

KOREAN STOP VOT PRODUCTION BY HERITAGE SPEAKERS IN THE LANGUAGE
CLASSROOM

Melinda Joy Johnson

A thesis submitted to the faculty at the University of North Carolina at Chapel Hill in
partial fulfillment of the requirements for the degree of Master of Arts in the
Linguistics Department in the University of North Carolina at Chapel Hill.

Chapel Hill
2015

Approved by:

Elliott Moreton

Jennifer Smith

Misha Becker

© 2015
Melinda Joy Johnson
ALL RIGHTS RESERVED

ABSTRACT

Melinda Joy Johnson: Korean stop VOT production by heritage speakers in the language classroom
(Under the direction of A. Elliott Moreton, Ph.D.)

Heritage speakers of Korean were recorded producing Korean and English stops and compared to the results of those produced by their non-heritage speaker peers in order to better understand the roles of first language attrition and language transfer, as well as the validity of the permanence hypothesis proposed in the Native Language Neural Commitment theory. The permanence hypothesis weighs age of first language acquisition as highly important in that after acquiring their sound categories, native speakers will persist in using those parameters throughout their lives. This study found that while the permanence hypothesis does have some credence in the participants' Korean aspirated stops, which were produced like their childhood input (i.e. their parents'), it also encountered signs of first language attrition in the form of their Δ VOT, which was not produced like their parents', and language transfer from Korean to English in the voiceless aspirated stop category, which was produced like a Korean stop in both languages. These results demonstrate the complexity of language acquisition, retention, and language transfer in bilinguals, and in heritage speakers in particular.

ACKNOWLEDGEMENTS

First, I would like to acknowledge my advisor, Dr. Elliott Moreton, without whose direction, this thesis would never have happened. He offered insightful advice and guidance, especially during theoretical discussions used in this thesis. I would also like to offer my gratitude and acknowledgement to my council, Dr. Jennifer Smith and Dr. Misha Becker, for their guidance and the opportunities they gave me to bring forth my study to various research groups in the department for feedback. The members of P-Side and AcqLab, students, faculty and staff of UNC-CH, also deserve my acknowledgement for their feedback on my presentations. This thesis began from a study conducted by Yina Ma and myself in November 2013. Without that initial study, this study would have not occurred. Another contributor I would like to acknowledge is the Odum Institute and in particular, Christopher Wiesen for his statistical analysis. I would also like to thank my participants, who were all very kind and willing participants. Finally I would like to acknowledge the support given to me by my family, friends, boyfriend and peers. I especially want to acknowledge the support of Laura Barnes, Stephanie Gardner, Joshua Poole, Melissa Johnson, Roy Johnson, Emily Meong, Breanne Jones and Rachel Broad. This group of people was supportive during this process in various ways.

TABLE OF CONTENTS

LIST OF FIGURES.....	vii
CHAPTER 1: INTRODUCTION.....	1
CHAPTER 2: BACKGROUND LITERATURE.....	3
2.1. Korean Stops.....	3
2.2. Native Language Neural Commitment.....	6
2.3. First Language Attrition.....	7
2.4. Language Transfer.....	8
2.4.1. Chang (2012) and Phonetic Drift.....	8
2.4.2. Flege et al. (2003) and Phonetic Categories.....	9
CHAPTER 3: METHODOLOGY.....	10
3.1. Questions.....	10
3.1.1. Big Question.....	10
3.1.2. Specific Questions.....	10
3.2. Hypothesis.....	10
3.2.1. Possible Outcomes.....	11
3.2.1.1. If Hypothesis Is Incorrect.....	11
3.2.1.2. If Hypothesis Is Correct.....	12
3.2.1.2.1. VOT.....	12
3.2.1.2.2. Δ VOT.....	16
3.3. Experiment.....	17

3.3.1. Participants.....	17
3.3.2. Materials.....	18
3.3.3. Design.....	19
3.3.3.1. Korean Target Sounds.....	19
3.3.3.2. English Target Sounds.....	21
3.4. Measurements and Analyses.....	22
CHAPTER 4: RESULTS AND DISCUSSION.....	25
4.1. Korean VOT by All Speaker Groups.....	25
4.2. Statistical Analysis of All Speaker Groups.....	27
4.3. English VOT by Heritage and Non-Heritage Speaker Groups.....	28
4.4. Heritage Speaker English and Korean VOT.....	29
4.5. Instructor VOT.....	32
CHAPTER 5: CONCLUSION.....	36
REFERENCES.....	38

LIST OF FIGURES

Figure 1: Bilabial minimal triplets.....	3
Figure 2: Hypothetical Heritage Results (Numbers based on Klatt (1975) and Silva (2006))	12
Figure 3: Hypothetical Heritage Results if Categories Merged (Flege et al. 2003) (Numbers based on Klatt (1975) and Silva (2006)).....	14
Figure 4: Hypothetical Heritage Results if Categories dissimilate (Flege et al. 2003)(Based on Klatt (1975) and Silva (2006))	15
Figure 5: Hypothetical Heritage Results if Categories Merged (Flege et al. 2003) (Numbers based on Klatt (1975) and Silva (2006)).....	16
Figure 6: Korean Tokens Used in Experiment with Transcription and Gloss	20
Figure 7: Korean Phrase Presentation Example.....	20
Figure 8: Korean Phrase Gloss	21
Figure 9: English Tokens Used in Experiment with Transcription and Gloss.....	22
Figure 10: English Phrase Presentation Example.....	22
Figure 11: An example of the landmarks used for measurement, using Subject 11's first Korean [t] repetition (Boersma 2001).....	23
Figure 12: An example of the landmarks used for measurement, using Subject 11's first Korean [t ^h] repetition (Boersma 2001).....	24
Figure 13: Korean VOT of Aspirated and Lax Stops Produced by All Speaker Groups. (Error bars based on Upper/Lower Values of a mixed-design Estimates).....	25
Figure 14: Statistical Results from a mixed-design.....	27

Figure 15: English VOT Results by Heritage and Non-Heritage Speaker Groups
(Error bars based on Upper/Lower Values of a mixed-design Estimates)..... 29

Figure 16: English and Korean VOT Results Produced by Heritage Speaker Group
(Error bars based on Upper/Lower Values of a mixed-design Estimates)..... 30

Figure 17: Korean Aspirated and Lax VOT Produced by Instructor Group (Error bars
based on Upper/Lower Values of a mixed-design Estimates) 33

CHAPTER 1: INTRODUCTION

Heritage speakers present a unique opportunity to better understand language systems, their acquisition, interactions, and attrition. Employing this opportunity, the purpose of this study is to test theories and phenomena in phonetic category acquisition and retention using an acoustic, empirical analysis of Korean heritage speakers' stop production of Korean and English.

Heritage speakers are defined in current literature as "bilinguals who usually come from immigrant and/or ethnic minority backgrounds" (Albrini 2014). These speakers grow up speaking the language of their parents but then shift to the dominant, host community's language, usually when entering school. For example, this study specifically looked at Korean American heritage speakers. More specifically, this study looked at second generation Korean Americans, who, while their parents were adult immigrants, have lived their whole lives in the United States and are enrolled in Korean language classes in college.

As mentioned above, heritage speakers provide an opportunity to better understand cross-linguistic theory. Historically, theoretical linguistics has been dominated by monolingual study. This allowed linguists to found theories since monolinguals, while having complicated linguistic systems, are much less complicated than bilinguals or multilinguals, who have multiple linguistic systems. Now that a foundation has been established and then tested by bilingual studies, it is

time to add the more complex bilinguals to the realm of theoretical linguistics to see if current theories hold true (Benmamoun et al. 2013). This study will look into the roles of first language attrition, language transfer, and the permanence hypothesis of the theory of Native Language Neural Commitment in heritage speakers' phonetic systems in order to better understand and test said hypothesis and phenomena.

CHAPTER 2: BACKGROUND LITERATURE

2.1. Korean Stops

This study focused on Korean stop production by heritage, non-heritage, and instructor speakers. These stops are a rare triplet of voiceless stops across three places of articulation: bilabial, alveolar, and velar. And, they are named lax, aspirated, and tense: [p, p^h, p*, t, t^h, t*, k, k^h, k*]. These three voiceless stops are phonemically different, as demonstrated in the minimal pairs below.

