Mandarin Speakers' Production of English and Mandarin Post-Vocalic Nasals: An Acoustic Approach

by

Ya Li

B.A., University of Western Ontario, 2006

A Thesis Submitted in Partial

Fulfillment of the Requirements for the

Degree of

MASTER of ARTS

in the Department of Linguistics

University of Victoria

Mandarin Speakers' Production of English and Mandarin Post-Vocalic Nasals: An Acoustic Approach

by

Ya Li

B.A., University of Western Ontario, 2006

Supervisory Committee

Dr. Hua Lin, Supervisor

Department of Linguistics

Dr. John Esling, Departmental member

Department of Linguistics

Supervisory Committee

Dr. Hua Lin, Supervisor, (Department of Linguistics)

Dr. John Esling, Departmental member, (Department of Linguistics)

ABSTRACT

The present study adopts an acoustic approach to analyze Mandarin Chinese speakers' production of English and Mandarin alveolar and velar nasal codas /n, ŋ/ in different preceding vowel contexts. Its purposes are to explore the interrelationship between nasal codas and the preceding vowels in both L1 (First Language) and L2 (Second language) production and to identify and explain similarities and differences between the L1 and L2 production.

Specifically, 20 native Mandarin Chinese speakers performed a word-list reading task involving 22 English and Mandarin test words with three types of rimes, VN (Vn or Vŋ, i.e., a monophthong vowel followed by /n/ or /ŋ/), VGn (a diphthong vowel followed by /n/), and VG (a diphthong vowel). In total, 88 tokens (22 words x 4 repetitions) were collected for each speaker, and all tokens were measured by using the phonetic software, Praat. First, mean F1-F0 and F3-F2 (differences between the first and fundamental formant frequencies and between the third and second formant frequencies) over the first and the second half of vowel duration were measured to estimate

vowel height/backness changes over the duration. Also, N1/N2/N3 (the first, second, and third nasal formants) at the midpoint of nasal murmur duration and the band energy difference (ΔdB) between 0-525 Hz and 525-1265Hz bands over the nasal murmur duration were calculated to predict the alveolar or velar nasal place. Last, the vowel and nasal murmur duration (V_D & N_D) in each token were used to indicate the degree of vowel-nasal coupling.

Two-tailed paired-wise t-tests and repeated measures one-way ANOVA tests were used to examine the statistical significance of the above acoustic measurements across test words. The main results show that there is a strong vowel-nasal coarticulation effect in Mandarin VN and English VGn production but not in English VN production; specifically, nasal place in Mandarin VN and English VGn rimes covaries with vowel quality change over the duration. In contrast, there is a significant durational difference among English VN rimes but not among Mandarin VN and English VGn rimes; specifically, Vŋ rimes are longer than Vn rimes in English. The strong vowel-nasal coarticulation effect in the Mandarin VN and English VGn production and the significant durational difference in the English VN production can be both related to rhythmic factors.

TABLE OF CONTENTS

LIST OF TABLES

page

LIST OF FIGURES

ACKNOWLEDGMENTS

First, I wish to extend my deepest gratitude to my supervisor, Dr. Hua Lin. As a major contributor to Chinese linguistics research at the University of Victoria, Dr. Lin used her invaluable expertise to guide me through every step of my graduate study. It was she who helped me to choose the thesis topic based on her keen observation that Chinese speakers have difficulty with English nasal coda production. Her insight also helped me to appreciate that second language phonological studies could be greatly enhanced by an acoustic approach. The supervising style of Dr. Lin is all about caring. Caring means that she has generously shared her books, articles, and corpus data with me; that she has scheduled weekly meetings to review my work, taking notice of even my smallest progress; that she has taken time from her busy agenda to revise my every conference abstract and research paper and write me reference letters in her eloquent writing style; and the list goes on. All of these have not been easy for Dr. Lin, a professor, the graduate advisor in the department, and a woman who has to take care of a 12-year-old daughter and an 80-year-old mother. In a word, Dr. Lin is not only my supervisor but also my role model. Needless to say, this thesis would have been impossible without her close supervision and strong support.

I would also like to express my sincere appreciation to my other supervisory committee member, Dr. John Esling, for his advice along various stages of my thesis completion, from the research design, thesis proposal, to the final draft. Dr. Esling and his unique sense of humour, for example, introduced me to the phonetic wonderland in his laughter-filled phonetics class. I still remember my first exciting yet scary experience in watching, live through a special video camera (fibre-optic laryngoscope), how a human throat (Dr. Esling's throat, of course) makes sounds. I also wish to thank Dr. Li-Shih Huang, once my supervisory committee member and now on maternity leave, for all her support and encouragement. Dr. Huang, for example, always offered me timely help whenever I needed, from reviewing my conference abstracts, course papers, and scholarship applications to writing me reference letters. I have also benefited from chatting with her on various occasions and learned useful tips for an academic such as how to write a research proposal.

My gratitude also goes to Dr. Hossein Nassaji, who taught me how to develop a thesis proposal and provided me with valuable feedback concerning statistics used in my study, and Dr. Tae-jin Yoon, who allowed me to sit in his phonetic class and helped me to clarify some technical questions related to Praat programming. I also wish to thank my fellow graduate students, such as Thomas Magnuson, Qian Wang, Shu-min Huang, and Scott Moisik, to name a few, for their various help and encouragement along my way from the start of my MA courses to the completion of my thesis.

I also thank the Social Sciences and Humanities Research Council of Canada (SSHRC) for granting me a one-year master's scholarship (Award No. 766-2006-4194) so that I was able to afford a full-time study at the graduate level.

To Jeff

TRANSCRITION NOTES

English words: *Italic* alphabets.

Mandarin Chinese words: The *Italic* Chinese *Pinyin* alphabets with tonal markers.

The phonemic or underlying transcription of a sound: / IPA symbol(s) / (IPA: International Phonetic Association)

The phonetic or surface transcription of a sound: [IPA symbol(s)].

Chapter 1

INTRODUCTION

1.1 Purpose of the study

The present study adopts an acoustic approach to examine Mandarin speakers' production of Mandarin and English post-vocalic nasals /n, ŋ/ in different vowel contexts. Its purposes are to explore the interrelationship between nasal codas and the preceding vowels in both L1 and L2 production and to identify and explain similarities and differences between the L1 and L2 production.

The acoustic approach involves comparisons among acoustic parameters measured in English and Mandarin test words produced by Mandarin speakers. The acoustic parameters include the first three vowel and nasal formant frequencies (F1/F2/F3 and N1/N2/N3), the mean band energy difference (ΔdB) between the low (0-525Hz) and mid (525-1265Hz) frequency bands over the nasal murmur duration, the vowel and nasal murmur duration (V_D & N_D) in a word.

The vowel contexts vary in vowel type (monophthong or diphthong) and vowel quality (height and/or backness).

L1 refers to Standard Mandarin Chinese and L2 refers to Canadian English.

1.2 Motivation for the study

According to Maddieson (1999), nasals are among the most common sounds in languages around the world. Based on Greenberg's (2005) typological markedness theory, they are also considered as typologically unmarked sounds. For example, English shares the same set of nasals, /m, n, ŋ/ even with the totally unrelated Chinese language. As the most common and unmarked sounds, these three nasals are presumably among the easiest to produce and acquire. However, Mandarin speakers seem to have difficulty articulating English /n, ŋ/ codas. For example, Mandarin speakers are found to produce them interchangeably in English words such as *sin/sing* and *son/song* (Hansen, 2001). Also, they are often heard to produce English *down* as a similar Mandarin syllable, *dàng*.

Some known L2 theories, from the early L1 transfer theory (thoroughly discussed in James, 1988), Eckman's (1977) Markedness Differential Hypothesis (MDH), to Flege's (1995) Speech Learning Model (SLM), identify that various linguistic and cognitive factors, such as L1 transfer, markedness, perceptual links between L1 and L2 sound categories can all play a role in shaping L2 coda production. In Mandarin speakers' production of *down* as *dàng* (/dɑŋ/), for example, the nasal place change from /n/ to /ŋ/ may be influenced by the Mandarin phonotactic constraint that rimes must agree in backness (hereafter it will be referred as "the same backness constraint", in contrast with "the backness assimilation rule" in which the same phenomenon is described in terms of feature spreading). Also, the deletion of the medial glide /w/ (i.e., G in VGn) may be influenced by the Mandarin syllable structure constraint that nasal codas are not allowed to follow a diphthong vowel.

However, the above impressionistic explanations suggest more questions. First, do the Mandarin phonological constraints work in a parallel case? For example, if Mandarin speakers are also found to produce *cone* as the Mandarin syllable, *kòng* (to parallel *down/dàng*), then we can be more ready to claim that the above Mandarin constraints are transferred to the English production. Second, do similar phonetic properties in two sounds necessarily lead to the substitution of one sound for another? For example, *down* may well be produced as *dan* instead of *dàng* because the rime *an* in *dan* agrees in backness; better yet, it retains the nasal place. However, *dan* does not seem to be a more preferred (or frequent) candidate by Mandarin speakers than *dàng* to replace *down*.

These questions lead to at least two broader issues to be considered in this study: how to find out the systematic patterning in L2 production and how to ensure the patterning will be adequately but not over-powerfully explained by a linguistic factor such as L1 transfer. To address the first issue, it is important to identify speech patterns in L2 production as clearly as possible. Thus, this study decides to adopt an acoustic approach, since it generally allows us to capture subtle phonetic cues that may not be easily perceived by an auditory judgment (Hagiwara, 2006). To address the second issue, it is necessary to determine what role and/or how much of a role a linguistic factor plays in the L2 production. Therefore, the first thing this study needs to do is to compare between L1 and L2 nasal production and between nasal productions in different contexts so that linguistic factors may be able to manifest themselves following the results of the comparisons.

Last, previous research on nasal production by Mandarin speakers seems scanty. Within the limited amount of research in the area, for example, in Chen's (1997) study of Mandarin VN production and Hansen's (2001) study of L2 coda production by Mandarin speakers, the nasal production is investigated from either a pure phonetic (e.g., Chen, 1997) or a pure phonological (e.g., Hansen 2001 & 2006) point of view, but not both. This study will be the first acoustically based phonological study, to my knowledge, of Mandarin speakers' L2 post-vocalic nasal production, so it hopes to bridge the phonetic and L2 phonological studies of nasal production by describing acoustic patterns of the L1 and L2 production and explaining the patterns from a phonological point of view.

1.3 Research questions and hypotheses

Research questions:

1) How do vowel context and nasal place interact respectively in L1 and L2 production?

2) Can systematic similarities and differences be identified between the L1 and L2 production? If yes, what linguistic factors may come into play?

3) Is there any acoustic evidence that can be used to qualify/quantify previous phonological claims such as the backness assimilation rule?

Research hypotheses:

If it is true that Mandarin speakers tend to alternate the nasal place in English word pairs, *sin/sing* and *son/song*, and to drop the glide in an English VGn rime, then this study hypothesizes that:

1) The actual nasal place in Mandarin speakers' production of English and Mandarin velar /ŋ/ is different.

The basis for this claim is that Mandarin speakers should be able to distinguish the two nasal codas /n/ and /ŋ/ in their Mandarin production, but that their ability to produce the two codas distinctively in Mandarin does not carry over to their English production.

2) English post-vocalic nasal production by Mandarin speakers is related to supra-segmental factors.

The basis for this claim is that if nasal codas /n/ and /ŋ/ by themselves are among the easiest segments to produce, then Mandarin speakers' ability to produce the two codas distinctively in English may be instead hampered by high-level constraints (such as syllabic and prosodic constraints in L1 or L2).

3) Nasal place co-varies with vowel backness in Mandarin speakers' production of both Mandarin and English post-vocalic nasals.

The basis for this claim is that Mandarin VN production is subject to the same backness constraint, so their English VN and VGn production may be subject to a similar constraint.

1.4 Scope of the study

This study focuses on finding and explaining systematic acoustic similarities and differences between L1 and L2 production. Given this focus, this study does not intend to be a technical guide to acoustics, but the general methodology involving acoustic measurement is provided.

1.5 Structure of the thesis

This thesis paper is divided into 5 chapters. Chapter 1 provides the rationale of this study and specify research purposes, questions, and hypotheses. Chapter 2 provides a survey of previous studies on L2 nasal production, followed by an introduction of acoustic properties of vowels and nasals. Chapter 3 describes the methodology adopted by this study. Chapter 4 presents the results from both acoustic comparisons and statistical analyses and generally discusses the related L1 and L2 phonological factors. Chapter 5 further discusses the key results and their theoretical implications for L2 speech production. Chapter 6 concludes this study by answering the research questions, evaluating the research hypotheses, identifying the limitations, and postulating future directions.

Chapter 2

LITERATURE REVIEW

This chapter will start with an overview of English and Mandarin vowels and nasal codas (Section 2.1). It will then review some influential studies on nasal production from three perspectives, L1 and L2 nasal coda production and perception (Section 2.2), vowel-nasal coarticulation (Section 2.3), and acoustic properties of vowels and nasals (Section 2.4). Last, Section 2.5 will review some important L2 acquisition theories relevant to this study.

2.1 English and Mandarin vowels and nasal codas

As mentioned in Chapter 1, English and Mandarin share the same set of nasals /m, n, ŋ/ phonemically, but unlike English in which the three nasals /m, n, ŋ/ can occur in basically all types of syllable positions (even a nucleus position in an unstressed syllable such as /bʌtn̩/, *button*), Mandarin /m/ cannot occur syllable-finally, /ŋ/ cannot occur syllable-initially, and neither /n/ nor /ŋ/ can occur after a diphthong vowel. Table 2-1 illustrates the difference between English and Mandarin in terms of the rime structure involving a single post-vocalic nasal. Note that hereafter the onset will be ignored in this study.

Table 2-1 English and Mandarin rime structures involving a single post-vocalic nasal

English	Mandarin
VN# (N = /m, n, ŋ/)	VN# (N = /n, ŋ/)
e.g., *im*, *in* and *ing*	e.g., *in* and *ing*
VGn# * (G = /j, w/)	
e.g., *own*	

* VGŋ# is ignored due to rare occurrence in English.

Interestingly, while Mandarin allows glides and nasals as the only two types of codas, they never co-occur in a Mandarin syllable; that is, the glide-nasal coda cluster is prohibited (Lin, 2001a).

In addition, Table 2-2 compares Canadian English and Standard Mandarin vowel inventories. Note that the <u>underlined</u> sounds are NOT shared by both languages; specifically, English has a tense/lax contrast in high and mid vowels, but Mandarin has only a rounded/unrounded contrast in these vowels. Also, the sounds in parentheses are allophones rather than phonemes.

*Table 2-2 A comparison of Canadian English and Standard Mandarin monophthong vowels**

		front		central	back	
		unrounded	rounded		unrounded	rounded
English	high	i ɪ (lax)				u ʊ (lax)
	mid	(e) ɛ (lax)		ə ʌ		(o) ɔ (lax)
	low	æ			ɑ	
Mandarin	high	i	y			u
	mid	(e, ɛ)		(ə)	ɤ	(o)
	low			a	(ɑ)	

*Data source: Canadian English is from O'Grady and Archibald (2000); Standard Mandarin is from Lin (2001a).

