THE ROLE OF PHONOLOGY AND PHONETICS IN THE ADAPTATION OF ENGLISH WORDS INTO STANDARD CHINESE

By

Li-jen Shih

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

Linguistics–Doctor of Philosophy

2019

ABSTRACT

THE ROLE OF PHONOLOGY AND PHONETICS IN THE ADAPTATION OF ENGLISH WORDS INTO STANDARD CHINESE

By

Li-jen Shih

We can distinguish three approaches in the literature as to how borrowed foreign words are adapted to comply with the host language sound system: the purely perceptual approach, which claims that the adaptation occurs during perception beyond the listeners' conscious awareness, the purely phonological approach, which claims that the underlying representations of the source words are the input to the adaptation and mapped in the host language lexicon to the structurally closest native representations, and the hybrid approach, which claims that the adaptation occurs in the host language production grammar and that the host language production grammar makes direct reference to the phonetic information in the source words. This dissertation evaluates English loanwords in Standard Chinese (SC) against these three approaches, and the results generally support the hybrid approach. While the hybrid approach is supported, it falls short of explaining some of the loanword data. Two problems are pointed out and the solutions are proposed.

SC evidence against the purely perceptual approach comes from the treatment of word-final stops or coda liquids in borrowed English words. These sounds may delete (e.g. *Flint* → [fʷo.lin.tʰɤ] vs. *Clint* → [kʰɤ.lin__], *Arkin* → [a.ər.tɕin] vs. *Starr* → [ʂz̩.ta__]), and according to the purely perceptual approach, they delete simply because the borrowers do not hear them in the source words. However, when the source words are monosyllabic and contain no other sounds that violate the SC syllable structure requirements, these sounds are always preserved with the place of articulation correctly identified (e.g. *Ed* → [ai.tɤ], *Pink* → [piŋ.kʰɤ], *Barr* → [pa.ər]).

These data indicate that the borrowers know exactly what sounds are present in the source words and that the deletion must occur in the SC production grammar.

Data from a number of languages also demonstrate that the adaptation does not occur in the host language lexicon and that information viewed by the source and host languages as phonetic details plays a critical role in determining the output. An example is the SC adaptation of /ʃ, tʃ, dʒ/ in borrowed English words. These sounds are found to map to [ɕ, tɕ^{h}, tɕ] much more frequently than to [ʂ, ʈʂ^{h}, ʈʂ]. First, [ɕ, tɕ^{h}, tɕ] are allophones, so they are not available in the SC lexicon as potential matches. Second, according to Flemming (2003), /ʃ, tʃ, dʒ/ and [ɕ, tɕ^{h}, tɕ] are produced with a front tongue body but [ʂ, ʈʂ^{h}, ʈʂ] are produced with a back tongue body. Thus, the strong tendency of /ʃ, tʃ, dʒ/ being adapted as [ɕ, tɕ^{h}, tɕ] can be interpreted as the result of the borrowers' intention of maintaining the front tongue-body position associated with /ʃ, tʃ, dʒ/, which both English and SC consider a phonetic detail.

This dissertation presents three detailed case studies of English loanwords in SC. The first case study concerns an asymmetry found in the adaptation of tautosyllabic sibilant+/w/ and velar stop+/w/ sequences (e.g. *Sweet*, *Quincy*), the second case study concerns several asymmetries found in the adaptation of vowel+nasal consonant+vowel sequences, and the third case study concerns an asymmetry found in the adaptation of vowel+coda liquid sequences. To account for these adaptation asymmetries, I show previous phonetic findings and argue that the SC production grammar makes direct reference to non-contrastive perceptual and/or articulatory information contained in the source words. Perceptual experiments are conducted for the first and third case studies, and the results are in line with my argument. These sets of data strongly support the hybrid approach. These three case studies are also formalized in the framework of the Optimality Theory.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER ONE

INTRODUCTION

A number of models have been proposed as to how borrowed foreign words are adapted to conform to the phonological system of the host language. These models can be classified into three approaches in terms of their view on where the adaptation occurs and whether the phonetic information in the source words plays a role in the adaptation process: The purely perceptual approach holds that the adaptation occurs during perception, the purely phonological approach holds that the adaptation occurs in the host language lexicon, and the hybrid approach holds that the adaptation occurs in the host language production grammar while making direct reference to the phonetic information in the source words. This dissertation evaluates English loanwords in Standard Chinese (SC) against the three approaches, and the results generally support the hybrid approach. Although the hybrid approach is supported, it falls short of explaining some of the loanword data. Two problems are pointed out and the solutions are proposed. The adaptation of tautosyllabic sibilant+/w/ and velar stop+/w/ sequences, word-internal vowel+nasal+vowel sequences, and coda liquids in English loanwords in SC are studied separately. The analyses are in line with the hybrid approach and formalized in the framework of the Optimality Theory. Loanword data from other languages are also discussed when appropriate.

This dissertation is organized as follows. I begin in Chapter One with an overview of the main chapters, the introduction of the SC sound inventory and syllable structure, and the discussion of issues such as how the SC loanword data are collected and whether the pronunciations of the English words in the corpus are based on General American English or Received Pronunciation. Chapter Two is the literature review, where I discuss the proposals and predictions of the three approaches to loanword adaptation. Chapters Three and Four present evidence against the purely perceptual approach and the purely phonological approach, respectively. In Chapter Five, I point out two problems for the hybrid approach and propose the solutions. Chapters Six–Eight are three case studies of English loanwords in SC. Chapter Six concerns an asymmetry in how tautosyllabic sibilant+/w/ and velar stop+/w/ sequences are adapted (e.g. *Sweet* → [ʂʐ̩.wei.tʰɤ] vs. *Quincy* → [kʰʷən.ʂʐ̩]), Chapter Seven concerns several asymmetries found in the adaptation of vowel+nasal consonant+vowel sequences regarding the nasal consonant-gemination frequency (e.g. *Cannon* → [kʰan.nuŋ] vs. *Obama* → [ou.pa.ma]), and Chapter Eight concerns an asymmetry found in the adaptation of vowel+coda liquid sequences regarding the coda liquid-deletion frequency (e.g. *Beard* → [pi.ər.tɤ] vs. *Mark* → [ma.kʰɤ]). Chapter Nine is the conclusion, future studies, and theoretical implications.

1.1 Overview of the Dissertation

In this section, I briefly introduce the three approaches to loanword adaptation and the three case studies of English loanwords in SC that I will be presenting.

1.1.1 The Purely Perceptual Approach

In the view of the purely perceptual approach, loanword adaptation occurs during perception, where the proponents of this approach claim that non-native sounds and structures are perceived as native ones that have the closest phonetic properties beyond the listeners' conscious awareness (unfaithful perception) (Peperkamp and Dupoux 2001, 2003; Vendelin and Peperkamp 2004; Peperkamp 2005; Peperkamp, Vendelin and Nakamura 2008; Boersma and Hamann 2009; Calabrese 2009; Broselow 2009; Crawford 2007; Takagi and Mann 1994). There is evidence that loanwords are adapted in the host language production grammar. First, my corpus shows that while word-final stops and coda liquids may be deleted, these sounds are always preserved with the place of articulation correctly identified when the source words are monosyllabic and contain no other sounds violating the SC syllable structure requirements (e.g. *E<u>d</u>* → [ai.<u>tɤ</u>], *Ba<u>rr</u>* → [pa.<u>ər</u>]). This observation implies that the borrowers know what sounds are present in the source words and that the deletion of these sounds should not be attributed to unfaithful perception but is the outcome of phonological processes taking place in the host language production grammar. Second, previous perceptual studies have shown that Japanese/Korean speakers are capable of distinguishing English word-initial /r/ and /l/ from the Japanese/Korean single liquid category (realized as [ɾ]), so the fact that the two English sounds in word-initial position are always mapped to [ɾ] in borrowed English words should not be attributed to unfaithful perception either.[1] Since unfaithful perception does not occur, the featural change must take place in the host language production grammar.

1.1.2 The Purely Phonological Approach

By maintaining that loanword adaptation is performed by bilinguals, the purely phonological approach argues that the underlying representations of the source words serve as the input to the adaptation (Hyman 1970; Paradis and Lebel 1994; Paradis 1995a, b, 2006; Paradis and LaCharité 1997, 2001; LaCharité and Paradis 2002, 2005; Ulrich 1997; Rose 1999; Jacobs and Gussenhoven 2000; Shinohara 2004; Herd 2005; Uffmann 2006; Rose and Demuth 2006; Paradis and Tremblay 2009). This view entails that information that is regarded as non-contrastive or phonetic in the source words plays no role in the adaptation process. Evidence from English loanwords in SC, however, indicates that it does. First, I will show that a /v/'s frication-noise level and the tongue-body position during the production of /ʃ, tʃ, dʒ/ determine how these sounds are adapted. Second, I will show that stops undergo deletion much more frequently than fricatives, affricates, and /m/ and that, when transliterating Korean, Southern Min,

[1] To avoid confusion, sounds in the source language are placed between slashes and sounds in the host language are placed between bracket squares throughout this dissertation.

or Cantonese words into SC, coda stops, which are strictly unreleased in the source languages, are almost always deleted. These observations demonstrate that perceptual saliency plays a critical role in the adaptation process.

1.1.3 The Hybrid Approach

The hybrid approach holds that while loanword adaptation occurs in the host language production grammar, unfaithful perception occurs (Silverman 1992; Yip 1993, Kenstowicz 2003) or the adaptation takes into consideration the phonetic details contained in the source words in determining the output[2] (Yip 1993, 2002, 2006; Steriade 2001, 2009; Kang 2003; Kenstowicz 2005, 2006, 2007, 2012; Ito, Kang, and Kenstowicz 2006; Shinohara 2006; Adler 2006; Davis and Cho 2006; Kubozono 2006; Rose and Demuth 2006; Vendelin and Peperkamp 2006; Hsieh, Kenstowicz and Mou 2009; Smith 2006, 2009; Dohlus 2005; among many others). Due to this position, the data posing challenges to the purely perceptual approach and the purely phonological approach can be explained in most of the models classified into this approach. However, two problems are observed. First, while word-medial stops preceding a homorganic nasal or another stop in English are normally considered unreleased and therefore expected to delete in English loanwords in SC, my corpus shows that these sounds are preserved frequently. On the basis of acoustic/perceptual studies and research on second language acquisition, I argue that these sounds are perceived as released by the SC borrowers sometimes. Second, word-final stops may be unreleased in English, but when these sounds are voiceless and must be preserved to meet some SC requirement in the adaptation process, they always surface with aspiration even though SC unaspirated stops are the better matches. A similar problem is also found in English loanwords in Korean. To resolve this problem, I propose that the borrowers must have heard two

---

[2] This implies that the phonetic details contained in the source words are included in the input to the production grammar. Yip (2006) calls this input “a non-native percept”. As sounds and structures in the source words are not perceived as native ones, I simply call it faithful perception.

versions of every learned English word ending in a voiceless stop and stored them in their long-term memory (see Calabrese 2009 for discussion of how the long-term memory functions in perception). In one version, the stop is released, and, in the other version, it is unreleased. When an unreleased word-final voiceless stop in a borrowed English word must be preserved to satisfy some native requirement, to avoid the situation in which the word-final stop and its voiced counterpart are mapped to the same SC sound (voiced stops always surface without aspiration when preserved in English loanwords in SC), the borrowers retrieve the released version of the same word from their long term memory and substitute it for the current word, resulting in the observation that every preserved word-final voiceless stop surfaces with aspiration.

1.1.4 Three Case Studies

Chapters Six–Eight are three case studies of English loanwords in SC. In these case studies, I demonstrate that the phonetic information in the source words plays a critical role in the adaptation and also provide formal analyses of the observed adaptation tendencies.

Chapter Six concerns how tautosyllabic velar stop+/w/ and sibilant+/w/ sequences are adapted. My corpus shows that the velar stop+/w/ sequences are mostly realized with the /w/ merging into the velar stop (e.g. *Quincy* → [kʰʷən.ʂʐ̩]), while the sibilant+/w/ sequences are always realized with the sibilant forming a separate syllable (e.g. *Sweet* → [ʂʐ̩.wei.tʰɤ]). Two perceptual similarity hierarchies are proposed. According to the first hierarchy, a syllable-initial sibilant+/w/ sequence (e.g. *Sweet*) is perceptually less similar to the labialized version of the sibilant (e.g. *Sweet* → *[ʂʷei.tʰɤ]) than to the sibilant+epenthesis+[w] sequence (e.g. *Sweet* → [ʂʐ̩.wei.tʰɤ]); according to the second hierarchy, a syllable-initial velar stop+/w/ sequence (e.g. *Quincy*) is perceptually less similar to the velar stop+epenthesis+[w] sequence (e.g. *Quincy* → *[kʰɤ.wən.ʂʐ̩]) than to the labialized version of the velar stop (e.g. *Quincy* → [kʰʷən.ʂʐ̩]). I will show that previous articulatory findings are in line with these two hierarchies. In addition, perceptual experiments aiming to test these two hierarchies are conducted, and the results support these two hierarchies as well.

In Chapter Seven, I look at the adaptation of word-medial vowel+nasal consonant+vowel (VNV) sequences. The nasal consonant is found to geminate sometimes (e.g. *C<u>anno</u>n* → [kʰ<u>an.nu</u>ŋ]), and several asymmetries are observed in terms of the gemination frequency: (i) The nasal consonant geminates more often after a low monophthong than after a non-low monophthong; (ii) the nasal consonant geminates more often after a non-low lax vowel than after a non-low tense vowel or a diphthong; (iii) the nasal consonant geminates more often after a stressed vowel than after an unstressed vowel; (iv) the nasal consonant geminates more often when it is /n/ than when it is /m/.

I argue that these asymmetries have a phonetic basis. First, it has been reported in the literature that low vowels exhibit greater extent of nasalization than non-low vowels before a nasal consonant in English (explaining the asymmetry in (i)). Second, in English, non-low lax vowels are shorter than non-low tense vowels, and, in SC, vowels in a closed syllable are shorter than the same vowels in an open syllable. Given that the borrowers would like to maintain vowel length, when the pre-nasal vowel in the source word is short, the nasal consonant tends to geminate to close the syllable (e.g. *K<u>inney</u>* → [tɕ<u>in.ni</u>]), but when the pre-nasal vowel in the source word is long, the motive for closing the syllable does not exist (e.g. *K<u>eene</u>r* → [tɕʰ<u>i.na</u>]) (explaining the asymmetry in (ii)). Third, previous articulatory studies have indicated that the pre-nasal vowels in word-medial VNV sequences in English exhibit more extensive nasalization when they are stressed than when they are not (explaining the asymmetry in (iii)). Fourth, I argue that the asymmetry in (iv) occurs because an intervocalic /m/ is more resistant to being mapped to [n.m] or [ŋ.m] than an intervocalic /n/ to being mapped to [ŋ.n]. Following Adler's (2006) proposal that coronals and dorsals are articulatorily more similar to each other than to labials, I suggest that the asymmetry in (iv) can be explained in terms of articulatory similarity between the three places of articulation.

Chapter Eight investigates the adaptation of coda liquids. My corpus shows that coda liquids tend to be preserved after front vowels but tend to delete after non-front vowels. To account for this asymmetry, I propose that, in English, coda liquids are perceptually more distinctive after a front vowel than after a non-front vowel.

Previous articulatory findings have indicated that the vocalic gestures of /l/ and /ɹ/ are very similar to the dorsal gestures of /ɔ/ and /ə/, respectively. Thus, a vowel+coda liquid sequence can be thought of as the vowel immediately followed by /ɔ/ or /ə/. It has also been found that the pharyngeal widths measured during the production of vowels correlate with the vowels' height and backness, and we can infer from the reported measurements that, generally, /ɔ/ (= vocalic gesture of /l/) and /ə/ (= vocalic gesture of /ɹ/) are more similar to non-front vowels than to front vowels. This inference suggests that the production of a front vowel+coda liquid sequence involves a greater change in tongue position than the production of a non-front vowel+coda liquid sequence. Because a great change in tongue position entails a great change in acoustic signal, the proposal that a coda liquid is perceptually more distinctive after a front vowel than after a non-front vowel is confirmed.

The proposal is also supported by the results of perceptual experiment, which show that, for native SC speakers, the contrast between an English consonant+non-front vowel+liquid syllable and an English or SC consonant+non-front vowel syllable (e.g. E. /hɑɹ.gi/–E. /hɑ.gi/, E. /tɔl.bi/–SC [tou.pi]) is generally smaller than the contrast between an English consonant+front vowel+liquid syllable and an English or SC consonant+front vowel syllable (e.g. E. /hiɹ.gi/–E. /hi.gi/, E. /dil.bi/–SC [ti.pi]).

1.2 Data Collection

The SC loanword data examined in this dissertation are collected from the web sites of two of the largest Taiwan-based newspapers United Daily and China Times (udn.com and chinatimes.com). While written Chinese consists of characters which do not constitute an alphabet, every Chinese character represents a syllable of spoken Chinese. As a result, the sequences of characters in the corpus represent the SC sound strings that the borrowers think are most similar to the corresponding source words in terms of pronunciation. For example, the proper name *Mark* is found to correspond to the two-character sequence 马克 in the corpus, which is pronounced [ma.kʰɤ] with a low tone (Tone Three) on the first syllable and a falling tone (Tone Four) on the second syllable. As we can see in this mapping, the word-final /k/ in the source word corresponds to the syllable [kʰɤ] and the rest of the source word corresponds to the syllable [ma], showing that, at least for the borrower that is responsible for this mapping, epenthesizing the vowel [ɤ] at the end of the output form and deleting the coda /ɹ/ in the source word would render the output form most similar to the source word in pronunciation. Since this dissertation does not concern how the tones in the output forms are determined, they are not discussed and not marked for ease of reading.

There is this question of whether the SC output forms are based the spelling pronunciations of the English source words. This is not impossible but there are reasons that the extent of the effect on the arguments made in this dissertation is limited. First, the data are collected from two largest newspapers in Taiwan and the journalists who are responsible for the transliteration are therefore expected to be highly proficient in English. Second, a large part of the arguments made in this dissertation are based on the adaptation of consonants, which are usually pronounced according to how they are spelled in English. Third, a closer look at the data that back the arguments made in this dissertation shows that the influence of spelling pronunciation should be limited. For example, Chapter Eight considers the factors that determine whether a word-internal intervocalic nasal consonant tends to geminate, and a number of cases have been found where gamination occurs but the nasal consonant in the source word is represented by a single grapheme (e.g. *Diana* → [tai.an.na], *Janis* → [ʈʂən.ni.sʐ̩], *Thomas* → [tʰɑŋ.ma.sʐ̩], *Samuel* → [ʂan.mʲou]) or gemination does not occur but the nasal consonant in the source word is represented by a double grapheme (e.g. *Shannon* → [ɕa.nuŋ], *Bannister* → [pa.ni.sʐ̩.tʰɤ], *Jimmy* → [tɕi.mi], *Emma* → [ai.ma]).

The corpus contains 1390 English words, all proper names.[3] Some English words are found to map to more than one SC form. For example, *Hudson* is found to map to [xa.tɤ.sən] and [xa.sən], and *Miller* is found to map to [mi.lɤ] and [mei.lɤ]. In addition, some English words are found to correspond to multiple SC words that have the same pronunciation but are not represented by the same characters. For example, [pʰan.ni], the SC correspondent of *Penny*, is represented by 潘尼 and 潘妮 (the second characters have the same tone). When these occur, they are considered separate mappings. There is a total of 1557 mappings in the corpus.

1.3 General American English or Received Pronunciation

There is this question of whether the pronunciations of the English words in the corpus are based on General American English (GA) or Received Pronunciation (RP). If it is RP, given that RP is non-rhotic (i.e. the vowel+coda /ɹ/ sequences /ɑɹ/, /ɔɹ/, and /əɹ/ are realized as [ɑː], [ɔː],

[3] They are all proper names because foreign words and phrases that are not proper names are mostly borrowed into SC on the basis of the features of the objects or concepts the words or phrases represent. That is, the SC correspondents of these words or phrases are mostly calques. For example, the SC word for *computer* is 电脑, which is pronounced [tian.nɑu] with a falling tone (Tone Four) on the first syllable and a low tone on the second syllable (Tone Three). The first character means ‘electric’ and second means ‘brain’.

and [ɜː], respectively (Kang 2003: 229), and coda /ɹ/ following other vowels is realized as [ə]), the tendency of coda /r/ being deleted after non-front vowels simply can be said to be due to the absence of the coda /r/ from the input (e.g. *Sta<u>rr</u>* /stɑː/ → [ʂʐ̩.ta__]). However, it is observed that the coda /ɹ/ may be preserved, and when this happens, it is always mapped to [ər] (e.g. *A<u>r</u>kin* → [a.<u>ər</u>.tɕin], *Ba<u>rr</u>* → [pa.<u>ər</u>]) (see 3.2.2 and Chapter Eight). These data are difficult to explain. Take the mapping *A<u>r</u>kin* → [a.<u>ər</u>.tɕin] for example. Since *Arkin* is pronounced [ɑːkɪn] in RP, we wonder why the syllable [ər] is present in the SC correspondent. The mapping *Ba<u>rr</u>* → [pa.<u>ər</u>] is another example. *Ba<u>rr</u>* is pronounced [bɑː] in RP, and it is augmented to disyllabicity due to the SC minimum-word requirement. The way in which it is augmented is unexpected. First, we wonder why it is augmented through adding the syllable [ər]. Second, this way is inconsistent with what is found in another set of data in the corpus, which shows that CV source words are augmented by lengthening the vowel (*Bee* → [pi.i], *Lay* → [lei.i], *May* → [mei.i], *Day* → [tai.i]). To solve this problem, we have to assume that the input to the adaptation is the underlying representations of the source words. However, as the overview in 1.1 has shown, the adaptation references the phonetic information in the source words, which implies that the phonetic information in the source words is included in the input. Because the evidence is strong and it seems to be the only way to explain the data, the assumption that the input is the underlying representations of the source words does not hold. These considerations lead us to the conclusion that the pronunciations of the English words in the corpus are based on GA, in which coda /ɹ/ is pronounced.

1.4 SC Sound Inventory and Syllable Structure

I follow Duanmu (2007) and assume that SC has the consonants in (1) at the surface level. Sounds in parentheses are allophones.

(1) SC consonants (based on Duanmu 2007)

	Labial	Dental	Post-Alveolar	Alveolo-Palatal	Velar
Stop	p pʰ	t tʰ			k kʰ
Fricative	f	s	ʂ ʐ	(ɕ)	x
Affricate		ts tsʰ	ʈʂ ʈʂʰ	(tɕ tɕʰ)	
Liquid		l	r		
Nasal	m	n			ŋ

[ɕ, tɕ, tɕʰ] are allophones of /s, ts, tsʰ/, /ʂ, ʈʂ, ʈʂʰ/, or /x, k, kʰ/. They occur when one of the three underlying series precedes /i/ or /y/ (e.g. /si/ → [ɕi], /tsy/ → [tɕy]). /r/ occurs only in the syllable /ər/. /s, ts, tsʰ/ are realized as [sz̩, tsz̩, tsʰz̩] and /ʂ, ʐ, ʈʂ, ʈʂʰ/ as [ʂʐ̩, ʐʐ̩, ʈʂʐ̩, ʈʂʰʐ̩] when they are the only sound in the syllable.

I follow Lin (2008b) and assume that SC has the vowels in (2) at the surface level. Again, sounds in parentheses are allophones. The subscripted “c” in /a_c/ denotes “central”.

(2) SC vowels (based on Lin 2008b)

	Front	Central	Back
High	i y		u
Mid	(e ɛ)	ə	(ɤ o)
Low	(a)	a_c	(ɑ)

Processes regarding the vowels are listed in (3).

(3) Processes regarding SC vowels

i. The high vowels /i/, /y/, and /u/

a. /i/, /y/, and /u/ are realized as [j], [ɥ], and [w], respectively, before a non-high vowel (/iə/ → [je], /yə/ → [ɥe], /ua/ → [wa]).

ii. The mid vowel /ə/

a. /ə/ is realized as [ə] in a closed syllable (e.g. /fən/ → [fən], /məŋ/ → [məŋ]).

b. /ə/ is realized as [ɤ] in an open syllable if it does not follow a high vowel (e.g. /ə/ → [ɤ], /tʰə/ → [tʰɤ]).

c. /ə/ is realized as [e] before or after a high front vowel (e.g. /təi/ → [tei], /iə/ → [je], /yə/ → [ɥe]).

d. /ə/ is realized as [o] before or after a high back vowel (e.g. /kəu/ → [kou], /uə/ → [wo]).

iii. The low vowel /a/

a. /a/ is realized as [a] before or after /i/ or before /n/ (e.g. /nai/ → [nai], /ia/ → [ja], /xan/ → [xan]).

b. /a/ is realized as [a_c] in an open syllable if it does not follow /i/ (e.g. /tʰa/ → [$t^h a_c$], /ua/ → [wa_c]). (For ease of reading, the subscripted "c" is not marked in the transcriptions throughout this dissertation.)

c. /a/ is realized as [ɑ] before /u/ or /ŋ/ (e.g. /ʂau/ → [ʂɑu], /laŋ/ → [lɑŋ]).

d. /a/ is realized as [ɛ] after a high front vowel and before /n/ (e.g. /ian/ → [jɛn], /yan/ → [ɥɛn]).

iv. Diphthongs

a. SC has the four diphthongs [ei], [ou], [ai], and [ɑu].

Underlyingly, a maximum SC syllable is CVVN and a minimum SC syllable is V (C=consonant, V=vowel, N=/n/ or /ŋ/). On surface, I follow Duanmu (2007) and assume the following.

(4) Assumptions regarding SC syllable structure on surface

i. There are two timing slots in the rime. As a result, a monophthong in an open syllable lengthens (e.g. /ti/ → [tii],) but one in a closed syllable does not (e.g. /tan/ → [tan]). (For ease of reading, vowel length is not marked in the transcriptions throughout this dissertation.)

ii. A syllable-initial consonant and a following pre-nuclear high vowel merge. The result of the merger depends on the two segments.

a. When the pre-nuclear high vowel is /u/, it becomes a secondary articulation of the syllable-initial consonant (e.g. /ʂuəi/ → [ʂʷei], /fuə/ → [fʷo], /kua/ →[kʷa]).

b. When the pre-nuclear high vowel is /i/ and the syllable-initial consonant is a member of one of the three series /s, ts, tsʰ/, /ʂ, ʈʂ, ʈʂʰ/, and /x, k, kʰ/,

- /s/, /ʂ/, or /x/ + /i/ → [ɕ] (e.g. /sia/ → [ɕa])
- /ts/, /ʈʂ/, or /k/ + /i/ → [tɕ] (e.g. /tsiə/ → [tɕe])
- /tsʰ/, /ʈʂʰ/, or /kʰ/ + /i/ → [tɕʰ] (e.g. /tsʰia/ → [tɕʰa])

c. When the pre-nuclear high vowel is /y/ and the syllable-initial consonant is a member of one of the three series /s, ts, tsʰ/, /ʂ, ʈʂ, ʈʂʰ/, and /x, k, kʰ/ (we can see /y/ as a combination of /i/ and /u/),

- /s/, /ʂ/, or /x/ + /y/ → [ɕʷ] (e.g. /syə/ → [ɕʷe])
- /ts/, /ʈʂ/, or /k/ + /y/ → [tɕʷ] (e.g. /tsyan/ → [tɕʷɛn])
- /tsʰ/, /ʈʂʰ/, or /kʰ/ + /y/ → [tɕʰʷ] (e.g. /tsʰyə/ → [tɕʷe])

d. When the pre-nuclear high vowel is /i/ or /y/ and the syllable-initial consonant is *not* a member of one of the three series /s, ts, tsʰ/, /ʂ, ʈʂ, ʈʂʰ/, and /x, k, kʰ/, then the pre-nuclear high vowel becomes a secondary articulation of the syllable-initial consonant (e.g. /piau/ → [pʲɑu], /lyə/ → [lᶣe]).

CHAPTER TWO

APPROACHES TO LOANWORD ADAPTATION

A number of models have been proposed as to how loanwords are adapted to conform to the phonological system of the host language. These models can be classified into three groups, which are referred to as the purely perceptual approach, the purely phonological approach, and the hybrid approach here. Table 2.1 shows how these three approaches differ from each other:

Table 2.1 Views of the purely perceptual approach, the purely phonological approach, and the hybrid approach to loanword adaptation on where loanwords are adapted and whether the phonetic information in the source words plays a role in the adaptation

	Where are loanwords adapted?	Does the phonetic information in the source words play a role in the adaptation?
The purely perceptual approach	During perception	Yes
The purely phonological approach	In the host language lexicon	No
The hybrid approach	In the host language production grammar	Yes

As shown in Table 2.1, the purely perceptual approach claims that loanword adaptation occurs during perception, the purely phonological approach claims that loanword adaptation occurs in the host language lexicon, and the hybrid approach claims that loanword adaptation occurs in the host language production grammar while making reference to the phonetic information in the source words.

2.1 The Purely Perceptual Approach

2.1.1 Proposals

The purely perceptual approach holds that adaptations found in loanwords are due to unfaithful perception, which automatically occurs when the borrowers hear source words containing sounds or structures that do not fit in their native languages. In other words, the adaptations are not computed within the host language production grammar; rather, they are the automatic outcomes of unconscious unfaithful perception of the source words (Peperkamp and Dupoux 2001, 2003; Vendelin and Peperkamp 2004; Peperkamp 2005; Peperkamp, Vendelin and Nakamura 2008; Boersma and Hamann 2009; Calabrese 2009; Broselow 2009; Crawford 2007; Takagi and Mann 1994).[4]

This view is based on the Perceptual Assimilation Model promoted by Best and her colleagues, according to which non-native sounds may be classified into the native phonemic categories that have the closet phonetic properties (Best, McRoberts and Sithole 1988; Best and Strange 1992; Best 1994, 1995). The idea of perceptual assimilation is originally proposed to account for the observations that adult speakers usually have difficulty discriminating segmental distinctions that are not employed phonemically in their native languages. A well-known example is Korean and Japanese speakers' difficulties with the English /r/-/l/ contrast (e.g. Goto 1971; Miyawaki, Jenkins, Strange, Liberman, Verbrugge, and Fujimura. 1975; MacKain, Best, and Strange 1981; Mochizuki 1981). According to the Perceptual Assimilation Model, these difficulties are due to Korean and Japanese speakers perceptually assimilating the two English sounds to the single liquid category of their native languages. Observing a great similarity between segmental adaptation patterns in loanwords and patterns of adult speakers' perceptual difficulties with non-native segmental contrasts, Peperkamp and her colleagues claim that segmental adaptations in loanwords result from the assimilation that takes place during the perception of non-native sounds. The English liquids case cited above exemplifies this correlation. In Korean and Japanese loanword adaptation, English /r/ and /l/ are mapped to the single liquid category of Korean and Japanese (E. *rail* → K. [ɾe.il], E. *line* → K. [ɾa.in]; E. *reef*

[4] While Peperkamp and her colleagues explicitly claim that unfaithful perception occurs whether the borrower is monolingual or bilingual, in the analyses of Boersma and Hamann (2009), Calabrese (2009), Broselow (2009), and Crawford (2007), it is implied that unfaithful perception only occurs in monolinguals' perception. The results of Takagi and Mann's (1994) perceptual experiments also show that unfaithful perception occurs in their bilingual participants' perception.

→ J. [ɾ̠iːɸu], E. *l̠etter* → J. [ɾ̠etaː]), reflecting the speakers' perceptual difficulties with this contrast.

The purely perceptual approach maintains that perceptual assimilation occurs not only at the segmental level but also at the prosodic and phonotactic levels.[5] Thus, perceptual assimilation is also held responsible for the adaptations found at these two levels. French exhibits a correlation between adaptations in loanwords and French adult speakers' perceptual difficulties with non-native prosodic contrasts. In French, stress is not contrastive and always falls on the last syllable of the word (e.g. Dupoux, Pallier, Sebastian, and Mehler 1997). Experimental studies show that French speakers have difficulty aurally discriminating multi-syllabic words that only differ in stress position (Dupoux, Pallier, Sebastian, and Mehler 1997; Dupoux, Peperkamp, and Sebastián-Gallés 2001; Dupoux, Sebastián-Gallés, Navarrete, and Peperkamp 2008). The phonological "deafness" of French speakers to stress position[6] mirrors the stress assignment patterns in loanwords. For example, in Spanish-based loanwords, the stress always falls on the last syllable, regardless of the location of the stress in the Spanish word.

Japanese provides an example of a correlation between adaptations in loanwords and Japanese adult speakers' perceptual difficulties with non-native phonotactic contrasts. Japanese does not allow consonant clusters in intervocalic position except for geminates and nasal+homorganic consonant sequences. Dupoux, Kakehi, Hirose, Pallier, and Mehler (1999) and Dupoux, Parlato, Frota, Hirose, and Peperkamp (2011) carried out perceptual experiments to test whether Japanese-speaking listeners would perceive an intervocalic obstruent+obstruent cluster as one that satisfies this native phonotactic requirement. In these experiments, Japanese participants listened to a continuum of stimuli in the form of /eb(u)zo/ where the inter-

[5] Best (1994) also assumes that perceptual assimilation can extend to the level of prosody.
[6] The results of the experiments conducted by Schwab and Llisterri (2011, 2014), however, do not support this hypothesis. Broselow (2009) also points out that this hypothesis is not consistent with Altmann's (2006) findings.

consonantal /u/ ranges from a full vowel (i.e. /ebuzo/) to nothing (i.e. /ebzo/) and indicated whether there was a vowel between the two obstruents. The results showed that the participants "heard" the vowel /u/ frequently, even if the stimuli contained no trace of that vowel at all.[7] In Japanese loanword adaptation, vowel epenthesis (usually /u/) is the predominant strategy for repairing illegal consonant clusters (Lovins 1975). Thus, the results of the perceptual experiments mirror the way in which illegal consonant clusters are adapted in foreign words borrowed into Japanese.

Although no individual cases are discussed, Peperkamp and her colleagues assume that segment deletion in loanword adaptation is also due to unfaithful perception. Peperkamp and Dupoux (2003: 367) state that 'the phenomenon of phonological "deafness"[,] that is, the inability or extreme difficulty to discriminate certain nonnative contrasts, involves segmental and suprasegmental contrasts, as well as contrasts based on the presence versus absence of a segment.'

2.1.2 Predictions

An important prediction of the purely perceptual approach is that phonologically identical inputs that have different phonetic realizations in the source language(s) may be mapped to different outputs. According to this approach, this is because the different acoustic signals associated with the inputs are mapped to different phonetic surface forms in the host language, and these phonetic surface forms are subsequently mapped to different phonological surface forms via perceptual assimilation.

Phonologically identical inputs that wind up as different outputs may come from a single source language. Use the Korean treatment of English word-final post-vocalic stops (Kang 2003)

[7] Other studies showing perceptual vowel epentheses include Matthews and Brown (2004), Kabak and Idsardi (2007); Yeon (2003), and Pitt (1998).

as an example.[8] Korean phonologists have noticed that, in English loanwords, word-final stops are more likely to be adapted with an epenthetic vowel if the preceding vowel is tense or a diphthong than if the preceding vowel is lax (e.g. *quick* → /kʰwik/ but *week* → /wikʰɨ/). Because a native Korean word can end in a stop (e.g. /kæk/ 'guest'), the observed vowel epenthesis seems unnecessary. Based on Parker and Walsh's (1981) investigation and her own survey of the TIMIT[9] corpus, Kang (2003) reported that, in English, word-final stops are more likely to release after a tense vowel or diphthong than after a lax vowel. Kang (2003) suggested that the seemingly unnecessary vowel epenthesis originate from the releases. Since stops in coda position are always unreleased in Korean, the adaptation patterns receive a natural account in the purely perceptual approach. While unreleased and released final stops are identically represented in English phonology, the acoustic signals associated with them are different. Upon perception of the acoustic signals, unreleased stops are mapped to unreleased stops, and released stops are perceptually assimilated to aspirated stops with a following epenthetic vowel.

Peperkamp, Vendelin, and Nakamura (2008) reported a case in which phonologically identical inputs from different source languages are mapped to different outputs. In Japanese loanword adaptation, while word-final /n/ in English source words is realized as a moraic nasal consonant, the same segment in the same position in French source words is realized with an epenthetic vowel (e.g. English loanwords: *pen* → /pen/, *walkman* → /wokuman/; French loanwords: /kan/ → /kannu/ 'Cannes', /paʁizjɛn/ → /parijennu/ 'parisienne'). Based on Tranel's (1987) studies and an examination of the stimuli used in their own perceptual experiments, Peperkamp, Vendelin, and Nakamura (2008) reported that word-final /n/ in French usually has a strong vocalic release. Their perceptual experiments tested whether Japanese-speaking listeners

[8] Takagi and Mann (1994) and Kim and Curtis (2002) report similar cases.
[9] TIMIT is a corpus of read speech of American English that is designed to provide data for acoustic-phonetic analyses.

would perceive word-final /n/ in French and English stimuli as /nu/, and the results showed that it only occurred with French stimuli. According to the purely perceptual approach, the "unnecessary" vowel epenthesis found in French-based loanwords ending in /n/ is attributed to the strong vocalic release of the /n/ being perceptually assimilated to a vowel during phonetic decoding.

The purely perceptual approach also predicts a possible asymmetry in the treatment of phonologically identical inputs when one of the source languages involved is the host language itself. That is, it predicts that different strategies may be used to repair the same structures in loanwords and in native words. Korean provides an example:[10] While a word-final obstruent with a release feature in a native input undergoes occlusivization (e.g. /nacʰ/ → [nat̲] 'face', /nas̲/ → [nat̲] 'sickle'), the same feature in the same position in an English input is salvaged by vowel epenthesis (e.g. *coa<u>ch</u>* → [kʰo<u>cʰɨ</u>], *bo<u>ss</u>* → [po<u>s'ɨ</u>]). An account in the purely perceptual approach is that the release feature in an English input is perceptually assimilated to a vowel, while a comparable perceptual assimilation process does not occur if the input comes from the native language (Peperkamp, Vendelin, and Nakamura 2008).[11]

2.2 The Purely Phonological Approach

2.2.1 Proposals

The purely phonological approach (Hyman 1970; Paradis and Lebel 1994; Paradis 1995a, b, 2006; Paradis and LaCharité 1997, 2001; LaCharité and Paradis 2002, 2005; Ulrich 1997; Rose 1999; Jacobs and Gussenhoven 2000; Shinohara 2004; Herd 2005; Uffmann 2006; Rose and Demuth 2006; Paradis and Tremblay 2009) is in sharp contrast to the purely perceptual approach. For example, by maintaining that bilinguals rather than monolinguals introduce and

[10] A list of languages showing this asymmetry is given in Peperkamp, Vendelin, and Nakamura (2008: 156).
[11] Kang (1996), Kenstowicz and Sohn (2001), and Kenstowicz and Suchato (2006) report similar cases.

adapt foreign words and that bilinguals are capable of accurately identifying the sounds and structures in the source languages like native speakers, Paradis and LaCharité (1997) suggest that foreign words enter the host languages in the form of the source language phonological representations.[12] Many researchers follow Paradis and her colleagues in assuming this phonological nature of input.[13] It follows from this stance that the borrower adapts the mental representation of the borrowed foreign word in the host language production grammar. During the adaptation, the borrower seeks a match that is phonologically closest to the mental representation. In the view of this approach, perceptual distortion does not occur.

Paradis and LaCharité (1997) analyze the adaptation of /v/ in French loanwords in Fula to demonstrate how a non-native sound is phonologically adapted. The analyses are couched in the Theory of Constraints and Repair Strategies (Pradis 1988a, b, 1990, 1993; Paradis and Prunet 1988; Paradis and LaCharité 1993, 1997). This theory is summarized in (5):

(5) Theory of Constraints and Repair Strategies (Paradis and LaCharité 1997)

i. Repair Strategy: A universal, non-contextual phonological operation that is triggered by the violation of a phonological constraint, and which inserts or deletes content or structure to ensure conformity to the violated constraint.

ii. Preservation Principle: Segmental information is maximally preserved within the limits of the Threshold Principle.

[12] Though Paradis and LaCharité (1997: 395) state that 'it has not yet been clearly established whether the input to L1 is the output of the L2 lexical or postlexical level', LaCharité and Paradis (2005: 224) claim that 'loanword adaptation is generally based on the L2- (not the L1-) referenced perception of L2 phoneme categories'. The latest view of this approach regarding the nature of the input is assumed in this dissertation.

[13] While Jacobs and Gussenhoven (2000) have the same suggestion regarding the nature of the input, they maintain that the phonological representation is created by the borrower using a universal phonological vocabulary, like a child is able to analyze any speech signal when acquiring his or her first language.

iii. Threshold Hypothesis/Principle

   a. All languages have a tolerance threshold to the amount of repair needed to enforce segment preservation.

   b. This threshold is the same for all languages: two steps (or two repairs) within a given a constraint domain.

iv. Minimality Principle

   a. A repair strategy must apply at the lowest phonological level to which the violated constraint refers.

   b. Repair must involve as few strategies (steps) as possible.

v. Phonological Level Hierarchy (PLH)

   metrical level > syllabic level > skeletal level > root node > feature with a dependent > feature without a dependent

vi. Precedence Convention: In a situation involving two or more violated constraints, priority is given to that constraint referring to the highest phonological level of the PLH.

French has /v/ but Fula does not. Paradis and LaCharité's (1997) data show that 76.5% of the /v/ instances are adapted as /w/, 17.3% as /b/, and 6.2% as /f/ (e.g. *avocat* /avɔka/ → [awɔka] 'lawyer', *avion* /avjɔ̃/ → [abijɔn] 'aeroplane', *élève* /elɛv/ → [ɛlɛf] 'student'). Assuming that underlying representations are radically underspecified, Paradis and LaCharité (1997) argue that while all of the three repair strategies satisfy the Threshold Hypothesis/Principle (5 iii), /v/ → /w/ is the most economical one and therefore is preferred (5 iv). Specifically, it is argued that /v/ → /w/ is a one-step task, only requiring insertion of [+sonorant], and that both /v/ → /b/ and /v/ → /f/ are a two-step task, requiring delinking of [+continuant] and insertion of [–continuant] for /v/

→ /b/ and delinking of [+voice] and insertion of [−voice] for /v/ → /f/. In addition, /v/ → /w/ fully respects the Preservation Principle while /v/ → /b/ and /v/ → /f/ do not (5 ii).[14]

Segment deletion occurs when a non-native sound is embedded in a structure that disobeys the native phonology. For example, /v/ is deleted if it is one of the consonants in a word-initial or -final CC cluster (e.g. *voyou* [vwaju] → waju *wuwaju 'bum') (Fula does not allow complex onset or coda). Adaptation by vowel epenthesis, on the other hand, is a three-step task – inserting a nucleus, providing the empty nucleus with segmental materials, and repairing the problematic segment – beyond the limit set by the Threshold Hypothesis/Principle (5 iii). Although deleting /v/ disrespects the Preservation Principle (5 ii), at least the Threshold Hypothesis/Principle (5 iii) is satisfied. In addition, while delinking the root node is not the best way of satisfying the segmental constraint, as much phonological information will be lost, it is the minimal way of satisfying the syllabic constraint, which, according to the Precedence Convention (5 vi) and the Phonological Hierarchy (5 v), must be taken care of first.

In summary, while there are multiple ways of repairing a problematic segment or structure, the strategy that is adopted by a language is highly predictable. The bottom line for a repair strategy to be at least considered adopting is that it must satisfy the Threshold Hypothesis/Principle. If more than one repair strategy satisfies the Threshold Hypothesis/Principle, the Preservation Principle and the Minimality Principle come into play and determine which strategy will be adopted (or adopted more often than the other strategies if the language is in the low community bilingualism period, as in the case of /v/ adaptation in French loanwords in Fula). Only when a problematic segment is embedded in a problematic syllable

[14] Paradis and LaCharité (1997) pointed out that Fula is still in the low community bilingualism period and that according to sociolinguistic studies, variation in loanword adaptation is common in languages that are in this period. While /v/ → /b/ and /v/ → /f/ violate the Preservation principle and are less economical than /v/ → /w/, they still satisfy the Threshold Hypothesis/Principle and thus are alternatives to the pervasive /v/ → /w/.

structure, is the Preservation Principle totally disregarded, as preserving the segment and repairing the syllable structure at the same time would exceed the limit imposed by the Threshold Hypothesis/Principle. When this is the case, according to the Precedence Convention and the Phonological Hierarchy, the language has no choice but deletes the problematic segment.

2.2.2 Predictions

The purely phonological approach makes some crucial predictions. First, if the source word contains a phoneme that also exists in the host language, this phoneme does not change its phonemic category even if the phonetic properties of this phoneme are more similar to those of a different phoneme in the host language (Category Preservation Principle, LaCharité and Paradis 2005: 226). LaCharité and Paradis (2002, 2005) observe several cases in their database and one of them is the adaptation of /r/ in English loanwords in Japanese. While /r/ is a phoneme in both English and Japanese, it is realized as a palato-alveolar central approximant ([ɹ]) in English (Best and Strange 1992) but as an alveolar tap ([ɾ]) in Japanese (Bloch 1950; Price 1981; Vance 1987). Several studies have indicated that English /r/ is in fact perceptually more similar to Japanese /w/ than to Japanese /r/ (e.g. Mochizuki 1981; Yamada and Tohkura 1991). In particular, Best and Strange (1992) tested Japanese-speaking listeners' identification and discrimination of English /w/ and /r/, and the results showed that Japanese monolinguals were more likely to confuse them than Japanese-English bilinguals, whose performance was very similar to that of American English controls. These results match LaCharité and Paradis' (2002, 2005) Japanese data, where English /r/ in onset position is always mapped to Japanese /r/ and never mapped to Japanese /w/. These results also support their assumption that bilinguals, rather than monolinguals, introduce and adapt foreign words.

The second prediction the purely phonological approach makes is that if the source word contains a phoneme that the host language does not have, this phoneme is substituted by the phonologically closest native phoneme even if the host language has another phoneme that is phonetically more similar to the one in the source word (Category Proximity Principle, LaCharité and Paradis 2005: 227). According to the Theory of Constraints and Repair Strategies, phonological closeness between two sounds is determined by the number of featural changes that one has to undergo to become the other. Thus, the English vowels /ɪ, ʊ/ are phonologically closer to /i, u/ than any other vowels in a language that does not have /ɪ, ʊ/. Based on Delattre's (1981) and Martin's (2002) acoustic measurements of vowels in several languages, LaCharité and Paradis (2005) report that English /ɪ, ʊ/ are acoustically more similar to /e, o/ than /i, u/ in Spanish and French. Their Mexican Spanish and French loanword data, however, show that English /ɪ, ʊ/ are adapted as /i, u/ 98% of the time.

The purely phonological approach also predicts that phonetic variants of a phoneme in the source or host language are irrelevant in the adaptation process. For example, while /æ/ may be pronounced as [ɛ] when followed by a nasal consonant in English (e.g. *bank* /bæ̲ŋk/ is pronounced [bɛ̲ŋk]), LaCharité and Paradis' (2005) loanword data show that /æ/ in this context is almost always accurately identified in the adaptation of English words borrowed into Quebec French, which has /ɛ/ in its phonemic inventory. Another example that LaCharité and Paradis (2005) provide concerns a phonetic variant of a phoneme in the host language. In some dialects of Quebec French, /ʃ, ʒ/ may be phonetically realized as [h] (e.g. /ʃ̲ãʒe/ → [h̲ãhe] 'to change'). In the adaptation of English words into Quebec French, English /h/ is consistently viewed as problematic and undergoes deletion (e.g. *hacker* /hækəɹ/ → [_akəɹ] *[hakəɹ] *[ʃ/ʒakəɹ]).[15]

The last prediction the purely phonological approach makes that is relevant to the analyses in this dissertation concerns segment deletion. Given that every segment in the source word is accurately identified (and hence must be perceived in the first place), the purely phonological approach predicts that segment deletion due to unfaithful perception does not occur or are rare. LaCharité and Paradis' (2005) loanword data show that segment deletion is very rare in general, accounting for only 2.6% of the cases that are problematic for the host language and require a repair. In addition, most of the deletions in their data can be explained phonologically, bringing the percentage of phonetically-driven deletion to a even lower level.

