Non-native Perception and Production of Foreign Sequences

A Dissertation Presented

by

Jiwon Hwang

to

The Graduate School

in Partial Fulfillment of the

Requirements

for the Degree of

Doctor of Philosophy

in

Linguistics

Stony Brook University

August 2011

Stony Brook University

The Graduate School

Jiwon Hwang

We, the dissertation committee for the above candidate for the

Doctor of Philosophy degree, hereby recommend

acceptance of this dissertation.

Ellen Broselow, Professor, Department of Linguistics

Marie K. Huffman, Associate Professor, Department of Linguistics

Christina Y. Bethin, Professor, Department of Linguistics

Nancy Squires, Professor, Department of Psychology

Arthur Samuel, Professor, Department of Psychology

This dissertation is accepted by the Graduate School

Lawrence Martin
Dean of the Graduate School

Abstract of the Dissertation

Non-native Perception and Production of Foreign Sequences

by

Jiwon Hwang

Doctor of Philosophy

in

Linguistics

Stony Brook University

2011

A major question in the study of second language learning is the extent to which mispronunciations originate in the inability to correctly perceive vs. the inability to correctly produce foreign language structures. The goal of this thesis is to determine the extent to which second language (L2) learners' pronunciation errors reflect errors in perception or gestural mistiming, by investigating Korean L2 learners' production and perception of English stop-nasal sequences. Such sequences are prohibited in Korean, where a stop before a nasal obligatorily undergoes nasalization (/kukmul/→[kuŋmul] 'soup'). In production experiments where Korean L2 learners pronounced English nonsense words containing those sequences, vowel insertion after the stop and devoicing of the stop were common errors even though nasalization is the native repair strategy. More importantly, two asymmetries in the choice of repair were that (1) vowel insertion occurred almost exclusively after voiced stops, especially after velar stops (tegnal→tegVnal) and (2) devoicing occurred most frequently with labial stops (tebnal→tepnal). These asymmetric repair choices are puzzling because neither of the languages in contact provides evidence for such repairs.

Investigation of Korean speakers' perception of these sequences employing both behavioral tasks and EEG revealed that the greater frequency of vowel insertion after voiced stops was rooted in misperception: Korean listeners tended to hear an illusory vowel after voiced stops and had difficulty distinguishing voiced stop-nasal sequences from voiced stop-vowel-nasal sequences. This misperception is an effect of the native language system, in which voiced stops occur only preceding a vowel. In contrast, the frequent devoicing of labial consonants was not reflected in perception. I propose that instead, this pattern has its origin in the articulatory timing patterns of Korean, in which bilabial stops are much more closely overlapped with a following consonant than are velar stops, causing devoicing of [b]. Results of this project show that second language phonology involves a complex interplay between the native language grammar, misperception through the filter of L1, and mastery of new articulatory programs.

Table of Contents

List of Figures

List of Tables

Acknowledgments

I would like to sincerely thank my advisor Dr. Ellen Broselow for her help, support, guidance, patience and understanding at every step of this dissertation process. She has been encouraging me to pursue new ideas and to develop them in a scientific and meaningful way ever since I wondered why Korean speakers sometimes put a vowel when they pronounce 'magnet' in Error Analysis, the very first class I took from her. I also learned the value of diligence and passion from her. It has truly been a privilege to be her student. I would also like to thank Dr. Marie Huffman for teaching me the joy of learning phonetics and for keeping me focused in writing this dissertation. Without her insightful comments and suggestions and positive motivations, it would have been impossible to reach this stage. I thank Dr. Christina Bethin for teaching me phonological theories and training me to analyze linguistic data in her classes and seminars. I thank Dr. Arthur Samuel for teaching me statistics and sharing psycholinguistic perspectives. I thank Dr. Nancy Squires for introducing me to the new world of ERPs, teaching me how to conduct ERP studies, and letting me be a member of her lab. I thank Jennifer Park for helping me write scripts for ERP data analysis; she helped me save tons of time. I am thankful to all of my research participants for their willingness to be a part of this research.

I thank my linguist friends for sharing my ups and downs during my graduate life at Stony Brook: Yunju Suh, Young-ran An, Susana Huidobro, Carlos de Cuba, Laura Spinu, Marianne Borroff, Miran Kim, Andrei Antonenko, Tomoko Kawamura, Yu-an Lu, Mark Lindsay, Sandra Brennan, Hisako Takahashi, Chih-hsiang Shu, Hyun-ju Kim, Yukiko Asano, Sophia Kao, Hijo Kang, Ivana Mitrović, Katharina Schuhmann, Poppy Slocum, Jisung Sun, Shawn Gaffney and Ling Wang.

I thank my husband Daniel An and my son Yujay An for their support and love. Observing Yujay's language development as a linguist has been a great joy! My gratitude also goes to my family for their love, support and prayers: my parents, my parents-in-law, my brother, my sisters-in-law, brothers-in-law and all of my nephews and nieces.

This dissertation was supported by NSF Grant BSC 0746027.

Chapter 1
Introduction

1.1 Sources of Errors in Non-native Production

In second language (L2) development, L2 learners often make errors. Many L2 researchers have argued that the source of the errors is the native language system (Ellis, 1997; Gass, 1996; Schwartz & Sprouse, 1994). Under this approach, foreign linguistic forms are altered to conform to the phonology of the first language. However, examination of L2 data reveals many patterns that appear to be incompatible with this hypothesis. For example, learners may perform better on one TL structure than another, even though both are equally illegal in the NL. Furthermore, learners may repair TL structures by means of strategies (such as vowel insertion, consonant deletion, devoicing) that are not supported by the NL grammar. If learners' error patterns cannot all be analyzed as the transfer of the NL phonology, what are the other sources of error patterns? Do production errors reflect the learners' misperception of TL structures?

This dissertation aims to provide answers to these questions by investigating Korean speakers' production and perception of English stop-nasal sequences. Such sequences are illegal in Korean, and the data in (1) show that in Korean, all obstruents before a nasal become nasalized.

(1) Obstruent nasalization in Korean

	Input	Output	Gloss
a.	/kuk + mul/	[kuŋmul]	broth
	/kuk + i/	[kugi]	broth, nominative
b.	/kap^h + nɨn + ta/	[kamnɨnda]	repay, present tense
	/kap^h + a/	[kap^{h}a]	repay, imperative
c.	/muk' + nɨn + ta/	[muŋnɨnda]	tie, present tense
	/muk' + ə/	[muk'ə]	tie, imperative
d.	/pic^h + nan + ta/	[pinnanda]	shine, present tense
	/pic^h + i/	[pic^{h}i]	light, nominative
e.	/s'is + nɨn + ta/	[s'innɨnda]	wash, present tense
	/s'is + ə/	[s'isə]	wash, imperative

This phonotactic constraint is obeyed within a word or a phrase unless an intonational phrase boundary intervenes (Jun, 1992). That is, if a pause occurs between the two consonants, a stop followed by a nasal can be produced without nasalization of the stop.

Unlike Korean, however, English permits a sequence of stop followed by a nasal, as in words like 'segment', 'webmail' or 'jackknife'. Thus we would expect Korean speakers to have difficulty with English words containing a stop before a nasal. Korean speakers are faced with learning a new phonotactic pattern as well as learning to produce coda voiced stops.

If second language learners take their native language grammar as a starting point, then we would expect that Korean L2 learners of English would use the same repairs in English as they use in Korean. However, as I show in Chapter 2, Korean learners of English use different strategies:

- Nasalization: tepman → te**m**man
- Devoicing: sebnan → se**p**nan
- Vowel insertion : segmal → seg**ɨ**mal

While they sometimes nasalize the stops, they often devoice the voiced stop. They also insert a vowel after voiced stops. This is surprising, given that Korean does not have a voicing contrast. Furthermore, the frequency of occurrence of each repair strategy depends on the voicing and/or the place of the stop, factors that are irrelevant in nasalization, the native repair process in Korean.

I argue in Chapter 2 that the learners' error patterns cannot be fully accounted for as an effect of NL transfer. In Chapter 3, I demonstrate, based on results from perception studies, that the frequent pattern of vowel insertion after voiced stops in production reflects misperception; Korean listeners interpret a voiced stop as a voiced stop and a following vowel. In Chapter 4, I argue that an interlanguage grammar, which is developed under the assumption that learning is error-driven and perception is accurate, cannot fully account for the place asymmetry found in devoicing and vowel insertion. I also demonstrate that it is not possible to develop a grammar that derives vowel insertion because such a grammar would not be learnable from the available data. This was confirmed by the perception experiment results, of which showed that a vowel was already inserted in perception rather than in production. In Chapter 5, I develop a separate perception grammar that allows a mapping from voiced stops to voiced stop+[ɨ]. In Chapter 6, I argue that the greater frequency of vowel insertion after velar stops and the greater frequency of devoicing of labial stops are a result of gestural mistiming, indicating the significant role of mastery of gestural phasing in L2. In Chapter 7, I conclude that L2 phonology reflects a complex interplay of reranking of the production grammar, perceptual filtering and articulatory difficulty.

In the following section I review the major approaches to the learning of second language phonology which form the background for the current study.

1.1.1 Second language production

Some analyses of language contact phonology take the view that L2 learners or recipient language speakers in loanword adaptation change foreign structures to conform to the native language grammar, known as a *Full Transfer* (Eckman, 1977; Gass, 1996; Schwartz & Sprouse, 1994; Tarone, 1987). The effect of the native language in foreign language production can be viewed as equivalent to the initial state of second language acquisition, where the first language grammar plays a dominant role (Broselow, 2004; Escudero & Boersma, 2004). Within the framework of Optimality Theory (OT), the initial NL constraints can be reranked in response to target language (TL) data, eventually reaching the target language grammar (Broselow, 2004). The learning mechanism in this approach is essentially error-driven in that the mismatch between the optimal output and the actual output, i.e., the error, drives the reranking of the constraints (Boersma, Dekkers, & van de Weijer, 2000; Tesar & Smolensky, 2000).

In these developmental stages, a process known as *Full Access* (to Universal Grammar) or an effect of markedness is shown to govern some interlanguage patterns (Broselow & Finer, 1991; Eckman & Iverson, 1993; Hansen, 2001). For instance, speakers of Mandarin Chinese produced voiced coda obstruents in English less accurately than voiceless obstruents, which is argued to be an effect of markedness (Hansen, 2001). In an Optimality Theory approach, a similar idea, the Emergence of the Unmarked (McCarthy & Prince, 1994) is also argued to play a significant role in shaping interlanguage patterns that cannot be fully accounted for by the initial grammar or the target language grammar. But in Optimality Theory approaches, the role of universal markedness is expressed in terms of a set of universal structural constraints (Hancin-Bhatt, 2008). For example, native speakers of Mandarin in Broselow et al.'s (1998) study had less accuracy in their production of English voiced codas than voiceless codas and furthermore tended to devoice voiced final obstruents regardless of the fact that neither of the voiced stops nor voiceless stops are possible codas in Mandarin. Broselow et al. argue that universal markedness constraints, which are masked in the native language, emerge in the developing interlanguage grammar.

In summary, under this approach the mechanism that drives the modifications in interlanguage phonology is solely phonological and is conducted at the production level, where the native language production grammar and the universal markedness may interact. Furthermore, the general assumption underlying the components of this approach is that the L2 input to the native grammar is the phonetic output of the target language grammar, which implies that L2 speakers accurately perceive the target structures (Jacobs & Gussenhoven, 2000; Paradis, 1996; Paradis & LaCharité, 1997).

However, substantial evidence exists that the perception of L2 learners is also significantly affected by their native language (Flege, McCutcheon, & Smith, 1987). Moreover, there are many interlanguage patterns that cannot be explained by the phonological grammar of the first language or any interlanguage grammar that can be learnable from the input data. These cases suggest a role for the effect of misperception (Broselow, 2009b). In the next section, I will review some of the recent work on the role of perception in interlanguage phonology.

1.1.2 Second language perception

Some analyses of foreign language modification have taken into account the role of perception (Brannen, 2002; Broselow, 2009b; Hancin-Bhatt, 2000; Y. Kang, 2003; Peperkamp & Dupoux, 2003; Silverman, 1992 among others). Peperkamp & Dupoux (2003) take the strong position that all production modification reflects perception errors while others (Broselow, 2009b; Escudero & Boersma, 2004; Y. Kang, 2003; Kenstowicz, 2001; Yip, 2006) take a moderate view that both perception and production contribute to the foreign language repair process. Peperkamp & Dupoux (2003) propose a model of loanword adaptations, in which *phonetic decoding*, a mapping from non-native acoustic signal to the closest phonetic categories in L1, filters out acoustic details that are irrelevant to the L1 sound system. All modifications including segmental and suprasegmental change, insertion, or deletion in adapting foreign words are argued to occur at this level rather than at the production level. However, studies have shown that misperception cannot be the only source that contributes to repairs of foreign structures, although it is an important one (Smith, 2004).

Some studies have incorporated perception effects into the phonological model. For example, Kang (2003) argues that the pattern of frequent vowel insertion after English final post-vocalic stops in Korean loanwords of English can be explained in relation to the perceptual similarity between the English surface forms and the Korean lexical representations. Miao (2005) also adopts the view that maximizing the perceptual similarity between the source form and the recipient form is one of the critical factors in loan adaptation. For instance, American English liquids /l/ and /r/ codas are generally adapted as Mandarin rhotic vowel /ɚ/ (Intel → /iŋ-tʰɤ-**ɚ**/, Deere →/ti-**ɚ**/) despite the fact that Mandarin has liquids /l/ and /ɻ/. Although these liquids are restricted to onset position, in principle, English coda /l/ and /r/ could be adapted as Mandarin /lV/ and /ɻV/, keeping the /l/-/r/ contrast (Dunhill → təŋ-ɕi-**lu**). But this adaptation is very rare compared to the change to the rhotic vowel /ɚ/. Miao (2005) proposes that this pattern can be attributed to perceptual similarity because American English coda /l/ or /r/, which can be often syllabic, is perceptually more similar to Mandarin rhotic vowel /ɚ/ than to Mandarin /lV/ or /ɻV/.

Other studies of loanword adaptation or second language research assume a separate level of perception from a production grammar (Broselow, 2009b; Escudero, 2005; Escudero & Boersma, 2004; Kenstowicz, 2001; Silverman, 1992; Yip, 2006). Under this approach, foreign language modification can arise through perceptual filtering as well as by the phonological production grammar. For instance, Hancin-Bhatt (2000) illustrates a puzzling pattern in the production of English liquid-fricative complex codas by Thai learners of ESL, in which a liquid is much more likely to be deleted than a fricative even though neither liquids nor fricatives are possible in coda position in Thai (e.g, nals → [nas], *[nal]). Based on the results of a perceptual identification task, which showed that Thai speakers tended to fail to hear a coda /l/ in liquid-fricative codas, she argues that the /l/ deletion occurs at the perceptual level, and therefore the representation to the production grammar simply does not contain /l/. Hancin-Bhatt also found that Thai speakers made errors in the production of English simple codas. While the accurate production of nasal codas was almost 100%, the accuracy of liquid codas and voiced stop codas was 83% and 67%, respectively. Hancin-Bhatt argues that this accuracy difference in Thai speakers' production of English simple codas reflects the fact that nasals are allowed to occur in coda whereas liquid codas are rare and voiced stops are never allowed in Thai. Hancin-Bhatt's study proposes that production errors could be a result of either misperception or transfer of NL grammar.

Another issue in discussion of the influence of L1 on perceptual processing is whether perceptual filtering is done at the abstract phonological level or at the acoustic level. One claim is that L2 listeners suppress what would be predictable and non-contrastive information in their NL when they process acoustic signals of the TL (Brown, 2000; Kabak & Idsardi, 2007; Matthews & Brown, 2004), whereas another claim is that L2 listeners do not dismiss phonologically redundant acoustic information, but match the phonetic properties of the TL signal to the NL phonetic categories that mostly closely approximate them (Flege, 1995; Peperkamp & Dupoux, 2003; Silverman, 1992). An argument that supports the view that even non-contrastive phonetic information is available in interlanguage mapping partly comes from studies demonstrating context-specific modification patterns (Y. Kang, 2008; Suh, 2010). For instance, the adaptation pattern of Japanese /k/ into Korean is mainly determined by the position of the /k/: it is adapted as Korean lax stop /k/ in word medial position (e.g., katuo → kas*ɨo 'bonito') but as the tense stop /k*/ in word medial position (e.g., mikaN → mik*aŋ 'tangerine') (Ito, Kang, & Kenstowicz, 2006). This is argued to reflect perceptual similarity: Japanese initial /k/ is more close to Korean lax /k/ in terms of the cues provided by the following vowel rather than closure duration, a less robust cue, while Japanese intervocalic /k/ is closer to Korean tense /k*/ because the closure duration of the two segments, a robust cue in intervocalic position, is similar. Such a pattern cannot be readily explained under the model where adaptation is expected to be categorical across the board, whereas it can be under the approach, in which contextual information is accessed in interlanguage mapping.

1.1.3 Gestural mistiming

Some studies have found that articulatory timing or coarticulation patterns are language-specific (An, Hwang, & Suh, 2008; Oh, 2008; Zsiga, 2003). For example, Oh (2008) has demonstrated that French and English CV (alveolar stop+/u/) syllables have different coarticulation patterns, as indicated by differences in target values of F2 as well as in F2 slope change shape in the whole CV trajectory, and that early L2 learners did not acquire the dynamic pattern of the articulation in the target language while more advanced learners did. This indicates that language-specific coarticulation patterns are gradually learned just as the target values are. Zsiga (2003) also has shown that English and Russian clusters pattern differently in gestural timing, and argued that both native language and universal articulation patterns were reflected in L2 production of target language clusters. Given that not all phasing relations are universal, L2 learners are faced with learning target gestures as well as gestural coordination patterns of L2. And it is reasonable to hypothesize that some L2 production errors may reflect the failure to acquire the L2 gestural patterning.

Some studies have argued that difficulty in mastering new patterns of articulatory timing is another source of non-native production (Colantoni & Steele, 2007, 2008; Davidson, 2010; Messing, 2008). Important evidence for articulatory factors in interlanguage production comes from studies on L2 speech of non-native clusters (Davidson, 2003; Davidson, 2006a, 2006b; Davidson & Stone, 2003). Davidson has shown, based on ultrasound imaging and acoustic investigation, that an inserted vowel in non-native sequences may have an articulatory and acoustic profile different from that associated with true vowels. She argues that the vowel percept in non-native clusters comes from gestural mistiming that arises when two consonant gestures are not closely coordinated. Similarly, Messing (2008) also reported that the acoustic properties of the inserted vowels found in the production of English stop-liquid clusters by Mandarin learners of English fit very well with the qualities of intrusive vowels described in Hall (2006). Colantoni & Steele (2007) argue that interlanguage patterns that cannot result from L1 transfer are primarily governed by articulatory principles. Their investigation of the acquisition pattern of French voiced uvular fricative /ʁ/ by English speakers demonstrated that frication of /ʁ/ was acquired earlier than voicing. These patterns suggest that English speakers had more difficulty in producing voicing than frication due to articulatory and perceptual reasons. They also found that voicing of /ʁ/ was acquired earlier in prevocalic position, where the articulation of voicing is favored for aerodynamic reasons, supporting their articulatory-based hypotheses.

The importance of articulatory factors in non-native production of foreign clusters has been argued to play a role not only in the quality of inserted vowels but also in the relative accuracy of production. For example, in Davidson's (2010) study, accuracy depended on the voicing of the clusters or the type of the clusters (stop-initial sequences vs. fricative-initial sequences). Both English and Catalan speakers, regardless of their different language backgrounds, produced voiceless clusters and fricative-initial clusters more accurately than voiced clusters and stop-initial clusters, respectively. Davidson argues that this pattern is captured better by the articulatory account than by native language effects or markedness scales[1]. Messing (2008) also shows that Mandarin speakers were more accurate in producing voiceless stop-liquid clusters than in voiced stop-liquid clusters, and homorganic clusters than in heterorganic clusters, even though no evidence for this differential difficulty is provided in the NL or in TL. Messing argues that this pattern reflects the effect of gestural mistiming.

[1] Davidson (2010) also demonstrates that the choice of repairs reflects the native language grammar.

Chapter 2
Errors in Production

The goal of this chapter is to examine the production of English stop-nasal sequences by Korean L2 learners of English and to identify common errors. I report on two production experiments in which Korean L2 learners of English produced nonsense English words containing a voiced or voiceless stop followed by a nasal. In the first experiment, Korean L2 learners of English read English sentences where the experimental words were embedded in initial position. In the second experiment, they heard English nonce words containing an unreleased stop followed by a nasal, and repeated what they heard, which was designed to control the potential effect of orthography and release of the stop. I demonstrate that, despite the different methods, both experiments display highly similar repair patterns and I argue that the *Full Transfer Hypothesis* cannot fully explain such patterns.

2.1 Korean Sound System

I begin by setting out the facts of the Korean sound system. Korean has three different sets of voiceless stops: lenis, aspirated and tense stops. The following is the phoneme inventory of Korean.

(1) Inventory of Korean phonemes (adopted from Y. Kang, 2003)

p	p^h	p’	t	t^h	t’	k	k^h	k’		i	ɨ	u
			ts	ts^h	ts’					ɛ	ə	o
			s		s’				h		a	
m			n									
			L[2]							j		w

As shown in (2), Korean has no voicing contrast; lenis stops become allophonically voiced in intersonorant position. Some examples of the predictable voicing process are given in (2):

(2) Intersonorant prevocalic voicing in Korean

	Input	Output	Gloss
a.	/papo/	[pabo]	fool
b.	/kilta/	[kilda]	be long
c.	/kamtok/	[kamdok]	a director

[2] /L/ is a representation of a liquid phoneme in Korean, which is realized as /l/ or /ɾ/ depending on the syllabic position (Kang 2003, Cho 1997).

As shown in (3), stops may not occur before a nasal consonant. All obstruents before a nasal become nasalized (Lee & Pater, 2008).

(3) Obstruent Nasalization

	Input	Output	Gloss
a.	/kuk + mul/	[kuŋmul]	broth
b.	/kap^h + nɨn + ta/	[kamnɨnda]	repay, present tense
c.	/muk' + nɨn + ta/	[muŋnɨnda]	tie, present tense
d.	/pits^h + nan + ta/	[pinnanda]	shine, present tense
e.	/s'is + nɨn + ta/	[s'innɨnda]	wash, present tense

2.2 Production Experiment I: Reading

This experiment was designed to examine how Korean native speakers would produce English stop-nasal sequences when reading a passage written in English. The English reading passage consisted of an introduction to a made-up RPG (Role Playing Game) containing descriptions of characters appearing in the game. The names of the characters, all nonce words, served as both the experimental and filler items. Nonce words were used to investigate production behaviors independent of effects of familiarity or lexical frequency. The introduction of the game did not include any items containing stop-nasal sequences.

2.2.1 Participants

20 Korean native speakers from the Stony Brook student population participated in the production task, 5 males and 15 females. All spoke Korean as their native language and reported that the language they used most frequently was Korean. The average age of first exposure to English study was 10.7 years old and the average age of arrival in an English-speaking country was 23.5 years old. The average length of residence in an English-speaking country was 12.8 months. None reported any speech or hearing disorders. As a control group, 6 English native speakers participated in the experiment, recruited from the Stony Brook University Psychology subject pool. All volunteered to participate in the experiment and were paid a small monetary compensation or given course credit for their time upon completing the language questionnaire.

2.2.2 Materials

The experimental materials used in this experiment were nonce English words of the form CVCCVC (e.g., Tibnan, Kedmal) which contained a stop-nasal sequence in word internal position. All test words appeared in phrase-initial position and contained one of the stops [p, t, k, b, d, g] followed by one of the nasals [m, n], for a total of 12 sequences. (No velar nasals were used because the velar nasal cannot occur in onset position in English). All test items were disyllabic with the orthographic vowel 'i' or 'e' in the first syllable and 'a' in the second syllable. It was assumed that these vowels would be interpreted as [ɪ] and [ɛ] in the first syllable and as a reduced vowel in the second syllable when the stress was put on the first syllable. The first consonant varied among [p, t, k, s, z] and the last consonant was either [n] or [l]. The total number of the experimental items was 48 per participant (12 sequences* 2 vowels*2 repetitions).

Filler items were nonsense words containing non-native consonants or [l], which were used to draw Korean participants' attention away from the experimental items. [l] was included particularly to contrast with [ɹ] because the [l]-[ɹ] pair is generally known to be difficult for Korean speakers. All fillers contained consonants from the set [θ, ɹ, z, v, f] or [l]. The total number of the experimental items was 20. The complete text can be found in Appendix 1. Individual test items and fillers can be found in Appendix 2.

2.2.3 Procedures

Subjects were given a few minutes to familiarize themselves with the materials. Any English words in the text that they were not familiar with were produced by the experimenter, with the meaning if needed. Participants were instructed to produce the experimental items (which were proper nouns) any way they wanted to. If they asked about stress, they were instructed to put a stress on the first syllable. The recording session began only when participants indicated they were ready to proceed. Participants read the introductory passage once and the main body of the passage twice. The recording was conducted in the sound treated room in the Linguistics Department at Stony Brook University, using a Marantz digital recorder with 44.1kH sampling rate. Participants completed a session by filling out a simple language questionnaire. This session took about 20 minutes.

2.2.4 Transcriptions and analysis

Participants' productions of the internal consonant sequences in the experimental items (e.g. ti**pn**al, te**gn**an) were transcribed by three English native speakers who were phonetically trained[3]. In addition to the participants' data, 6 English native speakers' productions of the same items were transcribed[4]. These were randomly presented to the transcribers in an attempt to avoid transcribers' possible accommodation to Korean-accented English.

Transcribers were not aware of the target segments but were told that each item consisted of the form CVCCVC. When transcribers disagreed on an item, the transcription shared by any two of the transcribers was taken as accurate. When 3 disagreed, the transcription by the transcriber whose transcription of the English data was overall most accurate was used. On the basis of the transcriptions, the data were coded as one of the following: correct production; nasalization of the stop; devoicing of the voiced stop; vowel insertion between the sequences; or other. Other production patterns, which occurred much less frequently, were fricativization of the stop; metathesis; change in place of articulation of the stop; voicing of the voiceless stop; change of the nasal into a stop; and voicing of the voiceless stop with a different place of articulation. In the next section, the results of Korean participants' production of English stop-nasal sequences are presented.

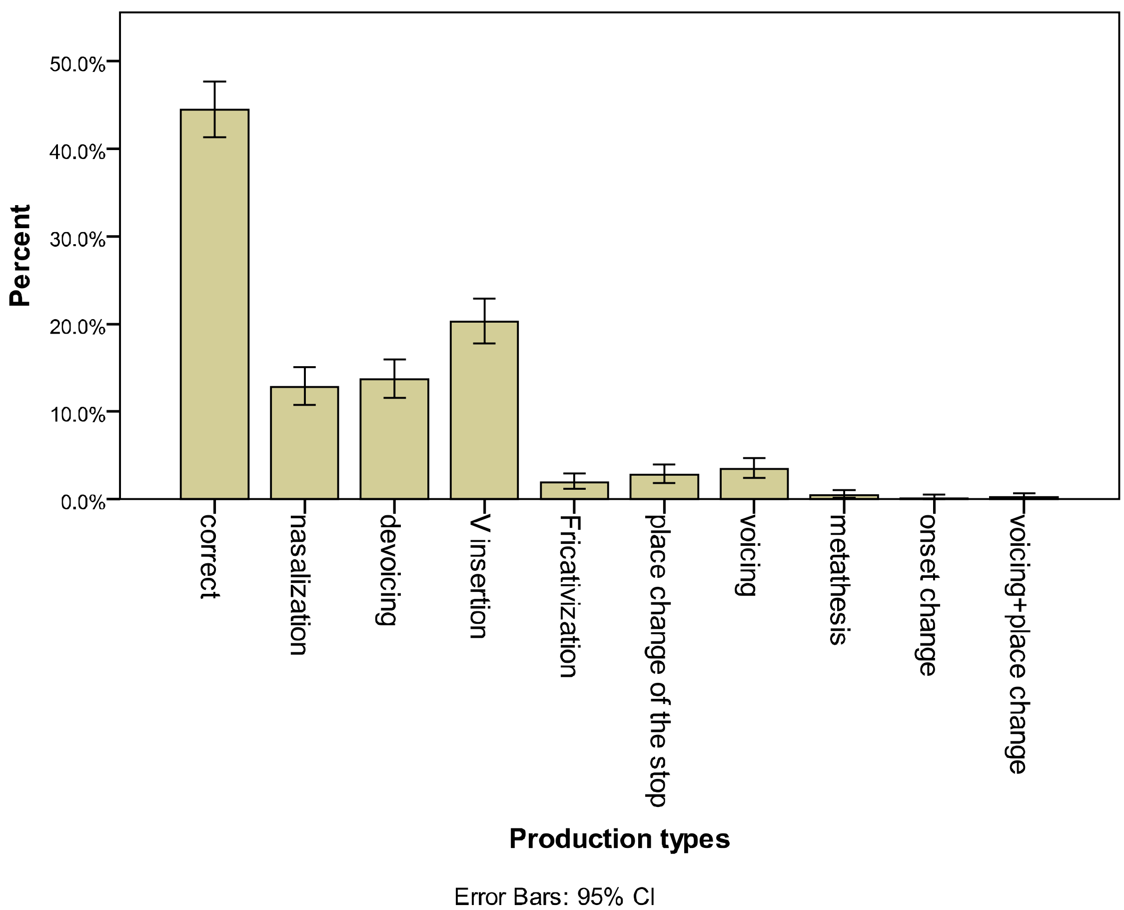
[3] 938 items produced by the Korean participants were transcribed.
[4] 288 items produced by the English speakers were transcribed.

2.2.5 Results

In this section, the rates of accurate production and repairs are presented. The predictions of the *Full Transfer Hypothesis* are evaluated, and the major repair types are statistically tested to determine whether they were significantly affected by the voicing or/and the place of articulation of the stop.

The proportion of production types of English stop-nasal sequences from the reading experiment is plotted in Figure 2.1. Overall, the percentage of accurate production in the Korean participants' production was over 40%, higher than the frequency of any other repair type. The major error patterns included nasalization, devoicing of the voiced stop and vowel insertion after the stop. The other minor error types were lower than 9% altogether and these were excluded from further analysis.

Figure 2.1. Proportion of production types of English stop-nasal sequences in the reading production experiment, averaged over all participants (error bars represent 95% Confidence Interval)



Recall that the predictions of the *Full Transfer Hypothesis* were that (i) nasalization should be the most favored repair for both sequence types (voiceless stop-nasal and voiced stop-nasal); and that (ii) accuracy rates should be generally low and similar for both sequence types. Although the production experiments present evidence that the native language process of nasalization was transferred, the predictions of the *Full Transfer Hypothesis*, which predicts that learners should repair stop-nasal sequences by nasalizing the stop, were not borne out.

First, nasalization was not the most common repair, either for voiceless stop-nasal sequences or voiced stop-nasal sequences. For voiceless stop-nasal sequences, participants were almost equally likely to insert a vowel between the clusters as to nasalize the stop, as illustrated in Table 2.1. Repeated measures ANOVA with *Repair type* (nasalization vs. vowel insertion) confirmed that the rates of nasalization and vowel insertion did not differ significantly ($F<1$). For voiced stop-nasal sequences, participants chose three main repair types: nasalization, vowel insertion, and devoicing of the stop. Table 2.1 shows that participants devoiced the stop or inserted a vowel frequently, but nasalized the stop much less frequently. An ANOVA revealed a *Repair type* effect ($F(2, 38)= 4.74$, $p<.05$), confirming this pattern. Bonferroni pairwise comparisons revealed that vowel insertion was significantly more frequent than nasalization ($p<.05$). Thus, for neither of the sequence types was nasalization preferred to other repair types.

Table 2.1. Production patterns found in the reading experiment

	Voiceless stop-nasal	Voiced stop-nasal
Nasalization	62 (13.5%)	58 (12.1%)
Vowel Insertion	54 (11.8%)	136 (28.3%)
Devoicing	-	128 (26.7%)

Second, the difference in the accuracy rates between voiceless stop-nasal and voiced stop-nasal sequences seems very large (See Table 2.2). An analysis of variance (ANOVA) on the accuracy rates was performed for participants with *Voicing* (voiceless vs. voiced) as an independent factor. A significant *Voicing* effect ($F(1, 19)= 50.34$, $p<.005$) was found, which confirms that voiceless stop-nasal sequences were more often accurately produced than voiced stop-nasal sequences. This was not an expected pattern under the *Full Transfer Hypothesis* either.

Table 2.2. Rates of accurate production of the sequences by voicing of the stop

	Voiceless stop-nasal	Voiced stop-nasal
Accurate Production	283 (61.8%)	134 (27.9%)

In this section, I have shown that the *Full Transfer Hypothesis* was not supported by the results of the reading production experiment. Instead, in addition to nasalization, which is the native repair strategy, two additional major repairs emerged in production of English stop-nasal sequences: devoicing and vowel insertion. In the following section, we examine in what context these repairs occurred.

2.2.5.1 Vowel insertion

As Table 2.3 shows, vowel insertion was much more frequent in voiced stop-nasal sequences than in voiceless stop-nasal sequences and in velar stop-nasal sequences than in other sequences. A by-subject ANOVA of vowel insertion frequency was carried out for *Voicing* (voiceless vs. voiced stop) and *Place*(bilabial, coronal vs. velar stop) and revealed a significant *Voicing* effect ($F(1, 19)= 21.742$, $p<.005$); vowel insertion occurred more frequently after voiced stops than voiceless stops. The analysis also showed a *Place* main effect ($F(2, 38)= 7.236$, $p<.005$). Pairwise comparisons showed that a vowel was inserted more often after velar stops than bilabial stops ($p<.005$). No interaction of the two factors was found.

Table 2.3. Rates of vowel insertion by voicing and place of the stop in the reading experiment

	Voiceless stop-nasal			Voiced stop-nasal		
	Bilabial	Coronal	Velar	Bilabial	Coronal	Velar
Vowel Insertion	9 (6.3%)	20 (12.7%)	25 (15.8%)	34 (21.3%)	39 (24.4%)	63 (39.4%)

2.2.5.2 Devoicing

Another major repair was devoicing of the stop. Table 2.4 reports devoicing frequency by the place of the stop. It appears that devoicing occurred more frequently with labial and coronal voiced stops than with a velar voiced stop. A by-subject ANOVA of devoicing frequency (from voiced stop-nasal sequences) with *Place* as a factor showed a significant effect (Place $F(2, 38)= 9.239$, $p=.001$), confirming the place effect. Pairwise comparisons indicated that devoicing was significantly less frequent after velar stops than bilabial stops ($p=.004$) and coronal stops ($p=.044$).

Table 2.4 The rates of devoicing by place of the stop in the reading experiment

	Voiced stop-nasal		
	Bilabial	Coronal	Velar
Devoicing	44 (27.5%)	62 (38.8%)	22 (13.8%)

2.2.6 Summary

The major repair types from the reading production experiment were nasalization, devoicing and vowel insertion. Interestingly, devoicing and vowel insertion occurred in specific contexts, a finding which requires further explanation. Below, I summarize the two effects and in the next section, I explore possible explanations.

(1) Voicing Effect

Vowel insertion was significantly more frequent in voiced stop-nasal sequences than in voiceless stop-nasal sequences.

(2) Place Effect

Devoicing occurred more frequently in labial stop-nasal sequences, whereas vowel insertion occurred more frequently in velar stop-nasal sequences.

2.2.7 Voicing effect

The fact that vowel insertion was more frequent in voiced stop-nasal sequences is surprising, given that voicing is not contrastive in Korean, and therefore Korean speakers are not expected to be sensitive to voicing contrasts in English. However, this pattern is highly similar to the pattern found in the adaptation of final stops in English words borrowed into Korean (Y. Kang, 2003; Park & de Jong, 2006; Rhee & Choi, 2001 among others). For example, the borrowed word ‘kick’ is pronounced as [k^hik] without a vowel whereas ‘mug’ is pronounced as [məgɨ] with a vowel inserted after the stop. Although the contexts are different, the role of voicing of the stop seems important in both cases.

