**Phonological and Phonetic Asymmetries of Cw Combinations**

A Dissertation Presented

by

**Yunju Suh**

to

The Graduate School

in Partial Fulfillment of the

Requirements

for the Degree of

**Doctor of Philosophy**

in

**Linguistics**

Stony Brook University

**August 2009**

**Stony Brook University**

The Graduate School

**Yunju Suh**

We, the dissertation committee for the above candidate for the
Doctor of Philosophy degree, hereby recommend
acceptance of this dissertation.

**Marie K. Huffman – Dissertation Advisor**
**Associate Professor, Department of Linguistics**

**Ellen Broselow**
**Professor, Department of Linguistics**

**Christina Y. Bethin**
**Professor, Department of Linguistics**

**Edward Flemming**
**Associate Professor, Department of Linguistics and Philosophy,**
**Massachusetts Institute of Technology**

This dissertation is accepted by the Graduate School

Lawrence Martin
Dean of the Graduate School

Abstract of the Dissertation

**Phonological and Phonetic Asymmetries of Cw Combinations**

by

**Yunju Suh**

**Doctor of Philosophy**

in

**Linguistics**

Stony Brook University

**2009**

This thesis investigates the relationship between the phonological distribution of Cw combinations, and the acoustic/perceptual distinctiveness between syllables with plain C onsets and with Cw combination onsets. Distributional asymmetries of Cw combinations discussed in this thesis include the avoidance of Cw combinations in the labial consonant context and before rounded or back vowels, and the preference for contrastive labialization on velar consonants.

Following the claim that the phonetic salience of phonological contrasts should be taken into account in the explanation of typologically common co-occurrence restrictions (Kawasaki 1982, Flemming 1995, et seq.), I focus on showing that the salience of the contrast between CV and CwV syllables is a major factor in the asymmetric distribution of Cw combinations, with supporting evidence from acoustic and perception experiments. The results of a study on the C-Cw comparison in Korean and Spanish show that the acoustic difference between plain C and Cw combination is weaker in contexts where Cw

combinations are avoided. In addition, cross-linguistic comparison reveals that the phonetic status of the /w/ component—a glide consonant or a secondary labialization—affects the acoustic distinctiveness of a Cw combination relative to the plain C counterpart. Perception experiment results confirm that the degree of acoustic distinctiveness between CV and CwV syllables in different consonant and vowel contexts is reflected in the strength of perceptual cues to the contrast.

Based on the findings from the experimental studies, I propose an analysis of the Cw asymmetries within the Dispersion Theory of Contrast (Flemming 1995, et seq.). The asymmetric distribution of Cw combinations emerges because the contrast between CV and CwV is allowed only when they are separated by sufficient perceptual distance. Optimality-Theoretic interaction between such requirements on perceptual distance (minimal distance constraints) and constraints governing the preservation of contrast gives rise to the asymmetries in the distribution of Cw combinations observed in languages like Korean, Mandarin, Cantonese and Dan. I compare the current analysis with an alternative identity avoidance (OCP) analysis, and argue that the former is superior in terms of both descriptive and explanatory adequacies. I also show the parallel between the distributional asymmetries of Cw and Cj combinations, and extend the Dispersion-Theoretic analysis to the Cj distribution for languages like Ukrainian.

## Table of Contents

**List of Figures**

## List of Tables

## Acknowledgements

I would like to thank the following individuals for their help in completing this dissertation. My deepest gratitude goes to my advisor Marie Huffman, for guiding me through every step of the research. Without her continuing encouragement and support this dissertation wouldn't have been possible. I am immensely indebted to Ellen Broselow for her insightful comments and suggestions on the dissertation, and for all I have learned from attending her seminars. Christina Bethin has also been a great help with her detailed comments on both the contents and organization of the dissertation, and acted as the chair of the defense. Edward Flemming kindly agreed to be in the dissertation committee as an outside member. His pioneering work in the Dispersion Theory had been a great influence in initiating this research. Richard Larson has given me financial support by appointing me as a co-organizer for the Morris Symposium on Language and Communication while I was writing this dissertation. Susana Huidobro provided Spanish data for the experiments and helped with the participant recruitment. Jiwon Hwang, Andrei Antonenko, and Giorgos Tserdanelis helped with speech analysis and statistics software. Yu-an Lu, Chih-hsiang Shu and Roksolana Mykhaylyk provided data and judgment of the Mandarin and Ukrainian languages.

This dissertation research was partly supported by NSF Grants #0325188 and #0746027.

# Chapter 1. Phonological and phonetic aspects of Cw asymmetries

Phonological studies have established that sequencing of sounds within a syllable follows a certain pattern, known as the Sonority Sequencing Principle, which requires sonority to rise from the onsets toward the nucleus, and to fall from the nucleus to codas (e.g., Selkirk 1982; Kenstowicz 1994:254). It has also been found that languages prefer to have a certain degree of sonority difference between the consonants preceding the nucleus (Minimum Sonority Distance Principle, Harris 1983; Sonority Dispersion Principle, Clements 1990). Since the sonority scale generally does not distinguish among places of articulation, consonant combinations that have the same degree of internal sonority distance form equally good clusters in terms of sonority distance or dispersion. However, it is often observed that consonant combinations of the same sonority profile are not equally represented across languages. For example, in English, French and many other languages, [tl] and [dl] sequences do not occur within a syllable onset, unlike [pl], [bl], [kl] and [gl]. Likewise, the [pw] combination is often not allowed in the syllable onset of languages in which [tw] and [kw] are legitimate complex onsets, e.g., English, Korean. Such *asymmetry* in the distribution of sound sequences of same sonority profile does not follow from the general principle of sonority dispersion, yet the fact that similar patterns prevail across many languages suggests that these may form an important cross-linguistic trend that calls for an explanation.

It has been claimed that the acoustic and perceptual salience of sounds or sound sequences gives rise to their asymmetric distributions found in many languages (Kawasaki 1982; Flemming 1995, 2002; Padgett 2001, 2003; Narayan 2008, among others). In this approach, co-occurrence restrictions emerge from the consideration of phonetic salience: a sound or a sound string is often absent if it lacks sufficient distinctiveness from others.

The previous studies highlight the relationship between phonetic and phonological asymmetries by reporting the results of acoustic or perceptual experiments (e.g., Kawasaki 1982, Padgett 2001, Narayan 2008) or proposing phonological models that are able to incorporate phonetic aspects into phonological analysis (e.g., Flemming 1995, 2002; Padgett 2003). In this thesis I integrate experimental results with a formal analysis, focusing on the distributional restrictions related to *Cw combinations*, which include both Cw clusters and consonants with the secondary articulation of labialization. I present an acoustic analysis and a perception experiment examining the asymmetry in the relative salience of the C-Cw contrast, and propose a phonological analysis built upon the results from the experiments, within the Dispersion Theory of Contrast (Flemming 1995, 2002).

The rest of this chapter is organized as follows. Phonological patterns of Cw combinations are discussed in 1.1, with particular interest in their asymmetric distribution found in many languages in terms of consonant places and vowels, as well as the phonetic status of the /w/ component (i.e., a glide consonant or secondary labialization). It is shown that labial consonants and back vowels are less likely to co-occur with a prevocalic /w/, whereas velar consonants are preferred in Cw combinations. Section 1.2 discusses how the phonetic salience of /w/ is expected to vary depending on the consonant and vowel contexts, and the status of /w/. It is argued that the phonological distribution of Cw combinations is related to the phonetic salience of /w/, as labial consonants and back vowels lower the phonetic salience of /w/, and velar consonants enhance it. Theoretical background for the integration of the phonological and phonetic asymmetries is introduced in 1.3, on the basis of which the analysis of the Cw distribution is developed in the following chapters.

## 1.1 Phonological asymmetries of Cw combinations

Because languages do not always have foolproof evidence regarding the status of consonant + glide combinations, this study is not confined to clear-cut cases of Cw *clusters*, but includes cases that may be considered consonants with contrastive secondary labialization. For example, controversy exists as to whether Mandarin CG combinations are to be considered single labialized consonants (e.g., Duanmu 1990, 2000) or clusters (e.g., Chan 1985). A more important reason to take into consideration both Cw clusters and labialized consonants in the description of Cw co-occurrence restrictions is that in some languages Cw combinations behave as a unit segment in certain contexts and as sequences in other contexts (See 1.1.2).

Strictly speaking, the term “secondary labialization” refers to the addition of lip rounding, and does not necessarily imply the presence of tongue back raising or velarization. It may therefore sound misleading to draw a parallel between Cw clusters and labialized consonants, considering their articulatory correlates. However, according to Ladefoged and Maddieson (1996:356), labialization often accompanies velarization, and they use the term "labialization" for labio-velarization. The frequent co-occurrence of labial and velar secondary articulations is not a coincidence if we consider their acoustic effects: the lowering of formant frequencies is the primary acoustic result of both lip protrusion and tongue backing (Ohala and Lorentz 1977:584). Studies investigating secondary labialization (Zue 1976, Gordon et al. 2002, de Jong and Obeng 2000, Ladefoged and Maddieson 1996) and velarization (Padgett 2001, Choi 1995) have shown the lowering of formants (especially F2) to be the major acoustic correlate of labialization. Therefore, secondary velarization may often accompany labialization as a contrast enhancement (Stevens, Keyser and Kawasaki 1986, Ladefoged and Maddieson 1996, 356).

### 1.1.1 Consonant place and vowel asymmetries

Based on a survey of 251 languages, Kawasaki (1982:16) finds that languages tend to disallow labial consonants before a labiovelar glide /w/; examples of such languages are English, Korean, Ronga, Tarascan, Urhobo, Vietnamese and Zulu. Is the distribution of contrastive labialization similar to that of Cw clusters? Kawasaki claims that this may not be the case, since secondary labialization often “seems to have originated as assimilatory allophonic labialization of a rounded vowel onglide because of the ‘inherently labialized’ nature of labial consonants” (Kawasaki 1982:17). Nevertheless, according to her survey of languages, the predominant place for phonemic labialization is dorsal, though a few languages have labialized labials (2 out of 40) and labialized dentals/alveolars (2 out of 40). Similar results are obtained when we look at the distribution of labialized consonants in the languages represented in UPSID (UCLA Phonological Segment Inventory Database: Maddieson and Precoda 1992). Table 1 shows the number of languages that have contrastive labialization according to UPSID, classified according to language families, and the percentage these languages occupy out of the total 451 languages in the database.

Table 1.1. Languages with contrastively labialized consonants in UPSID

| | Labial | Coronal | Velar |
|---|---|---|---|
| Afro-Asiatic | 2 | 1 | 8 |
| Australian | | | 1 |
| Austro-Tai | | | 6 |
| Caucasian | | 2 | 3 |
| East Papuan | | 1 | |
| Eskimo-Aleut | | | 1 |
| Indo-European | 1 | | |
| Khoisan | | | 1 |
| Na-Dene | | 1 | 5 |
| Niger-Kordofanian | | 1 | 7 |
| North American | | | 23 |
| Northwest Caucasian | | 1 | 1 |
| Papuan | 2 | | 5 |
| Sino-Tibetan | | | 1 |
| South American | 1 | 1 | 11 |
| Total | 6 | 8 | 73 |
| | 1.33% | 1.77% | 16.19% |

The most striking fact in Table 1.1 is the prevalence of labialized velar consonants. In other words, many languages employ phonemic labialization only on velar consonants. Therefore as Kawasaki (1982) suggests, the universal tendency in labialized consonants is best stated as preference for velar place (rather than the rarity of the labialized labials).

Though velars are the predominant labialized consonants in UPSID languages, it is not the case that velar is the "default" consonant place for labialization. In other words, the existence of labialization on labials or coronals does not necessarily imply the existence of labialized velars. There are three languages in UPSID (Ponapean, Lenakel, both Austro-Tai and Easy Malayo-Polynesian languages, and Irish) in which labialization is found only on labial consonants. Among them, Irish is the only language with a phoneme that UPSID categorizes as *labialized velarized* labial consonants. It is well known that the Irish non-palatalized labial and coronal consonants are velarized (Green

1997, Ní Chiosáin and Padgett 2001). However, according to Green (1997:42), additional lip protrusion on Irish velarized consonants occurs only on a velarized labial before a front vowel. Thus it seems that the Irish labialized velarized labial is essentially a velarized labial, with lip protrusion added as a contrast enhancement strategy vis-a-vis their palatalized counterpart. This leaves only Ponapean and Lenakel as languages with labialization contrast occurring only on labial consonants, but the labialized labials in these languages are also described as velarized labials by Rehg and Sohl (1981) and Lynch (1978). Furthermore, it is also reported that the labialization component of the secondary articulation does not occur in word-final position in Ponapean (Rehg and Sohl 1981: 28). Therefore it may be more appropriate to view the labialized labials of these languages as velarized labials, with an additional lip protrusion (See 4.5 for more discussion of Ponapean).

In addition to the consonant place-related asymmetry, Kawasaki also reports that Cw combinations are often avoided in certain vowel contexts (Kawasaki 1982, 19-21), such as rounded or back vowels (e.g., Yao, Zulu, Huichol, Toura, Amharic). Furthermore, she notes that some languages allow labialized consonants only before front vowels (e.g., Navaho, Gã), low vowels or front and low vowels (e.g., Kpelle, Nupe).

Below I look more closely at the distribution of Cw combinations in two languages that show asymmetrical distribution of the Cw combinations, namely Korean and Mandarin. These languages have Cw combinations whose status (clusters or labialized consonants) is not uncontroversial. They exhibit the place and vowel asymmetries similar to those observed in languages discussed above.

*Korean.* In the Korean syllable, a vowel can be preceded by a singleton consonant, or a combination of a consonant and a glide /w/ or /j/. Following are examples of minimal pairs that differ in the presence of a prevocalic /w/:

| | | | | |
|---|---|---|---|---|
| (1) | tedʒi | 'land' | twedʒi | 'pig' |
| | tsa | 'ruler' | tswa | 'left' |
| | ke | 'crab' | kwe | 'chest' |
| | kʌl | 'to hang' | kwʌl | 'palace' |

As for the status of a prevocalic glide, there have been two different phonological points of view: One position is to affiliate the glide to the nucleus of a syllable, forming a diphthong (e.g., Sohn 1987). The other position is to consider the glide a member of the onset. Among the studies that take the latter position, there are also two different analyses for the internal structure of the consonant + glide (CG) combination in onset: the glide is argued to form a consonant cluster (branching onset) by B.G. Lee (1982), and Y. Lee (1994), whereas Ahn (1985) proposes that CG forms a contour consonant, i.e., a consonant with a secondary articulation. Tables 2 and 3 are the consonant and vowel inventories of Standard Korean. Note that there are three voicing types for Korean obstruents: *aspirated* ($C^h$), *lenis* or *slightly aspirated* (C), and *fortis* (C').[1]

[1] The fortis stop series is produced with glottal tension evidenced by the intensity of the following vowel (Kim 1965). Among the three types of Korean stop consonants, fortis or tense stops are characterized by the shortest VOT and the longest duration of the following vowel. They also show the longest closure duration, more than twice as long as that of the lenis counterpart in

Table 1.2. Korean consonant phonemes

| | Bilabial | Dental/alveolar | Palatal | Velar | Glottal |
|---|---|---|---|---|---|
| Stop | pʰ p p' | tʰ t t' | | kʰ k k' | |
| Fricative | | s s' | | | h |
| Affricate | | t͡sʰ t͡s t͡s'[2] | | | |
| Nasal | m | n | | ŋ | |
| Liquid | | l | | | |
| Glide | (w) | | (j) | (w) | |

Table 1.3. Korean vowel phonemes

| | | |
|---|---|---|
| i (y) | ɨ | u |
| e (ø) | ʌ | o |
| (ɛ) | a | |

Standard (Contemporary) Korean has seven vowels (Kim and Heo 1999, Kwon 2001 among others). Front vowels /e/ and /ɛ/ are merged to /e/ in contemporary Standard Korean (Lee 1996: 123; Kwon 2001). Traditionally the Korean vowel system has been analyzed as having front rounded vowels /y/ and /ø/, which are still retained among older generation speakers. However, in contemporary Standard Korean, the high front rounded vowel /y/ became /ɥi/, and mid front rounded vowel /ø/ is merged with (a diphthong) /we/ (Kim 1975, Kim and Heo 1999, Kwon 2001; Ito, Kang and Kenstowicz 2006, fn 4).

The combination of /w/ and a vowel, often referred to as a diphthong, has limited distribution, as shown in (2): High back unrounded vowel /ɨ/ and rounded vowels /o/ and /u/ do not follow /w/.

(2) /w/ + vowel combination (or w diphthongs) of contemporary Standard Korean

| | | |
|---|---|---|
| wi | -- | -- |
| we | wʌ | -- |
| | wa | |

Moreover, when the prevocalic glide co-occurs with a consonant, there are additional restrictions on vowels (3). Note that the laryngeal feature difference is ignored in (3) for

intervocalic position. Han (1992) proposes that fortis stops are the geminates of lax (lenis) consonants, and the glottal tension is the phonetic implementation of gemination, with a loss of the length contrast in the initial position. However, Kim (2002) argues against this by showing that glottal tension and durational properties of fortis stops are consistent regardless of the position in a word.

[2] The place of articulation of the Korean coronal affricates is the same as that of the stop and fricative (Kim 1999, 2001). Thus they are transcribed as /t͡s/ rather than /tʃ/ in Table 2.

simplicity. Question marks indicate the combinations that only occur in onomatopoeic words.

(3) Combination of a consonant and /w/ found in the initial position of monomorphemic forms in Korean (National Institute of the Korean Language 2003)

| | **p** | **t** | **s** | **t͡s** | **k** | **h** |
|---|---|---|---|---|---|---|
| **wi** | -- | ✓ | ✓ | ✓ | ✓ | ✓ |
| **we/ɛ** | -- | ✓ | ✓ | ✓ | ✓ | ✓ |
| **wa** | -- | ✓ | ? | ✓ | ✓ | ✓ |
| **wʌ** | -- | -- | ? | -- | ✓ | ✓ |

Labial consonants do not occur in Cw combinations.[3] Kang (1997) analyzes the loss of /w/ after /p/ as a recently completed language change, as reflected in a recent orthographic change. For coronal consonants, the precise description of co-occurrence restrictions on CwV syllables requires a reference to the manner of articulation, as well as the vowel properties. Coronal fricatives have a limited occurrence in Cwa and Cwʌ syllables, occurring mainly in onomatopoeic words. Dental/alveolar stops and affricates are not found in Cwʌ forms. Velar and glottal consonants can co-occur with /w/ before all four vowels /i e a ʌ/. Thus labials have the most restricted occurrence in CwV syllables, whereas velar and glottal consonants are the least restricted.

Despite the non-occurrence or marginal status of certain CwV combinations in the Korean lexicon, these forms are not completely illegitimate in the Korean grammar. This is because the CwV syllables unattested in monomorphemic forms freely occur across morpheme boundaries. For example, verbal suffixes beginning with a vowel may cause contraction when they combine with a verb stem ending in a vowel: depending on the vowels, a suffix or stem vowel may delete or be realized as a glide. In particular, verb stem-final *u* or *o* becomes *w* before vowel-initial suffixes in contracted forms (4).

(4)

| | | | |
|---|---|---|---|
| po + as' | poat | pwat | 'see + past' |
| s'o + as' | s'oat | s'wat | 'shoot + past' |
| k'o + as' | k'oat | k'wat | 'twist + past' |
| tu + ʌs' | tuʌt | twʌt | 'put + past' |
| $ts^h$u + ʌs' | $ts^h$uʌt | $ts^h$wʌt | 'dance + past' |

The /pw/ combinations are also found word-internally in Sino-Korean words whose poly-morphemic status is not uncontroversial in contemporary Korean. A morpheme-final consonant /p/ is syllabified with the prevocalic glide in the following morpheme in (5).

(5)

| | | | | |
|---|---|---|---|---|
| a. happʌ**p** + hwa | *legal + conversion* | → | hap'ʌ.**p**$^h$wa | 'legalization' |
| b. pʌ**p** + wʌn | *law + institute* | → | pʌ.**b**wʌn | 'court' |

[3] Initial /pwe/ is in fact attested in word-initial position: pwe 'to meet a senior', but this verb is historically derived from *po* 'see' + *i*'.

Aspirated stops are not allowed in coda position in Korean, which tells us that the labial stop at the end of 'legal' in (5a) is in the onset position in the compounded form 'legalization'. Also the voicing of the stem final labial consonant in /pʌp/ 'law' in the compounded form shows that it is not in syllable final position, since coda stops all surface as voiceless lenis. Therefore, [$p^h$] and [b] in compound forms are in the onset position with the glide [w], but not in the coda of the preceding syllables. Because of the unrestricted occurrences of Cw combinations in polymorphemic forms, the discussion of asymmetric distribution is limited to the morpheme structure conditions (MSCs).

Another interesting asymmetry between consonants in Cw combinations is found in words borrowed from other languages. Following are examples of loanwords taken from *oyleye yonglyey phyokicip* (‘Loanword examples’), published by the National Institute of Korean Language (2008).

(6) Loanwords with Cw clusters

a. Coronal/ Labial + w

| | |
|---|---|
| Twain, (Mark) | **$t^h$ɨ.w**e.in |
| twist | **$t^h$ɨ.w**i.sɨ.tɨ$^h$ |
| Dwight (Eisenhower) | **tɨ.w**a.i.$t^h$ɨ |
| Swaziland | **sɨ.w**a.tsil.len.dɨ |
| sweater | **sɨ.w**e.$t^h$ʌ |
| suede | **sɨ.w**e.i.dɨ |
| Duero | **tu**.e.ɾo |
| Buenos (Aires) | **pu**.e.no.sɨ |

b. Velar + /w/

| | |
|---|---|
| Quaker | **$k^h$w**e.i.$k^h$ʌ |
| quiz | **$k^h$w**i.tsɨ |
| quota | **$k^h$w**ʌ.$t^h$ʌ |
| Gwyneth (Paltrow) | **kw**i.ne.sɨ |
| Le Guin, (Ursula) | ɾɨ.**kw**in |
| Guatemala | **kw**a.$t^h$e.mal.la |
| Guam | **kw**am |
| Quasimodo | **$k^h$w**a.tsi.mo.to |
| Kwashiorkor | **$k^h$wʌ.**si.o.$k^h$ʌ |

Based on (6a-b) the following generalizations can be made: first, even when the source words contain a CwV- syllable which is banned by Korean phonotactics, /w/ is not deleted in the borrowed forms. Second, it is only when the consonant is velar that Cw combinations in foreign words are adapted as the Korean Cw combinations. Otherwise, a Cw cluster is broken into two syllables, with an epenthetic vowel [ɨ] inserted, or /w/ is "promoted" to the syllable nucleus /u/. This seems to roughly match with the patterns found in the phonotactics of native Korean words, showing the preference for velar to other consonants in Cw combinations. However, it remains unclear why /tw/ or /sw/ clusters are never adapted as Korean /tw/ or /sw/, and why /kw/ clusters do not incur vowel epenthesis, which is a general process in loanword adaptation of consonant clusters into Korean.

*Mandarin.* Like Korean, Mandarin allows complex onsets only when the second member is a glide. The prevocalic glide /w/ and /j/ are traditionally categorized as "medial". Medial /w/ occurs before the mid vowels and the low vowel, as shown in (7). Note that the low vowel undergoes allophonic alternation between [a] and [ɑ].

(7) Distribution of wV- in Mandarin (adapted from Duanmu 2000:61)

| | | |
|---|---|---|
| weɪ | wən | wɔ |
| | wa(ɪ/n) | wɑŋ |

Controversy exists in the literature regarding the status of the "medial", as well as the syllable structure of Mandarin: The phonological status of the medial is not directly relevant to the current study, but I follow Duanmu (1990, 2000) and Li (1999) and assume that the medial is part of the onset rather than the nucleus.

The Cw combinations occur in various consonant contexts, but their distribution is more restricted in some consonant contexts than in the others:

(8) Distribution of the Mandarin CwV- syllables

| | Labial p/pʰ/f | Dental t/tʰ/s/ts/tsʰ | Retroflex ʂ/tʂ/tʂʰ | Velar k/kʰ/x |
|---|---|---|---|---|
| weɪ | -- | ✓ | ✓ | ✓ |
| wə | -- | ✓ | ✓ | ✓ |
| wɔ | ✓ | ✓ | ✓ | ✓ |
| wa-(ɪ/n) | -- | ✓ | ✓ | ✓ |
| wɑŋ | -- | -- | ✓ | ✓ |

(http://www.mandarintools.com/chardict.html)

Combinations of labial + /w/ are the most restricted, as they occur only before the rounded vowel /ɔ/.[4] As in Korean, the velar + /w/ combination occurs in the widest range of vowel environments, whereas the coronal + /w/ combinations are more restricted in their distribution. Dental (t, tʰ, s, ts and tsʰ) + /w/ combination is banned before the [ɑ] allophone, which occurs before the coda /ŋ/, but found before [a], which occurs elsewhere.[5] Among coronals, retroflex fricatives and affricates co-occur with /w/ before all four vowels, like velars.

The distribution of the Korean and Mandarin Cw combinations in some sense instantiates the reported co-occurrence restrictions of Cw clusters and labialized consonants shown above: First, both languages exhibit the most severe restrictions on labial + /w/ combinations: Korean bans them in all vowel contexts in which a Cw combination can occur, whereas Mandarin allows them only before /ɔ/. Second, though these languages allow dental + /w/ combinations, their distribution is more limited than velar + /w/, as they don't occur before back vowels /ʌ/ (Korean) and /ɑ/ (Mandarin). This

[4] Labial consonant + wɔ is written in Pinyin as Co, unlike other Cwɔ syllables which are written as Cuo (Miao 2005:42).

[5] Mandarin allows only the nasals /n/ and /ŋ/ in coda position. The front allophone of low vowel ([a]) occurs in open syllable, before coda /n/, and in a falling diphthong /aɪ/, as shown in (11).

pattern is a combination of two facts regarding the distribution of labialized consonants: that the secondary labialization occurs most often on velar consonants, and that it is avoided before back or rounded vowels.

### 1.1.2 Consonant place and Cw status asymmetries

As mentioned in the beginning of this section, it is not always easy to determine whether Cw combinations of a particular language are Cw clusters or labialized consonants. One criterion used by researchers to decide the status of Cw (or CG combinations in general) is the productivity of these combinations in the language (Ladefoged 1968 for West African languages; Chan 1985 for Mandarin Chinese): if Cw combinations occur on a relatively wide range of consonants, it is more plausible to view them as clusters than labialized consonants, considering economy of the description (i.e., to minimize the number of phonemes). In contrast, if a language has a limited number of Cw combinations, say, only velar + /w/, then it is more desirable to categorize it as a unit segment (rather than stating the gaps in other consonants + w clusters). However, even in languages which allow Cw combinations in various consonant places, there is evidence that /kw/ behaves as a unit segment, whereas other Cw combinations are classified as clusters. Following are some examples of such languages: Tarascan (South American, Chibchan), Dan (Niger-Congo, Mande) and Fox (North American, Algonquian).

*Tarascan.* In Tarascan, velar + /w/ combinations have special status, "somewhere between true unit morphophonemes and true clusters" (Friedrich 1975:32). The assignment of segmental status to /kw/ is based on both its phonetic nature (Friedrich 1975:30) and its phonological properties: Tarascan allows Cw combinations when C is dental/alveolar or velar (9a), but three-member consonant clusters found in this language always contain /k$^h$w/ (9b) (Foster 1968:31).

| | | | |
|---|---|---|---|
| (9) | a. | kwíni | 'bird' |
| | | k$^h$wíni | 'to sleep' |
| | | swánarini | 'to put a warm cloth on one's eye' |
| | b. | ck$^h$wániarini | 'to have spots on one's face' |
| | | čk$^h$ántirani | 'to lie' |
| | | tk$^h$wíšuni | 'to kneel in a canoe' |

*Dan (Santa).* In Dan, consonants of all three major places—labial, coronal and velar—combine with a glide /w/ or /j/. The status of these CG combinations is phonologically ambiguous between clusters and consonants with secondary palatalization or labialization. According to the auditory judgment of the authors, "the segment following C may sometimes be identified as a vowel-like sound" (Bearth and Zemp 1967:15). However, they observed that /kw/ and /gw/ are exceptions to this variation, and are always realized as labialized consonants.

Evidence that velar + /w/ is a unit segment unlike other CGs also comes from the phonological perspective: When a CG combination is followed by a long vowel, {w, j} + long vowel alternates with {u, i} + short vowel. However, when the CG combination is a velar + /w/, the glide cannot alternate with the vowel /u/ (10).

| | | | |
|---|---|---|---|
| (10) | ɓjəː ~ ɓiə | 'cord' | |
| | kwɛː ~ *ku.ɛ | 'loom' | (Bearth and Zemp 1967:22) |

Sagey (1986) argues that this is because /kw/ and /gw/ form contour segments, so /w/ is always under the same x-slot with /k/ or /g/ (11b), unlike other CG combinations such as (11a), which have two possible structures.

(11)

```
a.  σ                   σ
   / \                 / \
O   N                 O N
|    /\               | / \
x    x x      OR      x x x
|     \/              | |  |
ɓ^j   ə               ɓ j  ə          'cord'

b.  σ
   / \
O     N
|     /\
x    x x
|     \/
k^w   ɛ                               'loom'
```

(Adapted from Sagey 1986, (51))

*Fox.* Fox allows CG combinations in different consonant places, but Dahlstrom (1997) argues that only /kw/ is a single consonant, based on the evidence from reduplication. Examples are given in (12), with the reduplicants underlined. When a stem begins with a CG combination, monosyllabic reduplication copies the initial C or both C and G (12a-b), except when it is /kw/, as shown in (12c). In this case the full copy of /kw/ is obligatory. If all CG combinations have the same status of consonant clusters, this different is not expected.[6]

(12) Dahlstrom (1997:212)

| | | |
|---|---|---|
| a. kja:t-amwa | ka:-kja:t-amwa | 'he keeps it for himself' |
| | kja:-kja:t-amwa | |
| b. ʃwa:ʃika | ʃa:-ʃwa:ʃika | 'eight' |
| | ʃwa:-ʃwa:ʃika | |
| c. kwe:hta:nite:he:-wa | *ke:-kwe:hta:nite:he:-wa | 'he feels terrible' |
| | kwe:-kwe:hta:nite:he:-wa | |

[6] Kickapoo, a language closely related to Fox, also shows evidence that /kw/, unlike other CG combinations, forms a unit segment (Voorhis 1982).

In sum, the velar consonant + /w/ combination is special in two respects: 1) it is often the only Cw combination of a language; and 2) it tends to behave differently from other Cw combinations, i.e., as a unit segment, for which the evidence comes from both phonological patterning (e.g., phonotactic restrictions, reduplication, and vowel length alternation) and phonetic properties (auditory judgments).

Given these phonological asymmetries regarding Cw combinations, I now consider how the phonetic salience of Cw combinations is affected by phonological factors such as the place of articulation of the consonant, quality of the following vowel, and status of the /w/ component, i.e., a glide or a secondary labialization. I show that the phonological asymmetries we discussed above are closely related to the phonetic asymmetries of /w/. In other words, phonological contexts which favor Cw combinations, i.e., velar consonants and front vowels, provide stronger acoustic cues for the /w/ component, whereas labial consonants and back vowels, with which Cw combinations tend not to co-occur, render the /w/ less salient.

## 1.2 Relationship between phonetic salience of /w/ and Cw phonological asymmetries

### 1.2.1 Consonant place and vowel

Let us begin with the phonetic background for the avoidance of labial consonants in Cw combinations. Kang (1997) argues that there is a perceptual basis for the diachronic loss of /w/ after labial consonants in Korean. There is an "acoustic ambiguity" between 'labial + w + V' and 'labial + V', due to both having a rising second formant (F2) transition. This perceptual confusion leads to a sound change (Ohala 1993) from /pw/ to /p/, as the listeners may attribute the acoustic cue for /w/ to the preceding labial consonant rather than to /w/ itself.

In addition to the insufficient distinctiveness in F2 transition, the contrast between pV and pwV also lacks cues from the stop release burst, unlike the tV-twV or kV-kwV contrasts: spectra of /t/ and /k/ bursts are affected by the size of the cavity in front of the stop closure, with lip rounding favoring a lower frequency spectral energy concentration, but lip protrusion does not have such an effect on labial stop bursts, since the stop closure occurs at the lips and hence there is no front cavity.

Another possible reason for the avoidance of labial + /w/ combination is its similarity to velar + /w/ combinations in terms of formant transitions. Kawasaki's (1982) acoustic dissimilarity measurement (calculated from the formant frequency change between several points throughout the formant trajectory; see Kawasaki 1982:63-64 for the formula) revealed that labial + /w/ initial and velar + /w/ initial syllables are similar, whereas coronal + /w/ initial syllables are well distinguished from them. This leads to the prediction that the contrast between labial and velar is weak before /w/, and thus likely to be neutralized: this neutralization pattern is actually found in some dialects of peninsular Spanish (Penny 1972) and Mexican Spanish (Greenlee 1992), where /f/ is merged with /x/ before /w/ (e.g., *fwe* → *xwe* 'he went' in central Mexican Spanish (Greenlee 1992:174)).[7]

[7] The neutralization of /f/ and /x/ is not confined to the position preceding the glide /w/, but also found before back vowels /o/ and /u/ in Corrientes Spanish (Mazzaro 2005).

The analysis developed in this thesis essentially attributes the avoidance of labial + /w/ combinations in the Korean and Mandarin MSC to the perceptual similarity between plain and labialized labial consonants, rather than that between labial and velar before /w/. We may find support from historical change in Korean: /w/ is historically lost after a labial consonant, as is reflected in the change of orthography (e.g., pwe → pe 'hemp': Kang 1997:120), which suggests that it is because of their similarity to plain labials that the /pw/ combinations are avoided. Another reason to dispute the analysis based on the confusion between labial and velar, especially for Korean which lacks both labial and velar fricatives, is that in the Spanish dialects mentioned above, the labial-velar neutralization is reported only for fricatives, and not for stop consonants.

In contrast to the labial consonant context, the salience of the /w/ component is relatively stronger in the velar consonant context. The velar consonant release burst or frication has characteristically compact spectra (Blumstein and Stevens 1979, Stevens 1980), with spectral energy concentrated around a single prominent peak. The peak location of the velar stop burst is typically continuous with the F2 or F3 of the following vocoid (de Manrique and Massone 1981; Bonneau et al. 1996; Raškinis and Dereškeviciute 2007). This is because both the burst noise and F2/F3 reflect the resonance of the front cavity when the consonant is velar (Johnson 2003). Therefore, the velar consonant release burst may provide a strong cue for the lip rounding and protrusion: Bonneau's (2000) perception experiment results show that the French velar stop burst before /u/, but not the coronal and labial stop bursts in the same context, is able to lead listeners to correctly identify the following vowel as /u/.

Velar consonants are also special in that the F2 frequency at CV transition patterns differently depending on the backness of the vowel. The F2 of front vowels begins at a higher frequency after a velar than after an alveolar stop, and at the lowest frequency after a labial stop (Delattre et al. 1955; Johnson 2003:143). However, there is no constant F2 locus of back vowels after a velar consonant (Sussman et al. 1991), and it varies depending on the F2 of the vowel steady state. This is in contrast with a coronal stop, which exhibits relatively constant F2 locus regardless of the vowel context. Therefore, at least in the front vowel context, the difference in the F2 at CV transition between plain C and Cw onsets may be more prominent when C is a velar consonant, due to the tongue backing of the /w/ articulation.

The fact that some vowels tend not to co-occur with a preceding /w/ can also be attributed to their acoustic similarity. The articulation of /w/—lip rounding and tongue back raising—is associated with low F2 and F1 frequencies. Back vowels and rounded vowels are characterized by low F2, and non-low vowels by low F1. Therefore /w/ is acoustically very similar to non-low back rounded vowels like /u/ or /o/, and thus sequences like /wu/ or /wo/ are absent in many languages including Korean, due to their similarity with /u/ or /o/.

The acoustic cues for /w/ may not be very strong before the Korean back vowel /ʌ/ either, in comparison to front vowels /i/ or /e/. Even in the comparison with /a/, /ʌ/ is reported to have considerably lower F2,[8] as expected from the backer tongue position for /ʌ/ than /a/ (H.-B. Lee 1999, Kwon 2001). From this we may infer that the contrast between the Korean Cʌ and Cwʌ is perceptually weaker than the Ca-Cwa or Ce-Cwe

[8] According to Yang (1996), the average F2 frequencies of /a/ and /ʌ/ produced by male speakers are 1372 Hz and 1121 Hz respectively.

contrasts. Similarly, Mandarin Cɑ and Cwɑ may be closer to each other than Ca and Cwa are, due to the lower F2 of /ɑ/ in comparison to /a/ (Chen 2000, Liang 2002).

As both the consonant places and the vowels have influence on the salience of the C-Cw contrast, it is not surprising that languages like Korean and Mandarin would exhibit an interaction of consonant and vowel contexts in the co-occurrence restrictions of Cw combinations. Cw combinations are in general avoided in the labial place or before back vowels due to the impoverished cue for /w/ in the formant transition, but not when the consonant is a velar, as it has an effect of increasing the phonetic salience of /w/ due to the strong formant transition and consonant release cues.

### 1.2.2. Consonant place and the status of /w/

We also observed an interaction between consonant place and the status of the /w/ component in Cw combinations distribution, in languages like Dan and Tarascan (1.1.2). Recall that in these languages, it is only the velar consonant context that /w/ is realized as a secondary articulation of labialization rather than an independent glide consonant. The salience of /w/ may also be related to how it is phonetically realized, i.e., as a glide consonant or as a secondary articulation on the consonant. Kochetov and Goldstein (2006) investigated the relationship between articulatory overlap and the perceptibility of contrast between a plain C and a CG combination, though their experiments focused on the overlap between a consonant and a palatal glide /j/. In their perception experiment using synthesized stimuli, they found that the contrast between plain labial and labial + /j/ onsets was harder to perceive when [labial] and [palatal] gestures were more overlapped, i.e., when the /j/ component was realized as a secondary palatalization. Assuming that the same is true for Cw combinations with different statuses of the /w/ component, it is not surprising that there is an interaction between consonant place and the status of /w/ in the Cw combination, i.e., only velar stops, which provide inherently strong cues for the /w/ component, combine with a relatively weakened /w/ or the secondary labialization.

In sum, Cw combinations are avoided in contexts with relatively weak cues for the /w/ component (labial consonants and back/rounded vowels), and when Cw combinations do occur in an adverse condition for the salience of /w/, e.g., before back vowels (Korean and Mandarin) or as secondarily-labialized consonants (Dan and Tarascan), they are often restricted to the context with relatively strong cues for /w/, i.e., velar consonants.

This section examined the phonetic asymmetries of Cw combinations, and how they may be related to the phonological asymmetries introduced in 2.1. The salience of the /w/ component, or that of the contrast between syllables with plain C onsets and Cw combination onsets, tends to be stronger in the phonological contexts in which they frequently occur, and weakened where they are often avoided. I now turn to the theoretical background of a contrast-based analysis of Cw combinations developed in the following chapters.

## 1.3 Theoretical background

### 1.3.1 Dispersion Theory of Contrast

The Dispersion Theory of Contrast (Flemming 1995, 2002) emphasizes the role of contrast dispersion—auditory distance between contrastive strings—in the phonological patterns of languages. Analyses in Dispersion Theory (henceforth DT) focus on the evaluation of the phonological contrasts that languages employ ('systematic markedness': Ito and Mester 2003), rather than on the well-formedness ('syntagmatic markedness': *ib.*) of a word as in traditional OT phonology. In DT, phonological contrast is subject to the following three functional goals (Flemming 2002:24):

(13) a. Maximize the number of contrasts
b. Maximize the distinctiveness of contrasts
c. Minimize articulatory effort

The distinctiveness among sound strings is computed based on *auditory features* like transition F1 and F2, burst noise frequency (NF), or noise loudness (NL). Auditory features are scalar in Flemming (2002), as exemplified by the Noise Loudness scale and Transition F2 scale in (14).

(14) a. Noise Loudness scale (Flemming 2002:13)

| 5 | 4 | 3 | 2 | 1 |
|---|---|---|---|---|
| s ʃ tʃ | f θ x | ʈ c | t k | p b d g |

b. Transition F2 scale (Adapted from Flemming 1995: (11))[9]

| 5 | 4 | 3 | …… |
|---|---|---|---|
| ki | ti | pi | |

Distinctiveness between two words is expressed by the number of steps in the auditory scales by which they are separated. For example, based on the scales given in (14), the distinctiveness between /pi/ and /ki/ is expressed as (15): they differ by one step in the burst noise loudness scale, and by two steps in the transition F2.

(15) Δ(ki-pi)= Transition F2:2 and Noise Loudness:1

The balance between the conflicting functional goals in (13) is expressed in an Optimality-Theoretic model. Constraints that serve each goal interact with each other, to give rise to specific sound patterns. Since Dispersion Theory is mainly concerned with the part of grammar related to the presence (maintenance) or absence (neutralization) of contrasts between sounds or sound sequences in a language,[10] what is evaluated is not the possible output forms of a single lexical item, but the surface contrast between two or more strings. The following constraints are proposed regarding contrast dispersion in Flemming (1995, 2002). Constraints (16) and (17) take up the roles of (13b) and (13a), respectively.

(16) MINIMAL DISTANCE: F=*n*
Two contrasting strings differ by at least *n* steps in F auditory dimension.

[9] Transition F2 scale in (14b) is computed based on the binary feature system used in Flemming (1995).
[10] In DT, a total absence of a contrast and a contextual neutralization of a contrast that exist in other contexts are essentially parallel (Flemming 2002:45).

(17) MAXIMIZE THE NUMBER OF CONTRASTS (DON'T MERGE):
Allow as many contrasts as possible.

Minimal Distance constraints (16) require certain degrees of distinctiveness between sounds or sound sequences that are potentially in contrast, in terms of auditory features such as vowel formant transition or consonant release noise frequency. As shown above, the distinctiveness between segments is calculated with respect to their distance in auditory scales.

