DEVELOPMENT OF PITCH CONTRAST AND SEOUL KOREAN INTONATION

Sunghye Cho

A DISSERTATION
in
Linguistics

Presented to the Faculties of the University of Pennsylvania in Partial Fulfillment of the Requirements for the Degree of Doctor of Philosophy

2017

Supervisor of Dissertation

Mark Liberman
Christopher H. Browne Distinguished Professor in Linguistics

Graduate Group Chairperson

Eugene Buckley
Associate Professor of Linguistics

Dissertation Committee

Eugene Buckley, Associate Professor of Linguistics

Jianjing Kuang, Assistant Professor of Linguistics

Acknowledgements

I feel grateful to have this moment to thank all people I have been deeply indebted during my Ph.D. life. It has been a tough journey, sometimes feeling lost and disappointed in my own capacity, but many wonderful people supported and prayed for me, and now here I am, writing the acknowledgment to express my gratitude to them.

First, I would like to thank my advisor, Mark Liberman, who contributed his inimitable knowledge to my dissertation project and gave me a guidance to the right direction. He had brilliant answers for every question I asked, which can be found all over this dissertation. Whenever I had a meeting with him, he astonished me with his creative ideas and vast knowledge on any topic. This dissertation wouldn't have existed without his help and support.

My committee members, Gene Buckley and Jianjing Kuang, have played important roles in my dissertation, and also in my grad life. Meetings with Gene have always been joyful and inspiring, and he provided me with keen and practical feedback. Jianjing has served as a role model of an international, female, and academic figure to me, and her questions and feedback during the P-lab meetings were more than helpful.

My special thanks go to other teachers I had during my grad life, including Rolf Noyer, David Embick, Gillian Sankoff, and Georgia Zellou, and my previous advisors at Penn, Bill Labov and Jiahong Yuan. They were always available to have meetings with me, gave me guidance for my projects, and shared their knowledge. I greatly appreciate their help and support. I should also thank Sun-Ah Jun, who laid the foundation for my dissertation project and gave generous feedback. This dissertation is greatly indebted to her theoretical framework. In addition, my thanks go to Eun-jin Oh, my mentor in South Korea, who gave me continuous supports during my PhD study. I feel lucky to have these teachers in my life.

I should also thank my fellow graduate students, who have supported me both intellectually and emotionally. I have been grateful to have wonderful cohort-mates, Mao-Hsu Chen, Sabriya Fisher, Aaron Freeman, and Anton Ingason, who were willing to share their knowledge and have stress-relieving, small chats while going through the challenges and difficulties of the grad life. My thanks also go to my fellow Korean linguists at Penn, Soohyun Kwon, Yong-cheol Lee, Nari Rhee, and Gayeon Son, who have provided me unlimited love and supports. I was blessed to have them not only in my academic journey but also in my personal life, and I wouldn't have survived at Penn without their help.

I would like to express my gratitude to my colleagues and friends in the department, who provided me so much help and friendship. My thanks especially go to Akiva Bacovcin, Aletheia Cui, Amy Goodwin Davies, Aaron Dinkin, Kajsa Djaerv, Duna Gylfadottir, Helen Jeoung, Wei Lai, Marielle Lerner, Ruaridh Purse, Einar Freyr Sigurθsson, Betsy Sneller, Jingjing Tan, Jia Tian, Robert Wilder, and Hong Zhang. Like other linguists in the department, I also owe big thanks to Amy Forsyth, who helped me in everything, literally.

I am indebted to my friends in Philadelphia, Pastor Michael, Beth, Mimi and Michael, Elizabeth and Griffin, Pastor Taehoo, Susan, Delilah, Greta, Jackie, other OFRers, and Jiyu for their love and prayers for me. They made my life in Philadelphia more joyful and bright.

Lastly I wish to thank my family, both in US and in Korea. My husband, Haeman, provided me endless patience, love, and supports. He decided to move to U.S. with me, which was a huge transition in his personal life. He took many household responsibilities without any complaints, and tried to make me laugh whenever I felt depressed and nervous. I feel grateful to him more than I can express here. Also, my thanks go to my two furry sons, Oliver and Leo, who have taught me love for every living creature and opposition to all kinds of discrimination and violence. In addition, I would like to thank both my parent and parents-in-law for their love and prayers. Most importantly, I thank God, who always loves me and guides me in God's way.

ABSTRACT

DEVELOPMENT OF PITCH CONTRAST AND SEOUL KOREAN INTONATION

Sunghye Cho

Mark Liberman

Korean features a three-way contrast among voiceless stops: aspirated, tense, and lenis. Previous studies have shown that young Seoul Korean speakers produce vowels after aspirated and tense onsets with a higher pitch than those after a lenis onset and that the contrast in Voice Onset Time (VOT) between aspirated and lenis merges. This trade-off between VOT and pitch is suggested as an example of tonogenesis, in which an atonal language develops a tonal contrast. By examining 141 Seoul Korean speakers, this dissertation aims to provide a complete picture of the pitch contrast development and the intonation change, from their initial stage to completion.

The findings of this dissertation indicate that all syllables in an Accentual Phrase (AP) starting with a high-inducing consonant is produced with higher pitch values than those in an AP starting with a low-inducing consonant. Furthermore, the results show that females born in the 1950s are the ones who initiated this change, which has almost reached completion in the speech of speakers born in the 1990s. By comparing other languages that underwent tonogenesis, I argue that tonogenesis can affect not only syllables but also large prosodic units and the end result of tonogenesis varies depending on languages' phonological and morphological characteristics.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

It has been suggested that Seoul Korean is undergoing a tonogenetic sound change, where a phrase-initial consonant induces a pitch contrast on the following vowel depending on its laryngeal feature (Kim et al. 2002, Silva 2006, Wright 2007, Oh 2011, Kang 2014). Previous studies have demonstrated that an Accentual Phrase (AP) that starts with a [+spread glottis] or a [+constricted glottis] segment (aspirated and tense stops, coronal fricatives, and /h/) shows a high pitch in the AP-initial syllable, whereas an AP starting with other segments shows a low pitch in the first syllable.[1] Most previous studies have focused on the trade-off between the Voice Onset Time (hereafter VOT) and pitch contrasts among stop consonants, showing that younger speakers are more likely to rely on pitch than VOT in distinguishing lenis stops from tense and aspirated ones and the VOT contrast between aspirated and lenis is merging

[1] The laryngeal feature used to distinguish the three-way Korean stops may vary. Some researchers use [+stiff vocal fold] (adopted from Halle and Stevens 1971), and others use [+spread glottis] and [+constricted glottis]. This dissertation does not address differences between these features, but those who are interested in this topic may refer to Cho et al. 2002, Silva 2006, Wright 2007.

for younger speakers.

Although many questions regarding this change have been addressed and answered in previous studies, one of the most important questions remains unanswered: How does this change affect the intonation system of Seoul Korean? The complicated interaction between pitch contrast (including tone and accent) and intonation has been addressed in many languages from tonal languages like Chinese to less-tonal (so-called *pitch-accent*) languages like Japanese to non-tonal languages like English (Liberman and Pierrehumbert 1984, Pierrehumbert 1980, Poser 1984, Pierrehumbert and Beckman 1988, Shih 1988, Xu 1993, and among others). However, much less is known about how the tonogenetic sound change affects Korean intonation and what kind of prosodic system Seoul Korean is going to have after the tonogenetic change. This dissertation's broad aims are to provide a complete picture of the pitch contrast development and the change of the intonational melody in Seoul Korean from its initial stage to its completion and to shed light on the theory of tonogenesis by conducting both experimental and quantitative studies on Seoul Korean.

1.1 Overview of Korean phonology

1.1.1 Segmental overview

[2]The phonetic symbols used to transcribe the Korean affricates vary in the literature. Also, there has been much debate on the place of articulation of Korean affricates. In this dissertation, /t͡s, t͡s', t͡s^{h}/ (alveolar affricates) are consistently used, following experimental results of Kim (1997, 1999, 2001) and Ko (2013). See Section 2.2.2 for more information.

[3]Korean has other approximants, /j, w, ɰ/, but these are omitted in this table, because they are

Table 1.1: Korean consonant inventory

		Bilabial	Alveolar	Velar	Glottal
Stop	lenis	p	t	k	
	tense	p’	t’	k’	
	aspirated	p^h	t^h	k^h	
Affricate[2]	lenis		$\widehat{ts}$		
	tense		$\widehat{ts}$’		
	aspirated		$\widehat{ts}^h$		
Fricatives	non-tense		s		h
	tense		s’		
Nasal		m	n	ŋ	
Approximant[3]			l		

Table 1.1 shows the Korean consonant inventory. Korean features an unusual three-way contrast among voiceless stops and affricates. Although it is clear that Korean stops and affricates have three laryngeal categories, they are inconsistently described as follows:

(1) Lenis: plain, lax, lenis, unaspirated, slightly aspirated, unmarked, underspecified for laryngeal features

(2) Tense: tense, forced, reinforced, fortis, stiff glottis, constricted glottis, glottalized, geminate

(3) Aspirated: aspirated, heavily aspirated, spread glottis

(Wright 2007:3)

In this dissertation, ***lenis***, ***tense***, and ***aspirated*** are consistently used to describe the three laryngeal categories and to avoid any confusion.

traditionally considered as a part of diphthongs in the Korean phonology.

Furthermore, Korean has a two-way contrast among fricatives. While the alveolar tense fricative /s'/ is clearly described as a tense category, the phonological category of the other fricatives is unclear. In particular, /s/ has been variously described as 'plain,' 'aspirated,' 'lenis,' and 'non-tense.' I simply categorize /s/ (and /h/) as 'non-tense' for now, because the other terms may cause confusion due to sounding similar with one of the laryngeal categories for stops and affricates. This terminology issue is detailed in Section 2.2.1.

There are several phonological rules that apply to the obstruents. The aspirated and tense stops and all fricatives and affricates are neutralized to lenis stops in coda positions. This rule is known as Coda Neutralization, and it has been discussed in many previous studies (Martin 1951, Lee 1961, Kim-Renaud 1974, Kim 1979, Chung 1980, Kim and Jongman 1994, and among others). For example, Kim and Jongman (1994) conduct a production experiment and show that this phenomenon is an example of complete neutralization in regards to the phonetic values of stops and affricates in the coda position.

The lenis category is known to undergo voicing between sonorants, a process known as Lenis Voicing. Since the lenis category is voiceless when word-initially, there has been a debate regarding the underlying representation of the lenis category. While most previous studies have described it as voiceless, Kim and Duanmu (2004) argue that its underlying representation is voiced, which means that it is devoiced word-initially. However, Chang (2007) shows that the voicing of the lenis category is weak even in an intervocalic environment, suggesting that the nature of lenis voicing in Korean is phonetic, not phonological.

Another phenomenon is Post Obstruent Tensing, where a lenis obstruent onset becomes tense after an obstruent coda (Cho 1987, Kang 1992, Jun 1993, among others). Jun (1993) shows that this rule's domain is an AP(Accentual Phrase), which is the smallest prosodic unit in Korean.[4] These rules affect the Korean obstruents, changing their phonetic and/or phonological properties. However, it is important that they apply word-internally, and do not apply to phrase-initial obstruents. As the present dissertation's main focus is the obstruents in the phrase-initial position, the three rules are mentioned in the following chapters only when needed.

1.1.2 Prosodic properties

The most frequently cited theory on Korean prosody is Jun's intonation model (1993, 1996, 1998, 2006), in which there are three levels of prosodic units in the Korean prosodic hierarchy: Intonational Phrase (IP), Intermediate Phrase (ip), and Accentual Phrase (AP). IPs are the highest level in the hierarchy, marking the utterance's final boundary tones (%). For example, declarative sentences are marked with a falling tone (L%) and interrogative sentences are produced with a final rising tone (H%).[5] An ip is the domain of pitch resetting within an utterance. The smallest prosodic unit is an AP, which mostly consists of one or two content words optionally followed by a grammatical marker. The basic melody of an AP is TH-LH, where T stands for either

[4] APs are detailed in Section 1.1.2.

[5] *Tone* in Jun's theory does not refer to lexical tones, but refers to post-lexical, intonational tones like those in the ToBI (Tone and Break Indices) theory.

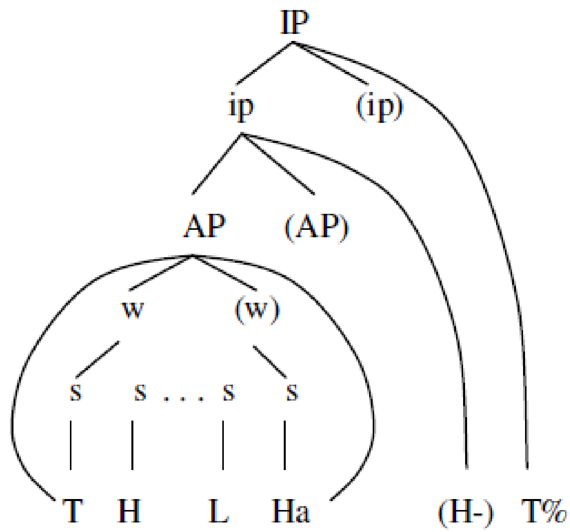


Figure 1.1: Intonation model of Seoul Korean (Borrowed from Jun and Cha 2015)

H (High) or L (Low), depending on the AP-initial onset consonant. If an aspirated or a tense obstruent is the onset of an AP-initial syllable, the AP melody is HH-LH; otherwise, the melody is LH-LH. The first two tones (**TH**-LH) are aligned with two AP-initial syllables, and the last two tones (TH-**LH**) are aligned with two AP-final syllables. When there are more than four syllables within an AP, AP-medial syllables receive pitch values via interpolation from the H tone on the second syllable (T**H**-LH) to the penultimate L tone (TH-**L**H). When an AP has less than four syllables, AP pitch patterns may vary. These patterns are explained in Section 2.3.2.

While most previous studies on Korean intonation and prosody have been based on carefully controlled production experiments (Silva 1992, Kang 1992, Jun 1993, among others), there are a few studies that have examined Korean intonation using a speech corpus. For example, Karpiński and Szalkowska-Kim (2012) compare the intonation pattern of interrogative sentences in Korean and Polish by using the PolnAsia corpus of map-task dialogues (Karpiński 2006), which employs 20 native speakers of Korean (young educated speakers; 13 females and 7 males). Wright (2007) also addresses the AP-initial pitch contrast with the Korean Telephone Conversation corpus (Ko et al. 2003) published by the Linguistics Data Consortium. His analysis of the corpus is mostly a description of several individual speakers and does not provide a dynamic picture of how the AP-initial pitch contrast affects Korean intonation.

1.2 Tonogenetic sound change

Tonogenesis refers to a process in which a previously toneless language develops a tonal contrast and a previously tonal language multiplies its tonal inventory (Haudricourt 1954, Matisoff 1973, Hyman 1978, Hombert 1978, Kingston and Diehl 1994, Thurgood 2002, Kingston 2011, among others). Tonogenesis involves several steps. The first step is that consonantal features, such as a voicing contrast or a laryngeal contrast, give rise to a pitch contrast. At this stage, the pitch contrast coexists with the segmental contrast as a secondary cue. Next, the pitch contrast is exaggerated to the extent that it can no longer be attributed to the phonetic perturbation of consonants. In the final stage, the pitch contrast develops as a tonal contrast when the consonantal contrast that coexisted with the pitch contrast is lost. (See Section 2.4 for examples of the languages that have undergone tonogenesis.)

Tonogenesis is one of the central concerns in this dissertation because many previous studies have described a recent sound change in Seoul Korean as an example of tonogenesis (Silva 2006, Wright 2007, Kingston 2011, Oh 2011, Kang 2014, and among others). Younger Seoul Korean speakers do not have the VOT contrast between aspirated and lenis stops; instead, they tend to rely on the following vowel's pitch to distinguish the two categories. Previous studies have shown that the VOT contrast between aspirated and lenis stops is merging in the AP-initial position for younger speakers (Silva 2006, Wright 2007, Oh 2011, Kang 2014). Furthermore, vowels following tense and aspirated stop onsets are produced with a higher pitch than those following lenis stops in the AP-initial position. (See Section 2.1.2 for more details.) Thus, the situation is similar to the tonogenetic process in that a segmental contrast (in this case, VOT) disappears and a new pitch contrast emerges to maintain the categorical contrast among the stops. The main goals of this dissertation are to investigate i) if this sound change is an example of tonogenesis, ii) how this change affects Seoul Korean intonation, and iii) what the tonogenetic sound change in Seoul Korean implicates in the theory of tonogenesis.

1.3 Research questions

This dissertation addresses two broad research questions:

RQ1) How does the emergence of a pitch contrast at the AP-initial position affect the Korean intonation system in general? When did the change start and what does the current status of Seoul Korean intonation look like? What does the tonogenetic sound change eventually result in and what does the Korean case implicate in the theory of tonogenesis?

RQ2) How far does the AP-initial H-segment's effect reach within an AP or within a large prosodic phrase? Silva (2006) and Kang (2014) show the effect of the AP-initial pitch contrast is also found in AP-second syllables. The second syllable's pitch is higher when the first syllable's onset is a high-pitch inducing type than when it is a low-pitch inducing type. Does the effect of the AP-initial H pitch extend farther than that?

While answering the above two groups of questions, I also aim to answer the following questions to shed a light on less clearly understood aspects of this sound change:

RQ3) Is there difference in tense and aspirated consonants? Several previous studies (Choi 2002, Kim 1994, Lee and Jongman 2012, Silva 2006) have found that aspirated consonants induce a higher pitch than tense consonants, while both of them are higher than lenis stops. If aspirated induces a higher pitch on the following vowel than tense, is this difference also found at the AP-second syllable or at later syllables of the AP?

RQ4) Do affricates and fricatives show the same pitch contrast with stops? A couple of previous studies find that fricatives and affricates also induce pitch contrast AP-initially (Jun 1993 and Cho et al. 2002 for fricatives, Kim 2004 and Perkins and Lee 2010 for affricates). However, it remains to be seen how the pitch contrast from fricatives and affricates interacts with Korean intonation and whether tense and aspirated fricatives and affricates pattern together with tense and aspirated stops.

RQ5) Besides age and gender, what controls the pitch contrast at the AP-initial position? What kind of linguistic elements are involved? For example, does the size of an AP, the AP position within an utterance, or the AP-initial segment type (such as manner of articulation) affect on the AP-initial syllable's pitch height?

The first two groups of questions are answered throughout Chapters 3, 4, 5, and 6. The third research question's answer is found in Section 4.3.1, and the fourth question's answer is found in Section 4.3.2. The fifth research question is answered in Sections 4.2 and 4.3.

1.4 Terminology

In this dissertation, I use *pitch* to refer to a phonetic height in sound, which mostly correlates with the fundamental frequency (f0), whereas *tone* is used as a post-lexical, intonational unit as in the ToBI theory. Please note that this meaning of *tone* is different from the standard definition of *tone* used in phonetics and phonology literature. When *tone* refers to a linguistic unit, where pitch is used in distinguishing the meaning of words, I use *linguistic tones* or *lexical tones* to clarify the meaning. Also, a *pitch contrast* is used to refer to the difference in sound height that comes from phrase-initial consonantal contrasts. A *tonal contrast* is not used in this dissertation because there is between-speaker variation in whether pitch is the primary cue in distinguishing lenis consonants from the others (tense and aspirated) or not.

The difference between a segment-induced pitch and an intonational melody must be mentioned here. This dissertation's use of intonational melody refers to the AP pitch pattern (TH-LH) in Jun's theory, and a segment-induced pitch refers to a slightly raised pitch due to pitch perturbation when articulating voiceless consonants, neither the AP-initial pitch contrast nor the AP intonational tones. Whenever a segment-induced pitch plays a role in the results' pitch patterns, I specifically use the term *segment-induced pitch* or *consonant-induced pitch* so that readers do not confuse it with the APs' intonational melody. When there is a need to distinguish them in writing, I use small letters (*h* or *l*) for a segment-induced pitch and capital letters (H or L) for the intonational melody.

Examples used in the dissertation are basically a phonemic transcription of the Korean alphabet (Hangul) in IPA. However, it should be noted that allophones are not considered when transcribing. For example, Korean /l/ is pronounced as [ɾ] intervocalically or word-initially; however, this alternation is not transcribed.

1.5 Outline of the dissertation

The remainder of this dissertation is organized as follows. Chapter 2 reviews previous studies on Korean consonants and prosody and on tonogenesis in general. In particular, previous findings about the tonogenetic sound change in Seoul Korean are detailed in Section 2.1.2. Chapter 3 describes an experimental study, and I provide the current picture of how the tonogenetic sound change affects the existing Korean intonational melody, employing phone-number strings and natural words. In Chapter 4, I analyze a large-scale speech corpus to provide a diachronic picture of the tonogenetic change development. Section 4.1 introduces the corpus used in Chapter 4, *The Speech Corpus of Reading-style Standard Korean* (The National Institue of the Korean Language 2005). Sections 4.2 and 4.3 use the corpus to examine 59 sentences produced by 118 speakers born in the 1930s–80s. Section 4.2 shows the AP pitch patterns by the speakers' year of birth, gender, and AP size with sentence-initial APs. Section 4.3 compares the laryngeal categories and different manners of articulation using all APs in the 59 sentences. In Chapter 5, I describe how this sound change develops among younger speakers born in the 1990s, using a new speech corpus built for this project. In Chapter 6, I discuss this sound change's phonological aspects and compare the case of tonogenesis in Seoul Korean to those of other languages. Lastly, Chapter 7 concludes this dissertation.

Chapter 2

Literature review

In this chapter, I introduce previous findings on topics related to the main theme of this dissertation: Korean consonants, Korean prosody, and tonogenesis. In Section 2.1, I briefly describe previous studies on VOT and pitch differences among three stop categories and provide other findings on Korean stops, including their aerodynamic and articulatory characteristics. Section 2.2 introduces acoustic and articulatory features of Korean fricatives and affricates. In Section 2.3, I summarize Jun's intonational model of Seoul Korean (1993, 1998, 2000, and in her subsequent works). Lastly, Section 2.4 demonstrates previous findings of languages that underwent tonogenesis.

2.1 Previous studies on Korean stops

2.1.1 Voice Onset Time (VOT)

Korean has a unique three-way contrast in voiceless stops (lenis vs. tense vs. aspirated), and previous studies examine their various features. As for the VOT contrast, previous studies show that older speakers have distinct VOT values for the three stop categories (Lisker and Abramson 1964, Han and Weitzman 1970, Silva 1992, Cho and Keating 2001, Oh 2011, Kang 2014). Figure 2.1 is reproduced from Lisker and Abramson's study (1964), which examines only one speaker. The participant's age and gender is unknown, but he or she is estimated to have been born in the 1930s or 40s. In this figure, it is clear that the participant produces the Korean stops with a three-way VOT contrast (Aspirated > Lenis > Tense). Tense stops have the shortest VOT values, whereas aspirated stops are produced with the longest VOT. Also, the mean VOT is longer for velar stops than bilabial or alveolar stops.

However, recent studies on the VOT of Korean stops reveal that the VOT difference between aspirated and lenis is neutralized for younger speakers (Kim 1994, Choi 2002, Silva 2006, Wright 2007, Kang and Guion 2008, Oh 2011, Kang 2014). Most of these studies find that the VOT of the aspirated category decreases in younger speakers' speech and shows a similar duration with that of the lenis category, confirming a reduction of the VOT difference between the aspirated and the lenis categories.

For example, Silva (2006) conducts a production experiment with 36 adult native

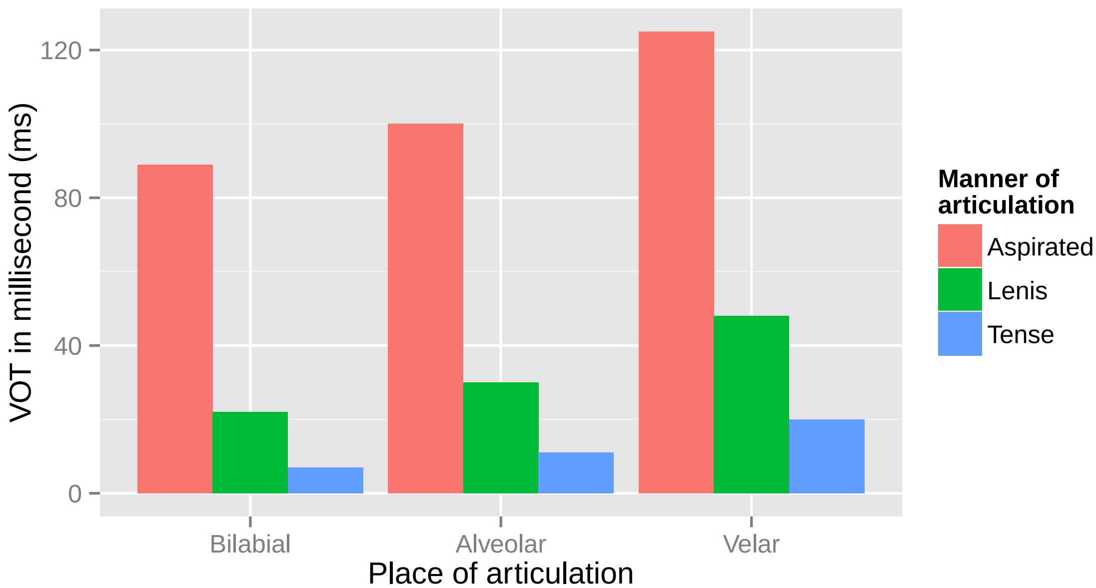


Figure 2.1: Mean VOT values of the three stop categories. This figure is reproduced with data from Lisker and Abramson (1964), whose speaker (age and gender unknown) was likely born in the 1930s or 1940s.

speakers of Seoul Korean (21 females and 15 males) born between 1943 and 1982. He shows that participants can be divided into two groups by their VOT values: traditionalists and innovators. Speakers born before 1965 produce the traditional VOT contrast among stops (VOT: Aspirated > Lenis > Tense), whereas those born after 1965 show the VOT merger between aspirated and lenis (VOT: Aspirated = Lenis > Tense), except one male speaker (born in 1970) who showed the traditional pattern.

Oh (2011) studies the gender difference in VOT values of the stop consonants by conducting a production experiment with 38 native speakers of Seoul Korean (19 males and 19 females, age range: 18–31 at the time of recording). She shows that the

VOT of aspirated stops is significantly longer for males than for females, but that of lenis or tense stops does not significantly differ by gender. Also, she finds that males show a distinctive VOT distribution among the stop categories, but female speakers exhibit an overlapping distribution between aspirated and lenis stops. She explains that the gender difference in VOT cannot be attributed to physiological differences between the genders as previously proposed for English, where females show a longer VOT for long-lag stops than males. In her conclusion, she attributes the gender difference in VOT to the tonogenetic sound change in Seoul Korean.

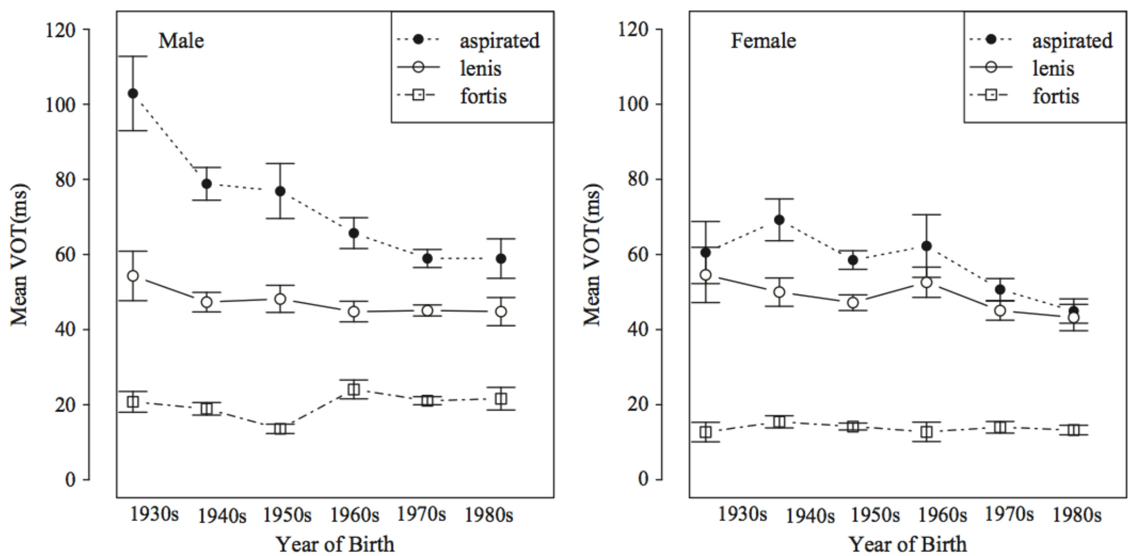


Figure 2.2: By-speaker difference in mean VOT between aspirated and lenis stops plotted against speakers' YOB (year of birth) and gender (Borrowed from Kang 2014). The y-axis shows mean VOT differences between aspirated and lenis stops.

Kang (2014) studies the VOT merger in depth using the Speech Corpus of Reading-style Standard Korean (The National Institue of the Korean Language 2005).[1] In

[1]This corpus is also used in this dissertation to examine the development of the AP-initial pitch contrast and the change of Seoul Korean intonation. See Chapter 4 for the results of the corpus analysis.

Figure 2.2, which is borrowed from Kang (2014, p. 80), the VOT difference between aspirated and lenis stops is smaller for younger male speakers than older male speakers and smaller for female speakers than male speakers. Figure 2.2 clearly shows that female speakers use a two-way contrast in VOT (Aspirated, Lenis vs. Tense), whereas old male speakers use the traditional three-way VOT contrast among stops. Also, it is shown that female speakers born in the 1930s produce aspirated and lenis stops with similar VOT values, suggesting that older female speakers are ahead of their male cohorts with regard to this merger. Kang's study demonstrates that older males mostly rely on the VOT contrast in distinguishing aspirated from lenis stops, whereas younger male speakers use both VOT and pitch cues with a tendency to rely more on pitch than on VOT. Kang suggests that females do not use VOT as a primary cue regardless of their age.

2.1.2 Pitch

There is a cross-linguistic tendency for vowels following voiceless consonants to be produced with a higher pitch than those following voiced consonants (Lehiste and Peterson 1961, Mohr 1971, Jun 1993, Hanson 2009, among others). This tendency suggests that consonants affect the pitch value of the following vowel. Korean stops are also known to have a direct influence on the pitch of the following vowels. However, what is unique about Seoul Korean is that vowels following stop consonants still show a large pitch difference depending on the laryngeal category of the stops, although all Korean stops are voiceless. Furthermore, younger speakers of Seoul Korean produce a much larger pitch difference in the AP-first syllable than those found in other languages.

Previous studies on the Korean stops find that vowels after aspirated or tense consonants are produced with a higher pitch than those after lenis ones at the phrase-initial position (Han and Weitzman 1970, Kagaya 1974, Cho et al. 2002, Silva 2006, Wright 2007, Oh 2011, Kang 2014). For example, Cho et al. (2002) examine four old male speakers of Seoul Korean in their late 50s and early 60s living in Los Angeles at the time of their study. Their study shows that the subjects produce vowels following lenis stops with a lower pitch than those following aspirated or tense stops.

Kang and Guion (2008) study the production of the stop consonants in 18 words produced by 22 native speakers of Seoul Korean. One group consists of 11 young speakers (5 males, 6 females) born after 1977 (age range from 20 to 29 at the time of recording) and the other group consists of old speakers (5 males, 6 females) born before 1966 (age range from 40 to 60 at the time of recording). Their result shows that both groups produce vowels after aspirated and tense stops with a higher pitch than those following lenis stops. However, both groups differ in that older speakers enhance the VOT contrast in clear speech compared to casual speech, whereas younger speakers increase the pitch contrast in clear speech. Their result suggests that older and younger speakers of Seoul Korean use different phonetic cues to enhance the phonological contrast among the stops. Jun (1993) states that the pitch difference between the stop categories is not due to phonetic perturbation of a preceding consonant as found in other languages (Lehiste and Peterson 1961, Mohr 1971, Hanson 2009, among others), but it is rather phonologically encoded in the prosodic system of Seoul Korean.

While most previous studies do not compare gender differences, a couple of the studies examine gender differences in the pitch difference between aspirated and lenis stops (Oh 2011, Kang 2014).[2] However, previous studies show inconsistent results in terms of gender differences. For example, Oh (2011) does not find any gender difference in the pitch values (in St) between lenis and aspirated stops and suggests that female speakers do not use a large pitch difference to compensate the VOT merger. On the other hand, Kang (2014) shows that females make a larger pitch difference than male speakers for the aspirated-lenis contrast, but the gender difference between tense and lenis stops does not reach significance.

While both aspirated and tense stops induce a higher pitch on the following vowel than lenis stops do, previous research also finds a pitch difference between aspirated and tense stops. Many previous studies show that the f0 of an aspirated-initial syllable is higher than that of a tense-initial one (Kim 1994, Choi 2002, Silva 2006, Lee and Jongman 2012). For example, Choi (2002) examines the stop consonants in Seoul Korean and Chonnam Korean with 6 speakers in their late 20s (2 males and 1 female for each dialect). Her result shows that 5 out of 6 speakers show a three-way contrast in pitch, where vowels after an aspirated stop have the highest f0 value, those following

[2]The reason that there are not many previous studies on the gender difference seems to be because it requires normalization of pitch values produced by two genders. It is difficult to come up with a reasonable method of normalizing the gender difference in pitch, because females and males have different pitch ranges. In Chapter 4, I normalize gender differences as well as interspeaker variation in pitch by using a relative semitone scale. See Chapter 4 for more information.

lenis have the lowest f0, and those following tense have an intermediate f0 (Aspirated > Tense > Lenis). In her study, only one Chonnam speaker produces a two-way division with the lowest f0 in the lenis context and a higher f0 in both aspirated and tense contexts (Aspirated, Tense > Lenis). Lee and Jongman (2012) compare the production of stop consonants in Seoul Korean to that of South Kyungsang Korean with 16 male speakers (8 males for each dialect). In their study, the age of speakers of Seoul Korean ranges from 21 to 32, and that of South Kyungsang Korean ranges from 24 to 48. Their result also suggests that the f0 is the highest for the aspirated context, intermediate for the tense context, and the lowest for the lenis context for both dialects.

On the other hand, in her corpus study, Kang (2014) shows that the pitch difference between aspirated and tense is not significant, but the interactions of the tense context with speakers' age and gender are significantly different from those of the aspirated context. This result means that the f0 increment of the tense context by speakers' age is not as large as that of the aspirated context and the gender difference in the tense-initial syllable is smaller than the one found in the aspirated-initial context. Based on her finding, she concludes that while both aspirated and tense are H-pitch inducing contexts, the degree of the f0 enhancement is larger for the aspirated context than for the tense one.

Previous studies also examine the pitch value of AP-second vowels (Lee 1999, Silva 2006, Kang 2014). They find that the pitch value of an AP-second vowel is significantly higher when the phrase-initial consonant is a high-pitch inducing type (aspirated and tense stops) than when it is a low-pitch inducing type (lenis stops). For example, Lee (1999) examines 4 speakers of Seoul Korean (2 males, 2 females, age unknown) and finds that the second syllable of APs starting with a H-pitch inducing stop is produced with a higher pitch than those starting with a L-inducing one, and there is interspeaker variation in the realization of the AP intonational melody.[3] Kang (2014) demonstrates that there is an interaction with speakers' age and the pitch values of AP-second syllables such that the pitch difference between H-initial and L-initial contexts is larger for younger speakers than older ones for both genders (Figure 2.3).

To summarize previous findings on pitch, vowels following tense and aspirated stops are produced with a higher pitch than those following lenis stops. Among the H-pitch inducing stop consonants, vowels following aspirate consonants are higher than those after tense stops. Also, previous studies show that the pitch difference between H-inducing and L-inducing contexts is larger for younger speakers than older ones. Lastly, it is found that the pitch value of the AP-second syllable is higher when an AP starts with a H-pitch inducing consonant than when it starts with a L-pitch inducing one.

[3] His second female speaker (F2) shows a similar pitch pattern to younger speakers (born after 1960), and the first female speaker (F1) shows a similar pattern to females born in the 1950s in Chapter 4 of this dissertation.

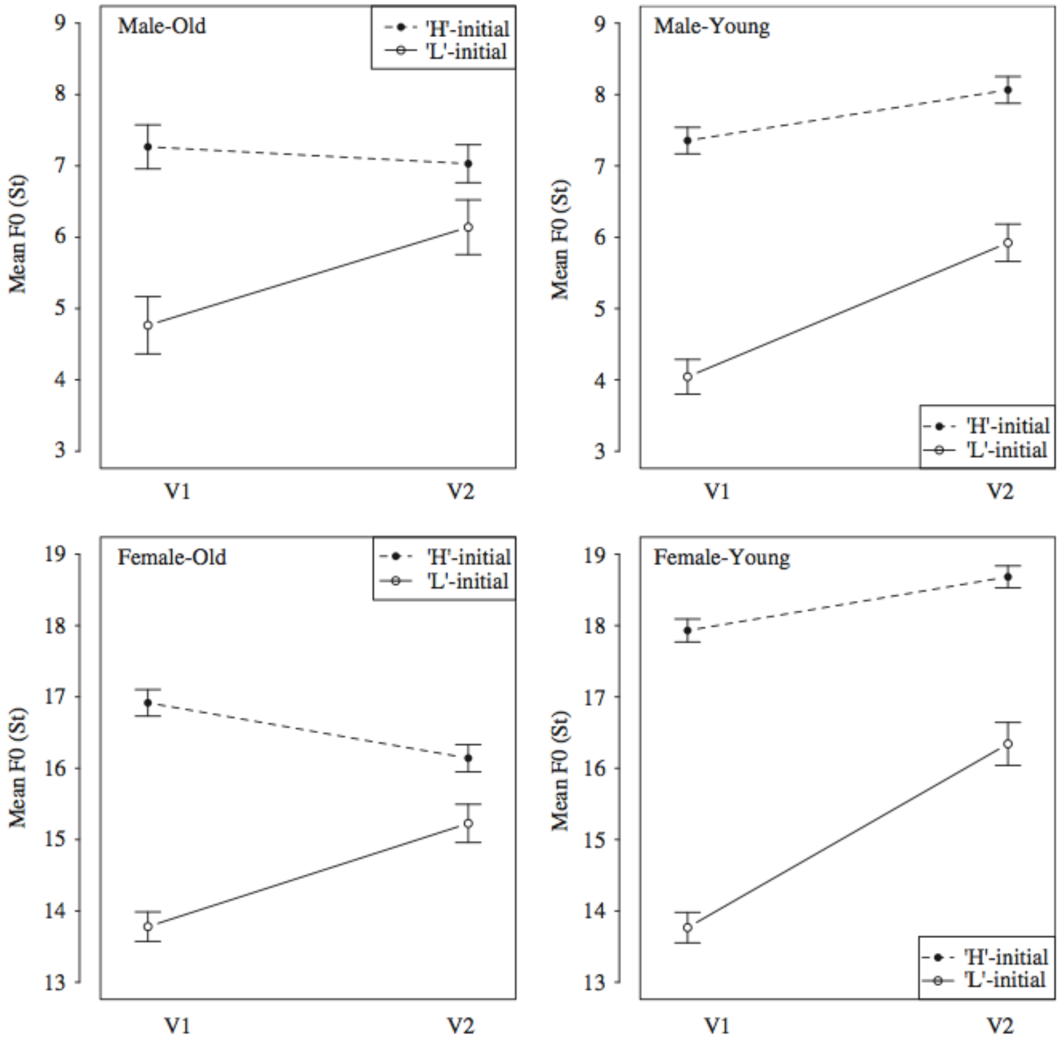


Figure 2.3: The mean f0 at vowel midpoint by word-initial consonant type for male (top row) and female (bottom row) speakers born before 1961 (left column) and after 1961 (right column). (Borrowed from Kang 2014)

2.1.3 Other characteristics

Previous studies find that the three stop categories differ from one another with respect to other acoustic features. For example, a fiberscope study by Kagaya (1974) and an electromyographic (EMG) study by Hirose et al. (1974) show that aspirated stops are characterized by a positive abduction of the vocal folds, heightened subglottal pressure, and suppression of the adductor muscles after the articulatory release, whereas tense stops show a greater vocal tract wall tension than aspirated or lenis and tightly adducted vocal folds before the articulatory release. None of these characteristics are found for the lenis stops.

Cho et al. (2002) also show these differences among the voiceless stops with acoustic measures. They find that H1–H2, which is frequently used to distinguish breathy and creaky phonation, is the largest for aspirated stops and the smallest for tense stops. It means that vowels after a tense stop are more likely to be produced with creaky phonation than those after a lenis or an aspirated stop. Also, they find that H1–F2 (spectral slopes), which is an indicator of the abruptness of vocal fold closure, is the largest for aspirated and the smallest for tense, indicating that a vowel is produced with a more rapid closure of the vocal folds after tense stops than after lenis or aspirated.

Lee and Jongman (2012) examine aerodynamic properties in Korean stop production by measuring both intraoral airflow after the release of the stop closure and the oral air pressure during the stop closure. Their result suggests that the airflow rate is the lowest after the release of tense stops, intermediate after the release of lenis, and the highest following aspirated stops. Also, the air pressure rate during the closure of lenis stops is lower than that of fortis ones, but it is not significantly different from that of aspirated ones.

Cho and Keating (2001) study articulatory properties of the alveolar stops /t, t', t^h/ and the alveolar nasal /n/ by conducting an electropalatography (EPG) study with 3 native speakers of Seoul Korean (2 males and 1 female, age range: 33–37). Their result suggests that the peak linguopalatal contact is greater for /t^h, t'/ (aspirated and tense) than /t, n/ (lenis and nasal) for all speakers. The difference in the peak linguopalatal contact among the three stop categories is the smallest in the utterance-initial position but the largest in the word-initial position of non-AP-initial words, suggesting that all stop categories are produced with a greater linguopalatal contact in the utterance-initial position due to a domain-initial strengthening effect (Fourgeron and Keating 1997). Also, their result demonstrates that the closure duration is longer for the aspirated and tense stops than the lenis and nasal stops. Based on their results, they conclude that both aspirated and tense stops are articulatorily stronger than lenis and nasal stops.

2.2 Previous studies on other Korean consonants

2.2.1 Fricatives

Korean has three fricatives: two coronal ones /s, s'/ and one glottal one /h/ (Table 1.1). Korean coronal fricatives show a two-way contrast: non-tense /s/ and tense /s'/. Previous studies find that the non-tense /s/ has a glottal opening configuration similar to that of aspirated stops and it has a larger glottal opening than /s'/ (Kagaya

1974, Jun et al. 1998).

Cho et al. (2002) compare the two coronal fricatives and find the following:

(1) Centroid frequency is higher for the tense /s'/ than for the non-tense /s/, indicating that the front cavity is smaller before the constriction of /s'/ than /s/.

(2) The pitch value is generally higher after /s'/ than after /s/.

(3) The non-tense /s/ has a longer duration (frication and aspiration) than the tense /s'/, although the tense /s'/ is longer than the non-tense /s/ when only frication is measured.

(4) H1 – H2 and H1 – F2 are higher after /s/ than after /s'/, indicating the non-tense /s/ is produced with a breathier phonation and a slower vocal fold closure than the tense /s'/.

(5) /s'/ is followed by glottalization more than 50% of the time.

Chang (2013) provides a table summarizing previous findings on acoustic differences between the two coronal fricatives (p. 12):

Table 2.1: Acoustic differences between Korean coronal fricatives (Borrowed from Chang 2013)

Property	**/s'/**	**/s/**
Centroid frequency	high	low
Constriction duration	long	short
Aspiration duration	short	long
f0 onset	high	high
F1 onset	low	high
Voice quality	pressed	breathy
Intensity buildup	quick	slow
Following vowel duration	long	short

It is agreed that /s'/ belongs to the tense series in that it is produced with an abrupt vocal fold closure and frequent glottalization before the vowel, and it also induces a higher f0 and tense phonation at the vowel onset (Cho et al. 2002). However, the phonological categorization of the non-tense /s/ has been controversial. For example, Cho et al. (2002) claim that /s/ seems to belong to the lenis series in that the onset of the following vowel exhibits a similar breathy voice quality to those after lenis stops and the f0 of the following vowel after /s/ is still lower than those after an aspirated stop.

Chang (2013) notes the dual nature of the non-tense /s/ in that it is similar to lenis stops with respect to its articulatory properties but also similar to aspirated with respect to timing of the glottal adduction and oral constrictions. Using two phonological features, [± tense] and [± spread glottis], he proposes a hybrid analysis where the non-tense /s/ belongs to a fourth laryngeal category (lenis-aspirated fricative): [– tense] and [+ spread glottis]. In his analysis, other aspirated stops have [+ tense] and [+ spread glottis] features and the tense /s'/ shows [+ tense] and [– spread glottis]. Lenis stops show [– tense] and [– spread glottis] features.

On the other hand, Jun (1993) argues that the non-tense /s/ belongs to the aspirated series because its phonetic values are similar to those of aspirated stops and it induces a high pitch in AP-initial positions. Kang (2014) compares /h/ to other aspirated stops in her corpus study and shows that /h/ patterns together with aspirated stops for the phrase-initial pitch contrast. Since the focus of this dissertation is the f0 pattern of AP-initial consonants, following Jun (1993) and Kang (2014), I also treat /s, h/ as aspirated fricatives in the following sections.[4]

2.2.2 Affricates

Korean affricates show a three-way contrast like Korean stops: lenis /t͡s/ vs. tense /t͡s'/ vs. aspirated /t͡s^h/. There has been much debate on the place of articulation of Korean affricates, and the IPA symbols for the affricates also vary in the literature.[5] Some studies argue that Korean affricates are post-alveolar (Pandeli 1993, Shin 1996). Pandeli (1993) examines one native speaker of Korean, and observes that Korean affricates are palato-alveolar in both word-initial and word-medial positions. Shin (1996) argues that Korean affricates are alveolo-palatals in intervocalic positions, based on her EPG study of one native speaker of Korean. Both studies suggest that /t͡ʃ, t͡ʃ', t͡ʃ^h/ are the accurate IPA symbols for Korean affricates.

On the other hand, there are other recent studies suggesting that Korean affricates are in fact alveolar, not post-alveolar (Kim 1997, 1999, 2001, and in her subsequent works; also in Ko 2013). For example, Kim (2001) examines four speakers of Korean and the results of her study of palatograms, linguograms, and LPC (F2 transition) analysis find that the place of articulation of Korean affricates are alveolar. Similarly, in Ko's study (2013) of static palatography, the results suggest that Korean affricates are alveolar sounds. In this dissertation, I adopt the proposals made by Kim (1997, 1999, 2001) and Ko (2013), and use /t͡s, t͡s', t͡s^h/ as the IPA symbols of the Korean

[4] The results of the corpus study in this dissertation also show that the pitch trajectory of APs starting with these fricatives (/s, h/) is not different from that of APs starting with an aspirated stops. See Chapter 4.