Given that participants in the reading experiment were asked to read English nonsense words presented in English orthography, the Korean participants may have been affected by the convention that English voiced stops are written with a following vowel in Korean. According to the standardized loanword adaptation rules suggested by National Academy of the Korean Language, all voiced stops must be written with the grapheme ‘ㅡ’ in addition to the graphemes of stops (Kwuklipkwukeyenkwuwen, 2009). In the Korean writing system, this symbol represents the vowel [ɨ]. Vendelin and Peperkamp (2006) present evidence that loanword pronunciation is affected by borrowers’ knowledge of the conventional correspondence rules between the source language orthography and the borrowing language phonemes. For example, French speakers are instructed to use their phoneme /œ/ for the English grapheme ‘u’ as in ‘but’ and /u/ for ‘oo’ as in ‘book’ (Vendelin & Peperkamp, 2006).

2.2.8 Place effect

Another widely known factor that may cause vowel insertion is an attempt to match the release of the stop (Chung, 2004; Jun, 2002; Y. Kang, 2003; Rhee & Choi, 2001). Coda stops are not released in Korean, in contrast to English coda stops, which may be released. The closest approximation to a released coda stop that a Korean speaker can readily produce is a stop followed by a vowel. Jun (2002) reports on an experiment in which Korean speakers inserted a vowel after released coda stops significantly more often than after unreleased ones. Similarly, Kang (2003) demonstrates that in loanwords from English to Korean, vowel epenthesis is more likely to occur when a word final stop follows a tense vowel than a lax vowel. She argues that this pattern reflects the fact that in English, word final stops are more likely to be released after tense vowels than after lax vowels.

Importantly, Kang (2003) also finds more frequent vowel insertion after dorsal than after labial stops in loanwords. She attributes this pattern as well to the release frequency of English stops, as in her own survey of an English corpus, Kang (2003) found that sentence-final stops in English are more often released when they are dorsal than labial. Additional evidence comes from Rositzke’s (1943) study, which also reported more frequent release of dorsal stops than labial stops, as also cited in Kang (2003). Thus, the fact that vowel insertion was much more likely after dorsal than after labial stop in the current study may reflect participants’ attempt to maintain perceptual similarity with the input language phonetic details. Since the participants in the current experiment were experienced with English inputs, as the high accuracy rates suggest, it may have been that they were able to mimic the release pattern of English by inserting a vowel.

I have suggested that orthography and release of the stop may have contributed to the voicing and place effects found in the reading task. An experiment that removes these two factors can help to determine whether these two factors are relevant to the voicing and place effects. In the next section, I report on another production experiment where these two factors were controlled.

2.3 Production II: Repetition Experiment

In the second experiment, Korean speakers first heard English nonsense words with unreleased stop and nasal sequences and then repeated them. They were asked to repeat an auditory stimulus presented through headphones. This task was conducted in an attempt to minimize the possible effects of orthography and release of the stop.

2.3.1 Participants

The participants in this experiment were 20 Korean speakers, 12 males and 8 females. All were from the Stony Brook student population: 16 graduate students and/or their spouses and four undergraduate students. All of the subjects had studied English in an EFL context for more than six years in Korea, beginning at an average age of 12.3 years. The duration of stay in the U.S. varied from six months to four and one half years, with an average duration of one year and seven months. None reported any speech or hearing disorders.

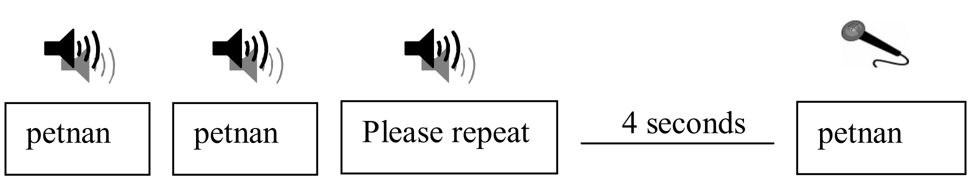
2.3.2 Materials

24 English nonce words were created as target items. Like the test items in the reading experiment, these items consisted of 12 bisyllabic minimal pairs with word internal stop-nasal sequences (e.g., [sɛbnəl] or [t^hɪkmən]). Again, the first vowel was either high front lax [ɪ] or mid front lax [ɛ], which are most similar to Korean front vowels. The total number of target words was 960 (12 sequences × 2 vowels × 2 repetition × 20 subjects = 960 words). 12 additional nonce words, containing word internal clusters of voiceless stops (which are licit sequences in Korean), were included as fillers (e.g., sɛpkən, tɪkpən).

To create auditory stimuli, a male English native speaker produced the experimental and filler items with stress on the first syllable at a natural speech rate. The native speaker was instructed to produce the items with no release of the coda stops before a nasal; there was neither an audible release burst nor any acoustic sign of it. The recording was conducted in a sound-treated booth in the Department of Linguistics at Stony Brook University. From three repetitions, the second repetition was taken as the auditory stimulus.

2.3.3 Procedures

20 Korean subjects were instructed to listen to a randomized set of the auditory stimuli produced by the English native speaker and to repeat what they heard. They were given no visual information (e.g., spelling of the stimuli) but only aural information through headphones. Each frame consisted of repetition of a stimulus followed by the phrase "Please repeat". After this they were given 4 seconds to produce the stimulus.



The subjects were familiarized with the experimental task by taking a practice trial round with 3 words that were picked from the filler items. Afterwards, they completed the first block with 24 experimental items and fillers and the whole block repeated. The recording was done at a 44100 Hz sampling rate with 16-bit resolution onto a laptop computer in a sound-treated booth in the Department of Linguistics at Stony Brook University, using Shure SM11 Dynamic Lavalier Microphone. The task took about 25 minutes to complete.

2.3.4 Transcriptions and analysis

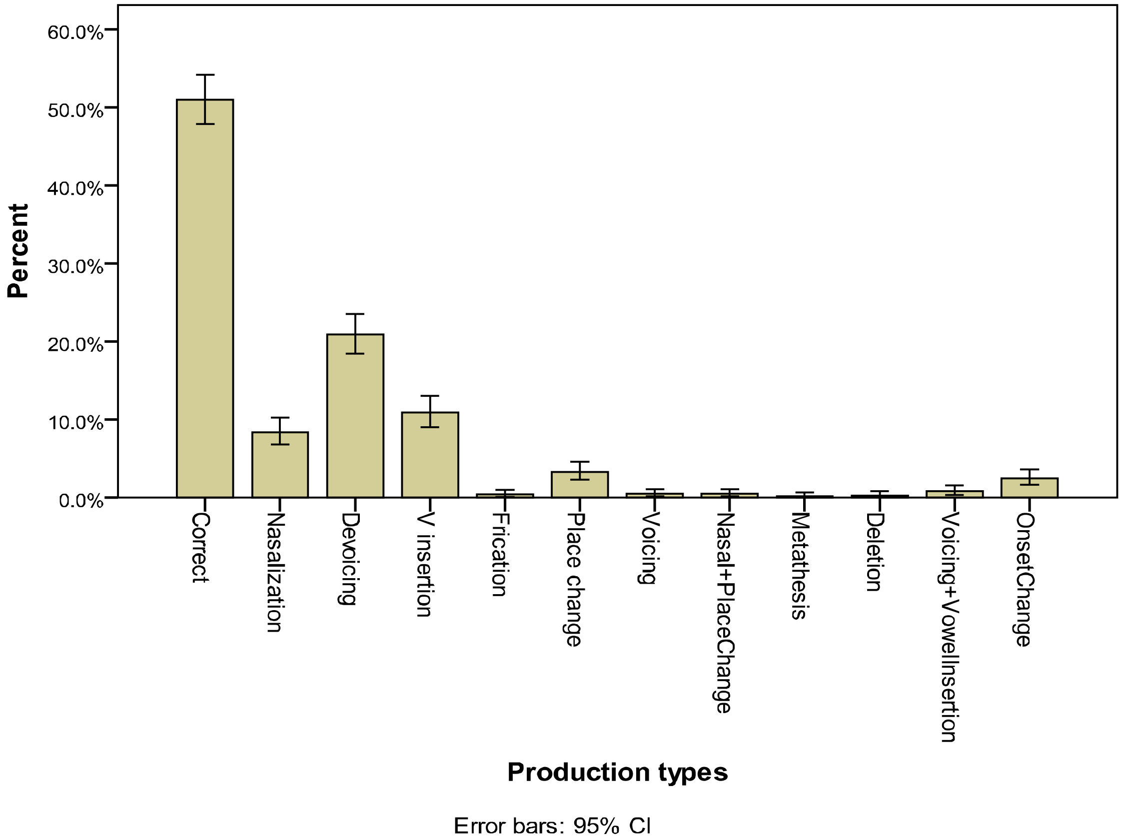
Two phonetically trained English native speakers transcribed the internal clusters of the target forms. They were unaware of the target words but were told that the words were of the form CVCCVC. When the transcribers disagreed, they discussed the form and agreed on a transcription. Based on the transcriptions, the data were coded as one of the following: correct production; nasalization of the stop; voicing change of the stop (devoicing for the voiced stops and voicing for the voiceless stops); vowel insertion between the two consonants; or other, which included fricativizing the stop, changing the place of articulation of the stop, changing the order of the stop and nasal (metathesis), deleting one of the consonants, voicing the stop with an inserted vowel or nasalizing the stop with a change of a place of articulation (See Appendix 4).

2.3.5 Results and discussions

The repetition experiment was carried out to see whether the voicing and place effects found in the reading experiment might have originated from orthography or from a desire to mimic the common release of final stops. If those factors contributed to the asymmetric patterns in experiment 1, they should not appear in the repetition experiment, where stimuli were presented only auditorily and the model contained no release of the stops. First, I present overall production types and their proportion, then examine whether the voicing and place effects still existed in the repetition experiment.

The proportion of the production types found in the repetition experiment is plotted in Figure 2.2. As in the reading experiment, accurate productions were more frequent than any other production types, and the three major repair types were nasalization, devoicing and vowel insertion.

Figure 2.2. Proportion of production types of English stop-nasal sequences in the repetition experiment, averaged over all participants[5] (error bars represent 95% Confidence Interval)



When the stop was voiceless, although nasalization occurred more frequently than vowel insertion, the difference did not reach significance ($F(1,19)=3.19$, $p=.09$), which suggests that there was large variation among the participants as to the frequency of nasalization and vowel insertion (See Appendix 5 for production pattern for individual participants). A by-subject ANOVA for voiced stop-nasal sequences showed that the frequencies of the repairs were significantly different from one another ($F(2, 38)=19.104$, $p<.005$). Pairwise comparisons showed that devoicing was significantly more often employed than vowel insertion ($p<.05$), which was significantly more often employed than nasalization ($p<.005$), as shown in Table 2.5. Again, as in the reading experiment, nasalization was not the preferred repair strategy.

5 'Place Change', the error that changed the place of the stop was quite frequent, and I suspect that it came from the fact that participants were given fewer cues to the place of the stop since the release bursts were not included in the auditory stimuli.

Table 2.5. The production patterns of the target CN in the repetition experiment

	Voiceless stop-nasal	Voiced stop-nasal
Nasalization	39 (8.1%)	42 (8.8%)
V insertion	7 (1.5%)	98 (20.4%)
Devoicing	-	201 (41.9%)

The frequency of accurate productions is listed in Table 2.6: the pattern in which voiceless stop-nasal sequences were more often accurately produced than voiced stop-nasal sequences in the reading experiment was found in the repetition experiment as well. A significant main effect of *Voicing* ($F(1, 29)=125.169$, $p<.005$) confirmed the difference.

Table 2.6. Accurate production of the sequences in the repetition experiment

	Voiceless stop-nasal	Voiced stop-nasal
Correct	369 (76.9%)	121 (25.2%)

Regardless of the task in the production experiments, the results consistently show that the Full Transfer Hypothesis cannot explain the patterns found in production of English stop-nasal sequences by Korean speakers.

2.3.5.1 Vowel insertion

Table 2.7 reports the frequency of vowel insertion in each context. As shown, vowel insertion was much more frequent in voiced stop-nasal sequences than in voiceless stop-nasal sequences, replicating the results of the reading experiment. A by-subject ANOVA of vowel insertion rates was performed with *Voicing* and *Place* as factors. The *Voicing* effect was significant ($F(1, 19) = 79.504$, $p<.005$), which confirms the pattern. The *Place* main effect was also significant ($F(2,38)=31.269$, $p<.005$); pairwise comparisons revealed that vowel insertion arose significantly more often in velar stop-nasal than in bilabial stop-nasal or in coronal stop-nasal sequences ($p<.005$ for both). This pattern is also comparable to the results of the reading experiment. The interaction of the two factors was significant ($F(2, 38)= 22.853$, $p<.005$), indicating that the voiced velar stop-nasal sequences were produced with vowel insertion significantly more often than the other sequences.

Table 2.7. Rates of vowel insertion by voicing and place of the stop in the repetition experiment

	Voiceless stop-nasal			Voiced stop-nasal		
	Bilabial	Coronal	Velar	Bilabial	Coronal	Velar
V insertion	0 (0%)	0 (0%)	7 (4.4%)	9 (5.6%)	17 (10.6%)	72 (45%)

2.3.5.2 Devoicing

More frequent devoicing in labial stop-nasal sequences was again found in the repetition experiment. A by-subject ANOVA on devoicing frequency was conducted with *Place*(bilabial, coronal vs. velar) and the effect was significant ($F(2, 28) = 22.081$, $p<.005$); pairwise comparisons showed that voiced bilabial stops were significantly more often devoiced than velar stops ($p<.005$).

Table 2.8. Rates of devoicing by place of the stop in the repetition experiment

	Voiced stop-nasal		
	Bilabial	Coronal	Velar
Devoicing	96 (60%)	79 (49.4%)	26 (16.3%)

2.4 Summary of the two production experiments

In the reading experiment, we found voicing and place effects with vowel insertion and devoicing. I hypothesized that these effects may have been due to the fact that the task used orthographic representation, and/or to participants' attempt to mimic the release pattern of English stops. In the repetition experiment, I controlled these two factors by presenting participants with auditory stimuli (rather than printed stimuli) containing no audible release bursts. The results of the repetition experiment were highly consistent with the results of the reading experiment, suggesting that neither orthography nor release bursts of the stimuli were a driving force in the voicing and place effects. Therefore, we still need explanations for the remaining questions listed below.

i. Accuracy Difference: In both experiments participants produced voiceless stop-nasal sequences more accurately than voiced stop-nasal sequences, even though neither sequence is legal in the native language.

ii. Voicing Effect: In both experiments, a vowel was inserted significantly more often after voiced stops than voiceless stops.

iii. Place Effect: In both experiments, vowel insertion was more common after velar stops than after labial stops. However, devoicing occurred significantly more often with labial stops than with velar stops.[6]

2.5 Summary

From the production experiments, we observed that the voicing and place of the stop in a stop-nasal sequence had a significant effect on the pronunciation of the sequence. Although

[6] Overall, the status of coronal stops was ambiguous. In some cases, coronals patterned with labial stops, but in other cases, they patterned with velar stops. The subsequent sections and chapters will include discussions of only labial and velar stops.

nasalization is explicable in terms of the NL production grammar, vowel insertion and devoicing are not. In Chapter 3, I investigate the possibility that the interlanguage patterns of devoicing and vowel insertion, which appear not to be learnable from the data of the languages in contact, are reflections of misperception rather than misproduction.

Chapter 3
Production Errors and Misperception

The production outcomes presented in Chapter 2 involved a range of repairs: nasalization, vowel insertion, stop devoicing and correct production. Of these repairs, only the first, nasalization, is attested in the native language. Furthermore, puzzling asymmetries were found in the choice of repair: (1) vowel insertion was more likely to occur when the stop was voiced and/or velar, but (2) devoicing was more likely to occur when the stop was labial than velar. Chapter 2 presented arguments that the voicing and place effects on vowel insertion and devoicing cannot be explained fully as transfer from the native language production grammar. The current chapter examines whether the asymmetries found in the production experiments were a result of misperception. The structure of this chapter is as follows: §3.1 reports on misperception hypotheses and several behavioral experiments; §3.2 reports on electrophysiological experiments; §3.3 summarizes the findings of the experiments.

3.1 Behavioral Experiments

In the preceding chapter we considered the facts of Korean production from the standpoint of the Phonological Grammar-Based Approach, assuming that learners accurately perceive foreign language forms. However, various studies have established that L2 learners' perception of foreign language structures is deeply affected by L2 learners' language background (Broselow, 2009b; Y. Kang, 2003; Kenstowicz, 2001; Peperkamp & Dupoux, 2003; Silverman, 1992; Yip, 2002 among others). These studies support the idea that non-native perception is not faithful and therefore affects the way foreign forms are produced or adapted.

Korean participants in the two production experiments employed devoicing and vowel insertion in stop-nasal sequences, in addition to the NL strategy of nasalization. We found the following patterns regarding devoicing and vowel insertion:

- Devoicing more often applied to labials (e.g., tebnal → tepnal).
- Vowel insertion more often applied to velars (e.g., teknal → tekɨnal) or in voiced context (e.g., tegnal → tegɨnal).

I will investigate these asymmetries in light of the following hypotheses:

(1) Misperception Hypotheses for Korean listeners

H1. **Voicing induces a bias toward perception of illusory vowels.**
This would account for the greater likelihood of vowel insertion after voiced than voiceless stops.

Prediction: [gN]-[gɨN] harder to distinguish than [kN]-[kɨN]

H2 **Bilabial closure induces a bias toward perception of the stop as voiceless.**
This would account for the greater likelihood for [b] than [g] to be devoiced.

Prediction: [bN]-[pN] harder to distinguish than [gN]-[kN]

H3 **(therefore) Velar closure induces a bias toward perception of an illusory vowel for Koreans.**
This would account for the greater likelihood of vowel insertion in [gN] than in [bN].

Prediction: [gN]-[gɨN] harder to distinguish than [bN]-[bɨN]

The basis for *H1* is the fact that voiced stops in Korean are always followed by a vowel. Voicing in Korean is fully predictable, with voiced stops occurring only as an allophonic variant of lenis stops surrounded by voiced segments (Y.-M. Cho, 1990; Jun, 1993; Silva, 1992). As shown in (2), voiced stops occur between vowels (a-b), or after a sonorant and before a vowel (c-f).

(2) Intersonorant voicing

a. /papo/	[pa**bo**] 'fool'	V_V
b. /sakoŋ/	[sa**g**oŋ] 'boatman'	V_V
c. /kilta/	[kil**d**a] 'be long'	l_V
d. /ilki/	[il**g**i] 'journal'	l_V
e. /kamtok/	[kam**d**ok] 'a director'	N_V
f. /panti/	[pan**d**i] 'firefly'	N_V

Voiced stops cannot surface before a voiced (sonorant) consonant, because Korean does not allow stops to surface before a nasal or a liquid. As discussed in Chapter 2, an obstruent will be nasalized before a nasal (3a). Similarly, a stop is nasalized before a liquid (3b-c)[7].

(3) Nasalization in intersonorant position

a. /kukmul/	[ku**ŋ**mul] 'soup'	after a vowel before a nasal
b. /hapli/	[ha**m**ni] 'rational'	after a vowel before a lateral
c. /jəkli/	[jə**ŋ**ni] 'irrationality'	after a vowel before a lateral

Thus, the only context in which a voiced stop can surface in Korean is preceding a vowel. Therefore, if Korean listeners hear the stop as voiced, they are likely to hear an illusory vowel following a voiced stop since voicing can serve as a cue to the presence of a following vowel. The vowel [ɨ], the preferred epenthetic vowel in loanwords such as [t^hɛnisɨ] 'tennis', is the likely candidate for this illusory vowel.

---

[7] /l/ also becomes nasalized following a nasal, shown in these examples: /simli/→ [si**mn**i] 'mind', /jəŋli/→ [jə**ŋn**i] 'smart'.

The second hypothesis addresses the finding that Korean speakers were more likely to devoice bilabial than velar stops. Cross-linguistic research on the interaction of place and voicing in production support the hypothesis that bilabial closure induces bias toward perception of a stop as voiceless whereas velar place induces bias toward perception of a stop as voiced. Benki (2001) found that English listeners classified bilabial or alveolar prevocalic stops as voiceless more often than velar stops, independent of F1 transitions. Similarly, English speakers exhibited different categorical perception boundaries for voicing depending on the place of the stop (Lisker and Abramson, 1964), a difference that may reflect properties of the mammalian auditory system, since a similar effect was found for chinchillas (Kuhl & Miller, 1975). Although these studies examined voicing contrasts in initial, prevocalic contexts (as opposed to the postvocalic, prenasal context considered here), at least some of the acoustic cues to voicing are potentially similar. If the Korean participants were less likely to perceive [b] in prenasal context as voiced than [g], then the greater likelihood of devoicing of [b] and of vowel insertion after [g] is expected.

3.1.1 Method

In order to assess the perception hypotheses, two behavioral perception tests were carried out. The first test involved a discrimination task using a speeded AXB paradigm. Participants heard 3 stimuli from a continuum consecutively and were directed to indicate whether the second stimulus was more similar to the first stimulus or to the third stimulus. The second test involved a categorization task using stimuli from the same continuum. Participants were presented with a stimulus and directed to identify the stimulus as either one or the other choice given to them. An English group was also recruited as a control to determine whether the Korean responses were an effect of native language or reflected universal principles.

3.1.1.1 Participants

Twenty Korean and twenty English native speakers participated in the experiment. Participants received either course credit (for those recruited from the Psychology Department subject pool at Stony Brook University) or a small monetary incentive. All of the English native speakers were monolingual. 19 of the Korean-speaking participants completed the reading production experiment before the perception task, but the other participant did not.

3.1.1.2 Materials

The vowel [ɨ], which appears as the epenthetic vowel after stops in loanwords such as [tɛk^{h}ɨnik] ‘technique’ (Y. Kang, 2003), is a likely candidate for an illusory vowel. Therefore, the -CVC- items contained the vowel [ɨ]. The auditory stimuli used for the CVC endpoint of the vowel-no vowel continuum were obtained from a balanced Korean-English bilingual speaker, who could produce the Korean epenthetic vowel [ɨ] while otherwise maintaining English phonology. The auditory stimuli used for the voicing continuum (e.g., *igna* to *ikna*) were recorded from a male English monolingual speaker. The recordings were done at the sound treated room in the Linguistics Department at Stony Brook University.

Voicing Continua

When creating two voicing continua, *ibna-ipna* [ɪbnə-ɪpnə] and *igna-ikna* [ɪgnə-ɪknə], the factors of vowel duration, consonant duration and glottal pulsing percentage during closure were manipulated from the naturally produced tokens. These three acoustic cues were chosen because they have been claimed to be the primary cues to the American English syllable-final stop voicing contrast. Vowel duration has been argued to be the most influential cue in a number of perception studies (Hogan & Rozsypal, 1980; Jones, 2003; Luce & Charles-Luce, 1985; Raphael, 1972; Rodgers, 2007), where longer duration of the vowel has been shown to induce perception of a following stop as voiced. Longer voicing percentages and shorter closure duration also serve as secondary cues to stop voicing (Rodgers, 2007). F1 transitions have also been shown to signal voicing contrasts (Jones, 2003; Kingston & Diehl, 1995; Moreton, 2004; Nittrouer, 2004; Thomas, 2000), but this cue is most useful in the context where the F1 offset frequency can be large, as in the vowel [æ] where an F1 steady state is high (Fischer & Ohde, 1990). Formant transitions were not manipulated in these stimuli, since the vowel of the stimuli for this experiment was [ɪ], which has a lower F1 steady state, where the spectral cue is minimized.

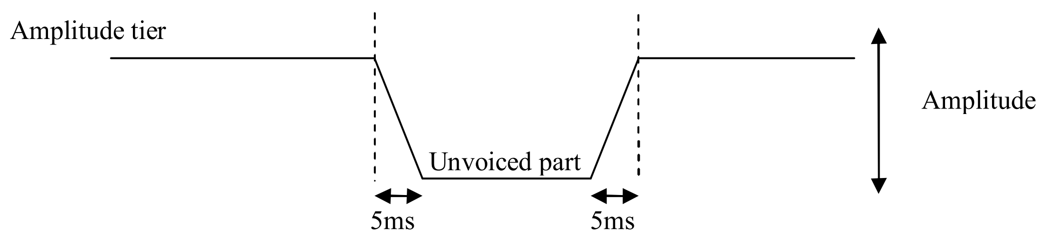
7 items were modified from a base file that was fully voiced. The total duration of vowel plus consonant was constant in all items but the proportion of vowel and closure length varied from 50%-50% (fully voiced stimulus 7) to 20%-80% (fully voiceless stimulus 1), as shown in (4). The percentage of glottal pulsing also changed from 100% (fully voiced stimulus 7) to 0% (fully voiceless stimulus 1).

(4) Acoustic cue manipulation on the voicing continua

Stimulus	Vowel	Closure	Glottal pulsing
1	20%	80%	0%
2	25%	75%	0%
3	30%	70%	20%
4	35%	65%	40%
5	40%	60%	60%
6	45%	55%	80%
7	50%	50%	100%

The manipulation of these cues was carried out in Praat (Boersma & Weenink, 2007). The base files were naturally produced 'igna' and 'ibna' where full voicing during the closure was observed. Vowel duration and consonant duration were manipulated using PSOLA, a resynthesis tool in Praat. The glottal pulsing percentage was modified by lowering the intensity of the voicing portion, rather than zeroing it out. When the amplitude of a desired duration of the closure was lowered, the amplitude of the first and the last 5 ms was linearly decreasing and increasing, as shown in (5). This method was chosen to avoid an unnatural abrupt change in amplitude, following the method in Hillenbrand et al. (1984).

(5) The amplitude manipulation for glottal pulsing in Praat



Finally, lengthening of the unvoiced portion during closure was edited by copying and pasting of the unvoiced part.

Vowel continua

The vowel continua *igna-igɨna* [ɪˈgnə-ɪˈgɨnə], *ikna-ikɨna* [ɪˈknə-ɪˈkɨnə], and *ibna-ibɨna* [ɪˈbnə-ɪˈbɨnə] were created by manipulating the vowel length of the items that contained an epenthetic full vowel. Dupoux et al. (1999) manipulated the vowel duration by splicing out the pitch periods at zero-crossings.[8] Following their method, six items for each continuum were created: stimulus 1 had no vowel (all transitions cut), and stimulus 6 had ten pitch periods. The glottal pulsing of the speaker who produced the materials was about 10ms, so the stimuli varied from no vowel to full vowel, in increments of about 20msec (two pitch periods). In addition to the vowel duration manipulation, the closure duration of the stop preceding the vowel was varied. With the vowel length modification only, there is a possibility that listeners could make a decision, especially in a discrimination task, based on the absolute length of the stimuli (which varied linearly in duration), treating them as acoustic noise rather than language. Also, the naturally spoken words *iCna* and *iCɨna* revealed that the closure duration of C in *iCna* was longer than the closure duration of C in *iCɨna*. The length of closure duration of stimuli 1, 2, and 3 was manipulated using PSOLA. The C duration of *iCna* was matched to the average duration of a corresponding *iCna* that was naturally spoken by the same speaker. Table 3.1 shows the actual duration of each consonant (b, g, k) on the continua. In order to examine whether this manipulation would induce significantly different responses from responses to the tokens with consonant duration manipulation, we included another 'no vowel' token (1a) that had no consonant duration change.

Table 3.1. Acoustic cue manipulation on the vowel continua

Stimulus	[b] duration [ɪˈbnə-ɪˈbɨnə]	[g] duration [ɪˈgnə-ɪˈgɨnə]	[k] duration [ɪˈknə-ɪˈkɨnə]	Epenthetic vowel
1a	60ms	50ms	90ms	0ms
1b	90ms	80ms	135ms	0ms
2	80ms	70ms	120ms	20ms
3	70ms	60ms	105ms	40ms
4	60ms	50ms	90ms	60ms
5	60ms	50ms	90ms	80ms
6	60ms	50ms	90ms	100ms

[8] Dupoux et al. (1999) examined the effect of vowel transitions that may have left in the 'no vowel' item in this experiment. They included another 'no vowel' token which was naturally spoken and compared it with the edited 'no vowel' token. They found no significant difference in results between them. Thus, we included the edited one only in our experiment.

3.1.1.3 Categorization tasks

For each continuum, participants were presented with 14 repetitions of each item on the continuum. There were a total of 490 trials (5 continua × 7 steps × 14 repetitions). For the voicing continua, participants were asked to classify the voicing of the stop before the nasal with a forced choice between ‘b’ and ‘p’ or between ‘g’ and ‘k’. For the vowel continua, they were asked to decide whether they did or did not hear a vowel between the stop and the nasal by pressing either ‘vowel’ or ‘no vowel’ buttons. The inter-trial interval was 1 second. After each block, participants were able to take a short break.

3.1.1.4 Discrimination tasks

Participants were tested on an AXB discrimination task, where the first and the third items in the AXB triplets were separated by 2 steps on the continuum. There were four possible combinations (AAB, ABB, BAA, and BBA). For the vowel continua there were 4 A-B pairs (0ms-40ms, 20ms-60ms, 40ms-80ms and 60ms-100ms) and for the voicing continua there were 5 A-B pairs (1-3, 2-4, 3-5, 4-6 and 5-7). The number of observations per triplet was 20. The inter-stimulus interval was 500 ms and the inter-trial interval was 2 seconds. Participants were run in 5 blocks, one block per continuum. The trials and blocks were randomized per participant. Participants had a short practice round before the actual task.

The experiments were carried out using SuperLab in a sound-treated room in the Department of Linguistics at Stony Brook University. Each experiment took about 40 minutes to complete.

3.1.1.5 Analysis

To assess the perception hypotheses, the accuracy data in the discrimination tasks and the classification rates in the categorization tasks were submitted to by-subject ANOVAs. Three analyses were conducted for the three critical comparisons based on the perception hypotheses and predictions from the discrimination and categorization task data.

As for *H1*, we hypothesized that voicing induces a bias toward perception of illusory vowels. This hypothesis predicted a lower percentage of ‘vowel’ responses in categorization and higher accuracy in discrimination for [kin-kɨn] than for [gn-gɨn]. To assess *H1*, the ANOVAs of percentage of ‘vowel’ responses in the categorization task was conducted with *Voicing* (gn-gɨn vs. kn-kɨn) and *Vowel Length* (0ms(a), 0ms(b), 20ms, 40ms, 60ms, 80ms vs. 100ms) as factors. Repeated measures ANOVAs of accuracy rates in the discrimination task were conducted with *Voicing* (gn-gɨn vs. kn-kɨn) and *Pair* (1-3, 2-4, 3-5 vs. 4-6) as within-subject factors.

For *H2,* we hypothesized that bilabial closure induces a bias toward perception of stops as voiceless. This predicted a lower percentage of ‘voiced’ responses in categorization and higher accuracy in discrimination for [pn-bn] than for [kn-gn]. To test this, the percentage of ‘voiced’ responses in the categorization task was submitted to a by-subject ANOVA with *Place* (pn-bn vs. kn-gn) and *Item* (1, 2, 3, 4, 5, 6, vs. 7) as factors. Accuracy in discrimination was submitted to a by-subject ANOVA with *Place* (pn-bn vs. kn-gn) and *Pair* (1-3, 2-4, 3-5, 4-6, 5-7, 6-8) as within-subject factors.

Given that *H1* and *H2* are supported, we predicted a lower percentage of 'vowel' responses in categorization and higher accuracy in discrimination for [bn-bɨn] than for [gn-gɨn], the percentage of 'vowel' responses in the categorization task was submitted to a by-subject ANOVA with *Place* (bn-bɨn vs. gn-gɨn) and *Vowel Length* (0ms(a), 0ms(b), 20ms, 40ms, 60ms, 80ms, 100ms) as factors. Accuracy was submitted to by-subject ANOVAs with *Place* (bn-bɨn vs. gn-gɨn) and *Pair* (1-3, 2-4, 3-5, 4-6) as factors.

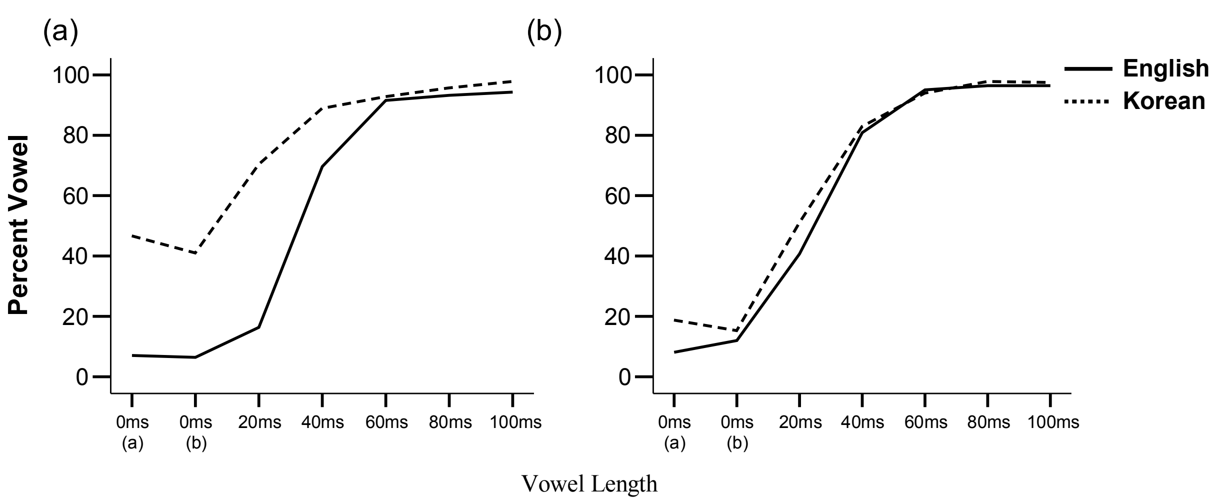
For all analyses, *Language* (Korean, English) was a between-subject factor. When sphericity (equal variances of the differences between levels of the repeated measures factor) was violated ($p < 0.05$), a Greenhouse–Geisser corrected degrees of freedom was used.

3.1.2 Results

3.1.2.1 *H1* : Voicing effect with vowel insertion ([gn-gɨn] vs. [kn-kɨn])

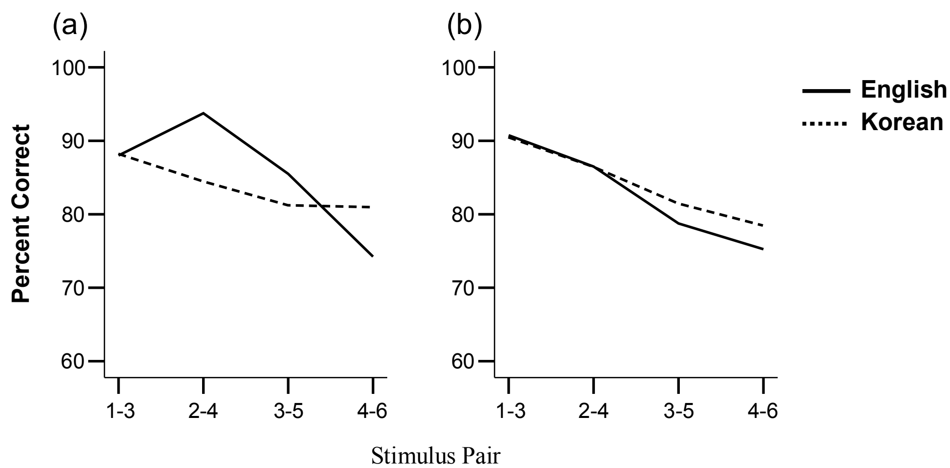
The mean percentages of 'vowel' responses are shown in Figure 3.1. The main effect of *Vowel Length* was significant ($F(3.10, 117.63)=310.48$, $p<.001$), indicating that as the vowel duration became longer, 'vowel' responses increased overall. The *Language* effect was also significant ($F(1, 38)=19.63$, $p<.001$), reflecting significantly more 'vowel' responses by Korean listeners than by English listeners. Similarly, the interaction between *Language* and *Voicing* was significant ($F(1, 38)=33.20$, $p<.001$). This was driven by the fact that both groups responded in a similar way to [kn-kɨn] whereas Korean listeners gave more 'vowel' responses to [gn-gɨn]. The three-way interaction between *Language*, *Voicing* and *Vowel Length* was significant as well ($F(6, 228)=8.41$, $p<.001$). This shows that for the continuum [gn]-[gɨn], but not [kn-kɨn], Korean participants tended to report that they heard a vowel even when there was no vowel in the stimulus, whereas English participants tended to report that they did not hear a vowel. The frequency of 'vowel' responses to the two different versions of [gn] and [kn] (1(a) or 1(b) on the continua) did not differ significantly. This suggests that manipulation of consonant duration did not affect the participants' responses.

