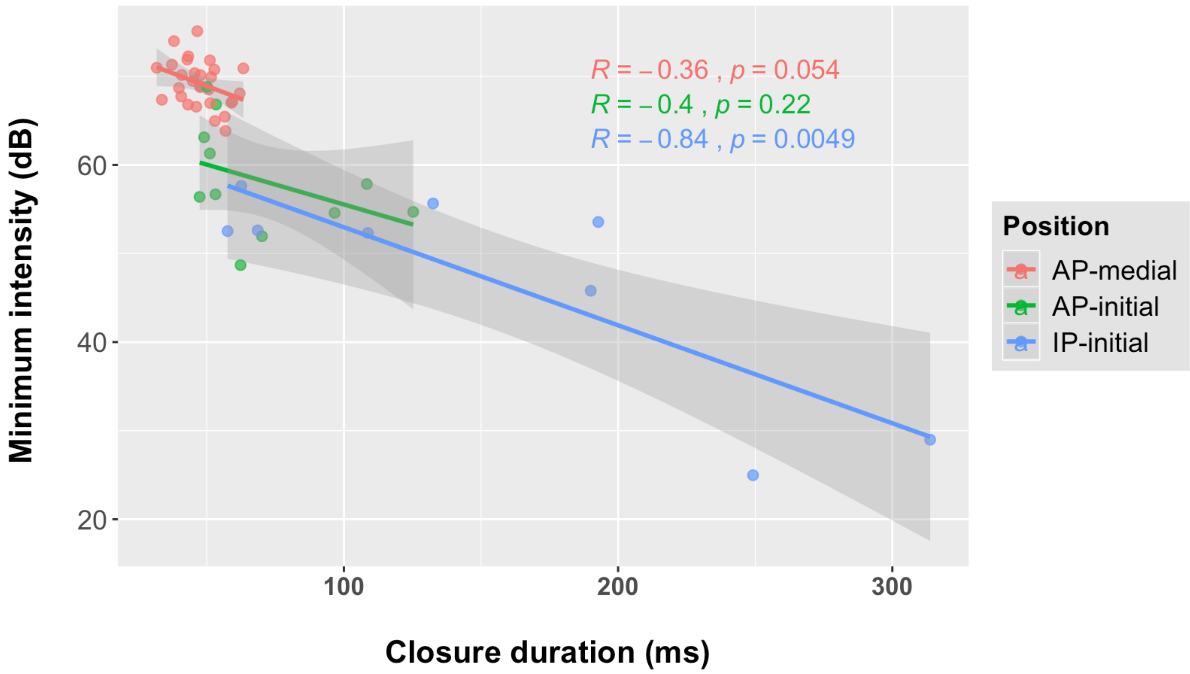


Figure H.39: Minimum intensity (dB) against closure duration (ms) for nasal stops by Seoul_9.

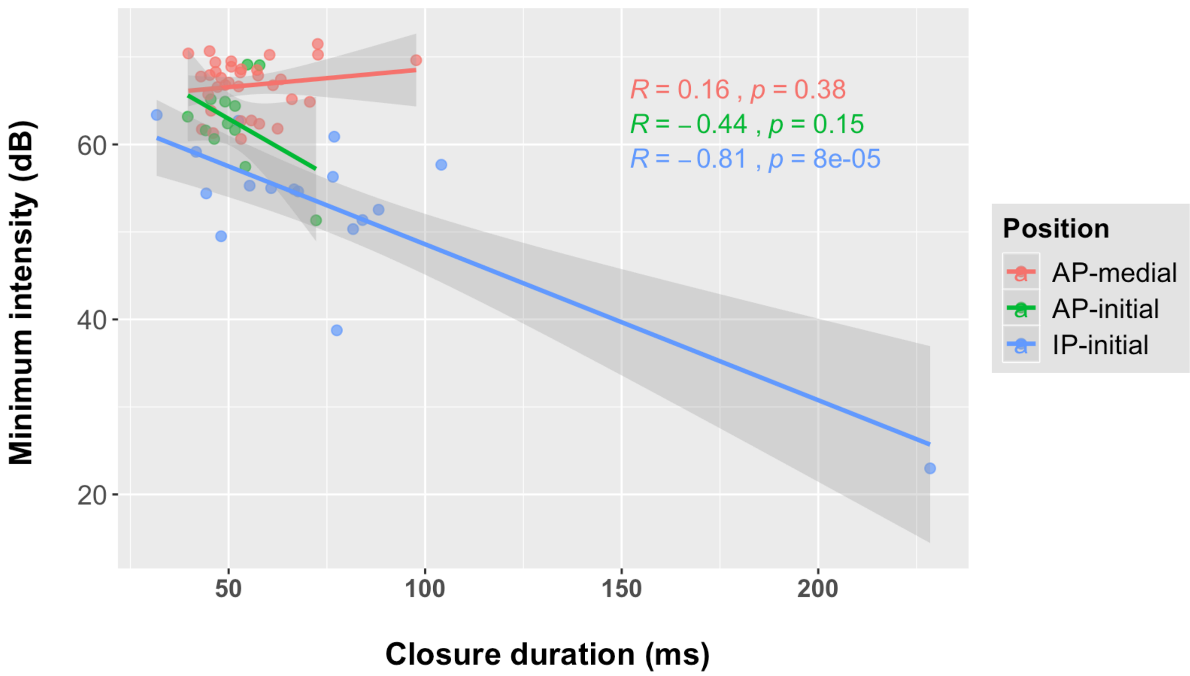


Figure H.40: Minimum intensity (dB) against closure duration (ms) for nasal stops by Seoul_10.

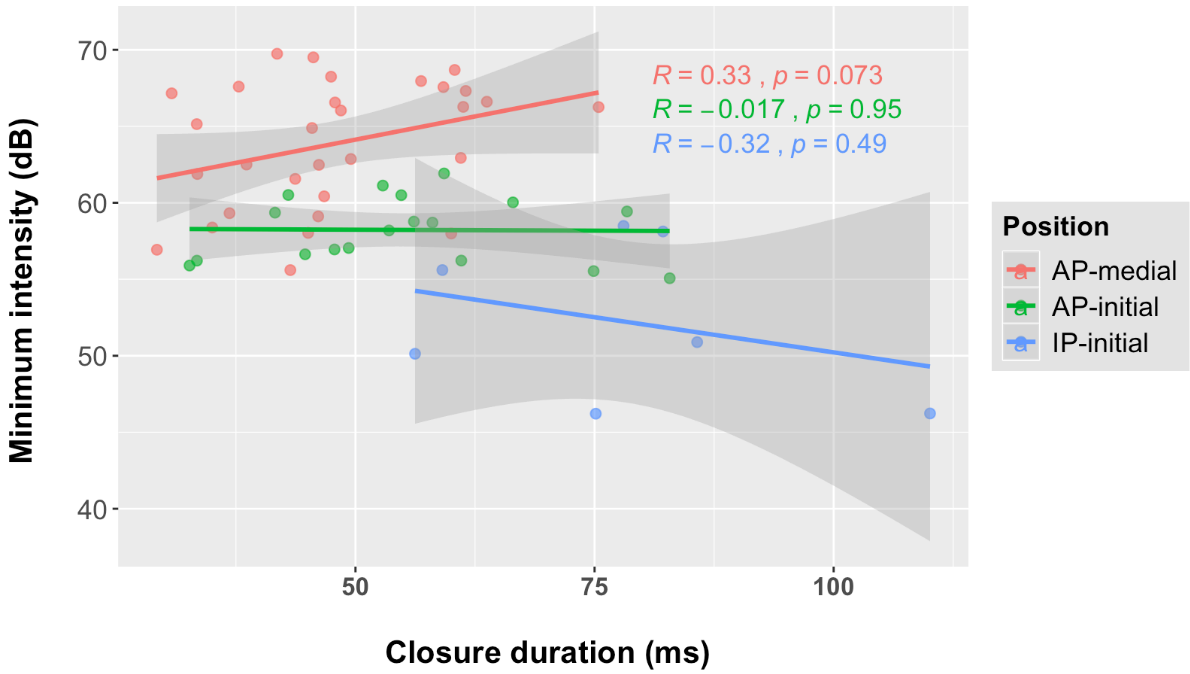


Figure H.41: Minimum intensity (dB) against closure duration (ms) for nasal stops by Busan_1.

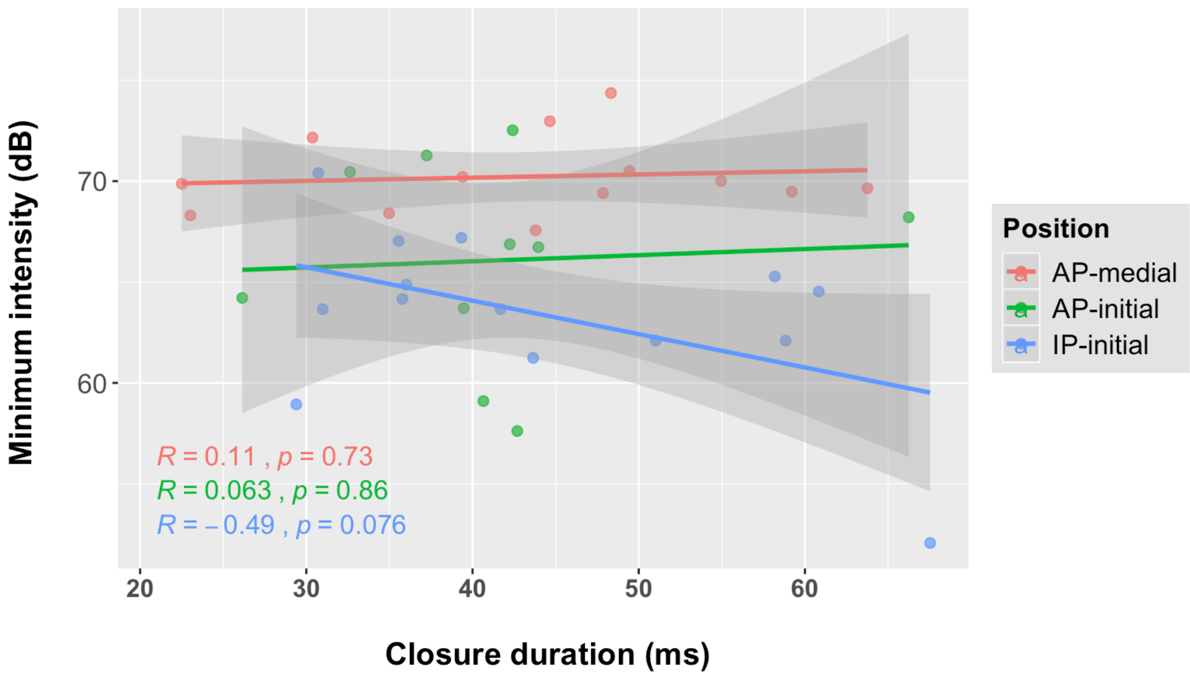


Figure H.42: Minimum intensity (dB) against closure duration (ms) for nasal stops by Busan_2.