Minimal distance constraints are subject to a strict dominance relationship. A constraint that requires greater distance (e.g., *n* steps in the dimension F) is more general than one that requires smaller distance (e.g., *n-1* steps in the dimension F), in the sense that the candidate contrasts which violate the latter necessarily violate the former. Therefore, it follows from the Pāṇinian Theorem on Constraint-ranking (Prince and Smolensky 1993:81-82) that Minimal Distance constraints are in a strict dominance relationship as shown in (18):

(18) MinimalDistance=F:n-1 >> MinimalDistance=F:n >> MinimalDistance=F:n+1
(where F: auditory feature, e.g., formant transition, noise frequency)

Minimal Distance constraints inherently conflict with Maximize Contrasts (Don't Merge), as the latter prefers contrast, regardless of the size of the auditory difference between contrasting structures. Thus, depending on the position of Maximize Contrasts constraint among the ranked Minimal Distance constraint, the distance requirement of contrastive items in a language is decided. Consider the ranking in (19):

(19) MinimalDistance=F:*n-1* >> MinimalDistance=F:*n* >> Maximize Contrasts
>> MinimalDistance=F:*n+1*

The result of this ranking is that the language allows a contrast between two strings that differ by *n* steps or *n+1* steps, but strings that differ in only *n-1* steps cannot be contrastive. Given the strict dominance relationship between Minimal Distance constraints, the presence of a contrast that is distinguished by *n* difference on the F dimension in a language implies the presence of a contrast distinguished by *n+1* difference on the same dimension, other things being equal.

Though the ranking in (19) dictates that two strings differing by *n-1* steps on dimension F cannot co-exist, it does not tell us which one of the two potential oppositions surfaces as a legitimate string in the language. The effort minimization principle (13c) may play a role here. This principle disfavors extra articulator movement or extended articulatory gesture.

### 1.3.2 Dispersion-Theoretic analysis of the distribution of Cw combinations

Let us now review a DT analysis of the co-occurrence restriction on the English Cw clusters, suggested in Flemming (1995). Flemming suggests that 'labial + /w/ + vowel' sequences are avoided in English because they do not yield sufficient acoustic distinction from 'labial + vowel', in particular in terms of the F2 at the transition. He also notes that /kw/ clusters are found before the rounded vowel /ɔ/ (unlike /tw/) in English, though there is only a small formant transition difference between /ɔ/ and /wɔ/, because in the velar context "the C-Cw contrast is sufficiently distinct due to additional differences in burst quality" (Flemming 1995:177).

According to the auditory representations given in Flemming (1995:177), the degree of auditory distinctiveness between a plain stop C and Cw clusters before a vowel V (i.e., Δ(CV-CwV)) vary as a function of the consonant places and vowels, as shown in (20):

(20) a. Δ(pi-pwi) = Transition F2: 2, Noise Frequency: 0
b. Δ(pɔ-pwɔ) = Transition F2: 0, Noise Frequency: 0
c. Δ(t-tw) = Transition F2: 3, Noise Frequency: 0
d. Δ(ki-kwi) = Transition F2: 4, Noise Frequency: 1
e. Δ(kɔ-kwɔ) = Transition F2: 1, Noise Frequency: 1

(Adapted from Flemming 1995: (11))

Flemming (1995) proposes that the English CV-CwV contrast is subject to the following minimal distance requirement:

(21) Minimal Distance= F2:3 or (F2:1 and NF:1) >> Maximize Contrasts

The Minimal Distance constraint in (21) requires a certain minimum degree of transition F2 difference, schematized as 3 steps in the auditory dimension of Transition F2, between CV and CwV, or a difference on both the Transition F2 and the Noise Frequency dimensions. Thus the pi-pwi contrast does not satisfy the Minimal Distance requirement by either measure, hence the gap in /pw/ combinations (22).

(22)

| pi pwi | Minimal Distance= F2:3 or (F2:1 and NF:1) | Maximize Contrast (Don't Merge) |
|---|---|---|
| a. pi pwi | *! | |
| ☞b. pi | | * |

Flemming (2002:125) criticizes this analysis by pointing out that the F2 frequency value at the beginning of the CV transition is "typically relatively high before front vowels", even after a labial consonant. Thus, the /pi/-/pwi/ difference is no smaller than /te/-/twe/ or /ke/-/kwe/ in terms of the F2 frequency at the transition. Considering the fact that sequences like /twe/ and /kwe/ exist in English, Flemming argues that the avoidance of /pw/ cluster in English cannot be attributed to the lack of sufficient distinctiveness between pV and pwV. Instead, he proposes that the avoidance of /pw/ clusters arises from the difficulty in producing labial release burst when the lips are rounded (*ibid.*), which would make /w/ and /pw/ hard to distinguish from each other. This would make the C-Cw distinctiveness of labial and coronal places comparable, given that the Noise Frequency difference is 0 for both /p/-/pw/ and /t/-/tw/ contrasts.

So far I have reviewed the DT analyses of the distributional asymmetries of the English Cw clusters suggested in Flemming (1995) and Flemming (2002). The analysis developed in this thesis adapts the basic architecture of DT, i.e., Optimality-Theoretic interaction of minimal distance requirement and contrast preservation, and the claim that the CV-CwV contrast salience is primarily responsible for the Cw distributional asymmetries (Flemming 1995). However, the present analysis challenges some assumptions on the assessment of the perceptual distance between the CV and CwV syllables on each scale (formant transition and noise frequency). Below I specify the changes that I propose, and why they are necessary.

*Formant transition scale.* In Flemming's auditory feature system, the formant transition scale is established based on the formant frequency at the transition between a consonant and a vowel. However, the precise nature of the formant transition cues to glide consonants like /w/ or /j/ requires more consideration. In particular, I argue that dynamic aspects of the transition, e.g., the direction of the F2 change, need to be taken into account in assessing the perceptual distinctiveness. It is the shared direction (rising) of the F2 change, not just the low F2 locus, that is responsible for the perceptual similarity between labial stops and labiovelar glides, as suggested in Liberman *et al.* (1956).[11] PV syllables exhibit typically rising formant transition (Delattre *et al*. 1955), unlike tV and kV which differ in their F2 transition direction depending on the vowel F2. Thus, though the F2 at the /p/ + front vowel transition may be as high as that of /t/ or /k/ + front vowel transition as Flemming (2002) argues, the shared direction of the F2 transition may render pV and pwV perceptually more similar to each other than tV and twV or kV and kwV.

The other aspects of formant transition as a perceptual cue for a glide are the duration and the extent of frequency change. According to Schwab et al. (1981), the perception of a rising CV F2 transition as a glide /w/ is dependent on the duration and the extent of frequency increase. In other words, F2 rising over a longer duration, or over a greater frequency range, is more likely to be heard as a labiovelar glide. Thus we would expect that more overlapped realization of a Cw combination, i.e., a labialized consonant as opposed to a Cw cluster, is related to reduced perceptibility of the /w/ component.[12] If the formant transition scale can incorporate the duration and extent of F2 change, the DT analysis may be extended to patterns found in languages like Dan, Tarascan and Fox. As discussed in 1.1.2, in these languages only the velar + /w/ combination forms a single segment (labialized velar), while other Cw combinations pattern as consonant clusters. In the DT analysis, the two seemingly unrelated patterns—Korean and Mandarin on the one hand and Dan and Tarascan on the other—are both understood as an effect of contrast dispersion: because of their strong consonant release cues (noise frequency), velar consonants are able to host the C-Cw contrast in contexts where the formant transition cues are impoverished, e.g., before back or rounded vowels, and when /w/ is realized as a secondary articulation rather than a glide consonant in a consonant cluster.

---

[11] While /bɛ/ and /wɛ/ are similar in the CV transition shape, Liberman et al. (1956) report that /dɛ/ and /gɛ/ transition is heard as /ɥɛ/ and /jɛ/ respectively, when the transition duration is elongated.

[12] According to Ladefoged and Maddieson (1996), the difference between a palatalized labial and a labial + /j/ cluster in Russian is the presence of /j/ steady state after the consonant release in the latter, which affects both the duration of vocoid and the extent of frequency change. The acoustic experiment in Chapter 2 reveals that the steady state of /w/ is also a distinguishing factor between Cw clusters and labialized consonants.

*Noise frequency scale.* Labial and coronal stops may differ in the effect of /w/ on the burst noise frequency as well. Flemming's (1995) analysis summarized above supposes that only velar stop noise frequency is affected when it is followed by /w/: noise frequency difference between /t/ and /tw/ is 0 (20c). However, according to Johnson (2003), the stop burst noise frequency is lowered by lip rounding if the constriction site (e.g., stop closure) is behind the lips, as the front cavity size is inversely proportional to the noise frequency. Thus both coronal and velar stop burst noise may carry perceptual cues for the /w/ component, though the strength may vary, unlike the labial stop burst whose noise frequency is not affected by lip protrusion.

In order to establish perceptual distance scales with the changes proposed above, I examine how the cues to the CV-CwV contrast present in the consonant release and the formant transition are affected by contexts, *i.e.,* consonant place, vowel quality, and the status of /w/, with both an acoustic study and a perceptual experiment. The acoustic study (Chapter 2) investigates differences between CV and CwV syllables at various consonant place and vowel contexts in two languages—Korean and Spanish—by examining acoustic correlates such as the formants measured at the transition, F2 frequency change through the vocoid, and the stop burst noise frequencies. The perception experiment in Chapter 3 tests the effects of consonant places and vowels on the perceptibility of the CV-CwV contrast, from which the scales for formant transition and noise frequency cues are established. Based on these scales, Chapter 4 develops DT analyses for the asymmetric Cw distributions of the languages introduced in this chapter. General conclusions are given in Chapter 5.

# Chapter 2. Acoustics of Cw combinations and the C-Cw contrast

This and the following two chapters develop a Dispersion-theoretic account of distributional asymmetries in the Cw combinations of Korean and other languages. As the first part of the analysis, this chapter provides an acoustic analysis of the contrast between plain consonants and Cw combinations in syllable onset position.

The main goal of the acoustic study is to investigate the relationship between the distributional asymmetries of Cw combinations and the relative acoustic salience of the CV-CwV contrast in different consonant place and vowel contexts. The hypothesis is that the acoustic distinctiveness between plain consonant onset and Cw combination onset is relatively smaller in contexts where Cw combinations are restricted, namely in the labial consonant context and the back vowel context. In addition, the current study takes into account another aspect, the timing relationship between the consonant and /w/ components. This is because as discussed in 1.1.2, some languages exhibit an interaction between consonant place and phonetic realization of a Cw combination, realizing the universally preferred velar + /w/ as a labialized consonant, and the other Cw combinations as clusters.

For these purposes, the acoustic properties of CV and CwV syllables with different consonants and vowels are considered. Two languages—Spanish and Korean—are investigated, which differ in the phonetic status of the glide component. I examine both the vocoid (glide and vowel) portion and the consonant release portion of the syllable, using acoustic measurements such as formants at vocoid onset, F2 change throughout the vocoid, peak location frequency and intensity difference between two frequency regions in the consonant release noise spectra.

The organization of this chapter is as follows: Section 2.1 introduces the background on the acoustic cues of the contrast between plain C and Cw combination. Then the methods of the experiment (2.2) and its results are presented (2.3), followed by the discussion in 2.4.

## 2.1 Acoustic cues to the C-Cw contrast

As stated in 1.2, the present study assumes that Cw combinations primarily contrast with their plain counterpart Cs, and that the distribution of the Cw combinations is influenced by the perception of the contrast between C and Cw. For example, the frequent gap in labial consonants before /w/ is due to the fact that labial consonants and labial + /w/ are perceptually less distinct than other C-Cw pairs. Therefore I focus on the acoustic comparisons between plain CV versus CwV syllables.

We first need to establish the main acoustic cues that play a role in distinguishing Cw combinations from their plain consonant counterparts. In both types of Cw combinations, the /w/ component involves lip protrusion and tongue body raising. The acoustic correlates of these articulatory gestures are the lowering of formants. The lip protrusion extends the size of the front cavity, lowering formants—in particular F2—of the following vocoid. Tongue back raising increases the back cavity, lowering F1 (Ladefoged and Maddieson 1996, de Jong and Obeng 2000).[13] Gordon et al. (2002) report that in Montana Salish both F1 and F2 of /a/ are lower after a labialized than after a plain dorsal fricative, with F1 lowering to a lesser extent than F2. Similar results from a comparison between Ponapean labialized and plain consonants are shown in Ladefoged and Maddieson (1996:358-360).

The effect of /w/ is also to be found on the acoustic properties of the consonant release, such as stop release burst or frication noise, since the spectral shape of consonant release noise reflects the frequencies of the vocal tract resonances following the release (Stevens 1980, Johnson 2003:143-144). In his study of English stop bursts in singleton and cluster onsets, Zue (1976:111) looked at the effect of labiovelar glide on stop burst spectra in Cw clusters, and found that the mean spectral frequency of alveolar and velar stop bursts is lower in Cw clusters than in singleton onsets. The spectral center of gravity of Hupa rounded velar fricatives was found to be lower than that of plain velar fricatives by Gordon et al. (2002).

In addition to the spectral properties such as vocoid formant and noise frequencies, the contrasting CV and CwV may also differ in the duration of the vocoid (glide + vowel) portion. However, the degree of durational difference between CV and CwV may vary depending on how the /w/ component is timed with the preceding consonant, that is, whether the glide forms a cluster with the consonant or is added as a secondary articulation. Though there appear to be no acoustic studies which compare Cw clusters and labialized consonants, I assume that they are in an analogous relationship to that between Cj clusters and palatalized consonants. The Cj-$C^j$ contrast has often been investigated, owing to its contrastiveness in languages like Russian: spectrogram comparison between Russian labial + /j/ clusters and palatalized labials (Ladefoged and Maddieson 1996: 364) shows that the F2 fall from the consonant release into the vowel begins immediately after the release in /$p^j$/, whereas there is a steady state of /j/ before the transition into the following vowel in /pj/ cluster. Thus /pj/ clusters show a longer total duration of vocoid portion in the latter.

The degree of overlap between a consonant and a glide affects not only the durational contrast between CV and CGV, but also the spectral contrast. Kochetov and Goldstein (2006:18) show that the greater overlap between the primary labial and the secondary palatal articulatory gestures leads to less distinct contrast in formant transition between plain labial and labial + /j/ combination.

Given these spectral and durational cues related to the C-Cw contrast, the acoustic study presented below compares CV and CwV syllables on several acoustic measures, such as the duration of vocoid portion, the F2 frequencies at the consonant release and how they change toward the vowel peak, and the consonant release spectral properties.

[13] F1 lowering in /w/ context is not solely due to the tongue back raising, but is also caused by the lip rounding. Lip protrusion has a lowering effect on not just F2 frequency, but on all formants, since it increases the total length of the vocal tract (Johnson 2003:97).

## 2.2 Methods

2.2.1 Language choice

In order to study the effect of timing variation as well as consonant place and vowel contexts, two languages—Korean and Spanish—are chosen as the subject languages of the study. I assume that the /w/ component in Korean Cw combinations is close to the secondary labialization in their phonetic realization (Ahn 1985), unlike that in Spanish Cw combinations, whose phonetic status as a glide rather than a secondary articulation is fairly uncontroversial.

Although there is no consensus in the phonological literature with respect to whether the Korean consonant + glide combinations are to be analyzed as clusters or labialized consonants, it is a relatively well-accepted assumption in Mandarin (Duanmu 1990, Li 1999) that Cw combinations are labialized consonants. Duanmu (1990:27) argues that the Mandarin Cw and Cj combinations are consonants with secondary articulations, rather than clusters, because if /lj/ or /nw/ are clusters in Mandarin, we would expect clusters with a bigger sonority difference such as */pl/ or */kr/ to be also legitimate in Mandarin, as predicted by the Sonority Dispersion Principle (Clements 1990, Kenstowicz 1994:283-284). The same argument can be applied to Korean, since liquids and nasals combine with a glide in Korean as well as obstruents, and consonant + glide combinations are the only possible potential clusters. If Duanmu's reasoning is on the right track, my assumption that the Korean Cw combinations are not sequences of consonants is supported.

An important reason why Spanish is chosen as a representative of 'cluster-like' Cw combinations is that they are not subject to consonant place-related co-occurrence restrictions in this language. Consonants of all three major places – labial, coronal and velar – are found before [w] word-initially, unlike languages like English which has a gap in labial + /w/ clusters.

2.2.2 Recording and stimuli

Four Argentinian Spanish speakers (2 males and 2 females) and four Standard Korean speakers (2 males and 2 females), aged 24 to 35, participated in the recording. A set of bisyllabic nonsense words was created for each language, in which the target syllable was the first syllable. Korean nonsense words were all of the form C(w)V*pa.* C was one of the voiceless lenis stops /p/, /t/ and /k/, voiceless fortis stops /p'/, /t'/ and /k'/, or fricative /s'/, and V was /e/, /a/ or /ʌ/. The voiceless aspirated stop series was not included, mainly because there are no corresponding aspirated stops in Spanish. There are no labial or velar fricatives in Korean, but the coronal fricative /s'/ was included in the stimuli, to allow cross-linguistic comparison of the temporal relationship between the fricative and the /w/ component. The choice of *pa* as the second syllable made all members of the stimuli set nonsense words in Korean. Though the vowel /i/ is allowed after Cw combinations in both Korean and Spanish, it was not included in the stimuli set, in order to avoid the effect of palatalization in Korean.

Forty-two Korean non-words were used for the recording (7 consonants * 2 *w* conditions * 3 vowels). The reading list was written in Korean orthography (Hangul). Speakers produced each token in the initial position of a carrier sentence [ ___ nɨn ʌpnɨn tanʌimnida] '___ is a nonexistent word'. The transcriptions of the Korean nonsense words are given in Table 2.1.[14]

Table 2.1. Korean nonsense word stimuli. Target syllables are underlined.

| | | LENIS | | | FORTIS | | | |
|---|---|---|---|---|---|---|---|---|
| | | STOPS | | | STOPS | | | FRICATIVES |
| | | LABIAL | CORONAL | VELAR | LABIAL | CORONAL | VELAR | CORONAL |
| | e | pepa | tepa | kepa | p'epa | t'epa | k'epa | s'apa |
| CV | a | papa | tapa | kapa | p'apa | t'apa | k'apa | s'epa |
| | ʌ | pʌpa | tʌpa | kʌpa | p'ʌpa | t'ʌpa | k'ʌpa | s'ʌpa |
| | e | pwepa | twepa | kwepa | p'wepa | t'wepa | k'wepa | s'wepa |
| CwV | a | pwapa | twapa | kwapa | p'wapa | t'wapa | k'wapa | s'wapa |
| | ʌ | pwʌpa | twʌpa | kwʌpa | p'wʌpa | t'wʌpa | k'wʌpa | s'wʌpa |

Korean morphemes do not have /pw/ combinations in the initial position, so one may question the naturalness of the Korean stimuli with initial /pw/ combinations. However, recall that /pw/ combinations occur as a syllable onset at least in word internal position, as shown in 1.1.1. In the recording there was no token in which the Korean subjects deleted /w/.

The Spanish nonsense words began with one of the following 9 sounds: voiced stops /b/, /d/ and /g/, voiceless stops /p/, /t/ and /k/, and voiceless fricatives /f/, /s/ and /x/. The initial consonant was either followed by a glide /w/ and then a vowel, or directly followed by a vowel, where the vowels were either /a/ or /e/. In total there were 36 syllables (9 consonants * 2 *w* conditions * 2 vowels). The words were in the form of C(w)V*spo* so that all words in the set are nonexistent in the Spanish lexicon. In Argentinian Spanish coda /s/ is deleted or becomes a fricative without oral constriction (i.e., [h]) (Hualde 2005:31). Thus the VC formant transition at the end of the target syllable was to be affected by the labial constriction rather than the coronal constriction, yielding a pattern similar to the CV transition in the Korean stimuli. Speakers produced each token in the initial position of a carrier sentence ___ *es lo que intentaba decir* '___ is what I was trying to say'. The written forms of the Spanish nonsense words and phonetic transcriptions are provided in Table 2.2.

[14] Note that the transcription in this table does not yet specify Korean Cw combination as labialized consonants. Also, allophonic intervocalic obstruent voicing is not included.

Table 2.2. The orthographic and phonetic forms of the Spanish nonsense word stimuli. Target syllables are underlined.

| | | | VOICELESS | | | VOICED | | |
|---|---|---|---|---|---|---|---|---|
| | | | LABIAL | CORONAL | VELAR | LABIAL | CORONAL | VELAR |
| STOP | CV | e | pespo<br>[pe(h).po] | tespo<br>[te(h).po] | quespo<br>[ke(h).po] | bespo<br>[be(h).po] | despo<br>[de(h).po] | guespo<br>[ge(h).po] |
| | | a | paspo<br>[pa(h).po] | taspo<br>[ta(h).po] | caspo<br>[ka(h).po] | baspo<br>[ba(h).po] | daspo<br>[da(h).po] | gaspo<br>[ga(h).po] |
| | CwV | e | puespo<br>[pwe(h).po] | tuespo<br>[twe(h).po] | cuespo<br>[kwe(h).po] | buespo<br>[bwe(h).po] | duespo<br>[dwe(h).po] | güespo<br>[gwe(h).po] |
| | | a | puaspo<br>[pwa(h).po] | tuaspo<br>[twa(h).po] | cuaspo<br>[kwa(h).po] | buaspo<br>[bwa(h).po] | duaspo<br>[dwa(h).po] | guaspo<br>[gwa(h).po] |
| FRICATIVE | CV | e | fespo<br>[fe(h).po] | sespo<br>[se(h).po] | jespo<br>[xe(h).po] | | | |
| | | a | faspo<br>[fa(h).po] | saspo<br>[sa(h).po] | jaspo<br>[xa(h).po] | | | |
| | CwV | e | fuespo<br>[fwe(h).po] | suespo<br>[swe(h).po] | juespo<br>[xwe(h).po] | | | |
| | | a | fuaspo<br>[fwa(h).po] | suaspo<br>[swa(h).po] | juaspo<br>[xwa(h).po] | | | |

Speakers repeated the list of sentences four times. Recording was done in a sound treated room, using Marantz digital recorder PMD 660 and Shure SM 48 microphone, at sampling rate 44100Hz.

### 2.2.3 Analysis

Praat version 4.4.16 (Boersma and Weenink 2006) was used for all analyses. F1 and F2 at the vocoid onset were measured. The term *vocoid onset* is used instead of vowel onset or glide onset since the initial consonants in target syllables are followed by either a glide /w/ or a vowel, depending on the syllable types. The second formant was also measured at a later point in the vowel to assess the degree and direction of F2 change (See 2.3.2.3 for details). The vocoid onset formants were measured by aligning the left edge of a 25 ms window to the beginning of two or more formants. F2 at the later point was measured from 25 ms window centered at the relevant time point. Maximum five formants were extracted from 0 to 5500 Hz for female speech, and from 0 to 5000 Hz for male speech. Formant analysis did not include the Spanish stimuli beginning with fricatives, to make the conditions parallel for the comparison between languages. Formants were primarily measured by a Praat script and checked by eye.

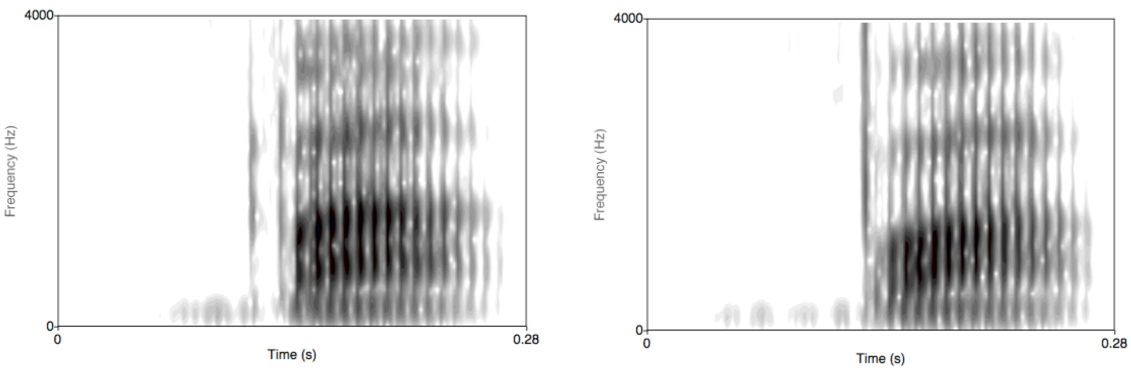
In addition to the vocoid formant cues, consonant release spectral cues were also measured. Stop release burst noise spectra were computed from a 12 ms Hanning window centered at the onset of the stop burst. Prior to the FFT the signal was downsampled to 22050 Hz to exclude higher frequencies that may not play an important role in speech sound perception. Fricative spectra were generated from a 30 ms Hanning window centered at three different places: the windows were centered at 1/4, 1/2 and 3/4 points of the frication duration.

### 2.3 Results

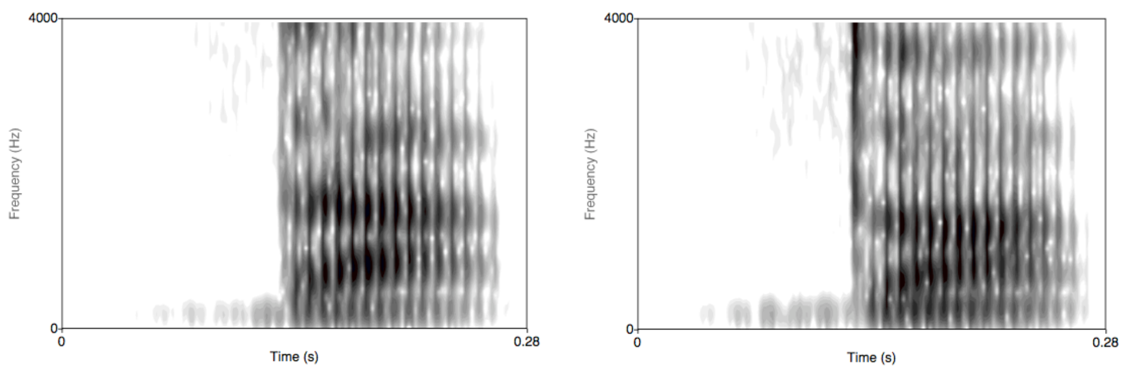
Figure 2.1 below provides sample spectrograms of Korean and Spanish Ca and Cwa syllables.

Figure 2.1. Sample spectrograms of the Korean and the Spanish Ca and Cwa syllables.

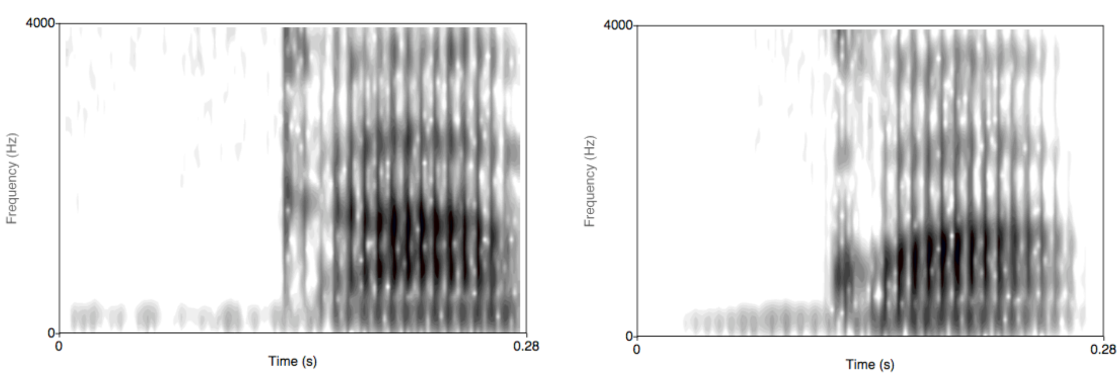
a. Korean *p'a* (L) and *p'wa* (R)



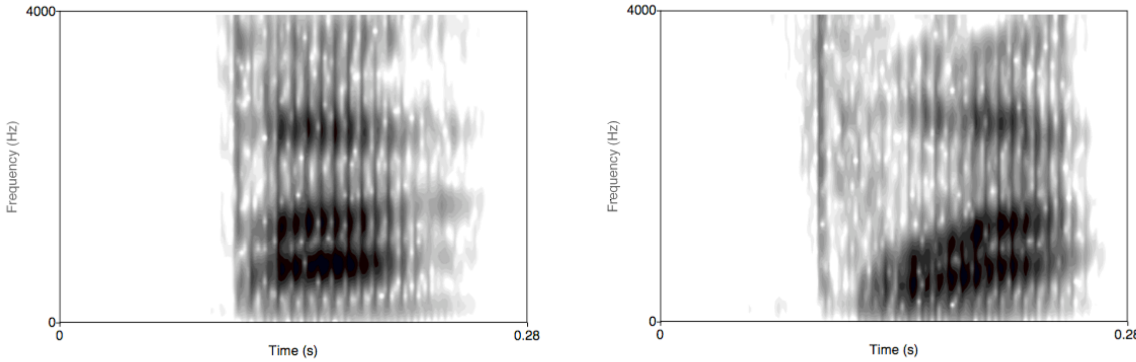

b. Korean *t'a* (L) and *t'wa* (R)



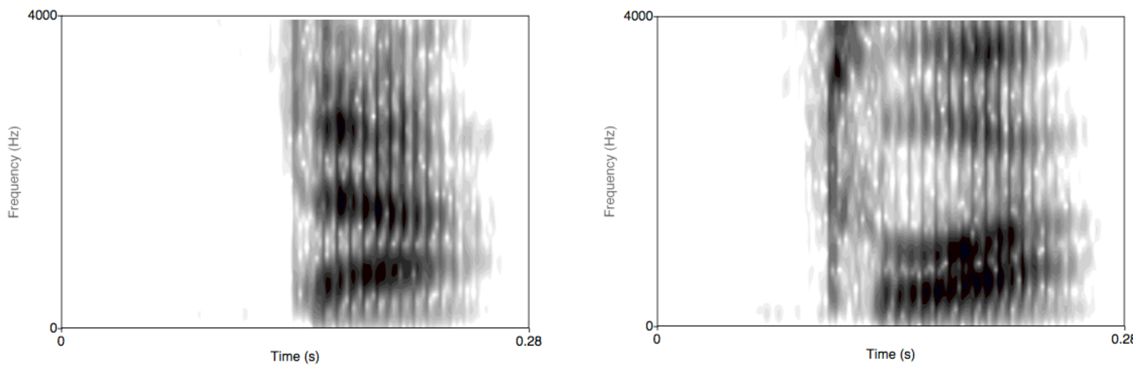

c. Korean *k'a* (L) and *k'wa* (R)



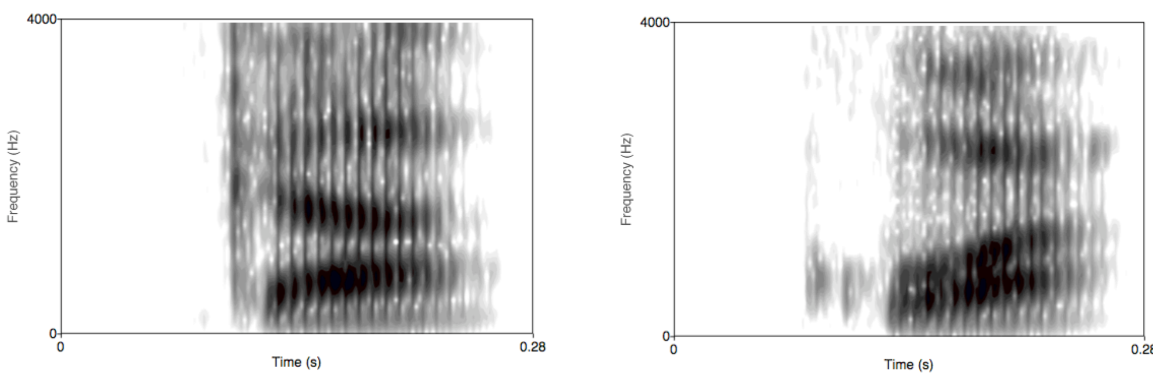
d. Spanish *pa* (L) and *pwa* (R)




e. Spanish *ta* (L) and *twa* (R)




f. Spanish *ka* (L) and *kwa* (R)




The spectrograms above provide an impression of how the members of each CV-CwV pair differ from each other, and which pair exhibits a greater effect of /w/. For example, the Korean CV and CwV in (a-c) look less distinctive from each other than their Spanish counterparts (d-f), in terms of the overall shape of formant trajectories. Also, in the velar context (c and f), the direction of F2 change clearly distinguishes between CV and CwV syllables, falling in the former and rising in the latter. The results of each acoustic measurement are presented below: Duration of vocoid portion (2.3.1), F1 (2.3.2.1) and F2 (2.3.2.2) at vocoid onset, F2 change (2.3.2.3), spectral peak location frequency (2.3.3.1) and intensity difference between mid- and high-frequency regions (2.3.3.2) of stop burst noise, and fricative noise center of gravity (2.3.3.3).

For each acoustic measure, the contextual effects on the CV-CwV difference — the consonant place and vowel – are discussed for each language, followed by the cross-linguistic comparison. Repeated measures and mixed model Analysis of Variance (ANOVA) tests were carried out on the results of each measurement, with within-subject factors *w, phonation type, place* and *vowel*, and a between-subject factor *language*. In order to focus on the effect of each linguistic factor (place, vowel and language) on the degree of the CV-CwV differences, only the main effect of the factor *w* and the interactions between *w* and other independent variables are reported. Throughout this chapter, P, T and K will be used to represent bilabial, coronal and velar stops, regardless of the laryngeal features (e.g., voiced, voiceless, lenis, fortis).

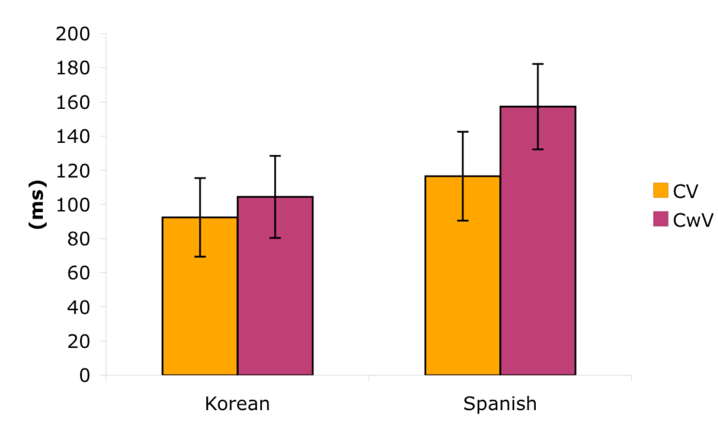
### 2.3.1 Duration of vocoid

The duration from the beginning of the voicing to the end of the vowel was measured. CwV syllables are expected to have a longer vocoid duration than their CV counterparts, especially when the /w/ component forms a consonant cluster with the preceding C. In contrast, if the glide is realized as secondary articulations on a consonant, the contribution of the /w/ component to the total vocoid duration may be smaller.

*Korean.* In a repeated measures ANOVA with between subject factors *w, phonation type, place* and *vowel*, the main effect of *w* did not reach significance ($F(1,3)=5.281$, $p=0.105$). The interactions between *w* and other factors such as *phonation type*, *place* and *vowel* were not significant either, which means that the vocoid duration of each CV-CwV pair was comparable, regardless of the phonation type, consonant place and vowel quality.

*Spanish.* Spanish showed a different pattern from Korean: There was a main effect of *w* ($F(1,2)=57.114$, $p<0.017$), showing that the addition of /w/ significantly increases the vocoid duration, unlike in Korean. However, no interactions between factors involving *w* were found significant in Spanish just as in Korean, which means that the duration increase in CwV syllables was constant in all consonant and vowel contexts. Figure 2.2 shows the average vocoid duration of CV and CwV syllables in Korean and Spanish, across all consonants and vowels.

Figure 2.2. Mean duration of vocoids of the Korean and the Spanish CV and CwV syllables. Error bars indicate standard deviation.



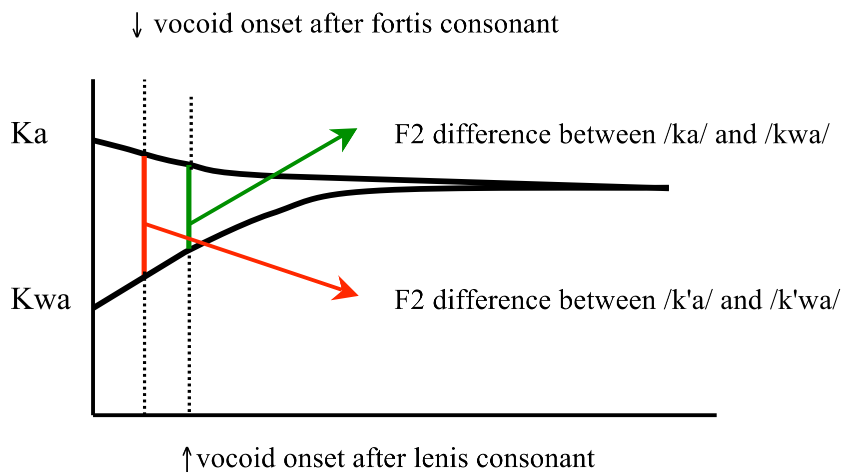

*Language comparison.* For the comparison between the two languages, only the most comparable groups of stimuli were selected from each language: The Korean vowel /ʌ/ context was excluded since there was no equivalent in Spanish. Further selection was made according to the consonant voicing type: only the Korean fortis and the Spanish voiceless contexts were chosen, since they have comparable voice onset times (VOT) according to the literature consulted (Korean /p'/=7 ms, /t'/=11 ms, /k'/=19 ms, Lisker and Abramson 1964; Spanish /p/=6.5 ms, /t/=10.4 ms, /k/=25.7 ms, Castañeda 1986). A mixed model ANOVA was carried out on these selected data, with a between-subject factor *language* and a within-subject factor *w*, and the results showed a significant interaction of *w* and *language* ($F(1,5)=9.062$, $p<0.03$). In other words, the durational difference between CV and CwV is significantly bigger in Spanish than in Korean.

In sum, the degree of the CV-CwV contrast in duration varied significantly between languages, but no language-internal factors (*place* or *vowel*) had any effect on it. Thus vocoid duration difference between CV and CwV may be an indicator of the contrast types (secondary articulation or clustering), but it does not sort out particular consonant places or vowels as providers of a greater CV-CwV contrast.

### 2.3.2 Formants

In the ANOVA test of the formant measures, *phonation type* was included as a within-subject factor. Korean lenis stops are slightly aspirated (e.g., VOT /p/=18 ms, /t/=25 ms, /k/=47 ms in Lisker and Abramson 1964), so a considerable portion of the vocoid is devoiced. Since I measured the formants at the onset of the voicing, the formant information at the release of the consonant, which contains an important cue for the consonant place and secondary articulation, may have been missed because of the considerable duration of aspiration after a lenis stop release. This discrepancy between consonant release and voice onset is expected to be weaker in the Korean fortis stop contexts since VOT is shorter. The stylized F2 trajectories of the Korean Ka and Kwa (Figure 2.3) show how the different VOT of the fortis and the lenis stops may lead to difference in the degree of the CV-CwV contrast. The difference between CV and CwV may be under-represented when the consonant has longer VOT.

Figure 2.3. Stylized F2 trajectories of Ka and Kwa, and the estimation of F2 difference between them in lenis (green) and fortis (red) contexts.

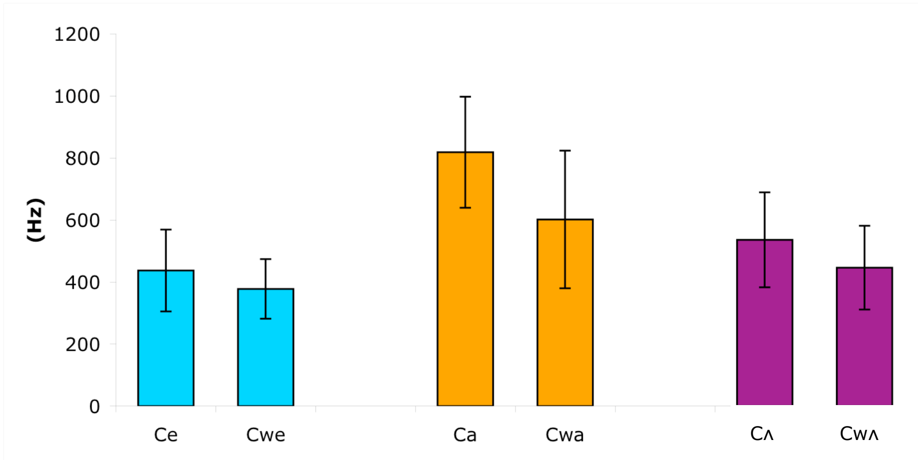



As discussed in 2.1, the lip rounding and tongue backing of the /w/ component have a frequency lowering effect on the first and second formants at the transition. Therefore the formants at vocoid onset, as well as the formant trajectory toward the vowel peak, are expected to differ between CV and CwV syllables (where C and V are identical). The results of the measurement match this prediction, as the main effect of the factor *w* was always significant. The interaction of *w* and other factors were also observed in some measures, as reported below.

2.3.2.1 F1 at vocoid onset

*Korean.* Figure 2.4 summarizes the F1 at vocoid onset data for Korean. As expected, F1 frequency was lower in CwV syllables than in CV syllables. In repeated measures ANOVA (within subject factors *w, place, vowel* and *phonation*), the main effect of *w* on F1 was found ($F(1,3)=41.568$, $p<0.008$). Among interactions involving the factor *w*, only the *w* * *vowel* interaction was at least marginally significant ($F(2,6)=4.898$, $p=0.055$). Pair-wise comparisons showed that the difference between /e/ and /a/ contexts ($p<0.043$) was significant, which suggests that when the vowel was /a/, the CV-CwV contrast was bigger than in /e/ vowel context. No significant interactions of *w* * *place* or *w* * *phonation* was found, which means that the CV-CwV contrast in F1 frequency is not influenced by consonant places or voicing types.

Figure 2.4. Mean F1 (±standard deviation) at vocoid onset of the Korean CV and CwV syllables.