Figure 3.1. Mean percentages of 'vowel' responses of English and Korean speakers in the two continua: (a) in [gn- gɨn] and (b) in [kn-kɨn].



The discrimination accuracy results pooled across participants are presented in Figure 3.2.

Figure 3.2. Mean discrimination accuracy of English and Korean in the two continua: (a) in [gn-gɨn] and (b) in [kn-kɨn].



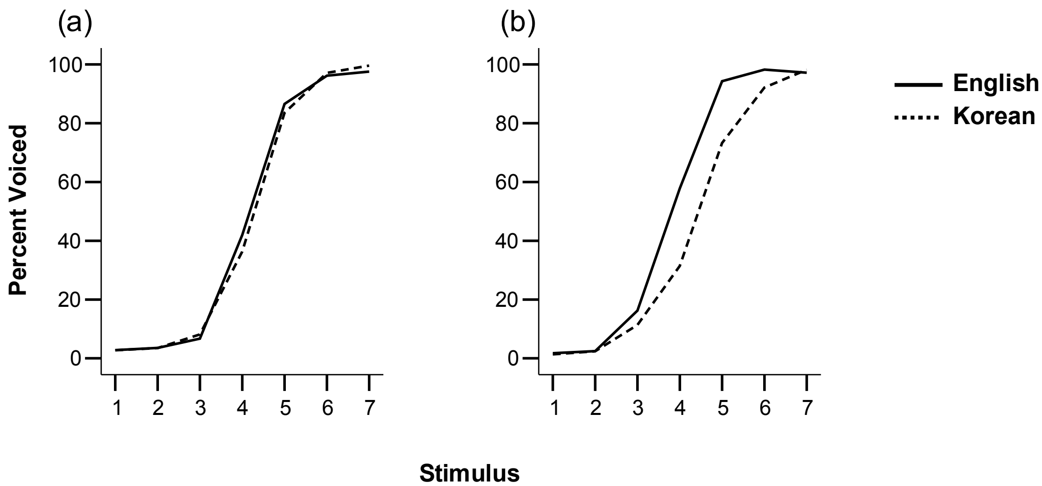
The main effect of *Language* was significant ($F(1, 38)=21.07$, $p<.001$), indicating that English listeners were more accurate than Korean listeners. The interaction effect between *Language* and *Pair* was also significant ($F(3, 114)=3.72$, $p<.05$). This suggests that Korean listeners were less accurate for certain pairs than were English listeners. The significant three-way interaction between *Language*, *Pair* and *Continuum* ($F(3, 114)=2.72$, $p<.05$) indicates that Korean native speakers were particularly less accurate for a particular pair in [gn-gɨn], as illustrated in Figure 3.2. Post hoc analyses revealed that pair 2-4 in [gn-gɨn] was the pair on which Korean listeners were significantly less accurate than English listeners ($p<.05$).

The results from both categorization and discrimination support *H1* in that Korean listeners' perception of the continua containing stimuli from no vowel to full vowel significantly depended on the voicing of the internal stop. Specifically, Korean listeners were less accurate than English listeners for [gn-gɨn] but not for [kn-kɨn]. Moreover, when the stop was voiced, Korean native speakers tended to report that they heard a vowel even when there was no vowel in the stimuli, whereas English native speakers did not.

3.1.2.2 *H2* : Place effect with devoicing ([pn-bn] vs. [kn-gn])

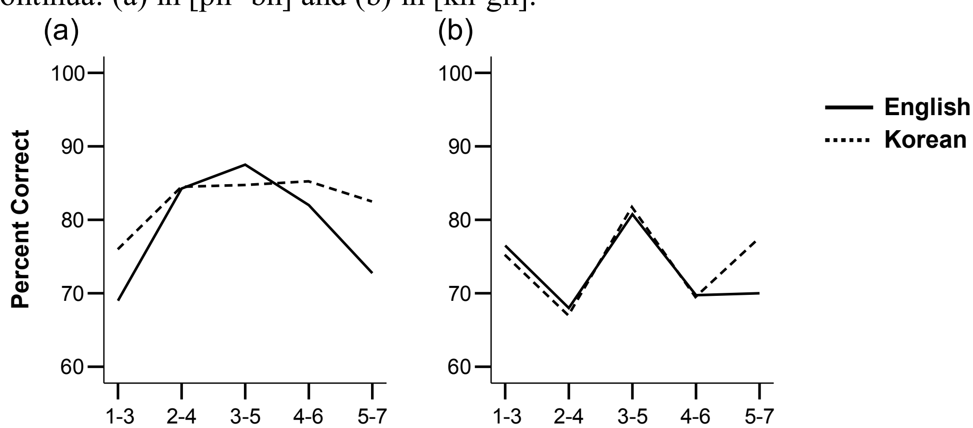
The mean percentage of vowel responses pooled across participants is plotted in Figure 3.3 for [pn-bn] and [kn-gn]. There was a significant main effect of *Item* ($F(2.50, 94.78)=995.41$, $p<.001$), indicating that participants provided more 'voiced' responses toward the end of the continua. A significant *Language* effect was found ($F(1, 38)=6.47$, $p<.05$), reflecting that Korean listeners gave fewer 'voiced' responses than English listeners. The significant interaction between *Language* and *Place* ($F(1, 38)=26.55$, $p<.001$) suggests that Korean listeners gave fewer 'voiced' responses than English listeners for the [kn-gn] continuum. The significant three-way interaction between *Language*, *Item* and *Continuum* ($F(6, 228)=5.33$, $p=.001$) indicates that Korean listeners' 'voiced' responses were fewer than English listeners' only for certain items on [kn-gn]. These results are actually in the opposite direction from the predictions made for *H2*, which predicts that listeners should be more likely to identify items on the [pn-bn] continuum as voiceless, leading to the more frequent devoicing in [bn]. In Chapter 6 I argue that this effect is a result of articulatory difficulty with voicing in labial stop-nasal sequences.

Figure 3.3. Mean percent of 'voiced' response by English and Korean in the two continua: (a) in [pn- bn] and (b) in [kn-gn].



In the discrimination task, the main effect of *Pair* was significant ($F(3.25, 123.33)=9.84$, $p<.001$). Pairwise comparisons (Bonferroni correction) revealed that the 3-5 pair was discriminated significantly more accurately than the rest of the pairs ($p<.005$ for all). The main effect of *Place* was significant ($F(1, 38)=35.06$, $p<.001$), indicating higher accuracy for [pn-bn] than for [kn-gn] overall. Critically, the interaction between *Language* and *Place* was not significant ($F(1, 38)=.88$, n.s.). Also, the three-way interaction of *Language*, *Place* and *Pair* was not significant ($F(4, 152)=.94$, n.s.). *Pair* and *Language* interaction was significant ($F(4, 152)=2.59$, $p<.05$), driven by Korean listeners' generally higher accuracy for the 1-3 pair and the 5-7 pair than for the other pairs ($p<.001$). Overall, English and Korean listeners did not present significantly different accuracy in any context except that Korean listeners were more accurate than English listeners for the pairs at the ends of the continua.

Figure 3.4. Mean discrimination accuracy of English and Korean participants in the two continua: (a) in [pn- bn] and (b) in [kn-gn].



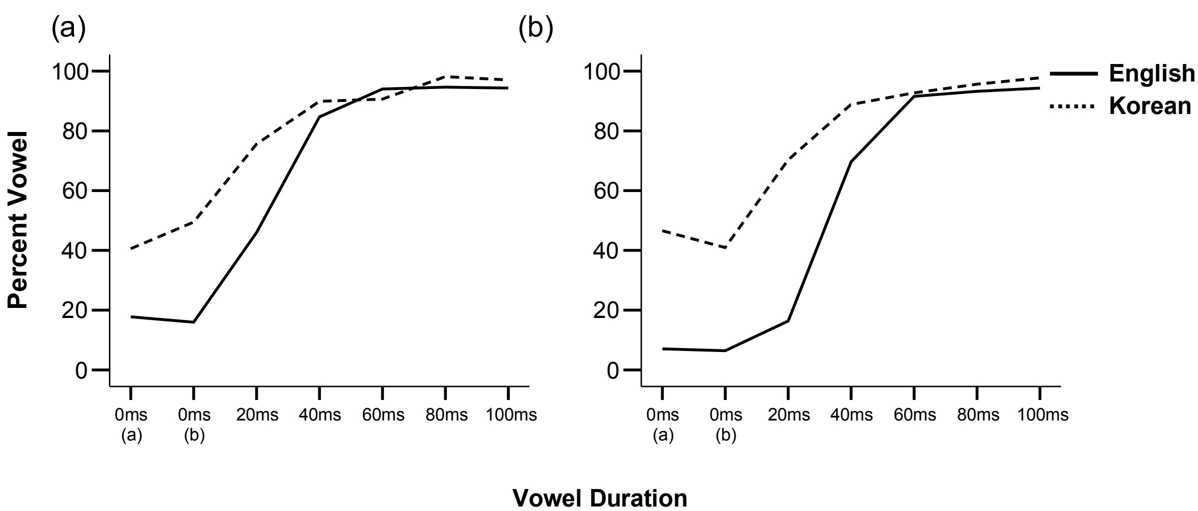
Stimulus Pair

H2, the hypothesis that bilabial closure induces a bias toward perception of the stop as voiceless, was not supported by the results we found in this section. [pn-bn] was not harder than [kn-gn] to distinguish for Korean listeners and between-items on the [kn-gn] continuum were more often classified as voiceless than were items on the [pn-bn] continuum. This is discussed further in Chapter 6.

3.1.2.3 *H3* : Place Effect with Vowel Insertion ([bn-bɨn] vs. [gn-gɨn])

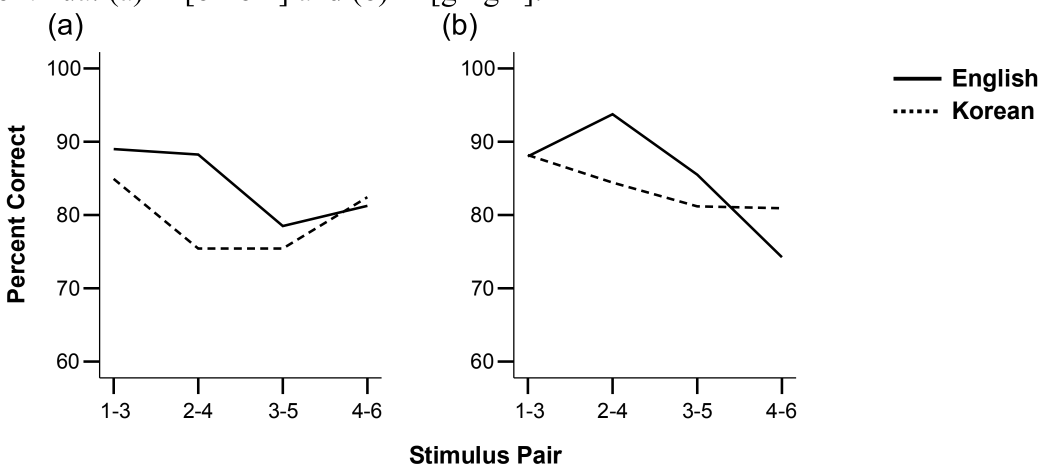
H3 hypothesized that velar closure induces a bias toward perception of an illusory vowel for Koreans. The categorization results pooled across participants are plotted in Figure 3.5. The main effect of *Vowel Length* was significant ($F(2.32,\ 88.2)=159.82$, $p<.001$), reflecting more vowel responses as the vowel length increased. *Language* was significant ($F(1, 38)=24.06$, $p<.001$), indicating that Korean listeners reported hearing a vowel significantly more often than English listeners. We found a significant two-way interaction between *Place* and *Language* ($F(1, 38)=6.37$, $p<.05$). This indicates that the difference between the two language groups was larger for [gn-gɨn] than for [bn-bɨn]. The three-way interaction between *Language*, *Place*, and *Vowel Length* was also significant ($F(6, 228)=2.91$, $p<.01$). This effect was driven by English listeners who gave significantly more 'vowel' responses to the stimuli with no vowel or the stimuli with 20 or 40 ms vowel length when they were from the [bn-bɨn] continuum than from the [gn-gɨn] continuum. On the other hand, Korean listeners did not show significantly different response patterns between the two continua. These results suggest that for English listeners, the category boundary is earlier for [bn-bɨn] than for [gn-gɨn], whereas for Korean listeners, it is about the same. But for both continua, Korean listeners gave significantly more 'vowel' responses than English listeners.

Figure 3.5. Mean percent of vowel response by English and Korean speakers in the two continua: (a) in [bn-bɨn] and (b) in [gn-gɨn].



The discrimination results pooled across participants are plotted in Figure 3.6. The main effect of *Pair* was significant ($F(3, 114)=10.84$, $p<.001$). Pairwise (Bonferroni) comparisons revealed Pairs 1-3 and 2-4 were discriminated more accurately than the other pairs ($p<.01$ for all). Neither the two-way interaction between *Language* and *Place* ($F(1, 38)=1.45$, n.s.) nor the three-way interaction between *Language*, *Place* and *Pair* ($F(3, 114)=43.54$, n.s.) was significant. This means that the accuracy of the two language groups for the two continua was not significantly different. The only significant interaction effect associated with *Language* was found in *Language* × *Pair* ($F(3, 114)=763.96$, $p<.001$). This indicates that regardless of the continuum, English listeners were more accurate than Korean listeners on some pair(s), and post hoc analysis showed that English listeners were significantly more accurate on Pair 2-4 than Korean listeners ($p<.001$).

Figure 3.6. Mean of discrimination accuracy of English and Korean participants in the two continua: (a) in [bn-bɨn] and (b) in [gn-gɨn].



The results from both categorization and discrimination experiments did not support *H3*, which was predictable given that *H2* was not supported.

3.1.3 Voicing effect but not place effect in perception

The results from the categorization and the discrimination tasks revealed a clear language-specific effect for voiced stop-nasal vs. voiced stop-ɨ-nasal stimulus pairs, supporting the hypothesis that voicing of the stop induces perception of an illusory vowel for Korean listeners but not for English-speaking participants. Korean participants reported they perceived a vowel significantly more often when the stop was voiced than voiceless, while English participants' perception of a vowel did not depend on the voicing of the stop. This language difference in categorization suggests that Korean listeners' perception is affected by their native language in terms of their NL phonetic categories and that the frequent vowel insertion in voiced stop-nasal sequences (but not voiceless stop-nasal sequences) in their production reflects their misperception. I conclude that even in the absence of a release burst for the stop, a voiced stop (but not a voiceless stop) is interpreted as a voiced stop followed by [ɨ]. In other words, because

Korean voiced stops are always followed by a vowel, stop voicing is interpreted as a cue for the presence of a vowel following the stop.

This voicing asymmetry in Korean L2 speakers' perception also informs us that simple illegality is not the source of the illusory vowel in perception, at least for Korean listeners, as Korean listeners' categorization of [kn-kɨn] was very similar to English native speakers'. In Dupoux et al.'s (1999) experiments, Japanese listeners were much more likely than French listeners to hear a vowel in stimuli such as [ebzo] or [egdo]. Dupoux et al. explain this language effect with reference to the fact that these consonantal contacts never occur in Japanese, and propose that L2 perception is deeply affected by knowledge of native language phonotactics. Similarly, Boersma and Hamann (2009) propose a constraint (SYLCON) that bans stop-nasal contact in the Korean perception grammar to explain why a vowel is inserted in 'picnic' but not in 'chapter' in loanword adaptation. Since [kn] sequences violate a structural constraint SYLCON, a vowel is inserted in 'pi[kn]nic' but not in 'cha[pt]er' in perceptual mapping. However, Korean listeners' categorization and discrimination of [kn-kɨn] stimuli pair in the current experiments do not support this proposal, as their perception (both discrimination and categorization) was highly similar to that of English native speakers. If it is Korean listeners' phonotactic knowledge that determines perception of an illusory vowel, it is expected that they should perceive a vowel in both [gn] and [kn], which was not the case. Then, what can explain the difference between 'picnic' and 'chapter'? I will discuss this in detail in Chapter 5 in relation to the perception of a release burst in the Korean perception grammar (Kabak & Idsardi, 2007).

With respect to the second and third misperception hypotheses, we failed to find evidence to support a perceptual explanation for the place asymmetry. We hypothesized that more frequent devoicing of /b/ and more frequent vowel insertion with /g/ in production indicated a perceptual bias to hear /b/ as voiceless more often than /g/. However, this was not supported by the experimental results. Instead, it was found that Korean speakers' categorization of the voicing contrast was very similar to that of English native speakers, especially for bilabial stops. In other words, the high rate of devoicing of /b/ in stop-nasal sequences cannot be a result of a tendency to misperceive /b/ as /p/. As a consequence, the source of the place effect with vowel insertion and devoicing cannot be Korean speakers' tendency to categorize /gN/ as /gɨN/ and /bN/ as /pN/. In Chapter 6, I explore another possible source of the place asymmetry.

The behavioral experimental results provided evidence for the hypothesis that voicing of the stop induces perception of an illusory vowel for Korean listeners but not for English-speaking participants. However, what is not clear is the level at which this misperception occurs. The next section reports on Event-Related Potentials (ERP) experiments designed to examine whether perception errors result from the failure of learners to hear the acoustic differences between different foreign language sounds, or simply from their failure to realize that these differences are relevant to linguistic categorization.

3.2 Electrophysiological Experiments

The event-related potential (ERP) technique is a non-invasive method of measuring electrical activity in the brain during cognitive processing (Luck, 2005). In contrast to the behavioral methods that require conscious decisions, ERPs can be used to investigate pre-attentive processes involved in cognitive tasks, such as speech perception. Cross-linguistic studies have used the ERP methodology to examine speakers' pre-conscious detection of stimulus change (Winkler et al., 1999; Sharma and Dorman, 2000).

To examine whether Korean listeners' illusory vowel perception in voiced stop-nasal sequences resulted from their failure to hear acoustic differences between [gn] and [gɨn] or from their miscategorization of [gn] as [gɨn], we employed an oddball paradigm in an ERP experiment. Mismatch negativity (MMN) is an auditory ERP component, typically elicited in an auditory oddball paradigm where repeated stimuli (the 'standard') are occasionally replaced with rare 'deviant' stimuli that differ from the standard in acoustic parameters such as frequency, duration or intensity (Näätänen, 2001; Winkler, Reinikainen, & Näätänen, 1993 among many others). It has been suggested that MMN is an attention-independent, automatic component (Winkler et al., 1993) that occurs 150-200 ms after the onset of deviance.

Early works on MMN mostly showed that MMN utilizes auditory sensory representations, but recent findings suggest that it also reveals language-specific effects of categories. For example, Nenonen et al. (2003) found that native speakers of Finnish displayed a larger MMN to a native vowel length contrast than non-native speakers of Finnish, while both groups showed similar-sized MMNs to non-speech stimuli that involved the same durational difference as the speech stimuli. Similarly, Näätänen et al. (1997) found that Finnish participants showed reduced MMN responses to a deviant that is a non-prototype(within-category) vowel in Finnish whereas Estonian participants showed larger MMNs to the same deviant that is a proto-type vowel (across-category) in Estonian. Similarly, language-specific MMNs have been found in more complex speech stimuli where the deviance was in non-initial position. Dehaene-Lambertz et al.'s (2000) cross-linguistic study employing an ERP method for perception of an illusory vowel between two consonants demonstrated an early impact of native language phonotactics, showing that Japanese listeners did not display any significant MMN response to the contrast (e.g., *igmo* vs. *ig**u**mo*) while French listeners did. A number of studies have shown similar language-specific effects where MMN responses are generated by contrasts that are linguistically significant in the listeners' native language, but not by non-contrastive changes (Dehaene-Lambertz, 1997; Dehaene-Lambertz, Dupoux & Gout, 2000; Sharma & Dorman, 2000).

Although the studies reported above present evidence that non-native (or within-) category differences are not perceived or are less robustly perceived than differences across native language categories, conflicting evidence also exists. For example, in Ylinen et al.(2006)'s cross-linguistic study, MMN responses to within-category changes and across-category changes in vowel duration did not differ significantly from each other in either Finnish native speakers or L2/non-native users of Finnish, although Finnish native speakers generally showed larger MMNs regardless of phonemic status of the categories. Ylinen et al. suggest that the difference in MMN amplitude may reflect a prototypicality effect, rather than a language effect—that is, Finnish native speakers showed overall enhanced MMN responses regardless because the stimuli were their native language prototypes. Another example comes from Rivera-Gaxiola et al.'s (2000) study where English native speakers' MMNs were obtained for pairs /ba/-/d̪a/(across-category), /d̪a/-/ɖa/ (across-category for Hindi speakers but not English speakers) and two /ba/ stimuli (within-category). All stimulus pairs involved the same acoustic distance on a continuum from labial /ba/ to retroflex /ɖa/ through dental /d̪a/. English-speaking listeners displayed significant MMNs to all pairs in the experiment including the /d̪a/-/ɖa/ pair, which is essentially within-category for English native speakers. But English speakers behaviorally categorized both /d̪a/ and /ɖa/ as alveolar.

Given inconsistent results across studies, it is not clear how non-native or within-category contrasts are represented at the auditory/low-phonetic level. Experimental evidence showing a loss of perceptual sensitivity to non-native contrasts suggests that language experience affects a low level of perceptual processing (Kuhl, 2004) while evidence showing intact perceptual sensitivities to non-native categories suggests language experience affects a higher level of speech perception. This may depend on the kinds of acoustic cues available to the phonetic contrasts, as suggested by Phillips (2001). Phillips suggests that brain representation of speech sounds differs depending on the kinds of phonetic categories: a native language effect at a higher level is supported better by contrasts involving difference in place of articulation than contrasts involving differences in vowel qualities or voicing (see Phillips, 2001 for detailed discussion).

The goal of this section is to examine whether brain responses to the contrasts involving absence/presence of a vowel in stop-nasal clusters can be found, and if so, whether native language experience affects the magnitude of these responses.

3.2.1 ERP 1

Following on the finding of a robust voicing effect for Korean listeners in the behavioral perception experiments, an ERP experiment was conducted to investigate whether a language effect would be found in neurophysiologic responses. Whether a language-specific effect is found or not provides an indication of the level at which misperception takes place for this particular contrast.

3.2.1.1 Participants

10 Korean and 10 English native speakers with normal hearing participated in the experiment. Korean participants' average length of stay in the U.S. was 8.9 months. Participants received a small payment or class credit for their participation.

3.2.1.2 Materials

To examine the MMN responses of the two language groups, we employed the pair (20ms-60ms) where the discrimination accuracy was maximal for the English group on the [gn-gɨn] continuum in the behavioral experiment. Another pair (20ms-60ms) from the [kn-kɨn] continuum was also employed; this pair involved the same acoustic distance as the [gn-gɨn] continuum, but was predicted to show no difference between English and Korean listeners, based on the behavioral experiment results. The item with 20ms vowel duration served as the repeated 'standard' (85%) in half of the blocks and as the rare 'deviant' (15%) in the other half of the blocks. Each item occurred randomly in three different versions from the same speaker: the stimuli were physically different, but shared acoustic parameters such as the duration of each segment in the stimulus. The inter-stimulus interval (ISI) was 500 ms and there were at least two standard stimuli preceding a deviant stimulus. The total of 8 blocks was presented randomly per subject.

3.2.1.3 EEG recording procedures and data analysis

During the electroencephalogram (EEG) recording, the subjects sat in a sound attenuating chamber and watched a silent movie while the auditory stimuli were presented binaurally via headphones (75dB). Continuous EEG was recorded using a 64 (Ag/AgCl) electrode cap and a Neuroscan Inc. data acquisition system with a fronto-central electrode as ground and electronically linked mastoid electrodes as reference. Eye activities were monitored from electrodes at the outer canthi of the eyes and from electrodes above and below the orbital region of the left eye. Impedances for all electrodes were maintained below 10KΩ. All recorded signals were bandpass filtered on-line between 0.1 Hz and 30 Hz, digitized at 500 Hz and amplified with a gain of 1000.

Non-typical extreme movements were inspected visually and removed from continuous EEG data. Then components indicating typical eye artifacts were identified and extracted via independent component analysis (ICA) algorithm (runica) of EEGLAB (Delorme & Makeig, 2004) for each participant. The corrected data were epoched[9] from -200 ms to 700 ms and baseline-corrected. Epoched data were separately averaged for standard and deviant stimuli in each pair. The first 5 trials in each block were excluded from averaging. Difference waveforms were obtained by subtracting the responses to the standard stimuli from the responses to those elicited by the identical stimuli when they were the deviant stimuli in other blocks. This was done to avoid subtracting responses to standards in one block from responses to physically different deviants in the same block (Jacobsen & Schröger, 2003). From those individual difference waveforms, mean MMN amplitudes were calculated using a 100 ms window centered at the negative peak latency, based on the grand average difference waveforms of each condition for each group. The presence of the MMN in each condition/group was tested by two-tailed one-group *t*-tests of the difference mean amplitude calculated at Fz[10].

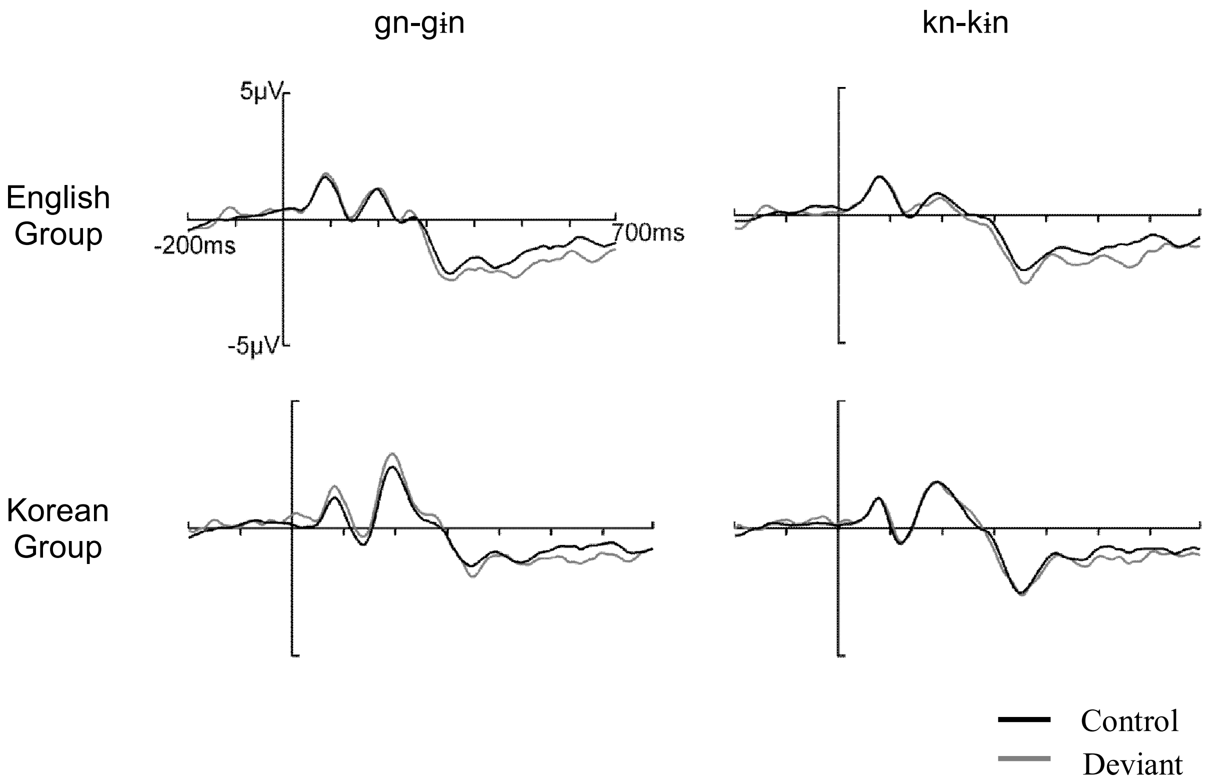
3.2.1.4 Results

Surprisingly, none of the *t*-tests of the difference mean amplitude at Fz was significant, displaying no significant MMNs in any condition or group. Figure 3.7 presents ERPs of standards and deviants in [gn-gɨn] and [kn-kɨn] from each language group. The black line represents responses to a stimulus when it served as standard, and the grey line represents responses to the same stimulus when it served as deviant. No MMN was elicited in [gn-gɨn] or [kn-kɨn] in any language group. Because no MMN was detected, no further analysis was carried out.

[9] A time interval from -200 ms to 700 ms around the onset of the stimuli is extracted from raw EEG signals for averaging.

[10] Fz is a middle-frontal electrode site.

Figure 3.7. ERPs of standards and deviants for change in *gn-gɨn* and *kn-kɨn* (20ms vs. 60ms)



3.2.1.5 Discussion

The lack of MMN responses in any condition is somewhat surprising, given that in the behavioral discrimination task, English and Korean listeners' accuracy rate for the pairs involving 20ms and 60ms vowels was above 80% at the lowest. As many studies have shown, behavioral discrimination responses generally parallel MMN responses (Jaramillo, Paavilainen, & Näätänen, 2000), with higher behavioral accuracy correlated with larger MMN amplitudes. There are several possible explanations for why the studies discussed here showed a difference between the results for the behavioral vs. the ERP studies. First, the participants in the behavioral and ERP studies discussed above were distinct. Second, in the discrimination/categorization experiments, participants had to pay attention to the stimuli to make a conscious decision about them, while in the ERP experiment, they were directed to watch a movie, without having to attend to the stimuli. Strange and Shafer (2008) suggest that some phonetic distinctions are so difficult that they may require attention to elicit MMN. And third, a possible reason for the absence of MMNs could be that a 40ms difference in vowel duration is not large enough to elicit MMN in a passive oddball paradigm, especially in an experiment in which complex speech sounds are used. Some other studies have reported similar differences in behavioral and ERP results (Hisagi, 2007; Shafer, Schwartz, & Kurtzberg, 2004). For example, in Shafer and colleagues' (2004) cross-linguistic study of Hindi and English speakers' perception of place of articulation features, Hindi speakers discriminated the dental vs. retroflex voiced stops in a behavioral task, as expected because it is their native contrast, but an MMN was not elicited to the same contrast in the ERP experiment.

The results here seem to contradict the findings of a similar study by Dehaene-Lambertz et al. (2000), who found a robust language-specific effect in MMN responses to pairs such as *igumo* and *igmo*. Japanese listeners whose language does not allow complex clusters did not display significant MMNs to the deviance between *igumo* and *igmo*, while French listeners showed significant MMN responses to the same deviance, reflecting the fact that French allows such clusters. However, the stimuli used in Dehaene-Lambertz et al.'s study contained either a full vowel or no vowel whereas the stimuli in the current study contained a 20ms vowel or a 60ms vowel. Thus, the acoustic difference in this study may simply have been too small to elicit a clear MMN.

In the second ERP experiment, two changes were made. First, to rule out the possibility that the acoustic difference in the first ERP experiment was simply too small to elicit a meaningful MMN response, the acoustic distance between standard and deviant stimuli was increased. Second, to facilitate comparison of behavioral and ERP responses, the same participants were included in both behavioral and ERP studies.

3.2.2 ERP 2

In the first ERP experiment, an acoustic difference of 40 ms between standard and deviant was used. In the second experiment, the acoustic distance between standards and deviants was increased to100 ms, in line with the difference in the Dehaene-Lambertz et al. study. To investigate the relationship between the behavioral and ERP responses, a categorization task was completed after the EEG recording.

The categorization experiment (reported in §3.1.2.1) showed the voicing effect in Korean listeners' perception of the stop-nasal sequences: for Korean listeners, a 'vowel' response was elicited for approximately 40% of vowelless (*igna*) stimuli, and about 97% of the stimuli containing a 100 ms-long vowel between the stop and the nasal. On the other hand, English listeners gave the 'vowel' response for only about 6% of the vowelless stimuli vs. about 94% for stimuli containing the 100 ms vowel. Since English listeners showed a larger categorization response difference (about 88%) between *igna* and *igɨna* than Korean listeners (about 57%), we expected that English listeners would show a larger MMN for the *igna-igɨna* pair than Korean listeners. But when the stimulus was *ikna-ikɨna*, the behavioral responses of the two groups were not significantly different: English listeners' showed 12% 'vowel' responses for *ikna* and 96% for *ikɨna* while Korean listeners' responded 'vowel' for 14% for *ikna* stimuli and 97% for *ikɨna*. Thus, we expect that the two groups should not show significantly different MMN amplitudes to the *ikna-ikɨna* difference in the ERP experiment.

3.2.2.1 Participants

11 English native speakers (3 female, 8 male) and 10 Korean native speakers (5 female, 5male) participated in the experiment. All were right-handed and none reported hearing problems. All of the English native speakers reported that they had taken Spanish classes for 2-3 years in high school but that the only language they speak and understand is English. Korean speakers had

resided in the United States for 11.4 months on average. Participants were given course credit or a small payment for their participation.

3.2.2.2 Materials

The experimental materials used for the ERP experiment were two pairs, *igna-igɨna* and *ikna-ikɨna*, as in the first ERP experiment. The stimuli differed from those in the first experiment in that one member of the pair had no vowel and the other contained a full vowel (100 ms). Except for this difference in the stimuli, the rest of the design was identical to the first ERP experiment.

3.2.2.3 EEG recording procedures and data analysis

The EEG recording was carried out as in the first ERP experiment. After the recording, non-typical extreme movements were inspected visually and removed from continuous EEG data. Then, the corrected continuous data were epoched from -200 ms to 700 ms and baseline-corrected. Next, components indicating typical eye artifacts were identified and extracted via the independent component analysis (ICA) algorithm (runica) of EEGLAB (Delorme & Makeig, 2004) for each participant. Next, any epochs with voltages exceeding ±100 μV at any electrode were excluded from averaging. Averaged data, difference waveforms and MMN amplitudes were obtained following the analysis procedure as in the first ERP experiment.

The presence of an MMN in each condition/group was tested by two-tailed one-group *t*-tests of the difference mean amplitude calculated at Fz. The mean MMN amplitudes from electrode sites F1-F4, FC1-FC4, CP1-CP4 and P1-P4 were submitted to repeated measure analysis of variance (ANOVA) with Anteriority (Anterior vs. Posterior), Laterality (Left vs. Right) and Category (Voiced pair vs. Voiceless pair) as within-subject factors and Language (Korean vs. English) as a between-subject factor.