[5] The most frequently used symbols for Korean affricates seem to be /č, č', č^h/.

affricates.

Much less is known about phonetic differences among the lenis, tense, and aspirated affricates. Kagaya (1974) shows that a tense affricate has similar characteristics to tense stops, such as the completely adducted and stiffened vocal folds before the articulatory release, abrupt closure of the vocal folds near the voice onset, increased subglottal pressure, and a lowered glottis right before the release, whereas aspirated stops and affricates are associated with a positive abduction of the vocal folds and heightened subglottal pressure. Kagaya also provides f0 values of the following vowels after stops and affricates, but those are not cited here because the values are too high, compared to other studies, and the context in the experiment is /VCV/, where the consonant position is not phrase-initial.

A couple of recent studies show that aspirated and tense affricates induce a high pitch on the following vowel (Kim 2004, Perkins and Lee 2010). Kim (2004) conducts both production and perception experiments to examine Korean stops and affricates. In her production experiment, 4 native speakers of Seoul Korean (2 females and 2 males; age unknown) read 12 monosyllabic nonsense words /ka, k^ha, k'a, pa, p^ha, p'a, ta, t^ha, t'a, t͡sa, t͡s^{h}a, t͡s'a/. The results of her production study demonstrate that aspirated and tense affricates induce a higher pitch on the following vowel than a lenis one. Although she does not point it out, an aspirated affricate also induces a higher pitch than a tense one in all subjects (Aspirated affricate > Tense affricate > Lenis affricate). The results of her perception experiment with 14 native speakers of Seoul Korean in their 20s suggest that the participants tend to identify stimuli as an aspirated affricate when the f0 is high and they are likely to identify the stimuli as a lenis affricate when the f0 is low. This pattern is also found in the results for stops. Similarly, Perkins and Lee (2010) show that vowels following aspirated and tense affricates are produced with a higher pitch than those following a lenis one, and the pitch values of the affricate context are comparable to those of the stop context.

However, the previous studies do not employ many speakers (2 females and 2 males in Kim 2004, and 2 females in Perkins and Lee 2010), so it is hard to see how the pitch contrast from the affricate consonants has developed over time, how gender and age are involved in the development of the phrase-initial pitch contrast from affricates, and how similar the pitch contrast from the affricates is to that of the stop consonants. Also, it remains to be seen how far the effect of phrase-initial affricates reaches within an AP.

2.3 Previous studies on Korean prosody

2.3.1 Prosodic structure of Korean

Two major approaches exist in the literature of Korean prosody: an indirect syntactic approach (Cho 1987, Kang 1992, Silva 1992, among others) and an intonational approach (Jun 1993, 1996, 1998, and in her subsequent works). Jun proposes that Korean prosody is better analyzed with the intonational approach than the syntactic one. In particular, she suggests that an Accentual Phrase (AP) is the domain of several phonological rules, which a Phonological Phrase (hereafter PhP) in the syntactic approach fails to explain. In this section, I summarize major points of her proposal and theory.

Jun (1993, 1998) shows that the domain of three phonological rules, post-obstruent tensing, lenis stop voicing, and vowel shortening, is an AP, not a PhP. She conducts several production experiments with native speakers of Seoul and Chonnam Korean. Two phonological rules, post-obstruent tensing and lenis stop voicing, out of three are examined for Seoul Korean because Contemporary Seoul Korean has lost the vowel length distinction. The two rules and her results are introduced here.

Jun (1998) examines the domain of the post-obstruent tensing rule in Seoul Korean, where a lenis stop consonant becomes a tense stop when preceded by another lenis stop. In this experiment, four native speakers of Seoul Korean are employed. For example, a verb phrase /mijʌk**k**uk **p**ʌljʌ/ 'Throw out the seaweed soup' has two target lenis stops that are subject to the tensing rule. The first target is the onset /k/ in the last syllable of /mi.jʌk.**k**uk/ 'seaweed soup', which is preceded by another lenis stop /k/, and the second target is the onset /p/ of the verb /**p**ʌ.ljʌ/ 'throw-out', which is preceded by the coda /k/ of the noun. While a verb phrase (Verb + its object Noun Phrase) is expected to form one PhP and the tensing rule always applies in both contexts in the syntactic approach, Jun shows that the tensing rule does not always apply to the second target segment /p/. She finds that the tensing rule only applies when the entire verb phrase is pronounced as one AP, {mijʌkkuk **p**ʌljʌ}, and the rule does not apply when the verb phrase is phrased with two APs, {mijʌkkuk} {pʌljʌ}. The result indicates that the tensing rule variably applies depending on whether an AP boundary intervenes between target and trigger segments or not.

In another experiment, Jun (1993) investigates the lenis stop voicing rule, where a voiceless lenis stop becomes voiced between two voiced segments. One female and two male speakers of Seoul Korean in their late 20s (at the time of her study) are examined.[6] For example, a noun phrase /kʌmɨn **k**ojaŋie **p**almok/ 'black cat's ankle' has two segments that are subject to the voicing rule. The first target consonant is /k/ in /kojaŋie/ 'cat's', which is between a sonorant and a vowel, and the second target segment is /p/ in /palmok/ 'ankle', which comes between two vowels. The noun phrase is expected to form one PhP in the syntactic approach and the voicing rule applies to both target segments. However, the result in Jun (1993) shows that the prosodic phrasing of the NP varies and when the phrase is produced with three APs, {kʌmɨn} {kojaŋie} {palmok}, the target segments, /k/ in /kojaŋie/ and /p/ in /palmok/, are not affected by the voicing rule and remain voiceless. This indicates that the prosodic phrasing of Korean and the voicing rule are not straightforwardly predicted by syntactic structures.

Based on the results, Jun argues that the syntactic approach has several limitations. Firstly, the mapping between the actual phrasing and the PhP boundaries is not perfect. While PhP boundaries are not flexible because they are dependent on syntactic structures, the actual phrasing of an utterance is flexible. Also, PhP phras-

[6]Although she does not state the exact age of the participants, it is likely that these speakers were born in the 1960s.

ing changes depending on which version of syntactic theory is used to predict PhP boundaries or which syntactic category is assumed for a prosodic phrase. Lastly, the syntactic approach explains the prosodic hierarchy only based on the distribution of allophones and segmental phenomena, not considering suprasegmental features.

2.3.2 Tonal patterns of Accentual Phrases

In Jun's framework (1993, 1996, 1998, 2000, and in her subsequent works) introduced in Chapter 1.1.2 of this dissertation, the smallest prosodic unit is an AP. She states that AP tonal patterns vary depending on the size of APs. The first two rows in Figure 2.4 show the possible tonal patterns of short APs. The first row exhibits 5 tonal patterns when the final AP boundary tone is H, and the second row illustrates the tonal patterns when the final boundary tone is L. The figure shows that the AP-medial +H and L+ tones can be variably undershot in short APs because there are not enough Tone Bearing Units for all four AP tones.

The last row shows 3 possible AP tonal patterns for longer APs. The first two patterns demonstrate AP tonal patterns when the final boundary tone is H, and the other pattern is when the final boundary tone is L. She also notes that the +H tone (T**H**-LH) is loosely linked with the second syllable and AP-medial syllables receive their tonal values via interpolation from +H to L+ (T**H**-**L**H) when an AP is longer than 4 syllables.

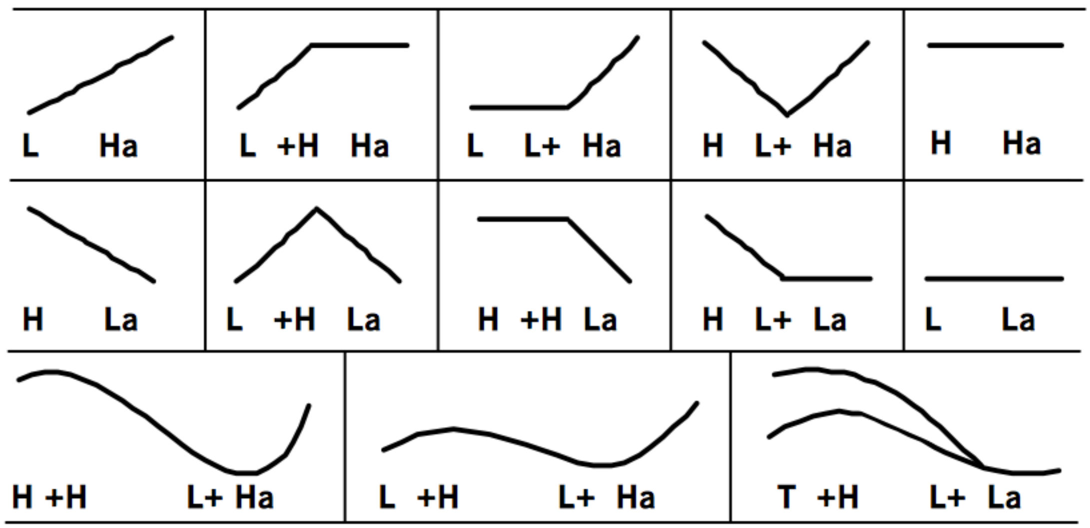


Figure 2.4: Schematic f0 contours of fourteen surface tonal patterns of APs (Borrowed from Jun 2000). "+H" is for the first H tone in the initial two tones (T**H**-LH)and "L+" represents the penultimate L tone (TH-**L**H). "a" stands for AP boundary tones.

2.4 Previous studies on tonogenesis

Tonogenesis refers to a process where previously toneless languages develop a tonal contrast and previously tonal languages multiply their tonal inventories by a tone split (Haudricourt 1954, Hombert 1978, Hyman 1978, Kingston 2011, Matisoff 1973, Ohala 1973, Ohala 1978, Thurgood 2002, among others). The first step of tonogenesis is that consonantal features, such as a voicing contrast or a laryngeal contrast, cause a pitch contrast. Next, the pitch contrast increases to the extent that it can no longer be considered as phonetic perturbation of consonants. In the final stage, the language acquires a tonal contrast when the consonantal contrast that coexisted with the pitch contrast is lost. The most well-studied pattern is the trade-off between a voicing contrast and a tonal contrast; voiceless consonants develop into a high tone and voiced consonants into a low tone.

Previous studies provide phonetic explanations of why consonants have an effect on following or preceding vowels. A crosslinguistic tendency, which is often called 'pitch skip', is that a vowel after a voiced consonant is produced with a lower pitch than the one after a voiceless consonant due to automatic phonetic mechanism (Lehiste and Peterson 1961, Mohr 1971, Jun 1996, Hanson 2009, among others). Kingston (2011) explains that the reason that voiced consonants induce a low pitch on the following vowel is because of an aerodynamic conflict when producing voicing and obstruents together. Producing a voiced sound requires intraoral air pressure (P_o) lower than subglottal air pressure (P_s) so that airflow keeps vibrating the vocal folds, flowing up from the lungs to the oral cavity ($P_o < P_s$). On the other hand, when producing an obstruent, P_o increases due to the noise characteristic of obstruents. To resolve this conflict, speakers lower the larynx to slow down the rise of P_o by enlarging the oral cavity. When the larynx is lowered, speakers tend to lower pitch (Ewan 1976, Honda et al. 1999), resulting in a lower pitch after voiced stops than after voiceless ones.

Ohala (1973) notes that the airflow rate is high after the release of voiceless aspirated consonants, because it faces little resistance due to the open glottis. This results in more vibrations of the vocal folds and a higher pitch. Kingston (2011) also states that the different status of the vocal folds may induce a pitch difference in neighborning vowels. With an electromyographic study, Löfqvist et al. (1989) show that voiceless consonants are associated with a higher level of cricothyroid muscle activity than voiced ones. They suggest that the increased activity of cricothyroid muscle induces a longitudinal tension of the vocal folds, increasing the frequency of glottal vibration. In a more recent study, Hanson (2009) shows that the pitch value of a vowel following a voiced stop is not as low as the one following a nasal. If the effect of 'pitch skip' stemmed from voicing, we would not expect to see such a difference. Based on this result, she suggests that the most potential cause of pitch skip (the pitch difference in vowels following voiced and voiceless consonants) is the status of the vocal folds (stiffened or not), rather than voicing itself.

Tonogenesis takes place when this intrinsic pitch difference of consonants is exaggerated and reinterpreted as a tonal contrast by speakers. Languages that underwent tonogenesis in their history differ from one another in i) which consonantal feature induces a tonal contrast, ii) whether a following or a preceding vowel is affected, and iii) which linguistic unit changes by tonogenesis. In the following sections, I provide examples of tonogenetic processes in several languages.

2.4.1 Vietnamese

Haudricourt (1954, 1972) proposes the process of tonogenesis in Vietnamese. His proposal is that consonants had played a central role in the process of Vietnamese tonogenesis (Table 2.2). In the initial stage, post-vocalic consonants gave rise to a three-way contrast of pitch contours. A final stop consonant developed a rising tone, and a final fricative induced a falling tone. A level tone developed from either open syllables or syllables with a nasal coda consonant. In a later stage, these three tones split into two different pitch heights depending on onset consonants. Syllables with a voiced onset developed a low tone, and those with a voiceless initial acquired a high tone, resulting in a six-way tonal contrast (2 pitch heights x 3 contours).

Table 2.2: Vitenamese tonogenesis in the consonant-based approach (borrowed from Kingston 2011). “T” stands for any following stops, and the Vietnamese tone names are in parentheses.

		Following consonants		
		CV, CVN	CVT	CVs, CVh
Preceding consonants	voiceless	*pa > pa high level (ngang)	*pak > pak high rising (sac)	pas > pa high falling (hoi)
	voiced	*ba > pa low level (huyen)	bak > pak low rising (nang)	bas > pa low falling (nga)

Table 2.3: Vitenamese tonogenesis in the laryngeal model (in Thurgood 2002). “T” and “S” stand for stops and fricatives respectively. (Vietnamese tone names)

Finals:	CVN, CV	CVN, CV	CVT	CVS
Register:	clear	creaky	? (> creaky)	?
voiceless (clear) onset	pa high level (ngang)	pa high rising (sac)	pak high rising (sac)	pa high falling (hoi)
voiced onset > breathy	pa low level, breathy (huyen)	pa low rising (nang)	pak low rising (nang)	pa low falling (nga)

On the other hand, Thurgood (2002, 2007) argues that the Vietnamese tonogenesis can be adequately explained only if a phonation distinction is considered (Table 2.3). He points out that there are examples where Haudricourt’s consonant-based model fails to explain. For example, in other Mon-Khmer languages, many words with a level tone have a glottal stop coda (CVʔ), neither a nasal (CVN) nor a zero coda (CV) as Haudricourt’s model expects. Also, a rising tone in Chong, a language spoken in

Cambodia and Southeastern Thai, shows unexpected final glottalization, not stop coda consonants as Haudricourt's consonant-based model expects.

Diffloth (1989) accounts for these unexpected discrepancies by reconstructing a phonation contrast between modal and tense voice. In his analysis, level tones developed from modal voice, whereas rising tones acquired tense voice. A final voiceless fricative with modal voice resulted in a falling tone. Also, as for the role of onset consonants, Ratliff (1997) suggests that the pitch height contrast is correlated with prevocalic voice-quality distinction (breathy vs. modal) in the proto language (cited in Thurgood 2002). Based on these observations, Thurgood (2002, 2007) argues that the contrast of laryngeal gestures associated with particular classes of consonants plays a crucial role in tonogenesis, not consonants per se.

2.4.2 Other languages

Previous research on Vietnamese tonogenesis has stimulated studies on tonogenesis in other languages. For example, previous studies on Khmu (also known as Kammu), a language spoken in northern Southeast Asia including parts of Thailand, Laos, and Northwestern Vietnam, show that a low tone in a tonal dialect corresponds to a word-initial voiced stop or a sonorant in a non-tonal dialect, whereas a high tone in the tonal dialects corresponds to a voiceless stop in the non-tonal dialect (Suwilai 2004, Kingston 2011). Table 2.4 illustrates this dialectal difference.

In Table 2.4, it is clear that the voicing contrast in the Eastern dialect corresponds

Table 2.4: The correspondences of tone and onset consonants in three dialects of Khmu. The *lb* dialect is spoken in Muang Hun, Udomsaj, and Laos, and the *ct* dialect is spoken in Om Kae village, Sipsongpanna, and Yunna (China). The data is from Suwilai 2004.

English gloss	Eastern Khmu (non-tonal)	Western Khmu (*lb*) (tonal)	Western Khmu (*ct*) (tonal)
'rice wine'	bu:c	pʰù:c	pù:c
'to take off clothes'	pu:c	pú:c	pû:c
'to cut down a tree'	bok	pʰòk	pòk
'to take a bite'	pok	pók	pók
'stone'	gla:ŋ	kʰlà:ŋ	klà:ŋ
'eagle'	kla:ŋ	klá:ŋ	klâ:ŋ
'to weigh'	ɟaŋ	cʰàŋ	càŋ
'astringent'	caŋ	cáŋ	câŋ

to the tonal contrast in the Western dialects.[7] The difference between a high tone and a rising-falling tone in the two Western dialects depends on whether a vowel is short and if a coda consonant is a nasal consonant or not. Long vowels developed a rising-falling tone as in [pû:c] 'to take off clothes' in the *ct* dialect, and a short vowel followed by a nasal consonant also developed a rising-falling tone as in [câŋ] 'astringent'.

Utsat (also known as Tsat or Hainan Cham) is another Southeast Asian language that underwent tonogenesis (Thurgood 1992). Thurgood (1992) states that the first tonal split was between voiceless and voiced coda consonants. The voiceless finals further split into three tone categories (55, 53, 35), depending on the category of the final consonants. Syllables ending in proto Chamic **-h* developed a high level tone

[7] Other Western dialects show a phonation contrast that corresponds to the voicing contrast of the Eastern dialect. For more information, see Suwilai 2004.

(55), whereas those ending in a stop developed a high falling tone (53), leaving a glottal stop reflex. Syllables that completely lost final stops acquired a mid rising tone (35). As for proto voiced coda consonants, when a syllable has a voiced onset consonant with a voiced coda, the syllable developed a low level tone (11). The other cases with a voiced coda evolved into a mid level tone (33). See Table 2.5 for examples.

Table 2.5: Correspondences of tone in Utsat and coda consonants in proto-Chamic. Cognates in Chru are also given for comparison. Examples are from Thurgood 1992.

	Proto-Chamic	Utsat	Chru	English gloss
Voiceless coda	*dilah	la^{55}	dəlah	‘tongue’
	*tanah	na^{55}	tənah	‘earth’
	*do:k	thoʔ^{53}	dò	‘live’
	*tikiʔ	kiʔ^{53}	təkì	‘few’
	*pa:ʔ	pa^{35}	pà	‘four’
	*sakiʔ	ki^{35}	-səkì	‘sick, painful’
	*ʔana:k	na^{35}	anà	‘child’
Voiced coda	*lapa	pa^{33}	ləpa	‘hungry’
	*ʔapui	pui^{33}	aphi	‘fire’
	*ʔikan	ka:n^{33}	akàn	‘fish’
	*ʔatas	ta^{33}	—	‘far’
	*dua	thua^{11}	dua	‘two’
	*batəu	tau^{11}	pətəu	‘stone’

A noticeable difference between proto-Chamic and Utsat is that disyllabic words in proto-Chamic became monosyllabic words in Utsat. Thurgood (1992) explains that this is due to the deletion of presyllables. In proto-Chamic, stress fell on the final syllable of disyllabic words, and presyllables (sometimes known as sesquisyllables) were unstressed and had a reduced vowel. Some presyllables dropped without any trace (e.g., *ʔ*apui* > *pui*^{33} 'fire'), whereas in other cases, the deleted voiced stop resulted in a Low level tone of the next syllable without changing the voicing of the following onset (e.g., **batəu* > *tau*^{11} 'stone'). The deletion of presyllables resulted in many monosyllabic words in Utsat (Thurgood 1992).

The cases of Bukawa and Yabem, Oceanic Austronesian languages spoken in Papua New Guinea, are also interesting. Both languages stemmed from Proto North Huon Gulf (herafter PNHG), and they share many similarities. For example, both languages have foot structures with iambic stress patterns. However, Ross (1993) states that the two languages differ from each other in that tones are predictable in Yabem, whereas tones are unpredictable in Bukawa. Ross explains that this is due to tonogenesis that PNHG underwent. During the development of PNHG from Proto Huon Gulf (PHG), vowels with a voiceless obstruent developed a H tone and those with a voiced obstruent acquired a L tone. Tone spread also occurred such that all vowels in a morpheme (a foot) had the same tone.[8] In Yabem, one of the daughter languages of PNHG, the voicing contrast and tonal contrast still coexist. However, in Bukawa, this consonant-tone harmony is destroyed because the voicing contrast is lost (Table 2.6).

In Athabaskan (Kingston 2005, 2011), following consonants and vowel length

[8] Ross (1993) states that the rule of tone spreading (or tone harmony) is that weak syllables of an iambic foot acquired the tone and the voicing of a strong syllable only when the strong one started with a stop in PHG. If a strong syllable did not start with a stop and a weak syllable started with a voiced stop, the tone and voicing of a weak syllable spread to the strong one. I discuss this point in Chapter 6.3 in detail.

Table 2.6: Tone and voicing in Bukawa and Yabem. Examples are from Ross 1993.

Bukawa	Yabem	English gloss
gá-kúŋ	ká-kúŋ	'I called out'
gà-kùŋ	gà-gùŋ	'I speared (something)'
gá-táŋ	ká-táŋ	'I weep'
gà-tàŋ	gà-dèŋ	'I move towards'

played a role in tonogenesis. Syllables with a long vowel developed a marked tone when the following consonant was glottalic in Proto-Athabaskan.[9] The only exception is when the final consonant was a glottalic stop, which developed an unmarked tone. Also, syllables with a reduced vowel developed a marked tone when the following consonant is glottalic in Proto-Athabaskan.

Table 2.7: Athabaskan tonogenesis. 'M' stands for a marked tone, and 'U' stands for an unmarked tone. Examples are from Kingston 2005.

Rime shapes	Proto-Athabaskan	Chipewyan (High-marked)	Gwich'in (Low-marked)	Hupa (Atonal)	English gloss
VK (> U)	*ɬəd	ɬə̀r	ɬád	ɬid	'smoke'
VK' (> M)	*wət'	bə́r	vàd	mət'	'belly'
VR (> U)	*ts'ən	tθ'ə̀n	tθ'án	ts'iŋ	'bone'
VR' (> M)	*qʊn'	kún	kǫ̀ʔ	xoŋ'	'fire'
V:R (> U)	*kʲa:n	tʃą̀	tsį́	kʲaŋ	'rain'
V:R' (> M)	*kʲʷa:n'	tsą́	trį̀ąʔ	tʃʷaŋ'	'excrement'

The last study introduced in this section is the tonogenetic change in Kurtöp, a

[9]It varies among daughter languages whether a marked tone is H or L. The tone developed from a final glottalic consonant is called 'marked', because the opposite tone developed in other phonetic environments (Kingston 2005). For example, Chipewyan is a high-marked language, and Gwich'in is a low-marked language.

Tibeto-Burman language spoken in northeastern Bhutan (Hysop 2009). In Kurtöp, tone is contrastive and unpredictable after sonorants and palatal fricatives (either H or L). Kurtöp has a three-way contrast among stops: (voiceless) aspirated, (voiceless) unaspirated, and voiced ones. Hysop (2009) conducts a production experiment with two male speakers of Kurtöp and finds that vowels following stops also show a large pitch difference, and this difference is larger for the young speaker than the old speaker. The results of Hysop's study reveal that vowels following voiceless aspirated and unaspirated stops are produced with a higher pitch than those following voiced stops. By showing that the pitch difference is maintained throughout the end of monosyllabic words, she suggests that the pitch difference among stops is not due to phonetic perturbation found in other languages. Also, her results of VOT measurements demonstrate that the VOT values of voiced stops show a bimodal distribution, indicating that voiced stops are merging into voiceless ones in terms of VOT, in favor of a tonal contrast. The case of Kurtöp reported in her study is quite similar to the tonogenetic change of Seoul Korean in that the VOT contrast is being traded off with a pitch contrast, although Korean does not have a voiced stop.

To sum up, previous studies on tonogenesis find the following:

(1) In general, a contrast in consonants or laryngeal gestures in producing consonants may develop into a tonal contrast over time.

(2) Dialects of the same language may show a correspondence between voicing of onset consonants and tonal contrast. A voiceless onset in Eastern Khmu corresponds to a high tone in tonal dialects of Western Khmu, and a voiced onset in Eastern Khmu corresponds to a low tone in Western Khmu.

(3) The voicing of coda consonants may develop into a tonal contrast as in Utsat.

Voiceless coda consonants in Proto-Chamic generally correspond to high tones in Utsat, and voiced finals in Proto-Chamic correspond to low tones in Utsat.

(4) Glottalic gestures in the coda position also induce a tonal contrast, as illustrated in Athabaskan tonogenesis.

(5) A tonal development in one language may accompany other sound changes, such as the loss of a voicing contrast (in Vietnamese and Bukawa), the deletion of sesquisyllables (in Khmu and Utsat) or tone harmony within a foot (in Yabem and Bukawa).

Bearing these findings in mind, I look at the tonogenetic sound change in Seoul Korean in the following chapters. Chapter 3 examines the synchronic pattern of the change with young speakers of Seoul Korean, and Chapter 4 investigates the diachronic pattern of the change by conducting an apparent-time study with a large-scale speech corpus. Chapter 5 provides the most recent picture of Seoul Korean intonation with speakers born in the 1990s. In Chapter 6, I discuss the implications that the findings of Chapters 3, 4, and 5 have for the theory of tonogenesis.

Chapter 3

Experiment

To investigate the current picture of the AP-initial pitch contrast in Seoul Korean intonation, a production experiment was conducted with 8 speakers of Seoul Korean in their 20s. In the experiment, phone-number strings and natural words were employed to investigate the overall effect of the pitch contrast and segmented-induced pitch within larger prosodic units.[1]

3.1 Speakers and stimuli

In this experiment, 8 speakers of Seoul Korean participated (4 female and 4 male speakers in their 20s; mean age, 23.88 years). At the time of recording (2014), they were all studying in the U.S. Most of them (7 out of 8) reported their length of residence in the U.S. as less than one year. One male speaker (27 years) reported he

[1] A substantial part of this chapter was published as a journal article in 2016: Cho, Sunghye and Yong-cheol Lee. 2016. The effect of the consonant-induced pitch on Seoul Korean intonation. *Linguistic Research 33*(2), 299–317.

had lived in the U.S. for three years, but used Korean daily to talk with his spouse at home.

In the experiment, phone-number strings and natural words were employed as stimuli. For both types, the onset consonants of all syllables of the stimuli were controlled for four contexts (*hh, hl, lh*, and *ll*) to examine the effect of the consonant-induced pitch from onset consonants (*h* or *l*) on Seoul Korean intonation (TH-LH).[2] In the high-high (*hh*) context, all syllables of the target strings/words started with high-pitch inducing consonants (aspirated).[3] In the low-low (*ll*) context, all syllables started with low-pitch inducing segments (lenis, sonorant, or vowels). In the other two contexts (*hl* and *lh*), low-pitch inducing types and high-pitch inducing types were alternated with each other so that *h-l-h-l*... or *l-h-l-h*... sequences were constructed. To rule out any morphological factors, 10 different phone-number strings were employed for each pitch context (*hh, hl, lh*, and *ll*) following the methods used in Lee (2015). The positions of the digits were controlled so that each digit (0–9) occurred similarly often in each position. In all phone-number strings, the first three numbers grouped together as an area code, and the last four digits also grouped together. The number of digits in the middle varied from 2 to 5 to see how the segment-induced pitch and the intonational melody are realized in different AP sizes. The high-pitch inducing digits for the phone-number strings include 3 /sam/, 4 /sa/, 7 /$\widehat{\text{ts}}^{\text{h}}$il/, and 8

[2]The segment-induced pitch *h* or *l* is marked with small (italicized) letters to differentiate them from the intonational tones (TH-LH) in Jun's theory, which are represented in capital letters.

[3]H-initial digits in Korean start with an aspirated onset, except 1 /il/, so natural words with aspirated onsets are selected as the target words (Table 3.2). The comparison between aspirated and tense is examined in detail in Sections 4.2.1 and 4.3.1.

/p^{h}al/, whose onset is an aspirated consonant, as well as 1 /il/, which is reported to be pronounced with a high pitch even though it is vowel-initial (Jun and Cha 2015).[4] The other digits, 2 /i/, 5 /o/, 6 /juk/, 9 /ku/, and 0 /koŋ/, are classified as low-pitch inducing digits. The total number of phone-number strings was 1,280 (4 contexts x 4 AP sizes x 10 digit strings x 8 speakers), and the digit strings were presented within a carrier sentence, /ne.pʌn.ho.nɨn 000–00...–0000 (i).ya/ 'My phone number is 000–00...–0000.'[5] Schematic forms and examples of the phone-number strings used in the experiment are presented in Table 3.1.

Table 3.1: Stimuli examples for the phone-number strings

	2-syllable	5-syllable
hh	*hhh–hh–hhhh* ex) 384–71–8148 /sam.p^{h}al.sa/–/t͡s^{h}il.il/–/p^{h}al.il.sa.p^{h}al/	*hhh–hhhhh–hhhh* ex) 331–43371–8841 /sam.sam.il/–/sa.sam.sam.t͡s^{h}il.il/ –/p^{h}al.p^{h}al.sa.il/
hl	*hlh–hl–hlhl* ex) 804–35–3585 /p^{h}al.koŋ.sa/–/sam.o/–/sam.o.p^{h}al.o/	*hlh–hlhlh–hlhl* ex) 803–70157–8539 /p^{h}al.koŋ.sam/–/t͡s^{h}il.koŋ.il.o.t͡s^{h}il/ –/p^{h}al.o.sam.ku/
lh	*lhl–lh–lhlh* ex) 230–58–6154 /i.sam.koŋ/–/o.p^{h}al/–/juk.il.o.sa/	*lhl–lhlhl–lhlh* ex) 215–64049–5357 /i.il.o/–/juk.sa.koŋ.sa.ku/ –/o.sam.o.t͡s^{h}il/
ll	*lll–ll–llll* ex) 955–26–0905 /ku.o.o./–/i.juk/–/koŋ.ku.koŋ.o/	*lll–lllll–llll* ex) 959–95929–9229 /ku.o.ku/–/ku.o.ku.i.ku/–/ku.i.i.ku/

[4]Non-tense fricatives were considered as aspirated, as they patterned together with other aspirated obstruents. See Section 4.3.2 for more information.

[5]A period is used to mark syllable boundaries.

Sixteen natural words employed as the target words are shown in Table 3.2. They were also controlled in terms of their segment-induced pitch patterns (*hh, hl, lh*, and *ll*) and the number of syllables (2–5 syllables). The target words do not contain low vowels, since there is a cross-linguistic tendency for low vowels to show lower f0 values than high or mid vowels (Whalen and Levitt 1995). All target words of *hh* and *hl* started with aspirated stops to directly compare *hh* to *hl*. The target words were presented in a carrier sentence /i.t͡sʰe.nɨn _____ mal.ha.se.jo/ 'Now say _____.' Each target word was read 10 times by each subject in a randomized order, and the total number of tokens of natural words was also 1,280 (4 contexts x 4 AP sizes x 10 repetitions x 8 speakers).

Table 3.2: Natural words employed in the experiment

	2-syllable	3-syllable	4-syllable	5-syllable
hh	/kʰʌ.pʰi/ 'coffee'	/kʰʌm.pʰu.tʰʌ/ 'computer'	/kʰʌm.pʰu.tʰʌ.t͡sʰɛk/ 'computer book'	/pʰi.s'i.kʰʌm.pʰju.tʰʌ/ 'PC computer'
hl	/pʰi.pu/ 'skin'	/kʰɛ.i.kʰɨ/ 'cake'	/pʰɨ.lʌ.pʰo.t͡sɨ/ 'propose'	kʰɨ.lim.kʰɛ.i.kʰɨ/ 'cream cake'
lh	/pu.pʰi/ 'mass'	/in.tʰʌ.nɛt/ 'internet'	/pi.hɛŋ.ki.pʰjo/ 'flight ticket'	/in.tʰʌ.nɛt.t͡sʰin.ku/ 'online friend'
ll	/u.pi/ 'raincoat'	/ki.lʌ.ki/ 'seagull'	/u.li.ko.mo/ 'my aunt'	/u.li.ʌ.mʌ.ni/ 'my mother'

Recordings were made in a sound-attenuated recording booth at the University of Pennsylvania, and were saved as .wav files at the sampling rate of 22.1 kHz, in mono. A headset condenser microphone (Shure WH30) was used, and the recordings were directly digitized into a desktop computer via *Praat* (Boersma and Weenink 2016). The subjects signed a consent form before they were recorded, and received 10 dollars for participating in the experiment. The recording took less than 30 minutes per speaker, and the total number of tokens collected was 2,560 (= 1,280 phone-number strings + 1,280 natural words). The onset and offset of syllables were forced-aligned by a Korean forced aligner (Yoon and Kang 2012), and the alignments were manually checked and corrected afterwards. The mean pitch value of each syllable was obtained with a Praat script. The obtained pitch values were double-checked for f0 tracking errors. The obtained f0 values were converted to semitones (St) with 100Hz as a reference f0, using the following formula: $\log_2(\text{Hz}/100) \times 12$.[6]

3.2 Results

3.2.1 Phone-number strings

Figure 3.1 shows the mean pitch values averaged across all 8 speakers. The most noticeable feature is that the pitch range of an entire AP is determined by the AP-initial digit. *hh* and *hl* generally pattern together in terms of their pitch contours, whereas *ll* and *lh* pattern together. There is a large pitch difference between APs starting with a low-pitch inducing digit (*ll* and *lh*) and those starting with a high-pitch inducing digit (*hh* and *hl*). Not only is this difference noticeable in the AP-initial

[6] I had 100 Hz as a reference f0 for Hz-to-St conversion in this chapter, since there were the same number of subjects for each gender and all of the speakers were in their 20s. However, I normalized pitch values (Hz) using a speaker's own baseline in Chapters 4 and 5. See Section 4.1.2.

position, but it is also retained until the end of an AP. The high-pitch inducing digit in non-AP-initial syllables of *lh* and *hh* is not as decisive as the digits in the AP-initial position. For example, the pitch contour of the *hh* context in Figure (3.1a) is slightly higher than that of the *hl* context, and that of the *lh* context is also slightly higher than that of the *ll* context, in particular, in the last AP of the carrier sentence. However, these differences between *hh* and *hl* (or between *ll* and *lh*) are much smaller than the difference between the *h*-initial and *l*-initial APs (*hh, hl* vs. *ll, lh*). Lastly, the pitch difference between *h*-initial APs (*hh* and *hl*) and *l*-initial APs (*lh* and *ll*) decreases in the phrase-final position to the extent that the mean pitch values overlap. This seems to be mainly due to the phrasal-final intonational H tone (TH-L**H**), and the effect of the phrase-final intonational H tone is observed in the second and third APs of the carrier sentence in all plots in Figure 3.1.

To determine whether the observations are statistically significant, five linear mixed-effects models (for each syllable position within an AP) were built with each pitch pattern within each speaker as a nested random effect. Instead of using all f0 values of the target sentences as the response variable in the analysis, only the third AP, where the number of syllables varied from two to five, was considered in the models. As for syllable positions, all AP-initial and AP-final syllables are coded as Initial and Final, respectively. All AP-penultimate syllables are defined as Penultimate for APs with more than three syllables. The rest of the syllables are defined relative to its position to the beginning of an AP (2-syllable AP: Initial–Final, 3-syllable AP: Initial–Penultimate–Final, 4-syllable AP: Initial–Second–Penultimate–Final, 5-syllable AP:

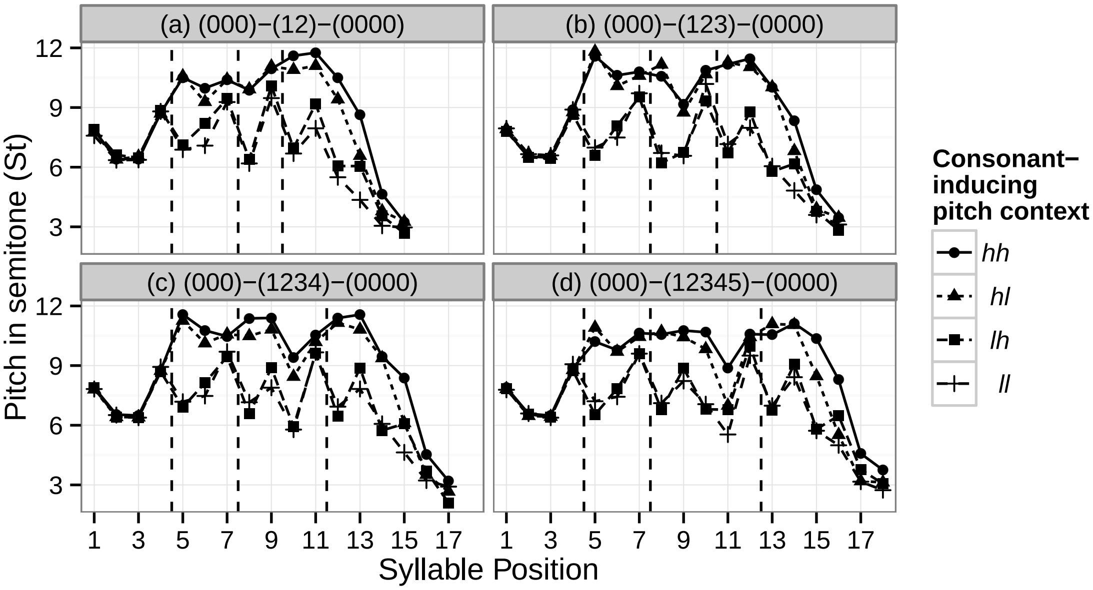


Figure 3.1: Mean pitch values of all speakers by the size of the second phone-number strnigs (where the number of digits varies from (a) 2 syllables to (d) 5 syllables). The x-axis shows syllable positions of the target sentences, and the dashed lines represent AP boundaries. The first AP is /ne.pʌ.ho.nɨn/ 'my number-TOPIC' in the carrier sentence and the other APs are the phone-number strings. The phone number string in the title of each plot (000–12...–0000) is for illustration.

Initial–Second–Third–Penultimate–Final). Table 3.3 summarizes the outputs of the models.

The models estimate that the pitch differences between *h*-initial and *l*-initial APs (***h**h* vs. ***l**l*, ***h**h* vs. ***l**h*, ***h**l* vs. ***l**l*, and ***h**l* vs. ***l**h*) are significant in all syllable positions (p = 0.003 for the comparison of *lh* vs. *hl* and p = 0.001 for the comparison of *ll* vs. *hl* in the AP-final position; $p < 0.001$ for all the other comparisons). This result indicates that the pitch context of AP-initial onset consonants affects the pitch values of the rest syllables of the same AP, confirming the observation that the entire pitch range of an AP is affected by an AP-initial onset consonant in Figure 3.1. The pitch

Table 3.3: The outputs of the linear mixed-effects models of the phone-number strings. Each table represents the results of group comparisons (of the pitch contexts) in a given syllable position. The first column of each table shows the estimated coefficient (i.e., the estimated pitch difference between two groups), and the second column is its standard error (SE). The z values are the estimates divided by the standard errors, and the probabilities in the fourth column are given based on the calculated z values.

	Initial				**Second**			
	Est.	SE	z value	Pr($>\|z\|$)	Est.	SE	z value	Pr($>\|z\|$)
$hl - hh$	-0.03	-.42	-0.08	0.999	-0.46	0.31	-1.46	0.46
$lh - hh$	-4.1	0.42	-9.76	< 0.001***	-2.21	0.31	-7.06	< 0.001***
$ll - hh$	-3.8	0.42	-8.956	< 0.001***	-3.02	0.31	-9.76	< 0.001***
$lh - hl$	-4.07	0.42	-9.58	< 0.001***	-1.75	0.31	-5.6	< 0.001***
$ll - hl$	-3.77	-.42	-8.88	< 0.001***	-2.57	0.31	-8.2	< 0.001***
$ll - lh$	0.3	0.42	0.7	0.896	-0.82	0.31	-2.6	0.045*

	Third				**Penultimate**			
	Est.	SE	z value	Pr($>\|z\|$)	Est.	SE	z value	Pr($>\|z\|$)
$hl - hh$	-0.75	0.55	-1.535	0.529	-1.06	0.39	-2.71	0.035*
$lh - hh$	-3.78	0.55	-6.84	< 0.001***	-2.64	0.39	-6.73	< 0.001***
$ll - hh$	-3.52	0.55	-6.37	< 0.001***	-3.16	0.39	-8.05	< 0.001***
$lh - hl$	-3.03	0.55	-5.49	< 0.001***	-1.58	0.39	-4.03	< 0.001***
$ll - hl$	-2.78	0.55	-5.02	< 0.001***	-2.1	0.39	-5.35	< 0.001***
$ll - lh$	0.26	0.55	0.46	0.967	-0.52	0.39	-1.32	0.55

	Final			
	Est.	SE	z value	Pr($>\|z\|$)
$hl - hh$	-0.15	0.23	-0.64	0.92
$lh - hh$	-0.97	0.23	-4.13	< 0.001***
$ll - hh$	-1.02	0.23	-4.37	< 0.001***
$lh - hl$	-0.82	0.23	-3.49	0.003**
$ll - hl$	-0.88	0.23	-3.73	0.001**
$ll - lh$	-0.06	0.23	-0.25	0.995

differences between *h*-initial and *l*-initial APs in the AP-final syllable (from 0.88 St to 1.02 St) are smaller than those of the other syllable positions, but they are still found to be significant ($p < 0.001$ for both comparisons of *hh* vs. *lh* and *hh* vs. *ll*, $p = 0.003$ for *lh* vs. *hl*, and $p = 0.001$ for *ll* vs. *hl*).

The models also estimate that the pitch differences between *hh* and *hl* are not significant in any syllable positions ($p < 0.999$ for the AP-initial position, $p = 0.46$ for the AP-second position, $p = 0.529$ for the AP-third position, $p = 0.92$ for the AP-final position), but in the penultimate syllable ($p = 0.035$). This result suggests that *hh* and *hl* generally pattern together, and the onset consonant of the AP-second syllable plays a less important role than that of the AP-initial syllable. However, this does not mean that high-pitch inducing consonants in non-AP-initial positions have no effect on the intonational melody. In the result of the penultimate syllable, the model estimates that the pitch value of *hh* is 1.06 St higher than that of *hl* and this difference is significant ($p =0.035$), suggesting that the high-pitch inducing segment in the penultimate syllable has an effect on the intonational L tone in TH-**L**H, making the pitch value of the penultimate syllable of *hl* higher than that of *hh* (see Figure (3.1d)). However, the model reveals that the pitch value of *hl* in the penultimate syllable is still 2.1 St and 1.58 St higher than those of *ll* and *lh* (APs starting with a low-pitch inducing segment), respectively, and these pitch differences are significant ($p < 0.001$ for both comparisons). This result suggests that even the pitch value of the penultimate syllable (where the intonational L tone in TH-**L**H falls) is higher when the AP-initial consonant is a high-pitch inducing segment than when it is a low-pitch inducing segment.

Similarly, the models estimate that the pitch difference between *ll* and *lh* is not significantly different in most syllable positions ($p = 0.896$ for the AP-initial syllable, $p = 0.967$ for the AP-third syllable, $p = 0.55$ for the AP-penultimate syllable, $p = 0.995$ for the AP-final syllable), except in the AP-second syllable ($p = 0.045$). This result also suggests that *ll* and *lh* generally pattern together in terms of their pitch values. Again, the significant difference between *ll* and *lh* in the AP-second syllable seems to be due to the consonant-induced pitch. The high-pitch inducing segment in the second syllable of *lh* seems to make the pitch value of its second syllable 0.82 St higher than that of *ll*, and this difference is significant in the model.

Figure 3.2 compares the mean pitch values of the digits in the target phone number strings. As expected, the mean pitch values of the digits with a high-pitch inducing onset (black) are higher than those with a low-pitch inducing onset (gray). One interesting result is that the mean pitch value of 1 /il/ is much higher than that of 2 /i/, even though both of them start with a vowel. This result agrees with Jun and Cha's study (2015) that 1 /il/ is produced with a higher pitch than 2 /i/. Similarly, the comparison between 3 /sam/ and 4 /sa/ is also interesting.[7] Recall that both digits are classified as high-pitch inducing digits in this chapter.[8] Both digits have the same onset /s/ and vowel /a/ and they only differ in the final coda consonant but their pitch values are considerably different. The mean pitch of 4 /sa/ is substantially lower

[7] I would like to thank Prof. Jun for pointing out the pitch difference between 3 and 4.
[8] See Section 4.3.2 for the result that /s/ and /h/ pattern together with other aspirated obstruents.

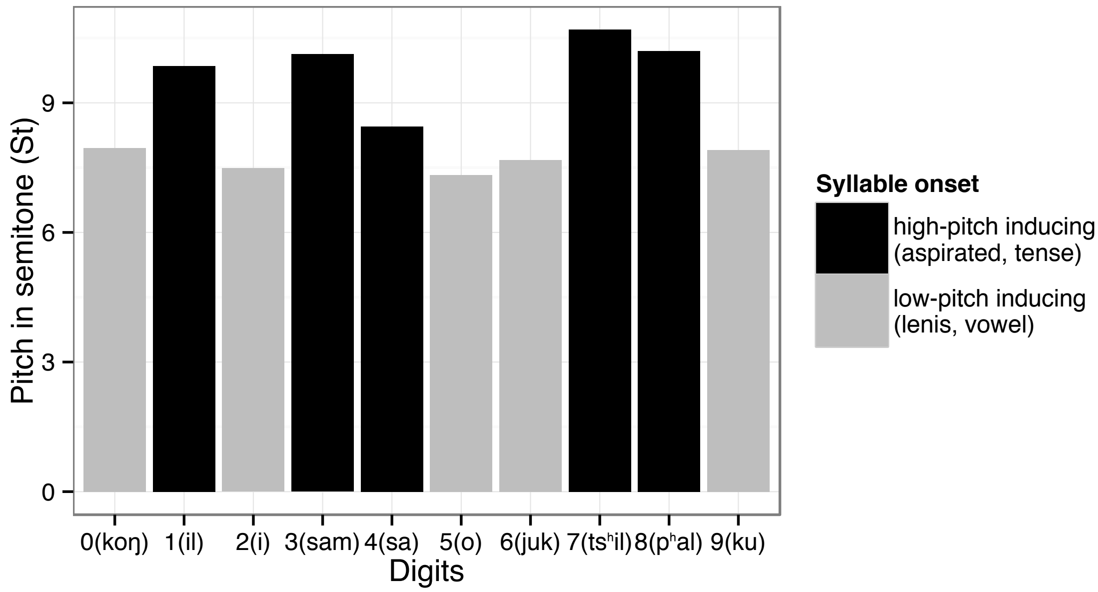


Figure 3.2: Mean pitch values of the digits in the phone-number strings. Black bars are the digits whose onset consonant is a high-pitch inducing type, and gray bars are the digits starting with a low-pitch inducing type. The y-axis is the mean pitch in semitone.

than the other high-digits including 3 /sam/, yet it is still higher than the low-pitch inducing digits (gray bars).

