PHONETICS AND PHONOLOGY INTERPLAY IN LOANWORD ADAPTATION: ENGLISH ALVEOLAR FRICATIVE INTO KOREAN

by

Miyeon Ahn

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Linguistics)
in the University of Michigan
2011

Doctoral Committee:

Assistant Professor Andries W. Coetzee, Chair
Professor Patrice Speeter Beddor
Professor Julie E. Boland
Professor San Duanmu

DEDICATION

To my dad, mom, and my husband

ACKNOWLEDGEMENTS

The writing of a dissertation was a trilled experience. This dissertation could not have been completed without the help, support and efforts of a lot of people. I sincere gratitude goes to my advisors, my colleagues and friends, and my family.

Firstly, my deepest gratitude is to my advisor, Dr. Andries W. Coetzee. It is not often that one finds an advisor who is already there for you. I have been incredibly fortunate to have an advisor who has been always there to listen and give advice. I am deeply grateful to him not only for giving me amazing academic advice but also for sharing his heart throughout my graduate studies. Andries taught me innumerable lessons for being an independent scholar, and I greatly appreciate amazing mentorship. One time, I asked Andries how he could be that excellent mentor, and he answered he had learned from his advisor. I hope I would be an advisor to my students as Andries has been to me.

I am also grateful to Dr. Patrice Speeter Beddor who has provided incredible patience and supports. She was readily available in every situation. She read and responded my chapters in timely manner, and warmly opened her office for discussion. Pam also concerned if I was facing any kind of difficulties during my studies, as she so generously is for all her students, and she was always willing to help. Pam's comments on each chapter were extremely perceptive and helpful. Her insightful comments along with technical and editorial advice were essential to the completion of this dissertation. I have learned being professional from her.

My thanks go out to the other committee members, Dr. Julie Boland and Dr. San Duanmu. Julie let me think interesting issues in various perspectives. I thank her encouraging words, thoughtful criticism, and attention during my studies. San's comments on my work always have been appropriate, through and helpful. I sincerely thank his advice and guidance over the past years.

My research for this dissertation was shared with Phondi group members: I thank my colleagues, Anthony Brasher, Susan Lin, Xinting Zhang, Kevin McGowan, Jonathan Yip, Yan Dong, Harim Kwon, Michael Opper and Cameron Rule for their helpful comments. I also wish to thank my experiment participants, mostly Ann Arbor Hope Church Young Adult group members. My graduate studies would not have been the same without social interactions with them, and this research would not able to been possible without their participation.

I also wish to thank my ex-advisors, Borim Lee and Jae-Young Lee. I appreciate their belief in me. I am grateful to my best friend, Soyeon Yoon, who given me faithful friendship.

Last, my thanks also go out to my family members: my parents, Hongchul Ahn and Sunhwa Lee, my sisters, Suyeon Ahn and Jiyeon Ahn, and my brother, Kichang Ahn, have been very supportive and proud of me. I have to give a special thanks to my parents-in-law, Taeik Cho and Geumja Lee, who have been the source of constant love.

Finally, my very special thanks to the one person whom I love most, my husband Kae Won Cho. I would like to thank him for his love and support. His faith and love have been with me during the past years, and they enabled me to complete this dissertation. Thank you. All errors and limitations of this dissertation are, of course, mine.

TABLE OF CONTENTS

CHAPTER IV.

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

The purpose of this investigation is to determine the general principles that govern Korean-speaking listeners' perceptual strategies regarding the adaptation of English alveolar fricatives into Korean. The English voiceless alveolar fricative is variously borrowed into Korean as /s*/ (as in*'say'* [s*ei]) or /s/(as in *'stay'* [sɨtei]). The systematic pattern for English onset /s/ is that prevocalic $/s/_E$are borrowed as $/s^*/_K$, and preconsonantal $/sC/_E$ is borrowed as $/s/_K$.

In investigating English /s/-initial loanwords in Korean, this dissertation postulated two hypothesis, Hypothesis 1, the consonantal information hypothesis, and Hypothesis 2, the vocalic information hypothesis. Hypothesis 1 predicted that /s/-internal properties (i.e., frication portion itself) play an important role in Korean listeners' English /s/ perception; Hypothesis 2 states that Korean listeners predominantly rely on /s/-external properties, particularly properties of the following vowel. Hypothesis 2 was postulated to apply to vowels that are acoustically present as well as vowels that are perceptually epenthesized.

Four production and three perception experiments were conducted to investigate aspects of these hypotheses. The experiments confirmed that Korean listeners were insensitive to consonantal information in their perception of the English fricative, in spite of temporal and spectral differences between English /s/s in prevocalic and preconsonantal positions. Korean listeners have difficulty discriminating /sC/ from /sɨ̯C/.

This difficulty is interpreted as an indication of a perceptual illusion, which in turn leads to speculate that the illusory vowel influences fricative perception. Phonotactic resolution (i.e., vowel epenthesis) is selectively applied in order for the epenthetic vowel to be perceptually least salient in favor of minimal perceptual consequences of an illusory epenthetic vowel.

These experimental results serve as information about listeners' perceptual modification strategies in facing phonotactically illegal sequences. They also suggest that phonetic and phonological knowledge about Korean interact in loanword adaptation, and that listeners incorporate this knowledge when they perceive non-native speech sounds.

CHAPTER I
INTRODUCTION

The speech process involves interactions between speakers and listeners. Speakers produce speech by engaging their articulatory systems with main task being to clearly convey their intent. The task of listeners, on the other hand, is to perceive speech in order to determine what the speaker has said. In perception, listeners assign a linguistic interpretation to the input acoustic signal.

In many communicative interactions, and in much of the research on speech perception and production, speakers and listeners share the same language. But this situation does not always hold. Loanword adaptation, which concerns borrowing of lexical items from foreign languages, usually involves at least two different languages. The particular type of cross-language situation that holds in this circumstance is that speakers and listeners do not share the same language, and listeners are incorporating into their language words from the speakers' native language.

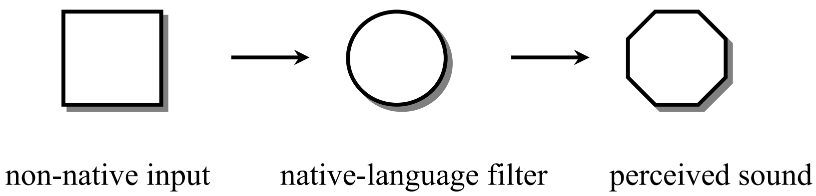
This dissertation is focused on the interplay between phonetics and phonology that occurs in loanword adaptation, with particular attention to the contribution of the listener to the process. The target of investigation is the adaptation of alveolar fricatives from English to Korean. The aim is to determine the general principles that govern Korean-speaking listeners' perceptual mechanisms or strategies regarding their adaptation of English alveolar fricatives into Korean. Specifically, I examine how Korean listeners deal with the (non-native) English fricatives on the basis of their native phonetic and phonological knowledge and how this knowledge determines the loanword adaptation patterns.

1.1 Native Phonetic/Phonological Knowledge as a Language Filter

In speech perception, an input sound is considered to be non-native when it is not included in listeners’ native language sound inventory. In perceiving a non-native speech sound in a word that will ultimately be added to the listeners’ lexicon, listeners have to balance two possibly conflicting demands. On the one hand, listeners presumably aim to perceive accurately what the speaker said. On the other hand, there is substantial evidence in the literature that listeners interpret the input signal within the framework of their native phonological system, which often involves perceptually modifying the input.

By way of illustration, let us assume that the shapes below represent sounds with the square being a non-native sound and the circle a similar-sounding native sound that operates as the native-language filter.

Figure I.1 Visual abstraction of foreign to native modification



When the square non-native phone enters a native inventory that allows only circles or circle-like phonemes, modifications may occur so that the new input is perceived so as to more closely conform to the native phonological inventory. The outcome of this process may be identical (or nearly so) to the input (the square), but it may also be the octagon, which represents a modified percept that is similar or possibly identical to a phone within the native inventory, allowing the listener’s lexicon to now encompass the input.

This modification, if any, must be minimal in that the modified sound needs to be sufficiently similar to the original input for perception to be successful (or sufficiently successful for the speaker’s intent to be understood). These two conflicting demands – (i) conforming to the native sound system, and (ii) minimizing the modification – force listeners to ‘negotiate’ in the process of non-native speech perception.

The native-language filter is the phonetic and phonological knowledge of the listener's native language, and it plays an important role in the negotiation strategies that listeners employ. The filter is shaped by language-specific experience that includes the phonetic details necessary for identification and discrimination of segmental contrasts. This filter is language specific because phonemic contrasts, the allophonic variants of phonemes, and phonotactic constraints are all language-specific. As discussed below, the filter can severely constrain the perception of non-native segments. In other words, the native language filter is imposed on the process of perceptual modification in non-native speech perception, resulting in relatively native-like sounds.

Experimental evidence accumulated over the last 40 years shows that listeners perceive non-native sounds through the filter of their native language. Non-native sounds that are highly distinct from native sounds may be relatively unaffected by the filter. For instance, Zulu clicks are accurately discriminated by English listeners, whose language does not use clicks linguistically (Best *et al.*, 1988). However, new sounds that are articulatorily and acoustically similar to one or more native sounds typically undergo perceptual modification (Best, 1995; Flege, 1995), as shown by the confusion of /l/ and /r/ by Japanese listeners, whose language lacks a phonological contrast between these two sounds (Miyawaki *et al.*, 1975; MacKain *et al.*, 1981; Sheldon and Strange, 1982). This perceptual confusion reflects how existing sounds influence the perception of new sounds in their acoustic vicinity.

Evidence shows that the native phonological knowledge of listeners can distort the input to the point that the listeners undergo a perceptual illusion, perceiving sounds that are absent from the acoustic signal. Berent *et al.* (2007) showed that English listeners systematically misperceived impermissible grammatical structures such that some consonant clusters that are absent in English are misperceived as containing an epenthetic vowel (e.g., [lbif] perceived as [ləbif]). Listeners perceived an illusionary extra vowel between the members of the consonant cluster so that the cluster structurally conforms to the phonological principles of their native language. This perceptual illusion suggests that listeners are sensitive to their native phonological system, that is, the native system influences non-native speech perception.

Korean loanword adaptation fits into this literature in that impermissible English consonant clusters are resolved with vowel epenthesis in Korean loanwords as in *'bright'* /bɹaɪt/ → [bɨɹaitɨ] and *'strike'* /stɹaɪk/ → [sɨtɨɹaikɨ]. It shows that native phonological principles (i.e., phonotactic constraints) modify the perceptual structures and yield perceptual illusions.

1.2 Voiceless Alveolar Fricatives in English and Korean

Given that the native phonological system plays a crucial role in cross-language speech perception, it is important to understand how listeners use their native phonetic and phonological knowledge during non-native perception. This dissertation approaches this issue by comparing voiceless alveolar fricatives in two languages, English and Korean. While English has only one voiceless alveolar fricative phoneme, /s/, Korean has a contrast between a lax /s/ and a tense voiceless alveolar fricative /s*/. Henceforth, when a distinction is necessary, the Korean sounds will be identified with a subscripted "$_K$" as in $/s/_K$ vs. $/s^*/_K$, and the English sounds with a subscripted "$_E$" as in $/s/_E$.

When foreign sounds are perceived as linguistic units, listeners often interpret the foreign units as similar-sounding native units. That is, they map the foreign sounds onto native ones. When English $/s/_E$ is heard by Korean listeners, they are expected to map $/s/_E$ onto a native Korean sound that is similar to $/s/_E$. However, that Korean has two English $/s/_E$-like sounds (i.e., $/s/_K$ and $/s^*/_K$), presents Korean listeners with a choice: when Korean listeners hear a production of $/s/_E$, they have to determine which Korean alveolar fricative to interpret this as, that is, whether the English utterance corresponds to $/s/_K$ or $/s^*/_K$. English listeners, on the other hand, are not faced with the same kind of choice when they hear Korean /s/ and /s*/. Since their phonological system has only one voiceless alveolar fricative, $/s/_E$, they should perceive utterances of both $/s/_K$ and $/s^*/_K$ as $/s/_E$.

Now, let us examine the actual loanword patterns of English /s/ containing words in Korean. As shown below, English /s/ is systematically borrowed into Korean as either /s/ or /s*/.[1] In order to delineate the general borrowing pattern of English /s/ in various positions, the data in Table I.1 show $/s/_E$ in both word-initial and final contexts, although this dissertation will focus on $/s/_E$ in word-initial context only.

There are five different phonological environments in which $/s/_E$ appears. Word-initially, $/s/_E$ can appear as a single onset consonant as in '*say*'(i.e., prevocalic $/s/_E$) and as a component of an onset consonant cluster as in '*stay*' (i.e., preconsonantal $/sC/_E$). Word-finally, $/s/_E$ can appear as a single coda consonant as in *'gas'* (i.e., postvocalic $/Vs/_E$). When it appears as a component of a coda consonant cluster, it can be either preconsonantal (i.e., $/\_sC/_E$) as in *'guest'* or postconsonantal (i.e., $/\_Cs/_E$) as in *'box'*. Note that word-final preconsonantal /s/ (i.e., $/\_sC/_E$) can also be postvocalic. A singleton $/s/_E$, regardless of position, is always borrowed as $/s^*/_K$, whereas a cluster $/s/_E$ is always borrowed as $/s/_K$. These borrowing differences are nearly invariant: $/s/_E$ in a given English word is borrowed as the same Korean sound with very few exceptions.

The loan word pattern is well-known, and has been the subject of previous experimental study (Kim, 1999; Kim and Curtis, 2002; Kang 2008, among others). Kim (1999) and Kim and Curtis (2002) argued that English to Korean /s/ adaptation is sensitive to fricative duration (i.e., duration-sensitive view). Kang (2008), however, argued that fricative duration is not a reliable cue for Korean listeners (i.e., duration-insensitive view). These views are discussed in detail in Chapter II and III.

[1] Refer to the national institute of the Korean language (www.korean.go.kr) for the notational system. Actual pronunciations are based on the usage in broadcast systems such as TV and radio.

Table I.1 English /s/ words Borrowed into Korean (top) and General Patterns (bottom)

English words	Korean orthography (Notational system)[2]		Pronunciation (Phonetic transcription)
say	세이	sei	[s*e.i]
site	사이트	sɑitɨ	[s*ɑ.i.tɨ]
Miss	미스	misɨ	[mi.s*ɨ]
stay	스테이	sɨtei	[sɨ.te.i]
smog	스모그	sɨmokɨ	[sɨ.mo.kɨ]
slow	슬로우	sɨlou	[sɨl.lo.u]
mist	미스트	misɨtɨ	[mi.sɨ.tɨ]

word-initial		word-final		
prevocalic (i.e., $_{\#}$[s])[3]	preconsonantal (i.e., $_{\#}$[sC)	postvocalic (i.e., s]$_{\#}$)	postvocalic *or* preconsonantal (i.e., sC]$_{\#}$)	postconsonantal (i.e., Cs]$_{\#}$)
say	*stay*	*gas*	*guest*	*box*
s*	s	s*	s	s*

2 Korean has prescriptive rules regarding the notational system of loanwords. The rules prescribe that English /s/ is written using a lax /s/$_K$ regardless of whether the /s/ is prevocalic or preconsonantal. Following the prescriptive notational system, therefore, all English /s/ words are written as illustrated in Table I.1 using the symbol for lax /s/. However, due to the difference between the system and the actual sounds, prevocalic /s/ is often written using the symbol for /s*/$_K$, which results in more than one possible spelling for prevocalic /s/ loanwords.

3 The symbol '$_{\#}$[' or ']$_{\#}$' indicates a word boundary.

1.3 Loanword Adaptation and Models of Cross-language Speech Perception

1.3.1 Loanword adaptation as a window on cross-language speech perception

The task of the listener in studies investigating perception of non-native speech sounds bears some similarities to the task of language users in loanword adaptation. When language users adapt the loans to bring them into accordance with their native phonological system, the borrowed linguistic forms (i.e., loanword representations) presumably serve as evidence about listeners' perception. Of course, perception need not be the sole determinant of loanword adaptation patterns. Other factors have also been found to be involved in loanword adaptation, although they are considered as being less influential than perceptual factors (Peperkamp and Dupoux, 2002; Peperkamp, 2005; Davidson, 2007). Factors such as orthography (Vendelin and Peperkamp, 2005) and proficiency of borrowers in the donor language (Paradis and LaCharité and; 2008, 2009) have been shown to influence how loanwords are adapted.

In this dissertation, I will assume loanword adaptation to be a window on non-native speech perception. I will assume that an orthographic influence is not relevant here because, orthographically, English /s/ is mostly represented as "s", with just a few classes of exceptions such as 'c' followed by 'I' (e.g., *'city'*, *'citation'* and *'suicide'*). Also, every English /s/-word that I discuss in this dissertation is represented only with the spelling "s". For these words, if different English /s/s are perceived differently by Korean listeners, orthography cannot be the reason for the difference.

In addition, I will investigate the possibility that the English proficiency of Korean listeners may influence how they perceive $/s/_E$. In all of the perception experiments reported, Korean listeners with different degrees of English proficiency and exposure will be compared to each other.

1.3.2 Cross-language speech perception model: Best's Perceptual Assimilation Model

The central concern of this dissertation is to examine how Korean listeners perceive English /s/ in relation to their native phonological system. Assuming that Korean listeners map English /s/ onto similar-sounding Korean fricatives, Korean listeners have to determine the phonetic representation in their native phoneme inventory that most closely matches the non-native speech signal.

The idea that listeners map a non-native speech sound onto the closest native representation is not new. A great deal of research has been devoted to this mapping relation. One of the most influential formal models to account for the mapping relation between the non-native speech sounds and the native phonemic category is the Perceptual Assimilation Model (PAM; Best, 1995). The main claim of PAM is that listeners detect gestural/acoustic similarities between native and non-native phones and perceptually *assimilate L2 sounds to the closest L1 category*. Specifically, non-native sounds "tend to be perceived according to their similarities to, and discrepancies from, the native segmental constellations that are in closest proximity to them in native phonological space" (Best, 1995:193).

In PAM, non-native sounds are grouped into three categories according to their degree of assimilation to native categories: assimilated to a native category (i-iii, below), assimilated as an uncategorizable speech sound (iv-v), and not assimilated to speech (v) (Best, 1995:194-195). These three categories generate six possible assimilation patterns (or types): (i) Two-Category assimilation, (ii) Category-Goodness Difference, (iii) Single-Category assimilation, (iv) Uncategorizable vs. Uncategorizable, (v) Uncategorizable vs. Categorizable, and (vi) Non-assimilable patterns. These six types result from the interactions of the degree of assimilation along with whether the native contrasts are preserved or not and with whether the assimilation is *ideal* or not.

According to this model, a Two-Category assimilation type for a non-native contrast is considered to be excellent assimilation because each non-native sound is mapped onto a different native category, thus, the phonemic contrast is preserved. If the two non-native sounds are assimilated to the same native category, and if these assimilations are equally good or poor (i.e., the two non-native sounds are equally similar to the one native category), it is a Single-Category assimilation type, whereas if one is a better match that the other, it is a Category-Goodness Difference type. These first three types consistently identify non-native sounds with native categories. The next two cases involve instances in which a non-native sound is not mapped onto any native category, but is still recognized as speech and hence mapped into the part of the acoustic space that represents speech. This type of mapping can involve either one non-native sound (i.e., Uncategorizable vs. Categorizable) or two (i.e., Uncategorizable vs. Uncategorizable). These uncategorizable mappings are possible only as long as the non-native sounds are perceived as speech sounds. If non-native sounds are perceived as non-speech, the non-native phones fall outside of speech domain, which would result in a Non-assimilable pattern. The following examples illustrate these patterns.

(i) Two-Category assimilation type: The Ethiopian ejectives [p’] and [t’] are assimilated by English listeners as English [p] and [t] respectively (Best, 1990).

(ii) Category-Goodness Difference type: The Farsi velar [g] and uvular [G] contrast is assimilated to one category [g] by English listeners, but Farsi [g] is considered to be a better match than Farsi [G] (Best *et al.*, 1988).

(iii) Single-Category assimilation type: Zulu aspirated voiceless velar [k^h] and ejective [k’] are assimilated into one phoneme category [k] by English listeners (Best *et al.*, 1988).

(iv) Uncategorizable vs. Uncategorizable type: No example of this type was found in the literature.

(v) Uncategorizable vs. Categorizable type: English listeners perceive Norwegian [i] as English [i]. They assimilate Norwegian high central [ʉ] either as [u] or [ʊ], however, neither [u] nor [ʊ] is similar to the Norwegian [ʉ] (Schultheiss, 2008).

(vi) Non-assimilable type: Clicks in Zulu share no or little speech-like properties with English (to English listeners) although the unfamiliar phonemic contrasts are easily discriminated by the listeners (Best *et al.*, 1988).

Concerning alveolar fricative adaptation from English to Korean, Korean listeners definitely assimilate English $/s/_E$ to their native categories $/s/_K$ or $/s^*/_K$. All of the assimilation types included in PAM have to do with the preservation (or loss) of a phonemic contrast during cross-language perception. In the case that I am studying in this dissertation, however, a single category, $/s/_E$, is split into two separate categories, $/s/_K$ and $/s^*/_K$, during perception. This, therefore, does not seem to fit into any of the types included in PAM.

Nevertheless, the prediction for the English-to-Korean situation in a model like PAM can be considered. This situation is perhaps best described as a "reverse Single-Category assimilation" type. Since one alveolar fricative of English /s/ is variously mapped into two alveolar fricatives in Korean $/s/_K$ and $/s^*/_K$, a single category is split into two distinct phoneme categories, that is, the reverse of Single-Category assimilation. Additionally, if the non-native perception were turned around to be Korean-to-English, it would be a Single-Category assimilation type, because the two different Korean phonemes $/s/_K$ and $/s^*/_K$ are likely to perceived as a single category, $/s/_E$.

Depending on listeners' language proficiency, English /s/ can be considered as two different categories. According to Best and Tyler (2007), naïve non-native listeners would not know whether two phones are contrastive in L1 or not. In this case, borrowing English $/sV/_E$ and $/sCV/_E$ as $/s^*/_K$ and $/s/_K$ would be viewed as a two-category assimilation. This approach assumes that the fricatives in the two $/sV/_E$ and $/sCV/_E$ contexts have systematically different properties.

1.4 Hypotheses

The purpose of this dissertation is to examine the nature of English loans that contain /s/. In the above section, I discussed English to Korean /s/ loans, and a model of loanword adaptation. That discussion generates the central research question of this dissertation: what are the principles that govern the adaptation of English loanwords into Korean? That is, how do Korean speakers decide which English /s/'s are borrowed as $/s/_K$, and which as $/s^*/_K$ in Korean? To what degree is loanword adaptation a phonetic (i.e., determined by articulatory and/or acoustic properties) or phonological (i.e., determined by constraints or rules governing structural markedness) process in Korean?

In this dissertation, I argue that the process by which /s/ loans are adapted into Korean is both phonetic and phonological (phonotactic): English /s/ adaptation is a phonetically motivated phonological category mapping process. In order to explain what motivates the borrowing process of $/s/_E$ to $/s^*/_K$ and $/sC/_E$ to $/s/_K$, I suggest and experimentally explore two possible hypotheses.

The fact that Koreans adapt the production of $/s/_E$ by English speakers in different phonological contexts as different phonemes suggests that $/s/_E$ may have systematically different acoustic properties in the two phonological contexts (a singleton $/s/_E$ and a cluster $/s/_E$), and that Korean listeners rely on these systematic acoustic differences. Hypothetically, if such differences in English correspond to the phonetic details of the Korean contrast between $/s/_K$ and $/s^*/_K$, then Korean listeners would perceive a prevocalic $/s/_E$ and a preconsonantal $/s/_E$ differently.

When considering this hypothesis about the influence of the acoustic properties of English /s/, we can consider both consonant-internal properties (i.e., the properties of the frication noise itself) and the properties of the vowel following the fricative (which may be partially determined by the properties of $/s/_E$ due to coarticulation). I consider these two issues separately.

The "Consonantal Information Hypothesis" (Hypothesis 1 or H_1) can be stated as follows: A prevocalic $/s/_E$ is acoustically similar to$/s^*/_K$, and a preconsonantal $/sC/_E$ is similar to $/s/_K$. In other words, English /s/ in a singleton [sV_] condition acoustically corresponds to Korean /s*/ and English /s/ in a cluster [sCV_] or [sCCV_] acoustically corresponds to Korean /s/. Here, 'acoustically corresponds' means 'shares similar acoustic properties'. Thus, this hypothesis predicts that, in terms of production, the acoustic properties of a prevocalic $/s/_E$ resemble $/s^*/_K$ and those of a preconsonantal $/s/_E$ resemble $/s^*/_K$ (H_1a). Consequently, Korean listeners perceive the voiceless frication in [sV_] as a realization of $/s^*/_K$, and in [sCV_] as a realization of $/s/_K$ (H_1b).

The "Vocalic Information Hypothesis" (Hypothesis 2 or H_2) claims that Korean listeners focus on vocalic information rather than consonantal information. Under this hypothesis, acoustic similarity is based on the vocalic properties such that, in production, the properties of the vowels following prevocalic $/s/_E$ acoustically resemble those following $/s^*/_K$ and the vocalic properties of English preconsonantal $/sC/_E$ acoustically correspond to Korean $/s^*V/_K$ (H_2a). Thus, in perception, Korean listeners perceive $/sV/_E$ as similar to $/s^*V/_K$ and $/sCV/_E$ to $/sV/_E$ (H_2b).

Although prevocalic $/s/_E$ is a permissible sequence in Korean, preconsonantal $/sC/_E$ is an illegal sequence in Korean because no tauto-syllabic consonant cluster is allowed on the surface. This phonotactic violation is resolved with vowel epenthesis, which results in a phonological repair. As shown above in Table I.1, the consonant cluster is broken into different syllables with an epenthetic [ɨ] (i.e., $/sCV/_E \rightarrow [\text{sɨCV}]_K$). The phonological repair emerges only when the phonotactic constraints are violated, thus, prevocalic $/s/_E$ to $/s^*/_K$ borrowing is unaffected by the phonological repair.

Due to vowel epenthesis in Korean perception of English speakers, the Vocalic Information Hypothesis applies to preconsonantal $/sC/_E$ as well as prevocalic $/s/_E$. While the vowel in prevocalic $/sV/_E$ is an actual full vowel, the vowel in a preconsonantal context is argued to be perceptually epenthesized by Korean listeners due to phonological modification. This vowel, in turn, produced by Korean listeners-turned-speakers. Under H_2, it is therefore necessary to investigate the vocalic properties of the epenthetic /ɨ/.

Both H_1 and H_2 postulate that Korean listeners are sensitive to the detailed acoustic properties of alveolar fricatives in Korean and English. The two acoustic similarity hypotheses suggest that Korean-speaking listeners' knowledge of the phonetic properties of their native sounds leads them to perceive a two-way distinction between context-specific characteristics of English $/s/_E$, which are systematically present in the production of English speakers.

The two hypotheses are summarized in Table I.2. These hypotheses are experimentally tested in the rest of this dissertation.

Table I.2 Summary of the Hypotheses

Hypothesis 1 (H_1): ***Consonantal Information Hypothesis***	**H_1a**	✓ $/sV/_E$: Frication properties of $/sV/_E$ are similar to those of $/s^*/_K$. ✓ $/sCV/_E$: Frication properties of $/sC/_E$ are similar to those of $/s/_K$.
	H_1b	✓ Korean listeners are sensitive to this consonantal information.
Hypothesis 2 (H_2): ***Vocalic Information Hypothesis***	**H_2a**	✓ $/sV/_E$: Acoustic vowel properties of $/sV/_E$ are similar to those of $/s^*V/_K$. ✓ $/sC/_E$: Perceptually epenthesized [ɨ] properties of $/sCV/_E$ (perceptually /sɨCV/) are similar to those of $/sV/_K$.
	H_2b	✓ Korean listeners are sensitive to this vocalic information.

1.5 Structure of the Dissertation and the Summary of the Chapter

This dissertation investigates cross-language speech perception and loanword adaptation, with a focus on the influence of listeners' native phonological system on non-native borrowings. Specifically, this dissertation investigates the following questions as applied to borrowing English /s/ into Korean: (i) What are the acoustic attributes of $/s/_K$ and $/s^*/_K$ on which Korean listeners rely in perception? (ii) What are the acoustic attributes of $/s/_E$ in difference contexts? (iii) Are Korean listeners sensitive to the potentially different acoustic properties in different contexts of the non-native sounds?

In this dissertation, I address these research questions, considering the implications for speech perception theories. In investigating these issues, this dissertation is organized as follows: Chapter II investigates the production and perception of voiceless alveolar fricatives in English (Experiment 1) and Korean (Experiment 2), focusing on an examination of Hypothesis 1. Using recorded productions of alveolar fricative utterances from each language, the consonantal properties of voiceless alveolar fricatives in English and Korean are acoustically examined. In addition, the role of consonantal information in /s/ borrowing is assessed in a perception experiment in which Korean listeners respond to English-like non-words in which the targeted consonantal properties of /s/ are systematically manipulated (Experiment 3).

In Chapter III, recordings from the two production experiments (i.e., Experiment 1 and 2) are reexamined in terms of their vocalic properties. In order to explore Hypothesis 2, the acoustic properties of the post-fricative vowels in English and Korean are investigated and compared. Chapter III also reports the results of a perception experiment that tested Korean listeners' sensitivity to consonantal/vocalic information. In the experiment, Korean listeners respond to English non-words whose vocalic properties are kept constant while consonantal portions of clusters, stops in particular, are slightly modified (Experiment 4).

Chapter IV focuses on the investigation of Hypothesis 2 as well. While addressing the durational properties of the full vowels in the two production experiments (i.e., Experiments 1 and 2) and the epenthetic vowel in non-words (Experiment 5) and English loans (Experiment 6), I show that the epenthetic vowel is minimally disruptive in perception. This claim about the perceptual modification caused by the epenthetic vowel is assessed in a perception experiment in which Korean listeners are asked to judge the perceptual similarity between tokens with and without an epenthetic vowel (Experiment 7).

Finally, Chapter V discusses the implications and consequences of the results of the experiments with reference to theories of speech perception. In this chapter, a theoretical model regarding cross-language speech perception is considered. The nature of English alveolar fricatives borrowing patterns, whether they are phonetic or phonological in nature, will be discussed. In the same chapter, I conclude by summarizing new findings about cross-language speech perception and by discussing the contribution of this dissertation.

CHAPTER II
CONSONANTAL PROPERTIES AND PERCEPTION OF CONSONANTAL INFORMATION

2.1 Introduction

This chapter examines whether consonantal information serves as a useful cue in English /s/ adaptation. The consonantal portion specifically refers to the frication itself. Thus, consonantal information, as considered here, is restricted to aperiodic properties of the signal. In this chapter, I investigate the temporal and spectral characteristics of the noise in English and Korean voiceless alveolar fricatives and whether these properties might play an important role in the adaptation patterns. Specifically, after reviewing how the acoustic properties of voiceless alveolar fricatives of English and Korean have been described in the literature, I present two production studies and one perception study.

Experiment 1 investigates the acoustic properties of English /s/ in different phonological contexts. Experiment 2 examines the acoustic properties of Korean voiceless alveolar fricatives. Based on the findings of Experiments 1 and 2, I investigate in Experiment 3 whether or not Korean listeners actively employ the consonantal information in their perception of English fricatives. The results of these experiments will provide us with an understanding of the acoustic characteristics of voiceless alveolar fricatives in the two languages, and of Korean listeners' sensitivity to consonantal information.

2.1.1 Hypothesis 1 (H_1): Consonantal Information Hypothesis

The hypotheses stated in Chapter I postulate that the loanword pattern is governed by acoustic similarity. H_1 states that the acoustic properties of the frication portion of a singleton $/s/_E$ (i.e., [sV]) are similar to the acoustic properties of $/s^*/_K$, and that cluster $/s/_E$ (i.e., [sCV]) has acoustics similar to $/s/_K$. Thus, Korean listeners perceive the singleton $/s/_E$ sounds as $/s^*/_K$, and the cluster $/s/_E$ sounds as $/s/_K$. H_1 is schematized in Figure II.1, which describes $[sV]_E$ to $[s^*]_K$ and $[sCV]_E$ to $[s]_K$ mappings.

Figure II.1 The mapping pattern of fricatives in English and Korean based on consonantal information

$$\begin{array}{lcc} \text{English:} & [\underline{\mathbf{s}}V]_E & [\underline{\mathbf{s}}CV]_E \\ & \downarrow & \downarrow \\ \text{Korean:} & [s^*]_K & [s]_K \end{array}$$

H_1 can be further divided into production and perception sub-hypotheses: in terms of production, the hypothesis is that the acoustic properties of English prevocalic $/s/_E$ are similar to Korean [s*] and those of English preconsonantal $/sC/_E$ are similar to Korean $[s]_K$ (H_1a), and in terms of perception, the hypothesis is that Korean listeners rely on the consonantal properties in their perception (H_1b).

In the experiments conducted here, H_1a is supported if the acoustic properties of the frication portion in $[sV]_E$ are similar to those of $/s^*/_K$, and the acoustic properties of the frication portion in $[sCV]_E$ are a better match for $/s/_K$ properties. H_1b is supported if Korean listeners actively use the consonantal information in their $/s/_E$ perception. Let us begin by examining the acoustic characteristics of /s/ in both languages.

2.1.2 Fricative measurements

The primary temporal measure is duration of the frication. Spectral characteristics typically analyzed for alveolar fricatives are the spectral shape of the fricative noise, the slope of F_2 transitions into following vowels, the CoG (center of gravity) of the fricative noise, and the breathiness of the following vowels (Jongman, 1989; Jongman *et al.*, 2000; Evers *et al.*, 1998; Ali *et al.*, 2001; Cho *et al.*, 2001, 2002; Hall *et al.*, 2006; cf. Maniwa *et al.*, 2009). CoG expresses the mean frequency of fricative noise for a given slice in time. Other acoustic measures regarding energy are normalized RMS (root mean square) amplitude, relative intensity of the fricative in relation to intensity of the following segment (Hendrick and Ohde, 1993), and overall normalized intensity.

The current study investigates fricative duration and fricative CoG for English and Korean alveolar fricatives. These are both fricative-internal properties. Fricative-external properties are considered in Chapter III.

2.2 Review of the Literature on the Acoustic Properties of English /s/

2.2.1 Duration of $/s/_E$

Results of early studies of consonant duration showed that the acoustic characteristics of a target segment vary significantly depending on its phonological context (e.g., Cooper *et al.*, 1952; Delattre *et al.*, 1955, among many others). For example, the acoustic characteristics of a segment may depend on its position in prosodic and morphological domains (Klatt, 1974), stress (Klatt, 1974), number of syllables in a word (Barnwell, 1971), phonological contexts (Schawartz, 1969) and syllable length (Schwartz, 1970), among others (Klatt, 1973).

Much research has been devoted to domain effects (Keating *et al.*, 1998; Byrd *et al.*, 2000, 2005; among others), and it has generally been found that sounds are longer in domain-final position (e.g., syllable-final, word-final or phrase-final) than in domain-initial position. This finding of final lengthening has consequences for our discussion in that $/s/_E$ in each domain-final position may be longer than the $/s/_E$ in domain-initial position. Regarding English /s/, Klatt (1974) showed that in primary stressed conditions, word-initial [s] are generally longer than intervocalic [s] (around 145 ms vs. 135ms).
in unstressed syllables was 112 ms.

Concerning the phonetic consequences of stress, it is widely acknowledged that a syllable that bears the main stress of a word is lengthened. D'Imperio and Rosenthall (1999), for instance, showed that syllables with stressed vowels are longer than ones with unstressed vowels. When the stress-bearing syllable becomes longer, it is primarily vowel duration that increases. However, it is not unusual for consonants in the stress-bearing syllable to become longer as well. With regard to English /s/, Klatt (1974) showed a stress effect wherein $/s/_E$ in the stressed syllable (e.g., '*sixteen*') was 15% longer than $/s/_E$ in an unstressed syllable (e.g., '*support*') in a word (around 127 ms vs. 112 ms).

The number of syllables in a word (i.e., word-length) has been recognized as a factor that influences the duration of a consonant. It is usually the case that the duration of each segment is shorter in a longer word than in a shorter word. Barnwell (1971), for instance, observed that the duration of a vowel in disyllabic words was 70% of that in monosyllabic words, and that the duration was reduced to 50% in multisyllabic words. This number-of-syllables effect was also found for $[s]_E$. According to Klatt (1974), within the stressed syllable condition, the duration of $[s]_E$ was reduced to 85% in disyllabic (around 127 ms) and 80% in trisyllabic words (around 119 ms), compared to monosyllabic words (around 145 ms).