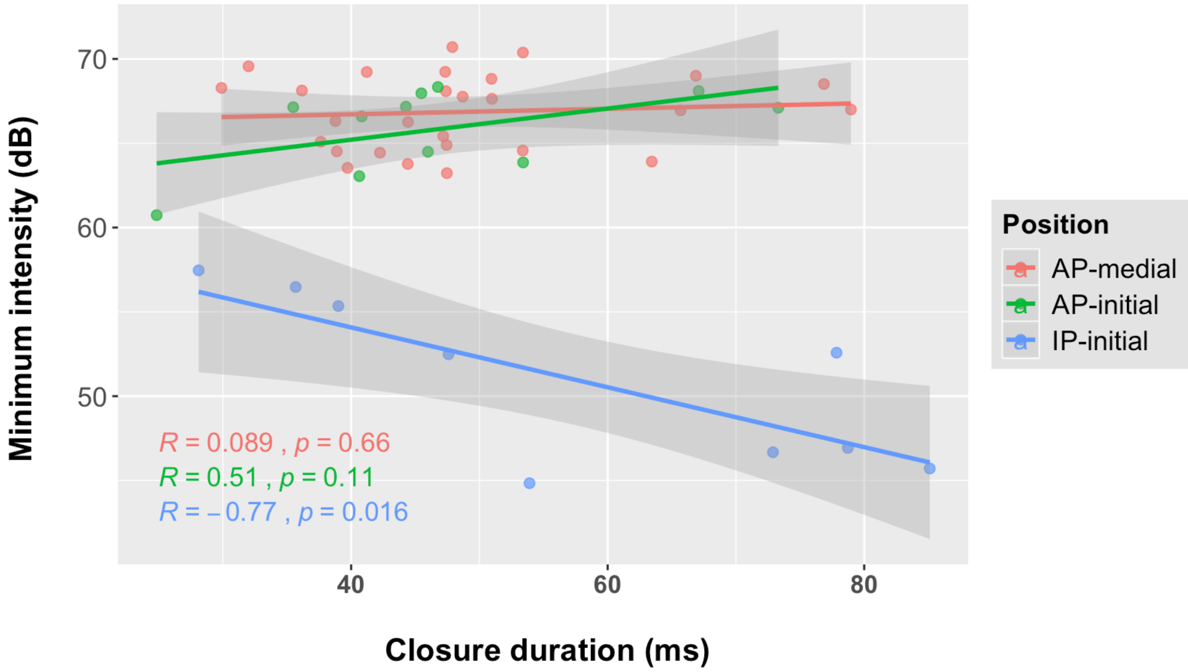


Figure H.43: Minimum intensity (dB) against closure duration (ms) for nasal stops by Busan_3.

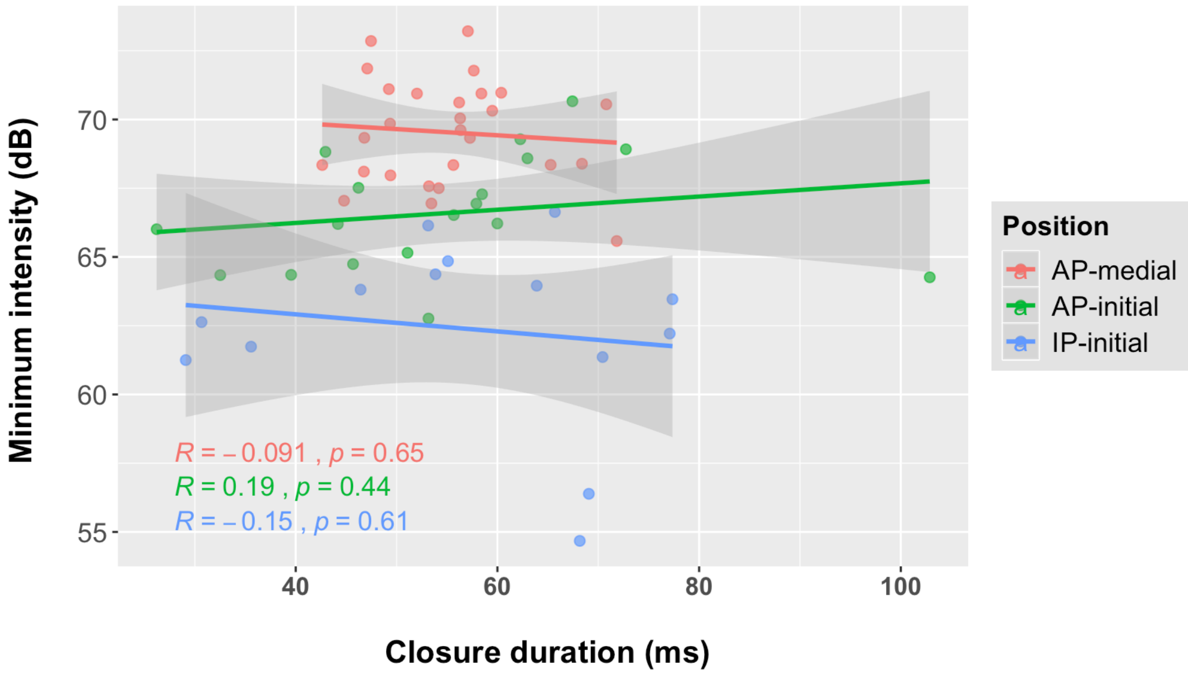


Figure H.44: Minimum intensity (dB) against closure duration (ms) for nasal stops by Busan_4.

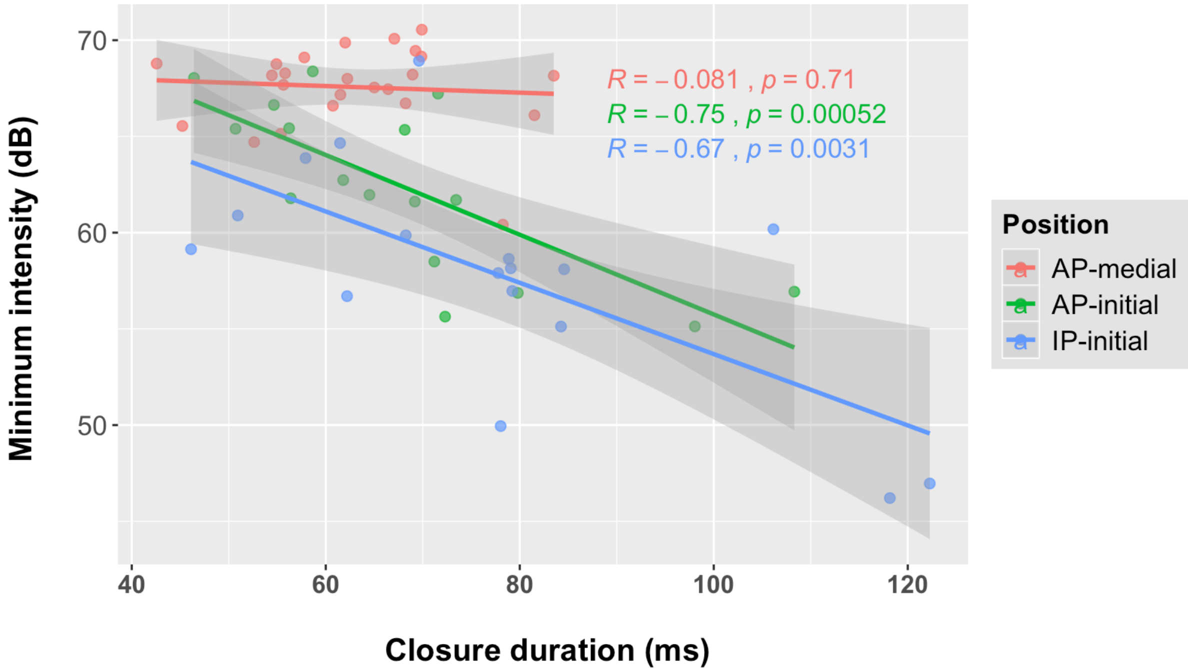


Figure H.45: Minimum intensity (dB) against closure duration (ms) for nasal stops by Busan_5.

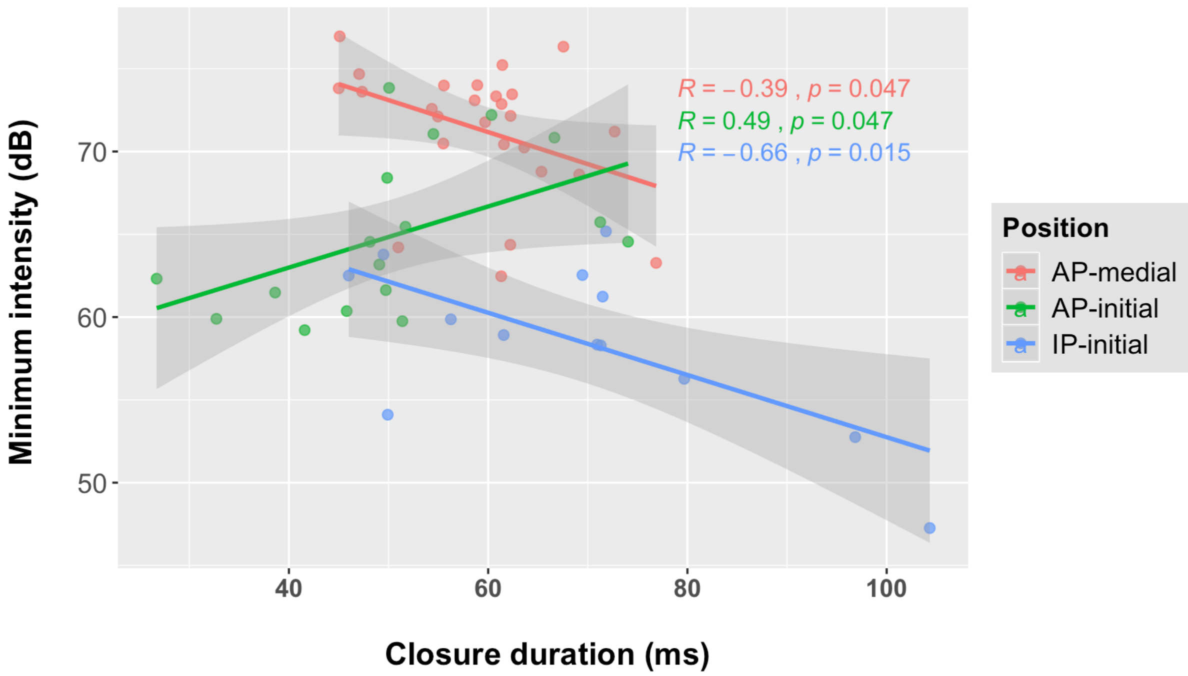


Figure H.46: Minimum intensity (dB) against closure duration (ms) for nasal stops by Busan_6.