	Lax	Aspirated	Tense
Transcription	[pul]	[p ^h ul]	[p'ul]
Gloss	'fire'	'grass'	'horn'

Figure 1: Bilabial minimal triplets

Acoustically, lax stops are characterized as having little aspiration or vocal fold tension in their production. While word-medially they are voiced, in word-initial position they are voiceless and are considered to be voiceless underlyingly (Cho et al. 2002, Silva 2006). Aspirated stops differ from lax stops in that they are produced with a puff of air called aspiration, and in that lax is debated to have a low tone underlyingly, and aspirated has a high one (Silva 2006). Acoustically, these tones, manifesting as fundamental frequency (F0) differ on average ~80Hz (Kim 2012). Tense stops are marked by tension of the vocal folds, which causes a long closure period during production, as well as a salient burst (Cho et al. 2002). This

study focused more specifically on word-initial lax and aspirated stops because they are undergoing acoustic change in Contemporary Standard Korean.

While Korean voiceless stops have long been considered to not have lexical tone, acoustic evidence has been presented to propose that Standard Korean has undergone tonogenesis in the last 80 years, leading to Contemporary Standard Korean having a lexical tone in the following vowel in lax and aspirated stops (Kang and Han 2013, Kang and Nagy 2013, Kim 2012, Silva 2006) due to a lessening of the differences between the lax and aspirated stops' voice onset times (VOTs), an acoustic parameter that is the duration between the release of a stop consonant and the beginning of vocal fold vibration. However, in the past, VOT was how these lax and aspirated phones were considered acoustically different (Silva 2006).

Silva (2006) conducted an empirical study comparing the VOTs of a spectrum of Seoul speakers born from the 1930s until the 1980s and later (as long as they were over the age of eighteen at the time of data collection). He found that the speakers born before 1965 had a larger difference in lax and aspirated stop VOT values than those born after 1965, who actually have collapsed the durations to the point of having little to no VOT difference. Silva (2006) thus proposes that a second acoustic parameter must exist for aspirated and lax stops, and that it must be lexical since aspirated and lax stops are both phonemes.

This second parameter is fundamental frequency (F0), which is the acoustic manifestation of the underlying tone feature. Lax and aspirated stops have longitudinally had different F0s in the following vowels, however, with VOT being a consistently different and salient acoustic parameter, F0 was considered non-lexical.

Due to the collapse of VOT differences, F0 has gained importance in Contemporary Standard Korean. This is called tonogenesis (Kang and Han 2013, Silva 2006).

While the existence of the F0 is important to this study, the main focus will be on the different possible VOT values of lax and aspirated stops in isolation and in comparison to each other (i.e. the difference between them). However, understanding that Korean has been undergoing a phonetic change for the last 3 to 5 generations is important for this study because it means that several different VOT values can occur (expanded upon in Section 3.2.1.2.). In particular, the possible and actual outcomes of Korean stop production by young instructors (born after 1965), heritage speaking adult students, and non-heritage speaking adult students will be the focus of this study.

Kang and Nagy (2012) began to look into whether this VOT merger of Contemporary Standard Korean is occurring in heritage speakers in Toronto, Canada. Using the Heritage Language Documentation Corpus (Nagy 2009, 2011, Kang and Nagy (2013) studied the VOT values from sociolinguistic interviews with heritage speakers. Their results showed that Toronto heritage speakers are indeed following the Contemporary Standard Korean trend. Thus, this study anticipated that heritage speakers in the United States might also be following this trend, however, considering possible interrupted acquisition and attrition, as well as language transfer, the trend might not have progressed as much as that in Toronto among these participants. The participants in the Toronto study were chosen based on being proficient enough in Korean to hold an hour-long interview in the language. These participants have varying levels of proficiency, being students in

Korean college classes and proficiency level not being held constant, but rather in a range of late beginner to intermediate, which will also affect their results in comparison to the Toronto participants.

2.2. Native Language Neural Commitment

The Native Language Neural Commitment (NLNC) theory claims that by eleven-months, infants' neural pathways are dedicated to their native language patterns and interfere with non-native language patterns (Benmamoun et al. 2013, Kuhl 2004). This commitment to native-language patterns means that, once commitment is complete, phonological learning is no longer needed for future language use and the period of phonological sensitivity is over. This means that during second language learning as adults, sensitivity to sound differences and parameters that do not exist in the speakers' native language are not perceived well (Benmamoun et al. 2013, Kuhl 2004, Zang et al. 2005). From this theory, two hypotheses arise when considering sequential bilinguals: permanence and contingency.

The permanence hypothesis states that, "once commitment has taken place in each optimal period, the resources dedicated to the original language cannot be reassigned, and the relevant knowledge persists throughout life" (Benmamoun et al. 2013). For this study, the permanence hypothesis is important since it will be tested in comparison to first language attrition and language transfer (discussed in 2.3. and 2.4.), which are factors in language acquisition and retention that have opposite effects to the permanence hypothesis due to the idea that continued exposure is

necessary for knowledge persisting throughout life. In regards to sequential bilinguals NLNC also has a contingency hypothesis.

Benmamoun et al. (2013) defines the contingency hypothesis as, “the persistence of learning is contingent upon continued exposure to the language; if this input ceases, reorganization of the resources can occur, optimizing the neural system to another language.” This hypothesis follows the proposals of this study much more closely than the permanence hypothesis. However, a purpose of this study is to look at these two hypotheses and factors such as first language attrition and language transfer in heritage speakers of Korean stop production in order to see if empirical, acoustic data follows the permanence or the contingency hypotheses.

2.3. First Language Attrition

First language attrition is the loss, deterioration, or reduced accessibility of linguistic knowledge over time, usually due to decreased exposure of the first language (Schmidt 2013). This decreased exposure is due to the acquiring and then domination of a second language. For example, a Korean speaker moves to the United States and, isolated from other Korean speakers, and surrounded by English speakers, the Korean speaker is no longer exposed to nor uses Korean, but rather uses their second language, English, in day-to-day interactions out of necessity. As the second language becomes the dominant language in their lives, their first language attrites out of neglect and lack of exposure (Benmamoun et al. 2013, Schmidt 2013). For heritage speakers, this phenomenon occurs in their first language to varying degrees depending on how much of their heritage language they

are exposed to throughout their lives. Strong heritage language speaking communities usually aid in heritage language retention (Albrini 2014).

In the case of this study, the heritage speakers in question spoke Korean until elementary school, in which they had to speak English. Finally, these speakers are enrolled in university Korean language classes, exposed to Korean stops that are different from the Korean spoken to and by them in infancy and childhood. Thus, possibly three languages are interacting, though the Korean from infancy might have attrited. This study aims to find out if it has indeed attrited or if the permanence hypothesis is at least partially correct. To do this, the phonetic categories of Korean heritage speakers will be analyzed acoustically in the manner of Kim (2012) and Silva (2006). First language attrition has several symptoms, including incomplete acquisition and second to first language transfer. This makes first language attrition and language transfer effects in heritage speakers' production impossible to separate as they are connected.

2.4. Language Transfer

When a speaker is acquiring or has acquired multiple languages, the linguistic systems can interact, causing language transfer (Chang 2013, Kim 2012, Flege et al. 2003). In this study, phonetic categories and the language transfer involved in their manifestations as shown through stop production are a major focus in order to better understand heritage speaker first and second language interactions.

2.4.1. Chang (2013) and Phonetic Drift

Chang (2013) performed an acoustic study to test English native speakers learning Korean and how Korean affects their English stop production. Chang recorded novice and advanced participants over the course of five weeks and found that the novice speakers' English was affected in the form of phonetic drift, a phenomenon where the phonetic parameters systematically change. Second languages can affect first languages and vice versa. Chang's (2013) study is important for this study because it shows language transfer in the classroom, as well as second to first language transfer, which are both things that this study will be analyzing. This study might encounter phonetic drift or the merging of phonetic categories between languages as hypothesized by Flege et al. (2003).

2.4.2. Flege et al. (2003) and Phonetic Categories

Flege has extensively studied phonetic category formation and drift due to the interactions between multiple languages acquired. These studies of bilinguals have led Flege to hypothesize that 4 phonetic outcomes are possible when sounds from different languages are similar phonetically. This is important because Korean and English stops as produced by the heritage speakers might experience these possibilities due to their phonetic similarities. Flege's hypotheses predict that English and Korean stops that are similar might merge, becoming one stop unlike Korean or English, or diverge, becoming extremes so that they are 2 exaggeratedly different stops, or just one stop that is Korean-like or English-like might form (Flege et al. 2003).