The phonological context in which prevocalic $/s/_E$ appears is also known to influence its duration. Schwartz (1969) measured the duration of intervocalic /s/ by varying the vowels flanking the fricatives and found that the duration of /s/ was generally longer before high vowels such as [i] than before low vowels such as [a] (e.g., the duration of [isi] was 181 ms and the duration of [isa] was 167 ms). The difference of the duration was up to approximately 15%.

With respect to syllable length (i.e., the number of segments in one syllable), it has been found that segments in a monosyllabic word with fewer segments tend to be longer than those in a word with more segments; the so called "cluster effect". Cluster effects have been found in several experimental studies (Haggard, 1971; Klatt, 1973, 1974; Lee and Iverson, 2007). For instance, Schwartz (1970) consistently found that $[s]_E$ in a cluster was shorter than $[s]_E$ in a singleton. He found that the average duration of [s] in /sV/ was 177 ms and that in /sCV/ was 143 ms. He also found that the fricative duration of a cluster $[sCV]_E$ was up to 40% shorter than that of a singleton $[s]_E$ showing that the duration of $[s]_E$ in /asi/ was 188 ms and that in /aspa/ was 117 ms. Kim (1999) also showed that the duration of $[s]_E$ in a singleton was 22% longer than the duration of $[s]_E$ in a cluster (i.e., the average of $[s]_E$ in [sV] was 170 ms and that in [sCV] 133 ms).

Kang (2008), however, argues that a general cluster effect may not be as important as the precise context in which $[s]_E$ appears. She suggests that the effects are most pronounced when one compares $[s]_E$ before a vowel and $[s]_E$ before a stop consonant, and that the duration of $[s]_E$ before other consonants such as [w], [n], or [l] may not be significantly different from that before a vowel.

The discussion of $[s]_E$ so far is has been concerned with the durational property of $[s]_E$, and it was shown that many previous studies found the duration of $[s]_E$ varies significantly according to its phonological environment. Let us now turn to our discussion of a spectral attribute, CoG, of $/s/_E$.

2.2.2 Center of gravity of $/s/_E$

Center of Gravity, or spectral mean (within a specified window of time), is the most common spectral measurement used by speech researchers studying fricatives (Nowak, 2006; Gordon *et al.*, 2002; Hamann and Avelino, 2007; Zygis, 2003). CoG correlates with the size of the cavity in front of the constriction: the farther forward the constriction, the smaller the cavity, and the higher the CoG. CoG has the advantage of characterizing a fricative's spectral shape in a single value. In a study of sibilant fricatives in Slavic languages, Zygis (2003) found CoG to be a reliable measurement for distinguishing between different fricative phonemes. In another cross-linguistic study of fricatives, Gordon *et al.* (2002:166) obtained similar results: "gravity center frequencies robustly differentiated many of the fricatives in the examined languages."

The CoG has been examined in several languages (e.g., Turkish Kanardian in Gordon and Applebaum, 2006; Aleut, Apache, Chockasaw, Gaelic, Hupa, Montana Salish, Toda in Gordon *et al.*, 2002). However, little attention has been given to English fricatives in various phonological contexts. A notable exception is Kang (2008) who examined the mean CoG of English word-initial /s/. In her measurements, the CoG values for prevocalic $/s/_E$ were slightly higher than those for preconsonantal $/sC/_E$, although both CoG values of a singleton /s/ and a cluster /s/ fell between 6000 Hz and 6500 Hz. As Kang pointed out, 'the difference is neither systematic nor significant' (Kang, 2008:10).

However, Kang's measurements are problematic. Firstly, the source data she used (American English Spoken Lexicon, AESL, found at the following URL: https://online.ldc.upenn.edu/aesl/) seem to have been recorded so that the highest analyzed frequency was 8000 Hz. Secondly, flanking vowel quality was apparently not controlled. However, the high F_2 of front vowels and low F_2 of back vowels may exert coarticulatory influences on the frequency of the preceding /s/. More rigorous study is needed to determine whether there is a mean frequency difference between a prevocalic and a preconsonantal /s/.

2.3 Review of the Literature on the Acoustic Properties of Korean /s/ and /s*/

The acoustic differences between [s] and [s*] in Korean have been examined in several studies. In the following, I review the acoustic properties of Korean /s/ and /s*/ focusing on the two /s/-internal cues, fricative duration and CoG.

2.3.1 Duration of $/s/_K$ and $/s^*/_K$

Some researchers argue that frication duration is one of the main acoustic characteristics that distinguishes Korean fricatives, with $[s^*]_K$ being longer than $[s]_K$. According to Kim and Curtis (2002), word-initially, there was around a 60 ms durational difference between the tense and lax fricatives (i.e., $[s^*]_K$ is 160 ms and $[s]_K$ is 100 ms long on average). Although the specific numbers vary among researchers, the overall finding is that /s*/ is longer than /s/. (See Kagaya (1974), Lee and Iverson (2007), Cheon and Anderson (2008) for similar results.) Given these robust differences, Korean speakers might be expected to use this acoustic cue as a crucial property to distinguish $[s^*]_K$ and $[s]_K$.

Unlike English, stress is not contrastive in Korean (Song, 2005). I therefore do not consider the potential effects of stress in the discussion. Also, since Korean does not allow tautosyllabic consonant clusters, the issue of cluster length is not relevant in Korean. However, as in English, it may be that the duration of $/s/_K$ and $/s^*/_K$ is influenced by domain, the number of syllables in a word, and syllable length. In one experimental study, Lee and Seong (1996) found that phrase-final syllables were significantly longer than phrase-initial syllables in Korean speakers' productions. However, it is not known whether the durations of $/s/_K$ and $/s^*/_K$ is systematically different according to the phonological contexts, that is, whether the number of syllables in a word and syllable length affect the durational properties of Korean consonants.

Cho *et al.* (2002) examined the acoustic and aerodynamic correlates of Korean fricatives (and stops) in the Seoul and Cheju dialects. Their measures of fricative durations of word-initial and word-medial $/s/_K$ and $/s^*/_K$ showed that $/s^*/_K$ was significantly longer than $/s/_K$, and that the difference was larger word-medially than word-initially (i.e., the duration of $/s^*/_K$ was around 150 ms word-initially and 180 ms word-medially, whereas the duration of $/s/_K$ was 110 ms word-initially and 90 ms word-medially). The significant durational difference was found in the Seoul dialect only.

2.3.2 Center of gravity of $/s/_K$ and $/s^*/_K$

Cho and his colleagues also examined the mean frequencies of /s/ and /s*/ in Korean. For speakers of the Seoul and Cheju dialects, /s*/ had a higher CoG value than /s/ (i.e., 6600 Hz and 6200 Hz, respectively, in the Seoul dialect), which they attributed to the relatively smaller front cavity in the articulation of $/s^*/_K$ compared to $/s/_K$.

Kang (2008) also found that the CoG of /s*/ was generally higher than that of /s/ in Korean. The difference was larger before /ɑ/ than before /i/ or /u/. Capitalizing on the understanding of the various acoustic properties of fricatives in the two languages, let us now shift our focus to the production studies, Experiments 1 and 2.

2.4 Experiment 1 – Consonantal Properties of Fricatives in English

Several previous studies have investigated acoustic motivations for the pattern of English /s/ adaptation into Korean. In this work, there are two points of view with regard to duration. One school of researchers, including Kim and Curtis (2002), Ahn and Iverson (2004), and Lee and Iverson (2007), argues that, among several acoustic characteristics, duration is the primary cue that determines how $/s/_E$ is borrowed into Korean. The other school of researchers, including Kang (2008), claims that /s/ borrowing is duration-insensitive.

These two opposite points of view – that the adaptation is duration-sensitive or insensitive – result from the disagreements among researchers regarding the durational properties of English /s/. Kim and Curtis (2002) reported that $/s/_E$ was longer when it was a singleton (e.g., *'say'*) than when it appeared as part of a consonant cluster (e.g., *'stay'*). Kang (2008), though, argued that the kind of post-fricative consonants in the cluster (i.e., [sCV]) matters. Specifically, she pointed out that, in the durational measures of Kim and Curtis (2002), the consonants that followed $/s/_E$ were always stops, which is problematic because the durational difference between a singleton $/s/_E$ and a cluster $/s/_E$ is greatest when the following consonants are stops.

Experiment 1 addresses this controversy through a thorough examination of $/s/_E$ in various consonantal contexts, including before a lateral [_l] and a labio-velar glide [_w] (e.g., '*slow*' [sloʊ] and *'sway'* [sweɪ]) These clusters, like clusters containing stops, are borrowed with $/s/_K$ (e.g., [sɨlou] and [sɨweɪ], respectively).

2.4.1 Design and methodology

2.4.1.1 Participants

Six native speakers of American English were recruited around the campus of the University of Michigan, Ann Arbor. They were native speakers of English and none of them reported hearing or speaking deficits. Participants received $15 for their participation.

2.4.1.2 Reading materials

Speech materials consisted of English monosyllabic /s/-initial words. Word-initial /s/ was followed either by a vowel or a consonant. The number of onset consonants in the target words varied from one to three as in [sV], [sCV] and [sCCV] (e.g., *say, stay, stray*). The C in [sC] included laterals (i.e., [_L]) as in *slow*, approximants (i.e., [_W]) as in *sway*, nasals (i.e., [_N]) as in *snow* or *smoke*, and stops (i.e., [_T]) as in *stay, ski* or *spy*. The C_1 in [sC_1C_2] was one of /p, t, k/ and C_2 was one of /l, r/. Example words are given in Table II.1, and Appendix 1 gives the complete wordlist. The wordlist contained some words that are not analyzed here because the recording was part of a larger experiment. These additional words did not contain word-initial [s], and therefore they also functioned as filler items for the purpose of the current experiment.

Table II.1 The target speech materials for Experiment 1

	structure		example tokens
singleton	[sV]		*say*
cluster	[sCV]	[sLV]	*slow*
		[sWV]	*sway*
		[sNV]	*snow*
		[sTV]	*spy, stay, sky*
	[sCCV]	[sTr(l)V]	*spring, split*
			stray
			scream

2.4.1.3 Procedure

The complete list contained 276 /s/-initial words in a randomized order. The words were produced in isolation. Participants read through the list five times. Across all participants, 8280 tokens were expected (i.e., 276 words × 5 repetitions × 6 speakers). However, only 8140 tokens were analyzed because (i) participants missed some words, (ii) incorrectly read tokens were excluded, and (iii) tokens that have ambiguous /s/ boundaries due to being preceded by /s/-final words (e.g., *bus stop*) were excluded as well. CoG measurements were taken only for fricatives in non-front vowel contexts, which were in 120 words.

The recording was made in a sound-attenuated booth in the sound laboratory at the University of Michigan, Ann Arbor, using an AKG microphone and the software program *Praat* (Boersma and Weenink, 2009), at a sampling rate of 44,100 Hz.

The acoustic measures of duration and CoG were made as follows: Fricative duration was the duration of the turbulent noise, measured from the onset to the offset of the friction as shown in the Figure II.2. The onset of the noise was decided by visually marking the initiation of the aperiodic portion of the sound, and the cessation of the noise was determined as the end of the aperiodic portion. The temporal span from the onset to the offset of the aperiodic portion was considered as the duration of the fricative. The fricative portion is marked with an arrow in Figure II.2. Spectral CoG is illustrated in Figure II.3. The frication portion was selected (top) and its spectral slice was framed in a 0-22050 Hz window (bottom). The CoG is the mean frequency over the entire frequency range weighted by the amplitude. Similar to other studies that measured CoG of fricatives (e.g., Hall *et al.*, 2006), the power for the weighting was 2, and CoG was measured over the full duration of frication, using *Praat* (Jassem, 1979; Jongman *et al*., 2000; Gordon *et al*., 2002).

Figure II.2 Duration measurement of friction

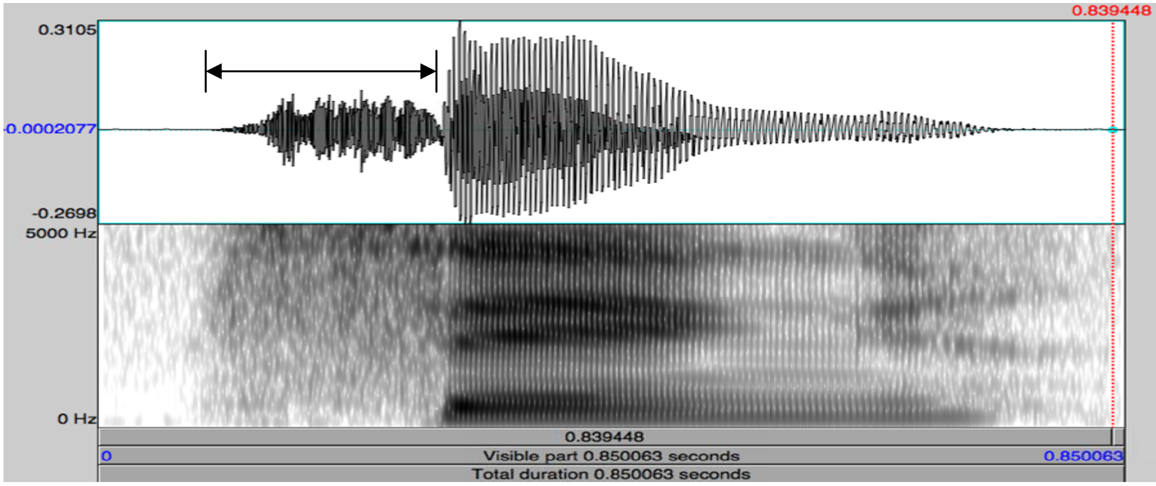
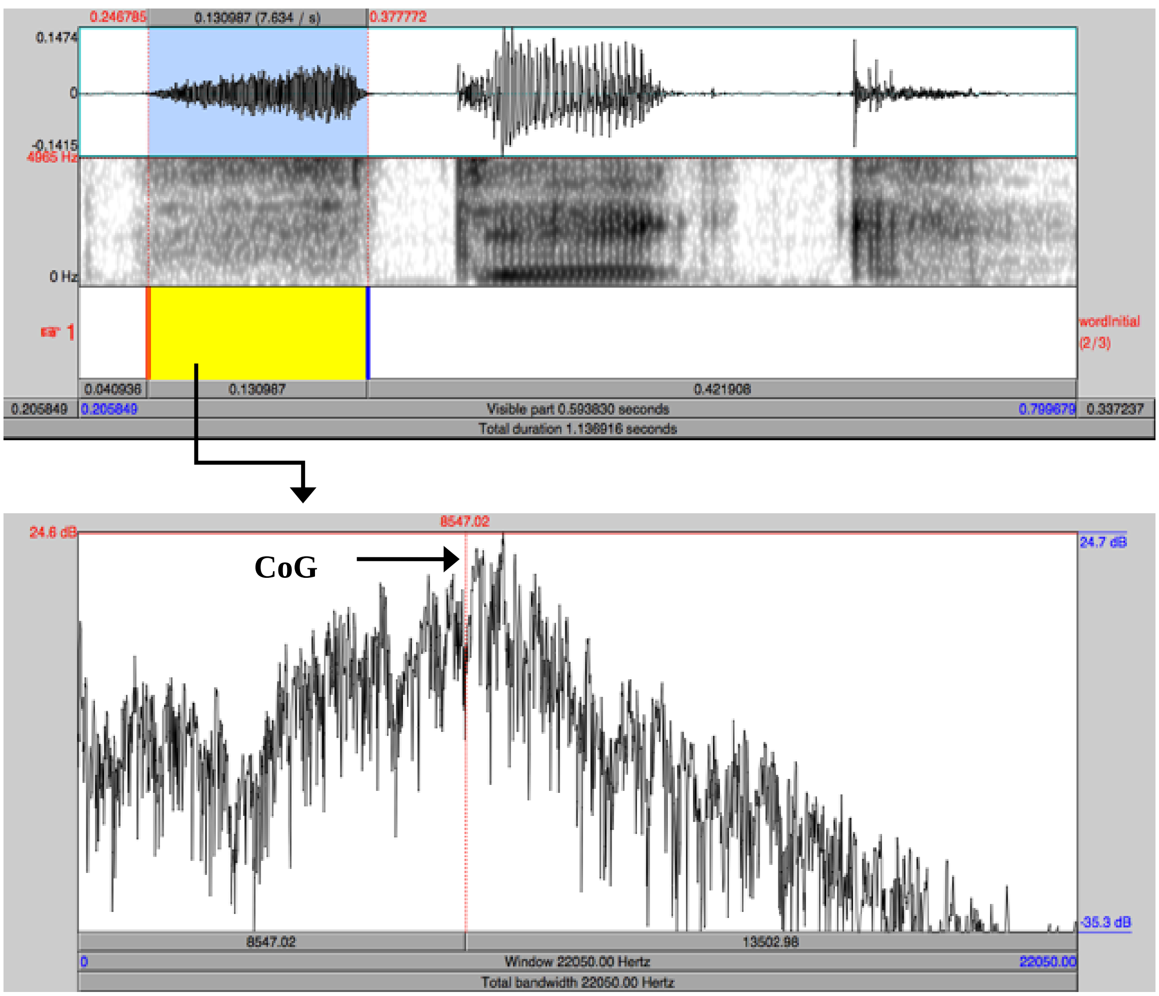


Figure II.3 CoG measurement of friction



2.4.2 Results

Table II.2 gives the average durational data for English [s] in various phonological contexts for the six participants.

Table II.2 Average duration of [s] in English in various contexts (ms)

subj	[sV]	[sCV] ([skV], [spV], [stV], [slV], [smV], [snV], [swV])	[sCCV]
S1	164	147 (140, 128, 146, 156, 149, 169, 165)	135
S2	152	138 (137, 127, 131, 146, 139, 147, 159)	128
S3	192	182 (181, 168, 188, 174, 177, 190, 205)	177
S4	202	173 (170, 151, 172, 182, 176, 192, 200)	156
S5	198	178 (175, 159, 183, 179, 182, 200, 188)	154
S6	154	142 (146, 124, 147, 136, 132, 146, 173)	131
Total	177	160	147

Prevocalic $/s/_E$ was found to be generally longer than the preconsonantal $/sC/_E$. On average, the number of onset consonants was inversely related to the duration of [s]. The prevocalic /s/ was longest (177 ms). Between the preconsonantal /s/s, [sCV] had longer duration (160 ms) than [sCCV] (147 ms). In order to examine the data by item, the duration means for each type of [s] were submitted to a repeated measures ANOVA with one three-level factor, the context of /s/ ([sV], [sCV], and [sCCV]). The main effect of context was significant (F (2, 273) = 74.051, $p < 0.001$). Tukey's post-hoc tests showed that all three categories were significantly different from each other. Due to the small number of participants (six), a by-subject analysis was not conducted.

A repeated measures ANOVA was also conducted on the /sCV/ subset of the data for the seven-level factor following /C/. A by-item analysis showed that duration differences across the [sCV] contexts (i.e., [skV], [slV], [smV], [snV], [spV], [stV] and [swV]) was significant (F (6, 138) = 38.707, $p < 0.001$). The average duration for each condition is given in Table II.3.

Table II.3 Mean Fricative Duration for Each [sCV] Condition

condition	Duration (ms)
[skV]	158
[slV]	162
[smV]	159
[snV]	174
[spV]	142
[stV]	161
[swV]	182
TOTAL	160

Tukey's post-hoc tests confirmed that the contrasts [sk~sn], [sk~sp], [sk~sw], [sm~sp], [sl~sp], [sl~sn], [sl~sw], [sm~sw], [sp~sw], [st~sw], [sn~sm], [sp~sn], [sn~st] and [sp~st] were all significantly different.

The CoG was also measured in the three contexts ([sV], [sCV], [sCCV]). As described above, because of the potential influence of vowel backness on the CoG measures, only words with non-front vowels were considered in the [sV] condition. Table II.4 shows the CoG results of individual speakers and averaged values across speakers.

Table II.4 Average CoG of [s] in English in various contexts (Hz)

subj	[sV]	[sCV] ([skV], [spV], [stV], [slV], [smV], [snV], [swV])	[sCCV]
S1	7469	7146 (7434, 7201, 6662, 7401, 7663, 6774, 6767)	6830
S2	7116	6895 (7108, 6841, 6777, 7379, 6913, 6489, 6024)	6636
S3	7837	7675 (7558, 8099, 7354, 7583, 8267, 7408, 7043)	7117
S4	8304	7867 (8217, 7470, 7941, 7849, 7847, 8009, 6964)	7132
S5	8634	8608 (8450, 8394, 8224, 8287, 8719, 8433, 8397)	7774
S6	7010	6933 (6791, 7106, 6857, 6892, 7094, 6879, 7162)	6793
Total	7302	6946 (7577, 7500, 7305, 7564, 7748, 7332, 7059)	6437

An ANOVA comparing the CoG of $[s]_E$ in the three contexts showed a significant difference between the contexts in a by-item analysis (F (2, 117) = 36.168, $p < 0.001$). (CoG measures were not taken for front vowel items which led to smaller degrees of freedom than for the duration analyses.) Tukey's post-hoc tests confirmed that all contrasts were significant. In these data, the average CoG value for a singleton $[s]_E$ was consistently higher than that for a cluster $[s]_E$. The average CoG decreased as the cluster became more complex. In general, the average CoG differences between [sV] and [sCV] held true for all participants, although the difference was much smaller for some speakers than for others.

A repeated measures ANOVA was conducted for the CoG values for [sCV] contexts. Since there was only one token for the [sWV] condition, [sWV] was excluded and the ANOVA was run on the remaining comparisons (i.e., the six-level factor [sCV] context). A by-item analysis showed that CoG differences across the [sCV] contexts (i.e., [skV], [slV], [smV], [snV], [spV] and [stV]) were significant (F (5, 52) = 3.762, p = 0.006). The average CoG for each condition is given in Table II.5. Tukey's post-hoc tests showed that the contrast [skV~smV] was significantly different (p = 0.005), all but all other contrasts were not significantly different ($p > 0.05$).

Table II.5 Fricative CoG for Each [sCV] Condition

condition	CoG (Hz)
[skV]	7578
[slV]	7559
[smV]	7749
[snV]	7331
[spV]	7497
[stV]	7301
TOTAL	7486

2.4.3 Discussion

The goal of this production experiment was to examine the acoustic properties of English /s/ in various phonological contexts. Regarding duration, although previous researchers disagree on the nature of the influence of different consonants in [sCV] clusters on [s] duration, they agree that a singleton $/s/_E$ is generally longer than a cluster $/s/_E$. It was, therefore, expected that there would be a significant durational difference between a singleton $/s/_E$ and a cluster $/s/_E$. The results of Experiment 1 confirmed that the duration of /s/ gradually decreased from [sV], [sCV] and [sCCV].

Although the mean duration of a cluster $/s/_E$ was overall shorter than that of a singleton $/s/_E$, the durational difference varied significantly depending on which consonant followed the $/s/_E$, as shown in Table II.3. Within the set of stop consonants (i.e., [sTV] = [stV], [skV] and [spV]]), [spV] was significantly shorter than [stV]. Similarly, within nasal consonants (i.e., [sNV] = [smV] and [snV]), [smV] was significantly shorter than [snV]. In terms of consonant manner, the duration of [s] in [sTV] was the shortest, so that the durational difference between [sTV] and [sV] was larger than for any other consonant clusters. The duration of [s] in [sNV] and [sLV] was still shorter than that in [sV]. The [s] in [sW] was slightly longer than [s] in [sV]. The durations of [sWV] and [sV] are similar presumably because English /w/ is an approximant that is articulatorily and acoustically similar to vowels.

Experiment 1 confirmed that the CoG values of $/s/_E$ in a singleton and in a cluster were significantly different. The CoG values were highest in the singleton /s/ and became lower when /s/ was embedded in a cluster. The values of clusters were consistently lower than the value of a singleton, regardless of what kind of consonants followed /s/. This experimental finding raises the possibility that the frequency difference might be an important acoustic cue that Korean listeners pay attention to in their English perception.

The following table shows the comparison of results between previous studies and Experiment 1 regarding duration and CoG. The ✓ marks indicate that a significant difference between prevocalic $/s/_E$ and preconsonantal $/sC/_E$ was found (in the same direction), and the × marks indicate that a significant difference was not found. Klatt (1973, 1977) and Kim (2008) showed a significant durational difference, which was also found in Experiment 1. CoG was examined only in Kang (2008) in which a significant difference was not found.

Table II.6 Comparison of Duration and CoG between Experiment 1 (English production) and Previous Studies

	Duration	CoG
Klatt (1973, 1977)	✓	NA
Kim (2008)	✓	NA
Kang (2008)	×	×
Ahn (Experiment 1)	✓	✓

2.5 Experiment 2 - Consonantal Properties of Fricatives in Korean

2.5.1 Design and methodology

Experiment 2 investigated the acoustic properties of Korean /s/ and /s*/.

2.5.1.1 Participants

Six native speakers of Korean were recruited from Ann Arbor, Michigan. All participants spoke the same dialect, Seoul Korean. In order to control for potential speaker variance due to different degrees of experience with English, participants were selected to have roughly the same amount of English exposure. Specifically, all participants moved to the US when they were 18 or older, and all had been in the US between 5 and 10 years at the time of recording. Participants had normal hearing and no speech disorders. Participants received $15 per hour for their participation.

2.5.1.2 Reading materials

Participants were asked to read a list of words written in Korean script. Sample target words are given in the Table II.6. Words were bisyllabic and their syllable structures were either C_1V_1.CV(C) or C_1V_1C.CV(C). C_1 was either $/s/_K$ or $/s^*/_K$, and V_1 was one of four non-front vowels, [ɑ], [ʌ], [o] or [u]. $/s/_K$ and $/s^*/_K$ before [i] were not included because their acoustic realizations corresponds to [ʃ] rather than [s] of English (i.e., $[ʃ]_K$ is an allophone of $/s/_K$ and $[ʃ^*]_K$ is an allophone of $/s^*/_K$). There were 32 words in total and the complete wordlist is given in Appendix 3.

Table II.7 The target speech materials for Experiment 2

	/s/ context word	Gloss	/s*/ context word	Gloss
[ɑ]	/sɑ.tɑ/	*'buy'*	/s*ɑ.tɑ/	*'wrap'*
	/sɑl.tɑ/	*'live'*	/s*ɑm.ci/	*'pocket'*
[ʌ]	/sʌ.tɑ/	*'stand'*	/s*ʌ.cim/	*'being written'*
	/sʌŋ.kjʌk/	*'personality'*	/s*ʌl.tɑ/	*'cut'*
[o]	/so.kæ/	*'introduction'*	/s*o.tɑ/	*'shoot'*
	/sok.tɑ/	*'be deceived'*	/s*on.sɑl/	*'rapidity'*
[u]	/su.rɑk/	*'agreement'*	/s*u.tɑ/	*'boil'*
	/suk.mjʌŋ/	*'destiny'*	/s*uk.kɑs/	*'crown daisy'*

2.5.1.3 Procedure

Participants were instructed to read the list of target words, presented in isolation. The 32 words were randomized and read 5 times by 6 speakers, which would yield 960 tokens in total (i.e., 6 speakers × 32 words × 5 repetitions). However, excluding incorrectly read items, 957 words were analyzed. The measurement techniques were the same as in Experiment 1.

2.5.2 Results

Table II.8 Average duration and CoG of [s*] and [s] in Korean (ms)
– by subject (top) and by context (bottom)

	Dur (ms)		CoG (Hz)	
	$[s^*]_K$	$[s]_K$	$[s^*]_K$	$[s]_K$
S1	137	108	5693	3974
S2	170	138	6948	6003
S3	124	110	5594	5043
S4	126	130	6936	6231
S5	109	102	7132	4511
S6	171	146	6207	4867
Total	140	122	6416	5105

	Dur (ms)		CoG (Hz)	
	$[s^*]_K$	$[s]_K$	$[s^*]_K$	$[s]_K$
[ɑ]	136	120	6827	4766
[ʌ]	141	116	6660	5283
[o]	133	113	6040	4735
[u]	148	141	6134	5636
Total	140	122	6416	5105

Although the duration and CoG of $[s^*]_K$ were longer and higher, respectively, than those of $[s]_K$ for every speaker, this outcome holds only for within-speaker comparisons. For instance, as shown in Tables II.8, the average duration of $[s^*]_K$ for S3 (124 ms) and S4 (126 ms) was shorter than the duration of [s] for S2 (138 ms). S3's CoG of $[s^*]_K$ (5594 Hz) is lower than that of $[s]_K$ for S2 (6003 Hz) and S4 $[s]_K$ (6231 ms). Also, S5's $[s^*]_K$ duration (109 ms) was even shorter than the average of $[s]_K$ (122 ms). Inspection of the data in terms of the following vowel contexts reveals greater consistency. $[s^*]_K$ is longer and has higher CoG than $[s]_K$ in every vowel context.

The duration and CoG measurements were analyzed in two-tailed paired t-tests. Due to the small number of subjects, by-subject analysis was not conducted. In by-item analysis, both duration and CoG for [s] and [s*] conditions were significantly different from each other. The average duration and the statistical values are given in Table II.9.

Table II.9 Summary of Experiment 2 – Consonantal properties

		[s*]	[s]	significance
Kor	Duration (ms)	140	122	$t(30) = 3.547, p < 0.001$
	Center of Gravity (Hz)	6416	5105	$t(30) = 8.357, p < 0.001$

The results of Experiment 2 therefore indicate that the fricative duration of $[s^*]_K$ was significantly longer than $[s]_K$, and that the CoG of $[s^*]_K$ was significantly higher than that of $[s]_K$.

2.5.3 Discussion

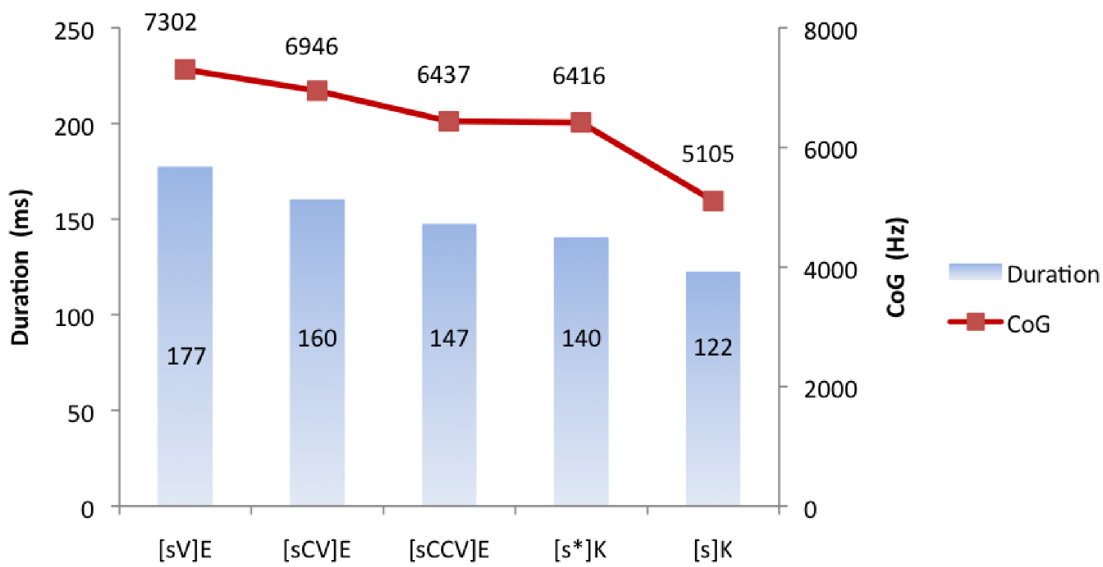
The fricative duration and the CoG of $[s^*]_K$ were longer and higher, respectively, than those of $[s]_K$. Table II.10 compares the duration and CoG differences between $[s^*]_K$ and $[s]_K$ found in Experiment 2 to the results of previous studies. The durational difference reported here was also found in all studies but Kang (2008). The CoG difference, on the other hand, was consistent in all studies that reported this measure.

Table II.10 Comparison of Duration and CoG between Experiment 2 (Korean Production) and Previous Studies

	Duration	CoG
Cho *et al*. (2002)	✓	✓
Chang (2008)	✓	NA
Kang (2008)	×	✓
Ahn (Experiment 2)	✓	✓

To conclude, the results of Experiments 1 and 2 show that $[sV]_E$ has longer duration and higher CoG than $[sCV]_E$, and that $[s^*]_K$ has longer duration and higher CoG than $[s]_K$. Figure II.4 summarizes the duration and CoG values of $[sV]_E$, $[sCV]_E$, $[sCCV]_E$, $[s^*]_K$ and $[s]_K$. In terms of Hypothesis 1a suggested in the previous chapter, it is challenging to argue that $[s^*]_K$ is similar to $[sV]_E$ and $[s]_K$ to $[sCV]_E$, because $[s^*]_K$ is more similar to $[sCV]_E$ than it is to $[sV]_E$.

Figure II.4 Duration and CoG of $[sV]_E$, $[sCV]_E$, $[sCCV]_E$, $[s^*]_K$ and $[s]_K$



2.6 Experiment 3 – The Role of Consonantal Information in Perception of English /s/

The purpose of Experiment 3 is to examine whether two specific acoustic properties of /s/$_E$, a temporal (i.e., duration) and a spectral property (i.e., CoG), are responsible for the adaptation pattern of English /s/ into Korean. Korean listeners, who are sensitive to at least some of the temporal and spectral cues of Korean alveolar fricatives, may attend to these two acoustic properties of English alveolar fricatives, and use these cues in their perception of English /s/-words (cf. Kim, 2002; Chang, 2007). In this experiment, I specifically examine the contribution of these two cues to Korean listeners' non-native perception.

In this experiment, the duration and CoG of English /s/ were systematically modified in order to examine Korean listeners' ability to identify different fricative durations and different CoG frequencies. If Korean listeners are sensitive to these temporal and spectral properties, they should perceive more /s*/$_K$ as duration increases and CoG raises.

To investigate interpretation of Korean listeners' use of temporal and spectral variations, the fricative duration and CoG of a naturally produced /s/$_E$ were varied while all other properties were held constant. Experiment 3 tests the identification of these various /s/$_E$'s by Korean listeners.

2.6.1 Design and methodology

2.6.1.1 Participants

Twenty six Korean-speaking listeners who were recruited around the University of Michigan's Ann Arbor campus participated in this experiment. All participants identified themselves as native speakers of Korean, and of the Seoul dialect in particular.

The participants consisted of three listener groups that differed in the duration of their residence in English-speaking countries, which was presumed to be more or less equivalent to their exposure to English. *Group A* had 9 listeners who have been in the U.S.

or another English-speaking country for 6-8 years (henceforth, *Group A*_ADVANCED or *Group A*_ADV). *Group B* included 8 listeners who have been in an English-speaking country for 2-3 years (*Group B*_INTER or *Group B*_INT). Finally, *Group C* consisted of 8 listeners who have resided in an English-speaking country for less than 6 months (*Group C*_NEW). These different listener groups were included in order to investigate the possible influence of English exposure on the perception of $/s/_E$. All participants were between 20 – 35 years old, and were paid for their participation.

The responses of four participants were eventually excluded from analysis due to relatively high rates of incorrect response time (IRT). For the purpose of this experiment, listeners were expected to respond after they had received all the relevant information from the consonant portions of the stimuli. Incorrect responses were judged based on response time and the durational property of the stimuli: Among various /s/ stimuli of Experiment 3, the shortest /s/ was 110 ms (i.e., the shortest /s/ ended 110 ms after the stimulus had started). Thus, any response time which was less than 110 ms was considered as an incorrect response (i.e., responses of which RT is less than 110 ms).

Listeners whose IRT rates were 10% or more of their total responses were excluded, leaving 7 listeners in each listener group. As shown in Table II.11, 2 listeners from *Group A*_ADV, 1 listener from *Group B*_INTER, 1 listener from *Group C*_NEW were excluded.