3.2.2.4 Categorization task

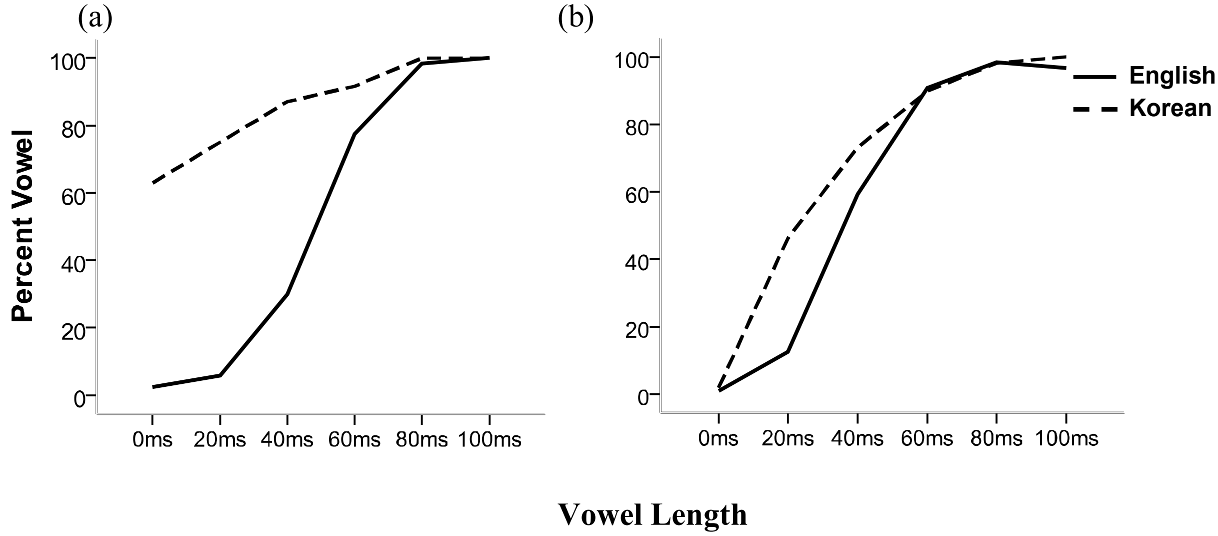
Directly following the EEG recording, subjects completed a categorization task in which they were asked to indicate whether they heard a vowel between the stop and the nasal. One member of each group failed to complete the task due to an equipment failure. The same materials were used as in the previous categorization experiment: two continua (*igna-igɨna* and *ikna-ikɨna*) varying in vowel length from 0 ms to 100 ms at 20 ms intervals. The procedure and the instructions in the categorization task were identical to those of the first categorization task.

3.2.2.5 Results

3.2.2.5.1 Categorization results

The mean percentages of 'vowel' responses as a function of vowel length and language for each continuum are plotted in Figure 3.8.

Figure 3.8. Mean percentages of 'vowel' response of English and Korean in the two continua: (a) in [gn- gɨn] and (b) in [kn-kɨn].



The results are very similar to the previous categorization results. The main effect of *Vowel Length* was significant ($F(2.14, 36.41)=137.14$, $p<.001$), indicating increased 'vowel' responses as the length of the vowel in the stimuli became longer. The *Language* effect was significant ($F(1, 17)=18.03$, $p<.005$), reflecting significantly more 'vowel' responses by Korean listeners than by English listeners. The interaction effect between *Language* and *Continuum* was significant ($F(1, 17)=36.25$, $p<.001$). As before, this was driven by the fact that both groups responded in a similar way to [kn-kɨn] whereas Korean listeners gave more 'vowel' responses to [gn-gɨn]. The three-way interaction between *Language*, *Continuum* and *Vowel length* was also significant ($F(3.18, 54.07)= 10.14$, $p<.001$). This again suggests that Korean participants heard a vowel even for the stimuli that had a shorter vowel or no vowel, especially when the continuum was [gn-gɨn]. The categorization experiment was replicated, showing a robust language-specific effect of voicing in perception of stop-nasal sequences.

3.2.2.5.2 ERP results

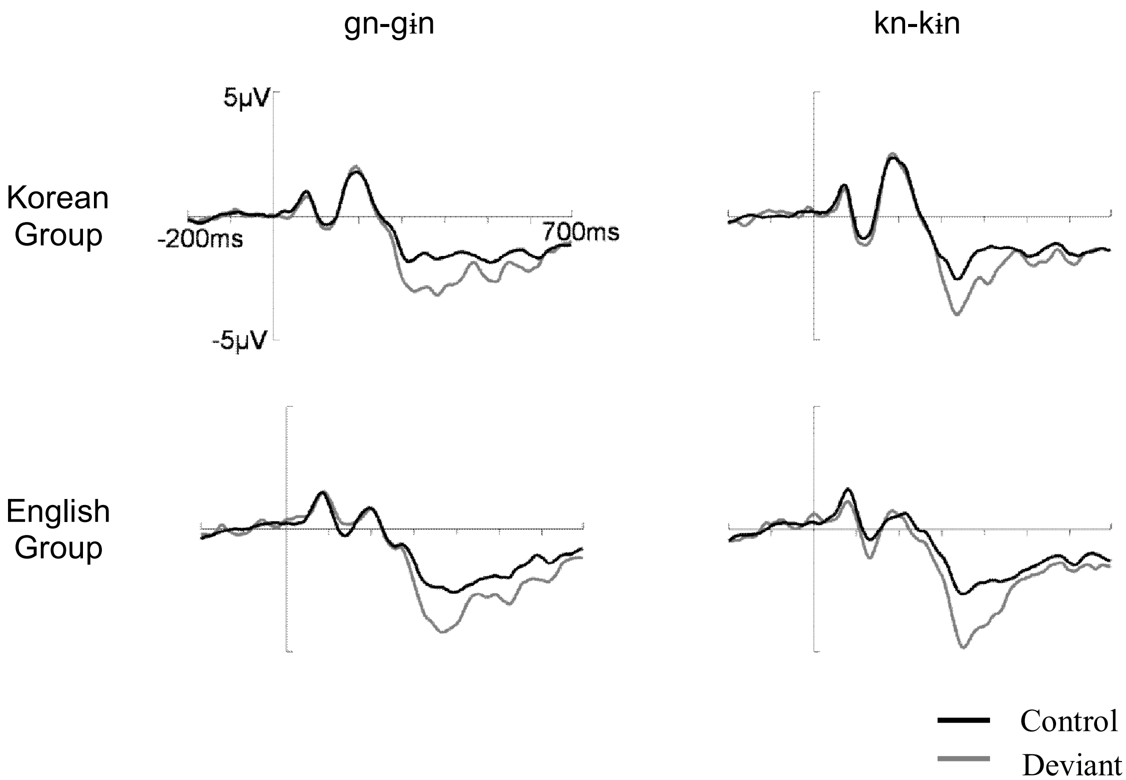
The grand-averaged ERPs are plotted in Figure 3.9. Both stimulus pairs elicited an MMN (Table 3.2) in both subject groups.

Table 3.2. Mean amplitude (μV) of MMN at Fz and p-value from a one-way *t*-test

	gn-gɨn (0 ms vs. 100 ms)	kn-kɨn (0 ms vs. 100 ms)
English group	-1.45 p<.005	-1.73 p<.005
Korean group	-1.43 p<.001	-1.15 p<.005

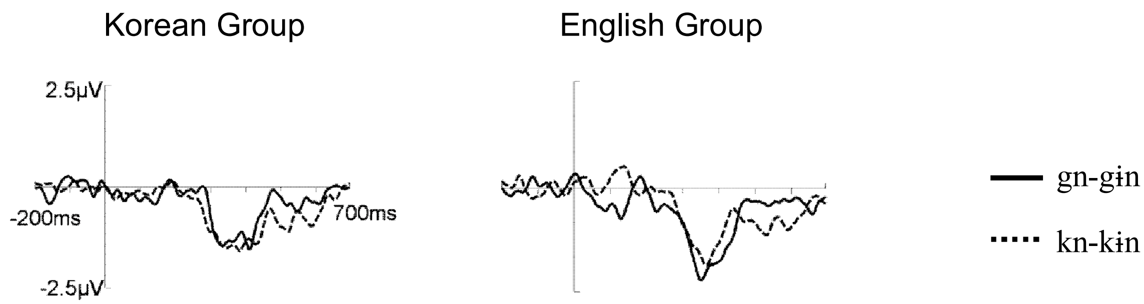
From the ANOVAs, a significant main effect of *Anteriority* was confirmed (F(1, 19)=51.06, p<.001), consistent with the typical MMN topography of the observed MMNs in the literature, where frontally recorded responses are stronger than parietally recorded ones. However, that was the only significant effect found in the statistical analyses. There was no significant effect of *Language* or significant interaction between *Language* and *Category*, indicating that the two groups' MMNs were very similar regardless of the stimulus type.

Figure 3.9. ERPs of controls and deviants in *gn-gɨn* and *kn-kɨn* stimulus pairs (0 ms vs. 100 ms) in Korean group and English group



Difference waveforms are presented in Figure 3.10. As shown in grand-averaged ERPs, the difference waveforms show neither significant differences between language groups nor stimulus types.

Figure 3.10. Difference waveform (grand-average deviant-minus-control difference waves) elicited by a 100 ms vowel duration difference in *ɪgnə- ɪgɨnə* and *ɪknə-ɪkɨnə* stimulus pairs in Korean group and English group



3.2.2.6 Discussion

Nonparallel results between categorization and MMN

MMN was elicited by larger acoustic distance (0 ms vs. 100 ms, or no V vs. full V) between standards and deviants in the second ERP experiment. But the size of MMN did not depend on the language group, voicing of the stop in the stimuli, or the interaction of these two factors. English native speakers displayed significant MMNs in all conditions, as expected. However, Korean speakers also displayed large MMNs even when the stimulus pair was *igna-igɨna*, suggesting that voicing of the stop did not affect their pre-attentive discrimination of the pair, unlike their categorization. This means that Korean and English participants' preattentive detection of acoustic differences between no vowel and vowel in stop-nasal sequences was fairly accurate for both language groups, regardless of the voicing of the stop in the stimulus pairs. What is striking then is that in their categorization of vowel vs. no vowel, language background played a significant role. In particular, Korean listeners tended to categorize stimuli with no vowel as containing a vowel only when the stop was voiced. But when the stop was voiceless, their response pattern was almost identical to that of English listeners, categorizing stimuli with no vowel correctly as containing no vowel.

However, recall that Dehaene-Lambertz et al. (2000) found a language-specific effect in Japanese and French listeners' perception of the same difference (full vowel vs. no vowel). I speculate that the difference in results between this study and their study may be due to stimulus differences. Dehaene-Lambertz et al. used standard stimuli (precursor set) created from six female speakers and deviant items (test items) from a synthesized male voice. Moreover, they used 6 different nonsense word pairs (*igno-iguno, igna-iguna, ikno-ikuno, ikma-ikuma, okna-okuna, ogma-oguma*) that differed in the vowels or consonants. Taken together, the experimental materials included a good deal of acoustic variability in testing listeners' discrimination of -CC- and -CuC-. On the other hand, in the current MMN paradigm, the materials were all from the same male voice and there was only one stimulus pair (*igna-igɨna or ikna-ikɨna*) for each condition. This difference points to the possibility that the level of the representation at which

MMN responses were measured may be different: the mismatch response in the current study may be explained with reference to acoustic differences while the MMN in Dehaene-Lambertz et al.'s study may reflect a difference in phonological categories.

The discrepancy between MMN and categorization responses suggests that misperception (or miscategorization) of coda voiced stop as a sequence of stop and vowel is not rooted in audibility, but rather in the categorization process. This mapping from acoustic streams to lexical representations has been argued by a number of researchers (Boersma, 1998; Broselow, 2009b; Chung, 2004; Kenstowicz, 2001; Silverman, 1992; Yip, 2006) to take place within a separate perception grammar, which classifies continuous acoustic signals into discrete perceptual categories, which then serve as inputs to the production grammar. Boersma (1999, p. 5) proposes that the perception grammar "is NOT about *audibility*". He further asserts that the fact that a listener maps certain acoustic forms to another perceptual category does not necessarily mean that they do not hear the acoustic difference between the two, but rather they categorize them according to their native language system. So, in the Korean case, the fact that Korean listeners tended to map *igna* with no vowel to *igɨna* with a full vowel does not necessarily mean that they did not hear the acoustic difference, as shown by the distinct results from behavioral categorization and MMN experiments.

Though the current results suggest that Korean listeners' misperceptions took place at the level of categorization rather than at the level of acoustic processing, I should note that I do not claim that the level of miscategorization is always later in the speech perception process. Given substantial neurophysiologic evidence that native language affects the very early level of auditory perception for certain phonetic categories, certain phonetic contrasts may be more susceptible to language effects at the pre-attentive level, although some other contrasts may be less so (see Phillips, 2001). As differences across studies show, considerable work remains to be done to examine which kinds of acoustic features are subject to total loss or insensitivity and which others are not.

In summary, Korean and English listeners' pre-attentive discrimination of no vowel vs. full vowel in stop-nasal sequences appeared to be very similar. This suggests that the difference in categorization between these two groups reflects a higher level of perceptual processing. In Chapter 5, I develop a perception grammar that encodes the voicing effect in the perception of stop-nasal sequences. In this grammar, [gn] is mapped to /kɨn/ but [kn] to /k^hn/, reflecting the categorization difference between the two groups.

3.3 Summary

In this chapter, we investigated Korean native speakers' perception of English stop-nasal sequences in comparison to the perception of these sequences by English native speakers, employing both behavioral and electrophysiologic methods. The experimental results showed that the frequent vowel insertion after voiced stop in stop-nasal sequences is rooted in Korean listeners' misinterpretation of voiced stops as being followed by a vowel, affected by their native language phonotactics. This result implicates that voicing of the stop, which is not contrastive in Korean, may be accessed by Korean listeners in their perceptual processing. However, we had no evidence that Korean listeners tended to categorize voiced stops as voiceless stops, especially when the stop is bilabial. Their perception of voicing contrast before nasals was similar to that of English listeners.

Chapter 4
Development of the Phonological Grammar

The production experiments discussed in Chapter 2 showed that Korean L2 learners of English commonly employed certain repairs in production of English stop-nasal sequences. The results revealed two asymmetries involving voicing and place: while Korean speakers sometimes nasalized the stop, they also tended to insert a vowel after a voiced stop or/and velar stop and to devoice labial voiced stops. This pattern is puzzling in that the L1 transfer account cannot explain why the repair patterns should be affected by the voicing or the place of the stop. The results of the perception experiments in Chapter 3 suggested that the voicing effect was a reflection of misperception but the place effect was not. In this chapter, I argue that an interlanguage grammar, which is developed under the assumption that learning is error-driven and perception is accurate, cannot fully account for the place asymmetry found in devoicing and vowel insertion. I also demonstrate that it is not possible to develop a grammar that derives vowel insertion because such a grammar would not be learnable from the available data. This was confirmed by the experimental results suggesting that a vowel was already inserted in perception rather than in production.

4.1 Korean Grammar and Nasalization

Nasalization patterns in Korean have been analyzed within the Optimality Theoretic framework using an undominated markedness constraint SYLCONT (Davis & Shin, 1999; I. K. Park, 2006), which defines preferred contact combinations across syllables. SYLCONT penalizes rising sonority over a syllable boundary, as defined in (1) (Bat-El, 1996; H. Kang, 2002; Vennemann, 1988).

(1) SYLCONT (Syllable Contact): The onset of a syllable must not be more sonorous than the last segment in the immediately preceding syllable and the greater the slope in sonority the better (Avoid rising sonority over a syllable boundary, C1 ≥ C2).

Korean provides cases where this constraint forces changes in a sequence of coda and onset in which the onset is more sonorous than the preceding coda:

(2) Syllable Contact Repairs
   a. coronal nasal-lateral (rising sonority) > lateral-lateral (equal sonority)
      (/nonli/ → [no**ll**i] ‘logic’)
   b. noncoronal nasal-lateral (rising sonority) > nasal-nasal (equal sonority)
      (/sam-lyu/→[sa**mn**yu] ‘third rate’)
   c. noncoronal obstruent-lateral (rising sonority) > nasal-nasal (equal sonority)
      (/pəp-li/→[pə**mn**i] ‘principle of law’)[11].

[11] Which segment is changed, onset or coda, is determined by an independent constraint (Kang H. 2002), which is outside the scope of this study.

In all these cases, the motivation for nasalization is to avoid rising sonority over a syllable boundary. In Korean, the preferred repair for an illegal syllable contact is change in the sonority of one or both segments, rather than insertion of a vowel to separate the adjacent consonants.

This grammar can correctly account for both nasalization and a faithful realization of underlying [ɨ] in Korean, as illustrated in (3).

(3) a. DEP-IO(V): Output segments must have input correspondents. ('No epenthesis')
b. Ranking : SYLLCONT >> DEP-IO(V) >> IDENT-IO(sonorant)

c. Nasalized production: /kukmul/ → [kuŋmul] 'soup'

Input: /kukmul/	SYLCONT	DEP-IO(V)	IDENT-IO[sonorant]
a. kuk.mul	*!		
b. ☞ kuŋ.mul			*
c. ku.gɨ.mul		*!	

d. Correct output within this grammar /jat^{h}ɨmjən/ → [jat^{h}ɨmjən] 'if it is shallow'

Input: / jat^{h}ɨmjən /	SYLCONT	DEP-IO(V)	IDENT-IO[sonorant]
a. jat.mjən	*!		
b. jan.mjən		*!	*
c. ☞ jat^{h}ɨmjən			

According to this grammar, which bans stop-nasal sequences and repairs such sequences by nasalizing the stop, language learners transferring their native language grammar should nasalize both voiceless stop-nasal and voiced stop-nasal sequences. (4) and (5) illustrate the predictions of the native language grammar for English input.

(4) English input /kɛtməl/ →[k^{h}ɛn.məl]: Nasalization

Input: / kɛtməl /	SYLCONT	DEP-IO(V)	IDENT-IO(nasal)
a. k^{h}ɛt.məl	*!		
b. ☞ k^{h}ɛn.məl			*
c. k^{h}ɛ.t^{h}ɨ.məl		*!	

(5)English input /kɛdməl/ →[k^{h}ɛn.məl]: Nasalization

Input: / k^{h}ɛdməl /	SYLCONT	DEP-IO(V)	IDENT-IO(nasal)
a. k^{h}ɛt.məl	*!		
b. ☞ k^{h}ɛn.məl			*
c. k^{h}ɛ.dɨ.məl		*!	

Note that in the two production experiments, nasalization was similarly frequent in both voiceless stop and voiced stop-nasal sequences, as summarized in (6). This is consistent with the native production grammar.

(6) Rates of nasalization in production experiments

	voiceless stop-nasal	voiced stop-nasal
Reading Experiment	62 (13.5%)	58 (12.1%)
Repetition Experiment	39 (8.1%)	42 (8.8%)

However, the occurrence of devoicing and vowel insertion in the production experiments is more difficult to explain as an effect of the grammars of either Korean or English. In the next section I turn to the discussion of vowel insertion.

4.2 Interlanguage Grammar

4.2.1 Vowel insertion and unlearnable rankings

On the assumption that Korean speakers have accurate perception, vowel insertion within the stop-nasal sequences is not motivated by English since DEP-IO(V) is undominated in the English grammar. The Korean grammar cannot derive such forms either, unless a reranking of two constraints occurs. (7) derives the output where a vowel is inserted by reranking IDENT-IO(nasal) and DEP-IO(V).

(7) English input /sɛgməl/ → [sɛgɨməl]: Epenthesis

Input: /sɛgməl/	SYLCONT	IDENT-IO(nasal)	DEP-IO(V)
a. sɛg.məl	*!		
b. sɛŋ.məl		*!	
→c. sɛ.gɨ.məl			*

In (8), three possible rankings of the constraints are presented. If we assume that the initial state for second language acquisition is the NL ranking, Korean speakers learning English begin with the Korean grammar as shown in (8a). (8b) is assumed to be an intermediate state between Korean and English in which a vowel is inserted after the stop by reranking F_1 and F_2. (8c) is the target language English, in which both F_1 and F_2 dominate the markedness constraint. All of the rankings in (8) assume that Korean speakers have accurate perception.

(8) Three possible rankings of the same constraints

Constraints : M = SYLCONT, F_1 = DEP-IO(V), F_2 = IDENT-IO(nasal)

	Ranking	Input	Output	Language
a.	M , F_1 >> F_2	Stop.Nasal	Nasal.Nasal	Korean
b.	M , F_2 >> F_1	Stop.Nasal	Stop V.Nasal	Interlanguage
c.	F_1, F_2 >> M	Stop.Nasal	Stop.Nasal	English

This analysis predicts three possible linguistic systems by reranking of the same constraints. However, if we assume that constraint reranking is error-driven, reranking takes place only if there is a mismatch between the correct form and the optimal form chosen by the current grammar. An error occurs if the optimal candidate and the observed form do not match. In this case, constraints may be reranked. If the two match, no error occurs, and so no reranking takes place (Tesar & Smolensky, 2000).

Vowel insertion is problematic in that it requires a reranking of two constraints which cannot be implemented by an error-driven learning algorithm. In other words, the intermediate state (8b) cannot emerge from the initial state (8a) if reranking is error-driven. (9) presents a model of error-driven learning in the acquisition of English word-internal stops before nasal. Assuming the English output [sɛgməl] is the input to the grammar, (9a) is the correct form and (9b) is the optimal candidate according to this ranking. Since there is a mismatch between them, learning occurs and SYLCONT needs to be demoted below IDENT-IO(nasal) so that (9a) can be optimal.

(9) Loser/Winner pair for English input [sɛgməl]

	[sɛgməl]	SYLCONT	DEP-IO(V)	IDENT-IO(nasal)
☺a.	sɛg.məl	*! →		
→b.	sɛŋ.məl			*
c.	sɛ.gɨ.məl		*!	

After learning occurs, the reranked grammar in (10) allows the optimal candidate to match the actual output. Since there is no error, no more learning takes place. So, by this learning algorithm there is no way to demote DEP-IO(V) below IDENT-IO(nasal) so that (10c) can be a winner. Error-driven learning cannot cause reranking of two faithfulness constraints (F_1 >> F_2 → F_2 >> F_1) in the absence of positive evidence. Consequently, the reranked grammar (8b) should not emerge in the learning process.

(10) Error-driven reranked grammar for [sɛgməl]

	[sɛgməl]	DEP-IO(V)	IDENT-IO(nasal)	SYLCONT
→ ☺ a.	sɛg.məl			*
b.	sɛŋ.məl		*!	
c.	sɛ.gɨ.məl	*!		

Some languages show similar patterns in which a native repair strategy is different from repairs used for foreign structures. In these cases, the interlanguage production grammar also appears unlearnable. For example, in Maori, a word-final consonant is deleted, but in loanwords a final consonant is retained and a vowel is inserted after it (Yip, 2002). In Japanese, the vowel insertion strategy is widely used for loanwords, while deletion is employed for native Japanese words (Paradis & LaCharité, 1997; Smith, 2004). Korean adaptation of English liquids also provides such a case: an intervocalic lateral is produced as a tap in native Korean while it is geminated in loanwords. Analysis of this pattern would require positing an unlearnable ranking in which faithfulness to duration is demoted below faithfulness to laterality (Broselow, 2009a; Kenstowicz, 2005). Word final /s/ in native Korean words is produced as /t/, but in loanwords it is realized with an inserted vowel (/pəs/ [pət] 'friend' vs. /bʌs/ [pəsɨ] 'bus'). On the assumption of accurate perception of the foreign structures, all the cases illustrated above, including the vowel epenthesis for stop-nasal sequences, will require rankings that are not learnable via an error-driven learning algorithm. Nevertheless, some studies have adopted the view that interlanguage phonology is independent from the native phonology (H. Cho, 1998; Kenstowicz, 2005; Ku, 1999; Lee, 1995), where the native ranking is reversed in loanword phonology or in which the grammar includes constraints or principles specific to foreign structures. However, Kang (2003) and Peperkamp (2005) provide convincing arguments that the arbitrary reranking for vowel insertion in loanwords is problematic.

Furthermore, recall that vowel insertion occurred most often after voiced stops or velar stops in both production experiments, as shown in (11). To account for these asymmetries, the grammar would need constraints specified in terms of voicing or place that are separately ranked, though neither Korean nor English (and possibly no language) provides any evidence for such rankings.

(11) Vowel insertion rates by the voicing and place of the stop

	Voiceless stop-nasal			Voiced stop-nasal		
	Bilabial	Coronal	Velar	Bilabial	Coronal	Velar
Reading Experiment	9 (6.3%)	20 (12.7%)	25 (15.8%)	34 (21.3%)	39 (24.4%)	63 (39.4%)
Repetition Experiment	0 (0%)	0 (0%)	7 (4.4%)	9 (5.6%)	17 (10.6%)	72 (45%)

Some of the inexplicable patterns that emerge in loanword adaptation or in interlanguage phenomena are argued to be reflections of universal markedness principles operating in foreign language production (Broselow et al., 1998; Shinohara, 2004 among others). Voiced obstruent codas are considered to be more marked than voiceless obstruent codas. Thus, devoicing of word final voiced codas by Mandarin speakers is analyzed as the Emergence of the Unmarked (TETU)

(Broselow, Chen and Wang 1998), due to the influence of a universal markedness constraint disfavoring voiced obstruent codas. However, as Steriade (1997) points out, the common way to avoid voiced codas in NL grammars is by devoicing them, not inserting a vowel, as in Dutch, German, Russian, etc. Steriade claims that no language inserts a vowel after voiced obstruents to avoid voiced obstruent codas. Therefore, an epenthetic vowel after voiced coda stops found in the present experiments cannot be explained by the constraint ranking inherited from UG.

So far, I have shown that the vowel insertion pattern cannot be derived in the interlanguage grammar on the assumption that the inputs to the grammar are English forms, i.e., that Koreans perceive the English forms accurately, and that reranking of the constraints is error-driven. I have also shown that the voicing effect in vowel insertion cannot be the effect of the Emergence of the Unmarked. I argue that these problems disappear once we take into account the possibility that the inputs to the production grammar may not be English forms. That is, if we find evidence that the input contains a vowel only after voiced stop, then we no longer need to posit novel rankings or constraints, because the Korean grammar can account for the pattern, as illustrated in (12). The input in (12a) contains a vowel and the faithful output is chosen under this native ranking.

(12) English input /sɛgɨməl/ → [sɛgɨməl]: Faithful output

Input: /sɛgɨməl/	SYLCONT	IDENT-IO(nasal)	DEP-IO(V)
a. sɛg.məl	*!		
b. sɛŋ.məl		*!	
→c. sɛ.gɨ.məl			

This analysis is supported by the results of the perception experiments reported in Chapter 3. Recall that Korean listeners, but not English listeners, tended to report hearing a vowel even though the stimulus contained no vowel--but, critically, only when the stop was voiced. This result confirms the analysis that the input to the grammar is filtered through L1. The mechanism of how the acoustic signal of the English form is filtered in perception is discussed in detail in Chapter 5. In the next section I turn to the discussion of devoicing.

4.2.2 Devoicing and TETU

Devoicing of voiced stops seems to be also problematic in that it is not motivated by the grammar of either English or Korean: a devoiced stop in this particular position will still form an illegal sequence in Korean, and both voiced and voiceless stops before a nasal are well-formed structures in English. The fact that the subjects adopted the devoicing strategy suggests that they have already developed an interlanguage grammar that allows a sequence of a voiceless stop and a nasal, approaching the grammar of English, which allows both voiceless and voiced stops before a nasal. The hypothesis that sequences of voiceless stop-nasal are easier to acquire than voiced stop-nasal is supported by the experimental results, where more tokens of voiceless stop-nasal were correctly produced (61.8% in the reading experiment and 76.9% in the repetition experiment) than voiced stop-nasal tokens (27.9% in reading and 25.2% in repetition). The following reranked constraints present the interlanguage grammar that describes the pattern of subjects who have mastered the faithful realization of stop-nasal inputs. For these participants, the markedness constraint SYLCONT has been demoted below the other faithfulness constraints by error-driven learning (Tesar & Smolensky, 2000). Once the demotion of SYLCONT takes place, the faithful candidate (14a) is chosen as optimal.

Demotion

(13) (SYLLCONT) >> DEP-IO(V) >> IDENT-IO[sonorant] >> SYLLCONT

(14) English input /pɪtmən/

Input /pɪtmən /	DEP-IO(V)	IDENT-IO[sonorant]	SYLCONT
a. ☞ p^hit.mən			*
b. p^hin.mən		*!	
c. p^hi.tɨ.mən	*!		

However, the current set of constraints cannot determine the output when an input has a voiced coda stop. To have a complete analysis, we first need to explain how voicing of stops is determined in Korean.

In Korean, voiceless obstruents become voiced between two sonorants. The constraints in (15) describe the distribution of voicing in obstruents in Korean. We have a contextual markedness constraint *NpN, which prohibits a voiceless stop between two sonorants, reflecting intersonorant voicing. *VD OBS bans voiced obstruents in general. IDENT-IO(voice) is a faithfulness constraint, prohibiting voicing change.

(15) Constraints for intersonorant voicing:
*NpN: Intersonorant voiceless plain stops are not allowed to surface.
* VD OBS: No voiced obstruent
IDENT-IO(voice): No voicing changes

(16) Ranking for intersonorant voicing:
*NpN >> * VD OBS >> IDENT-IO(voice)

The ranking in (16) chooses (17b) and (18b), which contain a voiced stop in intersonorant position, as optimal.

(17) Korean input /kompo/ 'pockmark'

input: /kompo/	*NpN	* VD OBS	IDENT-IO(voice)
a. kompo	*!		
b.☞ kombo		*	*

(18) Korean input /papo/ 'fool'

input: /papo/	*NpN	* VD OBS	IDENT-IO(voice)
a. papo	*!		
b.☞ pabo		*	*

Now, let us turn to the case where voiced stops before nasals were devoiced in the experiment. One possible explanation for the interlanguage pattern is that it represents an Emergence of the Unmarked Effect. Cross-linguistically, voiced obstruent codas are disfavored, with many languages allowing only voiceless obstruents in coda position. A positional markedness constraint *VOICEDOBSTRUENTCODA has been used to account for the devoicing of final obstruents in languages such as German, Dutch, Russian, Polish, and Catalan (Itô & Mester, 1994; Lombardi, 1995; Zoll, 1996, 1997, 2004 among others).

(19) * VD OBS CODA: No voiced obstruent coda
*VD OBS CODA >> *NpN >> * VD OBS >> IDENT-IO(voice)

Positional markedness constraints prevent marked forms in a weak position such as the coda. If markedness constraints such as * VD OBS CODA are assumed to be universal, this constraint must be present in the grammar of Korean, although it has no visible effect in this language, because inputs in Korean do not provide relevant forms in which * VD OBS CODA can play a role. However, * VD OBS CODA is effective in choosing the winner when an English input is provided. A similar interlanguage pattern has been reported (Broselow et al., 1998) in which speakers of Mandarin, whose native language allows no obstruent codas at all, first acquire voiceless obstruent codas (Broselow et al., 1998). We could describe the grammar of subjects who devoice obstruent codas as having achieved the following ranking as an intermediate stage between the Korean grammar (which bans all obstruent-nasal sequences) and the English grammar (which allows both voiceless and voiced obstruents to occur before nasals):

(20) English input [pɪdmən]

input: /pɪdmən/	*VD OBS CODA	*NpN	* VD OBS	IDENT-IO(voice)
a. p^hɪd.mən	*!		*	
b. ☞ p^hɪt.mən		*		*

Under this account, the devoicing pattern emerges as an intermediate stage of the acquisition of voiced stop-nasal sequences: producing voiceless stop-nasal sequences requires only that SYLCONT be demoted below the faithfulness constraints. Producing voiced stop-nasal sequences requires two demotions: both SYLCONT and * VD OBS CODA must be reranked below faithfulness constraints.

However, this account cannot explain the effect of place of articulation revealed in the data. As (21) shows, in both experiments devoicing was found much more commonly in bilabial and alveolar stops than in velar stops. A similar asymmetry regarding devoicing was found by Major & Faudree (1996), who found that Korean-speaking ESL learners devoiced 93% of final [b] in a word list reading task but only 53% of final [g]. This asymmetry is even greater in the text reading task: no instances of final [b] were devoiced, while every final [g] was devoiced[12]. This place asymmetry is not consistent with the TETU effect. If devoicing is solely a function of the

[12] Major & Faudree (1996) report accuracy in terms of voicing only. The correct production of voiced stops means that they were correctly produced in terms of voicing. Thus, the actual pronunciation of voiced stops when they were marked as 'correct' may not necessarily be English-like. In other words, they could have been produced with an inserted vowel.

demotion of constraints, a similar number of devoicing cases should be found across all of the voiced stops. To describe the pattern whereby labial but not velar stops are devoiced would require that the single constraint *VD OBS CODA be broken into separate constraints referring to place (*VD OBS CODA-LABIAL, *VD OBS CODA-VELAR) which could then be ranked independently. But to my knowledge, no language has been attested that would motivate assuming separate constraints relativized to place, and this pattern is not supported by either English or Korean data. The place asymmetry found in devoicing cannot be accounted for in terms of a simple reranking of constraints in the production grammar.

(21) Rates of Devoicing across Voiced Stops

	Bilabial	Coronal	Velar
Reading Experiment	44 (27.5%)	62 (38.8%)	22 (13.8%)
Repetition Experiment	96 (60%)	79 (49.4%)	26 (16.3%)

4.2.3 Summary

In this chapter I have shown that the greater frequency of vowel insertion after voiced than after voiceless stops cannot be explained by the reranking of production grammar constraints or as an Emergence of the Unmarked Effect. Instead, I argue that the assumption that the English input is filtered through the native language perception grammar, supported by the results of the perception experiments, can account for the voicing effect in vowel insertion. I have also shown that the greater frequency of devoicing with labial stops than with velar stops cannot be analyzed as an effect of the Emergence of the Unmarked. In Chapter 6 I argue that this effect reflects the articulatory timing patterns of Korean.

Chapter 5
Misinterpretation of Acoustic Cues

5.1 Misperception Approach

In the previous chapters, I showed that Korean listeners misinterpreted voiced stop-nasal sequences as voiced stop-vowel-nasal sequences. This suggests that Korean listeners must have perceived voicing. Given that voicing is not a contrastive feature in Korean, this raises the question of whether non-native features can be perceived by L2 learners, and if so, in what way they are mapped to phonological representations. According to the *Native Language Filter* (*NL filter*) model, it is suggested that perception is mediated by contrastive phonological features in NL. Therefore, when L2 listeners process acoustic signals of the TL, they suppress information that is predictable and non-contrastive in their native language (Brown, 1998, 2000; LaCharité & Paradis, 2005; Lahiri & Marslen-Wilson, 1991). For instance, Brown (2000) claims that perception is filtered through the distinctive feature system of the NL. This claim is based on Brown's finding that Mandarin native speakers perceived the English [r]-[l] contrast more successfully than did Japanese and Korean native speakers, although none of the three languages employs an [r]-[l] contrast. Brown claims that since Mandarin has a richer inventory of contrasts in coronals, Mandarin requires specification of a feature that can be used to distinguish the English [l]-[r] phonemes. Specifically, the feature that distinguishes Mandarin [s]-[ʂ] is available to Mandarin speakers in perceiving the English [r]-[l] contrast. The inventories of Korean and Japanese do not require such a feature, leading to greater difficulty in perceiving the English contrast.

In contrast, the *Cue Interpretation* model (Boersma & Hamann, 2009; Peperkamp & Dupoux, 2003; Silverman, 1992) assumes that all acoustic cues (whether contrastive or predictable) are available to listeners, but that the interpretation of these cues is guided by the NL perception grammar, which may induce greater attention to particular cues. L2 listeners map the phonetic properties of the TL signal to the NL phonetic representation that provides the best fit. Having briefly discussed two models that differ from each other in terms of how they incorporate misperception in the mapping from acoustic streams to phonological representation, I next review two studies about the mapping from English to Korean.

5.1.1 Native language filter: Kabak & Idsardi's model

Kabak and Idsardi (2007, henceforth K&I) tested Korean listeners' perception of English nonsense words with word-internal clusters such as [kC], [gC] or [ʧC]. Compared to [kC]-[kʊC] or [gC]-[gʊC] stimuli pairs, [ʧC]-[ʧiC] stimulus pairs were much less accurately discriminated by Korean listeners. Based on these results, they argue that not all illegal clusters induce perception of an illusory vowel. Rather, listeners process in terms of syllables: only when the first consonant would not be a possible coda do listeners hear an illusory vowel following it. Therefore, clusters containing [ʧ], which is not a legal coda consonant, were hard for Korean listeners to distinguish from the counterpart clusters containing a vowel. In contrast, clusters containing a velar [k] or [g], which may be a legal coda, were not as hard for Korean listeners to distinguish from the clusters containing a vowel, confirming their hypothesis that perception is processed in terms of legal NL syllables rather than legal syllable contacts.