2.3 The Hybrid Approach

The hybrid approach acknowledges the roles of both speech perception and the host language production grammar in loanword adaptation (Silverman 1992; Yip 1993, 2002, 2006; Kang 2003; Steriade 2001, 2009; Kenstowicz 2003, 2005, 2006, 2007, 2012; Ito, Kang, and Kenstowicz 2006; Shinohara 2006; Adler 2006; Davis and Cho 2006; Kubozono 2006; Rose and Demuth 2006; Vendelin and Peperkamp 2006; Hsieh, Kenstowicz and Mou 2009; Smith 2006, 2009; Dohlus 2005; among many others). In particular, this approach holds that while the loanword output is the artifact of the host language production grammar, unfaithful perception may occur and/or the host language production grammar takes into consideration the phonetic details contained in the source words. Most works taking this hybrid view see loanword adaptation as consisting of two stages, a perception stage followed by a production stage. Furthermore, some proponents of this approach demonstrate that factors other than speech

[15] Nevins and Braun (2009) report a counterexample concerning Brazilian Portuguese speakers of English. In Brazilian Portuguese, underlying /r/ surfaces as [h] in word-initial position. In these speakers' English, word-initial [h] is replaced by [r], so the English word *home* is pronounced [rom].

perception also help shape the output. These proposed factors include orthography (Smith 2006, 2009; Dong 2012; Vendelin and Peperkamp 2006), the borrower's explicit knowledge of the syntax and morphology of the source language (Silverman 1992; Smith 2006, 2009), the interlanguage or the local version of the source language (e.g. Hong Kong English) (Yip 2006; Kenstowicz 2005; Boersma and Hamann 2009; Dong 2012), and visual cues (Yip 2002, 2006; Kenstowicz 2006).

2.3.1 Proposals

In this subsection, I review the proposals of some of the proponents of this approach.

2.3.1.1 Silverman (1992), Yip (1993)

Silverman (1992) explicitly suggests a two-stage model of loanword adaptation. He analyzes English loanwords in Cantonese and proposes that the input is a nonlinguistic acoustic signal and has to go through a perceptual level and then a phonological level for adaptation. He makes several claims with respect to the perceptual level. First, the incoming acoustic signal is parsed into segment-sized chunks, which are replaced by the native segments that have the most similar articulatory/acoustic properties. As a result, some contrasts that are employed in the source language but not in the host language are neutralized and not represented at this level. For example, Cantonese lacks a voicing contrast in obstruents, so both /z/ and /s/ in borrowed English words are represented as Cantonese [s] at this level. Second, if a segment in the source word is imperceptible to the borrower due to its low degree of saliency, this segment is not represented at this level either. For example, word-final stops that follow a nasal consonant are not represented at this level (e.g. *band* → [pɛːn_] due to unfaithful perception). Third, the most salient components of the acoustic signal such as vowels and non-prevocalic sibilants are assigned a syllable node.[16] Fourth, a binary foot template is imposed at this level, resulting in disyllabic output forms. This segmental sequence with a preliminary prosodic structure enters the phonological level, at which native phonotactic and prosodic constraints hold and the raw linguistic material undergoes phonological processes. Silverman observes that the phonological processes found at this level are absent from the native Cantonese phonology. He argues that these processes are loanword-specific and exist in a system that is separate from the native phonology. Also analyzing English loanwords in Cantonese, Yip (1993) follows Silverman in assuming that Cantonese loanword adaptation consists of the two levels. In contrast to Silverman's rule-based analyses, she adopts a constraint-based framework and argues that the processes taking place at the phonological level do not constitute a separate system but are motivated in the native Cantonese phonology.

The adaptation of the English words *print* and *tips* into Cantonese exemplifies Silverman's two-stage model.

(6)	Acoustic Signal		Perceptual Level		Phonological Level		Output
	print	→	p^{h}lin	→	p^{h}i.lin	→	p^{h}i.lin
	tips	→	t^{h}ips	→	t^{h}ip.si	→	t^{h}ip.si

While Cantonese has [t], the /t/ in *print* appears word-finally and follows a nasal, a position rendering it difficult for the borrower to detect, and therefore is not represented at the perceptual level. Liquids are always perceived and represented at the perceptual level as Cantonese has [l], but they are usually found deleted at the phonological level if they appear in onset clusters (e.g. *printer* → [p^{h}ɛ.t^{h}a], *broker* → [puk.k^{h}a], *freezer* → [fi.sa]). The liquid in *print*, however, is preserved at the phonological level. This asymmetry is due to a binary foot template that has

[16] Non-prevocalic sibilants are always preserved by vowel epenthesis (e.g. *waste* → [wɐj.si], *cast* → [k^{h}a.si]), in contrast to non-sibilants in this context (e.g. *shaft* → [sap], *lift* → [lip]).

been imposed at the perceptual level. That is, preserving the liquid in *print* with vowel epenthesis yields a disyllabic output form (*print* → [pʰi.lin] *[pʰin]), and deleting the liquid in a word such as *printer*, *broker*, or *freezer* also yields a disyllabic output form. In the adaptation of *tips*, [s] is given a syllable node at the perceptual level as it is a sibilant and does not precede a vowel. [i] is epenthesized at the phonological level so that [s] can be syllabified.

2.3.1.2 Yip (2002. 2006), Kenstowicz (2005)

Still working on English loanwords in Cantonese, Yip (2002, 2006) considers the role played by the interlanguage or the local version of English (e.g. Hong Kong English) in determining the phonological representation that serves as the input to the Cantonese production grammar. For example, /z/ is always adapted as [s] in English loanwords (e.g. *raze* → [lej.siː]), and since /z/ is devoiced to [s] in the pronunciations of Cantonese learners of English and also in Hong Kong English (e.g. *zeal* and *raze* are pronounced [sil] and [ɹejs], respectively), Yip (2002, 2005) argues that the interlanguage of Cantonese learners or Hong Kong English is probably the source. Furthermore, Yip notes that at least some deletions are attributed to unfaithful perception. One example is the adaptation of *cast* as [kʰaː.siː], where the word-final /t/ in the source word is not seen in the output. Yip argues that the absence of [t] in the output cannot be attributed to a process that takes place in the production grammar, as [kʰaː.siː] always violates one more constraint (i.e. MAX-C) than [kʰaː.siːt] and hence is not preferred by any OT grammar.[17]

Kenstowicz (2005) also argues for the role of the interlanguage in deciding the input to the production grammar. He notes that while English does not have a phonemic aspirated-unaspirated contrast in voiceless stops, in English loanwords in Korean, word-initial voiceless

[17] This implies that *cast* is pronounced without the final /t/ in the interlanguage and/or Hong Kong English.

stops are consistently mapped to the voiceless aspirated series[18] (e.g. *poker* → [pʰoːkʰa], *talent* → [tʰalletʰɨ], *cola* → [kʰolla]). Moreover, in adapting French words, Korean maps voiceless unaspirated stops to the voiceless unaspirated tense series (e.g. *Paris* → [p*aɾi], *Toulouse* → [t*ullucɨ], *Cannes* → [k*anɨ]). These patterns demonstrate that the borrowers attend to the phonetic realizations of the source words. However, the unaspirated realizations of English voiceless stops in post-tonic position are equated by Korean borrowers with the voiceless aspirated series, rather than the phonetically closest voiceless unaspirated tense series (e.g. *happy* → [hɛpʰi] *[hɛp*i], *monitor* → [monitʰa] *[monit*a], *chicken* → [cʰikʰin] *[cʰik*in]). To provide an account, Kenstowicz suggests that Korean speakers may not be sensitive to stress and consequently processes associated with stress in English are absent from their interlanguage. Since a post-tonic position is also an onset position, Korean speakers simply use the native aspirated series in their interlanguage, and this is reflected in the adaptation of English words.

Both Yip (2002, 2006) and Kenstowicz (2005) give an account of the "divergent repairs" phenomena. Yip's (2002, 2006) analyses are provided here (see 6.2 for Kenstowicz's analyses). Yip (2002, 2006) suggests that the production grammar of the host language includes a set of loanword-specific faithfulness constraints which she calls MIMIC constraints. Like other constraints, MIMIC constraints can be freely ranked. The divergent repairs of word-final obstruents with a release feature in Korean exemplify how MIMIC constraints work. The markedness constraint that triggers repairing is CODACON (= coda condition) (Kenstowicz 2005). While occlusivization is used for native words (e.g. /nacʰ/ → [nat] 'face'), loanwords violating this constraint undergo vowel epenthesis (e.g. E. *coach* → [kʰocʰi] *[kʰot]). In the derivation of native words, DEP-V (= do not insert a vowel) dominates IDENT-[release] (= be faithful to a

[18] Korean has a three-way phonemic contrast in stops: voiceless unaspirated lax, voiceless unaspirated tense (marked with "*"), and voiceless aspirated.

release feature) (i.e. not inserting a vowel is more important than retaining a release feature in a native input), but in loanword adaptation, MIMIC-[release] (the loanword version of IDENT-[release]) is activated and outranks DEP-V (i.e. mimicking a release feature in a loanword input is more important than not inserting a vowel). The ranking MIMIC-[release] >> DEP-V >> IDENT-[release] accounts for the conflicting repair patterns.

2.3.1.3 Steriade (2001, 2009)

Steriade (2001, 2009) notes that to meet a phonotactic requirement, different languages may converge to a single strategy that has proven to cause a minimal perceptual change. For example, languages that disallow word-final voiced obstruents always employ the devoicing strategy, and evidence has shown that devoicing a word-final obstruent is perceptually less discernable than the same obstruent in the same position being nasalized, lenited, deleted, etc. Because constraints are freely ranked, OT predicts the existence of languages that employ strategies other than devoicing. To solve this “too-many-solutions” problem, Steriade (2001, 2009) proposes that speakers share the knowledge of relative perceptibility of different contrasts in various contexts, which she calls the P-map, and that this knowledge guides the speakers to choose the same strategy. The P-map can be described as perceptual similarity hierarchies such as those in (8). (“Δ” denotes differences.)

(7) Perceptual Similarity Hierarchies

i. Δ (x-z)/__α > Δ (y-z)/__α

In the context α, the contrast between x and z is perceptually more salient than the contrast between y and z.

ii. Δ (y-z)/__α > Δ (y-z)/__β

The contrast between y and z is perceptually more salient in the context α than in the context β.

These two perceptual similarity hierarchies can be combined into one:

(8) Δ (x-z)/__α > Δ (y-z)/__α > Δ (y-z)/__β

A perceptual similarity hierarchy projects faithfulness constraints that are ranked in the same order. If x and y stand for segments and z stands for silence, then the combined perceptual similarity hierarchy in (8) projects the ranked faithfulness constraints in (9):

(9) MAX-x/__α >> MAX-y/__α >> MAX-y/__β

"Do not delete x in the context α" is more important than "do not delete y in the same context α", which in turn is more important than "do not delete y in the context β".

Although the P-map is originally proposed for non-loan phonology, Steriade (2001) extends it to loanword adaptation. As an example, she discusses the Cantonese adaptation of the English words *tips* and *post* as [t^{h}ip.si] and [p^{h}ow.si], respectively. Cantonese does not allow consonant clusters, so the word-final /s/ and /t/ in the English words are problematic. As we can see in the Cantonese forms, the /s/ in *tips* is retained but the /t/ in *post* is deleted. She claims that this asymmetry is attributed to the relative perceptual similarity of the two contrasts /s-∅/ and /t-∅/ in that position. Specifically, due to /s/'s high-frequency noise, deleting the /s/ would cause a larger perceptual change than turning it into [si], so vowel epenthesis is the preferred strategy to meet the no-complex-coda requirement. Deleting the /t/, on the other hand, causes a smaller perceptual change than the vowel-epenthesis repair, as the /t/ lacks internal and contextual cues (Steriade 1997). As a result, deleting the /t/ is the preferred strategy. A post-consonantal non-prevocalic stop is not always deleted. For example, the /t/ in *st̲rawberry* is retained (*st̲rawberry*

→ [si.<u>t</u>^{h}aw.pe.lej]). According to the P-map, a post-consonantal pre-sonorant stop is preserved because it is perceptually too salient to delete. Specifically, given that the contextual cues of a pre-sonorant stop are present in the following sonorant, deleting it would cause a larger change in perception than preserving it via epenthesis.

The constraints in (10) and the tableau in (11) show how the Cantonese adaptation of the English words *tips*, *post*, and *strawberry* is formalized in P-map.

(10) Constraints

i. MAX-strident: Do not delete a strident.

ii. MAX-stop/$_\sigma$[C$_0$__R: Do not delete a stop if it appears before a tautosyllabic sonorant. (C$_0$ denotes a consonant that may or may not be present, and R denotes a liquid.)

iii. DEP-[i]: Do not insert [i].

iv. MAX-stop/C__]$_\sigma$: Do not delete a stop if it is syllable-final and follows a consonant.

(11) E. *tips* → Cant. [t^{h}ip.si], E. *post* → Cant. [p^{h}ow.si], E. *strawberry* → [si.t^{h}aw.pe.lej]

tips	MAX-strident	MAX-stop/$_\sigma$[C$_0$__R	DEP-[i]	MAX-stop/C__]$_\sigma$
→ a. t^{h}ip.si			*	
b. t^{h}ip	*!			
post	MAX-strident	MAX-stop/$_\sigma$[C$_0$__R	DEP-[i]	MAX-stop/C__]$_\sigma$
→ a. p^{h}ow.si			*	*
b. p^{h}ow.si.ti			**!	
strawberry	MAX-strident	MAX-stop/$_\sigma$[C$_0$__R	DEP-[i]	MAX-stop/C__]$_\sigma$
→ a. si.t^{h}aw.pe.lej			*	
b. saw.pe.lej		*!		
c. t^{h}aw.pe.lej	*!			

MAX-strident is undominated because a sibilant is always perceptually salient, regardless of its position in the word (*tip<u>s</u>* → [t^{h}ip.<u>s</u>i] *[t^{h}ip]; *<u>s</u>trawberry* → [<u>s</u>i.t^{h}aw.pe.lej] *[t^{h}aw.pe.lej]). Also undominated is MAX-stop/$_\sigma$[C$_0$__R, as a stop appearing before a tautosyllabic sonorant is also perceptually salient (*s<u>t</u>rawberry* → [si.<u>t</u>^{h}aw.pe.lej] *[saw.pe.lej]). DEP-[i] is ranked below MAX-strident and MAX-stop/$_\sigma$[C$_0$__R because [i] is epenthesized to preserve the /s/ in *tips* and both the /s/ and /t/ in *strawberry*. MAX-stop/C__]$_\sigma$ is ranked below DEP-[i] because a syllable-final stop in a cluster is perceptually non-salient and deleting it would not cause too much change in perception (*post* → [pʰow.si] *[pʰow.si.ti]).

2.3.2 Predictions

Like the purely perceptual approach, the hybrid approach predicts that phonologically identical inputs that have different phonetic realizations may be mapped to different outputs.

First, in the view of Silverman (1992), Yip (1993), Kenstowicz (2003), Broselow (2009), and Boersma and Hamann (2009), the input to loanword adaptation is an acoustic signal and has to go through a perceptual level at which the native phonology dictates the perception. If the inputs are phonologically identical but have different phonetic realizations, then at this level those with very weak phonetic cues are not perceived and hence unrepresented (e.g. E. *cast* → C. [kʰaː.siː_]; Yip 2006) and those that are perceived may be replaced by different native segments (e.g. E. *quick˺* → K. /kʰwik˺/ vs. E. *week˹* → K. /wikʰɨ/; Kang 2003).

Second, the models proposed by Yip (2002, 2006), Kenstowicz (2005, 2006, 2007), and Steriade (2001, 2009) assume that the output of the perceptual level is a non-native percept, indicating that the perception is basically faithful.[19] Thus, most of the cases showing the "divergent repairs" phenomenon are not due to unfaithful perception but result from applications of phonological processes in the host language production grammar, which these models have argued can access the phonetic information contained the source words. An example is the deletion of the /t/ in *post* but preservation of the /t/ in *strawberry* (*post* → [pʰow.si], *strawberry* → [si.tʰaw.pe.lej]) analyzed by Steriade (2001).

[19] Yip (2002, 2006) and Kenstowicz (2005, 2006, 2007) do not reject the possibility of occasional unfaithful perception. Since the P-map (Steriade 2001, 2009) is originally proposed for non-loan phonology, I assume that in this model perception is faithful.

Because adaptation is argued to take place in the production grammar other than sound substitution, all these models predict that repairing a word structure is totally phonological. For example, Silverman (1992) argues that the /r/ in *printer* is deleted but the /r/ in *print* is preserved via vowel epenthesis to meet the Cantonese disyllabic word requirement (*pr̲inter* → [pʰɛ.tʰa], *pr̲int* → [pʰi.l̲in]), which is held at the phonological level. The purely perceptual approach obviously cannot make the correct prediction because both /r/'s appear in identical contexts and therefore have identical phonetic realizations.

What is interesting about this case is the question of why it is the liquid, rather than the initial /p/, that gets deleted, as deleting the /p/ in *printer* also results in a disyllabic output form (*pr̲inter* → [pʰɛ.tʰa] *[l̲in.tʰa]) (Yip 1993). On the basis of Fay and Cutler's (1977) speech error data,[20] Yip (1993) argues that the liquid is perceptually less salient than the initial /p/ and hence is the target of deletion. This case demonstrates that while keeping every output form disyllabic is phonological, when there is a need to delete a sound, which sound to delete, the initial /p/ or the liquid, is related to perceptual saliency and the information is available in the grammar's perception component.

2.3.3 Summary of the Hybrid Approach

In the view of the hybrid approach, the input to loanword adaptation is an acoustic signal and therefore has to go through a perceptual level before the production grammar takes over.

At the perceptual level, non-native sounds may be perceived as the native segments that have the most similar auditory/articulatory properties, and some signals may not be detected, giving rise to deletion seen in the output. As a result, the output of the perceptual level consists

[20] Fay and Cutler's (1977) data show that when one of the segments in an onset C+liquid sequence is overlooked, speakers are more likely to ignore the liquid than the preceding consonant.

of native segments or a non-native percept with the possibility of unfaith perception taking place in some weak positions.

The host language production grammar has access to the perceptual information contained the source words and uses it to project and rank faithfulness constraints so that the output can be as perceptually similar to the source words as possible. A few models discuss how the production grammar uses the perceptual information. Yip (2002, 2006) and Kenstowicz (2005) propose that the production grammar uses the perceptual information through a set of faithfulness constraints that only apply to loanwords. Steriade (2001, 2009) proposes that the P-map, a distinct grammatical component in which speakers' shared knowledge of relative perceptibility of contrasts in various contexts is stored, projects faithfulness constraints and their rankings. In these models, the loanword output is determined on the basis of the way in which the perception-oriented faithfulness constraints interact with the relevant markedness constraints.

Like the purely perceptual approach, the hybrid approach predicts that languages may use different strategies to repair inputs that are phonologically equivalent but realized differently in the source language(s). Like the purely phonological approach, this approach also predicts that repairing can be triggered by a native phonological requirement.

CHAPTER THREE

ARGUMENTS AGAINST THE PURELY PERCEPTUAL APPROACH

In this chapter, I present evidence that loanword adaptation occurs in the host language production grammar rather than during perception. In 3.1, I show that while word-final stops and coda liquids may be deleted in English loanwords in SC (e.g. word-final stop: *Flint* → [f^{w}o.lin.t^{h}ɤ] vs. *Clint* → [k^{h}ɤ.lin__]; coda liquid: *Arkin* → [a.ər.tɕin] vs. *Starr* → [ʂz̩.ta__]), these sounds are always preserved with the place of articulation correctly identified when the source words are monosyllabic and contain no other sounds violating the SC syllable structure requirements (e.g. word-final stop: *Ed* → [ai.tɤ], *Pink* → [piŋ.k^{h}ɤ]; coda liquid: *Barr* → [pa.ər], *Hall* → [x^{w}o.ər]). These data indicate (i) that the deletion is not due to unfaithful perception, and (ii) that the borrowers know exactly what sounds are present in the source words even if some of the sounds may be perceptually non-salient (e.g. word-final stops following a nasal consonant). Evidence against the purely perceptual approach is also found in languages other than SC. In 3.2, I step away from SC and consider the adaptation of word-initial /r/ and /l/ as [ɾ] in English loanwords in Japanese and Korean. Phonetic data on Japanese speakers' perception and implications of Korean loanword data clearly show that Japanese and Korean speakers do not perceive the two English sounds in that position as [ɾ].

3.1 The SC Minimum-Word Requirement

Before I present the SC data, consider again the adaptation of word-initial obstruent+liquid sequences in English loanwords in Cantonese. As Silverman (1992) has shown, the liquids are either deleted or preserved (e.g. *freezer* → [fi.sa], *break* → [pik.lik]). Because the liquids appear in similar contexts, they are expected to have similar phonetic realizations. In addition, some of the English source words are derived from the same words (e.g. *printer* →

[pʰɛn.tʰa] vs. *print* → [pʰi.lin]), so it is implausible to attribute the deletion to unfaithful perception. These data demonstrate that the adaptation takes place in the Cantonese production grammar. Silverman (1992) argued that whether the liquids are deleted or preserved is determined by the constraint that requires every output form to be disyllabic. For example, the first /r/ in *printer* is deleted (*printer* → [pʰɛn.tʰa] *[pʰu.lɛn.tʰa]) but the /r/ in *print* is preserved (*print* → [pʰi.lin] *[pʰin]) because the results are disyllabic output forms. Yip (1993, 2002, 2006) has a similar argument.

Comparable phenomenon is found in English loanwords in SC. Segment deletion is rare in my corpus, and most of the deletion cases are deletion of word-final stops and coda liquids.[21] However, if the source words are monosyllabic and contain no other sounds violating the SC syllable structure requirements, then the word-final non-labial stops and coda liquids are always preserved. The data are given below.

Table 3.1 Frequencies of deletion and preservation of word-final non-labial stops and coda liquids in two types of English words borrowed into SC (stops: $\chi^2 = 27.85$, $p < .00001$; liquids: $\chi^2 = 83.84$, $p < .00001$)

		Type-A source words		Type-B source words	
		monosyllabic words where a word-final non-labial stop or liquid is the only sound that violates the SC syllable structure requirements		words with a word-final non-labial stop or coda liquid that do not meet the conditions of Type-A words	
word-final non-labial stop	del.	0	0%	72	39.78%
	pres.	48	100%	109	60.22%
	sum	48	100%	181	100%
coda liquid	del.	0	0%	483	76.18%
	pres.	30	100%	151	23.83%
	sum	30	100%	634	100%

[21] Labial stops are almost always preserved. In my corpus, the only mapping where a labial stop is deleted is *Hepburn* → [xɤ__pən] *[xɤ.pʰu.pən].

(12) Examples of deletion and preservation of word-final non-labial stops and coda liquids in two types of English words borrowed into SC

Preservation			Deletion		
i. Word-final stops in Type-A source words					
Pitt	→	[pi.tʰɤ]			
Hunt	→	[xɑŋ.tʰɤ]			
Matt	→	[mai.tʰɤ]			
Dick	→	[ti.kʰɤ]			
Pink	→	[piŋ.kʰɤ]			
Jack	→	[tɕe.kʰɤ]			
Reid	→	[lei.tɤ]			
Ed	→	[ai.tɤ]			
Rand	→	[lan.tɤ]			
Doug	→	[tɑu.kɤ]			
ii. Word-final stops in Type-B source words					
Sweet	→	[ʂʐ̩.wei.tʰɤ]	*Janet*	→	[ʈʂən.ni__]
Flint	→	[fʷo.lin.tʰɤ]	*Clint*	→	[kʰɤ.lin__]
Priest	→	[pʰu.li.sz̩.tʰɤ]	*Frist*	→	[fu.li.sz̩__]
Eric	→	[ai.ʐʷei.kʰɤ]	*Kodak*	→	[kʰɤ.ta]
Blake	→	[pu.lai.kʰɤ]	*Murdock*	→	[mei.tʷo__]
Hitchcock	→	[ɕi.tɕʰy.kʰɑu.kʰɤ]	*Swank*	→	[ʂʐ̩.wɑŋ__]
Floyd	→	[fʷo.lʷo.i.tɤ]	*Madrid*	→	[ma.tɤ.li__]
Fred	→	[fʷo.ʐʷei.tɤ]	*Hempstead*	→	[xan.pu.sz̩.ti__]

Thurmond	→	[sai.məŋ.tɤ]	*Raymond*	→	[lei.məŋ__]
Lundberg	→	[luŋ.pʷo.kɤ]	*Spielberg*	→	[ʂʐ̩.pʰi.pʷo__]
Kellogg	→	[kʰɤ.lʷo.kɤ]	*Kellogg*	→	[tɕa.lɤ__]
Greg	→	[kɤ.ʐʷei.kɤ]			

iii. Coda liquids in Type-A source words

Deere	→	[ti.ər]
Gere	→	[tɕi.ər]
Barr	→	[pa.ər]
Gore	→	[kɑu.ər]
Moore	→	[mʷo.ər]
Will	→	[wei.ər]
Bell	→	[pei.ər]
Hal	→	[xa.ər]
Bull	→	[pu.ər]
Hall	→	[xʷo.ər]

iv. Coda liquids in Type-B source words

Sears	→	[xi.ər.sz̩]	*Goodyear*	→	[ku.tʰɤ.i__]
Blair	→	[pu.lai.ər]	*Mayer*	→	[mei.je__]
Coors	→	[kʰu.ər.sz̩]	*Kurland*	→	[kʰu__lan]
Kirk	→	[kʰɤ.ər.kʰɤ]	*Burton*	→	[pʷo__tʷən]
Horton	→	[xʷo.ər.tʰəŋ]	*Horton*	→	[xɤ__tʷən]
Arkin	→	[a.ər.tɕin]	*Starr*	→	[ʂʐ̩.ta__]
O'Neil	→	[ou.ni.ər]	*Spielberg*	→	[ʂʐ̩.pʰi__pʷo]

Hilton	→	[xi.ər.tʷən]	*Wilkes*	→	[wei__kʰɤ.sz̩]
Cornell	→	[kʰɑŋ.nai.ər]	*Feldman*	→	[fei__tɤ.man]
Malcolm	→	[mai.ər.kʰən]	*Paltrow*	→	[pʰai__tɤ.lʷo]
Fulton	→	[fu.ər.tʷən]	*O'Toole*	→	[ɑu.tʰu__]
Bowles	→	[pɑu.ər.sz̩]	*Nicole*	→	[ni.kʰɤ__]
Golf	→	[kɑu.ər.fu]	*Hallmark*	→	[xɤ__ma__kʰɤ]
Mitchell	→	[mi.tɕʰi.ər]	*Adolf*	→	[a.tɑu__fu]

As we can see in Table 3.1 and (12), although the word-final non-labial stops and coda liquids in Type-B source words may be deleted, those in Type-A source words are always preserved via vowel epenthesis. Because the stops or liquids occur in identical contexts sometimes (e.g. *Deere* → [ti.ər] vs. *Goodyear* → [ku.tʰɤ.i__]; *Barr* → [pa.ər] vs. *Starr* → [ʂʐ̩.ta__]), the deletion should not be attributed to unfaithful perception and must take place in the SC production grammar. This argument is strengthened by another set of data in my corpus, which shows that CV and CVN source words are always augmented to CV.V and CV.ən/CVN.ən, respectively (e.g. *Bee* → [pi.i], *Lynn* → [lin.ən], *Kahn* → [kʰa.ən]/[kʰɑŋ.ən]). If word-final non-labial stops and coda liquids in the source words are not always perceived, and if unfaithful perception is the reason for the augmentation of CV and CVN source words, then we expect the Type-A source words to surface in the form of CV.V or CV.ən/CVN.ən at least occasionally (e.g. *Hunt* → *[xɑŋ.ən]). However, no such mapping is found in my corpus. In addition, when the word-final stops are preserved, the place of articulation is always maintained. These facts strongly suggest that the SC borrowers know what sounds are in the source words, indicating that the adaptation must take place in the SC production grammar.

Notice that when the word-final non-labial stops or coda liquids are deleted, the output forms are never monosyllabic, and that when they are preserved, the output forms always consist of at least two syllables. This fact is very similar to what Silverman (1992) has observed in English loanwords in Cantonese, but the constraint that calls for the preservation of the stop or liquid is obviously not the one that requires every output form to be disyllabic. Following Yip (1993), I propose that the constraint in effect here requires every output form to consist of at least two syllables. The purely perceptual approach falls short of accounting for these sets of data: If unfaithful perception were the cause of the deletion, we would see the stops or liquids in the Type-A source words delete at least occasionally.

3.2 Perceptual Assimilation does not Equal Unfaithful Perception

In the view of the purely perceptual approach, all adaptations found in loanwords are due to unfaithful perception. In 3.1, I have demonstrated that segment deletion in English loanwords in SC is not due to unfaithful perception. In what follows, I will step away from SC and show that featural changes in the adaptation of English words into other languages should not be attributed to unfaithful perception either. Specifically, I am considering the adaptation of word-initial /r/ and /l/ as [ɾ] in English loanwords in Japanese and Korean.

The view that unfaithful perception is the cause of loanword adaptation is developed on the basis of the Perceptual Assimilation Model (e.g. Best 1995), which holds that during perception non-native sounds may be assimilated to native categories that have the closest phonetic properties. It follows that if two sounds that form a non-native contrast are assimilated to the same native category, then the two sounds may not be easily discriminated. Neutralization of segmental contrasts is commonly observed in loanword adaptation, and the purely perceptual approach attributes it to perceptual assimilation. An example comes from Korean and Japanese.

Korean and Japanese speakers have great difficulty discriminating the contrast of /r/ and /l/ in word-initial position in English, and the loanword data show that all word-initial liquids are mapped to [ɾ] in English loanwords in Korean and Japanese. Thus, the purely perceptual approach suggest that the mappings result from the two English liquids in the position being perceptually assimilated to the single liquid categories of Korean and Japanese (Peperkamp and Dupoux 2003: 367; Peperkamp 2005: 346–347).

For the sake of clarity, the diagram in (13) shows the relation of the purely perceptual approach to the Perceptual Assimilation Model:

(13) The relation of the purely perceptual approach to the Perceptual Assimilation Model

Perceptual Assimilation Model		The purely perceptual approach
difficulty in contrast discrimination		segmental contrast neutralization in loanwords
↑		↑
perceptual assimilation	=	unfaithful perception

In this section, I demonstrate that perceptual assimilation should not be equated to unfaithful perception. Specifically, I will show that while Korean and Japanese speakers perceptually assimilate word-initial /r/ and /l/ in English to the single liquid category of their native language and therefore have difficulty discriminating the contrast, they do not perceive the two sounds in the position as [ɾ].

3.2.1 The Perceptual Assimilation Model Revisited

The Perceptual Assimilation Model distinguishes several types of perceptual assimilation. Two of them are relevant to the discussion here.

(14) Single-Category Assimilation and Category-Goodness Difference Assimilation in the Perceptual Assimilation Model (Best 1995: 195)

i. Single-Category Assimilation (SC)

Both sounds in a non-native contrast are assimilated to a native category, and they are equally distant from the norm of the native category. Discrimination of the two sounds is poor.

ii. Category-Goodness Difference Assimilation (CG)

Both sounds in a non-native contrast are assimilated to a native category, but one of the sounds is closer than the other to the norm of the native category. Discrimination of the two sounds is moderate to very good, depending on the difference in category goodness between the two sounds. If there is a big difference in category goodness, the contrast is easy to discriminate; if the difference is small, the contrast is difficult to discriminate.

In the SC type, both sounds are good or bad exemplars of a native category. In the CG type, one sound is a bad exemplar of a native category, but the other sound is a good exemplar of the same native category. Because Korean and Japanese speakers have shown poor performance in discriminating word-initial /r/ and /l/ in English, the purely perceptual approach must assume that the perception is of the SC type. This means that, for Korean and Japanese speakers, word-initial /r/ and /l/ in English are either good or bad exemplars of the single liquid category of their native language. Speakers are very sensitive to accents, and bad exemplars of a native category can be thought of as strong foreign accents.[22] It follows that while foreign sounds that are considered bad exemplars of a native category may be perceptually assimilated to the category, they should be immediately distinguished from the norm. Thus, we do not expect bad exemplars of a native

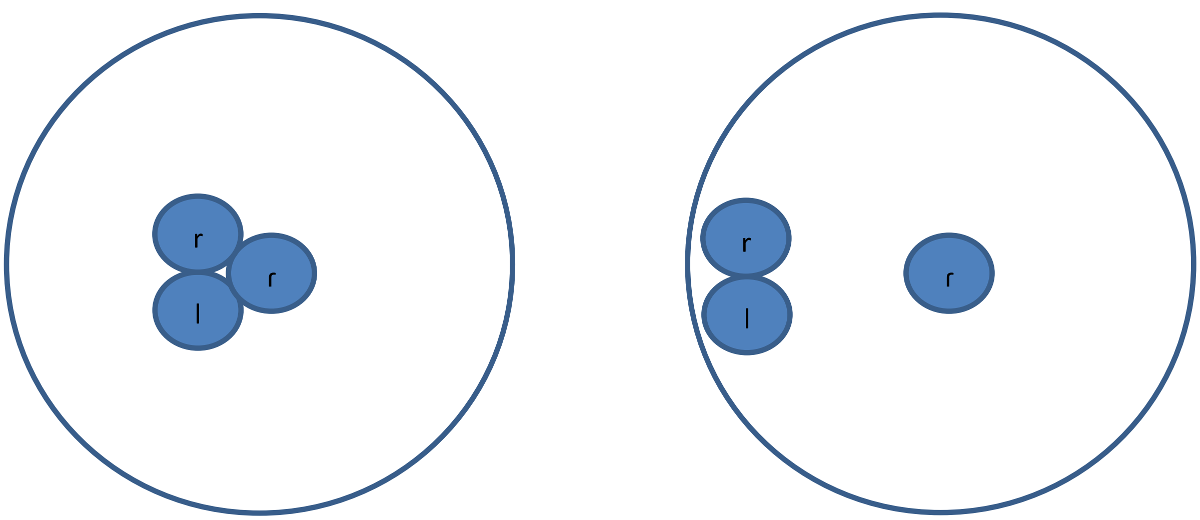
[22] For instance, LaCharité and Paradis (2005: 242) noted that 'native speakers of French are readily able to identify the region of origin of other speakers, on the basis of particularities of their accents.' See also the references cited there.

category to be perceived as the norm.[23] This argument is strengthened by the CG type of assimilation, in which two exemplars differing in category-goodness comprise a non-native contrast. Since contrast discrimination is moderate to very good, the differences between the two exemplars must be detectable, indicating that at least the worse exemplar is not perceived as the norm of the native category to which it is assimilated.

The purely perceptual approach maintains that unfaithful perception is responsible for the mapping of word-initial /r/ and /l/ to [ɾ] in English loanwords in Korean and Japanese. In order for the unfaithful perception to take place, this approach has to further assume that word-initial /r/ and /l/ in English are both good exemplars of the single liquid categories of the two languages (the norm is [ɾ] in this position). Figure 3.1 illustrates why word-initial /r/ and /l/ in English must be good exemplars of the Korean and Japanese liquid categories for the unfaithful perception to occur:

Figure 3.1 Single-Category Assimilation in Korean and Japanese speakers' perception of word-initial /r/ and /l/ in English

i. Good exemplars: Unfaithful perception occurs ii. Bad exemplars: No Unfaithful perception



[23] Whether a foreign sound is a good or bad exemplar of a native category is related to how the foreign sound and the native category are phonetically realized, which is sensitive to where they occur within the syllable. For example, the Korean liquid category is realized as [ɾ] in onset but as light [l] in coda, so Korean speakers probably consider English /l/ a worse exemplar when it is in coda than when it is in onset.

Korean and Japanese speakers have great difficulty distinguishing word-initial /r/ and /l/ in English because the two sounds are equally distant from the norm of the single liquid category of their native language. If the two sounds are good exemplars, they are close to the norm in the phonetic space and perceiving them as the norm occurs (i). If the two sounds are bad exemplars, they are far away from the norm in the phonetic space, and the same kind of unfaithful perception is not possible given that the listeners hear accents (ii).

The question now is whether word-initial /r/ and /l/ in English are good or bad exemplars of the Korean and Japanese liquid categories. Evidence shows that they are bad exemplars.

3.2.2 Japanese Speakers' Perception of Word-Initial /r/ and /l/ in English

Previous perceptual studies have shown that English /r/ sounds more similar to Japanese /w/ than to Japanese /r/ (see references cited in LaCharité and Paradis 2005: 249). Moreover, Best and Strange (1992) tested Japanese speakers' identification and discrimination of English /r/ and /w/, and one of the results obtained was that Japanese learners of English performed almost as well as American English monolinguals. Because Japanese learners of English are capable of discriminating the more difficult contrast E. /r/–E. /w/, we expect them to be able to discriminate the less difficult contrast E. /r/–J. /r/.

Perceptual evidence also shows that, for monolingual Japanese speakers, word-initial /r/ and /l/ in English do not sound similar to the Japanese liquid category. Guion, Flege, Akahane-Yamada, and Pruitt (2000) carried out a cross-linguistic perceptual experiment in which monolingual Japanese-speaking participants listened to English and Japanese consonant+/a/ stimuli and matched the stimuli with a Japanese sequence that they thought sounded the most similar. In addition, the participants were asked to give a goodness rating to each stimulus presented to them. The results relevant to the present discussion are given below:

Table 3.2 Monolingual Japanese speakers' perception of various English and Japanese consonant+/a/ sequences (Guion, Flege, Akahane-Yamada, and Pruitt 2000: 2715, 2717)

English or Japanese consonant+/a/ stimulus	most common match	proportion of responses the most common match receives	goodness rating (on a scale of 1–7, with 7 being the best rating)	fit index (good, fair, or poor match)	
E. ba	ba	84%	5.3	4.5	Good match
E. wa	ɰa	79%	3.5	2.8	Fair match
E. ra	ɯɾa	50%	3.3	1.7	Poor match
	ɾa	46%	3.4	1.6	Poor match
E. la	ɾa	50%	3.2	1.6	Poor match
	ɯɾa	37%	3.0	1.1	Poor match
J. ɾa	ɾa	99%	4.8	4.8	Good match

More than one Japanese form may appear as the most common match for an English stimulus (e.g. [ɯɾa] and [ɾa] are the most common matches for the English stimulus /ra/), but most of the time one Japanese form is dominating. The fit index scores are obtained by multiplying the proportion of responses by the goodness rating, and they are calculated to avoid the situation where a Japanese form receiving a low goodness rating ends up being the only most common match because there are no other good competitors. The matchings J. /ɾa/ → [ɾa] and E. /ba/ → [ba] receive the highest fit index scores, and these scores serve as reference points so that the fit index scores received by other matchings can be interpreted. [ɾa] is of course a good match for J. /ɾa/. [ba] is also a good match for E. /ba/. The matching E. /wa/ → [ɰa] receives a fit index score of 2.8, and the matchings E. /ra/ → [ɯɾa], E. /ra/ → [ɾa], E. /la/ → [ɯɾa], and E. /la/ → [ɾa] all receive a fit index score that is lower than 2.0. The authors view [ɰa] as a fair match for E. /wa/ and both [ɯɾa] and [ɾa] as poor matches for E. /ra/ and E. /la/.

[ɾa] being a poor match for E. /ra/ and E. /la/ indicates that word-initial /r/ and /l/ in English are not perceptually similar to [ɾ] and therefore are bad exemplars of the Japanese liquid category. Since the participants are monolingual speakers, who have very little or even no exposure to spoken English, it is reasonable that word-initial /r/ and /l/ in English are also bad exemplars of the Japanese liquid category for Japanese learners of English.

3.2.3 Korean Speakers' Perception of Word-Initial /r/ and /l/ in English

While phonetic data on Korean speakers' perception of word-initial /r/ and /l/ in English are unavailable to me, English words borrowed into Korean shed some light on whether Korean speakers perceive them as the Korean liquid category. Consider the loanword data in (15):

(15) Adaptation of /r/ and /l/ in word-initial position and in (s)+C+liquid+V clusters in English loanwords in Korean (data from Kang 2003 and Kenstowicz 2005)

i. word-initial

line	→	[ɾɑ.in]	*rail*	→	[ɾe.il]
lobby	→	[ɾo.bi]	*rope*	→	[ɾo.pʰɨ]
level	→	[ɾe.bel]	*raincoat*	→	[ɾɛ.in.kʰo.tʰɨ]

ii. (s)+C+liquid+V

slip	→	[sɨl.lip]	*grip*	→	[kɨ.ɾip]
slap	→	[sɨl.læp]	*scrap*	→	[sɨ.kʰɨ.ɾæp]
slope	→	[sɨl.lo.pʰɨ]	*microchip*	→	[mɑ.i.kʰɨ.ɾo.tsʰip]
classic	→	[kʰɨl.læ.sik]	*electric*	→	[il.lɛk.tʰɨ.ɾik]

In word-initial position, both /r/ and /l/ are mapped to [ɾ]. When occurring in a (s)+C+liquid+V cluster, /r/ is mapped to [ɾ] but /l/ geminates. Kenstowicz (2005) noted that in the cluster context while /l/ was found to map to [ɾ] sometimes (e.g. *glass* → [kɨl.lɑ.sɨ] or [kɨ.ɾɑ.sɨ]), /r/ never geminated. He concluded that Korean speakers are capable of distinguishing the two English liquids in (s)+C+liquid clusters. On the basis of this finding, we are confident to assume that Korean speakers can also distinguish the two English liquids in word-initial position, as word-initial position is a strong prosodic position and the acoustic cues are less likely to be skewed (see also Yip 2006: 953 for similar arguments and a similar case in English loanwords in Cantonese). Because the two English liquids are distinguishable in word-initial position, the contrast neutralization in loanwords, as shown in (15 i), cannot be attributed to *both* sounds being perceived as [ɾ].

A conservative position is that Korean speakers perceive one of the two liquids as [ɾ], indicating that the perception is of the CG type (if this is the case, the one that is perceived as [ɾ] is probably /l/; see below). This position, however, is weakened if we consider the fact that listeners are sensitive to foreign accents and that word-initial /r/ and /l/ in English are both regarded by Japanese speakers, whether monolingual or bilingual, as bad exemplars of the Japanese liquid category, the norm of which is also [ɾ]. In particular, note that, in Table 3.2, [ɯɾa] is one of the two most common matches for E. /ra/ and E. /la/. The authors argued that 'the /ɯ/ before the /ɾ/ might be the result of the approximate production of English /ɹ/, which is more vowel-like than a tap' (pp. 2716). Although /l/ is less vowel-like than /r/ in English,[24] it is still an approximant, and this explains why [ɯɾa] was frequently selected as the best match for English /la/ (37% of the time) in the experiment. Given this, it is very likely that word-initial /r/ and /l/ in English are not good exemplars of a liquid category that is phonetically realized as [ɾ].

3.2.4 Conclusion

Because the members of a non-native contrast cannot be distinguished from each other does not necessarily mean that they cannot be distinguished from the norm of the native category to which they are perceptually assimilated. If they are bad exemplars of the native category, the

[24] The matching E. /ra/ → [ɯɾa] receives a higher goodness rating than the matching E. /la/ → [ɯɾa]. Also, [ɯɾa] is chosen as the best match for E. /ra/ more often than it is chosen as the best match for E. /la/ (50% of the time vs. 37% of the time).

borrowers readily hear the accents, and the accents are expected to be handled in the production grammar even if neutralization of the contrast is what we see in the output.

3.3 Summary of Chapter Three

In this chapter, I have presented evidence arguing against the purely perceptual approach. First, I have shown that while word-final non-labial stops and coda liquids may be deleted in English loanwords in SC, they are always preserved with the place of articulation maintained when the source words are monosyllabic and contain no other sounds violation the SC syllable structure requirements. These data indicate that the deletion is not due to unfaithful perception but takes place in the SC production grammar. Second, I have shown evidence from Korean and Japanese that featural changes in borrowed foreign words may not be due to unfaithful perception either, as listeners are very sensitive to accents.

CHAPTER FOUR

ARGUMENTS AGAINST THE PURELY PHONOLOGICAL APPROACH

In the view of the purely phonological approach, loanword adaptation is performed by bilinguals, who have access to the underlying representations of the source words. It is held that when foreign words are borrowed, the underlying representations of the foreign words are mapped in the host language lexicon to the structurally closest native representations, which subsequently undergo phonological processes at the post-lexical level and eventually surface in conformity with the host language grammar. This view predicts that non-contrastive features and phonetic details of sounds in the source and host languages play no role in the adaptation processes. In this chapter, I will present evidence from English loanwords in SC and argue that this prediction is false. 4.1 concerns the role of frication noise in the adaptation of /v/, 4.2 concerns the role of tongue-body position in the adaptation of palato-alveolars (e.g. /ʃ/, /tʃ/, and /dʒ/), and 4.3 concerns stop deletion. 4.4 is the summary.

4.1 The Role of Frication Noise in the Adaptation of /v/ in English Loanwords in SC

Loanword data from various languages have demonstrated that non-contrastive features and/or phonetic details of sounds in the source words affect how the source words are adapted, giving rise to the "divergent repairs" phenomenon or adaptations that seem unnecessary. Some cases have been provided in 2.1.2. In this section, I present a divergent-repair case from English loanwords in SC. In this case, /v/ is found to be adapted as a function of its position relative to a vowel. On the basis of acoustic studies that demonstrate a correlation between a /v/'s position relative to a vowel and the /v/'s frication-noise level, I argue that the amount of frication noise contained in a /v/ determines how it is adapted. I first show the adaptation patterns in 4.1.1 and then the challenges these patterns may pose to the purely phonological approach in 4.1.2. A phonetic account is proposed in 4.1.3.

4.1.1 Adaptation Patterns

My corpus contains 93 instances of /v/ that have a correspondent in the SC adapted forms.[25] Of these instances, 86 occur before a vowel, five occur before a sonorant consonant, and two occur word-finally. Of the 86 instances that occur before a vowel, 59 are mapped to [w] and 27 to [f]. Of the seven instances that do not occur before a vowel, only one is mapped to [w], and all of the other six are mapped to [f]. A summary of the data is given in Table 4.1. Examples are shown in (19). (The numerals in Table 4.1 denote the numbers of instances.)

Table 4.1 A summary of the adaptation of /v/ as a function of the /v/'s position relative to a vowel in English loanwords in SC ($\chi^2 = 8.34$, $p < .01$)

	/v/ → [w]	/v/ → [f]	Sum
Prevocalic /v/	59	27	86
	68.60%	31.40%	100%
Non-prevocalic /v/	1	6	7
	14.29%	85.71%	100%

(16) Examples of the adaptation of /v/ as a function of the /v/'s position relative to a vowel in English loanwords in SC

i.	Prevocalic /v/ → [w]			Prevocalic /v/ → [f]		
	New Haven	→	[nou.xai.wən]	*Harvard*	→	[xa.fʷo]
	Ava	→	[ai.wa]	*Evans*	→	[i.fan]
	Davis	→	[tai.wei.sz̩]	*Shaver*	→	[ɕe.fʷo]
	Kevin	→	[kʰai.wən]	*Vanderbilt*	→	[fan.tɤ.pɑu]
ii.	Non-prevocalic /v/ → [w]			Non-prevocalic /v/ → [f]		
	Lovelock	→	[la.wei.lʷo.kʰɤ]	*Evelyn*	→	[ai.fu.lin]

[25] There is a total of 94 instances of /v/ in my corpus. One of the instances does not have a correspondent in the SC adapted form (*Albertville* → [a.pei.tʰɤ]).