Table II.11 Incorrect Response Time Rates of Each Listener in Experiment 3

Listener group	Subject index	Raw number of IRT	Total expected responses	% IRT
*Group A*_ADV	Subj-A1	3	1120	0.27%
	Subj-A2	16	1120	1.43%
	Subj-A3	42	1120	3.75%
	Subj-A4	21	1120	1.88%
	Subj-A5	41	1120	3.66%
	Subj-A6	10	1120	0.89%
	Subj-A7	13	1120	1.16%
	Subj-A8	164	1120	14.64%
	Subj-A9	638	1120	56.96%
*Group B*_INT	Subj-B1	2	1120	0.18%
	Subj-B2	25	1120	2.23%
	Subj-B3	19	1120	1.70%
	Subj-B4	5	1120	0.45%
	Subj-B5	33	1120	2.95%
	Subj-B6	19	1120	1.70%
	Subj-B7	9	1120	0.80%
	Subj-B8	107	1120	9.55%
*Group C*_NEW	Subj-C1	7	1120	0.63%
	Subj-C2	12	1120	1.07%
	Subj-C3	7	1120	0.63%
	Subj-C4	7	1120	0.63%
	Subj-C5	45	1120	4.02%
	Subj-C6	29	1120	2.59%
	Subj-C7	8	1120	0.71%
	Subj-C8	279	1120	24.9%

2.6.1.2 Stimulus materials

Stimuli were created in a two-step process: selection of the original /s/ stimuli and modification of /s/ to create fricative duration and CoG continua. I extracted the frication portions from two naturally spoken /s/-containing words from Experiment 1. The original stimuli were: (i) /s/ from *'stout'*, the second repetition of Subject 4, and (ii) /s/ from *'sight'*, the third repetition of Subject 4. These stimuli were selected because their durations and CoG values were close to the target endpoint values for the continua, so that they needed minimal modification. The /s/ from *'stout'* was 121 ms long and its CoG was 5448 Hz. /s/ from *'sight'* was 189 ms and its CoG was 7486 Hz.

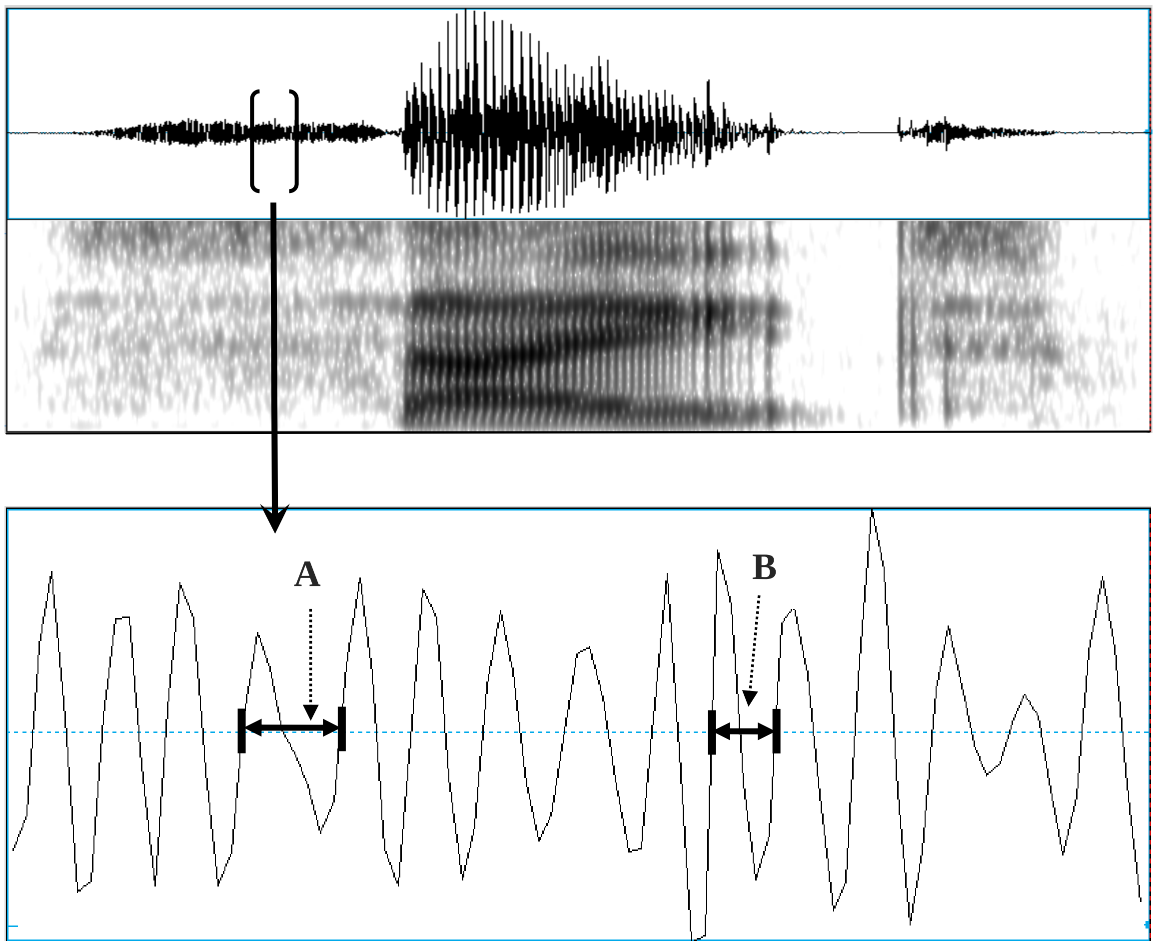
Additionally, these two words contained different vowels, [ɑɪ] or [ɑʊ] respectively. Although the different vowels may result in different coarticulatory effects, it is precisely these CoG values – lower in the [ɑʊ] than in the [ɑɪ] context – that motivated the stimulus choices (options were limited by the stimuli recorded in Experiment 1). *'stout'* was the only word that contained [ɑʊ] among various [stV] words, and no [stV] word contained [ɑɪ].

For /s/ modification, the frication portions of *'sight'* and *'stout'* were manipulated to create a 7-step duration continuum and 4-step CoG continuum that were orthogonally varied to yield 28 different /s/ tokens. Target durations were to range from 110 ms to 200 ms, while the target CoG range was 5200 Hz to 8200 Hz. The target endpoints of the continua were based on duration and CoG measures in Experiments 1 and 2, so that the continua included stimuli that had (roughly) the mean duration and CoG values found for $[sV]_E$, $[sCV]_E$, $[s]_K$, and $[s^*]_K$ in these experiments.

The /s/ from *'sight'* was used to create STEP 1 – STEP 14, and the /s/ from *'stout'* was used to create STEP 15 – STEP 28, with the aim of modifying the sounds as little as possible. To shorten or lengthen the duration, portions of frication were deleted or copied and inserted, respectively, at various places in the frication. In order to increase the CoG, either selected low-frequency frication portions were deleted, or high-frequency portions were copied and spliced in. Conversely, in order to reduce CoG, selected low-frequency frication portions were spliced in, or high-frequency portions were deleted. A low-frequency frication portion was identified as one with a relatively long temporal distance from one zero crossing to the next in the waveform, as shown in the 'A' portion of Figure II.5 for *'sight'*, and a high-frequency frication portion was one with a short distance between zero crossings, as shown in the 'B' portion of Figure II.5.

Creating each stimulus involved both insertion and deletion methods. Multiple of these edits to natural stimuli were required before reaching values in the required ranges. After these manipulations, all the modified tokens of /s/ still sounded like natural speech. After the experiment was conducted, it was realized that there was a substantial amplitude difference between the original frication portions. The amplitude of original /s/ from *'sight'* was 24.9 dB, whereas that of original /s/ from *'stout'* was 65 dB, which resulted in the manipulated /s/ sounds from STEP 1 – STEP 14 having higher amplitude than those from STEP 15 – STEP 28. See section 2.6.3 for a discussion of the possible influence that this may have had on the results of the experiment.

Figure II.5 Waveform and spectrogram of '*sight*' (top) illustrating lower frequency (A) portion and higher frequency (B) portions of the frication noise



These modifications for the increase and/or decrease of duration and CoG resulted in the 28 /s/-tokens whose relevant acoustic properties are given in Table II.12.

Table II.12 Stimulus Steps by Duration and CoG

Duration (ms)

CoG (Hz)

STEP 1 110 ms 8207 Hz	**STEP 2** 125 ms 8193 Hz	**STEP 3** 140 ms 8224 Hz	**STEP 4** 155 ms 8194 Hz	**STEP 5** 170 ms 8206 Hz	**STEP 6** 185 ms 8204 Hz	**STEP 7** 200 ms 8203 Hz
STEP 8 110 ms 7192 Hz	**STEP 9** 125 ms 7197 Hz	**STEP 10** 140 ms 7194 Hz	**STEP 11** 155 ms 7224 Hz	**STEP 12** 170 ms 7229 Hz	**STEP 13** 185 ms 7193 Hz	**STEP 14** 200 ms 7217 Hz
STEP 15 110 ms 6202 Hz	**STEP 16** 125 ms 6192 Hz	**STEP 17** 140 ms 6199 Hz	**STEP 18** 155 ms 6215 Hz	**STEP 19** 170 ms 6184 Hz	**STEP 20** 185 ms 6208 Hz	**STEP 21** 200 ms 6224 Hz
STEP 22 110 ms 5217 Hz	**STEP 23** 125 ms 5193 Hz	**STEP 24** 140 ms 5196 Hz	**STEP 25** 155 ms 5208 Hz	**STEP 26** 170 ms 5196 Hz	**STEP 27** 185 ms 5221 Hz	**STEP 28** 200 ms 5190 Hz

Each of the 28 modified tokens of /s/ was embedded in 10 different bisyllabic (CVCVC) /s/-initial English non-words. The speaker whose original /s/ tokens were modified (i.e., Subject 4), was recorded reading the carriers multiple times. The speaker was familiar with the IPA, and all the items of the reading list were presented in IPA transcription. From multiple recordings of each carrier, one token of each carrier was selected, and the /s/ portion of each was removed and replaced by the modified tokens. /s/ was prevocalic in five of the carrier non-words (e.g., *sailit*) and preconsonantal in the remaining five (e.g., *stailit*). In order to create the preconsonantal stimuli, a stop closure was inserted before every [t]. The duration of the closure was kept constant at 55 ms, which was the average duration of all preconsonantal carriers of the speaker. Table II.13 provides the list of English non-words in IPA and the possible spellings of each IPA representation. Each /s/ from the 28 kinds of /s/ from STEP 1 to STEP 28 was spliced into the 10 carriers, creating 280 stimuli.

Table II.13 English non-words for /s/-carriers

	/sV/		/sCV/	
	word	IPA	word	IPA
[_ɑɪ]	*sailit*	[sɑɪlɪt]	*stailit*	[stɑɪlɪt]
	silut	[sɑɪlət]	*stilut*	[stɑɪlət]
	sirem	[sɑɪɹəm]	*stirem*	[stɑɪɹəm]
[_ɑʊ]	*saulet*	[sɑʊlɛt]	*staulet*	[stɑʊlɛt]
	saurem	[sɑʊɹəm]	*staurem*	[stɑʊɹəm]

2.6.1.3 Procedure

Stimulus presentation and response collection were controlled with *Superlab*. Stimuli were differently randomized by *Superlab* for each presentation. Stimuli were presented to listeners over headphones in a sound-attenuated booth. Korean listeners were instructed to listen to the individual stimuli one at a time, and identify whether the initial segment of the given tokens sounded like /s/ or /s*/ in their native language. After each token was played, listeners pressed one of two buttons marked as /s/ or /s*/ in Korean orthography, and they were asked to press the button after each nonsense item was completed.

Listeners heard four randomizations of the 280 stimuli, so that the total number of potential responses was 1120. Thus, across all listeners, the total number of possible responses was 23520 (i.e., 280 stimuli × 3 listener groups × 7 listeners × 4 repetitions). However, as indicated in Table II.11, earlier listeners showed various IRT rates and these responses with incorrect response times were excluded in the data analysis. There were 1561 errors across listeners, which accounted for around 6.6% of the total number of possibly correct responses. Therefore, 21959 data points were analyzed.

2.6.2 Predictions

Assuming that exposure to English corresponds to listeners' proficiency in English, it can be assumed that listeners with less English exposure will be less proficient in English, while listeners with more English exposure will be more proficient. Less proficient listeners, such as *Group C*_NEW may not know that English has only one voiceless alveolar fricative, and may therefore interpret the phonetic differences between $/s/_E$ in different contexts in terms of the acoustic properties of the two Korean sounds, $/s/_K$ and $/s^*/_K$. On the other hand more experienced listeners such as *Group A*_ADV are likely to be aware that English has only one voiceless alveolar fricative, and in their responses may ignore the contextually determined acoustic properties of English /s/. It is therefore possible that *Group C*_NEW listeners may be more likely to interpret some $/s/_E$ as $/s/_K$ and some as $/s^*/_K$, while the *Group A*_ADV listeners may interpret all $/s/_E$ as only one the Korean sounds, i.e., as either $/s/_K$ or $/s^*/_K$.

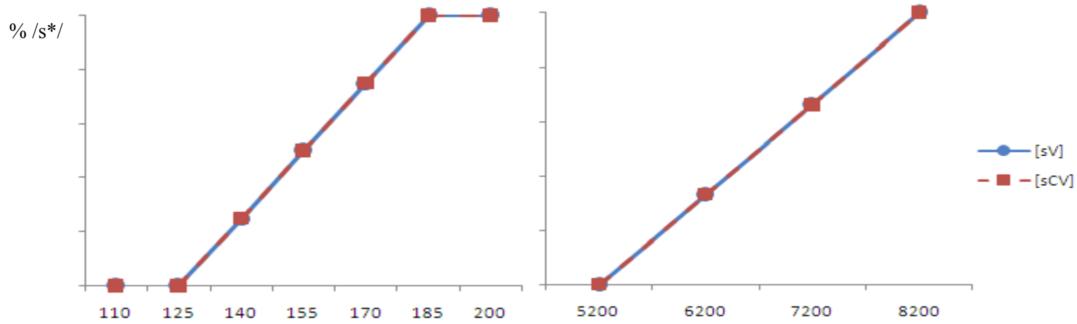
Regarding Korean listeners' perceptual sensitivity to duration and CoG, at least four potential outcomes are possible, as illustrated in Figure II.6. The X-axis indicates the properties of /s/ continua (duration continua on the left and CoG continua on the right), and the Y-axis shows the hypothetical percent /s*/-responses. In all schematic graphs in Figure II.6, percent predicted /s*/-responses to the duration continuum are on the left, and those to the CoG continuum are on the right. Note that, according to phonological contexts, [s] occurs either prevocalically or preconsonantally. The solid line represents /s*/-responses in the prevocalic condition, and the dashed line indicates /s*/-responses in the preconsonantal condition.

According to whether Korean listeners were sensitive to the two acoustic properties (i.e., duration and CoG) or not, and whether they were sensitive to phonological context (i.e., prevocalic and preconsonantal), four response patterns are possible. First, Korean listeners could be sensitive only to the acoustic properties but not to the phonological contexts. This prediction is indicated in Figure II.6A. The two graphs illustrate that, as the values of duration and CoG increase, the percentage /s*/-responses also increases. They also illustrate that the listeners' responses are not sensitive to phonological contexts. Thus, Figure II.6A can be understood as representing a scenario in which Korean listeners paid attention to the acoustic properties of English /s/, determining their identification purely on the basis of the acoustic properties of /s/, and ignoring any influence of phonological contexts.

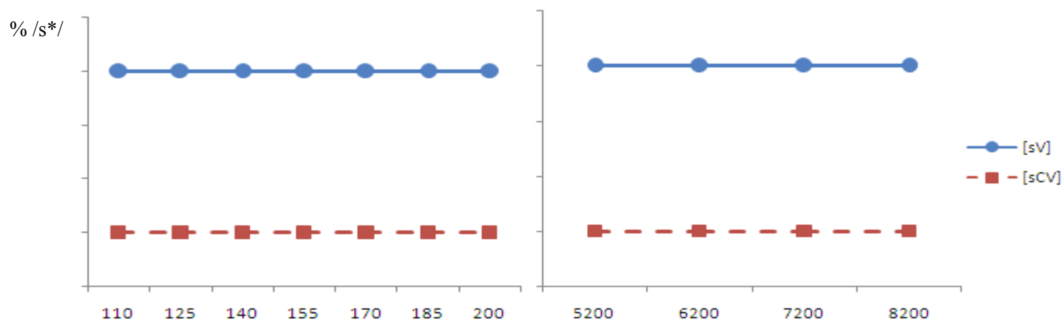
Second, if the listeners were sensitive only to the phonological contexts and not to the acoustic properties, their responses in the two contexts would be as in Figure II.6B. The two figures suggest that Korean listeners perceive only /s*/ in prevocalic contexts and only /s/ in preconsonantal contexts, regardless of the acoustic properties of the stimuli, that is, no matter how long or short the duration or how high or low the CoG, they predominantly perceive /s*/ in [sV] contexts and /s/ in [sCV] contexts. In other words, the figures suggest that the listeners' perception is completely independent of the acoustic information of the duration and CoG.

Figure II.6 Schematic Representations of Possible Outcomes

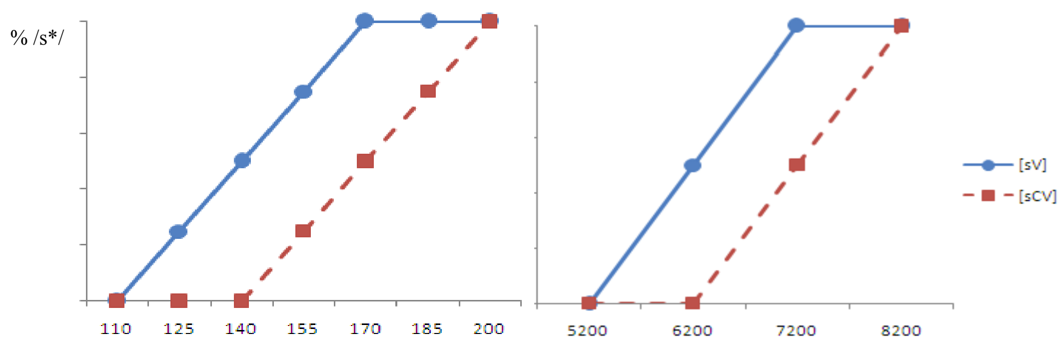
A. Duration/CoG sensitive & Context insensitive



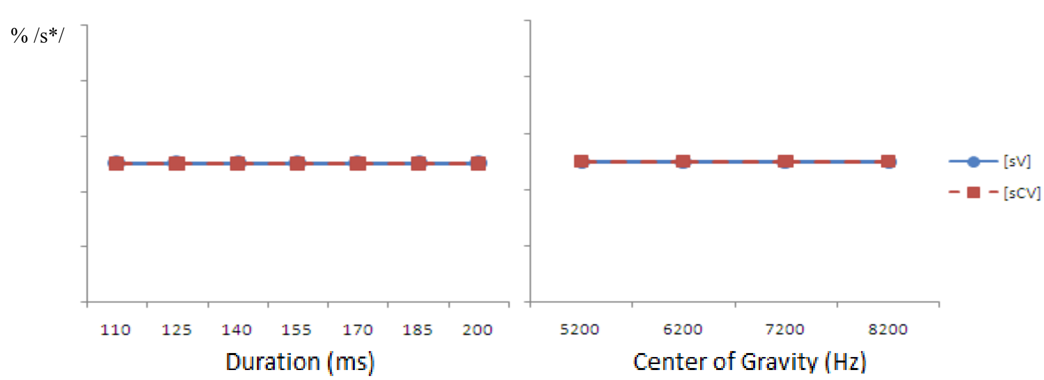
B. Duration/CoG insensitive & Context sensitive



C. Duration/CoG sensitive & Context sensitive



D. Duration/CoG insensitive & Context insensitive



Third, Korean listeners might be sensitive to both the acoustic properties and the phonological contexts as shown in Figure II.6C. Their /s*/-responses increase as fricative duration and CoG increase because of their sensitivity to acoustic properties. In addition, they report hearing more /s*/ in the prevocalic than in the preconsonantal condition because of phonological contextual effects. These results could possibly be interpreted as showing listeners' readiness to use all the information that is available to them and to combine all sources of information in perceiving speech.

Finally, Korean listeners may employ neither acoustic nor contextual information in their speech perception. In fact, this outcome can be subdivided into three categories: (i) listeners may show 0% /s*/-responses (i.e., 100% /s/-responses) irrespective of duration and CoG increments, (ii) they may show 100% /s*/-responses irrespective of duration and CoG increments, and (iii) as displayed in Figure II.6D, they may show 50% /s*/-responses (i.e., 50% /s/-responses). In all the three categories, listeners are not sensitive to acoustic manipulations or phonological context. In the first two categories, listeners consistently select one Korean fricative instead of judging based on duration and CoG or the phonological contexts. In the third category, listeners still do not depend on any consonantal cues, but they respond randomly.

To summarize, in Figure II.6A and II.6C, Korean listeners' perception of English /s/ is at least partially guided by their knowledge of the acoustic properties of Korean $/s/_K$ and $/s^*/_K$. Results like those in Figures II.6B and II.6C would suggest that phonological context matters to Korean listeners' perception so that the same (modified) English /s/ can be perceived differently depending on its phonological context.

2.6.3 Results

Across all conditions, listeners' /s*/-responses were slightly more frequent than their /s/-responses: total responses were 53.4% /s*/ and 46.6% /s/. The results of individual listeners according to phonological context are given in Table II.14.

In prevocalic conditions, the overall percentage of /s*/-responses was 91%, compared to 18% in preconsonantal contexts, indicating that listeners heard significantly more /s*/ before a vowel, thereby upholding the phonological contextual effects both in a by-item analysis ($t(54) = 1.67$, $p < 0.001$) and in a by-subject analysis ($t(40) = 13.606$, $p < 0.001$). Out of 21 listeners, 20 listeners' /s*/-responses were significantly higher in [sV] than in [sCV]. Exceptionally, one listener, Subj-B4, did not show a significant difference in /s*/-responses between [sV] and [sCV].

Table II.14 Percent /s*/-responses in [sV] and [sCV] contexts for individual listeners

Listener group	Subject index	% /s*/ in [sV]	% /s*/ in [sCV]	Significance	
Group A_ADV	Subj-A1	92.5	36.4	$t(54) = 26.3$	$p < 0.001$
	Subj-A2	99.6	0.2	$t(54) = 319.7$	$p < 0.001$
	Subj-A3	85.5	1.1	$t(54) = 44.6$	$p < 0.001$
	Subj-A4	97.8	11.0	$t(54) = 68.3$	$p < 0.001$
	Subj-A5	72.2	34.5	$t(54) = 7.0$	$p < 0.001$
	Subj-A6	95.3	6.6	$t(54) = 58.6$	$p < 0.001$
	Subj-A7	94.9	4.2	$t(54) = 59.0$	$p < 0.001$
Group B_INT	Subj-B1	95.7	24.8	$t(54) = 33.6$	$p < 0.001$
	Subj-B2	69.8	56.0	$t(54) = 4.1$	$p < 0.001$
	Subj-B3	99.8	1.6	$t(54) = 156.7$	$p < 0.001$
	Subj-B4	49.7	48.9	$t(54) = 0.065$	$p = 0.47$
	Subj-B5	88.5	6.5	$t(54) = 36.0$	$p < 0.001$
	Subj-B6	98.5	2.5	$t(54) = 119.5$	$p < 0.001$
	Subj-B7	98.5	8.8	$t(54) = 66.9$	$p < 0.001$
Group C_NEW	Subj-C1	99.6	27.0	$t(54) = 30.7$	$p < 0.001$
	Subj-C2	96.7	64.4	$t(54) = 18.2$	$p < 0.001$
	Subj-C3	64.4	2.1	$t(54) = 9.0$	$p < 0.001$
	Subj-C4	87.5	15.2	$t(54) = 17.3$	$p < 0.001$
	Subj-C5	93.6	19.4	$t(54) = 25.9$	$p < 0.001$
	Subj-C6	86.1	11.6	$t(54) = 26.0$	$p < 0.001$
	Subj-C7	92.1	2.7	$t(54) = 69.6$	$p < 0.001$

For each duration and CoG level, mean percentage /s*/-responses for all listeners are given in Table II.15.

Table II.15 Percent /s*/-responses by duration (A) and CoG (B) changes for each listener group

A. Duration

		110ms	125ms	140ms	155ms	170ms	185ms	200ms
Group	[sV]	91.8	92.1	92.2	91.9	92.9	88.4	88.4
*A*_ADV	[sCV]	12.6	13.0	12.4	14.9	17.2	12.5	10.5
Group	[sV]	85.1	86.0	86.0	91.4	85.5	82.2	83.9
*B*_INT	[sCV]	19.7	19.1	22.8	23.7	22.8	20.3	20.4
Group	[sV]	87.9	83.9	88.4	93.5	89.4	88.3	88.7
*C*_NEW	[sCV]	18.9	17.2	23.4	23.4	25.4	17.0	16.8
Pooled	[sV]	88.2	87.3	88.8	92.3	89.3	86.3	87.0
	[sCV]	17.1	16.4	19.5	20.7	21.8	16.6	15.9

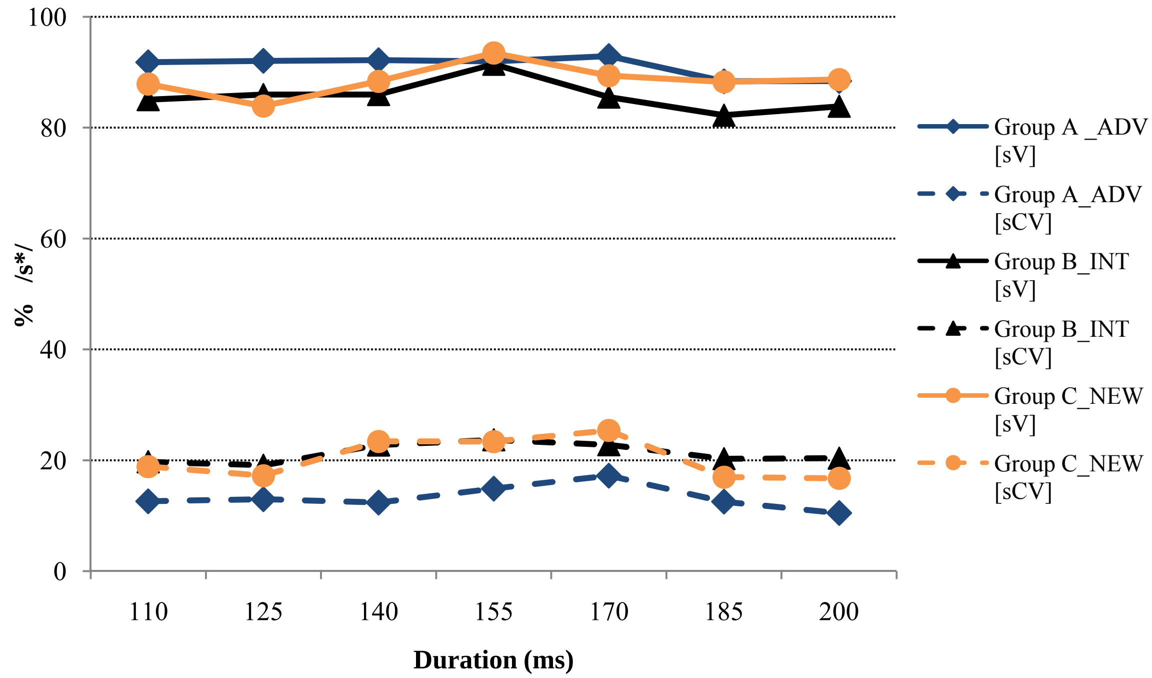
B. Center of Gravity

		5200 Hz	6200 Hz	7200 Hz	8200 Hz
Group	[sV]	92.8	88.7	90.3	92.8
*A*_ADV	[sCV]	14.3	12.1	14.1	13.1
Group	[sV]	94.6	80.5	79.3	88.8
*B*_INT	[sCV]	27.1	15.5	21.0	21.8
Group	[sV]	96.7	81.2	83.4	93.1
*C*_NEW	[sCV]	21.4	17.5	22.7	19.7
Pooled	[sV]	94.7	83.5	84.3	91.5
	[sCV]	21.0	15.0	19.2	18.2

Figure II.7 illustrates the percent /s*/-responses to each step of duration and CoG continua according to listener groups. An ANCOVA (Analysis of Co-variance) was conducted with phonological context (i.e., [sV] and [sCV]) and listener group (i.e., Group A, B and C) as discrete factors, and CoG and duration as continuous factors. The analysis was done by-item only. The analysis showed the main effect of phonological context to be significant (F (1, 166) = 5177.665, $p < 0.001$) but no main effect of listener group (F (2, 165) = 1.7, $p = 0.19$), duration (F (1, 166) = 0.11, $p = 0.74$) or CoG (F (1, 166) = 2.1, $p = 0.15$).

The overall pattern of results for Experiment 3 was nearly identical to Figure II.6B. The listeners did not give more /s*/ responses as the duration increased or the CoG raised, which indicated that Korean listeners were not sensitive to the acoustic properties of duration or CoG as manipulated here. Instead, they responded /s*/ significantly more often before a vowel than before a consonant, which suggests that their responses, at least in this task, are primarily determined by phonological contexts.

Figure II.7 Pooled responses of 21 listeners in 3 listener groups: percent /s*/-responses by duration (above) and CoG (below).



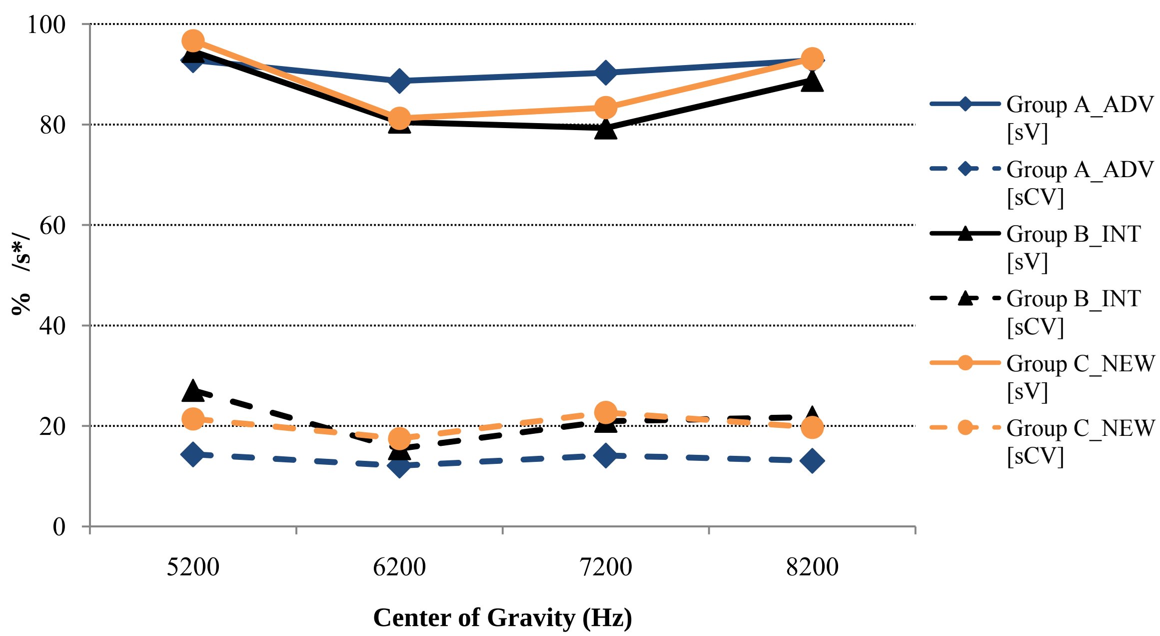


Table II.16 gives the results of responses according to each duration and CoG category (i.e., 28 steps). In each cell, the percentages on the top indicate the percentage /s*/-responses in prevocalic contexts, and those on the bottom show the percentage /s*/-responses in preconsonantal contexts. If Korean listeners had used temporal and spectral cues in their identifications, the percentage /s*/-responses would have been the highest in STEP 7 and the lowest in STEP 22. However, the listeners responded roughly the same way to the tokens represented by each of the different cells.

Table II.16 The percentage /s*/-responses for 28 steps: % /s*/ in prevocalic (top) and in preconsonantal contexts (bottom)

Duration (ms) →

CoG (Hz) ↓

STEP 1 94.9% 19.9%	**STEP 2** 95.5% 16.5%	**STEP 3** 95.2% 19.6%	**STEP 4** 96.9% 23.9%	**STEP 5** 94.8% 23.4%	**STEP 6** 82.4% 14.8%	**STEP 7** 83.8% 15.5%
STEP 8 84.3% 16.6%	**STEP 9** 82.1% 18.5%	**STEP 10** 82.2% 21.6%	**STEP 11** 93.2% 21.1%	**STEP 12** 84.3% 21.3%	**STEP 13** 84.5% 24.1%	**STEP 14** 77.3% 15.4%
STEP 15 78.5% 13.0%	**STEP 16** 77.0% 13.2%	**STEP 17** 80.6% 14.3%	**STEP 18** 82.7% 15.6%	**STEP 19** 81.5% 16.1%	**STEP 20** 85.5% 15.5%	**STEP 21** 94.0% 20.5%
STEP 22 94.9% 22.0%	**STEP 23** 94.2% 20.1%	**STEP 24** 96.4% 25.7%	**STEP 25** 97.4% 25.4%	**STEP 26** 97.2% 29.0%	**STEP 27** 91.9% 14.1%	**STEP 28** 91.5% 17.1%

To summarize, Korean listeners' response patterns to manipulated English [s] were not influenced by the duration or CoG of the frication noise. Instead, they were consistently influenced by phonological context.

As noted above, the stimuli had a substantial difference of amplitude between the original fricatives that were the source for STEP 1 – STEP 14 and STEP 15 – STEP 28. However, the above response pattern of participants showed no correlation between identification and amplitude (i.e., no matched change in response to SETPS 1 – 14 compared to15 – 28). Two-tailed t-tests showed that the response patterns of participants showed no correlation to the amplitude difference in both prevocalic ($t(26) = -0.317$, $p = 0.38$) and preconsonantal contexts ($t(26) = 0.462$, $p = 0.32$).

2.6.4 Discussion

The goal of Experiment 3 was to test H_1b that Korean listeners would be sensitive to the duration and CoG of frication noise, and to determine whether these properties are responsible for Korean listeners' perception of English /s/. The results obtained in this experiment revealed that Korean listeners were insensitive to both temporal and spectral cues (again, as manipulated here); rather, they were significantly influenced by the phonological context. No evidence was found in support of the frication cues hypothesis.

Little prior research has been conducted that has experimentally examined Korean listeners' perception of English /s/. In contrast, at least two studies have investigated Korean listeners' perception of $/s/_K$ and $/s*/_K$. Chang (2007) investigated Korean listeners' perception of $/s/_K$ and $/s*/_K$ before [_ɑ] and [_u]. In his study, he manipulated the frication of /s/ as well as cues realized on the following vowel (e.g., f_0 onset, F_1 onset), and presented them to Korean listeners. The listeners were encouraged to perceive the stimuli as Korean and, similar to the current experiment, their task was to identify the initial fricative as $[s]_K$ or $[s*]_K$ in a forced-choice task. It was found that listeners' responses were influenced by fricative duration before [_u] (i.e., [su] or [s*u]) but not before [_ɑ] (i.e., [sɑ] or [s*ɑ]). Although Chang's study showed that duration played a role in a certain context (i.e., only before [_u]), it is not possible to conclude that duration is the main cue used by Korean listeners in differentiating Korean $/s/_K$ and $/s*/_K$. Chang's result is at best inconclusive with regard to whether Korean listeners rely on the duration of frication noise to distinguish between native Korean $/s/_K$ and $/s*/_K$. Although Chang's results do not directly address Korean listeners' English /s/, it is reasonable to conclude that a cue that is not used in native perception will also not be used in the perception of foreign sounds.

The results of Kim (2002) are inconsistent with the current study. Kim (2002) reported that Korean listeners were significantly influenced by fricative duration in their perception of English alveolar fricatives. Specifically, she found that Korean listeners interpreted a long $/s/_E$ as $/s^*/_K$ and a short $/s/_E$ as $/s/_K$. However, her experimental design has at least two critical shortcomings. Korean listeners were presented with one type of stimulus (different utterances of a [sɑ]-token) which was a spliced out of *sock.* While listeners were not told whether they were listening to English or Korean, they were likely to perceive the stimuli as Korean rather than English, since [sɑ] is an existing Korean word meaning *'four'*. Thus, it is questionable whether the listeners perceived the tokens as English. Unlike the task in Experiment 3 where listeners perceived sounds that were in English "words" as Korean, the listeners in her task perceived sounds in words that could possibly be Korean.

Moreover, Kim's stimuli consisted of seven /s/ tokens whose durations ranged from 60 – 300 ms. The two endpoints were unnaturally long or short. It is also unclear whether the seven stimuli were created by manipulation of a single utterance of /s/. That is, instead of creating a duration continuum, she appeared to have used multiple tokens of /sɑ/ with various durations to get the continuum, which leads to the possibility of cues other than duration being involved. Thus, it is difficult to determine whether Korean listeners in Kim's study indeed relied on specifically the durational cue in their perception of English /s/.