According to Lin (2001a), Mandarin has only a low vowel /a/, but depending on the context, it has three allophones, [a], [ɑ], and [ɛ]; particularly, /a/ becomes [ɑ] when followed by /u/ or /ŋ/ (i.e., /a/+/u/ -> [ɑ w] or /a/ + /ŋ/ -> [ɑŋ]). In other words, Mandarin /a/ assimilates /u/ and /ŋ/ in backness.

As for mid vowels, Lin (2007) mentioned that Mandarin has only a mid vowel /ə/ (i.e., /ɤ/ in Figure 2-1). Depending on the context, /ɤ/ has four allophones, [ə], [e], [o], and [ɤ]; particularly, /ɤ/ becomes [ə] when followed by nasals. Lin (2007) further noted that [ə] followed by [n] is close to but not quite the same as [en] and [ə] followed by [ŋ] is close to but not quite the same as [ɤŋ]). In other words, [ə] is also assimilated to the backness of the nasal place.

Table 2-3 compares Canadian English and Mandarin diphthong vowels.

*Table 2-3 A comparison of Canadian English and Standard Mandarin diphthong vowels**

English	/aj/	/aw/	/ow/**	/ej/**	/oj/
Mandarin	/aj/	/aw/	/ow/	/ej/	

*Data source: Canadian English is from O'Grady and Archibald (2000); Standard Mandarin is from Lin (2007).
**English /ow/ and /ej/ are often treated phonemically the same as /o/ and /e/.

Note that Mandarin does not have /oj/ (as in *boy)*; also, Mandarin /aw/ becomes [ɑw] due to backness assimilation. According to Lin (2007), that Mandarin has [ɑŋ] but not [aŋ], [aj] but not [aw], and [ow] but not [oj] is based on the language-specific constraint on permissible syllable types; that is, Mandarin segments in rimes must have the same value for backness and/or roundness (i.e., "the same backness constraint").

Note also that Mandarin consonants preceding the high vowel /i/ undergo a phonological process, called "palatalization" (Lin, 2001a), so the Mandarin /s/ + /i/ + /n/ sequence, for example, becomes [ɕin] (*xin*) in which [ç] is palatal fricative. As for Mandarin /s/ + /i/ + /ŋ/, /s/ also becomes palatal [ç], but opinions differ as to how to represent /i/ phonemically and phonetically. For example, Lin (2007) and Cao (2000) treated *xìng* as having the phonemic form, /çiŋ/ (*xing*), but the phonetic form, [çjəŋ], where /i/ becomes a glide and schwa [ə] is inserted as the syllable nucleus, since /i/ and /ŋ/ can not be next to each other due to their difference in backness. However, Xu (1999) specifically mentioned that /çiŋ/ has an identical phonetic form [çiŋ]. Since /çiŋ/ will be included in this study as a test word, the production data will reveal its real phonetic status.

Finally, Lee and Zee (2003) considered that Mandarin has an extra vowel, the rhotacized /ɚ/ as in the pronunciation of the Mandarin word, *èr (“two”)*; however, according to Lin (2007), it is arguable whether it can be treated as /ɚ/, as the syllabic consonant /ɹ̩/ as in the English word *butter* (/bʌtɹ̩/), or as the diphthong vowel /əɹ/. This study leaves this vowel out due to its questionable status.

2.2 L1 and L2 nasal production and perception

Recasens (1988) provided a comprehensive account of nasal production and perception from a phonetic perspective. First, he attributed /ŋ/ being a less frequent and more marked nasal due to "the complexity of the articulatory manoeuvres required to produce a salient dorso-velar closure while holding the velum lowered" (Recasens, 1988, p. 230). Then, Recasens discussed nasal production in onset and coda positions. Generally, nasals are less phonetically salient in coda than in onset position because of the less abrupt oral release. In other words, nasal place in coda position is not easily detected by transitional cues (amplitude and spectral changes) between the preceding vowel and the nasal coda. Instead, nasal murmur duration is often used as a good perceptual cue to nasal place because nasal murmurs are longer syllable-finally than syllable-initially (Recasens, 1988). Further discussions of nasal place identification in terms of acoustic properties will be provided in Section 2.4.2.

Also, Recasens (1988) used both perception and production data to show that /n/ is more subject to consonantal assimilation (or less "co-articulation resistance", according to Harrington & Cassidy, 1999, p.111) than /m, ŋ/. For example, in the preceding high vowel /i/ context, /m, ŋ/ both are easily confused with /n/, but /m, ŋ/ are not subject to phonetic replacement between each other. Similarly, Hajek (1997) pointed out that coronal gestures (the alveolar /n/ is a coronal sound) are easily overlapped by non-coronal gestures and coronality is perceptually masked. In other words, /n/ is less marked than /m, ŋ/.

Further evidence for the unmarkedness of /n/ is from Zee's (1985) diachronic study of nasal coda changes in Chinese dialects including Mandarin. Basically, Zee's (1985) study identifies three major processes involving the diachronic development of Chinese syllable-final nasals; that is, /Vm/ -> /Vn/, /Vŋ/ -> /Vn/, and /Vn/ -> Ṽ (the diacritic ˜ above a symbol refers to a nasalized segment). The tendency of /m/ and /ŋ/ both becoming /n/ indicates that /n/ is easier to produce than /m, ŋ/.

As for L2 nasal coda production, Hansen's (2001) study reveals that Mandarin speakers tend to produce English /ŋ/ as /n/ in words such as *sing* and *song*, which renders *sing/song* sound similar to *sin/son*. Hansen (2001) reasoned that the nasal place fronting of /ŋ/ to /n/ might be due to the influence of the Beijing dialect of Mandarin spoken by her subjects. Since the Beijing dialect has a more prominent pronunciation of /ŋ/ than English, Beijing Mandarin speakers may under-compensate its production in English to signal that English /ŋ/ is less prominent than Beijing /ŋ/.

In addition, Hansen's (2006) study provides the hierarchy in terms of the target-like production of nasal codas, /n/ > /m/ > /ŋ/, indicating again that /n/ is easier to produce than the other two nasals. Also, her study finds that nasals are usually produced target-like by her Vietnamese participants, but /m, n/ sometimes are absent after a diphthong vowel such as /aj/. Note that Vietnamese is similar to Mandarin in that it does not allow consonant clusters.

Vietnamese speakers' nasal deletion may be explained by the sonority hierarchy (from the most sonorous to the least): vowels > glides > laterals > nasals > fricatives > stops (Hansen, 2006); that is, the less sonorancy of nasals than glides results in the nasal deletion. However, Mandarin speakers' glide deletion in the production of *down* as /dɑŋ/ can not be readily explained based on the sonority hierarchy.

Note that Mandarin speakers' production of *down* as /dɑŋ/ is not random but can be identified with a common phonological process, "glide-hardening", found in Northern Italian; that is, a nasalized offglide hardens (or consonantizes) to a velar nasal (Hajek, 1997). According to Hajek (1997), many Northern Italian dialects have undergone the following sequence of diachronic phonological change, [VN] > [Ṽ] > [ṼG̃] > [Ṽŋ]. For example, the pronunciation of the word, *spina* ('thorn'), has changed from [spẽj̃na] to [spẽɰ̃na] to [speŋna], and the word, *luna* ([lõŋna], 'moon'), used to be [lõw̃na] in these dialects.

Hansen also investigated the influence of task type on consonant coda production and stated that "there was a greater accuracy on the reading data (word-list and reading passage) compared with the interview data" (2006, p.118), which is consistent with previous findings that production is more accurate in formal styles such as word-list and text reading than in casual styles (e.g., Sato, 1985; Major, 1994). Since in fast and casual reading, the nasal murmur (i.e., a humming sound with clearly defined nasal formants following the vowel) tends to disappear and therefore is hard to measure acoustically, this study will elicit speech data through word-list reading in order to obtain accurate acoustic measurements of nasal murmurs.

As for nasal coda perception, Zee (1981) conducted a perceptual study of the effect of vowel quality on post-vocalic nasals and found that /ŋ/ tends to be identified as /n/ after the high vowel /i/ but can be correctly identified after the low vowel /a/ even in the noise condition. The explanation Zee (1981) provided is that /ŋ/ becomes the palatalized /ȵ/ when coarticulated with /i/. Since the palatal /ȵ/ is not a given choice for the subjects, /n/ is chosen as a close substitute for /ȵ/. /a/ as a low vowel, on the other hand, has no constriction above the pharyngeal area (note that the IPA uses "openness" to characterize low vowels), so the coarticulation effect with its following nasal is minimum. Consequently, nasal place perception will not be severely interfered by the preceding /a/. However, Zee's (1981) explanation is in conflict with Recasens' finding that "the degree of coarticulatory sensitivity for vowels decreases for /ə/> /a/> /i/" (1997, p. 546). Other studies (e.g., Chen, 2000; Clumeck, 1976) also found that low vowels are more subject to a strong vowel-nasal coarticulation effect than high vowels because they have a longer period of nasalization when followed by a nasal. Thus, Zee's (1981) finding that nasals are easy to identify following /a/ can not be explained in terms of the little coarticulation effect between a low (or open) vowel and a nasal. Nonetheless, the openness of the vowel /a/ still seems responsible for the easy identification of nasal place, perhaps because no stricture between oral articulators in the production of /a/ allows maximum flexibility to form a stricture appropriate for a following nasal and hence the accurate production of the nasal place.

Last, Aoyama's (2003) study of English nasal perception by Korean and Japanese speakers reveals that syllable-final /n/ - /ŋ/ contrast is poorly identified by Japanese speakers, but their identification of the /m/-/n/ and /m/-/ŋ/ contrasts is very good. Aoyama (2003) explained her findings based on Best's (1995) Perceptual Assimilation Model (PAM); that is, Japanese speakers classify /n/ - /ŋ/ as uncategorizable since their L1 does not distinguish the two (Japanese has only an arguable uvular /ɴ/ according to Yamane-Tanaka, 2008) and as a result, the uncategorizable /n, ŋ/ are subject to random categorization in the perception tests. Basically, Aoyama's study suggests that "the perceived relationship between L1 and L2 segments plays an important role in how L2 segments are perceived" (2003, p. 263).

2.3 Vowel-nasal coarticulation

Coarticulation broadly refers to the phenomenon that a phonological segment is realized differently in different environments (Kühnert & Nolan, 1999). According to Chafcouloff & Marchal (1999), vowel-nasal coarticulation has been studied extensively from the physiological rather than acoustic point of view, and the lack of acoustic studies of the coarticulation effect can be attributed to the technical difficulty of measuring the contextual influence of nasals on vowels acoustically and the conflicting data obtained by different studies (possibly due to differences in research methodology).

Despite the relative lack of acoustic evidence on nasal coarticulatory effects, previous physiological studies seem to agree that "there is strong interaction between oral and nasal sounds" (Chafcouloff & Marchal, 1999, p. 70). For example, Moll and Daniloff (1971) found that in a CVVn (i.e., CVGn, C: a consonant) syllable, the velum could be lowered as soon as the first consonant was produced, so the anticipatory effect of vowel-nasal co-articulation not only concerns the immediate context of a nasal but also extends over several vowels preceding the nasal. Sometimes, a nasal coda can even suppress a preceding vowel and become syllabic as in /bʌtn̩/(*button*). No wonder that Bladon and Nolan (1977) ranked nasals among sounds with the least coarticulatory resistance.

Furthermore, strong vowel-nasal coarticulation effects can be clearly shown from the perspective of vowel perception and production. In their discussion of vowel production, for example, Rosner and Pickering pointed out that "coarticulation can affect vowel height but has a larger impact on place of articulation along the front-back dimension" (1994, p. 272). Also, Chen's (2000) acoustic study of Mandarin VN production finds that when followed by /ŋ/, the three vowels /i, a, ə/ tend to move backward. Note that Chen's (2000) finding reflects the same backness constraint in Mandarin.

As for vowel perception, Beddor (1991) found that nasalization generally raises the perceived height of low vowels and lowers the perceived height of high vowels, but has a more pronounced lowering effect on front than on back vowels.

Furthermore, Kingston (1991) studied the impact of nasalization on perceived vowel height and found that perceptual height was realized not only by conventional tongue height, but by other articulations such as velum height (related to nasalization) as well, so he claimed that nasalization is an integrated part of perceived vowel height. Similarly, Esling also argued that "the tongue is not the only articulator that determines vowel quality" from a production point of view (2005, p.16). He proposed the alternative oral-laryngeal model instead of the traditional lingual oral model of distinguishing vowels in terms of the high-low and front-back dimensions of tongue movement (lip-rounding aside). Esling's (2005) new model integrates lingual-laryngeal articulations and adopts the front-open-raised-retracted dimension instead of the high-low-front-back dimension to reflect the contribution of larynx to vowel quality. Based on this new model, front, open, raised, and retracted vowels (formerly high-front, low-front, high-back, low-back vowels) are associated with the actions of different parts of articulators. The traditionally defined low, back vowel /ɑ/, for example, is primarily linked to laryngeal activities and the low lingual component of /ɑ/ is secondarily related to the laryngeal activities (Esling, 2005). Therefore, the traditional notion of vowel height/backness may not be adequate to explain complicated articulatory movements involving nasalized vowels due to the addition of velum lowering to the lingual-laryngeal gestures.

Rosner and Pickering (1994) also discussed the following two hypotheses used to capture the articulatory and auditory characteristics of co-articulated vowels: one is the reduction hypothesis that "vowels in context reduce" (p.73), and the other is the assimilation hypothesis that "co-articulation causes contextual assimilation" (p.271).

The vowel reduction hypothesis has its phonetic basis because vowels tend to be centralized under the nasal coda influence (to be further discussed in the next section). However, in Jha's (1986) acoustic study of nasal vowels in

Maithili, the front nasal vowels /ĩ, ẽ, æ̃/ are more fronted, and the central vowels /ã, ə̃/ and the back vowels /ɔ̃, õ, ũ/ are more backed than their oral counterparts. Admittedly, nasal vowels with distinctive (or phonological) nasalization may behave differently from nasalized vowels formed through vowel-nasal coarticulation. The assimilation hypothesis, on the other hand, can find its support from Chen's (2000) finding that Mandarin vowels move to the back when followed by velar /ŋ/.

To sum up, vowel-nasal coarticulation has a general effect on vowel quality change along the high-low and/or front-back dimensions.

2.4 The acoustic properties of vowels and nasals

This section will review previous studies on acoustic properties of vowels (Subsection 2.4.1) and nasals (Subsection 2.4.2). The last subsection, 2.4.3, will review previous studies of vowel and nasal duration.

2.4.1 Vowels

As mentioned in Section 2.3, vowels are traditionally described in terms of height, backness, and roundness. In the context of American vowels, vowel quality correlates with vowel formant frequencies in the following ways: the higher a vowel, the lower the F1; the more backed a vowel, the less the F2-F1 (the difference between the second and first formants) (Davenport and Hannahs, 1998). The high, front vowel /i/, for example, has the lowest F1 (around 300Hz) and the highest F2-F1 value (about 2000Hz).