*Spanish.* As in Korean, F1 was lowered by /w/ to a greater extent when the vowel was /a/ than when it was /e/, as we can see from Figures 2.5 and 2.6. Repeated measures ANOVA (*w* * *place* * *vowel* * *phonation*) results revealed that there was a main effect of *w* on F1 ($F(1,3)=28.423$, $p<0.013$), and the interaction of *w* and *vowel* was significant ($F(1,3)=74.506$, $p<0.003$), suggesting that the Ca-Cwa difference is greater than the Ce-Cwe difference. Unlike Korean, Spanish exhibited a consonant place influence on the degree by which CV and CwV differ, as shown by the significant interaction between *w* and *place* ($F(2,6)=47.384$, $p<0.001$). According to the subsequent pair-wise comparisons, the CV-CwV difference was distinguished between K and P ($p<0.003$) and K and T contexts ($p<0.005$). This suggests that the KV-KwV difference in F1 was smaller than the PV-PwV and TV-TwV differences in Spanish.

Figure 2.5. Mean F1 (±standard deviation) at vocoid onset of the Spanish Ce and Cwe syllables.

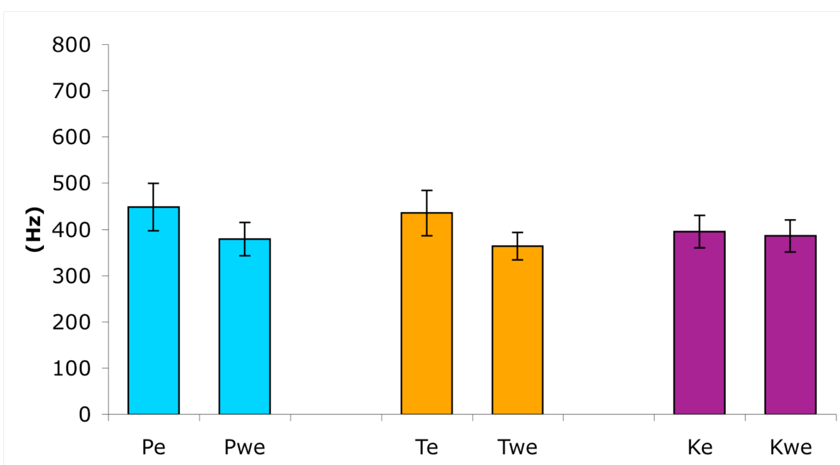
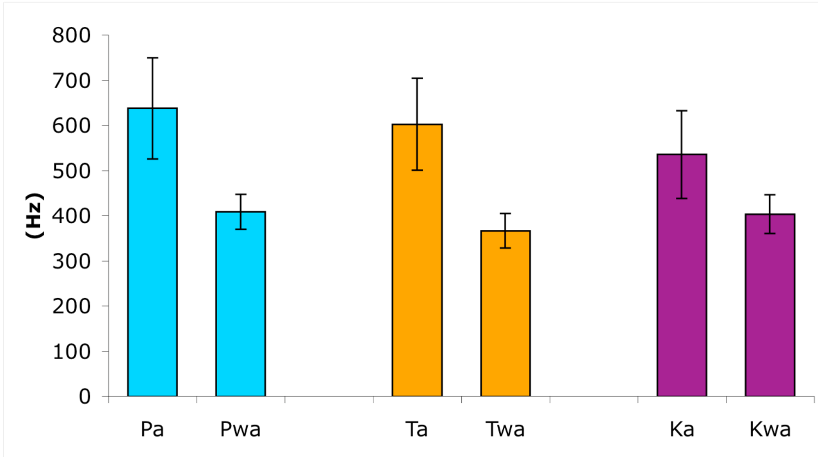


Figure 2.6. Mean F1 (±standard deviation) at vocoid onset of the Spanish Ca and Cwa syllables.




The results of F1 at vocoid onset measurement highlight the vowel height effect on the CV-CwV difference. A low vowel /a/ led to a greater degree of the CV-CwV difference than a mid vowel /e/ in both Korean and Spanish. The advantage in the degree of F1 difference may be responsible for the preference of low vowels after labialized consonants, reported in Kawasaki (1982) and summarized in 1.1.1.

*Language comparison.* As in the duration measurement, only the most comparable contexts between the two languages were selected for the language comparison: the Korean fortis and the Spanish voiceless consonants, and the vowels /a/ and /e/. The interaction between *w* and *language* was not significant ($F(1,6)=0.256$, $p=0.631$), which showed that the two languages did not differ in the degree of CV-CwV F1 contrast.

2.3.2.2 F2 at vocoid onset

*Korean.* F2 at vocoid onset of Korean CV and CwV is summarized in Figures 2.7, 2.8 and 2.9. As expected, the F2 at vocoid onset was significantly lower in CwV than in CV, as indicated by the main effect of *w* ($F(1,3)=180.562$, $p<0.001$). Interactions between *w* and *place* ($F(2,6)=6.79$, $p<0.029$) and *w* and *vowel* ($F(2,6)=25.957$, $p<0.011$) were significant, suggesting that the contrast was bigger in some consonant place and vowel contexts than others. *Phonation* did not have an influence on the overall size of the CV-CwV contrast, as shown by the insignificant interaction between *w* and *phonation*, though the three way interaction *w* * *place* * *phonation* was significant ($F(2,6)=5.607$, $p<0.042$). This was because the place effect (indicated by *w* * *place* interaction) was significant within the fortis context ($F(2,6)=12.803$, $p<0.008$), but not in the lenis context. As mentioned above, fortis is a better environment for capturing the CV-CwV difference at the formant onset, so the place effect was investigated within the fortis context. The *place* * *w* interaction was significant in the comparisons of the k'V, k'wV, p'V and p'wV syllables ($p<0.024$) and of the k'V, k'wV, t'V and t'wV syllables ($p<0.009$). This suggests that the k'V-k'wV contrast was greater than the p'V-p'wV and t'V-t'wV contrasts, considering that the mean F2 value difference was greater between k'V and k'wV syllables than between other CV and CwV syllables (Figures 2.7-2.9).

Among the three vowels, /ʌ/ provided the smallest CV-CwV contrast, as we can see from Figure 2.9 in comparison to Figures 2.7 and 2.8. Pair-wise comparisons revealed that the Cʌ-Cwʌ difference was significantly different from Ce-Cwe ($p<0.015$) and Ca-Cwa ($p<0.011$). The Ca-Cwa and Ce-Cwe contrasts were not distinguished.

Lastly, the three-way interaction *w* * *place* * *vowel* was significant ($F(4,12)=5.076$, $p<0.013$). This may be because the advantage of velars in the CV-CwV contrast was greater in the vowel /e/ context, as we can see by comparing Figure 2.7 with Figures 2.8 and 2.9. It was also the case that the k'e-k'we difference was considerably bigger than the k'a-k'wa difference, whereas in P and T contexts the Ce-Cwe and Ca-Cwa contrasts were comparable. This seems to be related to the inherently higher locus (i.e., origin of formant) of velar stops before front vowels (e.g., Sussman et al. 1991): Notice from Figures 2.7-2.9 that F2 at onset is considerably higher for /k'e/ than for /t'e/, whereas /t'a/ and /t'ʌ/ have higher F2 at onset than /k'a/ and /k'ʌ/, respectively. Because of the high F2 onset of front vowels after a velar consonant, the contrast between /k'e/ and /k'we/ was bigger than other CV-CwV contrasts.

Figure 2.7. Mean F2 (±standard deviation) at vocoid onset of the Korean fortis-initial Ce and Cwe syllables.

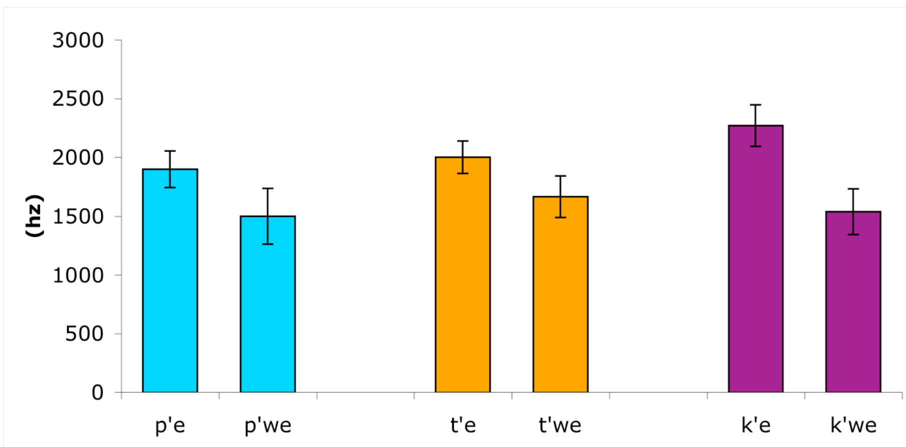



Figure 2.8. Mean F2 (±standard deviation) at vocoid onset of the Korean fortis-initial Ca and Cwa syllables.

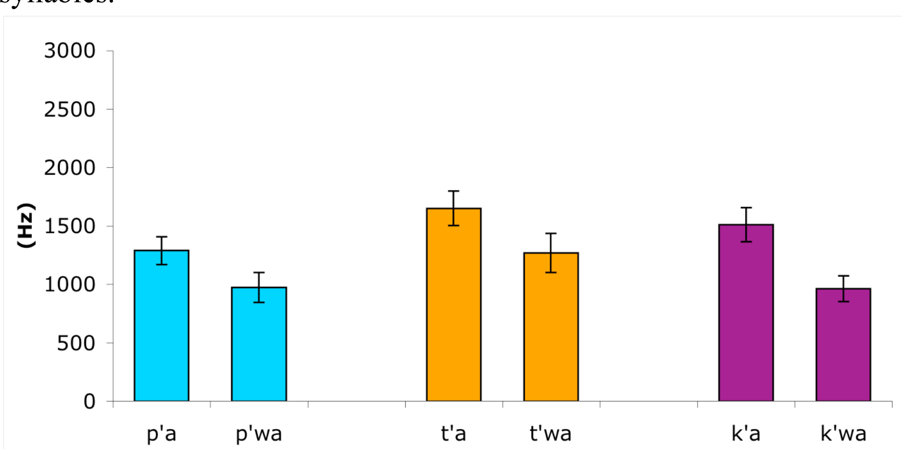
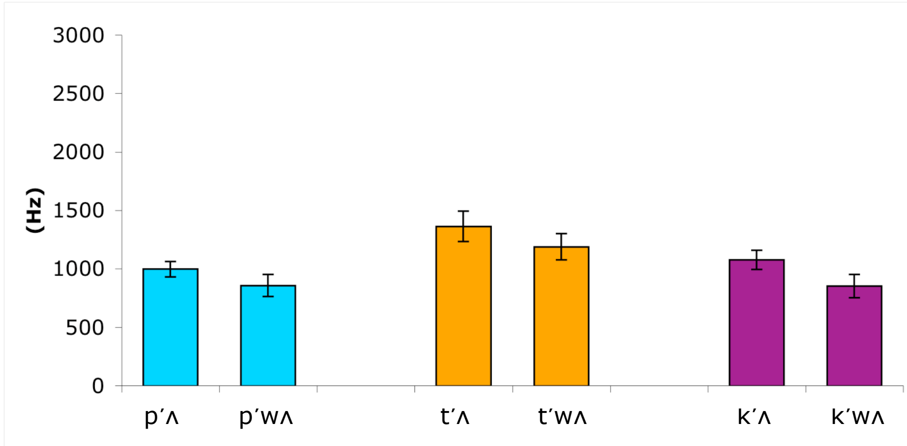


Figure 2.9. Mean F2 (±standard deviation) at vocoid onset of the Korean fortis-initial Cʌ and Cwʌ syllables.




*Spanish.* Figures 2.10 and 2.11 show the average F2 at vocoid onset of the Spanish CV and CwV syllables. As expected, a main effect of *w* was found (F(1,3)=1043.614, p<0.001). The interaction of *w* * *place* was also significant (F(2,6)=178.658, p<0.001), showing that the CV-CwV contrast differs in degree across consonant place. Post-hoc tests revealed that the comparison between K and P, and K and T contexts were significant (p<0.001), just as in Korean, which suggests that the CV-CwV difference was greater when C was velar than when it was labial or coronal. Between the two vowels /a/ and /e/, /e/ showed a significantly bigger CV-CwV difference, as revealed by the significant interaction of *w* * *vowel* (F(1,3)=4380.302, p<0.001). The interaction *w* * *place* * *vowel* (F(2,6)=28.25, p<0.001) was also significant. This was mainly because the Pe-Pwe contrast was bigger than Te-Twe, whereas Pa-Pwa and Ta-Twa were comparable, as shown in Figures 2.10 and 2.11. The advantage of the velar context in the size of the CV-CwV contrast was constant across the two vowel contexts.

Figure 2.10. Mean F2 (±standard deviation) at vocoid onset of the Spanish Ce and Cwe syllables.

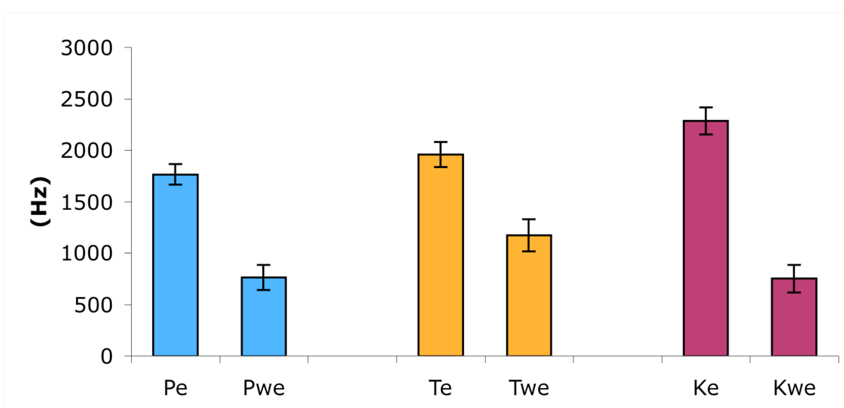
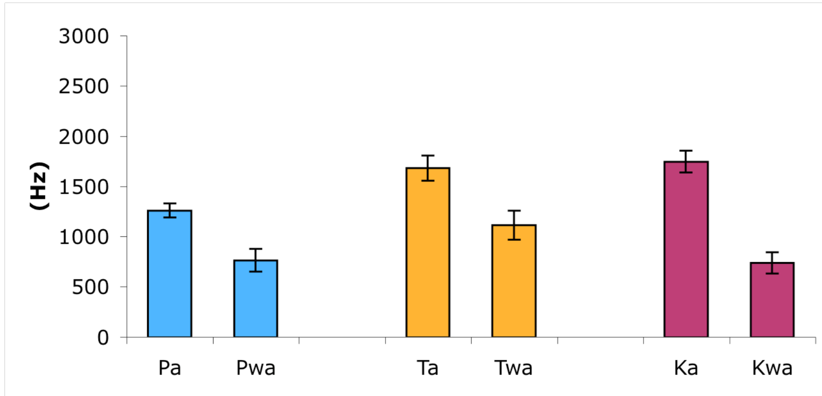


Figure 2.11. Mean F2 (±standard deviation) at vocoid onset of the Spanish Ca and Cwa syllables.

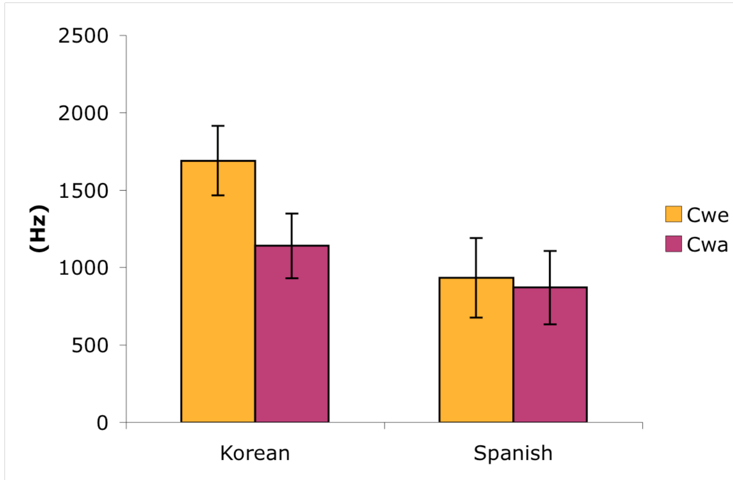



To summarize the results of the F2 at vocoid onset measure so far, the advantage of velars over other consonant places in the acoustic salience of the contrast was found for not only the Korean labialization contrast, but also the Spanish CV-CwV contrast. Among vowels it was the backer vowel /ʌ/ that conditioned smaller CV-CwV difference than /e/ and /a/ in Korean. Thus, the prediction on the relationship between phonological asymmetries of labialized consonants and the phonetic salience of the CV-CwV contrast discussed in 1.2 is supported by the results of the F2 at vocoid onset measurement.

*Language comparison.* In the comparison of the CV-CwV difference between languages, Spanish showed a bigger effect of *w* on F2 at vocoid onset than Korean: In a mixed model ANOVA conducted on the Korean fortis and the Spanish voiceless /a/ and /e/ vowel stimuli, the interaction of *language* and *w* was significant ($F(1,6)=61.618$, $p<0.02$).

From the language comparison we can also see that the Korean /w/ does not exhibit a constant target formant frequency in CwV syllables. It is clear from Figure 2.12 that the F2 at vocoid onset of the Korean CwV syllables is related to the following vowel's F2. Spanish, in contrast, shows rather constant F2 at onset between Cwa and Cwe syllables, indicating that the /w/ component is realized as a glide in a sequential relationship with the C. This is an expected difference between the two languages under the hypothesis that Cw combinations are more cluster-like in Spanish than in Korean. The results of a repeated measures ANOVA carried out only on the Cwa and Cwe syllables of each language support this observation. The effect of *vowel* on the F2 at onset of CwV syllables was significant in Korean ($F(1,3)=614.012$, $p<0.001$), but not in Spanish.

Figure 2.12. Mean F2 (±standard deviation) at vocoid onset of the Korean and the Spanish Cwa and Cwe syllables.




2.3.2.3 F2 change

In addition to the F2 frequency at vocoid onset, the frequency change from the onset to a steady state of F2 in the later part of the syllable is also calculated. This measure was included to capture not only the frequency range covered by the F2 trajectory, but also the direction of F2 change (i.e., rising, falling or level transition). The direction of F2 change is a crucial cue for the perception of glides, as well as for the perception of the consonant places. Rising CV F2 transition is associated with both labiovelar glide /w/ and labial consonants, whereas velar consonants and palatal glide /j/ are both characterized by falling F2 (Liberman et al. 1956, Schwab et al. 1981). Thus, shared F2 change direction of CV and CwV syllables (e.g., both rising) may be interpreted as a smaller phonetic difference between them, compared to divergent trajectories of CV and CwV (e.g., CV falling and CwV rising).

One Spanish speaker (female) pronounced the written *s* at the end of the target syllable as /s/ in a majority of the tokens, whereas in other speakers' tokens it was pronounced as /h/ or dropped, which is the standard pronunciation of Argentinian Spanish. This speaker's data were excluded from F2 change measure, because the coronal articulation of the coda consonant may have affected the formant trajectory throughout the vocoid.

In order to represent both the direction and the degree of F2 change, F2 change was computed as shown in (1). F2 ($t_i$) represents frequency of F2 at time point $t_i$.

(1) F2 change

a. CV syllables:
F2 change = F2 ($t_2$) – F2 ($t_1$)
where $t_1$ is the vocoid onset, and $t_2$ is the vowel midpoint.

b. CwV syllables:
F2 change = F2 ($t_2$) – F2 ($t_1$)
where $t_i \in$ {F2 maximum time point, F2 minimum time point}
and $t_1$ precedes $t_2$.

Let us describe how the F2 of the vocoid in each syllable changes over the vowel duration by looking at sample spectrograms. As shown in Figure 2.1, the F2 change was unidirectional in plain CV syllables, rising or falling toward the steady state of the vowel. Among CwV syllables, F2 was increasing from the vocoid onset onward when the consonant was P or K in both Korean and Spanish, whereas F2 change patterns in TwV differed between languages. In Spanish, F2 fell from the consonant release, and rose up to the following vowel peak (Figure 2.13). Korean TwV syllables did not show a clear F2 contour in majority of tokens (Figure 2.14).

Figure 2.13. Sample spectrogram of Spanish *twe*.

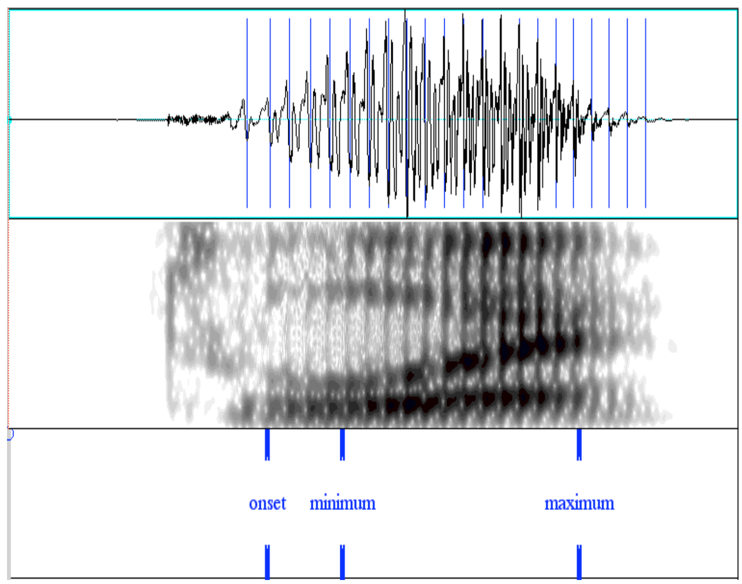
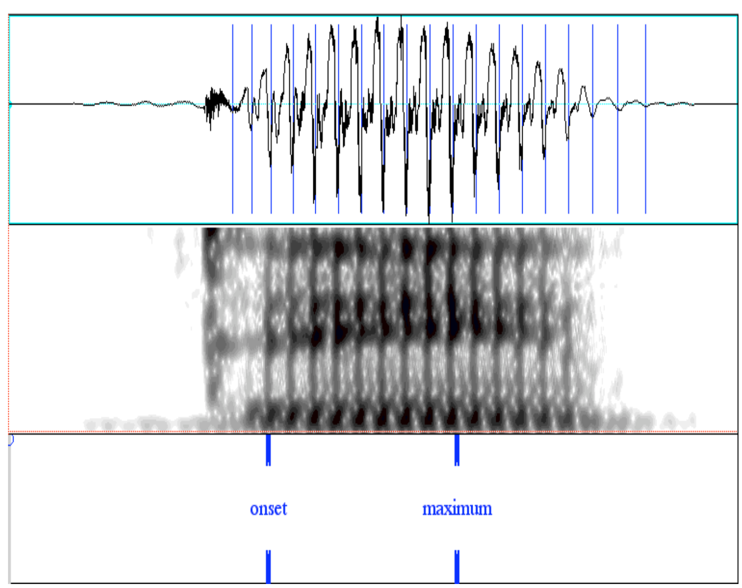



Figure 2.14. Sample spectrogram of Korean *t'we*.



The way the F2 change was calculated for CwV syllables (1b) ensured that when the F2 trajectory was U-shaped (falling and rising, as in Figure 2.13), the part of trajectory which travels across a bigger range of F2 (from minimum to maximum) was chosen.

*Korean.* Mean F2 frequency change and standard deviation of the Korean CV and CwV syllables are shown in Figures 2.15, 2.16 and 2.17. The main effect of *w* was found ($F(1,3)=58.728$, $p<0.005$), showing that in CwV syllables the F2 frequency changes more drastically than in CV syllables. The interactions *w * place* ($F(2,6)=6.091$, $p<0.036$) was significant, and the three-way interaction *w * place * phonation* was also significant ($F(2,6)=8.722$, $p<0.017$), suggesting that the CV-CwV contrast was influenced by the consonant place only in the fortis consonant context. The *w * place* interaction was indeed significant in the fortis context ($F(2,6)=16.367$, $p<0.004$), but not in the lenis context ($p=0.646$). Subsequent pair-wise comparisons revealed that the *w * place* interaction was significant between /k'/ and /t'/-initial syllables, and between /k'/ and /p'/-initial syllables ($p<0.029$), suggesting that the k'V-k'wV difference was significantly greater than the t'V-t'wV or p'V-p'wV difference (See Figures 2.15-17).

In addition to the F2 frequency change, the direction of F2 transition, roughly represented by the signs (+/-) of the mean F2 change value, also suggested that k'V and k'wV were more distinctive than other CV-CwV pairs: F2 of k'V decreased throughout the syllable regardless of the vowel (as shown by the negative mean F2 change), whereas the k'wV F2 change was always positive. This differentiated velar from labial and coronal, for which the CV and CwV syllables sometimes exhibited the same direction of F2 change. F2 trajectories of p'V and p'wV were similar in that they were both rising, except for the /ʌ/ vowel context: F2 change was close to 0 in Pʌ, indicating rather level F2 transition. Both TV and TwV exhibited clearly falling transition when V was /ʌ/, though in /a/ and /e/ vowel contexts they seem to diverge in the F2 transition direction (level for /t'e/ and /t'wa/, rising for /t'we/ and falling for /t'a/).

Figure 2.15. Mean F2 change in frequency (±standard deviation), from the onset to a later point in the Korean fortis-initial Ce and Cwe syllables.

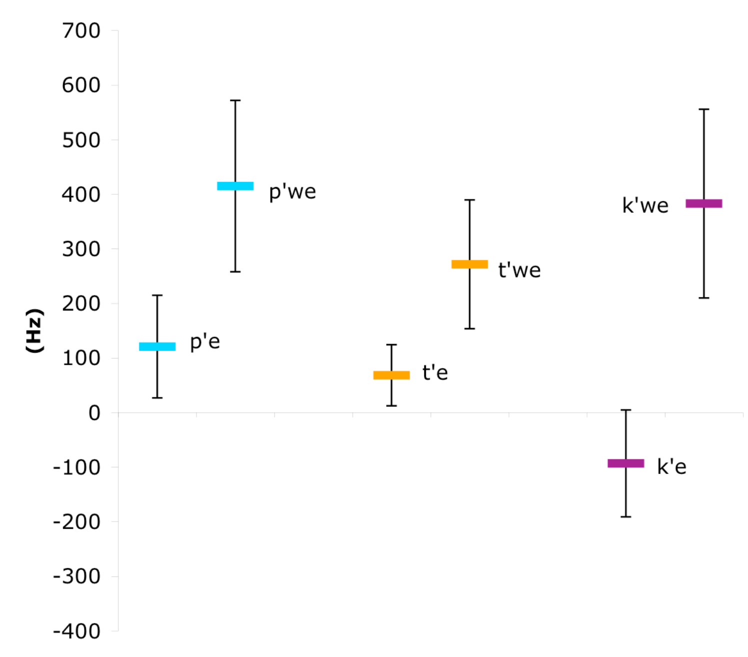



Figure 2.16. Mean F2 change in frequency (±standard deviation), from the onset to a later point in the Korean fortis-initial Ca and Cwa syllables.

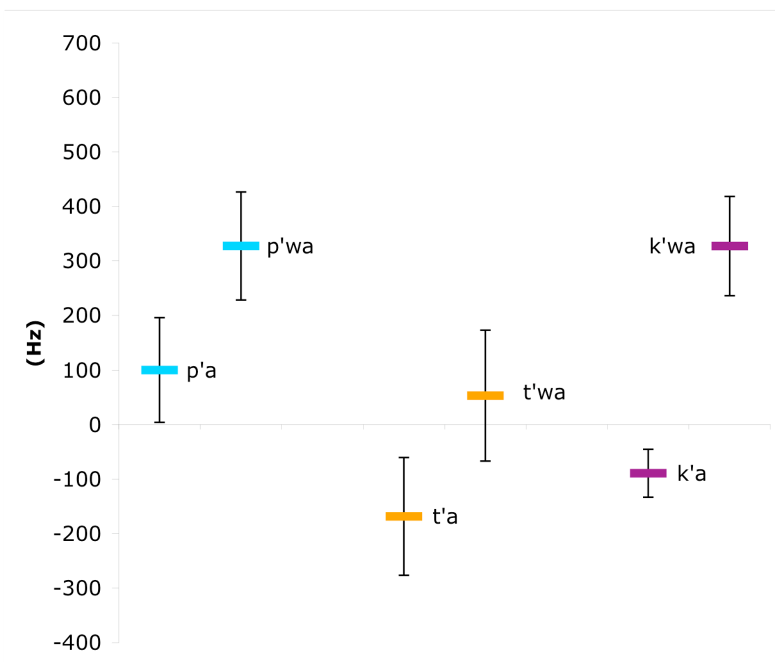
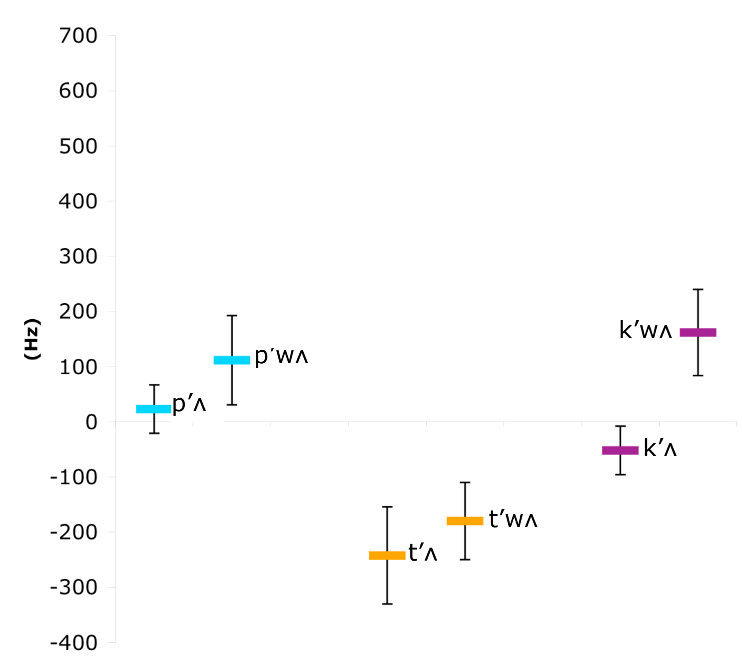


Figure 2.17. Mean F2 change in frequency (±standard deviation), from the onset to a later point in the Korean fortis-initial Cʌ and Cwʌ syllables.




The size of the CV-CwV F2 change difference was affected by vowels as well as consonant places, as shown by the significant interaction *w* * *vowel* (F(2,6)=33.328, p<0.001). Pair-wise comparisons revealed that this is due to the difference between /ʌ/ and other vowel contexts, just as in the F2 at onset measure: the Cʌ-Cwʌ contrast in F2 change was differentiated from the Ce-Cwe (p<0.006) and Ca-Cwa (p<0.005) contrasts, but the difference between Ce-Cwe and Ca-Cwa was not significant (p=0.381). Thus, the CV-CwV difference was significantly smaller in the /ʌ/ vowel context than in others, as we may see from the mean F2 changes in the Figures above. It is in fact expected that F2 at onset and F2 frequency change measures pattern together in terms of place and vowel effects. Assuming that the vowel peak is rather constant between CV and CwV, their difference at onset is bound to be reflected in the F2 change between the vocoid onset and the vowel peak.

Both the frequency difference between two points in a syllable and the direction of F2 change indicate that the velar consonant context is more advantageous than the others in terms of the CV-CwV contrast salience in Korean, again showing that the preferred consonant for contrastive labialization (i.e., velar) is also the one which provides a greater phonetic cue (in this case the F2 change) than other consonants for the CV-CwV contrast.

*Spanish.* In the results of the repeated measures ANOVA on Spanish F2 change, the main effect of *w* (F(1,2)=9769.975, p<0.001), and significant interaction of *w* * *place* (F(2,4)=20.481, p<0.008) were found. The pair-wise comparisons of consonant places (P-T, T-K, and P-K) revealed that T-K (p<0.046) and P-K (p<0.011) comparisons were significant. The K context provided a bigger CV-CwV contrast than P and T, just as in the Korean fortis context, as suggested by the mean F2 change of Spanish CV and CwV syllables in Figures 2.18 and 2.19. The interaction of *w * vowel* was marginally significant (F(1,2)=17.801, p=0.052); The Ce-Cwe difference was bigger than the Ca-Cwa difference.

Figure 2.18. Mean F2 change in frequency (±standard deviation), from the onset to a later point in the Spanish Ce and Cwe syllables.

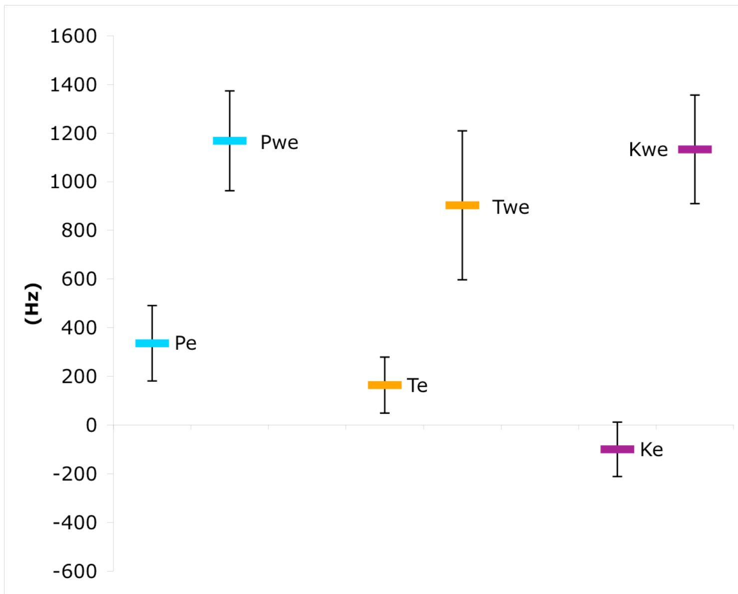
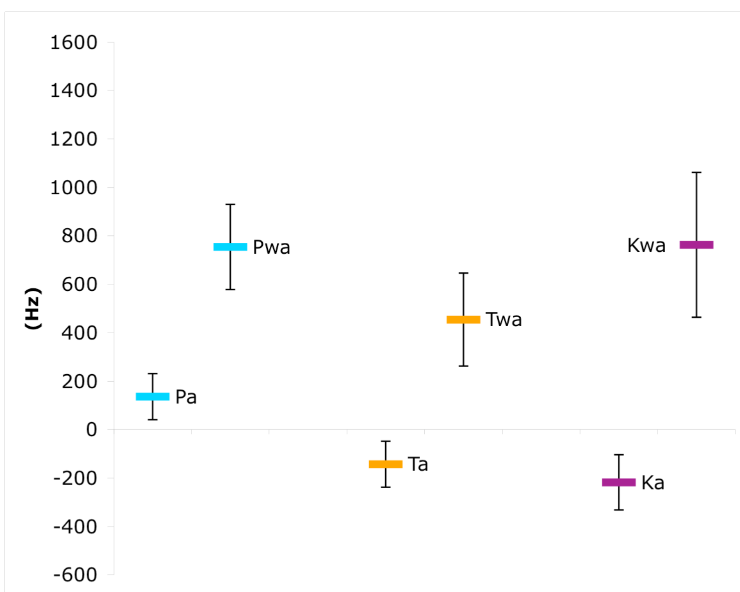



Figure 2.19. Mean F2 change in frequency (±standard deviation), from the onset to a later point in the Spanish Ca and Cwa syllables.



As in Korean, the sign of F2 change was differentiated between KV and KwV (negative for KV and positive for KwV) in both vowel contexts, unlike PV and PwV which share the rising F2 transition. As mentioned above, Spanish Twe and Twa F2 trajectories consisted of both falling and rising components in many tokens, and thus were distinguished from Te and Ta, which exhibited unidirectional formant change toward the vowel peak.

*Language comparison.* In a mixed model ANOVA with a within-subject factor *w* and a between-subject factor *language*, the interaction of *w* * *language* was significant ($F(1,5)=149.748$, $p<0.001$). This suggests that the Spanish CV-CwV difference was greater than the Korean counterpart.

### 2.3.3 Consonant release noise spectral properties

In addition to the formant frequency lowering, labialization or a prevocalic glide /w/ has an effect of lowering the frequency of energy concentration in noise spectra for consonants whose constriction is behind the lips (Zue 1976, Bonneau *et al.* 1996: 558, Gordon *et al.* 2002), as it increases the cavity in front of the consonant constriction. In this section I consider the effects of /w/ on the noise spectra of consonant release at different consonant places. The peak location frequency and the intensity difference between mid and high frequency regions are considered for stop bursts, and the center of gravity of spectra is measured for the frication noise of fricatives.

The overall shape of stop release burst noise spectra varied among places of articulation, and accordingly the way they were altered by the presence of /w/ component also differed. Velar spectra are characteristically compact, i.e., with an articulated intensity peak, in comparison to the diffuse spectra of coronal and labial stop bursts (Blumstein and Stevens 1979, Stevens 1980). As shown in Figures 2.20 and 2.21, both KV (red) and KwV (green) spectra had a single well-pronounced intensity peak below 3000 Hz, with intensity considerably greater than the peaks in higher frequency regions. Notice that in the KwV spectra the peak is located at a lower frequency than in the KV spectra.

Figure 2.20. Smoothed spectra of velar stop bursts, produced by a Korean male speaker. Vertical axis indicates the intensity (dB).

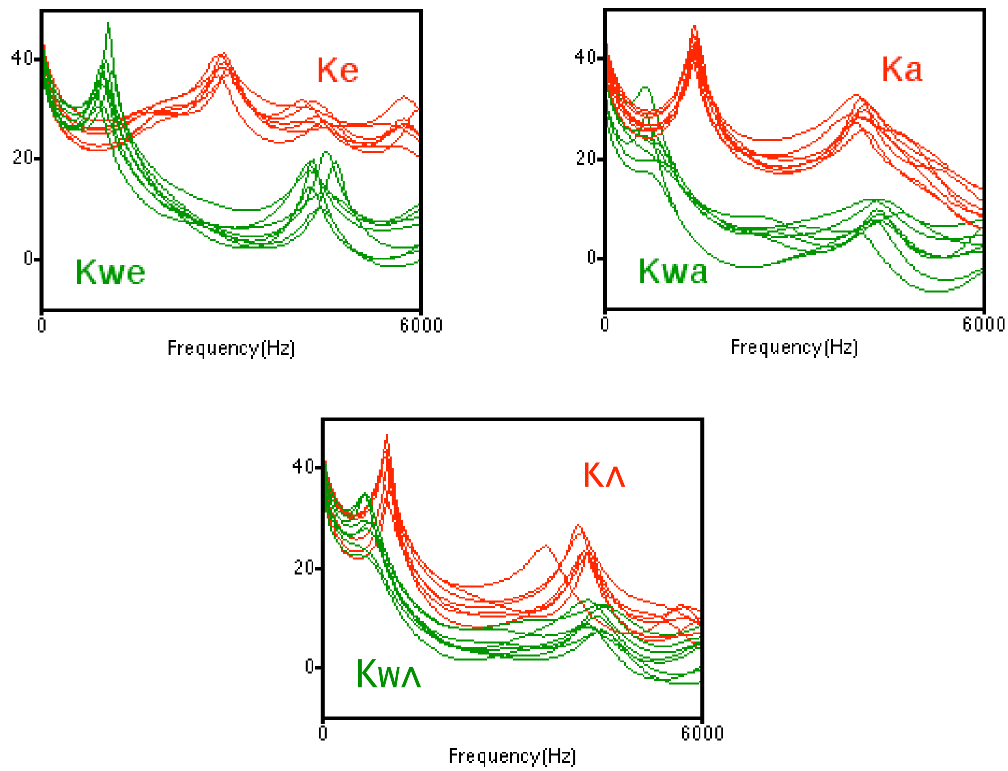
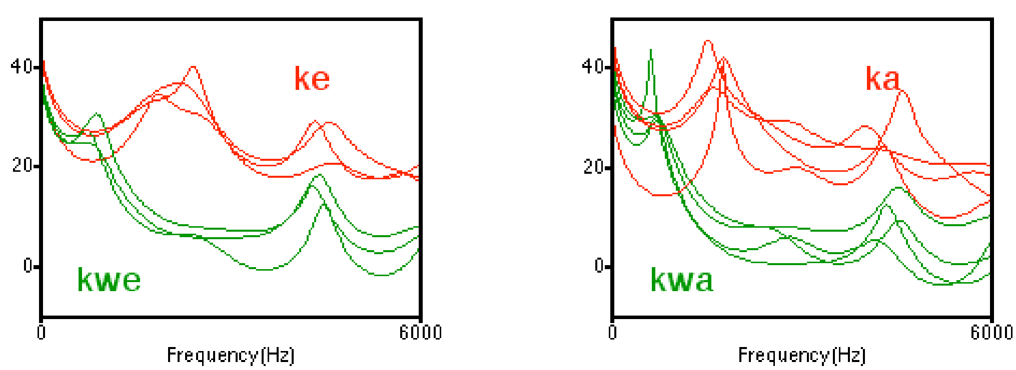



Figure 2.21. Smoothed spectra of velar stop bursts, produced by a Spanish male speaker. Vertical axis indicates the intensity (dB).




Both labial and dental stop burst spectra can be characterized as *diffuse-falling*, following Lahiri et al.'s (1984) characterization of burst spectra. In other words, the spectral energy was spread out over a wide frequency range in these spectra, as shown by multiple numbers of peaks whose amplitude tends to decrease as the frequency increases (Figures 2.22 and 2.23). The effect of /w/ was more obvious on dental burst spectra than labial spectra. The presence of /w/ altered the spectral shape of dental stop bursts, into more compact spectra with prominent amplitude peaks below 4 kHz. In contrast, labial stop spectra stayed diffuse-falling regardless of the syllable types (CV or CwV). This pattern was constantly observed throughout the tokens of both languages.

Figure 2.22. Smoothed spectra of dental and labial stop bursts, produced by a Korean male speaker. Vertical axis indicates the intensity (dB).