The three listener groups showed similar responses. Experience with English did not appear to influence on response in Experiment 1.

The current perception study showed that, when Korean listeners perceive non-native $/s/_E$ as Korean /s/ or /s*/, they do not focus on temporal or spectral properties of the frication noise. Rather, they used information that is extrinsic to the frication portion itself of /s/. Experimental findings of Korean listeners' tendency to perceive more /s*/ before a vowel than before a consonant support this conclusion. This pattern in Korean listeners' non-native speech perception suggests specific attention to cues external to /s/, and it is therefore appropriate to examine the contribution of the following context (i.e., consonant or vowel).

2.7 General Discussion and Summary of the Chapter

The aim of the two production experiments was to examine and compare acoustic properties of voiceless alveolar fricatives in English and Korean. The main hypothesis predicted (systematic) acoustic differences between a singleton $[s]_E$ and a cluster $[s]_E$, and that the [s] adaptation in Korean would be influenced by those acoustic properties of English /s/. In Experiments 1 (English /s/ production) and 2 (Korean /s/ and /s*/ production), I examined the acoustic properties of English prevocalic $/s/_E$, English preconsonantal $/sC/_E$, and Korean $/s/_K$ and $/s*/_K$. Two /s/-internal properties (i.e., consonantal information), a temporal (i.e., duration) and a spectral property (i.e., CoG), were measured. It was found that, in English, prevocalic $/s/_E$, has longer duration and a higher CoG than preconsonantal $/sC/_E$, and that in Korean, tense $/s*/_K$ has longer duration and higher CoG than lax $/s/_K$.

These results were of interest because the findings are compatible with how English /s/-containing words are borrowed into Korean: English /sV/ words are borrowed with $/s*/_K$ and /sCV/ words are borrowed with $/s/_K$. In view of the finding that a similar pattern of duration and CoG differences were found between English $[sV]_E$ and $[sCV]_E$ and Korean $/s/_K$ and $/s*/_K$. It could be hypothesized that Korean listeners borrow English /s/-containing words based on the acoustic information in consonants. Although the two production experiments in this chapter identified some of the acoustic properties of fricatives in the two languages, it was uncertain whether these acoustic cues are actually used by Korean listeners in the perception of $/s/_E$, which led to Experiment 3.

Experiment 3 examined Korean listeners' sensitivity to the consonantal information of fricative duration and CoG. By varying duration and CoG, I examined Korean listeners' perception of $/s/_E$ in preconsonantal and prevocalic contexts. It was found that Korean listeners predominantly perceived $/s*/_K$ before a vowel and $[s]_K$ before a consonant. The results of Experiment 3 suggest that listeners apparently do not use the frication-internal temporal and spectral cues, at least in the task (and for the stimuli) used here. Korean listeners were insensitive to the temporal and spectral cues provided by the fricative itself. Although, in general, $[sV]_E$ is significantly longer than $[sCV]_E$, and $[s*]_K$

is longer than $[s]_K$, Korean listeners did not report hearing more /s*/ as the duration of [sV] was lengthened. Also, they did not hear more /s*/ as the CoG was raised.

The results of Experiment 3 suggest that Korean listeners are insensitive to consonantal information in perceiving English /s/. This outcome may indicate that they rely on the acoustic properties of the following sound (C or V) rather than on the properties of the $/s/_E$ itself. Alternatively, the findings of Experiment 3 might be understood as showing that Korean listeners use an abstract mapping principle that is based on the phonological context of /s/ rather than on its acoustic properties. There is a large body of literature supporting the influence of grammar on speech perception (Massaro and Cohen, 1983; Pitt 1998; Ahn, 2009; amongst others). Korean listeners may have simply "learned" the loanword mapping principle – /s/ in $/sV/_E$ maps to $/s*/_K$ and /s/ in $/sCV/_E$ maps to $/s/_K$, which is independent from phonetic properties of their native language, and apply this principle in their perceptual task. This principle overrides potentially meaningful acoustic information. This would mean that, instead of listening to the acoustic properties, listeners apply the rule *by analogy*, based on their previous experience with the loanword pattern. Under this interpretation of the results, listeners do not rely on any acoustic differences between prevocalic $/s/_E$ and preconsonantal $/sC/_E$ because they rather employ an abstract rule stipulating that the [s] in prevocalic $/s/_E$ is borrowed as $[s*]_K$, and the [s] in preconsonantal $/sC/_E$ as $[s]_K$.

The problem with this interpretation, however, is that the source of such a rule cannot be explained easily. Where would listeners find evidence for such a rule, and how do they acquire this principle in the first place? For these reasons, the former interpretation (that listeners use acoustic information from the post-fricative context) should be considered. I explore this alternative interpretation in the following chapters.

The results of the three experiments suggest that Korean listeners' perception of English /s/ does not depend on what might have been expected to be relevant acoustic information in the fricative. A possible objection to this interpretation is that Korean listeners did not use temporal and spectral cues in these particular manipulated stimuli, but might have exploited these same cues in non-manipulated natural speech productions. Another objection is that listeners may use fricative internal cues, just not the ones manipulated here. These two considerations, (i) the possible limitation of manipulated /s/

and (ii) the possibility of other consonantal cues being involved, were therefore examined further with natural speech in the following chapter.

To conclude, the possible interpretations of Experiments 1, 2 and 3 can be summarized as follows:

Interpretation 1: Perception determined by acoustic properties of /s/.

a. Duration and CoG of frication are not the primary cues
b. Other acoustic properties determine perception

Interpretation 2: Perception not determined by acoustic properties of /s/.

a. By abstract mapping principle (analogy)
b. Alternative: neighboring sounds → *Hypothesis 2*

Bearing these interpretations in mind, I move on to the next chapter, in which I argue against the foregoing objections, explore alternative explanations to the abstract mapping principle, and look for a better account of Korean listeners' perception of English /s/.

CHAPTER III
POST-FRICATIVE VOWEL PROPERTIES AND PERCEPTION OF VOCALIC INFORMATION

3.1 Introduction

Based on the results of Experiment 3, it was speculated that /s/-internal cues are less important than /s/-external cues (i.e., cues outside the frication) to Korean listeners when they adopt English /s/ into their language. Although most previous studies have recognized consonants themselves as the source of Korean listeners' obstruent perception and have ignored the possibility of consonant external cues being involved, at least one previous study argued that external cues can be informative in obstruent perception. Kim *et al.* (2002) examined Korean listeners' word-initial stop perception. Kim and her colleagues showed that the characteristics of the following vowel, such as f_0 and $H1$-$H2$, are perceptually informative to Korean listeners when they perceive native stop contrasts. For lax stops, in particular, vocalic information, especially f_0 of the following vowel, overrode conflicting consonantal information in determining listeners' judgments. The remainder of this dissertation argues that /s/-external cues are more robust than /s/-internal cues for Korean listeners in their perception of English /s/, and shows experimentally that Koreans' non-native /s/ perception is better accounted for by /s/-external information than by /s/-internal information.

In this and the following chapters, I show the relevance of /s/-external cues and argue that the cues work differently in prevocalic $/s/_E$ and in preconsonantal $/sC/_E$. In prevocalic conditions (i.e., $/sV/_E$), /s/-external cues occur in the following vowel. By examining the vocalic information of English $/s/_E$ and Korean $/s/_K$ and $/s^*/_K$, I suggest that the vocalic information is relevant to Korean listeners' /sV/ perception.

On the other hand, in preconsonantal contexts (i.e., $[sCV]_E$), there is no actual vowel that follows the /s/ (i.e., no immediately following vocalic context). The cues could, though, be realized in the transition into consonant, and possibly on the vowel that follows the consonant. I will show in the next chapter that Korean listeners perform perceptual epenthesis in this context (i.e., they perceptually insert a vowel between /s/ and the following consonant), and I agree that the properties of this perceptually epenthesized vowel are used as cues in the perception of $[sCV]_E$ forms by Korean listeners. In this chapter, however, the discussion of the role of vocalic information is mainly focused on prevocalic $/s/_E$. I provide further evidence showing that the role of consonantal information is limited in /s/ perception. In the next chapter, I explore how Korean listeners deal with phonotactically illegal sequences (i.e., preconsonantal $/sC/_E$).

The present chapter has two aims. First, the acoustic properties of post-fricative vowels in English and Korean are examined. Second, Korean listeners' sensitivity to vocalic properties is examined. These two topics are addressed with respect to Hypothesis 2.

3.1.1 Hypothesis 2 (H_2): Vocalic Information Hypothesis

Similar to Hypothesis 1, Hypothesis 2 is concerned with similarity of acoustic information for English and Korean alveolar fricatives. As previously discussed in Chapter I, Hypothesis 2 concerns Korean listeners' perceptual sensitivity to the vocalic information. In terms of prevocalic $/sV/_E$ adaptation, this hypothesis states that the properties of the vowels that appear after prevocalic $/s/_E$ are more similar to the properties of vowels after $/s^*/_K$ than $/s/_K$ (H_2a) and that this acoustic fact plays an important role in Korean listeners' perception of English /s/ (H_2b). This hypothesis is represented graphically in Figure III.1.

Figure III.1 English to Korean prevocalic /s/$_E$ mapping based on vocalic information

English: [s**V̲**_]

Korean: [s***V̲**] [s**V̲**]

Hypothesis 2 can be understood in two ways: in terms of production, the acoustic properties of vowels after /s/$_E$ and vowels after /s*/$_K$ should be similar (i.e., H_2a), and in terms of perception, this vocalic information leads Korean listeners to perceive the fricative as tense (i.e., H_2b). In order to test H_2a, I revisit the data collected in Experiments 1 and 2, and examine the properties of the post-fricative vowels in these data. Note that the vocalic properties are /s/-external information; this information is not realized in the frication itself. H_2b is examined in Experiment 4, which was designed to indirectly diagnose Korean listeners' perception of /s/$_E$ by comparing Korean listeners' perception of /s*/$_K$ followed by either a vowel or a consonant.

3.1.2 Vocalic measurements

Vocalic cues (i.e., frication external cues) are defined here as the acoustic information realized on the vowels that follow the fricative. Some previous studies have noted that listeners find such external cues useful in their perception. As noted above, Kim *et al.* (2002) have shown that Korean listeners predominantly rely on vocalic information when they perceive native Korean stop consonants, and that vocalic information is useful in their native phoneme discrimination. Choi (2002) also showed that pitch information on the following vowel plays an important role in distinguishing stop contrasts in Korean.

In the previous chapter, it was found that /s/-internal cues such as the temporal (i.e., duration) and spectral (i.e., CoG) properties of the frication itself were not the deciding factors in the loanword adaptation pattern. Although most previous studies assumed that external cues were less directly relevant to the perception of phonological categories, Kim *et al.*'s findings for stops raise the possibility that Korean listeners' English /s/

perception may be done on the basis of fricative-external cues. That is, it is reasonable to speculate that, if Korean listeners depend on vocalic information in their native stop perception, they might also do so when they perceive native fricatives. Also, if they pay attention to the vocalic information in their native perception, they might depend on the same kind of cues in non-native speech perception. For the purpose of this study of Korean listeners' English /s/ perception, I examine and compare the post-fricative vocalic properties of the two languages. In the following discussion, two kinds of fricative-external (i.e., vocalic) properties are addressed, f_0 and $H1$-$H2$.

f_0 and $H1$-$H2$ were measured for the vowels that appear after fricatives in English and Korean. f_0 has been shown to be influenced by the phonological category of neighboring consonants (Kingston, 1986; Kingston and Diehl, 1994). In Korean in particular, the distinction among three-way stop contrasts is partially reflected in f_0 (Cho *et al.*, 2002; Kim *et al.*, 2002). $H1$-$H2$ is often used as a way to examine the breathiness of voice, which is relevant to the tense and lax distinction (Fisher-Jørgnesen, 1967 in Gujarati; Klatt and Klatt, 1990; Pierrehumbert and Talkin, 1992 in English; Ren, 1992 in Shanghai Chinese).

3.2 Vocalic properties of /sV/$_E$: previous research

3.2.1 f_0 of /sV/$_E$

It is well-known that, across languages, higher f_0 is found after voiceless consonants (e.g., [p], [t], [s]) than voiced consonants (e.g., [b], [d], [z]) (Ohde, 1984; Silverman, 1987; Kingston *et al.*, 2008). Löfqvist *et al.* (1989) measured f_0 values at the onset of English and Dutch vowels, and they found that f_0 values were generally higher after a voiceless than after a voiced consonant, regardless of the kind of preceding consonants. Similarly, Kingston and Diehl (1994) demonstrated that f_0 tends to be higher before after a voiceless consonant, thus, f_0 of the vowel [ɑɪ] as in *'pie'* is higher than f_0 of the same vowel in *'bye'*. Whalen *et al.* (1990) argued that f_0 values contribute to the voicing judgment. Kingston *et al.* (2008) suggested that f_0 information is integrated in the perception of voicing contrast.

While a large body of studies focused on f_0 with regard to voicing contrasts, relatively few studies have investigated the influence of other consonantal properties on f_0. However, Whalen (1991) examined the influence of a preceding /s/$_E$ on the f_0 of a following vowel, and found that vocalic transitions are important to perceive fricatives. Studies have not investigated f_0 of vowels after /sC/$_E$.

3.2.2 *H*1-*H*2 of /sV/$_E$

Voice quality of the following vowel is often measured as the difference between the amplitude of the first harmonic (*H*1) and the second harmonic (*H*2). In general, the greater the amplitude difference between *H*1 and *H*2, the more breathy the voice quality is. The voice quality of heavy breathiness (i.e., high *H*1-*H*2 value) marks lax obstruents while light breathiness (i.e., low *H*1-*H*2 value) correlates with tense obstruents (Jessen, 1998).

The *H*1-*H*2 of vowels after word-initial [s]$_E$ was examined by Kang (2008). (Since *H*1-*H*2 is measured at the onset of the following vowel, the values are not available for [sCV] words.) In her measurements, Kang compared the differences in *H*1-*H*2 at the vowel onset after [s]$_E$ to those of [s]$_K$ and [s*]$_K$. She reported that, while *H*1-*H*2 of [s]$_K$

was around 10 dB and $H1$-$H2$ of $[s^*]_K$ was 5 dB, that of $[sV]_E$ was around -3 dB, leading her to conclude, therefore, that English word-initial [s] was more similar to $[s^*]_K$.

3.3 Vocalic properties of $/s/_K$ and $/s^*/_K$: previous research

3.3.1 f_0 of $/sV/_K$ and $/s^*V/_K$

The acoustic characteristics of Korean post-fricative vowels have been examined in several studies. Cho and his colleagues (2002) examined f_0 of the vowels that follow $/s/_K$ and $/s^*/_K$ (e.g., [sɑ] vs. [s*ɑ]) and found that f_0 values at vowel onset were not significantly different. A similar result for f_0 was found by Chang (2007) and Kang (2008). In contrast, Ahn (1999) found that f_0 significantly differed after $[s^*]_K$ and $[s]_K$ in that f_0 is higher after $[s^*]_K$ than after $[s]_K$.

3.3.2 $H1$-$H2$ of $/s/_K$ and $/s^*/_K$

The above studies also examined $H1$-$H2$ values after $[s]_K$ and $[s^*]_K$. According to Cho *et al.* (2002), $H1$-$H2$ of $[s]_K$ and $[s^*]_K$ were significantly different. They found that $H1$-$H2$ values at vowel onset position were significantly higher after $[s]_K$ than after $[s^*]_K$. Both Ahn (1999) and Chang (2007) found similar results for $H1$-$H2$.

In the following sections, I examine the post-fricative vocalic properties in English and Korean for the production data from Experiments 1 and 2 of this dissertation. While the consonantal properties were examined in earlier in Chapter 2, I now revisit the experimental results in terms of vocalic properties.

3.4 Experiment 1 Revisited – Acoustic Properties of Post-fricative Vowels in English

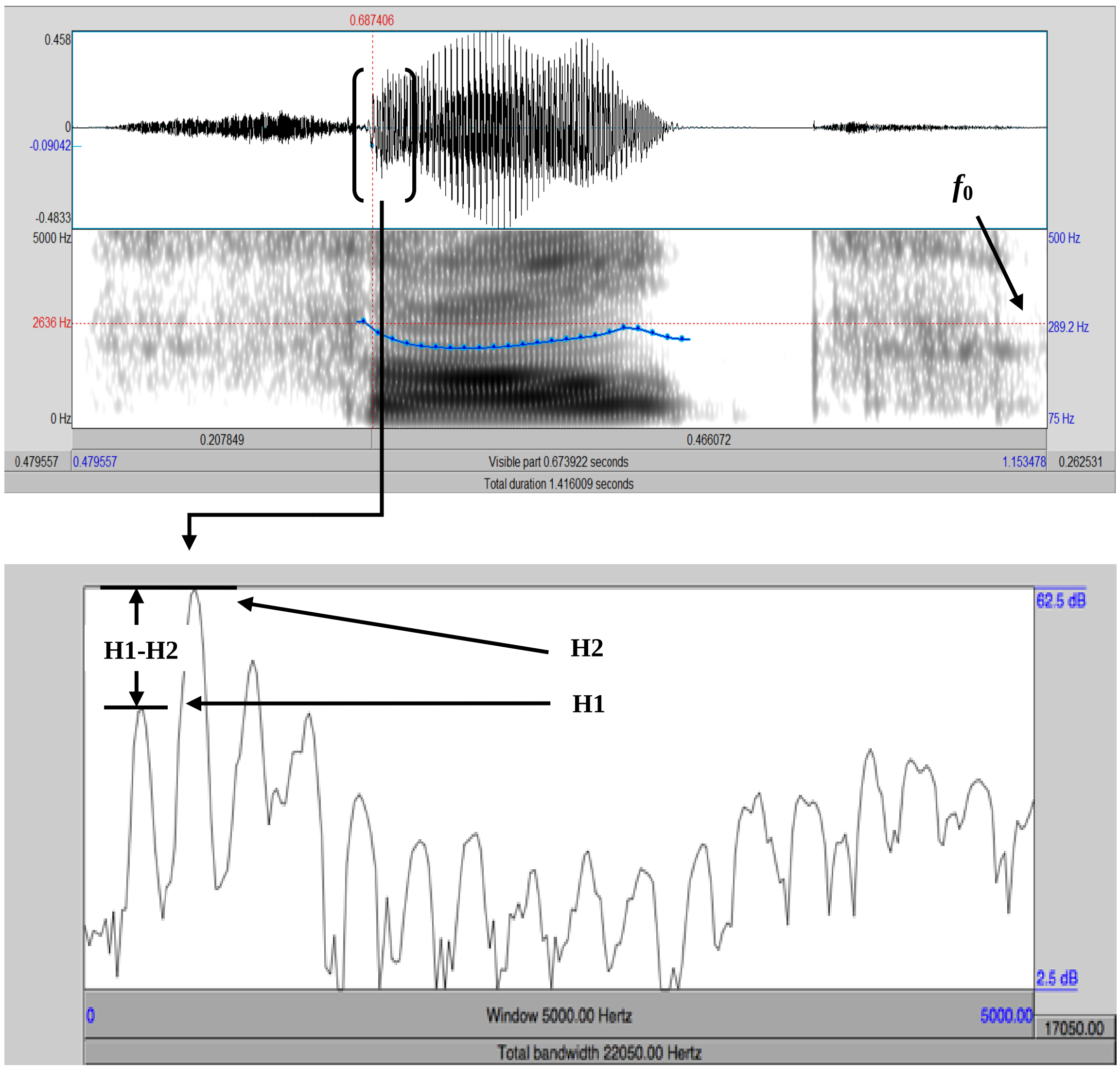
While vowels immediately follow fricatives in $[sV]_K$, $[s*V]_K$ and $[sV]_E$ (i.e., vowels are post-fricative), those in $[sCV]_E$ contexts appear after a stop. For $[sCV]_E$ clusters, although /s/ does not immediately precede a vowel, I still examine the properties of the vowel following the post-fricative consonant.

3.4.1 Methodology

The fundamental frequency (i.e., f_0) and the difference between the amplitude of the first and the second harmonics (i.e., $H1$-$H2$) of post-fricative vowels in English were examined. All measurements were taken in *Praat* (Boersma and Weenink, 2009). Measurements of f_0 and $H1$-$H2$ were taken at the onset of the vowel following $/s/_E$, $/sC/_E$, $/s/_K$ and $/s*/_K$. The criteria and procedures for measuring f_0 and $H1$-$H2$ are as follows: The f_0 values were measured 10 ms from the onset of voicing using the "Show pitch" option in *Praat*, as illustrated in Figure III.2. $H1$-$H2$ measurements were calculated from FFT spectra using a 10 ms window beginning at vowel onset. FFT spectra ranged from 0 to 5000 Hz. The amplitude difference of $H1$ and $H2$ was calculated by subtracting $H2$ values from $H1$. Figure III.2 illustrates a spectrum in which the difference of $H1$-$H2$ amplitudes is negative.

While data from six speakers were collected in Experiment 1, data from only the three female speakers were analyzed because of likely gender differences in f_0 and $H1$-$H2$. It is well established male speakers, on aveage, have lower f_0 than female speakers (Peterson and Barney, 1952; Diehl *et al.*, 1996), and that female speakers' voices are breathier than male speakers' voices (Henton and Bladon, 1985; Klatt and Klatt, 1990; Todaka, 1993; Blankenship, 1997; Hanson, 1997). There were 35 /sV/ words (e.g., *'salt'*) and 38 /sCV/ words (e.g., *'spice'*). Three female speakers read the words five times (i.e., 73 words × 3 speakers × 5 repetitions). Although 1095 data points were expected in total, only 1080 data points were analyzed, excluding 15 instances of incorrectly read tokens.

Figure III.2 Measurements of f_0 (top) and $H1$-$H2$ (bottom) at vowel onset in English *'salt'*



3.4.2 Results

Table III.1 displays the average f_0 and $H1$-$H2$ of /sV/$_E$ and /sCV/$_E$. In a by-item analysis, the average f_0 values for the two contexts were nearly identical, and a two-tailed paired sample t-test confirmed that the values were found not to be different from each other ($t(71) = 0.36$, $p = 0.359$). In the t-test, [s] and [s*] from a given speaker were paired each other. The average $H1$-$H2$ values of /sV/$_E$ and /sCV/$_E$ were also very similar. For statistical analysis, the two values were paired so that two values from a given speaker paired each other. The two-tailed paired sample t-test showed that the $H1$-$H2$ values of /sV/$_E$ and /sCV/$_E$ were not significantly different ($t(71) = -1.05$, $p = 0.147$).

Table III.1 Mean f_0 and $H1$-$H2$ of English post-fricative vowels measured at vowel onset

	/sV_/$_E$	/sCV_/$_E$
f_0 (Hz)	230.7	231.2
$H1$-$H2$ (dB)	-7.79	-8.61

Table III.2 gives f_0 and $H1$-$H2$ values for specific post-/s/ contexts in English. As shown in the table, the values in the /sCV/$_E$ contexts were similar regardless of the consonants (i.e., [spV], [stV] and [skV]).

Table III.2 Mean f_0 and $H1$-$H2$ of English vowel according to contexts

Contexts		f_0 (Hz)	$H1$-$H2$ (dB)
/sCV/$_E$ context	/spV/ (e.g., '*spice*')	227 (217 ~ 248)	-7.9
	/stV/ (e.g.,'*stall*')	233 (227 ~ 246)	-9.0
	/skV/ (e.g., '*scalp*')	234 (227 ~ 248)	-8.9
/sV/$_E$ context	/sV/ (e.g., '*salt*')	231 (224 ~ 240)	-7.8

3.4.3 Discussion

In this section, I examined the post-fricative vocalic properties from Experiment 1. In comparing f_0 and $H1$-$H2$ values in [sV]$_E$ and [sCV]$_E$ contexts, it is apparent that the two values are very similar so that no differences are found between the two contexts. These results suggest that the properties of the vowels that immediately follow /s/$_E$ are very similar to the ones that follow a stop consonant. Additionally, within [sCV] contexts, f_0 and $H1$-$H2$ were not influenced by the place of the stop consonants.

3.5 Experiment 2 Revisited – Acoustic Properties of Post-fricative Vowels in Korean

3.5.1 Methodology

Similar to the investigation of English vowels, the productions of three female Korean speakers were analyzed. There were 16 words in each context of /s*V/$_K$ and /sV/$_K$, as shown in Table III.4 (i.e., 32 words across contexts). /s/ or /s*/ was followed by one of [ɑ], [ʌ], [o] or [u]. All were bisyllabic words; the initial syllables were equally distributed between CV and CVC. Three female speakers read the 16 words five times, which could generate up to 480 data points (i.e., 32 words × 3 speakers × 5 repetitions). Some tokens with unnatural speech, creaky voice, incorrect reading or exceptionally low f_0 (e.g., they ranged 88 ~ 130 Hz) were excluded; thus, 403 data points were analyzed. Both f_0 and H1-H2 were examined, the measurement criteria were same as for English.

3.5.2 Results

As shown in Table III.3, the average values of f_0 of /s*/$_K$ and /sV/$_K$ were nearly identical, and a by-item two-tailed t-test confirmed that f_0 of /s*/$_K$ and /sV/$_K$ were not significantly different ($t(30) = 0.23$, $p = 0.41$). However, the H1-H2 values of /s*V/$_K$ and /sV/$_K$ were dissimilar, with H1 and H2 amplitudes being nearly the same after /s/ but H2 amplitude being greater than that of H1 after /s*/. The difference was confirmed as being significant by a by-item two-tailed paired sample t-test ($t(30) = 6.6$, $p < 0.001$).

Table III.3 Mean f_0 and H1-H2 of Korean post-fricative vowels measured at vowel onset

	/s*V/$_K$	/sV/$_K$
f_0 (Hz)	260.8	261.3
H1-H2 (dB)	-5.88	-0.51

Table III.4 gives a breakdown for individual words showing that, although the f_0 values were similar for /s/$_K$ and /s*/$_K$, $H1$-$H2$ values for /s/$_K$ and /s*/$_K$ were systematically differed. It appeared that, in most contexts, especially [ɑ], [ʌ] and [o] contexts, the $H1$-$H2$ values of /s/$_K$ were generally higher than that of /s*/$_K$. Exceptionally, the difference was much smaller in [u] contexts.

Table III.4 Mean f_0 and $H1$-$H2$ of Korean vowel according to contexts

	/s/$_K$ context			/s*/$_K$ context		
	words	f_0 (Hz)	$H1$-$H2$ (dB)	words	f_0 (Hz)	$H1$-$H2$(dB)
[ɑ]	/sɑ.tɑ/	254 (232 ~ 273)	-0.1	/s*ɑ.tɑ/	251 (237 ~ 290)	-7.0
	/sɑ.cin/	259 (225 ~ 273)	-1.6	/s*ɑ.um/	253 (230 ~ 274)	-8.7
	/sɑl.tɑ/	253 (231 ~ 281)	1.4	/s*ɑk.nun/	264 (216 ~ 356)	-6.5
	/sɑm.ʧon/	265 (244 ~ 284)	-1.3	/s*ɑm.ci/	254 (241 ~ 272)	-7.4
[ʌ]	/sʌ.tɑ/	258 (239 ~ 277)	-1.4	/s*ʌ.re/	257 (247 ~ 271)	-8.5
	/sʌ.ul/	254 (230 ~ 283)	-0.7	/s*ʌ.cim/	258 (240 ~ 284)	-7.1
	/sʌŋ.kjʌk/	259 (233 ~ 293)	-0.1	/s*ʌk.tɑ/	264 (241 ~ 285)	-7.8
	/sʌŋ.ɨi/	261 (253 ~ 283)	0.8	/s*ʌl.tɑ/	252 (237 ~ 283)	-8.7
[o]	/so.kæ/	266 (249 ~ 275)	-1.7	/s*o.tɑ/	266 (239 ~ 288)	-7.1
	/so.mɑŋ/	261 (249 ~ 275)	-3.5	/s*o.im/	261 (243 ~ 299)	-7.6
	/sok.tɑ/	271 (255 ~ 287)	-2.7	/s*on.sɑl/	262 (234 ~ 290)	-5.6
	/sol.cik/	265 (253 ~ 287)	-1.2	/s*ol.rim/	257 (233 ~ 287)	-6.4
[u]	/su.rɑk/	261 (248 ~ 278)	-0.7	/s*u.tɑ/	269 (246 ~ 300)	-1.3
	/su.jʌŋ/	262 (255 ~ 281)	-0.1	/s*u.sim/	265 (255 ~ 289)	0.0
	/suk.mjʌŋ/	271 (257 ~ 293)	1.7	/s*uk.kɑs/	268 (239 ~ 286)	-3.6
	/sun.su/	259 (235 ~ 291)	1.0	/s*uk.t*ɨm/	271 (236 ~ 313)	-1.0

3.5.3 Discussion

The reanalysis of the production data from Experiment 2 shows that /s*/$_K$ and /s/$_K$ have different acoustic effects on the following vowels. In terms of f_0, it is often the case that f_0 values are different after tense and lax consonants. However, as shown above, f_0 values after Korean fricatives do not differ as a function of the tenseness of the fricative. On the other hand, $H1$-$H2$ values of vowels after [s*]$_K$ and [s]$_K$ were different, similar to what has been found after stop consonants. The $H1$-$H2$ measurements between the tense and lax fricatives in Korean are in line with cross-linguistic studies such as Jessen (1998).

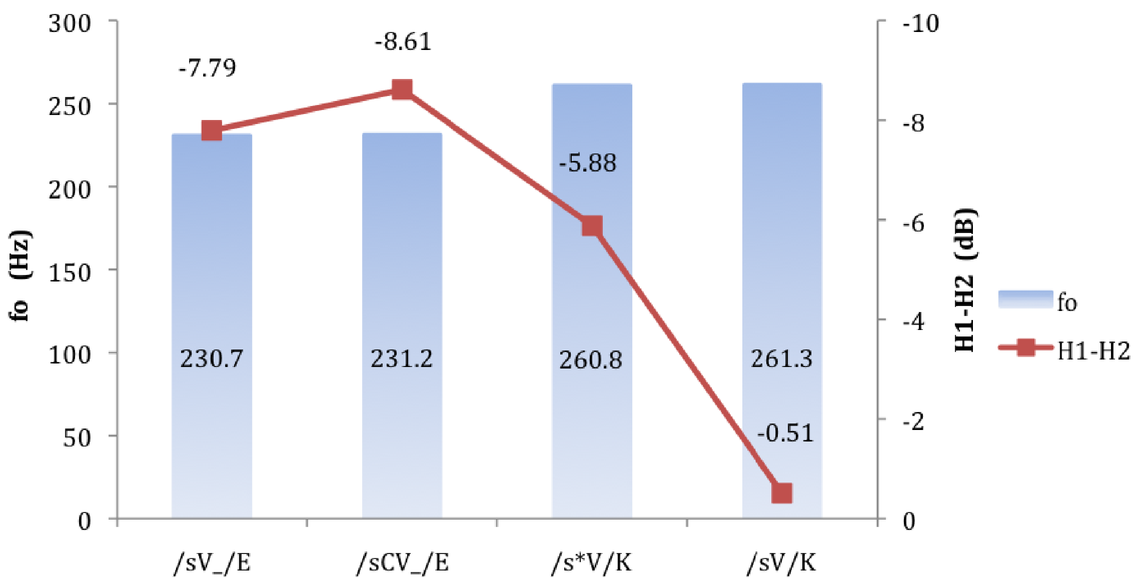
Table III.5 compares the above results with other studies that examined f_0 and $H1$-$H2$ of [s]$_K$ and [s*]$_K$. Similar to the previous chapter, the ✓ marks indicate that a significance difference between /s/$_K$ and /s*/$_K$ was found, and the × marks indicate that a significant difference was not found. Except Ahn (1999), most previous studies found that the f_0 values of vowels after [s]$_K$ and [s*]$_K$ were not significantly different. In every study, $H1$-$H2$ values were consistently different in that the difference is larger after [s*]$_K$ than [s]$_K$.

Table III.5 Comparison of f_0 and $H1$-$H2$ between Experiment 2 (Korean production) and Previous Studies

	f_0	$H1$-$H2$
Ahn (1999)	✓	✓
Cho *et al.* (2002)	×	✓
Chang (2007)	×	✓
Kang (2008)	×	✓
Ahn (Experiment 2)	×	✓

Figure III.3 illustrates the average f_0 and $H1$-$H2$ at vowel onset following English and Korean word-initial alveolar fricatives. In general, f_0 values were higher in Korean than in English, and $H1$-$H2$ values were lower in Korean than in English. English prevocalic /s/$_E$, which is borrowed as [s*]$_K$ in Korean, has lower f_0 but higher $H1$-$H2$ than [s*]$_K$, and English preconsonantal /sCV/$_E$ shows a similar pattern. The English-Korean difference in $H1$-$H2$ values is considerably greater for [sCV]$_E$ versus [s]$_K$ than for [sV]$_E$ versus [s*]$_K$.

Figure III.3 Mean f_0 and $H1$-$H2$ of English and Korean Post-fricative Vowels



The important outcome here is that the properties of the V in [sV]$_E$ are similar to the properties of V in [s*V]$_K$ than to the properties of V in [sV]$_K$, which agrees with prevocalic /sV/$_E$ to /s*/$_K$ adaptation. In the following section, Experiment 4 was designed to examine Korean listeners' sensitivity to this vocalic information.

3.6 Experiment 4 – The Role of Vocalic Information in English Perception

Revisiting Experiments 1 and 2 indicated that the vocalic properties of $[s^*]_K$ and $[s]_K$ are significantly different. The main purpose of Experiment 4 is to examine whether this acoustic difference is responsible for English /s/ adaptation into Korean. In this experiment, I explore whether Korean listeners employ vocalic information in their English perception. In order to address the reservation raised earlier concerning the use of heavily manipulated stimuli in Experiment 3, Experiment 4 uses non-manipulated natural /s/ tokens.

Experiment 4 tests H_2b, which suggests that Korean listeners' perception is based on vocalic information. To test this hypothesis, I compared /s/ in three contexts: [sV], [sCV] and [s_V], where [s_V] stands for an [sCV] token from which the [C] has been deleted. That is, /s/ was presented (i) before a vowel (i.e., [sV]), (ii) before a stop (i.e., [sCV]), or (iii) before a vowel which used to appear after a non-fricative consonant (i.e., the same as the [sCV] but with the [C] deleted).

3.6.1 Design and methodology

3.6.1.1 Participants

The participants from Experiment 3 also participated in Experiment 4.

3.6.1.2 Stimulus materials

The stimuli used in this experiment were similar to the carriers of Experiment 3. A female native speaker of American English read the list of English bisyllabic non-words provided in Table III.6 multiple times. No modification was involved in the stimuli of the [sV] and [sCV] conditions. The stimuli of the [s_V] condition were minimally modified in that the stop portions were completely removed. After the stop portions were removed, the stimuli sounded like a natural speech version of [sV].

Table III.6 Stimuli List of Experiment 4

	[sV]		[sCV]		[s_V]	
carriers	*silet*	[saɪlɛt]	*stilet*	[staɪlɛt]	*s_ilet*	[s_aɪlɛt]
	sailit	[saɪlɪt]	*stailit*	[staɪlɪt]	*s_ailit*	[s_aɪlɪt]
	silut	[saɪlət]	*stilut*	[staɪlət]	*s_ilut*	[s_aɪlət]
	sirem	[saɪɹəm]	*stirem*	[staɪɹəm]	*s_irem*	[s_aɪɹəm]
	saulet	[saʊlɛt]	*staulet*	[staʊlɛt]	*s_aulet*	[s_aʊlɛt]
	saurem	[saʊɹəm]	*staurem*	[staʊɹəm]	*s_urem*	[s_aʊɹəm]

The acoustic characteristics of the stimuli are provided in Table III.7.

Table III.7 Acoustic Characteristics of Each Stimulus in Experiment 4

Condition	Stimuli	Fricative characteristics		Vocalic characteristics	
		Dur (ms)	CoG (Hz)	f_0 (Hz)	$H1$-$H2$ (dB)
[sV]	*sailit*	168	9297	198	-9.1
	saulet	184	9313	201	-13.6
	saurem	207	9780	202	-12
	silet	194	9703	226	-5.5
	silut	172	8781	201	-8.3
	sirem	206	9793	215	-8.6
	average	**189**	**9445**	**207**	**-9.5**
[sCV] & [s_V]	*stailit*	166	9285	200	-9
	staulet	127	8581	196	-14.6
	staurem	180	8430	205	-9
	stilet	206	8997	197	-13
	stilut	114	8170	205	-13.4
	stirem	147	9325	211	-12.6
	average	**157**	**8798**	**202**	**-11.9**

In a by-item analysis, two-tailed *t*-tests of the twelve selected stimuli showed that, both fricative duration and CoG were significantly different between [sV] and [sCV] contexts ($t(10) = 2.048$, $p = 0.034$, and $t(10) = 2.561$, $p = 0.014$, respectively). Frication duration was longer, and CoG was higher in [sV] than in [sCV]. On the other hand, f_0 at the vowel onset and amplitude of $H1$-$H2$ were not significantly different ($t(10) = 0.956$, $p = 0.181$, and $t(10) = 1.588$, $p = 0.0717$, respectively.)