Sussman's (1990) study of the front/back vowel distinction further demonstrates that F3-F2 is a better indicator of vowel backness than F2-F1, because F3 and F2 for front vowels such as /i/ are very close but far apart for back vowels (Harrington & Cassidy, 1999). Syrdal and Gopal's (1986) perceptual study of American vowels, on the other hand, shows that F1-F0 is a better indicator of vowel height than F1 alone, because F1 is inversely correlated with F0. The high vowel /i/, for example, has a small F1-F0 due to its low F1 and high F0.

Generally, the higher a vowel, the lower the F1-F0; the further back a vowel, the greater the F3-F2. In fact, the use of F3-F2 and F1-F0 instead of F2-F1 and F1 to represent vowels in the auditory space is called "the spectral integration hypothesis", and according to Hayward (2000), vowels can be more clearly separated and thus better perceptually distinguished by the integrated F3-F2 and F1-F0 parameters. Thus, this study will use F1-F0 instead of F1 to correlate vowel height and F3-F2 instead of F2-F1 to correlate vowel backness. Another advantage of using the integrated acoustic correlates F1-F0/F3-F2 to infer vowel height/backness is that they are relative measurements, which should not be overly sensitive to speaker variations.

As for how to measure vowel formants (F1/F2/F3) accurately, van Son and Pols (1990) evaluated different methods of measuring Dutch vowels read at normal and fast rate. Their results show that whether measured at the midpoint of a vowel or by averaging the formants over the vowel duration, the vowel formant values obtained from the two methods are not significantly different. Therefore, van Son and Pols concluded that "when studying vowel target, the method that is most convenient can be used" (1990, p.1692).

Chen (2000), for example, measured F1, F2, and F3 throughout a vowel in order to determine the place of articulation of the nasal coda. Specifically, two types of measurements are taken: the time-averaged (every 10ms) F1, F2, and F3 values over the vowel duration and at the end point of the vowel. According to Chen (2000), the two types of measurements of vowel formant frequencies are able to complement with each other to detect nasal place, and the vowel formant measurement over the duration is especially useful in nasal place detection when a nasal murmur is not present but realized through vowel nasalization.

Note that for glides such as the off-glides /j, w/ in diphthong vowels, Ladefoged & Maddieson suggested that they be called vowel-like consonants or semivowels, because "they are produced with narrower constrictions of the vocal tract" than their corresponding vowels (1996, p. 323). In the spectrograms of semivowels, slow formant transitions can be observed between semivowels and their surrounding sounds (Hayward, 2000). As for acoustic correlates of glides, Espy-Wilson (1994) chose F1-F0 as an acoustic parameter to characterize the glide height, which further validates the use of F1-F0 in this study to correlate the height of both monophthong vowels (V) and diphthong vowels (VG).

Ladefoged & Maddieson (1996) also mentioned that vowels with an extra nasality feature (i.e., nasalized vowels) can be distinguished by a reduced intensity in F1 and a higher F3. The reduced intensity is due to the diversion of acoustic energy from the oral cavity into the nasal passage. From the power spectrum of a nasal sound, negative peaks, also called nasal zeros or antiformants, can be observed as a result of the diversion (Hayward, 2000). Moreover, Maeda (1982) found that the diversion flattens the spectral region between 300 and 2500 Hz.

Similarly, Fant found that nasalized vowels have “a distortion superimposed on the vowel spectrum”; specifically, N1 (the first nasal formant) occurs in the region usually below F1 (the first vowel formant) which consequently weakens and shifts up F1 (2004, p. 156). Beddor (1991) also found that nasal vowels have a broader and flatter spectral prominence in the low-frequency region and vowel height is determined both by the most prominent harmonics (i.e., F1, N1, and/or F2) in the low-frequency region and by the spectral slopes in the vicinity of these harmonics.

An acoustic approach to vowel nasalization detection was developed by Chen (1997) in her study of English and French nasalized vowels. She considered the reduction of the amplitude of F1 as the primary cue of vowel nasalization and successfully distinguished English/French nasalized vowels by employing the following two parameters: A1-P1 and A1-P0, where A1 is the vowel F1 amplitude, P0 is the amplitude of the nasal peak below the F1, and P1 is the amplitude of the nasal peak between the first two vowel formants, F1 and F2. Pruthi and Espy-Wilson (2007) also tested A1-P1 and A1-P0 on several corpus databases and achieved an acceptable accuracy rate on each database (the highest rate is 96.28%). Thus, A1-P1 and A1-P0 are qualified as major acoustic parameters for the automatic detection of vowel nasalization.

In addition, several other methods of detecting vowel nasalization were developed by Glass and Zue (1985), such as counting extra nasal peaks across a vowel spectrum and measuring spectral flattening at the low frequency region (0-1kHz). Given the scope of this study, vowel nasalization will not be directly measured using the above methods, but the nasal coda influence on vowel quality will be investigated in terms of the vowel formant change over the duration.

2.4.2 Nasals

Nasal place distinction is acoustically characterized by nasal zero location. According to Ladefoged and Maddieson (1996), nasal zeros have an inverse relationship with the volume of the cavity; that is, the more forward the tongue or the lower the tongue body, the larger the cavity and the lower the first nasal zero. For example, the first nasal zero has a value of 1780Hz for Catalan alveolar /n/ and 3700Hz for Catalan velar /ŋ/ (Recasens, 1983). Generally speaking, the first nasal zero is below 1000Hz for /m/, between 1000-2000Hz for /n/, and above 3000Hz for /ŋ/ (Recasens, 1988).

Although the first nasal zero seems to be a good cue to nasal place, Qi and Fox (1992) pointed out that conventional spectral analyses (e.g., Linear Predictive Coding, or LPC) could not detect the first nasal zero effectively and efficiently, because the significantly damped high frequency energy in nasal spectra (due to the presence of nasal zeros) would introduce non-linear equations into linear technique based analyses such as LPC. Thus, it is difficult to measure non-linear nasal zeros directly through conventional linear methods such as LPC.

Instead of nasal zero location, another parameter, the band energy difference, also seems to be a good cue to nasal place. According to Kurowski and Blumstein (1987), there is less change in energy in the region of Bark 5-7 (395-770Hz) relative to that of Bark 11-14 (1265-2310Hz) for /n/

than for /m/; within /n/, the energy change in Bark 11-14 is greater than in Bark 5-7. Since the Bark 5-7 and Bark 11-14 regions respectively encompass the first nasal zeros of /m, n/ (Kurowski & Blumstein, 1987), the energy reduction difference in the two nasals, /m, n/, and in the two frequency regions, Bark 5-7 and Bark 11-14, of /n/ is largely due to the first nasal zero influence. Inferred from Kurowski and Blumstein's (1987) findings, this study assumes a larger energy reduction for /n/ than for /ŋ/ in the low-mid frequency (<3000Hz) region due to the higher first nasal zero value for /ŋ/ (> 3000Hz) than for /n/ (<3000Hz). In other words, the first nasal zero should be absent for /ŋ/ but present for /n/ in the low-mid frequency region, so there should be less energy reduction in this region for /ŋ/ than for /n/.

In addition, Seitz, et al. (1990) detected nasal place by measuring rapid spectral changes over murmur-to-vowel transitions through subtracting a specific vowel spectrum from a specific murmur spectrum within the vowel-nasal boundary. Having compared several methods of detecting nasal place, Harrington (1994) concluded that methods having high nasal classification scores are those taking account of the contribution of both murmurs and vowels to nasal place distinction. In short, a nasal place is manifested by both the vowel context and the nasal murmur.

As for the nasal place difference in nasal formants, Recasens' (1983) perception study of Catalan alveolar, palatal, and velar nasals, /n, ȵ, ŋ/, preceding [a] find that N1 and its bandwidth value are higher for /ŋ/ than for /n, ɳ/; that the F1 transitions from the vowel to the murmur are falling more for /ɳ/ than for /n, ŋ/, and that the F2 transitions are more steadily rising for /ɳ/ than for /n, ŋ/. Also, Recasens (1983) investigated the perceptual role of transitions and murmurs in nasal place recognition and showed that transitions are better at detecting /ɳ/ than murmurs, but murmurs are better at detecting /n, ŋ/ than transitions. In addition, Recasens (1988) mentioned that N2 is between 1000-1500Hz for /m/, between 1500-2000Hz for /n/, and around 2000Hz for /ɳ/ and /ŋ/.

According to Ladefoged (2001), however, nasal formants are generally not good nasal place cues, because the first, second, and third nasal formants (N1, N2, & N3) of all nasals have a similar frequency level respectively at 250Hz, 2500Hz, and 3250Hz. Specifically in this study, N1/N2/N3 will still be measured just in case they do show some significance in nasal place distinction.

2.4.3 Vowel and nasal duration

According to Rosner and Pickering (1994), vowel duration can be both phonologically and phonetically significant in vowel perception and production. For example, in Japanese, vowels are distinguished phonologically in terms of long and short; in English, close (high) vowels are generally shorter than open (low) vowels. In Maithili, open nasal vowels tend to have a longer duration than close vowels, and the mid, central nasal vowel /ə̃/ has the shortest duration (Jha, 1986). Ainsworth (1981) even claimed that a formant frequency shift of 100Hz is perceptually equivalent to a durational difference of 250ms.

Nasalization is also found to have an impact on vowel duration depending on vowel context. For example, Clumeck (1976) found that the velum lowers earlier during a low vowel than during a high vowel in his study of vowel nasalization in six languages, so that low vowels have both a longer vowel duration and a longer duration of vowel nasalization than high vowels. Also, he found that the duration of vowel nasalization is relatively long in American English and Brazilian Portuguese but short in Hindi, French, Swedish long vowels, and Amoy Chinese. Clumeck (1976) thus claimed that the timing of velum-lowering is language-specific because it can be controlled precisely by the speakers of different languages.

Furthermore, Solé (1992, 1995) studied phonetic versus phonological nasalization and questioned the phonetic nature of vowel nasalization in some languages. Specifically, she showed that velar-lowering starts at the beginning of the vowels preceding a nasal consonant in American English and under all 3 conditions, slow, normal, and fast speech; whereas vowel nasalization in Spanish is timed with the beginning of the nasal. Also, vowel nasalization in a VVN (i.e., VGn) syllable is very long in American English and can take up the 80%-100% of the vowel duration. Since the vowel nasalization does not occur just near the vowel-nasal boundary but always co-varies with the vowel rather than the following nasal onset, Solé (1992, 1995) claimed that a nasalized vowel in American English is an allophonic variation of the corresponding oral vowel (or phonologized) rather than a mere result of coarticulation. In other words, the vowel preceding a nasal in American English should be underlyingly nasal not oral.

Regardless of the phonetic or phonological status of vowel-nasal coarticulation in different languages, Manuel (1999) claimed that languages differing in their coarticulation patterns may be associated with their individual prosody patterns. For example, in a syllable-timed language such as Mandarin Chinese, each syllable tends to have the same length so that the syllable duration is relatively fixed, whereas in a stress-timed language such as English, syllable duration varies with syllable length. The different rhythmic structures in different languages also imply the different ways of timing a segmental sequence or the different "temporal coordination of articulatory gestures", to borrow Manuel's term (1999, p.196). In fact, White and Mattys explicitly stated that speech rhythm implies "some form of top-down control of speech segment duration to regularise the language-specific rhythmic intervals" (2007, p.19). If rhythm indeed has a top-down influence on segmental production, then in Mandarin and English VN production, the coarticulation pattern and the associated segmental duration should be different.

In addition, Busà's (2007) acoustic study of coarticulatory nasalization and Beddor, Brasher, and Narayan's (2007) perceptual study of the vowel nasalization in VNC sequences show that nasal duration is inversely related to vowel nasalization in American English; that is, the longer the vowel nasalization, the shorter the duration of the following nasal. Also, open vowels such as /æ/ have a longer period of nasalization than close vowels such as /ɛ/.

Chen (2000) also found that in Mandarin VN rimes, the degree, rate, and duration of vowel nasalization vary with vowel height; specifically, low vowels have a larger, slower, and longer period of nasalization than high vowels. A piece of evidence also comes from Chinese loan translations of the English names, *Tom* and *Tim*, respectively as /taŋ.mu/ and /ti.mu/ (Lin, 2007). In /taŋ.mu/, the extra /ŋ/ is used to substitute /ɑm/ in *Tom*, but no extra nasal is used to substitute /ɪm/ in *Tim*, which suggests that the low vowel /ɑ/ in *Tom* is perceived as longer and more nasalized than the high vowel /ɪ/ in *Tim*.

As for the relationship between the place of articulation and duration, Lehiste (1976) claimed that the consonantal place tends to correlate inversely with the preceding vowel duration; generally, vowels are shorter when preceding labials (which have the longest consonantal place from the pharyngeal wall to the mouth) than preceding coronals and velars. In Recasens' (1983) study of Catalan VN#, for example, m is 78ms long (the preceding vowel is 75ms long), but n is only 62 ms long (the preceding vowel is 87ms long). Chen (1972, 1975) even went so far as to claim that Mandarin /ŋ/ is two times longer than /n/ and /ŋ/ in Vŋ is four times longer than V. Unfortunately, Chen (1972, 1975) did not provide acoustic evidence to support his claim. Nonetheless, his claim is not without a basis. For example, English surname *King* (/kɪŋ/) is transcribed as Mandarin *jīn.ēn* (/tɕin.ən/), in which the velar /ŋ/ is split into two syllables by the schwa insertion (Lin, 2007). This transliteration suggests that velar /ŋ/ is long enough to become at least two short /n/s. Note that di-syllabification is a preferred Mandarin prosody (Lin, 2001).

Because open (low) vowels are longer than close (high) vowels, and /ŋ/ is longer than /n/, V_{open}ŋ (an open vowel followed by ŋ) should have the longest duration, and V_{close}n (a close vowel followed by n) should have the shortest duration among all types of VN rimes. Figure 2-1 summarizes such a relationship between VN type and duration.

Figure 2-1 The relationship between VN type and duration

|← Total V+N Duration →|

V_{type}	Vowel nasalization period	Nasal place
V_{open}	long	ŋ
V_{close}	short	ŋ
V_{open}	long	n
V_{close}	short	n

Note that V_{open}n and V_{close}ŋ have a comparable duration between the maximum and the minimum, but V_{open}n is assumed to be a little shorter than V_{close}ŋ on the basis that the most part of /n/ is probably co-articulated with the preceding open vowel because open vowels have a higher degree of nasalization than close vowels.

2.5 L2 acquisition theories

The characteristics of L2 speech have been captured by various L2 acquisition theories. Some of the influential theories relevant to this study include the L1 transfer theory and Flege's (1995) Speech Learning Model (SLM).

James (1988) systematically reviewed the general patterns of L1 influence on L2 acquisition; specifically, he pointed out that native language influence has been regarded as "a major source of 'explanation' for the nature of second language learning in general" (p. 30). From a behaviourist perspective, when L2 has the same structures as the so-called L1 "habit structures", positive transfer occurs; otherwise, negative transfer or interference may occur. Furthermore, James (1988) claimed that L1 sounds which are associated between L1 and L2 may be assessed by L2 learners for their transfer potential at various levels of phonological/phonetic organization. In other words, the transferability of an L1 sound into L2 is determined not only by segmental factors but also by syllabic and prosodic factors.