Che<u>v</u>ron	→	[ɕʷe.<u>f</u>u.luŋ]
Ro<u>v</u>e	→	[lʷo.<u>f</u>u]
Li<u>v</u>e	→	[lai.<u>f</u>u]

As we can see in Table 4.1, prevocalic /v/ tends to be adapted as [w] and non-prevocalic /v/ tends to be adapted as [f].

4.1.2 Challenges to the Purely Phonological Approach

Paradis and LaCharité (1997) assumed that underlying representations are radically underspecified. Because [+continuant], [+voiced], and [Labial] are the only features in the radically underspecified underlying representation of English /v/ (Stemberger 1991), simply inserting [+sonorant] will turn English /v/ into SC [w]. As [w] is the SC sound that is phonologically closest to English /v/ according to Paradis and LaCharité's (1997) model, this model predicts that every instance of /v/ in the English source words is mapped to [w] in the SC lexicon. This model runs into a problem: It fails to account for the fact that most instances of non-prevocalic /v/ are realized as [f]. To exemplify this problem, consider the adaptation *Che<u>v</u>ron* → [ɕʷe.<u>fu</u>.luŋ], where the /v/ in the source word corresponds to the SC syllable [fu]. The presence of [f] cannot be explained if the /v/ in the source word is mapped to [w] in the SC lexicon, as [ɕʷe.<u>fu</u>.luŋ] is always a worse output candidate than [ɕʷe.<u>u</u>.luŋ] and will never be selected by the grammar.

The SC data reminds us of the adaptation of /v/ in French loanwords in Fula. Paradis and LaCharité (1997) showed that, of the 81 instances of /v/ that are not deleted and do not surface as non-adapted importations in their corpus, 62 are realized as [w] (76.5%), 14 as [b] (17.3%), and five as [f] (6.2%). Their analyses did not consider the /v/'s position in the source words; rather, they argued that /v/ tends to realize as [w] because, among the three minimal repairs, /v/ → [w]

satisfies the Preservation Principle (every distinctive feature in /v/ is retained) but the other two do not ([+continuant] is lost in /v/ → [b] and [+voiced] is lost in /v/ → [f]). Interestingly, two of the three examples they gave for /v/ → [f] involve a non-prevocalic /v/ (*élève* /elɛv/ → [ɛlɛf] 'student', *mouvement* /muvmã/ → [mufmaŋ] 'movement', *télévision* /televizjɔ̃/ → [tɛlɛfisjɔŋ] 'television'), and all of the three examples they gave for /v/ → [w] involve a prevocalic /v/ (*avocat* /avɔka/ → [awɔka] 'lawyer', *civil* /sivil/ → [siwil] 'civil', *verre* /vɛr/ → [wɛːr] 'glass') (pp. 400–401).[26] I suspect that the adaptation patterns will resemble those seen in Table 4.1 once the /v/'s position in the French source words is taken into consideration.

4.1.3 The Role of Frication Noise

I propose a phonetic account. Specifically, a /v/'s frication-noise level, which correlates with its position relative to a vowel, determines how it is adapted.

Steriade (2009) argued that devoicing is the preferred repair to voiced obstruents because the change from voiced to voiceless is perceptually minimal. Kenstowicz and Suchato (2006) reported that prevocalic /v/ in English loanwords in Thai is also realized as [w] (e.g. *visa* → [wīisâa], *vote* → [wòot]), and to account for the unexpected mapping, they argued that '[t]he voiced labio-dental /v/ is the least turbulent of the English fricatives, [and] if sounds are judged in terms of their distance in auditory space, then it is conceivable that /v/'s minimal turbulence is sufficient to push it into the sonorant region of the Thai speaker's P-Map and hence make /w/ the closest match' (pp.926). Hawaiian is another language in which prevocalic /v/ in English loanwords is realized as [w] (Clements 2001; Herd 2005; Adler 2006; e.g. *violin* → [waiolina], *drive* → [kalaiwa]), and Clements (2001: 86–87) had a similar argument: 'The only other

[26] Paradis and LaCharité (1997) gave three examples for the mapping /v/ → [b]; one contains a /v/ that occurs before a glide (*avion* /avjɔ̃/ → [abijɔn] 'aeroplane'), one contains a /v/ that occurs before a liquid (*livre* /livr/ → [liːbaːr] 'book'), and one contains a /v/ that occurs before a vowel (*vinaigre* /vinɛgr/ → [binɛːgara] 'vinegar'). They pointed out that the source words might be from Serer (pp. 404).

surprising element in the pattern […] is the interpretation of [v] as /w/ instead of the expected /p/[;] […] [s]ince English [v] itself has frictionless realizations, it would appear that Hawaiian speakers categorize English *v* as a sonorant sound[.]' Kenstowicz and Suchato' (2006) and Clements' (2001) comments on English /v/'s frication-noise level are supported acoustically. Olive, Greenwood, and Coleman (1993: 93–94) found that the waveforms of /v/ and /ð/ are much smoother and exhibit less frication noise than the waveforms of the other two voiced fricatives /z/ and /ʒ/ in English. Since voiced fricatives have weaker frication than their voiceless counterparts, as evidenced by the finding that the former show lower air pressure behind the constriction than the latter (Collier, Lisker, Hirose, and Ushijima 1979), /v/ and /ð/ are very likely the least turbulent fricatives in English. Other languages in which prevocalic /v/ in borrowed English words tends to realize as [w] include Cantonese (Silverman 1992; Yip 1993; Bauer and Benedict 1997; e.g. *volume* → [wɔlɐm], *PVC* → [p^hiːwiːsiː]) and NZ Maori (Herd 2005; e.g. *value* → [waariu], *level* → [reewara]). Kenstowicz and Suchato's (2006) and Clements' (2001) arguments, along with the acoustic evidence, offer a reasonable phonetic explanation for the commonly observed mapping /v/ → [w] in the adaptation of English words into a number of languages.

Acoustic studies also show that, in English, voiced fricatives are partially devoiced when followed by a voiceless consonant or occurring at the end of the utterance (Ladefoged and Johnson 2011; Pickett 1999: 128–130; Smith 1997). Ladefoged and Johnson (2011: 66) pointed out that this is particularly true of /v/. Although partially devoiced /v/ is not identical to /f/, it surely contains more frication noise than fully voiced /v/. I argue that this fact explains why word-final /v/ is not mapped to [w] but to [f] in English loanwords in SC. As for the cases in which pre-sonorant /v/ is mapped to [f] in my data, given that fricatives cannot occur in coda position in SC and that a syllable usually represents a word (L. Cheng 1991), I suggest that SC learners of English treat all pre-consonantal fricatives as if they were at the end of a word. Accordingly, pre-sonorant /v/ is mapped to [f] just like word-final /v/.

4.2 The Role of Tongue-Body Position in the Adaptation of Palato-Alveolars in English Loanwords in SC

In this section, I consider the adaptation of the three palato-alveolars /ʃ/, /tʃ/, and /dʒ/ in English loanwords in SC. The potential SC matches are the alveolo-palatals [ɕ], [tɕʰ], and [tɕ], the post-alveolars [ʂ], [ʈʂʰ], and [ʈʂ], and the dentals [s], [tsʰ], and [ts], and my data show that /ʃ, tʃ, dʒ/ are mostly mapped to [ɕ, tɕʰ, tɕ]. To account for this general tendency, I suggest that the SC adapters want to retain the front tongue-body position associated with /ʃ, tʃ, dʒ/. I show the general adaptation tendency in 4.2.1 and then the challenges to the purely phonological approach in 4.2.2. In 4.2.3, I propose an articulatory account on the basis of Flemming's (2003) research on the tongue position during the production of coronal consonants.

4.2.1 General Adaptation Tendency

Of the 165 instances of /ʃ, tʃ, dʒ/ that do not occur before a high front vowel in my corpus, 116 are adapted as [ɕ, tɕʰ, tɕ], 47 are adapted as [ʂ, ʈʂʰ, ʈʂ], and two are adapted as [tsʰ, ts].[27] A summary and examples of the adaptation are given below. (The numerals in the table denote the numbers of instances in my corpus.)

---

[27] There is a total of 192 instances of /ʃ/, /tʃ/, and /dʒ/ in my corpus (61 instances of /ʃ/, 49 instances of /tʃ/, 82 instances of /dʒ/, no instance of /ʒ/). Twenty-one of them precede a high front vowel and are excluded because (i) high front vowels in the source words are almost always mapped to [i] and (ii) [ʂ, ʈʂʰ, ʈʂ] and [s, tsʰ, ts] cannot precede [i]. Another six instances are excluded because the adaptations are obviously influenced by the orthography (e.g. *Sean* → [ɕi.ən], *Geter* → [kai.tʰɤ]). Instances of /ʃ/, /tʃ/, and /dʒ/ are combined as the adaptations do not show significant differences that might weaken my arguments. The SC matches may contain a labial feature, which probably originates from the labialization of /ʃ, tʃ, dʒ/. The fact that /tʃ/ is almost always mapped to an aspirated sound is problematic for the purely phonological approach, as aspiration of stops and affricates is predictable in English and therefore is not specified in the UR's. Comparable cases are also found in English loanwords in other languages. For example, /f/ is mapped to [pʰ] rather than [p] in Korean, and /ʃ/ is mapped to [tʃʰ] rather than [tʃ] in Thai. This issue is not discussed here. See Kenstowicz and Suchato (2006: 925) for a brief discussion of the Thai case.

Table 4.2 A summary of the adaptation of /ʃ, tʃ, dʒ/ that do not precede a high front vowel in English loanwords in SC ($\chi^2 = 119.89$, $p < .001$)

	[ɕ, tɕʰ, tɕ]	[ʂ, ʈʂʰ, ʈʂ]	[tsʰ, ts]	Sum
/ʃ, tʃ, dʒ/	116	47	2	165
	70.30%	28.48%	1.21%	100%

(17) Examples of the adaptation of /ʃ, tʃ, dʒ/ that do not precede a high front vowel in English loanwords in SC (symbols in parentheses denote the following vowels, and '#' denotes 'followed by silence')

i. Palato-alveolar → Alveolo-palatal

/ʃ/ → [ɕ]

Shuster (/u/) → [ɕou.sz̩.tʰɤ]

Shaver (/e/) → [ɕe.fʷo]

Chevron (/ɛ/) → [ɕʷe.fu.luŋ]

Sasha (/ə/) → [ʂa.ɕa]

Sharon (/æ/) → [ɕʷe.lʷən]

Charlotte (/ɑ/) → [ɕa.lʷo.tʰɤ]

Ashley (/l/) → [ai.ɕi.li]

Kushner (/n/) → [kʰɤ.ɕi.na]

Marsh (#) → [ma.ɕy]

Nash (#) → [nai.ɕy]

/tʃ/ → [tɕ(ʰ)]

Chambers (/e/) → [tɕʰɛn.pʷo.sz̩]

Cherry (/ɛ/) → [tɕʰʷe.li]

Mitchell (/ə/) → [mi.tɕʰi.ər]

Chandler (/æ/) → [tɕʰɛn.tɤ.lɤ]

ii. Palato-alveolar → Post-alveolar

/ʃ/ → [ʂ]

Schumer (/u/) → [ʂu.mʷo]

Shelly (/ɛ/) → [ʂa.li]

Keisha (/ə/) → [tɕi.ʂa]

Charlize (/ɑ/) → [ʂa.li.tsz̩]

Schmidt (/m/) → [ʂz̩.mi.tʰɤ]

Schwartz (/w/) → [ʂz̩.xʷa.tsz̩]

/tʃ/ → [ʈʂ(ʰ)]

Chula (/u/) → [ʈʂu.la]

Rochester (/ɛ/) → [lʷo.ʈʂʰɤ.sz̩.tʰɤ]

Richard (/ə/) → [li.ʈʂʰa]

Patchogue (/ɔ/) → [pʰa.ʈʂʰu.kɤ]

Hitchcock (/k/) → [ɕi.tɕʰy.kʰɑu.kʰɤ]

Richmond (/m/) → [li.tɕʰi.məŋ]

Coach (#) → [kʰɤ.tɕʰi]

Lynch (#) → [lin.tɕʰi]

/dʒ/ → [tɕ(ʰ)]

Jude (/u/) → [tɕʰou.tɤ]

Jake (/e/) → [tɕe.kʰɤ]

Jeff (/ɛ/) → [tɕe.fu]

Roger (/ə/) → [lʷo.tɕe]

Joanna (/o/) → [tɕɑu.an.na]

George (/ɔ/) → [tɕʰɑu.tʂʐ̩]

Jack (/æ/) → [tɕe.kʰɤ]

Jobs (/ɑ/) → [tɕʰɑu.pu.sz̩]

Ridgewood (/w/) → [ʐʷei.tɕi.u]

Cage (#) → [kʰai.tɕi]

Drudge (#) → [tʂʷo.tɕʰi]

Chattanooga (/æ/) → [tʂʰa.ta.nu.tɕa]

Charlie (/ɑ/) → [tʂʰa.li]

Greenwich (#) → [kɤ.lin.wei.tʂʐ̩]

/dʒ/ → [tʂ]

Judy (/u/) → [tʂu.ti]

James (/e/) → [tʂan.mu.sz̩]

Jennifer (/ɛ/) → [tʂən.ni.fʷo]

Benjamin (/ə/) → [pan.tʂɤ.miŋ]

Jodie (/o/) → [tʂu.ti]

Jordan (/ɔ/) → [tʂʷo.təŋ]

Janet /æ/ → [tʂən.ni]

Johnson (/ɑ/) → [tʂan.sən]

George (#) → [tɕʰɑu.tʂʐ̩]

iii. Palato-Alveolar → Dental

/ʃ/ → [s]

N/A

/tʃ/ → [tsʰ]

Child (/ɑɪ/) → [tsʰai.tɤ]

/dʒ/ → [ts]

Jersey (/ə/) → [tsɤ.ɕi]

4.2.2 Challenges to the Purely Phonological Approach

According to Stemberger (1991: 75), radically underspecified English /ʃ/, /tʃ/, and /dʒ/ are [+continuant, Coronal, −anterior], [Coronal, −anterior], and [+voiced, Coronal, −anterior], respectively. Thus, Paradis and LaCharité's (1997) model predicts that, in the SC lexicon, /ʃ/ is mapped to [ʂ], which is also [+continuant, Coronal, −anterior], and both /tʃ/ and /dʒ/ are mapped to [ʈʂ], which is [Coronal, −anterior]. While this model correctly predicts that /ʃ, tʃ, dʒ/ are rarely adapted as [s, ts^h, ts], which are [+anterior], it fails to predict that they are in the vast majority of cases adapted as [ɕ, $tɕ^h$, tɕ] when not preceding a high front vowel. [ɕ, $tɕ^h$, tɕ] are predictable in SC because they surface only when [s, ts^h, ts], [ʂ, $ʈʂ^h$, ʈʂ], or [x, k^h, k] precede a high front vowel (e.g. UR [sia] → [ɕa] 'down' with Tone 4). As allophones, [ɕ, $tɕ^h$, tɕ] are not present in the lexicon.

4.2.3 An Articulatory Account: The Tongue-Body Position Plays a Role

It has been reported in the literature that coronal consonants can condition fronting of adjacent vowels (e.g. Hume 1992). However, not all coronal consonants can front adjacent vowels; some coronal consonants, especially retroflexes, actually retract adjacent vowels. To account for this asymmetry, Flemming (2003) argued that different types of coronal consonants prefer different tongue-body positions, which facilitate the creation of the constrictions and hence are closely related to the positions of the tongue tip or blade during articulation. Based on articulatory and acoustic evidence, Flemming (2003) suggested that anterior coronals and non-anterior laminal coronals, both of which condition vowel fronting, favor a front tongue body, and that non-anterior apical coronals, which condition vowel retraction, favor a back tongue body. In addition, Flemming (2003) argued that while both anterior coronals and non-anterior laminal coronals favor a front tongue body, the tongue-body position associated with non-anterior laminal coronals is universally more front than that associated with anterior coronals. Flemming (2003) also argued that the fronting and retraction effects of coronal consonants on adjacent vowels result from effort minimization: If a coronal consonant precedes or follows a vowel that is produced with a similar tongue-body position, the tongue body does not have to quickly move when the sequence is produced.

Flemming's (2003) suggestions provide an explanation for the general adaptation tendency as shown in Table 4.2. As non-anterior laminal coronals, /ʃ, tʃ, dʒ/ favor a front tongue body. [ɕ, tɕʰ, tɕ] are basically palatalized palato-alveolars (Ladefoged and Maddieson 1996: 150–153, cited in Lin 2007: 47 and Flemming 2003: 338, 365), so they favor a front tongue body as well (Flemming 2003: 338).[28] [s, tsʰ, ts] also favor a front tongue body because they are anterior coronals. [ʂ, ʈʂʰ, ʈʂ], on the other hand, favor a back tongue body since they are non-anterior apical coronals.[29] The tongue-body positions preferred by the four sets of coronal consonants are summarized below:

Table 4.3 Tongue-body positions preferred by the palato-alveolars /ʃ, tʃ, dʒ/, the alveolo-palatals [ɕ, tɕʰ, tɕ], the dentals [s, tsʰ, ts], and the post-alveolars [ʂ, ʈʂʰ, ʈʂ]

	Palato-alveolars /ʃ, tʃ, dʒ/	Alveolo-palatals [ɕ, tɕʰ, tɕ]	Dentals [s, tsʰ, ts]	Post-alveolars [ʂ, ʈʂʰ, ʈʂ]
Type of coronal consonant	non-anterior laminal coronal	non-anterior laminal coronal	anterior coronal	non-anterior apical coronal
Tongue-body position preferred	front	front	front but not as front as a non-anterior laminal coronal	back

[28] Duanmu (2007: 31–34) has a similar idea. He analyzed SC [ɕ, tɕʰ, tɕ] as consisting of [Coronal] and [Dorsal] with the feature [–back] attached to [Dorsal].

[29] Lin (2007: 45) argued that the SC sounds that are usually transcribed as [tʂ], [tʂʰ], and [ʂ] (= [ʈʂ], [ʈʂʰ], and [ʂ] in Duanmu 2007 and in this dissertation) are not retroflexes because according to phonetic findings 'it is the upper side, not the underside, of the tongue tip that approaches the back of the alveolar ridge.' Nonetheless, Flemming (2003: 338–339) stated that 'retroflexes cover a range of articulations from full retroflexion, in which the underside of the tongue tip contacts the hard palate, to apical postalveolars, in which the tip of the tongue forms a constriction just behind the alveolar ridge.'

Thus, the fact that /ʃ, tʃ, dʒ/ are mostly adapted as [ɕ, tɕ^h, tɕ] in my data can be understood as the SC adapters' preference for retaining the front tongue-body position with which /ʃ, tʃ, dʒ/ are produced.

The proposal in this section implies that while the tongue-body positions characterizing the English and SC coronal fricatives/affricates are phonetic features/details, they must be specified in the representations used by the adapters.

Notice that [ɕ, tɕ^h, tɕ] are allophones, which are derived from [s, ts^h, ts], [ʂ, tʂ^h, tʂ], or [x, k^h, k] before a high front vowel (e.g. in SC, /si/ → [ɕi]). As instances of /ʃ, tʃ, dʒ/ followed by a high front vowel in the source words are not considered, the data shown in this section represent another case where a phoneme in the source language is directly mapped to an allophone in the host language. Because it is the preference for retaining the front tongue-body position that yields the adaptation tendency, the purely phonological approach, which claims mappings occur in the host language lexicon, is seriously challenged.

4.3 Deletion of Stops in English Loanwords in SC

Segment deletion does not occur often in loanword adaptation. Kang (2003: Footnote 15 on pp. 238) observed that deletion is very rare in adapting foreign words into Korean. In LaCharité and Paradis' (2005) loanword corpora, deletion accounts for only 2.6% of the adaptation of ill-formed sounds (i.e. those that the host language does not have) and ill-formed structures (i.e. those that do not meet the host language syllable requirements), and they claimed that most of the deletion cases can be explained phonologically. According to Paradis and LaCharité's (1997) model, segment deletion takes place only when an ill-formed segment occurs in an ill-formed structure; when a well-formed segment occurs in an ill-formed structure or an ill-formed segment occurs in a well-formed structure, the ill-formed segment or structure is repaired through vowel epenthesis or featural change rather than segment deletion. In this section, I present SC data showing that segment deletion mostly takes place when a non-word-initial stop, whether ill-formed or well-formed, appears in an ill-formed structure. SC loanword data are shown in 4.3.1. In 4.3.2, I give a different set of data I have collected online, which shows how words in Korean, Southern Min, or Cantonese pop songs are transliterated into SC. This set of data exhibits exactly the same patterns as the loanword data. 4.3.3 is the summary.

4.3.1 Loanword Data

My corpus also shows that deletion is seldom used as a strategy to repair ill-formed segments or structures compared to other strategies such as featural change and vowel epenthesis. Interestingly, most of the deletions found in my corpus are deletions of non-prevocalic liquids and stops. Detailed analyses of liquid deletion will be presented in Chapter Eight, so I will focus on stop deletion here.

4.3.1.1 Stops Delete More Frequently

Consider Table 4.4, which gives the frequencies of deletion of obstruents and /m/ in prevocalic and non-prevocalic position in my corpus. /m/ is included because it cannot end a syllable in SC.

Table 4.4 Frequencies of deletion and preservation of obstruents and /m/ in prevocalic and non-prevocalic position in English loanwords in SC

		Stop		Fricative		Affricate		/m/	
Prevoc.	Del.	0	0%	0	0%	0	0%	0	0%
	Pres.	1015	100%	424	100%	108	100%	246	100%
	Sum	1015	100%	424	100%	108	100%	246	100%
Non-prevoc.	Del.	104	18.31%	14	3.06%	0	0%	0	0%
	Pres.	464	81.69%	443	96.94%	18	100%	55	100%
	Sum	568	100%	457	100%	18	100%	55	100%

Unsurprisingly, prevocalic obstruents and /m/ are all preserved, whether they are well-formed or ill-formed from the perspective of SC. However, a clear asymmetry is observed between stops and the other sounds in non-prevocalic position. In non-prevocalic position, while instances of fricatives, affricates, and /m/ are almost always preserved, nearly one in every five instances of stops undergoes deletion (stop vs. non-stop: $\chi^2 = 70.17$, $p < .00001$).

This asymmetry poses a challenge to Paradis and LaCharité's (1997) loanword model. To better express the problem, consider first the frequencies of non-prevocalic /t/, /f/, and /m/ deletion, which are 27.15% (41/151), 2.17% (1/46), and 0% (0/55), respectively. Since SC has these sounds, Paradis and LaCharité's (1997) model predicts that they are preserved regardless of their position relative to a vowel. However, while non-prevocalic /f/ and /m/ are almost always preserved, non-prevocalic /t/ undergoes deletion more than one-fourth of the time. Another example concerns /d/, /z/, and /dʒ/, which SC does not have. Paradis and LaCharité's (1997) model predicts that these sounds would be deleted when occurring in an ill-formed structure. However, my corpus shows that while non-prevocalic /d/ is deleted at the very high frequency of 37.82% (45/119), non-prevocalic /z/ and /dʒ/ are almost always preserved, with non-prevocalic /z/ deleted at the frequency of 2.86% (3/105) and non-prevocalic /dʒ/ always preserved. The last example concerns the fricatives, affricates, and /m/. Among these sounds, only /f/, /s/, and /m/ are in the SC sound inventory. The fact that all these sounds, whether SC have them or not, are hardly deleted in non-prevocalic position is unexpected for Paradis and LaCharité's (1997) model. The data shown above clearly indicate that, when occurring in non-prevocalic position, the likelihood of a sound being deleted is not related to whether SC has it, but is related to whether it is a stop.

4.3.1.2 Word-Initial Stops Never Delete

Another challenge to Paradis and LaCharité's (1997) model concerns the position of stops within the word. While non-prevocalic stops are deleted frequently, only those occurring word-medially or word-finally may delete. Those occurring word-initially are always preserved even if they are voiced (hence ill-formed segments for SC) (e.g. *Blair* → [pu.lai.ər], *Drain* → [tɤ.lai.ən], *Grant* → [kɤ.lan.tʰɤ]). The data are given in Table 4.5:[30]

Table 4.5 Frequencies of deletion and preservation of non-prevocalic stops as a function of the stops' position within the word in English loanwords in SC (word-initial vs. non-word-initial: $\chi^2 = 37.84$, $p < .00001$)

		Word-initial				Word-medial				Word-final			
		Voiced		Voiceless		Voiced		Voiceless		Voiced		Voiceless	
Non-prevoc. stop	Del.	0	0%	0	0%	9	13%	23	17%	40	54%	32	28%
	Pres.	66	100%	48	100%	58	87%	115	83%	34	46%	83	72%
	Sum	66	100%	48	100%	67	100%	138	100%	74	100%	115	100%

According to Paradis and LaCharité's (1997) model, when occurring in a structure that violates the host language syllable constraints, segments that the host language does not have are deleted and segments that the host language have are preserved, whether the structure is word-initial, word-medial, or word-final. As we can see in Table 4.5, the factor that determines whether non-prevocalic stops may delete is the stops' position within the word, rather than whether they are voiced or voiceless. Again, Paradis and LaCharité's (1997) loanword model fails to account for this asymmetry.

4.3.2 Transliteration Data

The loanword data given above demonstrate that deletion may occur only when a non-word-initial stop does not precede a vowel. A different set of data shows the same pattern. Many web pages on the China-based website baidu.com provide the SC transliteration of Korean, Southern Min, and Cantonese pop-song lyrics for SC speakers who are fans of the pop cultures

[30] We are considering English words containing non-prevocalic stops that *may* delete. Thus, monosyllabic English words in the shape of (C)+V+stop are excluded from the count because deleting the final stops would result in monosyllabic output forms, which violate the undominated constraint requiring an output form to be at least disyllabic in length (e.g. *Jeep* → [tɕi.pʰu], *Webb* → [wei.pʷo], *Knight* → [nai.tʰɤ], *Rhode* → [lʷo.tɤ], *Mark* → [ma.kʰɤ], *Doug* → [tɑu.kɤ]).

and would like to learn the pop songs quickly. Here is an example, in which a sentence from a Korean pop song is transliterated into SC.

(18)	Korean source in characters	아직도 나를 그렇게 몰라
	Korean source in IPA	a.dʑik.to na.ɾɨl kɨ.ɾʌ.kʰɛ mol.la
	SC transliteration in characters	哈及多那屡扣劳该某拉
	SC transliteration in IPA	xa tɕi tʷo na ly kʰou lɑu kai mou la
		'(You) still don't understand me'

In this example, a Korean character, which is also a syllable, corresponds to a SC character. Relevant to our discussion here is the fact that the second and third Korean characters are pronounced with a syllable-ending /k/ and a syllable-starting /t/, respectively, suggesting that the syllable-final /k/ cannot be syllabified and needs repairing. Because the second Korean character corresponds to an SC character that is pronounced [tɕi], we know that the syllable-final /k/ is deleted in the transliteration.

Table 4.6 gives the numbers of deleted and preserved syllable-final stops found in the transliteration of Korean, Southern Min, and Cantonese words/phrases/sentences into SC in seven web pages on two China-based websites.[31] [32] As in Korean, syllable-final stops are unreleased in Southern Min and Cantonese.

[31] Links to the web pages: Korean 1- http://wenku.baidu.com/view/3abc75916bec0975f565e206.html, Korean 2 - http://wenku.baidu.com/view/a530eab3b4daa58da0114aaa.html?re=view, Korean 3 - http://wenku.baidu.com/view/2e56ec28fc4ffe473368ab71.html?re=view; Southern Min 1 - http://www.fyan8.com/minnan/wenzi.htm, Southern Min 2 - http://tieba.baidu.com/p/181585032, Southern Min 3 - http://zhidao.baidu.com/question/1539542019837288547.html?qbl=relate_question_1&word=%CC%A8%D3%EF%B8%E8%B4%CA%D3%C3%C6%D5%CD%A8%BB%B0%B5%C4%B7%A2%D2%F4%C0%B4; Cantonese - http://wenku.baidu.com/view/8b96d627af45b307e9719700.html?re=view

[32] Many thanks to Ok-Sook Park and Qian Luo for help with transcribing the data.

Table 4.6 Frequencies of deletion and preservation of syllable-final stops in the transliteration of Korean, Southern Min, and Cantonese words/phrases/sentences into SC found in seven web pages in two China-based sites ($\chi^2 = 60.28$, $p < .001$)

	Kor. 1	Kor. 2	Kor. 3	SM 1	SM 2	SM 3	Cant.	Sum	%
Delete	12	8	15	3	3	8	30	79	91.86%
Preserve[33]	3	2	2	0	0	0	0	7	8.14%
Sum	15	10	17	3	3	8	30	86	100%

As we can see, syllable-final stops are deleted at extremely high rates. Paradis and LaCharité's (1997) loanword model fails to account for this fact as SC possesses [p], [t], and [k] as phonemes.[34]

An asymmetry is observed between the two sets of data. It can be clearly seen that non-prevocalic stops in non-word-initial position in the English loanword data set are much more resistant to deletion than syllable-final stops in the other data set (a 23.58% deletion rate (32+72/205+189) for the English data set) ($\chi^2 = 128.25$, $p < .00001$). This asymmetry is apparently related to the fact that syllable-final stops are obligatorily unreleased in Korean, Southern Min, and Cantonese but may be released in English. I will argue in the next chapter that unreleased stops are deleted in English loanwords in SC.

4.3.3 Summary

To sum up, while it is true that segment deletion is uncommon in loanword adaptation, the SC data given above demonstrate that when it occurs, it targets non-prevocalic stops in non-word-initial position. Paradis and LaCharité's (1997) loanword model, which claims that segment deletion occurs only when a segmental constraint and a phonotactic constraint are violated simultaneously, fails to account for these data.

[33] /p/ accounts for five of the seven instances.
[34] Whether stops are released or unreleased is never contrastive in languages, so the fact that unreleased /t/ and /k/ are deleted at a near-100% rate in the Korean, Southern Min, and Cantonese data is particularly challenging for Paradis and LaCharité's (1997) loanword model.

4.4 Summary of Chapter Four

In this chapter, I have presented data arguing that non-contrastive features and phonetic details of sounds in the source and host languages affect how loanwords are adapted. In 4.1, I consider the adaptation of /v/ in English loanwords in SC and suggest that the extent of frication noise in a /v/ determines how it is adapted. 4.2 concerns the adaptation of the three palato-alveolars /ʃ/, /tʃ/, and /dʒ/ that do not precede a high front vowel in English loanwords in SC, and I have demonstrated that the tongue-body position associated with these sounds determine the adaptation patterns. 4.3 concerns stop deletion in English loanwords in SC. The data indicate that deletion takes place only when a non-word-initial stop, whether voiced or voiceless, precedes another consonant. The data are consistent with another set of data, which show that syllable-final stops are deleted at an extremely high frequency when Korean, Southern Min, and Cantonese words/phrases/sentences are transliterated into SC. Paradis and LaCharité's (1997) loanword model, on the other hand, falls short in accounting for all the observed patterns.

CHAPTER FIVE

PROBLEMS FOR THE HYBRID APPROACH

In the last two chapters I have provided evidence against the purely perceptual approach and the purely phonological approach and suggested (i) that loanword adaptation takes place in the host language production grammar and (ii) that loanword adaptation makes reference to the phonetic information in the source words. While the hybrid approach is generally supported, it is insufficient in explaining some of the loanword data. In this chapter, I will present the data and propose solutions. 5.1 concerns the adaptation of word-medial unreleased stops in English loanwords in SC. 5.2 concerns the adaptation of word-final unreleased stops in English loanwords in Korean and SC. 5.3 is the summary of this chapter.

5.1 Word-Medial Unreleased Stops may be Preserved in English Loanwords in SC

In this section I present the first problem for the hybrid approach and propose a solution. First of all, I follow Kang (2003) and argue in 5.1.1 that non-prevocalic released consonants elicit vowel epenthesis. This argument implies that unreleased stops are deleted. However, as I will show in 5.1.2, word-medial stops preceding another stop or a homorganic nasal in the English source words, which are generally considered strictly unreleased in phonetic literature, are preserved frequently. On the basis of acoustic and perceptual studies and research on second language acquisition, I propose in 5.1.3 that these stops may be perceived as released and therefore are preserved via vowel epenthesis. 5.1.4 is the section summary.

5.1.1 The Release-to-Vowel Epenthesis Hypothesis (Kang 2003)

I start with the release-to-vowel epenthesis hypothesis proposed by Kang (2003). In adapting English words into Korean, word-final postvocalic stops are realized with vowel epenthesis sometimes, even though Korean allows stops in coda position. Moreover, it has been reported that the vowel epenthesis occurs more frequently when the pre-stop vowels in the source words are tense than when they are lax (e.g. *qu<u>ick</u>* → /kʰw<u>ik</u>/, *w<u>eek</u>* → /w<u>ikʰɨ</u>/). Based on the findings that word-final stops are more likely to release after tense vowels than after lax vowels in English, Kang (2003) argued that the seemingly unnecessary epenthesis is caused by stop releasing. In what follows, I will show that the data in my corpus generally support Kang's (2003) release-to-vowel epenthesis hypothesis.

5.1.1.1 Word-Final Stops

First, my data show that word-final stops in the source words are preserved via vowel epenthesis or deleted (shown in 4.3), paralleling the fact that they are released or unreleased. Notice that they tend to delete when unreleased, as unreleased stops in Korean, Southern Min, and Cantonese words transliterated into SC almost always delete (also shown in 4.3).

5.1.1.2 Augmentation of CV+/n/ or CV+/ŋ/ Words

Second, my data show that when CV+/n/ or CV+/ŋ/ words are augmented to satisfy the SC minimum-word requirement, a vowel is almost always epenthesized before the word-final nasals, even though epenthesizing a vowel after the word-final nasals also makes the output forms disyllabic. The data are given in Table 5.1 with examples shown in (19):

Table 5.1 Two ways of augmenting CV+/n/ or CV+/ŋ/ words to disyllabicity and the frequencies of use in adapting English words into SC ($\chi^2 = 21.55$, $p < .001$)

	/n/	/ŋ/	sum	%
epenthetic vowel+nasal	25	2	27	93.10%
nasal+epenthetic vowel	1	1	2	6.90%
sum	26	3	29	100%

(19) Examples of two ways of augmenting CV+/n/ or CV+/ŋ/ words to disyllabicity in adapting English words into SC

	Epenthesis before the nasal			Epenthesis after the nasal		
i.	/n/					
	Deane	→	[ti.ən]	*Jeanne*	→	[ʈʂən.ni]
	Bain	→	[pei.ən]			
	Mann	→	[man.ən]			
	One	→	[wan.ən]			
	Cohn	→	[kʰɤ.ən]			
	Kahn	→	[kʰa.ən]/[kʰɑŋ.ən]			
ii.	/ŋ/					
	King	→	[tɕin.ən]	*Young*	→	[jɑŋ.kɤ]
	Long	→	[lɑŋ.ən]			

Because syllable-final nasals are unreleased in English, the fact that the epenthetic vowel is almost always placed before the word-final nasals lends support to Kang's (2003) release-to-vowel epenthesis hypothesis.

5.1.1.3 Word-Medial Pre-Liquid Stops

Third, my data show that while word-medial non-prevocalic stops may delete (shown in 4.3), they are almost always preserved when the following sound is a liquid. The data are given in Table 5.2 with examples shown in (20).[35] The data on the deletion vs. preservation of word-medial non-prevocalic stops are also shown for comparison.

[35] Instances followed by syllabic /l/ (e.g. *Trundle* → [tʰɤ.laŋ.tou]) are not considered.

Table 5.2 Frequencies of deletion and preservation of word-medial pre-liquid stops in English loanwords in SC ($\chi^2 = 9.2549$, $p = .002349$)

	word-medial pre-liquid stop		word-medial non-prevocalic stop	
delete	2	2.53	32	15.61
preserve	77	97.47	173	84.39
sum	79	100%	205	100%

(20) Examples of deletion and preservation of word-medial pre-liquid stops in English loanwords in SC

			Preservation			Deletion		
i.	Before /r/							
	/pr/	*Oprah*	→	[ou.pʰu.la]				
	/br/	*Bilbray*	→	[pi.pu.ʐʷei]				
	/tr/	*Contra*	→	[kʰɑŋ.tʰɤ.la]				
	/dr/	*Andrew*	→	[an.tɤ.lu]				
	/kr/	*Bancroft*	→	[pan.kʰɤ.lɑu.fu]				
	/gr/	*McGraw-Hill*	→	[mai.kɤ.lʷo.ɕi.ər]				
ii.	Before /l/							
	/pl/	*Kaplan*	→	[kʰa.pʰu.lan]				
	/bl/	*Dublin*	→	[tu.pʷo.lin]				
	/tl/	*Atlanta*	→	[ja.tʰɤ.lan.ta]		*Chipotle*	→	[tɕʰi.pʷo__li]
	/dl/	*Bradley*	→	[pu.lai.tɤ.li]		*Ridley*	→	[lei__li]
	/kl/	*Berkley*	→	[pʷo.kʰɤ.lai]				
	/gl/	*Douglas*	→	[tɑu.kɤ.la.sẓ]				

Since pre-liquid stops are released in English, Kang's (2003) release-to-vowel epenthesis hypothesis correctly predicts that they are preserved even in word-medial position.[36]

5.1.2 The Problem

While the loanword data and the transliteration data shown in 4.3.2 demonstrate that non-prevocalic released and unreleased stops tend to be preserved via vowel epenthesis and delete, respectively, a problem arises when we consider the adaptation of word-medial stops preceding a homorganic nasal or another stop. In this subsection I will show the data and why this is a problem for the models classified into the hybrid approach.

5.1.2.1 Word-Medial Unreleased Stops

Word-medial stops preceding a homorganic nasal or another stop are normally unreleased in English (Ladefoged 1975: 45, 49; MacKay 1978: 166; Ladefoged 1993: 56; Brinton 2000: 30), so we expect them to delete. However, my data show that they are preserved at a very high rate. The data are given in Table 5.3 with all the 17 mappings in my corpus shown in (21).[37] The transliteration data and loanword data on the deletion vs. preservation of word-final stops are repeated here for comparison.

---

[36] Word-medial pre-glide stops are also released in English and tend to be preserved in English loanwords in SC (e.g. *McGuire* → [mai.kʰʷei.ər]). These stops are not considered here because an SC syllable can start with a stop+high vowel+non-high vowel sequence (e.g. UR /kʰuan/ → [kʰʷan]). That is, the reason that they tend to be preserved is not because they are released but because they can be syllabified. However, these stops may be treated as if they were in syllable-final position (e.g. *Edward* → [ai.tɤ.xʷa]). When this happens, usually a morpheme boundary between the stop and the following glide is assumed by the adapters (e.g. deletion: *Eastwood* → [i.sz̩__wei.tʰɤ] *[i.sz̩.tʰʷei.tʰɤ]; preservation with vowel epenthesis: *Gateway* → [kai.tʰɤ.wei] *[kai.tʰʷei], *Goodwin* → [ku.tɤ.wən] *[ku.tʷən]). The adaptation of some obstruent+glide sequences (e.g. *swan*, *Guam*) in English loanwords in SC will be discussed and analyzed in Chapter Six.

[37] The word-medial /k/ in *Kennebunkport* (→ [kʰən.nʲi.pɑŋ.kʰɤ.kɑŋ]) and the /d/'s in *Lundberg* (→[lʷuŋ.tɤ.pɑu]) and *Goldberg* (→[kɤ.tɤ.pɑu]) are excluded from count because the *-port* and *-berg* components of these words are treated as separate morphemes and adapted based on their meanings: *-port* is mapped to the character 港 ([kɑŋ] with Tone 3), which means "port", and *-berg* is mapped to the character 堡 ([pɑu] with Tone 3), which means "fort". The /k/ and /d/'s in these words should be considered as if they were word-final.

Table 5.3 Frequencies of deletion and preservation of word-medial stops preceding a homorganic nasal or another stop in English loanwords in SC

English word-medial unreleased stop			Korean, SM, and Cantonese unreleased stop			English word-final variably released stop		
del.	pres.	sum	del.	pres.	sum	del.	pres.	sum
9	8	17	79	7	86	72	117	189
52.94%	47.06%	100%	91.86%	8.14%	100%	38.10%	61.90%	100%

(21) Deletion and preservation of word-medial stops preceding a homorganic nasal or another stop in English loanwords in SC

Deletion / Preservation

i. labial stop+stop

Hepburn → [xɤ__pən] *Hopkins* → [xʷo.pʰu.tɕin.sz̩]

ii. coronal stop+stop or homorganic nasal

Britney → [pu.lan__ni] *Putnam* → [pʰu.tʰɤ.nan]

Watkins → [xʷa__tɕin.sz̩] *Whitney* → [xʷei.tʰɤ.ni]

Putnam → [pʰu__nan] *Gardner* → [tɕa.tɤ.na]

Whitney → [xʷei__ni]

Lundberg → [luŋ__pʷo.kɤ]

iii. dorsal stop+stop

Victoria → [wei__tʷo.li.ja] *Rector* → [lei.kʰɤ.tʰɤ]

McDaniel → [mai__tan.ni.ər] *Specter* → [ʂʐ̩.pʰai.kʰɤ.tʰɤ]

McDonald → [mai__tʰɑŋ.na] *Stockton* → [ʂʐ̩.tʰʷo.kʰɤ.tʷən]

McBride → [mai.kʰɤ.pu.lai.tɤ]

As we can see, these stops are preserved 47% of the time, almost as high as the 53% rate at which they are deleted. Note that the difference in deletion rate between these stops and the word-final variably released stops is not significant ($\chi^2 = 1.44$, $p = .23$), whereas the difference in deletion rate between these stops and the unreleased stops in the transliteration data is (χ^2 = 17.28, p = .000032).

5.1.2.2 Challenges to the Hybrid Approach

Whether perception is claimed to be always faithful or not, the loanword models classified into the hybrid approach are compatible with the transliteration data, the loanword data on the adaptation of word-final variably released stops, and the data shown in Tables 5.1 and 5.2. For example, if unreleased stops are not perceived, then they are not included in the input to the production grammar, and deletion is what we see in the output; if they are perceived and included in the input to the production grammar, then their deletion is due to the constraint ranking in the production grammar (see below for the relevant constraints and rankings).

None of the models in the hybrid approach, however, can account for the data in Table 5.3. If the deletion is due to unfaithful perception, we wonder why preservation occurs frequently. If the perception is faithful and the deletion takes place in the production grammar, to explain the preservation, we will need to rank some of the constraints differently, resulting in a constraint ranking paradox.

5.1.3 The Proposal

To solve this problem, I propose that, despite being considered unreleased, word-medial stops preceding a homorganic nasal or another stop in English may be perceived as released by SC listeners. This explains why these stops are treated as if they were word-final.

5.1.3.1 Articulatorily Released

Henderson and Repp (1981) argued that, in sequences of two stops differing in place of articulation in English (e.g. *ca<u>ct</u>us*, *a<u>ct</u>*, *bi<u>g</u> <u>d</u>og*), the first stops must be released before the second stops; “otherwise, the second stop would be produced with an incorrect or dual place of articulation" (pp. 72). This argument is confirmed acoustically. In a series of production experiments in which the subjects produced two-stop sequences in word-internal position, Repp (1980, 1982) and Henderson and Repp (1981) found that the first stops showed release bursts in the oscillograms frequently. Kim (1998) reported a very similar finding (see below). Henderson and Repp (1981) carried out two follow-up perceptual experiments using some of the tokens collected in the production experiment as stimuli. In the stimuli, the first stops clearly showed release bursts in the oscillograms. The results of the perceptual experiments revealed that the release bursts were difficult to detect auditorily, even for subjects who had had phonetic training. Henderson and Repp (1981) concluded that the term "unreleased" commonly used in phonetic literature to describe the first stops in non-homorganic two-stop sequences in English must be interpreted in a perceptual rather than articulatory or acoustic sense. That is, the first stops in these sequences have been thought to be unreleased because the release bursts are difficult to detect by ear rather than being entirely absent acoustically or during articulation.

The stimuli used in Henderson and Repp's (1981) perceptual experiments were drawn from the utterances of one of the subjects participating in a previous production experiment, who were asked to utter at a normal conversational speed. According to the authors, the release bursts of the first stops in the stimuli, despite being clearly present in the oscillograms, were rather weak (i.e. of lower amplitude and shorter duration), and this explained why they were difficult to detect by ear. The authors also found that the release bursts usually occurred during the closures of the second stops, implying that the two stops usually overlapped (cf. Browman and Goldstein 1992; Ladefoged 1993; Zsiga 2003).

5.1.3.2 Features of Foreigner Talk Speech

Gass and Selinker (2008: 306–307) reported that patterns present in foreigner talk speech are usually not found in conversations between native speakers, and that the most common traits of foreigner talk speech include slow speaking rate, loud speech, and long pauses (see also Ellis 1994: 255–256). In addition, on the basis of Hatch's (1983) research, Gass and Selinker (2008: 306) pointed out that slow speaking rate results in clearer articulation, releasing of final stops, and fewer contractions. Thus, it is plausible that, in foreigner talk speech, the stops in non-homorganic two-stop sequences tend to not overlap or overlap to a lesser extent, and the first stops tend to be longer in duration and have higher amplitude. As a result, we expect (i) the first stops to show release bursts more frequently and expect (ii) the release bursts to be auditorily more salient and therefore more likely to be heard by the listeners.

Acoustic evidence confirms these expectations. First, in two production experiments where the subjects produced vowel+stop+stop+vowel nonsense English words in isolation (i.e. read a word list), Repp (1980, 1982) found release bursts belonging to the first stops in the oscilograms almost 100% of the time. These results are in contrast with the results of Henderson and Repp's (1981) production experiment. In Henderson and Repp's (1981) production experiment, the subjects produced the target words embedded in a frame sentence at a conversational speed, and the authors observed release bursts belonging to the first stops in the oscilograms only 58% of the time.

Second, Kim (1998) investigated the acoustic manifestations of word-final voiceless stops followed by another stop across a word boundary in English and Korean. She found three types of manifestations in English but only two in Korean. Her findings are shown in Table 5.4.

Table 5.4 Types of acoustic manifestations of word-final voiceless stops followed by another stop across a word boundary in English and Korean and the frequencies of occurrence (in percent) (Kim 1998: 352, 359)

	speaking rate	Type A	Type B	Type C
		aspiration with an oral burst	a brief low-amplitude oral burst after silence	silence corresponding to oral closure
English	fast	21.4	55.4	23.2
	normal	42.9	44.6	12.5
	slow	89.3	7.1	3.6
Korean	normal	0	83	17

Consider first the three manifestations in English. The Korean data will be discussed in the next subsection. The Type-C manifestation is simply silence. While both the Type-A and Type-B manifestations display a release burst, according to Henderson and Repp 1981, the Type-B manifestation is difficult to detect by ear, as it shows low amplitude and short duration. When the speaking rate goes from fast to normal to slow, fewer and fewer occurrences of the Type-C manifestation are observed, which means that there are more and more occurrences of the Type-A and Type-B manifestations combined. This fact confirms my first expectation.