The acoustic characteristics of these stimuli were compared with previous findings. Table III.8 gives fricative duration, CoG, f_0 and $H1$-$H2$ of the stimuli of Experiment 4, the pooled results of Experiment 1, and the results of the speaker (Subj4) who produced Experiment 4 stimuli.

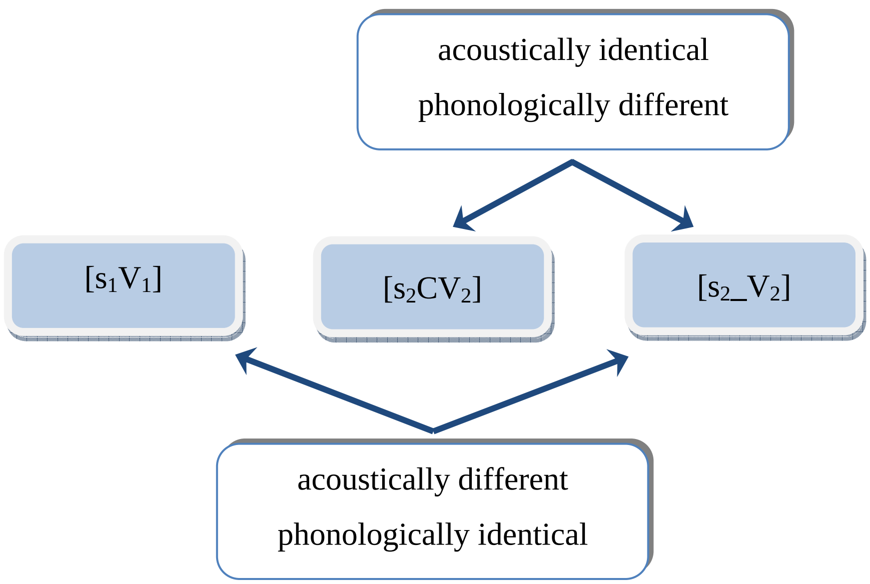
Table III.8 Comparisons of Acoustic Characteristics: fricative duration, CoG, f_0 and $H1$-$H2$ of Experiment 4 stimuli, pooled results of Experiment 1, and results of Subject 4

	Condition	Consonantal characteristics		Vocalic characteristics	
		Dur (ms)	CoG (Hz)	f_0 (Hz)	$H1$-$H2$ (dB)
Exp 4	[sV]	189	9445	207	-9.5
	[sCV]	157	8798	202	-11.9
Exp 1, pooled	[sV]	177	7302	230.7	-7.79
	[sCV]	160	6946	231.2	-8.61
Exp 1, Subj4	[sV]	202	8304	201	-6.22
	[sCV]	172	7941	194	-7.2

The average fricative durations, f_0 and $H1$-$H2$ of the Experiment 4 stimuli were similar to those of Experiment1 pooled and Subj4. However, the average fricative CoG values of the stimuli in Experiment 4 were about 900 ~ 1000 Hz higher than the same speakers' average CoG in Experiment 1. I will return to this issue below.

As shown in Figure III.4, in each triplet, the [s] and [V] portions of the [sCV] and [s_V] members were identical. These two tokens are therefore identical in terms of both their fricative internal and vocalic properties. They are also different from the [sV] member of the triplet in terms of these properties. On the other hand, the [sV] member and the [s_V] member are identical in their phonological pattern (both have [s] directly adjacent to a vowel), and these two members differ from [sCV] where [s] is adjacent to a consonant.

Figure III.4 Acoustic and Phonological Properties of [sV], [sCV] and [s_V] Triplets



3.6.1.3 Procedure

Six English non-words in each of the three [sV], [sCV] and [s_V] conditions generated a total of 18 tokens. These 18 tokens were presented to listeners in the same manner as in Experiment 3. Listeners were asked to identify whether each token started with [s] or [s*] in their native language. Similar to Experiment 3, in order to control for any English exposure effects, the listeners were separated into different groups based on duration of residence in an English speaking country.

Four randomizations of the 18 tokens were played to listeners. In line with the earlier experiment, the errors were determined based on the duration of /s/ and listeners' response time. In this experiment, the shortest /s/ ending points were 285 ms after the stimuli started to play; RTs shorter than 285 ms were considered incorrect.

Additionally, one data point was excluded for Subj-A2. This participant did not press the button correctly, thus, the next sound file was not played for a long period of time. This RT exceeded 10 times of this participant's average RT. IRT rates for each listener are given in Table III.9. Across all listeners, 1512 data points were expected to be collected (i.e., 6 words × 3 conditions × 4 repetitions × 3 listener groups × 7 listeners in each group). Across all listeners, 27 errors were found, which was around 1.8% of the total expected responses. Excluding the errors, 1485 data points were used in the data analysis.

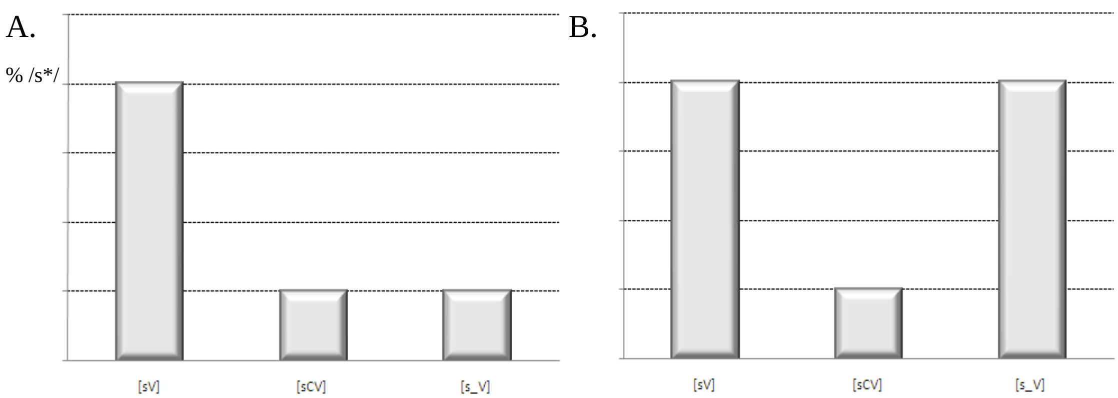
Table III.9 Incorrect Response Time Rates of Each Listener in Experiment 4

Listener group	Subject index	Raw number of IRT	Total expected responses	% IRT
*Group A*_ADV	Subj-A1	0	72	0%
	Subj-A2	2	72	2.78%
	Subj-A3	2	72	2.78%
	Subj-A4	5	72	6.94%
	Subj-A5	0	72	0%
	Subj-A6	2	72	2.78%
	Subj-A7	0	72	0%
	Subj-A8	8	72	11.11%
	Subj-A9	46	72	63.89%
*Group B*_INT	Subj-B1	0	72	0%
	Subj-B2	0	72	0%
	Subj-B3	3	72	4.17%
	Subj-B4	1	72	1.39%
	Subj-B5	1	72	1.39%
	Subj-B6	0	72	0%
	Subj-B7	1	72	1.39%
	Subj-B8	12	72	16.67%
*Group C*_NEW	Subj-C1	0	72	0%
	Subj-C2	2	72	2.78%
	Subj-C3	4	72	5.56%
	Subj-C4	0	72	0%
	Subj-C5	1	72	1.39%
	Subj-C6	2	72	2.78%
	Subj-C7	1	72	1.39%
	Subj-C8	12	72	16.67%

3.6.2 Predictions

With regard to Korean listeners' /s*/-responses in the three contexts, two patterns of outcomes seem probable. These possible outcomes are represented schematically in Figure III.5.

Figure III.5 Schematic Representations of Two Possible /s*/-response Patterns in [sV], [sCV] and [s_V] Contexts



In A and B, the listeners' /s*/-responses are similar in that the proportion /s*/ responses in [sV] is substantially higher than that in [sCV]. This is consistent with the results of Experiment 3, which showed that Korean listeners did not attend to the fricative-internal properties of duration and CoG. Listeners tended to identify prevocalic alveolar fricatives as $/s^*/_K$ and preconsonantal fricatives as $/s/_K$, irrespective of the acoustic properties of the fricatives. Such a response pattern can be explained in one of three ways, as discussed in the previous chapter. Listeners may rely on (i) acoustic differences in the frication portions that were not controlled for in the stimuli of Experiment 3, (ii) acoustic properties of the stimuli that are realized external to the fricative, most likely on neighboring vowels, or (iii) an abstract mapping rule.

However, A and B differ in the predictions for /s*/-responses to [s_V] stimuli. If Korean listeners' response pattern is as in Figure III.5A, it would eliminate the possibility of an abstract mapping rule because listeners perceive the acoustically identical /s/s in [sCV] and [s_V] as /s*/ the same proportion of the time. The response pattern for [sCV] and [s_V] contexts in Figure III.5B, however, allows the possibility that phonological contextual effects are involved.

Finally, the responses to [sV] and [s_V] contexts in Figure III.5A would indicate that listeners perceive acoustically different /s/$_E$'s differently in spite of the fact that they appear in the same phonological context. On the other hand, the responses to the same contexts in Figure III.5B would imply that listeners do not attend to consonantal properties but respond based on the phonological contexts in which /s/$_E$ appears.

In summary, a response pattern as in Figure III.5A would suggest that listeners rely on some phonetic properties of the stimuli rather than merely on the phonological context in which the /s/$_E$ appears. A response pattern as in Figure III.5B, on the other hand, may mean that listeners rely on an abstract mapping principle (i.e., their perception is based purely on the phonological context in which /s/$_E$ appears). However, I will show in Chapter 5 that an alternative explanation for a response pattern such as that in Figure III.5B is possible – specifically, an explanation that does not rely on a postulating an abstract mapping principle independent from acoustics and perception.

The predictions for the different listener groups are as follows: the more listeners are exposed to English, the more likely they are to know that there is only one category of voiceless alveolar fricative in English, and the more likely will be to ignore fine acoustic differences between different English /s/'s, and perceive all /s/$_E$'s the same (either as /s/$_K$ or as /s*/$_K$). Listeners with less English exposure and presumably lower English proficiency, may be more likely to interpret the fine acoustic differences between different English /s/'s in Korean terms, i.e., as evidence of either /s/$_K$ or /s*/$_K$. It is therefore possible that less experienced listeners are more likely to classify some /s/$_E$'s as /s/$_K$ and others as /s*/$_K$.

Another possible prediction is that the extent of listeners' exposure to English does not matter in accomplishing the given identification task, possibly because the experimental task (and loanword nativization in general) does not encourage listeners to attend to their English knowledge, but rather to utilize their native linguistic knowledge. In this case, there would be no effect of experience on response patterns.

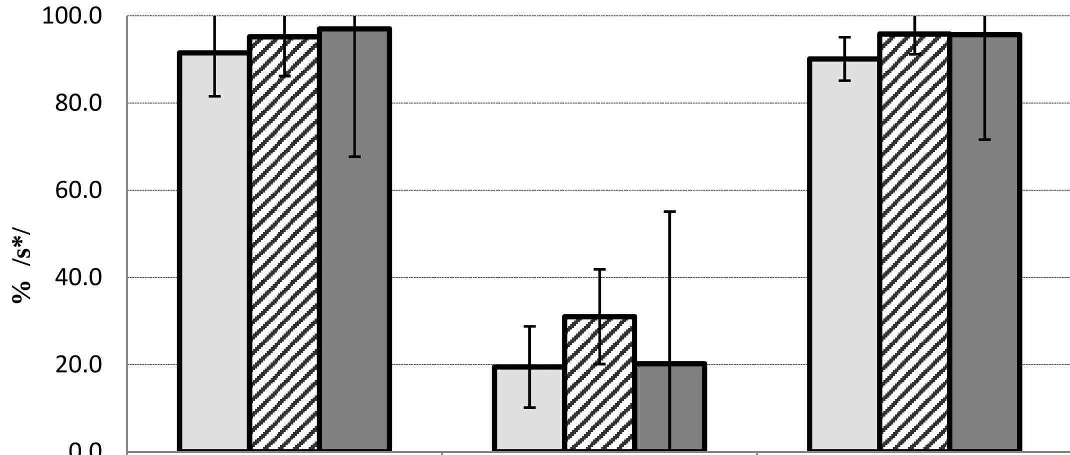
3.6.3 Results

The [sV] and [s_V] conditions represented the same phonological context (prevocalic). The [sCV] and [s_V] conditions represented the same acoustic information (identical [s] and [V] portions). The identification results were examined for the influence of condition, therefore testing both for the contribution of phonological context and acoustic information, and listener groups.

For the subject analysis, the data were entered into a univariate ANOVA in which listener group (i.e., *Group A*_ADV, *Group B*_INT and *Group C*_NEW) and condition (i.e., [sV], [sCV] and [s_V]) were fixed factors, and percent /s*/-response was the dependent variable. In the by-subject analysis, a significant main effect of condition was found (F (2, 61) = 101.916, p < 0.001), indicating that Korean listeners' identifications were significantly influenced by condition. Listener group was not found to be significant (F (2, 61) = 0.75, p = 0.477). The interaction between condition and listener group was not significant (F (4, 61) = 0.239, p = 0.915). Tukey's HSD post-hoc tests confirmed that the /s*/-response rate was significantly higher in [sV] than in [sCV] (p < 0.001), and also significantly higher in [s_V] than in [sCV] (p < 0.001). However, no significant difference was found between [sV] and [s_V] contexts (p = 0.99).

A similar pattern of results held for the item analysis. The results of a univariate ANOVA showed main effects of both condition (F (2, 52) = 955.549, p < 0.001) and listener group (F (2, 52) = 7.13, p = 0.02). The interaction between condition and listener group was not significant (F (4, 52) = 2.321, p = 0.71). Tukey's post-hoc pairwise comparisons of the item analysis showed that Group A_ADV and Group B_INT were significantly different (p = 0.01). The other comparisons among listener groups were not significant. Tukey's post-hoc pairwise comparisons also showed that [s_V] were significantly different from [sCV] (p < 0.001), and that [sCV] were different from [sV] (p < 0.001). Pooled identification results are displayed in Figure III.6.

Figure III.6   Korean listeners’ /s*/-response in [sV], [sCV] and [s_V]



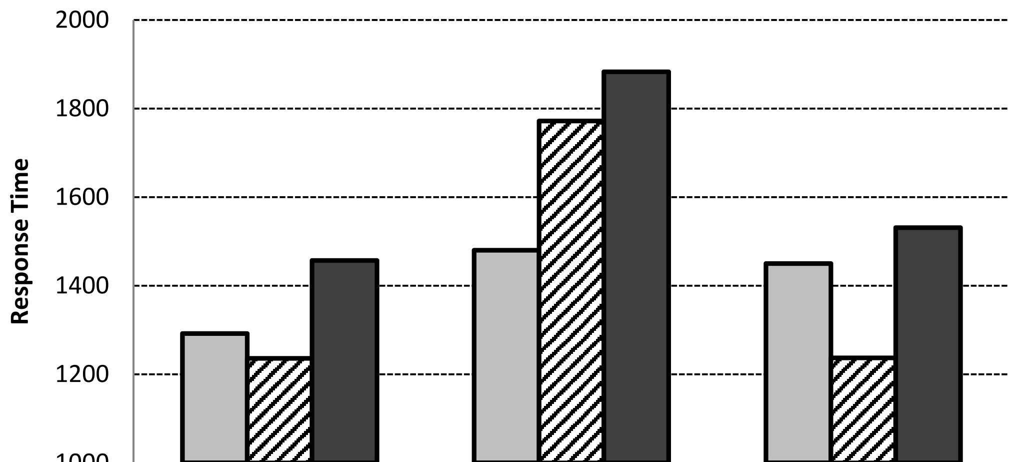
	sV	sCV	s_V
Group A	91.5	19.5	90.1
Group B	95.2	31.0	95.8
Group C	97.0	20.2	95.7

The overall pattern in Figure III.6 corresponds to the schematic outcome in Figure III.5B. Korean listeners /s*/-responses in preconsonantal contexts (i.e., [sCV]) were significantly lower than that in both prevocalic contexts (i.e., [sV] and [s_V]). The conditions with the same phonological context (prevocalic) patterned together in spite of the fact that these two conditions differed in terms of acoustic properties. On the other hand, the acoustically similar contexts (i.e., [sCV] and [s_V]) patterned differently. This outcome would appear to the interpretation that Korean listeners’ perception is likely determined by an abstract mapping principle rather than by the acoustic properties of the tokens.

Response time was examined as well. A univariate ANOVA was conducted in which listener group (i.e., *Group A*_ADV, *Group B*_INT and *Group C*_NEW) and condition (i.e., [sV], [sCV] and [s_V]) were fixed factors, and RT was a dependent factor. In a by-subject analysis, RT was significantly different according to conditions (F (2, 60) = 3.953, p = 0.025) but not according to listener groups (F (2, 60) = 1.465, p = 0.24). Tukey’s pair-wise comparisons showed that the [sV ~ sCV] contrast was significant, and that the other contrasts were not significantly different. In a by-item analysis, both listener group and condition showed main effects (F (2, 51) = 11.848, p < 0.001 and F (2,

51) = 32.222, $p < 0.001$, respectively). Figure III.7 describes the average RTs according to listener groups and conditions.

Figure III.7 Average Response Time by Listener Groups and Conditions in Experiment 4



	sV	sCV	s_V
Group A	1292	1480	1450
Group B	1236	1772	1237
Group C	1457	1883	1531

3.6.4 Discussion

The results of Experiment 4 point toward a phonological context account because listeners are responding differently to identical vowels and fricatives when they appear in different phonological contexts. In the [sCV] and [s_V] contexts, the acoustic properties of the fricative and vowel are exactly the same. Nevertheless, Korean listeners perceived significantly more /s*/ in [s_V] than in [sCV]. This finding confirms that their fricative perception is independent from /s/-internal properties, eliminating the possibility of involving any frication internal cues in their perception. Moreover, the fact that Korean listeners perceived /s*/ in [s_V] as often as they did in [sV] opens up the interpretation that phonological context counts.

The responses to the [sV] ~ [s_V] stimuli further confirm that Korean listeners are not basing their response patterns on /s/-internal information. The acoustic properties of [s] in [s_V] and [s] in [sV] were not the same, however, Korean listeners consistently showed dominant /s*/-responses in [s_V] as much as they did in [sV]. This pattern can be explained in two ways. One possibility is that the listeners are attending to phonological context because, in spite of the fact that the [s] portions of [sV] and [s_V] were not identical, the two [s]'s appear in the same surface phonological context (pre-vocalic) and the listeners perceived them as the same.

Now, let us return to examination of the acoustic characteristics of the stimuli used in Experiment 4. As raised earlier, while the acoustic properties of fricative duration, f_0 and $H1$-$H2$ of the stimuli were similar to previous findings, CoG of [sV] and [s_V] were slightly higher than in previous findings. In the following discussion, I examine whether or not the acoustic characteristics correlate with Korean listeners' /s/ or /s*/-responses.

Table III.10 repeats the acoustic characteristics of each stimulus used in Experiment 4, and illustrates percentages of Korean listeners /s/ and /s*/-responses for individual stimuli.

Table III.10 Acoustic Characteristics of Each Stimulus (partly repeated from Table III.7) and Percentages of /s/ and /s*/-responses for Each Stimulus

condition	stimuli	% /s/ -response	% /s*/ -response	Dur	CoG	f_0	$H1$-$H2$
[sV]	*sailit*	8.4	91.6	168	9297	198	-9.1
	saulet	2.5	97.5	184	9313	201	-13.6
	saurem	4.8	95.2	207	9780	202	-12
	silet	1.2	98.8	194	9703	226	-5.5
	silut	9.5	90.5	172	8781	201	-8.3
	sirem	7.2	92.8	206	9793	215	-8.6
[sCV]	*stailit*	76.2	23.8	166	9285	200	-9
	staulet	84.5	15.5	127	8581	196	-14.6
	staurem	75.9	24.1	180	8430	205	-9
	stilet	71.1	28.9	206	8997	197	-13
	stilut	78.8	21.3	114	8170	205	-13.4
	stirem	72.3	27.7	147	9325	211	-12.6
[s_V]	*s_ailit*	7.3	92.7	identical to above			
	s_aulet	9.5	90.5				
	s_aurem	2.4	97.6				
	s_ilet	3.8	96.3				
	s_ilut	9.8	90.2				
	s_irem	6.2	93.8				

Most importantly, given that CoG values of the stimuli were higher than the average of [sV] (of Experiment 1), and that high CoG is compatible with prevocalic /sV/ resulting in /s*/ percepts, it may be that this acoustic property of the stimuli lead to dominant /s*/-responses. However, dominant /s/-responses in [sCV] verify that listeners did not rely solely on CoG – if they did, then we would have expected an /s*/-dominant response pattern also in this condition.

Another concern that can be raised regarding the stimuli is the rather long fricative duration in the stimuli of the [sCV] contexts. Unfortunately, later in the experimental process, it was found that the fricative durations of *staurem* and *stilet* in the [sCV] context corresponded to prevocalic $/s/_E$ (i.e., 180 ms and 206 ms, respectively). Yet, as responses to individual stimuli in the above table show listeners did not give a particularly high proportion of /s*/-responses to the two stimuli that have longer [s]'s.

Listeners /s*/-responses were quite similar across the [sCV] contexts. As a result, it is unlikely that any direct correlation is exists between the consonantal properties of the stimuli and the response patterns.

As illustrated in Figure III.7, response time varied according to the phonological contexts. In general, it took longer for listeners to respond in [sCV] contexts. Response time in [sV] and [s_V] contexts were faster than in the [sCV] context. It is likely that [sCV] takes longer because it is against phonotactic constraints in Korean listeners' native language, which arguably implies that additional processing is required in processing [sCV]. Response time seems to correlate with English proficiency; as listeners' exposure to English increased, their response time became lower.

3.7 General Discussion and Summary of the Chapter

The additional analyses of the production data of Experiments 1 and 2 showed that the $H1$-$H2$ of $[s]_E$, in both preconsonantal and prevocalic contexts, is more similar to that of $[s^*]_K$ than to that of $[s]_K$. The results of Experiment 4 showed that listeners do use phonological context (in [sCV] vs. [s_V]). Given that the fricative properties of [s] in the [sV] and [s_V] contexts represent prevocalic $/s/_E$ and preconsonantal $/sC/_E$ respectively, the [sV] and [s_V] results can be interpreted as showing that listeners do not attend to fricative information. That is, the results confirmed that Korean listeners are insensitive to consonantal information, which provided further evidence for rejecting Hypothesis 1.

The experimental findings of this chapter can be understood as follows: First of all, Korean listeners may apply an abstract phonological mapping rule and acoustic cues play no role at all since Experiment 4 showed Korean listeners' sensitivity to phonological contexts. That is, listeners identify fricatives in identical vowel contexts very differently, defending on phonological contexts. This means that having acoustically different vowels is not a necessary condition for perceiving different fricatives. In Experiment 4, it can be understood that having different contexts is a sufficient for perceiving different fricatives. Experiment 4 does not rule out the possibility that systematically different vowels might also be a sufficient condition for perceiving different fricatives.

CHAPTER IV
PERCEPTUAL ADAPTATION AND NATIVE PHONOLOGY

4.1 Introduction

The results of Experiment 3 showed that Korean listeners predominantly perceive /s*/$_K$ in prevocalic contexts and /s/$_K$ in preconsonantal contexts, regardless of manipulated properties of the frication noise. This outcome suggests that Korean listeners do not exploit consonant-internal information as the major perceptual cue for the perception of alveolar fricatives. The results of Experiment 4 showed that listeners are more likely to perceive /s*/$_K$ in [s_V] (where C was deleted) than [sCV] contexts, even when the [s] and V are identical, suggesting that they attend to phonological context.

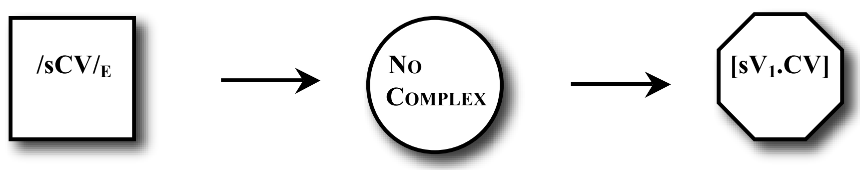
The purpose of this chapter is to investigate the possible contribution of temporal (rather than spectral, as in Chapter 3) vocalic cues in /s/$_E$ adaptation, and especially to provide a detailed account of preconsonantal /sC/$_E$ adaptation. To this end, I analyze the durational properties of vowels that follow /s/$_E$, /s/$_K$ and /s*/$_K$, and describe the properties of vowel duration in various contexts: Korean words (from Experiment 2), Korean non-words (Experiment 5) and English loans (Experiment 6). I also conducted a perception experiment that tested Korean listeners' perception of English preconsonantal /s/, i.e. /sC/. This experiment shows evidence that Korean listeners perform perceptual vowel epenthesis between the /s/ and /C/ in these tokens (Experiment 7). This suggests that, even in /sC/ tokens, the identification of /s/ may depend on the properties of a following vowel, even if the vowel in this case is not an acoustically present but a perceptually epenthesized vowel. The results of these experiments give evidence about the interplay between phonetics and phonology in loanword adaptation.

4.1.1 Hypothesis 2 (H_2) Revisited: Vocalic Information Hypothesis

The previous chapter, in testing Hypothesis 2, focused on prevocalic /sV/$_{E}$. In this chapter, the same hypothesis is concerned with preconsonantal /sC/$_{E}$. Regarding preconsonantal /sC/$_{E}$, H_2 predicts that the illusory vocalic properties of /sC/$_{E}$ are similar to /s/$_{K}$ (H_2a), and Korean listeners employ these acoustic properties in their English perception (H_2b).

I argue that, even though the /s/ in /sC/$_{E}$ is preconsonantal rather than prevocalic, Korean listeners' preconsonantal /sC/$_{E}$ adaptation is guided by vocalic information. Preconsonantal /sC/$_{E}$ is not phonotactically well-formed in Korean. Therefore, when presented with such input, Korean listeners perform perceptual epenthesis to create a form that is phonotactically legal in Korean, as illustrated in Figure VI.1.

Figure IV.1 Perceptual Modification of Illegal /sCV/$_{E}$ to [sV.CV]$_{K}$



A foreign input sequence /sCV/ is first evaluated and considered as being ill-formed. This input is processed according to native language grammatical constraints or rules. In other words, Korean listeners have learned that their native language does not allow consonant clusters on the surface. The illegal sequence then undergoes structural modifications so that it is perceived as a well-formed, permissible representation. In this modification, the native language filter is actively involved. This filter is shaped on the basis of native phonological knowledge (i.e., a phonotactic constraint, NoComplex for instance). Typically, in Korean loanword phonology, clusters are broken into two different syllables with vowel epenthesis as in [s$\underline{\underline{V}}$.CV] in which the underlined vowel is epenthesized. Vowel epenthesis results in the form [sV_1CV_2]$_{E}$ from preconsonantal /sC/$_{E}$, wherein V_1 is an epenthetic vowel.

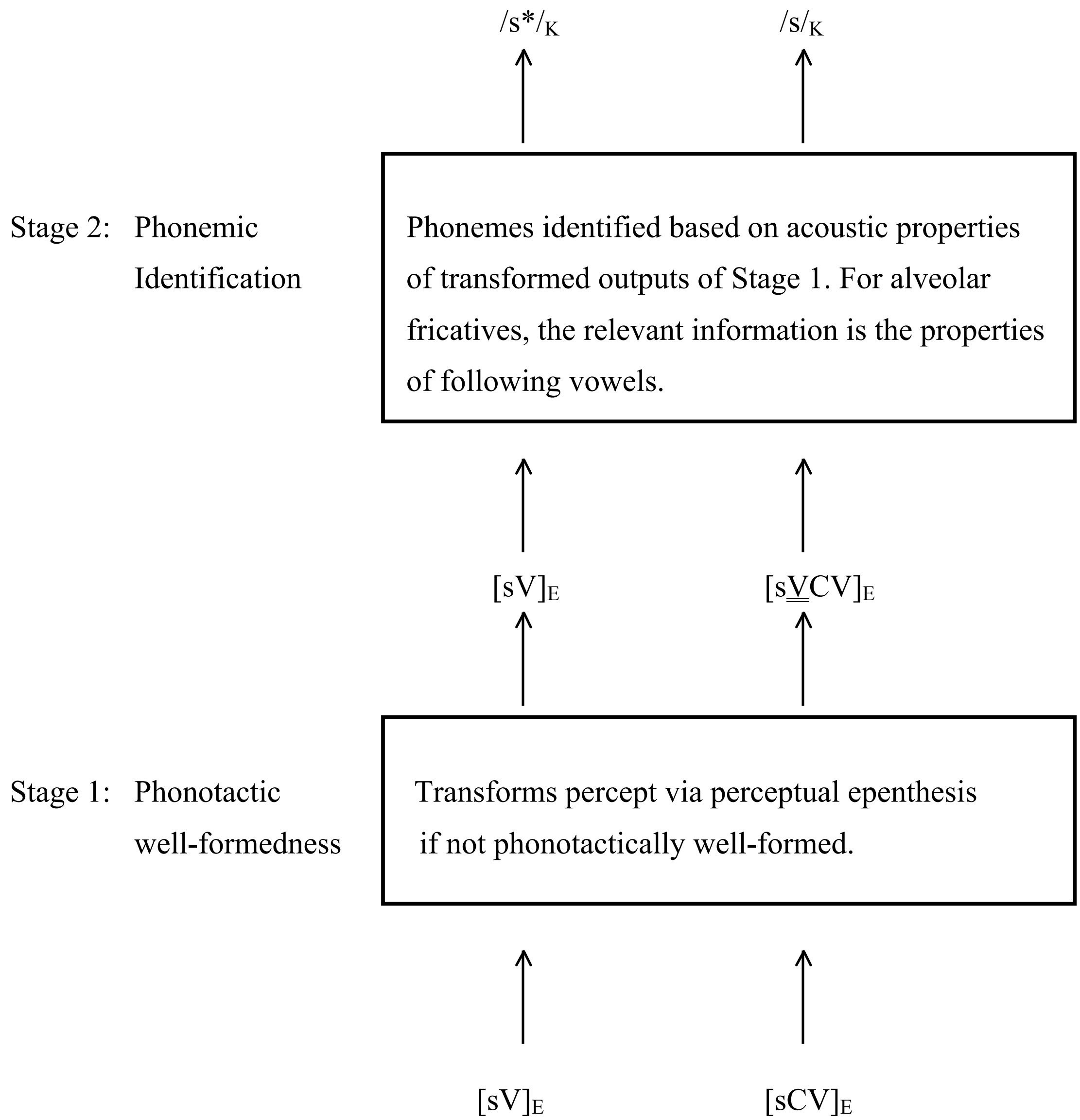
The epenthesized form is then submitted to the phoneme identification component of the perceptual grammar. From this point onward the perception works exactly the same as for inputs that actually contain an acoustically present vowel. That means that the property of the post-fricative vowel determines the identification of the fricative as either $/s/_K$ or $/s^*/_K$. The main difference between pre-consonantal and pre-vocalic contexts is hence that the post-fricative vowel is acoustically present in the pre-vocalic context, but is only perceptually present in the pre-consonantal context.

H_2 for preconsonantal $/sC/_E$ specifically considers the epenthetic vowel and the intrinsic nature of this vowel. The proposal is that the phonetic features of the epenthetic vowel are more similar to post-lax vowels (i.e., vowels that appear after lax consonants) than post-tense ones (i.e., vowels that appear after tense consonants). Consequently, the perceptually epenthetic vowel should be more like a post $/s/_K$ vowel than like a post $/s^*/_K$ vowel. If so, this would arguably explain why preconsonantal $/sC/_E$ – which is perceptually followed by a vowel with post-lax onset properties – is perceived as a $[s]_K$. In summary, H_2, the perceptual epenthesis hypothesis, is that the perceptual identification of an alveolar fricative in English /sC/ depends on the acoustic properties of the epenthetic vowel.

The above approach assumes a two stage perceptual processing module (Coetzee, 2010; Kingston, 2005; Poepel *et al.*, 2008).[4] In this model, Stage 1 is a stage of acoustic encoding during which a cognitive representation of the acoustic stimulus is created that is faithful to the actual acoustic stimulus. This initial acoustically encoded stimulus is compared to the phonotactic constraints of the language, and repaired via processes such as perceptual epenthesis, if the percept is found to violate phonotactic constraints. In Stage 2, phoneme identification is performed on the output of Stage 1. Crucially, Stage 2 does not differentiate the source of the inputs. Regardless of whether a vowel present at Stage 2 was present in the acoustic input or was supplied via perceptual epenthesis, it is evaluated equally. Figure IV.2 illustrates the schematic representation of this two-stage processing model.

[4] These assumptions are controversial. For instance, exemplar theory (Johnson, 1997; Pierrehumbert, 2001), which assumes that listeners store labeled exemplars of actual speech events, takes a different approach.

Figure IV.2 Schematic Representation of Two Stage Processing Model



4.1.2 Epenthetic vowels

Vowel epenthesis is a phonological process that inserts a vowel in an utterance, most often in order to convert the utterance into a phonotactically permissible string. It is widely agreed that a default vowel is employed as an epenthetic vowel and that the identity of the default vowel is determined within individual languages. In contrast, definitions of what counts as a default vowel vary among researchers, depending on how the nature of the vowel is approached. A phonetically based approach considers the vowel that is perceptually least salient in the language to be default (Kenstowicz, *in press*; Shinohara, 1997; Steriade, 2001). The default vowel is phonetically short and thus least disruptive. Others take a phonologically based approach to the nature of default vowels such that the universally least marked vowel (Lombardi, 2002) or an underspecified vowel (Abaglo and Archangeli, 1989) is the default. While these two approaches consider the default to be context-independent (i.e., the default vowel assignment is independent from the phonological context in which it appears), other approaches claim that the quality of default vowel is not fixed but rather that default assignment follows contextually dependent principles. For instance, an epenthetic vowel may be a copy of a preceding vowel (Kitto and Lacy, 1999), or the vowel may receive a place feature from an adjacent consonant (Uffmann, 2006). In these cases, the default is determined by its neighbors rather than being a fixed segment (i.e., context-dependent).

In Korean, vowel epenthesis is not attested as a productive native phonological process. It is only found in loans as a strategy to resolve phonotactic ill-formedness. In general, Korean employs [ɨ] as the epenthetic vowel to modify phonotactically ill-formed loans, so that this vowel is considered as the default vowel in Korean. Except for words that end with one of /ʃ, ʤ, c, c^h/, which are resolved with [i] epenthesis as in *'bush'* /bʊʃ/ → [pu.ʃi], *'judge'* /ʤəʤ/ → [cʌ.ci], '*mirage'* /məɹɑʒ/ → [mʌ.ɹɑ.ci] and *'switch'* /swɪtʃ/ → [sɨ.wi.c^{h}i], most English loans are resolved with the default [ɨ] vowel as in *'strike'* /stɹɑɪk/ → [sɨ.t^{h}ɨ.ɹɑ.i.k^{h}ɨ] and *'bright'* /bɹɑɪt/ → [bɨ.ɹɑ.i.t^{h}ɨ].

The vowel [ɨ] is not limited to loans. It is frequently found in native Korean words as a lexical vowel. Also, its usage is not restricted to certain environments; it can follow both lax and tense consonants as in [sɨ.mul] *'twenty'* and [s*ɨ.tɑ] *'write'*. Therefore, both epenthetic [ɨ] and lexical [ɨ] are attested in Korean vocabulary.