Flege's (1995) SLM, on the other hand, is more concerned with L2 sound acquisition from the perspective of perception. One of the most influential claims in SLM is that "category formation for English stops may be blocked by the continued perceptual linkage of L1 and L2 sounds (i.e., by equivalence classification)" (Flege, 1995, p. 258). This claim basically identifies similarities between L1 and L2 sounds as a source of interference in L2 sound acquisition.

In addition, Flege (1995) discussed the production and perception of word-final stop consonants and identified that Mandarin speakers distinguish /t/-/d/ mainly by closure voicing but rarely by the preceding vowel length. In other words, while native English speakers would also produce the vowel preceding /d/ longer than preceding /t/ in addition to voicing the final /d/, Mandarin speakers would attend only to the voicing distinction between /t/ and /d/. By the same token, Mandarin speakers may use different acoustic cues to differentiate the nasal place of /n, ŋ/ in their L1 and L2 production.

To sum up, previous studies concerning L1 and L2 nasal production and L2 acquisition provide both a theoretical and an experimental framework for this study; on the other hand, this study can provide a testing ground for previous theories and findings.

Chapter 3

THE EXPERIMENT

3.1 Participants

Twenty Mandarin Chinese speakers (10 females and 10 males) participated in this study. A one-page questionnaire (see Appendix 1) was administered to elicit participants' personal data such as gender, major, school year and English learning experience (see Appendix 2). The information gathered in this questionnaire was used to choose the right participants and to examine possible correlations between participants' background and their production patterns. The participants were mostly international students from the University of Victoria. All participants were between the ages of 19 and 40, and most of them were in the age group, 25-30. Eleven of them had received 10 years' formal English education before they came to Victoria, and 4 participants were ESL (English as Second Language) students.

3.2 Speech materials

Table 3-1 and 3-2 provide a total of 14 English and 8 Mandarin test words used in the word-list reading task. Table 3-1 includes 4 English and 4 Mandarin words with the VN type of rime, and the 4 words in each language contrast in vowel context (open vowel vs. close) and/or nasal place (alveolar vs. velar). The selection of these 8 words is for the purpose of investigating how vowel context and nasal place interact in both L1 and L2 production.

Table 3-1 Four English and four Mandarin CVN words

vowel context	English[1]		Mandarin[2]	
	/n/	/ŋ/	/n/	/ŋ/
close	*sin* (/sɪn/)	*sing* (/sɪŋ/)	*xìn* (/ɕin/)	*xìng* (/ɕiŋ/)
open	*son* (/sʌn/)	*song* (/sɔŋ/)	*sàn* (/san/)	*sàng* (/saŋ/)

[1] The English transcription is based on O'Grady & Archibald (2000).
[2] The Mandarin transcription is based on Lin (2001a).

Table 3-2 includes 5 English words with the VGn type of rime and the 5 corresponding English words with the VG type of rime. The selection of these 10 words is for the purpose of investigating how the nasal coda /n/ affects the production of a diphthong vowel and vice versa. In addition, 2 Mandarin words with the Vŋ type of rime, *gàng/gòng,* and 2 Mandarin words with the VG type of rime, *kào/gòu,* are included to contrast the 4 similar sounding English words, *gown/cone* and *cow/go,* respectively. The selection of these 4 words is for the purpose of investigating the vowel-nasal backness assimilation effect in both L1 and L2 production.

Table 3-2 Ten English[1]CVGn and CVG and four Mandarin[2] CVŋ and CVG words

V	VGn	VG	Vŋ	VG
/aj/	*pine* (/pajn/)	*pie* (/paj/)		
/oi/	*coin* (/kojn/)	*coy* (/koj/)		
/aw/	*gown* (/gawn/)	*cow* (/kaw/)	*gàng* (/gaŋ/)	*kào* (/kaw/)
/ej/	*pain* (/pejn/)	*pay* (/pej/)		
/ow/	*cone* (/kown/)	*go* (/gow/)	*gòng* (/guŋ/)	*gòu* (/gow/)

[1] The English transcription is based on O'Grady & Archibald (2000).
[2] The Mandarin transcription is based on Lin (2007).

Note that all the Mandarin test words bear the falling Mandarin fourth tone to simulate the natural falling pitch of the English test words, though the pitch fall is much more gradual in English than in the Mandarin 4th tone.

3.3 Data collection procedure

First, participants were instructed to read a consent form, and after they signed the consent form and filled out the questionnaire, they were shown the paper form of the word-list (see Appendix 4). Each English test word was listed in a separate row, with the Chinese gloss and a rime word (a very common word) in the same row. They were asked to go through the word-list to identify words that were unfamiliar to them. Most participants claimed that they knew all the words, so only a few participants used the help from the rime words. For example, if a participant did not know the word *coy*, I would point out *toy* to her so that she knew *coy* rhymes with *toy* and could produce *coy* easily.

Participants were further instructed to practise the on-screen reading of the test words presented randomly in a PowerPoint Window. Note that for Mandarin test words, Chinese characters were presented on screen instead of the *Pinyin* representation. Each test word successively appeared 4 times (hence 4 tokens for each word) in a slide. The successive appearance of the four tokens was intended to improve the chance for the word to be produced consistently. There was a 2-second interval following each appearance of the word and the participants were instructed to read each word according to the rhythm of its appearance. A total of 88 tokens (22x4) were collected for each participant. While participants were practising on-screen reading, a trial recording was carried out before the formal recording to ensure the recording quality.

The reading task was performed in a sound-attenuated room in the phonetics laboratory of the University of Victoria. A large diaphragm condenser microphone (M-Audio Lunar) was placed at about a 10 cm distance from the participant's mouth. The recording workstation was a Windows XP PC equipped with a Mic Preamp and A/D converter (M-Audio firewire 410), and the recording software was Audacity 1.2.4. The sampling frequency was 44100Hz. Each recording file was saved in .wav format. Then, speech data were acoustically analyzed using Praat 4.4.22.

3.4 Data analysis

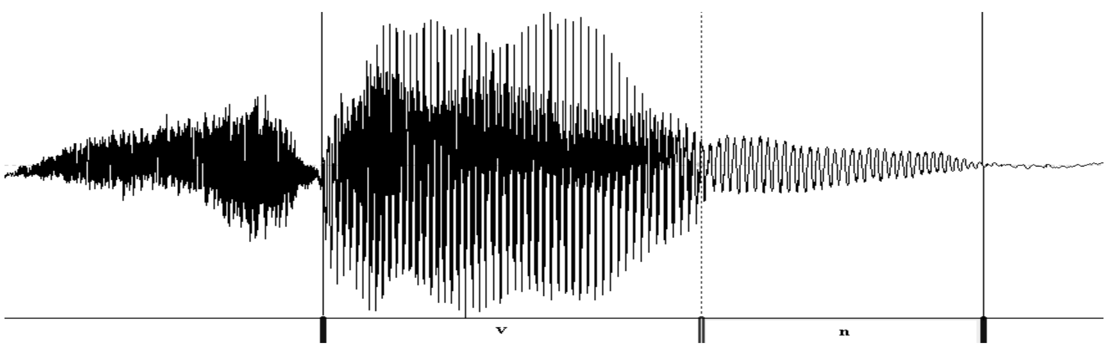
3.4.1 Segmentation

Before the data analysis, tokens were extracted from the initial .wav file by using Praat script 2 (all scripts used in this study were adapted by the researcher from a number of sources. See Appendix 3 for the script description and the source). A preliminary .TextGrid file was then automatically created by Praat script 3 to segment and label each extracted token in terms of the vowel and nasal murmur intervals.

When the .TextGrid file, along with the token's waveform and spectrogram, was read into the Praat object window by using Praat script 1, the preliminary segmenting points were manually adjusted by the researcher. During the manual segmentation process, the first zero-crossing point of the first glottal pulse of the vowel (can be clearly shown in the enlarged waveform) was marked as the start point of the vowel and the last zero-crossing point of the last glottal pulse of the nasal murmur as the end point of the nasal murmur.

The trickiest part for labeling the vowel and nasal murmur intervals is to determine the vowel and nasal murmur boundary (i.e., the end of the vowel and the start of the nasal murmur). Generally, the end of the vowel was chosen at a zero-crossing point where the vowel formants (indicated by the relatively complicated waveform) disappear and the glottal pulse abruptly changes in both the pattern and the amplitude. If a boundary point is not clear, the segmentation process was also augmented through the researcher's auditory judgment and the spectrogram display. Figure 3-1 illustrates the labeled waveform display of a token of *son* produced by a female speaker. V represents the vowel interval and N represents the nasal murmur interval.

Figure 3-1 An illustration of segmentation



3.4.2 Acoustic measurements

This study used Praat scrip 4-7 to measure the following acoustic parameters of vowels and nasals:

(1) Vowels

i) V_D: the vowel duration in a token (unit: s).

ii) F0_fh & _sh: the mean fundamental frequencies over the 1^{st} and 2^{nd} half of the vowel duration (unit: Hz).

iii) F1_fh & _sh: the mean 1^{st} formant frequencies over the 1^{st} and 2^{nd} half of the vowel duration (unit: Hz).

iv) F2_fh & _sh: the mean 2^{nd} frequencies over the 1^{st} and 2^{nd} half of the vowel duration (unit: Hz).

v) F3_fh & _sh: the mean 3^{rd} frequencies over the 1^{st} and 2^{nd} half of the vowel duration (unit: Hz).

(2) Nasals

i) N_D: the nasal murmur duration in a token (unit: s).

ii) N1: the 1^{st} formant frequency at the midpoint of the nasal murmur duration (unit: Hz).

iii) N2: the 2^{nd} formant frequency at the midpoint of the nasal murmur duration (unit: Hz).

iv) N3: the 3^{rd} formant frequency at the midpoint of the nasal murmur duration (unit: Hz).

v) ΔdB: the mean band energy difference between 0-525Hz and 525-1265Hz over the nasal duration (unit: dB).

The vowel F0/F1/F2/F3 frequencies were automatically measured by Praat Script 5. The nasal N1/N2/N3 frequencies and ΔdB were automatically measured by Praat Script 6 and 7, respectively. V_D and N_D were automatically measured by Praat Script 4.

The measuring process was monitored and the results were visually checked for any abnormality. If necessary, the researcher would manually calculate the acoustic parameters. For example, a female speaker's production of a token was rather creaky, the widely spaced glottal pulses made the automatic calculation impossible. Therefore, a manual calculation based on the spectrum display of the token was performed by the researcher.

In addition, the following acoustic parameters (some of which were derived from the above basic ones) were used to correlate with (⇔) a segmental feature of vowels or nasals:

i) F1-F0_fh & _sh ⇔ Vowel height: If F1-F0_fh > _sh, then the vowel height decreases over the duration.

ii) F3-F2_fh & _sh⇔ Vowel backness: If F3-F2_fh < _sh, then the vowel backness increases over the duration.

Note that a vowel is assumed to be less influenced by the nasal coda over the first half than over the second half of the duration. For a diphthong vowel (VG), V and G can be roughly represented by the first and second half of the vowel. Thus, a comparison of formant patterns between the first and second half of a vowel can be used to indicate the general direction of the vowel movement across the auditory space (i.e., vowel height/backness changes over the duration).

iii) N_D% ⇔ The degree of vowel-nasal coupling: The greater the N_D %, the less the vowel and the nasal overlap in time and the less the degree of vowel-nasal coupling.

iv) N1/N2/N3 and ΔdB ⇔ nasal place: the greater the N1 and N2 (N3 does not seem to be a useful predictor of nasal place), and the smaller the ΔdB, the more backed the nasal place.

3.5 Statistical analyses

Acoustic data measured and calculated by the Praat scripts 4-7 were fed first into Microsoft Office Excel 2003, then into the statistical software, SPSS 16.0 for Windows. The following sections illustrate two statistical analyses used by this study to find out whether or not there exist significant differences among the test words in terms of the above acoustic correlates.

3.5.1 Two-tailed paired samples t-test

Two t-tests are used to compare respectively between mean F1-F0_fh and _sh and between mean F3-F2_fh and _sh for each test word across tokens and speakers. If there is a significant change in vowel height over the duration, mean F1-F0_fh and _sh should be statistically different; similarly, if there is a significant change in vowel backness over the duration, mean F3-F2_fh and _sh should be also statistically different. Hence, the statistical results are further used to estimate the significance of nasal coda influence on the preceding vowel. If the nasal coda influence is significant, for instance, the height/backness of a monophthong vowel will change significantly over the duration; however, for a diphthong vowel, the height/backness change over the duration are expected, so the nasal coda influence may be indicated by the otherwise non-significant vowel changes.

3.5.2 Repeated measures one-way ANOVA

First, two ANOVA tests are used to compare respectively among mean N_D%s and mean Ds for *sin/sing/xìn/xìng/son/song/sàn/sàng* across tokens and speakers, and two more ANOVA tests are used to compare respectively among mean N_D%s and mean Ds for *pine/coin/gown/pain/cone/gàng/gòng* across tokens and speakers. The statistical results are used to indicate whether or not a word is significantly different from the remaining words in terms of N_D% and D.

Second, three ANOVA tests are used to compare respectively among mean N1s, mean N2s, and mean N3s for *sin/sing/xìn/xìng/son/ song/sàn/sàng* across tokens and speakers, and three more ANOVA tests are used to compare respectively among mean N1s, mean N2s, and mean N3s for *pine/coin/gown/pain/cone/gàng/gòng* across tokens and speakers. The statistical results are used to indicate whether or not a word is significantly different from the remaining words in terms of N1/N2/N3.

Third, two ANOVA tests are used to compare among mean ΔdBs respectively for *sin/sing/xìn/xìng/son/song/sàn/sàng* and *pine/coin/gown/pain/cone/gàng/gòng* across tokens and speakers, and the statistical results are used to indicate whether or not a word is significantly different from the remaining words in terms of ΔdB.

All of the above statistical results from the ANOVA tests are used to infer the nasal place difference among test words.

Chapter 4

RESULTS

4.1 Results on VN production

4.1.1 Vowel measurements

Figure 4-1 and 4-2 respectively illustrate vowel height/backness changes over the duration for the 4 English words, *sin/sing/son/song*, and the 4 Mandarin words, *xìn/xìng/sàn/sàng*. The start point of each arrowed line represents mean F3-F2 (the x-axis) and mean F1-F0 (the y-axis) over the first half of the vowel duration, and the end point (where the arrow head is) represents mean F3-F2 (the x-axis) and mean F1-F0 (the y-axis) over the second half of the duration. As discussed in Section 2.4.1, the greater the mean F3-F2 and the smaller the mean F1-F0, the more backed and higher the vowel. Thus, vowels in Figure 4-1 and 4-2 reflect their relative position in the auditory space, and the arrow points toward the general direction of the vowel movement across the space.