Kim (1998) observed that the Type-A manifestation is extremely similar to the acoustic manifestation of prevocalic stops. For example, she reported that all the prevocalic stops in the words *c̲at*, *sk̲y*, and *d̲ot* show in the corresponding energy shapes a near-90° rise followed by a high-energy plateau (pp. 363–364), and that a very similar energy curve is found in the energy shape for the /t/ in the word *kit* when it is realized with aspiration (Type A). Kim (1998: 364) claimed that this plateau-like energy shape is attributed to 'the explosion of compressed air when oral closure is removed.' Kim's (1998) observations suggest that word-final stops with the Type-A manifestation must be perceived as released. In addition, since the releasing of prevocalic stops is auditorily salient, the releasing of word-final stops displaying a similar acoustic manifestation should be auditorily salient as well. As we can see in Table 5.4, the slower the speaking rate, the more occurrences of the Type-A manifestation and the fewer occurrences of the Type-B manifestation. This confirms my second expectation that, in slow and more careful speech such as foreigner talk speech, the release bursts of the first stops in non-homorganic two-stop sequences are auditorily more salient and therefore more likely to be heard by the listeners.

5.1.3.3 The Transliteration Data Revisited

Table 5.4 shows that although word-final stops can be realized with aspiration (Type A) in English, this manifestation is not found in their Korean counterparts. This fact is consistent with the general agreement in the literature that word-final stops are either released or unreleased in English but always unreleased in Korean (at least perceptually). In addition, it suggests that word-final stops in Korean, which only show the Type-B and Type-C manifestations, be auditorily non-salient when produced at a normal speaking rate.

While the Korean data in Table 5.4 are based on tokens produced at a normal speaking rate, Kim-Renaud (1974: 116) claimed that 'even in a deliberately slow and careful speech, the releasing of obstruents in syllable final position does not occur' (cited in Kim 1998: 350). Following Kim-Renaud (1974), I assume that syllable-final stops are not realized with aspiration (Type A) in slow or careful speech in Korean. This means that syllable-final stops in Korean are never auditorily salient. Furthermore, we cannot exclude the possibility that the Korean-to-SC transliteration in the data shown in 4.3.2 were conducted by Korean-SC balanced bilinguals (there are many Korean-SC balanced bilinguals in China). Thus, the Korean input to the transliteration may be based on utterances produced at a normal, rather than a slow, speaking speed.

Acoustic or perceptual data on the releasing of syllable-final stops in Cantonese or Southern Min are unavailable to me. However, it has been repeatedly stated in the literature that syllable-final stops in these two languages are strictly unreleased (e.g. for Cantonese, see Cheung

1986, Bauer and Benedict 1997, Khouw and Ciocca 2006; for Taiwan Southern Min, see Chung 1996). Given Henderson and Repp's (1981) argument that the term "unreleased" used in the literature refers to auditorily undetectable release bursts or no release bursts, I assume that the release bursts of syllable-final stops in Cantonese and Southern Min are either acoustically absent or too weak to be audible. In addition, as a native Southern Min speaker, release bursts of syllable-final stops in Southern Min are always undetectable to me. I conclude that, like in Korean, syllable-final stops in Cantonese and Southern Min are always auditorily non-salient, which leads to deletion in transliteration.

5.1.4 Summary

The data in my corpus generally support the release-to-vowel epenthesis hypothesis, suggesting that unreleased stops be deleted. However, word-medial stops preceding a homorganic nasal or another stop, which are normally considered unreleased in English, are found preserved frequently. Research has shown that the first stops in word-internal non-homorganic two-stop sequences in English are in fact released articutorily, and that while the release bursts are often acoustically absent and, when present, usually auditorily undetectable when the utterances containing these sequences are produced at a conversational speed, in foreigner talk speech, which is more careful and slower, they are acoustically present more often and easier to be detected by the listeners. Thus, I propose that the first stops in these sequences are perceived as released by SC speakers sometimes, explaining my data that they may be preserved in adapting English words into SC.[38]

[38] On the basis of Gass and Selinker's (2008) and Hatch's (1983) observations regarding the features of foreigner talk speech, I assume that word-medial stops preceding a homorganic oral stop or nasal in English may be perceived as released by SC speakers as well. There are six instances of a word-medial stop preceding a homorganic nasal in my corpus. Three are deleted and three are preserved. There is only one instance of a word-medial stop preceding a homorganic stop in my corpus, and it is deleted.

5.2 Word-Final Unreleased Stops are Realized with Aspiration when Preserved via Epenthesis in Adapting English Words into Korean and SC

In this section I present the second problem for the hybrid approach. The big picture the Korean and SC loanword data project is that every sound in the source words is in the input to the adaptation. In addition, it is reasonable to assume that the borrowers perceive every feature that is contrastive in their native languages. For example, I assume that Korean- and SC-speaking listeners can easily distinguish aspirated and unaspirated stops in the English source words even though stop aspiration is a non-contrastive feature in English. With these assumptions in hand, the hybrid approach faces a problem when it comes to the adaptation of word-final stops in English loanwords in Korean and SC. The Korean and SC loanword data show that when these stops are preserved via vowel epenthesis, they are always realized with aspiration. After elaborating the assumptions in 5.2.1 and the challenges to the hybrid approach in 5.2.1, a solution is proposed in 5.2.3.

5.2.1 Faithful Perception

To start this section, I would like to emphasize that the deletion of word-final stops in English words borrowed into SC should not be attributed to unfaithful perception, as they are always preserved if the source words are monosyllabic and contain no other sounds that violate the SC syllable constraints (e.g. *Matt* → [mai.tʰɤ], *Pink* → [pʰiŋ.kʰɤ], *Ed* → [ai.tɤ], *Doug* →[tɑu.kɤ]) (see 3.1). Because the adaptation of word-medial stops preceding a homorganic nasal or another oral stop shows an identical pattern (i.e. there is no significant difference between the two sets of data), it is plausible that the SC adapters treat them as if they were word-final, suggesting that deletion of them is not due to unfaithful perception either. Segment deletion rarely occurs in loanwords in Korean (Kang 2003: 238), indicating that almost every sound in the source words is perceived.

The fact that even auditorily non-salient sounds in the English source words are perceived by the SC and Korean adapters serves as a strong argument against the models in the hybrid approach that acknowledge unfaithful perception (Silverman 1992; Yip 1993; Kenstowicz 2003). The models proposed by Yip (2002, 2006), Kang (2003), Kenstowicz (2005), and Steriade (2001, 2009), and many others, on the other hand, basically assume faithful perception. These scholars claim that the relative perceptual similarity between the input and potential output forms is an important factor in determining which potential output form is selected as the output. This claim entails that the input to the production grammar of the host language contains non-contrastive phonetic details.

To formalize the Korean adaptation of word-final postvocalic stops in English words, Kang (2003: 252– 253) proposed that these stops are marked as released or unreleased in the input to the Korean production grammar and that the constraint ranking BESIMILAR-[release] >> DEP-V determines whether vowel epenthesis applies (note that [release] is not a phonological feature in English or Korean). An example is given in the tableaux in (22):

(22) E. *jee**p*** → K. [tsi**pʰ˹ɨ**] ~ [tsi**p˺**] (Kang 2003: 253) (BESIMILAR-[F] means that sounds in correspondence should be similar regarding the feature)[39]

i. *jee**p**˹* (released)	BESIMILAR-[release]	DEP-V
→ a. [tsi**pʰ˹ɨ**]		*
b. [tsi**p˺**]	*!	
ii. *jee**p**˺* (unreleased)	BESIMILAR-[release]	DEP-V
a. [tsi**pʰ˹ɨ**]	*!	*
→ b. [tsi**p˺**]		

[39] Kang (2003) used BeSimilar-[release] instead of the traditional IDENT-[released] to highlight her point that the mappings are based on perceptual similarity rather than distinctive features.

Given the constraint ranking BESIMILAR-[release] >> DEP-V, when the word-final /p/ in the input is marked as released, it surfaces with vowel epenthesis (22 i); when it is marked as unreleased, it surfaces unchanged (22 ii). Another output candidate that is worth considering is [tsi**p˹ɨ**], where the [p] is released but unaspirated. I will include this candidate when discussing a problem for these models.

Following Kang (2003), I propose that word-final stops in borrowed English words are also marked as released or unreleased in the input to the SC production grammar. Consider the tableau below. For the sake of consistency, I use Kang's (2003) terminology BESIMILAR here.

(23) E. *Rober**t*** → SC [lʷo.pʷo.**tʰ˹ɤ**] ~ [lʷo.pʷo__]

i. *Rober**t**˹* (released)	MAX-rel. stop	DEP-V	MAX-unrel. stop	BESIMILAR-[rel]
→ a. [lʷo.pʷo.**tʰ˹ɤ**]		*		
b. [lʷo.pʷo__]	*!			
ii. *Robert**˺* (unreleased)**	**MAX-rel. stop**	**DEP-V**	**MAX-unrel. stop**	**BESIMILAR-[rel]**
a. [lʷo.pʷo.**tʰ˹ɤ**]		*		*!
→ b. [lʷo.pʷo__]			*	

When the word-final /t/ in the input is marked as released (23 i), the grammar, which says MAX-released stop >> DEP-V, selects Candidate (23 i a) as the output. When it is marked as unreleased (23 ii), MAX-released stop is irrelevant, and the violation of BESIMILAR-[release] incurred by Candidate (23 ii a) makes it lose to its competitor Candidate (23 ii b).[40]

The constraints and ranking in (23) also correctly predict the output when the input contains a word-medial stop that precedes a nasal or another oral stop, as the SC adapters treat a stop in these contexts as if it were word-final.

[40] We need to distinguish the two constraints MAX-released stop and MAX-unreleased stop. If we use the more general constraint MAX-stop, in order for the candidates (23 i a) and (23 ii b) to be selected as the output, it must outrank DEP-V in (23 i) but must not outrank DEP-V in (23 ii), presenting a constraint ranking paradox.

5.2.2 The Problem

Referencing the relative perceptual similarity between the input and potential output forms accounts for the variable realizations of word-final stops in English loanwords in Korean and SC. However, we encounter a problem when considering the fact that word-final voiceless stops in the source words are always mapped to aspirated stops when preserved via epenthesis. Specifically, if the stops are marked as unreleased in the inputs to the production grammars, then the preferred candidates, aspirated stops, always incur one more violation than the unaspirated candidates.

5.2.2.1 Korean Data

To see the problem in Korean, consider the tableau in (24):

(24) E. *min**t*** → K. [min.**tʰ˹ɨ**] (Henceforth, '←' means that the candidate is the actual word in the data but not selected by the constraint ranking.)

i. *min**tʰ˹*** (released)	NoComp	BeSimilar-[rel]	Dep-V	BeSimilar-[asp]
a. [min**t˺**]	*!	*		*
→ b. [min.**tʰ˹ɨ**]			*	
c. [min.**t˹ɨ**]			*	*!
ii. *min**t˺*** (unreleased)	NoComp	BeSimilar-[rel]	Dep-V	BeSimilar-[asp]
a. [min**t˺**]	*!			
← b. [min.**tʰ˹ɨ**]		*	*	*!
→ c. [min.**t˹ɨ**]		*	*	

Kim's (1998) acoustic report indicates that if English word-final voiceless stops are perceived as released, they are also perceived as aspirated (see also Flege 1989; Flege and Wang 1989; Ladefoged 1993, Kang 2003). This means that, in the English input to the production grammar of the host language, word-final voiceless stops that are marked as released are also marked as aspirated. In (24 i), the word-final /t/ in the input is marked as released and aspirated, so Candidate (24 i b), which satisfies both BeSimilar-[release] and BeSimilar-[aspiration], is selected over Candidate (24 i c), which satisfies BeSimilar-[release] but violates BeSimilar-

[aspiration]. In (24 ii), the word-final /t/ in the input is marked as unreleased. While both Candidates (24 ii b) and (24 ii c) violate BESIMILAR-[release], as an epenthetic vowel is found in these two candidates, Candidate (24 ii b) additionally violates BESIMILAR-[aspiration]. The violation of BESIMILAR-[aspiration] incurred by Candidate (24 ii b) renders Candidate (24 ii c) to be selected as the output.

In Korean, inter-sonorant voiceless unaspirated plain stops are phonetically voiced (Kang 2003: 247–248), so Candidates (24 i c) and (24 ii c) can be transcribed as [min.d˺ɨ], indicating that these two candidates also violate BESIMILAR-[voice]. If BESIMILAR-[voice] is ranked above BESIMILAR-[aspiration], then Candidate (24 ii b) would be selected as the output and the problem does not exist. This is illustrated in the tableau below, which is adapted from (24 ii).

(25) E. *min**t̚*** → K. [min.**tʰ˺ɨ**] (adapted from (24))

*min**t̚*** (unreleased)	NOCOMP	BESIMILAR-[rel]	DEP-V	BESIMILAR-[voi]	BESIMILAR-[asp]
a. [min**t̚**]	*!				
→ b. [min.**tʰ˺ɨ**]		*	*		*
c. [min.**d˺ɨ**]		*	*	*!	

Given the ranking BESIMILAR-[voi] >> BESIMILAR-[asp], while both Candidates (25 b) and (25 c) have the same counts of violation, Candidate (25 b) is selected as the output because it satisfies BESIMILAR-[voi] but its competitor Candidate (25 c) does not.

However, as I will discuss below, perceptual studies indicate that Korean learners of English have great difficulty distinguishing between word-final post-consonantal unreleased voiced and voiceless stops in English, suggesting that the adapters actually do not hear the differences between the two English forms /mɪnt̚/ and /mɪnd̚/. Because BESIMILAR constraints evaluate perceptual similarity, Candidate (25 c) does not violate BESIMILAR-[voice], and the problem remains. This argument is illustrated in the tableau below, which is adapted from the tableau in (25):

(26) E. *min**t̚*** (= *min**d̚*** perceptually for Korean listeners) → K. [min.**tʰ˺ɨ**]

*mɪn**t̚*** (unreleased) (perceptually confusable with /mɪn**d̚**/)	NOCOMP	BESIMILAR-[rel]	DEP-V	BESIMILAR-[voi]	BESIMILAR-[asp]
a. [min**t̚**]	*!				
← b. [min.**tʰ˺ɨ**]		*	*		*!
→ c. [min.**d˺ɨ**]		*	*		

In (26), the input word *min**t̚*** is perceptually confusable with /mɪnd̚/, so Candidate (26 c) does not violate BESIMILAR-[voice]. As a result, Candidate (26 c) is still selected over Candidate (26 b) as the output.

5.2.2.2 SC Data

A similar problem is observed in the SC adaptation of English monosyllabic words containing word-final voiceless stops, which is illustrated in the tableaux in (27):

(27) E. *Ka**t**e* → SC [kʰai.**tʰ˺ɤ**] ([σσ]MINWD means that the minimum word size is two syllables)

i. *Ka**tʰ˺**e* (rel.)	[σσ]MINWD	MAX-rel. stop	DEP-V	MAX-unrel. stop	BESIMILAR-[rel]	BESIMILAR-[asp]
a. [kʰai__]	*!	*				
b. [kʰai.i]		*!	*			
→ c. [kʰai.**tʰ˺ɤ**]			*			
d. [kʰai.**t˺ɤ**]			*			*!
ii. *Ka**t̚**e* (unrel.)	[σσ]MINWD	MAX-rel. stop	DEP-V	MAX-unrel. stop	BESIMILAR-[rel]	BESIMILAR-[asp]
a. [kʰai__]	*!			*		
b. [kʰai.i]			*	*!		
← c. [kʰai.**tʰ˺ɤ**]			*		*	*!
→ d. [kʰai.**t˺ɤ**]			*		*	

Like the tableaux in (24), when the word-final stop in the input is marked as released (27 i), we obtain the correct output, but when marked as unreleased (27 ii), the aspirated candidate (27 ii b), which is the actual SC form in my corpus, loses to the unaspirated candidate (27 ii c), as the former violates BESIMILAR-[aspiration] but the latter does not.[41]

5.2.3 The Proposal

I propose a solution: In adapting English words into Korean and SC, every word-final voiceless stop in the source words that is preserved via vowel epenthesis is marked as released (and aspirated) in the input to the Korean or SC production grammar even if they are perceived as unreleased (e.g. E. *min**t̚*** (perception by K. listeners) → E. *min**t**ʰ⌜* (input to K. grammar) → K. [min.**tʰɨ**] (K. output)).

Yip (2006: 952) pointed out that, 'we need to distinguish between the perception of the presence of a segment or property, the perception of a distinction of that segment and another, and the perception of the basis of that distinction.' As argued above, the SC and Korean adapters must be able to perceive the presence of unreleased stops in borrowed English words. In addition, because voiced stops are always mapped to unaspirated voiceless (plain) stops, the fact that word-final unreleased voiceless stops always surface with aspiration when preserved by vowel epenthesis strongly suggests that the adapters know a distinction between them and their voiced counterparts and want to keep this distinction. I provide some mappings below which can illustrate this idea.

(28) Mappings illustrating why word-final unreleased voiceless stops are not realized with aspiration in adapting English words into SC and Korean

voiceless	→	aspirated	voiced	→	unaspirated

i. SC adaptation of English words

*Ka**t̚**e*	→	[kʰai.**tʰ⌜ɤ**] *[kʰai.**t⌜ɤ**]	*Ca**d̚**e*	→	[kʰai.**tɤ**]

[41] In the tableaux in (24) and (27), how the constraint BESIMILAR-[aspiration] is ranked with respect to other constraints does not affect the results. I simply place it at the bottom of the rankings to make the tableaux easier to read.

ii. Korean adaptation of English words

*min**t̚*** → [min.**tʰ˺ɨ**] *[min.**t˺ɨ**] *poun**d̚*** → [pʰɑ.un.**tɨ**][42]

The unreleased /t/ in *Kat̚e* and the unreleased /d/ in *Cad̚e* have to be preserved via vowel epenthesis due to the SC minimum word requirement. The unreleased /t/ in *mint̚* and the unreleased /d/ in *pound̚* also have to be preserved via vowel epenthesis because (i) segments do not delete in Korean loanword adaptation (Kang 2003) and (ii) Korean does not allow complex coda. The two unreleased /t/'s are mapped to aspirated [t]'s, and the two unreleased /d/'s are mapped to unaspirated (plain) [t]'s. I propose that the two unreleased /t/'s are not mapped to their perceptually closest matches unaspirated (plain) [t]'s because the adapters have the knowledge that /t/ is different from /d/ in English and would like to maintain this distinction in the adaptation.

Interestingly, perceptual studies have shown that SC and Korean learners of English have great difficulty distinguishing between word-final unreleased voiced and voiceless stops in English.

5.2.3.1 Previous Perceptual Studies

Flege (1989) conducted an identification experiment in which native English speakers and SC learners of English listened to English monosyllabic words ending in /t/ or /d/ and determined the final stops' identities. Four types of stimuli were used in this experiment: (i) the "unedited" stimuli, where all the final stops had an audible release burst, (ii) the "voicing-removed" stimuli, where the voicing during the closure intervals of the voiced final stops was removed, (iii) the "burst-removed" stimuli, where the release bursts of the final stops were removed, and (iv) the "voicing+burst-removed" stimuli, where the closure voicing and release bursts of the final stops were removed. The results are summarized in Table 5.5:

[42] This mapping is found in Tranter (2000).

Table 5.5 Frequencies of correct identification of word-final postvocalic /t/ and /d/ in English monosyllabic words by SC learners of English, native English adults, and native English children in four stimuli conditions (Flege 1989)

	unedited	v-removed	b-removed	v+b-removed
SC learners of English	94%	93%	64%	64%
English adults	100%	99%	99%	98%
English children	97%	98%	96%	96%

A few observations are relevant here. First, the near-perfect identification rates for the two English groups in the "voicing+burst-removed" stimuli condition confirm that English speakers primarily rely on the duration of the preceding vowels to discriminate the voicing contrasts in final stops. Second, as long as the released bursts were present (in the "unedited" and "voicing-removed" stimuli conditions), the SC subjects performed almost as well as the English subjects. Third, as long as the release bursts were removed (in the "burst-removed" and "voicing+burst-removed" stimuli conditions), the SC subjects' performance dropped to a very low level. Flege and Wang (1989) carried out a very similar experiment, in which SC subjects listened to stimuli of the "voicing+burst-removed" type only. The authors found that the SC subjects performed poorly, correctly identifying the final stops as voiced or voiceless only 59% of the time. These experiment results clearly show that, to discriminate the voicing contrasts in word-final postvocalic stops in English, native speakers rely on the durational differences in the preceding vowels while SC learners of English rely on the release bursts and ignore other cues.

As for Korean learners' perception of the voicing contrasts in word-final stops in English, perceptual studies show that they also rely on the release bursts. In an identification task where native English speakers and Korean learners of English listened to English monosyllabic words ending in an obstruent and decided whether the final obstruents were voiced or voiceless, Chang and Idsardi (2001) found that the Korean subjects correctly discriminated the voicing contrasts in stops 81.67% of the time (90.87% for the English subjects). One of the reasons for the high discrimination rate is that Chang and Idsardi (2001) did not control the factor that the final stops in their stimuli might be released (see Chang and Idsardi 2001 and Flege and Wang 1989 for other reasons). If some of the stimuli contained released final stops, then the high rate is unsurprising. Notice that if the voicing contrasts in unreleased word-final postvocalic stops are not easy to discriminate, then the voicing contrasts in unreleased word-final stops following a consonant are conceivably more difficult to discriminate, as the vowel-duration cue is unavailable.

Put together, the SC and Korean loanword data and experiment results presented above demonstrate that while unreleased voiced and voiceless stops in word-final postvocalic or word-final post-consonantal position in English are difficult to distinguish, the SC and Korean adapters obviously know a distinction between them and manage to maintain the distinction during the adaptation.

5.2.3.2 Source of Voicing Contrasts in Word-Final Stops

What is the source of the distinction for the SC and Korean adapters then? I propose that it is the final release bursts contained in the same words that are pronounced with the final stops released. The adapters must have heard two versions of every learned stop-ending English word. In one version, the final stops are released; in the other version, the final stops are unreleased (e.g. *Ka**tʰ⌜**e* ~ *Ka**t˺**e*; *min**tʰ⌜*** ~ *min**t˺***). These two versions of learned stop-ending English words are stored in the adapters' long-term memory (see Calabrese 2009 for discussion of how the long-term memory functions in perception). When adapting English words ending in voiceless stops into their native language, if the final voiceless stops are marked as unreleased in the input to the production grammar and vowel epenthesis must apply to satisfy some native requirement, the adapters know that, although substituting the native unaspirated (plain) stops for the final unreleased voiceless stops is the best choice in terms of perceptual similarity, the native unaspirated (plain) stops have been used for voiced stops in other borrowed English words. The adapters also know that a win-win deal is available. In this deal, they can avoid using the native unaspirated (plain) stops for both the word-final unreleased voiceless stops in the source words and voiced stops in other borrowed English words, and at the same time they can avoid mapping the word-final unreleased voiceless stops in the source words to the native aspirated stops, which are perceptually disfavored (as shown in the tableaux in (24 ii) and (27 ii)). The way of reaching this deal is to use the "released" version of the words as the input. So the adapters go back to their long-term memory and retrieve it, and once the "released" version replaces the "unreleased" version as the input, the final voiceless stops in the input, now marked as released and aspirated, can surface unchanged.

This proposal is in contrast with the process proposed by Peperkamp and her colleagues (e.g. Peperkamp, Vendelin, and Nakamura 2008), which can be referred to as an "online" loanword adaptation process (see 2.1). In an online loanword adaptation process, unfaithful perception may be the cause, and Peperkamp and her colleagues claimed that it is the only cause.

The strongest argument for unfaithful perception being the cause of loanword adaptation is probably the "perceptual illusion" phenomenon reported by Dupoux, Kakehi, Hirose, Pallier, and Mehler (1999) and Dupoux, Parloto, Frota, Hirose and Peperkamp (2011). These scholars showed that when presented with stimuli containing word-internal two-stop sequences, their Japanese-speaking subjects, whether monolinguals or learners of English, perceptually epenthesized a vowel inside the sequences (/ebzo/ → [ebuzo]) (i.e. misheard /ebzo/ as [ebuzo]). This phenomenon does not serve as an argument against my proposal, as the test word /ebzo/ used in their experiments is a nonce word. Because it was the first time that their Japanese subjects heard the word, suggesting that there was no information regarding this word in their long-term memory they could use, an online adaptation process must occur, which resulted in mishearing. The fact that word-final voiceless stops are always realized with aspiration when preserved via vowel epenthesis even if they are perceived as unreleased in adapting English words into SC and Korean (e.g. E. *Ka**t**e* → SC [kʰai.**tʰ˺ɤ**] *[kʰai.**t˺ɤ**], E. *min**t*** → K. [min.**tʰ˺ɨ**] *[min.**t˺ɨ**]) demonstrates that the adaptation is not an online process.

5.3 Summary of Chapter Five

My data generally support the release-to-vowel epenthesis hypothesis, suggesting that unreleased stops be deleted. However, word-medial stops preceding a homorganic nasal or another stop, which are normally considered unreleased in English, are found preserved frequently. I have argued that these stops may be perceived as released by the adapters. This is because, in foreigner talk speech, which is slower and more careful, the release bursts of these stops are more likely to be acoustically present and, when present, perceptually more salient.

I have also shown that, in adapting English words into Korean and SC, word-final unreleased voiceless stops that must be preserved via vowel epenthesis to satisfy some native requirement are always realized with aspiration even though the native unaspirated (plain) voiceless stops are perceptually the best matches. I propose that the adapters retrieve the same words containing the released (and aspirated) version of the stops from their long-term memory and substitute them for the original words. This way, the stops in the input, which are released and aspirated now, can surface unchanged, and the adapters can avoid the situation in which the unreleased stops in the original words and voiced stops and in other borrowed English words are both represented by the same native sounds.

The data and arguments presented in this chapter suggest that when modeling loanword adaptation, we need to take into consideration the adapters' interlanguages (cf. Kenstowicz 2005; Yip 2006) and the possibility that the adapters have some control over the input.

CHAPTER SIX

ADAPTATION OF SIBILANT+/W/ AND VELAR STOP+/W/ SEQUENCES IN ENGLISH LOANWORDS IN SC

I have shown in 4.1 that when English words are borrowed into SC, /v/ is not uniformly adapted but adapted as a function of the /v/’s position relative to a vowel. In this chapter, I will present another divergent-repair case found in my corpus. In 6.1, I will show that while the sibilant in a sibilant+/w/ sequence is always realized as a separate syllable (e.g. E. /sw/ → SC [ʂʐ̩.w]), the two sounds in a velar stop+/w/ sequence are almost always merged as one single sound, with the /w/ realized as a secondary articulation of the velar stop (e.g. E. /kw/ → SC [k^{w}]). In 6.2, I adopt Kenstowicz’s (2005) proposal and formalize the adaptation patterns. My analyses are supported by articulatory evidence and the results of perceptual experiments, which I will provide in 6.3. 6.4 is the summary.

6.1 Divergent Repairs

In this section, I present the SC loanword data. As we will see, sibilant+/w/ sequences and velar stop+/w/ sequence in borrowed English words are adapted differently.

6.1.1 Adaptation of Sibilant+/w/ Sequences

In SC, a high vowel following a consonant and preceding a non-high vowel surfaces as a secondary articulation of the consonant (e.g. /piau/ → [p^{j}ɑu] “chart” with Tone 3, /lyə/ → [l$^{\text{ɥ}}$e] “to omit” with Tone 4, /ʂuəi/ → [ʂ^{w}ei] “water” with Tone 3) (Duanmu 2007). However, when English words containing a word-initial sibilant+/w/ sequence are borrowed into SC, the /w/ does not surface in the same fashion. The data in (29 i) show the adaptation of all the instances of the sequences in my corpus. The native derivation of comparable sequence is shown in (29 ii) for comparison.

(29) Adaptation of word-initial sibilant+/w/ sequences in English loanwords in SC and native derivation of comparable sequences in SC words

i.	English		SC[43]
	Swank	→	[ʂʐ̩.wɑŋ]
	Swartz	→	[ʂʐ̩.xʷa.t͡sz̩]
	Sweet	→	[ʂʐ̩.wei.tʰɤ]
	Swingle	→	[ʂʐ̩.wei.kɤ]
	Zwingli	→	[tsz̩.wən.lʲi]
	Schwartz	→	[ʂʐ̩.xʷa.t͡sz̩]
	Schwarz	→	[ʂʐ̩.wo.t͡sʰz̩]
	Schwarzenegger	→	[ʂʐ̩.wa.ɕin.kɤ]
ii.	SC		SC
	/suan/	→	[sʷan] "sour" with Tone 1
	/suə/	→	[sʷo] "to search" with Tone 1
	/suəi/	→	[sʷei] "to follow" with Tone 2
	/ʂuən/	→	[ʂʷən] "wink" with Tone 4
	/ʂuə/	→	[ʂʷo] "to speak" with Tone 1
	/tsuən/	→	[tsʷən] "trout" with Tone 1

In (29 i), every word-initial sibilant+/w/ sequence breaks up, with the sibilant realized as [ʂʐ̩] or [tsz̩] and the /w/ assigned to the onset of the following syllable. In (29 ii), the two sounds in a comparable sequence merge, with the /u/ surfacing as a secondary articulation of the preceding sibilant.

[43] As mentioned in 4.2.4.4, word-initial non-prevocalic sibilants in the source words are mostly mapped to the syllable [ʂʐ̩] rather than the more similar syllable [sz̩] due to a non-linguistic factor.

6.1.2 Adaptation of Velar Stop+/w/ Sequences

An obstruent+/w/ sequence does not always break up into two syllables in English loanwords in SC. My corpus shows that if the obstruent is a velar stop, the sequence is usually treated like a native sequence. The adaptation of all the instance of the velar stop+/w/ sequences in my corpus is shown in (30 i) and (30 ii), and the derivation of comparable sequences in SC words is shown in (30 iii) for comparison.

(30) Adaptation of velar stop+/w/ sequences in English loanwords in SC and native derivation of comparable sequences in SC words

i.	English		SC	
	Quantum	→	[kʰʷən.tʰəŋ]	*[kʰɤ.wən.tʰəŋ]
	Quarles	→	[kʰʷei.lei.sz̩]	*[kʰɤ.wei.lei.sz̩]
	Quincy	→	[kʰʷən.ʂʐ̩]	*[kʰɤ.wən.ʂʐ̩]
	Guantanamo	→	[kʷan.ta.na.mʷo]	*[kɤ.wan.ta.na.mʷo]
	Queen	→	[kʰʷən.ən]	*[kʰɤ.wən]
	Quinn	→	[kʰʷei.ən]	*[kʰɤ.wən]
	Cuomo	→	[kʷo.mʷo]	*[kɤ.wo.mʷo]
	Quayle	→	[kʰʷei.ər]	*[kʰɤ.wei.ər]
	Squibb	→	[ʂʐ̩.kʷei.pɑu]	*[ʂʐ̩.kɤ.wei.pɑu]
	McGuire	→	[ma.kʰʷei.ər]/[mai.kʰʷei.ər]	*[ma.kʰɤ.wei.ər]/*[mai.kʰɤ.wei.ər]
	Joaquin	→	[wa.kʰʷən]/[tɕa.kʰʷən]	*[wa.kʰɤ.wən]/*[tɕa.kʰɤ.wən]
ii.	English		SC	
	Gwyneth	→	[kɤ.ni.sz̩]	*[kɤ.wei.ni.sz̩]
	Aquinas	→	[ai.tɕʰi.na.sz̩]	*[ai. kʰɤ.wei.na.sz̩]

	Guarini	→	[ku.li.ni]	*[kɤ.wu.li.ni]
	Gwinnett	→	[ku.nei]	*[kɤ.wu.nei]
iii.	SC		SC	
	/kʰuan/	→	[kʰʷan]	"wide" with Tone 1
	/kʰua/	→	[kʰʷa]	"to praise" with Tone 1
	/kuan/	→	[kʷan]	"official" with Tone 1
	/kuə/	→	[kʷo]	"fruit" with Tone 3
	/kʰuəi/	→	[kʰʷei]	"sunflower" with Tone 2
	/kuai/	→	[kʷai]	"strange" with Tone 4

(30 i) shows that a velar stop+/w/ sequence is adapted as if it were a native sequence, with the /w/ merging into the preceding velar stop. (30 ii) shows the four instances in my corpus where the /w/ in a velar stop+/w/ sequence does not surface this way. As we can see, the /w/ is deleted or vocalized rather than being assigned to the onset of the following syllable. The adaptations in (30 i) and (30 ii) indicate that no matter how a velar stop+/w/ sequence is adapted, the two constituent sounds do not split into two syllables.

6.2 Formal Analyses: Two Sets of IDENT-[Feature] Constraints

To model the adaptation patterns, I adopt Kenstowicz's (2005) proposal that while a foreign input is subject to the same set of markedness constraints as a native input, it is evaluated by a different set of IDENT-[feature] constraints, specifically, a set of loan-sensitive output-to-output IDENT-[feature] constraints. Kenstowicz (2005) argued that, by default, a loan-sensitive IDENT-OO-[feature] constraint occupies the same place in the ranking as its input-to-output counterpart, but in some cases, a loan-sensitive IDENT-OO-[feature] constraint and its IO counterpart are ranked differently. Yip (2002, 2006) and Kang (2003) have similar proposals.

For example, Yip (2002, 2006) proposed a set of MIMIC constraints, which she argued are 'the OT instantiation of active loanword incorporation, and enforces faithfulness to the percept' (2006: 956). While all of the loan-sensitive constraints proposed in the literature are IDENT-[feature] in nature (e.g. MIMIC-length, Yip 2006; BESIMILAR-[voice], Kang 2003), I suggest that the set of loan-sensitive constraints include DEP and MAX constraints, as DEP and MAX constraints also demand identity between input and output. There is no reason that, in loanword adaptation, there is a force demanding replication of a feature (i.e. IDENT-OO-[feature]) whereas there is no force demanding replication of silence (i.e. DEP-OO-segment) and presence of a sound (i.e. MAX-OO-segment).

Because velar stop+/w/ sequences in the English source words are repaired the same way as velar stop+/u/ sequences in native words, the relevant OO faithfulness constraint must occupy the same place in the ranking as its IO counterpart. In contrast, because different strategies are used for repairing sibilant+/w/ sequences in the English source words and sibilant+/u/ sequences in native words, the relevant OO and IO faithfulness constraints must be ranked differently. This idea is illustrated in the tableaux in (43) and (44), with the relevant constraints shown in (31).

(31) Constraints

i. OK-σ: Only permissible syllables in SC are allowed. (Yip 1993)

ii. $*T^w$: No labialized coronals. (Bradley 2012)

iii. $*K^w$: No labialized dorsals. (Bradley 2012)

iv. DEP-IO-[ɤ]/velar stop __: Do not epenthesize [ɤ] after a velar stop in a native input.

v. DEP-EO-[ɤ]/velar stop __: Do not epenthesize [ɤ] after a velar stop in an English input.

vi. IDENT-IO-length-sibilant: A sibilant in a native input and its correspondent in the output have the same duration.[44]

vii. IDENT-EO-length-sibilant: A sibilant in an English input and its correspondent in the output have the same duration.

(32) SC /kuan/ → SC [kʷan] vs. E. *Guantanamo* → SC [kʷan.ta.na.mʷo]

SC /kuan/	OK-σ	DEP-IO-[ɤ]/velar stop __	*Kʷ
→ a. [kʷan]			*
b. [kɤ.wan]		*!	
c. [kwan]	*!		
E. *Guantanamo*	OK-σ	DEP-EO-[ɤ]/velar stop __	*Kʷ
→ a. [kʷan.ta.na.mʷo]			*
b. [kɤ.wan.ta.na.mʷo]		*!	
c. [kwan.ta.na.mʷo]	*!		

(33) SC /ʂuaŋ/ → SC [ʂʷɑŋ] vs. E. *Swank* → SC [ʂʐ̩.wɑŋ]

SC /ʂuaŋ/	OK-σ	IDENT-IO-length-sibilant	*Tʷ
→ a. [ʂʷɑŋ]			*
b. [ʂʐ̩.wɑŋ]		*!	
c. [ʂwɑŋ]	*!		
E. *Swank*	OK-σ	*Tʷ	IDENT-EO-length-sibilant
a. [ʂʷɑŋ.kʰɤ]		*!	
→ b. [ʂʐ̩.wɑŋ]			*
c. [ʂwɑŋ.kʰɤ]	*!		

In (32), *Kʷ is ranked below both DEP-IO-[ɤ]/velar stop __ and DEP-EO-[ɤ]/velar stop __, so to satisfy OK-σ, both the /k/+/u/ sequence in the SC input and the /g/+/w/ sequence in the English input are repaired by the labial vocoid merging into the preceding velar stop. In (33), *Tʷ is ranked below IDENT-IO-length-sibilant but above IDENT-EO-length-sibilant. As a result, while the /ʂ/+/u/ sequence in the SC input is repaired through merger, the /s/+/w/ sequence in the English input is repaired by the sibilant realized as a separate syllable.

[44] Following Duanmu (2007), the realization of /s/, /ts$^{(ɦ)}$/, /ʂ/, and /tʂ$^{(ɦ)}$/ as [sz̩], [ts$^{(ɦ)}$z̩], [ʂʐ̩], and [tʂ$^{(ɦ)}$ʐ̩], respectively, is considered lengthening of the sibilants.

6.3 Phonetic Motivations

On the grounds of the P-map theory (Steriade 2001, 2009), we can infer the perceptual similarity hierarchy in (34) from the constraint ranking DEP-EO-[ɤ]/velar stop __ >> *K^w and the perceptual similarity hierarchy in (35) from the constraint ranking *T^w >> IDENT-EO-length-sibilant. ("Δ" denotes differences, "K" a velar stop, "V" a vowel, and "S" a sibilant).

(34) Δ $_σ$[K+w+V vs. $_σ$[K+ɤ]+$_σ$[w+V > Δ $_σ$[K+w+V vs. $_σ$[K^w+V

Within a syllable, a velar stop+/w/+vowel sequence is perceptually less similar to a velar stop+[ɤ]+[w]+vowel sequence than to a labialized velar stop+vowel sequence.

(35) Δ $_σ$[S+w+V vs. $_σ$[S^w+V > Δ $_σ$[S+w+V vs. $_σ$[S+z/ʐ]+$_σ$[w+V

Within a syllable, a sibilant+/w/+vowel sequence is perceptually less similar to a labialized sibilant+vowel sequence than to a sibilant+[z/ʐ]+[w]+vowel sequence.

As I will be showing below, articulatory and perceptual evidence confirms these two perceptual similarity hierarchies.

6.3.1 Articulatory Evidence

To verify the two perceptual similarity hierarchies, consider first the way in which velar stops, the labiovelar approximant /w/, and sibilants are produced. Velar stops and /w/ are produced with the back of the tongue raised, and during the production of /w/, the lips are also rounded (Olive, Greenwood, and Coleman 1993: 27; Ladefoged and Johnson 2011: 70). Due to anticipatory coarticulation, velar stops that immediately precede /w/ also show lip rounding (e.g. *quick*; Ladefoged and Johnson 2011: 70–71). In addition, while velar stops show a more front tongue-body position than /w/ when produced alone, the exact tongue-body position for velar stops depends on the adjacent vowels (Ladefoged and Johnson 2011: 70; Olive, Greenwood, and Coleman 1993: 136–137, 143–144, 146–147). For example, the /k/ in /ki/ is produced with a more front tongue body than the /k/ in /ku/. Consequently, the production of a velar stop+/w/ sequence can be described as follows. To produce the velar stop, a closure is created by the back of the tongue contacting the point on the soft palate where the following /w/ is formed with the lips rounded, and to produce the following /w/, the closure made for the velar stop is released with the lips remaining rounded. These facts indicate that the articulatory gestures for the two sounds in the sequence overlap to a great degree. Because /w/ is very short in duration (Olive, Greenwood, and Coleman 1993), the sequence should sound very similar to the labialized version of the velar stop.

In contrast, sibilants are produced with the tongue tip or blade making a constriction against the alveolar ridge or hard palate. Since the tongue body is attached to the tongue tip and blade, the tongue body is not totally free during the production of sibilants (Flemming 2003). This suggests that we do not expect the tongue-body position for sibilants to exhibit much variation. Indeed, Ladefoged and Johnson (2011: 70–71) showed the traces of the movements of various articulators during the production of /b/, /d/, and /g/ in different vowel contexts, and one of the observations they made is that the tongue-body variation for /d/, which is produced at the same place as the sibilants /s/ and /z/, is much smaller than the tongue-body variation for /b/ and /g/. Moreover, Flemming (2003) argued that anterior sibilants (/s/ and /z/) and non-anterior laminal sibilants (/ʃ/ and /ʒ/) favor a front tongue-body position (see 4.2.3). This argument also suggests limited variation in tongue-body position during the production of the two types of sibilants. Because the tongue-body position for the two types of sibilants is sort of fixed, to produce a sibilant+/w/ sequence in English, the tongue body must move from a front position to a back position. I conclude that the inevitable movement of the tongue body makes the two sounds in the sequence less overlapped and hence perceptually more distinctive from each other, resulting in the sequence sounding more similar to the sibilant+[z/ʐ]+[w] sequence than to the labialized version of the sibilant.

6.3.2 Perceptual Experiments I, II, and III

Three experiments were conducted. Experiments I and II aimed to test the perceptual similarity hierarchies in (34) and (35), and Experiment III was designed to test the influence of English orthography.

6.3.2.1 Stimuli

Experiment I was an AXB discrimination task, in which the subjects listened to the stimuli and for each stimulus judged whether the A or B item sounded more similar to the X item. The AXB paradigm used in this task included seven triplets, in which the X items were disyllabic nonce English words and the A and B items were the X items' SC matches. Two example triplets are given below.

(36) Two triplets from the AXB paradigm used in Experiment I

A (SC)	X (E.)	B (SC)
[sẓ.wei.pi]	/swéɪ.bi/	[s^{w}ei.pi]
[k^{h}ɤ.wo.pi]	/kwóʊ.bi/	[k^{hw}o.pi]

The first syllable of the X item was either a /s/+/w/ sequence followed by one of the four vowels /ɔ/, /ɑ/, /u/, /eɪ/, or a /k/+/w/ sequence followed by one of the three vowels /oʊ/, /ɑ/, /eɪ/, and the second syllable was always /bi/. The stress was placed on the first syllable. One of the SC matches was trisyllabic but the other was disyllabic. Depending on the initial sound of the X item (i.e. /s/ or /k/), the trisyllabic SC match started with [sẓ.w] or [k^{h}ɤ.w] and the disyllabic SC match started with [s^{w}] or [k^{hw}]. Everything else in the two SC matches is equal. See Appendix A for the seven-triplet paradigm.

A male native American English speaker produced the X items and a male native Taiwan SC speaker produced the A and B items. Both talkers were doctoral students in linguistics at Michigan State University at the time of recording. They read the words in a carrier sentence ("I say __ twice" for the English words and "wo shuo __ san ci" [I say __ three times] for the SC words) at a normal speed in front of a microphone that was connected to a high-quality tape recorder. After the tapes were digitized into a computer, the target words were grouped into triplets, and 28 triplets were created: Seven triplets in the paradigm × two orders of the A and B items × two repetitions. A one-second silence was added at the beginning of every triplet, and a half-second silence was added between the A and X items and between the X and B items. The 28 triplets were mixed with 188 fillers, and the total 216 triplets were arranged in a random order. Notice that 112 of the 188 triplets that were used as the fillers also served as the stimuli in Experiment V, which tests the relative perceptual distinctiveness of a coda liquid after various vowels in English and will be introduced in 8.2.2.2.

Experiment II was also a discrimination task, which was similar to Experiment I. In this task, the subjects only heard the X items and were given the written forms of the SC matches (in characters) to choose from. The X items included the seven English words used in Experiment I and the three additional English words /swǽ.bi/, /swʌ́.bi/, and /kwǽ.bi/. See Appendix B for a list of the 10 English words and the corresponding written SC forms. The same English talker was recorded producing the three additional words in the same fashion. There was a total of 20 aurally presented stimuli arranged in a random order: 10 X items × two orders of the A and B items.

The 10 English words used in Experiment II were converted to written forms for Experiment III (e.g. *Swoobe* for /swú.bi/). That is, these were no aurally presented stimuli in this task. See Appendix C for the written forms of the 10 English words along with their written SC matches. There was a total of 20 triplets arranged in a random order: 10 X items × two orders of the A and B items.

6.3.2.2 Subjects

Twenty-nine native SC speakers (14 males, 15 females) volunteered to participate in the experiments. They were all from Taiwan and in their late twenties to early fifties. None of them was a native English speaker, and they had lived in the U.S. for an average of 5.5 years at the time of the experiments.

6.3.2.3 Procedure

The experiments were conducted in a quiet room. The subjects either came alone or in groups no bigger than three people. A brief instruction and practice were given before Experiment I started. In Experiment I, the subjects listened to the triplets and for each triplet judged which SC match, A or B, sounded more similar to the English word X by circling the letter A or B on an answer sheet. There was a short break after the 105^{th} triplet and after Experiment I. In Experiment II, the subjects listened to the 20 English words and for each of them circled one of the two SC matches presented in Chinese characters on the answer sheet that they thought sounded more similar to the aurally presented stimulus. The stimuli were played from two speakers about two meters away from the subjects. Experiment III was identical to Experiment II except that the subjects were given the written forms of the English words. The whole process lasted for 35-40 minutes.

6.3.2.4 Results and Discussion

The results of the three experiments are given below.[45] The numerals to the left of the percentages indicate the number of times the subjects made the selection.

Table 6.1 Results of three experiments testing the perceptual similarity hierarchies in (34) and (35) and the influence of English orthography (Experiment I: $\chi^2 = 179.66$, $p < .00001$; Experiment II: $\chi^2 = 301.76$, $p < .00001$; Experiment III: $\chi^2 = 215.05$, $p < .00001$)

English	SC	Experiment I		Experiment II		Experiment III	
sw	sz̩.w	350	75.43%	320	91.95%	301	86.49%
	s^{w}	114	24.57%	28	8.05%	47	13.51%
		464	100%	348	100%	348	100%
kw	k^{h}ɤ.w	98	28.16%	49	21.12%	61	26.29%
	k^{hw}	250	71.84%	183	78.88%	171	73.71%
		348	100%	232	100%	232	100%

The perceptual similarity hierarchies in (34) and (35) are confirmed. Table 6.1 shows that the trisyllabic SC correspondent is preferred when the English word starts with a sibilant but that the disyllabic SC correspondent is preferred when the English word starts with a velar stop. Moreover, a comparison of the results of Experiments II and III indicates that the orthography plays no role. If the English orthography plays a role, a trisyllabic SC correspondent should be selected more often in Experiment III than in Experiment II, whether the stimuli start with a sibilant or a velar stop. The reasons are (i) that when the subjects saw a written form starting with *sw* or *kw*, they might tend to insert a vowel between the two graphemes, as the these graphemes represent consonans, and (ii) that when the subjects saw a written form such as *Quaibe*, they might tend to think that each of the three grapheme sequences *qu*, *ai*, and *be* represent a syllable. However, for /sw/, a trisyllabic SC correspondent is selected more often in Experiment II (320 times, 91.95%) than in Experiment III (301 times, 86.49%) (shown in the four top cells in shade). Although a trisyllabic SC correspondent is selected more often in

[45] One of the subjects had to leave after the first experiment, so the results of the second and third experiments were based on the answers from the other 28 subjects.

Experiment III (61 times, 26.29%) than in Experiment II (49 times, 21.12%) for /kw/ (shown in the bottom four cells in shade), the difference is statistically insignificant ([kʰɤ.w+V.bi] vs. [kʰʷ+V.bi]: $\chi^2 = 1.72$, $p = .190225$). These facts indicate that the English orthography did not have effects on the subjects' judgments.

6.4 Summary of Chapter Six

When English words containing a sibilant+/w/ or velar stop+/w/ sequence are borrowed into SC, the sibilant+/w/ sequence splits into two syllables but the velar stop+/w/ sequence surfaces with the /w/ merging into the velar stop. To account for these patterns, I follow Steriade (2001, 2009) and propose two perceptual similarity hierarchies, which basically say that a sibilant+/w/ sequence is perceptually more similar to a sibilant+[z/ʐ]+[w] sequence than to the labialized version of the sibilant whereas a velar stop+/w/ sequence is perceptually more similar to the labialized version of the velar stop than to a velar stop+[ɤ]+[w] sequence.