A linear mixed-effects model was built to examine the pitch differences of the digits with the pitch values as the dependent variable and each digit as a fixed effect. The reference digit of comparison was 0 /koŋ/, which starts with a low-pitch inducing consonant (lenis). Each digit within each speaker was treated as a nested random effect. The output of the model is summarized in Table 3.4. As expected, the model estimates that most of the high-pitch inducing digits (four out of five) are significantly higher than 0 /koŋ/, a low-pitch inducing digit (reference). The model estmates that 3 /sam/, 7 /$\widehat{\text{ts}}^{\text{h}}$il/, 8 /$\text{p}^{\text{h}}$al/, and 1 /il/ are 2.18, 2.68, 2.26, and 1.9 St higher

Table 3.4: The output of a linear mixed-effects model of the digits. The reference (intercept) is 0 /koŋ/, whose onset is a low-pitch inducing consonant (lenis).

	Estimate	Standard Error	t-value	Pr(>| t |)
(Intercept)	7.95	2.27	3.5	0.009**
1 /il/	1.9	0.26	7.29	< 0.001***
2 /i/	-0.46	0.26	-1.78	0.081
3 /sam/	2.18	0.26	8.34	< 0.001***
4 /sa/	0.5	0.26	1.91	0.061
5 /o/	-0.62	0.26	2.36	0.021*
6 /juk/	-0.27	0.26	-1.03	0.306
7 /t͡sʰil/	2.68	0.26	10.24	< 0.001***
8 /pʰal/	2.26	0.26	8.64	< 0.001***
9 /ku/	-0.04	0.26	-0.16	0.876

than 0, respectively ($p < 0.001$ for the four comparisons). The largest estimated coefficient is found in 7 /t͡sʰil/ (2.68 St), which starts with an aspirated affricate. Also, 8 /pʰal/, which starts with an aspirated stop, shows a large coefficient (2.26 St), meaning that its mean pitch value is 2.26 St higher than that of 0 /koŋ/. The result that both aspirated affricate and stop cause high pitch values to the following vowel confirms Kang's finding (2014) that the sound change among voiceless stops is in fact a structural one, affecting all aspirated and tense categories in the Korean voiceless consonants.

The observation that 1 /il/ is higher than 2 /i/ in Figure 3.2 is also verified in the result of the linear mixed-effects model (Table 3.4). The model shows that 1 /il/ is 1.9 St higher than 0 /koŋ/ ($p < 0.001$) and this pitch difference is significant, whereas 2 /i/ is, rather, -0.46 St lower than 0 /koŋ/ ($p = 0.081$). Although the pitch difference between 2 /i/ and 0 /koŋ/ misses significance, it is worth to mention that the estimated pitch difference between 1 /il/ and 2 /i/ in the model is 2.36 St (= 1.9 + 0.46), which cannot be attributed to a coincidence. A one sample t-test conducted to examine this pitch difference also suggests that the pitch difference between 1 /il/ and 2 /i/ is significant ($t(2643.29)$=9.59, $p < 0.001$). Also, the pitch difference between 3 /sam/ and 4 /sa/ found in Figure 3.2 is borne out in the output of the linear mixed-effects model (Table 3.4). As expected, 3 /sam/ is significantly higher than 0 /koŋ/ ($p < 0.001$), but 4 /sa/ is not ($p = 0.061$). The model estimates that 3 /sam/ is 2.18 St higher than 0 /koŋ/, but 4 /sa/ is only 0.5 St higher than 0 /koŋ/. The pitch difference between 3 /sam/ and 4 /sa/ estimated from their coefficient is 1.68 St (= 2.18 – 0.5), which cannot be a coincidence. A one sample t-test finds that the pitch difference between 3 /sam/ and 4 /sa/ is significant ($t(2406.73)$=6.37, $p <$ 0.001). It seems to suggest that a pitch contrast is emerging between 3 /sam/ and 4 /sa/, as the case of 1 /il/ and 2 /i/.

The model in Table 3.4 also finds that 5 /o/ is 0.62 St lower than 0 /koŋ/, while both are classified as low-pitch inducing digits. This seems to be the cross-linguistic tendency for voiceless consonants to raise the pitch value of the following vowel (Ohala 1978, Hombert 1978, Hombert et al. 1979, Jun 1996, Hanson 2009, and among others). Because 0 /koŋ/ starts with a voiceless stop (lenis), it is produced with a higher pitch than 5 /o/, even though they have the same vowel. The other low-pitch inducing digits (2 /i/, 6 /juk/, and 9 /ku/) are neither significantly higher nor lower than 0 /koŋ/.

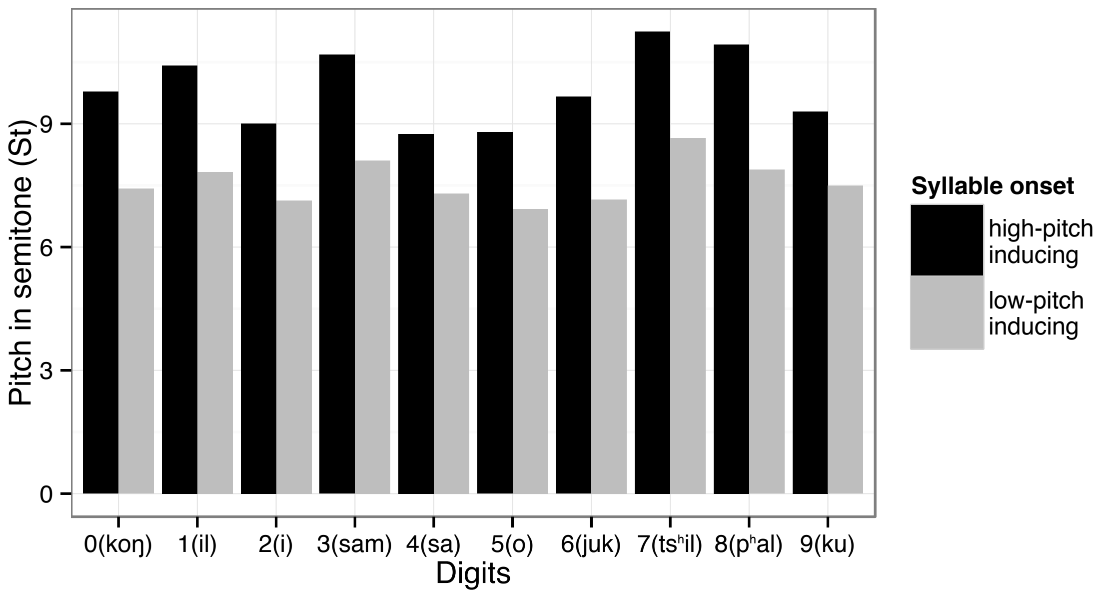


Figure 3.3: Mean pitch values of the digits by AP-initial onset consonants. The y-axis is f0 values in semitones, and the x-axis shows each digit from 0 to 9. Black bars represent the mean pitch value of each digit when an AP-initial consonant is a high-pitch inducing type (aspirated or 1 /il/), and gray bars show the mean pitch value of the same digit when an AP-initial onset is a low-pitch inducing type (lenis or vowel).

Figure 3.3 compares the mean pitch values of the digits in the high-initial APs to those in the low-initial APs. As expected, the pitch values are much higher when an AP starts with a high-pitch inducing digit than when it starts with a low-pitch inducing digit. A one-way repeated measures ANOVA was conducted to see if the pitch differences of the same digits in the high-pitch (*hh* and *hl*) and low-pitch contexts (*lh* and *ll*) were statistically significant. In the ANOVA test, the dependent variable was the pitch values of the digits, and the AP-initial consonant pitch (*h* vs. *l*) was used as a fixed within-subject factor. The test found that the effect of the AP-initial pitch context was significant ($F[1, 7] = 60.71$, $p < 0.001$), indicating that the pitch values of the digits in high-pitched APs (those starting with a high-pitch inducing consonant) are significantly higher than those in low-pitched ones (those starting with a low-pitch inducing consonant).

3.2.2 Natural words

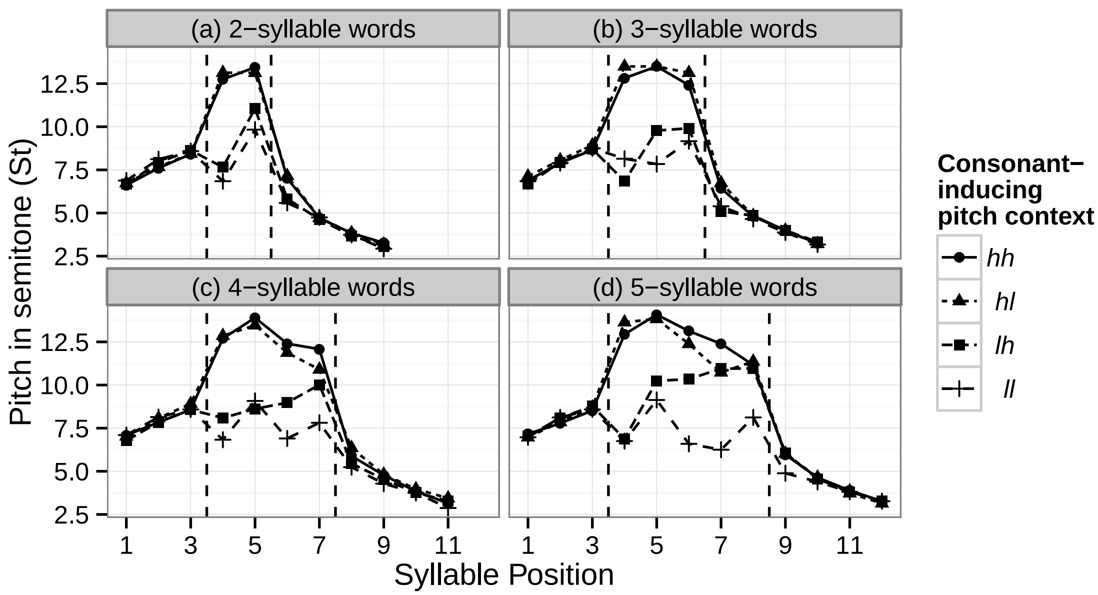


Figure 3.4: Mean pitch values of all speakers by the length of the target words. The x-axis shows syllable positions in the target sentences, and the y-axis shows the mean pitch values. The first three syllables are for /i.t͡se.nɨn/ 'now-TOPIC', the last four syllables are for /mal.ha.se.jo/ 'say' in the carrier sentence /i.t͡se.nɨn _____ mal.ha.se.jo/ 'Now say _____.', and the syllables in the middle (inside the dashed lines) are the target words.

Figure 3.4 shows the results of the natural words. The most important result is that the pitch range of an AP is mainly affected by the AP-initial onset consonant, as shown in the results of the phone-number strings. When an AP-initial onset is a high-pitch inducing segment (aspirated), the pitch range of the AP is much higher than when it is a low-pitch inducing segment (lenis, sonorant, or vowel). There is a large pitch difference between the target words starting with aspirated consonants (*hh* and *hl*) and those starting with sonorants or vowels (*lh* or *ll*). In addition, similar to the results of the phone-number strings, the effect of high-pitch inducing onsets in non-AP-initial syllables is much smaller than that of the AP-initial syllable. That is, *hh* and *hl* pattern together regardless of the size of the target words, and *ll* and *lh* are not much different from each other when the size of the target AP is 2 syllables, 3 syllables or 4 syllables. Yet, the target word of *lh* is noticeably higher than that of *ll* when the target AP is a 5-syllable word in Figure (3.4d). This seems to be because of the 5-syllable target word of *lh* /in.t^hʌ.net.$\widehat{ts}^h$in.ku/ 'online friend.' The speakers tended to produce /in.t^hʌ.net/ 'internet' and /$\widehat{ts}^h$in.ku/ 'friend' as two separate APs, raising /$\widehat{ts}^h$in.ku/ 'friend' to the same pitch range of *hh* and *hl* due to the high-pitch inducing onset consonant /$\widehat{ts}^h$/ (aspirated affricate). An example of the pitch track of this target word produced with two separate APs is shown in Figure 3.5.

For statistical analyses, five linear mixed-effects models (for each syllable position within an AP) were built with each pitch pattern within each speaker as a nested random effect. In these models, the f0 values of the target words were included as the dependent variable. Similar to the analysis of the phone-number strings, syllable positions within the target words were coded as the following: 2-syllable AP: Initial–Final, 3-syllable AP: Initial–Penultimate–Final, 4-syllable AP: Initial–Second–Penultimate–Final, 5-syllable AP: Initial–Second–Third–Penultimate–Final. Table 3.5 summarizes the outputs of the models.

Table 3.5: The outputs of the linear mixed-effects models of the natural words: each table represents the results of group comparisons (of the pitch contexts) in a given syllable position. The first column of each table shows the estimated coefficient (i.e, the estimated pitch difference between two groups), and the second column is its standard error (SE). The z values and the probabilities are given in the third and fourth columns.

	Initial				**Second**			
	Est.	SE	z value	Pr(> $\|z\|$)	Est.	SE	z value	Pr(> $\|z\|$)
$hl - hh$	0.3	0.23	1.33	0.543	-0.06	0.36	-0.16	0.999
$lh - hh$	-5.35	0.22	-23.82	< 0.001***	-4.31	0.36	-12.09	< 0.001***
$ll - hh$	-5.59	0.22	-24.89	< 0.001***	-4.6	0.36	-12.89	< 0.001***
$lh - hl$	-5.65	0.23	-25.1	< 0.001***	-4.25	0.35	-11.95	< 0.001***
$ll - hl$	-5.89	0.23	-26.18	< 0.001***	-4.54	0.36	-12.76	< 0.001***
$ll - lh$	-0.24	0.22	-1.06	0.714	-0.29	0.36	-0.8	0.853

	Third				**Penultimate**			
	Est.	SE	z value	Pr(> $\|z\|$)	Est.	SE	z value	Pr(> $\|z\|$)
$hl - hh$	-0.62	0.44	-1.42	0.489	-0.42	0.34	-1.22	0.613
$lh - hh$	-2.66	0.44	-6.09	< 0.001***	-2.44	0.34	-7.13	< 0.001***
$ll - hh$	-6.14	0.44	-14.68	< 0.001***	-5.32	0.34	-15.54	< 0.001***
$lh - hl$	-2.04	0.44	-4.67	< 0.001***	-2.02	0.34	-5.94	< 0.001***
$ll - hl$	-5.79	0.44	-13.27	< 0.001***	-4.9	0.34	-14.42	< 0.001***
$ll - lh$	-3.75	0.44	-8.6	< 0.001***	-2.88	0.34	-8.48	< 0.001***

	Final			
	Est.	SE	z value	Pr(> $\|z\|$)
$hl - hh$	-0.16	0.25	-0.66	0.912
$lh - hh$	-1.81	0.25	-7.33	< 0.001***
$ll - hh$	-3.54	0.25	-14.33	< 0.001***
$lh - hl$	-1.65	0.24	-6.67	< 0.001***
$ll - hl$	-3.38	0.25	-13.66	< 0.001***
$ll - lh$	-1.73	0.25	-6.99	< 0.001***

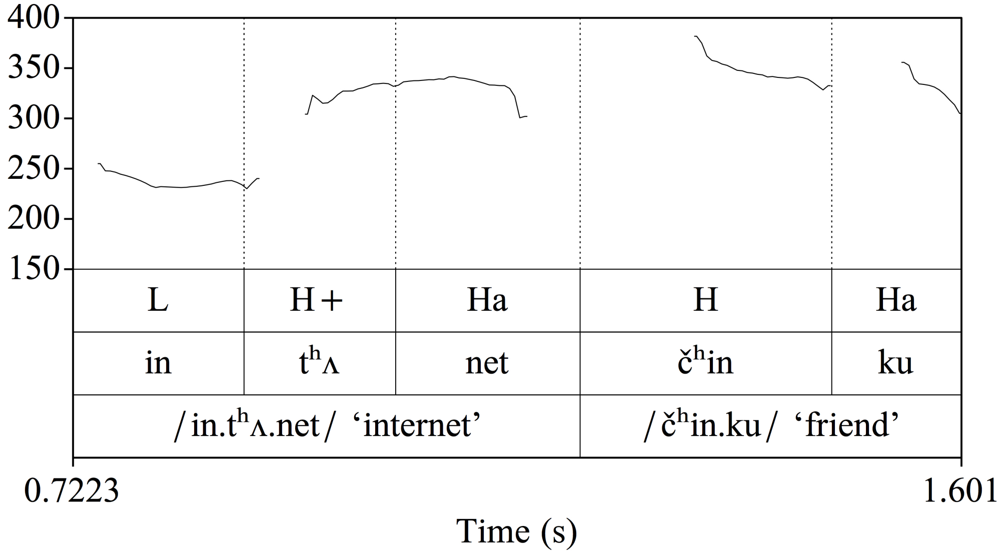


Figure 3.5: An example pitch track of the *lh* target word /in.tʰʌ.net.t͡sʰin.ku/ 'online friend' produced by a female speaker (JM). The pitch values are expressed in Herz (Hz). The first tier of the text grid shows the Korean intonational tones in the K-ToBI system (Jun 2006), and the second tier is for the syllable level. The third tier shows the AP boundary.

The models estimate that the pitch values of the target words starting with a high-pitch inducing consonant (*hh* and *hl*) are significantly different from those of starting with a low-pitch inducing consonant (*lh* and *ll*) in all syllable positions ($p < 0.001$ for all comparisons). In particular, the estimated pitch difference between *h*-initial and *l*-initial APs is the largest in the AP-initial syllable (from 5.35 St to 5. 89 St) and the second largest in the AP-second syllable (from 4.25 St to 4.6 St). The smallest pitch difference is found in the AP-final syllable position (from 1.65 St to 3.54 St), but the difference is still significant ($p < 0.001$ for the four pairwise comparisons).

However, it is notable that the estimated pitch differences between *h*-initial and *l*-initial words in the AP-third and penultimate syllables vary considerably. The pitch differences between the *h*-initial target words (*hh* and *hl*) and that of *ll* are large (6.41 St and 5.79 St for Third, 5.32 St and 4.9 St for Penultimate), whereas the pitch differences between the *h*-initial target words (*hh* and *hl*) and that of *lh* are relatively small (2.66 St and 2.04 St for Third, 2.44 St and 2.02 St for Penultimate). This result seems to be due to the 5-syllable target word of *lh*, /in.tʰʌ.net.t͡sʰin.ku/ 'online friend.' Because the participants tended to produce the target word with two separate APs, a 3-syllable AP (/in.tʰʌ.net/ 'online (LHH)') and a 2-syllable one (/t͡sʰin.ku/ 'friend (HH)'), the third syllable of this word in fact bears a final intonational H tone. Also, the penultimate and final syllables of this word are in a similar pitch range to those words starting with a high-pitch inducing consonant. This fact seems to be reflected in the models, estimating the pitch difference between the *h*-initial words and the target word of *lh* smaller than the one between the *h*-initial words and the *ll* target word in the third and penultimate syllables.

The models also show that the pitch values of *hh* and *hl* are not significantly different from each other in all syllable positions ($p = 0.543$ for Initial, $p = 0.999$ for Second, $p = 0.489$ for Third, $p = 0.613$ for Penultimate, $p = 0.912$ for Final). This result confirms that *hh* and *hl* pattern together, regardless of AP sizes. Similarly, the pitch values of *ll* and *lh* are not significantly different in the AP-initial and second syllables ($p = 0.714$ for Initial and $p = 0.853$ for Second). However, the models indicate that the pitch differences between *ll* and *lh* in the other syllable positions are significant ($p < 0.0001$ for all three syllable positions). Again, this result seems to be due to the 5-syllable target word of *lh*, /in.tʰʌ.net.t͡sʰin.ku/ 'online friend.' Because the participants tended to produce the target word with two separate APs, the last three syllables of that target word were produced with higher pitch values than those of *ll* and the models find the difference between *ll* and *lh* significant.

3.3 Discussion

The results of the phone-number strings and the natural words lead to the same conclusion: when an AP starts with a high-pitch inducing type (aspirated), all syllables in the AP are produced within a higher pitch range than when an AP starts with a low-pitch inducing segment (lenis or vowel). The results show that even the penultimate syllable, where the intonational L tone in TH-**L**H falls, is produced with a higher pitch in *h*-initial APs (*hh* and *hl*) than in *l*-initial ones (*lh* and *ll*). Also, it is found that there is a structural pitch difference between *h*-initial and *l*-initial APs. In the results of the phone-number strings and the natural words, *hh* and *hl* are not significantly different from each other in most of the AP positions, but *lh* and *ll* are considerably different from *hh* and *hl* in all AP positions. This result suggests that the pitch distinction in the AP-initial position is no longer due to automatic phonetic effects from consonantal perturbation, as it affects not only the initial but also non-initial syllables of APs.

Also, the pitch difference between *ll* and *lh* in the 5-syllable natural words provides a clue that an AP-initial onset consonant is a primary factor affecting the entire pitch range of an AP. The *lh* target word tends to be produced with two separate APs (Figure 3.5), and the pitch range of the second AP is similar to that of the *h*-initial APs because of the high-pitch inducing onset segment /$\widehat{\text{ts}}^{\text{h}}$/ of /$\widehat{\text{ts}}^{\text{h}}$in.ku/ ‘friend’. The high pitch values of the 5-syllable *lh* word in the penultimate and final syllables contribute to the significant pitch differences between *ll* and *lh* as demonstrated in Table 3.5. Again, this result suggests that the pitch distinction in the AP-initial position is the most important factor in deciding the pitch range of an AP.

In addition, the results of this chapter reveal that high-pitch inducing consonants in non-AP-initial positions seem to increase the pitch values of non-initial syllables, although their effect is much smaller than those in AP-initial positions. For example, in the phone-number strings, the pitch value of *hh* in the penultimate syllable is found to be slightly higher than that of *hl* and this seems to be because the penultimate syllable of *hh* always has a high-pitch inducing onset. (This effect is most clearly seen in Figure (3.1d).) Also, the results of the phone-number strings show that the pitch value of *lh* in the AP-second syllable is higher than that of *ll* due to the segment-induced *h* pitch of *l**h***.

One important question raised by these results is why we see such a large pitch difference between *h*-initial and *l*-initial APs (in all syllable positions). Considering the results of this chapter and other previous studies on the Korean stops, this seems to be due to the tonogenetic sound change in the AP-initial syllable position (Silva 2006, Wright 2007, Oh 2011, Kang 2014, among others). The results of this chapter indicate that the pitch difference in the AP-initial position extends to later syllables of the same AP, like a H-tone spreading phenomenon. Since the pitch of the AP-initial syllable with an aspirated onset consonant is much higher than the one with a lenis onset, the pitch of *h*-initial APs stays relatively high through the end of the APs.

Questions that are not answered in this chapter are when this change started to affect the intonational melody, how the intonational melody has changed over time, and who has led this change. Also, it remains to be seen if tense-initial APs show the same pitch pattern with aspirated-initial ones, as only aspirated-initial APs are examined in this chapter. These questions are further investigated in the next chapter.

Chapter 4

Corpus Analysis

This chapter is devoted to investigating how the AP-initial pitch contrast has developed over time and how different linguistic factors as well as speakers' age and gender affect the change of the intonational melody of Seoul Korean.

Chapter 3 illustrated that the AP-initial onset consonant has an effect on non-initial syllables of the same AP. However, much less is known about this effect. While previous studies (Lee 1999, Silva 2006, Kang 2014) show that a H-pitch inducing consonant in the AP-initial position affects pitch values of AP-second syllables, the exact pattern of the diachronic development of the AP melody and its interaction with age and gender remain to be seen. For example, it is not clear if we see the same effect of H-inducing consonants on non-initial syllables for older speakers, as only younger speakers are studied in Chapter 3. If older speakers do not show similar patterns, an in-depth investigation of the change is needed to fully understand this tonogenetic change. Also, it is not known if the pitch differences of non-initial syllables are affected with the same rate of the AP-initial pitch development. Did the AP-initial pitch contrast develop first and start to affect non-initial syllables later? Or did the pitch range of entire H-initial APs become high concomitantly? When did the AP-initial H-inducing consonants start to affect pitch values of later syllables? These questions remain to be answered. Lastly, it is not clear how the AP tonal pattern and the AP-initial pitch contrast are realized in different AP sizes. APs that are 2 to 5 syllables long are the most common in Korean, and I hypothesize that pitch patterns will differ depending on whether the AP melody TH-LH is fully realized, as in 4-syllable or 5-syllable APs; or whether the AP medial tones (T**H-L**H) are undershot, as in 2- or 3-syllable APs.

To observe the diachronic development of AP pitch patterns, a corpus study with a large number of speakers with a wide age range is necessary. Thus, in this chapter, I employ a large-scale speech corpus of Seoul Korean with 118 speakers to conduct an apparent-time study that investigates the diachronic trajectory of the change. With the above questions in mind, this chapter aims to provide a dynamic picture of the f0 dimension of the change by showing how the intonational melody of APs has changed over time due to the effect of the AP-initial pitch contrast; and how age, gender, and other linguistic factors affect the AP melody of Seoul Korean.

To address these questions, I draw speech data from *The Speech Corpus of Reading-Style Standard Korean* published by the National Institute of the Korean Language (The National Institue of the Korean Language 2005; hereafter the NIKL corpus).

4.1 The Speech Corpus of Reading-style Standard Korean

4.1.1 Speakers and materials

The NIKL corpus is based on the read speech of 60 male and 60 female speakers of Seoul Korean, residing in Seoul or the Seoul metropolitan area. The publisher of the corpus reports that the participants' parents are also speakers of Seoul Korean and have lived in the Seoul metropolitan area. The speakers' ages range from 19 to 71 at the time of recording (2003). Their age is converted to their year of birth (ranging from 1932 to 1984) by subtracting their age from 2003, following Kang's 2014 study, which employs the same corpus in investigating the trade-off between VOT and pitch cues in Seoul Korean. The analysis below is based on 118 speakers out of 120, because the sound files of two speakers have technical errors.[1] The breakdown of the speakers analyzed in this chapter by speakers' year of birth and gender is given in Table 4.1.[2]

The speech materials of the corpus are based on 19 well-known short stories, such as traditional folk tales and personal essays, and the total number of sentences in all

[1]The speakers whose sound files contain errors are fy15, a female speaker who was born in 1948, and mw12, a male speaker who was born in 1970. Their sound files contain no information (0 bytes), so I could not use those files for analysis. In the corpus, there were two speakers (mv15 and mv18) whose sound files could not be opened, because the files were headerless. I could recover them by converting them into .raw files, adding appropriate headers, and reconverting them into .wav files, so I included them in the analysis.

[2]Individual speakers vary in terms of their pitch ranges. Some speakers use a wide pitch range, whereas others have a monotonous voice. For this reason, I calculated each speaker's pitch range in Appendix A.

Table 4.1: Age and gender of the speakers in the NIKL corpus analyzed for this dissertation. The numbers in parentheses are the numbers of the speakers in the original corpus data.

	1930s	1940s	1950s	1960s	1970s	1980s	Total
Female	2	9 (10)	25	3	11	9	59 (60)
Male	4	12	4	7 (8)	27	5	59 (60)

19 short stories is 930. Younger speakers (under 50 at the time of recording) read all 19 stories, but older speakers (50 or over at the time of recording) read only 11 stories out of 19.[3] The number of sentences in those 11 stories is 404, and I draw the speech data only from those sentences to include all available speakers of the corpus in the data. Out of 404 sentences, 59 sentences are selected for analysis.[4]

4.1.2 Methods

The sampling rate of the sound files in the corpus was 16,000 Hz. For the 59 selected sentences, the onset and offset of syllables as well as the onset and offset of APs were forced-aligned with the Korean forced aligner (Yoon and Kang 2012). The alignments of syllables and the boundaries of the target APs were manually checked and corrected afterwards. As for the f0 measurement, I measured mean pitch values of all syllables within the target APs using *Praat* (Boersma and Weenink 2016) with a f0 range set at 75–300 Hz for male speakers and 100–500 Hz for female speakers. Since the AP-initial pitch contrast stemmed from the tonogenetic sound change of onset consonants,

[3] The 11 stories' IDs are from t09 to t19.
[4] The sentences used in this study are provided in Appendix B.

the f0 measurements were taken from the entire syllable duration, including onset consonants, unlike previous studies methods in which the f0 measurements were taken at the midpoint of vowels (Kang 2014), or onset, midpoint, and offset of vowels (Silva 2006).

The obtained pitch values were converted into a semitone scale (St), using speakers' baseline pitch values. Liberman and Pierrehumbert (1984) suggest that a natural scale for intonation appears to show exponential decay, decreasing proportionally to a speaker's baseline, where the baseline stands for the bottom of the speaker's pitch range. Following this idea, I first computed each speaker's baseline, which was defined as the 10th percentile of his or her own pitch range in this dissertation. Afterwards, St was calculated with each speaker's baseline as a reference f0, using the following formula: $St = \log_2(f0/baseline)*12$. One advantage of using this relative semitone scale over the one calculated with a fixed reference f0 is that it allows me to compare the pitch patterns of different age and gender groups directly, since the scale was calculated relative to each speaker's baseline.

4.1.3 Examples of AP phrasing

Before moving on to the results, examples of AP and syllable boundaries are provided in this section. Figure 4.1 shows a phrase /p'omnɛtɨsi t͡salaŋhanɨn t͡salaɯɭi jʌpɛsʌ/ 'Right next to the turtle who is showing off' produced by a male speaker born in 1979. The first AP clearly shows the AP tonal pattern HH-LH: The first syllable of the first AP /p'om/ is higher than the the first syllable (L) of the next AP /**t͡sa**laŋhanɨn/, but it is lower than the second syllable of the same AP /nɛ/ (+H). The pitch value of the third syllable in the first AP is slightly low (L+), yet the value increases again in the final syllable (Ha). Also, it is noticeable that the first AP, which starts with a tense stop /p'/, is produced in a relatively higher pitch range than the next AP, which starts with a lenis affricate /t͡s/.

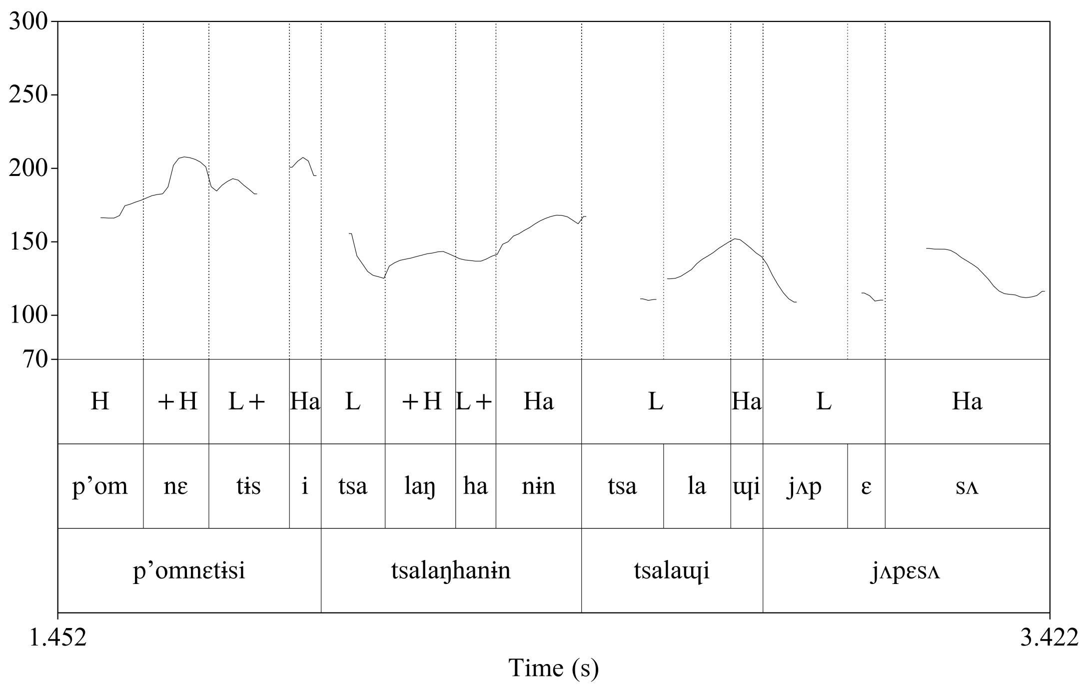
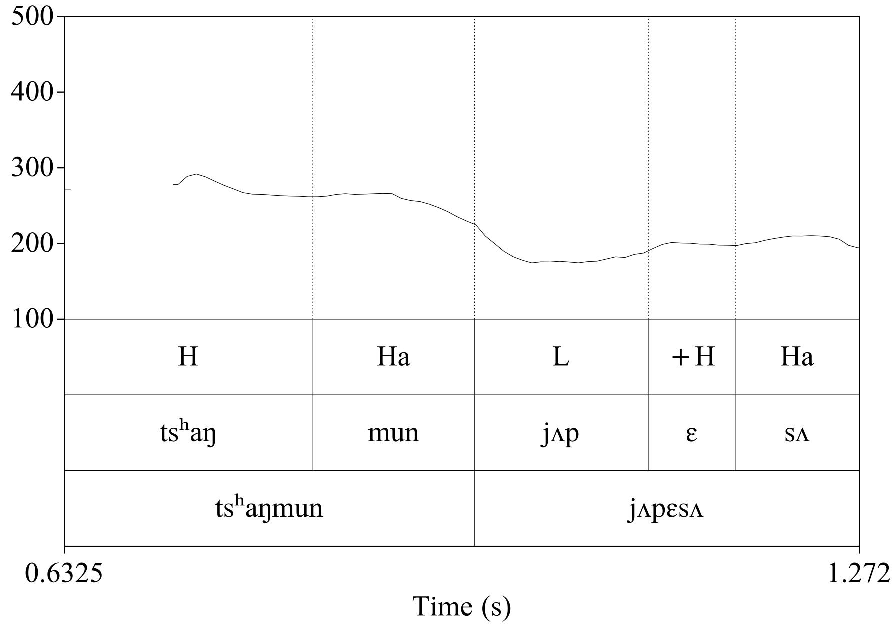


Figure 4.1: An example of a pitch track of syllable and AP boundaries produced by a male speaker born in 1979 (mv01). The phrase shown in the figure is /p'omnɛtɨsi t͡salaŋhanɨn t͡salaɰi jʌpɛsʌ/ 'Right next to the turtle who is showing off' (sentence ID: t09_s43). ToBI tones in the first tier are provided for reference. The second and third tiers show syllable and AP boundaries, respectively. The y-axis shows pitch values in hertz (Hz).

In general, AP phrasings are fairly consistent across speakers, but there are cases where a few speakers show a different AP phrasing from others. Figure 4.2 shows an example of such an instance of interspeaker variation in AP phrasing. The top figure



is a pitch track of a phrase /t͡sʰaŋ.mun.jʌp.ɛ.sʌ/ 'beside the window' produced by a female speaker born in 1976 (fv15). The phrase is grouped into two separate APs: /t͡sʰaŋ.mun/ 'window' and /jʌp.ɛ.sʌ/ 'beside'. It is clear that the first AP starts with an initial H tone due to an aspirated affricate /t͡sʰ/. Right after the second syllable of the first AP /mun/, the pitch value abruptly decreases to target the L-initial tone of the next AP /jʌp.ɛ.sʌ/ (**L**H-H). After the offset of the first syllable in the second AP, the pitch value increases so that the entire tonal pattern of the phrase becomes LH-(L)H, undershooting the penultimate L tone.

However, the same phrase exhibits a different AP phrasing in the bottom figure, where a pitch track of the same phrase produced by another female speaker (fx20, born in 1954) is shown. In this figure, the phrase is grouped into only one AP:

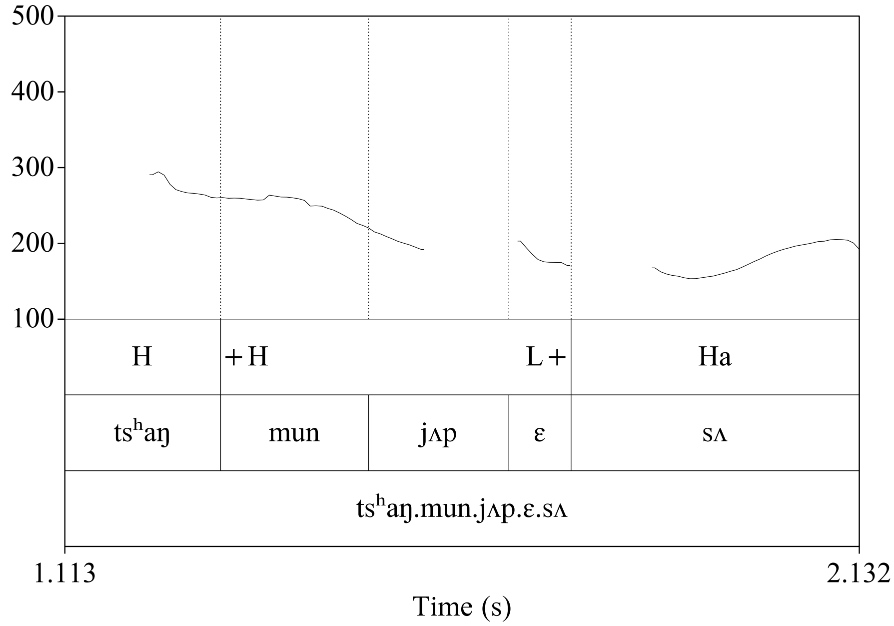


Figure 4.2: Examples of pitch tracks of the same phrase /t͡sʰaŋmunjʌpɛsʌ/ 'beside the window'. The top figure shows the phrase produced with two separate APs (Sentence ID: fv15_t11_s29). The bottom figure shows the same phrase produced with one AP (sentence ID: fx02_t11_s29). The first, second, and third tiers show ToBI tones, syllables, and AP boundaries, respectively. The y-axis shows pitch values in Hz.

/t͡sʰaŋ.mun.jʌp.ɛ.sʌ/ 'beside the window.'[5] After the offset of the second syllable /mun/, the pitch slope gradually decreases until the onset of the final syllable /sʌ/. This pattern suggests that the pitch values of AP-medial syllables are obtained via a linear interpolation from the H tone on the second syllable (T**H**-LH) to the L tone (TH-**L**H). The penultimate L tone shows a carry-over effect, being aligned with the onset of the final syllable.

When manually checking the alignments of syllable and AP boundaries, this type of interspeaker variation is considered. Sentences that have a large amount of inter-

[5]This pattern (the bottom one in Figure 4.2) is the AP tonal pattern most speakers show for the phrase.

speaker variation are not included in the data, and only those with consistent AP phrasings across speakers are selected. When only a few speakers show a different phrasing from other speakers as in Figure 4.2 (top), while the others show a consistent AP phrasing (the bottom one in Figure 4.2), only the pattern that most speakers produce are included in the analysis.

Lastly, there are cases where speakers make speech errors, resulting in an unusual AP phrasing. Figure 4.3 shows a pitch track of a phrase /pa.tat.ka.ɛ to.$\widehat{ts}^{h}$ak.han/ 'arriving at the seashore' produced by a male speaker born in 1970 (speaker: mw11, sentence ID: t09_s56). The speaker made a small pause (about 170ms) within the first phonological word /pa.tat.ka/ 'seashore', producing the first phrase 'at the seashore' with two separate APs: /pa.tat/ and /ka.ɛ/. The break within the first phonological word is unexpected and different from the usual grouping of the phrase, where the phonological word /pa.tat.ka/ and the following locative marker /ɛ/ form one single AP as in /pa.tat.ka.ɛ/. Tokens with a speech error as in this example are excluded from analyses in the following section.

4.2 Sentence-initial APs

The basic melody of an AP, TH-LH, is greatly affected if an ip or IP boundary is overridden with the AP-final tone. For example, the final boundary L tone of an IP, which is often found in a declarative sentence, may override the final boundary H tone of a sentence-final AP, obscuring the pitch pattern of the final AP. Since it is

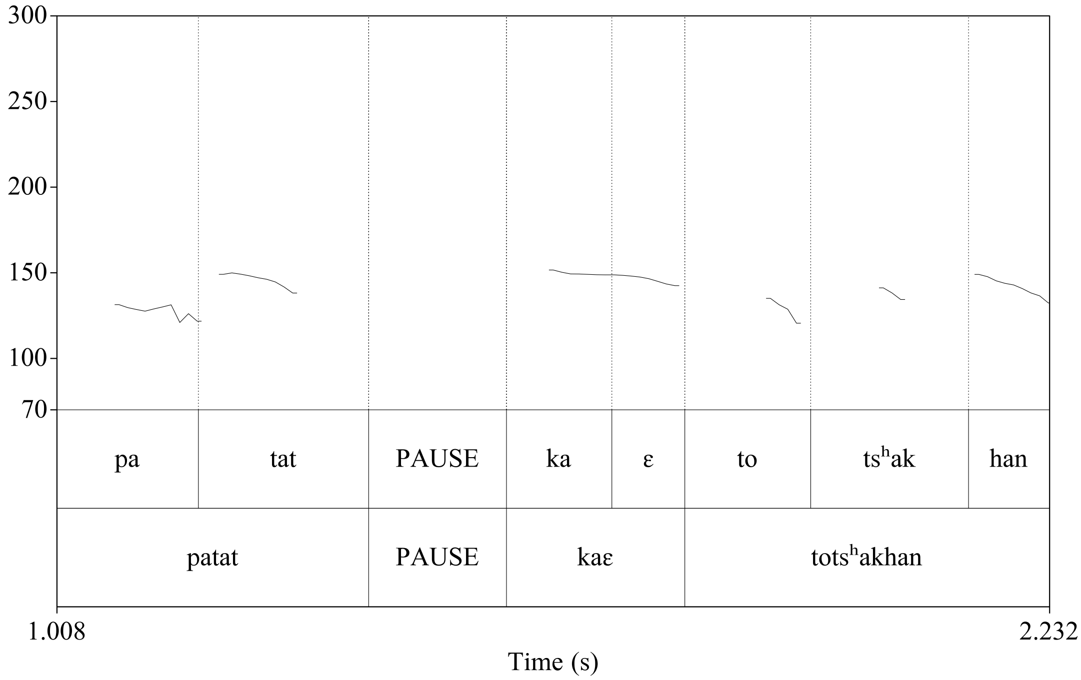


Figure 4.3: An example of a pitch track with an unusual AP phrasing /patat kaɛ/ 'seashore-LOC' produced by a male speaker born in 1970 (mw11). The y-axis shows pitch values in hertz (Hz).

possible for sentence-medial or -final APs to be affected, I first start with sentence-initial APs of the 59 target sentences, which are least likely to be affected by the boundary tones of larger prosodic units. Table 4.2 provides the number of sentence-initial APs stratified by manner of articulation of AP-initial onset consonants and Table 4.3 shows the number of target APs by size. Since there are few sentence-initial APs that are longer than 5 syllables, I only include 2- to 5-syllable APs in the analysis.

Although I have tried to include at least one AP for all onset categories and AP sizes, this is not impossible for the tense category, since only five out of the total

Table 4.2: Number of the target APs by manner of articulation of AP-initial onset consonants and their laryngeal category. Cells highlighted in gray indicate H-pitch inducing categories.

	Aspirated	Tense	Lenis	Total
Stop	5	5	12	22
Affricate	5	0	8	13
Fricative	14	0	NA	14
Sonorant or vowel	NA	NA	NA	10

Table 4.3: Number of the target APs by AP size. Cells highlighted in gray represent H-pitch inducing categories. Again, /s/ and /h/ are counted as aspirated.

	Aspirated	Tense	Lenis	Sonorant or Vowel	Total
2-syllable	2	1	4	2	9
3-syllable	10	1	5	2	18
4-syllable	10	3	7	4	24
5-syllable	2	0	4	2	8

404 sentences begin with a tense onset. All of them begin with tense stops, leaving accidental gaps for the tense affricate /t͡s'/, the tense fricative /s'/, and 5-syllable APs with a tense onset.[6] However, the main goal of this chapter is to investigate the trajectory of change in Seoul Korean over time due to the introduction of the AP-initial pitch contrast (H or L), not the contrast among the H-pitch inducing categories (aspirated and tense), so the current data suffice to serve this purpose despite the accidental gaps in the tense category.

[6] These accidental gaps seem to be inevitable, considering that there are not many words beginning with a tense affricate or a tense fricative in Korean in general.

4.2.1 Stop-initial 4-syllable APs

To begin with a controlled data set, this section only examines 4-syllable APs starting with a stop consonant. In the data, there are 11 stop-initial APs that are 4 syllables long. I select 9 APs out of 11 that do not have a low vowel in the AP-initial syllable. There are three APs starting with an aspirated stop, another 3 APs starting with a tense stop, and the other 3 APs starting with a lenis stop.[7] Table 4.4 shows the target APs and the number of syllables examined in this section.[8]

Table 4.4: Sentence-initial 4-syllable APs starting with a stop consonant (number of syllables examined). A dot represents a syllable boundary. In total, 4,138 syllables are examined.

Aspirated	Tense	Lenis
/pʰi.sʌ.t͡si.ɰi/ (467) 'resort-POSS'	/p'om.nɛ.tɨs.i/ (463) 'as if showing off'	/kʌ.ki.ɛ.nɨn/ (471) 'there-TOPIC'
/pʰi.la.mi.na/ (458) 'a minnow or'	/t'ok.kat.ɨn.s'al/ (466) 'the same rice'	/kɨ.lʌ.tʌ.ni/ (456) 'after that'
/kʰɨ.ko.t͡sak.ɨn/ (438) 'of various sizes'	/k'um.sok.ɛ.sʌ/ (468) 'in the dream'	/kɨ.lʌ.ta.ka/ (451) 'and then'

Figure 4.4 shows the mean pitch values (St) of the target APs by the laryngeal categories of the initial stops and speakers' YOB (aggregated in 10 year bands). In Figure 4.4, it is clear that the AP pitch contours of older speakers are quite different

[7]Since there were a limited number of sentences in the corpus, it was impossible to control the place of articulation of the initial stops.

[8]The maximum number of syllables for each target AP is 472 (118 speakers x 4 syllables per AP), but the target APs in Table 4.4 have fewer syllables than that because some files were excluded from the analysis and sometimes Praat failed to track the pitch value of certain syllables due to either reduced vowels or strong aspiration.