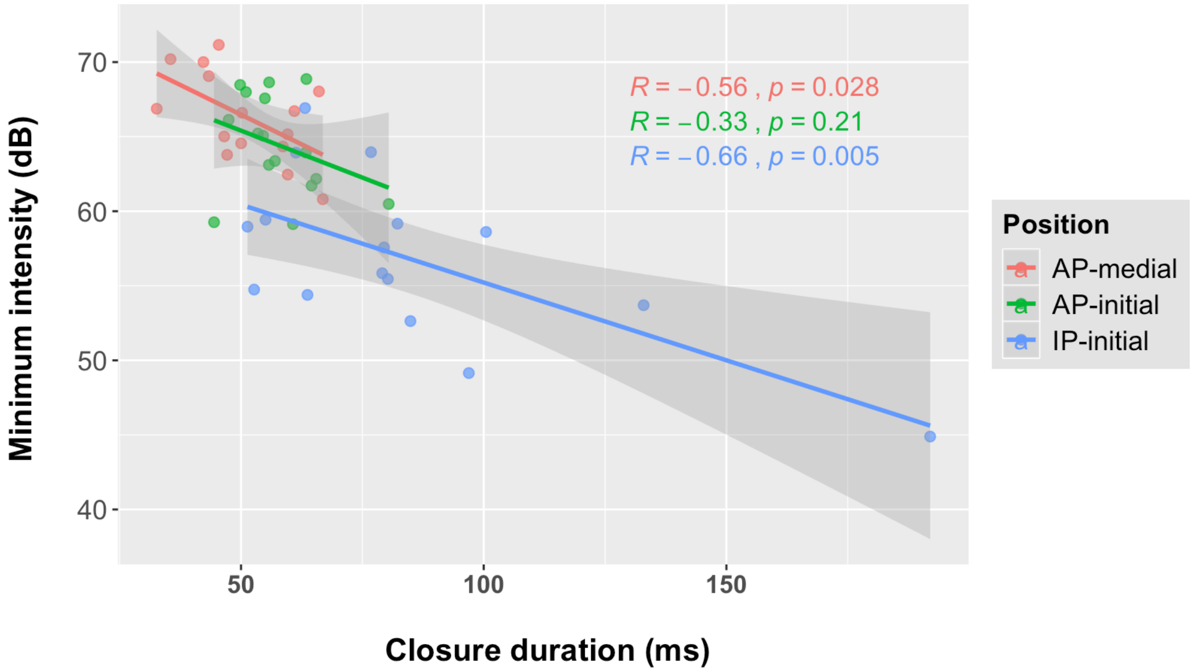


Figure H.47: Minimum intensity (dB) against closure duration (ms) for nasal stops by Busan_7.

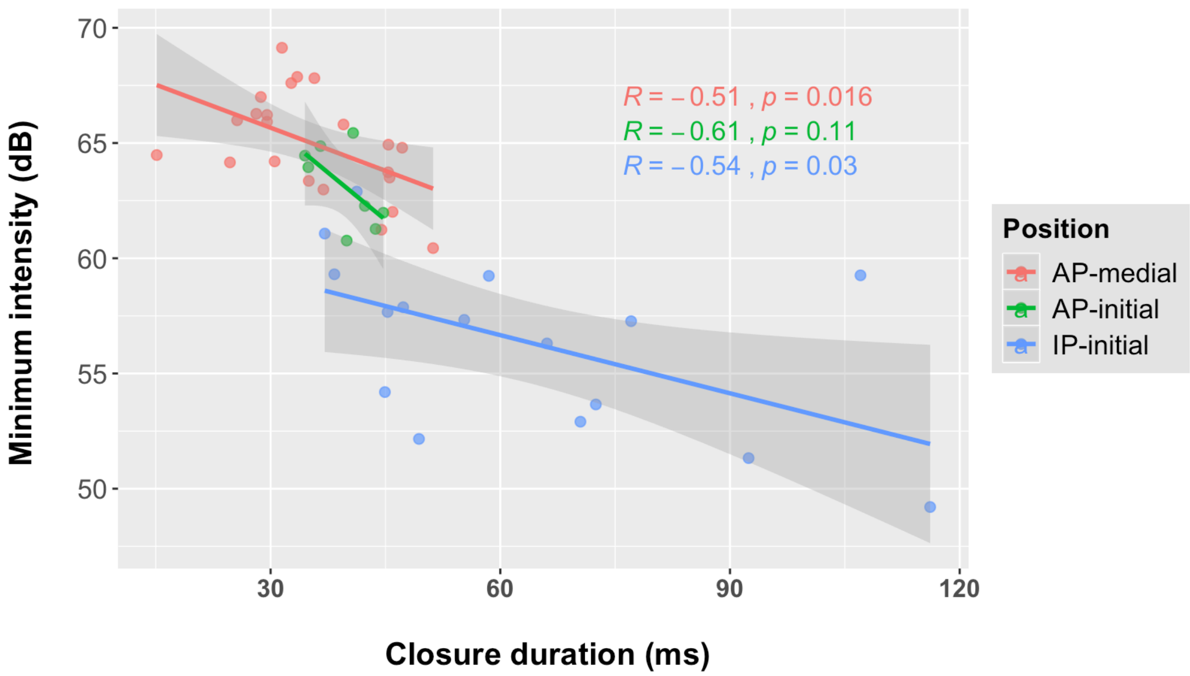


Figure H.48: Minimum intensity (dB) against closure duration (ms) for nasal stops by Busan_8.

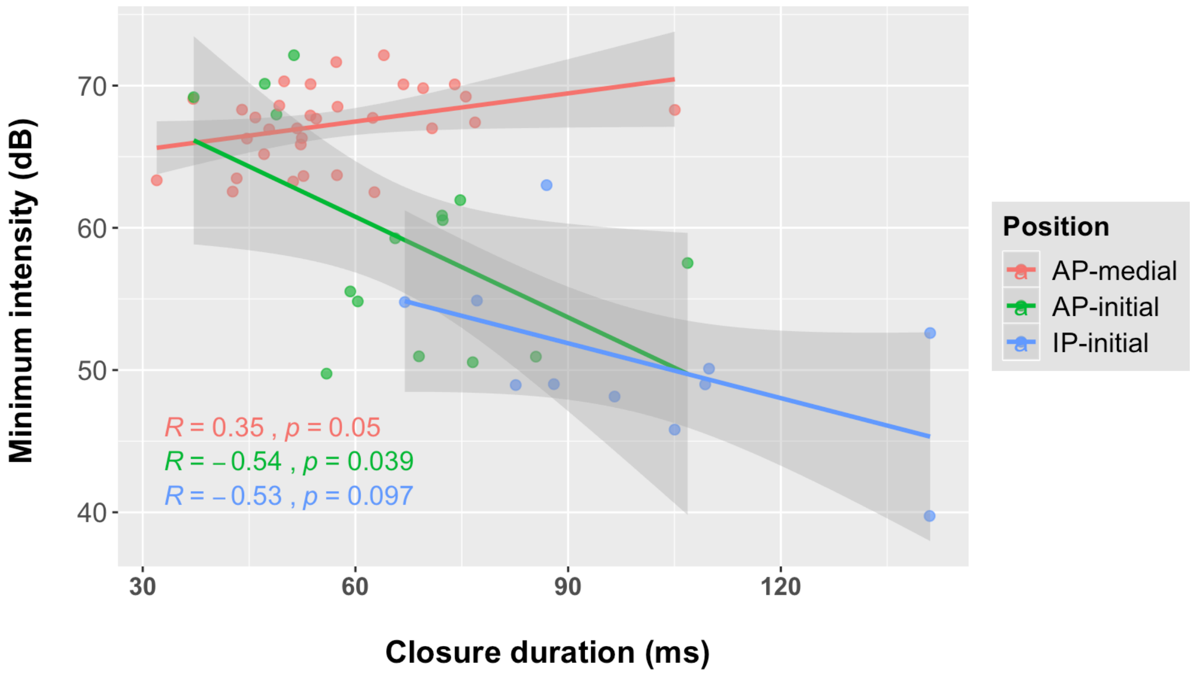


Figure H.49: Minimum intensity (dB) against closure duration (ms) for nasal stops by Busan_9.