However, this account is problematic for [gC]-[gʊC] pairs, which were fairly accurately discriminated, though voiced codas are also illegal codas in Korean. To resolve this, K&I argue that Korean listeners perceive only contrastive features in Korean. They therefore perceive [g] in [gC] clusters as /k/, as claimed by the *NL filter* model, because voicing in Korean is predictable. If Korean listeners perceived [gC] as /kC/, the [gC]-[gʊC] (/kC/-/kʊC/) pair could be easily discriminable because the legal coda /k/ in /kC/ would not induce perception of an illusory vowel as in /ʧC/. In summary, K&I argue that the reason Korean speakers can distinguish [gC]-[gʊC] is because (a) they perceive [gC] as /kC/; and (b) /VkCV/ is a legal sequence of syllables in Korean.

I now provide possible alternative analyses of their data. Steriade (2008) pointed out that the reason Korean listeners discriminated [gC]-[gVC] well could be because K&I used English [ʊ] . As shown in (1), the English vowel [ʊ] is nearly invariably adapted to the Korean vowel [u], which is not the default epenthetic vowel in Korean. Therefore, the successful discrimination of [gC]-[gVC] could have been simply due to the use of a non-epenthetic vowel in the materials; the presence (or absence) of the vowel [ʊ] in [gC] should not be difficult to perceive because it would have been represented as [guC] rather than [gɨC].

(1) English [ʊ] adapted as Korean [u] (Steriade 2008)

a.	full	[pʰul]
b.	hook	[huk, hukʰɨ]
c.	foot	[pʰut]
d.	look	[luk]

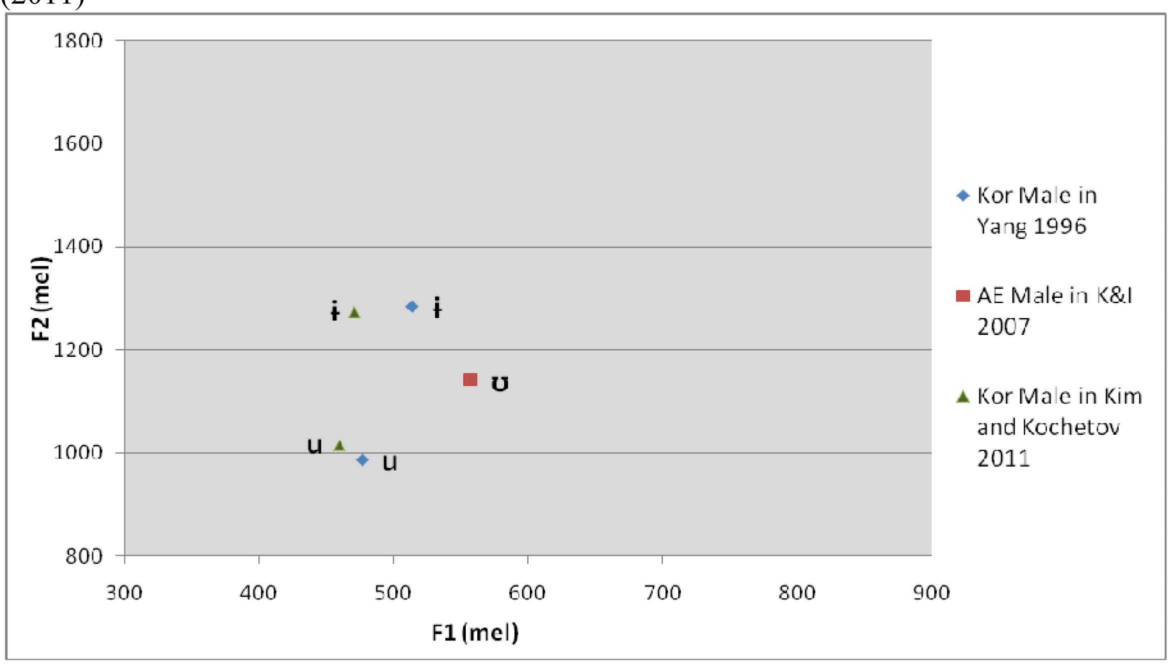
It is almost unarguably the case that English vowel [ʊ] is mapped to Korean vowel [u] in loan adaptation, as illustrated in (1). Yet, note that the examples given in (1) are all monosyllabic words bearing a primary stress on the vowel [ʊ], whereas K&I's stimuli contained the vowel [ʊ] in the second syllable following a stress on the first syllable (e.g., [pʰágʊma]). Considering the fact that the English vowel quality varies by stress position, the mapping from English [ʊ] to Korean [u] in monosyllabic words may not be sufficient evidence for Steriade's claim that Korean listeners mapped [ʊ] in the stimulus to their [u].

In fact, K&I argue that the vowel [ʊ] in their stimuli is acoustically closer to Korean vowel [ɨ] than to [u] based on an acoustic study of Korean vowels (Yang, 1996), and that therefore the choice of their vowel [ʊ] is not the primary source for Korean listeners' successful discrimination of [CVgCa] from [CVgʊCa]. However, this comparison still has the same limitation in that the vowels being compared occur in different contexts: the English vowels [ʊ] in K&I's materials are in the unstressed second syllable position whereas the vowels reported in Yang (1996) were in the first syllable following a consonant [h] as in [huda] or [hɨda]. A recent acoustic study of the Korean vowels by Kim & Kochetov (2011) demonstrates that the formants of the Korean vowel [ɨ], whether epenthetic or lexical, significantly vary depending on the preceding vowel and the place of the preceding consonant (Kim & Kochetov, 2011, p. 521).

Thus, unless we compare the vowels in the compatible contexts, it is hard to determine whether the English vowel [ʊ] in K&I's materials was indeed closer to the Korean vowel [ɨ] than to [u].

K&I's materials contained a vowel [a] in the first syllable, and the consonant preceding [ʊ] was either velar or alveolar (e.g., [pʰakʰʊma], [pʰalʊtʰa], etc). Kim & Kochetov (2011) provide the formants of [ɨ] measured in a comparable context: in the stimuli the vowel [ɨ] followed the same vowel [a] and a velar or alveolar consonant (e.g., [sapʰu], [satʰɨ] etc.)[13]. (2) shows superimposed F1 and F2 (in mel) of the vowel [ʊ] in K&I's materials, F1 and F2 (in mel) of the Korean vowels [u] and [ɨ] in Yang (1996) and F1 and F2 (in mel) of the Korean vowels [u] and [ɨ] in Kim and Kochetov (2011). All vowels were produced by male speakers. The chart illustrates that the vowel [ʊ] from the experimental materials in K&I's experiment is closer to the Korean vowel [ɨ] than to [u] in Yang (1996). Yet, it is not clear if the same holds for the comparison with the Korean vowels reported in Kim and Kochetov (2011), as the English vowel [ʊ] appears to be right in the middle of the Korean vowel [u] and [ɨ]. In other words, there is no clear reason to assume that the vowel [ʊ] in K&I's experimental materials was mapped onto the Korean [ɨ] instead of [u] because its acoustic/perceptual quality is intermediate between [ɨ] and [u].

(2) Superimposed F1 / F2 (in mel) of [ʊ] from the experimental materials in K&I's experiment and F1 /F2 (in mel) of [u] and [ɨ] reported in Yang (1996:251, Table III) and Kim & Kochetov (2011)



Another reason that K&I assume that the English vowel [ʊ] in their materials was mapped onto the Korean vowel [ɨ] instead of [u] is based on cross-linguistic perception results from Dupoux et al. (1999). In Dupoux et al.'s experiment, Japanese speakers had difficulty discriminating CC

[13] This is a subset of the data they provide. The averages of F1 and F2 were computed only for the items that contained the vowel [a] and velar and alveolar consonants (Kim and Kochetov 2010, Table 5)

from CVC, where the V was the French rounded back vowel [u]. The Japanese epenthetic vowel is [ɨ], which is equivalent to the Korean epenthetic vowel [ɨ], so the French vowel should be distinct from the Japanese epenthetic vowel. However, Korean differs from Japanese in having a contrast between [u] and [ɨ] (e.g., [k**u**ŋʤi] 'corner' [k**ɨ**ŋʤi] 'pride') –unlike Japanese, which has only one high back vowel. In other words, for Japanese speakers, there is no other vowel to map the French vowel [u] onto except their [ɨ], whereas for Korean speakers, there are two options: [u] or [ɨ]. Therefore, it is still not clear that Korean listeners mapped the English vowel [ʊ] onto their [ɨ] rather than [u].

So far, I have shown that there is no clear evidence for whether the use of the English vowel [ʊ] was appropriate or not, given that the English vowel [ʊ] falls between Korean vowels [u] and [ɨ]. Therefore, the reason why Korean speakers could distinguish [gC]-[gʊC] relatively well cannot be fully explained by the use of the non-epenthetic vowel [ʊ]. Instead, we need to examine whether Korean speakers indeed perceive [gC] as /kC/ because voicing is not contrastive in Korean, as K&I claimed. Note that this is a critical issue for K&I's argument that the perceptual epenthetic vowel is triggered by the syllable structure, not phonotactic restrictions, because perceiving [gN] as /kN/ makes the sequences legitimate in terms of syllables but not in terms of syllable contact. This is also related to the question of how non-native features are perceived. Since Korean has no voicing contrast, voicing should be difficult to perceive if Korean listeners' perception is only based on contrastive features in Korean, as claimed by the *NL filter* model. However, if their perception is based on subphonemic information, the voicing contrast should be perceived, although voiced stops would be interpreted as being followed by a vowel because voiced stops in Korean always appear before a vowel.

Now, recall the results of the perception experiments in Chapter 3. In the voicing classification task, Korean listeners reported over 95% of the time that the stop was voiced when the stop in *igna* or *ibna* was fully voiced. The Korean speakers' responses were almost identical to the English listeners' responses. However, in the classification of presence vs. absence of a vowel, Korean listeners reported they heard a vowel about 50% of the time in *igna* or *ibna,* unlike English listeners whose vowel response was less than 10%. These results clearly show that Korean listeners perceived voiced stops before a nasal as voiced, not as voiceless, and occasionally they tended to perceive a vowel after the stop.

These results are therefore not consistent with K&I's claim that Korean listeners' perception is phonological rather than phonetic. Instead, the results support the *Cue Interpretation* model that L2 learners have access to phonetic, non-contrastive acoustic cues.

5.1.2 Cue interpretation: Boersma & Hamann's model

In this section, I review Boersma & Hamann (2009, henceforth B&H)'s analysis of Korean listeners' perception of English stop-nasal sequences and show that their analysis captures the fact that Korean listeners' perception does not depend solely on contrastive features but rather makes use of available acoustic cues. However, I argue that their analysis does not account for the voicing effect found in the perception experiments, in which illusory vowel perception was induced only in voiced stop-nasal sequences but not in voiceless stop-nasal sequences. This fact still needs explanation under their analysis.

Using "cue constraints that evaluate the relation between the input of the perception process (the auditory-phonetic form) and the output of the perception process (the phonological surface form)" (p. 6), B&H propose a perception grammar in which a vowel is inserted. For instance, an English word 'tag' is adapted as [tʰɛ.gɨ] with an inserted vowel [ɨ] at the end of the word. The input to the perception grammar is an acoustic signal, represented in a narrow auditory transcription. (" _ ") in [_thæg̚ ˷^{g}] represents the silence during stop closure and superscript "th" is the release burst of voiceless aspirated alveolar stop. Then the vowel follows and [g̚] stands for the vowel transition from the preceding vowel into the voiced velar stop. " ˷ " means voicing during the voiced stop closure and the superscript "^{g}" represents the release burst of [g] after the closure.

Since Korean coda stops are never pronounced with a release burst, B&H (2009) propose a cue constraint *[burst]/C(.)/ that bans perception of an auditory release burst as a consonant in coda, motivated by the fact that released stops always occur in onset (that is, before a vowel) in Korean. To explain the occurrence of epenthetic vowel [ɨ] in L2 speech or loanword adaptation, a constraint *[ ]/ɨ/ is proposed, meaning that [ ](silence) should not be perceived as a phonological vowel [ɨ]. With this constraint ranked lower than *[burst]/C(.)/, the output (b) is chosen as optimal. The constraints summarized below encode the role of release burst and illusory vowel insertion in Korean listeners' perception:

(3) Perception Constraints (adopted from B&H, 2009)
 *[burst]/C(.)/= an auditory release burst should not be perceived as a phonological consonant in coda
 *[ ]/ɨ/= no auditory cue should be perceived as a phonological vowel ɨ

These constraints map the acoustic signal onto a phonological representation, as illustrated below:

(4) Korean perception of the English word 'tag'[14] (adopted from B&H, 2009)

[_thą́:g˺ ˷^{g}]	*[burst] /C(.)/	*[ ] /ɨ/
a. /.tʰæk./	*!	
→b. /.tʰæ.kɨ./		*

This analysis assumes that listeners are sensitive to acoustic cues in the input whether the cues are contrastive or not in the native language. These cues are interpreted through the cue constraints that are motivated by the native language data.

Now, to explain why the English word 'chapter' is adapted as 'cha[pʰ]er' but 'picnic' as 'pi[kʰɨn]ic', B&H rely on a structural constraint that bans stop-nasal sequences in perception. In their framework, structural constraints are in effect both in perception and production. They

[14] This tableau only shows the constraints and outputs that are relevant to the current discussion. Regarding the mapping of the vowel and stops, see Boersma and Hamann's full analysis.

argue that instead of nasalization, as in production, a vowel is inserted in perception to obey the high-ranked constraint SYLCON.

(5) Cue constraints and structural constraints (adopted from B&H, 2009)
SYLCON = no rising sonority over syllable boundary
*[_]/+nas/ = silence should not be perceived as nasal

This constraint chooses candidate (6a) over (6b), which violates *[ ]/ɨ/. This illustrates a case where a vowel is not inserted in perception because the syllable contact is legal.

(6) Korean perception of the English word *chapter* (adopted from B&H, 2009)

[_ʧáp̰˺_ᵗʰə]	SYLCON	*[burst] /C(.)/	*[_] /+nas/	*[ ]/ɨ/
→a. /.tsʰæp.tʰə./				
b. /.tsʰæ.pɨ.tʰə./				*!

In (7), a vowel is inserted because vowel insertion is preferable to violating the syllable contact law (b) or mapping silent closure to a nasal (c). This is a case in which vowel insertion occurs in perception because the syllable contact is illegal. Note that in both cases, the stop in the word internal clusters is unreleased in the input.

(7) Korean perception of the English word *picnic* (adopted from B&H, 2009)

[_ᵖʰɪk˺_nɪk˺_]	SYLCON	*[burst] /C(.)/	*[_] /+nas/	*[ ]/ɨ/
a. /.pʰik.nik./	*!			
→b. /.pʰi.kʰɨ.nik./				*
c. /.pʰiŋ.nik./			*!	

This analysis is somewhat problematic. First, if it is indeed the syllable contact constraint that motivates the vowel insertion in 'picnic' but not in 'chapter', we should be able to find other loanwords adapted with an inserted vowel when the syllable contact law is violated. In Korean, the syllable contact law is violated not only by stop-nasal sequences, but also by nasal-lateral sequences and stop-lateral sequences:

(8) Korean native words that violate SYLCON (data from Davis & Shin, 1999)

a.	nonli	[nolli]	'logic'
b.	kamli	[kamni]	'supervision'
c.	pəpli	[pəmni]	'principle of law'

In the native language, [n] becomes lateralized before /l/ (8a); [l] becomes nasalized after [m] (8b), and a stop becomes nasalized before /l/ (8c). Like obstruent nasalization before a nasal, these cases are also analyzed as motivated by the syllable contact law, resulting in level sonority across the syllable boundary (Davis & Shin, 1999). B&H's approach predicts that English words that violate the syllable contact constraint should be adapted with an inserted vowel, which is not the case as shown in (9):

(9) Loanwords that violate SYLLCON (data from the National Academy of the Korean language (Kwuklipkwukeyenkwuwen, 2001))

a.	o**nl**ine	[ollain] [onnain] [onlain] ***[onɨɾain]**
	Sta**nl**ey	[sɨt^{h}ɛʎʎi] [sɨt^{h}ɛnɲi] [sɨt^{h}ɛnɾi] ***[sɨt^{h}ɛnɨɾi]**
b.	ha**ml**et	[hɛmnit] [hɛmɾit] ***[hɛmɨlit]**

The words in (9) are common loanwords used in Korean. A survey of loanword pronunciations (Kwuklipkwukeyenkwuwen, 2001The National Academy of the Korean Language) revealed variants for each word. Strikingly, variants containing an inserted vowel were never found. If the syllable contact law is undominated in the perception grammar, B&H predict that any inputs that violate this constraint should be uniformly adapted with an inserted vowel like 'picnic'. Therefore, simple illegality over the syllable boundary cannot be the only source of the perception of an illusory vowel in the Korean perception.

Even more problematic for B&H's approach are the results from the categorization experiments discussed in Chapter 3, in which Korean listeners did not identify [k˺n] as containing a vowel between the stop and the nasal. Instead, the Korean speakers, like the English speakers, responded that there was no vowel. This is unexpected if SYLCON is undominated in the Korean perception grammar. On the other hand, Korean listeners categorized [g˺n] as containing a vowel between the consonants much more frequently than English listeners. Then, assuming that SYLCON is undominated in the Korean perception grammar, a question arises as to why one form [k˺n] did not undergo perceptual repair, that is, perceptual vowel epenthesis, while the other form [g˺n] did, when neither form is legal in terms of syllable contact. Thus, I argue that it is not SYLCON that motivates perceptual vowel insertion, but something else.

In this section, I have shown that while a *Cue Interpretation* model provides a better fit with my data than a *NL Filter model*, B&H's approach to Korean foreign word mapping does not fully account for the perception experiment results from Chapter 3. In the next section, I propose a different perception grammar, couched within their framework that can resolve this.

5.2 Perception Grammar for Illusory Vowels

First, I argue, following previous proposals (E. Jun, 2002; Y. Kang, 2003; Rhee & Choi, 2001), that what determines perceptual vowel insertion in large part is the release burst of the consonant, regardless of context (that is, whether the consonant appears pre-consonantally or word finally). Recall that the inputs to the perception grammar in (6) and (7) contain an unreleased stop in the internal cluster. However, given that in English these stops can be occasionally produced with a release burst in any context, the reason why 'picnic' is mapped to pi[k^{h}ɨn]ic could also be simply that variation exists in the auditory forms available to listeners. This predicts that we should find variants in learners' productions, because release bursts of the source language stops are variable. In fact, although /p^{h}i.k^{h}ɨ.nik/ is the more frequent pronunciation of 'picnic', /p^{h}ik.nik/ is a possible variant, and plausibly is likely in the deliberate speech often used in the foreign language classroom. Likewise, both /ts^{h}æp.t^{h}ə/ and /ts^{h}æ.p^{h}ɨ.t^{h}ə/ are used. In fact, many loanwords have variants usually with one form more dominant. For some of these, the dominant variant is vowelless, as shown in (10a), even when the stop is followed by a nasal.

(10) Variant loanwords (number of matches found in Google search, April, 2010)

		Dominant	Less dominant
a.	Pacman	/p^{h}ækmæn/ (282000)	/p^{h}æk^{h}ɨmæn/ (5720)
	Jacknife	/tsæknaip^{h}ɨ/(835000)	/tsæk^{h}ɨnaip^{h}ɨ/ (3630)
	nickname	/nikneim/ (21500000)	/nik^{h}ɨneim/ (2820)
	Chapman	/ts^{h}æpmæn/ (4100)	/ts^{h}æp^{h}ɨmæn/ (856)
	Tippmann	/t^{h}ipmæn/ (863)	/t^{h}ip^{h}ɨmæn/ (1)
b.	sickness	/sik^{h}ɨnisɨ/ (12900)	/siknisɨ/ (681)
	picnic	/p^{h}ik^{h}ɨnik/ (681000)	/p^{h}iknik/ (2920)

If we rely on release bursts of the consonant rather than on a highly ranked syllable contact constraint, the variant forms of the words in (10) (and also variant forms of the words with a word-final stop) can be directly explained: because the stops in these contexts are variably pronounced with an audible release burst, variant forms can arise. In addition, the words in (9), which had no vowel insertion despite violating the syllable contact law, can be explained as well: because coda nasals in English are never produced with an audible release burst, whether pre-consonantally or word-finally, a variant with an inserted vowel cannot emerge.

In the absence of release in the stimuli, then, I propose that voicing of the stop induces perception of a vowel. This can explain the contrast we found between the categorization of [k̚n] and [g̚n] in our experiments. The constraint in (11) prevents mapping of voiced stops to a coda position, motivated by the fact that in Korean, voiced stops always occur before a vowel.

(11) Cue constraint that prohibits voiced coda
*[̰]/C(.)/ = voicing should not be perceived as a phonological consonant in coda

The following tableaux illustrate the effect of incorporating this constraint into the grammar. In (12), since the input does not carry a release burst, *[burst] /C(.)/ is not in effect, but voicing of the stop violates *[̰] /C(.)/, and the candidate with a vowel inserted is chosen as optimal.

(12) Korean perception of English nonce word 'igna' without an audible release burst

	[ɪg˺ ̰nə]	*[burst] /C(.)/	*[̰] /C(.)/	*[_]/+nas/	*[ ]/ɨ/
a.	/.ik.nə./		*!		
→b.	/.i.kɨ.nə./				*
c.	/.iŋ.nə./			*!	

On the other hand, when the input is an unreleased voiceless stop, it does not violate any of the constraints, and is chosen as optimal, as shown in (13). Then, in the native production grammar, the stop will be nasalized. In the more English-like grammar where SYLCON has been already demoted, the stop will be correctly produced. This explains the voicing effect (i.e. more frequent vowel insertion after voiced stop-nasal sequences) found in the production experiments, which is in fact a result of the voicing effect in perception (i.e. categorization of [g˺n] as [gɨn]).

(13)Korean perception of English nonce word ‘ikna’ without an audible release burst

	[ɪk˺_nə]	*[burst] /C(.)/	*[̬] /C(.)/	*[_]/+nas/	*[ ]/ɨ/
→a.	/.ik.nə./				
b.	/.i.kɨ.nə./				*!
c.	/.iŋ.nə./			*!	

The constraint in (11) is included in B&H’s analysis, but its role is different in the analysis that I propose, in which this constraint motivates vowel insertion for unreleased voiced stops regardless of the context. In contrast, in B&H’s analysis this constraint forces vowel insertion only in word final position. My point is that this constraint is sufficient to explain vowel insertion between stop-nasal sequences as well as after a word-final stop, whereas B&H crucially rely on SYLCON in the former case.

The ranking of the constraints in the perception grammar presented here makes crucial reference to native language phonology as well as to (possibly universal) factors of perceptual salience. For instance, a constraint like *[burst]/C(.)/ is language-specific, in that a language that allows a release burst of a coda stop in production would have a perception grammar with this constraint ranked low. On the other hand, a language like Korean has a perception grammar with this constraint undominated. Likewise, a language that allows coda voicing would have *[̬]/C(.)/ ranked low in the perception grammar while a language like Korean has this constraint ranked high. But I assume a constraint like *[_]/+nas/ is probably universally undominated, given robust perceptual salience. Even Korean speakers who nasalize stops in this context in production were as accurate as English native speakers in discriminating stop-nasal sequences from nasal-nasal sequences (Kabak, 2003). I finally note that these cue constraints in the perception grammar can be expressed in more traditional formats like faithfulness constraints or markedness constraints. For instance, *[_]/+nas/ can be expressed as IDENT(nasal) and *[ ]/ɨ/ as DEP(V) while *[̬]/C(.)/ as NOCODAVOICE and *[burst]/C(.)/ as ASSUMEBURST-V(Assume a release burst is followed by a vowel). I leave this issue open but only assert that it is crucial that the ranking is not solely determined by the native phonology but also by perceptual salience.

5.3 Summary

I have shown in Chapter 4 that vowel insertion is particularly problematic because the assumption of accurate perception calls into question the learnability of the interlanguage production grammar. To address the learnability problem, I questioned whether errors in production may be a result of misperception. I have shown in Chapter 3 that some of the production patterns do appear to reflect misperception. In the production of English stop-nasal sequences by Korean L2 speakers of English, we found frequent vowel insertion, especially when the stop was voiced or/and velar. The perception experiments provided evidence for Korean listeners’ tendency to categorize [g˺n] as [gɨn] or [b˺n] as [bɨn], confirming the role of voicing with respect to the perception of an illusory vowel. Based on this result, I have argued in this chapter that Korean listeners' perception is explained better in a model that makes critical reference to acoustic cues rather than to minimally-specified features.

This analysis of vowel insertion proposed in the current chapter solves the learnability problem raised when accurate perception is assumed. Since the perception grammar supplies a vowel in the input to the production grammar, the input form no longer needs a repair to avoid illegal syllable contact, which in turn requires no unmotivated reranking of constraints.

However, recall that vowel insertion was frequent not only after voiced stops, but also after velar stops (whether voiced or voiceless) in the production experiments. Because voicing of the stop is the cue to the presence of a vowel after the stop (in the absence of a release burst), a form like 'sekmal' must be perceived accurately (that is, without a vowel) by Korean listeners. This means that a vowel must be inserted in the production grammar for those cases, which raises the exact same learnability problem because this pattern is not motivated by the data of either the target language or the native language. Therefore, misperception cannot be the only source of vowel insertion found in production. So, we are left with a remaining asymmetry: why is vowel insertion frequent after velar stops (even after voiceless velar stops) while devoicing occurs with bilabial stops? In Chapter 6, I propose an analysis of this puzzling production pattern, using gestural constraints.

Chapter 6
Asymmetries and Gestural Constraints

The production experiments discussed in the previous chapters identified two asymmetries: vowels were more frequently inserted after voiced stops or velar stops and devoicing was more frequent with labial stops than velar stops. These asymmetries in production were only partially explained by the perception-based explanations. While we found perceptual evidence for the tendency to insert vowels in voiced stop-nasal sequences, the place effect still remains puzzling: neither the phonological grammar nor misperception effects can explain why labial voiced stops were more likely to be devoiced while velar stops were more likely to be realized with an inserted vowel.

Given that in the production experiments Korean L2 learners were faced with producing new sequences, not just new segments, I consider another possible source of the place asymmetry: mistimed gestural phasing. In this chapter, I first provide a review of gestural overlapping in the framework of Articulatory Phonology and describe intrusive vowels compared to true phonological epenthetic vowels. In §6.2, I use the criteria outlined in previous research to classify inserted vowels in the production experiment either as true epenthetic vowels or as intrusive vowels, which arise from open transitions between consonants but which do not necessarily pattern as full vowels (Davidson, 2003; Davidson, 2006a; Hall, 2006). In §6.3 and §6.4, I present articulatory evidence to support the claim that asymmetric patterns of vowel intrusion and devoicing originate from specific gestural phasing. In §6.5, I provide an Optimality Theoretic analysis incorporating gestural constraints and argue that timing relations of gestures constitute an important conditioning factor for repair patterns in interlanguage phonology.

6.1 Phonological Vowel Epenthesis vs. Vowel Intrusion

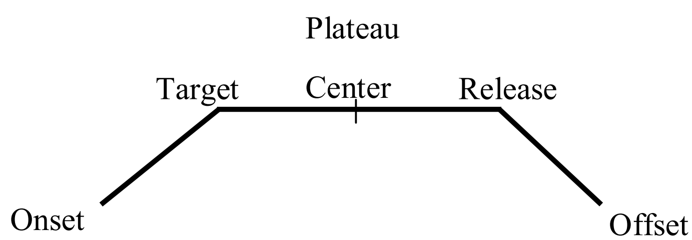
When a vowel is inserted in production, it may not always be a phonological vowel. Davidson (2003) and Davidson and Stone (2003) show, using ultrasound imaging, that the pronunciation of the illegal sequence /zC/ in English by native speakers of English did not have a tongue motion targeting a schwa gesture. Rather, the consonant gestures were pulled apart in time, causing a period where there is an open vocal tract. This open transition gives rise to a percept of a schwa-like vowel, previously termed a transitional vowel, an intrusive vowel, an excrescent vowel or a transitional vocalic element (Davidson & Stone, 2003; Gafos, 2002; Yanagisawa, 2005). Thus Davidson (2003) and Davidson and Stone (2003) argue that a vowel perceived between illegal clusters in native language is not necessarily a phonological epenthetic vowel, but a schwa-like vocoid caused by gestural mistiming.

Intrusive vowels have been observed not only in non-native production of foreign clusters but also in native production. Hall (2006) presents a list of languages in which vowel intrusion arises. She describes intrusive vowels as follows; they (i) tend to be shorter than lexical vowels, and to disappear in fast speech, (ii) have either schwa-like features or copy the features of a neighboring vowel, (iii) appear most frequently in heterorganic clusters and (iv) are often invisible to phonological patterns that depend on the presence of vowels (Davidson, 2006b; Hall, 2006; Matteson & Pike, 1958).

Hall argues that an intrusive vowel without its own articulation does not form a syllable nucleus, based on evidence that intrusive vowels do not pattern with full vowels. For instance, in the Siouan language Hocank, intrusive vowels occur in onset clusters where the second consonant is a sonorant(/ʃwaʃi/ → [ʃ**a**waʃi] ‘you dance’). In reduplication, the last syllable is normally copied (gihu ‘swing’, gihuhu ‘wagtail’, Hall, 2006, p.402). But if a stem contains an intrusive vowel in the first syllable, the entire form is reduplicated (p**a**ras ‘flat’, p**a**rap**a**ras ‘wide’), indicating that the whole form is treated as one syllable. On the basis of such evidence, Hall proposes that intrusive vowels are distinct from full vowels and are percepts resulting from a particular timing of articulatory gestures.

Following these studies, I now discuss the way in which gestural representations can represent the differences between intrusive vowels and phonological epenthetic vowels.

(1) Landmarks in gestural representation (from Gafos, 2002)

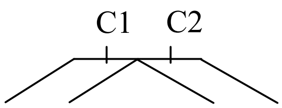


(1) illustrates abstract gestural landmarks as discussed by Gafos (2002). The onset is the beginning of the movement of an articulator toward a target constriction; the target is the point at which the intended constriction is accomplished; the center is a phase during which the constriction is maintained; the release represents the point when the articulator moves away from the constriction; and the offset is the point at which the articulation is completed.

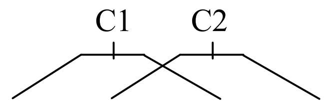
Rather than occurring as discrete units, speech gestures are continuously combined (coarticulated). When two consonant gestures are highly overlapped, as shown in (2a), the alignment between the release of the first consonant and the target of the second consonant gives rise to a lack of audible release of the first consonant (=close transition, following Catford, 1977, or release=target phasing, following Gafos, 2002). But when the phasing relation is looser, as shown in (2b), an acoustic release of the first consonant is present (open transition or center=onset phasing). Gafos (2002) demonstrated using the GEST model of gestural dynamics that a percept of a schwa-like vocalic element arises from a configuration in which the release of the first consonant is not obscured by the target of the second consonant, as shown in (2b). During this open transition, if the adjacent vowel gesture overlaps the period of release, the intrusive vowel is perceived as a copy vowel. If the tongue body is in a more neutral position during this period, the percept of the vowel is a schwa-like vocalic element (Hall, 2006).

(2) Inter-segmental timing relations

(a) Close transition, release=target



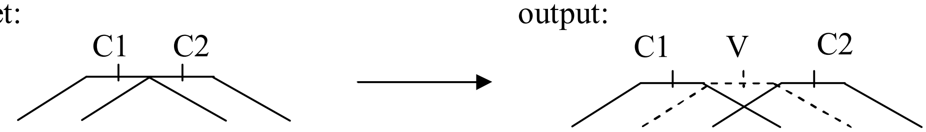
(b) Open transition, center=onset
Intrusive Vowel



On the other hand, a phonological epenthetic vowel is realized with its own gesture inserted, as shown in (3a). A vowel gesture illustrated with a dotted line is inserted between the two consonant gestures.

(3) Phonological epenthesis (true epenthetic vowel) from Davidson (2006b)

target: output:



Having described the differences between epenthetic vowels and intrusive vowels in gestural relations, I next investigate the actual status of the inserted vowels found in the current production experiments.

6.2 Comparison between Lexical and Inserted Vowels

Davidson (2006a) compared the acoustic characteristics of inserted schwas and lexical schwas produced by English native speakers, and found significant differences between the two types of vowels. Duration was longer and both F1 and F2 were higher for lexical schwas than for inserted schwas. These acoustic characteristics suggest that the inserted vowels did not have a gestural target, unlike lexical vowels. Based on these acoustic characteristics of inserted vowels, Davidson argues that the inserted vowels in the production of foreign clusters are not phonological but intrusive.

Relying on the fact that intrusive vowels and lexical vowels present significantly different acoustic characteristics, we define the status of the inserted vowels in the current production experiments by comparing them to the lexical vowels. In the next section, I compare the duration of lexical vowel [ɨ] to the duration of the inserted vowel [ɨ] in the same context in order to determine whether the inserted vowel is phonological or intrusive: if the duration of the inserted vowel is longer than the duration of the lexical vowel in the same context for the same speaker, the inserted vowel can be considered a product of true phonological vowel epenthesis, but if the inserted vowel is shorter than the lexical vowel, then it can be considered a reflection of vowel intrusion. Although it would be better if we could compare formants of the two types of vowels as well, it was impossible to control the surrounding vowels and consonants perfectly due to the limited existing lexical words. Thus we only rely on the duration of the vowel, which seems to be a good indicator of the status of inserted vowels (Davidson, 2006a).

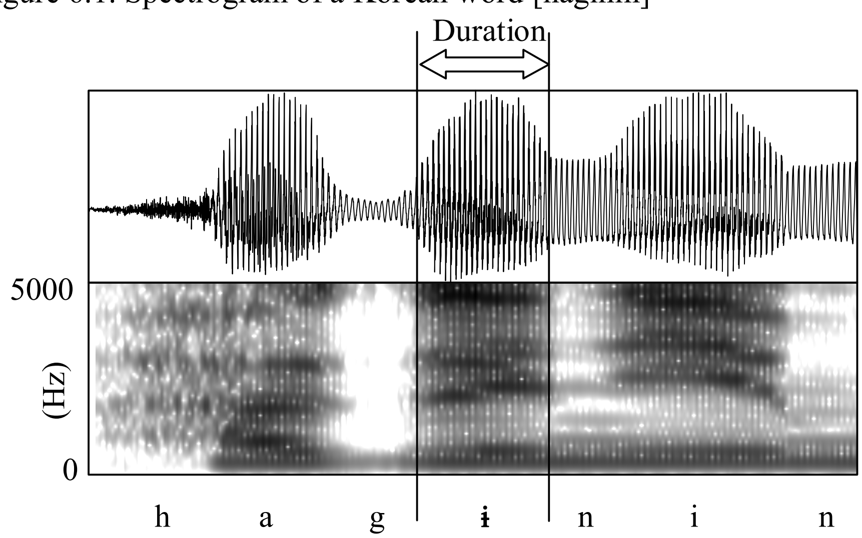
6.2.1 Data collection: Lexical [ɨ]

From 18 subjects, who participated in the reading experiment, I obtained acoustic data on the duration of the lexical vowel [ɨ] in Korean. The materials consisted of two sets of native Korean words and established loanwords, all of which contained the vowel [ɨ] in the second syllable. All words contained exactly 3 syllables. There were 24 words in total: 2 sets of 6 Korean native words that contained lenis stop ([p,t,k]) before the vowel and 2 sets of 6 loanwords that contained aspirated stops ([p^h,t^h,k^h]). The list of the words is in Appendix 6. Using these words, a short story was created and the words were embedded in phrase-initial position. The story was written in Korean orthography and participants were not aware of the purpose of the experiment. They were instructed to read the story quietly before the recording began, and the session did not begin until they became comfortable reading it out loud. They repeated the story twice. The rate of speech was comparable to the one used in the reading experiment of English passages (See Appendix 1).