Articulatory evidence confirms the two hierarchies. We can infer from Ladefoged and Johnson's (2011) observation and Flemming's (2003) argument that sibilants consistently show a fronted tongue body even in back-vowel contexts. As a result, to produce a sibilant+/w/ sequence, the tongue body must move from a front position to a back position, rendering the sibilant and /w/ perceptually distinctive from each other. In contrast, velar stops and /w/ are produced with the back of the tongue body raised, and when a velar stop precedes /w/, the two sounds show exactly the same tongue-body position. Due to anticipatory coarticulation and the short duration of /w/, a velar stop+/w/ sequence is expected to sound very similar to the labialized version of the velar stop.

The two hierarchies are also confirmed by the results of three perceptual experiments, which show that a sibilant+/w/ sequence and a velar stop+/w/ sequence in English are generally judged more similar to a sibilant+[z/ʐ]+[w] sequence and the labialized version of the velar stop, respectively, in SC. The results of the last two experiments also indicate that English orthography plays no role in the listeners' judgments.

CHAPTER SEVEN

ADAPTATION OF VOWEL+NASAL+VOWEL SEQUENCES IN ENGLISH LOANWORDS IN SC

When English words containing vowel+nasal+vowel (henceforth, VNV) sequences are borrowed into SC, the nasal consonants may geminate (e.g. *Cannon* → [kʰan.nuŋ], *Hamilton* → [xan.mi.tʷən]). Since no SC phonological requirement is violated if the nasal consonants do not geminate (e.g. *Shannon* → [ɕa.nuŋ], *Obama* → [ou.pa.ma]), gemination of the nasal consonants represents a case of "unnecessary repair". My corpus shows that gemination of the nasal consonants occurs more often in some conditions than in others. After showing the data in 7.1, I will provide evidence in 7.2 arguing that the adaptation asymmetries have a phonetic basis. Formal analyses will be provided in 7.3, and 7.4 is the summary.

7.1 Adaptation Asymmetries

The English words in my corpus contain 233 instances of VNV sequences. Because the VŋV sequence occurs only twice, this sequence is not considered.[46] Thus, the total number of instances analyzed here is 231.

The adaptations show three asymmetries: (i) The nasal consonants geminate most often when following a low monophthong, less often when following a non-low lax vowel, and least often when following a non-low tense vowel or diphthong;[47] (ii) the nasal consonants geminate more often when following a stressed vowel than when preceding a stressed vowel; (iii) the nasal consonants geminate more often when they are /n/ than when they are /m/. The data and examples are given below. The examples are listed with the pre-nasal vowels shown in front of

[46] The /ŋ/ in this sequence is realized as a nasal+velar stop sequence in both mappings (*Bingaman* → [pin.kɤ.man], *Springer* → [ʂʐ̩.pʰu.lin.kɤ]). Huang and Lin (2013), who also investigated the adaptation of VNV sequences in English loanwords in SC, reported the same adaptation pattern (e.g. *Singer* → [ɕin.kɤ]).

[47] There is no instance of /ɔɪ+N+V/ sequences in my corpus.

them, as the pre-nasal vowels play a significant role in whether gemination occurs. Two examples, one involving /n/ and the other involving /m/, are given for each vowel. If no example is given, it is because no instance is found in my corpus. All examples involve an English source word in which the pre-nasal vowel is stressed, except for those under the heading VNV́ in (38).

7.1.1 Quality of the Pre-Nasal Vowel

Table 7.1 Frequencies of nasal gemination in the adaptation of VNV sequences in English loanwords in SC as a function of the quality of the pre-nasal vowels in the source words (low monophthong vs. non-low lax vowel: $\chi^2 = 12.44$, $p < .0005$; non-low lax vowel vs. tense vowel and diphthong: $\chi^2 = 6.87$, $p < .01$; low monophthong vs. tense vowel and diphthong: $\chi^2 = 33.08$, $p < .00001$)

pre-nasal vowel in the source word	adaptation	frequency	percentage
low monophthong	VNV → VNNV	41	70.69%
	VNV → VNV	17	29.31%
non-low lax vowel	VNV → VNNV	36	40.91%
	VNV → VNV	52	59.09%
tense vowel or diphthong	VNV → VNNV	19	22.35%
	VNV → VNV	66	77.65%

(37) Examples of nasal gemination in the adaptation of VNV sequences in English loanwords in SC as a function of the quality of the pre-nasal vowels in the source words

i. The pre-nasal vowel in the source word = low monophthong

	VNV	→	VNNV	VNV	→	VNV
/æ/	*Cannon*	→	[kʰan.nuŋ]	*Shannon*	→	[ɕa.nuŋ]
	Hamilton	→	[xan.mi.tʷən]	*Miami*	→	[mai.a.mi]
/ɔ/	*Corning*	→	[kʰɑŋ.niŋ]			
/ɑ/	*Connie*	→	[kʰɑŋ.ni]	*Guantanamo*	→	[kʷan.ta.na.mʷo]
	Thomas	→	[tʰɑŋ.ma.ʂz̩]	*Obama*	→	[ou.pa.ma]
			[tʰɑŋ.mu.ʂz̩]			
			[tʰɑŋ.ma.sz̩]			

ii. The pre-nasal vowel in the source word = non-low lax vowel

	VNV	→	VNNV	VNV	→	VNV
/ɪ/	*Kinney*	→	[tɕin.ni]	*Gwyneth*	→	[kɤ.ni.sz̩]
/ɛ/	*Kenny*	→	[kʰən.ni]	*Lenox*	→	[lai.nʷo.sz̩]
/ʌ/	*Cunningham*	→	[kʰɑŋ.niŋ.xan]			
	Summers	→	[sɑŋ.mʷo.sz̩]	*Summers*	→	[sa.mʷo.sz̩]
/ʊ/	*Sunni*	→	[ɕyn.nʲi]			

iii. The pre-nasal vowel in the source word = tense vowel or diphthong

	VNV	→	VNNV	VNV	→	VNV
/i/	*Catalina*	→	[kʰai.tʰɤ.lin.na]	*Catalina*	→	[kʰa.tʰɤ.li.na]
				Freeman	→	[fei.li.man]
/eɪ/	*Cheney*	→	[tɕʰan.ni]	*Chaney*	→	[tɕa.ni]
				Jamie	→	[tɕe.mi]
/u/	*Clooney*	→	[kʰu.luŋ.ni]	*Spooner*	→	[ʂʐ̩.pʰu.na]
				Schumer	→	[ʂu.mʷo]
/oʊ/	*Stoner*	→	[ʂʐ̩.tuŋ.nʷo]	*Conant*	→	[kʰou.nan.tʰɤ]
				Naomi	→	[na.ou.mi]
/ɑɪ/				*Heineken*	→	[xai.ni.kən]
				Hyman	→	[xai.man]
/ɑʊ/				*Downey*	→	[tɑu.ni]

As shown in Table 7.1, the nasal consonants geminate 70% of the time when following a low monophthong, but the frequency drops to 41% when following a non-low lax vowel and to only 22% when following a tense vowel/diphthong.

7.1.2 Location of the Stress

Table 7.2 Frequencies of nasal gemination in the adaptation of VNV sequences in English loanwords in SC as a function of the stress in the source words ($\chi^2 = 8.45$, $p < .005$)

	adaptation	frequency	percentage
pre-nasal vowel stressed	V́NV → VNNV	79	47.02%
	V́NV → VNV	89	52.98%
post-nasal vowel stressed	VNV́ → VNNV	10	22.73%
	VNV́ → VNV	34	77.27%

(38) Examples of nasal gemination in the adaptation of VNV sequences in English loanwords in SC as a function of the stress in the source words

i.	V́NV	→	VNNV	V́NV	→	VNV
/i/	*Salinas*	→	[sa.lin.na.sẓ]	*Chino*	→	[tɕi.nʷo]
				Lehman	→	[li.man]
/ɪ/	*Finny*	→	[fən.ni]	*Gwyneth*	→	[kɤ.ni.sẓ]
				Jimmy	→	[tɕi.mi]
/eɪ/	*Trainor*	→	[ʈʂʰʷan.nʷo]	*Boehner*	→	[pei.na]
				Amy	→	[ai.mi]
/ɛ/	*Penny*	→	[pʰan.ni]	*Seneca*	→	[sai.nei.tɕa]
				Emily	→	[ai.mi.li]
/æ/	*Aniston*	→	[an.ni.sẓ.tʷən]	*Bannister*	→	[pan.ni.sẓ.tʰɤ]
	Samuel	→	[ʂan.mʲou]	*Cameron*	→	[kʰa.mai.luŋ]
/ʌ/	*Summers*	→	[sɑŋ.mʷo.sẓ]	*Summers*	→	[sa.mʷo.sẓ]
	Cunningham	→	[kʰɑŋ.niŋ.xan]			
/u/	*Clooney*	→	[kʰu.luŋ.ni]	*Luna*	→	[lu.na]
				Bloomington	→	[pu.lu.miŋ.tʷən]
/ʊ/	*Sunni*	→	[ɕyn.ni]			

/oʊ/	*Stoner*	→	[ʂʐ̩.tuŋ.nʷo]	*Leona*	→	[li.ɑu.na]
				Cuomo	→	[kʷo.mʷo]
/ɔ/	*Corning*	→	[kʰɑŋ.niŋ]			
/ɑ/	*Johnny*	→	[tɕʰɑŋ.ni]	*Guantanamo*	→	[kʷan.ta.na.mʷo]
	Thomas	→	[tʰɑŋ.ma.ʂʐ̩]	*Obama*	→	[ou.pa.ma]
/ɑɪ/				*Heineken*	→	[xai.ni.kən]
				Simon	→	[sai.man]
/ɑʊ/				*Downey*	→	[tɑu.ni]
ii.	VNV́	→	VNNV	VNV́	→	VNV
/ɪ/	*Sinatra*	→	[ɕin.na.tɕy]	*Gwinnett*	→	[ku.nei]
/ə/	*Denise*	→	[tan.ni.sz̩]	*Anita*	→	[ai.ni.tʰa]
				Sacramento	→	[ʂa.tɕa.mʲɛn.tu]
/oʊ/				*O'Neil*	→	[ou.ni.ər]
				Tomei	→	[tʰʷo.mei]
/ɑ/				*La Mesa*	→	[la.mei.sa]
/ɑɪ/				*Raimondi*	→	[ʐʷei.man.ti]

When following a stressed vowel, the nasal consonants geminate almost half of the time; when preceding a stressed vowel, they geminate less than one fourth of the time.[48]

[48] The pre-nasal and post-nasal vowels in a source word may be both unstressed (e.g. *Tiffany*). Of the 19 such instances in my corpus, seven are adapted with gemination (36.84%) and 12 are adapted without gemination (63.16%). These 19 instances are not considered because, when compared with the two data sets in Table 7.2, statistical significance at $p < .1$ is not reached (V́NV vs. VNV: $\chi^2 = 0.71$, $p = .40$; VNV́ vs. VNV: $\chi^2 = 1.34$, $p = .25$).

7.1.3 Place of the Nasal Consonant

Table 7.3 Frequencies of nasal gemination in the adaptation of VNV sequences in English loanwords in SC as a function of the place of the intervocalic nasal consonants ($\chi^2 = 44.91$, $p < .00001$)

	adaptation	frequency	percentage
/n/	VnV → VNNV	89	56.33%
	VnV → VNV	69	43.67%
/m/	VmV → VNNV	7	9.59%
	VmV → VNV	66	90.41%

(39) Examples of nasal gemination in the adaptation of VNV sequences in English loanwords in SC as a function of the place of the intervocalic nasal consonants

i.	VnV	→	VNNV	VnV	→	VNV
/i/	*Salinas*	→	[sa.lin.na.sz̩]	*Geena*	→	[tɕi.na]
/ɪ/	*Minnesota*	→	[miŋ.ni.su.ta]	*Gwyneth*	→	[kɤ.ni.sz̩]
/eɪ/	*Allegheny*	→	[ja.li.tɕan.ni]	*Boehner*	→	[pei.na]
/ɛ/	*Jennifer*	→	[ʈʂən.ni.fʷo]	*Lenox*	→	[lai.nʷo.sz̩]
/æ/	*Daniel*	→	[tan.ni.ər]	*Joanna*	→	[tɕɑu.y.na]
/ʌ/	*Cunningham*	→	[kʰɑŋ.niŋ.xan]			
/u/	*Clooney*	→	[kʰu.luŋ.ni]	*Spooner*	→	[ʂʐ̩.pʰu.na]
/ʊ/	*Sunni*	→	[ɕyn.ni]			
/oʊ/	*Antonio*	→	[an.tuŋ.ni.ɑu]	*Winona*	→	[wei.nʷo.na]
/ɔ/	*Corning*	→	[kʰɑŋ.niŋ]			
/ɑ/	*Giuliani*	→	[ʈʂu.li.an.ni]	*Guantanamo*	→	[kʷan.ta.na.mʷo]
/ɑɪ/				*Heineken*	→	[xai.ni.kən]
/ɑʊ/				*Downey*	→	[tɑu.ni]

ii.	VmV	→	VNNV	VmV	→	VNV
/i/				*Lehman*	→	[li.man]
/ɪ/				*Jimmy*	→	[tɕi.mi]
/eɪ/				*Damon*	→	[tai.məŋ]
/ɛ/				*Emory*	→	[ai.m^{w}o.li]
/æ/	*Hamilton*	→	[xan.mi.t^{w}ən]	*Samuelson*	→	[sa.m^{j}ou.sən]
/ʌ/	*Summers*	→	[saŋ.m^{w}o.sz̩]	*Summers*	→	[sa.m^{w}o.sz̩]
/u/				*Bloomingdale*	→	[pu.lu.miŋ.tai.ər]
/oʊ/				*Tacoma*	→	[t^{h}a.k^{h}ɤ.ma]
/ɑ/	*Thomas*	→	[t^{h}ɑŋ.ma.ʂʐ̩]	*Obama*	→	[ou.pa.ma]
/ɑɪ/				*Simon*	→	[sai.man]

When the nasal consonants are /n/, gemination takes place more than half of the time; when they are /m/, gemination takes place only one tenth of the time.

7.2 Phonetic Motivations

7.2.1 The Vowel Quality Condition

The data in Table 7.1 show that gemination occurs most often when the pre-nasal vowels in the source words are low, less often when non-low and lax, and least often when tense or diphthongal. To account for these asymmetries, I consider (i) the duration of the pre-nasal vowels, and (ii) the spatial extent of the anticipatory vowel nasalization.

7.2.1.1 Vowel Quality vs. Vowel Duration

Duanmu (1993, 1994, 2007) proposed that, on surface, the rime of an SC full syllable has two timing slots. Indeed, his production data (Duanmu 1994) reveal that a vowel in an open syllable is significantly longer than the same vowel in a closed syllable. Duanmu's (1993, 1994,

2007) SC data are consistent with Howie's (1976) Beijing Mandarin data. For example, Howie's (1976) measurements show that the [i] vowel in the syllable [in] is about 35% shorter than an [i] vowel in an open syllable (146.25 milliseconds for the former and 227.5 milliseconds for the latter on average). Accordingly, it is reasonable to hypothesize that when English words containing a VNV sequence are borrowed into SC, a short pre-nasal vowel tends to map to a short vowel, leaving one timing slot unfilled, whereas a long pre-nasal vowel tends to map to a long vowel, resulting in both timing slots being filled. When a timing slot is left unfilled, the adapters have three options. The first option is to syllabify the nasal consonant into the coda of the syllable headed by the pre-nasal vowel and leave the next syllable onsetless (i.e. VNV(V) → VN.VV, where V and VV stand for a short vowel and a long vowel, respectively). This strategy is unattested in my corpus. The other two options are to geminate the nasal consonant (i.e. VNV(V) → VN.NVV) and to lengthen the pre-nasal vowel (i.e. VNV(V) → VV.NVV). When both timing slots are filled, the adapters have two options. The first is to do nothing (i.e. VVNV(V) → VV.NVV), and the second is to shorten the pre-nasal vowel and geminate the nasal consonant (i.e. VVNV(V) → VN.NVV). Because which option is selected is also influenced by other factors, the role of the duration of the pre-nasal vowels in the source words may not be readily identified (see below).

7.2.1.1.1 Durations of American English Vowels

The durations of American English vowels have been investigated by many scholars, and the results of two of the studies are shown in Table 7.4. I chose these two studies mainly because in these studies the vowel tokens precede a tautosyllabic voiced stop. This reason will be clearer shortly.

Table 7.4 Duration of vowels in American English (in milliseconds) (Heffner 1937: 130; Hillenbrand, Getty, Clark, and Wheeler 1995: 3103)

Heffner (1937)	ɪ	ʊ	ʌ	ɛ	ɑ	i	u	oʊ	eɪ	æ	ɔ
	200	200	200	220	260	280	290	300	300	310	320
Hillenbrand et al. (1995)	ʌ	ɪ	ɛ	ʊ	u	i	oʊ	eɪ	ɑ	æ	ɔ
	216	226	227	230	273	283	301	301	301	311	319

The two sets of data in Table 7.4 generally show that, in American English, non-low lax vowels are shorter than tense vowels, which in turn are shorter than low vowels. However, we must consider the fact that many factors can affect vowel duration in English. For example, within a syllable a vowel preceding a voiced obstruent is longer than the same vowel preceding the voiceless counterpart of the obstruent, and a stressed vowel is longer than a vowel of the same quality that does not carry stress. As we are concerned with the duration of the pre-nasal vowels in the source words, which always occur in non-final syllables, some notes regarding the two sets of data in Table 7.4 are necessary.

First, although the two sets of data in Table 7.4 are based on measurements of vowels in monosyllabic words, Umeda's (1975) American English vowel duration data, where the target vowels occur in non-final stressed syllables of polysyllabic words, show the same upward shift in duration (shown in Figures 1 and 2 on pp. 435). Second, Peterson and Lehiste (1960) measured the duration of vowels in monosyllabic words ending in various consonants in American English and observed that vowels preceding a homorganic nasal consonant are almost as long as the same vowels preceding a homorganic voiced stop (pp. 700). Umeda's (1975) data also reveal that voiced stops and nasal consonants have very similar effects on the duration of preceding stressed vowels (shown in Figure 3 on pp. 435). Third, according to Umeda (1975: 436), consonants do not have consistent influences on the duration of following vowels in

American English.[49] On the basis of these studies, I conclude that the values in the two sets of data in Table 7.4 are able to represent the relative durations of pre-nasal stressed vowels in non-final syllables in American English.[50] [51]

In English, unstressed vowels fall into two categories: (i) Schwa vowels and (ii) vowels that have the original quality (e.g. *O'Neil*, *ancestor*) (Umeda 1975: 442). Although Umeda (1975) also measured the duration of unstressed vowels in American English, she only provided the duration of schwa vowels. Her measurements showed that schwa vowels in non-final syllables are mostly shorter than 50 milliseconds (Figures 24 and 25 on pp. 442). Since her data are based on text-readings (i.e. continuous speech), it is inappropriate to compare this duration with the values in the two sets of data in Table 7.4, which are based on word-list-readings (and on vowels in monosyllabic words). Nonetheless, when compared with stressed vowels in non-final syllables in the same data set (Umeda 1975: 435, Figures 1 and 2), it can be clearly seen that schwa vowels are shorter than stressed non-low lax vowels, which are at least 60 milliseconds long.

The English words in my corpus contain only 15 VNV instances in which the pre-nasal vowel is unstressed but does not reduce to schwa (e.g. *O'Neil* → [ou.ni.ər], *Sinatra* → [ɕin.na.tɕʰy]). These unstressed non-schwa vowels are /ɪ/ (4 instances), /ɛ/ (1 instance), /æ/ (1 instance), /oʊ/ (4 instances), /ɑ/ (4 instances), and /ɑɪ/ (1 instance). The non-low lax vowels must be short, as their stressed counterparts are short. While data on the duration of unstressed tense

[49] Umeda (1975) found an exception in her data: After /h/, /æ/ is the shortest vowel (pp. 436). Nonetheless, Hillenbrand, Getty, Clark, and Wheeler's (1995) data are based on vowels preceded by /h/, and as we can see in Table 7.4, /æ/ has the second longest duration.

[50] Peterson and Lehiste (1960) reported a separate set of vowel duration data. In this set of data, the values were computed on the basis of CVC minimal pairs, in which the two members differ only in the voicing status of the final consonants. Since suprasegmental factors such as tempo and the pitch pattern of the frame sentence had been carefully controlled, the authors claimed that the values could represent the vowels' intrinsic duration. The values showed the same pattern as the two sets of data in Table 7.4; that is, non-low lax vowels are shorter than tense vowels, which in turn are shorter than low vowels.

[51] The most straightforward data were provided by Umeda (1975). She showed a graph (Figure 4 on pp. 436) which clearly indicates that non-low lax vowels are shorter than tense and low vowels before a nasal consonant in non-final stressed syllables. Because she did not provide the duration values, the data are not included in Table 7.4.

and low vowels in American English are unavailable, due to a very small number of occurrences in my corpus, these vowels are simply regarded as having long duration here.

7.2.1.1.2 Controlling Other Factors

Because non-low lax vowels are shorter than tense vowels before a nasal consonant in American English, the hypothesis made earlier, that short and long pre-nasal vowels tend to map to short and long vowels, respectively, is upheld. However, gemination occurs most often when the pre-nasal vowels in the source words are low (71%), which, according to the data in Table 7.4, have the longest duration. In addition, gemination occurs less than half of the time when the pre-nasal vowels in the source words are non-low and lax (41%). These facts seem to contradict the hypothesis. As I will discuss below, other factors such as the extent of the anticipatory vowel nasalization and stress in the source words are also involved in the adapters' decision of whether to geminate the nasal consonants, and obviously the interference of these factors is the cause of the contradictions. The only way to see the role played by the duration of the pre-nasal vowels in the source words is to keep every other variable constant. A case with enough tokens to reach statistical significance is found in my corpus. The English words in my corpus contain six instances of the /ɪ́+n+V/ sequence and 23 instances of the /í+n+V/ sequence. In these instances, the pre-nasal vowels /ɪ/ and /i/ are stressed and followed by /n/. While both /ɪ/ and /i/ are considered high front vowels in English phonology, /ɪ/ is slightly lower and more back than /i/ phonetically. It is well known that the velar port does not completely close during the production of oral vowels (even when the oral vowels occur in totally oral environments) (e.g. Moll 1962; Ohala 1971; Clumeck 1976; Bell-Berti and Krakow 1991), and scholars have found a high correlation between vowel height and the spatial extent of vowel nasalization. The general finding is that the lower the vowel, the greater the nasalization, regardless of the context (e.g.

Moll 1962, Schourup 1973; Clumeck 1976; Beddor 1982; Henderson 1984; Krakow 1987, 1989, 1994; Bell-Berti 1993). Using cine films, Henderson (1984) measured the velar height and the velopharyngeal port opening during the production of several English vowels occurring in different contexts. The two subjects participating in her experiment spoke West Mid Scots English and North East Irish English. Both varieties have the low central vowel /aː/, and in both varieties /i/ and /e/ are not diphthongized. The results relevant to the present discussion show (i) that in the /t_n/ context the velar position for /ɪ/ is as high as that for /i/ (shown in Figure 6 on pp. 79) and (ii) that in the /n_n/ context while the velopharyngeal port opening for /e/ is substantially smaller than that for /aː/, it is almost as small as that for /i/ (shown in Figure 21 on pp. 94) (Henderson (1984) only showed the films of the velopharyngeal port opening for the three vowels /i/, /e/, and /aː/ in the /n_n/ context). Because /ɪ/ is closer to /i/ than /e/ in terms of vowel height, we can confidently say that the velopharyngeal port openings for /ɪ/ and /i/ are probably the same size. These pieces of evidence indicate that in English /ɪ/ is as nasalized as /i/ before tautosyllabic /n/. As vowels are less nasalized before a heterosyllabic nasal consonant than before a tautosyllabic one (see below for references), the question now is whether the /n/'s in the two English sequences /ɪ́+n+V/ and /í+n+V/ are syllabified into the same syllables as the preceding vowels. The answer is affirmative. Durvasula and Huang (2015) reported that in English intervocalic nasal consonants following a stressed vowel behave like word-medial codas, whether the preceding stressed vowel is tense or lax. With all these on hand, consider now how the two English sequences /ɪ́+n+V/ and /í+n+V/ are adapted in my corpus. It is found that five of the six /ɪ́+n+V/ instances (83.33%) and six of the 23 /í+n+V/ instances (26.09%) are adapted with gemination ($\chi^2 = 6.62$, $p = .010062$). The asymmetry in gemination rate between the two sequences, therefore, confirms the effect of the pre-nasal vowel durations on the adapters' decision as to whether or not to geminate the nasal consonants.

7.2.1.2 Vowel Quality vs. Spatial Extent of Vowel Nasalization

Due to coarticulation, vowels preceding a nasal consonant are nasalized. In addition, the extent of the anticipatory nasalization varies as a function of a number of factors. For example, everything else being equal, vowels are more nasalized before a tautosyllabic nasal consonant than before a heterosyllabic nasal consonant (for Beijing Mandarin, see Shi 2010; for English, see, e.g. Ohala 1971; Cohn 1990; Krakow 1989, 1993, 1994; Solé 1995; Byrd, Tobin, Bresch, and Narayanan. 2009), and, as already mentioned, low vowels are more nasalized than mid and high vowels. Thus, it is plausible that the extent of the anticipatory vowel nasalization in the source words also plays a role in whether gemination occurs. The hypothesis underlying the discussion in this section is that the more extensive the anticipatory vowel nasalization in the source words, the more frequently gemination occurs. A graphic illustration of the hypothesis is given below:

(40) A graphic illustration of the hypothesis that the more extensive the anticipatory vowel nasalization in the source words, the more frequently gemination occurs ("$_{strong}$" and "$_{weak}$" indicate strong and weak nasalization of the pre-nasal vowel)

English		SC
$\tilde{V}_{strong}$NV	→	$\tilde{V}_{strong}$N.NV
$\tilde{V}_{weak}$NV	→	$\tilde{V}_{weak}$.NV

When the pre-nasal vowel in the source word is greatly nasalized, the nasal consonant geminates because gemination would put the pre-nasal vowel in the output form in a syllable that ends in a nasal consonant, yielding great anticipatory vowel nasalization. When the anticipatory vowel nasalization in the source word is weak, in order for the pre-nasal vowel in the output form to be similar in nasality to its correspondent in the source word, the nasal consonant must not geminate; instead, the nasal consonant is syllabified into the onset of the following syllable, resulting in a heterosyllabic V.N sequence and hence weak anticipatory vowel nasalization. This hypothesis can be easily transformed to perceptual similarity hierarchies like those in 6.3, but it is illustrated the way shown above for ease of reading. This hypothesis also underlies the account of the stress condition in 7.2.2.

7.2.1.2.1 Articulatory and Perceptual Evidence

As mentioned in 7.2.1.1.2, Henderson's (1984) data show that in English the velopharyngeal port opening during a mid or high vowel is smaller than that during a low vowel when the vowels occur between two /n/s. Her data also reveal that, in every context she considered (/t_t/, /t_n/, /n_t/, /n_n/), the velar position during a low vowel is lower than that during a mid or high vowel (shown in Figure 6 on pp. 79 and Figure 9 on pp. 82). Utilizing different techniques, other investigators have made similar observations (e.g. Moll 1962; Bell-Berti, Baer, Harris, and Nimi 1979; Bell-Berti 1993). For example, in a study of the variation in velopharyngeal closure during the production of high and low vowels occurring in various contexts in American English, Moll (1962) found that low vowels are more likely than high vowels to be articulated with the velopharyngeal port opened. He also found greater velar height for high vowels than for low vowels and greater distance between the velum and the posterior pharyngeal wall for low vowels than for high vowels when the vowels are adjacent to /n/. It should be noted that in Moll's (1962) study all the differences between high and low vowels are statistically significant but those between front and back vowels of the same height are not.

On the perception side, Ali, Gallagherj, Oldstein, and Daniloff (1971) conducted a perceptual experiment in which native English speakers listened to the CV(V) components of English CV(V)C and CV(V)N sequences and predicted whether the missing consonants were nasals. The results showed that the subjects were more likely to predict the missing consonants, whether nasals or non-nasals, to be nasals when the preceding vowel was /ɑ/ than when it was /i/, /u/, or /eɪ/ (pp. 540). Similar findings are reported in Lintz and Sherman (1961), Hattori, Yamamoto, and Fujimura (1958), and Dickson (1962).

Put together, the results of these studies indicate that, regardless of the context, low vowels are spatially more nasalized than non-low vowels in English, which is also reflected in perception.

7.2.1.2.2 Controlling Other Factors

To demonstrate that the spatial extent of the anticipatory vowel nasalization in the source words does play a role in whether the intervocalic nasal consonants geminate, every factor should be kept constant except for the vowel height. Three sets of data with enough tokens are found in my corpus. In these data sets, the pre-nasal vowels in the source words are long and stressed, and the intervocalic nasal consonants are /n/. The backness status of the pre-nasal vowels is not controlled since, according to Moll (1962), front and back vowels of the same height do not differ significantly in the spatial extent of anticipatory nasalization in English. The data are presented in Table 7.5:

Table 7.5 Frequencies of nasal gemination in the adaptation of V́nV sequences in English loanwords in SC where the stressed pre-nasal vowel is /i/, /u/, /eɪ/, /oʊ/, /æ/, or /ɑ/ (high tense vs. mid diphthong: $\chi^2 = 3.63$, $p = .056681$; mid diphthong vs. low: $\chi^2 = 5.96$, $p = .014669$; high tense vs. low: $\chi^2 = 20.13$, $p < .00001$)

	pre-nasal vowel in the source word	adaptation	frequency	percentage
i	high tense (/i/ or /u/)	V́nV → VN.NV	7	**25.93%**
		V́nV → V.NV	20	74.07%
ii	mid diphthong (/eɪ/ or /oʊ/)	V́nV → VN.NV	12	**52.17%**
		V́nV → V.NV	11	47.83%
iii	low monophthong (/æ/ or /ɑ/)	V́nV → VN.NV	31	**81.58%**
		V́nV → V.NV	7	18.42%

As we can see, when every other factor that can affect the frequency of gemination is controlled, gemination occurs most often when the pre-nasal vowels in the source words are low, less often when they are mid diphthongs, and least often when they are high and tense. The data in Table 7.5 clearly indicate the effect of the spatial extent of the anticipatory vowel nasalization in the source words.

7.2.1.2.3 Some Issues

In the two varieties of English on which Henderson's (1984) study is based, /e/ does not exhibit more nasalization than /i/ in the context of /t_n/ (shown in Figure 6 on pp. 79 and Figure on pp. 82). As a matter of fact, the subject who spoke the North Irish variety produced /e/ with a higher velar position than s/he produced /i/ in this context (shown in Figure 6 on pp. 79). However, Table 7.5 (i) and (ii) show that gemination occurs more often when the pre-nasal vowels are mid than when they are high (although the difference is only marginally significant at $p < .05$). I do not exclude the possibility that mid vowels are more nasalized than high vowels in American English, but if American English and the two varieties of English Henderson (1984) studied behave the same way with respect to the extent of anticipatory vowel nasalization, a plausible explanation is as follows. Lin (2008a, b) examined vowel adaptation in English loanwords in SC and observed that high vowels are rarely mapped to low vowels but mid vowels sometimes are. This asymmetry is also found in my corpus. In my corpus, only 5.31% of the high vowel tokens (46/867) are mapped to low vowels, but when it comes to mid vowels, the rate rises to 36.51% (230/630) ($\chi^2 = 147.78$, $p < .00001$). Lin (2008a, b) also observed in her data that front vowels are not mapped to back vowels and vice versa. A similar pattern is found in my corpus. In my corpus, vowel-quality changes in the front-back dimension account for only 14.84% of all vowel-quality changes (347/2339) if reduced vowels in the source words are not considered. In Table 7.5 (ii), if the pre-nasal mid diphthongs are mapped to low vowels, there are two strategies the adapters can use to maintain the vowel frontness or backness. The first is to map /eɪ/ to [ai] or [ja] and /oʊ/ to [ɑu] (no gemination occurs; e.g. *Dana* → [tai.na], *Chaney* → [tɕʰa.ni], *Leona* → [li.ɑu.na]), and the second is to geminate the intervocalic nasal consonant (e.g. *Dana* → [tan.na], *Trainor* → [ʈʂʷan.nʷo], *Sonia* → [sɑŋ.ni.ja][52]). These observations can explain the difference in gemination frequency between Table 7.5 (i) and (ii): If mid and high vowels exhibit the same extent of anticipatory nasalization in American English, gemination occurs more frequently in Table 7.5 (ii) than in Table 7.5 (i) because (i) mid vowels are much more likely than high vowels to map to low vowels, and (ii) one of the two strategies to maintain the frontness or backness of lowered mid vowels is to geminate the nasal consonant.

This account, however, makes us wonder if the high gemination frequency in Table 7.5 (iii) (81.58%) mainly results from maintaining the frontness or backness of the pre-nasal vowels (/æ/ and /ɑ/) and has little to do with their extensive nasalization. The answer is no. First, there are many ways to maintain the vowel backness status of /æ/ and /ɑ/. For example, /æ/ can be mapped to [ai], [ja], or [ei], and /ɑ/ can be mapped to [ɑu], [ou], or [ɤ]. My corpus shows that

[52] In SC, the backness status of a low vowel preceding a nasal consonant is determined by the place of the nasal consonant (i.e. [an], [ɑŋ], *[aŋ], *[ɑn]). When /æŋ/ or /ɑn/ is present in the source word, the backness status of the vowel is usually maintained at the expense of the place of the nasal consonant (i.e. /æŋ/ → [an], /ɑn/ → [ɑŋ]) (Hsieh, Kenstowicz, and Mou 2009; cf. C. C. Lin 2002).

when /æ/ and /ɑ/ do not precede a nasal consonant, the two most common ways of maintaining the frontness of /æ/ are (i) /æ/ → [ai] (34/60) (e.g. *Matt* → [mai tʰɤ], *Gap* → [kai.pʰu]) and (ii) /æ/ → [ja] (17/60) (e.g. *Adam* → [ja.tɑŋ], *Seattle* → [ɕi.ja.tʰu]), and the two most common ways of maintaining the backness of /ɑ/ are (i) /ɑ/ → [wo] (46/90) (e.g. *Robert* → [lʷo.pʷo.tʰɤ], *Hopkins* → [xʷo.pʰu.tɕin.sz̩]) and (ii) /ɑ/ → [ɑu] (23/90) (e.g. *Scott* → [ʂʐ̩.kʰɑu.tʰɤ], *Oscar* → [ɑu.sz̩.kʰa]). It happens that none of the four SC correspondents can precede a tautosyllabic nasal consonant. This suggests that, in Table 7.5 (iii), if the adapters want to maintain the frontness or backness of the pre-nasal vowels, we see frequent use of the four strategies rather than gemination. In other words, the high gemination frequency itself indicates that it is not due to maintaining the frontness or backness of the pre-nasal vowels.

Second, the English words in my corpus contain 73 instances of VmV sequences but only seven of them are adapted with gemination (9.59%), suggesting that /m/ is resistant to gemination in SC (see 7.2.3). Interestingly, of the seven instances that are adapted with gemination, six have /æ/ or /ɑ/ as the pre-nasal vowel (e.g. *Samuel* → [ʂan.mʲou], *Thomas* → [tʰɑŋ.ma.sz̩]). This fact shows that /m/, which is resistant to gemination, is geminated occasionally when following a low vowel.

Based on these pieces of evidence, I conclude that, in Table 7.5 (iii), the nasal consonants geminate in order for the pre-nasal vowels in the SC forms to be heavily nasalized so that they can be similar in nasality to the pre-nasal vowels in the source words.

I now briefly discuss the adaptation of VNV sequences where the pre-nasal vowels are diphthongs. Given that /m/ geminates at a very low frequency (9.59%), the discussion will focus on the sequences containing /n/. The two diphthongs /ɑɪ/ and /ɑʊ/ are not classified as low vowels but form a separate group with /eɪ/ and /oʊ/ because in my corpus intervocalic /n/

following /ɑɪ/ or /ɑʊ/ never geminates (0/4), which is in sharp contrast to the 78.95% gemination frequency (15/19) for intervocalic /n/ following /ɑ/. I argue that the change in formant transition inside the two diphthongs (i.e. from a low vowel to a high vowel) is perceptually salient (cf. Chen, Kapatsinski, and Guion-Anderson 2012; Repp and Svastikula 1987). Thus, the two diphthongs tend to map to diphthongs, which cannot precede a tautosyllabic nasal consonant in SC. As for /eɪ/ and /oʊ/, because the more sonorous components are mid vowels, which are closer to high vowels than low vowels, the diphthong-internal formant transitions are perceptually less salient, suggesting that they be mapped to monophthongs more often. Accordingly, we expect more frequent gemination of intervocalic /n/ after /eɪ/ or /oʊ/. This is exactly what is found in my corpus. As shown in Table 7.5 (ii), 12 of the 23 instances of /eɪ+n+V/ and /oʊ+n+V/ sequences are adapted with gemination (52.17%).

The discussion above suggests that, for the adapters, being faithful to a diphthong-internal low vowel-to-high vowel formant change is more important than being faithful to heavy vowel nasalization, which in turn is more important than being faithful to a diphthong-internal mid vowel-to-high vowel formant change (see 7.3 for formal analyses). Without these constraints and the ranking, we cannot explain the data in which intervocalic /n/ geminates more frequently after a mid diphthong (/eɪ/ or /oʊ/) than after a low diphthong (/ɑɪ/ or /ɑʊ/) even though the latter is more nasalized than the former. The suspicion that the internal formant transition in a diphthong is a crucial factor in whether a following intervocalic /n/ geminates finds empirical support in my corpus. In my corpus, 63.12% of the /ɑɪ/ and /ɑʊ/ tokens (67/106) are mapped to diphthongs, but when it comes to /eɪ/ and /oʊ/, the diphthong-to-diphthong mapping rate drops to 41.35% (129/312) ($\chi^2 = 15.18$, $p < .0001$). These data indicate that when an intervocalic /n/

follows a diphthong, the /n/ has more chances to geminate when the diphthong is mid than when it is low.

A summary of the discussion is as follows. How diphthongs are adapted is influenced by the internal formant transitions. /ɑɪ/ and /ɑʊ/ are more likely than /eɪ/ and /oʊ/ to remain as diphthongs because /ɑɪ/ and /ɑʊ/ contain more salient internal formant transitions than /eɪ/ and /oʊ/. Because SC disallows rimes of diphthong+nasal consonant, the difference between the two sets of diphthongs in the likelihood of being mapped to an SC monophthong accounts for the data in which an intervocalic /n/ following /eɪ/ or /oʊ/ geminates more often than one following /ɑɪ/ or /ɑʊ/ even though the latter are more nasalized than the former before a nasal consonant.

7.2.2 The Stress Condition

The data in Table 7.2 show that nasal gemination occurs more often when the pre-nasal vowels are stressed than when the post-nasal vowels are.

7.2.2.1 Word-Initial Nasals vs. Word-Final Nasals

To account for this asymmetry, we need to consider first how the constituent gestures for English /m/ and /n/ are coordinated. Each of the two nasal consonants involves two articulatory gestures: A velum lowering gesture for both /m/ and /n/, a lower lip raising gesture only for /m/, and a tongue tip raising gesture only for /n/. In a study investigating the effect of syllable position on the timing coordination of the velum and lip gestures for English /m/, Krakow (1989, 1993, 1994) found that the velum and the lower lip reached their targets at the same time when the /m/ was word-initial but that the velum reached the target considerably before the lower lip did when the /m/ was word-final (suggesting that the maximum achievement of the velum gesture occurred during the segment preceding the nasal consonant). Byrd, Tobin, Bresch, and Narayanan (2009) and Tobin, Byrd, Bresch, and Narayanan (2006), utilizing real-time MRI movies, found a very similar timing pattern for the two articulatory gestures for English /n/. They observed that while the maximum achievements of the velum and tongue gestures cooccurred for word-initial /n/, the maximum achievement of the velum gesture significantly preceded that of the tongue gesture for word-final /n/.

Additionally, both Krakow (1989, 1993, 1994) and Byrd, Tobin, Bresch, and Narayanan (2009) found a lower velum when the nasal consonant was word-final than when it was word-initial.[53] Similar findings are also reported in many previous studies (see the references cited in Krakow 1994: 32–33), including those focusing on languages other than English (e.g. Clumeck 1976; Fujimura, Miller, and Kiritani 1975). Since the velum reaches the target during the preceding segment when the nasal consonant is word-final but during the nasal consonant itself when it is word-initial, and since a word-final nasal consonant is associated with a lower velum than the same nasal consonant in word-initial position, I argue that a vowel preceding a word-final nasal consonant is spatially more nasalized than a vowel of the same height preceding the same nasal consonant in word-initial position. For example, in each of the two pairs of phrases *p**a**l**m** aid–p**a m**ade* and *b**on**e oh–b**ow** k**n**ow* (examples from Krakow 1989 and Byrd, Tobin, Bresch, and Narayanan 2009, and henceforth), the pre-nasal vowel in the first phrase is spatially more nasalized than that in the second phrase.

7.2.2.2 Word-Medial Nasals

Krakow (1989, 1993, 1994) and Byrd, Tobin, Bresch, and Narayanan (2009) also explored the inter-gestural coordination patterns for word-medial intervocalic nasal consonants and found that the local stress context had a significant effect on the patterns. They observed that a word-medial intervocalic nasal consonant following a stressed vowel behaved as though it

[53] Byrd, Tobin, Bresch, and Narayanan (2009) measured the size of the aperture between the velum and the pharyngeal wall and claimed that the change from the velum gesture onset to the maximal aperture could represent the spatial magnitude of the velum lowering gesture.

were word-final (i.e. the maximum achievement of the velum gesture was timed to the beginning of the oral gesture) while one preceding a stressed vowel behaved as though it were word-initial (i.e. the velic and oral targets were reached almost synchronously) (cf. Durvasula and Huang 2015). For example, in each of the following triplets, the intervocalic nasal consonants in the first and second items were found to have similar behaviors but the one in the third item was found to behave very differently: *hó**m**ey=ho**m**e E≠hoe **m**e*, *séa**m**y=sea**m** E≠see **m**e*, *bó**n**afide=bo**n**e oh≠bow k**n**ow*, *po**m**áde=pa **m**ade≠pal**m** aid*, *de**n**óte=toe **n**ode≠to**n**e ode*.

As for the spatial magnitude of the velum gesture for word-medial intervocalic nasal consonants as a function of stress, Krakow (1989, 1993, 1994) did not directly compare word-medial V́mV and VmV́ sequences in which the pre-nasal vowels had the same height. However, she found that the pre-nasal vowel in a word-medial V́mV sequence was associated with a lower velum than the same vowel preceding a word-initial /m/ (i.e. V#mV), and that the pre-nasal vowel in a word-medial VmV́ sequence was associated with a higher velum than the same vowel preceding a word-final /m/ (i.e. Vm#V). For example, she observed that the velar position was lower during the /oʊ/ vowel in ***hó**mey* than during the /oʊ/ vowel in *h**oe** me* but higher during the /ɑ/ vowel in *p**o**máde* than during the /ɑ/ vowel in *p**a**lm aid*. Since a word-initial nasal consonant patterns with a word-medial intervocalic nasal consonant preceding a stressed vowel (i.e. V#mV=VmV́) but a word-final nasal consonant patterns with one following a stressed vowel (i.e. Vm#V=V́mV), it is very plausible that, in a word-medial VmV sequence, the pre-nasal vowel is associated with a lower velum and hence spatially more nasalized when it is stressed than when the post-nasal vowel is stressed. Using the words and phrases just mentioned as examples, the velar position is expected to be lower during the /oʊ/ vowel in ***hó**m.ey* than during the /oʊ/ vowel in the hypothetical word *h**o**e.mé* but higher during the /ɑ/ vowel in *p**o**.máde* than during the /ɑ/ vowel in the hypothetical word ***pá**lm.aid*.

Byrd, Tobin, Bresch, and Narayanan (2009), on the other hand, did not find a consistent pattern for VnV sequences. However, as we will see below, when having the same height, the pre-nasal vowel in a V́nV sequence is generally produced with a lower velum than the pre-nasal vowel in a VnV́ sequence. The utterances of three of the four subjects participating in Byrd, Tobin, Bresch, and Narayanan's (2009) study showed a significant effect of stress on the spatial magnitude of velum lowering for intervocalic /n/, and the results are summarized in (41). In (41), (i) the post-nasal vowel in a V́nV sequence may carry a secondary stress, (ii) "#" means "word boundary", and (iii) ">" indicates that a statistically significant larger velum aperture displacement is observed for the nasal consonant in the context(s) to the left than for the nasal consonant in the context(s) to the right.

(41) Summary of the maximum velum aperture displacement for intervocalic /n/ in various stress contexts produced by three of the four subjects participating in Byrd, Tobin, Bresch, and Narayanan's (2009: 106) study

Subject A	V́#nV́	>	V́nV VnV́
Subject E	V́#nV́ V́nV	>	VnV́
Subject K	V́nV	>	V́#nV́ VnV́

As we can see in (41), the VnV́ sequence always appears to the right of ">", and the pre-nasal vowels in the sequence(s) appearing to the left of ">" always carry stress. This suggests that, in English, intervocalic /n/ is generally produced with a greater maximum velum aperture displacement when following a stressed vowel than when preceding a stressed vowel. Because stress position also determines where the maximum velum aperture displacement occurs (i.e.

during the preceding vowel or during the nasal consonant itself), I conclude that generally a stressed vowel is spatiall1y more nasalized than an unstressed vowel of the same height before an intervocalic /n/ in English.

7.2.2.3 Controlling Other Factors

My argument that, in English, vowels preceding an intervocalic nasal consonant are spatially more nasalized when they are stressed than when the post-nasal vowels are stressed is compatible with the data in Table 7.2, which show that the nasal consonants geminate more frequently when stress is on the preceding vowels than when it is on the following vowels. However, Table 7.2 also shows that gemination occurs only about half of the time when the pre-nasal vowels are stressed. This low gemination frequency is of course due to the influences of other factors such as the duration and height of the pre-nasal vowels. Only one case with enough tokens is found in my corpus after all these other factors are controlled. There are 25 instances of VnV sequences where the pre-nasal vowels are stressed non-low lax vowels in my corpus, and 20 of them are adapted with gemination (80.00%). In contrast, of the 21 instances of VnV sequences where the pre-nasal vowels are non-low and lax and the post-nasal vowels are stressed in my corpus, only eight are adapted with gemination (38.10%). The difference in gemination rate is significant ($p < .005$, $\chi^2 = 8.41$). The sharp contrast in gemination rate between these two sets of data confirms my argument.

7.2.3 The Nasal Consonant Place of Articulation Condition

The data in Table 7.3 show that gemination occurs much more often when the nasal consonant is /n/ than when it is /m/. This asymmetry is of course due to the SC requirement that no syllable end in [m]. However, the /m/ always can be mapped to [n.m] or [ŋ.m], as the /n/ is found to map to [ŋ.n] frequently when following a back vowel in the source words (e.g. *S<u>on</u>ia* →

[sɑŋ.ni.ja]) (cf. Hsieh, Kenstowicz, and Mou 2009; C. C. Lin 2002). Thus, the question is why the /m/ is mapped to [n.m] or [ŋ.m] much less often than the /n/ is mapped to [ŋ.n].