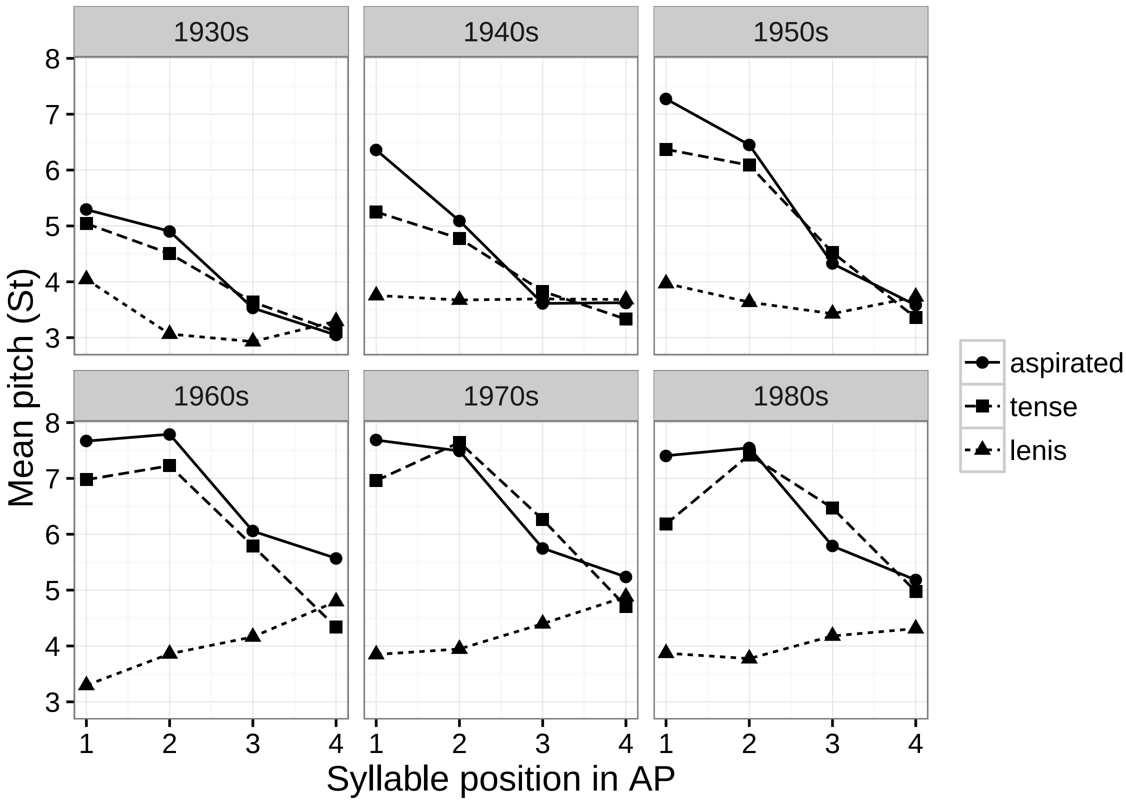


Figure 4.4: Mean pitch value (St) of each syllable in the target sentence-initial APs by the stop categories and speakers' year of birth, aggregated in 10 year bands. Syllable positions within the target APs are shown in the x-axis, where "1" means AP-initial and "4" means AP-fourth (final) syllables.

from those of younger speakers. The pitch difference between the H-pitch inducing categories (aspirated and tense) and lenis in the AP-initial syllable is much larger for younger speakers than older ones, confirming the results of previous studies on the tonogenetic sound change (Jun 1996, Silva 2006, Wright 2007, Kang 2014, among others). The AP-initial pitch difference is less than 1 St for speakers born in the 1930s, but the AP-initial pitch difference increases to about 3 St for speakers born after 1960, meaning that the pitch difference for younger speakers is three times larger than that of older speakers.

Figure 4.4 also shows that the pitch difference between H-initial and L-initial APs in the AP-second syllable increases in younger speakers, when compared to older ones. The difference among speakers born in the 1930s and 40s shows that only the AP-initial syllable is affected by the change, but for speakers born after 1950, it is clearly seen that the AP-second syllable is also produced with a remarkably higher pitch when an AP starts with a H-pitch inducing segment than when it starts with a L-pitch inducing one. This result is in line with previous studies (Silva 2006, Kang 2014), which also find a pitch difference in the AP-second syllable.

What is interesting in the result is that the pitch difference between H-initial and L-initial APs is found not only in the first two syllables of an AP but also in the remaining syllables for younger speakers. For speakers born in the 1930s, 40s, and 50s, the H-L pitch difference in the third and fourth syllables is not very noticeable. However, the pitch value of the AP-third syllable is higher in the H-initial APs than in the L-initial ones for speakers born after 1960. Considering that the AP-third syllable is the position where the intonational L of the final LH boundary tones occurs in the Korean ToBI model, it is noteworthy that younger speakers show a pitch difference of about 2 St in the AP-third syllable position.

There also exist differences in the AP-fourth syllable between older and younger speakers. Older speakers born in the 1930s, 40s, and 50s do not show a pitch difference between L-initial and H-initial APs in the AP-fourth position. The mean pitch values of the fourth syllables of tense-initial and lenis-initial APs are not that different for speakers born in the 1960s and 70s, although that of aspirated-initial APs is slightly higher in both groups. In addition, the youngest speakers, those born in the 1980s, produce the fourth syllables of both aspirated and tense-initial APs with a higher pitch than that of lenis-initial APs. However, it is clear that the effect of the AP-initial pitch context considerably decreases in the AP-fourth syllable even for younger speakers, when compared to its effect in the first syllable position.

To see if the observed differences are statistically significant, I built four linear mixed-effects models, one for each syllable position. In the analyses, the dependent variable is speakers' pitch values in each position and the independent variables are the three stop categories and speakers' YOB. Each speaker is included as a random effect. The reference category for the stops is Lenis, and YOB is centered at 1961, which is the median value in the corpus, following the method used in Kang 2014. Table 4.5 summarizes the results of the linear mixed-effects analyses.

The results confirm that Aspirated and Tense are produced with a significantly higher pitch in all AP positions except the final one ($p < 0.001$ for the AP-initial, second, and third positions). For example, the estimated pitch values of aspirated-initial and tense-initial APs are 3.31 St and 2.45 St higher than that of lenis-initial APs in the AP-initial position, respectively ($p < 0.001$ for both comparisons). While both Aspirated and Tense show a higher pitch value than Lenis in the first syllable, the models show that the pitch difference from Lenis is larger for Aspirated (3.31 St) than Tense (2.45 St), confirming the observation made in Figure 4.4. This result is in line with some of the previous studies (Kim 1994,Choi 2002, Silva 2006, Lee and Jongman 2012) that also find a pitch difference between Tense and Aspirated

Table 4.5: The outputs of linear mixed-effects models of stop-initial 4-syllable APs in the sentence-initial position by Syllable Position and YOB. Lenis is the reference category for stops, and YOB is centered at 1961. 'Est.' stands for estimated coefficients, and 'SE' represents standard errors.

	AP-initial				**AP-second**			
	Est.	SE	t-value	Pr(>|t|)	Est.	SE	t-value	Pr(>|t|)
(Intercept)	3.82	0.14	27.01	< 0.001***	3.73	0.13	28.47	< 0.001***
YOB	0.00	0.01	-0.33	0.739	0.01	0.09	1.16	0.248
Aspirated	3.31	0.14	23.81	< 0.001***	2.87	0.10	28.24	< 0.001***
Tense	2.45	0.14	18.05	< 0.001***	2.67	0.10	26.31	< 0.001***
YOB*Asp	0.04	0.01	4.16	< 0.001***	0.06	0.01	7.97	< 0.001***
YOB*Tense	0.04	0.01	4.03	< 0.001***	0.07	0.01	9.32	< 0.001***
	AP-third				**AP-final**			
	Est.	SE	t-value	Pr(>|t|)	Est.	SE	t-value	Pr(>|t|)
(Intercept)	3.86	0.13	30.91	< 0.001***	4.18	0.14	29.84	< 0.001***
YOB	0.03	0.01	3.09	0.002**	0.03	0.01	3.19	0.002**
Aspirated	0.97	0.10	9.23	< 0.001***	0.18	0.14	1.29	0.198
Tense	1.26	0.10	12.11	< 0.001***	-0.21	0.14	-1.53	0.127
YOB*Asp	0.04	0.01	5.21	< 0.001***	0.03	0.01	2.66	0.008**
YOB*Tense	0.05	0.01	6.24	< 0.001***	0.02	0.01	1.58	0.114

in the AP-initial syllable. Also, the models estimate that even the third syllables of Aspirated and Tense are 0.97 St and 1.26 St higher than that of Lenis ($p < 0.001$ for both comparisons), but the size of the pitch difference is larger for the Lenis-Tense comparison (1.26 St) than the Lenis-Aspirated in this case (0.97 St).

The effect of YOB on lenis-initial APs is not significant in the first two syllable positions, but it was significant in the last two positions ($p = 0.002$ for AP-third, $p = 0.002$ for AP-final). This seems to be because younger speakers tend to undershoot the intonational L tone of the AP-penultimate syllable in the lenis context. Speakers born in the 1930s, 40s, and 50s produce the third syllable of lenis-initial APs with a slightly lower pitch than the first and second syllables, but those born after 1960 do not show such a pitch lowering on the penultimate position, producing rather a LH-(L)H or even L-H pattern (Figure 4.4).[9] The estimated pitch value of the penultimate L tone of speakers born in 1932 is 2.99 St (= 3.86 + (0.03 x (1932 – 1961))), whereas that of speakers born in 1984 is 4.55 St (= 3.86 + (0.03 x (1984 – 1961))).[10]

Even though younger speakers produce the penultimate and final syllables of lenis-initial APs with a higher pitch than older ones, the models still estimate that the pitch difference between H-initial and L-initial APs significantly varies by speakers' YOB. The two-way interaction of Aspirated and YOB is significant in all syllable positions ($p < 0.001$ for AP-initial, second, and third syllables, $p = 0.008$ for the AP-final position). That is, the pitch difference between lenis- and aspirated-initial APs increases for younger speakers in all syllable positions, confirming the observation made in Figure 4.4. For example, the lenis-aspirated difference in the AP-initial syllable for speakers born in 1932 is 2.15 St (= 3.31 + (0.04 x (1932 – 1961))), but this difference increases to 4.23 St (= 3.31 + (0.04 x (1984 – 1961))) for speakers born in 1984. Similarly, the model estimates that the lenis-aspirated difference in the final syllable increases 0.03 St per year, showing that the pitch trajectory of

[9] Lee (1999) finds the same inter-speaker variation in the production of the intonational L tone on the penultimate syllable. He explains that the penultimate L tone undershoot is more frequently found in fast speech. Cho (2011) also finds a similar effect of speech rate on the realization of an AP pitch pattern. She shows that not all four intonational tones (LH-LH) are realized in fast speech.

[10] Since YOB is centered at 1961, the interaction coefficient is multiplied by the year difference from 1961 and the resulting value is added to the intercept in order to calculate the estimated pitch value of a speaker. A similar formula is used in Kang 2014, which uses the same corpus.

aspirated-initial APs is higher for younger speakers than older speakers.

Similarly, the lenis-tense difference significantly increases by a function of YOB in all syllable positions except the AP-final ($p < 0.001$ for the first three positions). For example, the difference increases by 0.04 St/YOB in the AP-initial, by 0.07 St/YOB in the AP-second, and 0.05 St/YOB in the AP-third positions. The lenis-tense difference in the AP-third syllable is estimated to -0.19 St (= 1.26 + (0.05 x (1932 – 1961))) for speakers born in 1932, meaning the pitch value of the third syllable in lenis-initial APs is higher than that of tense-initial ones. However, this difference increases to 2.41 St (= 1.26 + (0.05 x (1984 – 1961))) for speakers born in 1984, indicating that the pitch value of tense-initial APs is higher than that of lenis-initial ones. It confirms the observation in Figure 4.4 that tense-initial APs are produced in a higher pitch range for younger speakers than older speakers. The model also estimates that there is a 0.02 St increment in pitch per YOB in the AP-final position, but this difference is not found significant.

While both YOB x Aspirated and YOB x Tense are found to be significant in the first three syllables, it is interesting to note that the pitch increment per YOB is not consistently larger for one interaction than the other. For example, the estimated coefficient of Aspirated x YOB in the AP-initial position (0.041 St/YOB) is slightly larger than that of Tense (0.039 St/YOB). However, that of Tense is larger than that of Aspirated in the AP-second and third positions, suggesting that the pitch value of Aspirated increases faster than that of Tense in the AP-initial position, but not in the other positions.

4.2.2 Stop-initial 4-syllable APs by YOB and gender

In this section, I investigate how the AP pattern varies by age and gender, using the same data set (stop-initial 4-syllable APs) in the previous section (Table 4.4). Figure 4.5 shows the mean pitch (St) of each syllable position by YOB and Gender. Since the APs starting with an aspirated or a tense consonant show similar pitch patterns in the previous section, the two categories are combined together as the H-initial context in Figure 4.5.

Figure 4.5 shows that the gender difference for older speakers born in the 1930s, 40s, and 50s is not especially salient. The AP pitch patterns of female speakers born in the 1940s overlap with those of male speakers, and female and male speakers born in the 1950s also show somewhat overlapping pitch patterns. However, it is notable that the pitch difference between H- and L-initial APs increases rapidly for older female speakers, when compared to their male counterparts. Female speakers born in the 1930s do not show any pitch difference between H-initial and L-initial APs, whereas females born in the 1940s show a pitch difference in the AP-first and second syllables and those born in the 1950s show a pitch difference from the first to the third syllables. The pitch differences of H- and L-initial APs for older male speakers, on the other hand, do not seem to change that much.

As for younger speakers born after 1960, female speakers generally show a larger pitch difference between H-initial and L-initial APs than male speakers.[11] In par-

[11] The pitch difference between H- and L-initial APs for female speakers born in the 1960s is twice larger than that of their male cohort, especially in the first three AP positions. This seems to be

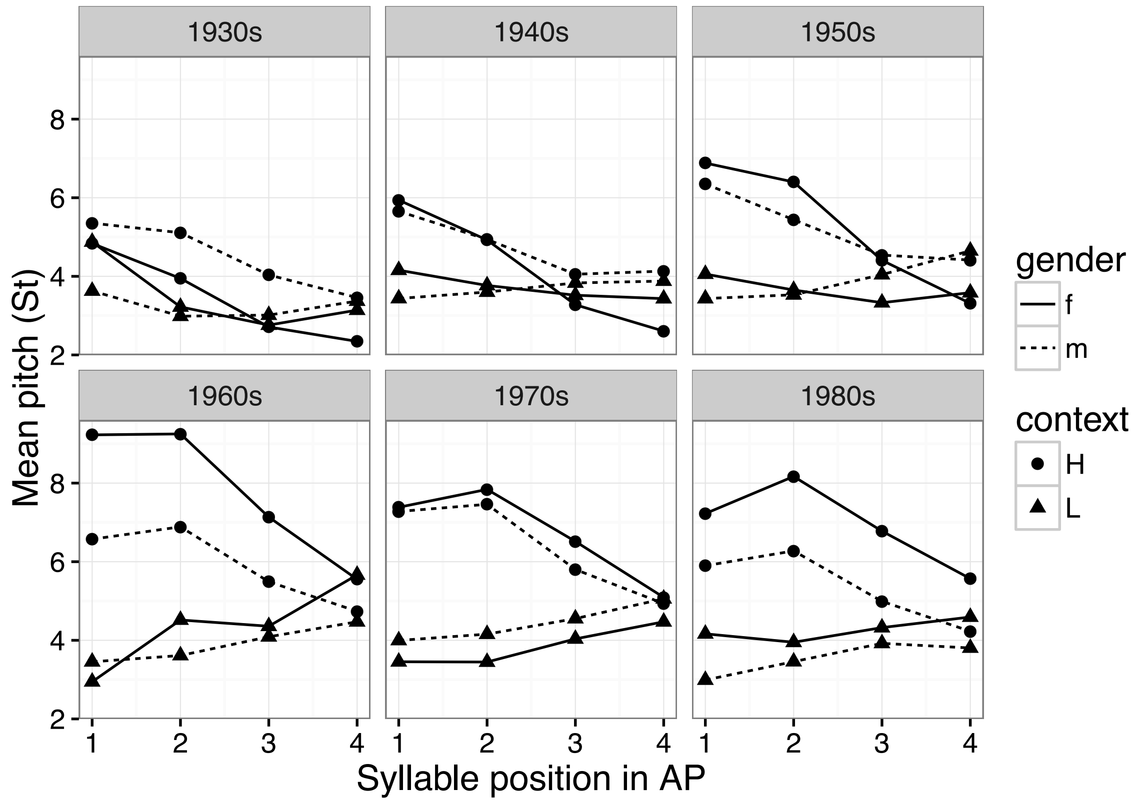


Figure 4.5: Mean pitch value (St) of each syllable in the target sentence-initial APs by speakers' gender (females: solid lines, males: dashed lines) and year of birth, aggregated in 10 year bands. Syllable positions within the target APs are shown in the x-axis, where"1" means AP-initial and "4" means AP-fourth syllables. The H-pitch inducing context includes APs starting with aspirated or tense consonants, and the L-pitch inducing context includes those starting with lenis stop consonants.

ticular, it is notable that younger female speakers produce larger pitch differences than younger male speakers in all AP positions. For example, the pitch difference for female speakers born in the 1970s is about 1.5 St in the AP third syllable, whereas that of their male cohort is less than 1 St. Similarly, females born in the 1980s exhibit a larger pitch difference between H- and L-initial APs than their male cohort in all

because two out of the three female speakers born in the 1960s have an unusually wide pitch range. One of them (speaker ID: fx08) has the widest pitch range among all 118 speakers in the data, and the other speaker (ID: fx20) has the second widest pitch range. See Appendix A for speakers' pitch ranges.

AP positions.

As for statistical analyses, I built four linear mixed-effects models (one for each syllable position) to examine how the effect of AP-initial consonants varies according to speakers' YOB and gender. In each model, the dependent variable is a pitch value (St) in a given syllable position, and independent variables include Gender, YOB, Context (H-initial vs. L-initial), and their interactions. The reference category for Gender is males, and the reference of Context is L-initial. YOB is centered at 1961, and each speaker is added as a random effect. Table 4.6 summarizes the outputs of the linear mixed-effect analyses.

The main effect of Context is significant in the AP-initial, second, and third positions, but not in the final position ($p < 0.001$ for the three positions). The models estimate that the H-L pitch difference is the largest in the AP-initial position (2.77 St), followed by the AP-second position (2.54 St). However, this pitch difference considerably decreases in the AP-third position (0.91 St), indicating that the effect of the AP-initial pitch context is weakened in the penultimate syllable due to the intonational L tone (TH-**L**H). Accordingly, the model does not find a significant effect of Context in the AP-final position.

Although the main effect of Gender is not found significant in all AP positions, the results show that the gender difference in pitch varies by Context (in the AP-second and third positions) and by the interaction of Context and YOB (in the AP-second, third, and final positions). The two-way interactions of Gender and Context in the AP-second and third syllables indicate that the pitch difference between the H and L

Table 4.6: The outputs of linear mixed-effects models of stop-initial 4-syllable APs by YOB, Gender and Context. L-initial is the reference category for the AP-initial context, and male speakers are the reference category for Gender. YOB is centered at 1961. 'Est.' stands for estimated coefficients, and 'SE' represents standard errors. Estimated coefficients are rounded at the second decimal place.

	AP-initial				**AP-second**			
	Est.	SE	t	Pr(>|t|)	Est.	SE	t	Pr(>|t|)
(Intercept)	3.66	0.20	17.86	< 0.001***	3.73	0.19	19.89	< 0.001***
YOB	0.01	0.01	0.44	0.664	0.02	0.01	1.34	0.183
Gender	0.29	0.28	1.03	0.305	-0.03	0.26	-0.11	0.915
Context	2.77	0.18	15.64	< 0.001***	2.54	0.13	20.26	< 0.001***
YOB*Gender	-0.02	0.02	-0.9	0.368	-0.01	0.02	-0.74	0.464
YOB*Context	0.03	0.01	2.74	0.006**	0.05	0.01	5.76	< 0.001***
Gender*Context	0.20	0.24	0.83	0.409	0.49	0.18	3.81	0.005**
YOB*Gen*Con	0.02	0.02	1.05	0.292	0.04	0.01	2.87	0.004**
	AP-third				**AP-final**			
	Est.	SE	t	Pr(>|t|)	Est.	SE	t	Pr(>|t|)
(Intercept)	4.07	0.18	22.94	< 0.001***	4.41	0.20	22.38	< 0.001***
YOB	0.02	0.01	2.05	0.042*	0.03	0.01	2.06	0.040*
Gender	-0.41	0.25	-1.66	0.099	-0.45	0.27	-1.64	0.102
Context	0.91	0.13	7.08	< 0.001***	0.06	0.17	0.36	0.720
YOB*Gender	0.00	0.02	0.24	0.814	0.01	0.02	0.35	0.715
YOB*Context	0.02	0.01	2.71	0.007**	0.00	0.01	-0.05	0.963
Gender*Context	0.46	0.18	2.52	0.012*	-0.09	0.24	-0.39	0.695
YOB*Gen*Con	0.05	0.01	3.67	< 0.001***	0.05	0.02	2.80	0.005**

contexts is 0.49 St and 0.46 St larger for female speakers than males in those positions, respectively ($p = 0.005$ for AP-second, $p = 0.012$ for AP-third). For example, the estimated difference in the AP-second syllable is 3.73 St for male speakers, but this difference increases to 4.22 St (= 3.73 + 0.49) for females. This effect is also shown in Figure 4.5; the H-L contrasts in the AP-second and third syllables are larger for females than males, except those born in the 1930s and 40s. Similarly, the estimated pitch difference in the AP-third position is 4.07 St for male speakers, but that of female speakers increases to 4.53 St (= 4.07 + 0.46).

The significant three-way interaction of YOB, Gender, and Context in the AP-second, third, and final positions ($p = 0.004$ for AP-second, $p < 0.001$ for AP-third, $p = 0.005$ for AP-final) suggests that the gender difference in the H-L contrast is modulated by YOB. That is, younger speakers show a larger gender difference in the H-L contrast than older ones. For example, the gender difference between H- and L-initial APs (Gender x Context) in the AP-second syllable is 0.49 St, indicating that females, in general, show a 0.49 St larger H-L contrast than males of the same age. However, the model estimates that this gender difference increases by 0.036 St/YOB. For example, the estimated gender difference for speakers born in 1932 is -0.55 St (= 0.49 + (0.036 x (1932 – 1961))), suggesting that female speakers show a 0.55 St smaller H-L difference than male speakers. However, the difference increases to 1.32 St (= 0.49 + (0.036 x (1984 – 1961))) for speakers born in 1984, suggesting that female speakers born in 1984 show a 1.32 St larger H-L contrast in the AP-second syllable than male speakers of the same age. This three-way interaction is also observed in Figure 4.5. For instance, male speakers born in the 1930s show a larger H-L contrast than their female counterparts in the AP-second syllable, but males born in the 1980s show a smaller H-L contrast than their female counterparts in that position. Similar three-way interactions are found in the AP-third and final positions, suggesting that female speakers in general produce a larger H-L contrast in non-initial positions than males, and this gender difference is larger for younger speakers than older ones.

4.2.3 All 4-syllable APs by manner of articulation

In this section, I examine different manners of articulation with all 4-syllable APs in the data to see if fricative-initial and affricate-initial APs show similar patterns to stop-initial ones observed in the previous sections. The corpus data include 24 APs that are 4 syllables long. Out of 24 APs, 10 APs start with an aspirated consonant, 3 of them with a tense consonant, 7 of them with a lenis, and the other 4 with a vowel or sonorant. Fricative-initial APs are all coded as H-initial, as previous studies show that a fricative-initial AP induces a H pitch on the AP-initial syllable (Jun 1993, Cho et al. 2002, Kang 2014, among others). The corpus does not have a sentence-initial AP starting with an aspirated or tense affricate. Accordingly, all affricate-initial APs in our data begin with a lenis affricate, so all of them are coded as L-initial in the analysis. In total, 11,199 syllables are investigated in this section. Table 4.7 shows the target APs and the number of examined syllables by their manner of articulation and their AP-initial pitch context.[12]

Figure 4.6 shows the mean pitch values of sentence-initial 4-syllable APs by the manner of articulation, the tonal context of the AP-initial consonant and speakers' YOB. What is noticeable is that the pitch trajectories of APs starting with an aspirated fricative, /s/ or /h/, do not greatly differ from those starting with a H-pitch

[12] There are three sentences starting with the same AP /ho.laŋ.i.nɨn/ 'tiger-TOPIC,' and two sentences starting with /na.mu.k'un.ɨn/ 'woodman-TOPIC.'

Table 4.7: All sentence-initial 4-syllable APs in the data (number of examined syllables) by AP-initial onset consonants. A dot represents a syllable boundary.

L-initial		
Lenis stop	Affricate (L-initial)	Sonorant or vowel
/ka.jʌp.kɛ.to/ (472) 'pitifully'	/t͡so.saŋ.ɛ.kɛ/ (469) 'ancestor-FROM'	/na.mu.k'un.ɨn/ (944) 'woodman-TOPIC'
/kʌ.ki.ɛ.nɨn/ (471) 'there-TOPIC'	/t͡sam.si.hu.ɛ/ (472) 'after a while'	/nal.kɛ.o.sɨl/ (471) 'celestial robe-ACC'
/pa.tat.ka.ɛ/ (463) 'seashore-LOC'		/ʌ.mʌ.ni.ka/ (468) 'mother-NOM'
/kɨ.lʌ.ta.ka/ (456) 'and then'		
/kɨ.lʌ.tʌ.ni/ (451) 'then'		

H-initial		
Aspirated stop	Tense stop	Fricative (H-initial)
/pʰi.sʌ.t͡si.ɰi/ (467) 'resort-POSS'	/k'um.sok.ɛ.sʌ/ (468) 'in the dream'	/ho.laŋ.i.nɨn/ (1414) 'tiger-TOPIC
/pʰi.la.mi.na/ (458) 'a minnow or'	/t'ok.kat.ɨn.s'al/ (466) 'the same rice'	/hon.t͡sa.nam.ɨn/ (472) 'left alone'
/kʰɨ.ko.t͡sak.ɨn/ (438) 'of various sizes'	/p'om.nɛ.tɨs.i/ (438) 'as if showing off'	/sa.njaŋ.k'un.ɨn/ (472) 'hunter-TOPIC
		/siŋ.siŋ.ha.ko/ (472) 'fresh'
		/sin.ha.tɨl.i/ (472) 'liege subjects-NOM'

inducing stop in all age groups. There seems to be less than a 1 St difference between those APs in the H context (solid lines in Figure 4.6), yet their pitch patterns look remarkably similar to each other. This result is in line with those of the previous studies, which suggest that this tonogenetic sound change is a structural one, affecting all AP-initial aspirated and tense consonants (Kang and Guion 2008, Kang 2014).

What is interesting in Figure 4.6 is that the pitch pattern of APs starting with

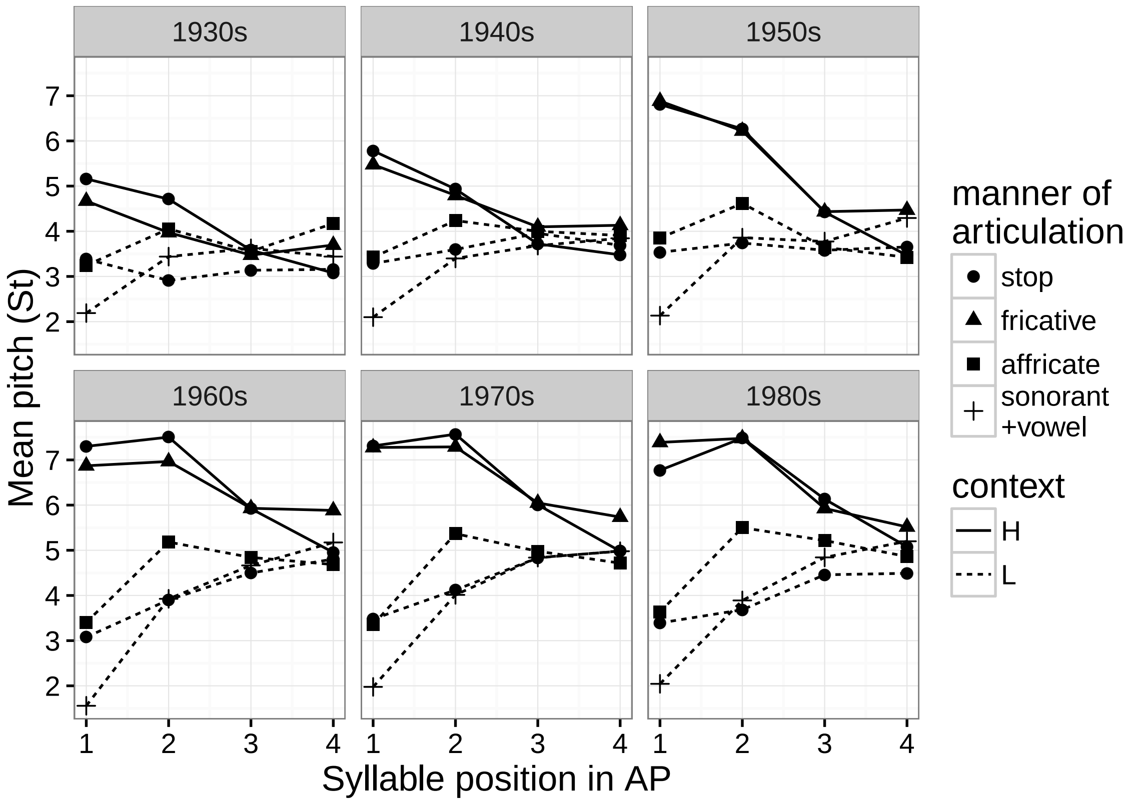


Figure 4.6: Mean pitch value (St) of each syllable in the target sentence-initial APs by manner of articulation and speakers' year of birth, aggregated in 10 year bands. The H-pitch inducing context (solid lines) includes APs starting with aspirated stops, tense stops, or aspirated fricatives, and the L-pitch inducing context (dashed lines) includes those starting with lenis stops, lenis affricates, sonorants, or vowels.

an aspirated fricative also seems to diverge from L-initial ones in younger speakers. For speakers born in the 1930s, the AP-initial syllables show a small pitch difference between the two tonal contexts, but the other syllables do not show such a difference. The speakers born in the 1940s show a larger and clearer AP-initial pitch difference than those born in the 1930s, and they also exhibit a small pitch difference in the AP second syllable. However, their pitch values of the third and final syllables in H-initial APs almost overlap with those in L-initial ones. From speakers born after

1950, the effect of H-inducing consonants is clearly seen in the AP-initial, second, and third syllables, yet again the pitch values of H-initial and L-initial APs overlap in the AP-final syllable position due to the effect of the intonational L tone in the penultimate syllable.

While the pitch patterns of L-initial APs do not differ that much among all age groups, I observe that the first syllables of sonorant or vowel-initial APs are constantly about 1 St to 1.5 St lower than those starting with a lenis stop or a lenis affricate in all age groups. This seems to be due to the cross-linguistic pattern that a vowel following a voiceless segment is produced with a higher pitch than those following a voiced segment word-initially (Lehiste and Peterson 1961, Mohr 1971, Jun 1996, Hanson 2009, among others). Since all obstruents in Korean are voiceless word-initially, it is reasonable to observe such a pitch difference between APs starting with a lenis stop or affricate and those starting with a sonorant or vowel. However, considering that this pitch difference in the first syllable position is equivalent to the effect size of physiological perturbation, the pitch difference between H-initial and L-initial APs among speakers born after 1950 (about 3 to 4 St) must not be attributed to automatic phonetic perturbation.

One more observation in Figure 4.6 is that APs starting with a lenis affricate show a higher pitch value in the AP-second syllable than the other L-initial APs. Also, this pitch difference in the AP-second position increases in younger speakers. For example, speakers born in the 1940s show about a 0.5 St pitch difference in the AP-second position, but this difference increases to about 1.5 St for speakers born in the 1980s. This seems to be because the AP-second syllable of APs starting with a lenis affricate has a fricative onset consonant. The two lenis affricate-initial APs, /tso.**saŋ**.ɛ.kɛ/ 'ancestor-from' and /t͡sam.**si**.hu.ɛ/ 'after a while,' happen to have an aspirated fricative /s/ in the second syllable. On the other hand, all second syllables of APs starting with other L-pitch inducing consonants do not have a H-inducing onset consonant (Table 4.7). This segmental difference seems to cause the observed pitch difference on the second syllable among L-initial APs. However, the size of the increment for younger speakers is quite small, when compared to the pitch contrast from the AP-initial consonant context, suggesting that it seems to be due to a segment-induced pitch.

As for statistical analyses, 4 linear mixed-effects models (one for each syllable position) were built to examine if APs starting with distinct manners of articulation and laryngeal gestures yield significant pitch differences. In the models, the dependent variable is the pitch value (St) of a given syllable position, and independent variables are manner of articulation of AP-initial onsets and speakers' YOB. As for the manner of articulation, stops are divided into H-inducing and L-inducing ones and APs starting with a lenis stop (L-initial Stop) are the reference category for manner of articulation.[13] Each speaker is added as a random effect. Table 4.8 summarizes the outputs of the linear mixed-effects models.

The models estimate that the pitch values of fricative-initial APs are significantly

[13] Note that fricatives are not divided by their tonal context, because all fricative-initial APs are in the H context. Also, the models do not include H-inducing Affricate, because all affricate-initial APs in the data start with the lenis one.

Table 4.8: The outputs of linear mixed-effects models of sentence-initial 4-syllable APs by YOB and manner of articulation of AP-initial onsets. A lenis stop is the reference category for Manner of articulation, and YOB is centered at 1961. 'Est.' stands for estimated coefficients, and 'SE' represents standard errors. 'Sonorant' includes APs starting with a sonorant or a vowel.

	AP-initial				**AP-second**			
	Est.	SE	t	Pr(>|t|)	Est.	SE	t	Pr(>|t|)
(Intercept)	3.39	0.12	29.41	< 0.001***	3.77	0.12	31.92	< 0.001***
L-Affricate	0.13	0.12	1.14	0.255	1.08	0.10	11.12	< 0.001***
H-Fricative	3.22	0.08	39.01	< 0.001***	2.50	0.07	36.72	< 0.001***
H-Stop	3.29	0.09	38.10	< 0.001***	2.73	0.07	38.56	< 0.001***
Sonorant	-1.36	0.09	-14.42	< 0.001***	0.03	0.08	0.33	0.740
YOB	0.00	0.01	-0.09	0.932	0.01	0.01	1.67	0.097
YOB*L-Affricate	0.00	0.01	-0.42	0.674	0.02	0.01	3.01	0.003**
YOB*H-Fricative	0.05	0.01	8.78	< 0.001***	0.06	0.01	12.58	< 0.001***
YOB*H-Stop	0.04	0.01	6.14	< 0.001***	0.06	0.01	11.77	< 0.001***
YOB*Sonorant	-0.01	0.01	-0.89	0.375	0.00	0.01	-0.47	0.641
	AP-third				**AP-final**			
	Est.	SE	t	Pr(>|t|)	Est.	SE	t	Pr(>|t|)
(Intercept)	4.14	0.11	36.24	< 0.001***	4.19	0.13	33.04	< 0.001***
L-Affricate	0.20	0.11	1.90	0.058	0.00	0.12	0.01	0.990
H-Fricative	0.89	0.08	11.89	< 0.001***	0.75	0.08	8.85	< 0.001***
H-Stop	0.83	0.08	10.69	< 0.001***	-0.02	0.09	-0.26	0.793
Sonorant	0.09	0.09	1.04	0.299	0.33	0.10	3.35	< 0.001***
YOB	0.03	0.01	3.92	< 0.001***	0.04	0.01	4.18	< 0.001***
YOB*L-Affricate	0.01	0.01	1.10	0.270	-0.01	0.01	-0.87	0.387
YOB*H-Fricative	0.03	0.01	5.80	< 0.001***	0.01	0.01	1.94	0.052
YOB*H-Stop	0.04	0.01	6.89	< 0.001***	0.01	0.01	2.40	0.017*
YOB*Sonorant	0.01	0.01	0.52	0.606	0.00	0.01	-0.24	0.808

higher than those starting with a lenis stop (the reference category) in all syllable positions ($p < 0.001$ for all comparisons). The estimated pitch value of the AP-initial syllable is 3.39 St for APs with a lenis stop onset, but this increases to 6.61 St (= 3.39 + 3.22) for fricative-initial APs. A similar trend is found even in the AP-final syllables, but the pitch difference between the two APs (0.75 St) is not as large as the one found in the AP-first syllable. The models also estimate that the effect of Fricative is quite similar to that of H-inducing Stop in all syllable positions, except the final. For example, the estimated coefficient of Fricative is 3.22 St and that of H-inducing Stop is 3.29 St for the AP-initial position, showing that the first syllables of Fricative and H-inducing Stop are not that different from each other. The two consonant categories are also similar in the AP-second (Fricative: 2.5 St, H-inducing Stop: 2.73 St) and AP-third positions (Fricative: 0.89 St, H-inducing Stop: 0.83 St). This result suggests that both aspirated fricatives and H-inducing stops belong to the same category (aspirated) and they pattern together with regard to this tonogenetic sound change.

In addition, the models find that the interaction of YOB and Fricative is significant in the first three AP-positions ($p < 0.001$ for AP-initial, second, and third syllables). This result indicates that the pitch difference between APs with lenis stop onsets and those with fricative onsets increases in younger speakers. The size of increment is the largest for AP-second syllables (0.06 St increment per year) and the smallest for AP-third syllables (0.03 St increment per year). For example, the estimated pitch difference in the AP-second syllables of Lenis Stop and Fricative is 1.48 St (= 3.22 + (0.06 x (1932 – 1961))) for speakers born in 1932, yet this difference increases to 4.6 St (= 3.22 + (0.06 x (1984 – 1961))) for those born in 1984. That is, speakers born in 1984 produce a 3.12 St larger pitch difference (= 4.6 – 1.48) on the AP-second position than those born in 1932. This difference between speakers born in 1932 and those born in 1984 decreases to 0.52 St (= 0.01 x (1984 – 1932)) in the AP-final position, indicating that the effect of AP-initial Fricative decreases in the final syllable, although the model finds this effect only marginally significant (p = 0.052). Also, it is notable that the interaction of YOB and Fricative has a similar size of effect to that of YOB and H-inducing Stop in all AP positions (Table 4.8), confirming again that this tonogenetic sound change is a structural one.

As for the comparison of Lenis Stop and Lenis Affricates, those two APs are not significantly different from one another in all syllable positions, except the AP-second. As observed in Figure 4.6, the estimated pitch value of Lenis Affricate is 1.08 St higher than that of Lenis Stop due to the effect of H-inducing consonant /s/ in that position ($p < 0.001$). Also, the interaction of Lenis Affricate and YOB in the AP-second syllable is found to be significant ($p = 0.003$), suggesting that the pitch difference between Lenis Stop and Lenis Affricate increases by 0.02 St/YOB. For example, speakers born in 1932 show a 0.5 St pitch difference (= 1.08 + (0.02 x (1932 – 1961))) between Lenis Affricate and Lenis Stop, but this difference is augmented to 1.54 St for speakers born in 1984. These results indicate that a H-pitch inducing consonant in non-initial syllables may have an effect on pitch for younger speakers, raising the pitch values of the given syllables. This result is in line with Chapter 3, which also finds that younger speakers (in their 20s) produce a small pitch increment in non-initial syllables with a H-pitch inducing onset consonant. However, the fricative /s/

in the AP-second position plays a less important role than H-inducing consonants in the AP-initial position, which show a global effect raising pitch values of all syllables in the given AP. The result indicates that H-inducing consonants in the AP-initial position have a global effect on the pitch range of the entire AP, whereas those in non-initial APs show a local effect on the pitch value of the given syllable only.

Lastly, APs starting with a sonorant or a vowel are found to be significantly different from those starting with a lenis stop in the AP-initial and AP-final syllables ($p < 0.001$ for both positions). The significant difference in the AP-initial position seems to be because lenis stops are voiceless in Seoul Korean. The crosslinguistic tendency for voiceless stops to induce a high pitch on the following vowel explains the fact that the initial syllable of APs starting with a lenis stop shows a 1.36 St higher pitch value than those starting with a sonorant or vowel. Although it is not clear why the pitch difference in the AP-final syllable is significant, it is noteworthy that the interaction between YOB and Sonorant is not significant in any syllable, indicating that the pitch differences between Lenis Stop and Sonorant are not due to the tonogenetic sound change.

4.2.4 All sentence-initial APs by YOB, gender, and AP size

In this section, I look at all sentence-initial APs in the data by their AP size to investigate if APs of different sizes show similar patterns observed in the previous sections. I combine all APs starting with a H-inducing consonant as H-initial and APs starting with a lenis consonant, sonorant, or a vowel as L-initial for the sake of simplicity.

4.2.4.1 Sentence-initial 2-syllable APs

I first begin with 2-syllable APs. The data contain nine 2-syllable APs; 2 of them start with aspirated consonants, 1 with a tense consonant, 4 with lenis consonants, and the other 2 with nasals. The total number of syllables examined is 2,080. Table 4.9 shows the target APs and their respective number of syllables by AP-initial pitch context.

Table 4.9: 2-syllable APs by their AP-initial pitch context (number of syllables examined). A dot marks a syllable boundary.

H-initial	/t͡sʰa.ga/ (229) 'car-NOM'	/t͡sʰʌŋ.so/ (236) 'cleaning'	/p'oŋ.p'oŋ/ (236) 'bang, bang!'
L-initial	/na.nɨn/ (236) 'I-TOPIC'	/na.ɰi/ (236) 'my'	/t͡sʌ.lɨl/ (234) 'me (humble)-ACC'
	/kɨ.t'ɛ/ (230) 'that time'	/kɨ.kʌ/ (220) 'that'	/t͡sʌ.nɨn/ (233) 'me (humble)-TOPIC'

Figure 4.7 shows the mean pitch values of the 2-syllable APs by Context, Gender, and YOB. Here, the main difference from 4-syllable APs is that the pitch difference between H-initial and L-initial APs does not dramatically decrease in the AP-final (second) position for younger speakers. Figure 4.7 clearly shows that the pitch differences between the H-initial and L-initial contexts in both AP-initial and second positions increase in younger speakers. For example, the pitch trajectories of the H and L contexts converge in the AP-second syllable for speakers born in the 1930s, but the pitch patterns do not converge in younger speakers, showing a large pitch

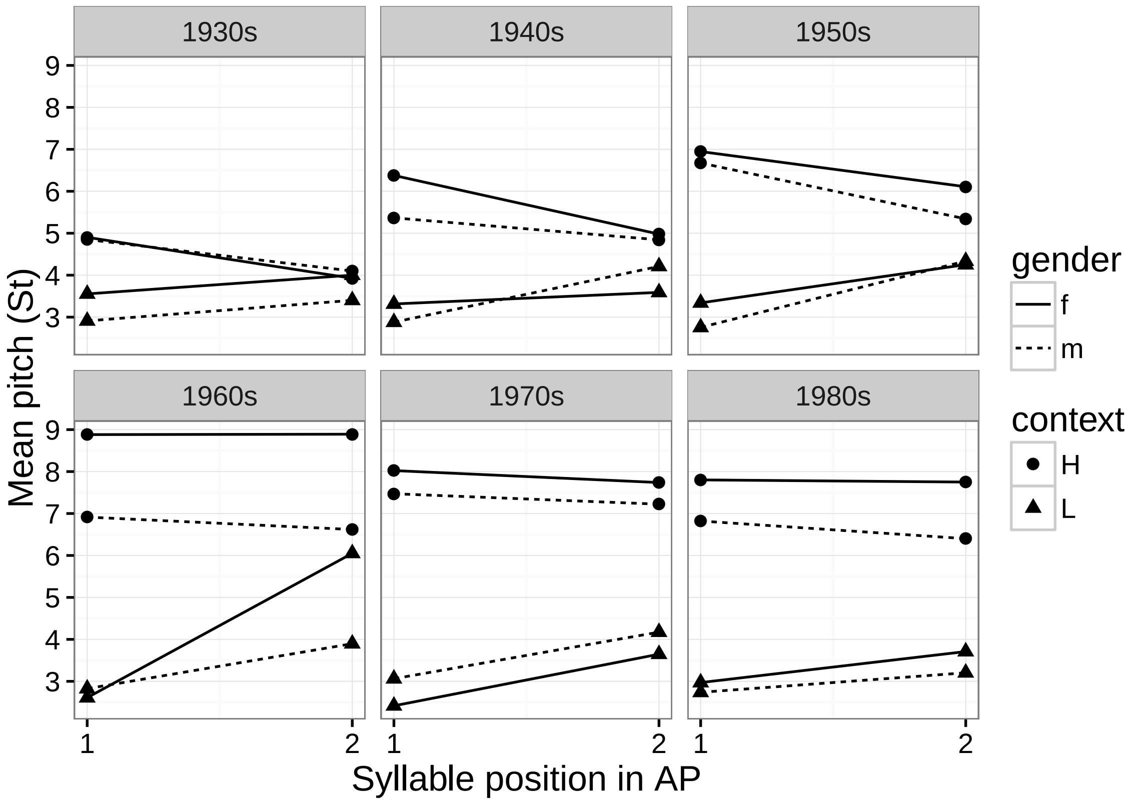


Figure 4.7: Mean pitch value (St) of each syllable in the 2-syllable APs by Context, Gender, and YOB. Solid lines show females, and dotted lines show males. H-initial includes all APs starting with aspirated or tense consonants, and the other APs are L-initial.

difference (about 3 to 4 St) in the second syllable. The gender difference is not as large as the context difference, but in general, female speakers seem to produce a larger pitch difference between H-initial and L-initial contexts than male speakers.

I conduct two linear mixed-effects analyses to examine the effect of Context, Gender, and YOB in 2-syllable APs. In the models, the dependent variable is the pitch values (St) of each syllable position and independent variables are their pitch contexts (H-initial vs. L-initial), Gender, and YOB. The reference category of Gender is male, and the reference of Context is L-initial. YOB is centered at 1961, and speakers are included as a random effect. Table 4.10 summarizes the outputs of the models.

Table 4.10: The outputs of linear mixed-effects models of sentence-initial 2-syllable APs by YOB, Gender, and Context. The reference category for Gender is males, and that of Context is L-initial. YOB is centered at 1961.

AP-initial	Estimate	Std. Error	t-value	Pr ($>\|t\|$)
(Intercept)	2.93	0.15	19.86	< 0.001***
Gender	0.15	0.21	0.73	0.469
YOB	0.00	0.01	0.00	0.998
Context	3.59	0.15	24.27	< 0.001***
YOB*Gender	-0.02	0.01	-1.60	0.111
Gender*Context	0.53	0.21	2.59	0.009**
YOB*Context	0.05	0.01	5.35	< 0.001***
YOB*Gender*Context	0.02	0.01	1.41	0.158
AP-second	Estimate	Std. Error	t-value	Pr ($>\|t\|$)
(Intercept)	4.03	0.16	24.87	< 0.001***
Gender	0.01	0.23	0.05	0.963
YOB	0.00	0.01	-0.27	0.792
Context	2.01	0.14	14.55	< 0.001***
YOB*Gender	-0.01	0.02	-0.42	0.679
Gender*Context	0.47	0.19	2.45	0.015*
YOB*Context	0.07	0.01	7.82	< 0.001***
YOB*Gender*Context	0.02	0.01	1.40	0.161

The models estimate that the effect of Context and two-way interactions with Context (Gender x Context and Context x YOB) are significant in both positions, while the others are not. The significant main effect of Context ($p < 0.001$ in both positions) indicates that the syllables in the H-initial context are produced with a higher pitch value (3.59 St for AP-initial, 2.01 St for AP-second) than those in the L-initial context in both AP-initial and second positions. Also, the significant interaction of Context with Gender (p = 0.009 for AP-initial, p = 0.015 for AP-second) indicates that the pitch difference between the H and L contexts is larger for female speakers than male speakers in both AP positions. The models estimate that the pitch difference between H-initial and L-initial APs is 0.53 St and 0.47 St larger for females than males in the AP-initial and second syllables, respectively.