4.2 Experiment 2 Revisited: Post-fricative Vowel Duration in Korean Words

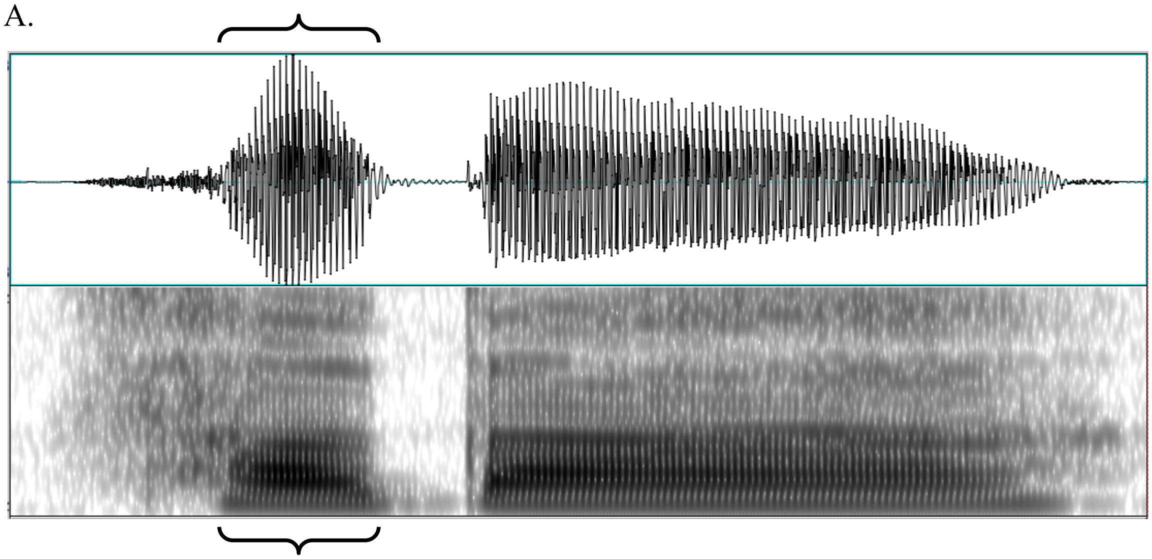
H_2 predicts that vocalic properties play a role in loanword adaptation. In this section, the vowel productions of the Korean speakers from Experiment 2 are re-examined to determine whether post-fricative vowel duration might be a relevant factor. Vowel durations in Korean non-words and those in English loans are also examined in Experiments 5 and 6, respectively.

4.2.1 Methodology

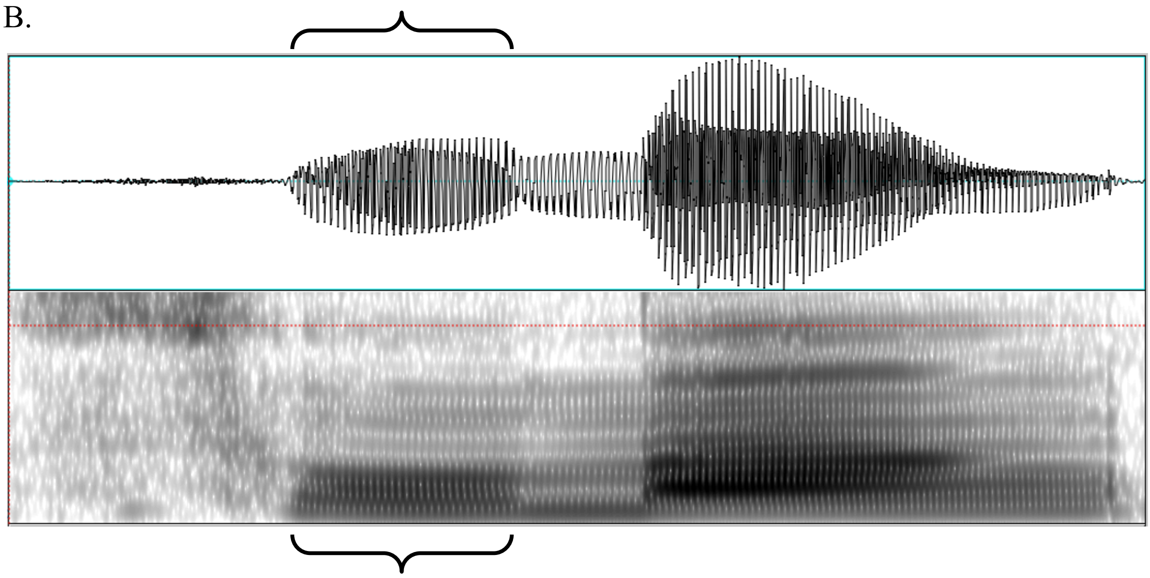
In Experiment 2, Korean-speaking participants produced 32 different bisyllabic words of the form [sV_1.CV_2(C)] or [s*V_1.CV_2(C)]. In the following, the temporal properties of V_1 (i.e., the duration of the first vowels) were examined. Between (voiceless) obstruents (e.g., [sɑtɑ]), vowel onset and offset were identified as the onset and offset, respectively, of periodic pulsing, as illustrated in Figure IV.3A. When the vowel is followed by a sonorant (e.g., [somɑŋ]), offset of the vowel was taken to be the point where a clear change in the waveform was observed, as in Figure IV.3B. The braces in each figure indicate the vowel portions.

Figure IV.3 Waveform of [sɑtɑ] *'buy'* (A) and [somɑŋ] *'hope'* (B) and V_1 Duration Measurements

A.



B.



Many factors can influence phonetic duration of a vowel. Vowel duration varies depending on syllable structure (Maddieson, 1985), vowel height, (Maddieson, 1997) prosody (Benguerel, 2002) and kinds of adjacent segments (Braunschweiler, 1997), for instance. It has been experimentally shown that vowels in closed syllables are usually shorter than vowels in open syllables. It is also shown that falling intonation results in vowel lengthening (Benguerel, 2002). Another contextual effect follows from the segments in adjacent positions. Vowels followed by voiceless consonants are usually shorter than vowels followed by voiced consonants. For instance, Braunschweiler (1997) experimentally showed that word-medially vowels become longer when followed by voiced stops rather than voiceless stops in German.

Although I acknowledge these issues, the design of the $/s/_K$ and $/s^*/_K$ stimulus lists was not balanced to control for these possible factors. However, several aspects of the design minimized the influence of these factors on duration. All tokens were read as isolated words, thus excluding many prosodic factors. $/s/_K$ and $/s^*/_K$ lists were balanced for syllable structure, with equal numbers of open and closed syllables. Voicing of the following consonant, however, was difficult to control. Within $/s/_K$ lists, 6 post-$/s/_K$ vowels were followed by voiceless consonants, whereas within $/s^*/_K$ lists, 9 post-$/s^*/_K$ vowels were followed by voiceless consonants. Phonetic vowel shortening, therefore, is likely to be more marked for $/s^*/_K$ lists. Nevertheless, the results of this experiment could be taken as suggestive only. This is the reason for the following experiments that better control for these confounding factors.

4.2.2 Results

The durations for post /s/$_K$ and post /s*/$_K$ vowels were compared using a two-tailed paired *t*-test. Durations for the vowels in these two contexts were found to be significantly different ($t(30) = -2.81$, $p < 0.001$) in a by-item analysis, with the post /s*/$_K$ vowels being longer than the post /s/$_K$ vowels. A by-subject analysis was not conducted due to the small number of speakers. The vowel durations in each word are given in Table IV.1.

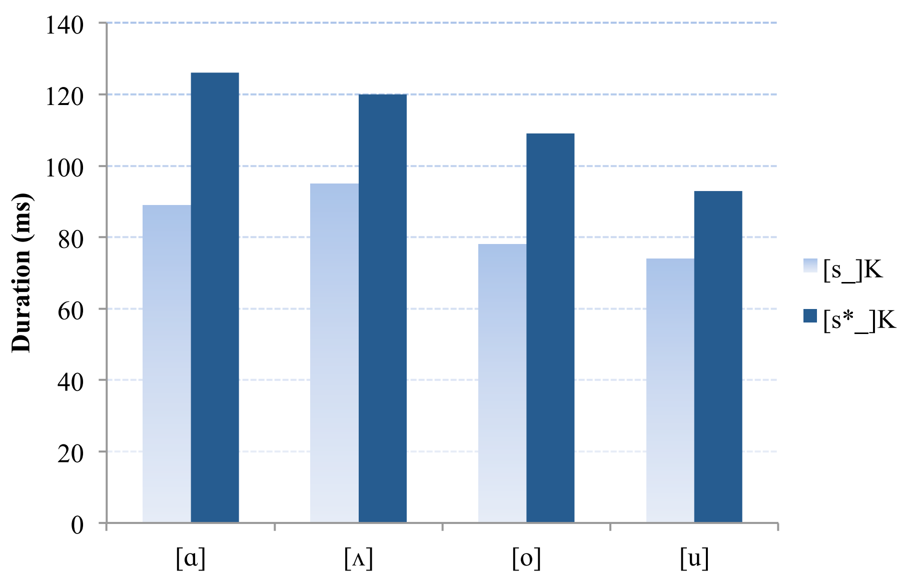
Table IV.1 Average Vowel Durations for Individual Words: post-/s/$_K$ and post-/s*/$_K$

	/s/$_K$ context			/s*/$_K$ context		
	words	Duration (ms)		words	Duration (ms)	
[ɑ]	/sɑ.tɑ/	106	(82 ~ 136)	/s*ɑ.tɑ/	145	(103 ~ 145)
	/sɑ.cin/	117	(91 ~ 142)	/s*ɑ.um/	179	(142 ~ 235)
	/sɑl.tɑ/	79	(61 ~ 111)	/s*ɑk.nun/	89	(75 ~ 107)
	/sɑm.cʰon/	54	(48 ~ 77)	/s*ɑm.ci/	92	(83 ~ 105)
	pooled	89	(48 ~ 142)	pooled	126	(75 ~ 235)
[ʌ]	/sʌ.tɑ/	114	(82 ~ 158)	/s*ʌ.re/	152	(114 ~ 203)
	/sʌ.ul/	111	(52 ~ 177)	/s*ʌ.cim/	141	(120 ~ 163)
	/sʌŋ.kjʌk/	60	(44 ~ 79)	/s*ʌk.tɑ/	83	(71 ~ 101)
	/sʌŋ.ɨi/	95	(68 ~ 145)	/s*ʌl.tɑ/	103	(85 ~ 114)
	pooled	95	(44 ~ 177)	pooled	120	(71 ~ 203)
[o]	/so.kæ/	103	(66 ~ 147)	/s*o.tɑ/	124	(94 ~ 164)
	/so.mɑŋ/	95	(55 ~ 153)	/s*o.im/	125	(94 ~ 194)
	/sok.tɑ/	52	(40 ~ 66)	/s*on.sɑl/	95	(80 ~ 111)
	/sol.cik/	62	(44 ~ 75)	/s*ol.rim/	94	(82 ~ 109)
	pooled	78	(40 ~ 153)	pooled	110	(82 ~ 194)
[u]	/su.rɑk/	81	(40 ~ 137)	/s*u.tɑ/	120	(93 ~ 171)
	/su.jʌŋ/	84	(62 ~ 117)	/s*u.sim/	133	(101 ~ 180)
	/suk.mjʌŋ/	66	(56 ~ 78)	/s*uk.kɑs/	59	(54 ~ 72)
	/sun.su/	62	(42 ~ 87)	/s*uk.t*ɨm/	62	(50 ~ 69)
	pooled	73	(40 ~ 137)	pooled	94	(50 ~ 180)
TOTAL		84	(40 ~ 177)		112	(50 ~ 235)

The average duration of the vowels in post-/s/$_K$ contexts was 84 ms, whereas the duration in post-/s*/$_K$ was 112 ms. That is, there was about a 25% difference between the average durations of vowels in the two contexts. The average durations of each token are given in the figure along with their range (numbers in the parenthesis). In spite of some voiceless consonants in the second syllables in the /s*/ list, the vowels were longer in that context. This shows that the voicing of the following consonant probably did not influence the durations of the preceding vowels very much.

The results for the different vowels are displayed in Figure IV.4. The largest difference was found [ɑ] where the difference in the two fricative contexts was about 30%. The smallest difference was found [u] where the difference was about 20%. Overall, Korean speakers' vowel productions are significantly longer following word-initial tense than following word-initial lax fricatives.

Figure IV.4 Vowel Duration after [s]$_K$ and [s*]$_K$ in Korean words



4.2.3 Discussion

The results of the additional analyses of Experiment 2 data for Korean show that the post-fricative vowel was significantly longer after a tense than a lax consonant irrespective of vowel quality, and this finding is consistent with previous literature showing that post-tense consonant vowels are longer in duration than post-lax vowels (Jessen, 1998). Although the durational differences vary slightly according to vowel quality, in all contexts post $[s^*]_K$ vowels were longer than post $[s]_K$ vowels by at least 20%.

According to Bochner *et al.* (1988), the difference is perceptible in principle, that is, the difference exceeds the just noticeable difference (JND). This 20% (or more) durational difference after tense and lax consonants in production may influence Korean speakers' perception, a possibility that is discussed further at the end of this chapter.

4.3 Experiment 5: Post-fricative Vowel Duration in Korean Non-words

Experiment 5 was designed to examine the temporal properties of post-fricative vowels in Korean non-words. Non-words were used in order to avoid any potential lexical effects. Note that, although the target tokens are non-words, they contain a lexically full vowel (i.e., [ɨ] is a lexical vowel rather than an epenthetic one in this experiment).

4.3.1 Design and methodology

4.3.1.1 Participants

Six native speakers of Korean were recruited around the University of Michigan's Ann Arbor campus. They were Seoul dialect speakers who were born and lived in Seoul or around Seoul (Gyungki province) until their 20s. None of them had speaking or hearing difficulties and they were paid for their participation.

4.3.1.2 Reading materials

The reading materials consisted of 90 bisyllabic Korean non-words of the form [$C_1V_1.C_2V_2C_3$]. The first syllable of each word was always either /sɨ/ or /s*ɨ/. The onset consonant of the second syllable (i.e., C_2) was one of [p], [t], [k], [m], [n], [r], [c^h], [s] or [h], and the consonants were followed by one of [ɑ], [i], [ʌ], [o] or [u] (i.e., V_2). In order to control for the influence of adjacent segments, the /s/$_K$ and /s*/$_K$ lists were balanced with regard to C_2 and V_2. C_3 was always [k]. These lists, therefore, controlled for the potential confounding factors that were present in Experiment 2.

4.3.1.3 Procedure

The wordlist, which was written in Korean orthography, was presented to the participants in random order. The participants were instructed to read the Korean non-words as if they were existing Korean words. They read the same randomization 5 times. Six participants read the 90 words 5 times, which generated 2700 non-words in total (i.e., 6 speakers × 90 words × 5 repetitions). The details of the recording procedure were same as Experiment 1 and 2. There was one test session, which took 10 ~ 20 minutes.

4.3.2 Results

Vowel duration was measured using the method described for the additional analyses of Experiment 2 data. The collected measurements for Korean non-words were subjected to a by-item two-tailed paired sample *t*-test which showed that the vowel duration was significantly different between the two contexts ($t(88) = -6.613$, $p < 0.001$). Due to the small number of participants, the by-subject analysis was not performed. The post tense vowel was significantly longer than the post lax. That is, [ɨ] in /s*ɨ/$_K$ was significantly longer than [ɨ] in /sɨ/$_K$. The average duration of post tense [ɨ] was 82 ms, and that of post lax [ɨ] was 58 ms, which was about a 30% difference. The average durations of post-tense and post-lax vowels (across tokens and speakers) for each non-word are provided in Table IV.2.

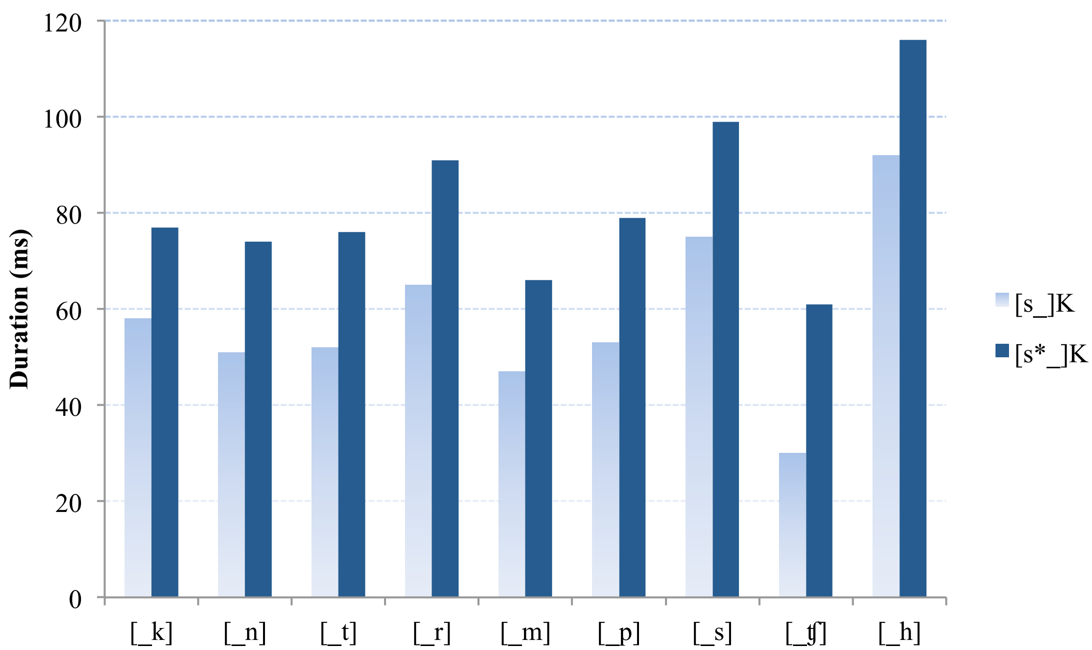
Table IV.2 Duration of Post-tense and Post-lax Vowels in Korean Non-words

consonant	vowel	/s/$_K$		/s*/$_K$	
		IPA	Duration (ms)	IPA	Duration (ms)
p	ɑ	/sɨ.pɑk/	49	/s*ɨ.pɑk/	78
	i	/sɨ.pik/	59	/s*ɨ.pik/	80
	ʌ	/sɨ.pʌk/	47	/s*ɨ.pʌk/	76
	o	/sɨ.pok/	48	/s*ɨ.pok/	79
	u	/sɨ.puk/	63	/s*ɨ.puk/	80

t	ɑ	/sɨ.tɑk/	40	/s*ɨ.tɑk/	68
	i	/sɨ.tik/	58	/s*ɨ.tik/	87
	ʌ	/sɨ.tʌk/	50	/s*ɨ.tʌk/	72
	o	/sɨ.tok/	54	/s*ɨ.tok/	73
	u	/sɨ.tuk/	60	/s*ɨ.tuk/	82
k	ɑ	/sɨ.kɑk/	50	/s*ɨ.kɑk/	76
	i	/sɨ.kik/	64	/s*ɨ.kik/	84
	ʌ	/sɨ.kʌk/	55	/s*ɨ.kʌk/	72
	o	/sɨ.kok/	59	/s*ɨ.kok/	73
	u	/sɨ.kuk/	63	/s*ɨ.kuk/	82
m	ɑ	/sɨ.mɑk/	46	/s*ɨ.mɑk/	67
	i	/sɨ.mik/	49	/s*ɨ.mik/	68
	ʌ	/sɨ.mʌk/	47	/s*ɨ.mʌk/	65
	o	/sɨ.mok/	44	/s*ɨ.mok/	62
	u	/sɨ.muk/	51	/s*ɨ.muk/	68
n	ɑ	/sɨ.nɑk/	51	/s*ɨ.nɑk/	78
	i	/sɨ.nik/	54	/s*ɨ.nik/	79
	ʌ	/sɨ.nʌk/	49	/s*ɨ.nʌk/	71
	o	/sɨ.nok/	48	/s*ɨ.nok/	66
	u	/sɨ.nuk/	53	/s*ɨ.nuk/	76
r	ɑ	/sɨ.rɑk/	62	/s*ɨ.rɑk/	87
	i	/sɨ.rik/	68	/s*ɨ.rik/	101
	ʌ	/sɨ.rʌk/	61	/s*ɨ.rʌk/	89
	o	/sɨ.rok/	65	/s*ɨ.rok/	86
	u	/sɨ.ruk/	69	/s*ɨ.ruk/	93
c^h	ɑ	/sɨ.c^hɑk/	29	/s*ɨ.c^hɑk/	62
	i	/sɨ.c^hik/	28	/s*ɨ.c^hik/	61
	ʌ	/sɨ.c^hʌk/	29	/s*ɨ.c^hʌk/	61
	o	/sɨ.c^hok/	33	/s*ɨ.c^hok/	58
	u	/sɨ.c^huk/	33	/s*ɨ.c^huk/	65
s	ɑ	/sɨ.sɑk/	63	/s*ɨ.sɑk/	99
	i	/sɨ.sik/	76	/s*ɨ.sik/	100
	ʌ	/sɨ.sʌk/	75	/s*ɨ.sʌk/	95
	o	/sɨ.sok/	78	/s*ɨ.sok/	101
	u	/sɨ.suk/	84	/s*ɨ.suk/	102
h	ɑ	/sɨ.hɑk/	88	/s*ɨ.hɑk/	114
	i	/sɨ.hik/	97	/s*ɨ.hik/	119
	ʌ	/sɨ.hʌk/	85	/s*ɨ.hʌk/	113
	o	/sɨ.hok/	92	/s*ɨ.hok/	118
	u	/sɨ.huk/	98	/s*ɨ.huk/	116

Figure IV.5 represents the average durations of post-tense and post-lax vowels according to C_2 context. The duration varied slightly according to the place and manner of C_2, such that vowels were shortest when they were followed by the affricate [c^h] (or, [⍰]). More importantly, post-tense [ɨ] were significantly longer than post-lax [ɨ] in all contexts. Additionally, it was found that the participants frequently deleted [ɨ] vowels in their production. Those tokens from whose vowels were deleted contributed duration of zero ms to the calculation of the mean. Out of 2700 words, 109 words had no vowels, and of these 98 instances were post-lax (i.e., $[s\_]_K$), and only 12 were post-tense (i.e., $[s^*\_]_K$).

Figure IV.5 Vowel Duration after $[s]_K$ and $[s^*]_K$ in Korean Non-words, according to C_2 context



4.3.3 Discussion

Experiment 5 found that vowels following tense fricatives were significantly longer than those following lax ones. Depending on the consonant that followed the vowel, the durational differences varied; the difference was the largest when the vowel preceded [c^h] where the difference was around 50%, and the difference was the smallest when the vowel appeared before [m]. Overall, the vowels were the longest when they are followed by [h] and the shortest when followed by [c^h]. In all conditions, a post-tense [ɨ] was longer than a post-lax [ɨ].

The results of Experiment 5 are consistent with the findings in the previous section where the results from Experiment 2 were analyzed. That is, similar to Experiment 2, post-/s*/$_K$ vowels were significantly longer than post-/s/$_K$ vowels. The overall difference between post-tense and post-lax [ɨ] found in Experiment 5 was larger than the corresponding difference in the data from Experiment 2.

Another interesting outcome of Experiment 5 is that vowel deletion is sometimes attested. That is, unlike other vowels such as [ɑ], [ʌ], [o] and [u] examined in Experiment 2, [ɨ] was sometimes dropped in the utterances, which can be taken as support for the argument that [ɨ] is the default vowel in Korean. At least according to the phonetic view of default vowels, default vowels tend to be weak vowels. The fact that only [ɨ] deletes suggests that it is weaker than the other vowels of Korean, and hence likely the default vowel. The deletion rate of [ɨ] was considerably higher after a lax [s]$_K$ than after a tense [s*]$_K$, meaning that a vowel is more easily deleted after a lax than a tense consonant.

4.4 Experiment 6 – Post-fricative Vowel Duration in English Loans

Similar to the previous experiment, Experiment 6 was designed to examine the temporal properties of post-fricative vowels. In this experiment, I examine the epenthetic vowel [ɨ] in English loans. These loans are originally English words, but were borrowed into Korean sufficiently long ago that they are now used as everyday Korean words.

4.4.1 Design and methodology

4.4.1.1 Participants

The six Korean participants in Experiment 5 also participated in Experiment 6. After Experiment 5 was over, they had 10 ~ 15 minutes break, and began Experiment 6 when they felt comfortable.

4.4.1.2 Reading materials

Twelve English loans were selected on the basis of English syllable structure and their borrowed form. All of the words started with an /s/-cluster in English [$C_1C_2V(C)(C)$] and were monosyllabic words in English. While C_1 was [s] in all cases, C_2 was one of 6 consonants: [t], [k], [p], [m], [n] or [l] (thus, /stV_/, /skV_/, /spV_/, /smV_/, /snV_/ or /slV_/ in English). There were 2 words for each consonant, thus, 12 words in total. In Korean, the consonant clusters are broken up into separate syllables with an epenthetic vowel, thus, the borrowed representations are /sɨtV_/, /sɨkV_/, /sɨpV_/, /sɨmV_/, /sɨnV_/ or /sɨlV_/. The original English words and their borrowed representations are given in Table IV.3.

Most of the loans were borrowed as trisyllabic words in Korean. Exceptionally, the loan *'snack'* is borrowed as a bisyllabic word [sɨ.næk], thus an accusative affix '-ɨl' was added to this loan so that all the target words were trisyllabic. In order to collect natural Korean productions of English loans, the target words were embedded in Korean sentences. Given that the position of a word in the larger prosodic structure of an utterance can have an influence on its durational properties, all target words were placed in sentence-initial position.

Table IV.3 Wordlists for English loans production in Experiment 6

	loans	transcription	Korean orthography and Gloss	
t	*stand*	/sɨ.ten.tɨ. k^hjʌ.se.jo/	스텐드켜세요	*'Turn on the lamp'*
	stove	/sɨ.to.pɨ. k^hjʌ.se.jo/	스토브켜세요	*'Turn on the stove'*
k	*sketch*	/sɨ.ke.ʧi. hɑ.se.jo/	스케치하세요	*'(Please) sketch out'*
	skirt	/sɨ.kʌ.tɨ. ip.ʌ.jo/	스커트입어요	*'(I) wear a skirt'*
p	*sport*	/sɨ.po.ʧɨ. cɑl.hæ.jo/	스포츠잘해요	*'(I) am good at sports'*
	speed	/sɨ.pi.tɨ. cɨl,kjʌ.jo/	스피드즐겨요	*'(I) enjoy speed'*
m	*smile*	/sɨ.mɑ.il. hɑ.se.jo/	스마일하세요	*'(Please) smile'*
	smog	/sɨ.mo.kɨ. p^hi.hæ.jo/	스모그피해요	*'(I) avoid smog'*
n	*snow*	/sɨ.no.u. coh.ɑ.jo/	스노우좋아요	*'(I) like snow'*
	snack	/sɨ.næk.ɨl. mʌk.ʌ.jo/	스낵을먹어요	*'(I) am having snacks'*
l	*slope*	/sɨl.lo.p^hɨ. nop.ɑ.jo/	슬로프높아요	*'The slope is high'*
	slow	/sɨl.lo.u. hɑ.se.jo/	슬로우하세요	*'Slow down'*

4.4.1.3 Procedure

The Korean-speaking participants were instructed to read the sentences, which were presented in Korean orthography. The participants were told to read the sentences as naturally as possible. Six participants read the randomized lists of the 12 words 5 times, which resulted in 360 tokens in total (i.e., 6 speakers × 12 words × 5 repetitions). There was one test session, and it took 10 ~ 20 minutes.

4.4.2 Results

Vowel duration was measured in the same manner as in Experiment 5. The measurements were averaged for each word, and the average durations are given in Table IV.4.

Table IV.4 Average Vowel Duration for Each English Loans: with zero duration tokens (A) and without zero duration tokens (B)

A.

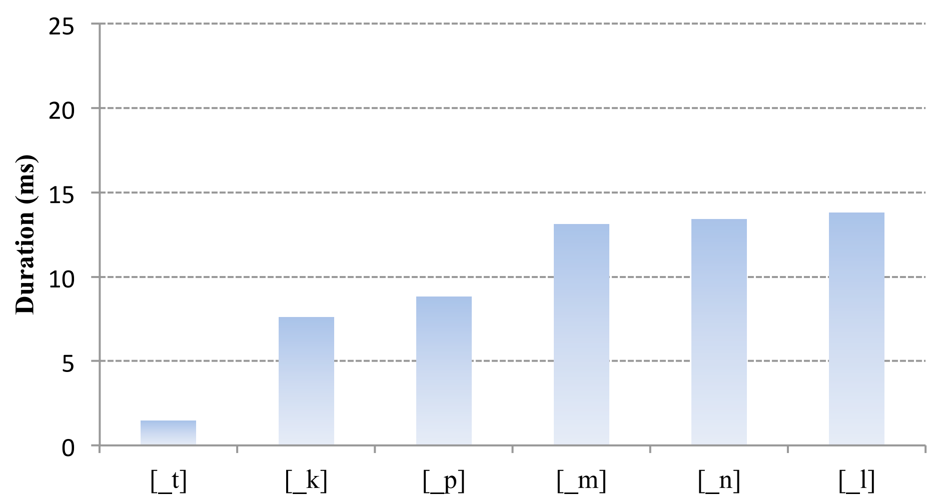
context	English loans	Duration (ms)	English loans	Duration (ms)
_t	/sɨ.ten.tɨ/	0.2 (0 ~ 6.5)	/sɨ.to.pɨ/	2.7 (0 ~ 26)
_k	/sɨ.ke.c^{h}i/	5.7 (0 ~ 39)	/sɨ.kʌ.tɨ/	9.5 (0 ~ 49)
_p	/sɨ.po.c^{h}ɨ/	8.9 (0 ~ 46)	/sɨ.pi.tɨ/	8.7 (0 ~ 49)
_m	/sɨ.mɑ.il/	12.6 (0 ~ 55)	/sɨ.mo.kɨ/	13.6 (0 ~ 56)
_n	/sɨ.no.u/	13.6 (0 ~ 43)	/sɨ.næk.ɨl/	13.3 (0 ~ 57)
_l	/sɨl.lo.p^{h}ɨ/	12.9 (0 ~ 63)	/sɨl.lo.u/	14.6 (0 ~ 48)

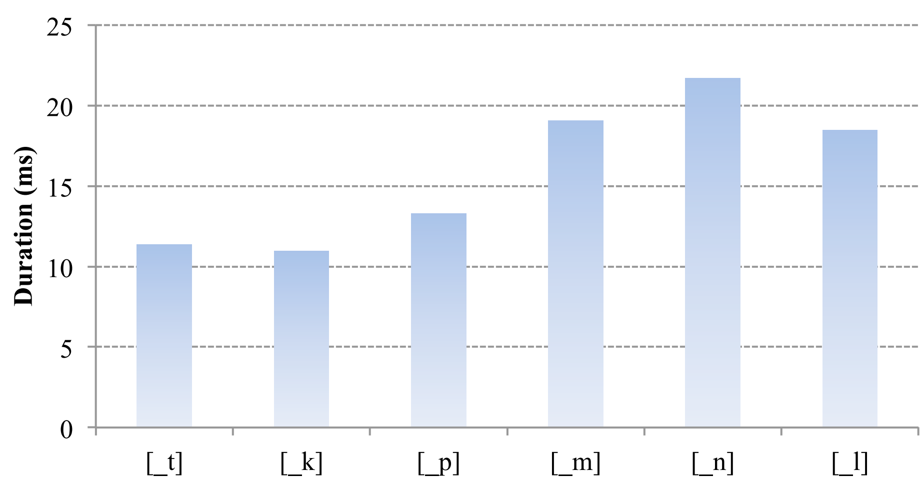
B.

context	English loans	Duration (ms)	English loans	Duration (ms)
_t	/sɨ.ten.tɨ/	2.2 (0.7 ~ 6.5)	/sɨ.to.pɨ/	20.5 (4.7 ~ 26)
_k	/sɨ.ke.c^{h}i/	9.5 (2.1 ~ 39)	/sɨ.kʌ.tɨ/	12.4 (0.6 ~ 49)
_p	/sɨ.po.c^{h}ɨ/	15.7 (2.4 ~ 46)	/sɨ.pi.tɨ/	10.9 (2.4 ~ 49)
_m	/sɨ.mɑ.il/	17.3 (2.5 ~ 55)	/sɨ.mo.kɨ/	20.8 (3.2 ~ 56)
_n	/sɨ.no.u/	20.5 (2.2 ~ 43)	/sɨ.næk.ɨl/	22.9 (2.8 ~ 57)
_l	/sɨl.lo.p^{h}ɨ/	16.1 (2.4 ~ 63)	/sɨl.lo.u/	20.8 (2.5 ~ 48)

Figure IV.6 displays the average vowel durations in each context with and without zero duration tokens respectively. In general, the epenthetic vowel was longer when the consonant clusters consisted of fricatives and sonorants (i.e., /sm/, /sn/ and /sl/) than when they consisted of fricatives and stops (i.e., /st/, /sk/ and /sp/). The vowel was the longest when it was inserted before [l], and the shortest before [t]. Across all contexts, the average duration was 9.7 ms, which included zero duration tokens. Zero-duration tokens were the most frequent before [t] (i.e., 52 times out of 360 tokens across all speakers) and the least frequent before [l] and [n] (i.e., 11 times out of 360 tokens across all speakers).

Figure IV.6 Vowel Duration after [s_] in English Loans:
with zero duration tokens (A) and without zero duration tokens (B)





Additionally, $H1$-$H2$ was measured. By definition, only non-zero duration tokens could be measured. Among non-zero duration tokens, those of which duration was over 10 ms were measured in order to be consistent (with previous experiments) in measuring $H1$-$H2$. There were 131 tokens that met the requirements for inclusion in this analysis.

Table IV.5 $H1$-$H2$ after [s_] in English Loans: 131 over-10 ms duration tokens only

context	$H1$-$H2$ (dB)
_t	-1.4
_k	-3.0
_p	-1.2
_m	-0.6
_n	-3.4
_l	1.6
Pooled	-1.3

4.4.3 Discussion

Experiment 6 investigated the duration and $H1$-$H2$ of the epenthetic vowel [ɨ] in English loans. Although the duration fluctuated depending on the components of consonant clusters, in general, it was found that the average duration of epenthetic vowel was comparatively short. The durational difference between an epenthetic and a lexical vowel is substantial; whereas the average duration of lexical [ɨ] was 82 ms in Experiment 5, the average duration of epenthetic vowels is 9.7 ms in Experiment 6 (16 ms excluding zero durations tokens). The average $H1$-$H2$ of epenthetic vowels was -1.3 dB, which was similar to that of $[sV]_K$.

Previous studies have noted that epenthetic and lexical vowels are not phonetically identical. Ladefoged *et al.* (1998) examined the pitch contour of epenthetic and lexical vowels in Scottish Gaelic, a language in which epenthesis is a native phonological process. They found that the pitch contour is more distinguishable in lexical vowels than epenthetic ones. Bosch and de Jong (1997) also studied that the durational properties of epenthetic vowels in Scottish Gaelic. In a monosyllabic word /CVCC/ with a short stressed vowel, a vowel is often added to the consonant cluster to create [CVCVC]. They reported that the duration of the epenthetic vowel is longer than the duration of a lexical vowel in similar contexts. On the other hand, Gouskova and Hall (2009) showed that in Lebanese Arabic the epenthetic vowel [i], which is added to some CCC sequences, is significantly shorter than a lexical [i], and that the former has lower F_2 values than the latter.

The durational difference found in Scottish Gaelic (Bosch and de Jong, 1997) and that found in Korean loans (Experiment 5 and 6) are opposite; epenthetic vowels are longer in duration (than lexical ones) in Scottish Gaelic, while they are shorter in Korean.

While the studies discussed above investigated various phonetic differences between epenthetic and lexical vowels in languages in which the epenthetic vowel emerged as the result of a native phonological process, Davidson (2006) reported the phonetic properties in which epenthesis arises in non-native sequences. She compared the durational properties of epenthetic schwas (i.e., Cᵊ C) to lexical schwas (i.e., CəC) in English, and found that lexical schwas are significantly longer than epenthetic ones. For instance, when the initial segment was /v/ as in [vᵊgalu] or [vəgalu], the duration of the epenthetic schwa [CᵊC] was 42.5 ms while that of lexical schwa [CəC] was 64.9 ms.

Experiments 6 and 7 further contribute to our understanding of the phonetic correlates of epenthetic as compared to lexical vowels.

4.5 Experiment 7 – Perceptual Epenthesis

H_2 predicts that Korean listeners are sensitive to vocalic information in perceiving English /s/, with particular focus on Korean listeners' /s/-cluster perception. /s/-cluster perception is argued to be influenced by native phonological knowledge. The native phonological knowledge that is of particular interest here is two-fold: (i) Korean uses vowel epenthesis in order to repair phonotactically ill-formed $/sC/_E$ clusters, and (ii) in Korean, post-lax consonant vowels are shorter in duration than post-tense consonant vowels. Given that the default epenthetic vowel is typically short (Steriade, 2001; Kenstowicz, 2003), it follows that the epenthetic vowel should be more similar to the vowel that appears after a lax fricative than the vowel that appears after a tense one. Korean listeners are assumed to employ their knowledge of the fine subphonemic phonetic properties of their language when they perform phonological epenthesis. To put it differently, perceptual vowel epenthesis is motivated by a phonological constraint, and the identity of the epenthetic vowel is phonetically determined based on the phonetic properties of vowels in Korean.