Before addressing the question, one may wonder if the /m/ geminates at such a low frequency simply because most of the vowel tokens preceding them happen to be tense/diphthongal or unstressed. Consider the data in Table 7.6:

Table 7.6 Frequencies of nasal gemination in the adaptation of VNV sequences in favor of gemination in English loanwords in SC ($\chi^2 = 7.50$, $p < .01$)

	adaptation		frequency	percentage
/n/	gemin.	$\acute{V}_{low\&back}$nV or $\acute{V}_{lax\&back}$nV → V_{back}ŋ.nV	12	75.00%
	non-gemin.	$\acute{V}_{low\&back}$nV or $\acute{V}_{lax\&back}$nV → V.nV	4	25.00%
/m/	gemin.	$\acute{V}_{low}$mV or $\acute{V}_{lax}$mV → V_{front}n.mV or V_{back}ŋ.mV	7	30.43%
	non-gemin.	$\acute{V}_{low}$mV or $\acute{V}_{lax}$mV → V.mV	16	69.57%

As discussed in previous sections, gemination tends to occur when the pre-nasal vowels in the source words are low or lax or carries stress. Thus, the data in Table 7.6, which show that intervocalic /m/'s are mapped to [n.m] or [ŋ.m] significantly less often than intervocalic /n/'s are mapped to [ŋ.n] in these contexts, confirms the asymmetry.

Why is mapping intervocalic /n/ to [ŋ.n] tolerated but mapping intervocalic /m/ to [n.m] or [ŋ.m] not? First of all, the adaptation VNV → VN.V is never attested in my corpus, indicating that the adapters avoid onsetless syllables when they can. Second, while the place of the first half of a geminate may change, the second half or the prevocalic nasal consonant in the SC output always has the same place of articulation as its English correspondent (i.e. V**n**V → Vn.**n**V, Vŋ.**n**V, V.**n**V; V**m**V → Vn.**m**V, Vŋ.**m**V, V.**m**V). This can be accounted for by the findings that the formant transition from a nasal consonant to a vowel is stronger than the formant transition from a vowel to a nasal consonant (e.g. Jun 1995; Repp and Svastikula 1987; the references cited in Zee 1981). In addition, Huang and Lin's (2013) observation that, in English loanwords in SC, intervocalic /ŋ/ is always realized as a nasal+velar stop sequence when gemination occurs (e.g.

Singer → [ɕin.kɤ]) receives an explanation. The question is now narrowed down to the syllable-final position. Namely, when gemination occurs, what is the reason that the place of the first half of the geminate changes from coronal to dorsal frequently but the change from labial to coronal or dorsal is uncommon?

7.2.3.1 Perceptual Account

A reasonable hypothesis is that, in syllable-final position, the contrast between /n/ and /ŋ/ is perceptually less discernable than a nasal place contrast involving /m/. That is, when the context is in favor of gemination but the place of the first half of the geminate cannot be maintained, the adapters are willing to change the place of the first half of the geminate from coronal to dorsal because the perceptual deviation caused by this change is small (English V_{back}**n**V → SC *V_{back}**n**.nV → SC V_{back}**ŋ**.nV), but because the change from labial to coronal or dorsal is perceptually quite noticeable, the adapters would rather not geminate (English V**m**V → SC *V**m**.mV → SC V**∅**.mV rather than SC *V**n**.mV/*V**ŋ**.mV)[54]. Research on perceptual similarity between the three nasal consonants in syllable-final position, however, has yielded conflicting results.

Malécot (1956) investigated the roles played by the nasal resonances and the formant transitions present in adjoining vowels in the identification of nasal consonants in English. He presented English-speaking listeners with N+/æ/ or /æ/+N syllables and found that as long as the stimuli contained an /m/ transition or resonance (e.g. æ(m)+(æ)n, æ(ŋ)+(æ)m; the symbols in parentheses represent the discarded components of the original production of a native English speaker), the stimuli were judged by most of the listeners as containing /m/. This finding is compatible with the hypothesis, as the finding suggests that, at least after /æ/, /m/ is perceptually

[54] See Kawahara and Garvey (2014) for a review of the proposal in the literature that phonological patterns are attributed to the relative perceptibility of phonetic contrasts.

more identifiable than the other two nasal consonants (i.e. an /m/–/n/ contrast and an /m/–/ŋ/ contrast are perceptually more discernible than an /n/–/ŋ/ contrast).

The results of Kawahara and Garvey's (2014) and Zee's (1981) studies, however, do not support the hypothesis. To test Jun's (1995, 2004) proposal that nasal stops are more likely to assimilate in place than oral stops in pre-consonantal position because in this position nasal place contrasts are less perceptible than oral place contrasts, Kawahara and Garvey (2014) conducted a series of perceptual experiments where English-speaking listeners were (i) presented with pairs of /ɑ+nasal stop/ or /ɑ+oral stop/ syllables (e.g. /ɑm/–/ɑn/ or /ɑb/–/ɑd/) and rated the similarity of the two syllables in each pair or (ii) presented with /ɑ+nasal stop/ or /ɑ+oral stop/ syllables and determined what the final consonants were. A careful check of the results showed that generally the /m/–/ŋ/ contrast was judged less discernable than the /m/–/n/ and /n/–/ŋ/ contrasts, suggesting /n/ probably being the most identifiable of the three nasals after /ɑ/. Zee (1981) examined the effect of vowel quality on the identification of syllable-final nasal consonants in English. In his experiments, English-speaking listeners were presented with VN syllables where V was /i/, /u/, /e/, /o/, or /ɑ/ and C was /m/, /n/ or /ŋ/ in three noise conditions and identified the nasal consonants. One of his findings was that all the three nasal consonants tended to be correctly identified after /ɑ/. This finding implies that, after /ɑ/, the three contrasts /m/–/n/, /m/–/ŋ/, and /n/–/ŋ/ are well discernable (cf. Chen, Kapatsinski, and Guion-Anderson 2012 and Chen and Guion-Anderson 2011 for Quanzhou Southern Min). Kawahara and Garvey's (2014) and Zee's (1981) findings not only do not support the hypothesis but also conflict with each other.

7.2.3.2 Articulatory Account

To account for why intervocalic /n/ is mapped to [ŋ.n] much more often than intervocalic /m/ is mapped to [n.m] or [ŋ.m], I follow Adler (2006) and argue in 7.2.3.2.1 that /n/ and /ŋ/ are articulatorily more similar to each other than to /m/. Other sets of data in my corpus, which will be presented in 7.2.3.2.2, also support this articulatory account.

7.2.3.2.1 Articulatory Similarity

As no conclusion can be drawn upon perceptual studies, I follow Adler (2006) and propose an articulatory account. Adler (2006) analyzed on-line adaptation of English words into Hawaiian, where [p], [k], [ʔ], and [h] are the only obstruents. What concerns us here is her observation that coronal obstruents were never mapped to [p] (i.e. E. coronal stop → H. [k] or [ʔ] and E. coronal fricative → H [k] or [h] unless deleted). This observation, as Adler (2006) argued, indicated that coronals and dorsals were more similar to each other than to labials. She reviewed perceptual and acoustical literature on the relative similarity of stops at the three places of articulation and found that the data contradicted each other. Since there was no evidence that coronals and dorsals were perceptually or acoustically more similar to each other than to labials, and since places of articulation had equal status in standard feature geometry, suggesting that traditional phonology plays no role, she concluded that the place similarity/dissimilarity observed in her data must be driven by some other forces. On the basis of Best's (1995) vocal tract model, in which the oral node splits into the lips and the tongue and the tongue further splits into the tip and the body, Adler (2006) argued that, as both coronals and dorsals were produced with the tongue, they were articulatorily more similar to each other than to labials, and that this articulator similarity/dissimilarity led to the adaptation patterns observed in her data. To support her idea, Adler (2006) presented findings from previous EPG studies tracking the movement of the tongue in alveolar-velar assimilation, which suggested that 'a sound that is perceived as a [t] or [k] might actually exhibit articulatory features of both types of segment' (pp. 1036). Another piece of evidence Adler (2006) presented was alveolar-velar substitutions found in children's speech. She argued that children made these substitutions to produce the closest approximation of the intended sounds. Adler (2006) also argued that the results of studies on visual cues were in the same line of her proposal, as visual cues provided listeners with information about how the intended sounds were articulated (cf. Yip 2002, 2006; Johnson, DiCanio, and McKenzie 2007) rather than auditory information.

Adler's (2006) proposal straightforwardly accounts for my data. Because SC has all the three nasal consonants, the adapters must know that /n/ and /ŋ/ are articulatorily more similar to each other than to /m/. As a result, when the context favors gemination but maintaining the place of the first half of the geminate is not possible due to some SC requirements (i.e. *V_{back}n.nV, *Vm.mV), the adapters choose to map /n/ to [ŋ.n] but choose to map /m/ to [Ø.m] rather than to [n.m] or [ŋ.m], and these decisions yield the asymmetry seen in the gemination rates.

7.2.3.2.2 Adaptation of /C+i/ Syllables

Other sets of data in my corpus are consistent with the data in Table 7.3 and therefore support Adler's (2006) proposal that the relative articulatory similarity of sounds motivates adaptation patterns. Here we are concerned with the adaptation of /i/ following various sounds.

Yip (2006) analyzed English loanwords in Cantonese and discovered that preserving vowel quality is less important than preserving consonant quality. A similar asymmetry is found in my corpus, which shows that vowel adaptation generally displays more variation than consonant adaptation. However, /i/ behaves differently. In my corpus, when /i/ appears in a CV syllable that cannot surface unchanged, it is usually the consonant rather than the /i/ vowel that undergoes modification (e.g. */ki/ → [tɕi]). /i/ being resistant to modification in loanword adaptation is not surprising, as /i/ is a palatal and therefore perceptually very salient (Kenstowicz 2005). Despite the stability of /i/, systematic exceptions are found in my corpus. In my corpus, /i/

is almost always mapped to [ei] when it serves as the vowel in a labial C+V syllable that requires modification to surface. Consider the data in Table 7.7:

Table 7.7 Adaptation of various /C+i/ syllables in English loanwords in SC

	English	SC					
	C+i	C+i	frequency	C+ɤ/ei	frequency	C+ai	frequency
non-labial C	θi	ɕi	2	sɤ	1		
	si	ɕi	16			sai	3
	zi	ɕi	4	sɤ	1	sai	1
		tɕʰi	1				
	ki	tɕi	9				
	gi	tɕi	1				
	hi	ɕi	2	xɤ	1		
ɹ	ɹi	li	30	lei	3	lai	1
				ʐʷei	12		
	t⁽ʰ⁾ɹi/dɹi	tɤ.li	1	tsʰʷei	3		
labial C	fi			fei	11		
	vi	mi	1	wei	4		
	wi			wei	4		

Adaptation of other /C+i/ syllables in my corpus is not included in Table 7.7 because either these syllables can surface unchanged (e.g. E. /pʰi/ → SC [pʰi]), or the adaptation involves only consonant devoicing (e.g. E. /di/ → SC [ti]), or the adaptation involves a negligible minor change to the consonant place (e.g. E. /ʃi/ → SC [ɕi]; /ʃ/ is a palato-alveolar fricative and [ɕ] is an alveolo-palatal fricative).

As we can see in Table 7.7, while /i/ is mostly mapped to [i] when following a non-labial sound, it is almost always mapped to [ei] when following a labial sound ($\chi^2 = 34.14$, $p < .00001$). The adaptation of /ɹi/, /t⁽ʰ⁾ɹi/, and /dɹi/ is worth noting. Some of the instances are adapted with the /i/ being mapped to [ei] accompanied with a labial secondary articulation. The source of the labial secondary articulation is the /ɹ/, as /ɹ/ is usually produced with rounded lips in English (e.g. Ladefoged and Johnson 2011; Campbell, Gick, Wilson, and Vatikiotis-Bateson 2010). Because SC does not allow [Cʷi] (SC does not allow /(C)ui/ underlyingly), to retain the lip rounding associated with the /ɹ/, the following /i/ is lowered to [ei] (e.g. E. /ɹ^{w}i/ → SC *[ʐ^{w}i] → SC [ʐ^{w}ei]), a strategy that is also adopted to adapt the labial C+/i/ syllables. There are two ways of explaining the absence of a labial secondary articulation from the output. First, not every English speaker produces /ɹ/ with their lips rounded (p.c. Dennis Preston, Grover Hudson). Second, lip rounding is a phonetic rather than phonological feature of /ɹ/ in English, and it is reasonable that phonetic features are not always retained even if they can be.

What is crucial to our discussion here is that whereas /i/ is mostly mapped to [i] at the cost of the place of a preceding non-labial consonant, it is almost always lowered to [ei] and therefore loses the perceptual saliency to some extent to maintain the place of a preceding labial consonant (even in the adaptation where /vi/ is mapped to [mi], the labial place of /v/ remains). In adapting /(C)ɹi/ syllables, the perceptual saliency of /i/ is also sacrificed to accommodate the labial feature of /ɹ/ if the /ɹ/ contains a labial feature and the adapters keep it.

In a perceptual account of these data, the hypothesis would be that labial sounds (including rounded /ɹ/) are more salient than /i/. That is, while /i/ is salient, the saliency of /i/ is sacrificed in order to maintain a more salient sound. This hypothesis, however, is not supported empirically. As we can see in Table 7.7, /si/ and /zi/ are mostly mapped to [ɕi]. /s/ and /z/ are sibilants, so this account would have to hypothesize that sibilants are less salient than /i/. If labial sounds are more salient than /i/, then they are more salient than sibilants, which does not seem to comply with acoustic findings. The adaptation of /θi/ also represents a problem. Like /si/ and /zi/, /θi/ tends to map to [ɕi], indicating that /θ/ is less salient than /i/. It is well known that /θ/ and /f/ are acoustically very similar and that, in child language acquisition, /θ/ is often confusable with /f/. Additionally, Kenstowicz and Suchato (2006) pointed out that while /θ/ was realized as [t] in English loanwords in Quebec French, Brannen's (2002) perceptual experiments had shown that

Quebec French speakers still judged [f] as a closer match for /θ/. In light of these facts, a perceptual account runs into a problem: Given that /θ/ and /f/ are very similar to each other perceptually and acoustically, there is no reason that /θ/ is less but /f/ more salient than /i/.

The data in Table 7.7 (and the mapping of /θ/ to [t] rather than to [f] in Quebec French) can be easily explained if we take the relative articulatory similarity of sounds into consideration. Articulatorily, labial sounds are not similar to sounds that are produced with the tongue. As a result, for labial C+/i/ syllables to surface, the labial consonants remain as labials and the /i/ vowel lowers to [ei] and loses some of the perceptual saliency.

7.3 Formal Analyses

A summary of the phonetic motivations is given below:

(42) Summary of phonetic motivations for the asymmetries in the adaptation of VNV sequences in English loanwords in SC

(When all other factors are equal,)

i. Gemination occurs more often when the pre-nasal vowels in the source words are short than when they are long because, in SC, vowels are shorter in closed syllables than in open syllables.

ii. Gemination occurs more often when the pre-nasal vowels in the source words are low than when they are non-low because, in English, low vowels are spatially more nasalized than non-low vowels before nasal consonants.

iii. Gemination occurs more often when the pre-nasal vowels in the source words are stressed than when they are unstressed because, in English, stressed vowels are spatially more nasalized than unstressed vowels before intervocalic nasal consonants.

iv. The place of the first half of a geminate changes from coronal to dorsal more often than from labial to coronal or dorsal because a coronal sound and a dorsal sound are articulatorily more similar to each other than to a labial sound.

To provide formal analyses of the data, I assume that the adapters perceive the anticipatory vowel nasalization in the source words as heavy, medium, or light.

Table 7.8 SC adapters' perception of the anticipatory vowel nasalization in VNV sequences in English

	The pre-nasal vowel is low	The pre-nasal vowel is non-low
The pre-nasal vowel is stressed	heavy	medium
The pre-nasal vowel is unstressed	heavy	light

I will explain later why I assume that the adapters hear heavy rather than medium nasalization when the pre-nasal vowel is low and unstressed.

As for the adapters' perception of the anticipatory vowel nasalization in VN.NV and V.NV sequences in SC, I assume the following:

Table 7.9 SC adapters' perception of the anticipatory vowel nasalization in VN.NV and V.NV sequences in SC

	The pre-nasal vowel is low	The pre-nasal vowel is non-low
VN.NV	heavy	medium
V.NV	light	light

One may put an SC V.NV sequence on par with an English VNV́ sequence because the nasal consonant in the English sequence behaves like an onset. However, unlike an English VNV́ sequence, an SC V.NV sequence contains a clear syllable boundary. Additionally, a syllable boundary is also a morpheme or even word boundary in SC. For these reasons, I assume that the adapters perceive the pre-nasal vowel in an SC V.NV sequence as lightly nasalized regardless of the vowel height.

On the grounds of the motivations and assumptions, I propose that the constraints in (43) and their relative rankings yield the asymmetries in Tables 7.1–7.3.

(43) Constraints

i. IDENT-EO-length/nasal consonant: A nasal consonant in the English input and its correspondent in the SC output have the same duration.

This constraint is violated if the intervocalic nasal consonant geminates.

ii. IDENT-EO-length/vowel: A vowel in the English input and its correspondent in the SC output have the same duration.

This constraint is violated if the pre-nasal vowel in the source word is long but shortened due to gemination or if the pre-nasal vowel in the source word is short but lengthened due to non-gemination.

iii. Vowel Nasality Distance < 1: The perceived distance in nasality between a vowel in the English input and its correspondent in the SC output should be smaller than one step.

This constraint is violated if the pre-nasal vowel in the source word is mapped to an SC vowel that is perceived as containing nasality of the next level. For example, a heavily nasalized vowel in the source word is mapped to an SC vowel that exhibits medium vowel nasalization.

iv. Vowel Nasality Distance < 2: The perceived distance in nasality between a vowel in the English input and its correspondent in the SC output should be smaller than two steps.

This constraint is violated if the pre-nasal vowel in the source word is perceived as heavily nasalized and mapped to an SC vowel that is perceived as lightly nasalized or the pre-nasal vowel in the source word is perceived as lightly nasalized and mapped to an SC vowel that is perceived as heavily nasalized. Violating this constraint entails violating Vowel Nasality Distance < 1.

v. IDENT-EO-articulator: A sound in the English input and its correspondent in the SC output are both produced with the lips or the tongue (Adler 2006: 1037)

This constraint is violated if /m/ in the source word is mapped to [n.m] or [ŋ.m].

vi. IDENT-EO-place/nasal consonant: A nasal consonant in the English input and its correspondent in the SC output have the same place of articulation.

This constraint is violated if /n/ in the source word is mapped to [ŋ.n]. Violating IDENT-articulator entails violating this constraint (Adler 2006: 1037).

vii. IDENT-EO-back: A vowel in the English input and its correspondent in the SC output are both back or non-back

This constraint is violated if the pre-nasal vowel in the source word is mapped to an SC vowel that differs in backness.

viii. IDENT-EO-internal formant transition/high-low F1: A diphthong in the English input and its correspondent in the SC output both contain an internal high-to-low F1 formant transition

This constraint is violated if the pre-nasal vowel in the source word is a low diphthong (i.e. /ɑɪ/ or /ɑʊ/) and mapped to an SC vowel that is not.

The tableaux below show the constraint rankings. In these tableaux, N represents [n] or [ŋ], the numerals in parentheses represent the numbers of instances in my corpus, $V_{low}V_{high}$ represents the low diphthongs /ɑɪ/ and /ɑʊ/, and IDNET-EO is shown as IDENT. Because we are not concerned with the lengths of the post-nasal vowels, they are simply represented as V. For ease of reading, I give one or two examples for a candidate if instances are found in my corpus.

In addition to the tableaux that correspond to the 12 combinations resulting from the three conditions (7.3.1 and 7.3.2), I will show a tableau in which the pre-nasal vowel in the source word is back (7.3.3) and a tableau in which it is a low diphthong (7.3.4).

7.3.1 VnV Sequences

The tableaux in (44)–(47) represent the adaptations where the nasal consonants are /n/ and the pre-nasal vowels in the source words are non-low. In these tableaux, the candidates with the largest numbers of instances are selected as long as IDENT-length/nasal consonant does not outrank both IDENT-length/vowel and Vowel Nasality Distance < 1. I simply do not rank these three constraints with each other.

(44) i. Nasal: /n/

ii. Pre-nasal vowel: non-low and short (e.g. /ɪ/, /ɛ/)

iii. Stress: pre-nasal vowel

$\acute{V}_{non\text{-}low}$nV (med. nas.) *Kénny*, *Línney*, *Bénnett*, *Lénox*	IDENT-length/nasal consonant	IDENT-length/vowel	Vowel Nasality Distance < 1
→ $V_{non\text{-}low}$N.nV (12) (med. nas.) *Kénny* → [kʰən.ni] *Línney* → [lin.ni]	*		
V_{low}N.nV (8) (heavy nas.) *Bénnett* → [pan.na.tʰɤ]	*		*!
Vː.nV (5) (light nas.) *Lénox* → [lai.nʷo.sz̩]		*	*!

(45) i. Nasal: /n/

ii. Pre-nasal vowel: non-low and long (e.g. /i/, /u/, /eɪ/)

iii. Stress: pre-nasal vowel

$\acute{V}ː_{non\text{-}low}$nV (med. nas.) *Cl<u>oóney</u>, D<u>ána</u>, G<u>éno</u>, Sp<u>oóne</u>r*	IDENT-length/nasal consonant	IDENT-length/vowel	Vowel Nasality Distance < 1
$V_{non\text{-}low}$N.nV (14) (med. nasa.) *Cl<u>oóney</u>* → [kʰu.l<u>uŋ.ni</u>]	*	*!	
V_{low}N.nV (5) (heavy nas.) *D<u>ána</u>* → [t<u>an.na</u>]	*	*!	*
→ Vː.nV (31) (light nas.) *G<u>éno</u>* → [tɕ<u>i.nʷo</u>] *Sp<u>oóne</u>r* → [ʂʐ̩.pʰʷ<u>u.na</u>]			*

In the two tableaux below, candidates containing a low vowel also violate Vowel Nasality Distance < 2. This constraint is not shown in these tableaux because the other three constraints are sufficient to select the optimal candidates.

(46) i. Nasal: /n/

ii. Pre-nasal vowel: non-low and short (e.g. /ɪ/, /ɛ/)

iii. Stress: post-nasal vowel

$V_{non\text{-}low}n\acute{V}$ (light nas.) *S<u>iná</u>tra*, *D<u>ení</u>se*, *Gw<u>inné</u>tt*, *Le<u>oná</u>rdo*	IDENT-length/nasal consonant	IDENT-length/vowel	Vowel Nasality Distance < 1
$V_{non\text{-}low}$N.nV (2) (med. nas.) *S<u>iná</u>tra* → [ɕ<u>in.na</u>.tɕy]	*		*!
V_{low}N.nV (6) (heavy nas.) *D<u>ení</u>se* → [t<u>an.ni</u>.sz̩]	*		*!
→ Vː.nV (13) (light nas.) *Gw<u>inné</u>tt* → [k<u>u.nei</u>] *Le<u>oná</u>rdo* → [li.<u>au.na</u>.tʷo]		*	

(47) i. Nasal: /n/

ii. Pre-nasal vowel: non-low and long (e.g. /i/, /u/, /eɪ/)

iii. Stress: post-nasal vowel

$V\!\!:_{non\text{-}low}n\acute{V}$ (light nas.) *<u>O'Néi</u>l*	IDENT-length/nasal consonant	IDENT-length/vowel	Vowel Nasality Distance < 1
$V_{non\text{-}low}$N.nV (0) (med. nas.)	*!	*	*
V_{low}N.nV (0) (heavy nas.)	*!	*	*
√ Vː.nV (1) (light nas.) *<u>O'Néi</u>l* → [<u>ou.ni</u>.ər]			

The tableaux in (48)–(50) represent the adaptations where the nasal consonants are /n/ and the pre-nasal vowels in the source words are low. Vowel Nasality Distance < 2 must outrank Vowel Nasality Distance < 1 since a two-step change is more noticeable than a one-step change and accordingly violating the former is more serious than violating the latter. The tableau in (48) also demonstrates that Vowel Nasality Distance < 2 must outrank IDENT-length/nasal consonant and IDENT-length/vowel.

(48) i. Nasal: /n/

ii. Pre-nasal vowel: low (and long) (e.g. /æ/, /ɑ/)

iii. Stress: pre-nasal vowel

V́ː$_{low}$nV (heavy nas.) *Jánet*, *Cánnon*, *Cónnie*, *Shánnon*	Vowel Nasality Distance < 2	IDENT-length/nasal consonant	IDENT-length/vowel	Vowel Nasality Distance < 1
V$_{non\text{-}low}$N.nV (2) (med. nas.) *Jánet* → [ʈʂən.ni]		*	*	*!
→ V$_{low}$N.nV (31) (heavy nas.) *Cánnon* → [kʰan.nuŋ] *Cónnie* → [kʰɑŋ.ni]		*	*	
Vː.nV (7) (light nas.) *Shánnon* → [ɕa.nuŋ]	*!			*

This ranking, however, does not select the candidate which we actually see in the data if the stress is on the post-nasal vowel and medium nasalization of the pre-nasal vowel in the input is assumed. Consider the tableau in (49):

(49) i. Nasal: /n/

ii. Pre-nasal vowel: low (and long) (e.g. /æ/, /ɑ/)

iii. Stress: post-nasal vowel

$V{:}_{low}n\acute{V}$ (med. nas.) *Annétte*	Vowel Nasality Distance < 2	IDENT-length/nasal consonant	IDENT-length/vowel	Vowel Nasality Distance < 1
$V_{non\text{-}low}N.nV$ (0) (med. nas.)		*	*!	
← $V_{low}N.nV$ (2) (heavy nas.) *Annétte* → [an.ni.tʰɤ]		*	*!	*
→ V:.nV (0) (light nas.)				*

In (49), there is no way for the candidate $V_{low}N.nV$ to be selected, no matter how the constraints are ranked with each other. To solve this problem, I suggest that the pre-nasal vowel in the input be perceived as heavily nasalized. Low oral vowels have been found to exhibit much more intrinsic nasalization than non-low oral vowels across-linguistically (e.g. Moll 1962; Clumeck 1976). This fact is consistent with the results of perceptual experiments. For example, in an experiment where English-speaking listeners listened to CV stimuli which had been edited from naturally produced CVC or CVN syllables (i.e. the final consonants were deleted) and determined whether the stimuli ended in a nasal or oral consonant, Ali, Gallagherj, Oldstein, and Daniloff (1971) found that when the vowel was /ɑ/, the missing consonants that were actually oral were frequently judged to be nasal. This finding demonstrates the very high level of intrinsic nasality of low vowels. Since low vowels exhibit extensive nasalization even in totally oral environments, it is very plausible that they are perceived as heavily nasalized before an ambisyllabic nasal consonant no matter which syllable the nasal consonant is affiliated with. The tableau in (50) shows that once the pre-nasal vowel in the input is marked as heavily nasalized, the actual output is selected.

(50) i. Nasal: /n/

ii. Pre-nasal vowel: low (and long) (e.g. /æ/, /ɑ/)

iii. Stress: post-nasal vowel

V:$_{low}$nV́ (heavy nas.) *<u>Ann</u>é<u>tte</u>*	Vowel Nasality Distance < 2	IDENT-length/nasal consonant	IDENT-length/vowel	Vowel Nasality Distance < 1
V$_{non-low}$N.nV (0) (med. nas.)		*	*	*!
→ V$_{low}$N.nV (2) (heavy nas.) *<u>Ann</u>é<u>tte</u>* → [<u>an.ni</u>.tʰɤ]		*	*	
V:.nV (0) (light nas.)	*!			*

7.3.2 VmV Sequences

Gemination rarely occurs when the nasal consonant is /m/, and, as I have argued, this is because /m/ is produced with the lips but /n/ and /ŋ/ are not. The tableau in (51) demonstrates that IDENT-articulator must outrank Ident-length/nasal consonant, Ident-length/vowel, and Vowel Nasality Distance < 1.

(51) i. Nasal: /m/

ii. Pre-nasal vowel: non-low and short (e.g. /ɪ/, /ɛ/)

iii. Stress: pre-nasal vowel

V́$_{non-low}$mV (med. nas.) *S<u>úmme</u>rs*, *J<u>ímmy</u>*, *<u>Émma</u>*	IDENT-articulator	IDENT-length/nasal consonant	IDENT-length/vowel	Vowel Nasality Distance < 1
V$_{non-low}$N.mV (0) (med.nas.)	*!	*		
V$_{low}$N.mV (1) (heavy nas.) *S<u>úmme</u>rs* → [s<u>ɑŋ.mʷo</u>.sz̩]	*!	*		*
→ V:.mV (9) (light nas.) *J<u>ímmy</u>* → [tɕ<u>i.mi</u>] *<u>Émma</u>* → [<u>ai.ma</u>]			*	*

The tableaux in (52)–(56) show the role of IDENT-articulator in adapting other VmV sequences.

(52) i. Nasal: /m/

ii. Pre-nasal vowel: non-low and long (e.g. /i/, /u/, /eɪ/)

iii. Stress: pre-nasal vowel

V́ː$_{\text{non-low}}$mV (med. nas.) *Ámy*, *Léhman*	IDENT-articulator	IDENT-length/nasal consonant	IDENT-length/vowel	Vowel Nasality Distance < 1
V$_{\text{non-low}}$N.mV (0) (med. nas.)	*!	*	*	
V$_{\text{low}}$N.mV (0) (heavy nas.)	*!	*	*	*
→ Vː.mV (23) (light nas.) *Ámy* → [ai.mi] *Léhman* → [li.man]				*

(53) i. Nasal: /m/

ii. Pre-nasal vowel: non-low and short (e.g. /ɪ/, /ɛ/)

iii. Stress: post-nasal vowel

V$_{\text{non-low}}$mV́ (light nas.) *Sacraménto*, *Des Móines*	IDENT-articulator	IDENT-length/nasal consonant	IDENT-length/vowel	Vowel Nasality Distance < 1
V$_{\text{non-low}}$N.mV (0) (med. nas.)	*!	*		*
V$_{\text{low}}$N.mV (0) (heavy nas.)	*!	*		*
→ Vː.mV (15) (light nas.) *Sacraménto* → [ʂa.tɕa.mʲɛn.tu] *Des Móines* → [ti.mʷo.in]			*	

(54) i. Nasal: /m/

ii. Pre-nasal vowel: non-low and long (e.g. /i/, /u/, /eɪ/)

iii. Stress: post-nasal vowel

Vː$_{\text{non-low}}$mV́ (light nas.) *Toméi*	IDENT-articulator	IDENT-length/nasal consonant	IDENT-length/vowel	Vowel Nasality Distance < 1
V$_{\text{non-low}}$N.mV (0) (med. nas.)	*!	*	*	*
V$_{\text{low}}$N.mV (0) (heavy nas.)	*	*	*	*
→ Vː.mV (1) (light nas.) *Toméi* → [t^{hw}o.mei]				

When the pre-nasal vowel in the source word is low and therefore perceived as heavily nasalized, non-gemination still occurs more often than gemination, suggesting that Vowel Nasality Distance < 2 does not outrank IDENT-articulator. I simply do not rank these two constraints with each other. This is shown in the two tableaux below.

(55) i. Nasal: /m/

ii. Pre-nasal vowel: low (and long) (e.g. /æ/, /ɑ/)

iii. Stress: pre-nasal vowel

V́ː$_{\text{low}}$mV (heavy nas.) *Hámilton*, *Thómas*, *Cámeron*, *Sámuelson*	Vowel Nasality Distance < 2	IDENT-articulator	IDENT-length/nasal consonant	IDENT-length/vowel	Vowel Nasality Distance < 1
V$_{\text{non-low}}$N.mV (0) (med. nas.)		*	*	*!	*
V$_{\text{low}}$N.mV (6) (heavy nas.) *Hámilton* → [xan.mi.t^{w}ən] *Thómas* → [t^{h}ɑŋ.ma.sẓ]		*	*	*!	
→ Vː.mV (7) (light nas.) *Cámeron* → [k^{h}a.mai.luŋ] *Sámuelson* → [sa.m^{j}ou.sən]	*				*

(56) i. Nasal: /m/

ii. Pre-nasal vowel: low (and long) (e.g. /æ/, /ɑ/)

iii. Stress: post-nasal vowel

Vː$_{low}$mV́ (heavy nas.) *L<u>a Mé</u>sa*	Vowel Nasality Distance < 2	IDENT-articulator	IDENT-length/nasal consonant	IDENT-length/vowel	Vowel Nasality Distance < 1
V$_{non-low}$N.mV (0) (med. nas.)		*	*	*!	*
V$_{low}$N.mV (0) (heavy nas.)		*	*	*!	
→ Vː.mV (3) (light nas.) *L<u>a Mé</u>sa* → [l<u>a.mei</u>.sa]	*				*

7.3.3 V$_{back}$NV Sequences

When the input contains a back pre-nasal vowel, it is mapped to V$_{back}$ŋ.NV more often than to V$_{non-back}$n.NV (N represents /m/ or /n/). This fact suggests that IDENT-back outranks IDENT-place/nasal consonant. I provide one tableau below. In this tableau, the two constraints in question are separated from other constraints.

(57) i. Nasal: /n/

ii. Pre-nasal vowel: non-low, long, and back (e.g. /u/, /oʊ/)

iii. Stress: pre-nasal vowel

V́ː$_{non\text{-}low\&back}$nV (med. nas.) *António*, *Tóny*, *Smithsónian*, *Spoóner*, *Leóna*	IDENT-length/nasal consonant	IDENT-length/V	VND < 1	IDENT-back	IDENT-place/nasal consonant
a. V$_{non\text{-}low\&back}$ŋ.nV (5) (med. nas) *António* → [an.tʷuŋ.nʲi.ɑu]	*	*!			*
b. V$_{low\&back}$ŋ.nV (2) (heavy nas.) *Tóny* → [tʰɑŋ.nʲi]	*	*!	*		*
c. V$_{non\text{-}low\&non\text{-}back}$n.nV (1) (med. nas.) *Smithsónian* → [ʂʐ̩.mʲi.sən.nʲi]	*	*!		*	
d. V$_{low\&non\text{-}back}$n.nV (0) (heavy nas.)	*	*!	*	*	
→ e. Vː$_{back}$.nV (8) (light nas.) *Spoóner* → [ʂʐ̩.pʰʷu.na] *Leóna* → [lʲi.ɑu.na]			*		
f. Vː$_{non\text{-}back}$.nV (0) (light nas.)			*	*!	

As we can see, the input is mapped to V$_{back}$ŋ.nV more often than to V$_{non\text{-}back}$n.nV even though the optimal candidate actually involves non-gemination. IDENT-back must outrank IDENT-place/nasal consonant or we cannot explain why more instances are observed for (57 a) and (57 b) than for (57 c) and (57 d), respectively. In other words, when everything else is equal, changing the place of the first half of the geminate is preferred to changing the vowel backness (cf. Hsieh, Kenstowicz, and Mou 2009; C. C. Lin 2002).

7.3.4 V$_{low\ diph}$NV Sequences

Gemination never occurs when the pre-nasal vowel in the source word is a low diphthong (i.e. /ɑɪ/ or /ɑʊ/), suggesting that Vowel Nasality Distance < 2 cannot outrank IDENT-internal formant transition/high-low F1. A tableau is provided in (58). The perceived nasality of the two low diphthongs is based on the vowel /ɑ/, as it more sonorous than /ɪ/ and /ʊ/ (cf. Lin 2008 b).

(58) i. Nasal: /n/

ii. Pre-nasal vowel: low diphthong (/ɑɪ/, /ɑʊ/)

iii. Stress: pre-nasal vowel

$\acute{V}_{low}V_{hi}nV$ (heavy nas.) *Héineken, Dówney*	IDENT-internal formant transition/high-low F1	VND < 2	IDENT-length/nasal consonant	IDENT-length/V	VND < 1
$V_{non\text{-}low}N.nV$ (0) (med. nas.)	*		*	*!	*
$V_{low}N.nV$ (0) (heavy nas.)	*		*	*!	
→ $V_{low}V_{hi}.nV$ (4) (light nas.) *Héineken* → [xai.ni.kən] *Dówney* → [tɑu.ni]		*			*

7.3.5 Constraint Rankings

The rankings of all the constraints used in the analyses close this section. Constraints in a box do not rank with each other

Figure 7.1 Constraint rankings proposed in Chapter Seven

IDENT-EO-internal formant transition/high-low F1 Vowel Nasality Distance < 2 IDENT-EO-articulator	>>	IDENT-EO-length/nasal consonant IDENT-EO-length/vowel Vowel Nasality Distance < 1
IDENT-EO-back	>>	IDENT-EO-place/nasal consonant

7.4 Summary of Chapter Seven

When English words containing VNV sequences are borrowed into SC, the nasal consonants may geminate. It is observed that the nasal consonants geminate more often in some conditions than in others, and I have argued that the asymmetries are phonetically driven.

First, I have shown that the gemination occurs more often when the pre-nasal vowels in the source words are non-low and lax than when they are tense or diphthongs because, in English, non-low lax vowels are shorter than tense vowels and diphthongs and, in SC, vowels are shorter in closed syllables than in open syllables.

Second, I have shown that the findings reported in the literature that the spatial extent of the anticipatory vowel nasalization in VN and VNV sequences in English differ as a function of the height of the pre-nasal vowel and stress location explains why gemination occurs more often when the pre-nasal vowels in the source words are low or stressed than when they are non-low or unstressed.

Third, following Adler (2006), I have argued that /n/ is mapped to [ŋ.n] much more often than /m/ is mapped to [n.m] or [ŋ.m] because a coronal sound and a dorsal sound are articulatorily more similar to each other than to a labial sound.

I have proposed a set of constraints and demonstrated that these constraints can account for the asymmetries once they are appropriately ranked with each other.

CHAPTER EIGHT

ADAPTATION OF CODA LIQUIDS IN ENGLISH LOANWORDS IN SC

When English words are borrowed into SC, generally coda liquids after front vowels tend to be preserved but those after central or back vowels tend to delete. To account for this asymmetry, I propose that, in English, coda liquids are perceptually more distinctive when following front vowels than when following non-front vowels. After showing the data in 8.1, I will provide articulatory and perceptual evidence in 8.2 to support the proposed perceptual difference. The adaptation asymmetry is formalized in 8.3, and 8.4 is the summary.

8.1 Adaptation Asymmetry

Segment deletion is uncommon in my corpus, and many of the deletions are observed in the adaptation of coda liquids. As I will be showing in this section, coda liquids in the source words tend to be preserved after front vowels but tend to delete after non-front vowels.

Before I present the data, note that not every instance of the vowel+coda liquid sequences in my corpus is considered. Word-final liquids in monosyllabic words where the liquids are the only segments violating the SC syllable constraints are always preserved (e.g. *Barr* → [pa.ər], *Hall* → [xʷo.ər]) (see 3.2). These mappings are omitted because it is the SC minimum-word requirement rather than the backness status of the preceding vowels that causes the liquid preservation. Additionally, sequences of vowel+coda /l/ are found to map to [ɑu], [ou], or [wo] sometimes (e.g. *Adolf* → [a.tɑu.fu], *Albany* → [ɑu.pa.ni], *Rachael* → [ʐʷei.tɕʰou], *Hazel* → [xai.ʐou], *McConnell* → [mai.kʰɑŋ.nʷo], *Wiesel* → [wei.sʷo]). When this occurs, the coda /l/ is obviously vocalized. Because we are interested in the adaptation of coda liquids as [ər] when preserved, these mappings are omitted. Other mappings that are also omitted include (i) those showing no adaptation (i.e. vowel+/ə/+/ɹ/ → vowel+[ər]; e.g. *Power* → [pɑu.ər], *Brewer* →

[pu.lu.ər]), (ii) those in which the English orthography obviously plays a role (e.g. *Powell* → [pɑu.wei.ər], *Cowell* → [kɑu.wei.ər]),[55] and (iii) those in which the preceding vowels in the source occur only a few times in my corpus (e.g. /ʊ/ occurs only twice before a coda liquid in my corpus). The last set of mappings is not considered because we cannot determine whether the liquids tend to be preserved or deleted.

8.1.1 Coda /ɹ/

Now consider the data in Table 8.1, which show the frequencies of preservation and deletion of coda [ɹ] after different vowels. Examples are provided in (59).

Table 8.1 Frequencies of preservation and deletion of coda /ɹ/ as a function of the backness status of the preceding vowel in English loanwords in SC (front vowel+/ɹ/ vs. central vowel+/ɹ/: χ^2 = 153.94, p < .00001; front vowel+/ɹ/ vs. back vowel+/ɹ/: χ^2 = 86.27, p < .00001; central vowel+/ɹ/ vs. back vowel+/ɹ/: χ^2 = 1.15, p = .28, not significant at p < .1)

	front vowel				central vowel		back vowel			
	/ɪ/	/ɛ/	sum		/ə/		/ɔ/	/ɑ/	sum	
preserve	9	8	17	85%	6	2.43%	2	4	6	4.62%
delete	1	2	3	15%	231	97.47%	48	76	124	95.38%

(59) Examples of preservation and deletion of coda /ɹ/ as a function of the backness status of the preceding vowel in English loanwords in SC

		Preservation	Deletion
i.	After a front vowel		
	/ɪ/	*Pierce* → [pi.ər.sz̩]	*Goodyear* → [ku.tʰɤ.ji]
		Beard → [pi.ər.tɤ]	
	/ɛ/	*Blair* → [pu.lai.ər]	*Javier* → [tɕa.wei.je]
		O'Hare → [ou.xai.ər]	*Mayer* → [mei.je]

[55] Many of the cases are adaptation of vowel+/ə/+/l/ sequences. It is reasonable to expect these sequences to map to vowel+[ər], just like vowel+/ə/+/ɹ/ sequences being mapped to vowel+[ər]. These mappings are excluded because I suspect that the adapters analyze words such as *Powell* and *Cowell* as *Po-well* and *Co-well*, respectively. My suspicion originates from the observation that the word *Powell* is also adapted as [pɑu.ər]. In this mapping, the adaptation is in line with the expectation. Other cases are simply judged by the author.

ii. After a central vowel

/ə/ *K<u>ir</u>k* → [kʰ<u>ɤ.ər</u>.kʰɤ] *Bak<u>er</u>* → [pei.kʰ<u>ɤ</u>]

B<u>er</u>nard → [p<u>ei.ər</u>.na] *Hepb<u>ur</u>n* → [xɤ.p<u>ə</u>n]

iii. After a back vowel

/ɔ/ *H<u>or</u>ton* → [xʷ<u>o.ər</u>.tʰəŋ] *N<u>or</u>ton* → [nʷ<u>o</u>.tʷən]

C<u>or</u>bett → [kʰ<u>ou.ər</u>.pei.tʰɤ] *P<u>or</u>tland* → [pʷ<u>o</u>.tʰɤ.lan]

/ɑ/ *<u>Ar</u>kin* → [<u>a.ər</u>.tɕin] *M<u>ar</u>k* → [m<u>a</u>.kʰɤ]

G<u>ar</u>th → [tɕ<u>a.ər</u>.ʂʐ̩] *St<u>arr</u>* → [ʂʐ̩.t<u>a</u>]

In American English, vowels may merge before /ɹ/ (Lehiste 1964). Specifically, /i/ and /ɪ/ merge, /e/, /ɛ/, and /æ/ merge, /u/ and /ʊ/ merge, and /o/ and /ɔ/ merge. In Table 8.1 and (59), I simply use /ɪ/, /ɛ/, and /ɔ/ to represent the high front vowels, the non-high front vowels, and the non-high back rounded vowels, respectively. As we can see, the coda /ɹ/ tends to be preserved after a front vowel but tends to delete after a central or back vowel.

8.1.2 Coda /l/

The data on coda /l/ are given in Table 8.2 with examples shown in (60).

Table 8.2 Frequencies of preservation and deletion of coda /l/ as a function of the backness status of the preceding vowel in English loanwords in SC (front vowel+/l/ vs. central vowel+/l/: $\chi^2 = 3.98$, $p < .05$; front vowel+/l/ vs. back vowel+/l/: $\chi^2 = 8.67$, $p < .005$; central vowel+/l/ vs. back vowel+/l/: $\chi^2 = .0672$, $p = .80$, not significant at $p < .1$)

		preserve	delete
front vowel	/i/	4	5
	/ɪ/	18	10
	/eɪ /	2	1
	/ɛ/	18	10
	/æ/	3	11
	sum	45	37
		54.88%	45.12%
central vowel	/ə/	6	14
		30%	70%

Table 8.2 (cont'd)

back vowel	/oʊ/	6	11
	/ɔ/	4	15
	/ɑ/	1	4
	sum	11	30
		26.83%	73.17%

(60) Examples of preservation and deletion of coda /l/ as a function of the backness status of the preceding vowel in English loanwords in SC

		Preservation	Deletion
i.	After a front vowel		
	/i/	*O'Neil* → [ou.ni.ər]	*Copperfield* → [kʰɑu.pʷo.fei]
		Steele → [ʂʐ̩.ti.ər]	*Spielberg* → [ʂʐ̩.pʰi.pʷo]
	/ɪ/	*Gilbert* → [tɕi.ər.pʷo]	*Stilwell* → [ʂʐ̩.ti.wei]
		Hilton → [ɕi.ər.tʷən]	*Wilks* → [wei.kʰɤ.sʐ̩]
	/e/	*Bloomingdale* → [pu.lu.miŋ.tai.ər]	*Lauderdale* → [lʷo.tɤ.tai.pɑu]
		Wales → [wei.ər.sʐ̩]	
	/ɛ/	*Elton* → [ai.ər.tʷən]	*Feldman* → [fei.tɤ.mən]
		Welch → [wei.ər.ɕy]	*Phelps* → [fei.pʰu.sʐ̩]
	/æ/	*Malcolm* → [mai.ər.kʰan]	*Calhoun* → [kʰa.xuŋ]
		Albion → [a.ər.pin]	*Albert* → [ja.pʷo]
ii.	After a central vowel		
	/ə/	*Trumbull* → [ʈʂʷan.pu.ər]	*Michael* → [mai.kʰɤ]
		Hagel → [xai.kɤ.ər]	*Nicholson* → [ni.kʰɤ.sən]
iii.	After a back vowel		
	/o/	*Bowles* → [pɑu.ər.sʐ̩]	*Goldie* → [kɤ.ti]
		Olson → [ou.ər.sən]	*Nicole* → [ni.kʰɤ]

/ɔ/ *B<u>al</u>timore* → [p<u>a.ər</u>.ti.mʷo] *<u>Al</u>tman* → [<u>a</u>.tʰɤ.man]

K<u>ol</u>be → [kʰ<u>ɤ.ər</u>.pi] *H<u>all</u>mark* → [x<u>ɤ</u>.ma.kʰɤ]

/ɑ/ *g<u>ol</u>f* → [k<u>ɑu.ər</u>.fu] *<u>Al</u>varado* → [<u>a</u>.wa.la.tʷo]

The data in Table 8.2 show a similar asymmetry except for the /i+l/ and /æ+l/ sequences. A careful check of the data indicates that four of the five /i+l/ instances where the /l/ is deleted are embedded in the morpheme *-field*. In addition, there are only five English words containing the morpheme *-field* in my corpus. Thus, it seems (i) that as long as a coda /l/ appears in the morpheme *-field*, it gets deleted, and (ii) the /l/'s in other /i+l/ instances are almost always preserved, which is consistent with most sets of the data. A possible explanation is that if the /l/ in the morpheme *-field* is preserved, the SC output form will be long in terms of syllable count.[56] Although I did not discuss the effect of output form length on the preservation vs. deletion of non-prevocalic stops in 4.3 and 5.1, I have observed that, in many cases, the SC output forms that are derived via deletion would consist of more syllables than other output forms if the deletion did not occur. The adaptation of words containing a /i/+coda /l/ sequence exhibits the same asymmetry. Consider the data in Table 8.3.