It is also found that the two-way interaction of YOB and Context is significant, suggesting that the H-L contrast increases in younger speakers in both syllable positions (0.05 St/YOB for AP-initial and 0.07 St/YOB for AP-second). For example, the estimated H-L pitch difference of the AP-initial syllable is 2.14 St (= 3.59 + 0.05 x (1932 – 1961)) for speakers born in 1932, but it increases to 4.74 St (= 3.59 + 0.05 x (1984 – 1961)) for speakers born in 1984. This suggests that the youngest cohort produces a pitch difference twice as large as the oldest cohort in the AP-initial position. The three-way interaction of YOB, Gender, and Context is not significant in either position, indicating that the gender difference in the H-L contrast neither increases nor decreases over time.

4.2.4.2 Sentence-initial 3-syllable APs

Now we turn our attention to sentence-initial 3-syllable APs in the data. There are eighteen 3-syllable APs, and 10 of them start with an aspirated stop or fricative consonant, 1 of them with a tense consonant, 5 of them with a lenis consonant, and the other 2 with a vowel. In total, 6,310 syllables (3,873 syllables in the H-initial context + 2,427 syllables in the L-initial context) are examined. Table 4.11 shows the target

APs and the number of examined syllables by their AP-initial pitch contexts.

Table 4.11: Sentence-initial 3-syllable APs by their AP-initial tonal context (number of syllables examined). A dot marks a syllable boundary.

H-initial	/sʌn.njʌ.nɨn/ 'angel-TOPIC (705)	/t͡sʰam.ki.lɨm/ 'sesame oil' (354)
	/ha.t͡si.man/ 'but' (352)	/hun.hun.han/ 'heartwarming' (351)
	/t'a.sɨ.han/ 'warm' (354)	/t͡sʰak.po.lɨl/ 'book wrapper-ACC' (346)
	/tʰo.k'i.wa/ 'rabbit-and' (346)	/tʰo.k'i.ka/ 'rabbit-NOM' (345)
	/sa.sɨm.ɨn/ 'deer-TOPIC' (354)	/sa.sɨm.ɰi/ 'deer-POSS' (354)
L-initial	/ʌm.ma.nɨn/ 'mom-TOPIC' (354)	/t͡sa.la.ɰi/ 'turtle-POSS' (354)
	/kɨ.k'ol.i/ 'that form-NOM' (352)	/t͡sa.la.nɨn/ 'turtle-TOPIC' (351)
	/i.lɨl.pon/ 'seeing this' (348)	/kɨ.lɛ.sʌ/ 'so' (352)
	/kɨ.lʌn.tɛ/ 'then' (336)	

Figure 4.8 shows the mean pitch values of the target 3-syllable APs by YOB and Gender. This figure exhibits a similar pattern to that of 2-syllable APs. An AP-initial pitch contrast is not clear for speakers born in the 1930s, but it is clearly seen for younger speakers (those born after 1950). Also, the pitch difference in the AP-second and third syllables between H-initial and L-initial contexts increases in younger speakers, when compared to older ones. As for the gender comparison, female speakers in general show a larger pitch difference between H-initial and L-initial APs than male speakers. Also, the development of the AP-initial pitch contrast is more clearly observed in older female speakers (from the 1930s to the 1950s) than their male counterparts, indicating that the tonogenetic sound change has been led by female speakers.

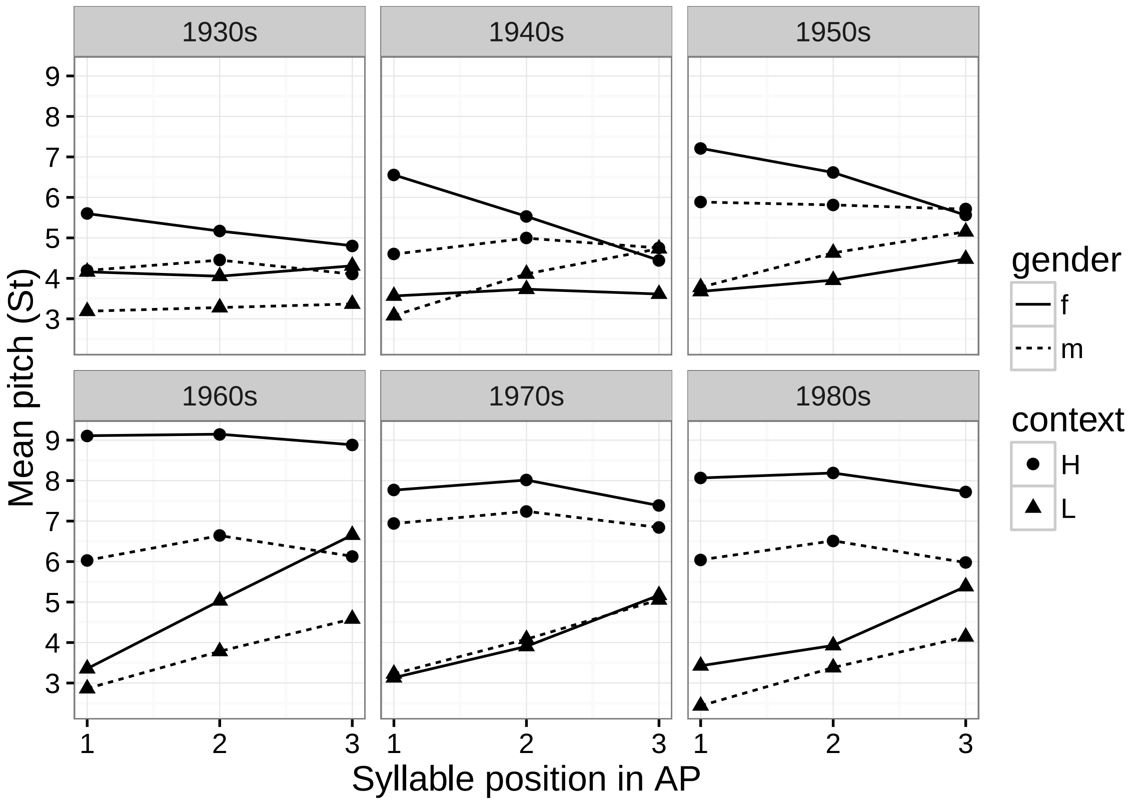


Figure 4.8: Mean pitch value (St) of each syllable in sentence-initial 3-syllable APs by Context, Gender, and YOB. Solid lines show females, and dotted lines show males. H-initial includes all APs starting with aspirated or tense consonants, and the other APs are L-initial.

I build 3 linear mixed-effects models (one for each syllable position) for statistical analyses with the pitch values (St) as a dependent variable and Gender, YOB, and Context as independent variables. Male speakers are the reference category for Gender, and L-initial is the reference category for Context. Each speaker is added as a random effect, and YOB is centered at 1961. Table 4.12 summarizes the outputs of the analyses.

The models show that the main effect of Context and the two-way interactions with Context (Context x Gender and Context x YOB) are significant in all syllable

Table 4.12: The outputs of the linear mixed-effect models of sentence-initial 3-syllable APs by Gender, YOB, and Context. The reference category for Gender is males, and that of Context is L-initial. YOB is centered at 1961.

AP-initial	Estimate	Std. Error	t-value	Pr (>$\|t\|$)
(Intercept)	3.15	0.16	19.09	< 0.001***
Gender	0.37	0.23	1.60	0.112
YOB	-0.01	0.01	-0.69	0.493
Context	2.68	0.10	26.70	< 0.001***
YOB*Gender	-0.01	0.02	-0.43	0.665
Gender*Context	1.16	0.14	8.25	< 0.001***
YOB*Context	0.07	0.01	9.91	< 0.001***
YOB*Gender*Context	-0.01	0.01	-0.88	0.380
AP-second	Estimate	Std. Error	t-value	Pr (>$\|t\|$)
(Intercept)	3.98	0.17	24.13	< 0.001***
Gender	-0.01	0.23	-0.06	0.949
YOB	0.00	0.01	-0.20	0.840
Context	2.18	0.09	23.84	< 0.001***
YOB*Gender	0.00	0.02	0.18	0.858
Gender*Context	0.85	0.13	6.66	< 0.001***
YOB*Context	0.06	0.01	10.68	< 0.001***
YOB*Gender*Context	0.01	0.01	0.79	0.428
AP-third	Estimate	Std. Error	t-value	Pr (>$\|t\|$)
(Intercept)	4.70	0.19	25.06	< 0.001***
Gender	0.00	0.26	0.01	0.991
YOB	0.01	0.01	1.14	0.256
Context	1.10	0.09	12.23	< 0.001***
YOB*Gender	0.02	0.02	1.26	0.211
Gender*Context	0.36	0.13	2.85	0.004**
YOB*Context	0.04	0.01	7.31	< 0.001***
YOB*Gender*Context	0.00	0.01	0.54	0.590

positions in 3-syllable APs ($p = 0.004$ for Gender x Context in the AP-third syllable, $p < 0.001$ for the other comparisons), strikingly similar to the results of 2-syllable APs. The main effect of Context indicates that the AP-initial tonal context is an important factor in deciding the pitch values of all syllable positions within 3-syllable APs. For example, the pitch values of syllables in the H context are 2.68 St, 2.18 St, and 1.1 St higher than those in the L context in the AP-first, second, and third positions, respectively.

Also, the models estimate that female speakers show a larger pitch difference between H-initial and L-initial APs than male speakers. In the AP-initial position, female speakers show a 1.16 St larger H-L contrast than males, and female speakers are estimated to produce a 0.36 St larger H-L contrast than male speakers even in the AP-third position. These results again suggest that the tonogenetic sound change has been led by female speakers.

Lastly, the models reveal that the two-way interaction of YOB and Context is significant ($p < 0.001$ in all comparisons), indicating that the pitch difference between H-initial and L-initial APs increases in younger speakers. The estimated pitch difference between those two contexts in the AP-initial position is 0.65 St (= 2.68 + (0.07 x (1932 – 1961))) for speakers born in 1932, but this difference increases to 4.29 St (= 2.68 + (0.07 x (1984 – 1961))) for speakers born in 1984, which is about six to seven times larger than that of speakers born in 1932. This effect is significant even in the AP-third syllable ($p < 0.001$), where the model estimates that the difference in the H-L contrast between speakers born in 1931 and 1984 is 4.74 St (= 1.1 + (0.07

x (1984 – 1932))). The three-way interactions of Gender, YOB, and Context are not significant in all AP positions.

4.2.4.3 Sentence-initial 4-syllable APs

Moving on to 4-syllable APs, Figure 4.9 shows the mean pitch value of each syllable position in 4-syllable APs by Context, Gender, and YOB, aggregated in 10 year bands. The 4-syllable APs used in this analysis are listed in Table 4.7.[14] Here, similar AP pitch patterns to those of 2-syllable or 3-syllable APs in the initial and second syllable positions are observed. That is, the AP-initial pitch difference increases in younger speakers, and it is larger for females than for males. Similarly, the pitch difference in the AP-second position is larger for younger speakers than older ones, and also larger for females than males.

The main difference between 3-syllable and 4-syllable APs is that the effect of intonational L tone in TH-**L**H is clearly seen in 4-syllable APs. The H-L contrast is quite large in the third syllable position of 3-syllable APs (Figure 4.8), yet it is relatively small in the AP-third position in 4-syllable APs (Figure 4.9). However, it is still notable that the pitch difference in the AP-third syllable between the two tonal contexts increases in younger speakers in Figure 4.9. Speakers born in the 1930s and 40s do not show a pitch difference in the AP-third position, but female speakers born in the 1950s do exhibit a pitch difference in that position, indicating that females are the ones who seem to have started to raise pitch values of later syllables in the H-

[14] Note that all 4-syllable APs are examined in this section, whereas only 9 APs (stop-initial) are included in Section 4.2.2.

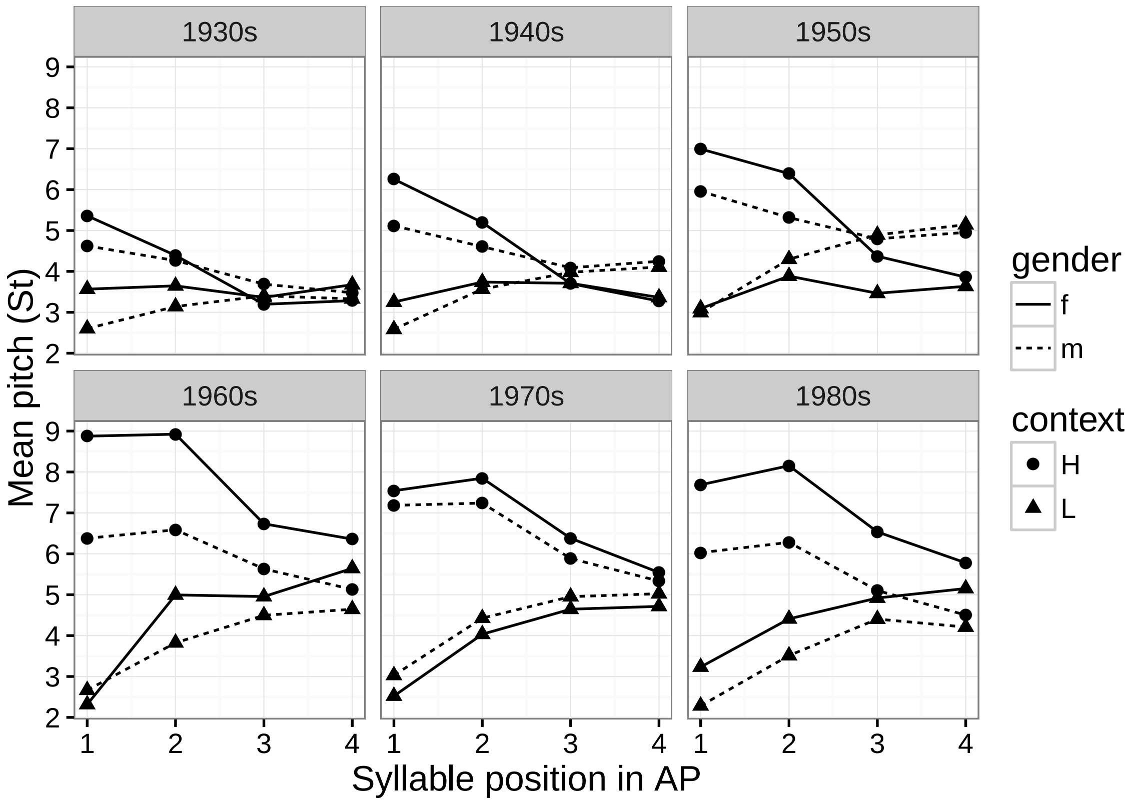


Figure 4.9: Mean pitch value (St) of each syllable in sentence-initial 4-syllable APs by Context, Gender, and YOB. Solid lines show females, and dotted lines show males. H-initial includes all APs starting with aspirated or tense consonants, and the other APs are L-initial.

initial context. Speakers born after 1960 show a pitch difference in the third syllable and it is larger for females than males. The effect of the AP-initial tonal contexts decreases in the final syllable in general, yet younger generations still show a larger pitch difference than older ones.

For statistical analyses, I build 4 linear mixed-effects analyses—one for each syllable position—with the pitch values (St) as the dependent variable and Gender, YOB, and Context as independent variables. Table 4.13 summarizes the outputs of the models.

Table 4.13: The outputs of the linear mixed-effect models of sentence-initial 4-syllable APs by Gender, YOB, and Context. The reference category for Gender is males, and that of Context is L-initial. YOB is centered at 1961. 'Est.' stands for estimated coefficients, and 'SE' represents standard errors. Estimated coefficients are rounded to the second decimal place.

	AP-initial				**AP-second**			
	Est.	SE	t	Pr (>|t|)	Est.	SE	t	Pr (>|t|)
(Intercept)	2.78	0.15	19.00	< 0.001***	3.94	0.16	24.86	< 0.001***
Gender	0.22	0.21	1.10	0.275	0.07	0.22	0.33	0.741
YOB	0.00	0.01	0.28	0.782	0.02	0.01	1.76	0.081
Context	3.35	0.09	38.80	< 0.001***	2.04	0.07	29.21	< 0.001***
YOB*Gender	-0.01	0.01	-0.95	0.345	0.00	0.02	-0.22	0.829
Gender*Context	0.75	0.12	6.26	< 0.001***	0.73	0.10	7.43	< 0.001***
YOB*Context	0.05	0.01	8.89	< 0.001***	0.05	0.00	11.87	< 0.001***
YOB*Gen*Con	0.00	0.01	0.06	0.955	0.01	0.01	1.49	0.137
	AP-third				**AP-final**			
	Est.	SE	t	Pr (>|t|)	Est.	SE	t	Pr (>|t|)
(Intercept)	4.45	0.15	29.40	< 0.001***	4.53	0.17	27.22	< 0.001***
Gender	-0.45	0.21	-2.14	0.034*	-0.42	0.23	-1.81	0.072
YOB	0.03	0.01	2.69	0.008**	0.02	0.01	2.28	0.024*
Context	0.59	0.07	7.82	< 0.001***	0.23	0.09	2.67	0.008**
YOB*Gender	0.01	0.01	0.87	0.385	0.02	0.02	1.21	0.230
Gender*Context	0.43	0.11	4.13	< 0.001***	0.11	0.12	0.93	0.354
YOB*Context	0.02	0.00	4.96	< 0.001***	0.01	0.01	1.54	0.124
YOB*Gen*Con	0.02	0.01	2.52	0.012*	0.02	0.01	1.81	0.070

The significant main effect of Context ($p = 0.008$ in AP-final, $p < 0.001$ in the other positions) indicates that the pitch values of syllables in H-initial APs are larger in all syllable positions than those in the L context. The estimated coefficient of Context is the largest in AP-initial (3.35 St) and the smallest in AP-final (0.23 St), confirming the observation in Figure 4.9 that the pitch difference between H and L-initial APs is the largest in the AP-initial position, but it decreases after the penultimate syllable.

The models also reveal that the two-way interactions with Context are significant in all AP syllables, except the final. The significant two-way interaction of YOB x Context ($p < 0.001$ in the AP-initial, second, and third positions) suggests that the effect of Context varies by a function of speakers' age. For example, the estimated pitch difference in the AP-initial position is 1.9 St (= 3.35 + (0.05 x (1932 – 1961))) for speakers born in 1932, but it increases to 4.5 St (= 3.35 + (0.05 x (1984 – 1961))) for speakers born in 1984. This effect is also found to be significant in the AP-third position, where the intonational L tone falls ($p < 0.001$). The model estimates that the pitch difference for speakers born in 1932 is 0.01 St (= 0.59 + (0.02 x (1932 – 1961))), meaning that the AP-third syllable in the H-initial context is almost the same with the one in the L-initial context. However, this difference increases to 1.05 St (= 0.59 + (0.02 x (1984 – 1961))) for speakers born in 1984. These results confirm the observation that the pitch difference between the H and L contexts increases in younger speakers.

As for the gender comparison, the models find that the two-way interaction of Gender and Context is significant in AP-initial, second, and third positions ($p < 0.001$ in all three cases), indicating that the pitch difference between the H and L contexts is larger for females than males. For example, in the AP-initial position, the model estimates that female speakers show a 0.75 St larger pitch difference between the two contexts than male speakers. Similarly, female speakers produce 0.73 St and 0.43 St larger pitch differences than male speakers in the AP-second and third syllables, respectively. This result confirms that females are more advanced than males in terms of this sound change.

4.2.4.4 Sentence-initial 5-syllable APs

Lastly, this section examines 5-syllable APs that are located sentence-initially. The data include eight 5-syllable APs. Out of 8, 2 APs start with an aspirated consonant, 4 of them with a lenis consonant, and the other with a vowel. The total number of syllables examined is 4,637. Table 4.14 shows the target APs by the AP-initial tonal contexts, and Figure 4.10 provides the mean pitch value of each syllable of the target APs by Gender, Context, and YOB.

Table 4.14: Sentence-initial 5-syllable APs by their AP-initial context (number of syllables examined). A dot marks a syllable boundary.

H-initial	
/t͡sʰu.u.sil.tʰɛn.tɛ/ (560) '(I suppose you feel) cold-honorific'	/ho.laŋ.i.ɛ.kɛ/ (590) 'to the tiger'
L-initial	
/wit.toŋ.nɛ.ɛ.sʌ/ (583) 'from an upper town'	/twɛ.to.lok.i.mjʌn/ (589) 'if possible'
/wit.ma.ɨl.ɛ.sʌ/ (585) 'from an upper village'	/t͡sin.nun.k'ɛ.pi.ka/ (590) 'sleet-NOM'
/koŋ.sa.ha.nɨ.la/ (580) 'while in construction'	/t͡sik.t͡saŋ.sɛŋ.hwal.ɨl/ (560) 'working-ACC'

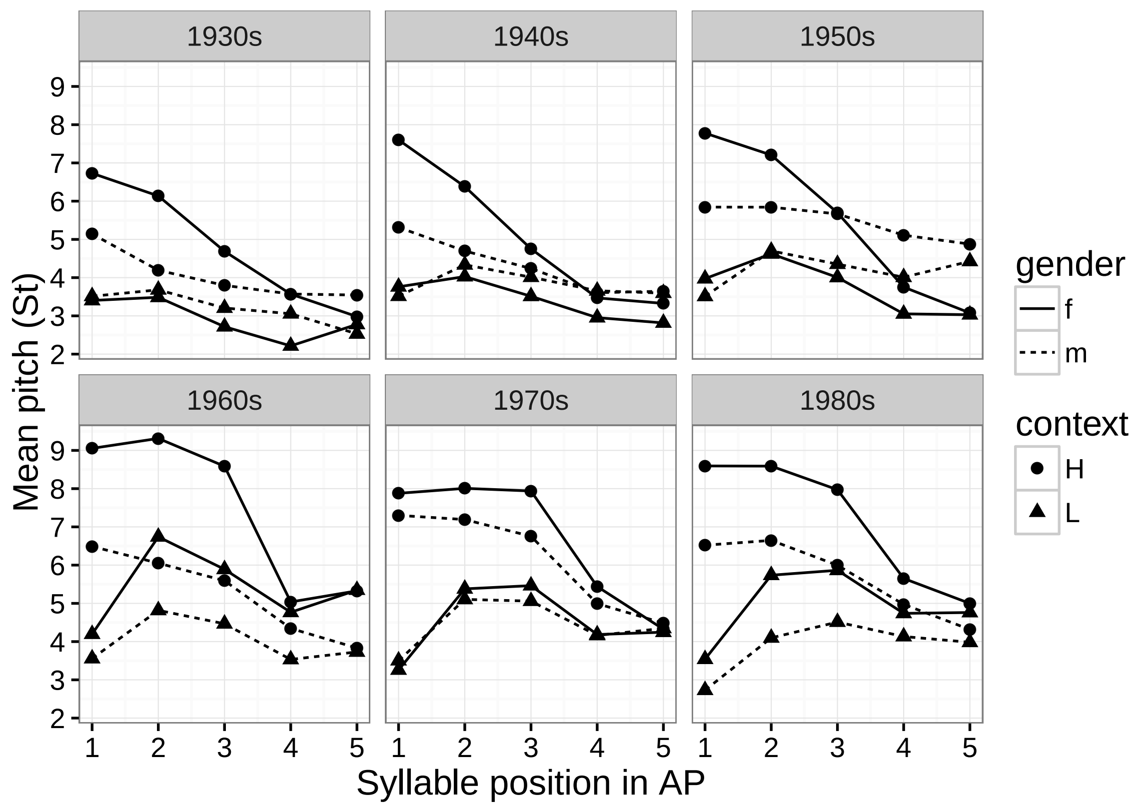


Figure 4.10: Mean pitch value (St) of each syllable in sentence-initial 5-syllable APs by Context, Gender, and YOB. Solid lines show females, and dotted lines show males. H-initial includes all APs starting with aspirated or tense consonants, and the other APs are L-initial.

In Figure 4.10, it is observed that 5-syllable APs show a quite similar pitch pattern to 4-syllable ones. Younger speakers show a large pitch difference in the AP-initial, second, and third syllables, but this pitch difference decreases after the intonational L tone on the penultimate syllable. Younger speakers born after 1970 seem to produce a pitch difference on the AP-fourth position, where the intonational L tone falls, but the pitch values of the final syllables almost overlap even for those younger speakers. This result is in line with the results in Chapter 3, where young speakers in their 20s show overlapping pitch values in H-initial and L-initial contexts due to the effect of the final H tone of the AP melody. Furthermore, similar to the results of the previous sections, female speakers show a larger pitch difference than males.

As for statistical analyses, 5 linear mixed-effects models were built to examine if the observations are statistically significant. In these models, pitch values are the dependent variable and Gender, Context, and YOB are added as independent variables. Each speaker is treated as a random effect. Table 4.15 summarizes the outputs of the models.

The significant main effect of Context from AP-initial to AP-fourth positions ($p < 0.001$ for all cases) confirms the observation that the effect of the H-initial context reaches up to the AP-penultimate (fourth) syllable, where the intonational L tone falls, and the effect dramatically decreases after that. The models estimate that the syllables in the H-initial APs are produced with 2.8 St, 1.32 St, 1.09 St, and 0.61 St higher pitch values than those in the L-initial APs in AP-initial, second, third, and fourth syllables, respectively. The model still estimates that the final syllable of H-initial APs is 0.25 St higher than that of the L-initial ones, but this effect is not found significant ($p = 0.166$).

The two-way interactions with Context (Context x Gender and Context x YOB) are found to be significant in AP-initial, second, and third positions, similar to the results of the previous sections. That is, the effect of the AP-initial tonal contexts is modulated by Gender or YOB. The models show that female speakers produce a larger pitch difference than male speakers in those positions ($p < 0.001$ for AP-initial and second, $p = 0.001$ for AP-third). For example, female speakers show a 1.38 St

Table 4.15: The outputs of the linear mixed-effect models of 5-syllable APs by Gender, YOB, and Context. The reference category for Gender is males, and that of Context is L-initial. YOB is centered at 1961.

	AP-initial				**AP-second**			
	Est.	SE	t	Pr (>|t|)	Est.	SE	t	Pr (>|t|)
(Intercept)	3.48	0.15	23.73	< 0.001***	4.64	0.16	28.96	< 0.001***
Gender	0.25	0.20	1.21	0.229	0.25	0.22	1.14	0.258
YOB	-0.01	0.01	-0.98	0.330	0.02	0.01	1.75	0.080
Context	2.80	0.14	19.69	< 0.001***	1.32	0.16	8.19	< 0.001***
YOB*Gender	0.00	0.01	-0.26	0.799	0.03	0.02	1.67	0.098
Gender*Context	1.38	0.20	6.99	< 0.001***	1.27	0.22	5.73	< 0.001***
YOB*Context	0.06	0.01	6.80	< 0.001***	0.05	0.01	5.04	< 0.001***
YOB*Gen*Con	-0.03	0.01	-1.86	0.063	-0.04	0.02	-2.79	0.005**

	AP-third				**AP-fourth**			
	Est.	SE	t	Pr (>|t|)	Est.	SE	t	Pr (>|t|)
(Intercept)	4.44	0.15	28.74	< 0.001***	3.82	0.15	25.08	< 0.001***
Gender	0.07	0.22	0.32	0.752	-0.28	0.21	-1.32	0.189
YOB	0.03	0.01	3.20	0.002**	0.02	0.01	1.86	0.066
Context	1.09	0.16	6.68	< 0.001***	0.61	0.17	3.53	< 0.001***
YOB*Gender	0.04	0.02	2.33	0.021*	0.03	0.02	2.25	0.026*
Gender*Context	0.79	0.23	3.47	0.001**	0.19	0.24	0.82	0.413
YOB*Context	0.04	0.01	3.43	< 0.001***	0.02	0.01	1.59	0.111
YOB*Gen*Con	-0.01	0.02	-0.78	0.439	-0.01	0.02	-0.44	0.658

	AP-final			
	Est.	SE	t	Pr (>|t|)
(Intercept)	3.88	0.16	24.90	< 0.001***
Gender	-0.30	0.22	-1.38	0.169
YOB	0.03	0.01	2.53	0.013*
Context	0.25	0.16	1.49	0.136
YOB*Gender	0.03	0.02	1.76	0.080
Gender*Context	-0.09	0.23	-0.39	0.695
YOB*Context	-0.01	0.01	-0.58	0.559
YOB*Gen*Con	0.00	0.02	0.17	0.865

larger pitch difference than males in the AP-initial position, and they also produce a 0.79 St larger pitch difference than males in the AP-third position.

Also, the significant interactions of YOB and Context in the first three syllable positions ($p < 0.001$ in AP-initial, second, and third positions) suggest that younger speakers produce a larger pitch difference than older speakers in those positions. These results confirm the observation that younger speakers raise the pitch values of syllables in the H context, but this effect decreases after the intonational L tone. The three-way interaction of YOB, Gender, and Context is significant for AP-second, but insignificant for the other positions.

4.3 All APs in the corpus data

In this section, I look at how distinct segmental and laryngeal categories of AP-initial onset consonants have affected the AP tonal pattern over time. Although examining sentence-initial APs provides the trajectory of the tonogenetic change with a controlled data set, it is impossible to look at various AP-initial obstruent categories with sentence-initial APs, as there are only 59 sentence-initial APs. For example, in Section 4.2.3, only 4-syllable APs are examined, and APs starting with an aspirated or tense affricate or those starting with a tense fricative are not included in the result.

For this reason, in this section I analyze all APs starting with obstruent onsets, regardless of their positions in sentences. Because there are not many APs longer than 5 syllables, the analyses below include 2-syllable to 5-syllable APs only: 286

APs in total (Table 4.16). This section includes comparisons of the laryngeal categories (Section 4.3.1) and the manners of articulation of AP-initial obstruents (Section 4.3.2).

4.3.1 Laryngeal categories

In this section, I examine how the AP intonational melody has changed over time depending on the laryngeal categories of AP-initial consonants (aspirated, tense, and lenis). More specifically, this section aims to investigate the difference between AP-initial H-inducing laryngeal categories (aspirated vs. tense). It is clear that APs starting with a lenis obstruent are produced with a lower pitch than those starting with aspirated and tense, but the difference between aspirated and tense has not been conclusive. Previous studies find that the pitch value of the first syllable in a tense-initial AP is generally lower than that of an aspirated-initial one (Kim 1994, Choi 2002, Lee and Jongman 2012), but it remains to be seen if this pitch difference is maintained in later syllables of the same AP. The results of Chapter 4.2.1 demonstrate that there is a small pitch difference between aspirated- and tense-initial APs in the first syllable, yet this difference is not shown in later syllables. However, since only 4-syllable APs starting with stop consonants are examined in Chapter 4.2.1, APs starting with affricates and fricatives, which also show the laryngeal contrast, need to be further investigated.

Table 4.16 shows the breakdown of all APs that start with obstruents by their laryngeal categories and manners of articulation. The corpus data contains a total of 286 obstruent-initial APs. However, it should be noted that there are not many tense-initial APs in the data set, compared to aspirated- or lenis-initial ones.[15] Also, the data has less 5-syllable APs than 2-, 3-, or 4-syllable ones for all laryngeal categories.

Table 4.16: Number of all APs that start with an obstruent by their AP size and the laryngeal category of AP-initial consonants. (Total: 286 APs)

	Aspirated	Tense	Lenis	Total
2-syllable	23	10	27	60
3-syllable	44	10	51	105
4-syllable	33	14	44	91
5-syllable	13	5	12	30

Figure 4.11 shows the mean pitch values of the obstruent-initial APs by speakers' YOB and the laryngeal categories. In this figure, it is clear that tense-initial and aspirated-initial APs are produced within a higher pitch range than lenis-initial ones in younger speakers' speech. Similar to the results of sentence-initial APs, speakers born in the 1930s do not show a large pitch difference in the AP-initial position, and their pitch values of lenis-initial APs overlap with or come close to those of aspirated- and tense-initial APs in the AP-final position. From speakers born after 1940, the pitch difference between the L-initial and H-initial contexts increases not only in the AP-initial position, but also in the AP-second syllable.

Speakers born after 1960 produce large pitch differences in non-initial positions,

[15] Note that this is expected because there are not many tense-initial words in Korean in general.

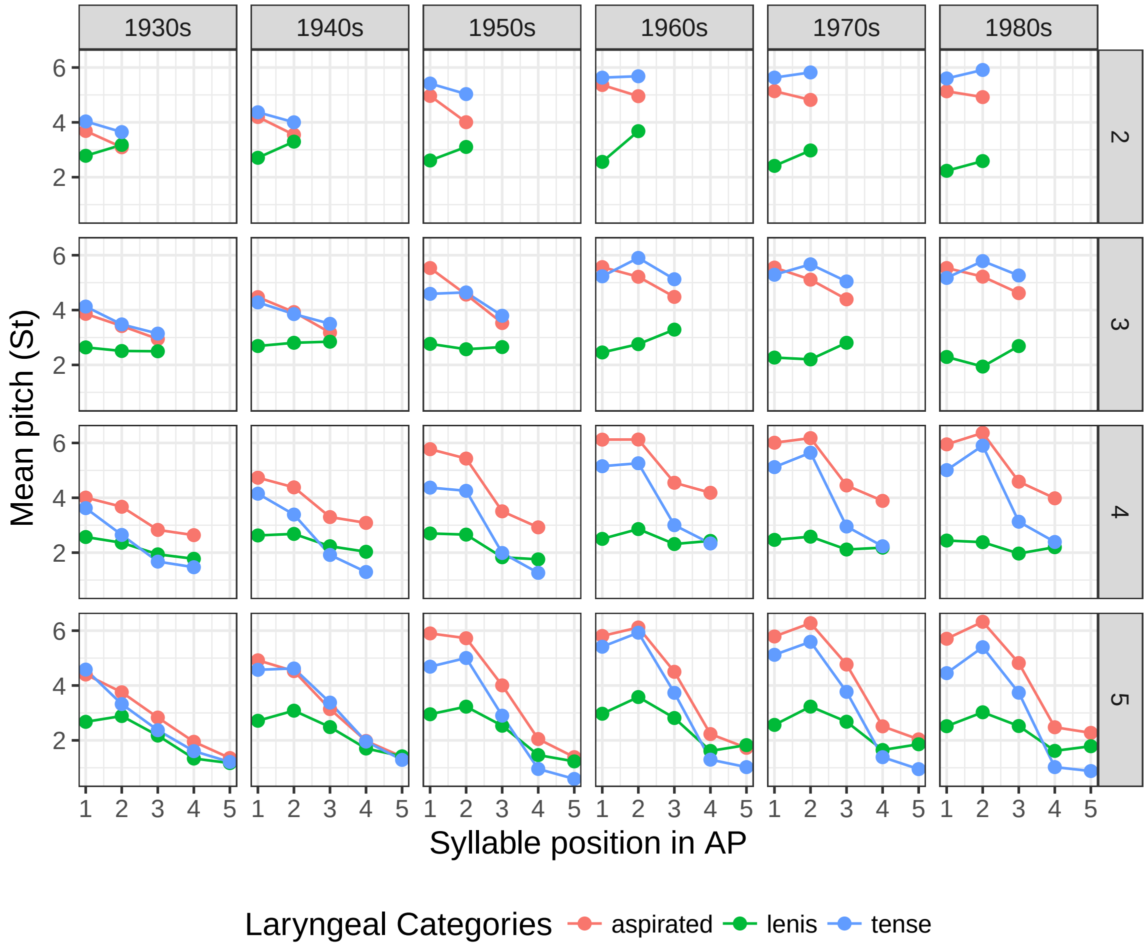


Figure 4.11: Mean pitch value (St) of each syllable in all obstruent-initial APs by AP size, laryngeal categories of AP-initial consonants, and speakers' YOB. Each row shows each AP size, from 2 syllables (the top panels) to 5 syllables (the bottom panels).

but the pitch values vary depending on the size of AP and the laryngeal categories of the AP-initial onsets. For example, the pitch patterns of H-initial and L-initial APs diverge and do not overlap from each other in 2-syllable and 3-syllable APs, similar to the results of Section 4.2 (sentence-initial APs). Younger speakers also show diverging pitch patterns for 4-syllable and 5-syllable APs by the laryngeal category of the AP-initial onsets. However, it is noticeable that tense-initial APs are lower than aspirated-initial ones in both sizes.[16] For example, 4-syllable APs starting with aspirated consonants are produced in a much higher pitch range than those staring with lenis in younger speakers' speech, but the pitch values of tense-initial APs steeply decrease after the second syllable, showing about the same pitch value with lenis-initial ones in the AP-final position. On the other hand, 2-syllable and 3-syllable APs do not show such a pattern in this figure. In 2-syllable APs, tense-initial ones are slightly higher than aspirated-initial ones, and in 3-syllable APs, both categories are produced with similar pitch values. Moreover, considering the results of sentence-initial APs in Section 4.2.1, where a tense-initial AP is produced with similar pitch values to an aspirated-initial one in most syllable positions, except the first syllable, the pattern of 4-syllable and 5-syllable APs found in Figure 4.11 is somewhat unexpected.

For statistical analyses, a linear mixed-effects regression model was built for each syllable position. In each model, the pitch values in St are the dependent variable, and the three laryngeal categories and speakers' YOB are considered as fixed-effect predictors. Aspirated is chosen as the reference of the laryngeal categories to see the statistical significance between Aspirated and Tense, and speakers' YOB is centered at 1961. Speakers and APs are treated as random effects.[17] Syllable positions are defined with the same method used in Chapter 3: Initial–Final for 2-syllable APs,

[16]The pitch values of 5-syllable APs starting with a H-inducing consonant decrease after the second syllable in this figure, unlike the result of 5-syllable APs in Section 4.2.4.4, and this seems to be because of sentence-final APs that are greatly affected by final lowering. See Figure 4.13.

[17]I add APs as a random effect, since their position within a sentence varies unlike sentence-initial APs in the previous section. One may think that AP positions within a sentence can be added as a random effect in the models. However, their positions within a sentence cannot be an appropriate factor, because they do not have a uniform meaning. For example, the second AP in a sentence that consists of three APs are not the same with the second AP in a sentence that has nine APs. For this reason, each AP is added as a random effect, not its position within a sentence.

Initial–Penultimate–Final for 3-syllable APs, Initial–Second–Penultimate–Final for 4-syllable ones, and Initial–Second–(Third)–Penultimate–Final for 5-syllable ones. The results of a mixed-effect model for the AP-initial position, for example, include all initial syllables of 2- to 5-syllable APs in the corpus data. The result of the AP-third position is not reported here because it only examines the pitch difference of the third position of 5-syllable APs. Table 4.17 summarizes the outputs of the results.

Table 4.17: The outputs of the linear mixed-effects models of all APs starting with obstruents in the corpus data. The reference category for the laryngeal category is Aspirated, and YOB is centered at 1961. 'Est.' stands for estimated coefficients, 'SE' for standard error, and 't' for t-values.

	AP-initial				**AP-second**			
	Est.	SE	t	$Pr(>\|t\|)$	Est.	SE	t	$Pr(>\|t\|)$
(Intercept)	5.34	0.13	39.65	< 0.001***	5.47	0.20	27.79	< 0.001***
Lenis	-2.75	0.14	-19.05	< 0.001***	-2.73	0.23	-11.47	< 0.001***
Tense	-0.40	0.22	-1.85	0.065	-0.60	-.34	-1.77	0.078
YOB	0.03	0.01	4.78	< 0.001***	0.05	0.01	8.46	< 0.001***
YOB*Lenis	-0.04	0.00	-30.46	< 0.001***	-0.05	0.00	-29.48	< 0.001***
YOB*Tense	-0.00	0.00	-1.72	0.085	0.01	0.00	3.29	0.001**
	AP-penultimate				**AP-final**			
	Est.	SE	t	$Pr(>\|t\|)$	Est.	SE	t	$Pr(>\|t\|)$
(Intercept)	3.30	0.24	13.52	< 0.001***	3.63	0.19	19.01	< 0.001***
Lenis	-1.38	0.32	-4.32	< 0.001***	-1.19	0.24	-4.99	< 0.001***
Tense	-0.86	-.45	-1.90	0.059	-0.35	0.35	-1.02	0.309
YOB	0.03	0.00	7.77	< 0.001***	0.03	0.00	7.07	< 0.001***
YOB*Lenis	-0.03	0.00	-16.60	< 0.001***	-0.03	0.00	-24.34	< 0.001***
YOB*Tense	-0.01	0.00	-3.49	< 0.001***	0.00	0.00	-0.15	0.879

The results illustrate that the difference between Aspirated and Tense is not significant in the AP-initial and -final positions, and their interactions with YOB are also not significant in these positions. The results of AP-second and -penultimate syllables reveal that Tense is not significantly different from Aspirated in general (p = 0.078 for AP-second, p = 0.059 for AP-penultimate), but the models find that their interaction with YOB is significant (p = 0.001 for AP-second, $p < 0.001$ for AP-penultimate). An interesting point about this result is that the direction of change is the opposite in the two positions. The model for the AP-second position estimates that the difference between Aspirated and Tense decreases by 0.01 St/YOB, whereas the model for the AP-penultimate position estimates that the difference increases 0.01 St per year.[18] For example, the model for the AP-second position estimates that the pitch difference between Aspirated and Tense for speakers born in 1932 is -3.02 St (= -2.73 + (0.01 x (1932–1961)), meaning that Tense is 3.02 St lower than Aspirated. The Aspirated-Tense difference for speakers born in 1984 is 2.5 St (= -2.73 + (0.01 x (1984 – 1961)), suggesting the difference between Tense and Aspirated decreases for younger speakers in the AP-second position. On the other hand, the model for the AP-penultimate estimates that the difference between Tense and Aspirated is -0.57 St (= -0.86 + (-0.01 x (1932 – 1961)) for speakers born in 1932, but this difference increases to -1.09 St (= -0.86 + (-0.01 x (1984 – 1961)) for speakers born in 1984.

This mixed result of Tense and Aspirated can be interpreted in several ways. It might be the case that the laryngeal contrast develops into a three-way tonal contrast

[18]Note that the estimated coefficient of Tense*YOB has a negative value in the AP-second but a positive value in the AP-final. This is because Aspirated is the reference category, not Tense.

(Aspirated > Tense > Lenis), even though the Aspirated-Tense difference is somehow reversed in 2-syllable APs or neutralized in 3-syllable APs under this interpretation. Alternatively, it might be because 4- and 5-syllable APs with tense onsets happen to be located in latter parts of sentences in comparison to those with aspirated onsets and produced with a lower pitch due to the declination effect. For example, in a sentence with 5 APs, if an aspirated-initial AP is the second one and a tense-initial APs is the fourth one, we would expect the tense-initial AP to be generally produced with a lower pitch than the aspirated initial one. Otherwise, it might be the case that other contextual differences, such as an uneven distribution of old/new information and phrase frequencies in the short stories used in the corpus, affect the pitch difference between Tense and Aspirated. Since the reading materials of the corpus are short stories, it is plausible for these factors to contribute to the pitch difference between aspirated- and tense-initial APs.

One possible way to examine if the pitch difference between Aspirated and Tense indicates a three-way tonal contrast is to plot APs by Sentence Position (sentence-initial, -medial, and -final) and see if they show a consistent pattern. If we find a similar pattern regardless of APs' position in the target sentences (Aspirated > Tense), it may support the possibility of the three-way tonal split (Aspirated > Tense > Lenis) as mentioned in Kang 2014. If we do not find a consistent pattern—for example, in some cases Tense is higher than Aspirated, and in other cases Aspirated is higher than Tense—it may support the possibility that tense-initial APs generally appear in latter positions of the target sentences than aspirated-initial APs or that other contextual factors, such as old/new information or phrase frequencies in the stories, affect the pitch patterns. Figures 4.12 and 4.13 provide the pitch patterns of 4-syllable and 5-syllable APs in the data by speakers' YOB, the laryngeal categories, and APs' position within the target sentences.

Figure 4.12 shows that both aspirated- and tense-initial APs have developed the phrase-initial H pitch concurrently and that they are produced with almost identical pitch patterns in the sentence-initial position (the leftmost column). This result is the same with the one found in Chapter 4.2.1. However, the AP pitch patterns are inconsistent in sentence-medial and -final positions: tense-initial APs are lower than aspirated-initial ones sentence-medially, whereas they are higher than aspirated-initial ones sentence-finally. A closer look at the sentence-medial APs seems to suggest that the pitch difference is not likely due to a three-way tonal split. For example, speakers born in the 1930s also show about 1.5 St of pitch difference between aspirated- and tense-initial APs. The pitch difference between aspirated and tense remains about the same (about 1.5 St) for the youngest generation, whereas the pitch difference between H-pitch inducing ones (aspirated, tense) and lenis increases.

We see a similar pattern in Figure 4.13. 5-syllable APs starting with tense and aspirated rarely show a pitch difference sentence-finally, which suggests that they have developed the AP-initial pitch contrast concurrently.[19] The two laryngeal categories show a pitch difference sentence-medially; aspirated-initial APs are produced with a

[19]The corpus data contains five 5-syllable APs starting with a tense consonant (Table 4.16), but unfortunately, all of them are either sentence-medial or sentence-final. Thus, it is impossible to compare aspirated- and tense-initial APs in the sentence-initial position.