H_2 is tested in Experiment 7, which examines Korean listeners' perceptual strategies involving vowel epenthesis. Korean listeners are assumed to perceptually insert a vowel because of a phonotactic constraint against tautosyllabic consonant clusters (i.e., NOCOMPLEX). The epenthetic vowel should be minimally disruptive and the least salient in perception. Since the [ɨ] that appears after $[s]_K$ is shorter than the one that appears after $[s^*]_K$, it follows that the post-$[s]_K$ vowel will be less disruptive perceptually, and hence that this is the vowel that is expected as an epenthetic vowel. This reasoning postulates that the epenthetic [ɨ] should be more like the [ɨ] that typically appears after a lax consonant than like the [ɨ] that typically appears after a tense consonant. In addition, it was found that the *H*1-*H*2 values of the epenthetic vowels are more similar to the post-lax than the post-tense lexical vowels. Henceforth, '[ɨ] after a lax consonant' is marked as [ɨ̺]. While [ɨ] marks a lexical full vowel, [ɨ̺] marks a shorter, reduced epenthetic vowel.

The hypothesis explored in Experiment 7 is, therefore, that [ɨ̺] should be perceptually less salient than [ɨ] for Korean listeners, and therefore also perceptually less disruptive. That is, a word with an epenthetic [ɨ̺] should be perceptually more similar to a word without an epenthetic vowel than to a word with a full lexical vowel. The following representation formally expresses the perceptual difference.

Δ [stɑʊɹəm] ~ [si̯tɑʊɹəm] < Δ [stɑʊɹəm] ~ [sɨtɑʊɹəm]

It claims that the perceptual difference of [stɑʊɹəm] ~ [si̯tɑʊɹəm] is smaller, and therefore less salient, than that of [stɑʊɹəm] ~ [sɨtɑʊɹəm]. That is, [stɑʊɹəm] ~ [si̯tɑʊɹəm] are predicted to be more similar to each other than [stɑʊɹəm] ~ [sɨtɑʊɹəm] are to each other.

In Experiment 7, I test whether Korean listeners' perceptual modification of phonotactically ill-formed /sC/$_E$ clusters is processed to minimize perceptual disruption. The results of the experiment will inform us about Korean listeners' perceptual modification strategies in perceiving illegal English clusters, and about the loanword adaptation patterns of English /s/ cluster words.

4.5.1 Design and methodology

4.5.1.1 Participants

Twenty-five Korean-speaking listeners were recruited around the University of Michigan Ann Arbor campus. Similar to Experiments 3 and 4, three listener groups were recruited according to their residence in English-speaking countries: *Group A*_ADVANCED, *Group B*_INTERMEDIATE and *Group C*_NEW. No speaking or hearing disorder was reported, and participants were paid $15 for their participation.

Out of the 25 participants, one was excluded from data analysis, using the same principle with regard to incorrect response time as in the previous experiments. Each stimulus consisted of two consecutive /s/ containing words (e.g., [stɑʊɹəm] ~ [sɨtɑʊɹəm]), and the shortest duration from the beginning of the token to the end of the second [s] was 980 ms. Those responses with response times shorter than 980 ms necessarily were made before the entire second /s/ was heard, and the responses were therefore excluded. Table IV.6 illustrates % IRT of each listener. The % IRT of Subj-A9 in *Group A*_ADV was 18.6 %, and this participant's responses were therefore excluded from the data analysis.

Table IV.6 Incorrect Response Time (IRT) Rates of Listeners in Experiment 7

Listener group	Subject index	Raw number of errors	Total expected responses	% IRT
*Group A*_ADV	Subj-A1	2	840	0.24%
	Subj-A2	1	840	0.12%
	Subj-A3	6	840	0.71%
	Subj-A4	2	840	0.24%
	Subj-A5	2	840	0.24%
	Subj-A6	5	840	0.60%
	Subj-A7	3	840	0.36%
	Subj-A8	1	840	0.12%
	Subj-A9	156	840	18.57%
*Group B*_INT	Subj-B1	2	840	0.24%
	Subj-B2	19	840	2.26%
	Subj-B3	6	840	0.71%
	Subj-B4	3	840	0.36%
	Subj-B5	2	840	0.24%
	Subj-B6	18	840	2.14%
	Subj-B7	11	840	1.31%
	Subj-B8	7	840	0.83%
*Group C*_NEW	Subj-C1	26	840	3.10%
	Subj-C2	7	840	0.83%
	Subj-C3	3	840	0.36%
	Subj-C4	0	840	0.00%
	Subj-C5	4	840	0.48%
	Subj-C6	1	840	0.12%
	Subj-C7	28	840	3.33%
	Subj-C8	12	840	1.43%

4.5.1.2 Stimulus materials

A naïve female English-Korean bilingual speaker was recorded in a sound-attenuated room in the University of Michigan Phonetics Lab reading English non-words and Korean words. The English tokens were bisyllabic non-words of the form [$C_1C_2V_1$.CVC]. C_1 was /s/ in all instances while C_2 was one of /p, t, k, l, m, n/. V_1 was either [ɑɪ] or [ɑʊ] and the second syllable was either [lɛt] or [ɹəm]. The combinations of the consonants and vowels generated 24 words, given in Table IV.7. The speaker was asked to read these non-words as if they were English words (i.e., pronouncing /s/-clusters faithfully without vowel epenthesis). The speaker read the non-words several times and one ideal token for each word was selected.

Table IV.7 English non-words read

C_1	C_2	V_	Words
s	*p*	*ilet*	spilet [spɑɪlɛt], spaulet [spɑʊlɛt], spirem [spɑɪɹəm], spaurem [spɑʊɹəm],
	t	*aulet*	stilet [stɑɪlɛt], staulet [stɑʊlɛt], stirem [stɑɪɹəm], staurem [stɑʊɹəm],
	k	*irem*	skilet [skɑɪlɛt], skaulet [skɑʊlɛt], skirem [skɑɪɹəm], skaurem [skɑʊɹəm],
	l	*aurem*	slilet [slɑɪlɛt], slaulet [slɑʊlɛt], slirem [slɑɪɹəm], slaurem [slɑʊɹəm],
	m		smilet [smɑɪlɛt], smaulet [smɑʊlɛt], smirem [smɑɪɹəm], smaurem [smɑʊɹəm],
	n		snilet [snɑɪlɛt], snaulet [snɑʊlɛt], snirem [snɑɪɹəm], snaurem [snɑʊɹəm]

The same speaker also read a list of Korean words such as *'for oneself'* [sɨsɨɹo], *'teacher'* [sɨsɨŋ], *'write'* [s*ɨtɑ] and *'trash'* [s*ɨɹeki]. Among several utterances, one token of [sɨsɨɹo] and one token of [s*ɨtɑ] were selected on the basis of their similarity to the average [ɨ] durations after lax and tense consonants, respectively. Table IV.8 compares the durations of these stimuli and their *H*1-*H*2 values to corresponding values of the Korean vowel productions analyzed in Experiments 2 and 5.

Table IV.8 Acoustic Properties of [ɨ] and [ɨ̰]

		[ɨ]	[ɨ̰]
Experiment 5	Duration (ms)	82	58
Experiment 2	*H*1-*H*2 (dB)	-5.88	-0.51
Experiment 7	Duration (ms)	89	53
	*H*1-*H*2 (dB)	-8	-2

Using *Praat*, the first [ɨ̰] from [sɨ̰sɨ̰ro] and the [ɨ] from [s*ɨta] (i.e., the entire periodic portions between the voiceless consonants) were excised. These two vowels were then spliced into the English non-words between the initial [s] and the following consonant. This therefore resulted in stimuli such as [sɨpɑɪlɛt] vs. [sɨ̰pɑɪlɛt] and [sɨtɑʊɹəm] vs. [sɨ̰tɑʊɹəm]. The created stimuli were paired with the original, unmanipulated tokens to create 'different' pairs for an AX task, as in Table IV.9. The order of pair members was counterbalanced across repetitions. 'Same' pairs were identity pairings of an original or a manipulated word. The members of each stimuli pair were separated by 200 ms silence. Since each word generated 4 pairs and there were 24 words (i.e., 6 $C_2 \times 4$ V_1), 168 stimuli were created in total.

Table IV.9 Illustrative Same and Different Paired Stimuli for Experiment 7

	inserted-[ɨ] stimuli (TNS)	inserted-[ɨ̯] stimuli (LAX)
SAME	[staʊɹəm] ~ [staʊɹəm] [sɨtaʊɹəm] ~ [sɨtaʊɹəm]	[staʊɹəm] ~ [staʊɹəm] [sɨ̯taʊɹəm] ~ [sɨ̯taʊɹəm]
DIFF	[staʊɹəm] ~ [sɨtaʊɹəm] [sɨtaʊɹəm] ~ [staʊɹəm]	[staʊɹəm] ~ [sɨ̯taʊɹəm] [sɨ̯taʊɹəm] ~ [sɨtaʊɹəm]

4.5.1.3 Procedure

Stimulus presentation was controlled with *Superlab*, and stimuli pairs were presented in random order. Korean listeners were instructed to listen to the stimuli and decide whether the two consecutive words in each stimulus sounded the same or different. They registered their response by pressing one of two response box buttons marked as 'same' or 'different'. Listeners were told to press the button after the stimuli completed playing, and they were aware that their response times were being collected. There was one test session and it took 30 ~ 40 minutes.

Each listener heard 5 randomizations of the 168 AX pairings and was expected to generate 840 responses (i.e., 168 × 5 times); thus, 20160 responses across the 24 listeners, with the responses of Subj_A9 in *Group A*_ADV being excluded from the analysis. Responses with incorrect response time (i.e., 171 responses) from other participants were also excluded, so that 19989 responses were included in the statistical analysis.

4.5.2 Predictions

Members of the ‘different’ pairs have the same acoustic properties except for the inserted vowel portions. If listeners perceive inserted-[ɨ̯] stimuli to be more similar to the original than inserted-[ɨ] stimuli are to the original, accuracy should be higher in inserted-[ɨ] stimulus pairs. That is, the perceptual similarity model (and any model of perception) predicts that listeners will be better at discriminating [stɑʊɹəm] ~ [sɨtɑʊɹəm] pairs than [stɑʊɹəm] ~ [sɨ̯tɑʊɹəm] pairs because the acoustic difference between [stɑʊɹəm] and [sɨtɑʊɹəm] is larger than that between [stɑʊɹəm] and [sɨ̯tɑʊɹəm]. The perceptual similarity model will be challenged if listeners’ discrimination of inserted-[ɨ̯] pairs is higher than or similar to that of inserted-[ɨ] pairs.

Differences between listener groups may be found. As listeners become more proficient in English, they may become more tolerant of tautosyllabic consonant clusters, even though these clusters are banned in Korean. Therefore, more advanced listeners such in the *Group A*_ADV may be better able to differentiate tokens with a cluster from a token with an epenthetic vowel. Less experienced listeners, however, may be more likely to perceptually repair tautosyllabic clusters. Listeners in *Group C*_NEW may therefore be more likely to confuse tokens with clusters and tokens with an epenthetic vowel.

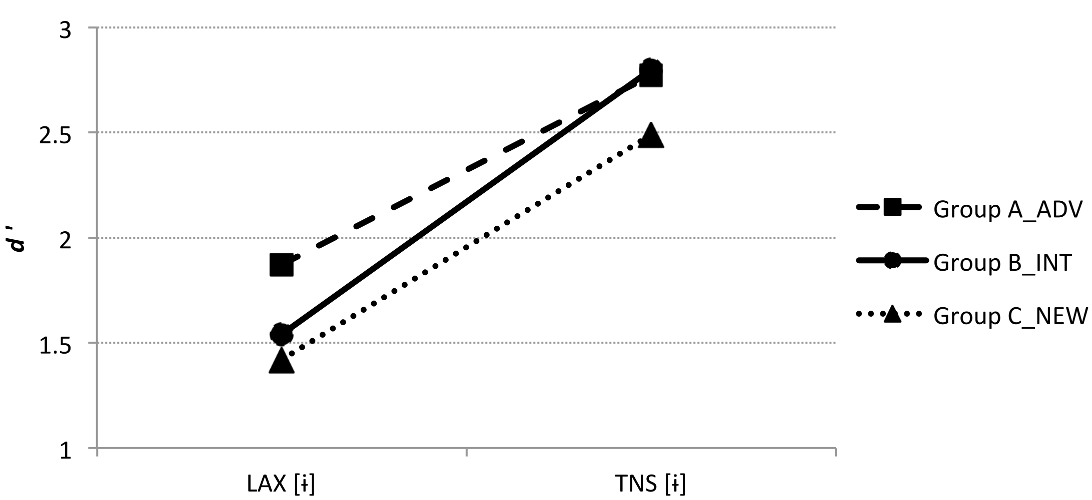
4.5.3 Results

Overall accuracies of all listeners are as follows: on “same” trials of the lax condition, accuracy was 97.6%, and on “different” trials of the lax condition, accuracy was 51.8%. On the other hand, in the tense condition, accuracy was 95.8% on “same” trials, and 86% on “different” trials.

The collected responses were transformed to *d’*-scores as in Signal Detection Theory (MacMillan and Creelman, 2005). Higher *d’*-scores correspond to higher discriminability; that is, the greater listeners’ sensitivity to stimulus differences, the higher the *d’*-scores. After *d’*-score transformation, a repeated measures ANOVA was conducted with group (i.e., *Group A*_ADV, *Group B*_INT, *Group C*_NEW), C_2 (i.e., /p, t, k, l, m, n/) and condition (i.e., tense and lax) as the independent variables and *d'*-score as a dependent variable.

The results of the ANOVA showed the main effect of condition (tense and lax) to be highly significant in both a by-item analysis (F (1, 142) = 428.771, $p < 0.001$) and a by-subject analysis (F (1, 46) = 34.204, $p < 0.001$), which means that listeners' discrimination of the tense condition was significantly more accurate than that of lax. In the by-item analysis, listener group showed a significant main effect (F (2, 142) = 15.482, $p < 0.001$) but not in the by-subject analysis (F (2, 46) = 1.235, $p = 0.301$). Tukey's HSD post-hoc pairwise comparison revealed that *Group C*_NEW group responded significantly less accurately than *Group A*_ADV ($p = 0.001$) and *Group B*_INT ($p < 0.001$). *Group B*_INT was also found to be significantly different from *Group A*_NEW ($p = 0.001$). Types of C_2 did not show any significant effect. Figure IV.6 gives *d'*-scores in lax inserted-[ɨ̥] and tense inserted-[ɨ] contexts, pooled across 8 listeners in each listener group. Across listener groups, the average *d'*-score in lax inserted-[ɨ̥] contexts was 1.6 and that in tense inserted-[ɨ] contexts was 2.7. The interaction between listener group and condition was significant (F (2, 142) = 4.179, $p = 0.017$) in a by-item analysis but not in a by-subject analysis.

Figure IV.7 *d'*-scores of 3 Listener Groups (averaged across 8 listeners per group) for the AX Discrimination Task



*Group A*_ADV had the highest *d'*-scores in both lax [ɨ̯]- and tense inserted-[ɨ] contexts, and the scores of *Group C*_NEW were the lowest in both contexts. In lax [ɨ̯] contexts, the scores of *Group B*_INT resembled that of *Group C*_NEW, while their scores resembled that of *Group A*_ADV in tense [ɨ] contexts.

The above patterns were representative of the overall responses of individual listeners within each group. In Table IV.10, a positive tense [ɨ] – lax [ɨ̯] value indicates that tense [ɨ] has the higher *d'*-score while a negative value indicates that tense [ɨ] has the lower *d'*-score. Most listeners showed positive tense [ɨ] – lax [ɨ̯] values, that is, they had higher *d'*-scores in tense inserted-[ɨ] contexts. Exceptionally, listener Subj-A6 in the *Group A*_ADV group showed a negative tense [ɨ] – lax [ɨ̯] value. Tense [ɨ] – lax [ɨ̯] differences also correspond to the steepness of the slope: the greater the difference, the steeper the discriminability incline. The discriminability incline was the highest in *Group B*_INT, whose tense [ɨ] – lax [ɨ̯] *d'* difference was 1.3. The value was the lowest for *Group A*_ADV, which was 0.9.

Table IV.10 *d′*-scores of Individual Listeners

Listener group	Subject index	lax [ɨ̯]	tense [ɨ]	tense [ɨ] – lax [ɨ̯]
*Group A*_ADV	Subj-A1	2.4	3.1	0.7
	Subj-A2	1.9	3.1	1.2
	Subj-A3	1.0	2.6	1.6
	Subj-A4	0.6	2.7	2.1
	Subj-A5	2.8	3.3	0.5
	Subj-A6	2.3	2.1	-0.2
	Subj-A7	0.9	2.1	1.2
	Subj-A8	3.1	3.2	0.1
*Group B*_INT	Subj-B1	2.4	3.1	0.7
	Subj-B2	1.0	2.9	1.9
	Subj-B3	1.6	2.6	1.0
	Subj-B4	1.7	2.7	0.9
	Subj-B5	2.6	3.2	0.6
	Subj-B6	1.2	2.6	1.3
	Subj-B7	1.7	2.9	1.2
	Subj-B8	0.1	2.4	2.3
*Group C*_NEW	Subj-C1	2.0	2.3	0.3
	Subj-C2	0.9	2.2	1.3
	Subj-C3	1.1	1.6	0.5
	Subj-C4	1.6	3.1	1.5
	Subj-C5	0.5	2.7	2.2
	Subj-C6	2.0	3.1	1.2
	Subj-C7	2.2	2.9	0.7
	Subj-C8	1.1	1.9	0.8
TOTAL		1.6	2.7	1.1

4.5.4 Discussion

The general response pattern in the discrimination task showed that listeners perceived the tense inserted-[ɨ] words (e.g., [stɑʊɹəm] ~ [sɨtɑʊɹəm]) as more different from the original than the lax inserted-[ɨ̯] words (e.g., [stɑʊɹəm] ~ [sɨ̯tɑʊɹəm]). This shows that insertion of [ɨ] is perceptually more disruptive than insertion of [ɨ̯]. Since smaller acoustic differences should, all else being equal, be harder to discriminate than larger differences. This effect is expected. More interesting thing is how poorly listeners did on the [ɨ̯] condition. The smaller difference is very difficult for listeners, and on average, they are at chance level.

In terms of the listener group effect, it can be said that the effect was found as expected. As listeners' exposure to English increased, their sensitivity increased in proportion to their exposure. *Group A*_ADV showed the highest *d'*-scores, and *Group C*_NEW showed the lowest. The scores of *Group B*_INT were in between.

4.6 General Discussion and Summary of the Chapter

In this chapter, I examined the temporal properties of vowels in various contexts. By revisiting the Korean production data from Experiment 2, it was found that post-tense vowels are generally longer than post-lax vowels. Of particular interest is the vowel [ɨ], which is adopted in vowel epenthesis, and also lexically present in Korean. The results of Experiment 5 showed that post-tense [ɨ] was longer than post-lax [ɨ]. Experiment 6 showed that epenthetic [ɨ] was even shorter than lexical [ɨ]. Based on the acoustic findings, I postulated that Korean listeners apply a modification strategy to the illegal English consonant cluster /sC/ by perceptually inserting a vowel between /s/ and /C/, and specifically inserting a vowel that is perceptually the least disruptive. By showing that [stɑʊɹəm] ~ [sɨtɑʊɹəm] pairs were more discriminable than [stɑʊɹəm] ~ [sɨ̯tɑʊɹəm] pairs by Korean listeners, it was concluded that the epenthetic [ɨ̯] is perceptually less disruptive than epenthetic [ɨ].

Since all the Korean listeners who participated in Experiment 7 speak English, it is likely that they know that English /sC/ clusters do not contain a vowel in between consonants. Nevertheless, due to the fact that the cluster is phonotactically illegal in Korean, the sequence is modified by their native language filter into /sVC/, which is phonotactically legal in Korean.

The duration of the epenthetic vowel is closer to that of the post-lax vowel than to the post-tense vowel, and the epenthetic vowel is arguably minimally disruptive. Consequently, post-lax vowels will be the better option than post-tense ones to adopt in vowel epenthesis. By adopting post-lax vowels that are perceptually the least disruptive vowel (e.g., [ɨ̥] rather than [ɨ]) as an epenthetic vowel, listeners take advantage of the fact that [ɨ̥] is perceptually less salient than [ɨ], and that the resulting form is hence more similar to the original form with the cluster.

Once the vowel is adopted as in /sɨ̥C/, the $/s/_E$ in this form is no longer perceptually preconsonantal; rather, it is processed as a prevocalic $/s/_E$. Due to the vocalic properties of the vowel [ɨ̥], the $/s/_E$ is perceived as a lax $[s]_K$, which is consistent with the results of Experiment 4.

In Korean, if one pronounces only the fricative /s/ or /s*/ continuously without adding any vowels, it is consistently perceived as a lax $[s]_K$ regardless of how long the frication is extended. Once the vowel is added, the syllable /sV/ is perceived as either $[sV]_K$ or $[s^*V]_K$, depending on the properties of the following vowel. In other words, when the listeners receive vocalic information, both $[s]_K$ and $[s^*]_K$ percepts are possible. However, if they do not receive vocalic information, only the $[s]_K$ percept is possible. These possible percepts are summarized in Table IV.11.

Table IV.11 Possible Percept Strategy for Korean Fricatives by Korean Listeners

	$[s]_K$	$[s^*]_K$
/sV/	✓	✓
/s/	✓	×

Alternatively, it can be argued that English /sC/ is borrowed with Korean lax /s/ because there is no following vowel, and whenever Korean listeners do not receive vocalic information, only the lax percept is possible. This can be argued against as follows: First, if Korean listeners index /s/ that is not followed by a vowel as a lax (i.e., /s_LC/, and always perceive it as a lax, preconsonantal /sC/$_E$ perception should not necessarily take longer than prevocalic /s/$_E$ perception. However, as illustrated in Figure III.7 in the previous chapter, /sC/$_E$ takes longer to process than /sV/$_E$, which argues against the only-lax-percept idea. Secondly, word-final /s/ as in '*gas*' is borrowed as tense in Korean. If Korean listeners marked /s/, which is not followed by a vowel, as a lax, word-final /s/ adaptation cannot be explained. Finally, the idea that lax /s/ is the default is not well motivated. Why a lax /s/ rather than a tense /s*/ should be default is hard to explain.

The same perceptual strategy applies in English perception. In principle, English prevocalic /s/$_E$ could be variously perceived (and borrowed) as [s]$_K$ and [s*]$_K$. However, due to the fact that the properties of the vowel that follows /s/ resemble post-tense vowels in Korean, the [s]$_K$ percept is blocked; thus, prevocalic /s/$_E$ is borrowed as [s*]$_K$. In preconsonantal /sC/$_E$, no vowel is acoustically present. However, such phonotactically ill-formed sequences in Korean are perceptually repaired via perceptual epenthesis by Korean listeners. Korean listeners thus apply vowel epenthesis, inserting the less salient [ɨ] which results in an [s]$_K$ percept.

Table IV.12 Possible Percept Strategy for English Fricatives by Korean Listeners

	[s]$_K$	[s*]$_K$
/sV/$_E$	(✓)	✓
/sCV/$_E$	✓	×

Perceptual vowel epenthesis is a well-known modification strategy in non-native speech perception (Berent *et al.*, 2007; Coetzee, 2010; Dupoux *et al.*, 1999; Kabak and Idsardi, 2007). These studies consistently show that listeners perceptually apply vowel epenthesis to resolve illegal sequences in their native language. Korean listeners' perceptual modification is of interest in that they selectively apply vowel epenthesis; they only insert a vowel that is closest to the lexical post-lax vowel. Put differently, Korean listeners' selective vowel epenthesis indicates that they perceive [sɨ̯] as a better choice than [sɨ]. Thus, Korean listeners may have a perceptual bias against [sɨ] in that [sɨ] is not allowed (i.e., *[sɨ]), and/or that [s*ɨ̯] is not allowed (i.e., *[s*ɨ̯]). For English /sC/$_{E}$ words, although both /sɨC/$_{K}$ and /sɨ̯C/$_{K}$ are possible candidates, because of the phonotactic constraint *[sɨ], /sɨ̯C/$_{K}$ would be considered as optimal.

In summary, Korean listeners incorporate phonetic and phonological knowledge in borrowing English /sC/ words into Korean. Native phonological knowledge (i.e., structural constraints) prevents illegal units being borrowed as they are and is the source of actual vowel epenthesis in loanwords and perceptual vowel epenthesis in the perception of English words with onset clusters. Perceptual knowledge encourages the epenthesis to be minimally disruptive. A vowel similar to a short post-lax consonant vowel is selectively epenthesized since that vowel is perceptually less disruptive than a post-tense vowel; thus, perceptual knowledge affects the selection of the epenthetic vowel. English /sC/ loans demonstrate that perceptual knowledge is incorporated into the selection of the perceptually epenthesized illusory vowel.

CHAPTER V
CONCLUSION

5.1 Introduction

The purpose of this dissertation was to investigate the perceptual principles that govern the process of loanword adaptation, and more generally to understand the principles that relate the cognitive representations formed by listeners to the acoustic reality of non-native speech.

In order to investigate these issues for English /s/-initial loanwords in Korean, this dissertation postulated two hypothesis, Hypothesis 1 (H_1), the consonantal information hypothesis, and Hypothesis 2 (H_2), the vocalic information hypothesis. While Hypothesis 1 predicted that /s/-internal properties (i.e., frication portion itself) played an important role in Korean listeners' English /s/ perception, Hypothesis 2 stated that Korean listeners predominantly rely on /s/-external properties, particularly properties of the following vowel. Hypothesis 2 was postulated to apply to vowels that are acoustically present as well as vowels that are perceptually epenthesized.

These two hypotheses were experimentally examined in four production and three perception experiments. The results of the experiments showed that (i) the perceptual role of consonant-internal information is limited, and that (ii) [sC] is difficult to discriminate from [sɨ̯C], which could explain why [s]$_K$ is the preferred result of borrowing [sC]. These results rejected the consonantal information hypothesis (H_1) and supported the vocalic information hypothesis (H_2). The detailed review of each hypothesis and experimental results are described below.

5.2 Review of Hypotheses and Summary of Results

The two hypotheses are re-stated below:

H_1 Consonantal Information Hypothesis

H_1a /s/ frication noise of $/sV/_E$ acoustically corresponds to that of $/s^*/_K$.
/s/ frication noise of $/sCV/_E$ acoustically corresponds to that of $/s/_K$.

H_1b Korean listeners are sensitive to consonantal information.

H_2 Vocalic Information Hypothesis

H_2a Vocalic properties of post-$/s/_E$ vowels acoustically correspond to those of post-$/s^*/_K$.
Vocalic properties of epenthetic $/sC/_E$ vowels perceptually correspond to those of post-$/s/_K$.

H_2b Korean listeners are sensitive to vocalic information.

H_1 was examined in Experiments 1, 2 and 3. Experiments 1 and 2 examined the temporal and spectral properties of /s/ frication noise. It was found that $/s/_E$ and $/s^*/_K$ have longer duration and higher CoG than their counterparts $/sC/_E$ and $/s/_K$, respectively. It was also found that both English $/s/_E$ and $/sC/_E$ were more similar to $/s^*/_K$ than to $/s/_K$ in terms of consonantal properties, which casts doubt on the hypothesis that the $/s/_E$ to $/s^*/_K$ and $/sC/_E$ to $/s/_K$ mappings are based on consonantal properties. Experiment 3 confirmed that Korean listeners were insensitive to these consonantal properties in their perception of the English fricative. Neither increasing duration nor raising CoG affected Korean listeners' perception of the identity of English $/s/_E$ and $/sC/_E$. Instead, Korean listeners consistently perceived (or interpreted the stimuli as) $/s^*/_K$ in prevocalic and $/s/_K$ in preconsonantal contexts. The results of Experiments 1, 2 and 3 suggested that, although the acoustic properties of /s/ frication differ according to phonological contexts, Korean listeners do not employ this information in their English perception: H_1 rejected.

H_2 was partly examined by revisiting the data collected for Experiments 1 and 2, and Experiment 4. In order to test H_2a, I examined two properties of post-fricative vowels in English and Korean. In English, the f_0 and $H1$-$H2$ values of post-/s/$_E$ and post-/sC/$_E$ were similar to each other. In Korean, while the f_0 values of post-/s/$_K$ and post-/s*/$_K$ were not different, $H1$-$H2$ values were significantly different between the two contexts. Experiment 4 explored the possibility of vocalic properties being involved by examining Korean listeners' /s*/-responses in [sV], [sCV] and [s_V] (i.e., deleted C) contexts. In [sCV] and [s_V] contexts, although the frication portions of /s/ were acoustically identical, listeners' /s*/-responses were significantly different, which indicated that listeners use phonological contexts. In [sV] and [s_V] contexts, Korean listeners' /s*/-responses were equally high, which suggested that Korean listeners are influenced by phonological contexts.

H_2 was further investigated through Experiments 2, 5, 6 and 7. Examination of Korean vowel duration in the productions from Experiment 2 showed that post-lax vowels were significantly shorter than post-tense vowels. Experiment 5 specifically examined the duration of Korean [ɨ], and found that, consistent with the results of Experiment 2, post-lax [ɨ̺] was significantly shorter than post-tense [ɨ]. Experiment 6 examined the duration of epenthetic [ɨ] in English loans, which confirmed that epenthetic [ɨ] was even shorter than lexical post-lax [ɨ̺]. Finally, Experiment 7 investigated Korean listeners' perceptual sensitivity to epenthetic [ɨ] and [ɨ̺]. They were better at discriminating post-tense [ɨ] epenthesis than post-lax [ɨ̺] epenthesis from stimuli with no epenthetic vowel (i.e., [sCV]), which showed that post-lax [ɨ̺] is perceptually less salient than post-tense [ɨ]. Given listeners' difficulty in discriminating [sCV] from [sɨ̺CV], it was concluded that Korean listeners experience a perceptual illusion because of phonotactic knowledge when presented with a phonotactically illegal sequence such as /sC/$_E$. Specifically, they perceive an illusory vowel between /s/ and /C/, and the identity of this vowel is chosen to be perceptually the least salient, i.e., [ɨ̺] rather than [ɨ]. The results of the series of experiments in Chapter IV led me to reject the abstract mapping hypothesis in favor of the contribution to listener's perception of an illusory epenthetic vowel: H_2 supported.

5.3 General Discussion

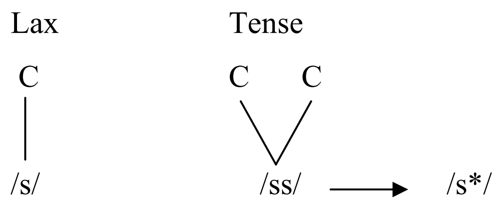
This dissertation investigated the principles governing cross-language speech perception and English loanword adaptation. I argued that English /s/ adaptation is a phonetically driven phonological phenomenon, and provided experimental accounts of borrowing patterns of prevocalic $/s/_E$ as tense $/s^*/_K$ and preconsonantal $/sC/_E$ as lax $/s/_K$. Across the different experiments, this dissertation established that the role of consonantal information is limited in English to Korean /s/ adaptation, and that Korean listeners exploit perceptual vowel epenthesis in their use of vocalic information.

5.3.1 Frication duration-sensitive view

In the English production experiment (Experiment 1), it was found that English prevocalic $/s/_E$ differed from preconsonantal $/sC/_E$ in duration and CoG, that is, consonantal information consistently varied depending on phonological context. Because of this difference in acoustic properties between prevocalic $/s/_E$ and preconsonantal $/sC/_E$, several researchers have claimed that Korean listeners rely on consonantal information in their adaptation of English /s/ into Korean. In particular, frication duration was claimed by Kim (1999), Kim and Curtis (2002) and Lee and Iverson (2007) to be responsible for English /s/ adaptation and, in fact, whether Korean listeners exploit frication duration has been a central issue in the study of /s/ adaptation.

Ahn and Iverson (2003) argued that, in Korean, a tense obstruent was a geminate version of a lax consonant, suggesting the phonological representation in Figure V.1.

Figure V.1 Phonological representation of /s/ and /s*/ (modified from Ahn and Iverson 2003:10)



They supported Kim (1999), claiming that "… unless Korean speakers are phonemically sensitive to differences in consonantal duration, it is a mystery why they should borrow English s sometimes as tense /ss/ (when phonetically long in English) and sometimes as lax /s/ (when phonetically short in English)" (Ahn and Iverson 2003:10).

Iverson and Lee (2006, 2007) pointed out that the durational differences of English /s/ were positionally conditioned, Korean listeners' English coda /s/ perception (e.g., *dance* and *desk*). In their experiment, the listeners perceived significantly more [s*] in long [_s] than in short [_sk] conditions (i.e., [s*] in a singleton and [s] in a cluster). They concluded that "Koreans perceive phonetically long English [s] as tense, phonologically geminate /ss/ but phonetically short English [s] as lax /s/" (Iverson and Lee, 2007:10), supporting the idea that English /s/ borrowing into Korean is dependent on frication duration.

These claims are inconsistent with the experimental findings of Experiments 3 and 4 in this dissertation. In Experiment 3, it was shown that, irrespective of duration, Korean listeners consistently perceived [s] before a consonant (i.e., a cluster [s]), and [s*] before a vowel (i.e., a singleton [s]). Experiment 4 showed that, even when the fricative noise is the same in [sCV] and [s_V], listeners perceived more [s*] in [s_V]. The two experiments strongly suggest that frication cues are not used in Korean listeners' /s/ perception, or minimally that the role of consonantal information is limited in loanword adaptation.

Moreover, despite their claims to the contrary, Iverson and Lee (2007)'s study does not necessarily show that Korean listeners rely on the duration of frication noise in their identification of English /s/. Korean listeners' /s*/ responses in their study can be interpreted in terms of phonological contexts. That is, Korean listeners may have mapped a singleton $[s]_E$ to $/s^*/_K$ and a cluster $[sC]_E$ to $/s/_K$ not because of duration but because of phonological rule application. Given that the stimuli of their study were word final, [_s] is no longer prevocalic but still a singleton. Therefore, it is likely that [_s] and [s_] pattern similarly in terms of both being a singleton.

5.3.2 Beyond word-initial /s/ and /s/-stop clusters

The discussion of this dissertation has focused on word-initial prevocalic $/s/_E$ and preconsonantal $/sC/_E$. This section considers how this study can be expanded with regard to /s/ in other contexts. The following table is repeated from Table I.1.

Table V.1 English /s/-words Adaptation Pattern in Korean

word-initial		word-final		
prevocalic (i.e., $_\#[s)$	preconsonantal (i.e., $_\#[sC)$	postvocalic (i.e., $s]_\#$)	postvocalic *and* preconsonantal (i.e., $sC]_\#$)	postconsonantal (i.e., $Cs]_\#$)
say	*stay*	*gas*	*guest*	*box*
s*	s	s*	s	s*

Word-final clusters with /s/ cases are either preconsonantal $/sC/_E$ (e.g., *'guest'*) or postconsonantal $/Cs/_E$ (e.g., *'box'*).[5] Word-final preconsonantal $/sC/_E$ can be explained in a similar way to word-initial preconsonantal $/sC/_E$, because they are also borrowed as $[s\text{ɨ̯}C]_K$ with vowel epenthesis. Similar to word-initial preconsonantal $/sC/_E$, by adding the epenthetic vowel, Korean listeners modify the cluster to be permissible but perceptually least different from the original cluster form. On the other hand, word-final postconsonantal $/Cs/_E$ works same as word-final singleton $/s/_E$, as discussed below.

Word-final postvocalic $/s/_E$ (i.e., a singleton), like word-initial prevocalic $/s/_E$, is borrowed as $[s^*]_K$. For instance, the English word *'gas'* is borrowed as [kɑs*ɨ]. Unlike word-initial prevocalic $/s/_E$, though, word-final singleton $/s/_E$ induces vowel epenthesis.

[5] The postconsonantal /s/ includes preconsonantal (e.g., *'text'*).

We have observed that vowel epenthesis applies when the cluster is a phonotactically ill-formed representation as in preconsonantal /sC/$_E$ adaptation (e.g., *'stay'* /steɪ/ → [si̯teɪ]). However, vowel epenthesis in word-final postvocalic /s/$_E$ results from a different restriction. Korean does not allow released consonants in coda positions in the surface representation. Korean listeners are sensitive to consonant release; thus, any consonant with release is perceptually parsed as an onset.[7] The consonant /s/ is necessarily released so that it is parsed as an onset followed by vowel epenthesis.