6.2.2 Duration measurements

The total number of words was 864 words (2 lexica (native and loanwords) * 2 words * 2 aspiration *3 place * 2 repetition * 18 subjects), but only 781 words were included for the duration analysis. The excluded words were either words in which the boundary of the vowel was hard to determine or words that were produced incorrectly by the participant. The duration measurement was done in Praat (Boersma & Weenink, 2007). The onset and offset of the vowel were mainly based on the onset and offset of F2. When F2 was not a clear cue, a point where more than two formants appeared or disappeared together was taken as the onset or the offset of the vowel.

Figure 6.1. Spectrogram of a Korean word [hagɨnin]

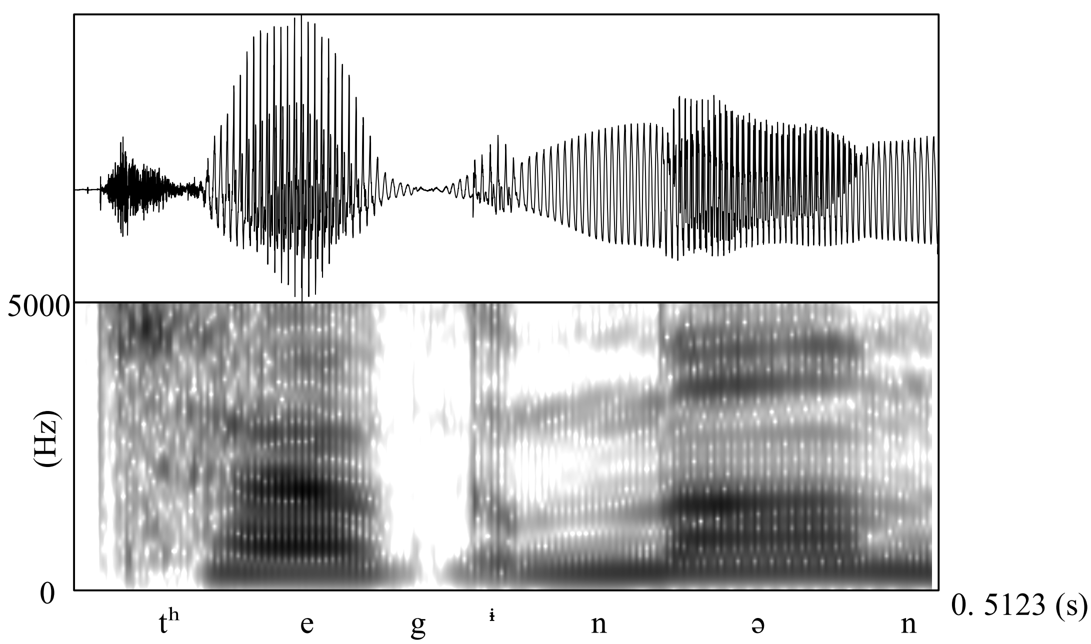


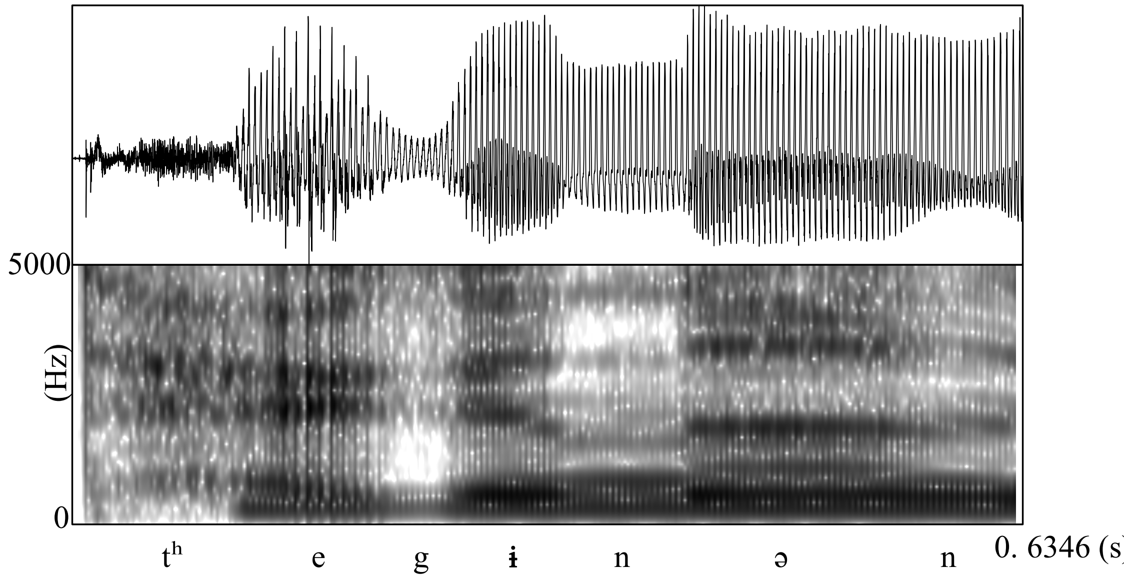
The items with inserted vowels from the reading experiment were also measured for vowel duration. There were 182 words in total from 14 subjects; I collected 18 subjects' data, but 4 of them did not insert a vowel in the reading experiment. Not every subject inserted a vowel in every context; some inserted a vowel only after voiced stops, some inserted a vowel only after velar stops, and some inserted a vowel only after voiced velar stops. The mean of the lexical vowels was calculated only for the words that had corresponding items that shared the preceding consonant with the inserted vowels. For instance, if a speaker inserted a vowel in 'tebnal' but not in 'tepnal', then the mean duration of the vowel in Korean words [pabɨɾo] and [sabɨɾo] was calculated and compared with the vowel inserted in 'tebnal', but not with the lexical words that contain [p^h] before [ɨ].

6.2.3 Results

Based on each subject's mean duration of the lexical vowel [ɨ] in 6 contexts, the inserted vowel in the English nonsense words from the reading experiment was defined as either an intrusive vowel or a true epenthetic vowel. If the duration of the inserted vowel was longer than the mean of the lexical vowel in the same context, it was coded as a true epenthetic vowel, but if it was shorter, it was coded as an intrusive vowel. This coding was done for each subject who completed the reading experiment and the Korean short story reading task. Figure 6.2 shows that an inserted vowel coded as intrusive is much shorter in duration than an inserted vowel coded as epenthetic.

Figure 6.2. Spectrograms illustrating an intrusive vowel and an epenthetic vowel from Korean speakers' production of English nonsense word [t^hegnən]





This coding makes two predictions regarding the inserted vowels coded as epenthetic. In Chapter 3, the perception experiments showed that Korean listeners perceived illusory vowels in certain contexts. I argue that these misperceived vowels are equivalent to true epenthetic vowels defined in this chapter because, if inserted in perception, they should be already present in the input of the production grammar, and should therefore manifest the properties of full lexical vowels. Now, recall that the voicing, but not the place, of the stop was likely to induce Korean listeners' perception of an illusory vowel in stop-nasal sequences. If the perceptually inserted vowels are indeed realized as phonological epenthetic vowels in production, the same pattern regarding the voicing and the place of the stop should arise in their production. In other words, we predict that the vowels coded as phonological epenthetic vowels (1) should be more frequent in voiced stop-nasal sequences than voiceless stop-nasal sequences and (2) should not be differentially frequent depending on the place of the stop (labial vs. velar).

Table 6.1 shows the number of epenthetic vowels and their mean duration. As predicted, epenthetic vowels were more frequent in voiced stop-nasal sequences than in voiceless stop-nasal sequences, but they were similarly frequent when the stop was either bilabial or velar. This result is consistent with the perception experiment results that demonstrated a voicing effect but not a place effect in perception of a vowel in stop-nasal sequences, supporting the two predictions made above. The distributional similarity between illusory vowels in perception and epenthetic vowels in production strongly suggests that the some errors in production are indeed a result of misperception.

Table 6.1. Vowel epenthesis rates and mean duration

	Voiceless stop-nasal		Voiced stop-nasal	
	Bilabial	Velar	Bilabial	Velar
Vowel Insertion	3	2	12	15
duration (SD)	55.9ms (16.6)	72.2ms (20.3)	82.7ms (18.4)	61.7ms (17.4)

Now, let's turn to intrusive vowels. Overall, inserted vowels coded as intrusive were more frequent than vowels coded as epenthetic. Table 6.2 presents the number of cases of intrusive vowels for each context and the mean duration of those vowels.

Table 6.2. Vowel intrusion rates and mean duration

	Voiceless stop-nasal		Voiced stop-nasal	
	Bilabial	Velar	Bilabial	Velar
Vowel Intrusion	6	21	22	43
duration (SD)	36.3ms (5.7)	27.8 ms (8.7)	48.2 ms (14.4)	39.1 ms (10.8)

These results reveal two distributional asymmetries. First, the number of intrusive vowels is much greater in velar stop-nasal sequences than in labial stop-nasal sequences. Second, intrusive vowels were found more frequently in voiced stop-nasal sequences than in voiceless stop-nasal sequences. These two asymmetries remain to be explained. However, having excluded the true epenthetic vowels associated with misperception enables us to examine the possible sources of those asymmetries in articulatory patterns. In the next section, I provide a gestural account for these two asymmetries.

6.3 Vowel Intrusion and Gestural Overlap

A crucial condition for the emergence of vowel intrusion is an intergestural phasing relation in which the release of the stop in a stop-nasal sequence is not obscured by the following nasal. This amounts to saying that the likelihood of the stop release in stop-nasal sequences correlates with the likelihood of vowel intrusion. In light of this connection between stop release and vowel intrusion, I investigate why vowel intrusion was more frequent (i) in voiced stop-nasal sequences than in voiceless stop-nasal sequences and/or (ii) in velar-nasal sequences than in labial-nasal sequences

6.3.1 Voicing effect in vowel intrusion

One reason for the greater frequency of intrusive vowels in voiced stop-nasal sequences than voiceless stop-nasal sequences could be that the gestural coordination of stop and nasal consonants likely differs depending on the voicing of the stop: open transition for voiced stop-nasal sequences but closed transition for voiceless stop-nasal sequences. But another possible reason could be that the two types of sequences involve the same gestural coordination but are different in terms of acoustic realization. In this section, I argue that the greater frequency of vowel intrusion in voiced stop-nasal sequences than in voiceless stop-nasal sequences is because the transition from a voiced stop to a nasal has the characteristics of a vocalic element, rather than because voiced stop-nasal sequences more frequently involve open transitions than do voiceless stop-nasal sequences.

Gafos (2002) puts forward a model where the same pattern of gestural coordination may not necessarily result in the same acoustic realization, depending on the nature of the consonants involved in the clusters. Gafos (2002) discusses cases in which variable surface forms such as aspiration or intrusive vowel may arise from the same timing relation. For instance, in a language like Sierra Popoluca, in which heterorganic clusters require open transition between the consonants, it has been claimed that an intrusive vowel arises after a nasal (voiced) consonant ([miɲə̠pa] ‘he comes’), whereas aspiration occurs after a voiceless stop ([kek^h̠paʔ] ‘it flies’) in the clusters (Elson, 1947).

According to this model, the period of release from the open transition may be perceptually similar to a vowel after a voiced stop, since this release will be voiced, and perceptually similar to aspiration after a voiceless stop. If this is true, the higher rate of vowel intrusion in English voiced stop-nasal sequences pronounced by Korean speakers may simply reflect the English-speaking transcribers’ judgment that the release of a voiceless stop constituted correct production rather than an intrusive vowel, while the release of a voiced stop was more likely to be perceived as an inserted vowel because of it shares the property of voicing with a true vowel. This would explain why we see more frequent vowel intrusion in voiced stop-nasal sequences.

This account can be tested by looking at both the items transcribed as correct and the items transcribed as containing an inserted vowel. If the pulled-apart gestural organization between the stop and the nasal results in variable acoustic realization depending on the voicing of the stop, we should find higher rates of release burst than of vowel intrusion after voiceless stops, but the reverse after voiced stops. To test this prediction, the frequency of release bursts was coded for the items classified as correctly produced.

The stop release burst was indeed often followed by a period of voicelessness (aspiration) after voiceless stops, as illustrated in Figure 6.3. In contrast, the release burst after voiced stops tended to be weaker in amplitude and shorter in duration.

Figure 6.3. Release burst in [k^{h}ik^{h}nəl]

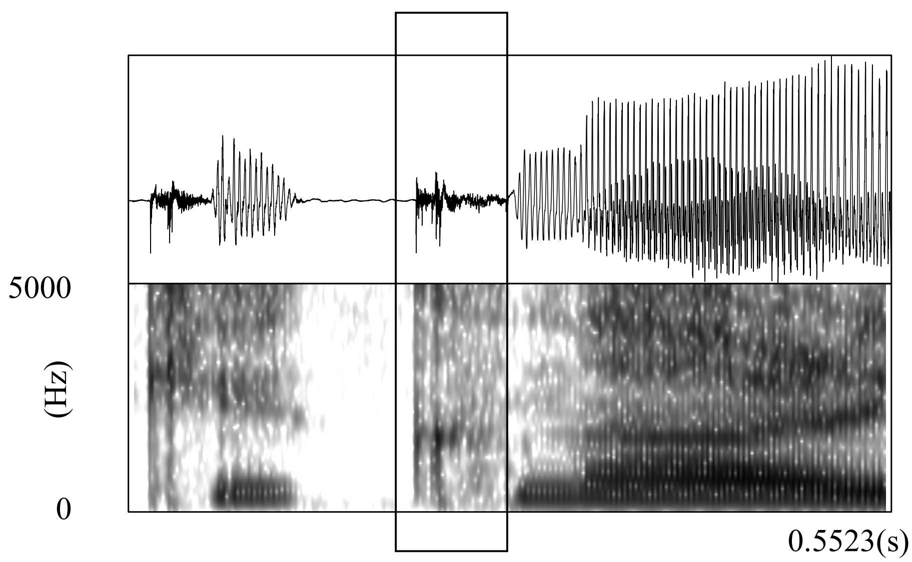
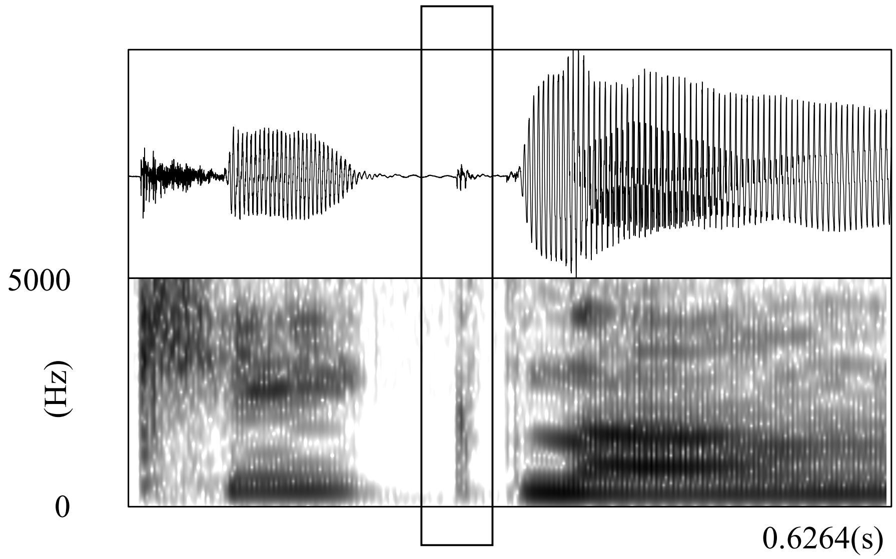
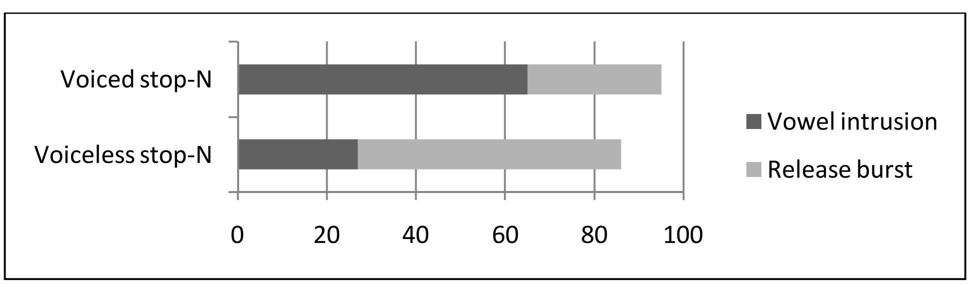


Figure 6.4. Release burst in [k^hignən]



In Figure 6.5, the frequency of release bursts and vowel intrusion is plotted together. The sum of the frequency of vowel intrusion and release burst in voiceless stop-nasal sequences is in fact comparable to the sum of the frequency of vowel intrusion and release burst in voiced stop-nasal sequences. This suggests that the open vocal tract between the two consonantal gestures produced transitional acoustic variants, and the quality of these variants largely depended on the voicing of the stop. A strong release burst with aspiration was more likely after a voiceless stop, whereas an intrusive vowel was more likely after a voiced stop. In other words, the reason we had more frequent vowel intrusion in voiced stop-nasal sequences than in voiceless stop-nasal sequences is not because voiced stop-nasal sequences more frequently involve open transition but because this open transition more frequently has the properties of a vocalic element (therefore, intrusive vowel). This suggests that the two kinds of sequences are different not at the level of gestural relation but at the level of acoustic realization.

Figure 6.5. Vowel intrusion and release burst rates from labial-nasal and velar-nasal sequences by voicing of the stop[15]



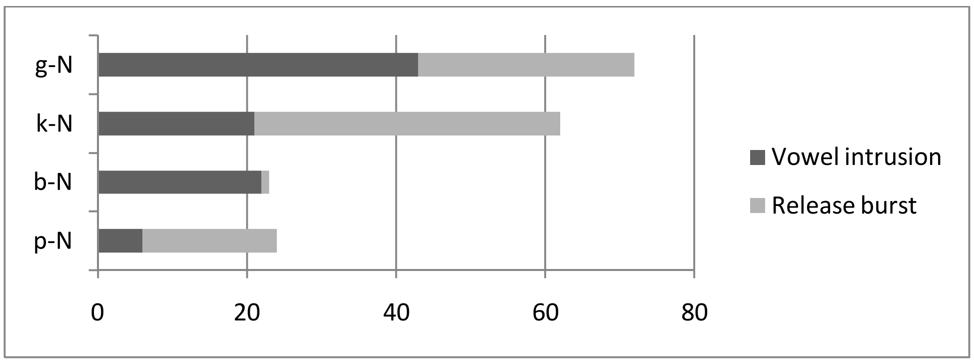
[15] The rates of the release burst were obtained from the eighteen participants whose data were coded for vowel intrusion vs. vowel epenthesis. The other three participants did not complete the task to collect lexical vowel [ɨ]. All the subsequent analyses were based on those 18 subjects.

In the following section, I move on to the second asymmetry, namely the place effect we found in vowel intrusion.

6.3.2 Place effect in vowel intrusion

Intrusive vowels were considerably more frequent in velar stop-nasal clusters than in labial stop-nasal clusters, regardless of the voicing of the stop. This still holds when we consider the frequency of release bursts as well, as shown in Figure 6.6. The overall frequency of vowel intrusion and release bursts was higher in velar stop-nasal sequences than in labial stop-nasal sequences.

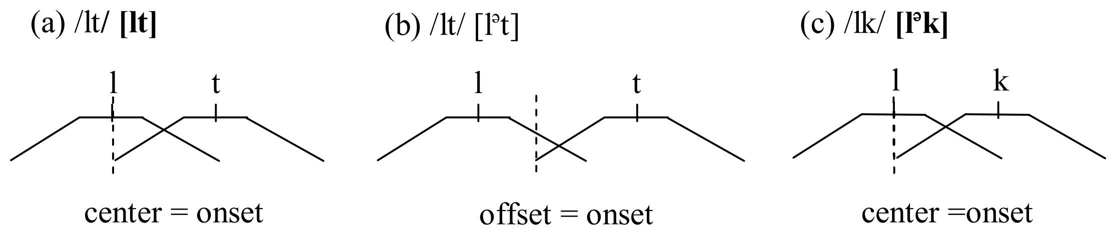
Figure 6.6. Frequency of vowel intrusion and release bursts by place and voicing of the stop



In this section, I propose that a gestural account can also explain the place asymmetry. I show that the likelihood of intrusive vowels is largely attributable to whether the sequences are homorganic or heterorganic. All the velar-initial sequences were heterorganic but only half the labial-initial sequences were heterorganic. Specifically, I argue that velar stop-nasal clusters are more likely to involve vowel intrusion or release burst because they are heterorganic clusters, which are more likely to be released than are homorganic clusters.

Given the same open transition, release of the first consonant is mainly determined by whether it shares the place of articulation with the following consonant or not (Catford, 1977; Côté, 2000; Gafos, 2002; Gafos, Hoole, Roon & Zeroual, 2010; Ladefoged, 1993). For instance, Gafos (2002) demonstrates that an open transition (specifically the phasing in which the center of the first consonant is aligned with the onset of the second consonant) produces a schwa-like element in heterorganic clusters such as /lk/, but not in homorganic clusters such as /lt/, using computational simulations with a model of gestural dynamics. This is so because in a center=onset phasing, by the time the articulator of the first gesture is released, the same articulator has already started its next constriction, as shown in (4a). In this case, the articulator stays in the same place. Thus, unless the gestures of /l/ and /t/ are much more temporally separated as shown in (4b), there will be no acoustic release between the homorganic consonants. But in heterorganic clusters, the same center=onset phasing produces an audible release or an intrusive vowel during the transition between the release of the first consonant and the target of the second consonant, as shown in (4c).

(4) Phasing relations in homorganic and heterorganic clusters



Based on the fact that the same phasing relations generate different acoustic results depending on the hetero- or homorganicity of the clusters, Hall (2006) argues that intrusive vowels should occur more frequently in heterorganic clusters, and she presents a list of languages including Sierra Popoluca and Dutch in which vowel intrusion occurs almost exclusively in clusters of heterorganic consonants. Zsiga (2000) also reports experimental results indicating that both English and Russian clusters at word boundaries were much more likely to be unreleased when they were homorganic than when heterorganic. Similarly, in her discussion of the audibility of stop release burst, Côté (2000) asserts that "articulatorily unreleased stops occur before homorganic nasal or oral stops" (p. 142).

Thus, the place asymmetry observed in the production data (the greater likelihood of a vowel in velar-nasal than in labial-nasal sequences) can be directly explained by the fact that half of the labial stop-nasal sequences were homorganic but all velar stop-nasal sequences were heterorganic. Homorganic labial stop-nasal sequences were much less likely to be released or to be produced with an intervening intrusive vowel. On the other hand, being heterorganic, velar stop-nasal sequences are much more likely to be realized with intrusive vowels or release bursts[16].

While the difference between heteroganic and homorganic sequences is clearly a factor, the data do allow us to compare velar-initial and labial-initial sequences that share the property of heterorganicity:

(5) Velar stop-nasal vs. labial stop-nasal sequences

	Homorganic sequences	Heterorganic sequences
Labial stop-nasal	-pm-, -bm-	-pn-, -bn-
Velar stop-nasal	-----------	-km-, -kn-, -gm-, -gn-

Excluding homorganic sequences, vowel intrusion or release burst is still more likely to occur in velar stop-nasal sequences than labial stop-nasal sequences. I will argue that this occurs because in the participants' native Korean, labial-nasal sequences show more gestural overlap

[16] Park (2006) also found that a coronal stop before a coronal nasal was mostly nasalized whereas a coronal stop before a labial nasal was more likely to be produced with an inserted vowel. Although he did not examine other places of articulation, this result is consistent with the findings that homorganic clusters are less likely to be released than heterorganic clusters.

than velar-nasal sequences. Different degrees of gestural overlap between back-to-front stop clusters and front-to-back stop clusters have been found in other languages; for example, EMMA (Electromagnetic Midsagittal Articulometer) magnetometer studies of Georgian revealed that back-to-front stop clusters [kt] exhibited less overlap than front-to-back [tk] (Chitoran, Goldstein & Byrd, 2002), and the same place order effect was also found in Georgian stop-liquid clusters (Chitoran & Goldstein, 2006). A similar tendency for greater overlap in front-to-back clusters has been found in other languages (Byrd, 1996; Gafos et al., 2010; Kühnert, Hoole & Mooshammer, 2006; Zhao & Stevens, 2003; Zsiga, 1994), although Zsiga (2000) found that the first consonant in clusters distributed across English word boundaries is more likely to have an audible release if the first consonant is further forward than the second consonant—that is, that release was more likely in a front-to-back sequence. Therefore, mixed results regarding the place order effect suggest that languages may differ in how place and order affect overlap, and we need language-specific evidence from articulatory data for possible different overlap patterns in Korean.

Temporal phasing patterns in Korean stop-stop sequences were found in a cross-linguistic examination of gestural temporal overlap in consonant clusters (Kochetov, Pouplier & Son, 2007). In this study, degree of overlap in Korean and Russian /pt/, /kt/ and /kp/ was assessed by measuring Plateau Lag (PL) from the release of the first stop to the achievement of target for the second stop. Positive PL values correspond to the open transition that arises when the target constriction of C2 is accomplished after the release of C1. PL values at 0ms correspond to a close transition in which the release of C1 is aligned with the target of C2. Negative PL values represent overlap, where the target of C2 is already attained before the release of C1. Thus, smaller numbers mean greater temporal overlap. Kochetov et al. found that the PL values depended on the type of cluster: PL values were significantly larger for /kp/ and /kt/ than for /pt/. This means that k-initial clusters were much less overlapped than p-initial clusters. Zsiga (2011) shows that the majority of the Korean speakers had a significant correlation between stop-stop gestural timing and nasalization in stop-nasal sequences in their production of English. That is, the Korean speakers with higher overlap in English stop-stop sequences are more likely to nasalize a stop before a nasal. On the assumption that nasalization is a process resulting from gestural overlapping, this result suggests that timing patterns in stop-stop sequences found in Kochetov et al's (2007) study can be extended to stop-nasal sequences.

In learning English, Korean L2 learners of English must learn how to produce stop-nasal sequences without nasalizing the stop. In doing so, gestural mistiming can occur. Given the different degrees of temporal overlap between the two types of native clusters in Korean, the same degree of gestural mistiming will have different resulting configurations: the stop from originally highly overlapped labial stop-/n/ sequences will still be less likely to release than the stop from originally less overlapped velar stop-N sequences. Thus, the fact that intrusive vowels and release bursts are more common in velar-nasal sequences simply reflects the different degrees of temporal overlap between the two types of clusters in Korean.

(6) Gestural overlapping in Korean word internal clusters

	Gestural Overlapping	Release of stop
-pn-, -bn-	More overlapped (Kochetov et al. 2007)	Less likely to release
-km-, -kn-, -gm-, -gn-	Less overlapped (Kochetov et al. 2007)	More likely to release

In the next section I turn to the discussion of devoicing in labial-nasal sequences.

6.4 Devoicing and Gestural Overlap

Another puzzle regarding the place of the stop in stop-nasal production is not only that vowel intrusion/release bursts arose commonly in velar-nasal sequences, but also that devoicing arose frequently in labial-nasal sequences. This pattern is again illustrated in Table 6.3. In Chapter 4, I showed that the phonological grammar cannot readily account for this fact since there appears no obvious evidence to motivate this pattern in either of the languages in contact. Furthermore, the perception experiment results of Chapter 3 did not support an analysis of this pattern as an effect of misperception.

Table 6.3. The rates of vowel intrusion and devoicing by voicing and place of the stop[17]

	Voiceless stop-nasal		Voiced stop-nasal	
	Bilabial	Velar	Bilabial	Velar
Vowel Intrusion	6	21	22	43
Release burst	18	41	1	29
Devoicing	-	-	34	22

I will argue that this pattern also follows from the fact that labial stop-nasal sequences are much more overlapped than velar stop-nasal sequences in Korean. In other words, vowel intrusion is preferred in /g.m/ and /g.n/ sequences because the stop is more likely to be released before a heterorganic nasal, due to the low degree of overlap. Devoicing is preferred in /b/-nasal sequences since the homorganic labial stop-nasal sequences are produced without release in the first place: the articulators involved in those sequences allow the clusters to be overlapped enough to decrease the chance of release. The difficulty of sustaining voicing during the closure of the stops that remained unreleased results in devoicing of the stop. To argue for this, I discuss the relationship between voicing and stop release as well as the relationship between devoicing and non-release of the stop.

6.4.1 Voicing and release

Voicing and release of a stop are closely related from an aerodynamic point of view, especially in a voiced cluster. Sustaining voicing requires sufficient difference in air pressure above and below the glottis. It is challenging to produce voicing when the segment is an

[17] Again, the rates of devoicing are only from the eighteen participants whose production data were classified into vowel insertion or vowel intrusion.

obstruent because intraoral air pressure quickly builds up during constriction, making it harder to maintain transglottal airflow (Johnson, 2003; Westbury, 1983). In voiced clusters, particularly when the two gestures are closely coordinated and the first consonant is unreleased, sustaining voicing for longer periods is especially challenging (Blevins, 2006; Ohala, 1997; Westbury & Keating, 1986). But a pulled-apart coordination in this type of cluster will facilitate maintaining voicing of the two consonants by allowing a release or an intrusive vowel. Otherwise, the stop is likely to become devoiced.

Davidson (2006b) proposes a gestural constraint that prohibits close temporal coordination between a voiced obstruent and a nonapproximant (obstruent or nasal) in English, precisely relying on this articulatory difficulty. She reported that non-native onset clusters for English native speakers such as /ft/, /zm/ or /vg/ were frequently produced with an intrusive vowel, and more often when the clusters were voiced than voiceless. She argues that vowel intrusion in voiced clusters helps maintain voicing in voiced clusters.

Now, considering the fact that a nasal stop is involved in the sequences in question, this explanation appears to be problematic at first when we look at the cluster with a reversed order of stop and nasal, because it is widely known that post-nasal stops tend to be voiced (Hayes, 1996; Pater, 1996). If our participants inserted intrusive vowels in voiced stop-nasal sequences because of articulatory difficulty with voicing, it seems counterintuitive that voiceless stops undergo voicing in nasal-stop sequences. In other words, in the same nasal neighboring context, it is not clear why stop voicing is not preferred in pre-nasal position whereas it is in post-nasal position. However, there is a cross-linguistic tendency that pre-nasal voicing is much rarer than post-nasal voicing (Pater, 1996).

Hayes and Stivers (1995) and Hayes (1996) provide a detailed explanation for this asymmetric patterning on articulatory grounds, essentially showing that coarticulation with nasal facilitates voicing of a post-nasal stop but not of a pre-nasal stop. They present two factors specific to obstruents adjacent to nasals. The first factor called ‘nasal leak’, occurs through nasal coarticulation in which the obstruent has a slightly lowered velum in transition to the neighboring nasal. Air leaks through the velar port, lowering the supraglottal air pressure and thus, encouraging voicing. The second factor is “velar pumping” which has two distinctive effects on the size of the supraglottal cavity during stop articulation, depending on the position of the nasal. For the post-nasal stop, *rarefactory velar pumping* occurs from a transition where the velum continues to rise toward the high target position for an obstruent (Bell-Berti, 1975). Even after the velum is high enough to block nasal leak, the velum rise enlarges the supraglottal cavity, preventing buildup of supraglottal pressure and thereby facilitating voicing. In contrast, the stop-to-nasal transition involves *compressive velar pumping*: the velum moves toward a lower position from the higher position made for the stop in anticipation of nasal articulation, compressing the supraglottal cavity. This results in higher supraglottal pressure, which impedes voicing. Putting all together, both nasal leak and velum raising in a nasal-to-stop transition favor voicing, supporting the cross-linguistic pattern of post-nasal voicing. In a stop-to-nasal transition, on the contrary, velum lowering that impedes voicing is immediately followed by an opposing effect, nasal leak, which makes this position not particularly suitable for voicing. (7) summarizes nasal leak and pumping effect for pre- and post-nasal obstruent. Hayes and Stivers (1995) also demonstrate that a vocal tract model (Westbury & Keating, 1986) produces post-nasal obstruents with full closure voicing whereas pre-nasal obstruents with considerably longer period of voiceless interval than voicing interval.

(7) Nasal leak and pumping effect for pre- and post-nasal obstruent from Hayes and Stivers (1996) and Hayes (1996)

<table>
<tr><td colspan="2"></td><td colspan="3">Time →</td></tr>
<tr><td rowspan="3">Post-nasal obstruent</td><td>Velum</td><td>slightly open</td><td colspan="2">closed but not all the way up ⇒ closed all the way up</td></tr>
<tr><td>Effect</td><td>Nasal Leak</td><td colspan="2">Rarefactory velar pumping</td></tr>
<tr><td>Voicing Preference</td><td>Yes</td><td colspan="2">Yes</td></tr>
<tr><td rowspan="3">Pre-nasal obstruent</td><td>Velum</td><td colspan="2">closed all the way up ⇒ closed but not all the way up</td><td>slightly open</td></tr>
<tr><td>Effect</td><td colspan="2">Compressory velar pumping</td><td>Nasal Leak</td></tr>
<tr><td>Voicing Preference</td><td colspan="2">No</td><td>Yes</td></tr>
</table>

Taken together, there are well-justified phonetic reasons why stop voicing is difficult before a nasal (but not after a nasal), especially when the cluster is closely coordinated. Without an intervening intrusive vowel, the voiced stop before a nasal may be devoiced to ease articulatory complexity.

Different phonological patterning depending on whether the consonant is released or not is certainly not limited to interlanguage phenomena. In fact, in some languages, released stops and unreleased stops in the exact same prosodic contexts behave distinctively. Seo (2006) presents two such cases. In Choctaw, for example, /k/ is always released before a nasal or a liquid and the release is often accompanied by a very short copy vowel of the preceding vowel, which can be seen as an intrusive vowel. On the other hand, /p/ and /t/ are not released in the same context (Nicklas, 1974). At the surface, while released /k/ is realized with an intrusive vowel, unreleased /t/ or /p/ are modified in a way to conform to the phonological constraints in the language (e.g., /ta**pl**i/ →[ta**bl**e] 'to cut in two', /aka**tl**i/ → [aka**ll**i] 'to patch' vs. /ha**kl**o/→[ha**k**^{a}**l**o] 'to hear'). Likewise, /p/ and /t/, which are always unreleased in Toba Batak, change to /k/ before a sonorant, while /k/, which is invariably released, does not undergo any change in the identical condition (Nababan 1981). These examples suggest that release bursts are a very active phonological conditioning factor in some languages, and the Korean-English interlanguage patterns revealed the important role of release burst in determining repair choices.

Similarly, Rhee (1998) points out that voiced stops should be licensed by release based on the fact that maintaining transglottal airflow can directly benefit from opening the supralaryngeal cavity sooner[18]. Release of a stop supports the voicing of that stop because the movement towards release has the effect of opening the supraglottal cavity, making it larger and thus lowering air pressure. Thus, he considers stop release as “an active articulatory gesture that enables air pressure to be differentiated between supra- and subglottal cavities” rather than “a passive product of relaxation to a more open configuration” (p. 123). He further argues that several languages in which post-vocalic stops are strictly unreleased devoice coda voiced stops, not because [voice] is not licensed by a following [+son] such as a vowel, as proposed by Lombardi (1995) but because voicing must be licensed by release. In other words, Rhee emphasizes the role of release, rather than the syllabic position of the consonant, to license stop voicing. His evidence for devoicing in nonreleasing languages is listed in (8). In Sino-Tibetan languages including Mishmi, Boro, Kinnauri, West Tarangan, Efik, and Ibibio, stops are strictly unreleased in coda, whether word-finally or pre-consonantally, and they become devoiced in those contexts.