CHAPTER 3: METHODOLOGY

3.1. Questions

3.1.1. Big Question

When recorded producing Korean and English stops, will heritage speakers who are enrolled in university Korean language adult classes demonstrate Korean and English stops that are completely native-like or will varying inputs (i.e. parents' Korean, peers' English, instructors'/course materials' Korean), first language attrition, and language transfer affect their stop phonetic categories, making them un-native-like, or will the permanence hypothesis hold true, which means that their Korean phonetic knowledge from childhood cannot be altered and persists throughout life?

3.1.2. Specific Questions

1. Will the heritage speakers' VOT and $\Delta VOT_{(asp-lax)}$ for Korean stop production trend like their parents' original input?
2. Will the heritage speakers' Korean and English stop categories affect one another, and if so, how?
3. How will the heritage speakers perform compared to their non-heritage peers?
4. Will heritage speakers follow the overlapping $VOT_{asp/lax}$ trend like the Toronto heritage speakers (Kang and Nagy 2013)?

3.2. Hypothesis

Korean heritage speakers in the adult Korean language classroom will produce Korean stops that are near native-like but not entirely native-like due to varying inputs, first language attrition, and language transfer. There are several ways in which this near native production can manifest, as outlined below in Section 3.2.1.2. They will be more native-like than their non-heritage Korean classroom peers since heritage speakers are known to outperform non-heritage speakers (Benmamoun et al. 2013).

3.2.1. Possible Outcomes

3.2.1.1. If Hypothesis Is Incorrect

If the hypothesis is incorrect, there is only one possible outcome to the experiment being used to test the hypothesis (detailed in 3.3.). The data collected will be the VOT of the Korean lax and aspirated stops and the English voiced and voiceless aspirated, as well as the difference in VOTs ($\Delta VOT_{(asp-lax/voiced-voiceless)}$), in the manner of Silva (2006). These data will be plotted in a scatterplot, and if the results indicate that the hypothesis is incorrect, the plot will look like Figure 2.

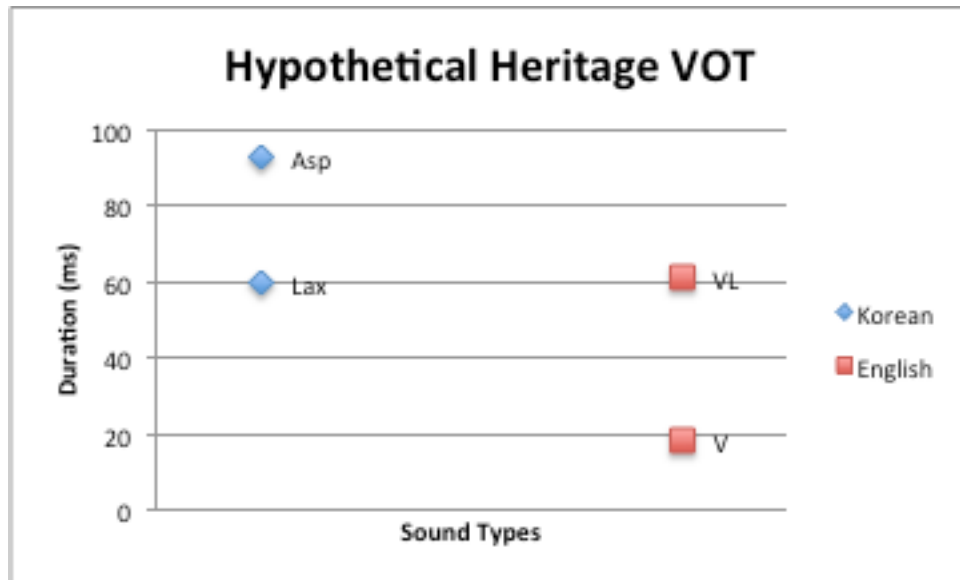


Figure 2: Hypothetical Heritage Results (Numbers based on Klatt (1975) and Silva (2006))

This figure is based on Klatt’s (1975) English and Silva’s (2006) Korean VOT results. The heritage speaker input would have been their parents, so the Korean VOT values and the Δ VOTs will mimic a Korean native speaker born between 1952 and 1968 if no language attrition or language transfer has occurred.

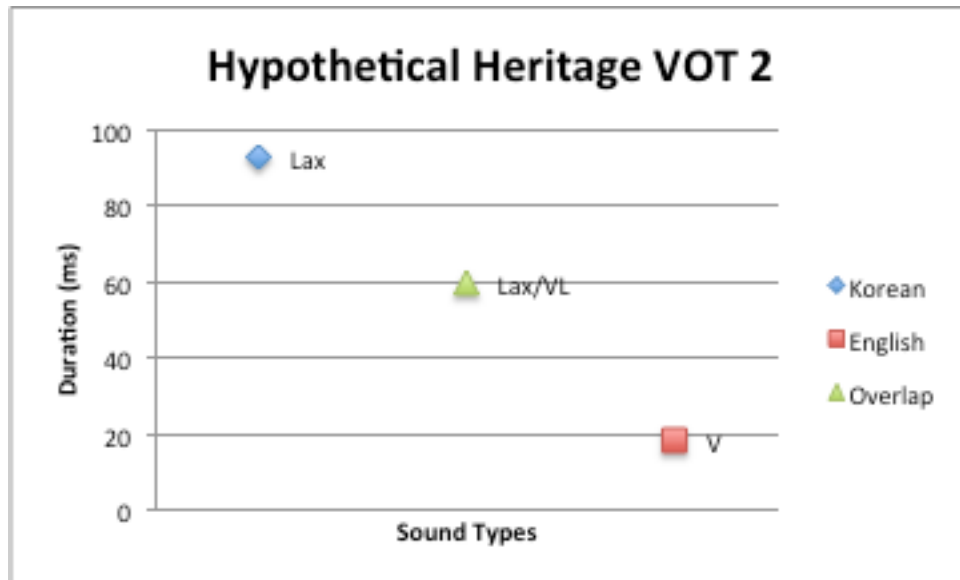
3.2.1.2. If Hypothesis Is Correct

If the hypothesis is correct, there are several possible ways in which the first language attrition and language transfer can affect the results. Firstly, there are 2 different types of analyses: VOT and Δ VOT. First language (L1) attrition and language transfer can affect one, the other, or both. Also, Flege et al. (2003) hypothesizes that when a second language (L2) (i.e. English) is acquired, there are several ways in which the phones of both languages can interact when they are similar:

3.2.1.2.1. VOT

The greater the perceived dissimilarity of an L2 sound from the closest L1 sound, the more likely a new category will be formed for the L2 sound (Flege et al. 2003).

The English and Korean aspirated VOTs are ~30ms apart in duration in native English and Korean studies (Klatt 1975, Silva 2006), thus it is likely that heritage speakers perceive them as dissimilar. Similarly, the Korean lax and English voiced stops are ~40ms different in duration. However, there is VOT overlap in Korean lax and English voiceless aspirated VOTs. Since these sounds are in pairs, however, it is unlikely that these two similar sounds will be perceived as the same sound. And, the pitch of the following vowel will be different, adding a second parameter of dissimilarity in their input. English aspirated stops have a raising pitch (Kim 2012), and Korean lax stops have a falling pitch (Cho et al. 2002, Kang and Han 2013, Kim 2012, Silva 2006). Though it is unlikely, this hypothesis predicts that Korean aspirated and English voiced will be separate phones, the Korean lax and English voiceless aspirated might interact.



**Figure 3: Hypothetical Heritage Results if Categories Merged (Flege et al. 2003)
(Numbers based on Klatt (1975) and Silva (2006))**

The heritage speakers may have an overlap in the form of a merged phonetic category of Korean lax and English voiceless according to a second Flege hypothesis of what can happen to similar categories: When a category is not formed for an L2 sound because it is too similar to an L1 counterpart, the L1 and L2 categories will assimilate, leading to a "merged" L1-L2 (Flege et al. 2003).

A similar possibility is that, if similar enough, the speaker can categorize two different languages' sounds as just one, however, instead of merging the two into a new category, they will just appear to be from one language. For example, in this study, the speakers might produce English voiceless stops and Korean aspirated stops that are completely like a native English or a native Korean stop (Flege et al. 2003).