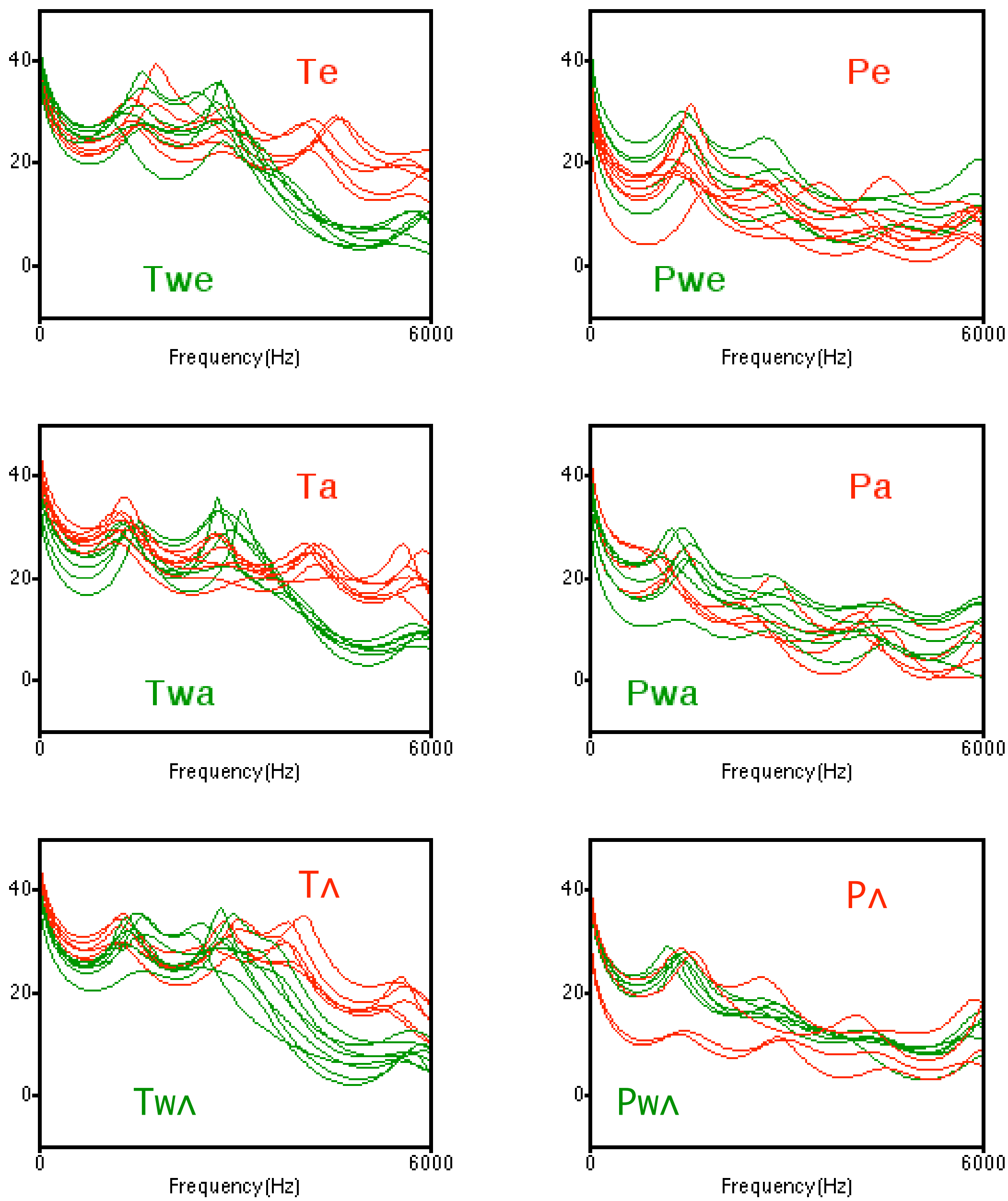
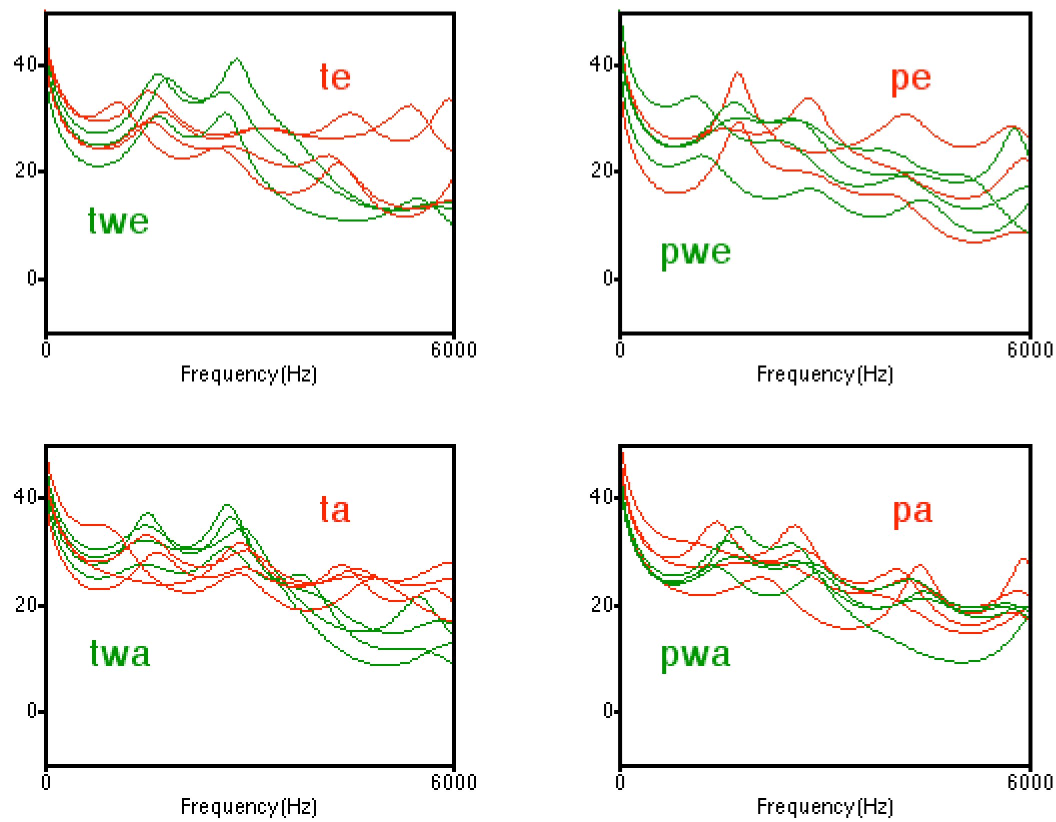


Figure 2.23. Smoothed spectra of dental and labial stop burst, produced by a Spanish male speaker. Vertical axis indicates the intensity (dB).




The average peak location frequency of each stop burst spectrum is reported in 2.3.3.1, which is generally taken as a primary indicator of spectral energy concentration in the studies of noise spectra (Stevens 1980, Bonneau *et al.* 1996, Gordon *et al.* 2002, Jongman *et al.* 2000, de Manrique and Massone 1981 among many others).

2.3.3.1 Stop burst spectral peak location frequency

Spectral peak location frequency was defined as the frequency of the highest amplitude peak among 6 poles in an LPC smoothed spectrum.[15] Figures 2.24 and 2.25 show the average peak location frequencies of P, T and K stop release burst spectra in Korean and Spanish.

[15] Since the sampling rate was 22.05 kHz, the number of LPC coefficients for a vowel analysis should be at least 22, and hence the number of peaks 11 (half of the number of coefficients). However, for voiceless sounds, "it is often possible to model adequately with far fewer coefficients" (Harrington and Cassidy 1999:222).

Figure 2.24. Mean spectral peak location frequency of the Korean stop bursts in CV and CwV. Error bars indicate standard deviation.

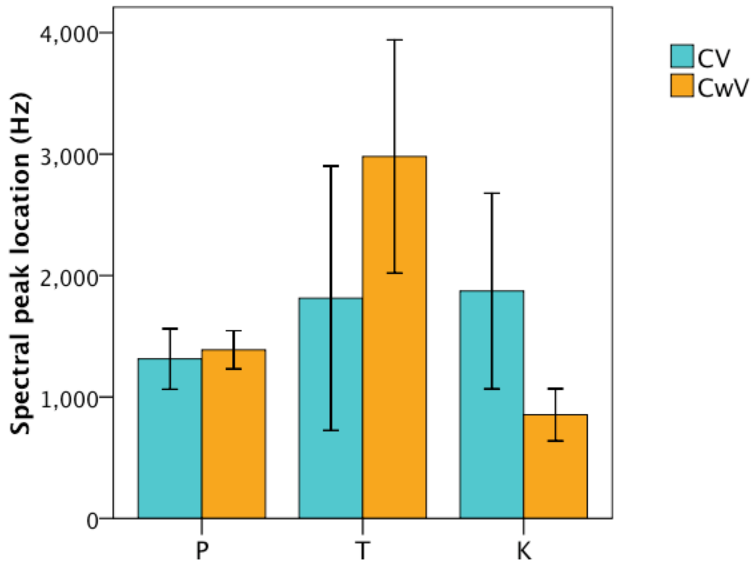
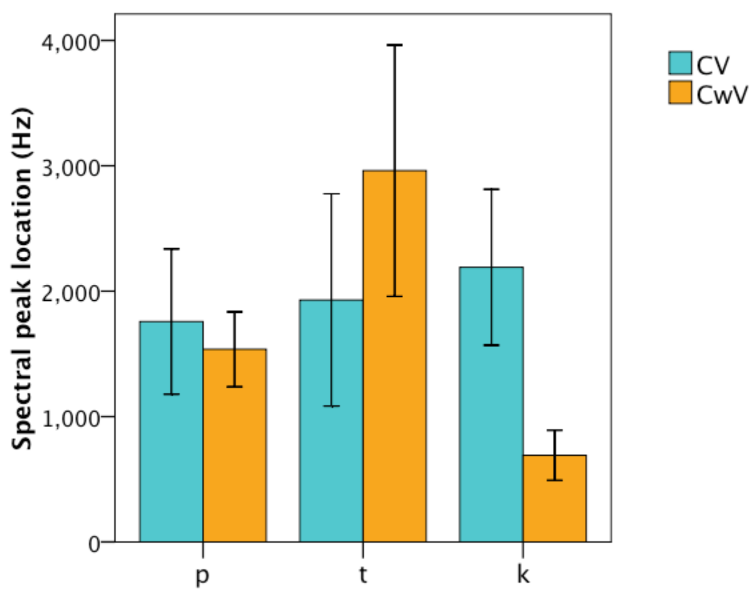



Figure 2.25. Mean spectral peak location frequency of the Spanish stop bursts in CV and CwV. Error bars indicate standard deviation.




The /w/ component had a frequency lowering effect on the peak location of K burst spectra. Because lip rounding and protrusion increase the cavity in front of the stop constriction, we might expect the T spectra to have lower frequency prominence in CwV than in CV, just like the velar stop burst spectra. However, there was a weak tendency for the mean frequency of spectral peak of the coronal stop burst to be higher in labialized context. The location of the labial burst spectral peak stayed fairly constant in CV and CwV.

Since the effect of /w/ on the stop burst spectra is realized in such different ways for the different places of articulation, repeated measures ANOVA (within-subject factors *w*, *phonation* and *vowel*) was done separately on the peak locations of P, T and K stop burst spectra. The main effect of *w* was found only for K in both Korean (F(1,3)=522.8, p<0.001) and Spanish (F(1,3)=36.56, p<0.009). Data from both languages showed significant interaction of *w* and *vowel* on the peak location for velar (Korean F(2,6)=488.184, p<0.001; Spanish F(1,3)=19.196, p<0.022). In pair-wise comparisons, the peak location differed to a greater degree when the vowel was /e/ than when it was /a/ in both languages. In Korean, the back vowel /ʌ/ context showed the smallest contrast (Figures 2.26 and 2.27).

Figure 2.26. Mean spectral peak location frequency (±standard deviation) of the Korean velar stop bursts.

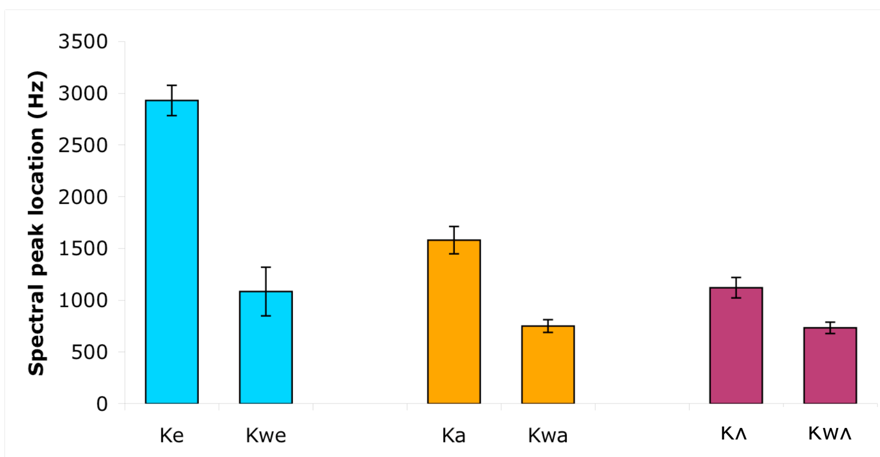
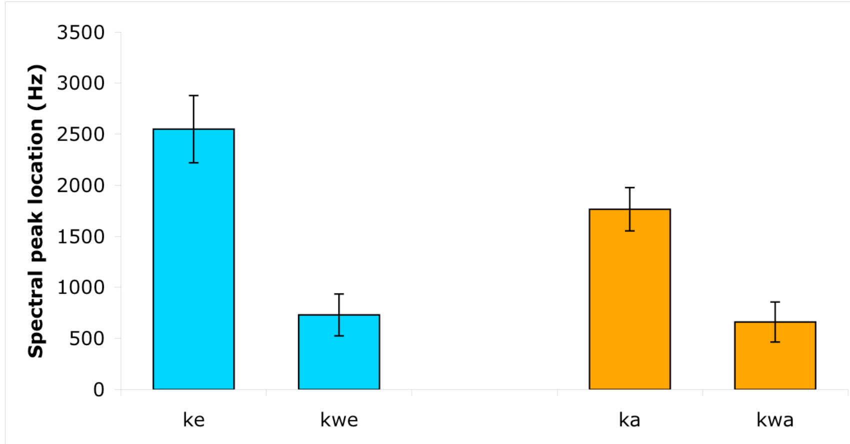


Figure 2.27. Mean spectral peak location frequency (±standard deviation) of the Spanish velar stop bursts.




As shown in Figures 2.20 and 2.21, velar stop burst spectra had the amplitude peak between 500 and 3000 Hz, with the precise location largely determined by whether there was a /w/ component, and what the following vowel was. It is in fact a typical characteristic of the velar stop burst or frication spectra that the peak of spectra is continuous with the F2 or F3 of the following vocoid (Raškinis and Dereškeviciute 2007, de Manrique and Massone 1981).[16] Thus, in the context of low F2 or F3 of the following vocoid, such as back vowel or glide /w/, the velar stop burst peak frequency is lower than in front vowel contexts. According to Johnson (2003:143-144), both the burst and the F2 locus of velar stops are associated with a front cavity resonance. In this they differ from alveolar/dental stops, in which the release burst is shaped by a front cavity resonance but F2 locus is associated with a back cavity resonance. Thus, the close correlation between vowel formants and the stop burst at velar consonant place in particular may be due to the shared resonating cavity of F2 locus and the stop release burst.

Let us now turn to the cross-language comparison of the stop burst peak frequency patterns. Mixed model ANOVA (Within-subject factors *w* and *vowel*, and between-subject factor *language*) revealed that there was no significant interaction of *w* and *language* in any place context.

Overall, the frequency lowering effect of burst spectra from /w/ was observed in the spectral peak location measure only for the velar stops in both languages. In the coronal context, the presence of /w/ increased the peak location frequencies, though the effect was not statistically significant. As shown by the standard deviation as large as 1000 Hz, the location of the coronal stop burst peak varied to a great extent. Notice that the dental stop burst spectra in Figures 2.22 and 2.23 had peaks at around 1000-1500 Hz and 3000-4000 Hz, and in plain T context the lower frequency peak tended to have higher amplitude, but in Tw spectra this tendency was weakened, and in many tokens the peak at 3000-4000 Hz had more energy. This increased the average peak location frequency of Tw spectra. And since it varied from token to token which of the two peaks (one at 1000-

[16] Bonneau et al. (1996: 557) report that before a front vowel, the velar stop burst peak is related to the F3 rather than F2 of the following vowel.

1500 Hz and the other at 3000-4000 Hz) had higher amplitude than the other, the standard deviation was larger than that of other stop bursts.

Although the average peak location frequency was higher in TwV than in TV, it was clear from the shape of the dental stop burst spectra in Figures 2.22 and 2.23 that presence of /w/ rendered the spectral energy more concentrated in the lower frequencies: The peak amplitude in the higher frequency area was lower in the TwV spectra than in TV spectra. Thus, the presence of /w/ did have a frequency lowering effect on the overall spectral energy concentration of the dental stop bursts. The labial stop burst spectral shape, in contrast, did not seem to be altered by the addition of /w/. Therefore, the highest amplitude peak location frequency may not be a reliable measure for capturing the shift of the spectral energy distribution by /w/ on dental stop bursts. As an alternative way of measuring the noise spectral energy distribution, the intensity difference between two different frequency ranges (Stevens *et al.* 1999, Flemming 2007) is measured in 2.3.3.2.

2.3.3.2 Intensity difference between mid- and high-frequency regions in stop burst spectra

To describe the spectral energy distribution, intensity was measured from two frequency regions of the spectra generated from the 12 ms window centered at the burst onset—mid frequency range (1250-4000 Hz) and high frequency range (4500-8000 Hz)—and the difference between them was calculated (Cf. Flemming 2007, Stevens et al. 1999, Li et al. 2007).[17] As /w/ decreases the amplitude of peaks in the relatively higher frequency range in dental stop burst spectra (Figure 2.22 and 2.23), it is expected that the energy difference between the mid and high frequency regions will be greater for the burst spectra of Tw than T. In contrast, the energy difference is expected to be similar between plain P and Pw burst spectra, as the labial stop burst spectra were not so much affected by the presence of /w/.

Figures 2.28-2.31 are the scatter-plots of the Korean and the Spanish dental and labial stop burst intensities, in which the mid-frequency intensity is represented on the horizontal axis, and the high-frequency intensity on the vertical axis. The dots closer to the diagonal represent tokens in which the intensity of mid and high frequencies are relatively comparable.

[17] The mid and high frequency ranges are defined somewhat differently from Flemming (2007) and Stevens et al. (1999), in which the mid frequency range is up to 3000 Hz, and high frequency from 3500 Hz. As we can see from the spectra in Figures 2.22 and 2.23, the TwV spectra consistently have two amplitude peaks below 4000 Hz. If the high frequency range is defined as 3500-8000 Hz, the intensity difference between this and the mid frequency region may not capture the overall frequency lowering effect of spectral energy concentration from /w/, as both mid and high frequency ranges would contain an amplitude peak in Tw spectra as well as in T spectra.

Figure 2.28. Scatter-plot of the Korean dental stop burst intensity in mid and high frequency ranges.

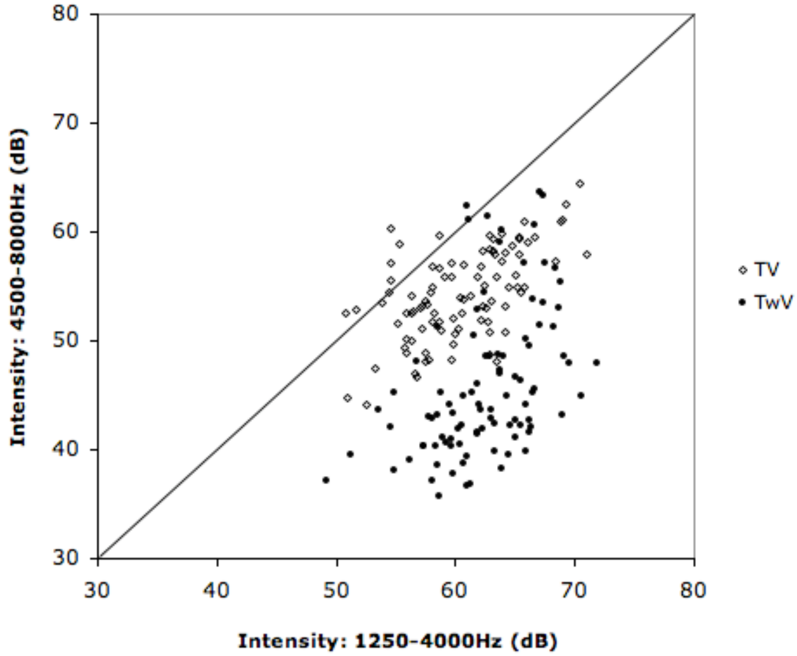



Figure 2.29. Scatter-plot of the Korean labial stop burst intensity in mid and high frequency ranges.

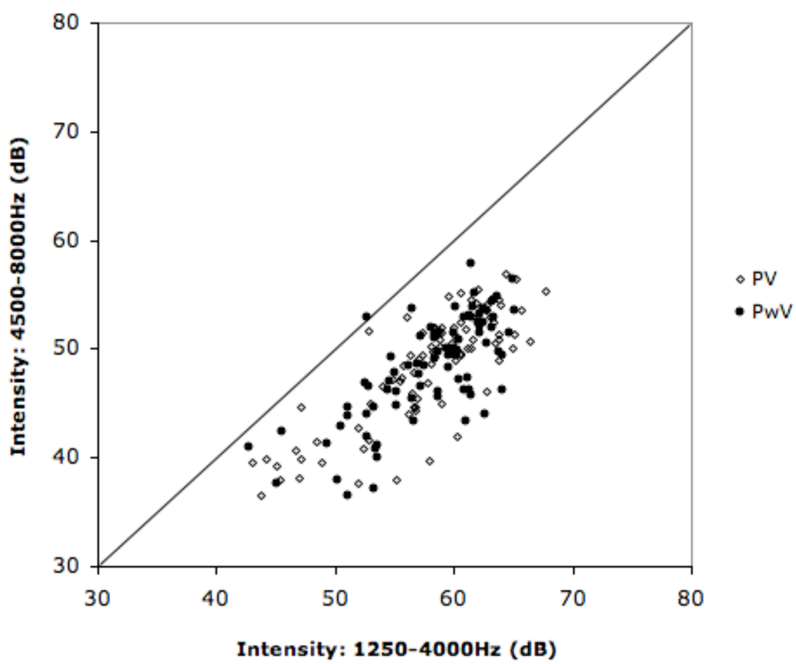


Figure 2.30. Scatter-plot of the Spanish dental stop burst intensity in mid and high frequency ranges.

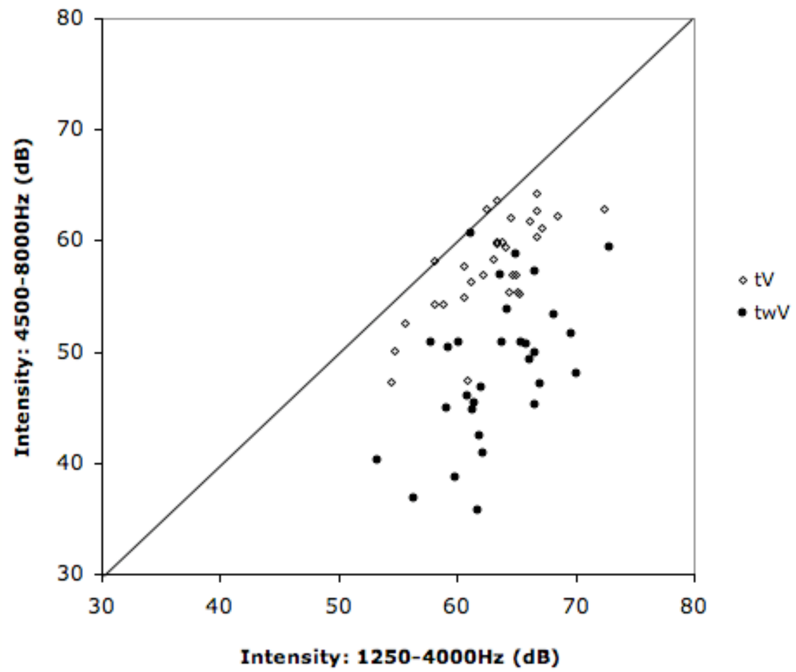
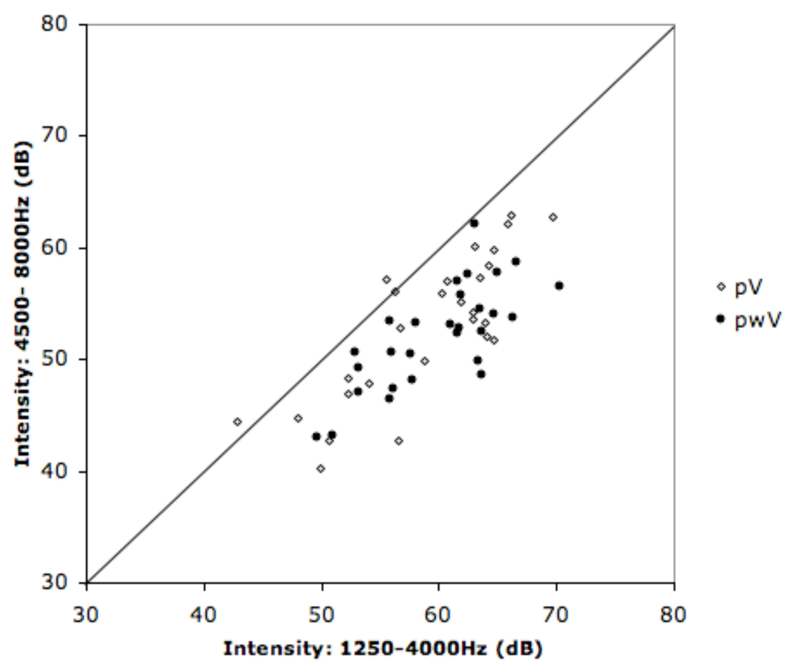



Figure 2.31. Scatter-plot of the Spanish labial stop burst intensity in mid and high frequency ranges.



In both Korean and Spanish, TV and TwV seemed to be fairly well distinguished: most of TV tokens are located closer to the diagonal than TwV tokens are in Figures 2.28 and 2.30, showing that spectral energy is more evenly distributed across mid and high frequencies in TV burst than TwV burst. PV and PwV are not well separated, as shown by their overlap in Figures 2.29 and 2.31.

Tables 2.3 and 2.4 show the intensity difference between the two frequency regions in the stop burst spectra of the Korean and the Spanish PV, PwV, TV and TwV. Greater value indicates more asymmetric, skewed distribution of spectral energy, with concentration in the mid-frequency region.

Table 2.3. Intensity difference between mid and high frequencies in the Korean labial and dental stop burst spectra.

| | Mean (dB) | Std. deviation | N | | Mean (dB) | Std.deviation | N |
|---|---|---|---|---|---|---|---|
| PV | 9.41 | 3.35 | 86 | PwV | 9.69 | 3.53 | 83 |
| TV | 6.13 | 3.92 | 93 | TwV | 16.56 | 6.22 | 94 |

Table 2.4. Intensity difference between mid and high frequencies in the Spanish labial and dental stop burst spectra.

| | Mean (dB) | Std. deviation | N | | Mean (dB) | Std.deviation | N |
|---|---|---|---|---|---|---|---|
| PV | 5.99 | 3.98 | 27 | PwV | 7.73 | 3.53 | 28 |
| TV | 5.16 | 3.11 | 31 | TwV | 14.66 | 5.64 | 30 |

As expected from the Figures 2.28 and 2.30, TwV had considerably bigger intensity difference than TV, whereas the intensity difference of PV and PwV are similar to each other. Repeated measures ANOVA (factors *w, vowel* and *phonation* for Korean; *w* and *vowel* for Spanish) was carried out on labial and coronal stop burst separately, and the results showed that there was a main effect of *w* on the intensity difference for dental stop burst in both Korean ($F(1,3)=19.97$, $p<0.021$) and Spanish ($F(1,3)=15.016$, $p<0.03$). The labial stop burst did not show an effect of /w/ in either Korean ($p=0.189$) or Spanish ($p=0.341$). Thus intensity difference distinguished between CV and CwV syllables with respect to the dental stop bursts, but not the labial stop bursts.

Vowel quality did not have an effect on the degree of the TV-TwV contrast in the intensity difference. The interaction of factors *w* and *vowel* was insignficant in both Korean ($p=0.146$) and Spanish ($p=0.903$).

Language comparison was made only for the dental context, which showed a significant in-language difference between CV and CwV. The interaction of *w* and *language* was insignificant ($F(1,6)=0.075$, $p=0.794$). Therefore the /w/ component had similar degree of frequency-lowering effect on the dental stop release burst spectra of Spanish and Korean.

To summarize the results from stop burst spectral measurements: the effect of /w/ on stop burst spectra was not uniform across consonant places: velar spectra were compact in both CV and CwV contexts, with the spectral peak location significantly lowered in frequency in the latter; dental spectra exhibited the change in the overall spectral shape, from diffuse-falling to more compact spectra. The frequency lowering effect of /w/ on dental stop burst spectra was better captured by the consideration of spectral shape and the distribution of spectral energy in different frequency ranges, than by the prominent peak location frequency. Labial spectra seemed to be the least distinctive between CV and CwV contexts, in both the spectral peak location and the energy distribution. These results partially support the idea that the phonological asymmetries in Cw combinations may be related to the phonetic salience of the CV-CwV contrast, as Cw combinations are preferred on velars and disfavored on labials.

There was no statistically significant effect of *language* on the degree of the CV-CwV contrast, in either the peak location frequency of velar or the intensity difference of dental stop burst spectra. This contrasts with the vocalic cues such as vocoid duration and F2 discussed in the previous section, in which Spanish exhibited significantly bigger CV-CwV contrast than Korean.

2.3.3.3 Fricative noise spectral center of gravity

Noise spectra of fricatives were examined to determine whether the C-Cw contrast showed any place, vowel or language effects. The spectral properties of the Spanish voiceless fricatives /f/, /s/ and /x/ and the Korean /s'/ were examined. Center of gravity (COG) of the spectra, defined as the mean of the frequency values weighted by their amplitudes, was used as the indicator of frication noise frequency (Jongman et al. 2000, Gordon et al. 2002). Since Korean does not have fricatives produced in the labial and velar regions, the language-internal consonant place effects were investigated only for the Spanish fricatives. Also, a language comparison was made only between the Spanish and the Korean alveolar fricatives. The spectra were generated at three different time points of the frication – the points 1/4, 1/2 and 3/4 into the total fricative duration from the beginning. COG was measured at these three different points, to see how the degree of C-Cw contrast changed throughout the fricative duration.

*Spanish /f/, /s/ and /x/.* Mean COG of the Spanish voiceless fricative spectra at three time points is shown in Figures 2.32, 2.33 and 2.34.

Figure 2.32. Mean center of gravity (±standard deviation) of the Spanish /f/ frication noise spectra in fV and fwV, measured at 3 time points.

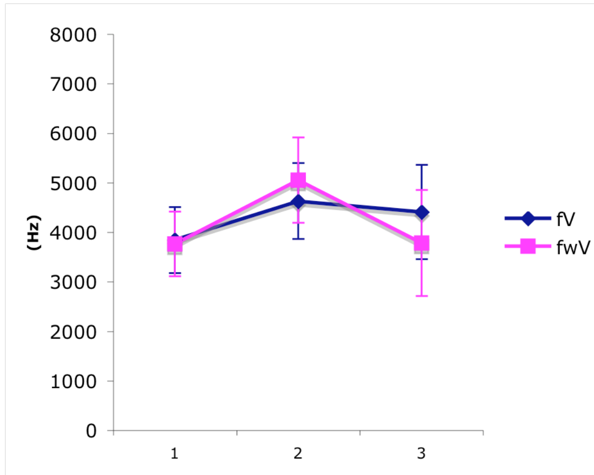



Figure 2.33. Mean center of gravity (±standard deviation) of the Spanish /s/ frication noise spectra in sV and swV, measured at 3 time points.

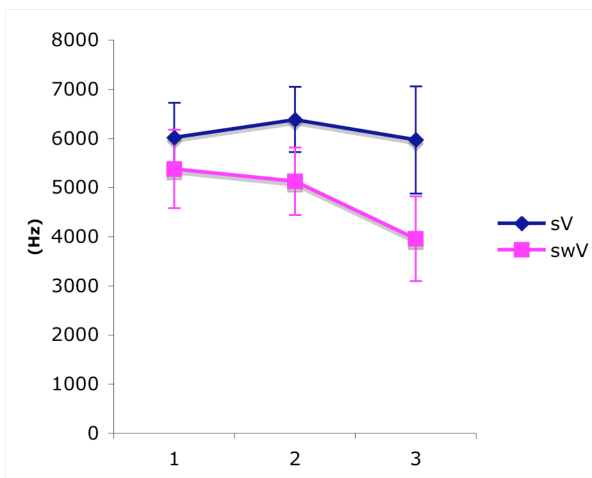
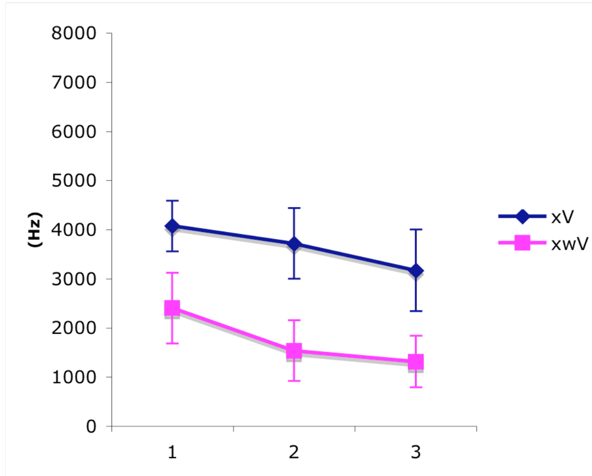


Figure 2.34. Mean center of gravity (±standard deviation) of the Spanish /x/ frication noise spectra in xV and xwV, measured at 3 time points.

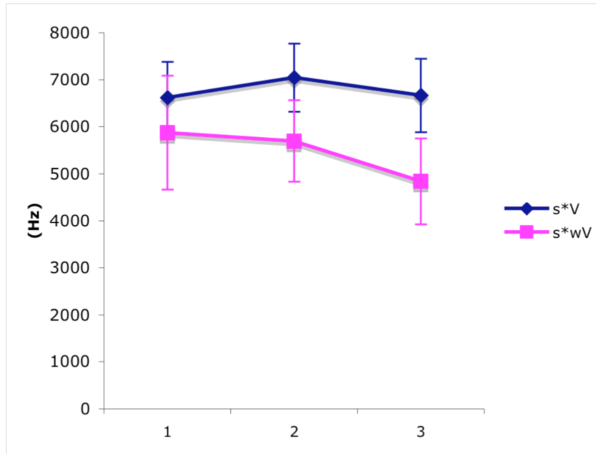



A repeated measures ANOVA (*w* * *place* * *vowel* * *point*) was done on the COG of the Spanish fricatives. The main effect of *w* was found (F(1,3)=104.451, p<0.002), showing that the COG was in general significantly lower in Cw than in C context. However, there was also a significant interaction of factors *w* and *place* (F(2,6)=25.845, p<0.001), which suggests that not all fricatives undergo the COG lowering by /w/: from Figures 2.32-2.34 we may see that the effect of /w/ on COG is clearer in /s/ and /x/ spectra than in /f/ spectra, mirroring the pattern for stops. This observation is supported by the ANOVA results within each fricative context: the main effect of *w* was significant in /s/ (p<0.01) and /x/ (p<0.001), but insignificant in /f/ spectra (p=0.644).

The interaction of factors *w, place* and *point* was also significant (F(4,12)=13.247, p<0.001). This suggests that the degree of C-Cw contrast differed between time points, but only in coronal place context: The /s/-/sw/ contrast increases over the course of the fricative duration (*w* * *point* significant: p<0.046). /x/-/xw/ difference was greater at time point 2 than at other points, but this difference was only marginally significant (p<0.057).

*Korean /s'/.* The overall pattern of COG change in the Korean fricative /s'/, shown in Figure 2.35 below, is similar to that of Spanish /s/ (Figure 2.33). Like in Spanish, there was a main effect of *w* (F(1,3)=56.275, p<0.005). The difference between /s'/ and /s'w/ seems to increase across the time points, though the interaction of *w* * *point* was not significant (p=0.094), unlike in Spanish.

Figure 2.35. Mean center of gravity (±standard deviation) of the Korean /s'/ frication noise spectra in s'V and s'wV, measured at 3 time points.




*Language comparison.* Repeated measures ANOVA with between-subject factor *language* was done on the COG of alveolar fricative spectra from both languages; the interaction of *w* and *language* (F(1,6)=0.075, p=0.793) and the interaction of *w, language* and *point* (F(1,6)=0.21, p=0.663) did not reach the significance level. This confirms the observation from the Figures 2.33 and 2.35 that the effect of /w/ on the COG of alveolar fricative spectra is similar in Spanish and Korean.

To summarize, the effect of /w/ on the frication noise spectral center of gravity was statistically not significant for the Spanish labial fricative /f/, showing the disadvantage of labials in the salience of CV-CwV contrast. There was no cross-linguistic difference with regard to the degree of frequency lowering of alveolar fricative frication noise, and the time point at which the effect of /w/ was the greatest in the fricative duration. The lack of cross-language difference in frication noise COG is paralleled with that in stop burst spectral measures discussed above.

## 2.4 Summary and discussion

The results of the acoustic study showed that CV and CwV in certain contexts exhibit greater acoustic differences than in other contexts. This section summarizes the findings of the acoustic measures, with particular focus on the context effects in the contrast between CV and CwV syllables, and discusses their implications for the distributional asymmetry of Cw combinations.

### 2.4.1 Consonant place effect on the CV-CwV contrast

The summary of the consonant place effects observed above is presented in Table 2.5. Throughout the table the notation "A > B" indicates that the CV-CwV difference was greater in the context of A than in B in terms of a particular acoustic measure.

Table 2.5. Effects of consonant places on the CV-CwV difference

| Measurements | | Δ(CV-CwV) | | |
|---|---|---|---|---|
| | | Korean | Spanish | Section # |
| Vocoid duration | | No place effect | No place effect | 2.3.1 |
| Formant | F2 | k' > {p', t'} | K>{P, T} | 2.3.2.2 – 2.3.2.3 |
| | F1 | No place effect | {P, T} > K | 2.3.2.1 |
| fricative noise spectra | Center of gravity | NA | {x, s} > f | 2.3.3.3 |

Consonant place effects were observed in the formants at vocoid onset and the F2 change. F2 contrast was bigger in the velar context than in other consonant contexts. A significant place effect on F1 contrast was observed in Spanish, but not in Korean. It was in the velar context that F1 contrast between the Spanish CV and CwV was smaller than in other consonant contexts, a result contrary to that from the F2 measures.

Overall, the claim that the distributional asymmetry of Cw combinations is related to the salience of /w/ in this context is partly supported by the acoustic study results: Velar place, the most common for contrastive labialization, had a clear advantage in the size of the CV-CwV F2 contrast in Korean, as all significant comparisons (F2 at onset and F2 change) pointed to that direction. The results of formant frequency measures do not uniformly show the contrast to be particularly weak in the labial context in comparison to coronal place. Nevertheless, the disadvantage of labial was partly suggested in the observation of F2 transition directions (2.3.2.3): PV and PwV syllables shared the rising F2 trajectories in the /a/ and /e/ vowel contexts, unlike TV and TwV syllables.

The frequency lowering effect of /w/ was found on the burst spectra of both velar and dental stops, but not on the labial stop burst spectra. The same pattern was observed in the fricative center of gravity measure for Spanish. Different measures were employed for velar and dental stop bursts to capture the effect of /w/, as the spectral shape was affected in fairly different ways. Thus, a more conclusive comparison between consonant places regarding the size of the cues for the C-Cw contrast needs to await the perception study in the next chapter.

### 2.4.2 Vowel effect on the CV-CwV contrast

Table 2.6 summarizes the vowel effects found in the degree of CV-CwV differences in each acoustic measure.

Table 2.6. Effects of vowels on the CV-CwV difference

<table>
<tr><td colspan="2" rowspan="2">Measurements</td><td colspan="3">Δ(CV-CwV)</td></tr>
<tr><td>Korean</td><td>Spanish</td><td>Section #</td></tr>
<tr><td colspan="2">Vocoid duration</td><td>No vowel effect</td><td>No vowel effect</td><td>2.3.1</td></tr>
<tr><td rowspan="2">Formant</td><td>F2</td><td>{e, a} > ʌ<br>(Ke > Ka > Kʌ)</td><td>e > a</td><td>2.3.2.2-<br>2.3.2.3</td></tr>
<tr><td>F1</td><td>a > e</td><td>a > e</td><td>2.3.2.1</td></tr>
<tr><td rowspan="2">Burst noise spectra</td><td>K peak</td><td>e > a > ʌ</td><td>e > a</td><td>2.3.3.1</td></tr>
<tr><td>T intesity difference</td><td>No vowel effect</td><td>No vowel effect</td><td>2.3.3.2</td></tr>
<tr><td>Frication noise spectra</td><td>Center of gravity</td><td>No vowel effect</td><td>No vowel effect</td><td>2.3.3.3</td></tr>
</table>

In Korean, for which three vowel contexts were considered, F2 at onset and F2 change difference between CV and CwV were bigger before fronter vowels /e/ and /a/ than before the back vowel /ʌ/. Recall that /ʌ/ is a less favored vowel after Cw combinations, in that Tw does not co-occur with it. This shows that there was a correlation between the CV-CwV acoustic distinctiveness and the distribution of Cw combinations at least in the F2 measure. The size of the F1 contrast, on the other hand, was bigger for /a/ than /e/ contexts in both Korean and Spanish. Thus the formant transition contrast seems to be greater in the Ca-Cwa contrast than in the Ce-Cwe contrast in Korean, as the former is advantageous in the F1 cue, whereas the size of the F2 cue was comparable.