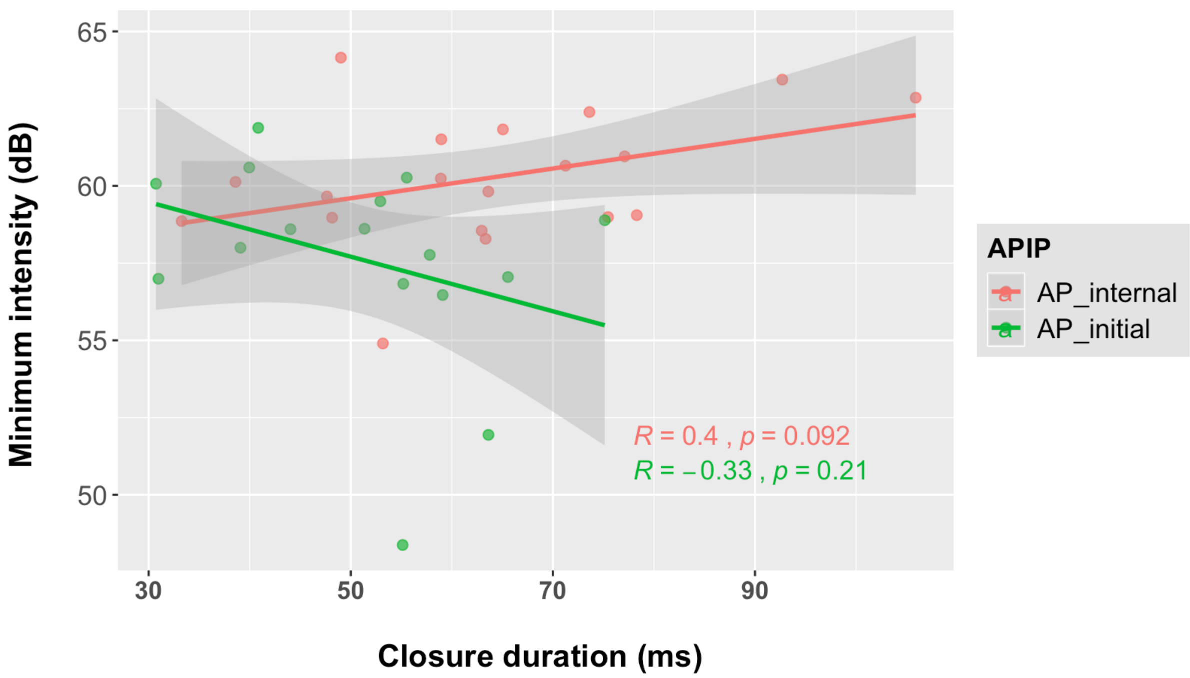


Figure H.50: Minimum intensity (dB) against closure duration (ms) for nasal stops by Busan_10.

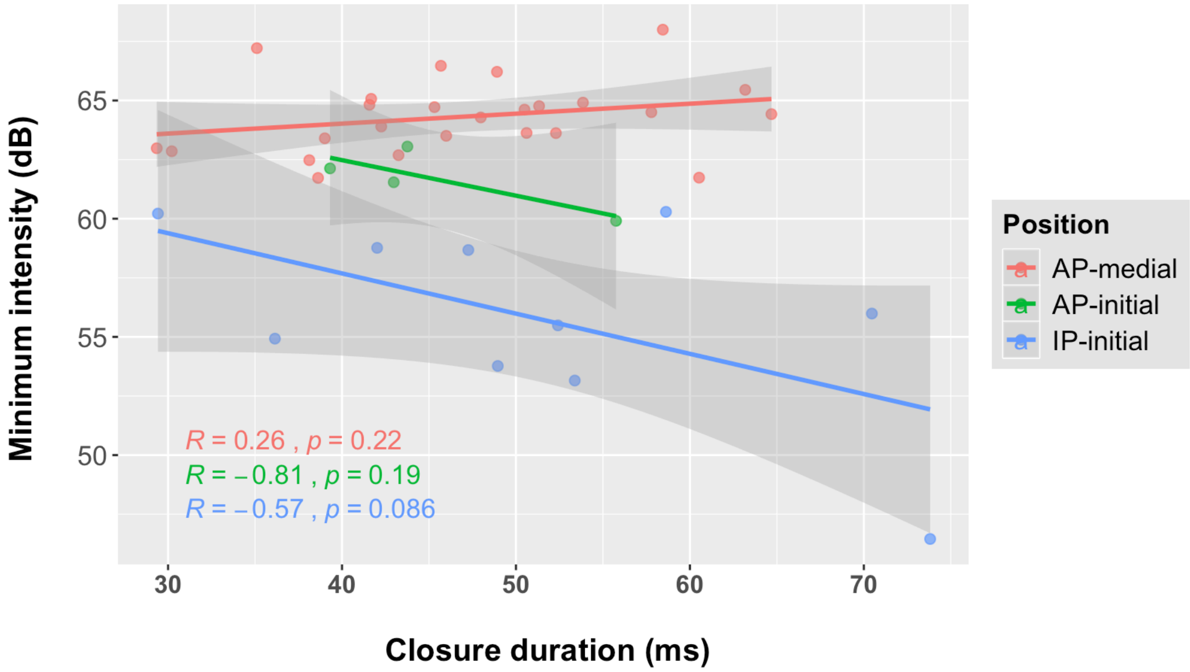


Figure H.51: Minimum intensity (dB) against closure duration (ms) for nasal stops by Ulsan_1.

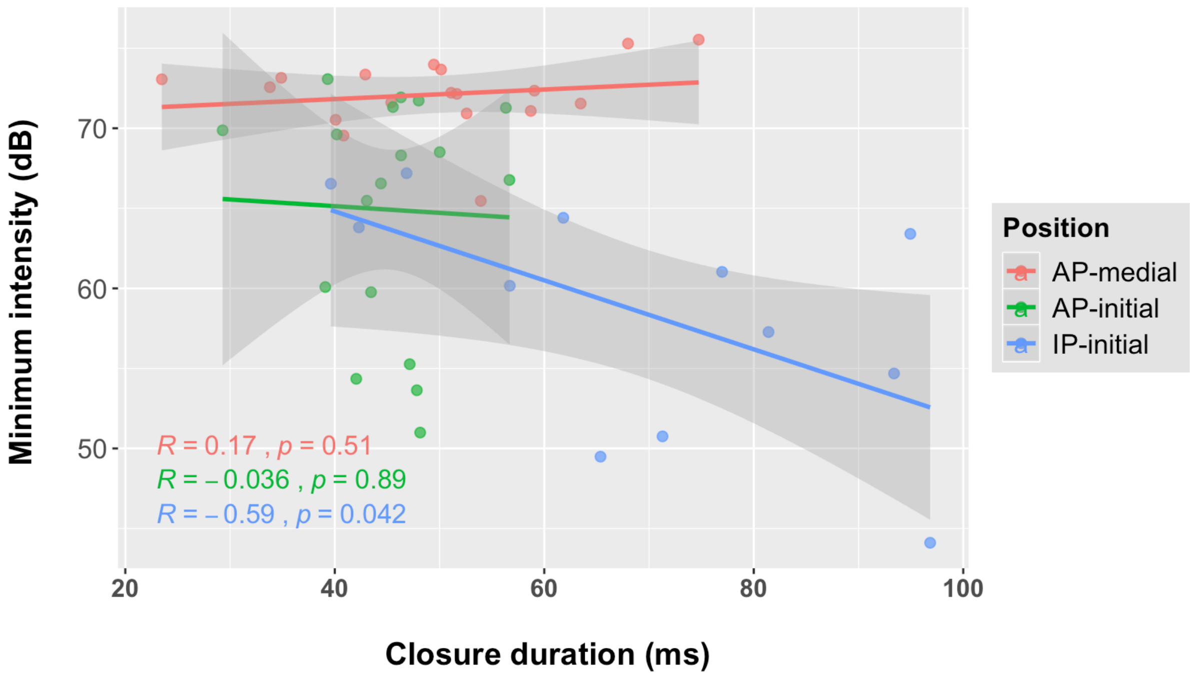


Figure H.52: Minimum intensity (dB) against closure duration (ms) for nasal stops by Ulsan_2.

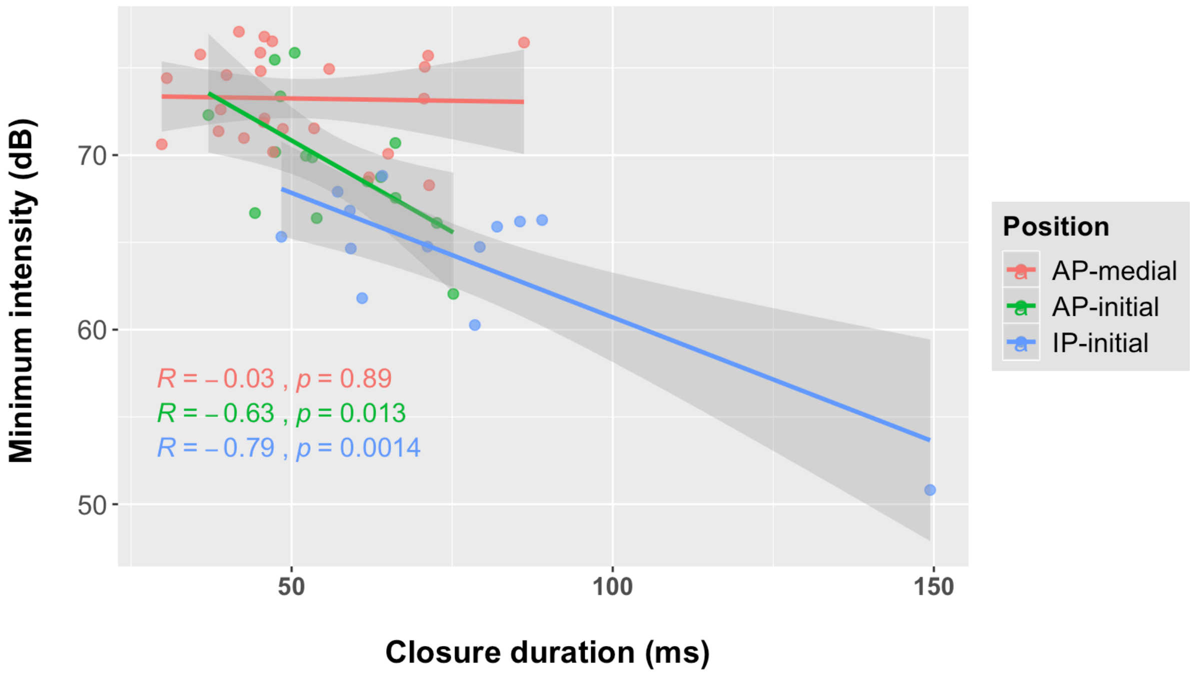


Figure H.53: Minimum intensity (dB) against closure duration (ms) for nasal stops by Ulsan_3.