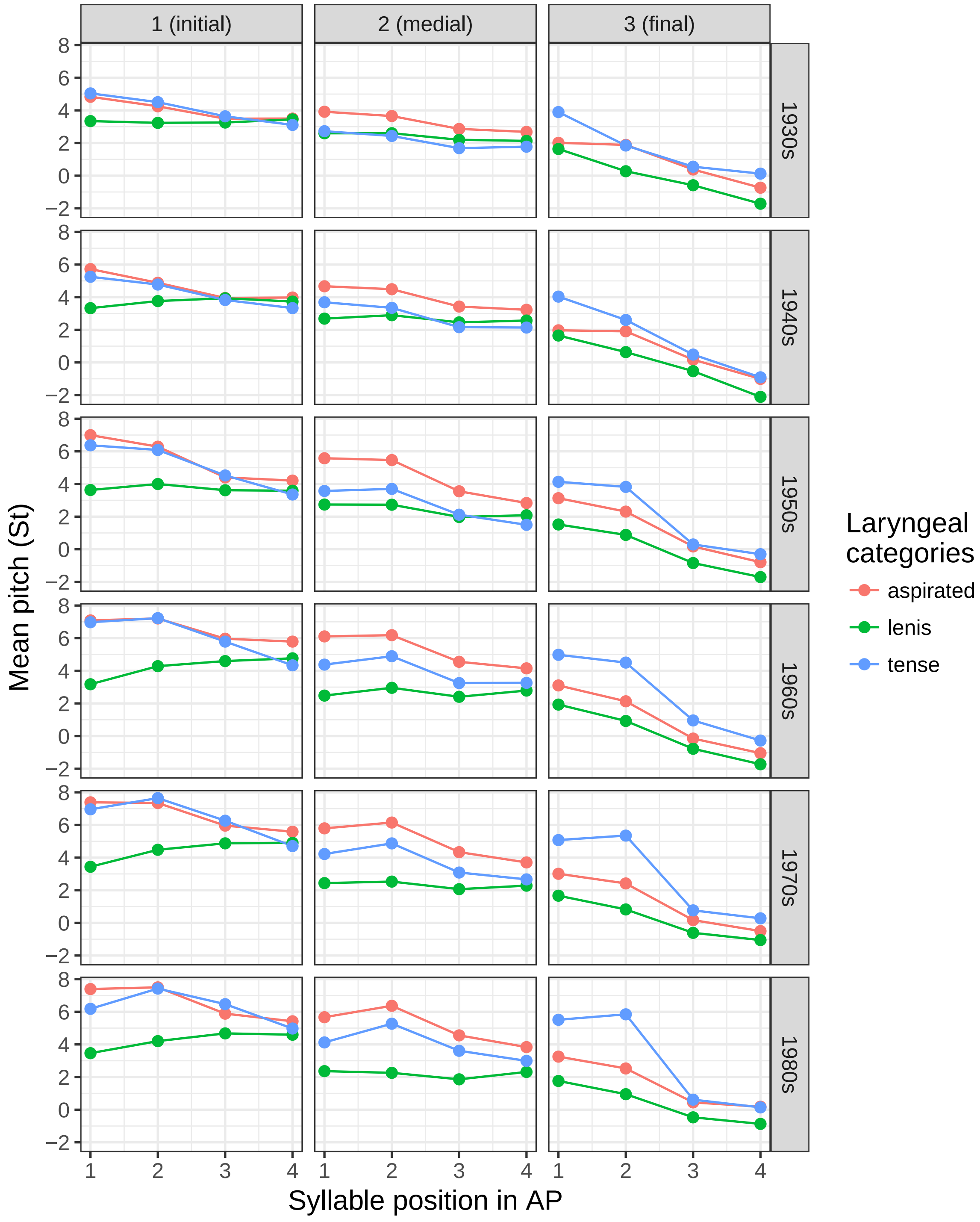


Figure 4.12: Mean pitch value (St) of each syllable in 4-syllable APs by speakers' YOB, the largyngeal categories, and positions within the target sentences.

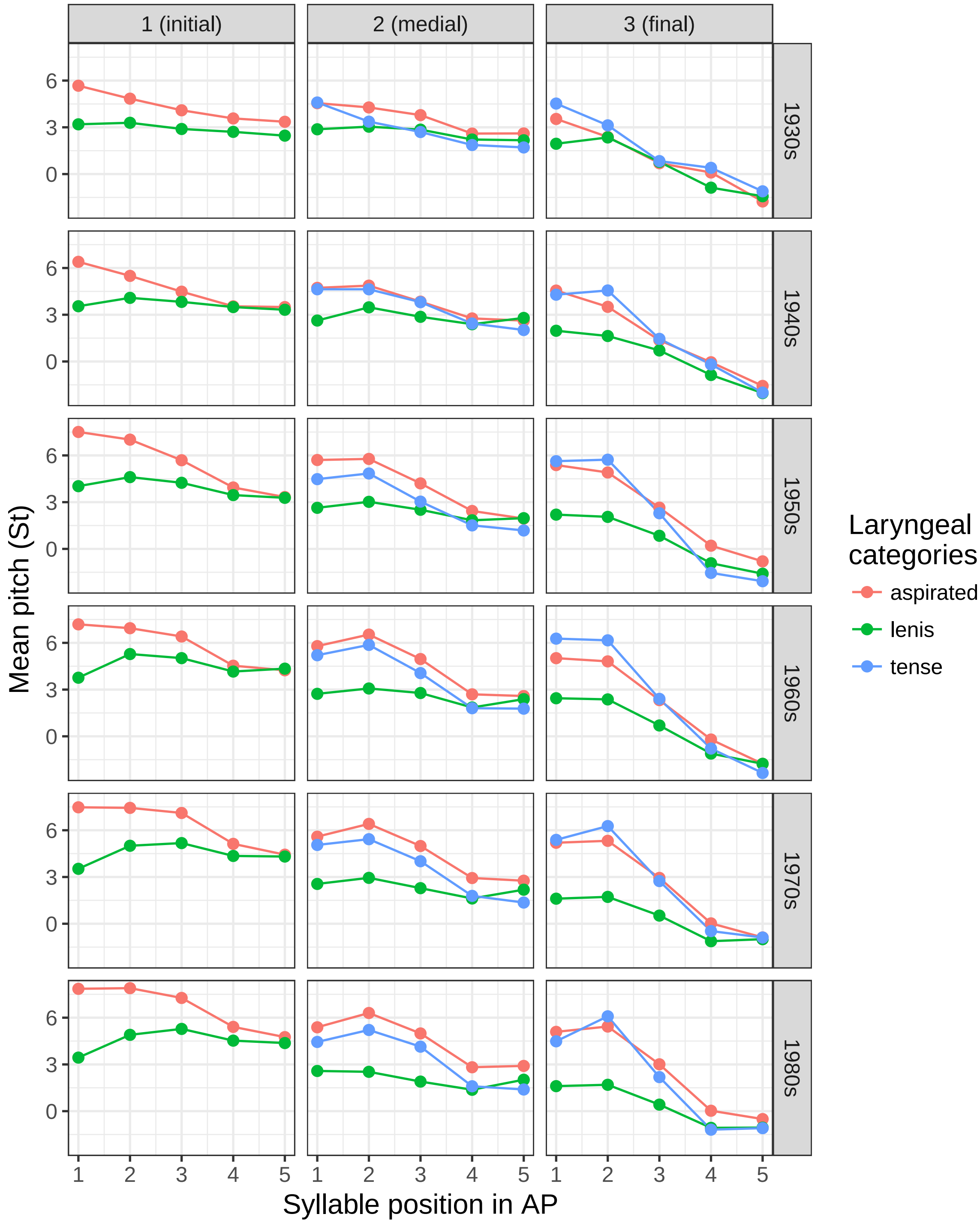


Figure 4.13: Mean pitch value (St) of each syllable in 5-syllable APs by speakers' YOB, the largyngeal categories, and positions within the target sentences.

higher pitch than tense-initial ones. However, similar to the results of 4-syllable APs, the pitch difference is quite similar across all generations, except those born in the 1940s. As the pitch patterns are not consistent both by AP size and sentence position, it is hard to conclude that the three laryngeal categories are likely to develop a three-way tonal split. Therefore, a plausible conclusion for now is that both tense-initial and aspirated-initial APs are produced with higher pitch values than lenis-initial ones, and the tonogenetic sound change is a structural one that affects all H-inducing obstruent categories.

4.3.2 Manner of articulation

This section aims to investigate if distinct manners of articulation of AP-initial consonants affect the AP-initial pitch contrast and intonational melody. Table 4.18 shows the number of all APs starting with an obstruent by AP size and the manner of articulation (the same data set used in the previous section). Although 286 APs are included in the analysis, the data lack 3-syllable and 4-syllable APs starting with a tense affricate. Also, in general, there are not many APs starting with a tense consonant, when compared to those starting with an aspirated or lenis. For example, there is only one AP starting with a tense fricative for each AP size. The intonational pattern of the categories with one target AP may vary due to other linguistic factors than AP-initial onset consonants, such as the position of a target AP within a sentence and consonant-induced pitch differences in non-AP-initial positions, whereas categories with many target APs are relatively free from those factors. Since the goal of this section is to compare different manners of articulation, not the three laryngeal categories, tense-initial and aspirated-initial APs are combined as H-initial APs in this section.

Table 4.18: Number of all APs starting with an obstruent by their manner of articulation, their laryngeal categoreis and AP size.

	Stop			**Affricate**			**Fricative**	
	Aspirated	Tense	Lenis	Aspirated	Tense	Lenis	Aspirated	Tense
2-syllable	3	9	21	6	1	7	16	1
3-syllable	7	6	40	5	0	10	30	1
4-syllable	8	13	32	4	0	12	21	1
5-syllable	2	2	8	3	2	6	8	1

Figure 4.14 shows mean pitch patterns of all APs by manners of articulation of the AP-initial onset, AP size, and speakers' YOB. The figure clearly illustrates that APs starting with affricates and fricatives pattern together with those starting with stops in that AP-initial aspirated and tense categories (H-inducing) induce higher pitch values than AP-initial lenis (L-inducing). Similar to the results of sentence-initial APs, the pitch difference is larger for younger speakers than older speakers. For example, speakers born in the 1930s and 40s show a small pitch difference in the first syllable, and their pitch difference decreases in the AP-final position, regardless of AP size. However, speakers born in the 1970s and 80s produce a large pitch difference in the AP-initial position, and this pitch difference is mostly maintained in the AP-final position for shorter APs and before the intonational L tone in the penultimate

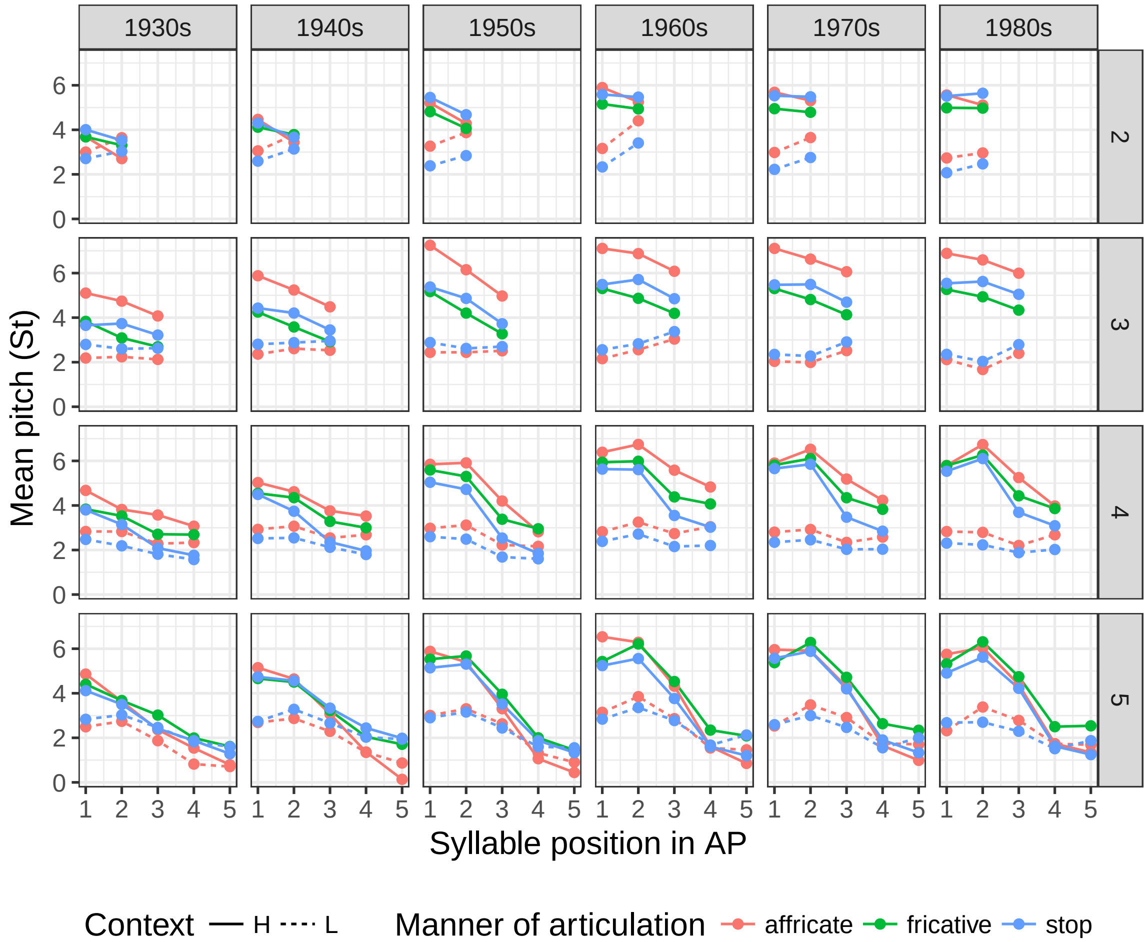


Figure 4.14: Mean pitch value (St) of each syllable in all APs starting with an obstruent by AP size, manner of articulation, and speakers' YOB. The top row shows 2-syllable APs, and the bottom row is for 5-syllable APs.

syllable for longer APs.

This result suggests that the effect of AP-initial onset consonants observed in younger speakers' speech in Section 4.2 is not limited to sentence-initial APs. For example, the pitch patterns of 2-syllable and 3-syllable APs in this figure are remarkably similar to those in Figures 4.7 and 4.8, which only examine sentence-initial APs. This suggests that the undershoot of AP-medial tones (T**H-L**H) is found in all short

APs regardless of their position within the target sentences, and the undershoot contributes to the diverging intonational patterns (H-initial vs. L-initial) in short APs in Seoul Korean. Also, similar to the results of 4-syllable and 5-syllable APs in the sentence-initial position, the pitch values of H-initial APs dramatically decrease due to the intonational L tone in longer APs. This, again, confirms that the results in Section 4.2 are the general trajectory of the tonogenetic change in Seoul Korean, found in all APs regardless of their position within the target sentences. This result also confirms the finding of the experiment with young speakers in Chapter 3, which also finds a global effect of AP-initial onset consonants on the pitch range of an entire phrase, regardless of AP positions within a sentence.

For statistical analyses, I built a linear mixed-effects model for each syllable position of each pitch context (H or L) with pitch values as the dependent variable and the manner of articulation and speakers' YOB as fixed-effect predictors. The reference category for Manner is Stop, and speakers' YOB is centered at 1961. Syllable positions are defined in the same method used in Chapter 3 and in the previous section, where the first syllable of an AP is defined as AP-initial, the last syllable as AP-final, and the second last syllable as AP-penultimate (2-syllable APs: Initial–Final; 3-syllable APs: Initial–Penultimate–Final; 4-syllable APs: Initial–Second–Penultimate–Final; 5-syllable APs: Initial–Second–(Third)–Penultimate–Final). Speakers and APs are included as random effects. Table 4.19 shows the results of **H**-initial APs.

The results in Table 4.19 are highly consistent across all syllable positions, except AP-second. The models for AP-initial, -penultimate, and -final positions mostly

Table 4.19: The outputs of the mixed-effects models for manner comparison in APs starting with a **H**-inducing obstruent. The reference category for Manner is Stop, and speakers' YOB is centered at 1961. 'Est.' stands for estimated coefficients, 'SE' for standard error, and 't' for t-values.

	AP-initial				AP-second			
	Est.	SE	t	Pr $(>\lvert z\rvert)$	Est.	SE	t	Pr $(>\lvert z\rvert)$
(Intercept)	5.20	0.22	23.67	< 0.001***	4.91	0.27	18.06	< 0.001***
Affricate	0.80	0.34	2.33	0.021*	0.93	0.46	2.01	0.046*
Fricative	-0.15	0.24	-0.62	0.538	-0.05	0.32	-0.15	0.880
YOB	0.03	0.01	3.64	< 0.001***	0.05	0.01	6.95	< 0.001***
Affricate*YOB	-0.00	0.00	-1.29	0.198	-0.01	0.00	-2.89	0.004**
Fricative*YOB	-0.00	0.00	-1.44	0.149	-0.01	0.00	-3.87	< 0.001***
	AP-penultimate				**AP-final**			
	Est.	SE	t	Pr $(>\lvert z\rvert)$	Est.	SE	t	Pr $(>\lvert z\rvert)$
(Intercept)	2.76	0.37	7.39	< 0.001***	3.41	0.31	10.89	< 0.001***
Affricate	0.13	0.70	0.19	0.849	0.35	0.55	0.64	0.523
Fricative	0.47	0.48	0.98	0.330	0.14	0.39	0.367	0.714
YOB	0.03	0.01	5.81	< 0.001***	0.04	0.01	6.29	< 0.001***
Affricate*YOB	-0.01	0.00	-1.68	0.093	0.00	0.00	1.07	0.285
Fricative*YOB	0.00	0.00	0.38	0.702	-0.00	0.00	-1.56	0.118

show that Affricate and Fricative are not significantly different from Stop (the reference category), nor their interactions with YOB.[20] The model for the AP-initial position estimates that Affricate is 0.8 St higher than Stop (p = 0.021), yet this difference neither increases nor decreases over time (p = 0.198), suggesting that the difference between Affricate and Stop is not due to the tonogenetic sound change. The insignificant results indicate that all AP-initial H-inducing obstruents have undergone the change concurrently, regardless of different manners of articulation.

[20]The highly significant effect of YOB in all models reflects that aspirated and tense stops (H-pitch inducing ones) are produced with higher pitch values in younger speakers than in older speakers.

The only case where the interaction of manner with YOB is significant is the AP-second position of 4-syllable and 5-syllable APs. However, both interactions with YOB indicate that the pitch difference rather decreases by 0.01 St/YOB ($p = 0.004$ for Affricate x YOB and $p < 0.001$ for Fricative x YOB). For example, the Stop-Affricate difference is 0.52 St smaller for speakers born in 1984 than speakers born in 1932 (-0.52 = -0.01 x (1984 – 1932)). This result also supports the finding that distinct manners of articulation are not a significant factor with regard to this tonogenetic change.

Table 4.20 shows the outputs of the mixed-effects models for L-initial APs. Here, again, the pitch values are added as a dependent variable, and Manner and YOB are included as independent variables. Speakers and APs are treated as random effects. The reference category of Manner is Stop, and speakers' YOB is centered at 1961. Since the three fricatives are defined as H-inducing ones (/s, h/: aspirated fricatives, /s'/: a tense fricative), the results of L-initial APs do not report the comparison of Stop and Fricative.

Similar to H-initial APs, the results of L-initial APs are highly consistent. Affricate is not significantly different from Stop in all syllable positions, and their interactions with speakers' YOB are also insignificant in most cases, except the AP-initial. In the AP-initial position, the Stop-Affricate difference increases 0.003 St/YOB ($p = 0.013$), estimating that younger speakers' Stop-Affricate difference is 0.156 St (= 0.003 x (1984 – 1932)) larger than that of older generations. However, the estimated increase is too small to presume that the lenis stops and affricates are indeed different

Table 4.20: The outputs of the mixed-effects models for manner comparison in APs starting with a **L**-inducing obstruent. The reference category for Manner is Stop, and speakers' YOB is centered at 1961. 'Est.' stands for estimated coefficients, 'SE' for standard error, and 't' for t-values.

	AP-initial				**AP-second**			
	Est.	SE	t	Pr ($>\|z\|$)	Est.	SE	t	Pr ($>\|z\|$)
(Intercept)	2.56	0.10	26.10	$<$ 0.001***	2.62	0.14	18.32	$<$ 0.001***
Affricate	0.15	0.16	0.97	0.334	0.14	0.27	0.53	0.596
YOB	-0.01	0.00	-2.93	0.004**	-0.02	0.00	-2.95	0.004**
Affricate*YOB	0.003	0.00	2.48	0.013*	0.00	0.00	1.68	0.094
	AP-penultimate				**AP-final**			
	Est.	SE	t	Pr ($>\|z\|$)	Est.	SE	t	Pr ($>\|z\|$)
(Intercept)	1.82	0.24	7.62	$<$ 0.001***	2.39	0.16	14.59	$<$ 0.001***
Affricate	0.31	0.45	0.70	0.487	0.17	0.30	0.57	0.568
YOB	-0.00	0.00	-0.34	0.736	0.00	0.00	9.29	0.774
Affricate*YOB	0.00	0.00	1.26	0.207	0.00	0.00	1.23	0.218

from each other. Considering the results of other positions, it is rather reasonable to conclude that Manner of articulation is not an important factor among L-inducing consonants and the tonogenetic sound change has affected all obstruent series.

4.4 Summary of the findings

Figure 4.15 shows mean pitch patterns of all APs that are 2 to 5 syllables long in the corpus data. The findings of Chapter 4 are summarized as:

- The pitch difference of the first syllable between H-initial and L-initial APs is only about 1 St for speakers born in the 1930s, but it increases to 3 St for younger generations.
- Female speakers born after 1950 show a large pitch difference on the second

syllable depending on the AP-initial onset consonant.

- The pitch difference between H-initial and L-initial APs increases in non-initial syllables for speakers born after 1960.
  - The shorter an AP is, the larger the pitch difference in non-initial positions is observed.
  - 4-syllable APs also show a pitch difference in non-initial positions by the AP-initial onset consonants, but the difference decreases from the intonational L tone.
  - 5-syllable APs do not show a pitch difference in the penultimate and final syllables due to the effect of the intonational L tone and final lowering. Yet, the pitch difference in the third syllable increases over time.
- The change has affected all consonants regardless of their manner of articulation. AP-initial stops, affricates, and fricatives have developed a pitch contrast concurrently.
- The change is a structural sound change, affecting all aspirated and tense consonants.
- Female speakers have led the change, which is a common pattern as for a sound change.[21]
- The AP-initial consonant has a global effect on the pitch values of the AP.

[21] A sound change below the level of (speakers') consciousness is often led by female speakers (See, for example, Labov 1990).

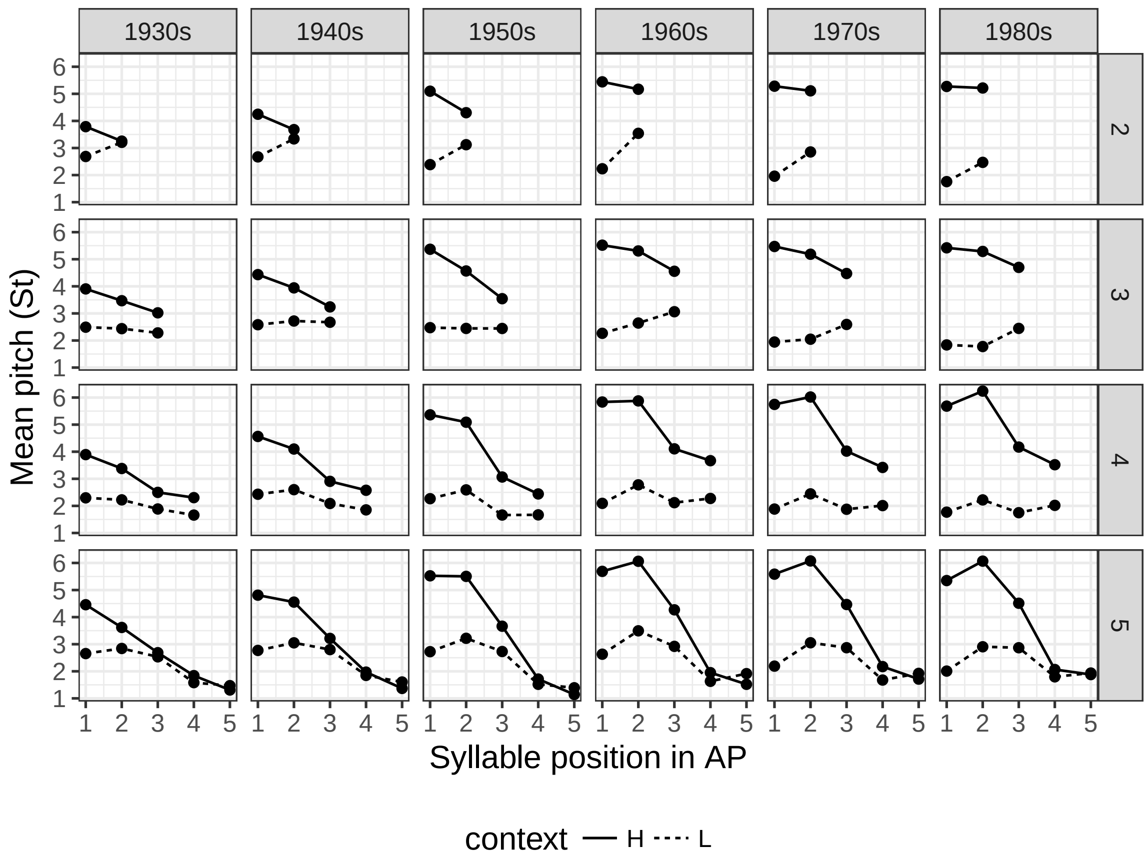


Figure 4.15: Mean AP patterns of all APs that are 2 to 5 syllables long by AP-initial pitch contexts and speakers' YOB, aggregated in 10 year bands.

Chapter 5

New corpus data

The results of Chapter 4 raise questions: what about speakers born after 1990? As the youngest cohort in the NIKL corpus is speakers born in the 1980s and most of them are in their 30s (in 2017), it is reasonable to ask what happened to speakers younger than that. For example, would they keep increasing the pitch difference between H-initial and L-initial APs? Would they show a smaller gender difference than younger speakers (born after 1960) in the NIKL corpus? Would they produce a smaller pitch difference between aspirated- and tense-initial APs than younger speakers in the NIKL corpus? This chapter aims to answer these questions by building a new speech corpus with Korean speakers who were born in the 1990s.

5.1 Speakers, materials, and procedures

For a new speech corpus, 15 speakers of Seoul Korean (8 males and 7 females), who were born in the 1990s, were employed. They have identified themselves as native speakers of Seoul Korean, residing in Seoul or the Seoul metropolitan area in most of their lives. Their year of birth (YOB) ranges from 1992 to 1996, with 1994 as the median YOB. All speakers reported that they had no hearing or speaking disorder. The information of the speakers analyzed in this chapter is provided in Table 5.1.

Table 5.1: YOB and gender of the 15 speakers in the new speech corpus.

Females		Males	
Speaker	YOB	Speaker	YOB
fu01	1994	mu01	1996
fu02	1992	mu02	1995
fu03	1995	mu03	1995
fu04	1996	mu04	1996
fu05	1996	mu05	1992
fu06	1996	mu06	1992
fu07	1996	mu07	1994
		mu08	1994

Recordings were made in a quiet room at a university in South Korea, and each story that was read by one speaker was saved as a .wav file at a sampling rate of 44.1 kHz. A head-mounted microphone with USB connection (Logitech H390) was used, and the recordings were directly digitized into a laptop computer. The same

11 short stories that all speakers in the NIKL corpus read were employed as reading materials for the new corpus. The recording took about one hour per speaker, and the participants received 10,000 won (about 10 USD) for their participation.

After the recordings were made, sentences in the recorded stories were saved as separate .wav files, following the file naming convention of the NIKL corpus.[1] Most speakers read all 11 short stories (404 sentences per speaker), with exceptions of fu02, who did not read one short story (story ID: t17) and fu04, who did not read another story (story ID: t12). The total number of sentences in the corpus is 6,013 (= (404 sentences x 15 speakers) - 47 sentences (for the two female speakers who did not read one short story each)), and the total size of the new corpus is about 2.7 Gigabytes.

The same 59 sentences analyzed in Chapter 4 were selected for this chapter.[2] Since Chapter 4 employed 5 sentences from the two short stories that the two female speakers did not read (1 sentence for t17 and 4 sentences from t12), the following analysis omits those 5 sentences.[3] The total number of sentences analyzed in this chapter is, therefore, 880 (= 59 sentences x 15 speakers - 5 sentences).

For f0 measurement, the same procedure in Chapter 4 was carried out in this chapter. I firstly forced-aligned the onset and offset of syllables and APs in the target sentences, using the Korean forced aligner (Yoon and Kang 2012). The boundaries

[1] The file names are a combination of 4 characters for speaker ID, followed by 3 characters for story ID and another 3 characters for sentence ID. For example, the file name of the second sentence of the first story read by a female speaker born in the 1980s in the NIKL corpus is fv01_t01_s02.wav. For speakers' ID in the new corpus, I used 'u' (ex: f**u**01, m**u**01), since 'z' was used for the oldest cohort in the NIKL corpus, followed by 'y', 'x', 'w', and 'v' for the youngest cohort (in reverse-alphabetical order).

[2] See Appendix B for the list of the sentences employed from the stories.

[3] Those missing sentences would not have much effect on the statistical results, since this chapter also employs linear mixed-effects regression, which allows a random intercept for each speaker.

of syllables and APs were manually checked and corrected in *Praat* (Boersma and Weenink 2016). After checking alignments, mean pitch values of all syllables in the target sentences were measured using a *Praat* script, with a f0 range set at 75–300 Hz for the male speakers and 100–500 Hz for the females. The obtained pitch values were manually inspected for f0 tracking errors, such as pitch doubling and pitch halving. After that, they were converted into a semitone scale, using a speaker's own baseline, which was defined as the 10th percentile of a speaker's pitch range in this dissertation.[4] Speakers' baselines were obtained with a Python script. Lastly, pitch values in Hz were converted into a St scale based on the obtained baseline values, using the following formula: $St = \log_2(f0/baseline)*12$.

5.2 Results

5.2.1 Laryngeal categories

I start with sentence-initial APs to compare the results of speakers born in the 1990s to those in the NIKL corpus in Section 4.2. Since 59 sentences are selected for analysis, there are 59 sentence-initial APs. Tables 4.2 and 4.3 are repeated here to show the number of sentence-initial APs in the target sentences by the laryngeal categories and AP size.

[4]Please note that the 10th percentile here means the 10th observation (token-wise) when the number of the observations (all pitch values of a speaker) is divided by 100, so the 10th percentile is the pitch value below which 10% of a speaker's pitch values are found. For example, when there are 700 observations (instances of pitch values) for one speaker, the 10th percentile is the 70th pitch value from the bottom.

Table 5.2: Number of the target APs by manner of articulation of AP-initial onset consonants and their laryngeal category. Cells highlighted in gray indicate H-pitch inducing categories. (Repeated from Table 4.2)

	Aspirated	Tense	Lenis	Total
Stop	5	5	12	22
Affricate	5	0	8	13
Fricative	14	0	NA	14
Sonorant or vowel	NA	NA	NA	10

Table 5.3: Number of the target APs by AP size. Cells highlighted in gray represent H-pitch inducing categories. Again, /s/ and /h/ are counted as aspirated. (Repeated from Table 4.3)

	Aspirated	Tense	Lenis	Sonorant or Vowel	Total
2-syllable	2	1	4	2	9
3-syllable	10	1	5	2	18
4-syllable	10	3	7	4	24
5-syllable	2	0	4	2	8

To see if the pitch difference among the laryngeal categories increases for speakers born in the 1990s, this section shows the results of 49 sentence-initial APs starting with an obstruent, including stops, affricates, and fricatives. Figure 5.1 illustrates the pitch patterns of sentence-initial APs produced by speakers born in the 1990s by AP size and the laryngeal category of AP-initial onset consonants.

Figure 5.1 shows that short aspirated- or tense-initial APs (2 syllables and 3 syllables long) are produced within a higher pitch range than lenis-initial ones. For example, there is a 4 St pitch difference between H-inducing and L-inducing contexts

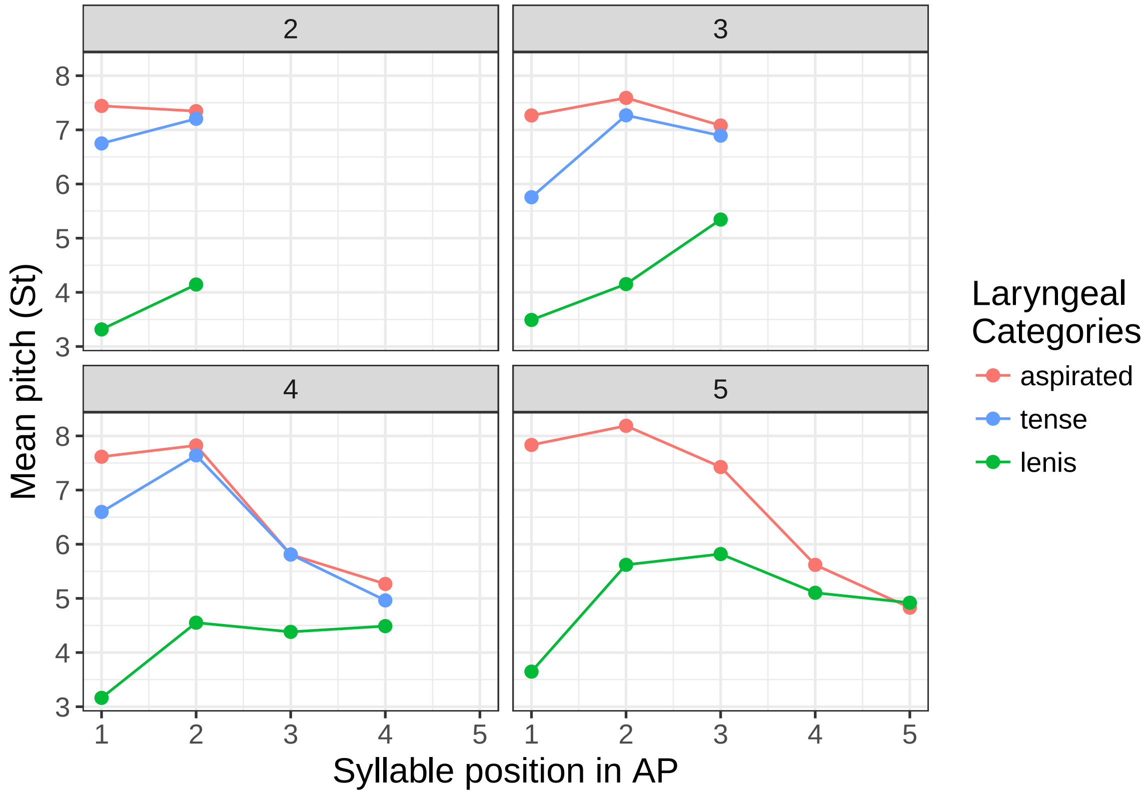


Figure 5.1: Mean pitch patterns in sentence-initial APs with an obstruent onset produced by speakers born in the 1990s. Please note that there is no 5-syllable sentence-initial AP starting with a tense consonant.

in the first syllable of 2-syllable APs. Also, the pitch value of the second syllable in 2-syllable APs is about 3 St higher for the aspirated- or tense-initial contexts than the lenis-initial ones. Similar to young speakers in the NIKL corpus, the H-L pitch difference decreases after the intonational L tone in the penultimate syllable for long APs (4-syllable and 5-syllable ones). For instance, the second syllable of 4-syllable APs in the H-inducing contexts is about 3 St higher than the one in the L-inducing context, but this difference decreases to 1.5 St in the penultimate syllable.

Since the pitch patterns of speakers born in the 1990s seem quite similar to those of young speakers in the NIKL corpus (born after 1960), it's worthwhile to compare them in one figure to see how much they are similar or different.[5] There are 62 speakers born after 1960 in the NIKL corpus, so the figure below includes a total of 77 speakers (62 speakers in the NIKL corpus + 15 speakers in the new corpus). See Table 5.4 for the number of speakers by their YOB and gender.

Table 5.4: Age and gender of the speakers from the two corpora. Speakers born in the 1960s to 80s are from the NIKL corpus, and those born in the 1990s are from the new speech corpus.

	NIKL corpus			New corpus	
	1960s	1970s	1980s	1990s	Total
Female	3	11	9	7	30
Male	7	27	5	8	47

In Figure 5.2, which shows mean pitch patterns produced by speakers born after 1960 in the NIKL corpus and the speakers in the new corpus, it is clear that the pitch patterns of speakers born in the 1990s resemble those of speakers born in the 1970s and 80s. This indicates that the change has almost reached completion. Also, it is observed that the difference between aspirated- and tense-initial APs has decreased in the speech of speakers born in the 1990s, when compared to the previous cohorts. For

[5] It should be noted that since the participants of the NIKL corpus were recorded in 2003, about 13 years before the speakers of the new corpus were recorded, there is a possibility that the speakers in the NIKL corpus have undergone linguistic change for 13 years. However, since investigating linguistic change across lifespan is beyond the scope of this chapter, I will leave this possibility for a future study and just assume that speakers in the NIKL corpus have not (greatly) changed since 2003.

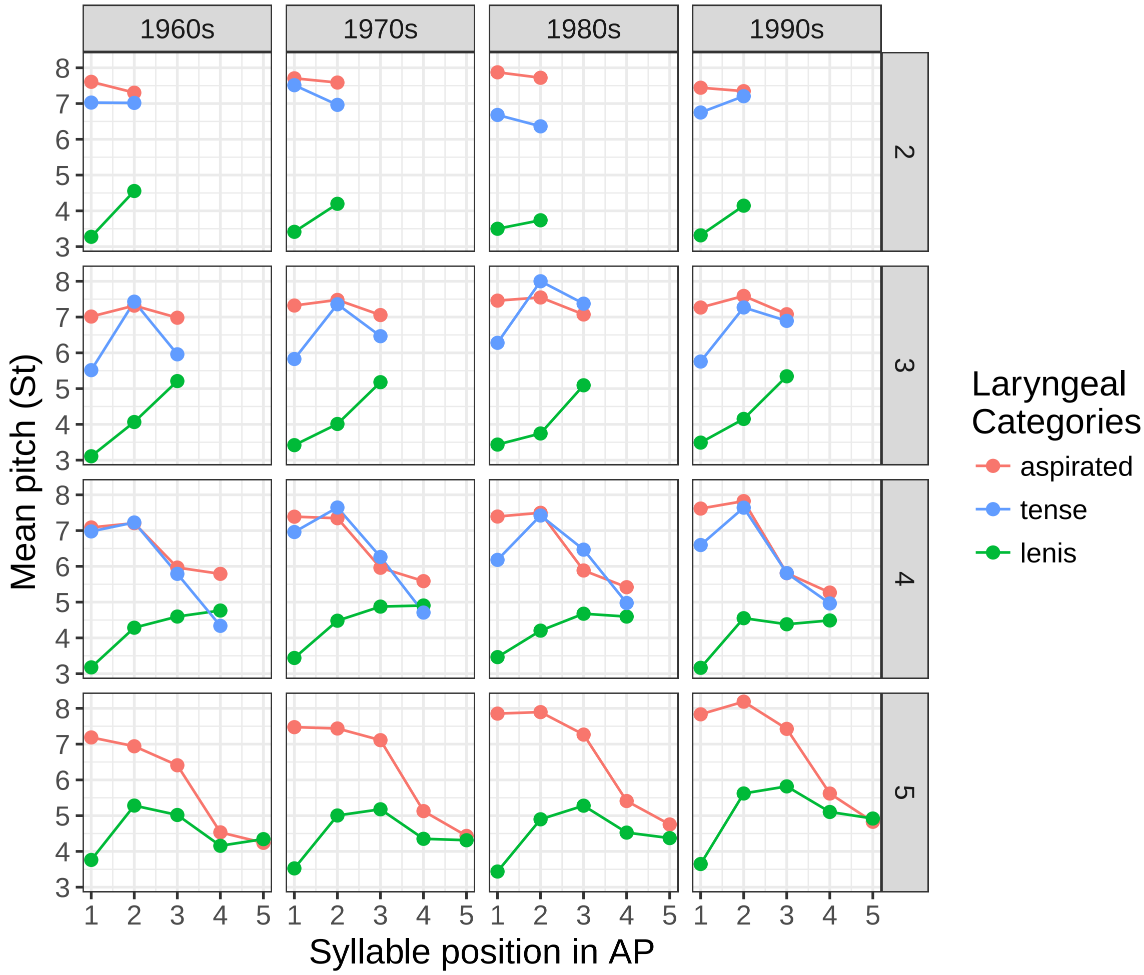


Figure 5.2: Mean pitch patterns of sentence-initial APs with an obstruent onset produced by speakers born in the 1990s in the new corpus and speakers born after 1960 in the NIKL corpus. YOB is aggregated in 10 year bands. Please note that there is no 5-syllable sentence-initial AP starting with a tense consonant.

example, speakers born in the 1990s still show a pitch difference between aspirated and tense in the AP-initial position, but not in other positions, regardless of AP size. This confirms the finding in Chapter 4 that there is a pitch difference between aspirated and tense in the AP-initial position, yet the tonogenetic change is a structural one affecting all APs with aspirated or tense onsets.

To examine the pitch difference of the three laryngeal categories produced by younger speakers in the two corpora, I build four linear mixed-effects models for each syllable position. Since there is no 5-syllable APs starting with a tense obstruent, pitch values of 5-syllable APs are not included in the models. Syllable positions are defined as the same with the previous sections (2-syllable APs: Initial–Final, 3-syllable APs: Initial–Penultimate–Final, 4-syllable APs: Initial–Second–Penultimate–Final). The reference for the laryngeal categories is Lenis, and YOB is centered at 1978, which is the median year of birth among the speakers analyzed in this section. The dependent variable in the models is pitch values (St) in the given syllable position, and the laryngeal categories and speakers' YOB are included as independent variables. Speakers are treated as a random effect. Table 5.5 summarizes the outputs of the models.

The results of the models show that Aspirated and Tense are produced with a higher pitch than Lenis in all syllable positions ($p < 0.001$ for all comparisons). For example, the model for the AP-initial position estimates that Aspirated and Tense are 3.98 St and 3.25 St higher than Lenis, respectively. Also, the model for the AP-final position shows that Aspirated and Tense are 1.57 St and 0.89 St higher than Lenis ($p < 0.001$ for both comparisons). This makes an interesting comparison to the results of Chapter 4, where most analyses find the pitch difference between Lenis and other H-inducing contexts is either insignificant or marginally significant. It confirms that younger speakers produce APs starting with an aspirated or tense consonant with a higher pitch than those starting with a lenis consonant in all syllable positions, and

Table 5.5: The outputs of the linear mixed-effects models for comparison of young speakers of the NIKL corpus and speakers in the new corpus. The reference for the laryngeal categories is Lenis, and YOB is centered at 1978, which is the median YOB for speakers included in the analysis.

	AP-initial				**AP-second**			
	Est.	SE	t-value	Pr($>\|t\|$)	Est.	SE	t-value	Pr($>\|t\|$)
(Intercept)	3.41	0.13	26.65	< 0.001***	4.47	0.14	32.13	< 0.001***
YOB	-0.01	0.01	-0.37	0.714	0.00	0.01	0.09	0.931
Aspirated	3.98	0.05	74.25	< 0.001***	3.00	0.06	54.56	< 0.001***
Tense	3.25	0.09	37.32	< 0.001***	3.06	0.09	32.62	< 0.001***
YOB*Asp.	0.01	0.01	1.21	0.228	0.01	0.01	1.81	0.070
YOB*Tense	-0.01	0.01	-1.528	0.127	-0.00	0.01	-0.05	0.956
	AP-penultimate				**AP-final**			
	Est.	SE	t-value	Pr($>\|t\|$)	Est.	SE	t-value	Pr($>\|t\|$)
(Intercept)	4.42	0.13	32.98	< 0.001***	4.69	0.13	36.96	< 0.001***
YOB	-0.01	0.01	-0.73	0.468	-0.01	0.01	-0.53	0.595
Aspirated	2.29	0.07	33.20	< 0.001***	1.57	0.06	26.58	< 0.001***
Tense	2.07	0.11	18.97	< 0.001***	0.89	0.10	9.01	< 0.001***
YOB*Asp.	0.01	0.01	1.27	0.205	0.00	0.01	0.01	0.995
YOB*Tense	0.00	0.01	0.33	0.740	0.03	0.01	2.50	0.013*

AP-initial consonants have a global effect on the pitch values of the entire AP.

Another interesting point is that most two-way interactions with YOB are not significant in the results, indicating that the change has almost reached completion. The only exception is the interaction of Tense and YOB in the final position ($p = 0.013$). The model estimates that the difference between Tense and Lenis increases by 0.03 St per year of birth. For example, the difference between Lenis and Tense in the final syllable is about 0.35 St (= 0.89 + (0.03 x (1960 – 1978)) for speakers born in 1960, but it increases to 1.43 St (= 0.89 + (0.03 x (1996 – 1978)) for speakers born in 1996. This confirms the observation in Figure 5.2 that speakers born in the 1990s show a smaller pitch difference between H-inducing contexts (Aspirated and Tense), as the pitch value of Tense comes close to Aspirated in the AP-final position and the difference between Tense and Lenis becomes larger for younger speakers.

5.2.2 Gender difference

This section investigates the gender difference of speakers born in the 1990s, and compares it to those of young speakers in the NIKL corpus. Recall that Chapter 4 shows that female speakers have led the change. Since the results of the previous section indicate that the tonogenetic change has almost reached completion, we would expect to see the gender difference has decreased in the speech of speakers born in the 1990s. Figure 5.3 illustrates the gender difference for the speakers in the new corpus data, using the same data set employed in the previous section (49 sentence-initial APs with obstruent onsets).

Figure 5.3 shows that the expectation is borne out. Male speakers (in dotted lines) born in the 1990s catch up to female speakers (in solid lines) on the tonogenetic change such that male speakers' pitch patterns greatly overlap with those of female speakers. For example, in 2-syllable APs, male speakers' pitch pattern of each laryngeal category overlaps with that of female speakers, although male speakers show a slightly larger pitch difference between the H-inducing contexts (Tense and Aspirated). The overlap of pitch patterns produced by both genders is also found in 3-syllable and 4-syllable APs. In 5-syllable APs, however, male speakers produce smaller pitch differences

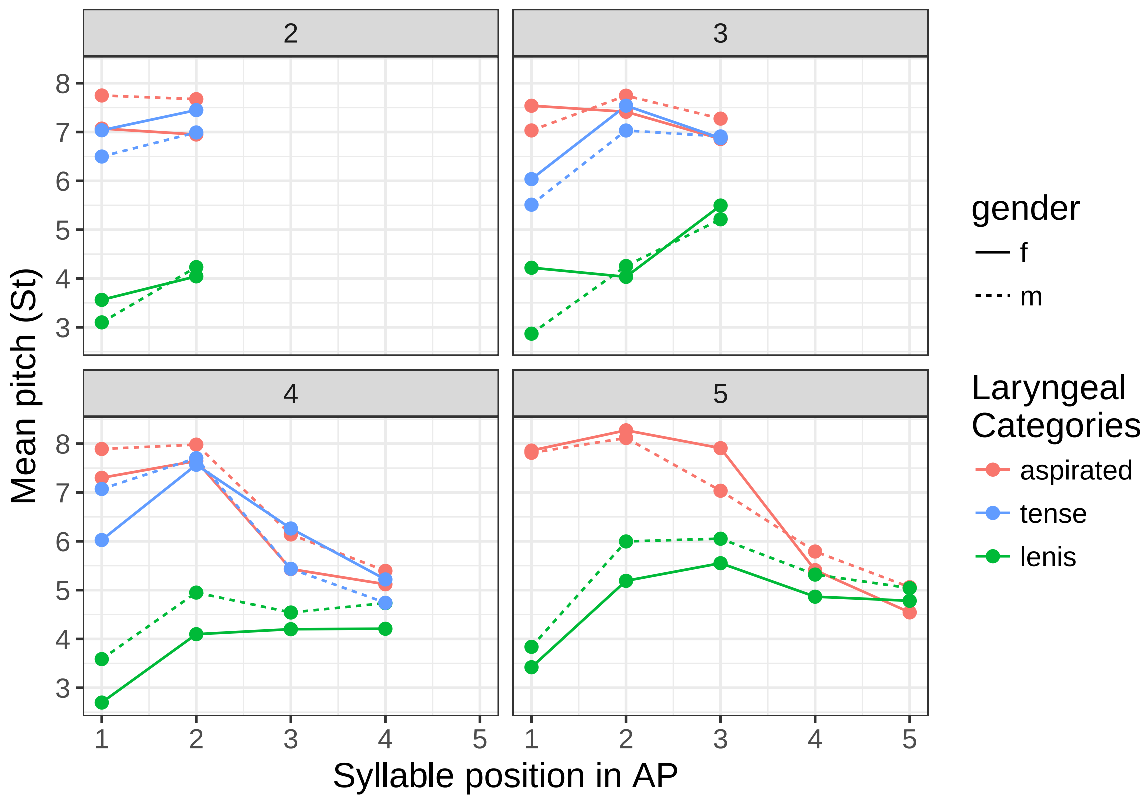


Figure 5.3: Mean pitch patterns of sentence-initial APs starting with an obstuent in the new corpus data by gender, laryngeal categories, and AP size.

betweeen Aspirated and Lenis in non-initial positions, in particular in the third and fourth syllables, when compared to their female counterparts.