(8) Devoicing in nonreleasing languages from Rhee (1998, p. 131)

Boro		
zab	[zap˺]	‘to heap up’
zab-kru	[zap˺kru]	‘to heap up everything’
West Tarangan		
RED + jaban	[jip˺jabən]	‘RED-dry’
RED + kudam	[kit˺kudəm]	‘cloud’
Ibibio		
deeb	[deep˺]	‘scratch’
ndeebe	[ndeebe]	‘I have not scratched’
Mishimi		
kab	[kap˺]	‘bulbul’
kab-welaŋ	[kabwelaŋ]	‘bulbuls’
Kinnauri		
əg	[ək˺]	‘cave’
əg-u	[əgu]	‘the cave’

Rhee (1998) also summarizes Parker’s (1981) argument relating final devoicing to nonrelease. Parker illustrates five stages regarding voicing and devoicing. The crucial stages for our purpose are Stage IV and V, which involve loss of [ə] in CVCə. This loss in turn causes final voiced stops to be unreleased, and consequently devoice. Parker’s account of devoicing of final stops as

[18] Although there is another strategy often used to sustain voicing during closure-- expansion of the pharyngeal cavity via larynx lowering, jaw lowering or velum elevation, which all serve to lower the supraglottal air pressure (McGowan and Saltzman 1995, Ohala 1997, Westbury 1983)-- Rhee relies on release only.

a result of losing stop release highlights Rhee (1998)'s point that release plays a critical role in coda stop devoicing.

(9) Nonrelease and devoicing stages from Parker (1981)

I – C V C V → unmarked stage
[-vce]

II – C V C V → intervocalic voicing
[+vce]

III – C V C ə → unstressed vowel reduction
[+vce]

IV – C V C˺ → loss of [ə] causes final voiced stop unreleased
[+vce]

V – C V C˺ → devoicing in unreleased position
[-vce]

Within these accounts, the devoicing in labial-nasal sequences is well motivated: the smaller likelihood of release of labial stops leads to frequent devoicing but the greater likelihood of release of velar stops leads to frequent vowel intrusion or release bursts. In the next section, I provide a formal analysis of the asymmetries found in the production experiment, employing gestural constraints that incorporate the timing dimension into the phonological grammar, and argue that timing relations in clusters determine whether the first consonant is released, and in turn, the pronunciation patterns seen in Korean learners' production of English stop-nasal sequences.

6.5 Interlanguage grammar with gestural constraints in OT

Several studies have adopted gestural constraints in the framework of Optimality Theory (Benus, Smorodinsky, & Gafos, 2004; Borroff, 2007; Davidson, 2003; Davidson, 2006b; Gafos, 2002; Hall, 2003). These studies borrowed from Articulatory Phonology (Browman and Goldstein 1986, *et seq.*) the assumption that phonological representations are specified in terms of gestures and the timing relations among them. In this section, I adopt Gafos (2002)'s ALIGN and Davidson (2006b)'s RELEASE constraints to show how intrusive vowels or release bursts occur between the stop and the nasal in the grammar.

Intergestural timing relations can be expressed via ALIGN constraints that specify landmarks of gestures and their organization, as shown in (10).

(10) Gafos 2002, p.284

ALIGN (Gesture$_1$, landmark$_1$, Gesture$_2$, landmark$_2$): Align landmark$_1$ of gesture$_1$ to landmark$_2$ of gesture$_2$.

I propose that in Korean, ALIGN (C$_1$, release, C$_2$, target) is active, based on the articulatory data displaying high interconsonantal overlap, compared to other languages such as Russian (Kochetov et al., 2007). However, I posit two separate ALIGN constraints to take into account the significant differences in overlap between back-to-front clusters ('kt') and front-to-back clusters ('pt') in Korean.

(11) ALIGN constraints in Korean
ALIGN (K, release, T, target): Align the release of velar stop with the target of coronal stop.
ALIGN (P, release, T, target): Align the release of labial with the target of coronal stop.

We have seen from Kochetov et al. (2007) that Korean /kt/ was much less overlapped than 'pt', having a positive Plateau Lag, while the transition between /p/ and /t/ was always closely reflected in a negative Plateau Lag. This difference can be expressed by ranking ALIGN (P, release, T, target) higher than ALIGN (K, release, T, target): it is more critical to have close transition between labial and coronal stops than between velar and coronal stops.

(12) The ranking between the two ALIGN constraints in Korean
ALIGN (P, release, T, target) >> ALIGN (K, release, T, target)

While ALIGN constraints govern the temporal patterning of gestures in a given language, RELEASE constraints force the first gesture to release (Davidson, 2006b). Davidson (2006b) proposed a RELEASE constraint, as illustrated in (13), which prohibits certain gestures from cooccurring by banning close alignment of such sequences. Thus, this constraint deals with phonotactic restrictions. This constraint differs from a phonotactic markedness constraint such as SYLCONT in traditional OT in that it prohibits close alignment of certain gestures, rather than their simple co-occurrence. In other words, as long as the two consonants are aligned in a way that allows release of the first consonant (releasing by itself or into an inserted vowel), this constraint is not violated. In contrast, traditional phonotactic constraints, which do not take timing relations into consideration, will be violated by an illegal sequence of consonants, whether the first consonant is released or not, necessitating a repair such as insertion of a true vowel between them, deletion of one of the consonants, or a change in one or both consonants.

(13) Davidson 2006b, p.15
RELEASE/Gα, Gβ: Do not overlap the release of a consonant containing gestural parameter(s) Gα with the plateau of a following consonant containing gestural parameter(s) Gβ.

In Korean, we need a RELEASE constraint that prohibits stop-nasal sequences since a stop and nasalmay not be closely aligned. In articulatory terms, obstruents can be labeled as [CD< NARROW, VEL=CLOSED], which refers to the gestures that involve constriction degree greater than that of a narrow or closed velum. [VEL=WIDE] refers to the gestures that involve wide velic aperture. See (14) below.

(14) [CD<NARROW, VEL=CLOSED]: Gestures with a constriction degree (CD) that is more constricted than [Narrow][19]
[VEL=WIDE]: Gestures with wide velic aperture.

---

[19] [CD<NARROW] without [VEL=Closed] refers to the gestures of a nonapproximant including obstruents and nasal stops, as proposed in Davidson (2006). In the current case, [VEL=Closed] should be added in order to refer to obstruents only.

A RELEASE constraint that militates against stop-nasal sequences is expressed as RELEASE/[CD< NARROW, VEL=CLOSED], [VEL=WIDE], which is defined in (15).

(15) RELEASE/[CD< NARROW, VEL=CLOSED], [VEL= WIDE]: Do not overlap the release of obstruents with the plateau of a nasal. (REL/[Obs],[Nas])

Additional faithfulness constraints banning change in velic gesture (nasality) or insertion of a gesture are listed in (16).

(16) DEP: Output segments must have input correspondents (Do not insert a gesture).
IDENT(NAS): Corresponding segments are identical in nasality (Do not change the CD of velum gesture)

With all these constraints defined, the ranking in the native grammar, which allows nasalization of the stop, is listed in (17).

(17) Korean Native Ranking:
DEP, REL/[Obs],[Nas], ALIGN (P, release, T, target) >> ALIGN (K,release, T, target) >> IDENT(NAS)

In tableau (18), candidate (18d), which changed the velic gesture (nasality), is the optimal output since IDENT(NAS) is ranked the lowest. Candidate (18a) is ruled out because it violates REL/[Obs],[Nas], which prohibits close coordination between an obstruent and a nasal. Note that (18b), whose internal consonantal gestures are pulled apart, has acoustic variants, as we discussed in section 6.4.1. It is realized either with an audible release burst ([sɛk^{h}nəl]) or with an intrusive vowel ([sɛk^{i}nəl)). Since [k] is voiceless, the dominant form is the one with a release burst. (18b) does not violate REL/[Obs],[Nas] but it violates ALIGN (K, release, T, target) because the stop is released, not satisfying the alignment between the release of the stop and the nasal target. Candidate (18c) with full vowel insertion is also ruled out because DEP is undominated in Korean.

(18) Nasalization in the native grammar, sɛkməl → sɛŋməl

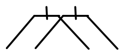
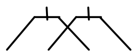
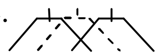
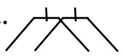
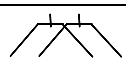
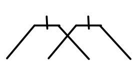
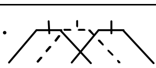
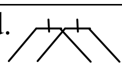
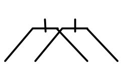
/sɛknəl/	DEP	REL/[Obs] ,[Nas]	ALIGN (P, release, T, target)	ALIGN (K, release, T, target)	IDENT (NAS)
a.  sɛknəl		*!			
b.  sɛk^{h}nəl/sɛk^{ɨ}nəl				*!	
c.  sɛkɨnəl	*!				
☞d.  sɛŋnəl					*

Tableau (19) illustrates the native grammar when the input contains a voiced stop. As in (18), the optimal candidate is (19d), which changed the stop to a nasal stop. (19a) and (19e) are ruled out because of the violation of REL/[Obs], [Nas]. Candidate (19b), which is coordinated loosely, is produced with an intrusive vowel or a release burst. Because the stop is voiced, (19b) is more likely to be realized with an intrusive vowel than with a release, but both variants are ruled out because ALIGN (K,release, T, target) is not satisfied. (19c) is ruled out because vowel insertion is prohibited.

(19) Nasalization in the native grammar, sɛgməl → sɛŋməl

/sɛgnəl/	DEP	REL/[Obs], [Nas]	ALIGN (P, release, T, target)	ALIGN (K, release, T, target)	IDENT (NAS)
a.  sɛgnəl		*!			
b.  sɛg^{ɨ}nəl/sɛg^{g}nəl				*!	
c.  sɛgɨnəl	*!				
☞d.  sɛŋnəl					*
e.  seknal		*!			

(18) and (19) demonstrate how the stop in stop-nasal sequences is nasalized in Korean following either a voiceless or a voiced stop.

Now I turn to the interlanguage grammar to derive the vowel intrusion/release bursts and devoicing found in the production experiment. Korean L2 learners of English need to learn that stop-nasal sequences are allowed in English and that in English, consonant clusters, including word-internal stop clusters, are generally closely coordinated (Anderson, 1974; Y. Kang, 2003; Matteson & Pike, 1958). Knowing this, they will need to demote REL/[Obs],[Nas]. At the same time, to maintain nasality, they must promote IDENT(NAS). Also, we need another RELEASE constraint that prohibits close coordination between voiced stop and nasal, as argued in the previous section. The constraint in (20) was proposed in Davidson (2006) for word initial clusters:

(20) RELEASE/[Pharyngeal volume],[CD<NARROW]: Do not overlap the release of a voiced obstruent with the plateau of a nonapproximant[20]. (Davidson 2006)
(REL/[VdObs],[NonApprox])

I assume that the same constraint applies to Korean clusters in word medial position, the only context where clusters occur in Korean. This constraint is more specific than Rel/[Obs],[Nas], in that it only prohibits close transition between a voiced stop and a nasal.

The reranking required to move from the Korean grammar to the interlanguage pattern is illustrated in (21). Within this grammar, nasalization is no longer optimal. Instead, vowel intrusion/a release burst occurs in velar stop-nasal sequences (back-to-front clusters) and devoicing occurs in voiced bilabial stop-nasal sequences(front-to-back or homorganic clusters).

(21) Reranked constraints in the interlanguage grammar
DEP, REL/[Obs],[Nas], (REL/[VdObs],[NonApprox]), **IDENT(NAS),**
ALIGN (P, release, T, target) >> **REL/[Obs],[Nas]** >> ALIGN (K,release, T, target) >> IDENT(NAS)

The critical ranking to derive this pattern is shown in (22). Note that REL/[Obs],[Nas], which prohibits close transition between stop and nasal, is ranked between the two ALIGN constraints. This ranking allows close transition between front-to-back sequences but not between back-to-front sequences. This will be illustrated below.

(22) Critical ranking to derive the place effect
ALIGN (P, release, T, target) >> **REL/[Obs],[Nas]** >> ALIGN (K, release, T, target)

Tableau (23) illustrates an intermediate grammar, in which a form in open transition is chosen as optimal because the sequences involve back-to-front coordination. Candidate (23b) does not violate REL/[Obs],[Nas] because the stop is released, but it violates ALIGN K, release, T, target). However, since ALIGN(K, release, T, target) is ranked below REL/[Obs],[Nas], the form in open transition is the actual output. This gestural representation can

[20] Nasal consonants are considered nonapproximant in Davidson (2006).

result in acoustic forms with an intrusive vowel or a release burst. Since the stop is voiced in this case, the dominant form is the form with an intrusive vowel. Candidate (23a) in close transition is ruled out because it violates REL/[Obs],[Nas]. Candidate (23c) is ruled out because of an inserted vowel gesture. Candidate (23d) is now out because nasalization is no longer optimal in this grammar.

(23) Interlanguage grammar, / sɛgnəl/ →[**sɛg^{i}nəl**/ sɛg^{g} nəl]

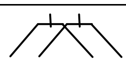
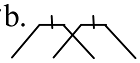
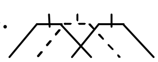
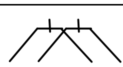
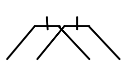
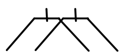
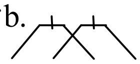
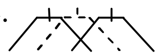
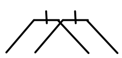
/sɛgnəl/	DEP	REL/[VdObs], [NonApprox]	IDENT (NAS)	ALIGN (P, release, T, target)	REL/[Obs] ,[Nas]	ALIGN(K, release,T, target)
a.  sɛgnəl		*!			*	
☞b.  **sɛg^{i}nəl**/ sɛg^{g} nəl						*
c.  sɛgɨnəl	*!					
d.  sɛŋnəl			*!			
e.  seknal					*!	

Tableau (24) illustrates a grammar in which candidate (24b) with the same gestural configuration as (23b) is optimal. But since the stop is voiceless, this candidate is more likely to surface with a release burst than an intrusive vowel. I argued that a single gestural representation can result in different acoustic variants (intrusive vowel or release burst) depending on the voicing of the first consonant, and (23) and (24) demonstrate this point. What is critical here is that since REL/[Obs],[Nas] is ranked higher than ALIGN (K, release, T, target), the candidate in open transition, whether with an intrusive vowel or a release burst, is optimal.

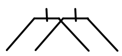
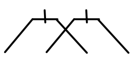
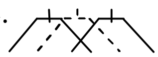
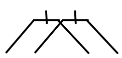
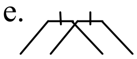
(24) Interlanguage grammar, sɛknəl/ →[**sɛk^{h}nəl** /sɛk^{i}nəl]

/sɛknəl/	DEP	REL/[VdObs], [NonApprox]	IDENT (NAS)	ALIGN (P, release, T, target)	REL/[Obs] ,[Nas]	ALIGN (K,release, T, target)
a.  sɛknəl		*!			*	
☞b.  **sɛk^{h}nəl** /sɛk^{i}nəl						*
c.  sɛkɨnəl	*!					
d.  sɛŋnəl			*!			

Now, let's turn to the case where the stop is labial. Within the same grammar, devoicing is favored over vowel intursion/release burst because ALIGN (P, release, T, target) is ranked higher than REL/[Obs],[Nas]: unlike in back-to-front sequences, aligning the release of the labial stop with the following segment is more important than avoiding close transition between stop and nasal in the front-to-back sequences.

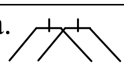
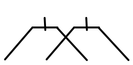
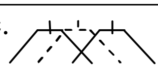
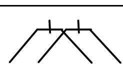
Therefore, candidate (25e), in which the stop is unreleased and devoiced, will be optimal: (25a) is out because keeping voicing in an unreleased stop is difficult, (25b) is out because the open transition in front-to-back order of articulation is prohibited, and (25c) and (25d) are ruled out due to the undominated faithfulness constraints.

(25) Interlanguage grammar, / sɛbnəl/ →[sɛpnəl]

/sɛbnəl/	DEP	REL/[VdObs], [NonApprox]	IDENT (NAS)	ALIGN (P, release, T, target)	REL/[Obs], [Nas]	ALIGN (K,release, T, target)
a.  sɛbnəl		*!			*	
b.  **sɛb^{i}nəl**/sɛb ^{b}nəl				*!		
c.  sɛbɨnəl	*!					
d.  sɛmnəl			*!			
☞e.  sɛpnəl					*	

And within the same grammar, an input containing a voiceless bilabial stop followed by a nasal will be correctly pronounced in close transition, as shown in (26).

(26) Interlanguage grammar, /sɛpnəl/ →[sɛpnəl]

/sɛpnəl/	DEP	REL/[VdObs], [NonApprox]	IDENT (NAS)	ALIGN (P, release, T, target)	REL/[Obs], [Nas]	ALIGN K,release, T, target)
☞a.  sɛpnəl					*	
b.  **sɛp^{h}nəl**/sɛp^{i}nəl				*!		
c.  sɛpɨnəl	*!					
d.  sɛmnəl			*!			

Finally, to reach the target language grammar, REL/[VdObs], [NonApprox] will eventually need to be demoted as ALIGN(K,release, T, target) is promoted, so that any sequences can be realized in close transition without change in nasality or voicing.

The interlanguage grammar proposed above accounts for the asymmetric repair patterns found in the production experiment. Vowel intrusion or a release burst was more common in velar-nasal sequences because the native language allows slower transition between velar-initial sequences. Devoicing was more frequent in labial-nasal sequences because the labial-initial sequences must have close transition, and maintaining voicing in an unreleased stop is not possible in Korean. This grammar suggests that the puzzling asymmetric pattern found in the L2 learners' production is rooted in the native language, but can be accounted for only by considering the timing relation between gestures.

6.6 Summary

In this chapter, I have shown that the interlanguage patterns found in the production experiment which were not clearly explained either as an effect of the phonological grammar or as an effect of misperception can be accounted for once we incorporate the timing relationships among gestures into the phonological representation. The greater frequency of vowel intrusion in voiced stop-nasal sequences was attributed to an effect of acoustic variability, i.e., given the same open transition, voiced stops were more likely to be produced with an intrusive vowel and voiceless stops with a release burst. However, the overall rate of release of the stop (both release burst and vowel intrusion) was similar whether the stop was voiceless or voiced. The greater frequency of vowel intrusion/release burst in velar stop-nasal sequences resulted from the fact that in Korean, velar-initial clusters have a less overlapped coordination, giving rise to higher possibility of release of the first consonant. Frequent devoicing in labial stop-nasal sequences was explained by the fact that labial-initial clusters are closely coordinated in Korean, resulting in devoicing due to the articulatory difficulty of maintaining voicing in an unreleased stop.

Chapter 7
Conclusion

English stop-nasal sequences, which do not occur in Korean, pose a serious problem for Korean speakers. According to the *Full Transfer Hypothesis*, Korean speakers are expected to nasalize a stop before a nasal. However, in the two production experiments discussed in Chapter 2, nasalization was not the dominant pattern in Korean L2 learners' production of nonce words containing stop-nasal sequences (e.g., *segnal*). Although nasalization was applied in some cases, other error types were at least as frequent as nasalization: devoicing of the stop, and insertion of a vowel between the stop and the nasal. Furthermore, systematic patterns emerged that are not explainable with reference to the native language: (i) accuracy was higher for voiceless stop-nasal sequences than for voiced stop-nasal sequences; (ii) vowel insertion occurred more frequently after voiced stops than voiceless stops; and (iii) devoicing occurred more frequently with bilabial stops whereas vowel insertion occurred more frequently after velar stops. In this dissertation I have considered several possible explanations of these emerging patterns: native language constraint reranking, universal markedness, misperception, and transfer of native language gestural timing patterns.

In Chapter 4, I argued that these asymmetric patterns cannot be explained by either the reranking of L1 constraints or the effect of universal markedness because these patterns cannot emerge in a model where learning is error-driven and accurate perception is assumed. Although devoicing might be explained as an emergence of the unmarked effect in the interlanguage, there is no evidence in either Korean or English that would favor devoicing more strongly in bilabial than in velar stops. Nor does the greater frequency of vowel insertion after voiced stops or/and velar stops appear to have its source in either of the languages in contact.

To determine whether vowel insertion and devoicing could be explained as a result of misperception, I conducted both behavioral and ERP (Event-Related Potentials) experiments (Chapter 3). The results of these experiments did provide some support for misperception, rooted in the native language. Korean listeners tended to hear an illusory vowel after voiced stops and to have more difficulty distinguishing voiced stop-nasal sequences from voiced stop-vowel-nasal sequences than did English native listeners. Similarly, Dupoux et al. (1999) have shown that Japanese listeners, but not French listeners, tended to perceive illusory vowels in sequences that are illegal in Japanese but not in French. Based on this result, Dupoux et al. proposed that the reason Japanese speakers frequently inserted a vowel in production is because they perceived a vowel in illegal sequences. However, the experimental results in this dissertation demonstrated that simple illegality of the sequence is not sufficient to induce perception of an illusory vowel. Korean listeners' perception of an illusory vowel depends on the voicing of the stop, rather than simple illegality of the sequence, as shown by the fact that Korean listeners had greater difficulty in discriminating [gn]-[gVn] than [kn]-[kVn]. Nonetheless, in the ERP experiment, Korean listeners were able to distinguish the absence or presence of a vowel between the stop and the nasal, regardless of the voicing of the stop. For both voiced and voiceless sequences, Korean and

English listeners displayed similar MMN (Mismatch Negativity) responses. This suggests that Korean listeners' misperception of voiced stops as voiced stops followed by a vowel does not arise simply from an inability to discriminate the contrast between stop and stop-vowel. Rather, Korean listeners accurately hear the voicing contrast, and then interpret voiced stops as being followed by a vowel, because in Korean, word medial voiced stops are possible only if followed by a vowel. A perception grammar that enforces this mapping was proposed, following the *Cue Interpretation* model (Boersma & Hamann, 2009).

However, the production data revealed that vowel insertion did not take place after all voiced stops at the same rate: vowels were inserted most frequently after voiced velar stops, while bilabial stops were frequently devoiced. If these patterns are rooted in perception, we would expect that Korean speakers should have more difficulty distinguishing [gn]-[gVn] than [bn]-[bVn], and [bn]-[pn] than [gn]-[kn]. While the tendency to insert a vowel more frequently after a voiced than a voiceless obstruent could be explained as a function of perception, the results of the perception experiments presented no evidence that the tendency to insert a vowel after velars or to devoice bilabial stops reflects a perceptual effect. Korean listeners did not report hearing a vowel more often after [g] than [b] and had no more difficulty distinguishing [bn] from [pn] than [gn] from [kn].

In Chapter 6, I argued that an examination of the gestural timing of the stop-nasal sequences can explain the asymmetric patterns for different places of articulation of the stop. Following research establishing that certain vowel percepts in consonant clusters represent not full phonological vowels but rather loose coordination of the consonantal gestures, I compared the acoustics of the vowels that had been transcribed as inserted vowels in my participants' productions with underlying lexical vowels, following Davidson's (2006) definitions of true phonological epenthetic vowels vs. intrusive (non-phonological) vowels. This resulted in a classification of added vowels either as true inserted vowels, which were similar in duration to lexical vowels, or as intrusive vowels, which were shorter than lexical vowels, and resulted from loose coordination of consonantal gestures. The vowels that were identified by these criteria as true epenthetic vowels were equally frequent in both *b-nasal* and *g-nasal* sequences, a finding that is consistent with the perception experiment results, in which Korean listeners reported hearing a vowel in [bn] or [gn] to the same degree. However, the vowels identified as intrusive showed a place asymmetry, occurring more frequently after voiced stops (voicing effect on vowel intrusion) and after velar stops (place effect on vowel intrusion). It was proposed that once we assume release bursts and intrusive vowels as acoustic variants resulting from the same gestural representation, the voicing effect disappears. That is, the combined rates of release bursts and intrusive vowels are comparable between voiceless stop-nasal sequences and voiced stop-nasal sequences. I also argued that the Korean articulatory pattern in which velar stops are much less overlapped with the following consonant than labial stops gives rise to greater opportunities to release the velar stop, causing frequent release bursts or intrusive vowels after velar stops. In contrast, the Korean gestural timing pattern, in which labial stops are much more closely overlapped with the following consonant than velar stops, makes release of the voiced bilabial stop less likely, causing devoicing of [b] due to aerodynamic reasons. I argued that incorporating timing relations into the phonological grammar can capture these asymmetric patterns.

The results of this study show that second language phonology involves a complex interplay between L1 grammar, misperception through L1, and mastery of new gestural coordination. Based on this, I proposed a phonological model in which perceptual and gestural constraints were incorporated into, and I explained the error patterns revealed in the production experiment results.

This model makes specific predictions about the developmental paths that Korean learners will follow in the acquisition of English. The present study was not designed to test these predictions, as it included no longitudinal component, and participants varied in terms of age of first exposure to English, years of study, and length of stay in an English-speaking country. However, an examination of the correlation between error patterns and overall accuracy is suggestive. The participants in each production experiment were divided into two groups according to their average accuracy rates. Those with accuracy higher than the group average (51.0% in the repetition experiment, 43.4% in the reading experiment) were labeled as advanced, while those whose accuracy fell below the group average were labeled less advanced. In both experiments, the groups with lower accuracy showed greater rates of nasalization and vowel insertion than the groups with higher accuracy. In the reading experiment, though not the repetition experiment, the group with higher accuracy showed greater rates of devoicing than the group with lower accuracy.

(1) Average rates of errors in groups with lower and higher accuracy in production experiments

	Repetition Experiment		Reading Experiment	
	Less advanced	Advanced	Less advanced	Advanced
Accuracy	34.64%	62.88%	29.58%	57.29%
Nasalization	19.27%	1.136%	20%	5%
Devoicing	20.83%	20.45%	7.708%	18.96%
Vowel Insertion	13.8%	8.902%	29.79%	9.792%

This pattern, although not examined on individual levels, is generally consistent with the predictions of the proposed phonological model. However, no clear correlation emerged between accuracy rate and background factors (Age, Years of study of EFL, Length of stay in an English-speaking country or Age at exposure to English).

A full test of the predictions of the proposed model will require the collection of longitudinal data on perception, production and articulation from individual L2 learners with different L1 backgrounds and the comparison of the data both across speakers of the same L1 and across language groups.

Much work on L2 speech has modeled the learning of a second language in terms of moving from the native language grammar to the foreign language grammar. It is clear that a full understanding of L2 acquisition must take account of all aspects of the L1 and L2 systems—the mapping from the acoustic signal to lexical representations (perception), the mapping from lexical representations to surface representations (production), and the articulatory coordination of gestures—as well as the interaction of these different components of the sound system.

References

An, Y., Hwang, J., & Suh, Y. (2008). Phonetic status of Cj combinations in Korean and Spanish. *Poster presented at the Journal of the Acoustical Society of America, 124*.

Anderson, S. R. (1974). *The Organization of Phonology*. New York: Academic Press.

Bat-El, O. (1996). Selecting the Best of the Worst: The Grammar of Hebrew Blends. *Phonology, 13*, 283-328.

Bell-Berti, F. (1975). Control of pharyngeal cavity size for English voiced and voiceless stops. *The Journal of the Acoustical Society of America, 57*, 456-461.

Benki, J. R. (2001). Place of articulation and first formant transition pattern both affect perception of voicing in English. *Journal of Phonetics, 29*, 1-22.

Benus, S., Smorodinsky, I., & Gafos, A. I. (2004). Gestural coordination and the distribution of English "geminates". In S. Arunachalam & T. Scheffler (Eds.), *U. Penn Working Papers in Linguistics 10.1. Proceedings of the 27th Annual Penn Linguistic Colloquium*. (pp. 33-46). Philadelphia: Penn Linguistics Club.

Blevins, J. (2006). Phonetically-based sound patterns: Typological tendencies or phonological universals? *LabPhon 10*, 201-226.

Boersma, P. (1998). *Functional Phonology.* Ph.D. Dissertation. University of Amsterdam.

Boersma, P. (1999). On the need for a separate perception grammar. Rutgers Optimality. In: Archive 358, http://roa.rutgers.edu.

Boersma, P., Dekkers, J., & van de Weijer, J. (2000). Introduction. In J. Dekkers, F. van der Leeuw & J. van de Weijer (Eds.), *Optimality Theory: Phonology, Syntax and Acquisition* (pp. 1-44). New York: Oxford University Press.

Boersma, P., & Hamann, S. (2009). Loanword adaptation as first-language phonological perception. In A. Calabrese & L. Wetzels (Eds.), *Loan Phonology* (pp. 11-53). Amsterdam: Benjamins.

Boersma, P., & Weenink, D. (2007). Praat: doing phonetics by computer In (4.6.29 ed., pp. Retrieved October 1, 2007, from http://www.praat.org/).

Borroff, M. L. (2007). *A landmark underspecification account of the patterning of glottal stop.* Ph.D. Dissertation. Stony Brook University.

Brannen, K. (2002). The role of perception in differential substitution. *Canadian Journal of Linguistics, 47*, 1-46.

Broselow, E. (2004). Unmarked Structures and Emergent Rankings in Second Language Phonology. *International Journal of Bilingualism, 8*, 51-65.

Broselow, E. (2009a). Asymmetries in Foreign Language Perception and Production. In *KNAW Colloquium on Language Acquisition & Optimality Theory*. Amsterdam.

Broselow, E. (2009b). Stress adaptation in loanword phonology: perception and learnability. In P. Boersma & S. Hamann (Eds.), *Phonology in perception* (pp. 191-234): De Gruyter Mouton.

Broselow, E., Chen, S.-I., & Wang, C. (1998). The Emergence of the unmarked in second language phonology. *Studies in Second Language Acquisition, 20*, 261-280.

Broselow, E., & Finer, D. (1991). Parameter setting in second language phonology and syntax. *Second Language Research, 7*, 35-59.

Brown, C. A. (2000). The interrelation between speech perception and phonological acquisition from infant to adult. In J. Archibald (Ed.), *Second language acquisition and linguistic theory* (pp. 4-65). Oxford, UK: Blackwell.

Brown, C. A. (1998). The role of the L1 grammar in the L2 acquisition of segmental structure. *Second Language Research, 14*, 136-193.
Byrd, D. (1996). Influences on articulatory timing in consonant sequences. *Journal of Phonetics, 24*, 209-244.
Catford, J. C. (1977). *Fundamental problems in phonetics*. Bloomington: Indiana University Press.
Chitoran, I., & Goldstein, L. (2006). Testing the phonological status of perceptual recoverability: Articulatory evidence from Georgian. *Paper presented at LabPhon X*. Paris.
Chitoran, I., Goldstein, L., & Byrd, D. (2002). Gestural overlap and recoverability: Articulatory evidence from Georgian. In C. Gussenhoven & N. Warner (Eds.), *Laboratory Phonology 7* (pp. 419-448). Berlin: Walter de Gruyter.
Cho, H. (1998). English loanwords and syllable structures. *Studies in Phonetics, Phonology and Morphology, 4*, 273-289.
Cho, Y.-M., Y. (1990). Syntax and phrasing in Korean. In S. Inkelas & D. Zec (Eds.), *The Phonology-Syntax Connection* (pp. 47-62). Chicago: Univeristy of Chicago Press.
Chung, K. (2004). Is loanword phonology separate from native phonology? *Studies in Phonetics, Phonology and Morphology, 10*, 543-566.
Colantoni, L., & Steele, J. (2007). Acquiring /ʁ/ in Context. *Studies in Second Language Acquisition, 29*, 381-406.
Colantoni, L., & Steele, J. (2008). Integrating articulatory constraints into models of second language phonological acquisition. *Applied Psycholinguistics, 29*, 489-534.
Côté, M.-H. (2000). *Consonant cluster phonotactics: A perception-based approach.* Ph.D. Dissertation. MIT.
Davidson, L. (2003). *The atoms of phonological representation: gestures, coordination and perceptual features in consonant cluster phonotactics.* Ph.D. Disseration. Johns Hopkins University.
Davidson, L. (2006a). Phonology, phonetics, or frequency: Influences on the production of non-native sequences. *Journal of Phonetics, 34*, 104-137.
Davidson, L. (2006b). Phonotactics and Articulatory Coordination Interact in Phonology: Evidence from Nonnative Production. *Cognitive Science, 30*, 837-862.
Davidson, L. (2010). Phonetic bases of similarities in cross-language production: Evidence from English and Catalan. *Journal of Phonetics, 38*, 272-288.
Davidson, L. (2011). Phonetic and Phonological Factors in the Second Language Production of Phonemes and Phonotactics. *Language and Linguistics Compass, 5*, 126-139.
Davidson, L., & Stone, M. (2003). Epenthesis versus gestural mistiming in consonant cluster production: an ultrasound study. In G. Garding & M. Tsujimura (Eds.), *WCCFL 22 Proceedings* (pp. 165-178). Somerville, MA: Cascadilla Press.
Davis, S., & Shin, S.-H. (1999). The Syllable Contact Constraint in Korean: An Optimality-Theoretic Analysis. *Journal of East Asian Linguistics, 8*, 285-312.
Dehaene-Lambertz, G. (1997). Electrophysiological correlates of categorical phoneme perception in adults. *NeuroReport, 8*, 919-924.
Dehaene-Lambertz, G., Dupoux, E., & Gout, A. (2000). Electrophysiological Correlates of Phonological Processing: A Cross-linguistic Study. *Journal of Cognitive Neuroscience, 12*, 635-647.
Delorme, A., & Makeig, S. (2004). EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics. *Journal of Neuroscience Methods, 134*, 9-21.

Dupoux, E., Kakehi, K., Hirose, Y., Pallier, C., & Mehler, J. (1999). Epenthetic vowels in Japanese: A perceptual illusion? *Journal of Experimental Psychology: Human Perception and Performance, 25*, 1568-1578.

Eckman, F. (1977). Markedness and the Contrastive Anlysis Hypothesis. *Language Learning, 27*, 315-330.

Eckman, F., & Iverson, G. K. (1993). Sonority and markedness among onset clusters in the interlanguage of ESL learners. *Second Language Research, 9*, 234-252.

Ellis, R. (1997). *Second language acquisition*. Oxford: Oxford University Press.

Elson, B. (1947). Sierra Popoluca Syllable Structure. *International Journal of American Linguistics, 13*, 13-17.

Escudero, P. (2005). *Linguistic perception and second language acquisition.* Utrecht: LOT.

Escudero, P., & Boersma, P. (2004). Bridging the gap between L2 speech perception research and phonological theory. *Studies in Second Language Acquisition, 26*, 551-585.

Fischer, R. M., & Ohde, R. N. (1990). Spectral and duration properties of front vowels as cues to final stop-consonant voicing. *The Journal of the Acoustical Society of America, 88*, 1250-1259.

Flege, J. E. (1995). Second-language speech learning: theory, findings, and problems. In W. Strange (Ed.), (pp. 233-277). Timonium, MD: York Press.

Flege, J. E., McCutcheon, M. J., & Smith, S. C. (1987). The development of skill in producing word-final English stops. *The Journal of the Acoustical Society of America, 82*, 433-447.

Gafos, A. I. (2002). A grammar of gestural coordination. *Natural Language and Linguistic Theory, 20*, 269–337.

Gafos, A. I., Hoole, P., Roon, K., & Zeroual, C. (2010). Variation in timing and phonological grammar in Moroccan Arabic clusters. In C. Fougeron, B. Kühnert, M. d'Imperio & N. Vallé (Eds.), *Laboratory Phonology X: Variation, Detail and Representation*. Berlin/New York: Mouton de Gruyter.

Gass, S. M. (1996). Second language acquisition and linguistic theory: the role of language transfer. In W. C. Ritchie & T. K. Bhatia (Eds.), *The handbook of second language acquisition* (pp. 317-345). San Diego: Academic Press.

Hall, N. (2003). *Gesutres and segments: vowel intrusion as overlap.* Ph.D. Dissertation. University of Massachusetts, Amherst.

Hall, N. (2006). Cross-linguistic patterns of vowel intrusion. *Phonology, 23*, 387-429.

Hancin-Bhatt, B. (2000). Optimality in second language phonology: codas in Thai ESL. *Second Language Research, 16*, 201-232.