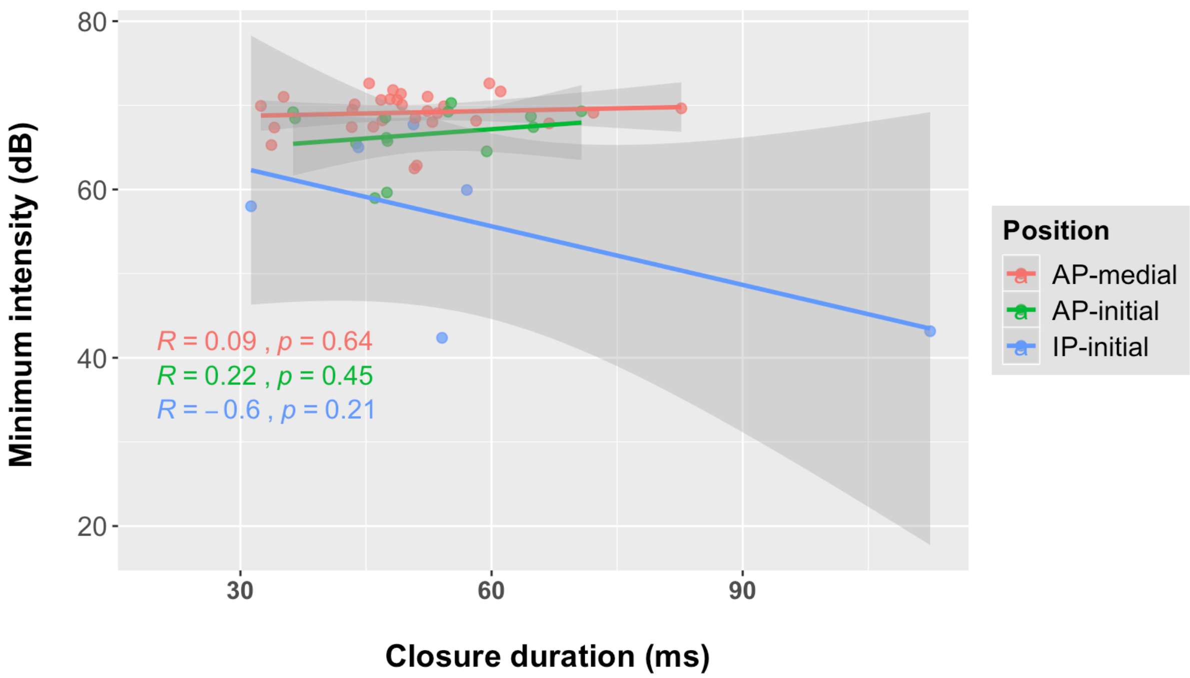


Figure H.54: Minimum intensity (dB) against closure duration (ms) for nasal stops by Ulsan_4.

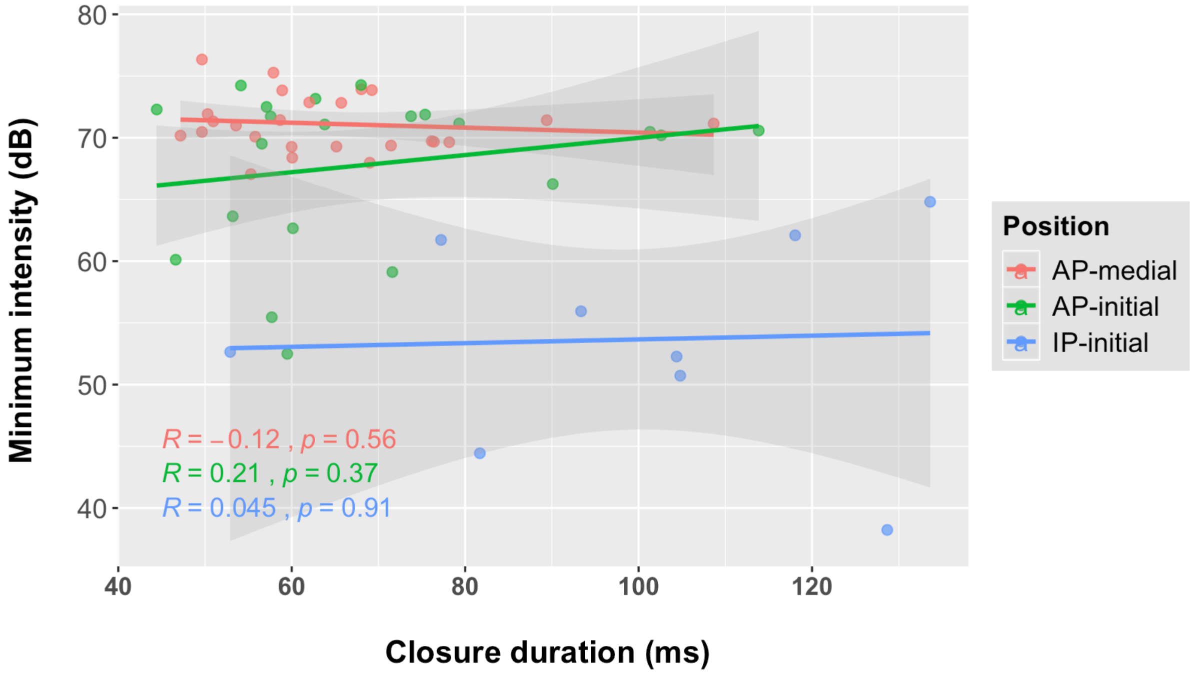


Figure H.55: Minimum intensity (dB) against closure duration (ms) for nasal stops by Ulsan_5.

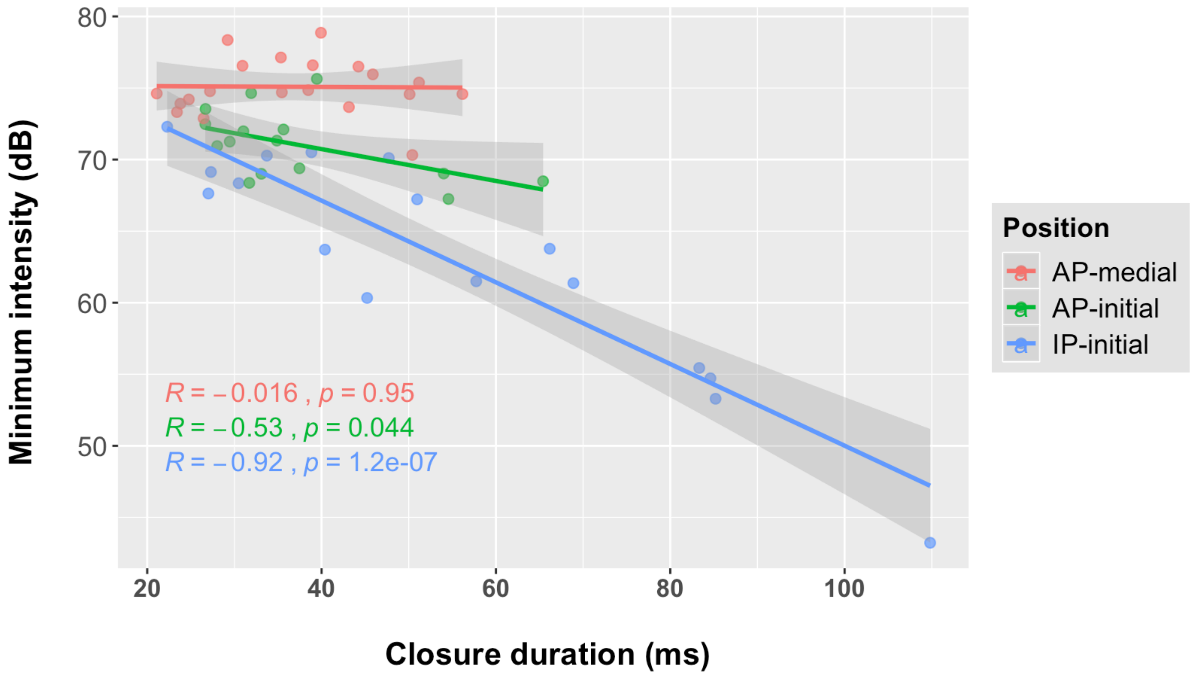


Figure H.56: Minimum intensity (dB) against closure duration (ms) for nasal stops by Ulsan_6.

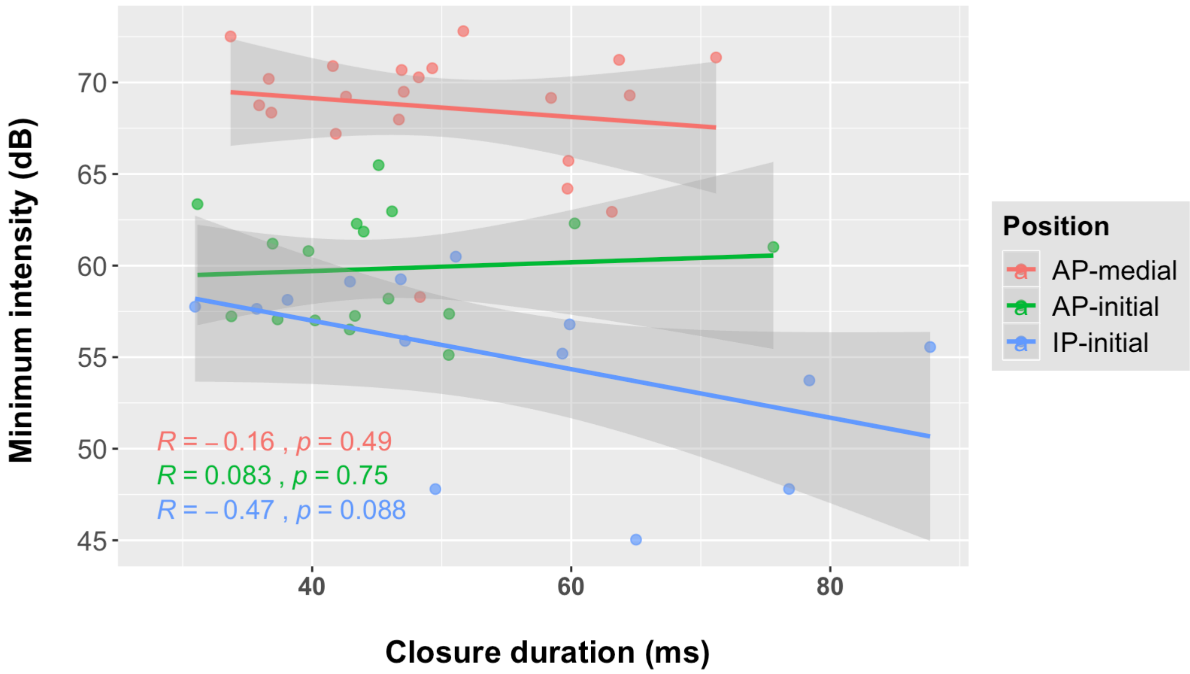


Figure H.57: Minimum intensity (dB) against closure duration (ms) for nasal stops by Ulsan_7.