The effect of vowels on the CV-CwV difference was found not only in formants but also in the stop release burst spectra when the consonant was velar: the /e/ vowel conditioned a bigger contrast in both Korean and Spanish, and the Korean back vowel /ʌ/ showed smaller contrast than other vowels. The F2 at onset and F2 change measures, as well as the peak location of velar stop bursts, make a correct prediction with regard to the universal avoidance of back vowels after labialized consonants: the CV-CwV difference is smaller in the Korean back vowel /ʌ/ context than fronter vowel contexts.

### 2.4.3 Language effect on the CV-CwV contrast

The language difference was also observed in the size of the CV-CwV contrast. Table 2.7 summarizes the language effects found in the size of the CV-CwV contrast in each measurement.

Table 2.7. Effects of languages on the CV-CwV difference

| Measurements | | Δ(CV-CwV) | Section # |
|---|---|---|---|
| Vocoid duration | | Spanish > Korean | 2.3.1 |
| Formant | F2 | Spanish > Korean | 2.3.2.2 - 2.3.2.3 |
| | F1 | No language effect | 2.3.2.1 |
| Burst/frication noise spectra | K peak | No language effect | 2.3.3.1 |
| | T intensity difference | No language effect | 2.3.3.2 |
| | /s/ center of gravity | No language effect | 2.3.3.3 |

Spanish had a bigger CV-CwV contrast in both durational and spectral aspects of the vocoid, especially the F2-related cues. However, there was no statistically significant language effect on the degree of CV-CwV contrast in the stop burst spectral properties: Velar stop burst peak frequency was located at similar frequencies in the two languages, and thus the difference between K and Kw was also comparable. Effect of /w/ on dental stop burst and alveolar fricative noise spectral energy distribution was also similar in Korean and Spanish, as shown by the intensity difference and center of gravity, respectively.

In the beginning of this chapter, I assumed that Korean and Spanish differ in the way Cw combinations are phonetically realized. In particular, Cw combinations were expected to be closer to labialized consonants in Korean, and to Cw clusters in Spanish. The results of the cross-linguistic acoustic comparison confirm this, given the relationship between the status of Cw combinations and the acoustic salience of the glide discussed in 2.1. The difference between CV and CwV, especially in the vocoid duration and the formants, was greater in Spanish than in Korean. In addition to the cross-linguistic difference in the size of the C-Cw contrast, I also found that the contribution of the /w/ component to the total vocoid duration was statistically not significant in Korean, and that there was no constant F2 target frequency of /w/ in the Korean CwV syllables. Based on these findings, I will categorize the Korean Cw combinations as labialized consonants ($C^w$) in the following chapters, as opposed to the Cw clusters of languages like Spanish.

The results of the acoustic study showed how the strength of the spectral cues for the CV-CwV contrast —formant transition and consonant release noise frequency— varies across consonant places, vowels, and the degree of overlap between C and the glide /w/ component. Before ending this chapter, let us summarize the implication of the findings from the acoustic study for the Dispersion-Theoretic analysis of the Cw distribution. Contextual effects on the size of the C-Cw difference were found in some acoustic measures, which are in accord with the relationship between phonological and phonetic asymmetries of Cw combinations discussed in 1.2: First, the formant transition cue of the CV-CwV contrast was weaker in the back vowel /ʌ/ context than in /e/ or /a/ contexts, and stronger in the velar stop context than in labial or coronal stop contexts. Recall that the back vowels are avoided after Cw combinations, and velar consonants are the most likely to be labialized. Second, stop burst noise frequency differed between CV and CwV when C was a coronal or velar stop, but not when it was a labial stop, which is often avoided in Cw combinations.

However, some predictions made in 1.2 were not directly supported by experimental results for some of the measures in this chapter. Though it was expected that the C-Cw contrast in the stop burst quality is greater when C is a velar stop, it was not possible to compare the degree of burst noise frequency lowering effect by /w/ on three consonant places with a single acoustic measure because of the inherently different nature of the velar, coronal and labial stop burst spectra. The formant transition frequency measures did not distinguish between labials and coronals, though the observation of F2 transition directions suggests that PV and PwV may be perceptually more similar to each other than TV and TwV in some vowel contexts.

The following chapter presents a perception experiment which investigates potential asymmetries in the perception of the CV-CwV contrast. By comparing the perceptibility of the C-Cw contrast in different contexts, a more conclusive comparison of cue strength will be made among consonant places and vowels. A more important reason to look into the perceptual asymmetry is to see how the place and the vowel effects on the acoustics of the CV-CwV contrast are reflected in perception; in other words, if they invoke similar place and vowel effects on the perceptibility of the CV-CwV contrast.

The perception experiment in the following chapter is designed on the assumption that the C-Cw contrast is "weaker" when the Cw combination is a labialized consonant than when it is a cluster, and thus is more difficult to perceive. Note that the first part of this assumption is made based on the acoustic study results; recall that at least in the formant transition cue, the Spanish C-Cw difference was greater than the Korean C-$C^w$ contrast. The perception of the "weaker" C-Cw contrast, namely the Korean labialization contrast, will be tested on speakers of Spanish, in which the CV-CwV contrast is "stronger", as well as on speakers of a language that lacks Cw combinations, namely Russian, to whom the C-Cw contrast is entirely non-native. The strength of each cue is tested separately, so that the results can be used to build the perceptibility scales of formant transition and noise frequency necessary for the DT analysis in Chapter 4.

# Chapter 3. Perception of the labialization contrast

In the previous chapter we saw that the acoustic distinctiveness between CV and CwV syllables was influenced by consonant place and vowel contexts, as shown by acoustic measures such as formant transition and stop burst noise frequencies. For example, velar consonants provided a greater degree of distinctiveness between CV and CwV in F2 formant transitions than labials and coronal consonants. Furthermore, the Korean back vowel /ʌ/ conditions a smaller CV-CwV contrast than the more front vowels /e/ and /a/ in both formant transition and burst peak frequency. These experimental results are in accordance with the prediction made in 1.2 regarding the phonetic asymmetries of the CV-CwV contrasts, and thus support the main claim that the phonological patterns of Cw distribution is related to the phonetic salience of the /w/ component. However, the effect of consonant place on the strength of the consonant release cue, especially between velar and coronal contexts, was difficult to assess with a single measurement like peak location frequency.

This chapter presents a perception experiment that tests the relative strength of perceptual cues for the CV-CwV contrast in different consonant and vowel contexts. The perception experiment will provide a way of comparing the size of cues to the CV-CwV contrast in different contexts, i.e., the performance in discriminating CV and CwV. The results of the experiment will be used as a basis for establishing the auditory scales relevant for the DT analysis of the distributional asymmetries in Cw combinations, developed in Chapter 4. Section 1 introduces previous research related to the perception of prevocalic glides and secondary articulation. Sections 2 and 3 present the procedure and results of the experiment and the general discussion of the results. Section 4 develops the distinctiveness scales necessary for the Dispersion-Theoretic analysis of the Cw distribution given in Chapter 4.

## 3.1 Perceptual correlates of prevocalic glide and the implications for the CV-CwV contrast

Studies have shown that both durational and spectral cues are important in the perception of prevocalic glides. Liberman et al.'s (1956) classic study on formant transition tempo suggests that the CV transition duration is relevant to the perception of a prevocalic glide. From synthetic speech stimuli with rising formant transition, the listeners perceived /bɛ/ when the transition duration was shorter, but as the duration reached 40-50 ms the perception changed to /wɛ/. From a similar set of stimuli with falling F2 transition the listeners heard /gɛ/ when the transition duration was shorter, and /jɛ/ when the duration reached 50-60 ms. Using stimuli in which F2 change duration, F2 change extent or the rate of F2 change is manipulated, Schwab et al. (1981) found that F2 change extent (in Hz) and F2 change transition duration contribute to the distinction between [ba] and [wa] syllables. Listeners tended to hear [wa] from a longer F2 transition duration or bigger F2 change extent.[18] Thus, different aspects of formant transition serve as a perceptual cue for a prevocalic glide, such as the direction (rising, falling, level), extent (frequency), and duration of the transition.

The effect of duration and degree of rising formant transition on the perception of /w/ suggests that labialized labial consonants may be perceptually closer to plain labials than labial + w clusters are. The acoustic measurements taken in Chapter 2 included the duration of vocoid, F1 and F2 frequencies at the onset of the vocoid, and the overall degree of F2 change in frequency. Recall that in the Korean CwV syllables the vocoid duration was not significantly longer than that in the CV syllables, and the F2 change degree was smaller than in the Spanish counterparts. Thus, rising F2 of the /w/ component is more likely to be mistakenly attributed to the labial stop in the Korean Pw combination than in the Spanish counterpart. The fact that shared transition direction is a factor causing perceptual confusion suggests that CV and CwV that share the falling formant transition direction, e.g., /tʌ/ and /$t^w$ʌ/ (See 2.3.2.3), may also be subject to a similar perceptual confusion.

In addition to the vowel formant cues, consonant release properties may also contribute to the perception of a prevocalic glide. Previous studies have found that the stop burst in CV syllables contributes to the identification of the vocalic context to a certain extent. In Ohde and Sharf's (1977) perception experiment, listeners identified the vowel context when only the aperiodic portion of the English stop + V syllables was presented, with an average 74% correctness. Cullinan and Tekieli (1979) found that English stop burst noise provides enough information for recognizing following vowel features such as backness, but not for recognizing other features such as height or tenseness. Note that the stimuli used in these works were English, so the burst included aspiration, which may provide the vowel formant information. In Bonneau's (2000) experiment, the French voiceless stop burst was used, which minimized the contribution from aspiration. The subjects recognized the vowel context from the burst at the correct rate of 86% or above in /ki/, /ku/ and /ti/, with the highest rate for /ku/ (98%).

According to Bonneau (2000:500-501), the reason for the high correct identification rate for /ku/ is the strong co-articulation between /k/ and /u/. This occurs because the lips can easily be protruded for the anticipation of /u/ at the release of the /k/ burst, and the tongue positions for /k/ and /u/ are close to each other. The /u/ vowel was not well recognized in the /t/ burst (39% identification rate), because the tongue body cannot anticipate the /u/ target position at the release of tongue tip closure. Bonneau's claim suggests that the /k/ burst noise may also provide more information for the perception of the prevocalic glide /w/ (i.e., the C-Cw contrast) than a /t/ burst does, given the similarity of /u/ and /w/.

Though no studies appear to have directly tested the role of stop burst in the perception of labialization, the consonant release noise quality seems to be an important source of cues for secondary palatalization. In palatalized dental/alveolars, the affrication at the release of the stop provides a strong cue, as does the formant transition. In his experiment on the perception of Russian /t/ and /$t^j$/ in coda position, Kochetov (2006) found that listeners were able to discriminate the two sounds even when presented with

[18] The F2 change rate (Hz/ms), however, had no significant relationship with the [ba]-[wa] distinctions. In other words, "any particular rate could signal either a stop or semivowel" (Schwab et al 1981: 121).

only the stop release portions. Padgett (2001:209) also mentions the importance of consonant release cues in the palatalization contrast. Padgett claims that labial consonants are dispreferred to coronals for the palatalization contrast, since their release burst quality is not affected by tongue movement, and thus the contrast is cued only by the formant transition difference.

The goal of the current experiment is to compare the relative perceptual distance between CV and CwV in each consonant and vowel context, when only the formant transition or the burst noise cue is provided. For the DT analysis in the following chapter, I will develop the formant transition and burst noise frequency scales in which the nine types of CV-C$^w$V contrasts (3 places * 3 vowels) are ordered in terms of their relative salience. Before that, I first look at the consonant place and vowel effects on the perceptibility of the contrast as I did in the acoustic study. This will show if the contextual asymmetries observed in the acoustic study are reflected in perception.

## 3.2 Methods

### 3.2.1 Stimuli

Korean CV and C$^w$V syllables were used as stimuli in this experiment. One male speaker's recordings from the acoustic study were used as stimuli for the perception experiment. The perception experiment was designed in such a way that the strength of the formant transition and the stop release burst cues could be assessed separately. For this purpose, the target syllables were divided into *vocoid* and *burst* portions at the nearest zero crossing in the waveform after the beginning of periodic pulses. Thus the burst portion of the syllable included aspiration, if there was any,[19] and the vocoid stimuli the voiced part of the glide and the vowel. Notice that the contents of the vocoid and the burst portions may vary greatly depending on the laryngeal features of the initial stop consonant, owing to the difference in the voice onset time. Since the VOT is close to 0 for fortis stops, the vocoid portion of a fortis stop-initial syllable may include most of the formant transition information. In contrast, in the lenis stop-initial syllables, the burst portion may include a considerable amount of information concerning the formant transition, due to the longer voice onset time. Therefore I focus on the fortis stop contexts in the perception experiment. IPA symbols /p/, /t/ and /k/ will be used to represent voiceless unaspirated (fortis) stops from here on, instead of the unconventional /p'/, /t'/ and /k'/.

Two representative tokens of each syllable type were chosen, e.g., $pa_1$ and $pa_2$, for the *pa* syllable type. The selection of tokens was based mainly on the duration of the signal: All 4 tokens of a CV syllable were first compared with the 4 tokens of its CwV counterpart, and two tokens of CV and two tokens of CwV syllables (4 in total) with the most comparable duration were selected, to reduce the effect of signal duration on the discrimination between CV and CwV. The pitch contour of the vocoid portion was

[19] Fant (1960, 1973) describes that stop burst as having three segments: transient, fricative and aspirative segments, which are difficult to separate from each other in the waveforms (Repp and Lin 1989:380). Thus, it seems natural to include the aspiration period in the burst portion rather than in the vocoid portion.

adjusted to minimize the pitch difference between CV and CwV tokens with the same consonant and vowel.

### 3.2.2 Participants

Since the hypothesis regarding the relationship between distributional asymmetry of Cw combinations and the CV-CwV contrast salience concerns "universal" tendencies observed in human languages rather than the phonotactics of specific languages, it was necessary to control transfer from the subjects' first language as much as possible. In other words, the first language of the participants of the experiment should not have an asymmetrical distribution of Cw combinations across consonant places. For example, monolingual English speakers would not be appropriate subjects for this study, since English has a gap in labial+/w/ clusters. Therefore, the subjects for this experiment were speakers of languages in which the Cw combinations are either represented at all three major consonant places (Spanish), or are completely lacking (Russian). Twenty-four Spanish speakers and eleven Russian speakers participated. All participants were undergraduate or graduate students at Stony Brook University. It was important that Russian participants' predominant language be Russian, but not English: the average age at which a Russian participant started living in the US was 16.4 (minimum 11), and all of them reported that they use Russian with their family and friends on a regular basis. For Spanish-speaking listeners English proficiency mattered less, as long as Spanish is one of their native languages. Nine of the Spanish-speaking participants were born in the US, but started speaking English only when they entered school. The other Spanish-speaking participants were from various Spanish-speaking countries such as Spain, Mexico, Dominican Republic, Columbia, Argentina, Peru, Ecuador, Bolivia, Guatemala and Puerto Rico. All of them reported using Spanish with their family and friends almost every day. No participants had had exposure to the Korean language.

### 3.2.3 Procedure

The experiment employed an AXB forced choice discrimination task model. The experiment consisted of two parts: the participants listened to vocoid stimuli in one part, and burst stimuli in the other part. In both parts, subjects listened to three consecutive demi-syllables with inter-stimulus interval of 1500 ms, and were asked to decide whether the second syllable (X) was similar to the first (A) or the third syllable (B). The relatively long inter-stimulus interval of 1500 ms was chosen to promote phonological (phonemic) rather than phonetic processing, following Werker and Logan (1985) and Brannen (2002).[20] There was a one-second interval between the button press and the beginning of the next trial.

As mentioned in 3.2.1, two tokens were used for a single stimulus type, e.g., $pa_1$ and $pa_2$ for the *pa* syllable and $p^wa_1$ and $p^wa_2$ for the $p^wa$ syllable. In each AXB trial, A and B stimuli were from syllables with the same phonation type, consonant place and vowel context, but one was a CV and the other was a $C^wV$ syllable. The target X stimulus was the same syllable as either A or B, but a physically different token. Mapping of

[20] Brannen (2002:22) states that "with a long duration between stimuli [(e.g., 1500 ms)], by the time a listener hears the next stimuli, the acoustic signal of the first stimulus has faded", especially "the non-distinctive phonetic features" of it.

syllables (CV or $C^wV$) to A and B positions was counterbalanced. There were 16 different combinations of AXB (4 X tokens * 2 A tokens * 2 B tokens) for each CV-$C^wV$ pair. The example set of trials testing *pa*-$p^wa$ discrimination is given in (1).

(1)

| A | X | B | correct answer |
|---|---|---|---|
| $p^wa_2$ | $p^wa_1$ | $pa_1$ | A |
| $p^wa_2$ | $p^wa_1$ | $pa_2$ | A |
| $pa_1$ | $p^wa_1$ | $p^wa_2$ | B |
| $pa_2$ | $p^wa_1$ | $p^wa_2$ | B |
| $p^wa_1$ | $p^wa_2$ | $pa_1$ | A |
| $p^wa_1$ | $p^wa_2$ | $pa_2$ | A |
| $pa_1$ | $p^wa_2$ | $p^wa_1$ | B |
| $pa_2$ | $p^wa_2$ | $p^wa_1$ | B |
| $pa_2$ | $pa_1$ | $p^wa_1$ | A |
| $pa_2$ | $pa_1$ | $p^wa_2$ | A |
| $p^wa_1$ | $pa_1$ | $pa_2$ | B |
| $p^wa_2$ | $pa_1$ | $pa_2$ | B |
| $pa_1$ | $pa_2$ | $p^wa_1$ | A |
| $pa_1$ | $pa_2$ | $p^wa_2$ | A |
| $p^wa_1$ | $pa_2$ | $pa_1$ | B |
| $p^wa_2$ | $pa_2$ | $pa_1$ | B |

In total there were 144 target trials (3 places * 3 vowels * 16 AXB combinations) and 72 filler trials for each part (vocoid and burst). The fillers consisted of CV tokens chosen from two different consonant places, e.g., $pa_1$ $ka_1$ $pa_2$.

Twenty-four (18 Spanish and 6 Russian) participants took the vocoid part of the experiment first, and eleven (6 Spanish and 5 Russian) participants started with the burst part. Participants were given a self-regulated break after every 40 trials, and a 5-minute break after the first part of the experiment was over. The entire experiment took approximately 50 minutes.

The experiment was run on SuperLab version 4.0, and presented using a MacBook OS X Version 10.5.2 through Koss R/80 headphones. The written instructions for the experiment were provided on the screen in the participants' native languages.[21]

### 3.3 Results

The relative perceptual salience of each CV-$C^wV$ contrast may be reflected in the participants' performance in the discrimination task, as well as the average time that the listeners take to make a decision on the similarity between stimuli within a trial. A negative correlation between the discrimination performance and the response time is expected, given the assumption that a subtler difference between sounds takes longer to

[21] Giving instruction in the native languages of the participants was considered to help them use their first language listening strategies (Narayan 2008:203). In the present study it may suppress the effect of English, which has bias against the C-Cw contrast in the labial context.

be perceived (e.g., Pisoni and Tash 1974). The results from each type of stimuli—*burst* and *vocoid*—were analyzed separately.

To assess the participants' performance in discrimination, both percent correct and d prime (d') value (MacMillan and Creelman 1991) were calculated for each CV-CwV contrast. Mixed model ANOVA results are reported only for d' below, though it was carried out on both percent correct and d', and the results exhibited the same patterns. Within-subject factors were *consonant place* and *vowel*, and between-subject factors were the *order* the stimuli were presented in (vocoid first or burst first), and the first *language* of the subject (Spanish or Russian).

As in the d' results, a mixed model ANOVA was carried out on the response time (RT). Only the correct responses were taken into account in RT measurement.

### 3.3.1 Vocoid stimuli

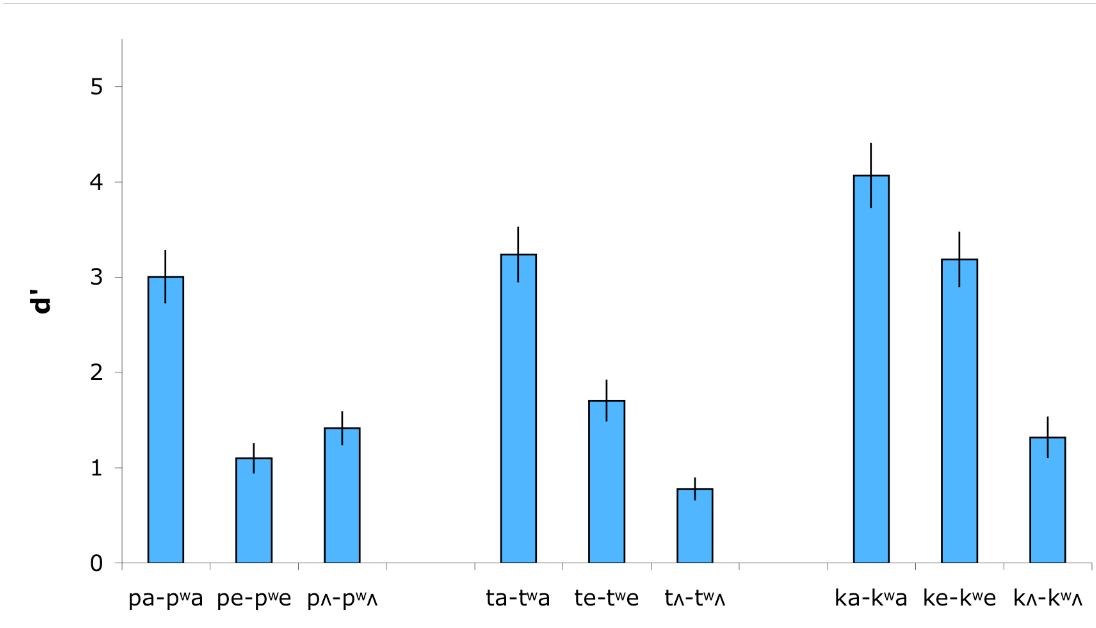
#### 3.3.1.1 d prime (d')

The main effects of *place* ($F(2,62)=14.225$, $p<0.001$) and *vowel* ($F(2,62)=40.969$, $p<0.001$) were significant. A Bonferroni post hoc test suggested that the d' was significantly higher in the velar context than in other consonant contexts ($p<0.001$). This patterns with the acoustic study results, in particular the F2 measurements. In vowel comparisons, all three vowel contexts were distinguished ($p<0.001$): /a/ showed the highest d', followed by /e/, and the d' of the /ʌ/ context was the lowest.

The fact that the d' was the highest in /a/ context is in parallel with the acoustic study results; recall that the F1 difference was the greatest in this context, and the F2 difference was bigger than /ʌ/, and no smaller than /e/ contexts (Table 2.6). Assuming that contrast perceptibility is determined by the combined effect of acoustic differences in F1 and F2, it seems natural that the Ca-$C^w$a contrast is easier to perceive than Ce-$C^w$e or Cʌ-$C^w$ʌ. The acoustic study results were in favor of the velar context as well: The CV-$C^w$V F2 difference was the greatest in the velar stop context, whereas F1 difference didn't differentiate consonant places (Table 2.7). Given that the contribution to the perception from F1 is the same across consonant places, the greater acoustic cue from F2 in the velar context may be responsible for the higher d' of the k-$k^w$ contrast.

The interaction of *place* and *vowel* was significant ($F(4,124)=5.038$, $p<0.001$), showing that the overall pattern described above was not always observed in the comparison within each consonant or vowel context. Figure 3.1 shows mean d' and standard error for each CV-CwV contrast.

Figure 3.1. Mean d’ (± standard error) of the vocoid stimuli.

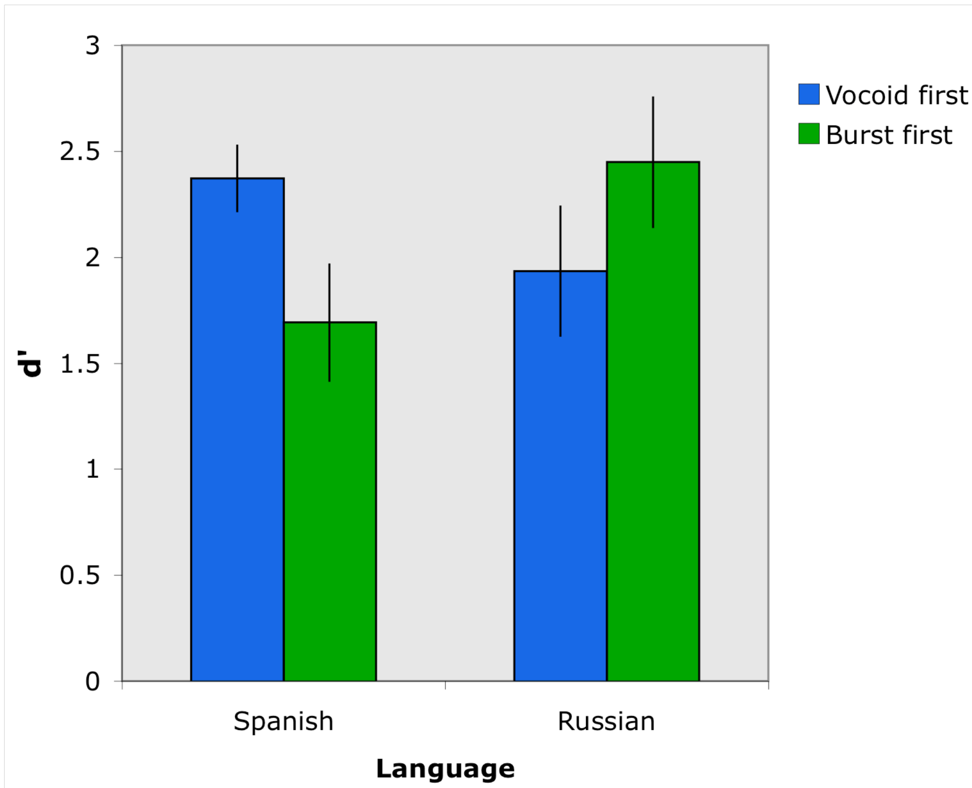



In the /e/ vowel context, d’ was higher for velar than for labial and coronal ($p<0.001$), matching the pattern found in the post-hoc comparisons of the *place* effect. Within the vowel /a/ context, the place effect was found only between velar and labial ($p<0.015$). Within the vowel /ʌ/ context, d’ was significantly lower in the coronal context than in the labial context ($p<0.01$). This is a somewhat unexpected result, since in the acoustic study no measurement pointed to an advantage of labial over coronal in the CV-CʷV contrast. A factor that seems to matter here is the direction of the F2 transition. F2 fell from the consonant release to the vowel peak in both /tʌ/ and /tʷʌ/ syllables, as indicated by the minus sign of F2 frequency change (Figure 2.17). In contrast, /pʌ/ and /pʷʌ/ were different in their F2 transition direction: F2 transition in /pʌ/ was more level than rising, with the frequency change close to 0, whereas it is clearly rising in /pʷʌ/. Thus, though the frequency measurements such as F2 onset and F2 change did not reveal any difference between the pʌ-pʷʌ and tʌ-tʷʌ contrasts, their perceptibility may vary.

In the coronal and velar place contexts, d’ was significantly higher for /a/ than for /e/, and higher for /e/ than for /ʌ/, just as in the post hoc test results of the main vowel effect. Within the labial context, however, there was no significant d’ difference between vowels /e/ and /ʌ/, and both were lower in d’ than /a/ ($p<0.001$). This may also be attributed to the effect of formant transition direction on the perception: Both /pe/ and /pʷe/ exhibited clearly rising F2 (Figure 2.15), unlike /pʌ/ and /pʷʌ/ which differed in the F2 transition direction, as described above. Due to the similarity in the transition direction of F2, the perceptual difference between /pe/ and /pʷe/ may not be greater than that between /pʌ/ and /pʷʌ/, despite the bigger difference in the frequency values of F2 change in the former.

There was no main effect of between-subject factors *language* and *order* on d’, and no significant interaction between them and within-subject factors *consonant* or *vowel*. However, the interaction between *language* and *order* was significant ($F(1,31)=6.417$, $p<0.017$). Interestingly, the Spanish and the Russian subjects showed opposite patterns vis-à-vis the order in which the stimuli are presented (Figure 3.2): the Spanish subjects did better on the vocoid stimuli when they were presented in the first part of the experiment (i.e., in vocoid-burst order), whereas the Russian subjects' performance was improved when the vocoid stimuli were presented in the second part.

Figure 3.2. Mean d' (± standard error) of the vocoid stimuli for the Spanish- and Russian-speaking listeners in vocoid-first and burst-first groups




It is not clear why this pattern occurs, but one possible explanation may come from the fact that familiarity with the stimuli differs between the Spanish- and Russian-speaking listeners, as summarized in (2): for Spanish-speaking listeners, the labialization contrast is "semi-native" in the sense that their L1 has Cw clusters. Between the two types of stimuli, the vocoid stimuli are less familiar to them, since the formant transition cue of /w/ is weaker than that in their native Cw combinations. Burst stimuli may be more native-like to them, as there was no acoustic difference between the Korean and the Spanish C-Cw contrasts in the burst frequencies (peak location or intensity difference; See 2.3.3). In contrast, for Russian listeners, Cw clusters or labialized consonants are non-native, and the vocoid and the burst stimuli are equally (un) familiar.

(2)

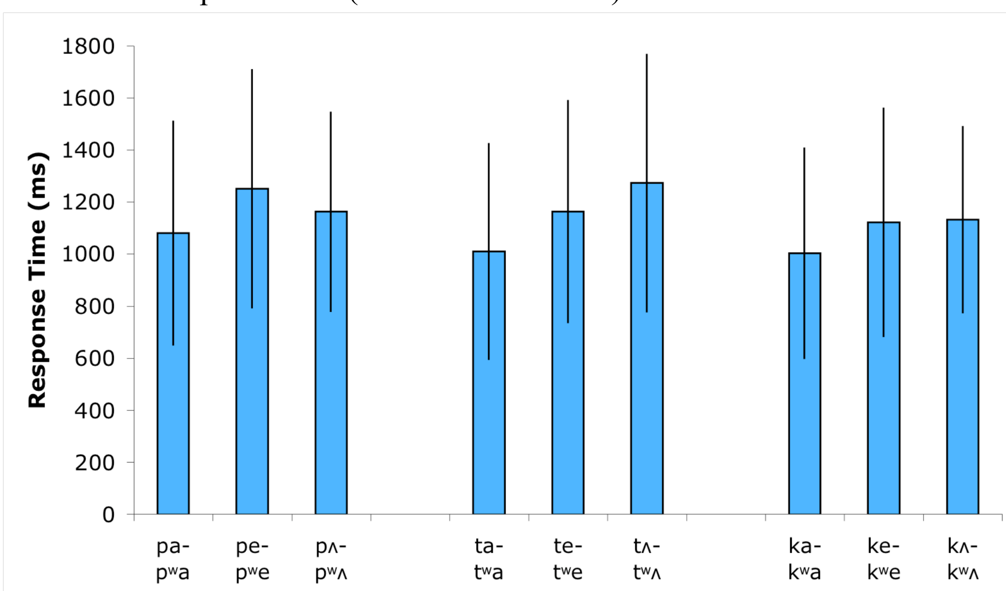
| **Order of stimuli presentation** | **Spanish** | **Russian** |
|---|---|---|
| vocoid-burst | semi-native → native | non-native → non-native |
| burst-vocoid | native → semi-native | non-native → non-native |

In the burst-vocoid order of stimulus presentation, the familiar burst stimuli in the first part might have set Spanish-speaking participants' perception for a native contrast of C and Cw, hindering their ability to hear the less familiar C-$C^w$ contrast of the vocoid stimuli in the second part. This is not expected for Russian speaking listeners; rather, the exposure to the C-$C^w$ burst noise contrast in the first part of the experiment may have helped facilitate the Russian listeners' perception in the second part. What is more important for the purpose of this study is that though the subjects' first language and the order of the two types of stimuli affected overall performance on vocoid stimuli, consonant place and vowel difference had a similar effect on the performance across the board. This is shown by the insignificance of interactions such as *language * place*, *language * vowel*, *order * place* or *order * vowel*.

3.3.1.2 Response Time (RT)

Average RT of the nine CV-$C^w$V contrasts from vocoid stimuli is given in Figure 3.3. Main effects of *place* ($F(2,62)=3.714$, $p<0.03$) and *vowel* ($F(2,62)=13.764$, $p<0.001$) were found. Post hoc test revealed that RT was significantly shorter in the velar context than in the labial context ($p<0.039$), and in the vowel /a/ context than /e/ and /ʌ/ ($p<0.001$).

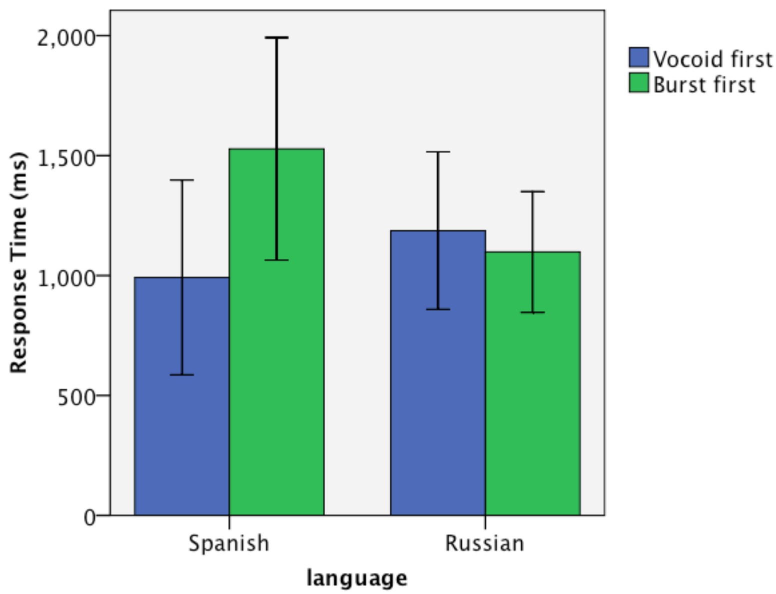
Figure 3.3. Mean response time (±standard deviation) of the vocoid stimuli.




Interaction of *place* and *vowel* ($F(3.568, 110.607)=3.534$, $p<0.009$) was significant, as in the d' measure, suggesting that the above stated main effects of place and vowel may not describe patterns found within a consonant or a vowel context. The overall results do not distinguish between /e/ and /ʌ/ vowel contexts, but it is obvious from Figure 3.3 that within the labial context, the pe-$p^w$e contrast was associated with the longest RT, whereas RT to the tʌ-$t^w$ʌ contrast was longer than to the te-$t^w$e contrast. Recall that it is also the /e/ vowel which elicited the lowest d' within the labial context. Thus, the relative disadvantage of the /e/ vowel in the labial context, which was attributed to the shared rising transition direction, is obvious in the RT as well as in the d' measure.

There was no main effect of *order* or *language*, but as with d', the interaction of *order* and *language* was significant (F(1,31)=5.666, p<0.024) (Figure 4). Spanish speaking listeners' RT was affected by the stimulus order, just as their d' was: Spanish subjects who listened to the burst stimuli first exhibited longer RT than the vocoid-first Spanish group. However, there was no such order effect among Russian speaking subjects. Again it is not clear why this pattern arose, and the different degrees of familiarity of the Spanish-speaking listeners with the vocoid and burst stimuli may be relevant, as I speculated for the *language* * *order* interaction in the d' of the vocoid stimuli. However, I dispute this idea in 3.3.3 below, for a reason made clear there.

Figure 3.4. Mean response time (±standard deviation) of the vocoid stimuli for Spanish- and Russian-speaking listeners in vocoid-first and burst-first groups.

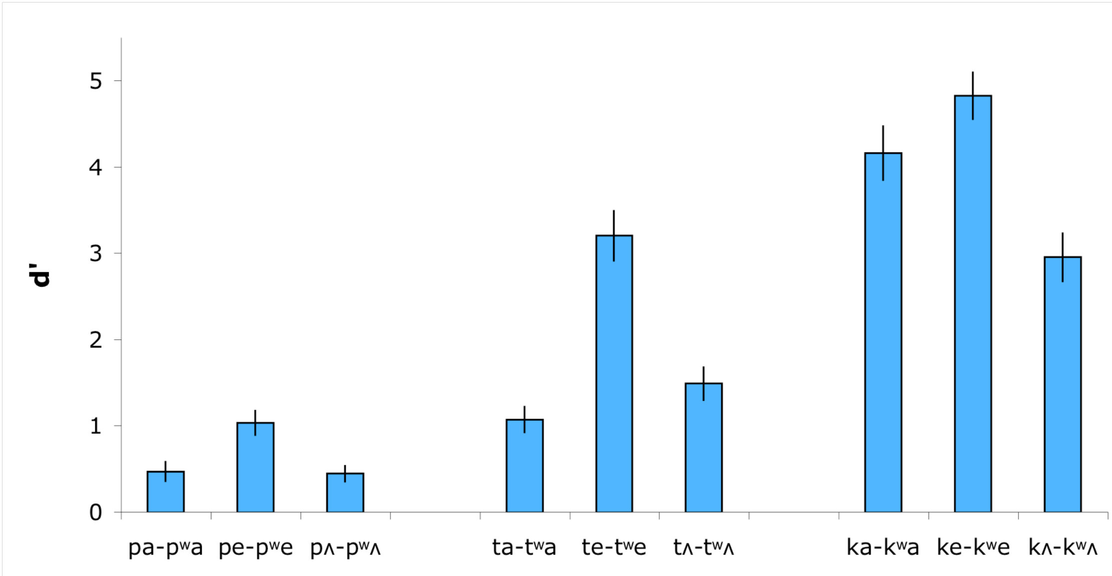



3.3.2 Burst stimuli

3.3.2.1 d prime

Average d' values of each CV-CwV contrast from the burst part of the experiment are shown in Figure 3.5. Repeated Measures ANOVA results revealed the main effects of *place* (F(2,62)=178.446, p<0.001) and *vowel* (F(2,62)=41.555, p<0.001). Bonferroni post-hoc test results suggested that the vowel /e/ invoked a significantly higher d' than the other vowels (p<0.001). Among consonant places, d' was the highest in K, followed by T, and the lowest in P contexts (p<0.001).

Figure 3.5. Mean d' (± standard error) of the burst stimuli.



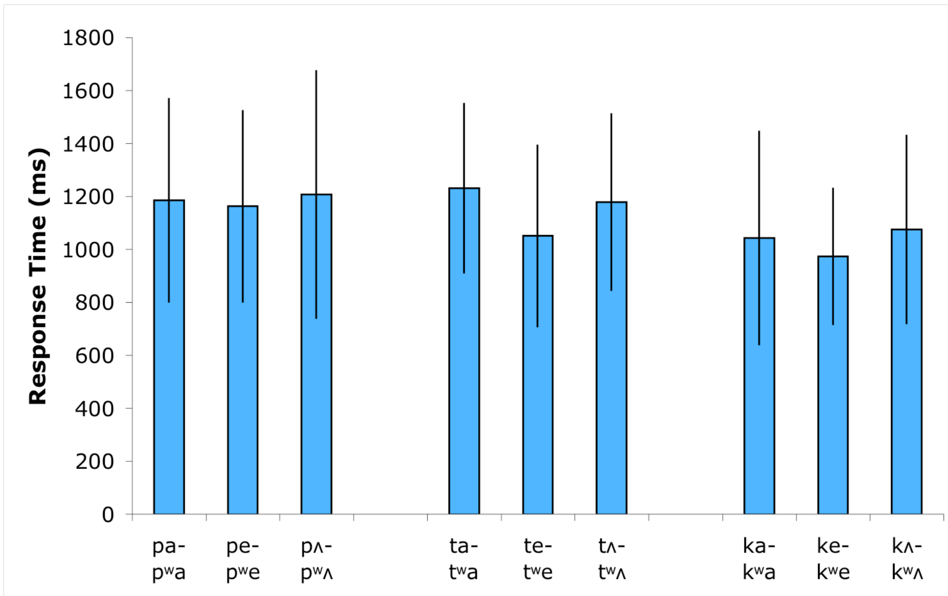

The vowel effect on the CV-$C^wV$ contrast salience was expected from the acoustic study, at least for the velar stop context; recall that the k-$k^w$ spectral peak location difference was greatest in the /e/ vowel context, and smallest in the /ʌ/ vowel context (Figure 2.26). The fact that a similar vowel effect (/e/ > /a/, /ʌ/) was also found in labial and coronal contexts is not expected, however, as no such effect on labial and coronal stop bursts was found in the acoustic study. This discrepancy may arise because the stimuli used in the burst spectral measurement in the acoustic study and the burst stimuli of the perception experiment do not exactly coincide. The burst stimuli of the perception experiment included the entire stop release burst (transient, frication and short aspiration). However, in the acoustic experiment the spectra were generated from a 12 ms window centered at the beginning of the burst, so the information from the later part, closer to the vowel, was not included.

Interaction of *place* * *vowel* was also significant ($F(4,124)=4.344$, $p<0.003$). This seems to be because /a/ and /ʌ/ vowels differed significantly in velar ($p<0.003$), but not in labial and coronal contexts. There was no effect of between-subject factors *language* or *order*, and their interaction was insignificant, unlike in vocoid stimuli.

3.3.2.2 Response Time (RT)

As with the vocoid stimuli, the main effects of *place* ($F(1.655, 51.302)=7.639$, $p<0.002$) and *vowel* ($F(2,62)=5.287$, $p<0.008$) were found in the response time when participants listened to stimuli containing bursts. According to Bonferroni post hoc test, RT was significantly shorter for velar than for labial ($p<0.011$) and for coronal ($p<0.001$). Among vowel contexts, /e/ conditioned a shorter RT than /a/ ($p<0.021$) and /ʌ/ ($p<0.013$). Notice that these consonant place and vowel effects match those observed in the d' measurement of the burst stimuli: the d' was higher for the velar context and /e/ vowel context. There was no significant interaction between *place* and *vowel*. Figure 3.6 shows the mean RT of the nine CV-$C^wV$ contrasts.

Figure 3.6. Mean response time (± standard deviation) of the burst stimuli.