Flege et al. (2003) has noted another possible manifestation: when a new category is established for L2 sound, it may dissimilate from neighboring L1 and/or

L2 sound - and vice versa - to preserve phonetic contrast (Flege et al. 2003). If the heritage speakers categorize Korean lax and English voiceless aspirated as different sounds, they are likely to make the differences in acoustic parameters more salient.

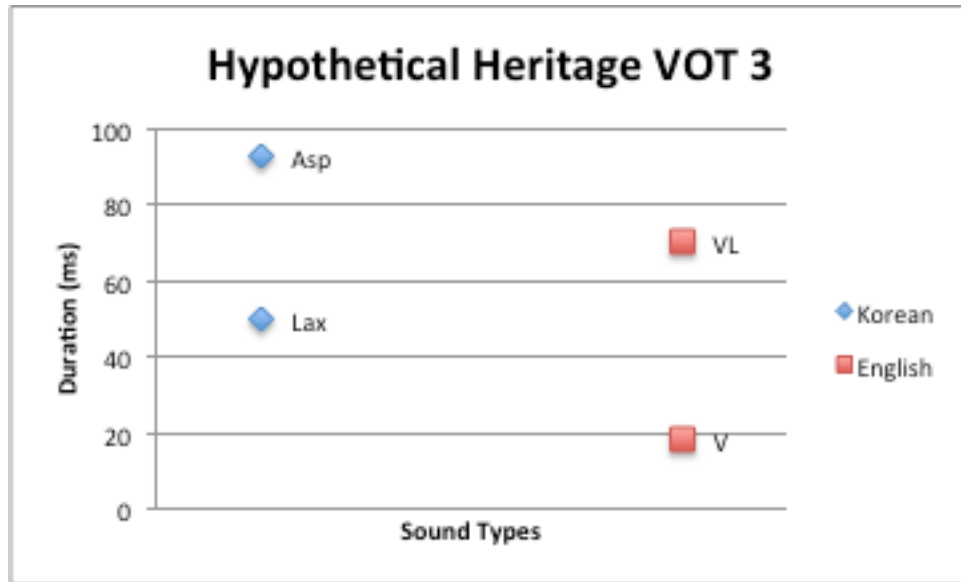
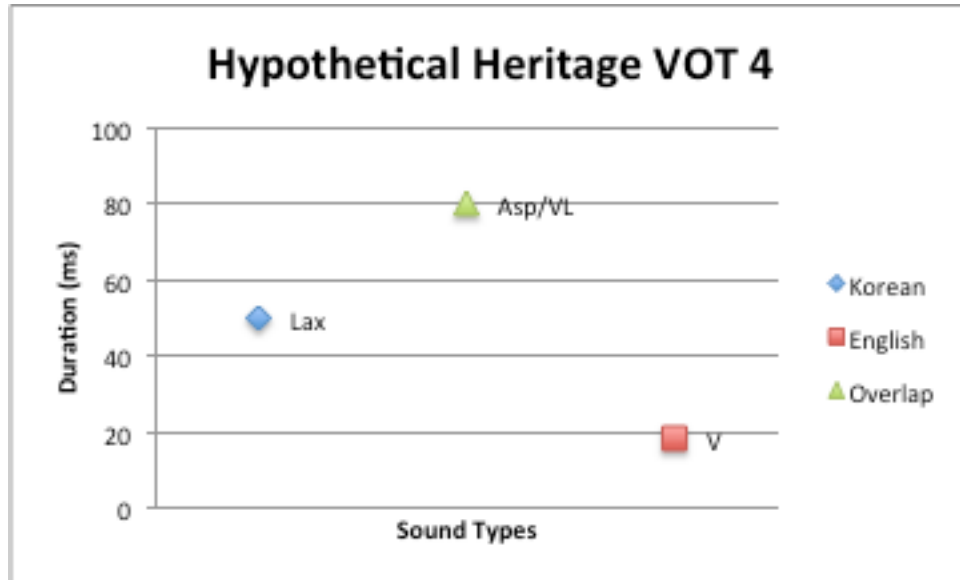


Figure 4: Hypothetical Heritage Results if Categories dissimilate (Flege et al. 2003)(Based on Klatt (1975) and Silva (2006))

However, these predictions are not made from isolated, empirical data. Korean and English stops do not exist in isolation, but rather in triplets and pairs. If the Korean lax and English voiceless aspirated interact and separate to become exaggeratedly different, this can affect the other sound categories in question. Korean aspirated and English voiceless aspirated are now more similar and could experience some language transfer of acoustic parameters. In fact, they might even overlap, forming one sound category, as shown below.



**Figure 5: Hypothetical Heritage Results if Categories Merged (Flege et al. 2003)
(Numbers based on Klatt (1975) and Silva (2006))**

3.2.1.2.2. Δ VOT

As mentioned above, these stops do not exist in isolation, as VOT data can imply. Δ VOT is also an acoustic parameter. Similar to a speaker exaggerating VOT to differentiate between two similar sounds in two different languages, VOT can be used to distinguish between sound pairs within the same language. This difference (i.e. Δ VOT) has been used by linguists to better understand Korean stops (Kang and Han 2013, Kang and Nagy 2013, Kim 2012, Silva 2006).

Above is outlined how VOT might be acoustically used for each phone, however, Δ VOT might also vary depending on how their phonetic categories have formed due to varying input, what has attrited over time, and language transfer, either English to Korean or Korean to English. If no transfer or attrition has occurred, the hypothesis of this study is incorrect, in which there can only be one result (see Figure 2). The heritage speakers will display Korean Δ VOTs of \sim 30ms

and English Δ VOTs of ~ 45 ms. If there has been language transfer and/or attrition, the Δ VOTs of both languages might merge somewhere between 30ms and 45ms (Flege et al. 2003), or diverge to create a stronger contrast with English Δ VOTs raising and Korean shrinking (Flege et al. 2003), or both languages' Δ VOT might trend like one languages' (i.e. both have Korean or both have English values). Finally, if the heritage speakers' current input is more similar to the latest VOT trend in Contemporary Standard Korean (i.e. Δ VOT ~ 10 ms), the heritage speakers' might reform their Korean Δ VOTs to match their current input or to fill in phonetic category gaps that have formed due to attrition (Flege et al. 2003).

3.3. Experiment

3.3.1. Participants

This experiment relied on the volunteered participation of 9 college students and 2 of their instructors. Among the nine students, 4 were heritage speakers and 5 were non-heritage speakers. All students recruited were enrolled in Korean late beginner or intermediate courses at the University of North Carolina at Chapel Hill, Duke University, and Wake Technical Community College. All participants were given a questionnaire in order to ascertain their English and Korean language usage histories and inputs. . These variables are important to know because the usage and exposure, as well as family history and origins can affect their results and can be used to better understand those results.

All of the heritage speakers began learning Korean before elementary school. They used it naturally at home. When beginning elementary school, they all reported that they stopped, lessened, or resisted speaking Korean in favor of English

until enrolling in their college Korean language courses, though all still favor English, except one participant who said she was comfortable using either.

All of their parents were born between 1952 and 1968 and speak Standard Korean, except one father, whose son reported spoke a regional dialect. Knowing the age of the parents is important since their original input will need to be cross-referenced with Silva's (2006) findings to see if the heritage speakers are maintaining that parental input or if other factors have played a bigger role in their Korean phonetic acquisition.

The non-heritage speakers also filled out the questionnaire. None of them spoke or were exposed to Korean in the home and were American English monolingual natives. They were all first taught Korean in the college classroom, though all had taken language courses in Spanish or French in high school, and were enrolled in late beginner to intermediate courses at the time of recording, like their heritage speaker classmates.

The instructors were the actual instructors of most of the student speaker participants. Two young (born between 1980-1985), female instructors were recorded. They both were graduate students or associate professors respectively teaching at the University of North Carolina at Chapel Hill and Wake Technical Community College. Both spoke Standard Contemporary Korean and were from Seoul or nearby, as reported by the instructor participants on their questionnaires.

3.3.2. Materials

Besides the questionnaire, which asked questions about who spoke Korean in the home and how frequently, as well as questions about family origins and

history, this experiment involved recording the participants producing the sounds in question in order to test the hypotheses. To test the hypotheses, Korean and English stops needed to be recorded in the same environments. To do this, a Logitech H390 USB Headset microphone (100Hz – 10kHz) was used to record each speaker reading phrases presented on a Macbook Air using Google Presentation in a sound proof chamber. The software used for the recordings was Praat (Boersma 2001) with mono sound recording and a sampling rate of 44kHz.