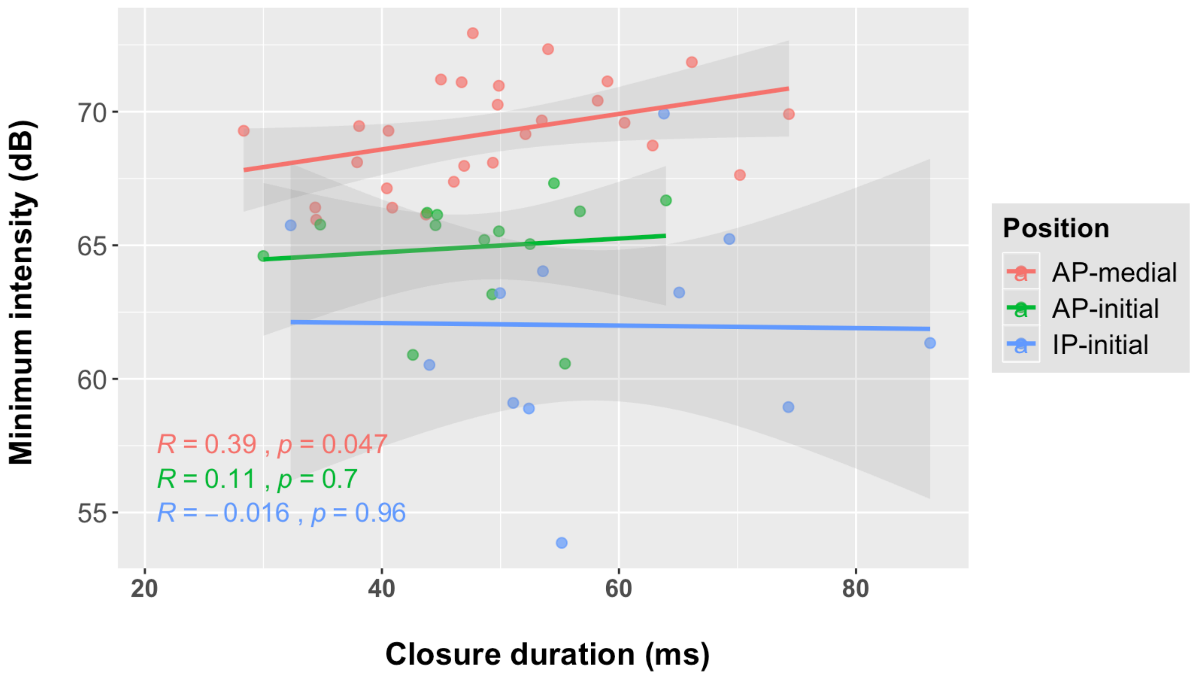


Figure H.58: Minimum intensity (dB) against closure duration (ms) for nasal stops by Ulsan_8.

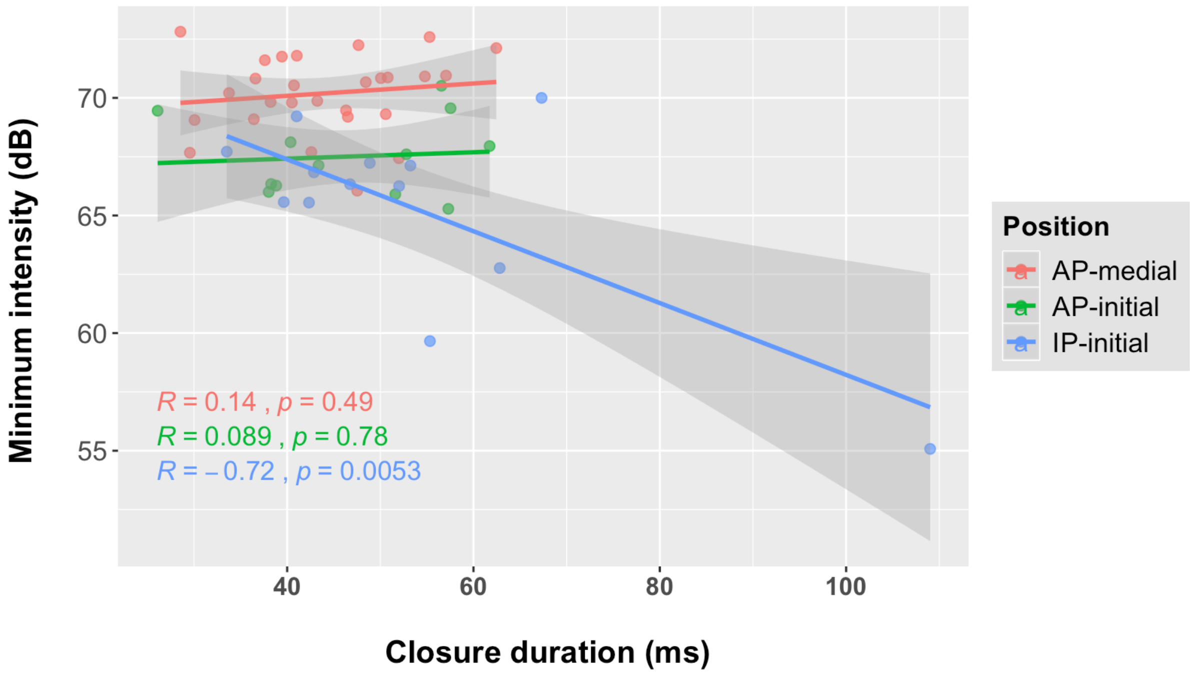


Figure H.59: Minimum intensity (dB) against closure duration (ms) for nasal stops by Ulsan_9.

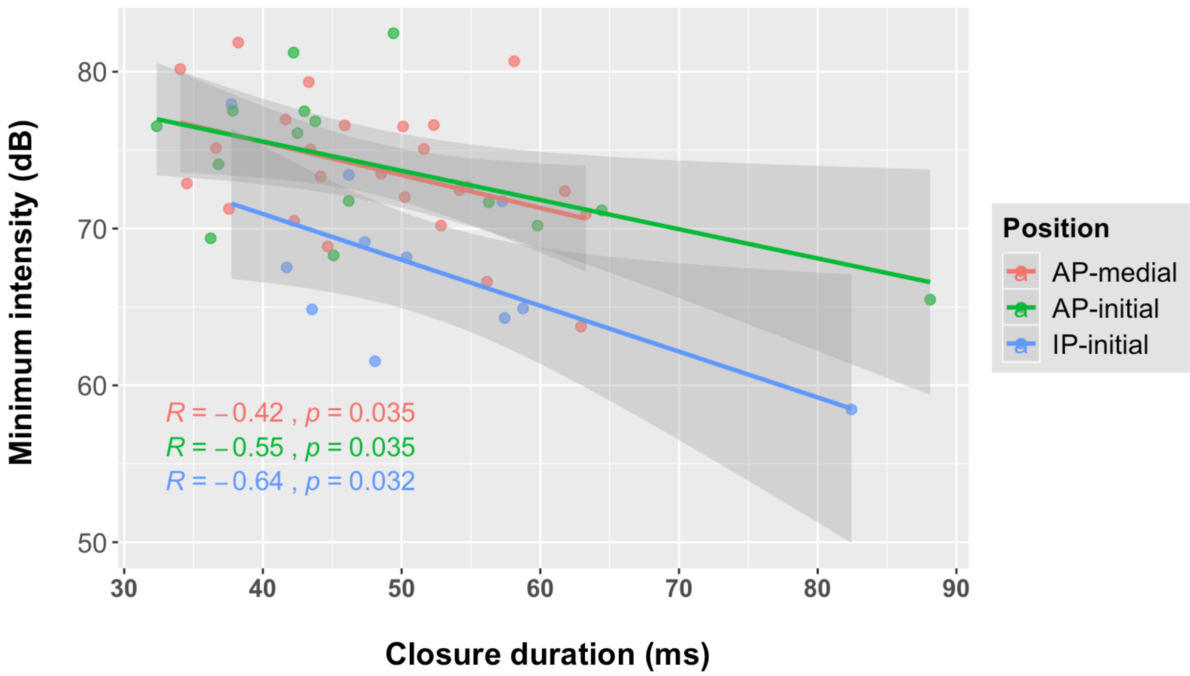


Figure H.60: Minimum intensity (dB) against closure duration (ms) for nasal stops by Ulsan_10.

Appendix I.

Model comparison

In the last three columns of each table, the figures show the F-statistic which reflects the difference between the overall significance of the two models being compared. The asterisks indicate the level of significance: $^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$. The shaded cells indicate the best-fitting model for each outcome, chosen using the cut-off of $p < 0.05$. For example, in Table I.1, the best-fitting model is Model 2 for % voicing because, although Model 3 is significantly better than Model 1, it is not significantly different from Model 2 or Model 4. Since Model 2 is simpler than Model 3 or Model 4, Model 2 is chosen by Occam's Razor. When neither Model 1 nor Model 2 was significantly different from Model 3 or Model 4, the model with a higher adjusted r^2 was chosen between Model 1 and Model 2, as in the case of Ulsan_4 in Table I.24.