*Figure 4-1   Mean F3-F2 and mean F1-F0 over the first and second half of vowel duration for **sin/sing/xìn/xìng** (Unit: Hz)*

	sin	*sing*	*xìn*	*xìng*
F3-F2_fh	838	781	891	732
F1-F0_fh	268	269	235	238
F3-F2_sh	1000	965	1014	908
F1-F0_sh	228	235	226	270

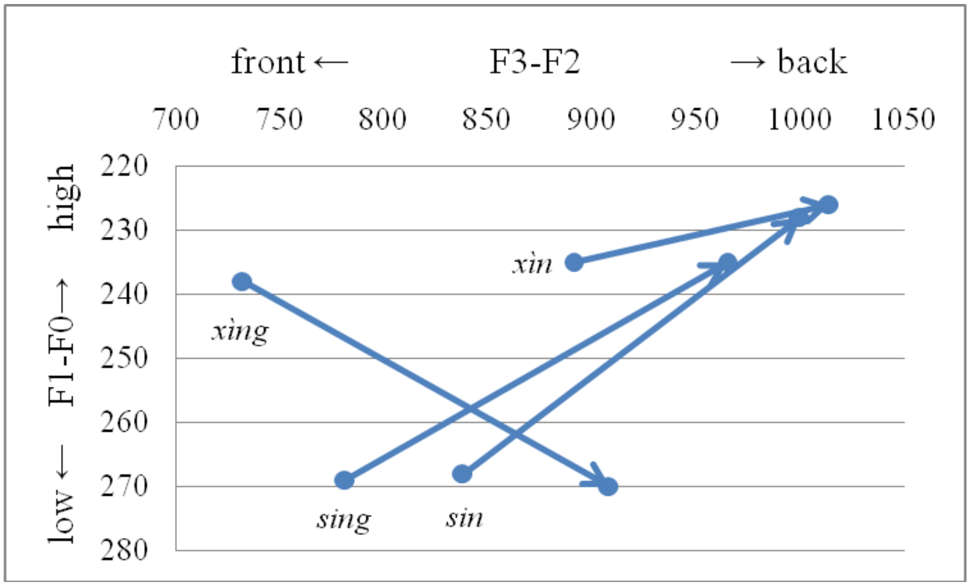


Figure 4-1 shows a general tendency for vowels in *sin/sing/xìn/xìng* to move backwards over the duration, indicating that a nasal coda has a general effect to render a high, front vowel more backed. Note that only for *xìng*, the vowel movement over the duration is towards the low back rather than high back, indicating that Mandarin /ŋ/ has a special effect on the preceding high, front vowel /i/.

A 2-tailed paired samples t-test revealed that the difference between mean F3-F2_fh and _sh is significant for *xìng* at the 5% level: t_{19} = -2.461, *p*

= .024, which shows that there is a significant change in vowel backness for *xìng* over the duration. Note that Xu's (1999) articulatory description of Standard Mandarin sounds treats the rime in Mandarin *xìng* as having identical phonemic and phonetic forms (i.e., /iŋ/ = [iŋ]). However, Lin (2007) treated the phonetic form of /iŋ/ as [jəŋ] based on the backness constraint. The above result basically supports Lin's (2007) treatment.

Figure 4-2 Mean F3-F2 and mean F1-F0 over the first and second half of vowel duration for ***son/song/sàn/sàng*** *(unit: Hz)*

	son	*song*	*sàn*	*sàng*
F3-F2_fh	1480	1767	1333	1526
F1-F0_fh	530	581	604	614
F3-F2_sh	1437	1714	1369	1703
F1-F0_sh	555	571	607	698

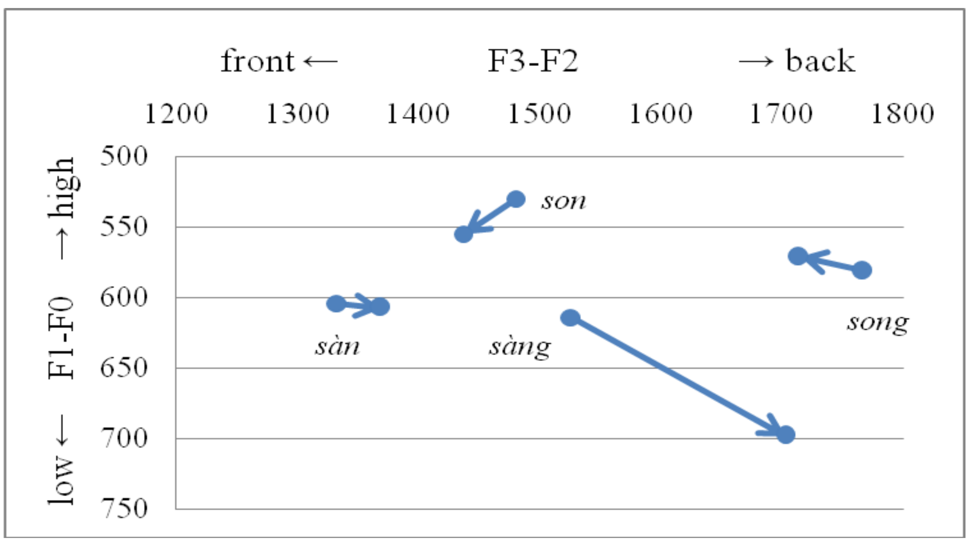


Figure 4-2 shows a general tendency for vowels in *sàn/sàng* to move towards the back and for vowels in *son/song* to move towards the front over the duration. Note that there are greater vowel height/backness changes over the duration for *sàng* than for the rest of the words, indicating a greater influence of Mandarin /ŋ/ on the open vowel /a/.

Also, vowels in the 4 words are distinctively different from one another in terms of their relative initial position: the vowel /a/ in *sàn* is more fronted and the vowel /ɔ/ in *song* is more backed than in the rest of the words. Also, the vowel /ʌ/ in *son* is higher and the vowel /a/ in *sàng* is more backed than in the rest words. The distribution of the 4 vowels in Figure 4-2 basically corresponds to their phonological representation described in Section 3.2.

A 2-tailed paired samples t-test revealed that the difference between mean F1-F0_fh and _sh is significant for *sàng* at the 5% level: $t_{19} = -2.370$, $p = .029$, which shows that there is a significant change in vowel height for *sàng* over the duration. Note that Lin (2001a) treated the rime in *sàng* as having the phonemic form, /aŋ/, and the phonetic form, [ɑŋ]. However, since /a/ is considered as a low, central vowel and /ɑ/ as a low, back vowel by the traditional definition, it does not seem plausible that the vowel change in *sàng* is mainly in height not in backness as presumed. A sensible way of understanding the vowel height change from /a/ to /ɑ/ is to adopt Esling's (2005) definition of vowels in terms of the front-open-raised-retracted dimension rather than the high-low-front-back dimension. According to Esling (2005), the difference between /a/ and /ɑ/ is not related so much to the static location of tongue body along the front-back dimension as to the degree of jaw opening and tongue retracting. In other words, /a/ moving to /ɑ/ involves at least two actions, jaw opening and tongue retracting, and each action can affect the vowel formant measurements in its own way. For example, increased jaw-opening raises F1 (Rosner& Pickering, 1994). Therefore, the significant F1-F0 change in *sàng* can be viewed as a result of the change in the two actions rather than in the tongue height.

Figure 4-1 and 4-2 together show that Mandarin /ŋ/ can be distinguished from alveolar /n/ in terms of its greater influence on the preceding vowel. In contrast, English /ŋ/ produced by the Mandarin speakers does not have such influence on the preceding vowel.

In addition, a visual examination of the beginning and end points of *sin/sing/xìn/xìng* (i.e., the first and second half of the vowel articulation) in Figure 4-1 and *son/song/sàn/sàng* in Figure 4-2 revealed that the vowels in *son/song/sàn/sàng* scatter more widely across the auditory space than in *sin/sing/xìn/xìng*, which conforms to Stevens' (1989) quantal theory; that is, as Sydal and Gopal (1986) put it, closely spaced vowels (i.e., high, front vowels) have minimal acoustic variability. Sussman (1990) also mentioned that the stop place is better discriminated in back vowel contexts than front vowel contexts due to larger acoustic variability. Thus, the nasal place should be easier to distinguish in *son/song/sàn/sàng* than in *sin/sing/xìn/xìng*.

4.1.2 Durational measurements

Figure 4-3 illustrates mean V_D (vowel duration), mean N_D (nasal murmur duration), and mean D (the total vowel and nasal duration) for each of the 8 words, *sin/sing/xìn/xìng/son/song/sàn/sàng*, across tokens and speakers. The x-axis represents the duration in second (s) and each bar along the y-axis represents each of the 8 words.

Figure 4-3 Mean V_D, mean N_D, and mean D for ***sin/sing/xìn/xìng/ son/song/sàn/sàng***

	mean V_D	mean N_D	mean D
sin	0.25	0.21	0.46
sing	0.26	0.23	0.49
xìn	0.17	0.12	0.29
xìng	0.17	0.13	0.30
son	0.23	0.19	0.42
song	0.33	0.19	0.52
sàn	0.20	0.11	0.31
sàng	0.22	0.11	0.33

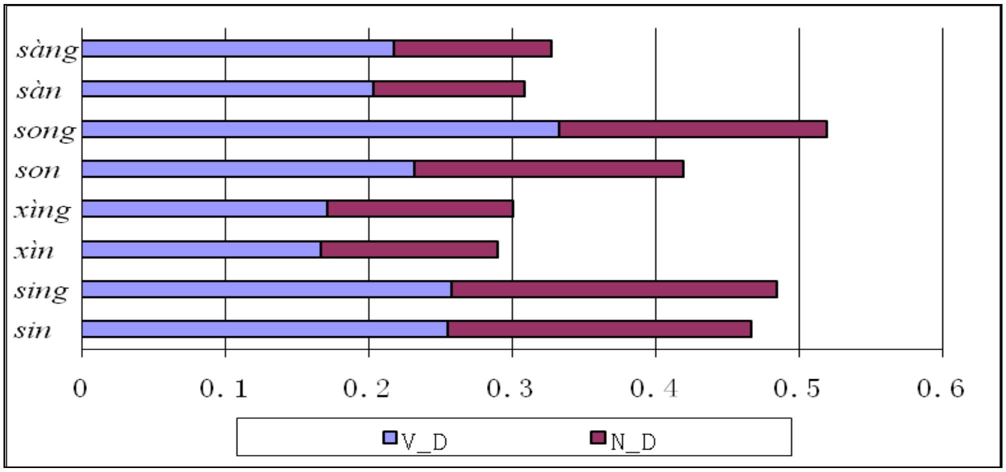


Figure 4-3 shows that the 4 Mandarin words, *xìn/xìng/sàn/sàng* (their average D = 0.31s), are shorter than the 4 English words, *sin/sing/son/song*, (their average D = 0.47s), which is expected because the 4th tone (hence the associated word) is the shortest among all 4 Mandarin tones when produced in isolation (Ho, 1976). Notice that all the 4 Mandarin words have a similar D around 0.3s, despite their difference in vowel context and/or nasal place. As mentioned in Section 2.4.3, Mandarin is a syllable-timed language, so the duration of the 4 Mandarin words (i.e., 4 syllables) is expected to be relatively equal. Although the 4 English words are all monosyllabic, their D is varied depending on the vowel context; specifically, *sin/sing* have similar duration but *son*/*song* respectively have the shortest and the longest duration among the 4 English words produced by the Mandarin speakers. A repeated measures one-way ANOVA test revealed that there is a significant difference among mean Ds for the 8 words, $F_{7,\,133}$ = 25.786, p < .001, and this is a medium effect size (partial eta-squared = .576). Specifically, the pairwise comparisons in the ANOVA test showed that D for *xìn/xìng/sàn/sàng* is significantly smaller than for *sin/sing/son/song* at the 5% level: $p < .001$. Also, D for *son* is significantly smaller than *song* at the 5% level: $p = .014$. Note that the difference in D between *son* and *song* mainly results from the difference in V_D; that is, V_D for *son* is much smaller than in *song*, suggesting that the vowel /ʌ/ in *son* is higher and thus shorter than the vowel /ɔ/ in *song*.

Figure 4-4 illustrates mean N_D% (the percentage of the nasal murmur duration over the total vowel and nasal murmur duration) for each of the 8 words, *sin/sing/xìn/xìng/son/song/sàn/sàng*, across tokens and speakers. The x-axis represents N_D/D in percentage (%), and each bar along the y-axis represents each of the 8 words.

Figure 4-4 Mean N_D% for ***sin/sing/xìn/xìng/son/song/sàn/sàng***

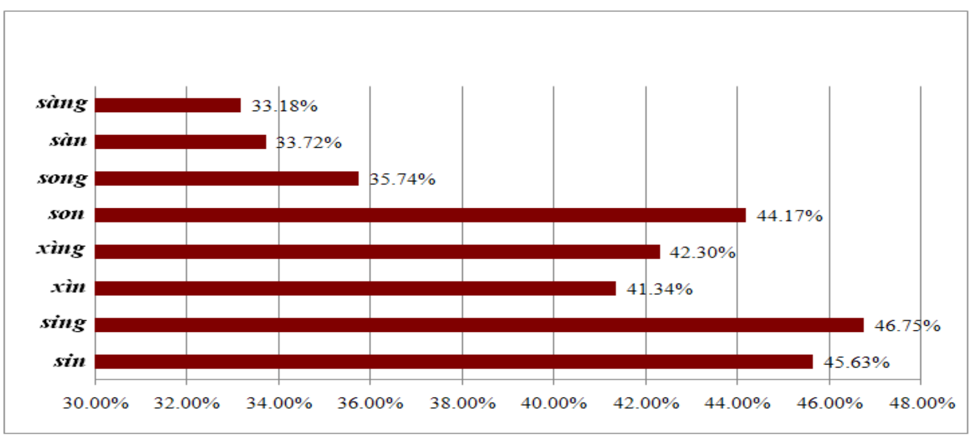


Figure 4-4 shows that mean N_D% is smaller (<40%) in *song*/*sàn/sàng* than in *sin/sing/xìn/xìng*/*son* (>40%), suggesting that vowels in *song*/*sàn/sàng* are lower and thus have a higher degree of nasalization than in *sin/sing/xìn/xìng*/*son*. As mentioned in Section 2.4.3, the degree of vowel nasalization correlates inversely with nasal murmur duration; that is, the longer the nasal murmur duration (hence the larger N_D%), the lower the degree of vowel nasalization since there is less overlap between the vowel and the nasal. Note that the vowel /ʌ/ in *son* is identified as being similar to high vowels in *sin/sing/xìn/xìng* rather than low vowels in *song*/*sàn/sàng* in terms of its large N_D% (44.17%). Recall in Figure 4-2, /ʌ/ in *son* is higher than in *song*/*sàn/sàng*, similar to a high-mid vowel, so it is not surprising that N_D% in *son* is comparable to that in *sin/sing/xìn/xìng* (all > 40%). Notice also that Mandarin *xìn/xìng* have a smaller N_D% than the corresponding English *sin/sing*, and *sàn/sàng* have a smaller N_D% than the corresponding *son/song*, suggesting that the Mandarin VN rimes have a higher degree of vowel nasalization and thus a smaller N_D% than their English counterparts.

A repeated measures one-way ANOVA test revealed that there is a significant difference among mean N_D%s for the 8 words, $F_{7,\,133} = 6.244$, $p = .022$, though this is a relatively small effect size (partial eta-squared = .247). Specifically, the pairwise comparisons in the ANOVA test showed that N_D% for *son* is significantly larger than for *song/sàn/sàng* at the 5% level: p = .033, .001, .001, respectively, which confirms that the vowel in *son* can be considered as a high-mid vowel rather than the presumed low-mid vowel /ʌ/.