3.3.3 Summary and discussion

Overall, the place and vowel asymmetries found in the acoustic study results were reflected in the d prime value of the perception experiment results. In particular, the d' from the vocoid stimuli matched the results of the formant transition measures in the fortis stop context (3.3.1.1). The d' was the highest in the velar stop context, in which the acoustic difference between CV and $C^wV$, i.e., formant transition frequency change or direction, was greater than in other consonant place contexts. Recall from Chapter 2 that in the vocoid formant measurements F1 and F2 favored different vowels: the F1 measure suggested the vowel /a/ as the provider of the most salient contrast between CV and $C^wV$, whereas F2 was indecisive between /a/ and /e/ with respect to the salience of the contrast. Taken together this may suggest that the Ca-$C^w$a difference is bigger than Ce-$C^w$e in Korean. The results of the perception experiment confirmed this speculation: the d' was the highest in /a/, followed by /e/, which in turn was higher than /ʌ/.

Other contextual effects observed in the d' may not be directly related to the acoustic study results. For example, vowel context did not influence acoustic measures such as the intensity difference between two frequency regions in the coronal or labial stop burst, but the d' of the burst stimuli, containing stop bursts, was significantly higher in /e/ context than in /a/ or /ʌ/ (3.3.2.1). This was attributed to the different nature of the stimuli in the acoustic and perception experiments. In the acoustic experiment, the measurements were taken from only the first 6 ms of the burst, whereas the perception experiment stimuli included the entire burst noise, including a short aspiration period. A pattern somewhat similar to this, i.e., confusability between labial or coronal stop bursts in /a/ vowel context and in lip rounding context, is reported in Bonneau (2000: 499), who found that when the burst noise from the French /pu/ and /tu/ was presented, listeners wrongly identified the following vowel as /a/ for 32% and 51%, but as /i/ for only 1% and 10 %, respectively. In Cullian and Tekieli's (1979) experiment on vowel identification from the stop burst, the backness of the vowel was more easily identified than vowel height. Thus, /p/ and /t/ bursts may vary only slightly between /a/ and /u/ (/w/) contexts, which differ in tongue height, whereas the burst quality may be more distinctive between front vowel contexts (/i/ or /e/) and back rounded vowel contexts.

Recall that in the acoustic study different spectral measures were used to capture the energy distribution shift by /w/ in velar stop burst and dental stop burst spectra, so it was impossible to directly compare velar and dental stops with respect to the strength of the acoustic effect from /w/ on their burst properties. Based on the results from the portion of the experiment involving burst stimuli, it is now possible to make a direct comparison between coronal and velar stops in the release burst: the burst provides the highest perceptibility of the labialization contrast in the velar stop context.

The interaction between listeners' first language and the order of stimulus type was significant in the d' of the vocoid part (3.3.1.1), which may be attributed to the existence of a similar contrast (C-Cw cluster) in Spanish. Though the overall performance may have been influenced by the listeners' first language and the experiment order, there was no interaction between *place* or *vowel* on the one hand, and *language* or *order* on the other. This shows that the asymmetries in the perceptibility of the CV-C$^w$V contrasts related to the consonant place and vowel are not language-specific, or by-products of prior experience in the experiment, but originate from the acoustic asymmetries.

The response time measure for both vocoid (3.3.1.2) and burst (3.3.2.2) stimuli shows that the perceptibility of the labialization contrast is not symmetrical, a similar conclusion to the one from the d' measure. RT was shorter in the velar context than in the labial context (vocoid), or in both labial and coronal stop contexts (burst), indicating that the CV-C$^w$V difference was relatively easier to hear in the velar context. In vowel /ʌ/ contexts the contrast was more difficult to hear than in /e/ (burst) or /a/ (vocoid) contexts. These results show that the participants' performance in discrimination and their response time were inversely correlated ($R^2$=0.8036 for vocoid; $R^2$=0.9196 for burst): the participants tended to take longer to respond when they listened to the CV-C$^w$V contrasts with a lower d'.

The order effect found in the Spanish-speaking subjects' response time when they listened to vocoid stimuli is most likely to reflect simple individual differences. The 6 Spanish subjects who took the experiment in burst-vocoid order showed longer average RT than the 18 vocoid-first Spanish subjects in both parts. The interaction of *language * order* was significant only in the RT of the vocoid trials, but even in the burst trials the effect was at least marginally significant (p=0.053). This was not the case for d' measure, in which the *language * order* effect was restricted to the vocoid stimuli: Spanish burst-first group's d' in the vocoid stimuli was lower than vocoid-first group's, but the two groups' d' in the burst stimuli did not differ (p=0.765).

### 3.4 Relative strength of formant transition and burst frequency cues to the CV-CwV contrast

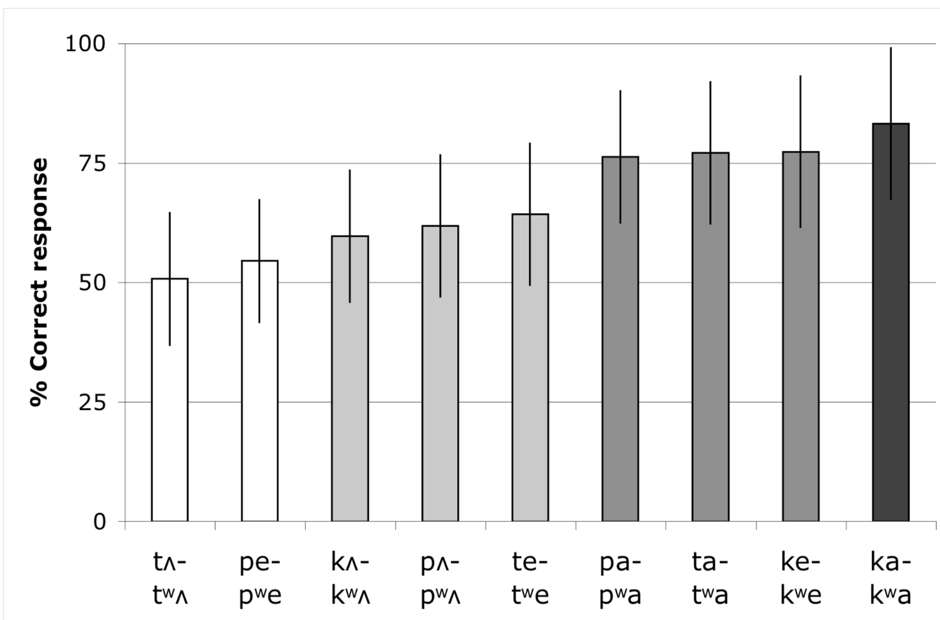
Let us now consider how the results from the perception study may contribute to the analysis of distributional asymmetries in Cw combinations within Dispersion Theory.

Recall that in DT analyses, legitimacy of a contrast in a language is judged by minimal distance requirements between two contrasting items in each auditory/perceptual dimension, e.g., formant transition or burst frequency. Thus I will establish scales on which the size of the perceptual cue for each CV-CwV contrast is compared, and the minimal distance requirement is stated. My assumption is that the perceptibility of the difference between each CV and CwV in the vocoid and burst portions of the syllable reflects the salience of the formant transition cue and the burst noise frequency cue for that particular CV-CwV contrast, respectively.

3.4.1 Formant transition cues for the labialization contrast

The formant transition cues in different CV-CwV contrasts are compared based on the vocoid stimulus discrimination performance. In Figure 3.7, the CV-CwV contrasts at the nine different environments (3 places * 3 vowels) are ordered from the contrast with the lowest percent correct to the one with the highest percent correct.[22]

Figure 3.7. Correct response rate (± standard deviation) of the vocoid stimuli for each CV-C$^{w}$V contrast. Shades of the bars group the contrasts according to the t-test results.




One-sample t-test (two-tailed) revealed that the correct response rate of pe-p$^{w}$e and tʌ-t$^{w}$ʌ were not different from chance level (50%). Excluding these two, a paired-sample t-test was done for each CV-CwV contrast. According to the results, the 7 consonant + vowel combinations were divided into three groups, differentiated by the shades of the columns

[22] The percent correct measure, rather than d', was used for establishing the perceptibility scales. This was mainly to sort out the contrasts for which the performance was not different from chance level (50%), which are assigned 0 value on the scales. T-tests on the d' also result in the grouping identical to that based on the percent correct, except that the difference between ke-k$^{w}$e and ka-k$^{w}$a was only marginally significant (p=0.06) for burst stimuli.

in Figure 3.7: The CV-CwV contrasts within a group did not significantly differ in their correct response rate, whereas those that belong to different groups were distinguished by the t-test results (Table 3.1): In other words, the differences between te-t$^w$e and pa-p$^w$a, and ke-k$^w$e and ka-k$^w$a were significant, but the correct response rate did not vary significantly within the group of the pa-p$^w$a, ta-t$^w$a and ke-k$^w$e contrasts, and the group of kʌ-k$^w$ʌ, pʌ-p$^w$ʌ and te-t$^w$e.

Table 3.1. Paired-sample t-test results of the correct response rate of the vocoid stimuli.

| | t | df | Sig. (2-tailed) |
|---|---|---|---|
| kʌ-k$^w$ʌ and pʌ-p$^w$ʌ | -0.96 | 34 | 0.344 |
| pʌ-p$^w$ʌ and te-t$^w$e | -0.807 | 34 | 0.425 |
| kʌ-k$^w$ʌ and te-t$^w$e | 1.606 | 34 | 0.118 |
| te-t$^w$e and pa-p$^w$a | -4.465 | 34 | **0.001** |
| pa-p$^w$a and ta-t$^w$a | -0.343 | 34 | 0.734 |
| ta-t$^w$a and ke-k$^w$e | -0.082 | 34 | 0.935 |
| pa-p$^w$a and ke-k$^w$e | -0.435 | 34 | 0.666 |
| ke-k$^w$e and ka-k$^w$a | -2.541 | 34 | **0.016** |

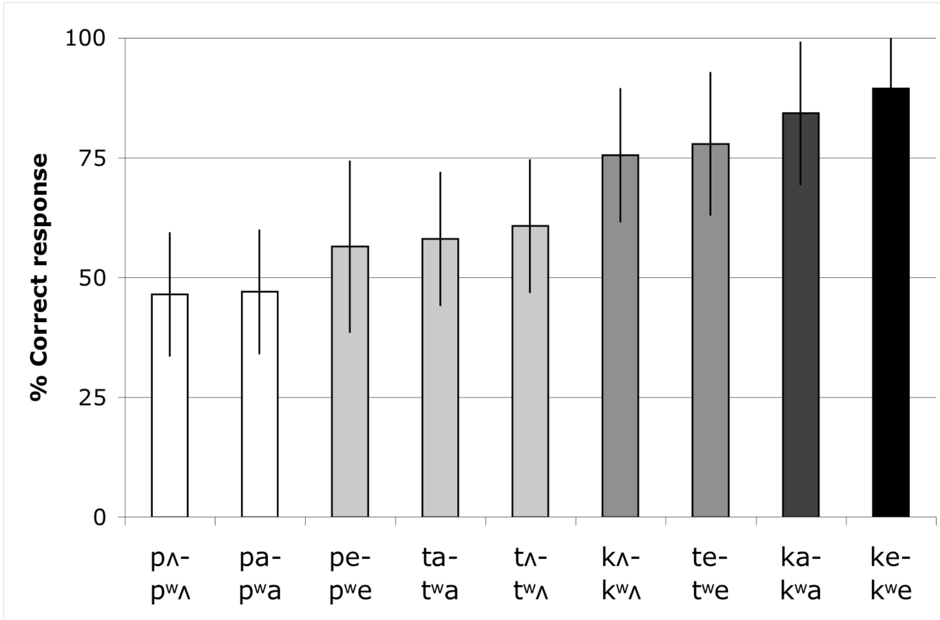
The strength of the formant transition cues for /w/ in each context is schematized as the perceptual distance between CV and CwV in (3):

| | | | | |
|---|---|---|---|---|
| (3) | ka | | | k$^w$a |
| | ke | | k$^w$e | |
| | ta | | t$^w$a | |
| | pa | | p$^w$a | |
| | te | t$^w$e | | |
| | pʌ | p$^w$ʌ | | |
| | kʌ | k$^w$ʌ | | |
| | pe p$^w$e | | | |
| | tʌ t$^w$ʌ | | | |

3.4.2 Burst noise frequency cue for the labialization contrast

The salience of the stop burst noise frequency cue is assessed based on the performance in the burst stimulus part of the experiment, in a similar way with the formant transition scale. Figure 3.8 shows the nine CV-CwV combinations ordered by their correct response rate from the burst stimuli part.

Figure 3.8. Correct response rate (± standard deviation) of the burst stimuli for each CV-$C^wV$ contrast. Shades of the bars group the contrasts according to the t-test results.




The correct response rate of pa-pʷa and pʌ-pʷʌ was not distinguished from 50% according to the one-sample t-test. By their average correct response rate, the seven remaining CV-CwV contrasts were divided into 4 groups, as shown by the different shades of bars in Figure 3.8. Again, the CV-$C^wV$ contrasts in different groups are distinguished by paired sample t-test ($p<0.05$), but the ones within the same group are not. See Table 3.2 for the details.

Table 3.2. Paired-sample t-test results of the correct response rate of the burst stimuli.

| | t | df | Sig. (2-tailed) |
|---|---|---|---|
| pe-pʷe and ta-tʷa | -0.465 | 34 | 0.645 |
| ta-tʷa and tʌ-tʷʌ | -0.861 | 34 | 0.395 |
| pe-pʷe and tʌ-tʷʌ | -0.234 | 34 | 0.226 |
| tʌ-tʷʌ and kʌ-kʷʌ | -4.923 | 34 | **0.001** |
| kʌ-kʷʌ and te-tʷe | -0.944 | 34 | 0.352 |
| te-tʷe and ka-kʷa | -2.698 | 34 | **0.011** |
| ka-kʷa and ke-kʷe | -2.62 | 34 | **0.013** |

Based on this grouping the relative size of the burst noise frequency cue is schematized by the distance between each CV and CwV in (4).

(4) ke $k^{w}$e
ka $k^{w}$a
te $t^{w}$e
kʌ $k^{w}$ʌ
tʌ $t^{w}$ʌ
ta $t^{w}$a
pe $p^{w}$e
pa $p^{w}$a
pʌ $p^{w}$ʌ

The relative distance between CV and CwV in formant transition and burst noise frequency, shown in (3) and (4), will be used as the basis for calculating the perceptual distinctiveness between plain and labialized consonants in the analysis of the Cw distributional asymmetries in Chapter 4.

## Chapter 4. A Dispersion-Theoretic analysis of the distribution of Cw combinations

This chapter presents an analysis of the distributional asymmetries in Cw combinations, developed within a theory that incorporates the functional goals of speech perception and production into co-occurrence restrictions, namely the Dispersion Theory of Contrast (DT; Flemming 1995, 2002). The asymmetries found in the distributions of Cw combinations are attributed to the interaction of functional goals such as a minimal distance requirement between contrasting items, maintenance of contrasts, and articulatory gesture coordination. Relative strength of the cues to the CV-CwV contrasts in various C and V contexts is assessed along two auditory-featural scales, namely Formant Transition and Noise Frequency. The scales are developed based on the Korean CV-CwV perception test results from Chapter 3, and it is shown that they can be effectively applied to other languages as well. Section 4.1 provides a DT analysis for Korean, Mandarin and Cantonese. This analysis is compared with an alternative feature identity avoidance account in 4.2. Section 4.3 extends the DT analysis to the Dan language, in which both labialized consonants and Cw clusters exist, but in different consonant contexts. Section 4.4 extends the current proposal, introducing distributional asymmetries found in Cj combinations (palatalized consonants and Cj clusters), and proposing a DT analysis. In 4.5, a seemingly problematic case for the current analysis on labialization is presented, namely Ponapean, in which only labial consonants co-occur with a labiovelar glide, and a solution is suggested. Section 4.6 shows that the scale of auditory features such as noise frequency plays a role in the grammar beyond the evaluation of contrasts involving labialization, acting also in output-output correspondence relationships such as arise with loanword adaptation.

### 4.1 Contrast dispersion and the asymmetries in labialization

The basics of the Dispersion Theory of Contrast were introduced in 1.3. Formant transition and noise frequency scales of labialization contrast are developed in 4.1.1, based on the acoustic and perception studies of the Korean CV-CwV contrasts from the previous chapters. These scales will be used for the analysis of co-occurrence restrictions of Mandarin (4.1.3) and Cantonese (4.1.4) as well as Korean (4.1.2). Since Mandarin and Cantonese have different vowel inventories from Korean, I will make inferences about the relative distance between contrasts on the scales based on the Korean data. This is possible because the relationship between vowel articulatory properties such as backness and height and the perceptual distance between CV and CwV syllables is rather straightforward, as discussed below.

### 4.1.1 Perceptibility scales of labialization contrast

In the original proposal of DT (Flemming 1995, 2002), auditory scales are posited for each dimension (F1 Transition, F2 Transition, Noise Frequency, etc.), and the perceptual difference between two sound strings is calculated from the number of steps that separate the relevant sounds on each scale. Padgett (2003) proposes an alternative distinctiveness assessment mechanism within Dispersion Theory, which does not directly refer to auditory features such as F1 or F2 transition. Instead, contrast between vowels is represented on a more abstract dimension of "vowel color" (backness and roundness), with /u/ and /i/ occupying the two extremes of the dimension. Thus vowel color encompasses various acoustic dimensions such as F1, F2 and F3. I will adapt Padgett's strategy of integrating formant information into a single perceptual dimension. In (1) I present a (labialization) Formant Transition scale, which represents the strength of the perceptual cues to the labialization from the vocoid portion of a C(w)V syllable, instead of two separate scales of F1 and F2 transitions. This integrated scale incorporates the durational aspects of the formant transition as well as the spectral aspects. Recall from the perception study results that the nine CV and CwV pairs were divided into 4 groups based on the percentage of correct responses for the vocalic portion (Figure 3.7). These four steps of perceptual distinctiveness are used here to indicate the distance between each CV and the corresponding CwV, and are assigned the values of 3, 2, 1 or 0. The value 0 is assigned to pe-$p^w$e and tʌ-$t^w$ʌ contrasts because their correct response rate was not different from chance level in the perception test, while the contrast between /ka/ and /$k^w$a/ is assigned 3 because they were the most distinct.

(1)

| CV-CwV distance scale for Formant Transition | | | | |
|---|---|---|---|---|
| 3 | ka | | | $k^w$a |
| 2 | ke<br>ta<br>pa | | $k^w$e<br>$t^w$a<br>$p^w$a | |
| 1 | te<br>pʌ<br>kʌ | $t^w$e<br>$p^w$ʌ<br>$k^w$ʌ | | |
| 0 | pe $p^w$e<br>tʌ $t^w$ʌ | | | |

Although the perceptual cues from the vocoid portion of the syllables are integrated into a single scale, a separate scale is postulated for the stop release burst noise frequency. Based on the correct response rate from the burst stimuli, the nine CV-CwV contrasts were grouped into 5 steps of perceptibility (Figure 3.8). Here, each group is assigned a perceptual distance of 0, 1, 2, 3 and 4, as shown in (2). The discrimination rate of pa and $p^w$a, and that of pʌ and *$p^w$ʌ* were not above chance level, hence the value 0 on the Noise Frequency scale.

(2)

| CV-CwV distance scale for Noise Frequency | | | | | |
|---|---|---|---|---|---|
| 4 | ke | | | | $k^w e$ |
| 3 | ka | | | $k^w a$ | |
| 2 | kʌ<br>te | | $k^w ʌ$<br>$t^w e$ | | |
| 1 | ta<br>tʌ<br>pe | $t^w a$<br>$t^w ʌ$<br>$p^w e$ | | | |
| 0 | pa $p^w a$<br>pʌ $p^w ʌ$ | | | | |

The overall distinctiveness of CV and CwV is expressed by the number of steps they differ on each scale, as in (3).

(3) $\Delta$(ka-$k^w$a) = Formant Transition: 3, Noise Frequency: 3
$\Delta$(ta-$t^w$a) = Formant Transition: 2, Noise Frequency: 1
$\Delta$(pa-$p^w$a) = Formant Transition: 2, Noise Frequency: 0
$\Delta$(ke-$k^w$e) = Formant Transition: 2, Noise Frequency: 4
$\Delta$(te-$t^w$e) = Formant Transition: 1, Noise Frequency: 2
$\Delta$(pe-$p^w$e) = Formant Transition: 0, Noise Frequency: 1
$\Delta$(kʌ-$k^w$ʌ) = Formant Transition: 1, Noise Frequency: 2
$\Delta$(tʌ-$t^w$ʌ) = Formant Transition: 0, Noise Frequency: 1
$\Delta$(pʌ-$p^w$ʌ) = Formant Transition: 1, Noise Frequency: 0

Having established the relative distinctiveness of the labialization contrast for each consonant place and vowel context, I now turn to the DT analysis of the asymmetrical distribution of the labialization contrast in specific languages. For each language, the relative distinctiveness of a contrast can be evaluated by considering a two-dimensional space defined by the Formant Transition and Noise Frequency.

4.1.2 Korean

In (4), the labialization contrast in Korean is represented along the two dimensions of perceptual distinctiveness—Formant Transition (FT) and Noise Frequency (NF), respectively. The number on each scale indicates the relative degree of distinctiveness between a particular CV syllable and its $C^w$V counterpart in that dimension. Any potential contrasts that are banned by the Korean morpheme structure conditions (pe-$p^w$e, pa-$p^w$a, pʌ-$p^w$ʌ, and tʌ-$t^w$ʌ) are in the shaded area and in smaller fonts.

(4) The CV-CwV contrast distinctiveness in Korean

| | | Formant Transition (FT) scale | | | |
|---|---|---|---|---|---|
| | | 0 | 1 | 2 | 3 |
| Noise Frequency (NF) scale | 0 | | **Δ(pʌ-p$^{w}$ʌ)** | **Δ(pa-p$^{w}$a)** | |
| | 1 | **Δ(pe-p$^{w}$e)**<br>**Δ(tʌ-t$^{w}$ʌ)** | | **Δ(ta-t$^{w}$a)** | |
| | 2 | | **Δ(kʌ-k$^{w}$ʌ)**<br>**Δ(te-t$^{w}$e)** | | |
| | 3 | | | | **Δ(ka-k$^{w}$a)** |
| | 4 | | | **Δ(ke-k$^{w}$e)** | |

The scale in (4) shows that the contrasts that are allowed in Korean morphemes have at least 2 NF difference, or 1 NF difference and 2 FT difference. This requirement of perceptual distance is expressed by the Minimal Distance constraint such as (5):

(5) MINIMALDISTANCE= {NF:1 and FT:2} or {NF:2 and FT:1}

Two contrasting items differ at least by one step on the FT scale and two steps on the NF scale, or by one step on the NF and two steps on the FT scales.

As noted in 1.3, an important notion in DT is that the Minimal Distance constraints that require a different degree of distance are in a strict dominance relationship. A constraint that requires a greater distance than {NF:2 and FT:1} or {NF:1 and FT:2}, *e.g.,* (6a), must be ranked lower than (5), as in (6b). Note that constraint (6a) may only be satisfied by labialized velar consonants before vowels /a/ and /e/, which are the only contexts in which the contrastive labialization occurs in languages like Cantonese (See 4.1.4):

(6) a. MINDIST={NF:3 and NF:2} or {NF:2 and FT:3}
b. MINDIST= {NF:2 and FT:1} or {NF:1 and FT:2}
>> MINDIST= {NF:3 and FT:2} or {NF:2 and FT:3}

Minimal Distance constraints interact with a constraint that requires contrasts in the input to be maintained in the output (7).

(7) MAXIMIZECONTRASTS

Maximize the number of contrasting items.
(i.e., Preserve input contrasts in the output.)

In Korean, MinDist= {NF:2 & FT:1} or {NF:1 & FT:2} dominates MAXIMIZECONTRASTS as in (8), so only the CV-CwV contrasts that satisfy the former surfaces.

(8) MINDIST= {NF:2 & FT:1} or {NF:1 & FT:2} >> MAXIMIZECONTRASTS
>> MINDIST={NF:3 & FT:2} or {NF:2 & FT:3}

Tableaux (9) and (10) demonstrate how the ranking in (8) gives rise to the Cw co-occurrence restrictions in Korean. The contrast between /te/ and /tʷe/ exists in Korean, as candidate (9a) satisfies the high ranking Minimal Distance constraint (5), though it violates lower ranked MinDist={NF:3 & FT:2} or {NF:2 & FT:3}. Candidate (9b) is ruled out due to the violation of MAXIMIZECONTRASTS.

(9)

| te tʷe | MinDist={NF:2&FT:1} or {NF:1 & FT:2} | MAXIMIZE CONTRASTS | MinDist={NF:3&FT:2} or {NF:2&FT:3} |
|---|---|---|---|
| a. ☞te – tʷe | | | * |
| b. te | | *! | |

Unlike /te/ and /tʷe/, the contrast between /pe/ and /pʷe/ does not exist in Korean. Candidate (10a), which preserves the contrast between /pe/ and */pʷe/, is ruled out since it violates the high-ranking Minimal Distance constraint. Although candidate (10b), which neutralizes the contrast between /pe/ and */pʷe/, violates MAXIMIZECONTRASTS, it is chosen as the optimal output.

(10)

| pe pʷe | MinDist={NF:2&FT:1} or {NF:1&FT:2} | MAXIMIZE CONTRASTS | MinDist={NF:3&FT:2} or {NF:2&FT:3} |
|---|---|---|---|
| a. pe - pʷe | *! | | * |
| b. ☞pe | | * | |

The constraint ranking correctly chooses to preserve the Cʌ-Cwʌ contrast when C is a velar stop (11), and to rule it out when C is a coronal stop (12). The kʌ-kʷʌ contrast satisfies the requirement of NF:2 and FT:1. In contrast, when C is a coronal stop, (5) is violated as the contrast tʌ-tʷʌ satisfies neither NF:2 & FT:1 nor NF:1 & FT:2. Thus the contrast between tʌ and tʷʌ does not surface (12).

(11)

| kʌ kʷʌ | MinDist={NF:2&FT:1} or {NF:1&FT:2} | MAXIMIZE CONTRASTS | MinDist={NF:3&FT:2} or {NF:2&FT:3} |
|---|---|---|---|
| a. ☞kʌ - kʷʌ | | | * |
| b. kʌ | | *! | |

(12)

| tʌ tʷʌ | MinDist={NF:2&FT:1} or {NF:1&FT:2} | MAXIMIZE CONTRASTS | MinDist={NF:3&FT:2} or {NF:2&FT:3} | *C$^G$ |
|---|---|---|---|---|
| a. tʌ- tʷʌ | *! | | * | * |
| b. ☞tʌ | | * | | |
| c. tʷʌ | | * | | *! |

Candidate (12c) also satisfies the minimal distance requirement, but it violates an additional constraint against secondary articulation, $*C^G$. This may be understood as a kind of articulatory effort minimization constraint within DT, in the sense that the secondary articulation involves additional articulator movement, in this case lip protrusion and tongue back raising. More discussion of this constraint is given in 4.2.

Before I close this section, it is necessary to recall that the /i/ vowel context, in which the C-Cw contrast is allowed for coronals and velars in Korean, is excluded from the experiments as well as the discussion so far. This was mainly in order to avoid the influence of palatalization in the Korean /i/ vowel context. I can only infer from the closeness between /e/ and /i/ in terms of the tongue backness and height that the distance between Ci and Cwi is also similar to that between Ce and Cwe.

To summarize, the MSCs regarding the Korean Cw combinations follow from the consideration of their perceptual contrast from plain consonants. A certain degree of perceptual distance between CV and CwV is required in order for these two types of strings to be contrastive: they must differ substantially in the frequency of the consonant release noise, and/or in the formant transition of the vocoid portion. The DT analysis captures this with the perceptual distance space (4) and the constraint ranking in (8).

### 4.1.3 Mandarin

As mentioned in 2.2.1, the combinations of a consonant and a glide in Mandarin are argued to be consonants with secondary labialization or palatalization, rather than clusters (Duanmu 1990, 2000; Li 1999). Mandarin shows an asymmetric distribution of labialized consonants similar to Korean. Labialized velar stops are attested before vowels [ɑ], [a], [ə] and [eɪ], but the contrast between plain and labialized dental stops exhibits a more restricted distribution, as the low back vowel [ɑ] does not co-occur with labialized dental stops. Labialized labial consonants, which do not occur before other vowels, occur before the mid back rounded vowel /ɔ/. In this particular context, however, plain consonants are not allowed, which means that there is no contrast between plain and labialized consonants before /ɔ/.

In fact there is another case in Mandarin in which the labialization contrast is neutralized to labialized consonant, rather than a plain consonant. Labialized dental stops exist before [ən] ($t^{hw}$ən, $t^{w}$ən; Pinyin *tun* and *dun*), but plain dental stops are absent in this particular context (*$t^{h}$ən, *tən; Pinyin **ten*, **den*).[23] In contrast, both plain and labialized velar stops occur before [ən] (Pinyin *ken-kun, gen-gun*). Thus the [ən] context is similar to the [ɑŋ] context in that they host the labialization contrast on velar stops, but not on dental stops. (13) summarizes the distribution of plain and labialized stops in Mandarin with examples in Pinyin (in italics). Numbers next to the Pinyin transcription indicate the tone.

[23] Dental stop + [ən] syllables are absent in Taiwanese Mandarin and in certain dialects of Mandarin spoken in Mainland China (e.g., Shanghai dialect), but are present in Beijing dialect (e.g. tən; Pinyin *den4* ‘yank’).

| (13) | Labial | Coronal | Velar |
|---|---|---|---|
| eɪ | peɪ *bei1* ‘sorrow’ | teɪ *dei1* ‘should’<br>tʷeɪ *dui1* ‘pile’ | keɪ *gei3* ‘reprimand’<br>kʷeɪ *gui3* ‘ghost’ |
| ə | pən *ben3* ‘root’ | tʷən *dun4* ‘shield’ | kən *gen3* ‘willing’<br>kʷən *gun3* ‘roll’ |
| a | pan *ban1* ‘class’ | tan *dan3* ‘gall bladder’<br>tʷan *duan3* ‘short’ | kan *gan1* ‘sweetness’<br>kʷan *guan1* ‘official’ |
| ɑ | pɑŋ *bang1* ‘defend’ | tɑŋ *dang4* ‘swing’ | kɑŋ *kang1* ‘hard’<br>kʷɑŋ *kuang1* ‘light’ |
| ɔ | pʷɔ *bo1* ‘wave’ | tʷɔ *duo1* ‘many’ | kʷɔ *guo3* ‘fruit’ |

As in Korean, coronal stop labialization is more restricted in terms of following vowels than velar labialization, as the t-tʷ contrast is not found before [ɑ] and [ə].

To apply the DT analysis to the Mandarin Cw distribution, we must locate the low back vowel [ɑ] and mid central vowel [ə] in the auditory distinctiveness scale. The tongue position for [ɑ] is backer than [a], and thus [ɑ] exhibits lower F2 than [a] (Chen 2000, Li 2006), so the F2 transition difference between Cɑ and Cʷɑ is expected to be smaller than that between Ca and Cʷa. This may result in a perceptibility difference of the C-Cw contrast in [a] and [ɑ] contexts on the Formant Transition scale, similar to that between Korean [a] and [ʌ]. But the Cɑ-Cʷɑ contrast in FT dimension may not be as small as the Korean Cʌ-Cʷʌ, because of the contribution from the F1 difference. Therefore I will posit that the Mandarin Cɑ-Cʷɑ contrast falls between Ca-Cʷa and Cʌ-Cʷʌ: i.e., Δ(tɑ-tʷɑ)=FT:1 and Δ(kɑ-kʷɑ)=FT: 2 on the FT scale, as in (14). The central vowel [ə] is fronter than [ʌ], so the distance between tə and tʷə as well as that between kə and kʷə are also assigned values greater than the Korean tʌ-tʷʌ and kʌ-kʷʌ distances on the Formant Transition scale.

Unlike the FT distance, the distance on the Noise Frequency scale might differ between the two allophonic contexts [a] and [ɑ] only when the consonant is velar. Again I am making an inference from the [a] vs. [ʌ] comparison: NF distance was 1 for both ta-tʷa and tʌ-tʷʌ contrasts, whereas it was 3 for ka-kʷa and 2 for kʌ-kʷʌ, reflecting the relationship between vowel backness and the velar burst noise frequency cue (See (4)). Therefore I postulate that the NF distance for tɑ-tʷɑ and tə-tʷə contrasts is 1, just as that for ta-tʷa. For kɑ-kʷɑ, the NF distance may be smaller than the Korean ka-kʷa, i.e., NF:2.

The partial perceptual distance scale of the Mandarin C-Cw contrasts, based on (4) with modification in particular for the back vowel [ɑ] and central vowel [ə], is given in (14). The CV-CwV contrasts outside the shaded area are allowed in Mandarin. I assume that the distance between Ceɪ and Cʷeɪ is comparable to the Ce-Cʷe contrasts of Korean at each place.

(14) The CV-CwV contrast distinctiveness in Mandarin

| | | Formant Transition (FT) scale | | | |
|---|---|---|---|---|---|
| | | 0 | 1 | 2 | 3 |
| Noise frequency (NF) scale | 0 | | **Δ(pə-$p^w$ə)** | **Δ(pa-$p^w$a)** | |
| | 1 | **Δ(peɪ-$p^w$eɪ)** | **Δ(tɑ-$t^w$ɑ)** **Δ(tə-$t^w$ə)** | **Δ(ta-$t^w$a)** | |
| | 2 | | **Δ(teɪ-$t^w$eɪ)** | **Δ(kɑ-$k^w$ɑ)** | |
| | 3 | | | **Δ(kə-$k^w$ə)** | **Δ(ka-$k^w$a)** |
| | 4 | | | **Δ(keɪ-$k^w$eɪ)** | |

Let us now turn to the evaluation of the CV-$C^w$V contrasts in Mandarin. The constraint ranking that I used for the Korean MSC can also distinguish between the allowed and disallowed C-Cw contrasts in Mandarin. MinDist= {NF:2 and FT:1} or {NF:1 and FT:2} is satisfied by the tan-$t^w$an contrast, but is violated by the contrasts between tɑŋ and $t^w$ɑŋ, and between tən and $t^w$ən. Therefore only the tan-$t^w$an contrast is legitimate in Mandarin. Tableaux in (15) and (16) illustrate how the labialization contrast on coronal stops is allowed before [a] but not before [ɑ].

(15)

| tan $t^w$an | MinDist={NF:2&FT:1} or {NF:1&FT:2} | MAXIMIZE CONTRASTS | MinDist={NF:3&FT:2} or {NF:2&FT:3} |
|---|---|---|---|
| a. ☞ tan-$t^w$an | | | * |
| b. tan | | *! | |

(16)

| tɑŋ $t^w$ɑŋ | MinDist={NF:2&FT:1} or {NF:1&FT:2} | MAXIMIZE CONTRASTS | MinDist={NF:3&FT:2} or {NF:2&FT:3} |
|---|---|---|---|
| a. tɑŋ-$t^w$ɑŋ | *! | | * |
| b. ☞tɑŋ | | * | |

Unlike the coronal stop context, the labialization contrast is allowed even before [ɑ] when the consonant is a velar stop, since the contrast satisfies the requirement of MinDist= {NF:2 and FT:1} (17).

(17)

| kɑŋ $k^w$ɑŋ | MinDist={NF:2&FT:1} or {NF:1&FT:2} | MAXIMIZE CONTRASTS | MinDist={NF:3&FT:2} or {NF:2&FT:3} |
|---|---|---|---|
| a. ☞kɑŋ-$k^w$ɑŋ | | | * |
| b. kɑŋ | | *! | |

Though not shown in (14), labialized consonants, including labialized labials, occur before vowel /ɔ/ as well, but they do not contrast with plain consonant counterparts in this context, as shown in (13). This is another case of neutralization of labialization contrast, in a context that decreases the salience of the contrast (a back rounded vowel), and thus can be given an analysis analogous to (16). A remaining question is why labialized consonants, rather than their plain consonant counterparts, occur before /ɔ/ vowel. Perhaps this is because Mandarin has another mid back rounded vowel, a falling diphthong /ow/. Adding an on-glide /w/ before /ɔ/ makes the two mid back vowels more distinct, one a rising diphthong and the other a falling diphthong, as we may see from the schematic comparison between (18a) and (18b).

(18) a. /ɔ/ vs. /ow/      b. /wɔ/ vs. /ow/

| a. | | b. | |
|---|---|---|---|
| | w | w | w |
| ɔ | o | ɔ | o |

So far in this section the Cw co-occurrence restrictions in two languages—Korean and Mandarin—are given a DT account. The constraint ranking relevant for Cw combinations in Korean and Mandarin is repeated in (19):

(19) MinDist= {NF:2 and FT:1} or {NF:1 and FT:2} >> MAXIMIZECONTRASTS
>> MinDist={NF:3 and FT:2} or {NF:2 and FT:3}

Let us now look at a language in which labialization contrast is more restricted than Korean and Mandarin, namely Cantonese. We will see that the analysis so far can easily account for the Cantonese facts as well.

4.1.4 Cantonese

In Cantonese, the C-Cw contrast is allowed only when the consonant is velar. Moreover, the labialized velar consonants are restricted in terms of the following vowel: they occur only before the mid front vowel /ε/ and the low central vowel /a/ (Bauer and Benedict 1997: 20-22).

(20) kε: '(possessive marker)'      $k^w$ε: 'die'[24]
ka: 'add'      $k^w$a: 'melon'

[24] In fact the velar-labialized velar contrast is mainly found before /a/ and /a:/ in Cantonese, though $k^w$ε: 'to die' is listed as a colloquial Cantonese word in Bauer and Benedict (1997). Cantonese has front rounded vowels /y/ and /œ/, so the contrast between /$k^w$i/ and /ky/, or /$k^w$ε/ and /kœ/ may be avoided due to their perceptual similarity. Kochetov (2008) reports that "the set of languages with contrastive secondary articulation [e.g., languages with contrastive labialization] hardly overlaps with the set of those that distinguish backness and rounding contrasts [e.g., languages with rounding contrast in front vowels]". Cantonese belongs to both sets, but it seems that the environment in which labialized velars occur hardly overlaps with that of roundness contrast (front vowels), and hence the rarity of the k-$k^w$ contrast before front vowels.

In the DT analysis, the difference between Cantonese on the one hand and Korean and Mandarin on the other is that the former requires greater distinctiveness between plain and labialized consonants. According to (4), it is only the labialization contrast on velar stops before non-back vowels that satisfies the minimal perceptual distance requirement of {NF:3 and FT:2} or {NF:2 and FT:3}. Thus MinDist= {NF:3 and FT:2} or {NF:2 and FT:3}, as well as the higher-ranking MinDist= {NF:2 and FT:1} or {NF:1 and FT:2}, dominates MAXIMIZECONTRAST in Cantonese, unlike in Korean and Mandarin. This ranking correctly distinguishes between the labialization contrast on coronal stops (22), which is illegal in Cantonese, and the legitimate velar labialization contrast (21).

(21)

| ka $k^w$a | MinDist={NF:2&FT:1} or {NF:1&FT:2} | MinDist={NF:3&FT:2} or {NF:2&FT:3} | MAXIMIZE CONTRASTS |
|---|---|---|---|
| a. ☞ka-$k^w$a | | | |
| b. ka | | | *! |

(22)

| ta $t^w$a | MinDist={NF:2&FT:1} or {NF:1&FT:2} | MinDist={NF:3&FT:2} or {NF:2&FT:3} | MAXIMIZE CONTRASTS |
|---|---|---|---|
| a. ta - $t^w$a | | *! | |
| b. ☞ta | | | * |

The constraint ranking in (21-22) also predicts labialization contrast on velar stops to be possible before mid central vowel /ə/ as shown in (14), but this vowel does not exist in Cantonese.

This section presented a DT analysis for the asymmetries in the distribution of labialized consonants. In languages like Korean, Mandarin and Cantonese, the Cw combinations exist only in the contexts where they are different to a certain degree from their plain C counterparts. Constraints that apply to phonological contrasts, such as Minimal Distance={NF:2 and FT:1} or {NF:1 and FT:2} and Minimal Distance={NF:3 and FT:2} or {NF:2 and FT:3}, the former universally ranked higher than the latter, were proposed, and shown to account for the asymmetric distribution of Cw combinations.

### 4.2 An alternative analysis: feature identity avoidance

In this section an alternative analysis of the co-occurrence restrictions on Cw combinations in Korean and Mandarin is discussed, using well-known constraints against repetition of identical features (OCP). It is shown that an account based on feature identity avoidance constraints faces serious difficulties in integrating consonant place and vowel asymmetries of Cw combinations, as well as in providing an explanation as to why such distributional asymmetries are found across languages.

Clements and Keyser (1983) attribute the non-occurrence of clusters consisting of a labial consonant and a labiovelar glide in English to the identity-avoidance constraint OCP (Obligatory Contour Principle, Leben 1973), applied at a segmental feature level (McCarthy 1986).

(23) OCP (labial): Two adjacent labial consonants are not allowed.

Gick (2003) claims that English bans labial + labiovelar glide combinations, but not velar + labiovelar clusters, because the primary articulation of English /w/ is the lip gesture, and OCP (place) applies only to the primary place of articulation. He suggests that in a consonant involving multiple articulation, *e.g.*, a velarized allophone of /l/ or a labiovelar glide /w/, one gesture is consonantal while the other is vocalic, and the former is equivalent to the primary place of articulation. The consonantal gesture is associated with properties such as reduction in final position, intermediate magnitude under resyllabification, and tendency to be away from the peak of vowel gesture. The results of Gick's (2003) experiment show that in English /w/, the lip gesture precedes the tongue dorsum gesture in syllable onset position, and also undergoes reduction in postvocalic positions, leading to the conclusion that the consonantal gesture of /w/ is labial.

In contrast, Moreton (2002) argues that there is no evidence that the constraint OCP (labial) is active in the grammar of English speakers, and therefore the absence of /pw/ is a lexical gap in English. In a series of experiments, Moreton found no bias against /bw/ clusters in English listeners' perception, unlike /dl/ clusters, which he claims to be a true phonological gap. He also argues that the few borrowed words containing /pw/ and /bw/, mostly of Spanish or French origin, are repaired in ways that he claims to "[be] unsystematic and look suspiciously like spelling pronunciation" (Moreton 2002:50), and thus cannot provide evidence for the illegality of /pw/ and /bw/ in English. Thus, there is little support for a constraint like (23) as an explanation of Cw patterning in English.