3.3.3. Design

Each student participant, heritage and non-heritage, read ten Korean phrases four times in a randomized order. They also read ten English phrases four times in a randomized order. Whether they read English or Korean first was alternated so the order would not somehow affect the results. Among the phrases were 12 target sounds (6 Korean, 6 English) and 8 misleading tokens so the participants would not guess the purpose of the study.

3.3.3.1. Korean Target Sounds

The target sounds of the Korean experiment were lax and aspirated stop pairs across three places of articulation: [p, p^h, t, t^h, k, k^h]. They were all presented in phrase-medial and word-initial position, as well as followed by [a].

Target	Hangul	Transcription	Gloss
[pa]	바보	[papo]	fool
[p ^h a]	파다	[p ^h ata]	(to) dig
[ta]	다방	[tapang]	coffee shop
[t ^h a]	타다	[t ^h ata]	(to) board
[ka]	강	[kang]	river
[k ^h a]	칸	[k ^h an]	room

Figure 6: Korean Tokens Used in Experiment with Transcription and Gloss

Each was presented in Hangul, the Korean alphabet. Partway through the experiment, it was noticed that there were some spacing issues of the Korean characters at the end of each phrase, so that [i gon TOKEN-i ra go hadzijo] read [i gon TOKEN i ra go ha dzijo]. However, because it was after the target token, which did not have spacing issues, it did not affect the results of this experiment.

이 건 강 이 라 고 하 지요.

Figure 7: Korean Phrase Presentation Example

The phrase reads, “This thing is called a _____,” with the target sound embedded in the token in the blank. Due to sentence structure, that token is actually phrase-medial. In this particular example above, the phrase reads, “This thing is called a river.” It uses the lax velar stop, [k], followed by [a]. For a better

understanding of the structure of this phrase, see Figure 8, which contains several levels of glosses.

- (1) Sentence Gloss
- “This thing is called a _____.”
 - “이 건 _____ (이) 라 고 하 시 요.”
 - [i gɔn _____ (i) ra ko ha dʒi jo]
- (2) Word-by-Word Gloss
- *i*
 - this
 - *kɔn*
 - thing
 - _____
 - token
 - *(i) ra ko ha dʒi jo*
 - is called
- (3) Morpheme-by-Morpheme Gloss
- *i*
 - this
 - *kes*
 - thing
 - *eun*
 - topic particle
 - _____
 - token
 - *(i) ra ko*
 - (i) added if token ends in a consonant
 - stem
 - *ha da*
 - present tense, plain
 - *dʒi jo.*
 - formal

Figure 8: Korean Phrase Gloss

3.3.3.2. English Target Sounds

The target sounds of the English experiment were voiced and voiceless aspirated stop pairs across three places of articulation: [b, p^h, d, t^h, g, k^h]. They were all presented in phrase-medial and word-initial position, as well as followed by [a]. They are considered the closest English-Korean stop correlates (Kim 2012) since in word-initial, the voiced stops are usually devoiced.

Target	Transcription	Gloss
[ba]	[barə]	bottle
[pa]	[pat]	pot
[da]	[dat]	dot
[ta]	[tarlə]	toddler
[ga]	[gad]	god
[ka]	[kat]	cot

Figure 9: English Tokens Used in Experiment with Transcription and Gloss

The presentation of the English tokens was in the English alphabet and the carrier-phrase was similar to the Korean in that the token is in phrase-medial position, as shown below.

The dot is here.

Figure 10: English Phrase Presentation Example

3.4. Measurements and Analyses

For each token produced, the voice onset time (VOT) was recorded in milliseconds (ms). VOT is the duration between the releasing of a stop closure and the beginning of vocal fold vibration. VOT is often used to define stop and affricate phones when acoustic properties are needed. This acoustic parameter that is produced due to physiological and aerodynamic elements in stops can also be used to better understand stop pairs, like in the Cho et al. (2002) sketch of Korean stops, affricates, and fricatives.

For this study, the landmarks were the transient burst in the spectrogram and the first voicing pulse, as shown in Figures 11 and 12. The difference between the aspirated and lax stop VOT values shows how salient the VOT parameters are. The difference can also show how Korean or English-like the phones are (Kim 2012, Silva 2006). This difference is defined as $\Delta VOT_{(asp-lax)}$. ΔVOT has been traditionally used in these analyses of Korean stops (Cho et al. 2002, Cho et al. 2011, Kang and Han 2013, Kang and Nagy 2013, Kim 2012, Silva 2006).

As mentioned above, these landmarks can be seen in Figures 11 and 12. In Figure 11, a Korean lax dental stop ([t]) can be seen. Subject 11's [t^h] is shown in Figure 12. This subject, a non-heritage speaker, has a longer highlighted duration (61.5ms vs. 121ms). These durational differences (ΔVOT) are one way in which the results of this experiment will be used to test the hypotheses, along with the actual VOTs themselves. In this example, ΔVOT is 59.5ms, which is a typical ΔVOT for a non-heritage speaker.

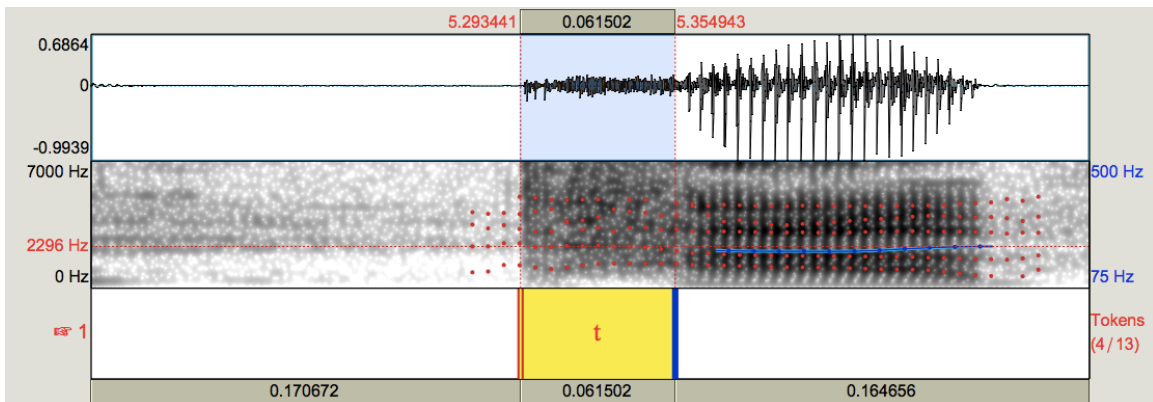


Figure 11: An example of the landmarks used for measurement, using Subject 11's first Korean [t] repetition (Boersma 2001)

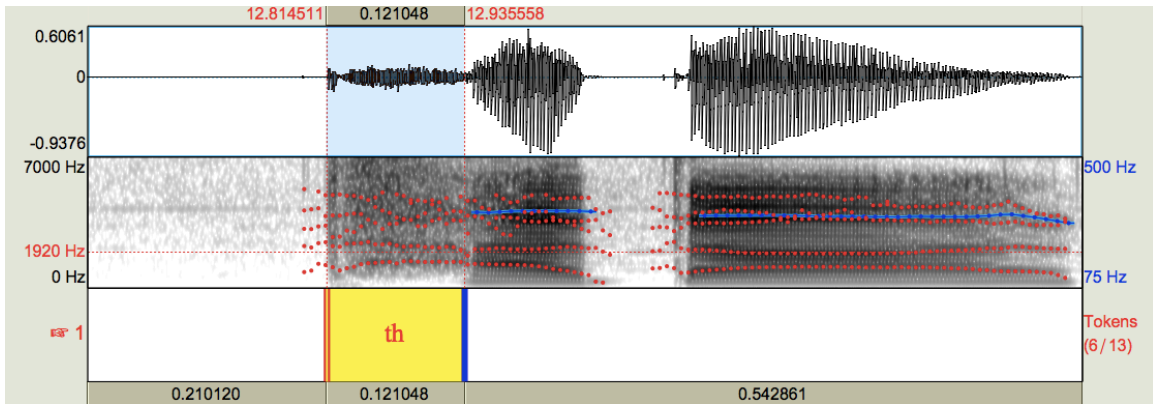


Figure 12: An example of the landmarks used for measurement, using Subject 11's first Korean [t^h] repetition (Boersma 2001)

The VOTs of each group were averaged according to which phone they are: Korean lax, Korean aspirated, English voiced, and English voiceless aspirated. These averages were then compared in several different ways: lax versus aspirated, across speaker groups, and across languages. The comparison portion of the analysis will be discussed in Chapter 4 – Results and Discussion. This analysis also included a statistical analysis in the form of A mixed-designs Model, which will also be further defined in Chapter 4 – Results and Discussion.