Figure 5.4 compares speakers in the new corpus to speakers born after 1960 in the NIKL corpus. In this figure, all sentence-initial APs are included; aspirated- and tense-initial APs are combined as a H-inducing context, and lenis- and sonorant-initial APs are represented as a L-inducing context for the sake of simplicity.

Figure 5.4 shows that the pitch difference between the H-inducing and L-inducing contexts increases for male speakers born in the 1990s, when compared to males born in the 1960s. For instance, male speakers born in the 1960s produce a 3 St pitch

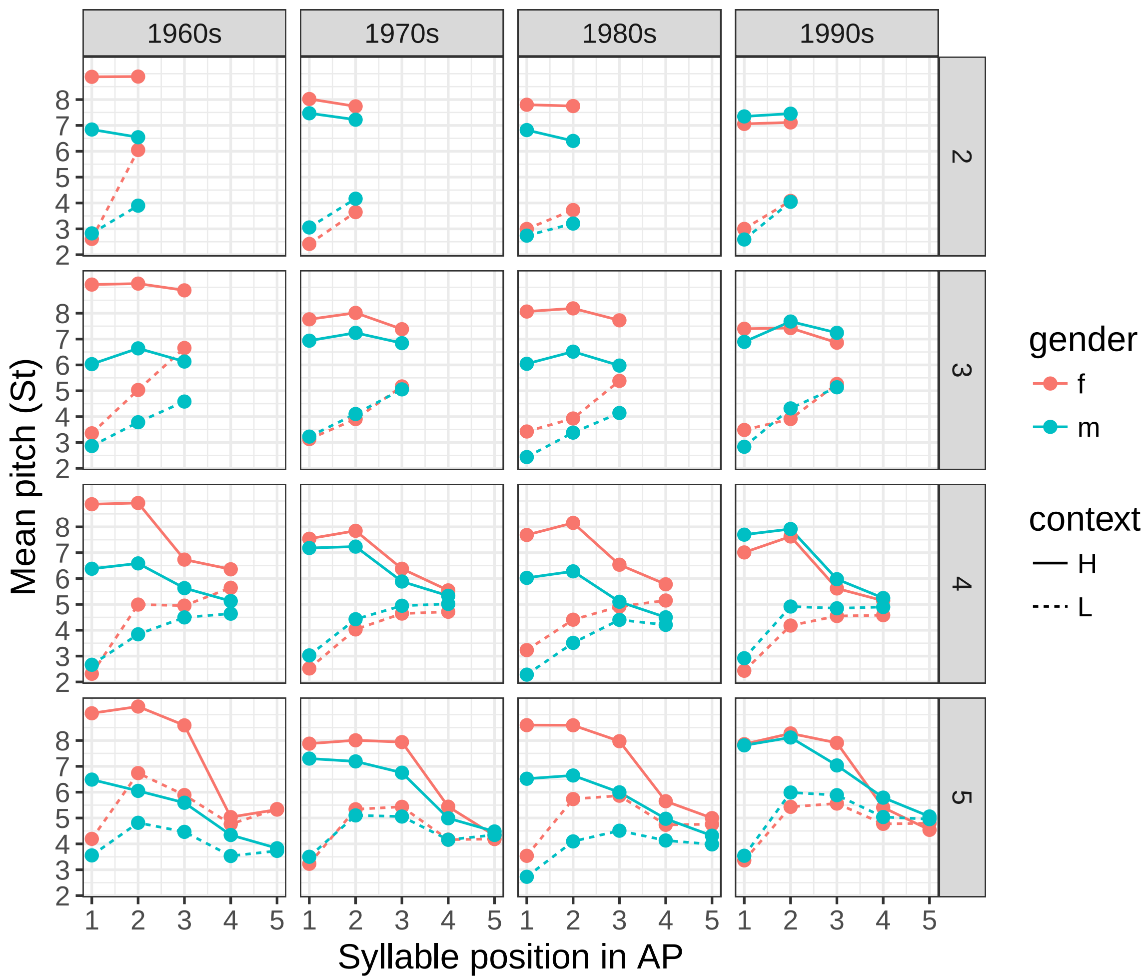


Figure 5.4: Mean pitch patterns of sentence-initial APs starting with an obstruent in the two corpora by gender, AP size, and YOB, aggregated in 10 year bands.

difference in the AP-initial position in 3-syllable APs, whereas the pitch difference in the same position produced by males born in the 1990s is about 4 St. Similarly, male speakers born in the 1960s show less than a 3 St pitch difference between the H-L contexts in the AP-initial position of 5-syllable APs, but those born in the 1990s have about 4 St difference in the same position. The observations suggest that male speakers born in the 1990s show similar pitch patterns to those of their female counterparts, indicating that the change is nearly complete.

To examine if this observation is statistically significant, I build four linear mixed-effects models for the AP-initial, -second, -penultimate, and -final positions. Syllable positions are defined in the same way in the previous sections (2-syllable APs: Initial–Final, 3-syllable APs: Initial–Penultimate–Final, 4-syllable APs: Initial–Second–Penultimate–Final, 5-syllable APs: Initial–Second–(Third)–Penultimate–Final), and the dependent variable is the pitch value (St) in a given syllable position. Gender, YOB, and Context are included as independent variables, where the reference category for Gender is male, and the reference for Context is the L-inducing context. YOB is centered at 1978, the median year of birth of the speakers examined in this section. Each speaker is added as a random effect. Table 5.6 summarizes the outputs of the mixed-effects regression models.

An important point in the results in Table 5.6 is that the effect of Context and two-way or three-way interactions with Context (YOB*Context, Context*Genter, and YOB*Context*Gender) are mostly found significant in all models. As expected, the results suggest that the effect of Context is highly significant in all syllable positions ($p < 0.001$ in all comparisons), indicating that the effect of an AP-initial H-inducing consonant has a global effect on the pitch value of the entire AP. The models estimate that the effect size is the largest in the AP-initial position, where the H-inducing context is 3.91 St higher than the L-inducing context. The estimated pitch difference between the two contexts is the smallest in the AP-final position (1.34 St), but it is still found significant ($p < 0.001$).

Table 5.6: The outputs of the linear mixed-effects models for gender comparisons in the two corpora. The reference category for Context is L-inducing and the reference of Gender is male. YOB is centered at 1978.

	AP-initial				**AP-second**			
	Est.	SE	t	Pr(> |t|)	Est.	SE	t	Pr(> |t|)
(Intercept)	3.02	0.14	20.86	< 0.001***	4.43	0.16	27.33	< 0.001***
YOB	-0.01	0.02	-0.816	0.417	0.02	0.02	1.06	0.292
Context	3.91	0.06	65.93	< 0.001***	2.72	0.06	45.66	< 0.001***
Gender	-0.04	0.24	-0.15	0.879	0.02	0.26	0.65	0.517
YOB*Context	0.03	0.01	4.77	< 0.001***	0.01	0.01	2.15	0.032*
YOB*Gender	0.01	0.02	0.60	0.552	-0.04	0.03	-1.42	0.161
Context*Gender	1.00	0.10	10.30	< 0.001***	0.82	0.10	8.43	< 0.001***
YOB*Con*Gen	-0.08	0.01	-7.84	< 0.001***	-0.03	0.01	-3.27	0.001**

	AP-penultimate				**AP-final**			
	Est.	SE	t	Pr(> |t|)	Est.	SE	t	Pr(> |t|)
(Intercept)	4.43	0.16	28.36	< 0.001***	4.58	0.15	30.12	< 0.001***
YOB	0.02	0.02	0.90	0.373	0.01	0.02	0.415	0.680
Context	1.86	0.07	27.51	< 0.001***	1.34	0.06	21.04	< 0.001***
Gender	0.08	0.26	0.32	0.750	0.26	0.25	1.05	0.297
YOB*Context	0.00	0.01	0.61	0.542	0.01	0.01	1.57	0.115
YOB*Gender	-0.03	0.03	-0.97	0.333	-0.03	0.03	-1.235	0.220
Context*Gender	0.69	0.11	6.17	< 0.001***	0.43	0.10	4.08	< 0.001***
YOB*Con*Gen	-0.03	0.01	-2.92	0.003**	-0.03	0.01	-2.68	0.007**

The significant effects of the two-way interaction between YOB and Context in the AP-initial and -second positions confirm the observation made in Figure 5.4 that the pitch difference between the H-L contexts increases for younger male speakers ($p < 0.001$ in the AP-initial position, $p = 0.032$ for the AP-second position).[6] For example, the model for the AP-initial position estimates that the H-L difference increases 0.03

[6]Note that the reference category for gender is male.

St per year of birth, meaning that male speakers born in 1960 produce a 3.37 St pitch difference (= 3.91 + (0.03 x (1960–1978)), whereas males born in 1996 show a 4.45 St difference (= 3.91 + (0.03 x (1996–1978)). The two-way interaction of YOB and Context is not significant in the AP-penultimate or -final positions, suggesting that male speakers are still in the process of developing the H-L pitch difference in non-initial positions.

Also, the two-way interaction of Context x Gender is found to be significant in all syllable positions ($p > 0.001$ for all comparisons), indicating that female speakers produce a larger H-L pitch difference than males in all syllable positions. For example, the model estimates that female speakers show a 1 St higher pitch difference than males in the AP-initial position. The effect size is the smallest in the AP-final position, where the estimated pitch difference is 0.43 St larger for females than males, but it is still significant ($p < 0.001$).

However, the significant three-way interaction of YOB x Gender x Context reveals that the gender difference in the H-L contrast is modulated by YOB ($p < 0.001$ for AP-initial, $p = 0.001$ for AP-second, $p = 0.003$ for AP-penultimate, $p = 0.007$ for AP-final). In other words, females show a larger H-L difference than male speakers in all syllable positions, but this gender difference decreases for younger speakers.[7] For example, it is estimated that female speakers born in 1960 produce a 2.44 St larger H-L difference (= 1.00 + (-0.08 x (1960–1978)) than their male counterparts,

[7] The estimated coefficients of the three-way interactions are negative values, meaning that the gender difference *decreases* by speakers' YOB.

whereas the H-L difference that females born in 1996 produce is 0.44 St smaller than that of their male counterparts (= 1.00 + (-0.08 x (1996–1978)) in the AP-initial position. The models estimate that the gender difference in the H-L pitch contexts also decreases (0.03 St per year of birth) in non-initial positions. This result confirms the observation in Figure 5.4 that male speakers born in the 1990s have caught up to female speakers in terms of the tonogenetic sound change, resulting in a smaller gender difference in pitch. This, in turn, supports the finding in the previous section (Section 5.2.1) that the tonogenetic change has nearly completed.

5.2.3 All APs by Gender and Context

The previous sections demonstrate that the tonogenetic change has almost completed, showing that the AP pitch patterns produced by speakers born in the 1990s are highly similar to those of young speakers in the NIKL corpus (born after 1960) and that male speakers born in the 1990s catch up to their female counterparts in terms of the tonogenetic change. However, the results of the previous sections are based on sentence-initial APs. In this section, I examine all APs that are 2 to 5 syllables long, regardless of their position within the target sentences, to see if the same results are found. The same 77 speakers in Table 5.4 are examined (30 females and 48 males from the two speech corpora), and the breakdown of all APs examined in this section by AP size and AP-initial onset type is shown in Table 5.7. In total, 399 APs are analyzed.

Table 5.7: Number of all APs that are 2 to 5 syllables long in the target sentences by AP size and AP-initial onset type. In total, 399 APs are examined.

	2-syllable	3-syllable	4-syllable	5-syllable
Aspirated	23	44	33	13
Tense	10	10	14	5
Lenis	27	51	44	12
Sonorant (or vowel)	18	39	42	14

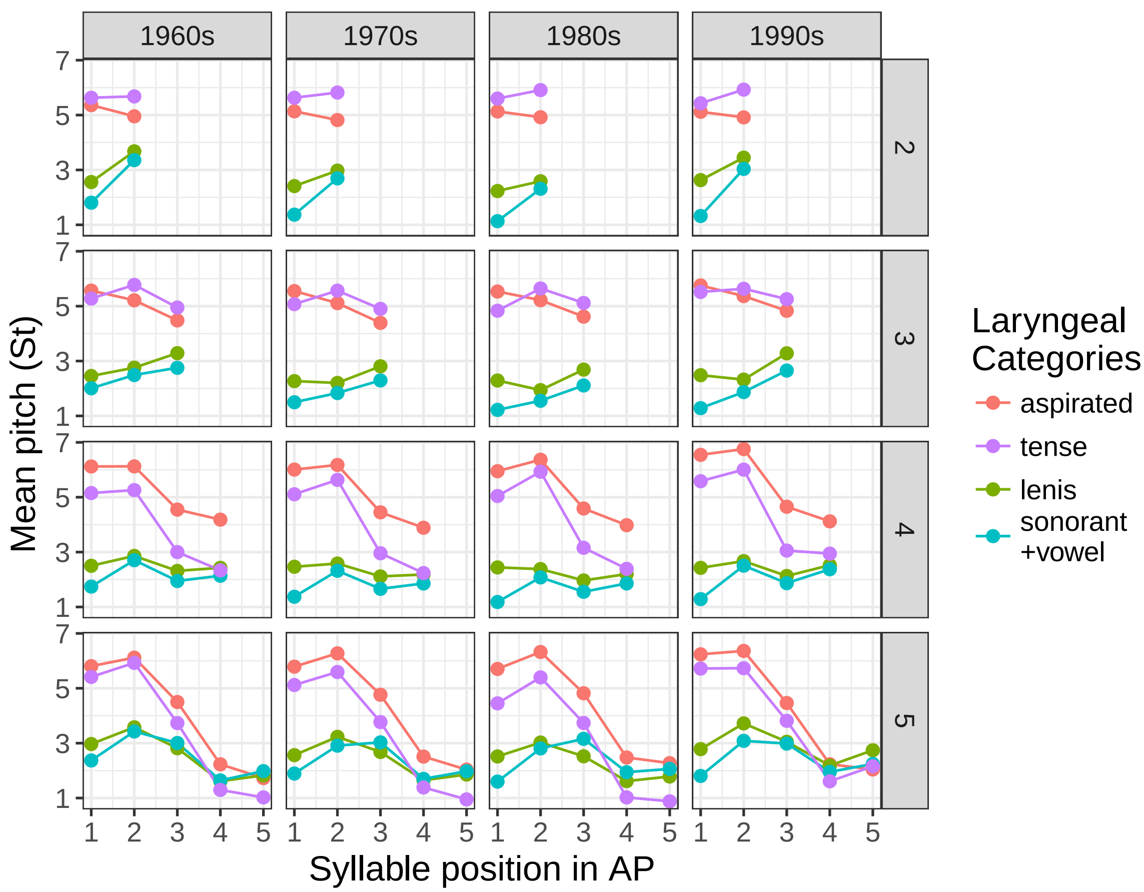


Figure 5.5: Mean pitch patterns of all target APs produced by speakers born after 1960 in the NIKL corpus and speakers born in the 1990s in the new corpus by AP-initial onset type and speakers' YOB.

Figure 5.5 shows the mean pitch patterns of all APs that are 2 to 5 syllables long in the target sentences. It is clear that aspirated- and tense-initial APs pattern together, whereas lenis- and sonorant-initial APs group together. Also, it is notable that the pitch patterns of speakers are quite similar, whether born in the 1960s, 70s, 80s, or 90s, although the pitch difference between H-inducing and L-inducing contexts is slightly larger for speakers born in the 1990s than those born in the 1960s. In particular, speakers born after 1970 show highly similar pitch patterns, indicating that the result in the previous sections that the change is almost complete is consistently found when all APs are examined.

Moving on to the gender difference, Figure 5.6 illustrates the mean pitch patterns of all APs by AP size, AP-initial pitch context, and speakers' YOB, aggregated in 10 year bands. This figure demonstrates that the H-L pitch difference has increased in male speakers' speech such that males born in the 1990s show almost identical pitch patterns to their female counterparts. Male speakers born in the 1990s make an interesting comparison to males born in the 1960s, who produce only half of the H-L pitch difference of their female counterparts. This is partly because female speakers born in the 1960s have unusually wide pitch ranges (See Appendix A.1), but the male speakers indeed have a smaller H-L pitch difference even when compared to males born in the 1970s. Considering that the H-L pitch difference increases for younger male speakers and also that speakers born in the 1990s do not show any gender difference when all APs are examined, it seems clear that the change has reached completion.

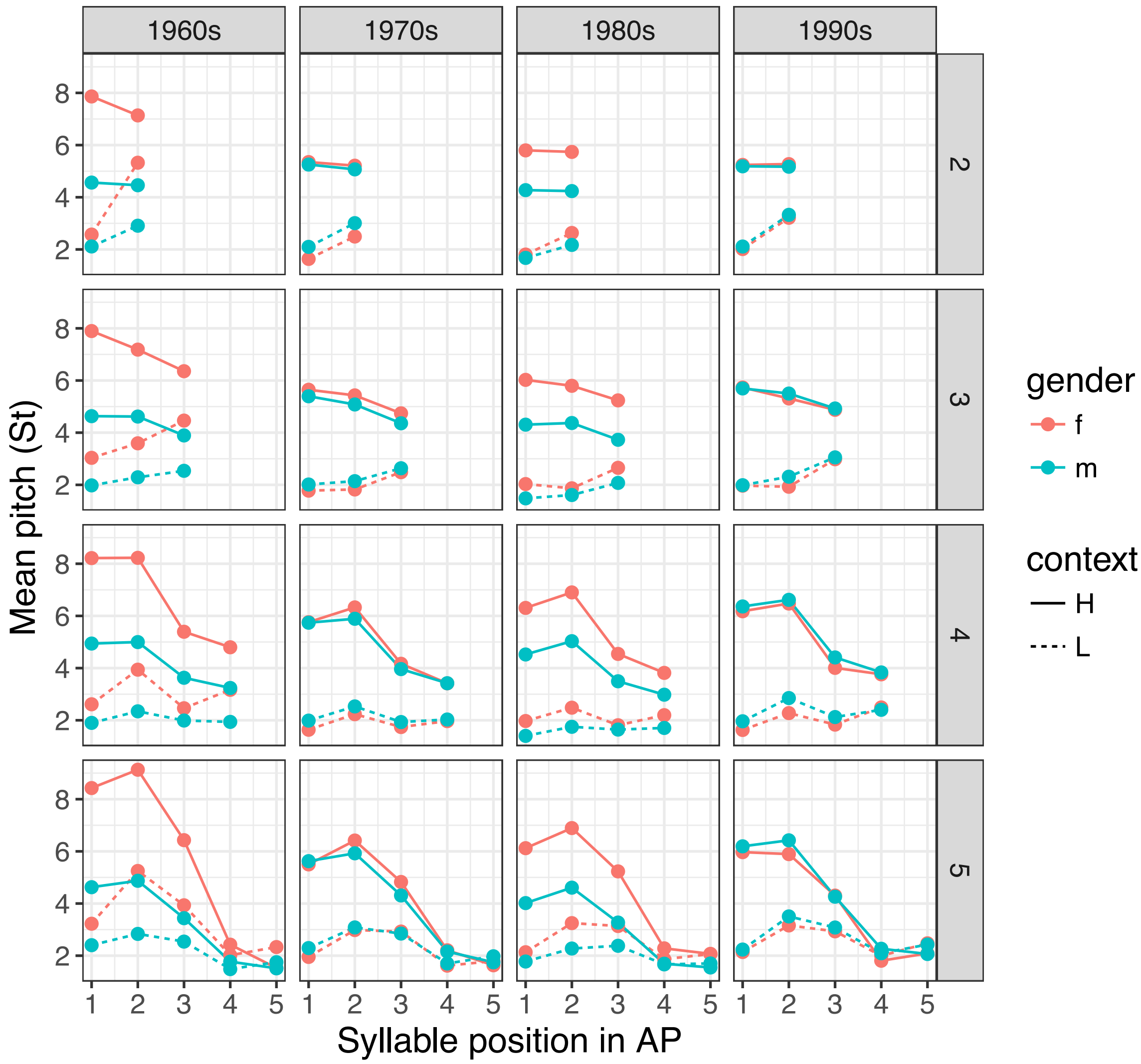


Figure 5.6: Mean pitch patterns of all target APs produced by speakers born after 1960 in the NIKL corpus and speakers born in the 1990s in the new corpus by speakers' gender and AP-initial pitch context.

Chapter 6

Discussion

This chapter brings the results of Chapters 3, 4, and 5 together and discusses several important points in the findings. Section 6.1 discusses both the merger of the VOT contrast in Kang (2014) and the trajectory of the pitch development in Chapters 4 and 5, as well as how these two changes progressed over time in Contemporary Seoul Korean. Section 6.2 describes the modes of the tonogenetic change and its phonetic and phonological aspects. Section 6.3 compares the case of Seoul Korean to other languages which previously underwent tonogenesis. Based on the discussion in Section 6.3, I suggest the implications that the tonogenetic change in Seoul Korean has for the theory of tonogenesis (Section 6.4).

6.1 VOT merger and development of pitch contrast

The trade-off between the VOT contrast and the pitch contrast in Seoul Korean is examined by Kang (2014), who uses the same corpus used in Chapter 4. Her study measures VOT values of word-initial consonants in sentence-initial APs, so it is worthwhile to consider the findings in Section 4.2 (sentence-initial APs) with the results of her study. Consider Figure 6.1 below, which is repeated from Section 2.1.

Kang's study shows that female speakers do not produce a large VOT difference between aspirated and lenis stops regardless of their age, indicating that females born in the 1930s are already in the process of the VOT merger. In particular, female speakers born in the 1980s show a fully overlapping VOT range for aspirated and lenis stops, which suggests that the VOT contrast is no longer used by young females. On the other hand, older males (born before 1960) show the most conservative VOT contrast of the stop categories among all age x gender groups (Aspirated > Lenis > Tense), whereas younger males (born after 1960) show a smaller VOT contrast between aspirated and lenis stops than older ones. However, it is noticeable that the VOT of aspirated and lenis stops is still distinctive even for the young male speakers born in the 1980s, although they have a wider interspeaker variation in VOT than males born in the 1970s.

Now consider Figure 6.2, which shows the mean pitch pattern of sentence-initial

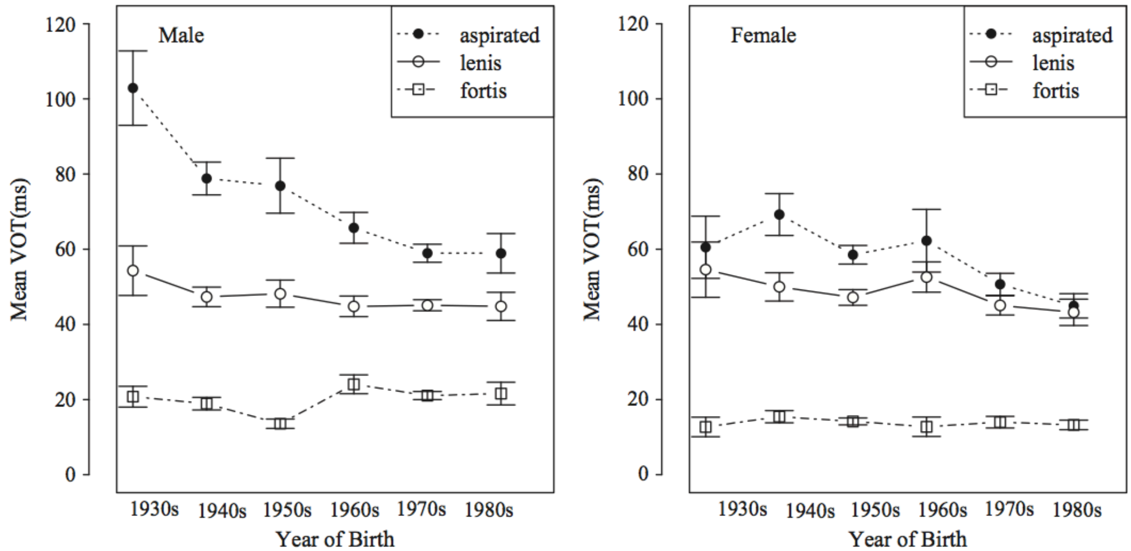


Figure 6.1: By-speaker difference in mean VOT between aspirated and lenis stops plotted against speakers' YOB and gender (Borrowed from Kang 2014 and repeated from Section 2.1).

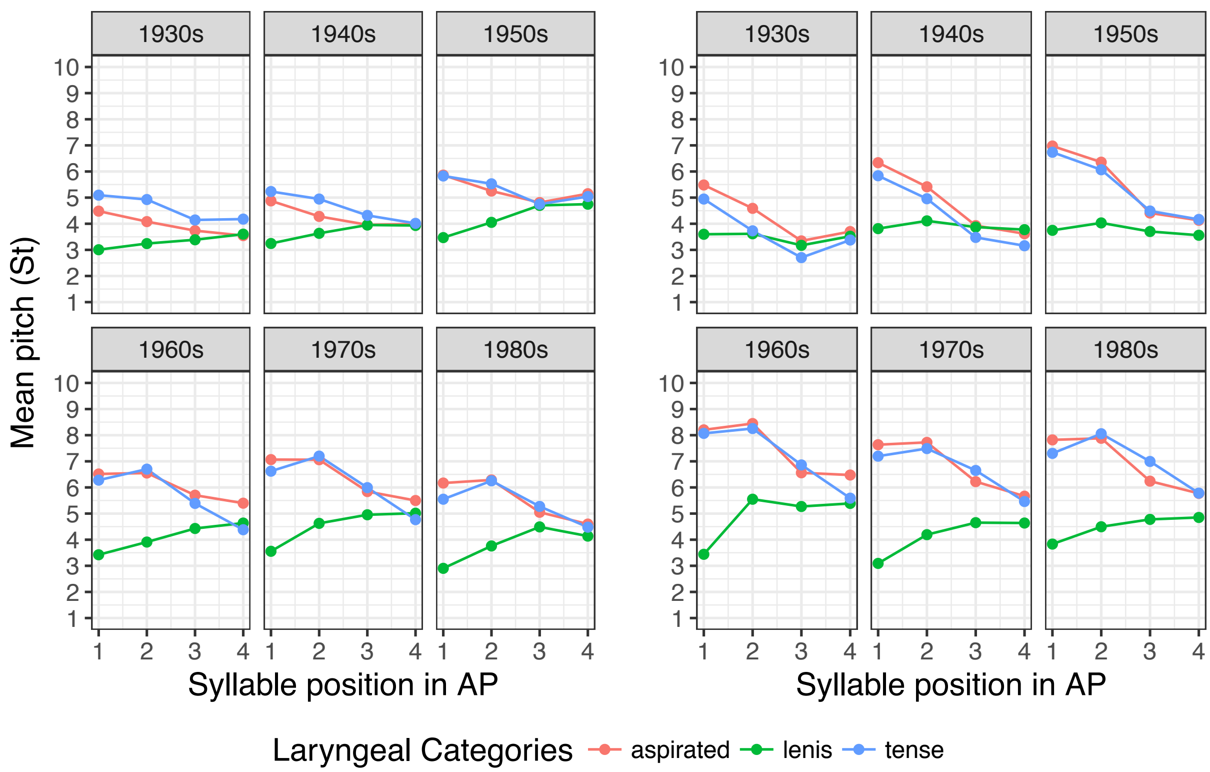


Figure 6.2: Mean pitch patterns of sentence-initial 4-syllable APs by the three largyneal categories and speakers' YOB and gender (Left: males, Right: females).

4-syllable APs by the laryngeal categories, speakers' YOB and gender. This figure is plotted with the same data set used in Section 4.2.4.3. Figure 6.2 illustrates that the pitch difference in the laryngeal categories originates in the speech of females born in the 1950s. They show a large pitch difference in the AP-initial as well as the AP-second syllables. The pitch difference in non-initial positions increases in female speakers born after 1960, whose VOT values between lenis and aspirated start to overlap with each other.

On the other hand, for male speakers, the pitch difference in the AP-initial and second positions is observed in speakers born after 1960, although their pitch difference in non-initial positions is not as large as the one found in young female speakers. Considering that young male speakers still show a small VOT difference between aspirated and lenis and that they do not show a large pitch difference in non-initial positions, it seems that the pitch difference in the AP-initial syllable accompanies the reduction of the VOT contrast between lenis and aspirated and the H-L pitch difference in non-initial positions further facilitates a complete merger of the VOT contrast. A rough timeline of the VOT merger and the development of the pitch contrast is shown in Figure 6.3.

Figure 6.3 illustrates these points: The VOT contrast is not used for young female speakers born after 1980. Females born after 1950 show the pitch difference between H-initial and L-initial APs, indicating that speakers born after 1950 and before 1980 use both cues in distinguishing the stop categories. Male speakers are about 10 years behind female speakers with regard to the pitch contrast development and the VOT

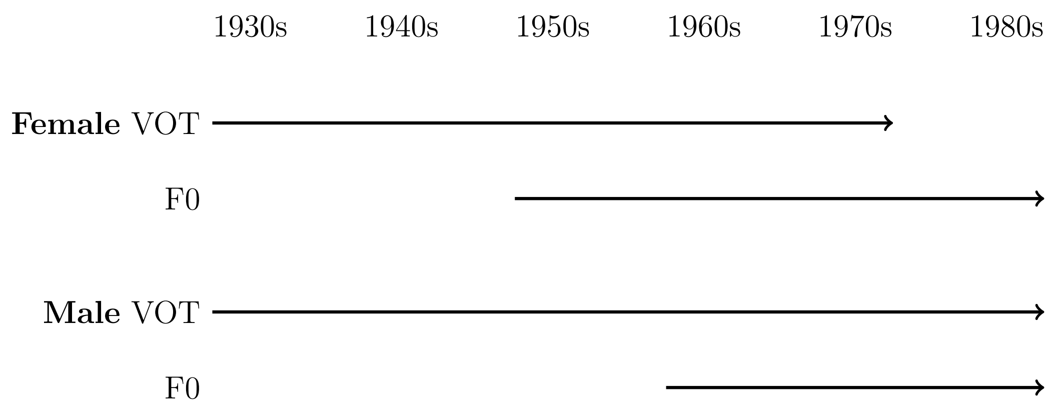


Figure 6.3: Schematic timeline of the VOT merger and the change of the intonational melody.

difference is still used in young male speakers.

The trajectory of the two sound changes in Figure 6.3 has an implication for the results of Chapter 5. Speakers born in the 1990s show little gender difference in terms of AP pitch patterns, and the male speakers also show a pitch difference in non-initial APs depending on AP-initial onset consonants. This seems to suggest that the VOT merger has further progressed in the speech of males born in the 1990s. They may show considerably overlapping VOT values between aspirated and lenis stops, because the H-L pitch difference in non-initial APs in female speakers' speech has facilitated the merger of the VOT contrast. However, investigating the trade-off between the VOT merger and the development of the AP-initial pitch difference in speakers born in the 1990s is beyond the scope of this dissertation, so I leave this as a possibility for a future study.

6.2 H-pitch spreading and the mode of change

The results of Chapters 3, 4, and 5 show an interesting interaction between the AP-initial H context and the intonational L tone in the penultimate syllable for younger speakers. That is, for short APs, the effect of the H-initial context lasts to the end of the phrase, showing a much higher pitch value on the final syllable. However, for longer APs, it is observed that the effect of the H-initial context decreases substantially after the intonational L tone, only showing a small pitch difference in the final syllable.

Figure 6.4: Direction of change in two approaches (H-initial APs): the H-tone spreading approach (left) and the interpolation approach (right). The direction of change is indicated with an arrow.

This pattern can be interpreted in two ways as schematized in Figure 6.4. The first possible interpretation is that the AP-initial H tone spreads to later syllables up until the intonational L tone in the penultimate syllable. In other words, the intonational L tone can be analyzed as blocking the spread of the H tone. This approach explains why we observed such a difference between short APs and long ones. In the results of

Chapters 3, 4, and 5, 2-syllable and 3-syllable APs were produced with a T-H or TH-H melody, meaning that the intonational L tone was undershot and did not block the spread of the AP-initial H tone. Therefore, the H tone seemed to spread to the end of the phrase, resulting in a large pitch difference even in the final syllable position. On the other hand, 4-syllable and 5-syllable APs were produced with a TH-LH melody, and the effect of the AP-initial H context was hardly seen in the final syllable. It can be interpreted that the intonational L tone blocks the H tone spread.

An alternative interpretation is that the style of interpolation from the intonational H tone of the second syllable to the intonational L tone of the penultimate syllable has changed. In Jun's original theory of the Korean prosodic system (1993, 1996, 2006, and in her other works), she explains that AP-medial syllables that are not the first two or the last two receive their pitch values via linear interpolation from the second H (T**H**-LH) to the penultimate L (TH-**L**H) in longer APs (Section 1.1.2). Considering the AP-initial pitch context is highly predictable based on the AP-initial onset consonant, except the lexically specified H tone on [il] reported in Jun and Cha (2015), it can be analyzed that the style of interpolation has been changed from a linear one to an angled one. That is, the H pitch of the AP-initial syllable is maximally maintained until it has a reason to lower down. For now, it is hard to tell which approach is the correct interpretation of the phenomenon.

However, it seems clear that the modes of the pitch change in the AP-initial position and non-initial positions are different. Consider Table 6.1, which is borrowed from Bermúdez-Otero 2007. Bermúdez-Otero states that there are three modes of implementation with regard to sound change. Lexical diffusion is implemented with an abrupt change in the phonetic dimension but with a gradual spread in the lexical dimension. For example, the H-toned [il] in Jun and Cha (2015) can be considered as an example of lexical diffusion, where the pitch value of the H-toned [il] is considerably high when compared to other syllables beginning with an /i/ vowel. However, the H tone on [il] is most frequently observed in [il] meaning ‘1’, and less frequently when [il] means ‘day’ or ‘work’.

Table 6.1: Modes of implementation in sound changes. (Borrowed from Bermúdez-Otero 2007)

Mode of implementation		**Possible?**	**Innovation in which component of grammar?**
phonetic dimension	lexical dimension		
abrupt	gradual	Yes	lexical representation
abrupt	abrupt	Yes	phonological rules
gradual	abrupt	Yes	phonetic rules
gradual	gradual	No	

As for the phonetic implementation, the AP-initial pitch difference investigated in this dissertation also seems to abruptly change. For example, female speakers born in the 1950s suddenly show a large pitch difference in the AP-initial and second syllables depending on the laryngeal category of an AP onset consonant, and the change from males born in the 1950s to 60s is also quite abrupt (e.g., see Figure 6.2). Also, the corpus study in Section 4.3 demonstrates that the tonogenetic change is a structural one, affecting all APs starting with aspirated or tense obstruents. This means that the change in the lexical dimension is also abrupt. Altogether, it suggests that the development of the pitch contrast in the AP-initial position is a phonological change.

In contrast, the change in non-initial positions seems to be gradual in the phonetic dimension in that it is less noticeable than the development of the pitch contrast in the AP-initial syllable. Yet, it is also a structural change, i.e., abrupt in the lexical dimension, since the pitch value is largely affected by whether the AP-initial consonant is a H-inducing one or not, regardless of the meaning of the AP. It suggests that the development of the pitch difference in non-initial syllables is a phonetic change, in contrast to the phonological change of the AP-initial pitch contrast. This indicates that the two interpretations in Figure 6.4 need to be considered as interpretations of a phonetic change.

6.3 Comparison to other languages

For better understanding of the tonogenetic change in Seoul Korean, it is necessary to compare it to the process of tonogenesis in other languages. In the case of Seoul Korean, it is found that the AP-initial H-inducing consonants affect pitch values of later syllables. However, for various reasons, this effect is rarely seen in other languages that previously underwent tonogenesis.

For example, tones seem to have developed based on syllables in Vietnamese. In Vietnamese, one syllable corresponds to a single morpheme in most cases (sometimes called ‘syllabemes’, which is coined from ‘syllable’ and ‘morpheme’) and many words are monosyllabic ones or disyllabic compounds with monosyllabic morphemes (Thompson 1987). Furthermore, unlike Korean, Vietnamese is an isolating language without grammatical markers for case, tense, (grammatical) gender, and agreement. Accordingly, previous studies on Vietnamese tonogenesis lack an explanation of polysyllabic words and the effect of consonants on the pitch of later syllables within the same word (Matisoff 1973, Thurgood 2002, Kingston 2011, among others). For instance, Haudricourt's model of tonogenesis in Vietnamese (1954) in Table 6.2 explains that a syllable-final consonant determined a pitch contour and a syllable-initial onset decided the pitch height of the syllable, resulting in six tones (two pitch heights x three contours). This suggests that the domain of the development of tonal contrasts was a syllable (or a morpheme) in this language.

Table 6.2: Vitenamese tonogenesis in the consonant-based approach (borrowed from Kingston 2011 and repeated from Chapter 2.4). "T" stands for any following stops, and the Vietnamese tone names are in parentheses.

		Following consonants		
		CV, CVN	CVT	CVs, CVh
Preceding consonants	voiceless	*pa > pa high level (ngang)	*pak > pak high rising (sac)	pas > pa high falling (hoi)
	voiced	*ba > pa low level (huyen)	bak > pak low rising (nang)	bas > pa low falling (nga)

In other Southeast Asian languages, such as Khmu (also known as Kammu) or Utsat (also known as Tsat or Hainan Cham), the loss of a word-initial syllable in a disyllabic (sesquisyllable) word resulted in many monosyllabic words in the languages

(Thurgood 1992, Suwilai 2004, Kingston 2011), making it difficult to see what happened in polysyllabic words in these languages. For example, in the case of Khmu, most words are monosyllabic or disyllabic (sesquisyllabic). In monosyllabic words, a High tone in tonal dialects of Western Khmu (e.g., /p**ó**k/ 'to take a bite') corresponds with a voiceless onset in Eastern Khmu (/**p**ok/ 'to take a bite'); A Low tone in Western Khmu (e.g., /p**ò**k/ 'to cut down a tree') corresponds with a voiced onset in Eastern Khmu (/**b**ok/ 'to cut down a tree'). In disyllabic words, the entire presyllable or the initial onset of a presyllable was deleted in tonal dialects of Western Khmu (e.g., 'husked rice' in Eastern Khmu is /r^ɨŋ**koʔ**/, but /**kóʔ**/ in Western Khmu), because presyllables were unstressed and had reduced vowels. As a result, both monosyllabic and disyllabic words developed into monosyllabic ones in the tonal dialects, whereas presyllables in non-tonal dialects were preserved (Table 6.3).

Table 6.3: Examples of disyllabic words in East Khmu, which correspond to monosyllabic words in West Khmu in Suwilai 2004.

East Khmu	West Khmu (tonal)	English gloss
h^imma:l	ma:l	'soul'
k^əndrɔːŋ	ntrɔːŋ/trɔːŋ	'back part of the body'
h^ərlɔʔ	rlɔ́ʔ/lɔʔ	'language'
h^əʔeʔ	ʔéʔ	'firewood'
h^iʔiər	ʔíər	'chicken'
h^uʔuʔ	ʔúʔ	'bad smell'
h^ɨʔɨr	ʔɨ́r	'fragrance'
r^ɨŋkoʔ	kóʔ	'husked rice'

Similarly, monosyllabic words in Utsat were developed from what were previously disyllabic in proto-Chamic, due to the deletion of presyllables (Thurgood 1992). Consider the examples in Table 6.4, which is repeated from Section 2.4.

Table 6.4: Correspondences of tone in Utsat and coda consonants in proto-Chamic. Cognates in Chru are also given for comparison. Examples are from Thurgood 1992. (Table 2.5 is repeated here.)

	Proto-Chamic	Utsat	Chru	English gloss
Voiceless coda	*dilah	la^{55}	dəlah	'tongue'
	*tanah	na^{55}	tənah	'earth'
	*do:k	$thoʔ^{53}$	dò	'live'
	*tikiʔ	$kiʔ^{53}$	təkì	'few'
	*pa:ʔ	pa^{35}	pà	'four'
	*sakiʔ	ki^{35}	-səkì	'sick, painful'
	*ʔana:k	na^{35}	anà	'child'
Voiced coda	*lapa	pa^{33}	ləpa	'hungry'
	*ʔapui	pui^{33}	aphi	'fire'
	*ʔikan	$ka{:}n^{33}$	akàn	'fish'
	*ʔatas	ta^{33}	—	'far'
	*dua	$thua^{11}$	dua	'two'
	*batəu	tau^{11}	pətəu	'stone'

Some presyllables in proto-Chamic dropped without any trace as in **ʔra**tas*** 'far' in proto-Chamic > /ta^{33}/ in Utsat, where the final coda consonant **-s* of the second syllable **-tas* was lost and the second syllable acquired a mid level tone like other syllables ending in a vowel or a sonorant. In contrast, in other cases, the deleted voiced stop resulted in a Low level tone of the next syllable without changing the voicing of the following onset as in **batǝu* 'stone' in proto-Chamic > /tau^{11}/ in Utsat. Whether the presyllable deleted with a trace or not, the result of tonogenesis brought monosyllabism in Utsat; As a result, it is also hard to see the effect of tonogenesis in polysyllabic words in this language.

The case of Proto North Huon Gulf (PNHG) in the Oceanic Austronesian language family is different from those of Southeast Asian languages. Ross (1993) states that tonogenesis had occurred during an interstage from Proto Huon Gulf (PHG) to PNHG. It is similar to tonogenesis of other languages in that firstly, it resulted in monosyllabism in many words and secondly, vowels with a voiced Post-PHG obstruent acquired a L tone, while those after a voiceless Post-PHG obstruent developed into a H tone as in the examples in (1).

(1) Examples of reconstructed words in PHG and PNHG

   a. PHG *p > PNHG *p with high tone
      i. *-pipi > *-píʔ 'squeeze'
      ii. *-peka > *péʔ 'defecate'
   b. PHG *b > PNHG *b with low tone
      i. *buak > *bùʔ 'areca nut'
      ii. *bol > *bɔ̀ʔ 'pig'
   c. PHG *t > PNHG *t with high tone
      i. *tete > *téʔ 'ladder'
      ii. *ate- > *áté- 'liver'
   d. PHG *d > PNHG *d with low tone
      i. *dala > *dàʔ 'blood'

ii. *dili > *-dìʔ ‘stand’

e. PHG *k > PNHG *k with high tone

i. *kurit > *kúlíʔ ‘octopus’

ii. *wakac > *wákáʔ ‘root’

f. PHG *g > PNHG *g with low tone

i. *gac > *-gàʔ ‘open (sg)’

ii. *jagiŋan > *zàgìŋ ‘partition wall’

However, unlike tonogenesis in other languages, vowels after sonorants acquired a H tone in PNHG, and more importantly, tone and voicing harmony also occurred such that all vowels in a morpheme (usually, a foot) had the same tone. Ross (1993, p.146) proposes the process had taken place as follows:

(2) Pre-PNHG harmony
where the weak and strong syllables of a foot differ in tone and voicing

a. if the onset of the strong syllable is a Post-PHG obstruent, the weak syllable acquires the tone and voicing of the strong;

b. otherwise, if the onset of the weak syllable is a Post-PHG (voiced) obstruent, the strong syllable acquires the tone and voicing of the weak.

According to Ross (1993), PNHG had an iambic foot structure, and when tonogenesis resulted in a disagreement in tone within a foot, tone harmony occurred as in the following examples:

(3) Examples of tone and voicing harmony in PNHG (PHG > PNHG)

a. *labi > *làbì ‘sago palm’

b. *ta-dil > *dà-dìʔ ‘stand (1st inclusive pl.)’

c. *ga-pip > *ká-píp ‘squeeze (1st sg.)’

d. *guluk > *gùlù- ‘nasal mucus’

In (3a), the weak syllable *la-*, which was supposed to show a H tone since it did not start with a voiced obstruent, acquired the L tone of the strong syllable *-bi* as explained in (2a). Similarly, the weak syllable *ta-* in PHG agreed with the voicing and the L tone of the strong syllable *-dìl*, resulting in *dà-* in (3b). In (3c), the spreading of voicing and tone from the strong syllable to the weak one was similar to the previous examples, but in this case, the voicing and the H tone of the strong syllable *-píp* spread to the weak syllable, making it *ká-*. The example in (3d) was the case where (2b) was applied. The onset of the weak syllable was a voiced obstruent **g* in PHG, the strong syllable *-luk*, which was supposed to show a H tone, acquired the L tone and voicing of the weak syllable *gà-* via the tone and voicing harmony.

The tone harmony is still observed in two daughter languages of PNHG: Yabem and Bukawa, where tone is a feature of a foot, not a syllable (Ross 1993). In Yabem, the voicing and tonal contrasts still coexist such that obstruents in all syllables in a low-toned foot are voiced. For example, in Table 6.5, all forms for the voiceless stem *-téŋ* ‘weep’ are high-toned, and obstruents in the subject pronominal prefixes are all voiceless. For the voiced stem *-dèŋ* ‘move towards’, all syllables are low-toned and the obstruents in the prefixes are voiced.

However, in Bukawa, the voicing harmony has been lost via diachronic changes so that voicing and tone within a foot no longer correlate. For example, the 1st sg. realis

Table 6.5: Examples of the realis paradigm in Yabem and Bukawa from Ross 1993. Note that Bukawa does not distinguish the 2nd and 3rd person singular forms.