Figure 4-3 and 4-4 together show that English /n, ŋ/ produced by the Mandarin speakers can be distinctively distinguished by the difference in N_D% and D in the open (low) vowel context, but this is not the case in the Mandarin /n, ŋ/ production.

4.1.3 Nasal measurements

Figure 4-5 illustrates mean N1/N2/N3 for each of the 8 words, *sin/sing/xìn/xìng/son/song/sàn/sàng*, across tokens and speakers. Each dot along the x-axis successively represents each of the 8 words, and the y-axis represents the formant frequency value in Hz.

Figure 4-5 Mean N1/N2/N3 for ***sin/sing/xìn/xìng/son/song/sàn/sàng***

	mean N1	*mean N2*	*mean N3*
sin	286	1670	2695
sing	279	1584	2623
xìn	266	1800	2879
xìng	294	1529	2681
son	285	1601	2699
song	278	1381	2676
sàn	258	1811	2845
sàng	368	1289	2814

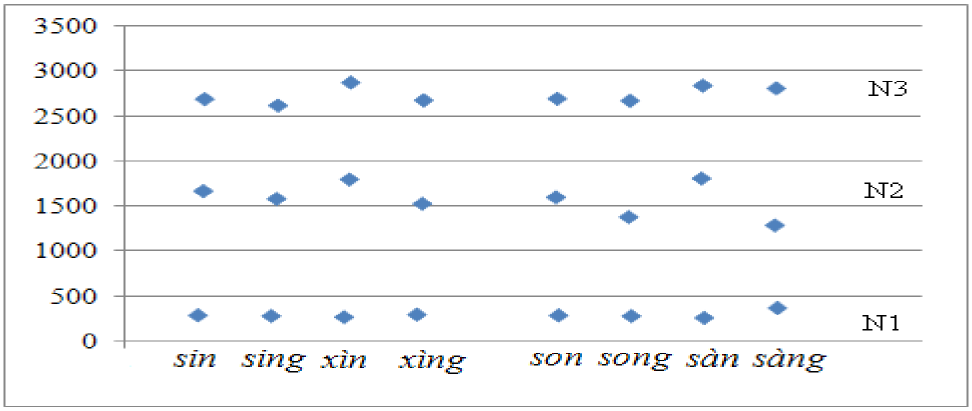


Figure 4-5 shows that mean N1 for *xìng/sàng* is higher than for the rest of the words, suggesting that the nasal place in *xìng/sàng* is at the back as expected, because N1 increases as the nasal place moves towards the back (Recasens, 1983). Note that, however, N1 for *sing/song* is not as high as expected given that the nasal place in *sing/song* is presumably at the back, similar to that in Mandarin *xìng/sàng*. Next, mean N2 for *sing/song* and *xìng/sàng* is respectively lower than for *sin/son* and *xìn/sàn*, which is consistent with the finding that velar articulation (as in /ŋ/) generally has a low or medium F2 (Fant, 2004).

A repeated measures one-way ANOVA test revealed that there is a significant difference among mean N2s for the 8 words, $F_{7, 133} = 6.139$, $p = .023$, though this is a relatively small effect size (partial eta-squared = .244). Specifically, the pairwise comparisons in the ANOVA test showed that N2 for *xìng/sàng* is significantly lower than for *xìn/sàn* at the 5% level: $p = .034$ between *xìng* and *xìn*, and $p < .001$ between *sàng* and *sàn*. Thus, the nasal place is distinctively different in the L1 production of Mandarin *xìn/sàn* and *xìng/sàng* but not in the L2 production of English *sin/son* and *sing/song*.

Figure 4-6 illustrates mean ∆dB for each of the 8 words, *sin/sing/xìn/xìng/son/song/sàn/sàng*, across tokens and speakers. The y-axis represents the band energy difference in Decibels (dB), and each pair of bars along the x-axis represents two words that contrast in nasal place.

Figure 4-6 Mean ∆dB for ***sin/sing/xìn/xìng/son/song/sàn/sàng***

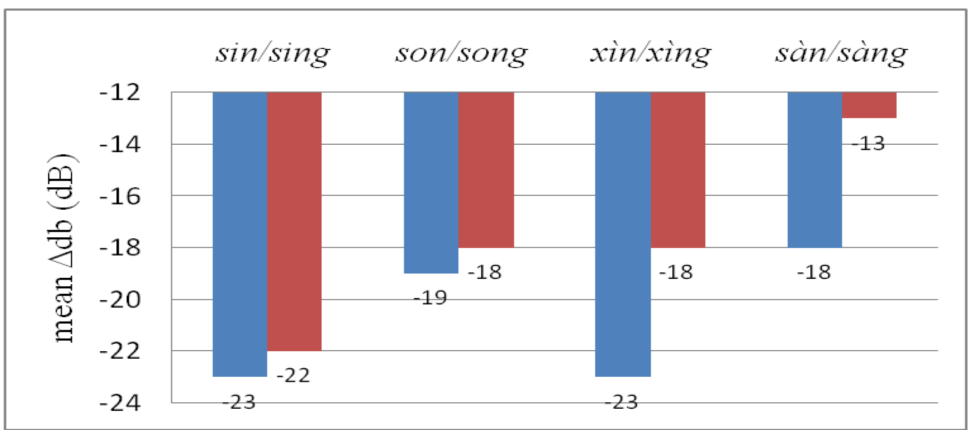


Figure 4-6 shows that mean ∆dB for *sàng* is the smallest (the closer to zero the negative number, the smaller the ∆dB) among the 8 words, suggesting that the nasal place in *sàng* is at the further back than in the rest of the words.

A repeated measures one-way ANOVA test revealed that there is a significant difference among mean ∆dBs for the 8 words, $F_{7,\,133} = 9.189$, $p = .007$, though this is a relatively small effect size (partial eta-squared = .326). Specifically, the pairwise comparisons in the ANOVA test showed that mean ΔdB for *xìng/sàng* is significantly smaller than for *xìn/sàn* at the 5% level: p = .009 between *xìng* and *xìn*, and p < .001 between *sàng* and *sàn*, confirming that *xìn/sàn* are significantly contrasted with *xìng/sàng* by nasal place.

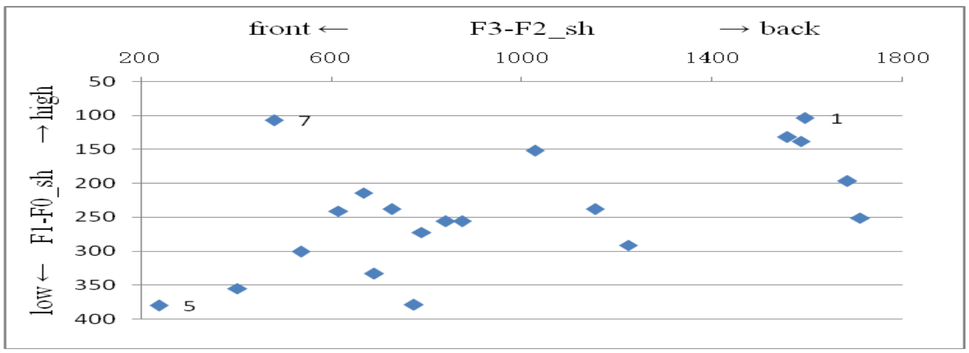
Figure 4-5 and 4-6 together show that the Mandarin nasal place in *xìng/sàng* is distinctively different from *xìn/sàn* in terms of both N2 and ΔdB, but this is not the case for the English nasal place in *sin/sing/son/song* produced by the Mandarin speakers.

4.1.4 Production variations for ***sin/sing/son/song*** *across speakers*

The above sections have shown the general patterns of L2 VN production; this section will show production variations across speakers for each of the 4 English words, *sin/sing/son/song* (the 4 tokens of each word were averaged for each speaker), in terms of vowel height/backness over the second half of the duration (assuming that the second half is subject more to the nasal coda influence than the first half) and mean ΔdB (assuming a major indicator of nasal place).

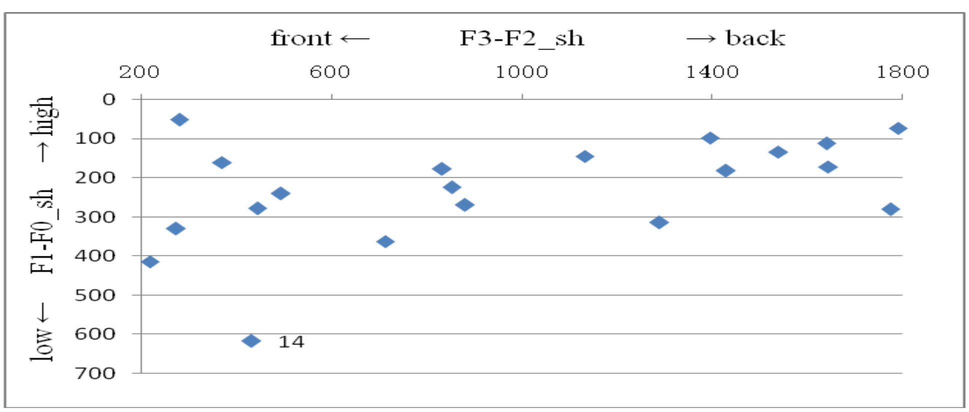
Figures 4-7 to 4-10 show the vowel distribution over the second half of the duration respectively in *sin/sing/son/song* across speakers. The x-axis represents mean F3-F2 over the second half of the duration (the greater the F3-F2_sh, the further back the vowel). The y-axis represents mean F1-F0 over the second half of the duration (the greater the F1-F0_sh, the lower the vowel). Each dot represents a vowel produced by each speaker (20 dots or speakers in total).

Figure 4-7 Vowel plots over the second half of the duration for ***sin*** *across speakers (Unit: Hz)*



In Figure 4-7, most dots fall within 100-400Hz in F1-F0 and 400-1800Hz in F3-F2. Speaker 7 produced a relatively high, front vowel, and Speaker 5 produced a relatively low, front vowel in *sin*.

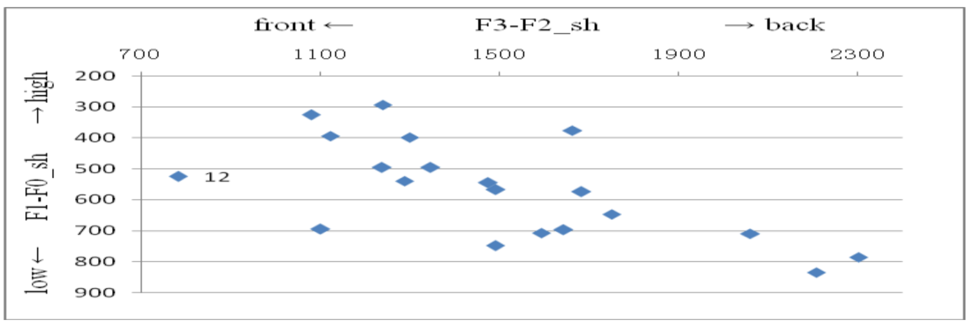
Figure 4-8 Vowel plots over the second half of the duration for ***sing*** *across speakers (Unit: Hz)*



In Figure 4-8, most dots fall within 100-300Hz in F1-F0 and 200-1800Hz in F3-F2. Speaker 14 produced a particularly low, front vowel in *sing*.

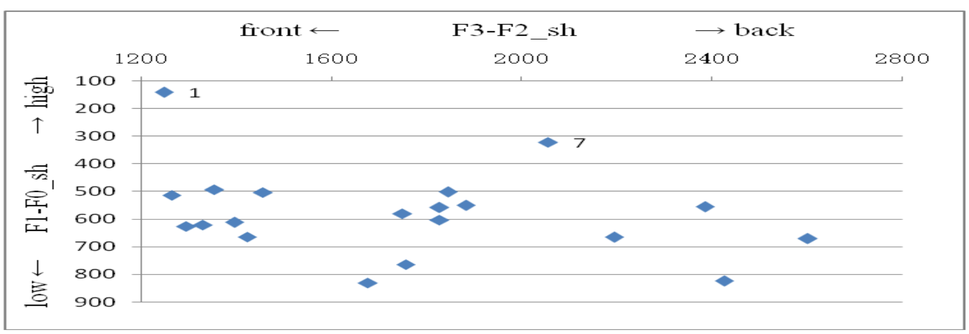
Figure 4-7 and 4-8 together show that dots in *sing* spread more widely along the x-axis than in *sin*, indicating that the vowel in *sing* has more variations in backness than in *sin*.

Figure 4-9 Vowel plots over the second half of the duration for ***son*** *across speakers (Unt: Hz)*



In Figure 4-9, most dots fall within 300-700Hz in F1-F0 and 1100-1700Hz in F3-F2. Speaker 12 produced a particularly front vowel in *son*.

Figure 4-10 Vowel plots over the second half of the duration for ***song*** *across speakers (Unit: Hz)*

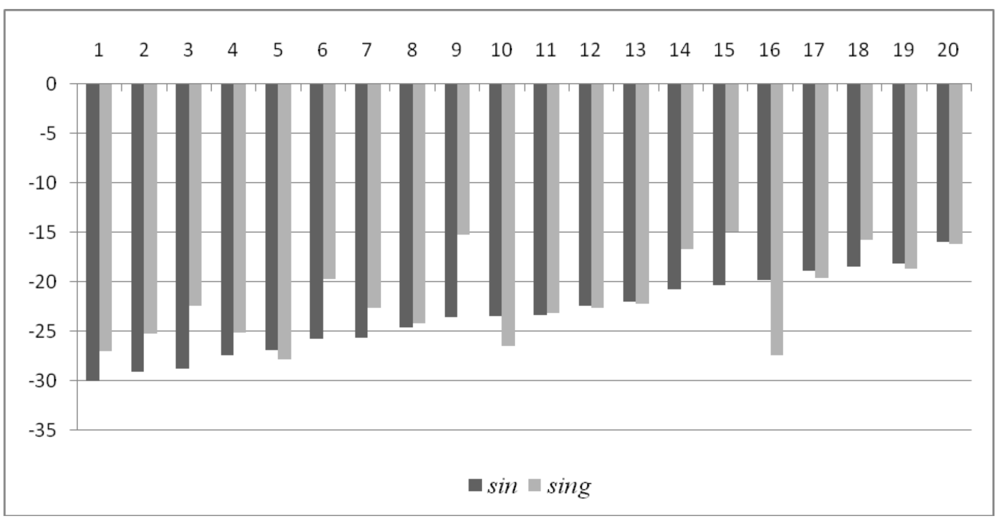


In Figure 4-10, most dots fall within 500-700Hz in F1-F0 and 1200-2000Hz in F3-F2. Speaker 1 produced a particularly high, central vowel in *song*.

Figure 4-9 and 4-10 together show that dots in *son* spread more widely along the y-axis than in *song*, indicating that /ʌ/ in *son* has more variations in height than /ɔ/ in *song*.