CHAPTER 4: RESULTS AND DISCUSSION

4.1. Korean VOT by All Speaker Groups

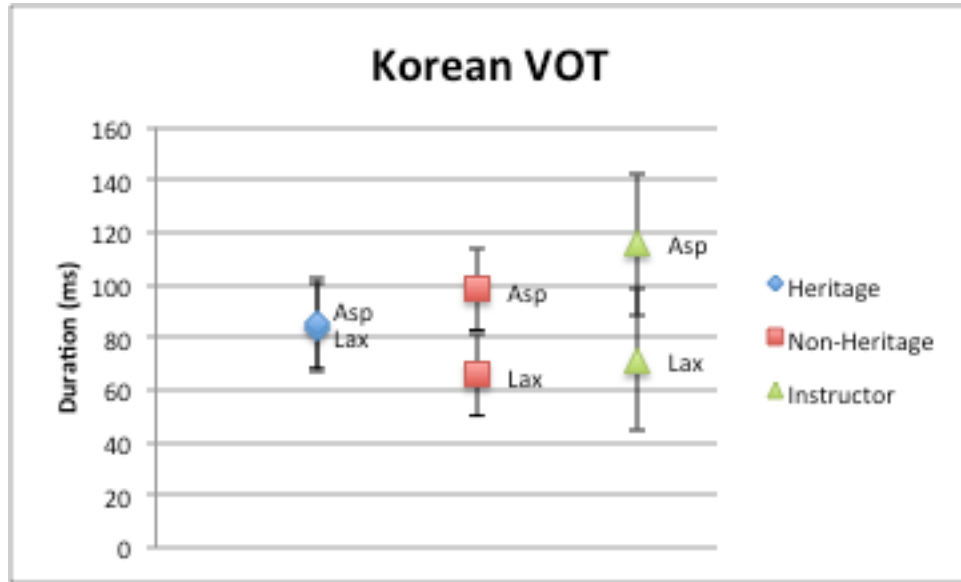


Figure 13: Korean VOT of Aspirated and Lax Stops Produced by All Speaker Groups. (Error bars represent upper and lower-tailed confidence test with Alpha = 0.05)

In Figure 13, the lax and aspirated Korean stops of each speaker group are represented with the duration of their respective VOTs in milliseconds on the y-axis. The distance between the aspirated point and the lax point of each speaker group represents the difference in their VOT values, or their Δ VOT. While the non-heritage speakers and the instructors display a difference in their lax and aspirated VOT production, the heritage speaker results for lax and aspirated VOT durations overlap, showing that these speakers do not use VOT to differentiate between Korean lax and aspirated stop production.

While Benmamoun et al. (2013) found in their survey of heritage speaker phonological and phonetic studies that phonological competence seems to be a well-preserved linguistic system among heritage speakers (i.e. compared to other systems such as syntax), they also proposed that it is still not entirely native-like. Thus, this study anticipated a larger difference between the Korean lax and aspirated VOTs of the heritage speaker group. However, that anticipation was incorrect as these Δ VOT values correspond to those of native Korean speakers in their age category (Silva 2006).

In regards to the Korean native-like Δ VOT and the proposals of this study's theoretical aspects, the comparison of the permanence hypothesis and first language attrition/language transfer should seem to favor the permanence hypothesis since they have native-like results, however in reality, first language attrition is the more salient factor. The Korean to which the heritage speakers were exposed as young children would have included a larger Δ VOT (Silva 2006) of ~20ms. If the permanence hypothesis were indeed at work here, the original acoustic parameters acquired during childhood would not have changed (Benmamoun et al. 2013). This manner of distinguishing between the Korean stop phones underwent attrition over time, or those are the results this study would have found. That the heritage speaker production is similar to Korean monolingual speakers of their own age group means that the old Δ VOT attrited, and these speakers acquired a new one to replace the lost one when restarting their Korean acquisition as adults. What is interesting is that their input (i.e. the instructors) does not match their output. After seeing these results, the VOT of audio course materials

was looked at and this study did not find large Δ VOTs. More careful, empirical analysis is needed in future studies however.

The heritage speakers could be following the trend of that audio input or possibly have input from Korean media, such as Korean television dramas, which some of the participants mentioned they enjoyed. Further testing of these hypotheses is needed to ascertain whether these students are indeed weighing these other inputs as more important than the instructors’.

4.2. Statistical Analysis of All Speaker Groups

<i>Label</i>	<i>Contrasts</i>			
	<i>Num</i>		<i>Den</i>	
	<i>DF</i>	<i>DF</i>	<i>F Value</i>	<i>Pr > F</i>
All Means Equal	9	556	90.56	<.0001
Sound among Language*Speaker	4	556	40.64	<.0001

Figure 14: Statistical Results from a mixed-effects ANOVA

The results were submitted to a mixed-effects (2.5) speaker type (heritage, non-heritage, instructor) X (2) language (Korean, English) X (2) sound type (lax/voiced, aspirated/voiceless aspirated) ANOVA. Because the instructors only have 1 language result, they lower the speaker type from (3) to (2.5) due to their fewer results. Also, the place of articulation of each sound type was neglected in the manner of Silva (2006). The model accounted for random effects as well, which were the individual speakers’ means. The all means equal hypothesis was disproven, with $F = 90.56$ and $Pr > F < 0.0001$, which means that, when looking at all the populations of data, they were significantly different overall in their means.

A second ANOVA was run, this time the distribution of the result populations was correlated, not isolated. Each sound type duration was looked at with dependencies on speaker type and language. This test did not look at individual comparisons (e.g. heritage Korean aspirated vs. heritage English aspirated), some of which were not significantly different in duration, however, it did look at the populations of results as a whole and found that sound type durations varied significantly depending on speaker type and language with $F = 40.64$ and $Pr > F < 0.0001$.

Individual populations comparisons were also done, such as the example above of heritage Korean and English aspirated stops. Some of these individual populations were found to be significantly different, such as the instructor lax and aspirated stop VOT durations, and others, such as the heritage speakers' Korean and English aspirated stops, were not found to have significantly different mean durations.

4.3. English VOT by Heritage and Non-Heritage Speaker Groups

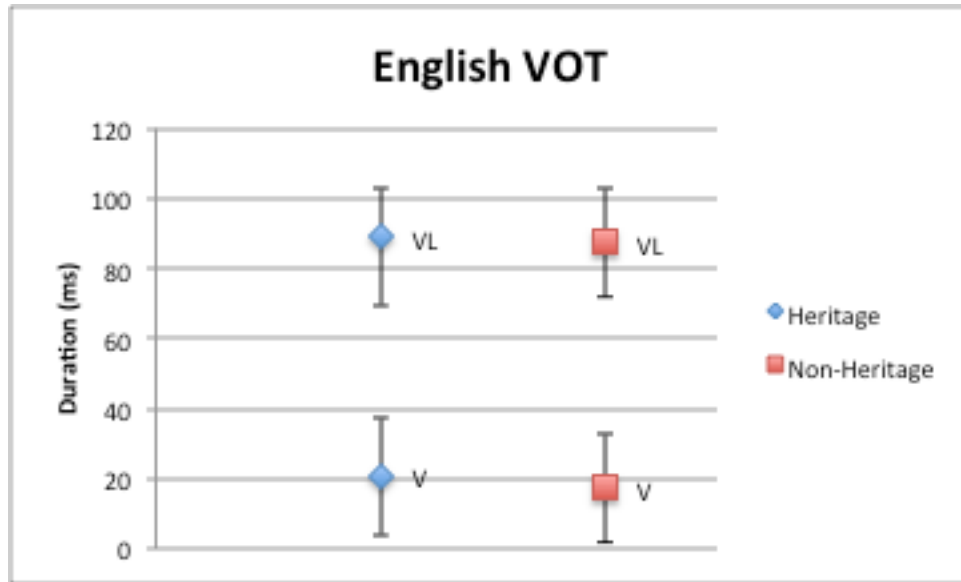


Figure 15: English VOT Results by Heritage and Non-Heritage Speaker Groups (Error bars represent upper and lower-tailed confidence test with Alpha = 0.05)

The heritage and non-heritage speakers were also recorded producing English stops: the voiceless aspirated stop, which corresponds to the Korean aspirated stop, and the voiced stop, which corresponds to the Korean lax stop in phonetic parameters. There is no significant difference between the productions of English stops by the heritage speakers compared to the non-heritage with $t = 0.10$ and $Pr > |t| = 0.9188$. Both groups differed in their VOT durations for voiceless aspirated and voiced stops, and the heritage speakers used VOT to differentiate between their English stops, unlike in their Korean stop production. These results were anticipated in this study since heritage speakers' had young ages of English acquisition as well as continuous exposure and usage of English, as reported in their questionnaires. Young exposure and continued usage of a language are ways in which native-like production can be achieved and maintained (Benmamoun et al. 2013, Krashen et al. 1979, and Schmidt 2009).