Table 8.3 Average numbers of syllables in SC forms that are derived from two sets of English words containing a /i/+coda /l/ sequence

		Average number of syllables in the SC output forms if the /l/'s are deleted	Average number of syllables in the SC output forms if the /l/'s are preserved
i	English words containing the morpheme *-field* e.g. *Copperf<u>iel</u>d*	3.8 *Copperf<u>iel</u>d* → [kʰɑu.pʷo.f<u>ei</u>] (attested)	4.8 *Copperf<u>iel</u>d* → [kʰɑu.pʷo.f<u>ei.ər</u>] (hypothetical)
ii	English words containing a /i/+coda /l/ sequence that is not in the morpheme *-field* e.g. *O'N<u>eil</u>*	2.25 *O'N<u>eil</u>* → [ou.n<u>i</u>] (hypothetical)	3.25 *O'N<u>eil</u>* → [ou.n<u>i.ər</u>] (attested)

[56] Notice that if the /l/ is deleted, usually the word-final /d/ is also deleted.

As we can see, in each column the number in the upper cell is significantly larger than the number in the bottom cell, and only when the /l/'s in the English words in (i) are deleted and the /l/'s in the English words in (ii) are preserved, are the average numbers of syllables in the two sets of output forms closer to each other (shown in the two grey cells). The data demonstrate that output form length is a factor in the preservation vs. deletion of coda /l/. This argument is strengthened by the mapping ***F**ie**l**d*** → SC [fei.ər.tɤ].[57] In this mapping, the /l/ in the source word is preserved and the SC output form comprises three syllables. I conclude that if the factor of output form length is controlled, coda /l/ will tend to be preserved after /i/.

Phonologically, /æ/ is a front vowel. However, /æ/ is not as front as other front vowels acoustically, as the F2 space for low vowels is more limited than the F2 spaces for non-low vowels. Moreover, as I will argue in 8.2, whether a coda liquid tends to be preserved or delete correlates with the tongue shape with which it is produced. Specifically, if the tongue shape is similar to the tongue shape for the preceding vowel, it tends to delete; if not, it tends to be preserved. The articulatory and perceptual evidence I will present indicates that, among the front vowels, /æ/ exhibits a tongue shape that is most similar to that for /l/. This is probably the reason why the /l/ in a /æ/+coda /l/ sequence does not join the /l/'s in other front vowel+coda /l/ sequences in their tendency to be preserved.

8.2 Phonetic Motivations

I argue that the adaptation asymmetry shown in Tables 8.1 and 8.2 has a perceptual basis. On the grounds of the P-map theory (Steriade 2001, 2009), I hypothesize the two perceptual similarity hierarchies in (61) and (62). ("Δ" denotes difference, "V" a vowel, and "L" a liquid consonant.)

[57] This mapping is not included in my corpus. I found this mapping after I had completed the statistical analyses.

(61) Δ V_{front}+L]$_\sigma$ vs. V]$_\sigma$ > Δ V_{front}+L]$_\sigma$ vs. V+ər]$_\sigma$

Within a syllable, a front vowel+liquid sequence is perceptually less similar to a vowel than to a vowel+[ər] sequence.

(62) Δ $V_{non\text{-}front}$+L]$_\sigma$ vs. V+ər]$_\sigma$ > Δ $V_{non\text{-}front}$ +L]$_\sigma$ vs. V]$_\sigma$

Within a syllable, a non-front vowel+liquid sequence is perceptually less similar to a vowel+[ər] sequence than to a vowel.

I propose that, in English, coda liquids generally are more distinctive after a front vowel than after a central or back vowel in perception, and that this difference leads to the two perceptual similarity hierarchies. A good example of the proposed perceptual difference in distinctiveness is the percept of an extra syllable in words such as *heel* and *hire* (Gick and Wilson 2006) as opposed to the monosyllabic perception of words such as *hall* and *hard*. Obviously, an underlyingly unattested schwa is heard between the vowel and liquid in the former set of words. The relevance of this schwa percept to the proposed perceptual difference in distinctiveness will be briefly discussed at the end of this section.

I will argue in 8.2.1 that the proposed perceptual difference in distinctiveness is the results of the tongue moving from a vowel to a following tautosyllabic liquid. In 8.2.2, I will report the results of two perceptual experiments that aim to test the proposal.

8.2.1 Articulatory Gestures and Perceptual Distinctiveness

After summarizing the phonetic literature on the component gestures of /l/ and /r/ and the inter-gestural coordination patterns in 8.2.1.1, I discuss the syllable /Cəɹ/ in 8.2.1.2 and other vowel+coda liquid sequences in 8.2.1.3 and present evidence from previous research showing why the liquid is perceptually salient or non-salient. In 8.2.1.4, I discuss the relevance of my proposal to the schwa percept during the articulation of high tense vowel+coda liquid sequences (e.g. *heeᵊl*, *hiᵊre*).

8.2.1.1 Inter-Gestural Coordination

The articulation of /l/ involves a tongue tip raising gesture and a tongue root/dorsum retraction gesture, and the relative timing of these two gestures is found to be dependent on the /l/'s position in the syllable (Sproat and Fujimura 1993; Narayanan, Alwan, and Haker 1997; Gick 1999, 2002, 2003; Gick, Kang, and Whalen 2000, 2002; Gick and Wilson 2006). Sproat and Fujimura (1993), for example, reported that while the tongue root/dorsum retraction gesture temporally follows the tongue tip raising gesture when /l/ is in onset, a reversed timing occurs when /l/ is in coda. In addition, in a study which attempts to argue that a more efficient cognitive model can be created if a single gesture can be shown to be shared by a number of segments, Gick, Kang, and Whalen (2000, 2002) reported that the tongue root/dorsum retraction gesture of /l/ is very similar to the dorsal gesture of /ɔ/.

The articulation of /ɹ/ involves multiple gestures as well (Delattre and Freeman 1968; Hagiwara 1995; Alwan, Narayanan, and Haker 1997; Gick 1999; Gick, Kang, and Whalen 2000, 2002; Gick and Campbell 2003; Campbell, Gick, Wilson, and Vatikiotis-Bateson 2010), which, according to Gick and his colleagues, include a tongue blade/body raising gesture and a tongue root retraction gesture (and possibly an additional lip constriction gesture; e.g. Gick and Campbell 2003). Gick and his colleagues also found that, like /l/, the relative timing of the composing gestures of /ɹ/ varies as a function of the /ɹ/'s position within the syllable. For example, Campbell, Gick, Wilson, and Vatikiotis-Bateson (2010) reported that while the tongue root retraction gesture takes place slightly later than the tongue blade/body raising gesture in the onset allophone, the tongue root retraction gesture takes place significantly earlier than the tongue blade/body raising gesture in the coda allophone.[58] Moreover, Gick, Kang, and Whalen (2000, 2002) reported a connection between /ɹ/ and /ə/. First, they argued that /ə/ is not targetless by showing that the pharyngeal region of the vocal tract is in fact constricted (resulting from the retraction of the tongue root) during the production of /ə/. Second, they compared the pharyngeal/uvular configurations during the production of /ɹ/ and /ə/ and found that the two configurations are very similar to each other. The connection between /ɹ/ and /ə/ was also observed by McMahon (1996, 2000), who stated that "Spectrograms for schwa and approximant [r] indicate that the spectral shapes for the two sounds are rather similar, except that F3 for [r] is kept low by some complex articulatory manoeuvres" (1996: 80).

These articulatory observations have implications regarding the perceptual difference proposed earlier. In what follows, I will discuss the adaptation of the syllable /Cəɹ/ and then other vowel+coda liquid sequences.

8.2.1.2 /Cəɹ/ and Final Reduction

In English, the syllable /Cəɹ/ is realized as [Cɹ̩] (e.g. Keyser and Stevens 1994: 218). Since the articulation of /ɹ/ involves a tongue blade/body raising gesture and a tongue root retraction gesture, with the latter temporally preceding the former in the coda allophone, the realization /Cəɹ/ → [Cɹ̩] can be regarded as the underlying schwa merging with the schwa component of the following /ɹ/. As shown in the data in Table 8.1, /Cɹ̩/ (=/Cəʳ/, with /ə/ and /ʳ/ representing the schwa component and the tongue blade/body raising gesture of /ɹ/) in the source words is almost always mapped to [Cɤ] (231/237; 97.47%) (e.g. *Bak<u>er</u>* → [pei.kʰ<u>ɤ</u>]). In these mappings, the schwa component of the syllabic /ɹ/ apparently corresponds to [ɤ]. There are reasons why the tongue blade/body raising gesture [ʳ] rarely has a correspondent in the SC output forms. First,

[58] Gick and Campbell's (2003: 1913) data showed that all the three composing gestures take place almost simultaneously in the coda allophone.

when a multi-gesture segment occurs in final position, the gesture occurring at the periphery is reduced in magnitude (i.e. final reduction; Browman and Goldstein 1992, 1995; Sproat and Fujimura 1993). Sproat and Fujimura (1993) called this gesture "the consonantal gesture", because this gesture is found to always result in a narrower constriction than the one that is not reduced/occurs more medially (i.e. the vocalic gesture). When the consonantal gesture is reduced to an extreme degree, the segment is vocalized (Gick 1999). I assume that final reduction happens in all coda consonants in the speech of SC talkers (cf. He 2014), as, in SC, a syllable is also a morpheme and a morpheme possibly represents a word. My assumption is in line with Z. Wang's (1997) argument that, in Mandarin, coda nasals are not stops but glides. In the present case, the gesture that is reduced is the tongue blade/body raising gesture, suggesting that the tongue blade/body raising gesture may not be in the input or very weak. The second reason is that the tongue blade/body raising gesture is not perceptually salient. As opposed to other consonantal gestures such as those found in fricatives and stops, the constriction resulting from the tongue blade/body raising gesture is relatively open. Consequently, neither frication noise nor release bursts are created. This gesture is simply the tongue blade/body approaching the hard palate, yielding the percept of an r-color superimposed on the schwa component (cf. Delattre and Freeman 1968).

8.2.1.3 Other Vowel+Coda Liquid Sequences

Now we consider the adaptation of other vowel+coda liquid sequences. Different vowels are produced with different tongue positions (precisely, tongue shapes). For example, /i/ and /ɑ/ are said to have a high front tongue position and a low back tongue position, respectively. When two vowels with similar tongue positions are produced one after the other, the acoustic change occurring at the boundary is subtle, as the tongue shape specified for the first vowel does not undergo much transformation to become the one required by the second vowel. On the other hand, the acoustic change will be great if the two vowels have very different tongue positions, since, for the tongue to acquire the shape required by the second vowel, it must transform a great deal from the original shape.

The magnitude of acoustic change occurring at the boundary of two adjacent sounds affects the perceptual distinctiveness of the second sound. Arguably, the greater/smaller the change, the more/less distinctive the second sound. Take sibilants for example. Although sibilants are perceptually salient, when one immediately follows another (e.g. [s]+[ʃ]), the second sibilant probably cannot be distinguished from the first sibilant easily, and the result could be merger. In contrast, if the first sibilant is replaced by silence, then the second sibilant becomes very distinctive due to a large acoustic change.

Turning back to vowel+coda liquid sequences. While liquids are consonants, /ɹ/ can be thought of as an r-colored schwa and /l/ can be thought of as /ɔ/ produced with an additional tongue tip raising gesture. In addition, in these sequences, the /ə/ and /ɔ/ components of the liquids immediately follow the vowels, with the consonantal gestures occurring at the periphery and likely undergo reduction. As a result, the perceptual distinctiveness of the liquids in these sequences is closely related to the degree of change in tongue shape from the vowels to the vowel components of the liquids. If the change is large, the liquids are distinctive and likely preserved; if the change is small, the liquids are less distinctive, and the result may be deletion.

I do not have direct quantitative data on the tongue shapes for vowels. However, the values for the three tongue position-related features of vowels, namely, [high], [low], and [back], shed some light on how different /ə/ and /ɔ/ are from other vowels. The values and "distances" are given in Table 8.4. (/i/, /æ/, and /u/ do not appear before /ɹ/ so the distances between these vowels and /ə/ are marked as n/a. The numerals in the last two rows represent the distances, which are based on counts of change in the three values.)

Table 8.4 Distances from various vowels to /ə/ and /ɔ/ in terms of the values for [high], [low], and [back] in English

	i	ɪ	e	ɛ	æ	ə	u	ʊ	o	ɔ	ɑ
[high]	+	+	–	–	–	–	+	+	–	–	–
[low]	–	–	–	–	+	–	–	–	–	+	+
[back]	–	–	–	–	–	+	+	+	+	+	+
distance from /ə/ (=/ɹ/)	n/a	2	1	1	n/a	0	n/a	1	0	1	1
distance from /ɔ/ (=/l/)	3	3	2	2	1	1	2	2	1	0	0

The distances in last two rows indicate (i) that the vowel that is most different from /ə/ is /ɪ/, and four of the six vowels that are similar to /ə/ are back vowels, and (ii) that /ɔ/ is most different from /i/ and /ɪ/ and most similar to /ɑ/, then /o/, /ə/, and /æ/. Though not perfect, the bigger picture is that /ə/ and /ɔ/ are generally more similar to non-front vowels than to front vowels. Notice that /æ/ is the front vowel that is most similar to /ɔ/. Also notice that /æ/ is even more similar to /ɔ/ than the two back vowels /u/ and /ʊ/. These are in line with the loanword data which show that coda /l/'s tend not to be preserved after /æ/.

Also, Whalen, Kang, Magen, Fulbright, and Gore (1999) assigned height and frontness values to some vowels. The values and "distances" from /ʌ/ and /ɔ/ to other vowels are given in Table 8.5. I assume that /ʌ/ can represent /ə/ (see below). (The distances are calculated by adding up the differences in height and frontness.)

Table 8.5 Distances from /ʌ/ and /ɔ/ to other vowels in terms of the height and frontness values assigned by Whalen, Kang, Magen, Fulbright, and Gore (1999: 597)

	i	e^j	ɑ	ʌ	ɔ	o^w
height	5	3	1	2	2	3
frontness	1	1	2	2	3	3
distance from /ʌ/	4	2	1	0	1	2
distance from /ɔ/	4	3	2	1	0	1

Table 8.5 shows that the vowel that is most different from /ʌ/ and /ɔ/ is /i/, and that the vowels that are most similar to /ʌ/ and /ɔ/ are all non-front vowels (/ɑ/ and /ɔ/ for /ʌ/; /ɑ/, /ʌ/, and [oʷ] for /ɔ/). We see a similar picture.

Some type of quantitative data is available. It has been claimed in the literature that the pharyngeal widths measured during the production of vowels correlate with the vowels' height and backness. For example, Whalen, Kang, Magen, Fulbright, and Gore (1999) measured the midsagittal widths of the air space along the entire vocal tract of two American English speakers during their production of eleven vowels, and a correlation analysis showed that the widths in the pharyngeal region could be predicted from the widths measured at the four spots on the tongue where receiver coils are usually placed in electromagnetometor experiments (the positions of the four spots are used by phoneticians to represent the tongue shapes). A similar suggestion was made in an earlier report. MacKay (1977) measured the distance between the external neck wall and the anterior pharyngeal wall during the production of ten English vowels using ultrasonic imaging and found that the measurements correlate with the vowels' height and backness. Table 8.6 gives MacKay's (1977) measurements:

Table 8.6 Distance between the external neck wall and the anterior pharyngeal wall during the production of ten English vowels (MacKay 1977: 341)

vowel	i	u	e	ɪ	ɛ	æ	ʊ	o	ʌ	ɑ
mean	4.03	4.30	4.35	4.69	4.91	4.93	5.14	5.15	5.35	5.58

Table 8.6 shows that a vowel generally has a smaller value than another vowel that is lower but has the same specification for backness (e.g. /i/ vs. /e/), and that a front vowel always has a smaller value than a non-front vowel that has the same specification for height (e.g. /e/ vs. /o/). These indicate that the pharyngeal widths indeed correlate with the vowels' height and backness (the larger the value, the smaller the pharyngeal width). Though the data do not include the measurements for /ɔ/, given that /ɔ/ is a low back vowel, the value for /ɔ/ probably falls

between those for /o/ and /ɑ/. Moreover, I assume that the value for /ʌ/ represents that for /ə/. /ə/ is usually substituted by /ʌ/ in phonetic measuring, and the reason is that it is easier for the talkers to sustain /ʌ/ as imaging requires at least a certain amount of time to complete. According to Gick, Kang, and Whalen (2000, 2002), there is no significant difference between /ə/ and /ʌ/ when it comes to phonetic measuring.

As we can see in the table, the vowels that appear on the right side are all non-front vowels and the vowels that appear on the left side are all front vowels except for /u/, indicating that /ə/ and /ɔ/ are generally more similar to non-front vowels than to front vowels in terms of tongue shape.

As mentioned above, while /æ/ is considered a front vowel phonologically, it is not as front as other front vowels phonetically. In addition, the tongue position is quite close to that for /ɔ/. These are compatible with the data in Table 8.6, which shows that /æ/ is the front vowel that is closest to the back vowels in terms of tongue shape.

In Table 8.6, /u/ appears near the left end and /ʊ/ appears sort of in the middle. I have conducted two experiments to examine the relative perceptual distinctiveness of coda liquids after various vowels. Unfortunately, the /ʊ/+coda liquid sequences are not included in the experiment design, but the results of the experiments, which I will show in the next section, indicate that coda /l/ is more distinctive after /u/ than after other non-front vowels.

8.2.1.4 Gick and Wilson (2006) and my Proposal

Before closing this section, I would like to briefly discuss the relevance of the schwa percept during the articulation of high tense vowel+coda liquid sequences (e.g. *heeᵊl*, *hiᵊre*) to my proposal. Gick and Wilson (2006) found that the schwa percept is ‘the incidental result of the tongue passing through a schwa-like configuration or “schwa space” during the transition between opposing tongue root targets' (pp. 635). This finding is entirely compatible with my proposal. First, we can infer from MacKay's (1977) data that the vowels that are most different from /ə/ in tongue shape are high tense vowels. Such a large difference in tongue shape makes the excrescent schwa in a high tense vowel+coda liquid sequence perceptually very distinctive, and this is probably the reason why the sequence is perceived as containing an extra syllable. When an English word containing a high tense vowel+coda liquid sequence is borrowed into SC, the perceptually distinctive excrescent schwa is preserved along with the following liquid, yielding the SC syllable [ər]. Second, tongue root targets correlate with tongue shapes (MacKay 1977; Whalen, Kang, Magen, Fulbright, and Gore 1999). When an English word containing a vowel+coda liquid sequence that does not give rise to an excrescent schwa is borrowed into SC, two situations happen. If the liquid tends to be preserved (e.g. after /ɪ/ or /ɛ/), the tendency is attributed to the high degree of perceptual distinctiveness of the liquid, which is the outcome of the conflicting tongue root targets for the liquid and the preceding vowel (as opposed to the conflicting tongue root targets for the excrescent schwa and the preceding vowel). If the liquid tends to delete (e.g. following /ɑ/), then it is because the tongue root retraction gestures of the liquid and the preceding vowel are similar (i.e. having similar targets), which makes the liquid perceptually indistinctive. In summary, whether an excrescent schwa occurs, it is the extent of perceptual distinctiveness of the component after the vowel (i.e. /ə/+liquid or simply the liquid) that determines the adaptation tendency.

8.2.2 Perceptual Experiments IV and V

Two perceptual experiments are conducted. One (Experiment IV) focuses on /ɹ/ and the other (Experiment V) on /l/.

8.2.2.1 Experiment IV: Perceptual Distinctiveness of Coda /ɹ/ in English

The goal of this experiment is to discover the relative perceptual distinctiveness of coda /ɹ/ after various vowels in English.

I assume that, perceptually, the extent of distinctiveness of a sound can be represented by the extent of saliency of the contrast between the sound and silence. Since the extent of distinctiveness of a sound depends on the context, the extent of saliency of the contrast between the sound and silence depends on the context as well. For example, if a consonant C is more distinctive in V_1C than in V_2C, then the contrast between V_1C and V_1 is more salient than the contrast between V_2C and V_2. The design of this experiment is based on this assumption.

8.2.2.1.1 Stimuli

This experiment was a classic same-different AX task. The AX paradigm used in this task included five pairs of nonce disyllabic words in the form of /hVɹ.gV/–/hV.gV/. /i/, /e/, /o/, /ɑ/, or /ə/ was used in the first syllables. When the vowel in the first syllables was /i/, /e/, or /o/, it was also used in the second syllables (i.e. /hiɹ.gi/–/hi.gi/, /heɹ.ge/–/he.ge/, /hoɹ.go/–/ho.go/). When the vowel in the first syllables is /ɑ/ or /ə/, /i/ is used in the second syllables (i.e. /hɑɹ.gi/–/hɑ.gi/, /həɹ.gi/–/hə.gi/).

All the words, including those used in the fillers, were written in IPA symbols. A female native American English speaker was recruited for recording. She grew up in the suburb of Chicago and, at the time of the recording, was a doctoral student in linguistics at Michigan State University. She was asked to read the words in the frame sentence “I say __ twice” at a normal speed in front of a microphone that was connected to a high-quality tape recorder. The stress was put on the first syllables unless the vowel is /ə/. When the vowel is /ə/, the stress was put on the second syllables. The recording took place in a quiet room.

The tape was digitized into a computer. The rimes of the first syllables were adjusted until they had the same duration. The resulting words were paired up, and 68 pairs were created. These pairs were classified into three categories: Category I included five pairs of /hVɹ.gV/–/hV.gV/ and five pairs of /hV.gV/–/hVɹ.gV/, Category II included five pairs of /hVɹ.gV/–/hVɹ.gV/ and five pairs of /hV.gV/–/hV.gV/, and Category III included the other 48 pairs, which served as the fillers. A one-second silence was added between the two words in each pair, and white noise was also added. The noise started 0.5 second before the first word and ended 0.5 second after the second word. I made four repetitions of the pairs classified into Categories I and II, amounting to a total of 128 sound files (20 pairs × 4 repetitions + 48 fillers). The sound files were saved in a random order.

8.2.2.1.2 Subjects

The subjects were 22 Taiwan SC speakers. Their age ranged from 18 to 46 years, and most of them were students at Michigan State University at the time of the experiment. They learned English as a foreign language in Taiwan before coming to the U.S. and had stayed in the U.S. from a few weeks to 11 years.

8.2.2.1.3 Procedure

The experiment was conducted in a quiet room. The subjects either came alone or as a group no bigger than 6 people. A brief instruction and practice were given before the experiment started. The subjects were instructed to listen carefully to the stimuli, which were played from two speakers about two meters away from them, and judge whether the two words in each pair sounded identical by selecting *same* or *different* on an answer sheet. Each session lasted about 30 minutes.

8.2.2.1.4 Results

The subjects are expected to make more errors with a pair in the paradigm if the /ɹ/ in the word /hVɹ.gV/ is perceptually less distinctive from the preceding vowel. Thus, the numbers of errors made by the subjects with the five pairs represent the relative perceptual distinctiveness of English coda /ɹ/ after the five vowels. A subject heard a pair in the paradigm eight times, amounting to a total of 880 tokens to be analyzed (5 pairs × 2 orders × 4 repetitions × 22 subjects). I gave one point to a token if the correct answer *different* was selected but no point if the wrong answer *same* was selected. Thus, the highest score a subject can get for a pair in the paradigm is eight points (s/he heard a pair in the paradigm eight times). The mean scores for the five test pairs are shown in Table 8.7. (The higher the score, the more distinctive the /ɹ/ after the vowel.)

Table 8.7 Mean scores for the test pairs in an AX discrimination task examining the relative perceptual distinctiveness of coda /ɹ/ after various vowels in English

vowel used in the first syllables	i	e	o	ə	ɑ
mean score	7.86	7.82	7.64	4.14	4.77

The results of an ANOVA analysis show a significant difference ($p < .05$), so multiple comparisons are carried out. The results of the comparisons are given below. (An asterisk indicates a significant difference between the mean scores associated with the vowel contexts at a .05 confidence level.)

Table 8.8 Multiple comparisons of the mean scores in Table 8.7

	i	e	o	ə	ɑ
i				*	*
e				*	*
o				*	*
ə	*	*	*		
ɑ	*	*	*		

Table 8.8 shows (i) that the mean scores associated with the /i/, /e/, and /o/ contexts are significantly different from the mean scores associated with the /ə/ and /ɑ/ contexts, (ii) that the

mean scores associated with the /i/, /e/, and /o/ contexts are not significantly different from each other, and (iii) that the mean scores associated with the /ə/ and /ɑ/ contexts are not significantly different from each other.

8.2.2.1.5 Discussion

The fact that the mean score associated with the /o/ context is not significantly different from those associated with the /i/ and /e/ contexts is not expected, as we can infer from MacKay's (1977) data that English coda /ɹ/ is less distinctive after /o/ than after /i/ or /e/.

Lehiste (1964) reported that /o/ and /ɔ/ are neutralized before coda /ɹ/ in the northern Midwestern dialect of American English and that, in this dialect, /o/ followed by coda /ɹ/ is in fact realized as a sound that is very similar to /ɔ/. Furthermore, the grapheme *o* is almost always transcribed as /ɔ/ before coda /ɹ/ in English dictionaries, especially when the sequence appears in a syllable that is not on the right edge of the word (e.g. Webster's Third New International Dictionary 1971; Chambers English Dictionary 1988). For example, although the vowels in *store* and *port* are transcribed as /o/ and /ɔ/ in some dictionaries, the second vowel in *coordinator* and the first vowel in *forward* are only transcribed as /ɔ/ in all dictionaries. The talker who I recruited for recording pointed out that the sequence /o+ɹ/ was somewhat unnatural to produce. Note that she grew up in the suburb of Chicago (therefore a speaker of the northern Midwestern dialect of American English) and was asked to read words written in IPA symbols. Also, this sequence appears in the first syllable of a disyllabic word (/hoɹ.go/). As a result, the talker probably was not comfortable with this sequence and produced it in an unnatural way. I checked the stimuli and found that the stimulus /hoɹ.go/ was obviously louder than the other stimuli and that the /ɹ/ in the stimulus sounded discrete, as though it occurred word-finally.

Although the next experiment is about the relative perceptual distinctiveness of coda /l/ following various vowels, I have included stimuli that contain the /o/+/ɹ/, /ɪ/+/ɹ/, /ɛ/+/ɹ/, and /ɔ/+/ɹ/ sequences, as in the present experiment the stimulus /hoɹ.go/ sounds unnatural and the extent of perceptual distinctiveness of coda /ɹ/ after /ɪ/, /ɛ/, and /ɔ/ is not examined.

8.2.2.2 Experiment V: Perceptual Distinctiveness of Coda /l/ in English

The main goal of this experiment is to discover the relative perceptual distinctiveness of coda /l/ after various vowels in English. As mentioned in 6.3.2.1, this experiment is combined with Experiment I. That is, the stimuli in one experiment serve as the fillers in the other experiment. As a result, the two sets of stimuli are produced by the same talkers in the same fashion, the subjects who participate in one experiment also participate in the other experiment, and the two experiments are conducted using the same procedure. For the sake of convenience, how the stimuli are constructed and the description of the subjects and experiment procedure are repeated below.

8.2.2.2.1 Stimuli

This experiment is an AXB discrimination task, in which the subjects listen to the stimuli and for each stimulus judge whether the A or B item sounds more similar to the X item. The AXB paradigm used in this task includes 14 triplets, in which the X items are disyllabic nonce English words and the A and B items are the X items' SC matches. An example triplet is given below.

(63) A triplet from the AXB paradigm used in Experiment V

A (SC)	X (E.)	B (SC)
[ti.pi]	/díl.bi/	[ti.ər.pi]

The first syllable of the X item is stressed and comprises an initial obstruent and a /il/, /ɪl/, /el/, /ɛl/, /æl/, /ul/, /ol/, /ɔl/, /ɑl/, /ʌl/, /ɪɹ/, /ɛɹ/, /oɹ/, or /ɔr/ rime, and the second syllable is always /bi/. One of the SC matches is disyllabic and the other is trisyllabic with [ər] as the second syllable. Everything else in the two SC matches is equal. See Appendix D for 14-triplet the paradigm.

A male native American English speaker produced the X items and a male native Taiwan SC speaker produced the A and B items. Both talkers were doctoral students in linguistics at Michigan State University at the time of recording. They read the words in a carrier sentence ("I say __ twice" for the English words and "wo shuo __ san ci" [I say __ three times] for the SC words) at a normal speed in front of a microphone that was connected to a high-quality tape recorder. After the tapes were digitized into a computer, the target words were grouped into triplets, and 112 triplets were created: 14 triplets in the paradigm × two orders of the A and B items × two repetitions. A one-second silence was added at the beginning of each triplet, and a half-second silence was added between the A and X items and between the X and B items. The 112 triplets were mixed with 104 fillers, amounting to a total of 216 sound files arranged in a random order.

It is worth noting that the American English talker who I recruited for recording grew up in the suburb of Cincinnati, Ohio, and he claimed that he was comfortable producing a word-medial /o/+/ɹ/ sequence. I checked the stimuli and the stimulus containing this sequence sounded as normal as the other stimuli.

8.2.2.2.2 Subjects

Twenty-nine native SC speakers (14 males, 15 females) volunteered to participate in the experiments. They were all from Taiwan and in their late twenties to early fifties. None of them was a native English speaker, and they had lived in the U.S. for an average of 5.5 years at the time of the experiment.

8.2.2.2.3 Procedure

This experiment was conducted in a quiet room. The subjects either came alone or in groups no bigger than three people. A brief instruction and practice were given before the experiment started. In the experiment, the subjects listened to the stimuli and for each stimulus judged which SC match, A or B, sounded more similar to the English word X by circling the letter A or B on an answer sheet. There was a short break after the 105th stimulus. This experiment, combined with Experiment I, was conducted before Experiments II and III (see 6.3.2), and the whole process took 35-40 minutes to complete.

8.2.2.2.4 Results and Discussion

I assume that the more distinctive the coda liquid in the English word, the more similar the English word to the trisyllabic SC correspondent. The results of the experiment are shown in Tables 8.9. (The numerals to the left of the percentages indicate the times the subjects made the selection.)

Table 8.9 Results of an AXB discrimination task examining the relative perceptual distinctiveness of English coda liquids after various vowels

	V	σ+[ər]+σ		σ+σ	
CV+/l/.bi	ɪ	221	95%	11	5%
	i	215	93%	17	7%
	e	209	90%	23	10%
	ɛ	207	89%	25	11%
	æ	190	82%	42	18%
	u	152	66%	80	34%
	ɔ	117	50%	115	50%
	ʌ	116	50%	116	50%
	o	109	47%	123	53%
	ɑ	97	42%	135	58%
CV+/ɹ/.bi	ɪ	226	97%	6	3%
	ɛ	225	97%	7	3%
	o	189	81%	43	19%
	ɔ	64	18%	168	72%

The results indicate that coda /l/ is (i) more distinctive after a front vowel than after a non-front vowel, (ii) less distinctive after /æ/ than after any other front vowel, and (3) more distinctive after /u/ than after any other non-front vowel. In general, the results are consistent with MacKay's (1977) data, which imply that, /ɔ/ (representing the vowel component of /l/) is (i) more similar to a non-front vowel than to a front vowel (with the exception of /u/), (ii) more similar to /æ/ than to any other front vowel, and (iii) less similar to /u/ than to any other non-front vowel.

The results also indicate that coda /ɹ/ is (i) most distinctive after /ɪ/ and /ɛ/, (ii) a bit less distinctive after /o/, and (iii) least distinctive after /ɔ/. These results are not entirely consistent with MacKay's (1977) data, because MacKay's (1977) data imply that the tongue shapes for /ə/ (representing the vowel component of /ɹ/), /o/, and /ɔ/ are similar.

To demonstrate that the pharyngeal/uvular configurations associated with /ɹ/ and /l/ are similar to those associated with /ə/ and /ɔ/, respectively, but dissimilar to those associated with other vowels, Gick, Kang, and Whalen (2002) provided plots of superimposed tract shapes for /l/,

/ɹ/, /ɔ/, /o/, and /ʌ/ for one of their subjects (the plot for /o/ is used to illustrate contrast). The plots clearly show similarity of /ɹ/ to /ɔ/ and dissimilarity of /ɹ/ to /o/ in the pharyngeal and uvular regions (Figure 4 on pp. 367), suggesting that the vowel component of /ɹ/ is associated with a tongue shape that is similar to that for /ɔ/ but dissimilar to that for /o/.

Although pharyngeal widths have been claimed to be able to predict the tongue shapes, we do not expect the predictions to be perfect. On the other hand, MRI data such as those Gick, Kang, and Whalen (2002) provided allow us to see the shapes of the vocal tract and make comparisons. On the basis of the different natures of the two types of data, Gick, Kang, and Whalen's (2002) MRI data are probably more reliable than MacKay's (1977) pharyngeal-width data. Accordingly, we could say that the results of the second experiment reflect the ways in which the sounds are produced.

8.3 Formal Analyses

8.3.1 Summary and Constraints

The adaptation asymmetry, the predictions of the articulatory evidence, and the results of the two perceptual experiments are summarized below.

Table 8.10 Summary of the adaptation asymmetry (shown in Tables 8.1 and 8.2), the predictions of the articulatory evidence (shown in Table 8.6) and the results of the two perceptual experiments (shown in Tables 8.7–8.9)

		prec. vowel	loanword	articulation	Experiment I	Experiment II
ɹ	i.	ɪ, ɛ	pres.	pres.	n/a	pres.
	ii.	ə, ɔ, ɑ	del.	del.	del.	del.
	iii.	i, e	n/a	pres.	pres.	n/a
	iv.	æ	n/a	pres.	n/a	n/a
	v.	o	n/a	pres.	pres.	pres.
	vi.	u	n/a	pres.	n/a	n/a
l	vii.	ɪ, ɛ	pres.	pres.	n/a	pres.
	viii.	ə, ɔ, ɑ	del.	del.	n/a	del.
	ix.	i, e	pres.	pres.	n/a	pres.
	x.	æ	del.	pres.	n/a	pres.
	xi.	o	del.	del.	n/a	del.
	xii.	u	n/a	pres.	n/a	pres.

As we can see, the four sets of data are consistent with each other except for the /æ+l/ sequence. This sequence will be discussed at the end of this section.

To formalize the data in Table 8.10, I propose the constraints in (64):

(64) Constraints

i. MAX-EO-liquid / i, ɪ, e, ɛ, u__]$_\sigma$: Do not delete a liquid if it appears in coda position and follows /i/, /ɪ/, /e/, /ɛ/, or /u/.

ii. MAX-EO-liquid / ə, ɔ, ɑ__]$_\sigma$: Do not delete a liquid if it appears in coda position and follows /ə/, /ɔ/, or /ɑ/.

iii. DEP-EO-ə: Do not insert [ə].

These constraints are insufficient in explaining the data in Table 8.10. I will propose a few more constraints in the next subsection.

8.3.2 Constraint Rankings and Tableaux

The constraint ranking in (65) reflects that coda liquids (i) tend to be preserved at the cost of schwa epenthesis after /i/, /ɪ/, /e/, /ɛ/, and /u/ and (ii) tend to delete after /ə/, /ɔ/, and /ɑ/.

(65) MAX-EO-liquid / i, ɪ, e, ɛ, u__]$_\sigma$ >> DEP-EO-ə >> MAX-EO-liquid / ə, ɔ, ɑ__]$_\sigma$

The tableaux in (66) and (67) illustrate the tendencies in Table 8.10 (i) and (vii), and the tendencies in Table 8.10 (ii) and (viii), respectively.

(66) /ɪ/ or /ɛ/+coda liquid (e.g. /ɹ/: *Pierce* → [pi.ər.sz̩], *Blair* → [pu.lai.ər]; /l/: *Gilbert* → [tɕi.ər.pʷo], *Elton* → [ai.ər.tʷən])

/ɪ/ or /ɛ/+coda liquid	MAX-EO-liquid / ɪ, ɛ__]$_\sigma$	DEP-EO-ə
→ V+[ər]		*
V	*!	

(67) /ə/, /ɔ/, or /ɑ/+coda liquid (e.g. /ɹ/: *Baker* → [pei.kʰɤ], *Norton* → [nʷo.tʷən], *Mark* → [ma.kʰɤ]; /l/: *Michael* → [mai.kʰɤ], *Hallmark* → [xɤ.ma.kʰɤ], *Alvarado* → [a.wa.la.tʷo])

/ə/, /ɔ/, or /ɑ/+coda liquid	DEP-EO-ə	MAX-EO-liquid / ə, ɔ, ɑ__]$_\sigma$
V+[ər]	*!	
→ V		*

Although /i/, /u/, and /e/ are laxed before /ɹ/ in American English, they are not in some other varieties. Based on the articulatory data and experiment results, we expect preservation if SC borrows words containing /i+ɹ/, /u+ɹ/, or /e+ɹ/ from these varieties. In addition, because the /i/+coda liquid and /u/+coda liquid sequences give rise to a schwa percept, the schwa in the outputs is not epenthetic. The tableaux in (68) and (69) illustrate the tendencies in Table 8.10 (iii), (vi), (ix) and (xii):

(68) /i/ or /u/+/ə/+coda liquid (e.g. /ɹ/: no loanword example; /l/: *O'Neil* → [ou.ni.ər])

/i/ or /u/+/ə/+coda liquid	MAX-EO-liquid / i, u, e__]$_\sigma$	DEP-EO-ə
→ V+[ər]		
V	*!	

(69) /e/+coda liquid (e.g. /ɹ/: no loanword example; /l/: *Bloomingdale* → [pu.lu.miŋ.tai.ər])

/e/+coda liquid	MAX-EO-liquid / i, u, e__]$_\sigma$	DEP-EO-ə
→ V+[ər]		*
V	*!	

Another sequence that is absent from my corpus but is tested in the experiments is /o+ɹ/. Gick, Kang, and Whalen's (2000, 2002) MRI data and the results of the second experiment indicate that /ɹ/ in this sequence is distinctive. Thus, if this sequence is adapted into SC, the /ɹ/ should tend to be preserved. To formalize this expectation, the constraint in (70) outranks DEP-ə. This expectation or the tendencies in Table 8.10 (v) are illustrated in the tableau in (71):

(70) MAX-EO-ɹ / o__]$_\sigma$: Do not delete /ɹ/ if it appears in coda position and follows /o/.

(71) /o/+coda /ɹ/ (no loanword example)

/o/+coda /ɹ/	MAX-EO-ɹ / o__]$_\sigma$	DEP-EO-ə
→ V+[ər]		*
V	*!	

In contrast, coda /l/ tends to delete after /o/ (Table 8.10 (xi)). Thus, the constraint in (72) should be ranked below DEP-ə. This tendency is illustrated in the tableau in (73).

(72) MAX-EO-l / o__]$_\sigma$: Do not delete /l/ if it appears in coda position and follows /o/.

(73) /o/+coda /l/ (e.g. *Goldie* → [kɤ.ti])

/o/+coda /l/	DEP-EO-ə	MAX-EO-l / o__]$_\sigma$
V+[ər]	*!	
→ V		*

Now we consider the /æ/+coda liquid sequences. The articulatory findings and the results of the second experiment seem to indicate that coda liquids are distinctive to some degree after /æ/. However, they also show that /æ/ is the front vowel that is most similar to /ɔ/ and /ə/. In Table 8.4, where distances from various vowels to /ɔ/ and /ə/ are calculated in terms of the three features [high], [low], and [back], /æ/ is even found more similar to /ɔ/ than the two back vowels

/u/ and /ʊ/. Given these pieces of evidence, it is not surprising that, in the loanword adaptation, coda /l/'s do not exhibit strong preservation tendency or even tend to delete after /æ/.

To formalize the deletion tendency of coda liquids after /æ/, DEP-ə outranks the constraint in (74).

(74) MAX-EO-liquid / æ__]$_{\sigma}$: Do not delete a liquid if it appears in coda position and follows /æ/.

The tableau in (84) illustrates the tendency:

(75) /æ/+coda liquid (e.g. /ɹ/: no example; /l/: *C<u>al</u>houn* → [kʰa̲.xuŋ])

/æ/+coda liquid	DEP-EO-ə	MAX-EO-liquid / æ__]$_{\sigma}$
V+[ər]	*!	
→ V		*

If MAX-EO-liquid / ə, ɔ, ɑ__]$_{\sigma}$ and MAX-EO-liquid / æ__]$_{\sigma}$ are combined, then we obtain the constraint hierarchy in (76):

(76) Rankings

MAX-EO-liquid / i, ɪ, e, ɛ, u__]$_{\sigma}$, MAX-EO-ɹ / o__]$_{\sigma}$

>>

DEP-EO-ə

>>

MAX-EO-liquid / ə, ɔ, ɑ, æ__]$_{\sigma}$, MAX-EO-l / o__]$_{\sigma}$

8.4 Summary of Chapter Eight

When English words are borrowed into SC, generally coda liquids after front vowels tend to be preserved but those after central or back vowels tend to delete. To account for this asymmetry, I propose that, in English, coda liquids are generally more distinctive after front vowels than after non-front vowels in perception.

The articulation of an English liquid involves a vocalic gesture and a consonantal gesture, and the vocalic gesture temporally precedes the consonantal gesture when the liquid occurs in coda position. The relative timing of the two gestures in a coda liquid suggests that the extent of perceptual distinctiveness of a coda liquid depends on the similarity of the tongue shapes required for the preceding vowel and the vowel component (i.e. vocalic gesture) of the liquid. Arguably, the more similar the tongue shapes, the less distinctive the liquid. It has been reported in the literature that the pharyngeal widths measured during the production of vowels can predict the tongue shapes with which the vowels are produced, and MacKay's (1977) pharyngeal-width measurements imply that, in general, the vowel components of liquids are more similar to non-front vowels than to front vowels in tongue shape.

Two discrimination experiments are conducted to examine the relative perceptual distinctiveness of coda liquids following various vowels in English. The results show that they are generally more distinctive when following front vowels than when following non-front vowels.

I have proposed a set of ranked constraints and demonstrated that the constraints and the rankings can account for the observed asymmetry.

CHAPTER NINE

CONCLUSION, FUTURE STUDIES, AND THEORETICAL IMPLICATIONS

Three approaches to loanword adaptation are distinguished: (i) The purely perceptual approach, which claims that loanword adaptation occurs during perception beyond the borrowers' conscious awareness, (ii) the purely phonological approach, which claims that the input to loanword adaptation is the underlying representations of the source words and that the adaptation occurs in the host language lexicon, and (iii) the hybrid approach, which claims that loanword adaptation occurs in the host language production grammar with direct reference to the information in the source words that is considered non-contrastive from the viewpoints of both the source and host languages.

I have shown evidence against the purely perceptual approach. The first piece of evidence comes from the adaptation of word-final stops and coda liquids in English loanwords in SC (3.1). While these sounds may be deleted, it is found that they are always preserved with the place of articulation correctly identified if deleting them would result in a violation of the SC minimum-word size requirement. This observation indicates that the borrowers know what sounds are present in the source words and that the decision to preserve or delete these sounds is made in the SC production grammar. The second piece of evidence comes from the adaptation of word-initial liquids as [ɾ] in English loanwords in Japanese and Korean (3.2). Previous perceptual studies and loanword data imply that Japanese and Korean speakers can distinguish English liquids from [ɾ], indicating that the adaptation should not be attributed to unfaithful perception but must occur in the Japanese/Korean production grammar.

I have also shown evidence against the purely phonological approach. The first piece of evidence comes from the observation that, in English loanwords in SC, prevocalic /v/ tends to be adapted as [w] but non-prevocalic /v/ tends to be adapted as [f] (4.1). These data are problematic for the purely phonological approach because this approach predicts /v/ to be adapted as [w] regardless of its position relative to a vowel. As non-prevocalic /v/ is partially devoiced and therefore contains more frication noise than fully voiced /v/ in English, the tendency of non-prevocalic /v/ to be adapted as [f] receives a natural explanation and demonstrates that the phonetic details in the source words play a role in determining how the source words are adapted. The second piece of evidence comes from the tendency of /ʃ, tʃ, dʒ/ to be adapted as [ɕ, tɕ^h, tɕ] rather than the phonologically closer [ʂ, ʈʂ^h, ʈʂ] in English loanwords in SC (4.2). Previous research has suggested that /ʃ, tʃ, dʒ/ and [ɕ, tɕ^h, tɕ] be produced with a front tongue body but that [ʂ, ʈʂ^h, ʈʂ] be produced with a back tongue body. The adaptation tendency, therefore, can be viewed as the borrowers' intention to retain the front tongue-body position associated with /ʃ, tʃ, dʒ/, which is considered a phonetic detail. The third piece of evidence is the observation that, in English loanwords in SC, stops are deleted frequently while deletion of fricatives and affricates rarely occurs (4.3). The purely phonological approach falls short of explaining these data because it predicts these sounds to be treated in the same fashion in terms of whether they should be deleted or preserved. These data can be easily explained if the adaptation makes reference to the phonetic details in the source words: Deletion targets stops because they lack internal cues and therefore are perceptually less salient than other sounds.

While the evidence generally supports the hybrid approach, this approach faces two problems. The first problem is that, in English loanwords in SC, word-medial stops preceding a homorganic nasal or another stop, which are normally considered unreleased and therefore are predicted to delete by the hybrid approach, are preserved frequently (5.1). On the basis of previous phonetic studies and research on foreigner talk speech, I have proposed that these stops may be perceived as released by SC speakers. The second problem is the observation that every word-final unreleased voiceless stop is realized with aspiration when preserved via epenthesis in English loanwords in SC and Korean even though the SC/Korean unaspirated stops are better matches (5.2). To solve this problem, I have proposed (i) that the borrowers must know that word-final stops are variably released in English and (ii) that, to avoid the situation in which unreleased voiceless stops and their voiced counterparts are mapped to the same sounds, the borrowers retrieve the released version of the same source word from the long-term memory and substitute it for the original word.

Three cases of English loanwords in SC are also analyzed in detail. The first case study concerns the tendency of sibilant+/w/ sequences to be adapted as sibilant+epenthesis+[w] sequences (e.g. /sw/ → [ʂʐ̩.w]) and the tendency of velar stop+/w/ sequences to be adapted as the labialized version of the velar stops (e.g. /kw/ → [k^w]) (Chapter Six). I propose that a sibilant+/w/ sequence sounds more similar to a sibilant+epenthesis+[w] sequence (e.g. /sw/ → [ʂʐ̩.w]) than to the labialized version of the sibilant (e.g. /sw/ → [$ʂ^w$]) and that a velar stop+/w/ sequence sounds more similar to the labialized version of the velar stop (e.g. /kw/ → [k^w]) than to a velar stop+epenthesis+[w] sequence (e.g. /kw/ → [k^hɤ.w]). Articulatory evidence and the results of perceptual experiments support my proposal.

The second case study concerns the adaptation of VNV sequences, which shows that the nasal consonant geminates more frequently in some conditions than in other (Chapter Seven). First, the nasal consonant geminates more frequently after a low monophthong (/æ/, /ɔ/, and /ɑ/) than after a non-low lax vowel (/ɪ/, /ɛ/, /ʌ/, and /ʊ/); second, the nasal consonant geminates more frequently after a non-low lax vowel (/ɪ/, /ɛ/, /ʌ/, and /ʊ/) than after a tense vowel or a diphthong (/i/, /u/, /eɪ/, /oʊ/, /ɑɪ/, /ɑʊ/); third, the nasal consonant geminates more frequently after a stressed vowel than after an unstressed vowel; fourth, the nasal consonant geminates more frequently when it is /n/ than when it is /m/. I propose that three factors contribute to the observed asymmetries: (i) The extent of the anticipatory vowel nasalization (accounting for the first and third asymmetries) (7.2.1.2 and 7.2.2), (ii) the duration of the pre-nasal vowel (accounting for the second asymmetry) (7.2.1.1), and (iii) the relative articulatory similarity between labials, coronals, and dorsals (7.2.3). My proposal is supported by articulatory and acoustic evidence and is in agreement with a different set of loanword data.