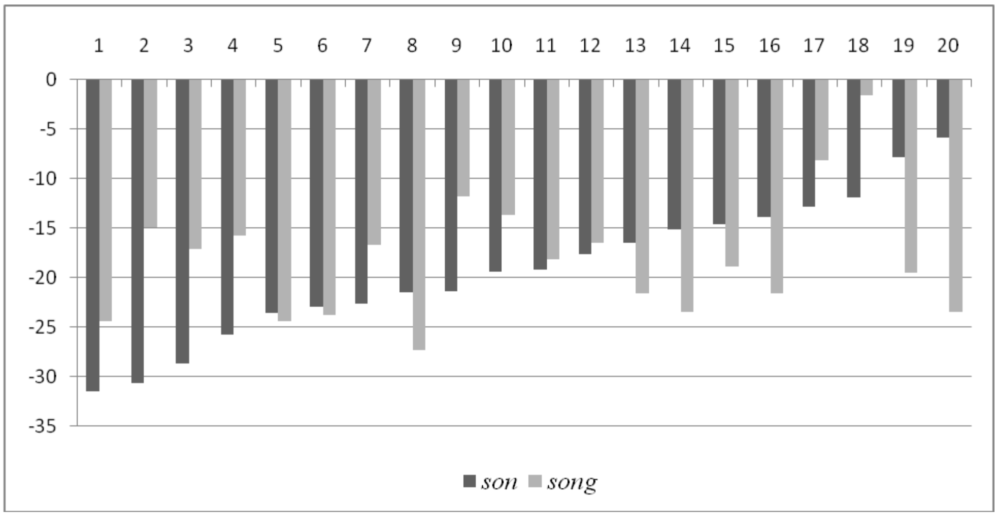
In the following Figure 4-11 and 4-12, the x-axis represents 20 speakers and the y-axis represents mean ΔdB (the lower the ΔdB, the more fronted the nasal place). Dark/grey bars in each figure represent mean ΔdBs for 2 words that contrast in nasal place.

Figure 4-11 Mean ΔdB for ***sin/sing*** *across speakers (Unit: dB)*



In Figure 4-11, one grey bar (for Speaker 16) is clearly higher than the corresponding dark bar, indicating that in this speaker's production, the nasal place in *sing* is probably less backed than in *sin* because there is a greater ΔdB decrease in *sing* than in *sin*.

Figure 4-12 Mean ΔdB for ***son/song*** *across speakers (Unit: dB)*



In Figure 4-12, seven grey bars (for Speaker 8/13/14/15/16/19/20) are clearly higher than the corresponding dark bars, indicating that in these speakers' production, the nasal place in *song* is probably less backed than in *son* because there is a greater ΔdB decrease in *song* than in *son*.

Figure 4-11 and 4-12 together indicate that in most speakers' production, velar /ŋ/ in *sing/song* has a smaller ΔdB than alveolar /n/ in *sin/son* as expected. Also, mean ΔdBs for *sin/sing* are generally less varied than for *son/song* (especially evident in Speaker 20's production of *sin/sing* and *son/song*), suggesting that the speakers interchangeably produced /n, ŋ/ in *sin/sing* more often than in *son/song*.

4.2 Results on VGn production

4.2.1 Vowel measurements

Figure 4-13, 4-14, and 4-15 illustrate vowel height/backness change over the first and second half of the duration for the 5 pairs of English words, *pine/pie, cone/go, coin/coy, gown/cow*, and *pain/pay,* and the 2 pairs of Mandarin words, *gàng/kào* and *gòng/gòu*. The start point of each arrowed line represents mean F3-F2 (the x-axis) and mean F1-F0 (the y-axis) over the first half of the duration, and the end point (where the arrow head is) represents mean F3-F2 (the x-axis) and mean F1-F0 (the y-axis) over the second half of the duration. Again, the greater the mean F3-F2 and the smaller the mean F1-F0, the more backed and higher the vowel. Thus, vowels presented in Figure 4-13, 4-14, and 4-15 reflect their relative position in the auditory space, and the arrow points toward the general direction of vowel movement across the space.

Figure 4-13 Mean F3-F2 and mean F1-F0 over the first and second half of vowel duration for ***pine/pie/coin/coy/pain/pay*** *(Unit: Hz)*

Mean value	*pine*	*pie*	*coin*	*coy*	*pain*	*pay*
F3-F2_fh	1699	1496	1918	1863	845	805
F1-F0_fh	672	743	464	482	520	468
F3-F2_sh	1047	867	1298	911	827	715
F1-F0_sh	556	743	379	335	319	272

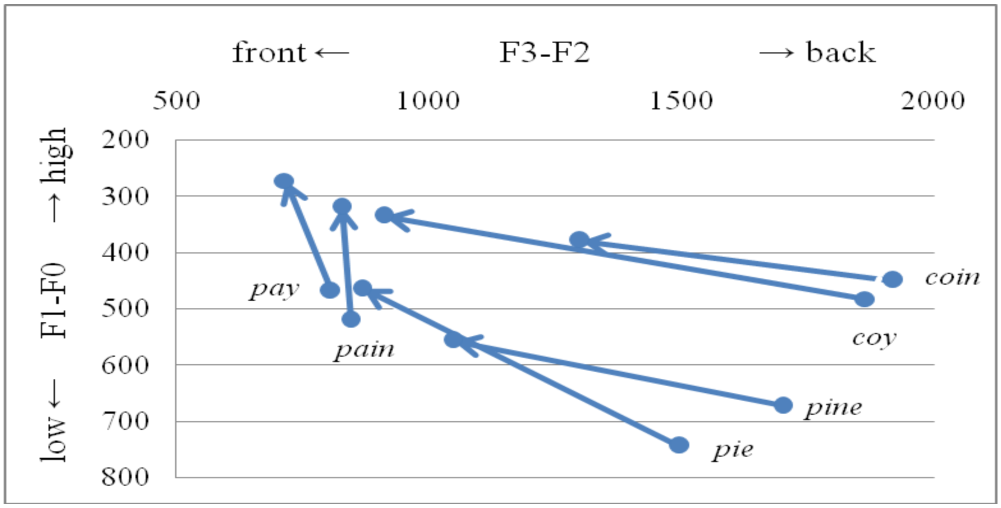


Figure 4-13 shows that the diphthong vowels, /aj, ej, oj/, in *pine/pain/coin* each have a reduced height over the second half of the duration compared to the corresponding diphthong vowels in *pie/pay/coy*. The reduced vowel height is expected because nasal codas have a general lowering effect on the preceding high vowels (or high glides in these cases).

Figure 4-14 Mean F3-F2 and mean F1-F0 over the first and second half of vowel duration for ***gown/cow/gàng/kào*** *(Unit: Hz)*

Mean value	*gown*	*cow*	*gàng*	*kào*
F3-F2_fh	1422	1492	1415	1691
F1-F0_fh	631	727	623	665
F3-F2_sh	1767	1803	1753	1893
F1-F0_sh	587	509	606	526

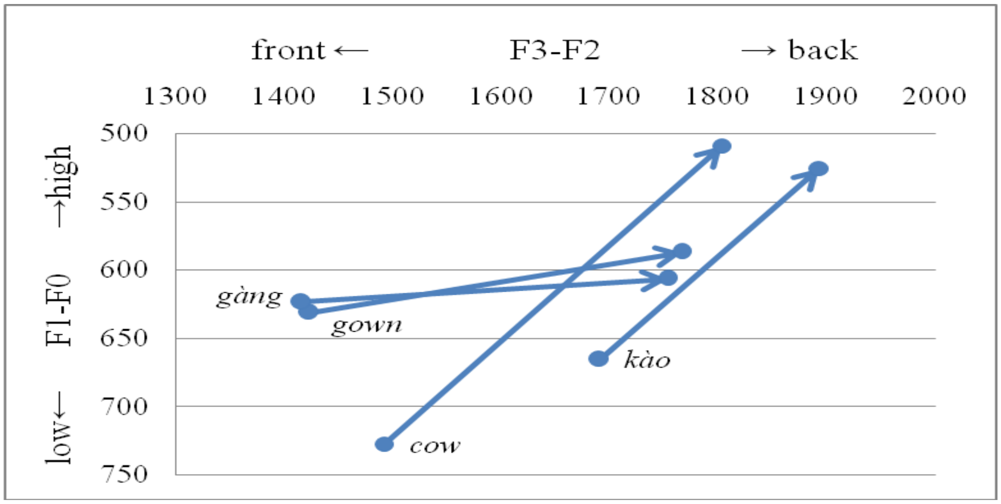


Figure 4-14 shows that the monophthong /a/ in Mandarin *gàng* and the diphthong /aw/ in English *gown* have a comparable movement in terms of both height and backness. Specifically, /aw/ in *gown* is greatly centralized in height; that is, the low vowel /a/ increases its height and the high glide /w/ reduces its height under the influence of /n/. Also, /a/ in Mandarin *gàng* moves to the further back under the influence of /ŋ/, which renders the monophthong /a/ in *gàng* similar to the height-reduced diphthong /aw/ in *gown*. In addition, Mandarin /aw/ in *kào* starts higher and further back than English /aw/ in *cow*, which reflects the same backness constraint.

Figure 4-15 Mean F3-F2 and mean F1-F0 over the first and second half of vowel duration for ***cone/go/gòng/gòu*** *(Unit: Hz)*

Mean value	*cone*	*go*	*gòng*	*gòu*
F3-F2_fh	1861	1595	1846	1718
F1-F0_fh	426	427	367	435
F3-F2_sh	1707	1861	1593	2013
F1-F0_sh	398	296	383	344

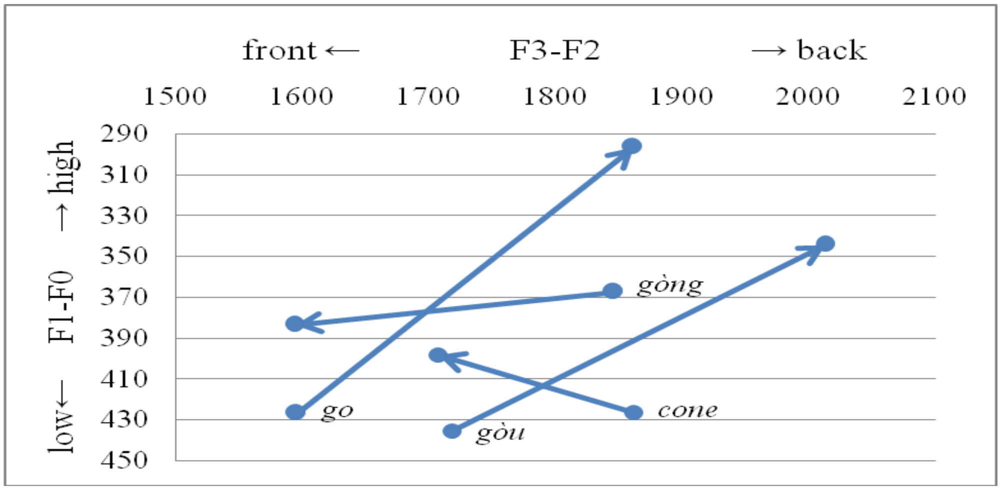


Figure 4-15 shows that Mandarin *gòu* and English *go* both have a comparable upward and backward vowel movement towards the presumed high, back position of the glide /w/, though /ow/ in Mandarin *gòu* is more backed than /ow/ in English *go* due to the same backness constraint. On the other hand, vowels in English *cone* and Mandarin *gòng* move towards the opposite direction, from the back to the front. The forward vowel movement in *cone/gòng* reflects a general centralization effect of nasal codas on the preceding back vowels.

Two 2-tailed paired samples t-tests were carried out for each word across tokens and speakers in order to compare respectively between mean F1-F0_fh and _sh (indicating vowel height change over the duration) and between mean F3-F2_fh and _sh (indicating vowel backness change over the duration). Table 4-1 lists the statistical results for the 10 English words, *pine/pie/coin/coy/gown/cow/pain/pay/cone/go*. The results show that vowel height/backness changes over the duration are significant at the 5% level for all but the 4 words, *gown/pain/pay/cone.* Note that the significant vowel quality change over the duration is expected for a diphthong vowel.

Table 4-1 Statistical results for vowel height/backness changes in the ten English words, ***pine/pie/coin/coy/gown/cow/pain/pay/cone/go***

	F1-F0_fh vs. _sh (vowel height change)		F3-F2_fh vs. _sh (vowel backness change)	
	t_{19}	p	t_{19}	p
pine	-2.984	= .007	5.903	< .001
pie	5.159	< .001	7.970	< .001
coin	2.342	= .030	7.152	< .001
coy	5.573	< .001	10.732	< .001
gown	1.388	= .181*	-6.255	< .001
cow	6.546	< .001	-5.052	< .001
pain	7.779	< .001	.211	= .835*
pay	6.744	< .001	1.561	= .135*
cone	.796	= .436*	1.955	= .065*
go	5.393	< .001	-4.076	< .001

* non-significant.

The non-significant height change for *gown* confirms that the glide /w/ in *gown* does not reach a presumed high position due to the nasal coda influence. Similarly, the non significant backness change for the glide /j/ in *pain/pay* suggests that the glide /j/ does not reach a presumed front position. In addition, both the non-significant vowel height and backness changes in *cone* indicate that the vowel in *cone* is a monophthong.

For the 2 Mandarin words, *gàng/gòng*, the vowel backness change over the duration is significant at the 5% level: $t_{19} = -6.127$, $p < .001$ for *gàng* and $t_{19} = 3.747$, $p < .001$ for *gòng.* The significant vowel backness change in these 2 Mandarin words confirms that velar /ŋ/ has a great influence on the preceding vowel, so that Mandarin VN rimes are subject to the backness assimilation rule.

In order to find out how diphthong vowels differ with and without the nasal coda /n/, four 2-tailed paired samples t-tests were also carried out for each of the 5 pairs of English words, *pine/pie*, *coin/coy, gown/cow, pain/pay, cone/go,* to compare respectively between the mean F1-F0_fh, F3-F2_fh, F1-F0_sh, and F3-F2_sh values of the 2 words in each pair. The results are used to indicate further whether or not there is a significant nasal coda influence on the preceding diphthong vowel. The statistical results are listed in Table 4-2.

Table 4-2 Statistical results for vowel height/backness differences between each of the five pairs of English words, ***pine/pie, coin/coy, gown/cow, pain/pay, cone/go***

	height (F1-F0_fh)	backness (F3-F2_fh)	height (F1-F0_sh)	backness (F3-F2_sh)
pine vs. pie	$t_{19} = -2.354$ $p = .029$	n.s.*	n.s.	n.s.
pine vs. pie	n.s.	$t_{19} = 3.321$ $p = .004$	n.s.	n.s.
pine vs. pie	n.s.	n.s.	n.s.	$t_{19} = 2.358$ $p = .029$
coin vs. coy	n.s.	n.s.	n.s.	$t_{19} = 3.982$ $p = .001$
gown vs. cow	$t_{19} = -2.648$ $p = .016$	n.s.	n.s.	n.s.
pain vs. pay	$t_{19} = 2.44$ $p = .025$	n.s.	n.s.	n.s.
cone vs. go	n.s.	$t_{19} = 3.656$ $p = .002$	n.s.	n.s.
cone vs. go	n.s.	n.s.	$t_{19} = 2.439$ $p = .025$	n.s.

*non-significant.