Table I.1: Model comparison for Seoul_1.

model	adjusted r^2	model 1 vs 3	model 2 vs 3	model 3 vs 4
1. % voicing ~ closure	0.2296			
2. % voicing ~ position	0.3679			
3. % voicing ~ closure + position	0.3919	5.8037 ***	2.3832	
4. % voicing ~ closure * position	0.379			0.646
1. Intensity ~ closure	0.3584			
2. Intensity ~ position	0.6871			
3. Intensity ~ closure + position	0.7504	42.62 ***	14.194 ***	
4. Intensity ~ closure * position	0.7461			0.572

Table I.2: Model comparison for Seoul_2.

model	adjusted r^2	model 1 vs 3	model 2 vs 3	model 3 vs 4
1. % voicing ~ closure	0.3351			
2. % voicing ~ position	0.4799			
3. % voicing ~ closure + position	0.622	11.629 ***	11.153 ***	
4. % voicing ~ closure * position	0.7443			7.2164 ***
1. Intensity ~ closure	0.3342			
2. Intensity ~ position	0.5678			
3. Intensity ~ closure + position	0.6799	23.672 ***	15.357 ***	
4. Intensity ~ closure * position	0.7241			4.2081 **

Table I.3: Model comparison for Seoul_3.

model	adjusted r^2	model 1 vs 3	model 2 vs 3	model 3 vs 4
1. % voicing ~ closure	0.3789			
2. % voicing ~ position	0.6186			
3. % voicing ~ closure + position	0.6274	11.002 ***	1.685	
4. % voicing ~ closure * position	0.6184			0.6719
1. Intensity ~ closure	0.4336			
2. Intensity ~ position	0.5372			
3. Intensity ~ closure + position	0.6281	14.336 ***	13.229 ***	
4. Intensity ~ closure * position	0.6686			3.9944 **

Table I.4: Model comparison for Seoul_4.

model	adjusted r^2	model 1 vs 3	model 2 vs 3	model 3 vs 4
1. % voicing ~ closure	0.2601			
2. % voicing ~ position	0.1288			
3. % voicing ~ closure + position	0.4153	5.2469 **	16.185 ***	
4. % voicing ~ closure * position	0.4404			2.347
1. Intensity ~ closure	-0.007316			
2. Intensity ~ position	0.1798			
3. Intensity ~ closure + position	0.2006	6.722 ***	2.1207	
4. Intensity ~ closure * position	0.2056			1.1308

Table I.5: Model comparison for Seoul_5.

model	adjusted r^2	model 1 vs 3	model 2 vs 3	model 3 vs 4
1. % voicing ~ closure	0.5502			
2. % voicing ~ position	0.6883			
3. % voicing ~ closure + position	0.8189	27.704 ***	26.248 ***	
4. % voicing ~ closure * position	0.874			8.4391 ***
1. Intensity ~ closure	-0.006154			
2. Intensity ~ position	0.8949			
3. Intensity ~ closure + position	0.8936	246.13 ***	0.3083	
4. Intensity ~ closure * position	0.8912			0.4029

Table I.6: Model comparison for Seoul_6.

model	adjusted r^2	model 1 vs 3	model 2 vs 3	model 3 vs 4
1. % voicing ~ closure	0.07075			
2. % voicing ~ position	0.683			
3. % voicing ~ closure + position	0.6728	26.764 ***	0.1584	
4. % voicing ~ closure * position	0.6562			0.3726
1. Intensity ~ closure	0.1324			
2. Intensity ~ position	0.7628			
3. Intensity ~ closure + position	0.8089	103.65 ***	14.731 ***	
4. Intensity ~ closure * position	0.8615			11.636 ***

Table I.7: Model comparison for Seoul_7.

model	adjusted r^2	model 1 vs 3	model 2 vs 3	model 3 vs 4
1. % voicing ~ closure	0.3251			
2. % voicing ~ position	0.3965			
3. % voicing ~ closure + position	0.5613	11.23 ***	14.899 ***	
4. % voicing ~ closure * position	0.58			1.8002
1. Intensity ~ closure	0.6612			
2. Intensity ~ position	0.5736			
3. Intensity ~ closure + position	0.8771	42.269 ***	114.58 ***	
4. Intensity ~ closure * position	0.8886			3.3279 **

Table I.8: Model comparison for Seoul_8.

model	adjusted r^2	model 1 vs 3	model 2 vs 3	model 3 vs 4
1. % voicing ~ closure	0.2993			
2. % voicing ~ position	0.5689			
3. % voicing ~ closure + position	0.5595	11.632 ***	0.2544	
4. % voicing ~ closure * position	0.5719			1.4901
1. Intensity ~ closure	0.4848			
2. Intensity ~ position	0.5944			
3. Intensity ~ closure + position	0.7172	20.31 ***	20.969 ***	
4. Intensity ~ closure * position	0.7571			4.6925 **

Table I.9: Model comparison for Seoul_9.

model	adjusted r^2	model 1 vs 3	model 2 vs 3	model 3 vs 4
1. % voicing ~ closure	0.4206			
2. % voicing ~ position	0.8383			
3. % voicing ~ closure + position	0.8398	54.621 ***	1.3768	
4. % voicing ~ closure * position	0.8446			1.6064
1. Intensity ~ closure	0.7054			
2. Intensity ~ position	0.6654			
3. Intensity ~ closure + position	0.8339	19.169 ***	47.636 ***	
4. Intensity ~ closure * position	0.8268			0.0779

Table I.10: Model comparison for Seoul_10.

model	adjusted r^2	model 1 vs 3	model 2 vs 3	model 3 vs 4
1. % voicing ~ closure	0.2263			
2. % voicing ~ position	0.2495			
3. % voicing ~ closure + position	0.3634	6.0642 ***	9.2315 ***	
4. % voicing ~ closure * position	0.411			2.8184 *
1. Intensity ~ closure	0.4302			
2. Intensity ~ position	0.4987			
3. Intensity ~ closure + position	0.699	28.251 ***	40.942 ***	
4. Intensity ~ closure * position	0.7382			5.4155 ***

Table I.11: Model comparison for Busan_1.

model	adjusted r^2	model 1 vs 3	model 2 vs 3	model 3 vs 4
1. % voicing ~ closure	0.1347			
2. % voicing ~ position	0.06067			
3. % voicing ~ closure + position	0.1038	0.4309	2.5396	
4. % voicing ~ closure * position	0.1291			1.4499
1. Intensity ~ closure	0.1224			
2. Intensity ~ position	0.5359			
3. Intensity ~ closure + position	0.5317	24.158 ***	0.5307	
4. Intensity ~ closure * position	0.5551			2.3459

Table I.12: Model comparison for Busan_2.

model	adjusted r^2	model 1 vs 3	model 2 vs 3	model 3 vs 4
1. % voicing ~ closure	0.3859			
2. % voicing ~ position	0.1304			
3. % voicing ~ closure + position	0.3512	0.3048	9.5073 ***	
4. % voicing ~ closure * position	0.418			2.3769
1. Intensity ~ closure	-0.003284			
2. Intensity ~ position	0.3498			
3. Intensity ~ closure + position	0.3515	10.576 ***	1.0899	
4. Intensity ~ closure * position	0.3695			1.4703

Table I.13: Model comparison for Busan_3.

model	adjusted r^2	model 1 vs 3	model 2 vs 3	model 3 vs 4
1. % voicing ~ closure	0.03061			
2. % voicing ~ position	0.2685			
3. % voicing ~ closure + position	0.2531	5.9165 ***	0.3441	
4. % voicing ~ closure * position	0.3241			2.627 *
1. Intensity ~ closure	0.07645			
2. Intensity ~ position	0.7189			
3. Intensity ~ closure + position	0.7251	55.257 ***	2.0008	
4. Intensity ~ closure * position	0.7811			6.6371 ***

Table I.14: Model comparison for Busan_4.

model	adjusted r^2	model 1 vs 3	model 2 vs 3	model 3 vs 4
1. % voicing ~ closure	0.3568			
2. % voicing ~ position	0.08119			
3. % voicing ~ closure + position	0.3307	0.2373	15.164 ***	
4. % voicing ~ closure * position	0.3386			1.2222
1. Intensity ~ closure	-0.01535			
2. Intensity ~ position	0.5861			
3. Intensity ~ closure + position	0.5786	41.175 ***	0.0095	
4. Intensity ~ closure * position	0.5723			0.5921

Table I.15: Model comparison for Busan_5.

model	adjusted r^2	model 1 vs 3	model 2 vs 3	model 3 vs 4
1. % voicing ~ closure	0.1489			
2. % voicing ~ position	0.06195			
3. % voicing ~ closure + position	0.1183	0.3395	3.3625 *	
4. % voicing ~ closure * position	0.2118			3.137 *
1. Intensity ~ closure	0.428			
2. Intensity ~ position	0.4818			
3. Intensity ~ closure + position	0.67	21.523 ***	32.348 ***	
4. Intensity ~ closure * position	0.6926			2.9937 *

Table I.16: Model comparison for Busan_6.

model	adjusted r^2	model 1 vs 3	model 2 vs 3	model 3 vs 4
1. % voicing ~ closure	0.1238			
2. % voicing ~ position	0.03508			
3. % voicing ~ closure + position	0.1006	0.458	3.9878 *	
4. % voicing ~ closure * position	0.1051			1.0993
1. Intensity ~ closure	0.05205			
2. Intensity ~ position	0.5692			
3. Intensity ~ closure + position	0.5765	34.43 ***	1.9143	
4. Intensity ~ closure * position	0.6634			7.7154 ***

Table I.17: Model comparison for Busan_7.

model	adjusted r^2	model 1 vs 3	model 2 vs 3	model 3 vs 4
1. % voicing ~ closure	0.6552			
2. % voicing ~ position	0.5009			
3. % voicing ~ closure + position	0.6831	2.0585	14.224 ***	
4. % voicing ~ closure * position	0.776			5.5637 **
1. Intensity ~ closure	0.5365			
2. Intensity ~ position	0.4894			
3. Intensity ~ closure + position	0.6558	8.7961 ***	22.269 ***	
4. Intensity ~ closure * position	0.6424			0.1951

Table I.18: Model comparison for Busan_8.

model	adjusted r^2	model 1 vs 3	model 2 vs 3	model 3 vs 4
1. % voicing ~ closure	0.2825			
2. % voicing ~ position	0.8117			
3. % voicing ~ closure + position	0.8023	25.975 ***	0.1421	
4. % voicing ~ closure * position	0.8			0.9037
1. Intensity ~ closure	0.5626			
2. Intensity ~ position	0.6598			
3. Intensity ~ closure + position	0.7455	16.812 ***	15.486 ***	
4. Intensity ~ closure * position	0.7392			0.4917

Table I.19: Model comparison for Busan_9.

model	adjusted r^2	model 1 vs 3	model 2 vs 3	model 3 vs 4
1. % voicing ~ closure	0.3191			
2. % voicing ~ position	0.3047			
3. % voicing ~ closure + position	0.3505	1.8961	3.5397 *	
4. % voicing ~ closure * position	0.4074			2.6799 *
1. Intensity ~ closure	0.4219			
2. Intensity ~ position	0.6234			
3. Intensity ~ closure + position	0.6467	18.813 ***	4.6303 **	
4. Intensity ~ closure * position	0.7045			6.277 ***

Table I.20: Model comparison for Busan_10.

model	adjusted r^2	model 1 vs 3	model 2 vs 3	model 3 vs 4
1. % voicing ~ closure	0.07984			
2. % voicing ~ position	0.2984			
3. % voicing ~ closure + position	0.2776	4.6956 **	0.2506	
4. % voicing ~ closure * position	0.2743			0.9425
1. Intensity ~ closure	0.01364			
2. Intensity ~ position	0.4675			
3. Intensity ~ closure + position	0.4525	15.026 ***	0.0658	
4. Intensity ~ closure * position	0.4896			2.1984

Table I.21: Model comparison for Ulsan_1.