### 4.2.1 Korean

The avoidance of the labial + /w/ combination in Korean might be attributed to OCP (labial) constraint (23), which prohibits the repetition of two [labial] features. Following Gick's (2003) claim that /w/ is primarily labial in languages like English which ban /pw/ clusters, we may also assume that /w/ is primarily labial in Korean. Although both labial and velar consonants share a place feature with labiovelar /w/, it is labials that share the C-place feature with /w/. The [dorsal] features of /k/ and /w/ are on different nodes, C-place and V-place, respectively, so they may not incur an OCP violation.

According to Kang (1997), the loss of /w/ after a labial consonant is a historical sound change of the Korean language, as shown by recent spelling changes, e.g., 뵈 *pwe* → 베 *pe* 'hemp cloth'. The fact that /pw/ is not allowed in monomorphemic forms of contemporary Korean suggests that the identity avoidance constraint OCP (labial) dominates a faithfulness constraint MAX (w), as illustrated by the tableau in (24).

(24) /pwe/ → /pe/ 'hemp cloth'

| pwe | OCP (labial) | MAX (w) |
|---|---|---|
| pwe | *! | |
| ☞pe | | * |

Unlike the complete ban on /pw/ combinations, vowel quality matters in the distribution of /tw/ combinations. To fully address the effect of vowel quality on the distribution of /tw/, it is necessary to have a closer look at the co-occurrence restrictions between /w/ and a vowel, repeated in (25).

(25) /w/ + vowel combination of contemporary Seoul Korean

| | | |
|---|---|---|
| wi | *wɨ | *wu |
| we | wʌ | *wo |
| | wa | |

Example (26) shows featural representations of Korean vowels. Following the unified feature theory (Clements 1991, Clements and Hume 1995), the traditional vowel features [back] and [round] are replaced by [dorsal] and [labial], respectively. Note also that in this feature system, high vowels and corresponding glides (/i/ and /j/, and /u/ and /w/) have the same place feature specifications, but are distinguished by the value of the [consonantal] feature or presence of a mora.

(26) Feature representation of the Korean vowels

| | i/j | e | a | ɨ | ʌ | o | u/w |
|---|---|---|---|---|---|---|---|
| high | + | - | - | + | - | - | + |
| low | - | - | + | - | - | - | - |
| dorsal (back) | | | | + | + | + | + |
| labial (round) | | | | | | + | + |

The fact that vowels /u/, /o/ and /ɨ/ are not allowed after /w/ may also be related to identity avoidance, i.e., the similarity in the overall featural contents between /w/ and the vowel. Notice that the number of features shared by /w/ and a vowel is related to the legitimacy of the combination of /w/ and that vowel: /w/ shares with /u/ the same value for all four features, with /o/ the same value for three features [dorsal], [-low] and [labial], and with /ɨ/ the same value for three features [dorsal], [high] and [-low]. These three are the vowels which do not co-occur with /w/. The mid back unrounded vowel /ʌ/, which has limited distribution with respect to the preceding consonants, has two features, [-low] and [dorsal] in common with /w/. Therefore we need a pair of relativized OCP constraints (Selkirk 1991, Moreton 2002) that penalize the same value for [low] in two adjacent [dorsal] segments (back vocoids), or two [dorsal] and [labial] segments (back rounded vocoids) (27a-b):[25]

(27) a. OCP (dorsal & low):
Adjacent [dorsal] segments cannot have the same value for [low].

b. OCP (lab & dorsal & low):
Adjacent [labial] and [dorsal] segments cannot have the same value for [low].

A ranking in which constraint (27b) dominates a faithfulness constraint correctly rules out nonexisting sequences like **wo* and **wu*. Constraint (27a) bans the combination of /w/ and a low back vowel /ʌ/, so it must be ranked lower than faithfulness constraints as

[25] To rule out the combination *wɨ, another constraint—OCP (dor & high &low)—is required, which is also ranked above Faithfulness.

in (28), as /wʌ/ is a legimate sequence in Standard Korean (29). Though the actual output of an input like /wo/ cannot be decided, as we lack the evidence for a ranking between relevant faithfulness constraints, I assume that /w/ is deleted in the output form (30). The deletion of /w/ from wV sequences is partly evidenced by dialectal variation of /w/ + V combinations: in Kyungsang dialect /wa/ is the only /w/-diphthong, and words with /wʌ/ in other dialects surface without /w/ (Kang 1997: 43).

(28) OCP (labial & low & dorsal) >> MAX >> OCP (low & dorsal)

(29) /wʌ/

| /wʌ/ | OCP (lab & low & dor) | MAX | OCP (low & dor) |
|---|---|---|---|
| ☞wʌ | | | * |
| ʌ | | *! | |

(30) hypothetical */wo/

| /wo/ | OCP(labial & low & dor) | MAX | OCP (low & dor) |
|---|---|---|---|
| wo | *! | | * |
| ☞o | | * | |

Let us now incorporate the place asymmetry of Cwʌ into the constraint system: a coronal onset does not co-occur with /wʌ/, unlike dorsal consonants. The OCP constraint that bans /wʌ/ (27a) is ranked lower than the faithfulness constraint, as /wʌ/ can occur by itself or after a velar consonant in monomorphemic forms. Likewise, a markedness constraint against a coronal stop (i.e., *t) is ranked lower than the faithfulness constraint. It is only the co-occurrence of a coronal onset and /wʌ/ that is absent in monomorphemic forms. To rule out sequences such as */twʌ/, but to allow /wʌ/ and /twe/, we need a local conjunction of two constraints (Prince and Smolensky 1993, Smolensky 1995, Lubowicz 1999) as (31), ranked higher than each component constraint (32).

(31) *t & OCP (dor & low):
No coronal onset and adjacent segments sharing values for [dorsal] and [low].

(32) OCP (dor & low) & *t >> MAX >> OCP (dor & low), *t

A tableau for a hypothetical input */twʌ/ is given in (33). The ranking in (32) correctly predicts that this input cannot surface faithfully.

(33) Hypothetical /twʌ/

| /twʌ/ | *t & OCP (dor & low) | MAX (W) | OCP (dor & low) | *t |
|---|---|---|---|---|
| twʌ | *! | | * | * |
| ☞ tʌ | | * | | |

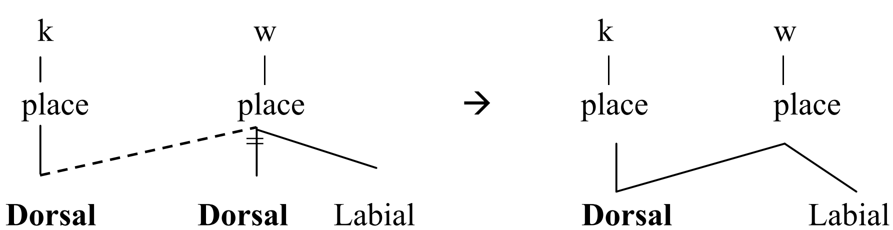
As /kw/ is allowed before /ʌ/, the conjoined constraint *k & OCP (dor & low) must be ranked lower than the faithfulness constraint, and by transitivity, lower than *t & OCP (dor & low).

(34) OCP (dor & low) & *t >> Max (w) >> OCP (dor & low) & *k

The analysis developed so far works, but has some problematic features. First, the ranking between two conjoined constraints in (34) is possible only if *k is ranked lower than *t, which seems to go against the universal markedness hierarchy among consonant places. Alternatively, one may argue that the markedness constraint involved here is *tw and *kw, rather than *t and *k, and *tw is ranked higher than *kw. However, a real challenge to this analysis lies in how to justify the supposedly universal ranking *tw >> *kw (given that Korean exemplifies the universal preference for /kw/), and why these constraints interact with the constraints of identity avoidance (OCP), which are seemingly not related to them. Notice that this interaction is natural in the contrast-based analysis such as the one proposed in 4.1: both the preference for /kw/, and the avoidance of /w/ before a non-low back vowel (and consequently the preference for front and/or low vowels after /w/), give rise to a greater phonetic salience between plain C and Cw combination onsets.

Let us now look at a somewhat different role that OCP (place) may play regarding the Korean Cw combinations. Recall from Chapter 1 (example (6)) that preference for a /kw/ combination in Korean was found not only in morpheme structure conditions, but also in loanword adaptation patterns: English velar + /w/ clusters are mapped to Korean /kw/ combinations, unlike labial + /w/ or coronal + /w/ clusters which incur epenthesis of vowel /ɨ/, like other consonant clusters. One possible analysis is that /k/ and /w/ build a tighter bond with each other in Korean, so that an epenthetic vowel cannot intervene, as shown in (35). This was also suggested by Silva (1991) on independent grounds that /k/ and /w/ may form a unit in Korean on an independent ground. In his study of optional /w/ deletion in speech, Silva found that the deletion rate of /w/ was the lowest when the preceding consonant was velar (Cf. Kang 1997). Based on this result, he argues that the two adjacent [back] features of dorsal consonants and /w/ incur an OCP violation, and become multiply linked in Korean. As a consequence, /kw/ "maintains the integrity of the two segments as a unit", hence the resistance to /w/ deletion (Silva 1991: 165).

(35)




In order to allow the merger of two dorsal place nodes, it is necessary for the dorsal features of /w/ and /k/ to be adjacent. Note that this analysis conflicts with the assumption I made to rule out */pw/ by OCP (labial) above. There, following Gick (2003), I assumed that the C-place of /w/ is labial in Korean, rather than dorsal, so that the OCP-place constraint applies to */pw/ but not to /kw/. If the deletion of /w/ from /pw/ and the merger of [dorsal] nodes associated with /k/ and /w/ are triggered by adjacency of the two place features, both [labial] and [dorsal] features of /w/ should be adjacent with the preceding consonant's place feature, and incur a potential OCP violation. And if both labial and dorsal consonants are subject to OCP (place) when followed by /w/, the question is why the strategies to fix the ill-formed structures with vary depending on which feature is involved: for [labial] it is the deletion of the segment, whereas for [dorsal] it is the deletion of one feature.

Thus, we find that an appeal to OCP does not suffice in explaining the consonant place and vowel asymmetries observed in the distribution of the Korean Cw combinations.

### 4.2.2 Mandarin

The identity avoidance analysis of Cw combinations faces an even more serious problem when we look at Mandarin co-occurrence restrictions. First, the fact that Mandarin labial + /w/ combinations are found only before a rounded vowel calls for attention. If the ban on labial + /w/ combinations in other vowel environments is ascribed to an OCP (labial) violation, the emergence of this marked combination before another [labial] segment—vowel /ɔ/—is surprising. Thus, the lack of *pwa- or *pweɪ syllables cannot be attributed to OCP (labial), since /pw/ combinations do occur before the vowel /ɔ/.

The important difference between Cwɔ on the one hand and other CwV syllables on the other is that Cwɔ does not contrast with Cɔ, while other CwV syllables contrast with CV. In other words, the vowel /ɔ/ is always preceded by an onglide /w/, as shown in (13) above. Duanmu (1990: 66-68) argues that the vowel [ɔ] is underlyingly /ə/ after a labial, and becomes [ɔ] by the spreading of the [labial] feature from the preceding consonant. In his analysis, labial consonants in turn become rounded by the insertion of [+rounded] under the labial node. Because Duanmu treats Cw combinations as labialized consonants, rather than clusters, no violation of OCP (labial) is incurred in [pwɔ]. However, this analysis requires the assumption that [ɔ] is underlyingly /ɔ/ after coronal or velar consonants, unlike after labial consonants, and that /w/ exists underlyingly in /twɔ/ or /kwɔ/. This does not explain why /twɔ/ and /kwɔ/ exist while */tɔ/ and */kɔ/ do not, just as /pwɔ/ exists but */pɔ/ does not.

The vowel asymmetry in Mandarin /tw/ combinations is also problematic for applying a similar analysis suggested for Korean in 4.2.1. Recall that only [a] co-occurs with Tw combinations in Mandarin. It is not clear how to differentiate between them in the formal feature representation, since [a] and [ɑ] are allophones of a single low vowel, given their complementary distribution ([ɑ] before /ŋ/, and [a] elsewhere).[26] Therefore, even if we posit a constraint against /tw/ before a low vowel, both [twan] and *[twɑŋ] would violate it. To distinguish these two, we need to assume that the allophone [ɑ] is specified for [back], and have *tw & OCP (back) rule out *[twɑŋ]. However, as shown in (13) above, back vowel /ɔ/ does co-occur with /w/ following a coronal stop in Mandarin.

To sum up this section, an alternative account was considered for the morpheme structure conditions regarding Cw combinations in Korean. Constraint ranking in which MAX (w) is dominated by OCP (labial), and a conjoined constraint *t & OCP (dorsal & low) (or *tw & OCP (dorsal & low)) was suggested for Korean, but it remains unclear why these two conjuncts interact. Moreover, additional difficulties emerge in extending

[26] Wang (1993) argues that the low vowels in Mandarin are unspecified for backness, and thus underlyngly represented as [+low, -round].

this analysis to the seemingly similar phenomena in Mandarin. Thus, the identity avoidance analysis attempted in this section faces the problems in not only its explanatory but also descriptive adequacies.

The distribution of Korean and Mandarin Cw combinations resembles that of Cw combinations in many other languages, in that the presence of /kw/ and the absence of /tw/ before a back vowel seem to be a combinatory effect of two tendencies—frequent occurrence of contrastive labialization on velar consonants, and avoidance of back/rounded vowels after labialized consonants. Thus, it is appropriate to view the language-specific patterns of Korean and Mandarin as cases of universal phonological trends, and to look for explanatory principles behind the emergence of these patterns. The current analysis succeeded in this by incorporating acoustic/ perceptual salience of contrastive labialization into the explanation of phonological patterns.

### 4.3 Interaction of contrast dispersion and articulatory gesture coordination in Cw combinations

In this section I consider a special case of distributional asymmetry in Cw combinations: the case in which the phonetic realization of the Cw combinations varies depending on the consonant place, i.e., labialized consonants in the velar context and Cw clusters in others.

The acoustic study results of Chapter 2 suggest that the Formant Transition cues of the prevocalic glide /w/ are stronger in Spanish than in Korean: Spanish CV-CwV difference was greater than Korean CV-CwV difference in terms of both the formant and vocoid duration measurements. The perception experiment did not directly test the effect of the degree of overlap between C and /w/—that is, whether the combination is close to a sequence of a consonant plus a glide, or a secondarily-articulated consonant—on the perceptibility of the /w/. Yet it seems plausible to assume that there is correlation between acoustic distinctiveness and perceptibility of contrast between plain CV and CwV. Kochetov and Goldstein (2006) found a related result regarding the relationship between C-G overlap and the perceptibility of the glide. In their perception experiment using synthesized stimuli of /bɛ/ and /bʲɛ/, they found that the difference between plain /bɛ/ and /bʲɛ/ was harder to perceive when [labial] and [palatal] gestures were more overlapped. This result supports their hypothesis that the clustering of a consonant and a glide (CG) is better than secondary articulation ($C^G$) in terms of the perceptibility of the glide component, as indicated schematically in (36).

(36) $\Delta(C_1V_1\text{-}C_1GV_1) > \Delta(C_1V\text{-}C_1{}^GV_1)$

When we apply this relationship to Cw clusters and labialized consonants, the following perceptibility scale is obtained.

(37) $\Delta(C_1V_1\text{-}C_1wV_1) > \Delta(C_1V\text{-}C_1{}^wV_1)$

Let us assign a hypothetical value *n* to the perceptual distance $\Delta(C_1V_1\text{-}C_1{}^wV_1)$, and *m* to $\Delta(C_1V_1\text{-}C_1wV_1)$. Minimal Distance=*n* dominates Minimal Distance=*m* (38), by the Pāṇinian Theorem of constraint ranking, as the latter is more specific than the former.

(38) Minimal Distance=*n* >> Minimal Distance=*m*

If this is the case, why do some languages employ a less effective way of combining a consonant and a glide, *i.e.,* the secondary articulation, while other languages choose cluster-like, perceptually optimized timing patterns? Given the perceptibility scale above, DT predicts that a C-Cw cluster contrast would always be a better option in a language that contrasts plain and labialized consonants, since the former satisfies more Minimal Distance constraints. Compare the tableaux in (39) and (40): (39b) violates MinDist=*m*, whereas (40b) satisfies all constraints:

(39)

| $C_1V_1$ $C_1{}^wV_1$ | MinDist=*n* | MAXIMIZECONTRASTS | MinDist=*m* |
|---|---|---|---|
| a. $C_1V$ | | *! | |
| b. ☞$C_1V$-$C_1{}^wV_1$ | | | * |

(40)

| $C_1V_1$ $C_1wV_1$ | MinDist=*n* | MAXIMIZECONTRASTS | MinDist=*m* |
|---|---|---|---|
| a. $C_1V$ | | *! | |
| b. ☞ $C_1V_1$-$C_1wV_1$ | | | |

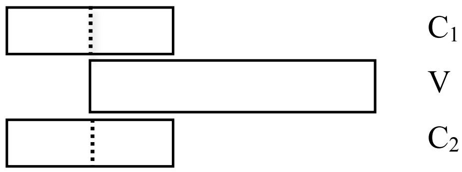
That some languages choose more “overlapped” articulation of CG combinations can be attributed to another functional principle, namely articulatory effort minimization. To understand the precise nature of effort minimization in the present case, we turn to the gesture coordination preference of Articulatory Phonology. According to Browman and Goldstein (1988, 2000), prevocalic C gestures form a strong bond with the vowel gesture (*C-V relation*): the center of consonant gestures (c-center) tends to be synchronous with a specific point in the vowel gesture (C-center effect).[27] However, when there is more than one prevocalic consonant gesture, the perfect C-V phasing of all C gestures would result in complete overlap between the C gestures (41a). This is problematic for the *Recoverability* of gestures (Browman and Goldstein 2000). The result is that the two consonant gestures are shifted from the optimal C-V relation (41a), as in (41b) or (41c).[28] The boxes represent the duration of each gesture, and the dashed lines indicate the c-center point of $C_1$ and $C_2$ gestures.

[27] Browman and Goldstein (1988) do not exactly specify which point of the vowel gesture is synchronous with the c-center, but show that the center of C gesture is always at the same distance from an anchor point, which is the attainment of the target of the coda consonant. In Gafos (2002: 318), in which gestures consist of key points of articulator movement (landmarks), the c-center is aligned with the onset of the vowel gesture.

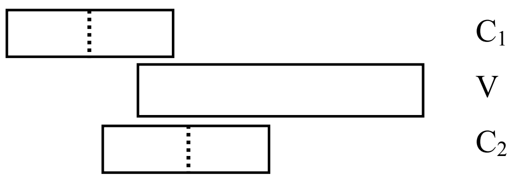
[28] Notice that though the displacement of C gestures is greater in (41b) than in (41a), the two C gestures within each organization are displaced by an equal distance from their optimal position. According to Browman and Goldstein (2000:29), phasing between two C gestures is the most optimal if each C gesture is shifted by the same distance from the optimal C-V relation.

(41) Simplified gesture organization of a multi-gestural onset

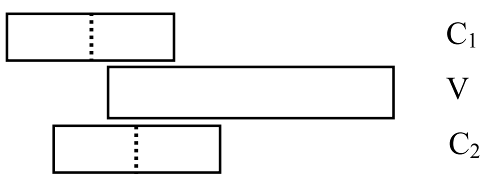
a. Optimal C-V relation (complete overlap): Unattested




b. Greater displacement of gestures




c. Less displacement of gestures




CG clusters and secondarily-articulated consonants differ in the timing relationship between the C and G components; in Articulatory-Phonological terms, C and G gestures overlap more in the latter. We may use the simplified gestural coordination patterns in (41b) and (41c) to illustrate CG clusters and secondarily-articulated consonants, respectively. Assuming that $C_1$ is a stop or fricative gesture and $C_2$ is a glide component, gestures in CG cluster (41b) deviate for a longer distance from the perfect c-center alignment (41a) than those in secondarily-articulated consonant (41c). In other words, secondary articulation is better than CG clustering, in terms of the universal gesture coordination requirement.

Browman and Goldstein (2000) argue that the functional principle related to the C-V relation is the parallel transmission of consonants and vowels (Liberman et al. 1967), which ensures the efficiency of speech communication. Following Browman and Goldstein I assume that the C-center effect is a universal property of multi-gestural onsets; that is, the gestural coordination in (41c) is universally preferred to that in (41b) in terms of their C-V relation. I also assume that this universal preference gives rise to a strict dominance relationship (42c) between the constraints that ban consonant clusters (42a) and secondarily-articulated consonants prevocalically (42b).

(42) a. *CGV: No CG clusters at syllable onset
b. *$C^GV$: No secondarily-articulated consonants at syllable onset
c. *CGV >> *$C^GV$

Thus, it is the role of the prevocalic articulatory gesture coordination preference that may force secondary articulation realization, at the cost of reducing the perceptual distance from the plain C counterpart. In languages that allow the perceptually weaker C-CG contrast (secondary articulation), *CGV is ranked higher than the Minimal Distance constraints that require *m* auditory contrast, and rules out the cluster realization of the CG combination. Note that I am expanding the perceptual space to include Cw clusters, and thus in (43), the contrast involving Cw clusters is evaluated along with that of labialization.

(43) MinimalDistance=*n*>> *CGV >> CONTRAST >> MinimalDistance=*m*

| $C_1V$ $C_1^wV$ $C_1wV_1$ | MinDist=*n* | *CGV | MAXCONTRASTS | MinDist=*m* |
|---|---|---|---|---|
| $C_1V_1$-$C_1wV_1$ | | *! | | |
| ☞$C_1V$-$C_1^wV_1$ | | | | * |
| $C_1V$ | | | *! | |

Let us now turn to a language in which the realization of the glide /w/ component varies among the consonants of different places of articulation, e.g., Dan (Santa). In this language Cw combinations are found at all three consonant places of articulation, but only velar + /w/ was considered a unit segment, *i.e.,* a labialized velar, both phonetically and phonologically. (44) provides example minimal pairs or near-minimal pairs that involve the C-Cw contrast.

(44) (Adapted from Bearth and Zemp 1967)

| | | | |
|---|---|---|---|
| baŋ | 'machete' | bwaŋ | 'basin' |
| tə: | 'palm tree' | twə: | 'domestic animal' |
| dɛ | 'spirit of divination' | dwɛ | 'place, rank' |
| kɛ | 'custom' | $k^w$ɛ | 'year' |
| kə | 'hoe' | $k^w$ə | 'together' |

I suggest that this language-internal timing variation with regard to the realization of Cw combination is a result of the interaction between the minimal distance requirement between plain C and Cw combinations, and the constraint that avoids CG clusters at onset (*CGV). The perceptibility of the CV-$C^w$V and CV-CwV contrasts was schematized as *n* and *m* respectively (*m*>*n*) above, without specifying which dimension—formant transition or noise frequency, or both—contributes to this difference. However, the acoustic study results in Chapter 2 showed that unlike the Formant Transition cues, the Burst Frequency cues of /w/ did not differ between Korean and Spanish. In other words, the degree of the frequency lowering by the /w/ on the stop burst spectra was not affected by the degree of articulatory overlap of the stop and the glide. Therefore, the auditory distance between $C^w$V and CwV exists only in the Formant Transition dimension. The two-dimensional auditory space (Formant Transition x Noise Frequency) in (45) illustrates how a plain consonant, a labialized consonant and a Cw cluster may differ: $C^w$ and Cw are distinguished only along the formant transition scale.

(45)

| | Formant Transition | | |
|---|---|---|---|
| NF | CV | | |
| | | $C^wV$ | CwV |

The fact that only velar consonant + /w/ combinations surface as a labialized consonant in Dan suggests that this language requires a fairly large difference in the consonant release NF dimension, like Cantonese. The difference in the release burst NF was bigger for the velar stop (3-4 steps) than coronal and labial stops (0-2 steps), according to the perceptibility scale (2) based on the Korean data. Interestingly, it is in the context where the consonant release NF differs less, i.e., dental and labial in comparison to velar, that the CG combination is realized as a cluster. As schematized in (45), clustering of C and /w/, compared to superimposing /w/ as a secondary articulation on C, increases the perceptual difference of Cw combination from plain C on the Formant Transition scale. Thus, by realizing /t + w/ and /p + w/ combinations as clusters rather than labialized consonants, perceptual distance between plain C and Cw combination onset may become compatible to that between plain and labialized velars (46).

(46)

| | | Formant Transition (FT) scale | | | |
|---|---|---|---|---|---|
| | | 0 | 1 | 2 | 3 |
| Noise frequency (NF) scale | 0 | | **Δ(pə-$p^w$ə)** | **Δ(pa-$p^w$a)** | |
| | 1 | **Δ(pe-$p^w$e)** | **Δ(tɑ-$t^w$ɑ)**<br>**Δ(tə-$t^w$ə)** | **Δ(ta-$t^w$a)** | **Δ(pV-pwV)**<br>**Δ(tV-twV)** |
| | 2 | | **Δ(te-$t^w$e)** | | |
| | 3 | | | **Δ(kə-$k^w$ə)** | **Δ(ka-$k^w$a)** |
| | 4 | | | **Δ(ke-$k^w$e)** | |

In order for the contrast to surface in Dan, the distance between CV and CwV may not fall within the shaded area in the table (46) above. Assuming that CwV syllables differ from their CV counterpart by at least 3 steps on FT scale, minimal distance requirement between plain and Cw onsets in Dan may be expressed by a constraint such as:

(47) MINIMAL DISTANCE= FT:3 or NF:3

This constraint is ranked higher than MAXIMIZE CONTRASTS in Dan, so that contrasts that do not satisfy either condition (FT:3 or NF:3), e.g., a labialization contrast on labial and coronal stops, do not exist, as illustrated for coronals in (48). Instead of neutralizing the contrast (48c), the /t + w/ combination is realized as a cluster (48b), which shows that the articulatory gesture coordination constraint *CGV is ranked lower than MAXIMIZE CONTRASTS.

(48) Dan

| ta $t^w a$ twa | MD=NF:3 or FT: 3 | MAXIMIZECONTRASTS | *CGV |
|---|---|---|---|
| a. ta-$t^w$a | *! | | |
| b. ☞ta-twa | | | * |
| c. ta | | *! | |

As shown in (49), though the ka-kwa contrast is more salient than the ka-$k^w$a contrast, /kw/ cluster is avoided. The constraint *CGV, ranked higher than *$C^G$V, rules out the /kw/ cluster in favor of /$k^w$/.

(49) Dan

| ka $k^w$a kwa | MD=NF:3 or FT: 3 | MAXCONTRASTS | *CGV | *$C^G$V |
|---|---|---|---|---|
| a. ☞ka-$k^w$a | | | | * |
| b. ka-kwa | | | *! | |
| c. ka | | *! | | |

If the degree of articulatory overlap between a consonant and a glide may vary within a language in order for the C-CG contrast to satisfy the minimal distance requirement, it is necessary to ensure that the analysis does not predict the same pattern for languages like Cantonese, where labialization contrast is limited to velars, and no Cw clusters are allowed. As shown in (50), contrast between /t/ and /tw/ cluster may satisfy the minimal distance constraint, unlike that between /t/ and /$t^w$/.

(50) Cantonese

| te $t^w$e twe | MinDist={NF:2&FT:3} or {NF:3&FT:2} | MAXIMIZE CONTRASTS |
|---|---|---|
| a. te - $t^w$e | *! | |
| ☜b. te - twe | | |
| c. te | | *¡ |

By altering the ranking of *CGV with respect to the MAXIMIZE CONTRASTS constraint, the two different patterns found in Dan and Cantonese are derived. Both Dan and Cantonese allow labialization contrast only on the velar consonants, since labialized coronals or labials do not satisfy the minimal distance requirement between plain C and Cw combination onsets. Cw combinations that do not differ enough from plain C do not exist in Cantonese, but exist as clusters in Dan. Whether or not Cw clusters are allowed in the language depends on the ranking of *CGV with respect to MAXIMIZE CONTRASTS: When *CGV is ranked above MAXIMIZE CONTRASTS, the Cantonese pattern where no Cw clusters are allowed is obtained (51-52); when *CGV is ranked lower than MAXIMIZE CONTRASTS, the patterns found in Dan emerge (48-49).

(51) Cantonese

| te $t^w$e twe | MinDist={NF:3&FT:2} or {NF:2&FT:3} | *CGV | MAXIMIZE CONTRASTS | *$C^G$V |
|---|---|---|---|---|
| a. te- $t^w$e | *! | | | * |
| ☞b. te | | | * | |
| c. te - twe | | *! | | |

(52) Cantonese

| ka $k^w$a kwa | MinDist={NF:3&FT:2} or {NF:2&FT:3} | *CGV | MAXIMIZE CONTRASTS | *$C^G$V |
|---|---|---|---|---|
| ☞ a. ka- $k^w$a | | | | * |
| b. ka | | | *! | |
| c. ka - kwa | | *! | | |

In this section I considered the language-internal variation in the realization of consonant + /w/ combinations in Dan (Santa), in which velar + /w/ is realized as a labialized consonant, whereas other Cw combinations surface as consonant clusters. The fact that the velar + /w/ combinations are realized as labialized consonants unlike labial or coronal + /w/ in this language was attributed to the inherently stronger burst noise frequency cue of labialized velar consonants, and the role of articulatory gesture coordination requirement that prefers secondarily-articulated consonants to CG clusters.

In the next section I extend the present analysis to the distribution of Cj clusters and palatalized consonants, showing that the general principles developed on the basis of Cw combinations also hold for other CG combinations.

### 4.4 Interaction of contrast dispersion and articulatory gesture coordination in Cj combinations

If the analysis presented in 4.3 for the CV-CwV contrast is on the right track, we would expect a similar relationship between the strength of the consonant release cue and the combination type (secondary articulation or cluster) in the distribution of palatalized consonants and Cj clusters as well. In this section I examine the distribution of the contrastive palatalized consonants and their cluster counterparts (Cj clusters), and test the possibility of extending the current analysis.

Takatori (1997) reports that there is an implicational relationship between palatalized coronal and palatalized labial stops within the Slavic language family: if a language has contrastive palatalized labial consonants, then it also has a palatalization contrast on coronals. Based on this implicational relation, Takatori (1997) proposed a universal markedness relationship between palatalized labials and palatalized coronals, shown in (53):

(53) $*p^j >> *t^j$

Kochetov (2002) presents a survey of contrastive palatalization in Slavic, Celtic and Uralic languages, which shows similar results. All languages reported in his survey have a palatalization contrast on coronal stops, though the phonetic realization may vary depending on the language (e.g., tʲ, tʃ, tɕ, c, tsʲ). However, only some of these languages contrast palatalized labials with plain labial consonants. This is in accordance with the implicational relationship between pʲ and tʲ, and subsequently the ranking of markedness constraints in (53) proposed by Takatori.

This implicational relationship between palatalized labials and coronals, however, does not extend to the Cj clusters. In fact the preference for coronal + /j/ combination is quite the opposite of what has been known as an OCP (coronal) effect against coronal + /j/ clusters (Borowsky 1986, Clements and Hume 1995). For example, Standard American English does not allow coronal + /j/ clusters, but it does allow labial + /j/ clusters, as in 'beauty'.[29] Kawasaki (1982) reports more examples of languages in which /j/ is not found after dental, alveolar or palatal consonants, and states the avoidance of the coronal + /j/ clusters as a universal tendency.

Interestingly, both patterns of co-occurrence restrictions have been given perception-based accounts. Flemming (1995:172-176; 2002:120-125) provides a Dispersion-Theoretic analysis of the neutralization of the t-tj and d-dj contrasts in Standard American (neutralization to /t/ and /d/) and Southern Californian English (neutralization to /tj/ and /dj/), based on the lack of sufficient F2 contrast between simplex coronal stops and coronal + /j/ clusters. Palatalization, or the prevocalic glide /j/, has an acoustic effect of increasing the F2 frequencies. Since labial consonants have characteristically low F2 at vowel onset, there must be considerable auditory difference between PV and PjV syllables. Coronal consonants, in contrast, condition high F2 transition by themselves, so the addition of /j/ may be harder to detect than in the labial context.

Despite the relative closeness in their formant transition, plain and palatalized coronal stops are well differentiated because of affrication: according to Ladefoged (1971), the release of palatalized coronal obstruents is accompanied by a fricative off-glide, since it is difficult to withdraw the tongue blade quickly. Labial constriction release, in contrast, is not affected by the tongue articulation. Therefore, coronal stops inherently provide more distinctive consonantal release cues for contrast related to palatalization than labial stops. Padgett (2001:209) suggests that the universal preference for the contrastive palatalization on coronals is due to the additional perceptual cues from the affrication. Thus, vocalic and consonantal cues of palatalization or /j/ glide point to different directions: the vowel formant cue is stronger when the consonant is labial, whereas the additional consonant release cue is present only for coronals.

Let us now address the question of why the preferred place for palatalized consonants and Cj clusters differ: coronal for the former, and labial for the latter. Padgett (*ibid.*) argues that the perceptual dispersion of the palatalization contrast is dependent on the number of cues, rather than the absolute degree of acoustic difference: a palatalization contrast on labials depends solely on the F2 transition cues, whereas on coronals it is aided by the affrication at the consonant release (though the transition cue is smaller than in the labial context). However, this leaves the avoidance of coronal + /j/ clusters unexplained. Affrication of coronal obstruents in the context of /j/ is common to both

[29] The fact that Cj cluster is only found before the vowel /u/ in English is also understood as a contrast dispersion effect: the high transition F2 cues of /j/ component is the clearest before /u/, which has the lowest F2 among English vowels.

types of CG combinations, so coronal + /j/ clusters are also advantageous over labial + /j/ clusters in terms of cue counting. Therefore, the relative strength of cues in each type of contrast, rather than their availability itself, may be responsible for the divergent co-occurrence restrictions on Cj combinations. The formant transition difference between $C^jV$ and CV is smaller than that between CjV and CV, in terms of both the frequency and the duration of the transition (see Ladefoged and Maddieson 1996:364, Fg.10.23 for the comparison between the Russian /pj/ and /$p^j$/ in prevocalic positions). In Cj clusters, due to the independent status of /j/, the vocalic cues (e.g., formant frequencies, vocoid duration) may be predominant for the perception of the C-Cj cluster contrast, more than they are for the C-$C^j$ contrast perception. Therefore the labial consonants are perceptually better for this type of contrast, since formant transition differences between /p/ and /pj/ are greater than those between /t/ and /tj/.

The consonant release cues become relatively more important in the perception of the palatalization contrast, because of the weaker formant transition cues signaling the presence of a palatal component, in comparison to the cluster realization. The realization of /j/ as a secondary articulation of palatalization can more easily occur when the consonant is coronal, since the consonantal cue, i.e., affrication, is present. In contrast, the weak realization of /j/ (i.e., secondary articulation of palatalization) is disfavored for labial consonants, whose release burst noise is not affected by palatal co-articulation.

This is reminiscent of the argument in the previous section, regarding the place asymmetry in the realization of Cw combinations in Dan. Velar consonant + /w/ combinations are realized as labialized consonants, whereas labials and coronals form consonant clusters with /w/. The contrast between the plain and the labialized velars satisfies the auditory distinctiveness required in the language, despite the weaker formant transition cues, due to the strong consonant release cues. If this parallelism between C+ /w/ and C + /j/ combinations is valid, we would expect a language in which coronal + /j/ combinations are realized as palatalized coronals, whereas labial + /j/ surfaces as a cluster. In fact this is the case in Slavic languages like Ukrainian, Polish and Czech. Ukrainian allows contrastive palatalization on coronals, but not on labials (Shevelov 1993). Instead, the language has labial + /j/ clusters, whereas Russian, a language that also belongs to the East Slavic branch, has palatalized labials. (54) exemplifies the difference between the two languages:

(54) Russian $p^ja t^j$ Ukrainian $pja t^j$ 'five'

In Polish (Bartek Czaplicki, p.c.) and Czech (Short 1993), palatalized /t/ historically became affricates /tɕ/ and /c/, respectively, but the palatalized version of the labial consonant is in the form of Cj clusters. This pattern is not confined to Slavic languages: according to Kochetov (2002:23), the Celtic language Manx also shows such asymmetry. Thus, there exists a parallelism between the distributions of the Cw and Cj combinations: when the consonantal cue is relatively strong for the C-CG contrast (velar + /w/ and coronal + /j/), the glide component is more likely to be weakened, presumably due to the avoidance of CG clusters, resulting in the secondary articulation. When the consonantal cue for the glide is relatively weak (labial + /w/, coronal + /w/ and labial + /j/), it tends to be an independent segment, and thus satisfies the minimal auditory distance requirement between C and CG.

Given these parallels, it is easy to envision a Dispersion-Theoretic OT account for the asymmetries of Cj combination types, parallel to the case of Cw in 4.3. Since I do not have any experimental results on the acoustics and perception of palatalized consonants and Cj clusters, I will rely on what has been reported in the literature about the acoustic correlates of palatalization, in order to estimate the auditory distance among CV, $C^jV$ and CjV. Plain and palatalized coronal stop bursts are both characterized by high-frequency strident noise, but the latter has longer noise duration (Kochetov 2002:60), due to the additional affrication. Stops with longer burst noise, or the partially affricated stops, have greater perceived loudness (Flemming 2002:23). Therefore we expect the palatalized coronal stop release burst to be louder than its plain counterpart. Such loudness difference is not expected between the plain and the palatalized labial stop bursts. The scale in (55) schematizes the relative distinctiveness between plain and palatalized consonants on the Noise Loudness dimension.

(55) Stop burst noise loudness scale for the C-Cj contrast

| Noise Loudness | |
|---|---|
| 0 | 1 |
| $\Delta$(p-$p^j$) | $\Delta$(t-$t^j$) |

The formant transition difference between plain and palatalized consonants, as mentioned above, is bigger for the labials than for the coronals (Padgett 2001, Kochetov 2001:93, Figure 15). From this we posit a distinctiveness scale (56): plain and palatalized coronals are closer to each other than plain and palatalized labials. The exact numbers in scales (55) and (56) are not based on any experimental results, but are hypothetically assigned for the DT analysis. It is the comparison between labial and coronal contexts that matters rather than the precise values.

(56) Formant transition scale for the C-Cj contrast

| Formant Transition | | | |
|---|---|---|---|
| 0 | 1 | 2 | 3 |
| | $\Delta$(t-$t^j$) | $\Delta$(p-$p^j$) | $\Delta$(t-tj)<br>$\Delta$(p-pj) |

Note that I am making an assumption that the difference between $C^j$ and Cj mainly resides in the vocalic formant cues, rather than the consonantal cues, which is in line with the observation in the comparisons between the labialized consonants and the Cw clusters. Thus the bigger contrast between C and Cj than that between C and $C^j$ is incorporated in the scale of Formant Transition here.

Given the distinctiveness scales for the C-Cj contrast, we now turn to the consonant place asymmetries in Cj combinations. In Ukrainian, where a palatalization contrast does not exist on labials, the required distance between plain and palatalized consonants must exceed that between /p/ and /$p^j$/, as shown by the constraint ranking in (57). The contrast between /p/ and /pj/ satisfies the Minimal Distance constraint, despite its violation of the lower-ranked *CGV (58).

(57) Minimal Distance= NL:1 or FT:3 >> MAXIMIZECONTRASTS >> *CGV >> *$C^GV$

(58) Ukrainian

| pa $p^ja$ pja | MinDist=NL:1 or FT:3 | MAXIMIZECONTRASTS | *CGV | *$C^GV$ |
|---|---|---|---|---|
| a. pa-$p^ja$ | *! | | | |
| b. pa | | *! | | |
| ☞c. pa-pja | | | * | |

On the other hand, coronal stops, due to the difference in the consonant release noise loudness, are allowed to host a palatalization contrast. Though candidates (59a) and (59c) tie in the two top ranking constraints, the articulatory gesture coordination constraint chooses /t/-/$t^j$/ over /t/-/tj/.

(59) Ukrainian

| ta $t^ja$ tja | MinDist=NL:1 or FT:3 | MAXIMIZECONTRASTS | *CGV | *$C^GV$ |
|---|---|---|---|---|
| ☞a. ta-$t^ja$ | | | | * |
| b. ta | | *! | | |
| c. ta-tja | | | *! | |

Languages in which both labial and coronal stops contrast with their palatalized counterparts, such as Bulgarian, have different constraint ranking, in which MAXIMIZECONTRASTS and *CGV dominate MinDist=NL:1 or FT:3 (60).

(60) Bulgarian: /p/-/$p^j$/

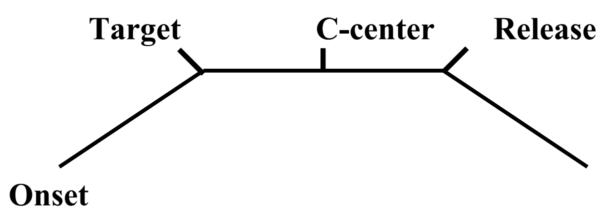
| pa $p^ja$ pja | MinDist=NL:1 or FT:2 | MAXIMIZE CONTRASTS | *CGV | MD=NL:1 or FT:3 |
|---|---|---|---|---|
| ☞a. pa-$p^ja$ | | | | * |
| b. pa | | *! | | |
| c. pa-pja | | | *! | |

To summarize: the interaction of the perceptual distance requirement between contrastive sounds and the principles of articulatory gesture coordination in the current analysis led us to expect asymmetric patterns of CG combinations depending on the consonant places not only in Cw cluster/labialization but also in Cj cluster/palatalization distribution. This prediction is borne out, as shown by the Ukrainian Cj combination distribution.