4.4. Heritage Speaker English and Korean VOT

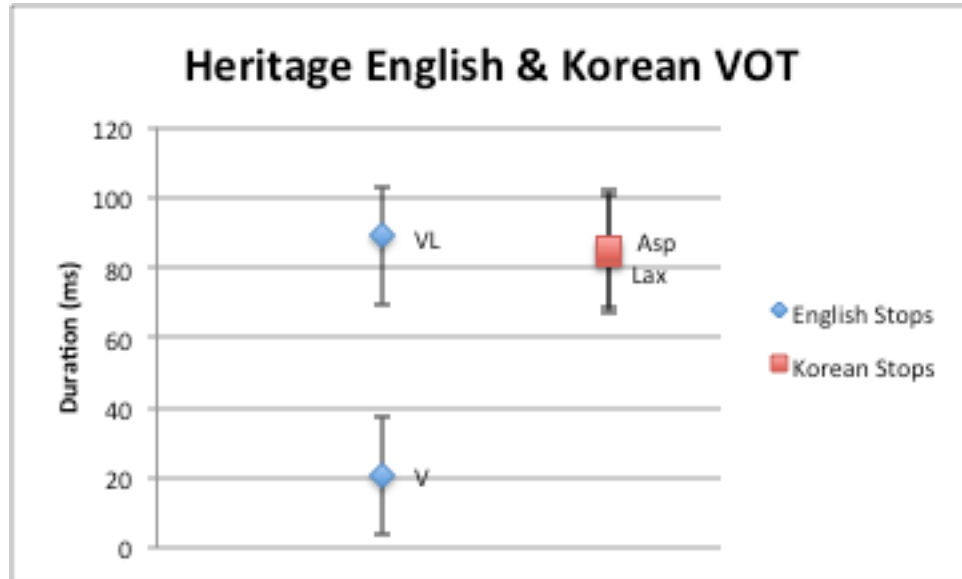


Figure 16: English and Korean VOT Results Produced by Heritage Speaker Group (Error bars represent upper and lower-tailed confidence test with Alpha = 0.05)

Another necessary analysis for this study is the VOT values of the heritage speakers' stops cross-linguistically in order to ascertain whether the speakers are producing their Korean stops differently than their English stops. If the English and Korean VOT values of corresponding stops (i.e. English voiceless aspirated and Korean aspirated / English voiced and Korean lax) are different, this would mean that they are in separate phonetic categories (Flege et al. 2003). When comparing the VOTs of the heritage speakers' English voiced and Korean lax stops, there is a significant difference ($t = 14.11$, $Pr > |t| < 0.0001$), meaning that these stops are separate phonetic categories (Flege et al. 2003). However the heritage speaker group did not have a significant difference between the VOT values of their English voiceless aspirated and Korean aspirated stop productions with $t = 0.22$ and $Pr > |t| = 0.8265$. Silva (2006) found that young speakers of Korean are lessening the VOT of

their aspirated stops while lengthening those of their lax stops to create the phonetic overlap. Thus, native Korean speakers' aspirated stops should have VOTs between 50 and 60ms (Silva 2006). However, these speakers have VOTs between 80 and 90ms.

This study anticipated non-native-like productions of Korean stops by heritage speakers. While the difference between Korean lax and aspirated stops (i.e. Δ VOT) was entirely Korean native-like, the actual VOTs of the stops may not be. The Korean aspirated stops were produced with a lengthened VOT and the lax stops were also lengthened in their VOT, making the Δ VOT native-like and the VOT itself un-native-like. Thus, this study seems to agree with Benmamoun et al.'s (2013) findings of near native-like production that is not quite native-like. However, Benmamoun et al.'s (2013) survey was not just looking at Korean, which has more than one native-like production, depending on speaker age. The seeming un-native-like Korean aspirated VOT can be explained when considering all aspirated stop inputs, such as American English and the Korean spoken by the participants' parents.

Klatt (1975) measured American English voiced and voiceless initial stops and found that voiced stops averaged ~18ms in VOT and voiceless aspirated stops averaged ~61ms. The English voiced stops above are similar to Klatt's findings, however the voiceless aspirated stops are ~30ms longer than those in Klatt's findings. While it is possible that the American English sub-dialect spoken by these participants is different from Klatt's participants or that American English voiceless aspirated stops have undergone a phonetic change since 1975, it is also likely that,

with two languages interacting in this study's participants' minds, there is Korean to English transfer as well. During acquisition of an L2, new L2 sound categories may or may not form (Flege et al. 2003). The VOT of the Korean and English aspirated stops are trending together, forming just one sound category instead of two. Thus, there is only one sound category for aspirated stops regardless of the language. It is difficult to determine to which language these values are trending; however the age of the participants' parents is key.

Korean native speakers born between 1955-1965, like their parents, produce Korean aspirated stops with longer VOTs than those produced by Korean monolingual speakers of similar age to the participants with values of ~90-95ms. Thus, it appears that for these participants, voiceless aspirated stops, regardless of language, have the VOT of their initial input, their parents, demonstrating how important age is as a factor in language acquisition, even in more complicated linguistic situations such as those of heritage speakers'. The voiceless aspirated category formed based on that Korean input and did not change due to first language attrition or language transfer, giving the permanence hypothesis some credence.

4.5. Instructor VOT

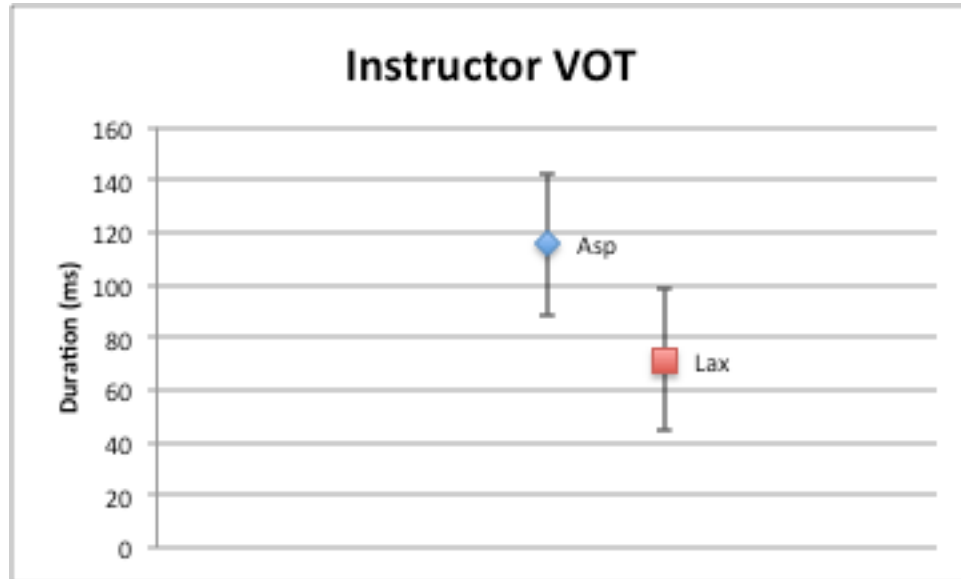


Figure 17: Korean Aspirated and Lax VOT Produced by Instructor Group (Error bars represent upper and lower-tailed confidence test with Alpha = 0.05)

The instructor VOT and Δ VOT results were the least anticipated in the proposals of this study, which assumed the instructors would display completely Korean native-like production. However, native Korean speakers of their age category (i.e. born between 1980-1985) had lax and aspirated VOT values of 50-60ms, with Δ VOT values of <10ms (Silva 2006). In this study, the instructors produced Δ VOTs of ~44ms ($t = 6.15$, $Pr > |t| < 0.0001$). And, the VOTs themselves are much greater than the Korean native VOTs of 50-60ms (Silva 2006) with lax being ~71ms ($t = 5.22$, $Pr > |t| < 0.0001$) and aspirated being ~115ms ($t = 8.42$, $Pr > |t| < 0.0001$). While their VOT and Δ VOT values were much greater than their age group's (Silva 2006), the pitch change in the following vowel was audibly different between the lax and aspirated stops. There are three possible explanations for this data trend of using VOT as a strong acoustic parameter based on similar studies

(Chang 2012, Cho et. al 2011, Kang and Han 2013): L2 to L1 transfer, prestige dialect, or hyper-articulation.