Now, although both word-initial preconsonantal /sC/$_E$ and word-final postvocalic /s/$_E$ undergo vowel epenthesis, they are borrowed differently: one with a lax consonant and the other with a tense consonant as in /sɨC/$_E$ → [sɨC] (e.g., *'stay'* /steɪ/ → [si̯tei]) and /sɨ/$_E$ → [s*ɨ]$_K$ (e.g., *'gas'* /gæs/ → [kɑs*ɨ]). It is likely that the distinct borrowing patterns result from the phonological contexts of /s/. Unlike preconsonantal /sC/$_E$, postvocalic /s/$_E$ is word-final, a position in which a segment undergoes final lengthening. Presumably, then, the epenthetic vowel is phonetically longer in *'gas'* than in *'stay'*. It is possible that the lengthened vowel [ɨ] causes the preceding /s/$_E$ to be perceived as a tense [s*]$_K$. Word-final postconsonantal /Cs/$_E$ (e.g., *'box'* [pɑks*ɨ]) receives the same explanation. Verifying this speculative account is left to future study.

The accounts suggested in this dissertation can be applied to various word-initial (e.g., *'slow'* [sɨlou] and *'smog'* [sɨmogɨ]) and word-final consonant clusters (e.g., *'dance'* [tæns*ɨ]).

---

[7] This explanation applies to loanwords only. In native Korean words, word-final /s/ is neutralized to [t] as in /kos/ → [kot] *'place'*. The underlying /s/ emerges only when it is followed by a vowel as in /kos/ + /e/ → [kose] *'place* + locative *to'*.

5.3.3 Comparison with results of other studies

There are at least two previous accounts of English to Korean /s/ adaptation: Kang and Kang (2004) and Kang (2008). In the following, I address these studies and compare how they differ from the accounts of this dissertation.

Kang and Kang (2004) examined the perception of English /s/ by Korean listeners. To this end, they synthesized several English /s/ words (e.g., *'sale'*), and spliced the frication portion onto a different word to make a /s/ cluster word. For instance, they spliced the /s/ of *'sale'* onto *'mile'* to create *'smile'*. When these cross-spliced words were presented to Korean listeners, they perceived the frication noise as $[s^*]_K$ in *'sale'* but $[s]_K$ in *'smile'*. Kang and Kang argued that English prevocalic /s/ was marked by high f_0 on the following vowel and that Korean listeners were sensitive to the cue. According to their account, [s] in *'sale'* was marked by high f_0, but when it is followed by a consonant as in *'smile'*, it was no longer so.

As shown in Experiment 1 in Chapter II, however, f_0 values were not found to be significantly different between $[s]_K$ and $[s^*]_K$ in Korean. It would be surprising if Korean listeners exploited f_0 information in identifying English $/s/_E$ and $/sC/_E$ when f_0 does not systematically distinguish $/s/_K$ and $/s^*/_K$ in their native language. However, their results are in line with the results of Experiment 4 in that the same [s] portions were perceived differently according to their phonological context, pointing to the limited role of frication noise.

Kang (2008) suggested that $H1$-$H2$ is responsible for prevocalic $/s/_E$ adaptation and that, when this information is not available, as is the case in preconsonantal $/sC/_E$, Koreans insert a devoiced vowel. According to Jun and Beckman (1994), vowel devoicing commonly occurs after aspirated stops and lax fricatives in Korean. Kang, thus, suggests that Koreans insert the vowel [ɨ] in order to resolve consonant clusters and devoice the vowel (e.g., *'stop'* /stɑp/ → [sɨtɑp] → [sɨ̥tɑp]). Although Kang's approach to the phonetic nature of the epenthetic vowel differs from this dissertation, it is similar to this study in that a non-salient segment is adopted in vowel epenthesis.

5.4 Future Directions

In this dissertation, I examined how non-native sounds are perceived and borrowed into another language. It was suggested that phonetics and phonology interact in the perception process in which listeners attend to detailed phonetic information in the input while simultaneously processing the input in terms of their native phonetic and phonological structures, exploiting both to determine the identity of a perceptually illusory vowel. This dissertation is consistent with previous findings that nonnative contrasts are severely and systematically affected by native phonotactics (Flege, 1989; Hallé *et al.*, 1998; Hallé *et al.*, 2003), phonetic variation (Flege and Hillenbrand, 1987), coarticulatory information (Beddor *et al.*, 2002), phonetic detail (Allen and Miller, 2001) and listeners' native language (Best and Strange, 1992).

A related study that I intend to pursue is native English listeners' perception of the Korean $/s/_K$ vs. $/s^*/_K$ contrast. In this dissertation, it was shown in that the two Korean phonemes have different acoustic characteristics, which is crucial to Korean listeners. These acoustic differences may not be meaningful to English listeners because the difference is not relevant in their language, or they are meaningful because there are systematic, subphonemic differences in production. Following the perceptual assimilation model (PAM), English listeners may assimilate both $/s/_K$ and $/s^*/_K$ phonemes to their one phonological category $/s/_E$. However, it may be the case that English listeners are able to perceive the phonetic differences. English listeners' perception of Korean /s/ and /s*/ contrast is likely an interesting area for a future study.

This dissertation has implications for Best's (1995) PAM. As discussed in Chapter I, PAM explains various patterns of how non-native sounds are assimilated into native sound categories. Naïve non-native listeners do not know whether two systematically different phones contrast in L1 or not. Thus, $[sV]_E$ and $[sC]_E$ borrowed as Korean /s/ and /s*/ could, at least for naïve listeners, be viewed as two-category assimilation (Best and Tyler, 2007).

APPENDIX

1. Experiment 1:English /s/ Production Wordlists

/sV/	1	*sack*	11	*sand*	21	*seam*	31	*sell*
	2	*sad*	12	*sash*	22	*seat*	32	*sense*
	3	*safe*	13	*sass*	23	*sect*	33	*set*
	4	*sag*	14	*sauce*	24	*seed*	34	*sick*
	5	*sage*	15	*save*	25	*seek*	35	*side*
	6	*sail*	16	*saw*	26	*seem*	36	*siege*
	7	*sake*	17	*say*	27	*seep*	37	*sigh*
	8	*sale*	18	*scene*	28	*seethe*	38	*sight*
	9	*salt*	19	*sea*	29	*seize*	39	*sign*
	10	*same*	20	*seal*	30	*self*	40	*sill*
	41	*sin*	51	*soak*	61	*song*		
	42	*since*	52	*soap*	62	*soon*	71	*suck*
	43	*sing*	53	*soar*	63	*soothe*	72	*sue*
	44	*sink*	54	*sob*	64	*sore*	73	*suit*
	45	*sip*	55	*sock*	65	*soul*	74	*sulk*
	46	*sir*	56	*sod*	66	*soup*	75	*sum*
	47	*sit*	57	*soil*	67	*sour*	76	*sun*
	48	*site*	58	*sole*	68	*south*		
	49	*six*	59	*some*	69	*sow*		
	50	*size*	60	*son*	70	*such*		

/sCV/	/swV/		/slV/		/smV/		/snV/	
	77	*swan*	88	*slab*	109	*smack*	120	*snack*
	78	*sway*	89	*slack*	110	*small*	121	*snail*
	79	*swear*	90	*slang*	111	*smart*	122	*snake*
	80	*sweat*	91	*slap*	112	*smash*	123	*snap*
	81	*sweep*	92	*slash*	113	*smear*	124	*snarl*
	82	*sweet*	93	*slave*	114	*smell*	125	*snatch*
	83	*swell*	94	*slay*	115	*smile*	126	*sneak*
	84	*swim*	95	*sleep*	116	*smite*	127	*sneer*
	85	*swing*	96	*sleeve*	117	*smith*	128	*sneeze*
	86	*Swiss*	97	*slice*	118	*smoke*	129	*sniff*
	87	*switch*	98	*slick*	119	*smooth*	130	*snore*
			99	*slight*			131	*snout*
			100	*slim*			132	*snow*
			101	*sling*				
			102	*slip*				
			103	*slope*				
			104	*slot*				
			105	*slow*				
			106	*slug*				
			107	*slum*				
			108	*sly*				
	/skV/		/spV/		/stV/			
	133	*scale*	156	*space*	184	*stab*	212	*stone*
	134	*scalp*	157	*spade*	185	*stack*	213	*stool*
	135	*scan*	158	*Spain*	186	*staff*	214	*stoop*
	136	*scar*	159	*span*	187	*stag*	215	*stop*
	137	*scheme*	160	*spare*	188	*stage*	216	*stout*
	138	*school*	161	*spate*	189	*stain*	217	*stove*
	139	*scoff*	162	*speak*	190	*stake*	218	*stow*
	140	*scoop*	163	*spear*	191	*stale*	219	*stud*
	141	*scope*	164	*speck*	192	*stall*	220	*stuff*
	142	*scorn*	165	*speech*	193	*star*	221	*stun*
	143	*scot*	166	*speed*	194	*state*		
	144	*scout*	167	*spell*	195	*stay*		
	145	*scum*	168	*spice*	196	*stead*		
	146	*skate*	169	*spike*	197	*steak*		
	147	*sketch*	170	*spill*	198	*steal*		
	148	*ski*	171	*spin*	199	*steam*		
	149	*skid*	172	*spine*	200	*steel*		
	150	*skill*	173	*spire*	201	*steer*		
	151	*skim*	174	*spit*	202	*stem*		
	152	*skin*	175	*spite*	203	*step*		
	153	*skip*	176	*spoil*	204	*Steve*		
	154	*skull*	177	*spool*	205	*stew*		
	155	*sky*	178	*spoon*	206	*stick*		
			179	*sport*	207	*stiff*		
			180	*spot*	208	*still*		
			181	*spouse*	209	*sting*		
			182	*spur*	210	*stock*		
			183	*spy*	211	*stoke*		

	/skCV/		/spCV/		/stCV/	
/sCCV/	222	*scrap*	236	*splash*	254	*strafe*
	223	*scrape*	237	*splat*	255	*strain*
	224	*scratch*	238	*splay*	256	*strand*
	225	*scream*	239	*spleen*	257	*strap*
	226	*screed*	240	*splice*	258	*straw*
	227	*screen*	241	*split*	259	*stray*
	228	*screw*	242	*sprain*	260	*streak*
	229	*scribe*	243	*sprat*	261	*stream*
	230	*scroll*	244	*spray*	262	*street*
	231	*scrawl*	245	*spread*	263	*stress*
	232	*scrub*	246	*spree*	264	*strew*
	233	*scruff*	247	*sprig*	265	*stride*
	234	*scrum*	248	*spring*	266	*strike*
	235	*scrip*	249	*sprit*	267	*string*
			250	*sprite*	268	*strip*
			251	*sprout*	269	*strive*
			252	*spruce*	270	*stroke*
			253	*sprung*	271	*stroll*
					272	*strong*
					273	*strop*
					274	*strum*
					275	*strung*
					276	*strut*

2. Experiment 1:English /s/ Production Wordlists (words used for CoG analysis)

/sV/	1	*salt*	11	*size*	21	*son*	31	*such*
	2	*sand*	12	*soak*	22	*song*	32	*suck*
	3	*sauce*	13	*soap*	23	*soon*	33	*sue*
	4	*saw*	14	*soar*	24	*soothe*	34	*suit*
	5	*side*	15	*sob*	25	*sore*	35	*sulk*
	6	*sigh*	16	*sock*	26	*soul*	36	*sum*
	7	*sight*	17	*sod*	27	*soup*	37	*sun*
	8	*sign*	18	*soil*	28	*sour*		
	9	*sir*	19	*sole*	29	*south*		
	10	*site*	20	*some*	30	*sow*		
/sCV/	**/swV/**		**/slV/**		**/smV/**		**/snV/**	
	38	*swan*	39	*slight*	46	*smack*	53	*snarl*
			40	*slope*	47	*small*	54	*snore*
			41	*slot*	48	*smart*	55	*snout*
			42	*slow*	49	*smile*	56	*snow*
			43	*slug*	50	*smite*		
			44	*slum*	51	*smoke*		
			45	*sly*	52	*smooth*		
	/skV/		**/spV/**		**/stV/**			
	57	*scalp*	69	*spice*	81	*stall*		
	58	*scar*	70	*spike*	82	*star*		
	59	*school*	71	*spire*	83	*state*		
	60	*scoff*	72	*spite*	84	*stew*		
	61	*scoop*	73	*spoil*	85	*stock*		
	62	*scope*	74	*spool*	86	*stoke*		
	63	*scorn*	75	*spoon*	87	*stone*		
	64	*scot*	76	*sport*	88	*stool*		
	65	*scout*	77	*spot*	89	*stoop*		
	66	*scum*	78	*spouse*	90	*stop*		
	67	*skull*	79	*spur*	91	*stout*		
	68	*sky*	80	*spy*	92	*stove*		
					93	*stow*		
					94	*stud*		
					95	*stuff*		
					96	*stun*		
/sCCV/	**/skCV/**		**/spCV/**		**/stCV/**			
	97	*screw*	104	*splice*	109	*straw*		
	98	*scribe*	105	*sprite*	110	*strew*		
	99	*scroll*	106	*sprout*	111	*stride*		
	100	*scrawl*	107	*spruce*	112	*strike*		
	101	*scrub*	108	*sprung*	113	*strive*		
	102	*scruff*			114	*stroke*		
	103	*scrum*			115	*stroll*		
					116	*strong*		
					117	*strop*		
					118	*strum*		
					119	*strung*		
					120	*strut*		

3. Experiment 2:Korean /s/ and /s*/ Production Wordlists

	token	glossary	token	glossary
[_a]	[sacin]	*'picture'*	[s*ata]	*'to wrap'*
	[sata]	*'to buy'*	[s*aum]	*'fight'*
	[samc^{h}on]	*'accident'*	[s*aknun]	*'sprout'*
	[salta]	*'to live'*	[s*amci]	*'pocket'*
[_o]	[somaŋ]	*'hope'*	[s*ota]	*'to shoot'*
	[sokta]	*'to be deceived'*	[s*onsal]	*'rapidity'*
	[sokæ]	*'introduction'*	[s*olrim]	*'being leaned'*
	[solcik]	*'honesty'*	[s*oim]	*'being stung '*
[_u]	[surak]	*'approval'*	[s*usip]	*'poking'*
	[sukmjʌŋ]	*'doom'*	[s*ukkas]	*'a crown daisy'*
	[sujʌŋ]	*'swimming'*	[s*uta]	*'to boil'*
	[sunsu]	*'innocence'*	[s*ukt'ɨm]	*'wormwood therapy'*
[_ʌ]	[sʌta]	*'to stand'*	[s*ʌlta]	*'to cut'*
	[sʌŋkjʌk]	*'personality'*	[s*ʌre]	*'a harrow'*
	[sʌul]	*'Seoul'*	[s*ʌkta]	*'to rot'*
	[sʌŋɨi]	*'effort'*	[s*ʌcim]	*'being written'*

4. Experiment 5: Korean Non-words Production Wordlists

consonant	vowel	/s/$_K$		/s*/$_K$	
		IPA	Korean Orthography	IPA	Korean Orthography
p	ɑ	/sɨ.pɑk/	스박	/s*ɨ.pɑk/	쓰박
	i	/sɨ.pik/	스빅	/s*ɨ.pik/	쓰빅
	ʌ	/sɨ.pʌk/	스벅	/s*ɨ.pʌk/	쓰벅
	o	/sɨ.pok/	스복	/s*ɨ.pok/	쓰복
	u	/sɨ.puk/	스북	/s*ɨ.puk/	쓰북
t	ɑ	/sɨ.tɑk/	스닥	/s*ɨ.tɑk/	쓰닥
	i	/sɨ.tik/	스딕	/s*ɨ.tik/	쓰딕
	ʌ	/sɨ.tʌk/	스덕	/s*ɨ.tʌk/	쓰덕
	o	/sɨ.tok/	스독	/s*ɨ.tok/	쓰독
	u	/sɨ.tuk/	스둑	/s*ɨ.tuk/	쓰둑
k	ɑ	/sɨ.kɑk/	스각	/s*ɨ.kɑk/	쓰각
	i	/sɨ.kik/	스긱	/s*ɨ.kik/	쓰긱
	ʌ	/sɨ.kʌk/	스걱	/s*ɨ.kʌk/	쓰걱
	o	/sɨ.kok/	스곡	/s*ɨ.kok/	쓰곡
	u	/sɨ.kuk/	스국	/s*ɨ.kuk/	쓰국
m	ɑ	/sɨ.mɑk/	스막	/s*ɨ.mɑk/	쓰막
	i	/sɨ.mik/	스믹	/s*ɨ.mik/	쓰믹
	ʌ	/sɨ.mʌk/	스먹	/s*ɨ.mʌk/	쓰먹
	o	/sɨ.mok/	스목	/s*ɨ.mok/	쓰목
	u	/sɨ.muk/	스묵	/s*ɨ.muk/	쓰묵

n	ɑ	/sɨ.nɑk/	스낙	/s*ɨ.nɑk/	쓰낙
	i	/sɨ.nik/	스닉	/s*ɨ.nik/	쓰닉
	ʌ	/sɨ.nʌk/	스넉	/s*ɨ.nʌk/	쓰넉
	o	/sɨ.nok/	스녹	/s*ɨ.nok/	쓰녹
	u	/sɨ.nuk/	스눅	/s*ɨ.nuk/	쓰눅
r	ɑ	/sɨ.rɑk/	스락	/s*ɨ.rɑk/	쓰락
	i	/sɨ.rik/	스릭	/s*ɨ.rik/	쓰릭
	ʌ	/sɨ.rʌk/	스럭	/s*ɨ.rʌk/	쓰럭
	o	/sɨ.rok/	스록	/s*ɨ.rok/	쓰록
	u	/sɨ.ruk/	스룩	/s*ɨ.ruk/	쓰룩
ʧ	ɑ	/sɨ.ʧɑk/	스착	/s*ɨ.ʧɑk/	쓰착
	i	/sɨ.ʧik/	스칙	/s*ɨ.ʧik/	쓰칙
	ʌ	/sɨ.ʧʌk/	스척	/s*ɨ.ʧʌk/	쓰척
	o	/sɨ.ʧok/	스촉	/s*ɨ.ʧok/	쓰촉
	u	/sɨ.ʧuk/	스축	/s*ɨ.ʧuk/	쓰축
s	ɑ	/sɨ.sɑk/	스삭	/s*ɨ.sɑk/	쓰삭
	i	/sɨ.sik/	스식	/s*ɨ.sik/	쓰식
	ʌ	/sɨ.sʌk/	스석	/s*ɨ.sʌk/	쓰석
	o	/sɨ.sok/	스속	/s*ɨ.sok/	쓰속
	u	/sɨ.suk/	스숙	/s*ɨ.suk/	쓰숙
h	ɑ	/sɨ.hɑk/	스학	/s*ɨ.hɑk/	쓰학
	i	/sɨ.hik/	스힉	/s*ɨ.hik/	쓰힉
	ʌ	/sɨ.hʌk/	스헉	/s*ɨ.hʌk/	쓰헉
	o	/sɨ.hok/	스혹	/s*ɨ.hok/	쓰혹
	u	/sɨ.huk/	스훅	/s*ɨ.huk/	쓰훅

REFERENCES

Abaglo, P. and Archangeli, D. 1989. Language-particular underspecification: Gengbe /e/ and Yoruba /i/. *Linguistic Inquiry* 20. 457-480.

Ahn, H. 1999. An acoustic investigation of the phonation types of post-fricative vowels in Korean. *Harvard Studies in Korean Linguistics* 7:57–71.

Ahn, Sang-Cheol and Iverson, Gregory K. 2004. Dimensions in Korean laryngeal phonology. *Journal of East Asian Linguistics* 13. 345-379.

Ali, Ahmed M. A., Jan Van d. Spiegel and Paul Mueller. 2001. "Acoustic-phonetic features for the automatic classification of fricatives" *Journal of the Acoustical Society of America* 109, 5(1) 2217-2235.

Allen and Miller. 2001. Contextual influences on the internal structure of phonetic categories: A distinction between lexical status and speaking rate. *Perception and Psychophysics.* 63. 798-810.

Barnwell. T. P. 1971. Relation of Syllable Boundaries to Vowel Durations. *Journal of the Acoustical Society of America* 50. 117.

Benguerel, André-Pierre. 2002. Duration of French Vowels in Unemphatic Stress. *Language and Speech.* 45:141-178.

Beddor, Patrice S., Harnsberger, J. D. and Lindemann, Stephanie. 2002. Language-specific patterns of vowel-to-vowel coarticulation: acoustic structures and their perceptual correlates. *Journal of Phonetics.* 30, 591-627.

Berent, I., Steriade, D., Lennertz, T. and Vaknin, V. 2007. What we know about what we have never heard: Evidence from perceptual illusions. *Cognition* 104. 591-630.

Best, Catherine. T. 1995. A direct realist view of cross-language speech perception. In W. Strange (ed.) *Speech Perception and Linguistic Experience: Issues in cross-language research*. 171-204. New York.

Best, Catherine, McRonerts, G. W. and Sithole, N. N. 1988. The phonological basis of perceptual loss for non-native contrasts: maintenance of discrimination among Zulu clicks by English-speaking adults and infants. *Journal of Experimental Psychology: Human Perception and Performance* 14. 345-360.

Best and Strange. 1992. Effects of phonological and phonetic factors on cross-language perception of approximants. *Journal of Phonetics* 20. 305-330.

Best and Tyler, 2007. Nonnative and second-language speech perception: Commonalities and complementarities. In: Bohn OS, Munro M, editors. *Second-language Speech Learning: The Role of Language Experience in Speech Perception and Production. A Festschrift in Honour of James E. Flege.* John Benjamins; Amsterdam: 2007. 13–34.

Blankenship, Barbara. 1997. *The Time Course of Breathiness and Laryngealization in Vowels.* Doctoral dissertation, University of California at Los Angeles.

Bochner, Joseph H., Karen B. Snall and Douglas J. MacKenzie. 1988. Duration and discrimination of speech and tonal complex stimuli by normally hearing and hearing-impaired listeners. *Journal of Acoustical Society of America* 84(2). 493-500.

Boersma, Paul and David Weenink. 2009. Praat: doing phonetics by computer (Version 5.1.05) [Computer program], http://www.praat.org/

Bosch, Anna and Kenneth J. de Jong. 1997. The prosody of Barra Gaelic epenthetic vowels. *Studies in the Linguistic Sciences* 27(1): 2-15.

Braunschweiler, Norbert. 1997. Integrated Cues of Voicing and Vowel Length in German: A Production Study. *Language and Speech* 40: 353-376.

Byrd, Dani, Abigail Kaun, Shrikanth Narayanan and Eliot Saltzman. 2000. Phrasal signatures in articulation. *Acquisition and the lexicon* (Papers in Laboratory Phonology 5). In Michael B. Broe and Janet B. Pierrehumbert (eds.). 70–87. Cambridge: Cambridge University Press.

Chang, Charles B. 2007. Korean fricatives: Production, perception and laryngeal typology. University of California, Berkeley, Phonology Lab Annual Report. 20-70.

Cheon, Sang Yee and Anderson, Victoria B. 2008. Acoustic and perceptual similarities between English and Korean sibilants: Implications for second language acquisition. *Korean Linguistics,* 14: 41-64.

Choi, H. 2002. Acoustic cues for the Korean stop contrast. ZAS Papers in Linguistics 28, 1-12.

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

Cooper, Franklin S., Pierre C. Delattre, Alvin M. Liberman, John M. Borst and Louis J. Gerstman. 1952. Some experiments on the perception of synthetic speech sounds. *Journal of the Acoustical Society of America* 24: 597-606.

Coetzee, Andries W. 2010. Syllables in speech processing: evidence from perceptual epenthesis. In Chuck Cairns and Eric Raimy, eds. *Handbook of the Syllable.* Leiden: E.J. Brill. 295-328.

Davidson, Lisa. 2006. Phonology, phonetics, or frequency: Influences on the production of non-native sequences. *Journal of Phonetics*, 34:1, 104-137.

Davidson, Lisa. 2007. Relating perception of non-native phonotactics to loanwords. *Phonology* 24. 261-286. Cambridge University Press.

Delattre, P. C., Liberman, A. M. and F. S. Cooper. 1955. Acoustic loci and transitional cues for consonants. *Journal of the Acoustical Society of America* 27.769-773.

Diehl, Randy L., Björn Lindbloom, Kathryn A. Hoemeke and Richard P. Fahey. 1996. On explaining certain male-female differences in the phonetic realization of vowel categories. *Journal of Phonetics* 24, 187-208.

D'Imperio, Maiapaola and Sam Rosenthall. 1999. Phonetics and phonology of main stress in Italian. *Phonology* 16.1-28. Cambridge University Press.

Dupoux, Emmanuel, Kazuhiko Kakehi, Yiki Hirose, Christophe Pallier and Jacques Mehler. 1999. Epenthetic vowels in Japanese: A perceptual illusion? *Journal of Experimental Psychologoy: Human Perception and Performance* 25, 1568-1578.

Evers, Vincent, Henning Reetz and Aditi Lahiri. 1998. Crosslinguistic acoustic categorization of sibilants independent of phonological status. *Journal of Phonetics* 26, 345-370.

Fisher-Jørgnesen, E. 1967. Phonetic analysis of breathy (murmured) vowels in Gujarati. *Indian Linguistics* 28:71-139.

Flege, J. E. 1989. Differences in inventory size affects the location but not the precision of tongue positioning in vowel production, *Language and Speech* 32, 2:123-147.

Flege, J. E. 1995. Second language speech learning: theory, findings, and problems. In *Speech perception and linguistic experience.* W. Strange (ed) 233–277. Timonium, MD: York Press.

Flege J. E., and Hillenbrand J. 1987. Limits on phonetic accuracy in foreign language production. In *Interlanguage Phonology: the Acquisition of a Second Language Sound System*, edited by Ioup G. and Weinberger S. Newbury House, Cambridge. 176-201.

Gordon, Mattew, Paul Barthmaier and Kathy Sands. 2002. A Cross-linguistic acoustic study of fricatives. *Journal of the International Phonetic Association* 32. 141-174.

Gordon, Mattew and Ayla Applebaum. 2006. Phonetic structures of Turkish Kabardian. *Journal of the International Phonetic Association* 36(2). 159-186.

Gouskova, Maria and Nancy Hall. 2009. Acoustics of epenthetic vowels in Lebanese Arabic. In Steve Parker (ed.) *Phonological argumentation: Essays on evidence and motivation.* 203-225. London:Equinox.

Haggard, M. P. 1971. Effects of clusters on segment durations. *Speech Synthesis and Perception* 5, 1-50. Cambridge, England: Psychological Laboratory.

Hall, T. A., Silke Hamann and Marzena Zygis. 2006. “The phonetic motivation for phonological stop assibilation. *Journal of the International Phonetic Association.* 36(1). 59-81.

Hallé, P., Segui, J., Frauenfelder, U. and Meunier, C. 1998. Processing of illegal consonant clusters: A case of perceptual assimilation? *Journal of*

Experimental Psychology: Human Perception and Performance, 24(2), 592–608.

Hallé, P., Best, C. T. and Bachrach. A. 2003. Perception of /dl/ and /tl/ clusters: A coss-linguistic perceptual study with French and Israeli listeners. *Proceedings of the 15th International Congress of Phonetic Sciences*. Barcelona, Spain. August 3-9, 2003. Universitat Auto`noma de Barcelona.

Hanson, Helen. 1997. Glottal characteristics of female speakers: Acoustic correlates. *Journal of the Acoustical Society of America* 101, 466-481.

Henton, Caroline G. and R. Anthony W. Bladon. 1985. Breathiness in normal female speech: inefficiency versus desirability. *Language Communication* 5, 221-227.

Jessen, Michael. 1998. *Phonetics and Phonology of Tense and Lax Obstruents in German*. In František Daneš, Jacek Fisiak, Leonid L. Iomdin, Philip Luelsdorff, Jim Miller, Olga Tomić, Charles Townsend, Josef Vachek and Olga Yokoyama (eds.) Studies in Functional and Structural Linguistics (SFSL) scholarly series 44. John Benjamins Publishing Company, Amsterdam and Philadelphia.

Johnson, Kieth. 1997. The auditory/perceptual bias for speech segmentation. Phio State University Working Papers in Linguistics 50: 101-113.

Jongman, Allard. 1989. Duration of frication required for identification of English fricatives. *Journal of the Acoustical Society of America* 85.1718-1725.

Jongman, Allard, Ratree Wayland and Serena Wong. 2000. Acoustic characteristics of English fricatives. *Journal of the Acoustical Society of America.* 108(3), 1252-1263.

Kabak, Bariş and William J. Idsardi. 2007. Perceptual distortions in the adaptation of English consonant clusters: Syllable structure or consonantal contact constraints? *Language and Speech* 50, 23-52.

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

Kang, H. and Kang, S.-K. 2004. Processing English [s] into Korean: prosodic influence on sound change. Poster paper presented at Ninth Conference Laboratory Phonology, University of Illinois, Urbana, IL.

Kang, Yoonjung. 2008. The adaptation of English /s/ in Korean. *Inquiries into Korean Linguistics III,* 1-14.

Keating, Patricia, T. Cho, Fougeron and C-S. Hsu. 1998. Domain-initial articulatory strengthening in four languages. In: J. Local, R.A. Ogden and R. Temple, Editors, *Papers in laboratory phonology VI: Phonetic interpretation*, Cambridge University Press, Cambridge. 143-161.

Kenstowicz, Michael. *in press*. Salience and similarity in loanword adaptation: case study from Fijian. *Language Sciences*.

Kim, Soohee. 1999. *Sub-phonemic Duration Difference in English /s/ and Few-To-Many Borrowing from English to Korean.* Ph.D. dissertation. University of Washington.

Kim, Soohee and Curtis, E. 2002. Phonetic duration of English /s/ and its borrowing into Korean. *Japanese/Korean Linguistics* 10, 406-419.

Kingston, John. 1986. Are f_0 differences after stops accidental or deliberate? *Journal of the Acoustical Society of America* Suppl. 1, 79, S27.

Kingston, John. 2005. Ears to categories: New argument for autonomy, S. Frota, M. Vigario and M. J. Freitas (Eds.) *Proceedings of the First Conference on Phonetics and Phonology in Iberia* (PaPI), 177-222. Berlin: Mouton de Gruyter.

Kingston, John and Diehl, R.L. 1994. Phonetic knowledge. *Language* 70, 419-454.

Kingston, John, Diehl Randy L., Cecilia J. Kirk, Wendy A. Castleman. 2008. On the internal perceptual structure of distinctive features: The [voice] constrast. *Journal of Phonetics* 36.28-54.

Kitto, C. and de Lacy, Paul. 1999. Correspondence in epenthetic quality. In: Kitto, C. Smallwood, C. (eds.) *Proceedings of ALFA* VI. University of Toronto Press, Toronto.

Klatt, Dennis. 1973. Duration characteristics of word-initial consonant clusters in English. *Journal of Acoustical Society of America.* 53:1, 379.

Klatt, Dennis. 1974. The duration of [s] in English words. *Journal of Speech, Language, and Hearing Research* 17.51-63.

Klatt, D.H. and L.C. Klatt. 1990. Analysis, synthesis and perception of voice quality variations among male and female talkers. *Journal of the Acoustical Society of America* 87: 820–856.

Ladefoged, Peter, Jenny Ladefoged, Alice Turk, Kevin Hind and St. John Skilton. 1998. Phonetic structures of Scottish Gaelic. *Journal of the International Phonetic* Association 28. 1-41.

Lee, Ahroung and Gregory K. Iverson. 2007. Variation in Korean Integration of English Word-Final /s/. *Language Research* 42, 239-251.

Lee, Hyunbok and Cheol-jae Seong. 1996. Experimental phonetic study of the syllable duration of Korean with respect to the positional effect. *Proceedings of Fourth International Conference on Spoken Language Processing.*

Lombardi, Linda. 2002. Markedness and the typology of epenthetic vowels. *Proceedings of Linguistics and Phonetics*.

Löfqvist, A., Baer, T., McGarr, N.S. and Seider Story, R. 1989. The cricothyroid muscle in voicing control. *Journal of the Acoustical Society of America* 85:1314-1321.

MacKain, Kristine, Catherine T. Best and Winifred Strange. 1981. Categorical perception of English /r/ and /l/ by Japanese bilinguals. *Applied Psycholinguistics* 2.369-390.

MacMillan, Neil A. and C. Douglas Creelman. 2005. Detection Theory: A User's Guide. Mahwah, NJ: Lawrence Erlbaum.

Maddieson, Ian. 1985. Phonetic cues to syllabification. In V. Fromkin (ed.), *Phonetic linguistics: essays in honor of Peter Ladefoged*, 203-221. New York: Academic Press.

Maddieson, Ian. 1997. Phonetic Universals. In William J. Hardcastle and John Laver (eds.), *The Handbook of Phonetic Sciences,* 619-639. Oxford and Cambridge: Blackwell.

Maniwa, Kazumi, Allard Jongman and Travis Wade. 2009. Acoustic characteristics of clearly spoken English fricatives, *Journal of the Acoustical Society of America*, 125 (6).3962-3973.

Massaro, D.W. and Cohen, M. M. 1983. Phonological constraints in speech perception. *Perception and Psychophysics* 34, 338-348.

Miyawaki, K., Strange, W., Verbrugge, R.R., Liberman, A.M., Jenkins, J.J., Fujimura, O. 1975. An effect of linguistic experience: the discrimination of [r] and [l] by native speakers of Japanese and English. *Perception and Psychophysics* 18.5, 331-340.

Ohde, R. 1984. Fundamental frequency as acoustic correlate of stop consonant voicing. *Journal of Acoustical Society of America* 75.224-230.

Peperkamp, S. and Emmanuel Dupoux. 2002. A typological study of stress 'deafness'. In C. Gussenhoven and N. Warner (eds.) *Laboratory Phonology 7*. Berlin: Mouton de Gruyter.

Peperkamp, S. 2005. A psycholinguistic theory of loanword adaptations. In M. Ettlinger, N. Fleischer, and M. Park-Doob (Eds.), *Proceedings of the 30th Annual Meeting of the Berkeley Linguistics Society* 341-352. Berkeley, CA: The Society.

Peterson, G. E. and Barney, H. L. 1952. Control methods used in a study of the vowels. *Journal of the Acoustical Society of America* 24, 175-184.

Pierrehumbert, Janet. 2001. Exemplar dynamics: Word frequency, lenition and contrast. In Joan L. Bybee and Paul Hopper (eds.), *Frequency and the Emergence of Linguistic Structure*. 137-157. John Benjamins, Amsterdam.

Pierrehumbert, Janet B. and Talkin, D. 1992. Lenition of /h/ and glottal stop. In Docherty, G.J. and Ladd, D.R. (eds.) *Papers in Laboratory Phonology II. Gesutre, segment, prosody*, 90-117. Cambridge, etc.: Cambridge University Press.

Pitt, M. A. 1998. Phonological process and the perception of phonotactically illegal consonant cluster. *Perception and Psychophysics* 60, 941-951.

Ren, N. 1992. *Phonation Types and stop consonant distinctions: Shanghai Chinese*. Ph.D. Dissertation, University of Connecticut.

Schwartz, M. F. 1970. Duration of /s/ in /s/-plosive blends. *Journal of the Acoustical Society of America* 47.1143(L).

Sheldon, A. and Strange, W. 1982. The acquisition of /r/ and /l/ by Japanese listeners of English: Evidence that speech perception can precede speech perception. *Applied Psycholinguistics* 3, 243-261.

Shinohara, S. 1997. Analysis phonologique de I'adaptation Japonaise des mots étrangers. Doctoral dissertation. Paris III (ROA).

Silverman, K.E.A. 1987. *The structure and processing of fundamental frequency contrours.* Doctoral dissertation, University of Cambridge.

Song, Jae Jung. 2005. *The Korean Language: structure, use and context*. London: Routledge.

Steriade, Donca. 2001. The Phonology of Perceptibility Effects: the P-map and its Consequences for Constraint Organization. UCLA ms.

Todaka, Yuichi. 1993. *A Cross-language Study of Voice Quality: Bilingual Japanese and American English Speakers.* Doctoral dissertation, University of California at Los Angeles.

Uffmann, Christian. 2006. Epenthetic vowel quality in loanwords: Empirical and formal issues. *Lingua* 116,1079-1111.

Vendelin, Inga and Sharon Peperkamp. 2005. Evidence for phonetic adaptation of loanwords: an experimental study. *Actes des Journées d'Etudes Linguistiques* 129-131.