model	adjusted r^2	model 1 vs 3	model 2 vs 3	model 3 vs 4
1. % voicing ~ closure	0.1765			
2. % voicing ~ position	0.1388			
3. % voicing ~ closure + position	0.1745	0.9611	2.3844	
4. % voicing ~ closure * position	0.2323			2.1673
1. Intensity ~ closure	0.04923			
2. Intensity ~ position	0.6865			
3. Intensity ~ closure + position	0.6964	0.426 ***	2.1661	
4. Intensity ~ closure * position	0.7479			4.5775 **

Table I.22: Model comparison for Ulsan_2.

model	adjusted r^2	model 1 vs 3	model 2 vs 3	model 3 vs 4
1. % voicing ~ closure	0.3651			
2. % voicing ~ position	0.1827			
3. % voicing ~ closure + position	0.4133	2.6429 *	16.325 ***	
4. % voicing ~ closure * position	0.4668			2.9075 *
1. Intensity ~ closure	0.1631			
2. Intensity ~ position	0.437			
3. Intensity ~ closure + position	0.4632	13.859 ***	3.1998 *	
4. Intensity ~ closure * position	0.4784			1.6412

Table I.23: Model comparison for Ulsan_3.

model	adjusted r^2	model 1 vs 3	model 2 vs 3	model 3 vs 4
1. % voicing ~ closure	0.6128			
2. % voicing ~ position	0.538			
3. % voicing ~ closure + position	0.6342	1.9668	9.4123 ***	
4. % voicing ~ closure * position	0.6167			0.2925
1. Intensity ~ closure	0.4258			
2. Intensity ~ position	0.516			
3. Intensity ~ closure + position	0.6293	15.001 ***	16.277 ***	
4. Intensity ~ closure * position	0.6801			4.8919 **

Table I.24: Model comparison for Ulsan_4.

model	adjusted r^2	model 1 vs 3	model 2 vs 3	model 3 vs 4
1. % voicing ~ closure	0.04011			
2. % voicing ~ position	0.07214			
3. % voicing ~ closure + position	0.05829	1.2607	0.6177	
4. % voicing ~ closure * position	-0.006195			0.1989
1. Intensity ~ closure	0.07821			
2. Intensity ~ position	0.4612			
3. Intensity ~ closure + position	0.4868	20.111 ***	3.3485 *	
4. Intensity ~ closure * position	0.5626			4.9833 **

Table I.25: Model comparison for Ulsan_5.

model	adjusted r^2	model 1 vs 3	model 2 vs 3	model 3 vs 4
1. % voicing ~ closure	0.3153			
2. % voicing ~ position	0.4411			
3. % voicing ~ closure + position	0.5096	7.9336 ***	5.7449 **	
4. % voicing ~ closure * position	0.6095			5.2203 **
1. Intensity ~ closure	0.1285			
2. Intensity ~ position	0.5536			
3. Intensity ~ closure + position	0.5489	25.702 ***	0.459	
4. Intensity ~ closure * position	0.5379			0.3918

Table I.26: Model comparison for Ulsan_6.

model	adjusted r^2	model 1 vs 3	model 2 vs 3	model 3 vs 4
1. % voicing ~ closure	0.3552			
2. % voicing ~ position	0.6894			
3. % voicing ~ closure + position	0.7017	22.489 ***	2.484	
4. % voicing ~ closure * position	0.7063			1.2728
1. Intensity ~ closure	0.6016			
2. Intensity ~ position	0.5235			
3. Intensity ~ closure + position	0.8091	28.171 ***	74.284 ***	
4. Intensity ~ closure * position	0.8762			14.027 ***

Table I.27: Model comparison for Ulsan_7.

model	adjusted r^2	model 1 vs 3	model 2 vs 3	model 3 vs 4
1. % voicing ~ closure	0.5697			
2. % voicing ~ position	0.4408			
3. % voicing ~ closure + position	0.637	5.1712 ***	24.776 ***	
4. % voicing ~ closure * position	0.6322			0.7228
1. Intensity ~ closure	0.02034			
2. Intensity ~ position	0.7005			
3. Intensity ~ closure + position	0.7135	61.493 ***	3.2209 *	
4. Intensity ~ closure * position	0.7155			1.1652

Table I.28: Model comparison for Ulsan_8.

model	adjusted r^2	model 1 vs 3	model 2 vs 3	model 3 vs 4
1. % voicing ~ closure	0.4624			
2. % voicing ~ position	0.4427			
3. % voicing ~ closure + position	0.5829	5.187 **	10.407 ***	
4. % voicing ~ closure * position	0.6453			3.3748 *
1. Intensity ~ closure	-0.01252			
2. Intensity ~ position	0.5632			
3. Intensity ~ closure + position	0.5652	34.879 ***	1.2241	
4. Intensity ~ closure * position	0.5561			0.5003

Table I.29: Model comparison for Ulsan_9.

model	adjusted r^2	model 1 vs 3	model 2 vs 3	model 3 vs 4
1. % voicing ~ closure	0.3776			
2. % voicing ~ position	0.6842			
3. % voicing ~ closure + position	0.7824	28.895 ***	14.081 ***	
4. % voicing ~ closure * position	0.8202			3.9447 **
1. Intensity ~ closure	0.2141			
2. Intensity ~ position	0.3971			
3. Intensity ~ closure + position	0.4816	13.906 ***	8.9902 ***	
4. Intensity ~ closure * position	0.5811			6.7011 ***

Table I.30: Model comparison for Ulsan_10.

model	adjusted r^2	model 1 vs 3	model 2 vs 3	model 3 vs 4
1. % voicing ~ closure	0.2909			
2. % voicing ~ position	0.09232			
3. % voicing ~ closure + position	0.3042	1.3054	10.438 ***	
4. % voicing ~ closure * position	0.3552			2.188
1. Intensity ~ closure	0.252			
2. Intensity ~ position	0.2174			
3. Intensity ~ closure + position	0.4184	8.0086 ***	17.585 ***	
4. Intensity ~ closure * position	0.4012			0.3244

Appendix J.

Perception study modelling results

Table J.1: Predicted probabilities of a correct answer for each combination of factor levels.

Condition	Consonant	Region	Host-sentence	Mean	Lower Quantile (0.025)	Upper Quantile (0.975)
same-spliced	t	Seoul	PW	0.71142	0.57799	0.82064
cross-spliced	t	Seoul	PW	0.38023	0.19827	0.60074
same-spliced	t	Seoul	AP	0.80547	0.55003	0.94940
cross-spliced	t	Seoul	AP	0.40814	0.12122	0.75467
same-spliced	t	Busan	PW	0.64668	0.49369	0.77659
cross-spliced	t	Busan	PW	0.42597	0.22300	0.64870
same-spliced	t	Busan	AP	0.70819	0.38315	0.91649
cross-spliced	t	Busan	AP	0.39205	0.10378	0.75242
same-spliced	t	Ulsan	PW	0.76740	0.62953	0.87301
cross-spliced	t	Ulsan	PW	0.40352	0.20309	0.64006
same-spliced	t	Ulsan	AP	0.79752	0.51121	0.95418
cross-spliced	t	Ulsan	AP	0.31236	0.06874	0.67016
same-spliced	n	Seoul	PW	0.73133	0.60063	0.83607
cross-spliced	n	Seoul	PW	0.40505	0.21630	0.61357
same-spliced	n	Seoul	AP	0.79277	0.51159	0.95146
cross-spliced	n	Seoul	AP	0.48354	0.15837	0.82884
same-spliced	n	Busan	PW	0.71812	0.56721	0.83591
cross-spliced	n	Busan	PW	0.43756	0.22816	0.66233
same-spliced	n	Busan	AP	0.71307	0.32517	0.94367
cross-spliced	n	Busan	AP	0.51571	0.16332	0.85776
same-spliced	n	Ulsan	PW	0.84724	0.73365	0.92796
cross-spliced	n	Ulsan	PW	0.58714	0.34725	0.79841
same-spliced	n	Ulsan	AP	0.67473	0.33188	0.91455
cross-spliced	n	Ulsan	AP	0.40014	0.10400	0.78110

Table J.2: Differences between the same- and the cross-spliced condition for the predicted probabilities of a correct answer.

Consonant	Region	Host-sentence	Mean	Lower Quantile (0.025)	Upper Quantile (0.975)
t	Seoul	PW	0.33119	0.13278	0.50581
t	Seoul	AP	0.39733	0.11330	0.63915
t	Busan	PW	0.22070	0.00570	0.41830
t	Busan	AP	0.31614	0.00891	0.58374
t	Ulsan	PW	0.36388	0.14947	0.55413
t	Ulsan	AP	0.48516	0.19078	0.71142
n	Seoul	PW	0.32628	0.12283	0.50576
n	Seoul	AP	0.30923	0.01957	0.60230
n	Busan	PW	0.28056	0.06549	0.47626
n	Busan	AP	0.19736	-0.14884	0.54381
n	Ulsan	PW	0.26010	0.06903	0.46404
n	Ulsan	AP	0.27459	-0.04429	0.57419

Table J.3: Pairwise comparison of region for the same-cross condition differences in the predicted probabilities of a correct answer.

Consonant	Region	Host	Mean	Lower Quantile	Upper Quantile	Probability < 0	Probability > 0
t	Seoul - Busan	PW	0.11049	-0.03253	0.24894	0.0615	0.9385
t	Seoul - Ulsan	PW	-0.03268	-0.16397	0.10689	0.686	0.314
t	Busan - Ulsan	PW	-0.14318	-0.3058	0.01526	0.96125	0.03875
t	Seoul - Busan	AP	0.08119	-0.08198	0.24889	0.16575	0.83425
t	Seoul - Ulsan	AP	-0.08783	-0.24277	0.06394	0.87575	0.12425
t	Busan - Ulsan	AP	-0.16902	-0.3524	0.01077	0.9665	0.0335
n	Seoul - Busan	PW	0.04573	-0.09299	0.18615	0.2595	0.7405
n	Seoul - Ulsan	PW	0.06618	-0.07712	0.20221	0.17725	0.82275
n	Busan - Ulsan	PW	0.02045	-0.14173	0.17459	0.39675	0.60325
n	Seoul - Busan	AP	0.11187	-0.13498	0.41329	0.20625	0.79375
n	Seoul - Ulsan	AP	0.03465	-0.14039	0.2091	0.346	0.654
n	Busan - Ulsan	AP	-0.07723	-0.38483	0.21363	0.70425	0.29575