Before ending this section, let us compare the current analysis with the Articulatory-Phonological analyses (Browman and Goldstein 2000, Gafos 2002) in terms of the articulatory gesture phasing of CG combinations. In 5.3 I posited constraints against prevocalic CG combinations and their universal ranking (*CGV >> *$C^GV$) based on the coordination requirement between consonant and vowel gesture (C-V relation) of Articulatory Phonology. These constraints are essentially the conflicting force to the contrast dispersion requirements in DT. Articulatory Phonology also acknowledges another phasing specification in gestural coordination, namely the C-C relation, whose functional goal is to ensure the recoverability of gestures. Recall that this was why the C and G gestures in (41b) and (41c) are displaced from the optimal C-V relation (41a). The actual phasing patterns between two consonant gestures emerge at the point where the balance between C-V relation and recoverability is achieved (*Self-organization*, Browman and Goldstein 2000:30). The recoverability of a stop gesture will suffer more from overlap with another stop gesture than with an approximant. So the balance between C-C and C-V relations may be achieved when the two C gestures are farther from each other in stop-stop sequences than in stop-approximant sequences.

Though this suggests that the gestural phasing patterns may differ depending on the types of consonant gestures involved, it is not entirely clear how the cross-linguistic variation in the coordination of similar gestures emerges, as pointed out by Gafos (2002). Gafos argues that the C-C relation should be separate from the recoverability requirement, as languages show different phasing patterns for the same type of cluster: for example, some languages have audible release of the first consonant in a cluster (e.g., Piro, Moroccan Arabic), whereas languages like English do not. To capture the cross-linguistic difference in phasing relation between two C gestures, Gafos postulates CC-Coordination constraints that are parameterized to languages. These constraints specify inter-gestural coordination as the synchronicity of one *landmark* within the first gesture and another landmark within the second gesture, where landmarks are the dynamic states that a gesture goes through as the articulator attains a constriction and moves away from it. The diagram in (61), adapted from Gafos (2002), illustrates the landmark specification of a gesture.

(61) Gafos (2002: 276, 279)

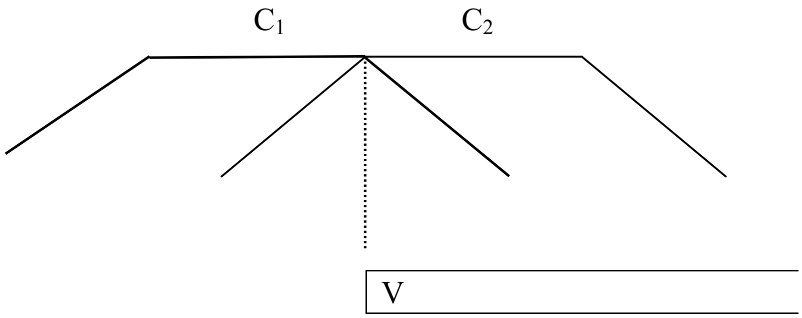



**Onset:** The onset of movement toward the target of the gesture
**Target:** The point in time at which the gesture achieves its target
**C-center:** The mid-point of gestural plateau
**Release:** The onset of movement away from the gestural target

Language-specific C-C coordination is expressed in the form of alignment constraints, such as (62a) and (63a). In the gesture coordination shown in (62), the release of the first gesture is synchronous with the target of the second gesture, and in (63) the former is synchronous with the c-center of the second gesture.

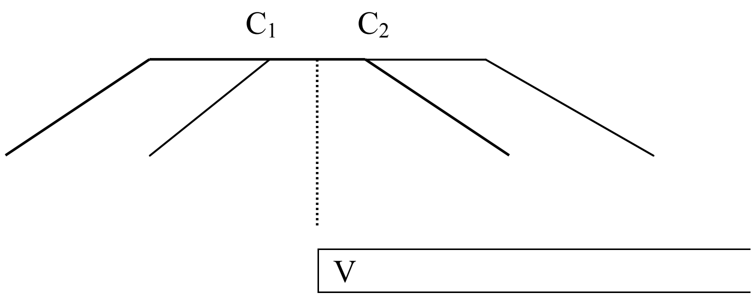
(62) a. Align ($C_1$, release, $C_2$, target)  (Gafos 2002:320, (30a))

b.




(63) a. Align ($C_1$, release, $C_2$, c-center)  (adapted from Gafos 2002:282)

b.




The two gestural landmark coordination patterns in (62b) and (63b) seem to be plausible representations of the Spanish Cw clusters and the Korean labialized consonants, respectively. Let us assume that $C_1$ is the tongue gesture for a coronal or velar stop, and $C_2$ is a lip rounding gesture. In (62b) the target of the lip rounding gesture is achieved at the release of the stop closure and the vowel onset. The acoustic result of this is a rather constant period of formant frequencies for the /w/ component (i.e., not affected by the vowel formants), which is exactly what we found in the F2 at onset of the Spanish Cw clusters. In contrast, half of the /w/ gesture plateau—from target to c-center—is overlapped with the stop closure in (63b). The glide resonance starts only after the release of the stop ($C_1$), i.e., at the c-center of lip rounding gesture, at which point the vowel target may have been almost achieved. The acoustic consequence of this gestural configuration is the formants at vocoid onset varying depending on the vowel quality, the result observed in the F2 at onset measure of the Korean Cw combinations.

Although the gestural landmark coordination suggested in Gafos (2002) provides a means to schematize the difference between CG clusters and secondarily-articulated consonants, I argue that the cross-linguistic variation in CG combinations are not imposed by the gestural coordination constraints themselves. If languages employ parameterized CC-Coordination constraints such as (62a) and (63a), language-internal variations in C-G coordination, e.g., the place asymmetries in the status of CG combination in Dan and Ukrainian, are not expected to occur. An additional problem is the stipulation of a separate Recoverability constraint, a universal constraint banning complete overlap of gestures, in addition to the CC-Coordination, which controls the language-particular degree of gestural overlap.

Instead of positing language-particular constraints specifying the coordination between gestural landmarks, the current analysis captures the cross-linguistic as well as language-internal variations in the phonetic status of CG combinations, with the different ranking relations between universal constraints (MINIMAL DISTANCE and MAXIMIZE CONTRASTS constraints). Since the acoustic-auditory distance between C and Cw varies between consonant places as well as C-w gesture coordination types, the language-internal place asymmetries in the status of CG combinations do not raise any problem. Furthermore, minimal distance constraints encompass both recoverability and the C-C relation, as 0 distance between $C_1$ and $C_1C_2$ corresponds to the lack of recoverability. So there is no need to postulate an additional Recoverability constraint as a primitive.

### 4.5. An additional complex case: Ponapean labialized labials

This section discusses a potentially problematic case for the Dispersion-Theoretic view of the asymmetry in the Cw distribution. While the labials are expected to be the least preferred place for the labialization contrast in the current analysis because of the weak acoustic/ perceptual cues, there exist languages that employ a labialization contrast exclusively on the labial consonants, such as Ponapean (Rehg and Sohl 1981, Goodman 1995). In other words, though a labialization contrast occurs most often on velar consonants, there is no predictability (implicational relationship) regarding the presence of labialization at different consonant places. However, we see that in these special cases, the contrast is enhanced by additional differences between the syllables.

According to Goodman's (1995) measurements, the F2 of the vowel steady state is considerably lower after labialized labials than after plain labials in Ponapean. For example, the average second formant at the steady state of the vowel /e/ was 1796 Hz after /p/, but 1379 Hz after /$p^w$/ (Goodman 1995:286).[30] The vowel peak formant difference between CV and CwV in Ponapean can be understood as a contrast enhancement, to ensure that PV and PwV are sufficiently distinctive from each other. In other words, the labialized labials do not yield sufficient acoustic contrast from the plain labials, so the vowel is produced in backer tongue position if it is next to a labialized labial consonant, as a way of increasing the perceptual distance from the plain labials. The importance of steady state vowel formants for the p-$p^w$ contrast in Ponapean is also evidenced by the consonant labialization harmony of Ponapean: adjacent labial consonants in a Ponapean morpheme agree in labialization (*P V $P^w$). If two labials differ in labialization, the flanked vowel would bear ambiguous cues between labialized and plain labial consonants, so it would be very difficult to maintain the difference between *PV$P^w$ versus *$P^w$VP. Thus, to improve the weak perceptibility of labialization contrast on labials, Ponapean employs other means such as backing of the vowel and consonant harmony.

Interestingly, in Ulithian, a language related to Ponapean (both Micronesian), a labialized counterpart of /p/ exists as a labialized labiodental fricative /$v^w$/ (Sohn and

[30] The F2 of the vowel steady state after the Korean plain and labialized labials is not as different as in the the Ponapean case: /p'e/ 2018 Hz, /$p'^we$/ 1925 Hz; /p'a/ F2 1372 Hz and /$p'^wa$/ F2 1298 Hz.

Bendor 1973). Such contrast-enhancing strategies (e.g., vowel quality, consonant manner) are not reported for the contrastive labialization on velar consonants. Given that the labialization contrast on velars is inherently more salient than that on labials, these strategies are less necessary for the velar labialization contrast.

### 4.6 The role of distinctiveness scales in the loanword adaptation of Cw clusters in Korean

In the previous sections I argued that the commonly observed CG co-occurrence restrictions or the gaps in the inventories of the secondarily-articulated consonants emerge because languages avoid contrasts between strings that are too similar. The role of the perceptual distance scales such as (1) and (2) was important in the evaluation of the contrast. This section presents an analysis for the loanword adaptation patterns of the Cw clusters into Korean introduced in Chapter 1, which shows that the perceptual distance scale that I postulated for the labialization contrast evaluation also play a role in a correspondence relationship of traditional OT phonology, namely the mapping between the source word and the loanword form.

Korean does not have consonant clusters in syllable onsets or codas, and in loanword adaptation consonant clusters in the source word are generally broken by the epenthetic vowel /ɨ/ (e.g., *ski* → sɨ.k$^{h}$i). The same strategy is employed in the adaptation of coronal + /w/ clusters: in the loanword forms of words containing /tw/ clusters, the consonant and the glide /w/ are separated into two syllables by an epenthetic vowel /ɨ/. English /kw/ clusters, in contrast, are always mapped to Cw combinations in Korean.[31] Data in (64) are repeated from Chapter 1.

| | | | | |
|---|---|---|---|---|
| (64) | *Twain* [t$^{h}$weɪn] | t$^{h}$ɨ.we.in | *t$^{hw}$e.in | |
| | *Quayle* [k$^{h}$weɪl] | *k$^{h}$ɨ.we.il | k$^{hw}$e.il | |

Given that both coronal and velar consonants are allowed to occur in Cw combinations (t$^{w}$edʒi 'pig', t'$^{w}$aɾi 'coiled object'; k$^{w}$e 'chest', k'$^{w}$aɾi 'ground cherry') in Korean, the discrepancy between coronal and velar in (64) is unexpected.

Let us start with the question of why Cw clusters are not always mapped to Cw combinations of Korean, even in the environments where the Korean phonotactics allow them. This may be related to the different temporal relationship of C and /w/, and the consequent acoustic divergence between the Cw combinations of the two languages. To preserve the drastic F2 change from /w/ to the following vowel, or the considerable duration of the low F2 in /w/ of the source language, the glide must not be preceded by a tautosyllabic consonant in the borrowing form, since Korean /w/ is realized as labialization on the preceding consonant.

[31] There is a possibility that the non-uniform adaptation of Cw clusters is purely a result of orthographic adaptation. It is true that in the English /kw/ is almost always spelled as *qu* in English, suggesting the possibility that the source words spelled with *w* are adapted with an inserted vowel between C and *w* (*Twain, sweet*), whereas the ones spelled with *u* (*Quinn, Quaker*) are not. However, it still remains unclear why spelling CuV should map to CwV in Korean. Furthermore, loanwords from languages other than English do not always fall under this spelling-based dichotomy: *suede* → sɨ.we.i.tɨ, *Kwashiorkor* → k$^{h}$wʌ.si.o.k$^{h}$ʌ.

However, vowel insertion between C and /w/ results in the alteration of the consonant release noise: the low frequency concentration of the stop burst noise before /w/ in the source word is lost when the stop is before the epenthetic vowel /ɨ/ in the adapted form. Thus there is a trade-off between the preservation of the burst noise frequency and that of the formant transition; the latter is chosen for Tw, whereas Kw opts for the former. I argue that this pattern arises because the difference in the noise quality of the burst between plain C and Cw/$C^w$ is more noticeable in the velar stop. Recall that the plain and the labialized velars are farther apart in the noise frequency scale than plain and labialized coronals, as shown in (2), and summarized in (65):

(65) Noise Frequency
$\Delta$(k-kw) > $\Delta$(t-tw)

I follow Steriade's (2001) proposal that the ranking between faithfulness constraints is projected from the degree of perceptual difference between the faithful output and the output that violates each faithfulness constraint. The application of this hypothesis to the present case predicts that the alteration of the noise frequency of /kw/ to that of /k/ is punished by a higher-ranking constraint than the change from /tw/ to /t/.

Fleischhacker (2001, 2005) applies Steriade's (2001) p-map hypothesis to the asymmetric vowel insertions in loanword adaptation. In loanword patterns observed in languages like Egyptian Arabic, Hindi and Kazakh, ST clusters (/s/ + voiceless stop), are adapted with a vowel added in front of the cluster, whereas a vowel is inserted between two consonants in TR (voiceless stop + liquid) clusters. In Fleischhacker (2001, 2005) this discrepancy is attributed to the fact that the perceptual distance between ST and VST is smaller than that between ST and SVT, whereas TR and TVR are perceptually closer than TR and VTR. The relative difference in the perceptual distance projects rankings of faithfulness constraints: the bigger the perceptual distance is between the faithful output and the epenthesized form, the higher the DEP constraint against the epenthesis is ranked (66).

(66) DEP-V/S_T >> DEP-V/_ST
DEP-V/_TR >> DEP-V/T_R

I propose that the perceptual distance scale given in (65) projects the ranking of DEP constraints (67).

(67) DEP-V/k_w >> DEP-V/t_w

The case at hand differs from the ST-TR epenthesis asymmetry in that the borrowing language (Korean) does allow Cw combinations, but still the vowel epenthesis sometimes occurs when words containing Cw clusters are borrowed from English or other languages. Since the Korean prevocalic /w/ is phonetically realized as secondary labialization on the preceding consonant, the mapping of the English Cw clusters to the native labialized consonants may be penalized by a faithfulness constraint in favor of preserving the cluster-like phonetic status of Cw, e.g., FAITH (w).[32] The DEP-V constraints in (67) can be in different ranking relations with respect to this faithfulness constraint (68).

(68) DEP-V/k_w >> FAITH (w) >> DEP-V/t_w

As a result of the ranking in (68), the loanword candidate with epenthetic vowel is worse than that with the "weak" glide (i.e., the realization of /w/ as labialization) in the velar stop context (e.g., *Quayle*), but the opposite holds for the coronal stop context (e.g., *Twain*). Tableaux (69) and (70) show how the different adaptation strategies are chosen for words containing /tw/ or /kw/ clusters.

(69) *Twain*

| tʰweɪn | DEP-V/k_w | FAITH (w) | DEP-V/t_w |
|---|---|---|---|
| a. tʰʷe.in | | *! | |
| ☞b. tʰɨ.we.in | | | * |

(70) *Quayle*

| kʰweɪl | DEP-V/k_w | FAITH (w) | DEP-V/t_w |
|---|---|---|---|
| ☞a. kʰʷe.il | | * | |
| b. kʰɨ.we.il | *! | | |

If this analysis is on the right track, it serves as another piece of evidence that perceptual distance plays an important role in the cluster resolution in the loanword adaptation (Fleischhacker 2001, 2005), and that the perceptual distance scales such as (65) are not only necessary for the Dispersion-theoretic evaluation of contrasts, but also for a more familiar output-output correspondence relation, such as loanword adaptation.[33]

## 4.7 Summary

This chapter presented an analysis of the distributional asymmetries of Cw combinations (Cw clusters and labialized consonants), within the Dispersion Theory of Contrast (Flemming 1995, 2002). Based on the experimental results from the previous chapters, the perceptual distance between plain C and Cw combinations of each consonant and vowel context was calculated along two different auditory dimensions that are known to be relevant for the perception of postconsonantal/prevocalic glide /w/, i.e.,

[32] The precise nature of FAITH (w) is not crucial to the analysis, but one possibility is the UNIFORMITY constraint (McCarthy and Prince 1999), which rules out many-to-one correspondence or coalescence ($C_1w_2 \rightarrow C^w_{12}$).

[33] Kang (2006) proposes a different explanation for the asymmetry between English /tw/ and /kw/ borrowed into Korean. She proposes that it is due to the clearly U-shaped formant transition in English twV syllables, which was also observed in Spanish (Figure 2.13), that leads Korean listeners to perceive English twV as t+ /ɨ/ + wV. In contrast, the F2 transition in English kwV is "not as prominently U-shaped" as in twV, so Korean listeners map it to the kwV syllable in their L1.

formant transition and the consonant release noise frequency. The dimension of formant transition incorporates frequency change, transition direction and the durational aspects of formant trajectory, so it reflects the perceptibility differences related to the timing relation of the consonant and the glide component—i.e., whether it is overlapping or sequential—as well as those related to the consonant places and vowels.

The co-occurrence restrictions on Cw combinations were attributed to the functional goal of maintaining the perceptual distance between a plain consonant and a Cw combination. More specifically, the contrasts between a Cw combination and its plain C counterpart, between which the perceptual distance does not exceed certain degree, are ruled out as a result of an Optimality-Theoretic interaction between minimal distance requirement constraints, and a constraint against merging potentially contrastive items. An alternative feature identity avoidance analysis was suggested for Korean, and rejected on the basis of the arbitrary conjunction of unrelated constraints, and its failure to extend to Mandarin, which shows a strikingly similar pattern in Cw distribution to Korean.

The DT analysis was also able to account for the consonant place asymmetries related to the phonetic status of the /w/ component, i.e., secondary articulation or clusters. Cw combinations are realized differently across the consonant places—/$k^w$/ but /tw/ and /pw/—in languages like Dan and Tarascan. The contrast between /k/ and /$k^w$/ is inherently more salient than the labialization contrast on other consonants, due to the strong consonant release cue. Thus, the realization of /t + w/ and /p + w/ as clusters in these languages can be understood as an enhancement of contrast from the plain /t/ and /p/ onsets, respectively, which makes it possible for them to satisfy the minimal distance requirement. The fact that such contrast enhancement in labial and coronal contexts does not occur in languages like Cantonese, and the contrast is simply absent, is attributed to the high ranking articulatory constraint against the realization of the /w/ component as a glide consonant rather than the secondary labialization.

It was also argued that the minimal distance requirements imposed on the C-CG contrast also give rise to the different C-/j/ timing relationships between labial and coronal in languages like Ukrainian, as well as the general preference of coronal place for contrastive palatalization. The availability of the inherent consonant release cue—affrication—makes coronal place a better context for the plain-palatalized consonant contrast, i.e., a contrast in which the /j/ glide component can be realized as a secondary palatalization without violating the minimal perceptual distance requirement.

Lastly, the perceptual distance scales put forth for the DT analysis of the Cw combination distribution, in particular the stop burst noise frequency scale, was argued to play a role in another part of the Korean phonology, namely the loanword adaptation of the English Cw clusters. The consonant place-specific adaptation strategies are attributed to the different degrees of the perceptual distance between plain and labialized stop burst noise, from which the ranking of DEP (V) constraints in different consonant place contexts is projected under the P-map hypothesis (Steriade 2001).

# Chapter 5. Conclusion

In this thesis I presented a contrast-based account of the asymmetric distribution of Cw combinations found in many languages (Kawasaki 1982, Maddieson and Precoda 1992), *i.e.*, the avoidance of Cw combinations in labial consonant and back vowel contexts, and the preference for secondary labialization on velar consonants. I showed that phonetic salience is an important factor leading to commonly found co-occurrence restrictions. In particular, I argued, following Flemming (1995), that the distribution of Cw combinations in languages like Mandarin and Korean is related to the acoustic and perceptual distinctiveness of plain vs. labialized consonants, which varies depending on the consonant places and vowels.

Experimental studies of the acoustic and perception of the C-Cw contrasts were conducted to test the hypothesis that there is a correlation between acoustic/perceptual salience of the plain C versus Cw contrast and the distribution of Cw combinations. The results showed that Cw combinations found in Korean morpheme initial position differ from their plain onset counterparts to a greater extent than those that are not attested in that position. A Dispersion-Theoretic analysis was developed, in which a ranking of minimal distance constraints and a contrast maintenance constraint evaluates well-formedness of the labialization contrasts, rather than labialized consonants per se. I showed that the current contrast-based analysis is advantageous over an alternative analysis based on the feature identity avoidance principle (OCP), because 1) it can account for the gap in plain consonants and the presence of labialized consonants in some vowel contexts in Mandarin, and 2) there is no need to postulate an arbitrary conjunction of two constraints such as OCP and *Cw. The current analysis also differs from Kawasaki's (1982) account of co-occurrence restrictions based on measures of acoustic salience from formant trajectories, which does not predict the predominance of labialized velar consonants. Cues from both formant transition and consonant release were considered in the current analysis, capturing the importance of consonant release cues, *e.g.,* stop burst noise frequency, in signaling secondary labialization contrasts.

Based on the acoustic comparison between Spanish and Korean Cw combinations, I also argued that asymmetries in the phonetic realization of Cw combinations, found in languages like Dan, are also related to the distinctiveness requirement of contrasts. More specifically, I proposed that a trade-off between articulatory gesture coordination preferences and contrast distinctiveness requirements gives rise to the pattern in which velar consonants, which provide an inherently strong consonant release cue, host contrastive labialization, a realization preferred by articulatory gesture coordination principles. In contrast, other consonants, which provide less strong contrastive cues, form less-overlapped combinations, i.e., Cw clusters, thereby increasing distinctiveness from their plain C counterparts. I showed that a similar analysis is applicable to the distribution of contrastive palatalization and Cj clustering in

Ukrainian, as coronal stops, which are associated with strong affrication noise before palatal release, host secondary palatalization, while labial consonants form Cj clusters.

While Dispersion Theory is successful in accounting for asymmetries in inventory structures and co-occurrence restrictions by appealing to principles of speech perception and production, whether or not such knowledge is part of the human language faculty still remains as a question. From another perspective, the synchronic co-occurrence restrictions may be merely the results of phonetically natural language changes (Blevins 2004, Ohala 1992), and specifying them as part of the cognitive computation system would thus be redundant. In order to address the question of whether contrast dispersion is driven by grammatical principles existent in human cognition, a necessary step would be to examine the learnability of the distinctiveness requirement through artificial language learning experiments suggested in research such as Wilson (2006), Moreton (2008) and Moreton et al. (2008).

## References


Ahn, S.-C. 1985. The interplay of phonology and morphology in Korean. Doctoral dissertation, University of Illinois at Urbana-Champaign.

Avery, P., E. Dresher and K. Rice (eds.). 2008. *Contrast in Phonology: Perception and Acquisition.* Berlin: Mouton de Gruyter.

Bakovic, E. 2000. *Harmony, Dominance, and Control.* Ph.D. dissertation, Rutgers University.

Bauer, R. and P. Benedict. 1997. *Modern Cantonese Phonology.* Berlin/New York: Mouton de Gruyter.

Bearth, T. and H. Zemp 1967. The phonology of Dan (Santa). *Journal of African Languages* 6, 9-29.

Blevins, J. 2004. *Evolutionary Phonology*. Cambridge: Cambridge University Press.

Blumstein, S. E. and K. N. Stevens. 1979. Acoustic invariance in speech production: evidence from measurements of the spectral characteristics of stop consonants. *Journal of Acoustical Society of America* 66, 1001-1017.

Boersma, P. and D. Weenink. 2006. Praat: doing Phonetics by computer (Version 4.4.16).

Bonneau, A. 2000. Identification of vocalic features from French stop bursts. *Journal of Phonetics* 28, 495-502.

Bonneau, A. L. Djezzar and Y. Laprie. 1996. Perception of the place of articulation of French stop bursts. *Journal of Acoustical Society of America* 100, 555-564.

Borden, G., K. Harris and L. Raphael. 1994. *Speech Science Primer: Physiology, Acoustics, and Perception of Speech.* Baltimore: Williams & Wilkins.

Borowsky, T. 1986. Topics in the lexical phonology of English. Doctoral dissertation, University of Massachusetts, Amherst.

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

Browman, C. and L. Goldstein. 1988. Some notes on syllable structure in Articulatory Phonology. *Phonetica* 45, 140-155.

Browman, C. and L. Goldstein. 1989. Articulatory gestures as phonological units. *Phonology* 6, 201-251.

Browman, C. and L. Goldstein. 2000. Competing constraints on gestural coordination and self-organization of phonological structures. *Les Cahiers de l'ICP, Bullletin de la Communication Parleé* 5, 25-34.

Castañeda, M. L. 1986. 'El VOT de las oclusivas sordas y sonoras españolas'. *Estudios de Fonética Experimental* 2, 91-110.

Chan, M. 1985. Fuzhou Phonology: a non-linear analysis of tone and stress. Doctoral dissertation, University of Washington, Seattle.

Chen, M. 2000. Acoustic analysis of simple vowels proceeding a nasal in standard Chinese. *Journal of Phonetics* 28, 43-67.

Choi, J. D. 1995. An acoustic-phonetic underspecification account of Marshallese vowel allophony. *Journal of Phonetics* 23, 323-347.

Clements, G. N. 1990. The role of the sonority cycle in core syllabification. In J. Kingston and M. Beckman (eds.), *Papers in laboratory phonology 1: Between the grammar and physics of speech,* 283-333. Cambridge: Cambridge University Press.

Clements, G. N. 1991. Place of articulation in consonants and vowels: a unified theory. Working papers of the Cornell Phonetics Laboratory 5, 77-123. Ithaca: Cornell University.

Clements, G. N. and E. Hume. 1995. Internal organization of speech sounds. In J. Goldsmith (ed.), *The Handbook of Phonological Theory*, Cambridge: Blackwell.

Clement, G. N. and S. J. Keyser. 1983. *CV Phonology: a Generative Theory of the Syllable.* Cambridge, MA: MIT Press.

Comrie, B. and G. Corbett (eds.). 1993. *The Slavonic Languages.* New York: Routledge.

Cullinan, W. and M. Tekieli. 1979. Perception of vowel features in temporally-segmented noise portions of stop-consonant CV syllables. *Journal of Speech and Hearing Research* 22, 122-131.

Dahlstrom, A. 1997. Fox reduplication. *International Journal of American Linguistics* 63, 205-226.

Delattre, P., A. Liberman and F. Cooper. 1955. Acoustic loci and transitional cues for consonants. *Journal of the Acoustical Society of America* 27, 769-73.

de Jong, K. and S. Obeng 2000. Labio-palatalization in Twi. Contrastive, quantal, and organizational factors producing an uncommon sound. *Language* 76, 682-703.

de Manrique, A. and M. Massone. 1981. Acoustic analysis and perception of Spanish fricative consonants. *Journal of Acoustical Society of America* 69, 1145-1153.

Duanmu, S. 1990. A formal study of syllable, tone, stress and domain in Chinese languages. Doctoral dissertation, Massachusetts Institute of Technology.

Duanmu, S. 2000. *The Phonology of Standard Chinese*. Oxford: Oxford University Press.

Fant, G. 1960. *Acoustic theory of speech production*. The Hague: Mouton.

Fant, G. 1973. *Speech Sound and Features.* Cambridge: MIT Press.

Fleischhacker, H. 2001. Cluster-dependent epenthesis asymmetries. In A. Albright and T. Cho (eds.), Papers in Phonology 5, UCLA Working Papers in Linguistics, 71-116.

Fleischhacker, H. 2005. Similarity in phonology: evidence from reduplication and loan adaptation. Doctoral dissertation, University of California, Los Angeles.

Flemming, E. 1995. Auditory representations in Phonology. Doctoral dissertation, University of California, Los Angeles.

Flemming, E. 2002. *Auditory Representation in Phonology.* New York: Routledge.

Flemming, E. 2006. The role of distinctiveness constraints in phonology. Ms, Massachusetts Institute of Technology.

Flemming, E. 2007. Stop place contrasts before liquids. Proceedings of International Congress on Phonetic Science 16, 233-236. Saarbrücken: Saarland University.

Foster, M.L. 1968. *The Tarascan Language.* Berkeley and Los Angeles: University of California Press.

Friedrich, P. 1975. *A Phonology of Tarascan.* Series in social, cultural and linguistic anthropology 4. University of Chicago studies in anthropology.

Gafos, A. 2002. A grammar of gestural coordination. *Natural Language and Linguistic Theory* 20, 269-337.

Gick, B. 2003. Articulatory correlates of ambisyllabicity in English glides and liquids. In Local, J., R. Ogden and R. Temple (eds.), *Papers in Laboratory Phonology* VI, Cambridge: Cambridge University Press.

Goldsmith, J. 1976. *Autosegmental Phonology.* New York: Garland.

Goodman, B. 1995. Features in Ponapean Phonology. Doctoral dissertation, Cornell University.

Gordon, M., P. Barthmaier and K. Sands. 2002. A cross-linguistic acoustic study of voiceless fricatives. *Journal of International Phonetic Association* 32, 141-174.

Green, A. 1997. The prosodic structure of Irish, Scots Gaelic, and Manx. Doctoral dissertation, Cornell University.

Greenlee, M. 1992. Perception and production of voiceless Spanish fricatives by Chicano children and adults. *Language and Speech* 35, 173-187.

Han, J.-I. 1992. On the Korean Tensed Consonants and Tensification. *CLS* 28, 206-223.

Harrington, J. and S. Cassidy. 1999. *Techniques in Speech Acoustics*. Dordrecht: Kluwer Academic Publishers.

Harris, J. 1983. *Spanish Phonology.* Cambridge: MIT Press.

Hayes, B. 1986. Inalterability in CV phonology. *Language* 62, 321-351.

Hualde, J. I. 2005. *The Sounds of Spanish*. Cambridge: Cambridge University Press.

Hume, E. and K. Johnson (eds.). 2001. *The Role of Speech Perception in Phonology.* San Diego, CA: Academic Press.

Ito, J. and A. Mester. 2003. Systemic markedness and faithfulness. *CLS* 39, 665-689.

Ito, C., Y. Kang and M. Kenstowicz. 2006. The adaptation of Japanese loanwords into Korean. MIT Working Papers in Linguistics 52, 65-104.

Jongman, A., R. Wayland and S. Wong. 2000. Acoustic characteristics of English fricatives. *Journal of Acoustical Society of America* 108, 1252-1263.

Johnson, K. 2003. *Acoustic and Auditory Phonetics* (Second Edition). Oxford: Blackwell.

Kang, H.-S. 1997. Phonological variation in glides and diphthongs of Seoul Korean: its synchrony and diachrony. Doctoral dissertation, Ohio State University.

Kang, H. 2006. An acoustic study of the perceptual significance of F2 transition of /w/ in English and Korean. *Speech Science* 13 (4), 7-21.

Kawasaki, H. 1982. An acoustic basis for universal constraints on sound sequences. Doctoral dissertation, University of California at Berkeley.

Kenstowicz, M. 1994. Phonology in Generative Grammar. Oxford: Blackwell.

Kim, C.-W. 1965. On the Autonomy of the Tensity Feature in Stop Classification. *Word* 21, 339-359.

Kim, C.-W. 1968. The vowel system of Korean. *Language* 44, 516-527.

Kim, H. 1999. The place of articulation of Korean affricate revisited. *Journal of East Asian Linguistics* 8, 313-347.

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

Kim, H. 2002. Closure Duration, Glottal Width and Glottal Tension in Korean Three-way phonation consonants. Paper presented at the 8$^{th}$ Conference on Laboratory Phonology, Yale University.

Kim, S. and Y. Heo. 1999. Sunginceyyak cokenkwa hankuwke moumcheykyey. (Licensing conditions and the Korean vowel system.) *Enehak* 25, 55-75.

Kochetov, A. 2002. *Production, perception and emergent phonotactic patterns: a case of contrastive palatalization.* New York: Routledge.

Kochetov, A. 2006. Licensing by cue: a case of Russian palatalized coronals. *Phonetica* 63, 113-148.

Kochetov, A. 2008. Perception of gestural overlap and self-organizing phonological contrasts. In P. Avery, E. Dresher, and K. Rice (eds.), 173-196.

Kochetov, A. and L. Goldstein. 2006. Optimal timing patterns and recoverability of the palatal gesture. Paper presented at 2006 Meeting of Canadian Linguistic Association, York, ON.

Kwon, K. 2001. Hyentaykuwkeeyseuy moumcheykyey pyenhwauy uwmcikimey tayhayye. (On the movement of the vowel system change in the contemporary Korean.) *Enehak* 30, 29-48.

Ladefoged, P. 1968. *A Phonetic Study of West African Languages.* Cambridge: Cambridge University Press.

Ladefoged, P. 1971. *Preliminaries to Linguistic Phonetics.* Chicago: University of Chicago Press.

Ladefoged, P. and I. Maddieson. 1996. *The Sounds of the World's Languages.* Cambridge, MA: Blackwell.

Lahiri, A., L. Gewirth and S. E. Blumstein. 1984. A reconsideration of acoustic invariance for place of articulation in diffuse stop consonants: evidence from a cross-language study. *Journal of Acoustical Society of America* 76, 391-404.

Leben, W. 1973. Suprasegmental phonology. Doctoral dissertation, Massachusetts Institute of Technology.

Lee, B.-G. 1982. A well-formed condition on syllable structure. In I. Yang (ed.), *Linguistics in the Morning Calm.* Seoul: Hanshin.

Lee, H.-B. 1999. Korean. In *Handbook of the International Phonetic Association,* 120-123. Cambridge: Cambridge University Press.

Lee. H.-Y. 1996. *Korean Phonetics* [In Korean]. Seoul, Korea: Thaehaksa.

Lee. Y.-S. 1994. Onset analysis of Korean on-glides. In Y. Kim-Renaud (ed.), *Theoretical Issues in Korean Linguistics.* Stanford Linguistic Society.

Li, W.-C. 1999. *A Diachronically-motivated Segmental Phonology of Mandarin Chinese.* New York: Peter Lang Publishing Inc.

Li, Y. 2008. Mandarin speakers' production of English and Mandarin post-vocalic nasals: an acoustic approach. M.A. Thesis, University of Victoria.

Liang, C.-Y. 2002. Acoustic analysis and synthesis of nasal codas in Mandarin Chinese. Master's thesis, Massachusetts Institute of Technology.

Li, F., J. Edwards, M. Beckman. 2007. Spectral measures for sibilant fricatives of English, Japanese, and Mandarin Chinese. In J. Trouvain and W. Barry (eds.), Proceedings of the 16th International Congress on Phonetic Science, 917-920. Saarbrücken: Saarland University.

Liberman, A., P. Delattre, L. Gerstman and F. Cooper. 1956. Tempo of frequency change as a cue for distinguishing classes of speech sounds. *Journal of Experimental Psychology* 52, 127-137.

Liberman, A. F. Cooper, D. Shankweiler and M. Studdert-Kennedy. 1967. Perception of the speech code. *Psychological Review* 74, 431-461.

Lisker, L. and A. Abramson 1964. A cross-linguistic study of voicing in initial stops: acoustical measurements. *Word* 20, 384-422.

Lubowicz, A. 2005. Locality of conjunction. In J. Alderete et al. (ed.), Proceedings of the 24th West Coast Conference on Formal Linguistics, 254-262. Somerville, MA: Cascadilla Proceedings Project.

Lynch, J. 1978. *A Grammar of Lenakel.* Pacific Linguistics Series 55. Canberra: Australian National University.

MacMillan, N. and C. Creelman. 1991. *Detection Theory: a User's Guide*. Cambridge: Cambridge University Press.

Maddieson, I. 1984. *Patterns of Sounds*. Cambridge: Cambridge University Press.

Maddieson, I. and K. Precoda. 1992. UPSID and PHONEME. University of California, Los Angeles.

Mazzaro, N. 2005. Aspiration and velarization of /f/ in Argentine Spanish. Toronto Working Papers in Linguistics 25, 58-67.

McCarthy, J. 1986. OCP effects: Gemination and antigemination. *Linguistic Inquiry* 17, 207-63.

McCarthy, J. and A. Prince. 1999. Faithfulness and identity in prosodic morphology. In R. Kager, H. van der Hulst, and W. Zonneveld eds., *The Prosody-Morphology Interface.* Cambridge: Cambridge University Press.

Miao, R. 2005. Loanword adaptation in Mandarin Chinese: perceptual, phonological and sociolinguistic factors. Doctoral dissertation, Stony Brook University.

Moreton, E. 2002. Phonological grammar in speech perception. Doctoral dissertation, University of Massachusetts, Amherst.

Moreton, E. 2008. Analytic bias and phonological typology. *Phonology* 25(1), 83-127

Moreton, E. G. Feng and J. Smith. 2008. Syllabification, sonority, and perception: new evidence from a language game. Proceedings from the main session of the 41st meeting of the Chicago Linguistic Society, 341-355.

Narayan, C. 2008. The acoustic-perceptual salience of nasal place contrasts. *Journal of Phonetics* 36, 191-217.

National Institute of the Korean Language. 2003. *Standard Korean Dictionary.*

National Institute of the Korean Language. 2008. *Oelaye phyoki yonglyeycip* (Examples of foreign word orthographic notations).

Ní Chiosáin, M. and J. Padgett. 2001. Markedness, segment realization, and locality in spreading. In L. Lombardi (ed.), *Segmental Phonology in Optimality Theory*. Cambridge: Cambridge University Press.

Ohala, J. 1992. What's cognitive and what's not in sound change. Paper A 297. Linguistic Agency of the University of Duisburg (LAUD), Universität Duisburg Gesamtochschule.

Ohala, J. 1993. The phonetics of sound change. In C. Jones (ed.), *Historical Linguistics: Problems and Perspectives.* New York: Longman.

Ohala, J. and J. Lorentz. 1977. The story of [w]: An exercise in the phonetic explanation for sound patterns. Proceedings of Annual Meeting of Berkeley Linguistics Society 3, 577-599.

Ohde, R. and D. Sharf. 1977. Order effect of acoustic segments of VC and CV syllables on stop and vowel identification. *Journal of Speech and Hearing Research* 20, 543-554.

Padgett, J. 2001. Contrast dispersion and Russian palatalization. In E. Hume and K. Johnson (eds.), 181-218.

Padgett, J. 2003. Contrast and postvelar fronting in Russian. *Natural Language and Linguistic Theory* 21, 39-87.

Penny, R. 1972. *The History of the Spanish Language.* Cambridge: Cambridge University Press.

Perkell, J. and D. Klatt (eds.). 1986. *Invariance and Variability in Speech Processes.* Hillsdale: Lawrence Erlbaum Associates.

Pisoni, D. and J. Tash. 1974. Reaction times to comparisons with and across phonetic categories. *Perception and Psychophysics* 15, 285-290.

Prince, A. and P. Smolensky. 1993. Optimality Theory: Constraint interaction in Generative Grammar. Ms., Rutgers University and University of Colorado at Boulder.

Raškinis, A. and S. Dereškeviciute. 2007. An analysis of spectral attributes, characterizing the interaction of Lithuanian voiceless velar stop consonants with their pre- and postvocalic context. *Information Technology and Control* 36, 68-75.

Rehg, K. and D. Sohl. 1981. *Ponapean Reference Grammar.* Honolulu: The University Press of Hawaii.

Repp, B. and H.-B. Lin. 1989. Acoustic properties and perception of stop consonant release transients. *Journal of Acoustical Society of America* 85, 379-396.

Sagey, E. 1986. The representation of features and relations in non-linear phonology. Doctoral dissertation, MIT.

Schwab, E., J. Sawusch and H. Nusbaum. 1981. The role of second formant transitions in the stop-semivowel distinction. *Perception and Psychophysics* 29, 121-128.

Selkirk. E. 1982. The syllable. In H. van der Hulst and H. Smith (eds.), *the Structure of Phonological Representations.* Dordrecht: Foris Publications.

Selkirk, E. 1991. Dependency, place and the notion “tier”. Ms, Department of Linguistics, University of Massachusetts, Amherst.

Shevelov. G. 1993. Ukrainian. In B. Comrie and G. Corbett (eds.), 999-1036.

Short, D. 1993. Czech. In B. Comrie and G. Corbett, eds., 455-532.

Silva, D. 1991. Phonological variation in Korean: the case of the “disappearing w”. *Language Variation and Change* 3, 153-170.

Smolensky, P. 1995. On the internal structure of constraint component in UG. ROA-86.

Sohn, H. 1987. On the representation of vowels and diphthongs and their merger in Korean. *CLS* 23, 307-323.

Sohn, H. and B. Bendor 1973. *A Ulithian Grammar*. Pacific Linguistics Series C-27. Department of Linguistics, The Australian National University.

Steriade, D. 2001. The phonology of perceptibility effects: the P-map and its consequences for constraint organization, Ms. MIT.

Stevens, K. N. 1980. Acoustic correlates of some phonetic categories. *Journal of Acoustical Society of America* 68, 836-842.

Stevens, K. N., S. J. Keyser and H. Kawasaki. 1986. Toward a phonetic and phonological theory of redundant features. In J. Perkell and D. Klatt, 426-449.

Stevens, K. N., S. Manuel and M. Matthies. 1999. Revisiting place of articulation measures for stop consonants: Implications for models of consonant production. Proceedings of International Congress on Phonetic Science 14, 1117-1120. San Francisco.

Sussman, H., H. McCaffrey and S. Matthews. An investigation of locus equations as a source of relational invariance for stop place categorization. *Journal of Acoustical Society of America* 90, 1309-1325.

Takatori, Y. 1997. A study of constraint interaction in Slavic phonology. Doctoral dissertation, Yale University.

Voorhis, P. 1982. The Kickapoo sound system and Kickapoo standard Orthography. W. Cowan (ed.), Papers of the Thirteenth Algonquian Conference, 81-89. Ottawa: Carleton University.

Wang, J. Z. 1993. The geometry of segmental features in Beijing Mandarin. Doctoral dissertation, University of Delaware.

Werker, J. and J. Logan. 1985. Cross-language evidence for three factors in speech perception. *Perception and Psychophysics* 37, 35-44.

Wilson, C. 2006. Learning phonology with substantive bias: an experimental and computational study of velar palatalization. *Cognitive Science* 30 (5), 945-982.

Yang, B. 1996. A comparative study of American English and Korean vowels produced by male and female speakers. *Journal of Phonetics* 24, 245-261.

Zue, V. W. 1976. Acoustic characteristics of stop consonants: a controlled study. Doctoral dissertation, MIT.