<table>
<tr><th>-táŋ ‘weep’</th><th>Yabem</th><th>Bukawa</th></tr>
<tr><td>1st sg.</td><td>ká-táŋ</td><td>gá-táŋ</td></tr>
<tr><td>2nd sg.</td><td>kó-táŋ</td><td rowspan="2">∅-táŋ</td></tr>
<tr><td>3rd sg.</td><td>ké-táŋ</td></tr>
<tr><td>1st incl. pl.</td><td>tá-táŋ</td><td>dá-táŋ</td></tr>
<tr><td>1st excl. & 2nd pl.</td><td>á-táŋ</td><td>á-táŋ</td></tr>
<tr><td>3rd pl.</td><td>sé-táŋ</td><td>sé-táŋ</td></tr>
<tr><th>-d/tèŋ ‘move towards’</th><th>Yabem</th><th>Bukawa</th></tr>
<tr><td>1st sg.</td><td>gà-dèŋ</td><td>gà-tèŋ</td></tr>
<tr><td>2nd sg.</td><td>gò-dèŋ</td><td rowspan="2">∅-tèŋ</td></tr>
<tr><td>3rd sg.</td><td>sè-dèŋ</td></tr>
<tr><td>1st incl. pl.</td><td>dà-dèŋ</td><td>dà-tèŋ</td></tr>
<tr><td>1st excl. & 2nd pl.</td><td>a-dèŋ</td><td>a-tèŋ</td></tr>
<tr><td>3rd pl.</td><td>se-dèŋ</td><td>se-tèŋ</td></tr>
</table>

form of ‘to weep’ is *gátáŋ*, with a voiced onset *g* and a H tone in the first syllable, and the 1st sg. realis form of ‘to move forward’ is *gà-tèŋ*, with a voiceless onset *t* and a L tone in the second syllable (Table 6.5). However, it is noticeable that the same tone is shared within a foot in both Yabem and Bukawa so that tone is considered as a feature of a foot, not a syllable.

6.4 Implications in tonogenesis

Considering previous studies, the findings in this dissertation suggest that Seoul Korean is a unique case as of a tonogenetic sound change. It is similar to the cases of other languages in that a consonantal contrast is traded off for a pitch contrast.

However, the newly emerged pitch contrast brings a different consequence in Seoul Korean. In most cases in the previous studies, the domain of tonogenesis is a syllable, only affecting the pitch value of the given syllable (Vietnamese, Khmu, Utsat). There are few cases of tonogenesis that have affected up to two syllables (PNHG). However, in Seoul Korean, the effect of an AP-initial H-inducing consonant reaches to the final syllable of the same AP for short APs and up to the penultimate syllable for longer APs. Although we have observed the effect of consonant-induced pitch in non-initial positions (Chapter 3), the effect was not as large as the one found in the AP-initial position. The effect of consonant-induced pitch in non-initial positions should be considered as phonetic perturbation, whereas the pitch difference in the AP-initial position should be considered as a phonological phenomenon.

The question of why we see such a different tonogenetic pattern in Seoul Korean arises. One possible explanation is that an AP is a strongly-tied prosodic unit, which is rarely separated. This possibility is supported from the results of previous studies that examine prosodic focus in Seoul Korean. For example, previous studies show that the effect of a prosodic focus in Seoul Korean spans over an entire phrase (Jun 2011, Lee 2012, Lee and Xu 2010, Lee et al. 2015). In particular, Lee et al. (2015) state that an AP is the domain of a prosodic focus in Seoul Korean, showing that when one numeric digit in a phone-number string receives a corrective focus, not only the focused digit but also the entire phrase containing the focused digit are affected. Consider Figure 6.5, which is borrowed from Lee et al. (2015).

Figure 6.5 shows that the same phone number strings have different pitch values

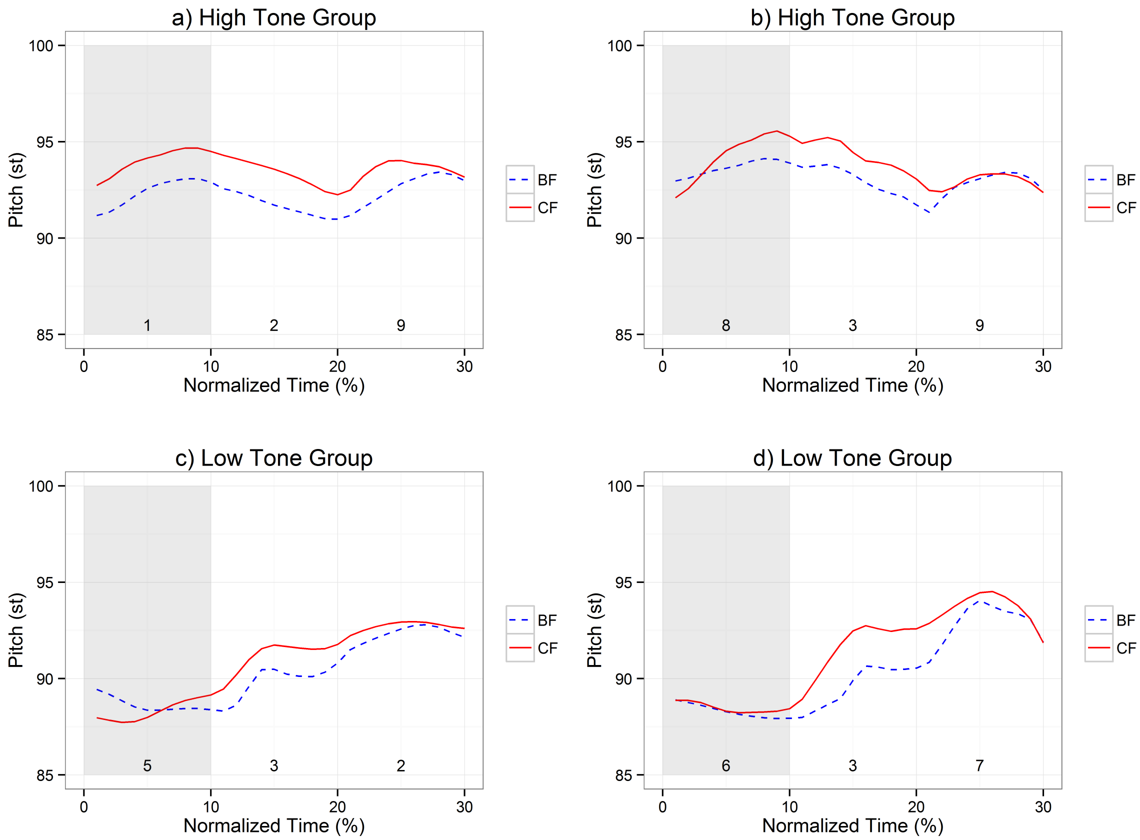


Figure 6.5: Example of pitch contours of H-initial and L-initial phone-number strings by two focus conditions (borrowed from Lee et al. 2015). 'BF' stands for broad focus condition, and 'CF' is for corrective focus. In all plots, the initial digits, which are highlighted in grey, are the corrected one in the CF condition. The digits in the bottom of the plots show the actual digits pronounced.

depending on the focus condition. In general, the phone number string with a corrective focus is higher than the one in the broad focus condition. An interesting point about this figure is that the entire phone number in corrective focus string is produced with high pitch values, even though the only corrected digit is the first one (shaded in grey in the figure). Also, it is noticeable that the effect of prosodic focus is not salient in Seoul Korean. That is, prosodic focus raises the pitch values of a phone-number string, but it does not change the intonational melody of an AP. H-initial APs still show H(H)-LH (the top panels in Figure 6.5) and L-initial strings are produced with a LH-(L)H melody whether in the broad focus or corrective focus condition. These points seem to support the explanation that an AP is one prosodic unit that cannot be separated.

This, in turn, has the broad implication for the theory of tonogenesis that the domain of tonogenesis may be language-specific, varying depending on the phonological and morphological systems of a language. For languages where one syllable corresponds to one morpheme, such as Vietnamese, a tonal contrast emerged on syllables. In Southeast Asian languages with sesquisyllables, such as Khmu and Tsat, tonogenesis resulted in monosyllabism. For languages where a foot is an important unit as in Yabem and Bukawa, tone became a feature of a foot. In the case of Seoul Korean, where an AP is a basic prosodic unit, the tonogenetic change has affected pitch values of an AP. Altogether, a plausible conclusion is that the end result of a tonogenetic change varies depending on the phonological and morphological features of a given language.

Chapter 7

Conclusion

This dissertation aimed to fully investigate the tonogenetic sound change in Seoul Korean from its initial stage to completion. This dissertation's findings can be summarized as follows. First, the results of Chapter 3 showed that young speakers in their 20s produced aspirated-initial APs with a higher pitch range than lenis-, sonorant-, or vowel-initial APs. Chapter 4 demonstrated that female speakers born in the 1950s were the ones who first started to show a large H-L pitch difference in the AP-initial position, and females born after 1960 produced non-initial syllables in the H-initial context with a higher pitch than those in the L-initial context. Male speakers were about 10 years behind in regards to this change. Furthermore, Chapter 4 revealed that all aspirated and tense obstruents were subject to the tonogenetic change, regardless of manner of articulation. Lastly, Chapter 5 illustrated that the change has almost reached completion, as speakers born in the 1990s did not show any gender difference in the H-L pitch contrast.

Although this dissertation fills a gap in literature by fully investigating the pitch change in Seoul Korean, there remain open questions that I hope to answer with future research. First, while this dissertation examines the roles of gender and age in the change of Seoul Korean intonation, other social factors must be investigated to better understand the change. For example, sociolinguistic studies in English show that middle-working and upper-working classes tend to lead changes of which native speakers are not consciously aware (Labov 1990, 2001). However, our understanding of these classes' roles in the Seoul Korean pitch contrast is incomplete. Questions such as "Which social class has led this change?" and "How does education level play a role in this change?" remain to be answered in future research.

Furthermore, much less is known about what happens in other Korean dialects. Considering Seoul Korean is the most influential and prestigious dialect in South Korea, it may be the case that the H-L pitch patterns are being diffused in other dialects. For example, Holliday and Kong (2011) show that young speakers of Jeju Korean are quite similar to young Seoul speakers in terms of VOT values, as well as pitch values, of the stop consonants. However, their results show that speakers of Daegu Korean, which is known as a pitch-accent language (North Kyungsang Korean), produce the most conservative VOT distinction between the three stop categories. It remains to be seen in future research what the intonation patterns of these dialects look like.

This dissertation's results have an important implication for the theory of tonogenesis: the end result of tonogenesis varies depending on the phonological and morphological characteristics of a given language. Most previous studies on tonogenesis have been about Southeast Asian languages, and many of these languages share similar linguistic characteristics or belong to the same language family. Therefore, it has been presumably assumed that tonogenesis would affect only syllables. However, this dissertation demonstrates that a tonogenetic sound change can affect a large prosodic unit, such as an AP, indicating that not only syllables but also larger prosodic units can change due to tonogenesis.

Another implication of this dissertation for tonogenesis literature is that it is the first case study that examines a tonogenetic sound change and its effect on intonation from its initial stage to completion. Most of the previous studies on tonogenesis have been focused on the reconstruction of proto languages or the description of a tonogenetic process, without a phonetic analysis. By combining both experiment and corpus studies with many speakers of a wide age range, this dissertation serves as an example of how to study the complete process of a tonogenetic sound change and its effect on intonation.

An implication for intonation literature is that this dissertation provides a pitch normalization method for a large number of subjects. When it comes to comparisons of different ages and gender groups, pitch normalization is a challenging issue. In this dissertation, I provide a new way of Hz-to-St conversion by using each speaker's own baseline, and the results show that the method is quite successful in normalizing pitch. This method can be readily applied in many other studies that investigate pitch differences and variation in different ages and gender groups.

Lastly, this dissertation has an implication for Korean linguistics literature in that it fully examines the tonogenetic change's effect on Seoul Korean intonation, answering questions that have not been previously investigated. This dissertation's results are highly consistent in the experiment as well as in the two corpus studies, providing a solid picture of the pitch contrast development. I believe this dissertation broadens our understanding of the tonogenetic change and Seoul Korean intonation.

Appendix A

Speakers' pitch range in the NIKL corpus

Individual speakers vary in terms of the pitch range of their voice. Some speakers use a wide pitch range, whereas others have a relatively monotonous voice. To estimate speakers' pitch range, I computed the median value of the absolute deviation from speakers' median pitch value (MAD) for each speaker in the NIKL corpus, which is a single numeric value that represents how far speakers' pitch values are apart from their own median pitch value. That is, this value can be considered as representing how broad a speaker's pitch range is.

This value was calculated as the following. After each speaker's median pitch value (Hz) was obtained, each f0 value in Hz was converted to a semitone value with the median f0 as the reference, using $\log_2$(f0/median f0)*12. The obtained semitone values were taken into absolute values, and the median value among the absolute deviations was obtained for each speaker. The table below shows the MAD value of all speakers in the NIKL corpus.

Male speakers						Femle speakers					
ID	YOB	MAD	ID	YOB	MAD	ID	YOB	MAD	ID	YOB	MAD
mz03	1930s	0.79	mv05	1970s	1.63	fz05	1930s	1.19	fy05	1950s	1.22
mz04	1930s	0.98	mv06	1970s	1.24	fz06	1930s	1.43	fy09	1950s	1.06
mz05	1930s	1.61	mv08	1970s	1.59	fy01	1940s	0.76	fy10	1950s	1.37
mz09	1930s	1.61	mv11	1970s	1.15	fy02	1940s	1.38	fy11	1950s	1.67
my01	1940s	2.37	mv12	1970s	1.8	fy06	1940s	0.8	fy12	1950s	1.47
my02	1940s	1.13	mv13	1970s	2.16	fy07	1940s	1.88	fy13	1950s	2.04
my03	1940s	1.35	mv15	1970s	1.7	fy08	1940s	0.94	fx08	1960s	3.12
my06	1940s	1.42	mv16	1970s	1.6	fy14	1940s	1.83	fx13	1960s	2.03
my09	1940s	1.3	mv17	1970s	1.41	fy16	1940s	1.38	fx20	1960s	2.61
my10	1940s	1.39	mv18	1970s	1.68	fy17	1940s	1.61	fv03	1970s	1.83
my11	1940s	0.87	mv19	1970s	2.47	fy18	1940s	1.6	fv05	1970s	1.29
mz01	1940s	1.29	mv20	1970s	1.63	fx01	1950s	1.63	fv06	1970s	1.11
mz02	1940s	1.5	mw01	1970s	1.97	fx02	1950s	1.34	fv10	1970s	1.8
mz06	1940s	1.59	mw02	1970s	1.99	fx03	1950s	2.08	fv11	1970s	2.29
mz07	1940s	1.21	mw03	1970s	1.81	fx04	1950s	1.6	fv12	1970s	1.92
mz08	1940s	0.88	mw04	1970s	1.76	fx05	1950s	2.03	fv13	1970s	2.1

my04	1950s	1.4	mw05	1970s	2.1	fx06	1950s	1.85	fv14	1970s	1.9
my05	1950s	1.18	mw06	1970s	1.95	fx07	1950s	1.61	fv15	1970s	1.43
my07	1950s	1.95	mw07	1970s	1.65	fx09	1950s	1.19	fv16	1970s	2.06
my08	1950s	1.26	mw08	1970s	1.39	fx10	1950s	1.12	fv20	1970s	1.88
mw13	1960s	1.45	mw09	1970s	1.66	fx11	1950s	1.85	fv01	1980s	1.9
mw14	1960s	1.9	mw10	1970s	1.75	fx12	1950s	1.41	fv02	1980s	2.19
mw15	1960s	1.72	mw11	1970s	1.23	fx14	1950s	1.05	fv04	1980s	1.7
mw16	1960s	1.37	mv04	1970s	1.48	fx15	1950s	1.58	fv07	1980s	2.54
mw17	1960s	1.34	mv02	1980s	1.47	fx16	1950s	1.42	fv08	1980s	1.89
mw18	1960s	1.54	mv07	1980s	1.41	fx17	1950s	1.21	fv09	1980s	1.97
mw19	1960s	1.3	mv09	1980s	1.63	fx18	1950s	1.53	fv17	1980s	1.55
mw20	1960s	1.5	mv10	1980s	0.81	fx19	1950s	1.4	fv18	1980s	1.63
mv01	1970s	2.39	mv14	1980s	1.73	fy03	1950s	1.16	fv19	1980s	1.81
mv03	1970s	1.91				fy04	1950s	1.72			

Table A.1: The median value of the absolute deviation from the median pitch value (MAD) for the 118 speakers in the NIKL corpus.

Appendix B

Target sentences of the corpus study

Sentence ID	IPA transcription and English translation
t09_s08	/kɨkʌ t͡sʰam kʰɨnilinɛ/ 'That's a big problem.'
t09_s10	/sinhatɨli kʌkt͡sʌŋsɨlʌn pʰjot͡sʌŋɨlo malhɛsʌjo/ 'With a look of worry, Liege subjects said.'
t09_s17	/kɨlɛsʌ t͡salanɨn t'ok'i kɨlimɨl katkosʌ put͡silʌnhi jukt͡silɨl hjaŋhajʌ hɛʌmt͡sʰjʌ katsɨpnita/ 'So, the turtle swam hard to the land with a picture of a rabbit.'
t09_s19	/t'asɨhan pomnalila sanɛnɨn ulkɨtpulkɨthan k'ottɨli pʰiʌitko sɛtɨli pankapkɛ t͡sit͡sʌkwiko itʌtʌjo/ 'It was a warm spring day with a riot of colors, and birds were singing for joy.'
t09_s25	/tʰok'iwa t͡salanɨn palo swipkɛ t͡sʰinkuka twɛʌtsʌjo/

	‘The rabbit became a friend of the turtle right away.’
t09_s29	/t͡salanɨn tʰok’ilɨl salsal k’wøʌtsʌjo/ ‘The turtle lured the rabbit out.’
t09_s32	/t͡salaɰi sokmaɨmɨl molɨnɨn tʰok’inɨn sʌllɛinɨn kasɨmɨl anko t͡sala ɰi tɨŋɛ ollatʰatsʌjo/ ‘The rabbit who didn’t realize the turtle’s real intention mounted the turtle with a leap of his heart.’
t09_s35	/tʰok’ika joŋkuŋɛ tot͡sʰakhat͡smat͡sa joŋwaŋnimɰi sinhaka natʰana tʰok’ilɨl patt͡sullo k’oŋk’oŋ mukʌtsʌjo/ ‘As soon as the rabbit arrived to the Sea God’s Palace, God’s subjects came to tie the rabbit with a rope.’
t09_s37	/t͡sʌlɨl ʌsʌ pʰulʌt͡susɛjo/ ‘Please let me go as soon as possible.’
t09_s43	/p’omnɛtɨsi t͡salaŋhanɨn t͡salaɰi jʌpʰɛsʌ tʰok’inɨn pʌlpʌlpʌl t’ʌlkiman hɛtsʌjo/ ‘The rabbit was trembling with fear while the turtle was showing off.’
t09_s44	/hat͡siman tʰok’iɛkɛ t͡soɨn k’øka sɛŋkaknatsʌjo/ ‘However, the rabbit came up with a good idea.’
t09_s56	/patatkaɛ tot͡sʰakhan tʰok’inɨn k’aŋt͡sʰuŋ t’wiʌnɛlimjʌ t͡salaɛkɛ malhɛtsʌjo/ ‘Hopping off on the seashore, the rabbit told to the turtle.’
t10_s08	/kɨlʌntɛ ʌnɨnal iʌtsʌjo/ ‘But it was one day.’
t10_s12	/t͡sʌnɨn sanjaŋk’unɛkɛ t͡s’otkiko itsʌjo/ ‘A hunter is chasing after me.’
t10_s13	/kajʌpkɛto sasɨmɨn otɨlotɨl t’ʌlko itsʌtsʌjo/

	‘The poor deer was trembling with fear.’
t10_s15	/t͡samsihuɛ sanjaŋk’uni s’iksi’kkʌlimjʌ twit͡s’ott͡sʰawatsʌjo/ ‘After a while, the hunter came after the deer, huffing and puffing.’
t10_s16	/sanjaŋk’unɨn sapaŋɨl tulipʌn kʌlitʌni namuk’unɛkɛ mulʌtsʌjo/ ‘After looking around for the deer, the hunter asked to the woodman.’
t10_s18	/namuk’unɨn sit͡sʰimilɨl t’uk t’ɛʌtsʌjo/ ‘The woodman pretended not to know about the deer.’
t10_s27	/kɨ t’ɛ at͡sʌs’ika sʌnnjʌɰi nalkɛosɨl han pul kamt͡sʰusɛjo/ ‘By then, hide the celestial robe of the fairy.’
t10_s29	/sasɨmɰi malɨl tɨtt͡sa namuk’unɨn nʌmu sinkihan iliʌsʌ nuni hwituŋkɨlɛ t͡sjʌtsʌjo/ ‘After hearing what the deer said, the woodman was surprized with his eyes wide open.’
t10_s32	/sasɨmɨn namuk’unɛkɛ mjʌtpʌnina tat͡simɨl hakonɨn supsokɨlo t’wiʌka pʌljʌtsʌjo/ ‘The deer ran into the woods after confirming it to the woodman for several times.’
t10_s34	/kʌkiɛnɨn t͡sʌŋmal alɨmtapko malkɨn jʌnmosi itsʌsʌjo/ ‘Indeed, there was a pond, which is very beautiful and clear.’
t10_s43	/sʌnnjʌnɨn palɨl toŋtoŋ kullʌtt͡siman nalkɛosɨl t͡sʰat͡sɨl suka ʌpsʌtsʌjo/ ‘The fairy stamped her feet with worry because she could not find her celestial robe.’
t10_s47	/nalkɛosɨl t͡sʰatkʌtɨn p’alli ollaonʌla/ ‘Come back as soon as you find your robe.’
t10_s48	/hont͡sa namɨn sʌnnjʌnɨn nʌmu sɨlpʰʌsʌ hɨnik’ʌ ulko itsʌsʌjo/

	'The fairy sobbed with grief, because she left alone.'
t10_s51	/t͡sʰuusiltʰɛntɛ usʌn i osilato ipɨsɛjo/ 'You must feel cold, why don't you wear this garment first?'
t10_s53	/sʌnnjʌnɨn namuk'uni kʌnnɛt͡sunɨn osɨl ipko hanɨn su upsi namukunɨl t'alakasʌ salkɛ tøʌtt͡sijo/ 'The fairy ended up living with the woodman after taking on the clothes the woodman gave her.'
t10_s69	/namuk'unɨn t'aŋpatakɛ tʰʌls'ʌk t͡sut͡sʌant͡sa hansumɨl swimjʌ sasɨmi putʰakhan malɨl sɛŋkakhɛtsʌjo/ 'The woodman flopped down on the floor and reminded what the deer asked him to do.'
t11_s02	/ʌmʌnika amuli tallɛto pot͡sʰɛmjʌnsʌ ulɨmɨl kɨt͡sʰit͡si anatsɨpnita/ 'The child kept crying even though his mother tried to soothe him.'
t11_s26	/ʌmmanɨn ʌlinɛlɨl tallɛtaka t͡sit͡sʰjʌtsɨpnita/ 'The mother was tired out from soothing her child.'
t11_s29	/holaŋinɨn t͡sʰaŋmun jʌpɛsʌ nɨktɛlan malɨl tɨtko piŋkɨlɛ usʌtsɨpnita/ 'The tiger smiled after overhearing "wold" beside the window.'
t11_s34	/holaŋinɨn t͡saki ilɨmɨl malhanɨn solilɨl tɨtko hɨmt͡sʰit twilo mullʌsʌtsɨpnita/ 'The tiger stepped back with surprise after they called his name.'
t12_s14	/t͡sʰamkilɨm palɨko olawatt͡si hamjʌ toŋsɛŋɨn holaŋilɨl nolljʌt͡suʌtsɨpnita/ 'His sister teased the tiger, saying "We climbed up this tree by applying sesame oil."'
t12_s22	/ilɨl bon holaŋinto t'okkat͡sʰi hananimk'ɛ kitolɨl hɛtsʌjo/ 'After seeing this, the tiger also prayed to God.'

t12_s25	/holaŋinɨn kʰolɨl pʌllɨmkʌlimjʌ toŋat͡sulɨl tʰako onuilɨl t͡s'ot͡sʰakatsʌjo/ 'The tiger also climbed up with the rope, huffing and puffing, in order to chase after the siblings, huffing and puffing.'
t12_s27	/holaŋiɛkɛ nɛljʌt͡susin toŋat͡sulɨn s'ʌkɨn toŋat͡sul iʌtkʌtɨnjo/ 'The rope that God gave to the tiger was a rotten one.'
t13_s11	/witmaɨlɛsʌ pɛk'otnipʰi t'ʌlʌt͡sjʌ mulɨl t'ala t'ʌnɛljʌomjʌn ponɨn t͡sinakako jʌlɨmi t͡sʰat͡saopnita/ 'When petals of pear flowers fall down along the brook from an upper village, spring passes by and summer comes.'
t13_s28	/wittoŋnɛɛsʌ k'otnipʰi hɨllʌ nɛljʌomjʌn pomi katɨsi namunipʰi t'ʌnɛljʌomjʌn kaɨlto kapnita/ 'As if spring passes by when petals fall down from the upper village, fall passes by when leaves fall down.'
t14_s02	/kʰɨko t͡sakɨn t͡swapʰanwiɛ katkat͡si mulkʌnɨl s'aanoko t͡sinakanɨn itɨlɰi nunkilɨl k'ɨnta/ 'Displaying goods of various sizes attracts passersby's attention.'
t14_s05	/t͡sikt͡saŋsɛŋhwalɨl hanɨn toŋan hoŋpona t͡sʰulpʰankwankɛɰi ilɨl olɛttoŋan hɛon tʰʌjʌsʌ ʌnt͡sɛnkaputʰʌ t͡sipɨlo tolakanɨn kilɛ taɨmnal t͡sokanɨl mili sasʌ ponɨn kʌtsi sɨpkwani tøʌtta/ 'Because I have been working on public relations and publications for a long time, it became a habit to buy the next morning's newspaper on the way back home.'
t14_s12	/nanɨn nant͡sʰʌhan pʰjot͡sʌŋɨlo kokɛlɨl k'ɨtʌkjʌtta/ 'I nodded yes, with a troubled look.'
t14_s14	/kɨlʌtʌni at͡sumʌninɨn kukjʌt͡sin t͡sʰʌnwʌnt͡s'ali han t͡saŋɨl kʌnnɛt͡suʌtta/ 'Then, the old lady handed over a 1000-won bill to me.'

t16_s01	/t͡sinnunk'ɛpika ɨls'injʌnsɨlʌpkɛ nɛlinɨn sipilwʌl jʌnkɨk jʌnsipsikanɛ nɨt͡sɨn nanɨn ant͡sʌlput͡sʌl mothamjʌ t͡sʰalɨl molatta/ 'On one gloomy, sleeting day in November, I restlessly drove my car because I was late for the practice meeting of a play.'
t16_s06	/koŋsahanɨla pʰanoɨn kutʌŋiɛ p'at͡sjʌ pʌlin kʌsiʌtta/ 'One tire of my car fell into the hole people dug for construction.'
t16_s10	/t͡sʰaka p'at͡sjʌtsʌjo/ 'Did your car fall into the hole?'
t17_s08	/t'okkatʰɨn s'al t'okatʰɨn sotʰɨlo papɨl t͡siʌmʌkʌtnɨnte ʌt͡s'ɛsʌ hankuk salammani suŋnjuŋɨl mantɨlʌ mʌkʌtsɨlk'a/ 'Why did only Korean people make *Sungnyung* (a traditional Korean infusion made with boiled scorched rice) while Asians cook rice with the same rice and the same kind of pot?'
t18_s06	/hunhunhan jʌlkiwa kot͡sotwøn sʌt͡sʌŋi kjosilɛ katɨk t͡sʰako aitɨlɰi nuntoŋt͡sanɨn tʌuk pitnako t͡sʰoloŋt͡sʰoloŋ hɛt͡sinta/ 'When the classroom filled with enthusiasm in a warm and nice atmosphere, children's eyes become clearer and more limpid.'
t18_s09	/t͡sʰɛkpolɨl s'asɛjo hanɨn naɰi mali t'ʌlʌt͡sit͡sa kjosilɨn t'ʌtɨls'ʌkhɛt͡siko ʌsusʌnhɛt͡sinta/ 'When I say "Pack your book wrapper," the classroom becomes loud and cluttered.'
t18_s12	/kɨk'oli hato usɨwʌ tolasʌsʌ t͡sʰilpʰanɨl hjaŋhɛ utkonani k'alɨlɨlɨ pʰoksoka tʰʌt͡sinta/ 'There was a burst of laughter when I laughed for seeing that, looking back toward the blackboard.'
t18_s13	/p'oŋp'oŋ p'aŋt͡sʰa watsʌjo ha nɨnɨ ɨmsʌŋɛ jʌpɨl poni pak'ɛsɨɛ himʌlkʌn kɨpsik p'aŋɨl tamakat͡siko wasʌ nɛ apʰɛ nɛminta

	‘When I looked around, after hearing somebody says, “Bang, bang”, a student handed a loaf of bread over to me.’
t18_s16	/t͡sʰʌŋso kʌmjʌlɨl hal t’ɛjʌtta/ ‘That was when I was checking if the classroom was clean enough.’
t18_s21	/t͡sosaŋɛkɛ kananɨl mʌnt͡sʌ pɛun aitɨl/ ‘Children who learned poverty first from their ancestors.’
t18_s25	/siŋsiŋhako tʰɨntʰɨnhakɛ t͡salatolok simko kak’uʌja hal i sipt͡samokɛ t͡sʌŋsʌŋkwa t͡sihɛlɨl pat͡sʰiko sipta/ ‘I would like to devote my wisdom and sincerety to children, who I need to plant and cultivate to grow strong and fresh.’
t18_s26	/k’umsokɛsʌ kak’ɨm pukhanɰi ʌlinitɨlɨl mannapoko antʰak’awa kasɨm altʰaka k’ɛkon hanta/ ‘Sometimes I see children in North Korea in my dream and wake up, feeling heartbroken for them.’
t19_s03	/pʰisʌt͡siɰi sisʌli kɨlʌkʰo okokanɨn kjotʰoŋpʰjʌni kɨlʌkʰo tʌkuna kɨ kosɛ mojʌtɨn salamtɨli kɨlʌhata/ ‘Everything in summer vacation spots is such a pain, including facilities, transportation and traffic jams, and many people who also visit the places for vacation.’
t19_s11	/kɨlʌtaka t͡som tʌ mutʌwʌt͡simjʌn nɛ kohjaŋ t͡sant͡sanhan kɛulɨl t͡sʰat͡sa nɛljʌkanta/ ‘And then, if it gets too hot, I visit my hometown to spend a summer vacation in a brook.’
t19_s19	/pʰilamina soŋsali t͡sʰɛŋkilɨmt͡sɛŋi katʰɨn ʌt͡sin kɛul kokitɨli nɛ momɨl sɨt͡sʰjʌkakito hako t͡s’oŋ t͡s’oŋ t͡sutuŋilo t͡s’okito hako t͡saŋnanto t͡sʰinta/ ‘Fresh-water fish, such as pale chub, killifish, and pond loach, pass me by, pecking my body and playing around.’

t19_s34	/tøtolokimjʌn t͡sʌkkɛ mʌknɨnta/ 'I eat as little as possible.'

Table B.1: Target sentences of the corpus study

Bibliography

Bermúdez-Otero, Ricardo. 2007. Diachronic phonology. In *The Cambridge Handbook of Phonology*, ed. Paul de Lacy, 497–517. Cambridge: Cambridge University Press.

Boersma, Paul, and David Weenink. 2016. Praat: Doing phonetics by computer. [computer program]. `http://praat.org/`.

Chang, Charles B. 2007. The status of voicing and aspiration as cues to Korean laryngeal contrast. In *UC Berkeley Phonology Lab Annual Report*, 330–344.

Chang, Charles B. 2013. The production and perception of coronal fricatives in Seoul Korean. *Korean Linguistics* 15:7–49.

Cho, Hyesun. 2011. The timing of phrase-initial tones in Seoul Korean. *Phonology* 28:293–330.

Cho, Taehong, Sun-Ah Jun, and Peter Ladefoged. 2002. Acoustic and aerodynamic correlates of korean stops and fricatives. *Journal of Phonetics* 30:193–228.

Cho, Taehong, and Patricia Keating. 2001. Atriculatory and acoustic studies of domain-initial strengthening in Korean. *Journal of Phonetics* 29:155–190.

Cho, Young-mee Yu. 1987. Phrasal phonology of Korean. In *Harvard Studies in Korean Linguistics*, ed. Susumu Kuno et al., volume 2, 328–340. Cambridge, MA: Harvard University Press.

Choi, Hansook. 2002. Acoustic cues for the Korean stop contrast: Dialectal variation. In *Papers on Phonetics and Phonology: The Articulation, Acoustics, and Perception of Consonants (ZAS Papers in Linguistics 28)*, ed. Tracy A. Hall, 1–12. Berlin: ZAS.

Chung, Kook. 1980. Korean Neutralization. Doctoral Dissertation, University of Texas, Austin.

Diffloth, Gérard. 1989. Proto-Austroasiatic creaky voice. *Mon-Khmer Studies* 15:139–154.

Ewan, William G. 1976. Laryngeal behavior in speech. Doctoral Dissertation, University of California, Berkeley.

Fourgeron, Cécile, and Patricia Keating. 1997. Articulatory strengthening at edges of prosodic domains. *Journal of Acoustical Society of America* 101:3728–3740.

Halle, Morris, and Kenneth Stevens. 1971. A note on laryngeal features. *Quaterly Progress Reports, MIT Research Laboratory of Electronics* 101:198–212.

Han, Meiko S., and Raymond S. Weitzman. 1970. Acoustic features of korean /p, t, k/, /p', t', k'/ and /ph, th, kh/. *Phonetica* 22:112–128.

Hanson, Helen M. 2009. Effects of obstruent consonants on fundamental frequency at vowel onset in English. *Journal of Acoustical Society of America* 125:425–441.

Haudricourt, André-Georges. 1954. De l'origine des tons en vietnamien. *Journal Asiatique* 242.

Haudricourt, André-Georges. 1972. *Two-way and three-way splitting of tonal systems in some far- eastern languages*, 58–86. Bangkok: Central Institute of English Language.

Hirose, Hajime, Charles Y. Lee, and Tatsujiro Ushijima. 1974. Laryngeal control in Korean stop production. *Journal of Phonetics* 2:145–152.

Holliday, Jeffrey, and Eun Jong Kong. 2011. Dialectal variation in the acoustic correlates of Korean stops. In *Proceedings of the 17th ICPhS*, 878–881. Hong Kong, China: International Phonetic Association.

Hombert, Jean-Marie. 1978. Consonant types, vowel quality, and tone. In *Tone: A linguistic survey*, ed. Victoria A. Fromkin, 77–112. New York: Academic Press.

Hombert, Jean-Marie, John J. Ohala, and William G. Ewan. 1979. Phonetic explanations for the development of tones. *Language* 55:37–58.

Honda, Kiyoshi, Hirojuki Hirai, Shinobu Masaki, and Yasuhiro Shimada. 1999. Role of vertical larynx movement and cervical lordosis in F0 control. *Language and Speech* 42.

Hyman, Larry M. 1978. Historic tonology. In *Tone: A linguistic survey*, ed. Victoria A. Fromkin, 257–269. New York: Academic Press.

Hysop, Gwendolyn. 2009. Kurtöp tone: A tonogenetic case study. *Lingua* 119:827–845.

Jun, Sun-Ah. 1993. The Phonetics and Phonology of Korean Prosody. Doctoral Dissertation, Ohio State University.

Jun, Sun-Ah. 1996. The influence of the microprosody on the macroprosody: A case of phrase initial strengthening. In *UCLA Working Papers in Phonetics*, volume 92, 97–116.

Jun, Sun-Ah. 1998. The Accentual Phrases in the Korean prosodic heirarchy. *Phonology* 15:189–226.

Jun, Sun-Ah. 2000. Korean ToBI, Version 3. In *UCLA Working Papers in Phonetics*, volume 99, 149–173.

Jun, Sun-Ah. 2006. Intonational phonology of Seoul Korean revisited. In *Japanese-Korean Linguistics*, ed. T. K. Vance and K. Jones, volume 14, 15–26. Stanford: CSLI.

Jun, Sun-Ah. 2011. Prosodic marking of complex NP focus, syntax, and the pre-/post-focus string. In *Proceedings of the 28th West Coast Conference on Formal Linguistics*, ed. Mary Byram Washburn, Katherine McKinney-Bock, Erika Varis, Ann Sawyer, and Barbara Tomaszewicz, 214–230. Somerville, MA: Cascadilla Proceedings Project.

Jun, Sun-Ah, Mary E. Beckman, and Hyuck-Joon Lee. 1998. Fiberscopic evidence for the influence on vowel devoicing of the glottal configurations for Korean obstruents. In *UCLA Working Papers in Phonetics*, volume 96, 43–68.

Jun, Sun-Ah, and Jihyeon Cha. 2015. High-toned [il] in Korean: Phonetics, intonational phonology, and sound change. *Journal of Phonetics* 51:93–108.

Kagaya, Ryohei. 1974. A fiberscopic and acoustic study of the Korean stops, affricates, and fricatives. *Journal of Phonetics* 2:161–180.

Kang, Kyoung-Ho, and Susan G. Guion. 2008. Clear speech production of Korean stops: Changing phonetic targets and enhancement strategies. *Journal of Acoustical Society of America* 124:3909–3917.

Kang, Ongmi. 1992. Korean Prosodic Phonology. Doctoral Dissertation, University of Washington.

Kang, Yoonjung. 2014. Voice Onset Time merger and development of tonal contast in Seoul Korean stops: A corpus study. *Journal of Phonetics* 45:76–90.

Karpiński, Maciej. 2006. Struktura i intonacja polskiego dialogu zadaniowego (The structure and intonationa of Polish task-oriented dialogues). Poznań: Wydawnictwo Naukowe UAM.

Karpiński, Maciej, and Emilia Szalkowska-Kim. 2012. On intonation of questions in Korean and Polish task-oriented dialogues. *Speech and Language Technology* 14.

Kim, Chin-Woo. 1979. Neutralization in Korean revisited. *Studies in the Linguistic Sciences* 9:147–155.

Kim, Hyunsoon. 1997. The Phonological Representation of Affricates: Evidence from Korean and other languages. Doctoral Dissertation, Cornell University.

Kim, Hyunsoon. 1999. The place of articulation of the Korean plain affricates revisited. *Journal of East Asian Linguistics* 8:313–347.

Kim, Hyunsoon. 2001. The place of articulation of the Korean plain affricate in intervocalic position: an articulatory and acoustic study. *Journal of the International Phonetic Association* 31:229–257.

Kim, Hyunsoon, and Allard Jongman. 1994. Complete neutralization of manner of articulation in Korean. In *Working papers of the Cornell Phonetics Laboratory*, volume 9, 137–154.

Kim, Mi-Ran. 1994. Acoustic characteristics of korean stops and perception of english stop consonants. Doctoral Dissertation, University of Wisconsin-Madison.

Kim, Mi-Ryoung, Patrice Speeter Beddor, and Julie Horrocks. 2002. The contribution of consonantal and vocalic information to the perception of korean initial stops. *Journal of Phonetics* 30:77–100.

Kim, Mi-Ryoung, and San Duanmu. 2004. “Tense” and “lax” stops in Korean. *Journal of East Asian Linguistics* 13:59–104.

Kim, Midam. 2004. Correlation between VOT and F0 in the perception of Korean stops and affricates. In *Interspeech 2004—8th International Conference on Spoken Language Processing*, 49–52. Jeju, Korea.

Kim-Renaud, Young-Key. 1974. Korean Consonantal Phonology. Doctoral Dissertation, University of Hawaii.

Kingston, John. 2005. The phonetics of Athabaskan tonogenesis. In *Athabaskan prosody*, ed. Sharon Hargus and Keren Rice, 137–184. Amsterdam & Philadelphia: John Benjamins.

Kingston, John. 2011. Tonogenesis. In *The blackwell companion to phonology*, ed. Marc van Oostendorp, Colin J. Ewen, Hume Elizabeth, and Keren Rice, 2304–2333. Blackwell Publishing.

Kingston, John, and Randy L. Diehl. 1994. Phonetic knowledge. *Language* 70:419–454.

Ko, Eon-Suk, Na-Rae Han, Alexandra Caravan, and George Zipperlen. 2003. Korean Telephone Conversations Speech. Philadelphia: Linguistic Data Consortium.

Ko, Insung. 2013. The Articulation of Korean Coronal Obstruents: Data from Heritage Speakers and Second Language Learners. Doctoral Dissertation, University of Hawaii.

Labov, William. 1990. The intersection of sex and social class in the course of linguistic change. *Language Variation and Change* 2:205–254.

Labov, William. 2001. *Principles of linguistic change*, volume 2. Oxford: Blackwell Publishing.

Lee, Hyuck-Joon. 1999. Tonal realization and implementation of the Accentual Phrase in Seoul Korean. Master's thesis, University of California, Los Angeles.

Lee, Hyunjung, and Allard Jongman. 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* 42:145–169.

Lee, Ki-Moon. 1961. *Kwukosakesul (an introduction to the history of the korean language)*. Seoul: Tower Press.

Lee, Yong-cheol. 2012. Prosodic correlation between the focusing adverb *ozik* 'only' and focus/GIVENness in Korean. *Journal of Speech Sciences* 2:85–111.

Lee, Yong-cheol. 2015. Prosodic focus within and across languages. Doctoral Dissertation, University of Pennsylvania.

Lee, Yong-cheol, Sunghye Cho, and Mark Liberman. 2015. An interaction effect between tonogenesis and focus in Seoul Korean. Paper presented at the 39th Penn Linguistics Conference (PLC), University of Pennsylvania, Philadelphia, Pennsylvania.

Lee, Yong-cheol, and Yi Xu. 2010. Phonetic realization of contrastive focus in Korean. In *Proceedings of Speech Prosody 2010*, ed. Mark Hasegawa-Johnson.

Lehiste, Ilse, and Gonden E. Peterson. 1961. Some basic considerations in the analysis of intonation. *Journal of Acoustical Society of America* 33:419–425.

Liberman, Mark, and Janet Pierrehumbert. 1984. Intonational invariance under changes in pitch range and length. In *Language sound structure*, ed. R. Oehrle M. Aronoff, 157–233. Cambridge, MA: MIT Press.

Lisker, Leigh, and Arthur S. Abramson. 1964. A cross-language study of voicing in initial stops: Acoustic measurements. *Word* 20:384–422.

Löfqvist, Anders, Thomas Baer, Nancy S. McGarr, and Robin Seider Story. 1989. The crico-thyroid muscle in voicing control. *Journal of Acoustical Society of America* 85:1314–1321.

Martin, Samuel E. 1951. Korean phonemics. *Language* 27:519–533.

Matisoff, James A. 1973. Tonogenesis in Southeast Asia. In *Consonant types and tone*, ed. Larry M. Hyman, Southern California Occasional Papers in Linguistics, 71–95. Los Angeles: Linguistics Program, University of Southern California.

Mohr, Buurckhard. 1971. Intrinsic variations in the speech signal. *Phonetica* 23:65–93.

Oh, Eunjin. 2011. Effects of speaker gender on voice onset time in Korean stops. *Journal of Phonetics* 39:59–67.

Ohala, John J. 1973. The physiology of tone. In *Consonant types and tone*, ed. Larry M. Hyman, 1–14. Linguistics Program, University of Southern California.

Ohala, John J. 1978. Production of tone. In *Tone: A linguistic survey*, ed. Victoria A. Fromkin, 5–39. New York: Academic Press.

Pandeli, Helen. 1993. The Articulation of lingual consonants: An EPG Study. Doctoral Dissertation, Cambridge University.

Perkins, Jeremy, and Seunghun Julio Lee. 2010. Korean affricates and consonant-tone interaction. In *Proceedings of the 6th Workshop on Altaic Formal Linguistics (MIT Working Papers in Linguistics 61)*, ed. Hiroki Maezawa and Azusa Yokogoshi, 277–286.

Pierrehumbert, Janet. 1980. The Phonology and Phonetics of English Intonation. Doctoral Dissertation, Massachusetts Institute of Techonology.

Pierrehumbert, Janet, and Mary E. Beckman. 1988. *Japanese tone structure*. MIT Press.

Poser, William J. 1984. The Phoneitcs and Phonology of Tone and Intonation in Japanese. Doctoral Dissertation, Massachusetts Institute of Techonology.

Ratliff, Martha. 1997. Discussion of tone and register papers. Paper presented at the 6th Annual Workshop on Comparative Linguistics: Prosody and Language Change, Wayne State University.

Ross, Malcolm D. 1993. Tonogenesis in the North Huon Gulf Chain. *Oceanic Linguistics Special Publications* 24.

Shih, Chi-lin. 1988. Tone and intonation in Mandarin. In *Working papers, Cornell Phonetics Laboratory*, volume 3, 83–109.

Shin, Jiyoung. 1996. Consonantal Production and Coarticulation in Korean. Doctoral Dissertation, London University.

Silva, David J. 1992. The Phonetics and Phonology of stop lenition in Korean. Doctoral Dissertation, Cornell University.

Silva, David J. 2006. Acoustic evidence for the emergence of tonal contrast in Contemporary Korean. *Phonology* 23:287–308.

Suwilai, Premsrirat. 2004. Register complex and tonogenesis in Khmu dialects. *Mon-Khmer Studies* 34:1–17.

The National Institue of the Korean Language. 2005. *A Speech Corpus of Reading-style Standard Korean. [DVDs]*. Seoul, Korea: The National Institute of the Korean Language.

Thompson, Laurence C. 1987. *A Vietnamese Reference Grammar*. Hawaii: University of Hawaii Press.

Thurgood, Graham. 1992. From atonal to tonal in Utsat (A Chamic Language of

Hainan). In *Proceedings of the 18th Annual Meeting of the Berkeley Linguistics Society: Special Session on the Typology of Tone Languages*, 145–156.

Thurgood, Graham. 2002. Vietnamese and tonogenesis: revising the model and the analysis. *Diachronica* 19:333–363.

Thurgood, Graham. 2007. Tonogenesis revisited: Revising the model and the analysis. In *Studies in Tai and Southeast Asian Linguistics*, ed. J. G. Harris, S. Burusphat, and J. E. Harris, 263–291. Bangkok: EK Phim Thai Co.

Whalen, D. H., and Andrea G. Levitt. 1995. The universality of intrinsic F0 of vowels. *Journal of Phonetics* 23:349–366.

Wright, Jonathan D. 2007. Laryngeal Contrast in Seoul Korean. Doctoral Dissertation, University of Pennsylvania.

Xu, Yi. 1993. Contextual Tonal Variation in Mandarin Chinese. Doctoral Dissertation, The University of Connecticut.

Yoon, Tae-Jin, and Yoonjung Kang. 2012. A forced-aligned-based study of declarative sentence-ending 'ta' in Korean. In *Proceedings of the 6th International Conference on Speech Prosody 2012*. Shanghai: Tongji University Press.