The third case study concerns the adaptation of coda liquids, which shows that they generally tend to be preserved after a front vowel but tend to delete after a non-front vowel (Chapter Eight). I propose (i) that, in the production of a two-sound sequence, the more the tongue shape changes, the more distinctive the second sound in perception and (ii) that perceptually distinctive sounds tend to be preserved but perceptually indistinctive ones tend to delete. My proposal, again, is supported by articulatory evidence and the results of perceptual experiments.

This dissertation makes some contributions. First, the inviolable SC minimum-word size requirement (i.e. every SC form in the corpus is at least disyllabic in length) and the finding that the SC borrowers could retrieve the released version of learned English words ending in a voiceless stop demonstrate that the adaptation of English words into SC does not take place in perception but takes place in the SC production grammar while making direct reference to the phonetic information in the source words (e.g. stop releasing), arguing against the purely perceptual and purely phonological approaches and supporting the hybrid models that claim faithful perception. Second, the observation that, unlike native sibilant+/u/ sequences, sibilant+/w/ sequences in borrowed English words are always realized with the sibilants forming separate syllables in SC supports the proposal in the literature that foreign input and native input are subject to different sets of faithfulness constraints (e.g. IDENT-OO-[feature] by Kenstowicz (2005); MIMIC-[feature] by Yip (2002, 2006); BESIMILAR-[feature] by Kang (2003)). Third, since many of the rankings of the loan-sensitive faithfulness constraints in this dissertation have a perceptual basis and are confirmed by the results of perceptual experiments, the P-map theory proposed by Steriade (2001, 2009), which suggests that there be a distinct grammatical component housing the knowledge of relative perceptibility of contrasts in various contexts, is also supported.

In this dissertation, the influence of the orthography is not discussed (see Vendelin and Peperkamp (2006) for an experimental study of the effects of English orthography on on-line adaptation of English vowels into French). Because evidence indicates that the borrowers know what sounds are present in the source words, it would be intriguing to see if the orthography makes a contribution. I suggest a future study in this regard, which concerns a possible effect of English orthography on the mapping of unreleased stops in English words borrowed into SC, Korean, and Cantonese (and any other language which allows unreleased stops and provides loanword data). Unreleased stops, which lack release bursts, are aurally non-salient. Consequently, it is not impossible that even listeners whose native languages allow strictly unreleased stops such as Korean and Cantonese do not detect them if the words are foreign. In line with Vendelin and Peperkamp (2006), what we can do is to have SC, Korean, and Cantonese speakers of English produce English nonce words in their native languages. In the first task, the subjects listen to disyllabic stimuli ending in vowel+unreleased stop sequences, while in the second task similar stimuli are presented along with the written forms. Since I have claimed in 5.1 that unreleased stops are deleted in English loanwords in SC (unless the deletion results in violation of the minimum-word requirement), I predict that, in the first task, the unreleased stops in the stimuli will have no correspondents in the SC subjects' production. If the result of the second task shows preservation of some of the unreleased stops, then (i) it must be due to the presentation of the written forms, and (ii) the claim made in 5.1 needs to be amended, that is, unreleased stops may be preserved if the adapters see the orthography. In addition, if preservation is attested in the SC subjects' production in the second task, I predict that the preserved unreleased stops are realized with aspiration when they are voiceless but without aspiration when voiced. If this is the case, then it demonstrates that the SC subjects rely on the orthography to distinguish unreleased voiceless stops from unreleased voiced stops.[59] On the other hand, as Korean and Cantonese syllables can end in an unreleased stop, we do not expect the results of the second task to show any deletion of the unreleased stops. However, if the results of the first task show deletion of some of the unreleased stops, then (i) the absence of the written forms is the cause, (ii) Korean and Cantonese speakers do not aurally detect every unreleased stop in foreign words even though coda stops are unreleased in their native languages, and (iii) the preservation of *all* post-vocalic word-final stops in English loanwords in Korean (Kang 2003) and Cantonese (Silverman 1992; Yip 1993, 2002, 2006; Bauer and Benedict 1997) is partially due to the orthography.

In the literature, the role of orthography refers to the influence of the source language orthography. It is possible that the host language orthography also plays a role. For example, if coda stops are strictly unreleased in the host language and the adapters do not know the host language orthography, then unreleased stops in borrowed English words may not be always preserved. To test this possibility, we can present Korean- and Cantonese-English bilinguals who

[59] Recall that, according to Flege (1989) and Flege and Wang (1989), SC speakers have great difficulty aurally discriminating voicing contrasts in word-final unreleased stops.

know the English orthography but do not know the host language orthography[60] (e.g. Korean/Cantonese Americans who speak Korean/Cantonese at home) with oral stimuli of real English words ending in vowel+unreleased stop sequences and have them produce the most similar correspondents in Korean or Cantonese. If the host language orthography influences the adaptation, then the unreleased stops in the stimuli should be deleted at least occasionally rather than being always preserved as shown in the loanword data.[61]

The data and analyses presented in this dissertation have theoretical implications. First, the claim that the SC borrowers could retrieve information about learned English words from the long-term memory implies that at lease the adaptation of English words into SC is not an online adaptation process but is based on mental representations. This does not mean that the perceptual information contained in the source words plays no role in the adaptation process; instead, it means that the borrowers have acquired the source words and make use the information they know about the source words, which could be perceptual, articulatory, or phonological, in the adaptation process. Because the SC loanword data and analyses presented in this dissertation have demonstrated that the phonetic information in the source words are critical in determining the output, the mental representations of the source words that the borrowers have constructed must contain the information that is normally considered phonetic.

The second implication concerns a possibility that the perceptual and articulatory information contained in the source words may conflict each other in the adaptation process. As

[60] While characters are used to represent Cantonese words, in order to type on computers, Cantonese speakers have learned a romanization system called Jyutping (粵拼). For example, if the tone is ignored, the character 濕, which means 'wet' and is pronounced [sap̚] in Cantonese, is spelled *sap* in Jyutping.

[61] This idea originates from personal experience. Although I am a native speaker of Southern Min, in which syllable-ending stops are strictly unreleased, I did not know that a Southern Min syllable could end in a stop until I studied linguistics. For me, a vowel that precedes a syllable-ending unreleased stop sounds a little different from the same vowel that does not, and I use this difference to make discrimination. Interestingly, this strategy fails if the words are foreign. For example, I have difficulty aurally distinguishing pairs such as *Kay* vs. *Kat̚e*, *can* vs. *can't̚*, *car* vs. *card̚*, etc.

shown in Chapters Four and Eight, palato-alveolars (/ʃ, tʃ, dʒ/) are mostly adapted as alveolo-palatals ([ɕ, tɕʰ, tɕ]) (4.2) and intervocalic /m/ is resistant to gemination (7.2.3), and I have argued for an articulatory account. If future acoustic/perceptual research leads us to the conclusion that /m/ is not the most identifiable of the three nasal consonants and that palato-alveolars (/ʃ, tʃ, dʒ/) are actually more similar to post-alveolars ([ʂ, ʈʂʰ, ʈʂ]) than to alveolo-palatals ([ɕ, tɕʰ, tɕ]), then the loanword data and analyses I have presented in this dissertation strongly suggest dominance of faithfulness to the articulator/tongue-body position over perceptual similarity. Recall that whereas /θ/ is adapted as [t] in English loanwords in Quebec French (Kenstowicz and Suchato 2006), Quebec French speakers still judge [f] as a closer match for /θ/ perceptually (Brannen 2002). Indeed, loanword data from SC and other languages all show that /θ/ is adapted as [s] or [t] and never adapted as [f]. On the other hand, it is possible that, in other cases, perceptually similarity is preferred over articulatory similarity. The adaptation of prevocalic /v/ in English words borrowed into SC and many other languages is an example. /v/ and [f] are produced in the same fashion. [w], however, is produced quite differently. As mentioned in 6.3.1, /w/ is produced with the back of the tongue raised and the lips rounded. The strong tendency of prevocalic /v/ to be adapted as [w] instead of the articulatorily identical sound [f] in English loanwords in SC seems to suggest that, in this particular case, having a match that is perceptually similar be more important than having one that is articulatorily similar or even identical. Of course what articulatory information and what perceptual information count as salient and therefore tend to be maintained by the borrowers awaits future research, but the discussion of this issue implies that not only is perceptual and articulatory saliency important for the study of loanword adaptation, but the interaction of these two types of information should be also taken into consideration.

The third implication concerns what phonological and phonetic information makes up the mental representation that serves as the input to loanword adaptation. Lin (2008b) investigated vowel adaptation in English loanwords in SC, and her analyses require 'assumptions on what vowel quality is present in the "perceived" input' (pp. 377). The loanword data and analyses presented in this dissertation are in line with Lin's suggestion. Namely, they imply that some phonetic information in the source words should be specified and therefore present in the mental representation that serves as the input (4.1). The phonetic information that this dissertation suggest be specified and present in mental representation that serves as the input include (i) the extent of frication noise in /v/ (4.1), (ii) the tongue-body position for palato-alveolars (4.2), (iii) release bursts of stops (5.1), (iv) vowel duration (7.2.1.1), (v) the extent of anticipatory vowel nasalization (7.2.1.2 and 7.2.2), and (vi) diphthong-internal formant transitions (7.2.1.2.3). However, Lin (2008b) also argues that, in order to account for the great mapping variability in her data, some features of English vowels must be unspecified (e.g. mid central vowels in the source words were argued to be unspecified for [low], [back], and [round]). While the SC loanword data presented in this dissertation exhibit some level of variability, clear adaptation tendencies are observed. These tendencies seem to imply that the relevant phonetic details are salient enough and therefore are usually present in the mental representation that serves as the input.

APPENDICES

APPENDIX A

Paradigm for Experiment I (6.3.2.1)

A (SC)	X (E.)	B (SC)
[sẓ.wu.pi]	/swú.bi/	[sʷu.pi]
[sẓ.wei.pi]	/swéɪ.bi/	[sʷei.pi]
[sẓ.wo.pi]	/swɔ́.bi/	[sʷo.pi]
[sẓ.wa.pi]	/swɑ́.bi/	[sʷa.pi]
[kʰɤ.wei.pi]	/kwéɪ.bi/	[kʰʷei.pi]
[kʰɤ.wo.pi]	/kwóʊ.bi/	[kʰʷo.pi]
[kʰɤ.wa.pi]	/kwɑ́.bi/	[kʰʷa.pi]

Tone 3 is assigned to the SC syllable [pi], and Tone 2 is assigned to the other SC syllables.

APPENDIX B

Paradigm for Experiment II (6.3.2.1)

A (SC)	X (E.)	B (SC)
斯屋比 ([sz̩1.wu1.pi3])	/swú.bi/	蘇比 ([sʷu1.pi3])
斯維比 ([sz̩1.wei2.pi3])	/swéɪ.bi/	隨比 ([sʷe2i.pi3])
師娃比 ([ʂʐ̩1.wa1.pi3])	/swʌ́.bi/	刷比 ([ʂʷa1.pi3])
斯窩比 ([sz̩1.wo1.pi3])	/swɔ́.bi/	梭比 ([sʷo1.pi3])
斯維比 ([sz̩1.wei2.pi3])	/swǽ.bi/	隨比 ([sʷe2i.pi3])
師娃比 ([ʂʐ̩1.wa1.pi3])	/swɑ́.bi/	刷比 ([ʂʷa1.pi3])
科維比 ([kʰɤ1.wei2.pi])	/kwéɪ.bi/	奎比 ([kʰʷei2.pi3])
科沃比 ([kʰɤ1.wo4.pi3])	/kwóʊ.bi/	闊比 ([kʰʷo4.pi3])
科維比 ([kʰɤ1.wei2.pi])	/kwǽ.bi/	奎比 ([kʰʷei2.pi3])
科娃比 ([kʰɤ1.wa1.pi3])	/kwɑ́.bi/	誇比 ([kʰʷa1.pi3])

Numerals represent tones.

APPENDIX C

Paradigm for Experiment III (6.3.2.1)

A (SC)	X (E.)	B (SC)
斯屋比 ([sz̩1.wu1.pi3])	*Swoobe* (/swú.bi/)	蘇比 ([sʷu1.pi3])
斯維比 ([sz̩1.wei2.pi3])	*Swabe* (/swéɪ.bi/)	隨比 ([sʷe2i.pi3])
師娃比 ([ʂʐ̩1.wa1.pi3])	*Swube* (/swʌ́.bi/)	刷比 ([ʂʷa1.pi3])
斯窩比 ([sz̩1.wo1.pi3])	*Swobe* (/swɔ́.bi/)	梭比 ([sʷo1.pi3])
斯維比 ([sz̩1.wei2.pi3])	*Swabby* (/swǽ.bi/)	隨比 ([sʷe2i.pi3])
師娃比 ([ʂʐ̩1.wa1.pi3])	*Swabbie* (/swɑ́.bi/)	刷比 ([ʂʷa1.pi3])
科維比 ([kʰɤ1.wei2.pi])	*Quaibe* (/kwéɪ.bi/)	奎比 ([kʰʷei2.pi3])
科沃比 ([kʰɤ1.wo4.pi3])	*Quobe* (/kwóʊ.bi/)	闊比 ([kʰʷo4.pi3])
科維比 ([kʰɤ1.wei2.pi])	*Quabe* (/kwǽ.bi/)	奎比 ([kʰʷei2.pi3])
科娃比 ([kʰɤ1.wa1.pi3])	*Kwabbie* (/kwɑ́.bi/)	誇比 ([kʰʷa1.pi3])

Numerals represent tones.

APPENDIX D

Paradigm for Experiment V (8.2.2.2.1)

A (SC)	X (E.)	B (SC)
[ti.ər.pi]	/díl.bi/	[ti.pi]
[ti.ər.pi]	/díl.bi/	[ti.pi]
[xei.ər.pi]	/héɪl.bi/	[xei.pi]
[xei.ər.pi]	/hέl.bi/	[xei.pi]
[xei.ər.pi]	/hǽl.bi/	[xei.pi]
[tʰu.ər.pi]	/túl.bi/	[tʰu.pi]
[kou.ər.pʲi]	/gól.bi/	[kou.pi]
[tʰou.ər.pi]	/tɔ́l.bi/	[tʰou.pi]
[tɑu.ər.pi]	/dɑ́l.bi/	[tɑu.pi]
[tɑu.ər.pi]	/dʌ́l.bi/	[tɑu.pi]
[ti.ər.pi]	/díɹ.bi/	[ti.pi]
[xei.ər.pi]	/hέɹ.bi/	[xei.pi]
[sou.ər.pi]	/sóɹ.bi/	[sou.pi]
[tʰou.ər.pi]	/dɔ́r.bi/	[tʰou.pi]

Tone 3 is assigned to the SC syllable [pi], Tone 2 is assigned to the SC syllable [ər], and Tone 1 is assigned to the other SC syllables.

BIBLIOGRAPHY

BIBLIOGRAPHY

Adler, Allison. 2006. Faithfulness and perception in loanword adaptation: A case study from Hawaiian. *Lingua* 116: 1024–1045.

Ali, Latif, Tanya Gallagher, Jeffrey Goldstein, and Raymond Daniloff. 1971. Perception of coarticulated nasality. *Journal of the Acoustical Society of America* 49: 538–540.

Altmann, Heidi. 2006. *The Perception and Production of Second Language Stress: A Cross-Linguistic Experimental Study*. Doctoral Dissertation, University of Delaware.

Alwan, Abeer A., Shrikanth S. Narayanan, and Katherine Haker. 1997. Toward articulatory-acoustic models for liquid approximants based on MRI and EPG data. Part II. The rhotics. *Journal of the Acoustical Society of America* 101: 1078–1089.

Bauer, Robert S., and Paul K. Benedict. 1997. *Modern Cantonese phonology*. Berlin: Mouton de Gruyter.

Beddor, Patrice S. 1982. *Phonological and Phonetic Effects of Nasalization on Vowel Height*. Doctoral Dissertation, University of Minnesota, Minneapolis.

Bell-Berti, Fredericka, and Rena A. Krakow. 1991. Anticipatory velar lowering: A coproduction account. *Journal of the Acoustical Society of America* 90: 112–123.

Bell-Berti, Fredericka, Thomas Baer, Katherine S. Harris, and Seiji Nimi. 1979. Coarticulatory effects of vowel quality on velar function. *Phonetica* 36: 187–193.

Bell-Berti, Fredericka. 1993. Understanding velic motor control: Studies of segmental context. In *Nasals, Nasalization, and the Velum*, ed. M. K. Huffman and R. A. Krakow, 63–85. New York, NY: Academic Press.

Best, Catherine T. 1994. The emergence of native-language phonological influences in infants: A perceptual assimilation model. In *The Development of Speech Perception: The Transition from Speech Sounds to Spoken Words*, ed. J. C. Goodman and H. C. Nusbaum, 167–224. Cambridge, MA: MIT Press.

Best, Catherine T. 1995. A direct realist view of cross-language speech perception. In *Speech Perception and Linguistic Experience: Issues in Cross-Language Research*, ed. W. Strange, 171–204. Baltimore: York Press.

Best, Catherine T., and Winifred Strange. 1992. Effects of phonological and phonetic factors on cross-language perception of approximants. *Haskins Laboratories Status report on Speech Research* 109/110: 89–108.

Best, Catherine T., Gerald W. McRoberts, and Nomathemba M. Sithole. 1988. Examination of perceptual reorganization for nonnative speech contrasts: Zulu click discrimination by English-speaking adults and infants. *Journal of Experimental Psychology: Human Perception and Performance* 14: 345–360.

Bloch, Bernard. 1950. Studies in colloquial Japanese IV: Phonemics. *Language* 26: 86–125.

Boersma, Paul, and Silke Hamann. 2009. Cue constraints and their interactions in phonological perception and production. In *Phonology in Perception*, ed. P. Boersma and S. Hamann, 55–110. Berlin: Mouton de Gruyter.

Bradley, Travis. 2012. Labialization and palatalization in Judeo-Spanish phonology. In *Romance Linguistics: Selected Papers from the 42nd Linguistic Symposium on Romance Languages (LSRL)*, ed. J. Smith and T. Ihsane, 77–98. Cedar City, Utah.

Brannen, Kathleen. 2002. The role of perception in differential substitution. *Canadian Journal of Linguistics* 47: 1–46.

Brinton, Laurel J. 2000. *The Structure of Modern English: A Linguistic Introduction*. Amsterdam: John Benjamins.

Broselow, Ellen. 2009. Stress adaptation in loanword phonology: perception and learnability. In *Phonology in Perception*, ed. P. Boersma and S. Hamann, 191–234. Berlin: Mouton de Gruyter.

Browman, Catherine P., and Louis Goldstein. 1992. Articulatory phonology: An overview. *Haskins Laboratories Status Report on Speech Research* 111/112: 23–42.

Browman, Catherine P., and Louis Goldstein. 1995. Gestural syllable position effects in American English. In *Producing Speech: Contemporary Issues. For Katherine Safford Harris*, ed. F. Bell-Berti and L. J. Raphael, 19–33. Woodbury, NY: AIP Press.

Byrd, Dani, Stephen Tobin, Eric Bresch, and Shrikanth Narayanan. 2009. Timing effects of syllable structure and stress on nasals: A real-time MRI examination. *Journal of Phonetics* 37: 97–110.

Calabrese, Andrea. 2009. Perception, production and acoustic inputs in loanword phonology. In *Loan Phonology*, ed. A. Calabrese and W. L. Wetzels, 59–114. Philadelphia, PA: John Benjamins.

Campbell, Fiona, Bryan Gick, Ian Wilson, and Eric Vatikiotis-Bateson. 2010. Spatial and temporal properties of gestures in North American English /r/. *Language and Speech* 53: 49–69.

Chang, Woohyeok, and William Idsardi. 2001. Speech perception and production of the English final voicing contrast by Korean speakers. *Harvard Studies in Korean Linguistics* 9: 117–129.

Chen, Ying, and Susan Guion-Anderson. 2011. Perceptual confusability of word-final nasals in Southern Min and Mandarin: Implications for coda nasal mergers in Chinese. In *Proceedings of the 17th International Congress of Phonetic Sciences*, ed. W.-S. Lee and E. Zee, 464–467. Hong Kong: City University of Hong Kong.

Chen, Ying, Vsevolod Kapatsinski, and Susan Guion-Anderson. 2012. Acoustic cues of vowel quality to coda nasal perception in Southern Min. In *Proceedings of the 13th Annual Conference of the International Speech Communication Association (Interspeech 2012)*, 1410–1413.

Cheng, Lilly L-R. 1991. *Assessing Asian Language Performance: Guidelines for Evaluating Limited-English Proficient Students* (2nd ed.). Oceanside, CA: Academic Communication Associates.

Cheung, Kwan-hin. 1986. *The Phonology of Present-Day Cantonese*. Doctoral Dissertation, University College London.

Chung, Rong-fu. 1996. *The segmental phonology of Southern Min in Taiwan*. Taipei: Crane Publishing.

Clements, G. Nick. 2001. Representational economy in constraint-based phonology. In *Distinctive Feature Theory*, ed. T. A. Hall, 71–146. Berlin: Mouton de Gruyter.

Clumeck, Harold. 1976. Patterns of soft palate movements in six languages. *Journal of Phonetics* 4: 337–351.

Cohn, Abigail C. 1990. Phonetic and phonological rules of nasalization. *UCLA Working Papers in Phonetics* 76: 1–224.

Collier, Rene, Leigh Lisker, Hajime Hirose, Tatsujiro Ushijima. 1979. Voicing in intervocalic stops and fricatives in Dutch. *Journal of Phonetics* 7: 357–373.

Crawford, Clifford. 2007. The role of loanword diffusion in changing adaptation patterns: A study of coronal stops in Japanese borrowings. *Working Papers of the Cornell Phonetics Laboratory* 16: 32–56.

Delattre, Pierre C., and Donald C. Freeman. 1968. A dialect study of American r's by X-ray motion picture. *Linguistics* 44: 29–68.

Delattre, Pierre. 1981. An acoustic and articulatory study of vowel reduction in four languages. In *Studies in Comparative Phonetics*, ed. P. Delattre, 63–93. Heidelberg: Julius Groos Verlag.

Dickson, David R. 1962. An acoustical study of nasality. *Journal of Speech and Hearing Research* 5: 103–111.

Dohlus, Katrin. 2005. Phonetics or phonology: Asymmetries in loanword adaptations – French and German mid front rounded vowels in Japanese. *ZAS Papers in Linguistics* 42:117–135.

Dong, Xiaoli. 2012. *What Borrowing Buys Us: A Study of Mandarin Chinese Loanword Phonology*. Doctoral Dissertation, Utrecht University.

Duanmu, San. 1993. Rime length, stress, and association domains. *Journal of East Asian Linguistics* 2: 1–44.

Duanmu, San. 1994. Syllabic weight and syllabic duration: A correlation between phonology and phonetics. *Phonology* 11: 1–24.

Duanmu, San. 2007. *The Phonology of Standard Chinese*. Oxford: Oxford University Press.

Dupoux, Emmanuel, Christophe Pallier, Nuria Sebastian, and Jacques Mehler. 1997. A destressing "deafness" in French? *Journal of Memory and Language* 36: 406–421.

Dupoux, Emmanuel, Erika Parlato, Sonia Frota, Yuki Hirose, Sharon Peperkamp. 2011. Where do illusory vowels come from? *Journal of Memory and Language* 64: 199–210.

Dupoux, Emmanuel, Kauhiko Kazehi, Yuki Hirose, Christophe Pallier, and Jacques Mehler. 1999. Epenthetic vowels in Japanese: A perceptual illusion. *Journal of Experimental Psychology: Human Perception and Performance* 25: 1568–1578.

Dupoux, Emmanuel, Núria Sebastián-Gallés, Eduardo Navarrete, and Sharon Peperkamp. 2008. Persistent stress 'deafness': The case of French learners of Spanish. *Cognition* 106: 682–706.

Dupoux, Emmanuel, Sharon Peperkamp, and Núria Sebastián-Gallés. 2001. A robust method to study stress 'deafness'. *Journal of the Acoustical Society of America* 110: 1606–1618.

Durvasula, Karthik, and Ho-Hsin Huang. 2015. Word-internal ambisyllabic consonants are codas in American English. Ms., Michigan State University.

Fay, David, and Anne Cutler. 1977. Malapropisms and the structure of the mental lexicon. *Linguistic Inquiry* 8: 505–520.

Flege, James E. 1989. Chinese subjects' perception of the word-final English/t/-/d/ contrast: Performance before and after training. *Journal of the Acoustical Society of America* 86: 1684–1697.

Flege, James E., and Chipin Wang. 1989. Native-language phonotactic constraints affect how well Chinese subjects perceive the word-final English /t/-/d/ contrast. *Journal of Phonetics* 17: 299–315.

Flemming, Edward. 2003. The relation between coronal place and vowel backness. *Phonology* 20: 335–74.

Fujimura, Osamu, Joan E. Miller, and Shigeru Kiritani. 1975. An X-ray observation of movements of the velum and tongue. *Journal of the Acoustical Society of America* 58: S40.

Gass, Susan. M., and Larry Selinker. 2008. *Second Language Acquisition: An Introductory Course* (3rd ed.). New York, NY: Routledge.

Gick, Bryan, A. Min Kang, Douglas H. Whalen. 2000. MRI and X-ray evidence for commonality in the dorsal articulations of English vowels and liquids. In *Proceedings of the 5th Seminar on Speech Production: Models and Data*, 69–72. Munich: Ludwig-Maximilians-Universität.

Gick, Bryan, A. Min Kang, Douglas H. Whalen. 2002. MRI evidence for commonality in the post-oral articulations of English vowels and liquids. *Journal of Phonetic*s 30: 357–371.

Gick, Bryan, and Fiona Campbell. 2003. Intergestural timing in English /r/. In *Proceedings of the 15th International Congress of Phonetic Sciences*, ed. M. J. Solé, D. Recasens, and J. Romero, 1911–1914. Barcelona: Universitat Autonoma de Barcelona.

Gick, Bryan, and Ian Wilson. 2006. Excrescent schwa and vowel laxing: Cross-linguistic responses to conflicting articulatory targets. In *Papers in Laboratory Phonology VIII: Varieties of Phonological Competence*, ed. L. Goldstein, D. H. Whalen, and C. T. Best, 635–660. Berlin, New York: Mouton de Gruyter.

Gick, Bryan. 1999. A gesture-based account of intrusive consonants in English. *Phonology* 16: 29–54.

Gick, Bryan. 2002. The American intrusive /l/. *American Speech* 77: 16–83.

Gick, Bryan. 2003. Articulatory correlates of ambisyllabicity in English glides and liquids. In *Papers in Laboratory Phonology VI: Constraints on Phonetic Interpretation*, ed. J. Local, R. Ogden, and R. Temple, 222–236. Cambridge: Cambridge University Press.

Goto, Hiromu. 1971. Auditory perception by normal Japanese adults of the sounds “L” and “R.” *Neuropsychologia* 9: 317–323.

Guion, Susan G., James E. Flege, Reiko Akahane-Yamada, and Jesica C. Pruitt. 2000. An investigation of current models of second language speech perception: The case of Japanese adults' perception of English consonants. *Journal of the Acoustical Society of America* 107: 2711–2724.

Hagiwara, Robert. 1995. Acoustic realizations of American /r/ as produced by women and men. *University of California Working Papers in Phonetics* 90: 1–187.

Hatch, Evelyn M. 1983. *Psycholinguistics: A Second Language Perspective*. Rowley, Mass: Newbury House.

Hattori, Shiro, Kengo Yamamoto, and Osamu Fujimura. 1958. Nasalization of vowels in relation to nasals. *Journal of the Acoustical Society of America* 30: 267– 274.

He, Yunjuan. 2014. Production of English syllable final /l/ by Mandarin Chinese speakers. *Journal of Language Teaching and Research* 5: 742–750.

Heffner, R-M. S. 1937. Notes on the length of vowels. *American Speech* 12: 128–134.

Henderson, Janette B, and Bruno H. Repp. 1981. Is a stop consonant released when followed by another stop consonant? *Haskins Laboratories Status Report on Speech Research* 67/68: 71–82.

Henderson, Janette B. 1984. *Velopharyngeal Function in Oral and Nasal Vowels – A Cross-language Study*. Doctoral Dissertation, University of Connecticut.

Herd, Jonathon. 2005. Loanword adaptation and the evaluation of similarity. *Toronto Working Papers in Linguistics* 24: 65–116.

Hillenbrand, James, Laura A. Getty, Michael J. Clark, and Kimberlee Wheeler. 1995. Acoustic characteristics of American English vowels. *Journal of the Acoustical Society of America* 97: 3099–3111.

Howie, John M. 1976. *Acoustical Studies of Mandarin Vowels and Tones*. Cambridge: Cambridge University Press.

Hsieh, Feng-fan, Michael Kenstowicz, and Xiaomin Mou. 2009. Mandarin adaptations of coda nasals in English loanwords. In *Loan Phonology*, ed. A. Calabrese and W. L. Wetzels, 131–154. Philadelphia, PA: John Benjamins.

Huang, Ho-Hsin, and Yen-Hwei Lin. 2013. To geminate or not to geminate: The case of nasals in Standard Chinese loanwords. Paper presented at the 25th North American Conference on Chinese Linguistics (NACCL-25), University of Michigan, Ann Arbor.

Hume, Elizabeth. 1992. *Front vowels, coronal consonants and their interaction in non-linear phonology*. Doctoral Dissertation, Cornell University.

Hyman, Larry M. 1970. The role of borrowings in the justification of phonological grammars. *Studies in African Linguistics* 1: 1–48.

Ito, Chiyuki, Yoonjung Kang, and Michael Kenstowicz. 2006. The adaptation of Japanese loanwords into Korean. *MIT Working Papers in Linguistics* 52: 65–104.

Jacobs, Haike, and Carlos Gussenhoven. 2000. Loan phonology: Perception, salience, the lexicon and OT. In *Optimality Theory: Phonology, Syntax, and Acquisition*, ed. J. Dekkers, F. van der Leeuw, and J. van de Weijer, 193–210. Oxford: Oxford University Press.

Johnson, Keith, Christian T. DiCanio, and Laurel McKenzie. 2007. The acoustic and visual phonetic basis of place of articulation in excrescent nasals. In *UC Berkeley Phonology Lab Annual Report* 2007, 529–561.

Jun, Jongho. 1995. *Perceptual and Articulatory Factors in Place Assimilation: An Optimality Theoretic Approach*. Doctoral Dissertation, UCLA.

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

Kang, Hyunsook. 1996. English loanwords in Korean. *Studies in Phonetics, Phonology and Morphology* 2: 21–47. Seoul: Hankwukmunhwasa.

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

Kawahara, Shigeto, and Kelly Garvey. 2014. Nasal place assimilation and the perceptibility of place contrasts. *Open Linguistics* 1: 17–36.

Kenstowicz, Michael, and Atiwong Suchato. 2006. Issues in loanword adaptation: A case study from Thai. *Lingua* 116: 921–949.

Kenstowicz, Michael, and Hyang-Sook Sohn. 2001. Accentual adaptation in North Kyungsang Korean. In *Ken Hale: A Life in Language*, ed. M. Kenstowicz, 239–270. Cambridge, MA: MIT Press.

Kenstowicz, Michael. 2003. The role of perception in loanword phonology. A Review of Les emprunts linguistiques d'origine europeenne en Fon. *Studies in African Linguistics* 32: 95–112.

Kenstowicz, Michael. 2005. The phonetics and phonology of Korean loanword adaptation. In *Proceedings of the First European Conference on Korean Linguistics*, ed. S.-J. Rhee, 17–32.

Kenstowicz, Michael. 2007. Salience and similarity in loanword adaptation: A case study from Fijian. *Language Sciences* 29: 316–340.

Kenstowicz, Michael. 2012. Cantonese loanwords: Conflicting faithfulness in VC rime constraints. *Catalan Journal of Linguistics* 11: 65–96.

Keyser, Samuel J., and Kenneth N. Stevens. 1994. Feature geometry and the vocal tract. *Phonology* 11: 207–236.

Khouw, Edward, and Valter Ciocca. 2006. An acoustic and perceptual study of final stops produced by profoundly hearing impaired adolescents. *Journal of Speech, Language, and Hearing Research* 49: 172–85.

Kim, Hyunsoon. 1998. A phonetic characterization of release and nonrelease: The case of Korean and English. *Language Research* 34: 347–368.

Kim, Soohee, and Emily Curtis. 2002. Phonetic Duration of English /s/ and its borrowing in Korean. In *Japanese/Korean Linguistics* 10, ed. N. Akatsuka and S. Strauss, 406–419. Palo Alto: CSLI.

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

Krakow, Rena A. 1987. Stress effects on velic gestures. Journal of the Acoustical Society of America 82: S17.

Krakow, Rena A. 1989. *The Articulatory Organization of Syllables: A Kinematic Analysis of Labial and Velar Gestures*. Doctoral Dissertation, Yale University.

Krakow, Rena A. 1994. Nonsegmental influences on velum movement patterns: Syllables, sentences, stress, and speaking rate. *Haskins Laboratories Status Report on Speech Research* 117/118: 31–48.

Kubozono, Haruo. 2006. Where does loanword prosody come from? A case study of Japanese loanword accent. *Lingua* 116: 1140–1170.

LaCharité, Darlene, and Carole Paradis. 2002. Addressing and disconfirming some predictions of phonetic approximation for loanword adaptation. *Langues et Linguistique* 28: 73–91.

LaCharité, Darlene, and Carole Paradis. 2005. Category preservation and proximity versus phonetic approximation in loanword adaptation. *Linguistic Inquiry* 36: 223–258.

Ladefoged, Peter, and Ian Maddieson. 1996. *The Sounds of the World's Languages*. Cambridge, MA: Blackwell.

Ladefoged, Peter, and Keith Johnson. 2011. *A Course in Phonetics* (6th ed.). Boston, MA: Wadsworth/Cengage Learning.

Ladefoged, Peter. 1975. *A Course in Phonetics* (2nd ed.). Orlando: Harcourt Brace.

Ladefoged, Peter. 1993. *A Course in Phonetics* (4th ed.). Orlando: Harcourt Brace.

Lehiste, Ilse. 1964. *Acoustical Characteristics of Selected English consonants*. Bloomington: Indiana University.

Lin, C. Charles. 2002. Nasal endings of Taiwan Mandarin: Production, perception, and linguistic change. Paper presented at the 35th International Conference on Sino-Tibetan Languages and Linguistics, Arizona State University, Tempe.

Lin, Yen-Hwei. 2007. *The Sounds of Chinese*. Cambridge: Cambridge University Press.

Lin, Yen-Hwei. 2008a. Patterned vowel variation in Mandarin loanword adaptation: evidence from a dictionary corpus. In P*roceedings of the 20th North American Conference on Chinese Linguistics (NACCL-20),* ed. M. Chan and H. Kang, 175–187. Columbus, Ohio: The Ohio State University.

Lin, Yen-Hwei. 2008b. Variable vowel adaptation in Standard Mandarin loanwords. *Journal of East Asian Linguistics* 17: 363–380.

Lintz, Lois B., and Dorothy Sherman. 1961. Phonetic elements and perception of nasality. *Journal of Speech and Hearing Research* 4: 381–396.

Lovins, Julie. 1975. *Loanwords and the Phonological Structure of Japanese*. Bloomington: Indiana University Linguistics Club.

MacKain, Kristine S., Catherine T. Best, and Winifred Strange. 1981. Categorical perception of English /r/ and /l/ by Japanese bilinguals. *Applied Psycholinguistics* 2: 369–390.

MacKay, Ian R. A. 1977. Tenseness in vowels: An ultrasonic study. *Phonetica* 34: 325–351.

MacKay, Ian R. A. 1978. *Introducing Practical Phonetics*. Boston: Little, Brown.

Malécot, André. 1956. Acoustic cues for nasal consonants: an experimental study involving a tape-splicing technique. *Language* 32: 274–284.

Martin, Pierre. 2002. Le système vocalique du français du Québec: De l'acoustique à la phonologie. *La Linguistique* 38: 71–88.

Matthews, John, and Cynthia Brown. 2004. When intake exceeds input: Language specific perceptual illusions induced by L1 prosodic constraints. *International Journal of Bilingualism* 8: 5–27.

McMahon, April. 1996. On the Uuse of the past to explain the present: The history of /r/ in English and Scots. In *English Historical Linguistics 1994: Papers from the 8th International Conference on English Historical Linguistics (8. ICEHL)*, ed. D. Britton, 73–89.

McMahon, April. 2000. *Lexical Phonology and the History of English*. Cambridge, New York: Cambridge University Press.

Miyawaki, Kuniko, James J. Jenkins, Winifred Strange, Alvin M. Liberman, Robert Verbrugge, and Osamu Fujimura. 1975. An effect of linguistic experience: The discrimination of [r] and [l] by native speakers of Japanese and English. *Perception and Psychophysics* 18: 331–340.

Mochizuki, Michiko. 1981. The identification of /r/ and /l/ in natural and synthesized speech. *Journal of Phonetics* 9: 283–303.

Moll, Kenneth L. 1962. Velopharyngeal closure on vowels. *Journal of Speech and Hearing Research* 5: 30–37.

Narayanan, Shrikanth S., Abeer A. Alwan, and Katherine Haker. 1997. Toward articulatory-acoustic models for liquid approximants based on MRI and EPG data. Part I. The laterals. *Journal of the Acoustical Society of America* 101: 1064–1077.

Ohala, John J. 1971. Monitoring soft palate movements in speech. Paper presented at the 8lst Meeting of the Acoustical Society of America, Washington, D.C.

Olive, Joseph, Alice Greenwood, and John Coleman. 1993. *Acoustics of American English Speech: A Dynamic Approach*. Springer Verlag.

Paradis, Carole, and Antoine Tremblay. 2009. Nondistinctive features in loanword adaptation: The unimportance of English aspiration in Mandarin Chinese phoneme categorization. In *Loan Phonology*, ed. A. Calabrese and L. Wetzels, 211–224. Amsterdam: John Benjamins.

Paradis, Carole, and Caroline Lebel. 1994. Contrasts from segmental parameter settings in loanwords: Core and periphery in Quebec French. *Toronto Working Papers in Linguistics* 13: 75–94.

Paradis, Carole, and Darlene LaCharité. 1993. Constraint-based theories in multilinear phonology. *Canadian Journal of Linguistics* 38: 127–303.

Paradis, Carole, and Darlene LaCharité. 1997. Preservation and minimality in loanword adaptation. *Journal of Linguistics* 33: 379–430.

Paradis, Carole, and Darlene LaCharité. 2001. Guttural deletion in loanwords. *Phonology* 18: 255–300.

Paradis, Carole, and Jean-Francois Prunet. 1988. Locality in a theory of constraint violations. *GLOW Newsletter* 22: 46–48.

Paradis, Carole. 1988a. On constraints and repair strategies. *The Linguistic Review* 6: 71–97.

Paradis, Carole. 1988b. Towards a theory of constraint violations. *McGill Working Papers in Linguistics* 5: 1–43.

Paradis, Carole. 1995a. Native and loanword phonology as one: the role of constraints. In *Proceedings of the 13th International Congress of Phonetic Sciences*, ed. K. Elenius and P. Branderud, 74–82. Stockholm University, Stockholm.

Paradis, Carole. 1995b. Derivational constraints in phonology: Evidence from loanwords and implications. In *Proceedings of the 31st Chicago Linguistic Society Meeting*, ed. A. Dainora, R. Hemphill, B. Luka, B. Need, and S. Pargman, 360–374.

Parker, Frank, and Thomas Walsh. 1981. Voicing cues as a function of the tense/lax distinction in vowels. *Journal of Phonetics* 9: 353–358.

Peperkamp, Sharon, and Emmanuel Dupoux. 2003. Reinterpreting loanword adaptations: The role of perception. In *Proceedings of the 15th International Congress of Phonetic Sciences*, ed. M. J. Solé, D. Recasens, and J. Romero, 367–370. Adelaide: Causal Productions.

Peperkamp, Sharon, Inga Vendelin, and Kimihiro Nakamura. 2008. On the perceptual origin of loanword adaptations: Experimental evidence from Japanese. *Phonology* 25: 129–164.

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

Peterson, Gordon E., and Ilse Lehiste. 1960. Duration of syllable nuclei in English. Journal of the Acoustical Society of America 32: 693–703.

Pickett, James M. 1999. *The Acoustics of Speech Communication: Fundamentals, Speech Perception Theory, and Technology*. Boston: Allyn and Bacon.

Pittt, Mark A. 1998. Phonological processes and the perception of phonotactically illegal consonant clusters. *Perception and Psychophysics* 60: 941–951.

Price, Patti J. 1981. *A Cross-Linguistic Study of Flaps in Japanese and in American English*. Doctoral Dissertation, University of Pennsylvania.

Repp, Bruno H. 1980. Perception and production two-stop-consonant sequences. Haskins Laboratories Status Report on Speech Research 63/64: 177–194.

Repp, Bruno H. 1982. Perceptual assessment of coarticulation in sequences of two stop consonants. *Haskins Laboratories Status Report on Speech Research* 71/72: 131–166.

Repp, Bruno H., and Katyanee Svastikula. 1987. Perception of the [m]-[n] distinction In VC syllables. *Haskins Laboratories Status Report on Speech Research* 91: 157–175.

Rose, Yvan, and Katherine Demuth. 2006. Vowel epenthesis in loanword adaptation. *Lingua* 116: 1112–1139.

Rose, Yvan. 1999. A structural account of root node deletion in loanword phonology. *Canadian Journal of Linguistics* 44: 359–404.

Schourup, Lawrence C. 1973. A cross-language study of vowel nasalization. *Ohio State University Working Papers in Linguistics* 15:190–221.

Schwab, Sandra, and Joaquim Llisterri. 2011. Are French speakers able to learn to perceive lexical stress contrasts? In *Proceedings of the 17th International Congress of Phonetic Sciences*, ed. W.-S. Lee and E. Zee, 1774–1777. Hong Kong: City University of Hong Kong.

Schwab, Sandra, and Joaquim Llisterri. 2014. Does training make French speakers more able to identify lexical stress? In *Proceedings of the International Symposium on the Acquisition of Second Language Speech, Concordia Working Papers in Applied Linguistics* 5: 624–636.

Shi, Xiujuan. 2010. Beijinghua yinjie yinlian zhong yuanyin de bihua du (Zhongwen). *Dijiu jie Zhongguo yuyinxue xuehui huiyi lunwenji*. [Extent of vowel nasalization in syllables across a word boundary (Chinese). In *Proceedings of the 9th Phonetic Conference of China*.] http://cpfd.cnki.com.cn/Article/CPFDTOTAL-EGYO201005001137.htm

Shinohara, Shigeko. 2004. Emergence of universal grammar in foreign word adaptation. In Constraints in Phonological Typology and Acquisition, ed. R. Kager, J. Pater, and W. Zonneveld, 292–320. Cambridge: CUP.

Shinohara, Shigeko. 2006. Perceptual effects in final cluster reduction patterns. Lingua 116: 1046–1078.

Silverman, Daniel. 1992. Multiple scansions in loanword phonology: Evidence from Cantonese. *Phonology* 9: 289–328.

Smith, Caroline L. 1997. The devoicing of /z/ in American English: Effects of local and prosodic context. *Journal of Phonetics* 25: 471–500.

Smith, Jennifer L. 2006. Loan phonology is not all perception: Evidence from Japanese loan doublets. In *Japanese/Korean Linguistics* 14, ed. Timothy J. Vance and Kimberly A. Jones, 63–74. Palo Alto: CSLI.

Smith, Jennifer L. 2009. Source similarity in loanword adaptation: Correspondence Theory and the posited source-language representation. In *Phonological Argumentation: Essays on Evidence and Motivation*, ed. Steve Parker, 155–177. London: Equinox.

Solé, Maria-Josep. 1995. Spatio-temporal patterns of velopharyngeal action in phonetic and phonological nasalization. Language and Speech 38: 1–23.

Sproat, Richard, and Osamu Fujimura. 1993. Allophonic variation in English /l/ and its implications for phonetic implementation. *Journal of Phonetics* 21: 291–311.

Stemberger, Joseph P. 1991. Radical underspecification in language production. *Phonology* 8: 73–112.

Steriade, Donca. 1997. Phonetics in phonology: The case of laryngeal neutralization. Ms., UCLA.

Steriade, Donca. 2001. Directional asymmetries in place assimilation: A perceptual account. In *The Role of Speech Perception in Phonology*, ed. E. Hume and K. Johnson, 219–250. San Diego: Academic Press.

Steriade, Donca. 2009. The phonology of perceptibility effects: The P-map and its consequences for constraint organization. In *The Nature of the Word: Studies in Honor of Paul Kiparsky*, ed. K. Hanson and S. Inkelas, 151–180. Cambridge, MA: MIT Press.

Takagi, Naoyuki, and Virginia Mann. 1994. A perceptual basis for the systematic phonological correspondences between Japanese loan words and their English source words. *Journal of Phonetics* 22: 343–356.

Tobin, Stephen, Dani Byrd, Eric Bresch, and Shrikanth Narayanan. 2006. Syllable structure effects on velum-oral coordination evaluated with real-time MRI. *Journal of the Acoustical Society of America* 119: 3302.

Tranel, Bernard. 1987. *The Sounds of French: An Introduction*. Cambridge: Cambridge University Press.

Tranter, Nicolas. 2000. The phonology of English loan-words in Korean. *Word* 51: 377–404.

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

Ulrich, Charles H. 1997. Loanword adaptation in Lama: testing the TCRS model. *Canadian Journal of Linguistics* 42: 415–463.

Umeda, Noriko. 1975. Vowel duration in American English. *Journal of the Acoustical Society of America* 58: 434–445.

Vance, Timothy J. 1987. *An Introduction to Japanese Phonology*. Albany, NY: State University of New York Press.

Vendelin, Inga, and Sharon Peperkamp. 2004. Evidence for phonetic adaptation of loanwords: an experimental study. In *Actes des Journées d'Etudes Linguistiques de l'Université de Nantes*, 127–131.

Vendelin, Inga, and Sharon Peperkamp. 2006. The influence of orthography on loanword adaptations. *Lingua* 116: 996–1007.

Whalen, Douglas H., A. Min Kang, Harriet S. Magen, Robert K. Fulbright, and John C. Gore. 1999. Predicting midsagittal pharynx shape from tongue position during vowel production. *Journal of Speech, Language, and Hearing Research* 42: 592–603.

Yamada, Reiko A., and Yoh'ichi Tohkura. 1991. Age effects on the acquisition of non-native phonemes: Perception of English /r/ and /l/ for native speakers of Japanese. In *Proceedings of the 12th International Congress of Phonetic Sciences*, Vol. 4, 450–453.

Yeon, Sang-Hee. 2003. Perception of English palatal codas by Korean speakers of English. Paper presented at the 145th Meeting of the Acoustical Society of America, Nashville, Tennessee.

Yip, Moira. 1993. Cantonese loanword phonology and Optimality Theory. *Journal of East Asian Linguistics* 2: 261–291.

Yip, Moira. 2002. Perceptual influences in Cantonese loanword phonology. *Journal of the Phonetic Society of Japan* 6: 4–21.

Yip, Moira. 2006. The symbiosis between perception and grammar in loanword phonology. *Lingua* 116: 950–975.

Zee, Eric. 1981. Effect of vowel quality on perception of post-vocalic nasal consonants in noise. *Journal of Phonetics* 9: 35–48.

Zsiga, Elizabeth C. 2003. Articulatory timing in a second language: Evidence from Russian and English. *Studies in Second Language Acquisition* 25: 399–432.