Chang (2012) found that even a small amount of study of a second language affects the production of the first language. He found that learning Korean caused phonetic drift in the participants' L1 English. This L2 to L1 transfer is a possible explanation for why the VOT values are not conforming to Silva's (2006) findings. However, this is unlikely an unlikely cause of the un-native-like results when compared to the heritage speaker results. If L2 to L1 transfer were the main cause, the heritage speakers and the instructors would be more similar. However, the heritage speakers are behaving native-like in pronunciation, at least in their Δ VOTs. While the instructors were not native-like in that they did not conform to their age bracket, they could be considered native-like in that they were speaking like an older person possibly due to respect or prestige given to the older dialect. It might be possible if the instructor results were compared to the VOT and Δ VOT values of Silva (2006)'s results for speakers born between 1945-1949. Silva found that Korean native speakers born in these years produce aspirated stops with VOTs \sim 120ms and lax stops with VOTS \sim 80ms, with a Δ VOT of \sim 40ms. These values are similar to those produced by the instructors.

Prestige dialects are dialects that are held at a higher value than other dialects. One possible explanation for the instructor results is that they are speaking the older, possibly prestige dialect. However, there is no evidence that the older generations' pronunciation is considered better or more standard than the younger speakers' of Contemporary Standard Korean. Kang and Han (2013) performed a

longitudinal case study on a man who had been recorded as a child. They found, when they re-recorded him, that his VOT had shifted towards the younger pronunciation compared to when he was a child. This means that the younger speakers are not trying to conserve the older speech, but rather that older speakers are changing their pronunciation to fit the new standard. Due to a prestige dialect being unlikely with Kang and Han (2013)'s results, this leaves hyper-articulation.

Hyper-articulation in Korean has several phonetic cues. Cho et al. (2011) looked at VOT changes of aspirated stops in different types of Korean speech. They found that the VOT of aspirated stops is much greater in "clear" speech instead of conversational speech. Instructors, in order to best give their students salient cues for perception of sounds in Korean, are possibly hyper-articulating in their teaching, or "clear" speech, register. Because Cho et al. (2011) did not look at lax stops in their study, this hypothesis is not entirely evidenced. But, further testing could include lax stops in order to better understand these instructors' results. To do this, instructors would need to be recorded speaking to a student about some task (e.g. a reading comprehension activity) and speaking to another native speaker of similar age so they can reach a more conversational register. To ascertain whether hyper-articulation is indeed the cause of this study's instructor results, this further testing would be necessary. However, this study's results do not disagree with Cho et al. (2011), and with further study, it is quite possible that hyper-articulation is the cause of this study's results.

CHAPTER 5: CONCLUSION

Using an empirical study of the acoustic properties of stop production by heritage speakers and their non-heritage peers, as well as their instructors, this study strove to analyze the roles of first language attrition, language transfer, and the validity of the permanence hypothesis from NLNC. In order to do so, the participants were recorded reading Korean and English phrases and the stops' VOTs were measured and compared.

Heritage speakers were found to produce Korean lax and aspirated stops with an average Δ VOT of <10ms, which is in accordance to Korean native speakers their age (Silva 2006). These results imply that the ~20ms Δ VOT of their original input (i.e. their parents) are no longer part of their phonetic system. It is not due to language transfer either, since native English Δ VOT is much greater. Rather, it is due to first language attrition. The original manner of differentiating lax and aspirated Korean stops was possibly lost, and further research could be done to ascertain whether this new Δ VOT is due to more recent input in the form of Korean media, which some participants reported enjoying.

However, the VOTs of the heritage speakers' Korean stops are greater than a Korean native speaker their age. In particular, the Korean aspirated stop VOT is similar to those reported by Silva (2006) to be produced by Korean speakers their parents' ages. In fact, this supports the permanence hypothesis, which states that

the input received from their parents would persist throughout life, unchanging. While the whole stop phonetic system did not persist unchangingly, the heritage speaker participants retained their aspirated VOT from childhood. And, their English voiceless aspirated stops experienced language transfer from Korean, forming just one voiceless aspirated sound category for both languages, showing that language transfer plays a role in heritage speaker phonetic systems.

This study furthers theoretical linguistic testing of first language attrition, language transfer, and the role of age in acquisition and retention presented in the form of the permanence hypothesis. Each plays a role in the phonetic category formation, retention, and attrition of the heritage speakers.

REFERENCES

- Albrini, Abdulkafi (2014). "Toward understanding the variability in the language proficiencies of Arabic heritage speakers." *International Journal of Bilingualism*, 18, 6: 730-765.
- Benmamoun, Elabbas, Silvina Montrul and Maria Polinsky (2013). "Heritage languages and their speakers: Opportunities and challenges for linguistics." *Theoretical Linguistics*, 39, 3-4: 129-181.
- Boersma, Paul (2001). Praat, a system for doing phonetics by computer. *Glott International* 5:9/10: 341-345.
- Chang, Charles B. (2012). "Rapid and multifaceted effects of second-language learning on first-language speech production." *Journal of Phonetics*, 40: 249-268.
- Cho, Taehong, Sun-Ah Jun and Peter Ladefoged (2002). "Acoustic and Aerodynamic Correlates of Korean Stops and Fricatives." *Journal of Phonetics*, 30, 2: 193-228.
- Cho, Taehong, Yoonjeong Lee and Sahyang Kim (2011). "Communicatively driven versus prosodically driven hyper-articulation in Korean." *Journal of Phonetics*, 39, 3: 344-361.
- Flege, James E., Carlo Schirru, Ian R.A. MacKay (2003). "Interaction between the native and second language phonetic subsystems." *Speech Communication* 40: 467-491.
- Kang, Yoonjung and Naomi Nagy (2013). VOT Merger in Heritage Korean in Toronto. *Proceedings of the the annual meeting of the Canadian Linguistic Association 2012*.
- Kang, Yoonjung and Sungwoo Han (2013). Tonogenesis in early Contemporary Seoul Korean: a longitudinal case study. *Lingua*: 134, 62-74.
- Kim, Mi-Ryoung (2012). "L1-L2 Transfer in VOT and f0 Production by Korean English Learners: L1 Sound Change and L2 Stop Production." *말소리학 음성과학 제 4 권 3 호*: 31-41.

- Krashen, Stephen D., Michael A. Long, Robin C. Scarcella (1979). "Age, Rate and Eventual Attainment in Second Language Acquisition." *TESOL Quarterly* 13, 4: 573-582.
- Kuhl (2004). "Early Language Acquisition: Cracking the Speech Code." *NeuroScience* 5: 831-843.
- Nagy, N. 2009. The challenges of less commonly studied languages: Writing a Sociogrammar of Faetar. In J. Stanford & D. Preston, eds. *Variation in Indigenous Minority Languages*. Philadelphia: John Benjamins. *Impact series* 25. 397-417.
- Nagy, N. 2011. A multilingual corpus to explore geographic variation. *Rassegna Italiana di Linguistica Applicata* 43.1-2:65-84.
- Silva, David J. (2006). "Acoustic Evidence for the Emergence of Tonal Contrast in Contemporary Korean." *Phonology* 23, 2 (2006): 287-308.
- Schmid, Monika S. (2009). "On L1 attrition and the linguistic system." *Eurosla Yearbook* 9: 212-244.
- Zhang, Yang, Patricia K. Kuhl, Toshiaki Imada, Makoto Kotani, and Yoh'ichi Tohkura (2005). "Effects of language experience: Neural commitment to language-specific auditory patterns." *NeuroImage* 26: 703-720.