Hancin-Bhatt, B. (2008). Second language phonology in optimality theory. In J. G. Hansen Edwards & M. L. Zampini (Eds.), *Phonology and second language acquisition* (pp. 117-146). Amsterdam, Philadelphia: John Benjamins.

Hansen, J. (2001). Linguistic constraints on the acquisition of English syllable codas by native speakers of Mandarin Chinese. *Applied Linguistics, 22*, 338-365.

Hayes, B., & Stivers, T. (1995). Postnasal voicing. In. Los Angeles: University of California, Los Angeles.

Hayes, B. P. (1996). Phonetic mechanisms of post-nasal voicing. *The Journal of the Acoustical Society of America, 100*, 2693-2693.

Hillenbrand, J., Ingrisano, D. R., Smith, B. L., & Flege, J. E. (1984). Perception of the voiced--voiceless contrast in syllable-final stops. *The Journal of the Acoustical Society of America, 76*, 18-26.

Hisagi, M. (2007). *Perception of Japanese temporally-cued phonetic contasts by Japanese and American English listeners: Behavioral and Electrophysiological Measures.* Ph.D. Dissertation. City Univeristy of New York-Graduate School and University Center.

Hogan, J. T. , & Rozsypal, A. J. (1980). Evaluation of vowel duration as a cue for the voicing distinction in the following word-final consonant. *The Journal of the Acoustical Society of America, 67*, 1764-1771.

Ito, C., Kang, Y., & Kenstowicz, M. (2006). The adaptation of Japanese loanwords into Korean. *Studies in Loanword Phonology (MITWPL 52), 52*, 65-104.

Itô, J., & Mester, R. A. (1994). Reflection on CodaCond and Alignment. *Phonology at Santa Cruz (PASC), 3*, 22-46.

Jacobs, H., & Gussenhoven, C. (2000). Loan phonology: perception, salience, the lexicon, and OT. In J. Dekkers, F. van der Leeuw & J. van de Weijer (Eds.), *Optimality Theory: Phonology, Syntax and Acquisition* (pp. 193-210). Oxford: Oxford University Press.

Jacobsen, T., & Schröger, E. (2003). Measuring duration mismatch negativity. *Clinical Neurophysiology, 114*, 1133-1143.

Jaramillo, M., Paavilainen, P., & Näätänen, R. (2000). Mismatch negativity and behavioural discrimination in humans as a function of the magnitude of change in sound duration. *Neuroscience Letters, 290*, 101-104.

Johnson, K. (2003). *Acoustic and Auditory Phonetics* (2nd ed.): Wiley-Blackwell.

Jones, C. (2003). *Development of Phonological Categories in Children's Perception of Final Voicing in English.* Ph.D. Dissertation. University of Massachusetts at Amherst.

Jun, E. (2002). An Experimental study of the effect of release of English syllable final stops on vowel epenthesis in English loanwords. *Studies in Phonetics, Phonology and Morphology, 8*, 117-134.

Jun, S.-A. (1992). The Domain of Nasalization and the Prosodic Structure in Korean. In H.-M. Sohn (Ed.), *Korean Linguistics 7* (pp. 11-29). Seoul, Korea: Hanshin Publishing Co.

Jun, S.-A. (1993). *The Phonetics and Phonology of Korean Prosody.* Ph.D. Dissertation. Ohio State University.

Kabak, B. (2003). *The perceptual processing of second language consonant clusters.* Ph.D. Dissertation. University of Delaware.

Kabak, B., & Idsardi, W. J. (2007). Perceptual Distortions in the Adaptation of English Consonant Clusters: Syllable Structure or Consonantal Contact Constraints? *Language and Speech, 50*, 23-52.

Kang, H. (2002). On the Optimality-Theoretic Analysis of Korean Nasal-Liquid Alternations. *Journal of East Asian Linguistics, 11*, 43-66.

Kang, Y. (2003). Perceptual similarity in loanword adaptation: English postvocalic word-final stops in Korean. *Phonology, 20*, 219-273.

Kang, Y. (2008). Interlanguage Segmental Mapping as Evidence for the Nature of Lexical Representation. *Language and Linguistics Compass, 2*, 103-118.

Kenstowicz, M. (2001). The role of perception in loanword phonology. *Linguistique Africaine, 20*.

Kenstowicz, M. (2005). Phonetics and phonology of Korean loanword adaptation. In S.-J. Rhee (Ed.), *Proceedings of First European Conference on Korean Linguistics.*

Kim, K., & Kochetov, A. (2011). Phonology and phonetics of epenthetic vowels in loanwords: Experimental evidence from Korean. *Lingua, 121*, 511-532.

Kingston, J., & Diehl, R. L. (1995). Intermediate properties in the perception of distinctive feature values. In B. Connell & A. Arvaniti (Eds.), *Phonology and phonetics: Papers in laboratory phonology IV* (pp. 7-27). Cambridge: Cambridge University Press.

Kochetov, A., Pouplier, M., & Son, M. (2007). Cross-language differences in overlap and assimilation patterns in Korean and Russian. *Proceedings of the 16th International Congress of Phonetics Science (ICPhS)*, 1361-1364.

Ku, P.-S. (1999). English loanwords and vowel epenthesis in Optimality Theory. *Studies in Phonetics, Phonology and Morphology, 5*, 59-77.

Kuhl, P. K. (2004). Early language acquisition: Cracking the speech code. *Nature Reviews Neuroscience, 5*, 831-843.

Kuhl, P. K., & Miller, J. D. (1975). Speech perception by the chinchilla: Voiced-voiceless distinction in alveolar plosive consonants. *Science, 190*, 69-72.

Kühnert, B., Hoole, P., & Mooshammer, C. (2006). Gestural overlap and C-center in selected French consonant clusters *Proceedings of the 7th International Seminar on Speech Production* (pp. 327-334). Ubatuba, Brazil.

Kwuklipkwukeyenkwuwen. (2001). Oylaye palum silthay cosa [A survey of loanword pronunciation] In. Seoul: Kwuklipkwukeyenkwuwen [The National Academy of the Korean Language].

Kwuklipkwukeyenkwuwen. (2009). Oylaye Pyokipeop [The Orthography of Loawords]. In: Kwuklipkwukeyenkwuwen [The National Academy of the Korean Language].

LaCharité, D., & Paradis, C. (2005). Category Preservation and Proximity versus Phonetic Approximation in Loanword Adaptation. *Linguistic Inquiry, 36*, 223-258.

Ladefoged, P. (1993). *A course in phonetics* (3rd edition ed.). Fort Worth, TX: Harcourt, Brace, and Jovanovich.

Lahiri, A., & Marslen-Wilson, W. D. (1991). The mental representation of lexical form: a phonological approach to the recognition lexicon. *Cognition, 38*, 245-294.

Lee, P. (1995). Korean loanword phonology: an Optimality perspective (In Korean). *Ene, 20*, 121-151.

Lee, S., & Pater, J. (2008). Phonological inference and word recognition: Evidence from Korean. Manuscript. Korea University and University of Massachusetts, Amherst.

Lisker, L., & Abramson, A. S. (1964). A cross-language study of voicing in initial stops: Acoustical measurements. *Word, 20*, 384-422.

Lombardi, L. (1995). Why Place and Voice are Different: Constraint-specific Alternations in Optimality Theory. Manuscript. Univeristy of Maryland, College Park.

Luce, P., A., & Charles-Luce, J. (1985). Contextual effects on vowel duration, closure duration, and the consonant/vowel ratio in speech production. *The Journal of the Acoustical Society of America, 78*, 1949-1957.

Major, R. C., & Faudree, M. C. (1996). Markedness universals and the acquisition of voicing contrasts by Korean speakers of English. *Studies in Second Language Acquisition, 18*, 69-90.

Matteson, E., & Pike, K. (1958). Non-phonemic transition vocoids in Piro (Arawak). *Miscellanea Phonetica, 3*, 22-30.

Matthews, J., & Brown, C. A. (2004). When intake exceeds input: language specific perceptual illusions induced by L1 prosodic constraints. *International Journal of Bilingualism, 8*, 5-23.

McCarthy, J., & Prince, A. (1994). The emergence of the unmarked: optimality in prosodic morphology. *NELS, 24*, 333-379.
Messing, C. L. (2008). *Intrusive Segments as Instances of Gestural Mistiming: Mandarin L1 English Learners' Production of English Consonant Clusters.* Master's Thesis. University of California, Davis.
Miao, R. (2005). *Loanword Adaptation in Mandarin Chinese: Perceptual, phonological and sociolinguistic factors.* Ph.D. Dissertation. Stony Brook University.
Moreton, E. (2004). Realization of the English Postvocalic [voice] contrast in F1 and F2. *Journal of Phonetics, 32*, 1-33.
Näätänen, R. (2001). The Perception of Speech Sounds by the Human Brain as Reflected by the Mismatch Negativity (MMN) and Its Magnetic Equivalent (MMNm). *Psychophysiology, 38*, 1-21.
Näätänen, R., Lehtokoski, A., Lennes, M., Cheour, M., Huotilainen, M., Iivonen, A., Vainio, M., Alku, P., Ilmoniemi, R. J., & Luuk, A. (1997). Language-specific phoneme representations revealed by electric and magnetic brain responses. *Nature, 385*, 432-434.
Nenonen, S., Shestakova, A., Huotilainen, M., & Näätänen, R. (2003). Linguistic relevance of duration within the native language determines the accuracy of speech-sound duration processing. *Cognitive Brain Research, 16*, 492-495.
Nicklas, T. D. (1974). *The elements of Choctaw.* Ph.D. Dissertation. University of Michigan.
Nittrouer, S. (2004). The role of temporal and dynamic signal components in the perception of syllable-final stop voicing by children and adults. *The Journal of the Acoustical Society of America, 115*, 1777-1790.
Oh, E. (2008). Coarticulation in non-native speakers of English and French: An acoustic study. *Journal of Phonetics, 36*, 361-384.
Ohala, J. J. (1997). Aerodynamics of phonology. *Proceedings of the Seoul International Conference on Linguistics*, 92-97.
Paradis, C. (1996). The Inadequacy of Filters and Faithfulness in Loanword Adaptation. In J. D. B. Laks (Ed.), *Current Trends in Phonology: Models and Methods* (pp. 509-534). Salford: University of Salford Publications.
Paradis, C., & LaCharité, D. (1997). Preservation and minimality in loanword adaptation. *Journal of Linguistics, 33*, 379-430.
Park, H., & de Jong, K. J. (2006). Native Koreans' perception of voicing in VC position: Prosodic restructruing effects on consonant identification. In *Indiana University Working Papers in Linguistics- Online* (Vol. 06).
Park, I. K. (2006). Production of stop-nasal sequences by Korean learners of English: an Optimality Theoretic approach. *Language Research, 42*, 1-18.
Parker, F. (1981). A functional-perceptual account of final devoicing. *Journal of Phonetics, 9*, 129-137.
Pater, J. (1996). Austronesian nasal substitution and order NC effects. In R. Kager, H. van der Hulst & W. Zonneveld (Eds.), *The prosody morphology interface* (pp. 310-343): Cambridge University Press.
Peperkamp, S. (2005). A psycholinguistic theory of loanword adaptations. In *Proceedings of the 30th Annual Meeting of the Berkeley Linguistics Society*. Berkeley, CA: Berkeley Linguistics Society.

Peperkamp, S., & Dupoux, E. (2003). Reinterpreting loanword adaptations: the role of perception. *Proceedings of the 15th International Congress of Phonetics Science (ICPhS)*, 367-370.

Phillips, C. (2001). Levels of representation in the electrophysiology of speech perception. *Cognitive Science, 25*, 711-731.

Raphael, L., J. (1972). Preceding Vowel Duration as a Cue to the Perception of the Voicing Characteristic of Word-Final Consonants in American English. *The Journal of the Acoustical Society of America, 51*, 1296-1303.

Rhee, S.-C. (1998). *Aspects of release and nonrelease in phonology.* Ph.D. Dissertation. University of Illinois at Urbana-Champaign.

Rhee, S.-C., & Choi, Y.-K. (2001). A satistical observation of vowel epenthesis in English loanwords in Korean and its significance. *Studies in Phonetics, Phonology and Morphology, 7*, 153-176.

Rivera-Gaxiola, M., Csibra, G., Johnson, M. H., & Karmiloff-Smith, A. (2000). Electrophysiological correlates of cross-linguistic speech perception in native English speakers. *Behavioural Brain Research, 111*, 13-23.

Rodgers, B. (2007). Obstruent voicing perception cues. *LSO Working Papers in Linguistics 7: Proceedings of WIGL 2007*, 135-143.

Rositzke, H. A. (1943). The articulation of final stop sin General American speech. *American Speech, 18*, 39-42.

Schwartz, B. D., & Sprouse, R. A. (1994). Word order and nominative case in non-native language acquisition: A longitudinal study of (L1 Turkish) German interlanguage. In T. Hoekstra & B. D. Schwartz (Eds.), *Language Acquisition Studies in Generative Grammar: Papers in Honor of Kenneth Wexler from the 1991 GLOW Workshops (Language Acquisition & Language Disorders, vol. 8)* (pp. 317-368). Philadelphia: John Benjamins.

Seo, M. (2006). Motivation for phonological modifications of stop plus sonorant sequences. *Studies in Phonetics, Phonology and Morphology, 12*, 103-126.

Shafer, V. L., Schwartz, R. G., & Kurtzberg, D. (2004). Language-specific memory traces of consonants in the brain. *Cognitive Brain Research, 18*, 242-254.

Sharma, A., & Dorman, M. F. (2000). Neurophysiologic Correlates of Cross-Language Phonetic Perception. *The Journal of the Acoustical Society of America, 107*, 2697-2703.

Shinohara, S. (2004). Emergence of Universal Grammar in foreign word adaptation. In R. Kager, J. Pater & W. Zonneveld (Eds.), *Constraints in phonological acquisition* (pp. 292-320). Cambridge: Cambridge University Press.

Silva, D. (1992). *The Phonetics and Phonology of Stop Lenition in Korean.* Ph.D. Dissertation. Cornell University.

Silverman, D. (1992). Multiple scansions in loanword phonology: evidence from Cantonese. *Phonology, 9*, 298-328.

Smith, J. (2004). Loan phonology is not all perception: evidence from Japanese loan doublets. In T. Vance (Ed.), *Japanese/Korean Linguistics* (Vol. 14). Stanford: CSLI.

Steriade, D. (1997). Phonetics in Phonology: The Case of Laryngeal Neutralization. Manuscript. UCLA.

Strange, W., & Shafer, V. L. (2008). Speech perception in second language learners:The re-education of selective perception. In J. G. Hansen Edwards & M. L. Zampini (Eds.),

Phonology and second language acquisition (pp. 153-191). Amsterdam, Philadelphia: John Benjamins.
Suh, Y. (2010). English Sibilants before /i/ in Korean Loanwords: A Contrast-specific Role of Phonetic Similarity *Korean Journal of Linguistics, 35*, 941-967.
Tarone, E. (1987). Some influences on the syllable structure of interlanguage phonology. In G. Ioup & S. Weinberger (Eds.), *Interlanguage Phonology: The Acquisition of a Second Language Sound System*. Cambridge: Newbury House Publishers.
Tesar, B., & Smolensky, P. (2000). *Learnability in optimality theory*. Cambridge, Mass: MIT Press.
Thomas, E. R. (2000). Spectral differences in /ai/ offsets conditioned by voicing of the following consonant. *Journal of Phonetics, 28*, 1-25.
Vendelin, I., & Peperkamp, S. (2006). The influence of orthography on loanword adaptations. *Lingua, 116*, 996-1007.
Vennemann, T. (1988). *Preference Laws for Syllable Structure*. Berlin: Mouton de Gruyter.
Westbury, J. R. (1983). Enlargement of the supraglottal cavity and its relation to stop consonant voicing. *The Journal of the Acoustical Society of America, 73*, 1322-1336.
Westbury, J. R., & Keating, P. A. (1986). On the Naturalness of Stop Consonant Voicing. *Journal of Linguistics, 22*, 145-166.
Winkler, I., Reinikainen, K., & Näätänen, R. (1993). Event-Related Brain Potentials Reflect Traces of Echoic Memory in Humans. *Perception and Psychophysics, 53*, 443-449.
Yanagisawa, K. (2005). *A cross-linguistic study of the transitional vocalic element in plosive + /r/ clusters.* Master's Thesis. University College London.
Yang, B. (1996). A comparative study of American English and Korean vowels produced by male and female speakers. *Journal of Phonetics, 24*, 245-261.
Yip, M. (2002). Perceptual influences in Cantonese loanword phonology. *Journal of the Phonetic Society of Japan, 6*, 4-21.
Yip, M. (2006). The symbiosis between perception and grammar in loanword phonology. *Lingua, 116*, 950-975.
Ylinen, S., Shestakovaa, A., Huotilainen, M., Alku, P., & Näätänen, R. (2006). Mismatch negativity (MMN) elicited by changes in phoneme length: A cross-linguistic study. *Brain Research, 1072*, 175-185.
Zhao, S., & Stevens, K. N. (2003). Gestural overlap of stop-consonant sequences: Evidence from analysis and synthesis. *Acoustical Society of America Journal, 113*, 2253.
Zoll, C. (1996). *Parsing Below the Segment in a Constraint Based Framework.* Ph.D. Dissertation. University of California, Berkeley.
Zoll, C. (1997). Conflicting directionality. *Phonology, 14*, 263-286.
Zoll, C. (2004). Positional asymmetries and licensing. In J. J. McCarthy (Ed.), *Optimality Theory in Phonology: A Reader*. Malden, MA and Oxford: Blackwell.
Zsiga, E. (1994). Acoustic evidence for gestural overlap in consonant sequences. *Journal of Phonetics, 22*, 121-140.
Zsiga, E. (2000). Phonetic alignment constraints: consonant overlap and palatalization in English and Russian. *Journal of Phonetics, 28*, 69-102.
Zsiga, E. (2003). Articulatory timing in a second language. *Studies in Second Language Acquisition, 25*, 399-432.
Zsiga, E. (2011). External Sandhi in a second language: the phonetics and phonology of obstruent nasalization in Korean-accented English. *Language, 87*, 289-345.

Appendix 1. The reading passage in reading experiment (Chapter 2)

A description of the story:
"Secret of Miden"

Miden was a peaceful world filled with human and magical beings like fairies, knights, and magicians. It was a prosperous and rich land until a race of evil spirits invaded the land. To survive, all races of this world banded together and fought against their ultimate devastation. Working together, they also found the secret of the Ancient City.

Experimental items embedded in sentence:

Petnal is a charming and somewhat mysterious spellcaster.
Tibnan met Frenny while travelling on the streets of Miden.
Kiknal is one of the most powerful spells that she casts.
Rubert can fly from his hometown to Miden.
Tepman is his very good friend.
Pednan is the Captain of the Black Fire, Ficky's air force. He is also a knight, but acts nothing like one.
Segmal is a member of the Black Fire. She is very cooperative and quick to jump into new requests.
Tigmal found Anth wounded in the forest.
Pitman Village is where they first met.
Among the troops of soldiers, Kipman is known to be the best.
Tecknal is a giant one-handed sword made by Jinth.
Jegamal is Flinker's sword and it was stolen.
Kedmal is the twin brother of Kinth, with a spirit as pure as snow.
Manth is a knight and Ketamal is a fairy.
Pidmal can unleash very powerful magic with Reckton.
Kignan has a strained relationship with Reckton and Stonell.
Sebnan Church is where Stonell was anointed as a knight.
Tebanal is his hometown, a well-known city in Miden.
Rath is more interested in developing his love triangle between himself, Veth and Rell.
Ketman is a young human, and Kibmal fell in love with him. However, they know their destiny cannot be the same.
Tegnan is the Ancient city in Miden and Rimart used to research the secrets of this city.
Zepanan is Linth's old friend and she has abundant wisdom.
Tipnal is a lonely warrior who seeks love.
Tekaman, her father, always looks after her.
Kitnal and Linth are Leckton's childhood friends. They still care about Leckton.
Vicker and Filmart are fairies but long to be human.
Seckman met Vell in the mountain Heth.
Tebmal started a great fight with them to save Miden there.
Kidnan is a legendary magician.
Tikman is Kidnan's sould mate, who is warm and caring.
Sepnal lost most of his power after he casted spells to Zenth.
Pedaman saved him.

Appendix 2.The experimental and filler items used in the reading experiment (Chapter 2)

Experimental items:
Note: words are written in orthography

Vowel [ɪ]		Vowel [ɛ]	
Voiceless	Voiced	Voiceless	Voiced
Kipman Tipnal	Kibmal Tibnan	Tepman Sepnal	Tebmal Sebnan
Kitnal Pitman	Kidnan Pidmal	Petnal Ketman	Pednan Kedmal
Tikman Kiknal	Kignan Tigmal	Seckman Tecknal	Segmal Tegnan

Filler items:

Veth	Leckton	Jinth	Rubert
Anth	Rell	Rath	Ficky
Manth	Rimart	Zenth	Vicker
Linth	Flinker	Kinth	Vell
Heth	Frenny	Filmart	Reckton

Zepanan	Tebanal
Ketamal	Pedaman
Tekaman	Jegamal

Appendix 3. The experimental and filler items in the repetition experiment (Chapter 2)

Experimental Items (auditory):

Vowel [ɪ]		Vowel [ɛ]	
Voiceless	Voiced	Voiceless	Voiced
[kʰɪpmən] [tʰipnən]	[kʰɪbmən] [tʰibnən]	[tʰɛpmən] [sɛpnəl]	[tʰɛbmən] [sɛbnəl]
[pʰɪtmən] [kʰɪtnəl]	[pʰɪdmən] [kʰɪdnəl]	[kʰɛtməl] [pʰɛtnən]	[kʰɛdməl] [pʰɛdnən]
[tʰɪkmən] [pʰɪknəl]	[tʰɪgmən] [pʰɪgnəl]	[sɛkməl] [tʰɛknən]	[sɛgməl] [tʰɛgnən]

Filler items (auditory):

[sɛpkən]	[ɛpkəl]	[ʃɪptən]	[ɪptəl]
[sɛtkən]	[ɛtkəl]	[ʃɪtpən]	[ɪtpəl]
[sɛkpən]	[ɛkpəl]	[ʃɪktən]	[ɛktəl]

[sɛpəməl]	[sɛbəməl]	[tʰɛpənəl]	[tʰɛbənəl]
[kʰɛtəməl]	[kʰɛdəməl]	[kʰɛtənəl]	[kʰɛdənəl]
[tʰɛkəməl]	[tʰɛgəməl]	[pʰɛkənəl]	[pʰɛgənəl]

Appendix 4. Other error types found in the production experiments

The errors from the reading experiment:

	Voiceless				Voiced			
	Bilabial	Coronal	Velar	total	Bilabial	Coronal	Velar	total
Fricativization	3	5		8	7	3		10
Place change of the stop	3	6	6	15	1	1	9	11
Voicing	11	7	14	32	-	-	-	-
Metathesis	1			1	3			3
Onsetchange		1		1				0
Voicing+Place Change		1	1	2				0

The errors from the repetition experiment:

	Voiceless				Voiced			
	Bilabial	Coronal	Velar	total	Bilabial	Coronal	Velar	total
Fricativization	0	0	1	1	1	0	2	3
Place change	13	9	1	23	8	1	0	9
Nasal+Place Change	1	1	1	3	1	1	0	2
Voicing	0	0	5	5	-	-	-	-
Metathesis	0	2	0	2	0	0	0	0
Deletion	1	1	0	2	1	0	0	1
Voicing+Vowel Insertion	0	3	4	7	0	0	1	1
Change nasal to stop	17	3	2	22	0	1	1	2

Appendix 5. Production patterns from individual participants

The errors from the reading experiment by subjects:

Participant		Voiceless stop-nasal			Voiced stop-nasal		
		Bilabial	Coronal	Velar	Bilabial	Coronal	Velar
1	Correct	5	7	5	0	2	5
	Nasalization	1	0	1	0	1	0
	Devoicing	0	0	0	1	3	0
	Vowel Insertion	2	1	2	6	2	3
2	Correct	6	6	6	1	0	5
	Nasalization	0	1	0	4	0	0
	Devoicing	0	0	0	2	1	1
	Vowel Insertion	0	0	0	0	7	1
3	Correct	8	8	7	4	5	3
	Nasalization	0	0	1	2	1	0
	Devoicing	0	0	0	1	2	1
	Vowel Insertion	0	0	0	0	0	4
4	Correct	6	1	3	1	0	0
	Nasalization	0	0	0	0	0	0
	Devoicing	0	0	0	0	0	0
	Vowel Insertion	1	7	4	6	7	8
5	Correct	6	3	7	1	3	3
	Nasalization	0	2	0	0	0	0
	Devoicing	0	0	0	7	4	5
	Vowel Insertion	0	0	0	0	1	0
6	Correct	3	4	1	2	5	1
	Nasalization	5	2	6	3	0	0
	Devoicing	0	0	0	0	0	0
	Vowel Insertion	0	0	0	2	3	6
7	Correct	3	6	5	5	6	5
	Nasalization	1	1	0	2	0	0
	Devoicing	0	0	0	0	2	2
	Vowel Insertion	0	0	0	0	0	1
8	Correct	5	7	5	4	2	4
	Nasalization	0	0	0	0	0	0
	Devoicing	0	0	0	2	6	0
	Vowel Insertion	0	1	3	0	0	4
9	Correct	2	4	3	1	1	4
	Nasalization	4	1	3	4	1	1
	Devoicing	0	0	0	2	6	1
	Vowel Insertion	0	0	0	0	0	1
10	Correct	2	7	1	1	3	3

	Nasalization	5	0	6	4	1	0
	Devoicing	0	0	0	0	3	1
	Vowel Insertion	0	1	0	3	1	4
11	Correct	5	5	7	1	1	4
	Nasalization	1	1	0	3	0	0
	Devoicing	0	0	0	4	7	3
	Vowel Insertion	0	0	0	0	0	1
12	Correct	0	4	3	0	4	3
	Nasalization	4	2	4	8	4	2
	Devoicing	0	0	0	0	0	0
	Vowel Insertion	0	0	0	0	0	3
13	Correct	5	4	5	0	0	4
	Nasalization	0	0	0	1	0	0
	Devoicing	0	0	0	5	6	0
	Vowel Insertion	0	0	3	1	0	4
14	Correct	6	8	6	2	2	4
	Nasalization	0	0	0	0	0	0
	Devoicing	0	0	0	6	6	4
	Vowel Insertion	0	0	0	0	0	0
15	Correct	6	8	6	5	3	4
	Nasalization	0	0	0	0	0	1
	Devoicing	0	0	0	2	5	1
	Vowel Insertion	0	0	2	1	0	2
16	Correct	1	1	1	0	0	1
	Nasalization	1	1	1	1	0	0
	Devoicing	0	0	0	0	0	0
	Vowel Insertion	4	6	6	7	8	7
17	Correct	2	3	6	0	1	0
	Nasalization	2	1	1	3	2	1
	Devoicing	0	0	0	0	0	0
	Vowel Insertion	2	2	1	5	5	7
18	Correct	5	5	3	1	2	1
	Nasalization	1	0	0	4	2	0
	Devoicing	0	0	0	0	0	0
	Vowel Insertion	0	2	3	3	4	6
19	Correct	6	7	7	1	3	5
	Nasalization	0	0	0	0	0	0
	Devoicing	0	0	0	7	4	2
	Vowel Insertion	0	0	1	0	1	1
20	Correct	7	7	2	0	0	2
	Nasalization	1	1	0	2	0	0

Devoicing	0	0	0	5	7	1
Vowel Insertion	0	0	0	0	0	0

The errors from the repetition experiment:

Participant		Voiceless stop-nasal			Voiced stop-nasal		
		Bilabial	Coronal	Velar	Bilabial	Coronal	Velar
1	Correct	2	6	5	0	1	0
	Nasalization	5	1	1	5	1	0
	Devoicing	0	0	0	3	5	4
	Vowel Insertion	0	0	0	0	1	3
2	Correct	0	0	2	0	0	2
	Nasalization	4	7	5	7	6	0
	Devoicing	0	0	0	0	0	0
	Vowel Insertion	0	0	1	1	1	6
3	Correct	2	1	5	0	0	4
	Nasalization	4	4	0	5	0	0
	Devoicing	0	0	0	0	2	0
	Vowel Insertion	0	0	1	2	6	3
4	Correct	7	7	7	0	0	2
	Nasalization	1	1	0	3	0	0
	Devoicing	0	0	0	5	8	2
	Vowel Insertion	0	0	0	0	0	4
5	Correct	1	5	8	0	0	0
	Nasalization	4	1	0	1	4	2
	Devoicing	0	0	0	6	1	2
	Vowel Insertion	0	0	0	0	2	3
6	Correct	5	7	7	3	0	2
	Nasalization	1	0	0	1	0	0
	Devoicing	0	0	0	2	8	0
	Vowel Insertion	0	0	1	0	0	6
7	Correct	5	8	7	0	0	1
	Nasalization	0	0	0	0	0	0
	Devoicing	0	0	0	7	8	3
	Vowel Insertion	0	0	0	0	0	4
8	Correct	6	6	7	2	4	3
	Nasalization	0	0	0	0	1	0
	Devoicing	0	0	0	5	3	2
	Vowel Insertion	0	0	1	1	0	3

9	Correct	8	7	4	3	3	2
	Nasalization	0	0	0	0	0	0
	Devoicing	0	0	0	4	5	0
	Vowel Insertion	0	0	0	1	0	6
10	Correct	7	7	7	0	4	3
	Nasalization	0	0	0	1	0	0
	Devoicing	0	0	0	5	3	0
	Vowel Insertion	0	0	0	0	1	4
11	Correct	6	8	8	1	2	1
	Nasalization	0	0	0	0	2	0
	Devoicing	0	0	0	7	4	4
	Vowel Insertion	0	0	0	0	0	3
12	Correct	8	8	7	2	6	3
	Nasalization	0	0	0	0	0	0
	Devoicing	0	0	0	4	2	0
	Vowel Insertion	0	0	0	1	0	5
13	Correct	8	8	8	2	2	5
	Nasalization	0	0	0	0	0	0
	Devoicing	0	0	0	6	6	3
	Vowel Insertion	0	0	0	0	0	0
14	Correct	6	5	8	0	3	5
	Nasalization	0	0	0	1	0	0
	Devoicing	0	0	0	7	4	2
	Vowel Insertion	0	0	0	0	0	1
15	Correct	5	8	8	1	6	6
	Nasalization	0	0	0	0	0	0
	Devoicing	0	0	0	5	2	0
	Vowel Insertion	0	0	0	1	0	2
16	Correct	5	6	7	0	2	1
	Nasalization	0	0	0	0	0	0
	Devoicing	0	0	0	8	6	0
	Vowel Insertion	0	0	1	0	0	7
17	Correct	7	7	7	0	1	3
	Nasalization	0	0	0	1	0	0
	Devoicing	0	0	0	5	5	3
	Vowel Insertion	0	0	1	0	2	2
18	Correct	7	7	7	0	5	5
	Nasalization	0	0	0	1	0	0
	Devoicing	0	0	0	7	3	0
	Vowel Insertion	0	0	0	0	0	3

19	Correct	8	8	7	3	6	7
	Nasalization	0	0	0	0	0	0
	Devoicing	0	0	0	3	0	0
	Vowel Insertion	0	0	0	2	2	1
20	Correct	6	8	7	1	2	1
	Nasalization	0	0	0	0	0	0
	Devoicing	0	0	0	7	4	1
	Vowel Insertion	0	0	1	0	2	6

Appendix 6. The items that contain Korean lexical vowel [ɨ] (Chapter 6)

The reading passage written in Korean:

하근이에 대한 이야기.
하근인 몇일전 여자친구와 헤어진 후 괴로워한다.허브티 한잔이면 추운 겨울 몇시간이고 수다를 떨며 그녀와 함께 시간을 보냈었다. 시큰한 느낌이 코끝에 몰려온다. 그녀와의 추억은 참 많다. 히트곡, ‘비가 내리면’을 함께 들으며 비가 오는 창밖을 바라보기도 했고, 머그컵 너머로 풍기는 커피향에 함께 취하기도 했다. 그녀의 강아지 보듬이도 생각난다. 보듬인 늘 꼬리를 흔들며 하근이를 따라다녔었다. 머드팩 한아름 사들고 만나러 나와 하근이의 안좋아진 피부를 걱정했던 그녀였다. 리프트 위에서 ‘사랑해’라며 외쳐 스키장에 같이간 친구들마저 놀래켰던 그녀, 그녀가 떠난것이다. 다그쳐 왜 떠났냐고 묻는 것은 부질없는 짓이다.

하근이의 가장 친한 친구 의석이는 하근이에게 전화를 걸었다. 깊으면 병이 되는게 실연의 아픔인거라며 털고 일어나라고 위로해준다. 밥으로 빈가슴을 채우라고 하근이를 음식점에도 데리고 갔다. 갑숙이 누나는 하근이 누나인데 같이 나왔다. 식순이 누나라며 하근이는 누나를 많이 놀려댔었는데 힘든일이 있을땐 늘 하근이 편이 되어준다. 합심이 잘되는 누나랑 의석이는 오늘도 하근이를 위해 함께 나온것이다. 시큼해 더 풍미가 좋은 샐러드를 꼭꼭 씹어 먹는다. 개그맨 ‘왕수혁’이 텔레비젼에 나와 청중을 웃기고 있다. 하근이는 다리를 꼬고 앉아있다가, 삭신이 쑤시는것 같다. 다리를 비틀면 좀 시원할까. 비트니 좀 나아진것 같다. 겨우 배를 채운 하근이는 계산대로 간다. 사실 누나가 매번 밥을 사주어서 오늘도 누나가 밥 사주는게 괜히 미안해진다. 빚이라는 생각이 들어, 갚으면 좀 속이 시원해질 것 같다.

하근이는 해질무렵 집에 돌아왔다. 코브라 두마리가 주인을 알아보고 인사하자 하근이는 씨익 웃으며 방에 들어가 컴퓨터를 켠다. 피크닉 가자며 봄을 기다리던 그녀 사진이 보인다. 노트북 화면에 떠있는 그녀사진이 이제 낯설어진다. 다듬이 소리처럼 가슴을 울리는 슬픔이 밀려온다. 별의 별 생각이 다 든다. 밭으로 내려가 농사나 지으며 살까 하는 생각. 삽으로 땅을 일구며 땀을 흘리면 좀 나아지려나 하는 생각. 헤드폰 끼고서 큰 소리로 헤비메탈 음악을 몇날 몇일 들어볼까 하는 생각. 하지만 이제 그녀를 그만 잊기로 한다. 스크롤 막대를 아래로 내려 ‘삭제’를 누른다. 스프링 노트에 적혀있는 그녀와의 추억도 이제 그만 서랍 깊숙이 묻어 버린다. 하근이는 한결 가벼워진 마음으로 잠자리에 든다.

Korean native words				Loanwords			
깊으면 [kip^{h}ɨmjən]	갚으면 [kap^{h}ɨmjən]	밥으로 [pabɨɾo]	삽으로 [sabɨɾo]	리프트 [ɾip^{h}ɨt^{h}ɨ]	스프링 [sɨp^{h}ɨɾiŋ]	코브라 [k^{h}obɨɾa]	허브티 [həbɨt^{h}i]
비트니 [pit^{h}ɨni]	밭으로 [pat^{h}ɾo]	다듬이 [tadɨmi]	보듬인 [podɨmin]	히트곡 [hit^{h}gok]	노트북 [not^{h}ɨbuk]	헤드폰 [hɛdɨp^{h}on]	머드팩 [mədɨp^{h}ɛk]
시큼해 [ʃik^{h}ɨmhɛ]	시큰한 [ʃik^{h}ɨnhan]	다그쳐 [tagɨtʃə]	하근인 [hagɨnin]	피크닉 [p^{h}ik^{h}ɨnik]	스크롤 [sɨk^{h}ɨɾol]	개그맨 [kɛgɨmɛn]	머그컵 [məgɨk^{h}əp]