The statistical results show that F1-F0_fh is significant at the 5% level for the 3 pairs, *pine/pie, gown/cow*, *pain/pay*, that F3-F2_fh is significant at the 5% level for the 2 pairs, *pine/pie and cone/go*, that F1-F0_sh is significant at the 5% level for the pair, *cone/go,* and that F3-F2_sh is significant at the 5% level for the 2 pairs, *pine/pie* and *coin/coy*. The significant difference between the 2 words in these pairs confirms that /aj/ in *pine* starts from a higher position than in *pie*, that /oj/ in *coin* ends with a more backed position than in *coy*, that /aw/ in *gown* starts from a higher position than in *cow*, and that /ej/ in *pain* starts from a lower position than in *pay*. In all of these cases, vowels are centralized in terms of height/backness. Finally, /ow/ in *cone* starts from a more backed position and ends with a lower position than in *go*, confirming that /ow/ in *cone* is a monophthong.

Table 4-1 and 4-2 together show that diphthong vowels in VGn rimes are subject to a significant change in height and/or backness. Also, Mandarin /ŋ/ can cause a significant backward movement for the preceding low, central vowel /a/ in *gàng* and a significant forward movement for the preceding high, back vowel /u/ in *gòng*. The backward moving /a/ in *gàng* happens to resemble the height-reduced /aw/ in *gown*.

4.2.2 Durational measurements

Figure 4-16 illustrates mean V_D (vowel duration), mean N_D (nasal murmur duration), and mean D (the total vowel and nasal murmur duration) for each of the 7 words, *pine/coin/gown/pain/cone/gàng/gòng*, across tokens and speakers. The x-axis represents the duration in second (s), and each bar along the y-axis represents each of the 7 words.

Figure 4-16 Mean V_D, mean N_D, and mean D for ***pine/coin/gown/pain cone/gàng/gòng***

	mean V_D (s)	mean N_D (s)	mean D (s)
pine	0.35	0.18	0.53
coin	0.36	0.18	0.54
gown	0.36	0.21	0.57
pain	0.33	0.20	0.53
cone	0.32	0.21	0.53
gàng	0.23	0.14	0.37
gòng	0.22	0.15	0.37

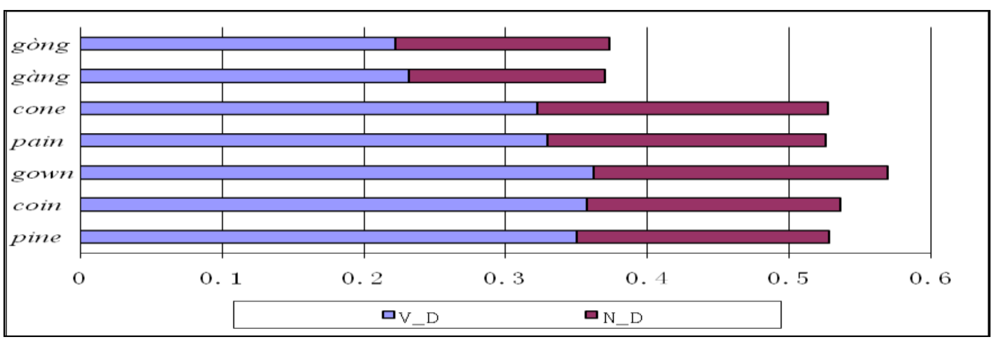


Figure 4-16 shows that the 2 Mandarin words, *gàng/gòng*, have similar duration (the average D = 0.37s) and are shorter than the 5 English words, *pine/coin/gown/pain/cone* (D > 0.5s). The short and almost identical duration for *gàng/gòng* is again expected due to the findings that Mandarin

4th tone words are normally produced with a very short duration and that Mandarin as a syllable-timed language favors syllables with similar length. Note that the 5 English words also have similar duration (the average D = 0.54s) except that *gown* is a little longer than the rest of the English words.

A repeated measures one-way ANOVA test revealed that there is a significant difference among Ds for the 7 words, $F_{6, 114} = 16.772$, $p < .001$, and this is a medium effect size (partial eta-squared = .469). Specifically, the pairwise comparisons in the ANOVA test showed that mean Ds for *gàng/gòng* are significantly smaller than for *pine/coin/gown/pain/cone* at the 5% level: $p < .001$.

Figure 4-17 illustrates mean N_D% (the percentage of the nasal murmur duration over the total vowel and nasal murmur duration) for each of the 7 words, *pine/coin/gown/pain/cone/gàng/gòng*, across tokens and speakers. The x-axis represents N_D/D in percentage (%), and each bar along the y-axis represents each of the 7 words.

Figure 4-17 *Mean N_D% for* ***pine/coin/gown/cone/pain/gàng/gòng***

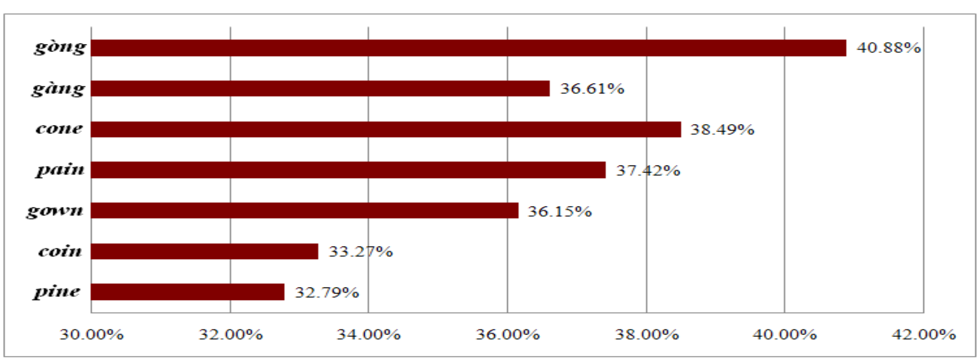


Figure 4-17 shows that *gòng* has the greatest N_D% (40.88%) and *pine/coin* have the smallest N_D% (< 34%) among all the 7 words. The great N_D% for *gòng* is expected because the high vowel /u/ in *gòng* is supposed to have a low degree of nasalization and thus a long period of nasal murmur, let alone the velar nasal /ŋ/ in *gòng* is inherently long. The small N_D% for *pine* may be due to a high degree of vowel nasalization, given that the vowel /a/ in *pine* is a low vowel and the glide /j/ in *pine* is greatly reduced in height; the small N_D% for *coin*, on the other hand, is probably due to the great V_D% in *coin*. It makes sense for the vowel /oj/ in *coin* to be long because /o/ and /j/ contrast so greatly in height/backness/roundness that it takes time for articulators to reach their individual positions.

A repeated measures one-way ANOVA test revealed that there is a significant difference among N_D%s for the 7 words, $F_{6,\ 114} = 2.576$, $p = .022$, though this is a small effect size (partial eta-squared = .119). Specifically, the pairwise comparisons in the ANOVA test showed that mean N_D% for *pine* is significantly smaller than for *pain/cone/gòng* at the 5% level: p = .016, .008, .015, respectively, suggesting that the degree of vowel nasalization in *pine* is higher than in *pain/cone/gòng*.

4.2.3 Nasal measurements

Figure 4-18 illustrates mean N1/N2/N3 for each of the 7 words, *pine/coin/gown/pain/cone/gàng/gòng*, across tokens and speakers. Each dot along the x-axis successively represents each of the 7 words, and the y-axis represents the formant frequency value in Hz.

Figure 4-18 Mean N1/N2/N3 for ***pine/coin/gown/pain/cone/gàng/gòng***

	mean N1 (Hz)	mean N2 (Hz)	mean N3 (Hz)
pine	268	1568	2681
coin	256	1620	2641
gown	274	1389	2648
pain	266	1539	2675
cone	250	1513	2657
gàng	266	1164	2476
gòng	287	1452	2718

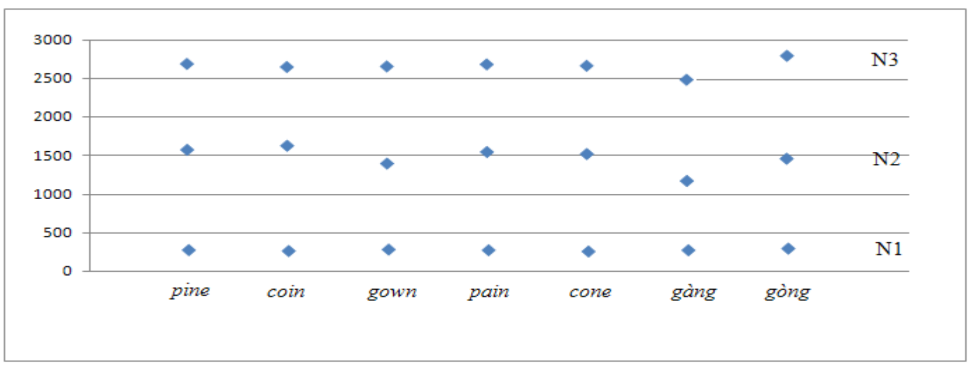


Figure 4-18 shows that mean N1/N2/N3 for the 5 English words, *pine/coin/gown/pain/cone*, are respectively similar except that N2 for *gown* is smaller than for the rest of the English words. The small N2 for *gown* suggests that the nasal place in *gown* is at the further back than in the rest of the English words. In Mandarin words, *gàng/gòng*, mean N1/N2/N3 are all respectively smaller for *gàng* than for *gòng*, suggesting that the nasal place in the 2 words is different.

Two repeated measures one-way ANOVA tests revealed that there is a significant difference respectively among the mean N2 and N3 values of the 7 words, $F_{6, 114} = 5.816$, $p < .001$, and $F_{6, 114} = 2.607$, $p = .042$, respectively, though both are a small effect size (partial eta-squared = .234 and .121, respectively). Specifically, the pairwise comparisons in the ANOVA tests showed that mean N2 for *gàng* is significantly smaller than for *pine/coin/gown/pain/cone/gòng* at the 5% level: $p < .001$, $< .001$, $= .022$, $= .002$, $< .001$, $< .001$, respectively; and that mean N3 for *gàng* is significantly smaller than for *gòng* at the 5% level: $p = .026$, indicating that the nasal place in *gàng* is significantly different from that in *gòng*, even though the 2 Mandarin words share the same velar nasal /ŋ/ phonemically.

Figure 4-19 illustrates mean ΔdB for each of the 7 words, *pine/coin/gown/pain/cone/gàng/gòng*, across tokens and speakers. The y-axis represents the band energy difference in Decibels (dB), and each bar along the x-axis represents each of the 7 words.

Figure 4-19 Mean ΔdB for ***pine/coin/gown/pain/cone/gàng/gòng*** *(Unit: dB)*

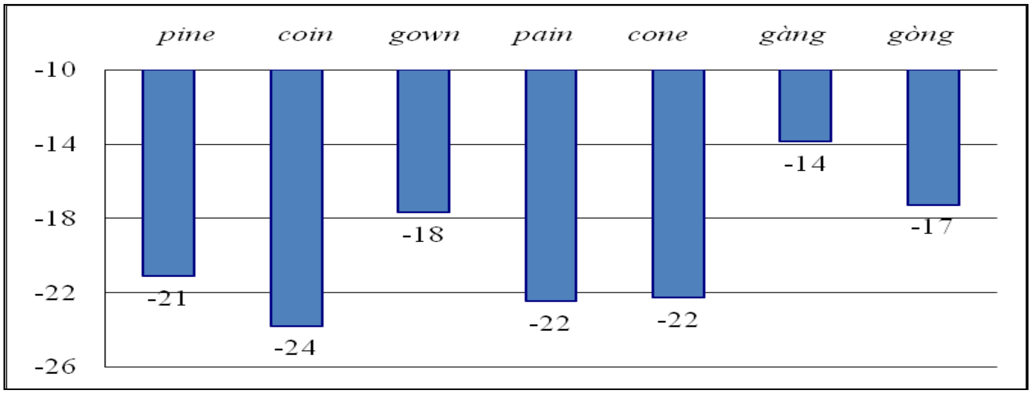


Figure 4-19 shows that mean ΔdB for the Mandarin words, *gàng/gòng*, is smaller than for the English words, *pine/coin/gown/pain/cone*, which is expected because velar /ŋ/ (presumably in the 2 Mandarin words) is subject to less energy loss than alveolar /n/ in the low frequency region. Note that *gown* has the least ΔdB (-18dB) among all the 5 English words and is comparable to that in *gòng* (-17dB), suggesting that the nasal place in *gown* is at the further back than in the rest of the 5 English words.

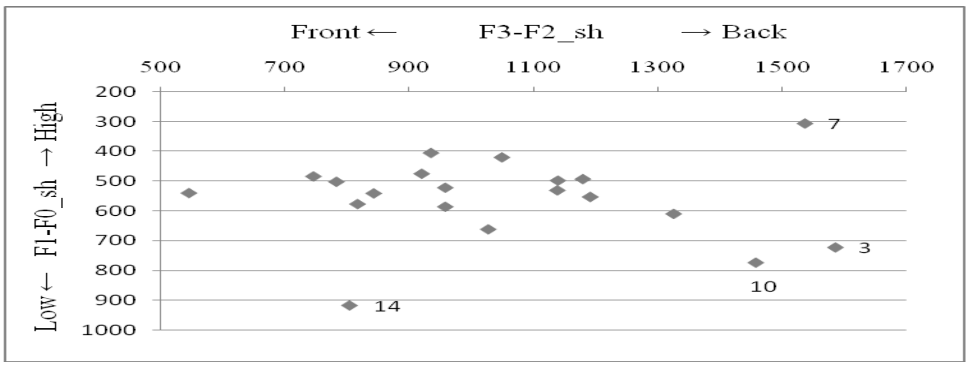
A repeated measures one-way ANOVA test revealed that there is a significant difference among the mean ΔdB of the 7 words, $F_{6, 114} = 14.071$, $p < .001$, and this is a medium effect size (partial eta-squared = .425). Specifically, the pairwise comparisons in the ANOVA test showed that mean ΔdB for *gown* is significantly different from those for *pine/coin/pain/cone/gàng* at the 5% level: p < .001, = .003, < .001, < .001, = .016, respectively, but not from that for *gòng*: $p = .763$, indicating that the nasal place in *gown* can be identified with that in *gòng*, but distinctly different from those in the remaining words. Also, mean ΔdB for *gàng* is significantly lower than for *pine/coin/gown/pain/cone/gòng* at the 5% level: $p < .001$, $< .001$, $= .016$, $< .001$, $< .001$, $= .039$, respectively, indicating that the nasal place in *gàng* is more backed than in the remaining words and may be treated as uvular /ɴ/. A parallel treatment can be found in Yamane-Tanaka's (2008) recent ultrasound study of the moraic nasal place in Japanese. Yamane-Tanaka found that the Japanese moraic nasal is basically a velar but a uvular in the low vowel /a/ context (i.e., /ŋ/ → [ɴ] / a_#).

4.2.4 Production variations for ***pine/coin/gown/pain/cone*** *across speakers*

The above sections have shown the general patterns of L2 VGn production; this section will show production variations across speakers for each of the 5 English words, *pine/coin/gown/pain/cone* (the 4 tokens of each word were averaged for each speaker), in terms of vowel height/backness over the second half of the duration (assuming that the second half is the glide part of a vowel and subject more to the nasal coda influence than the first half) and mean ΔdB (assuming a major indicator of nasal place).

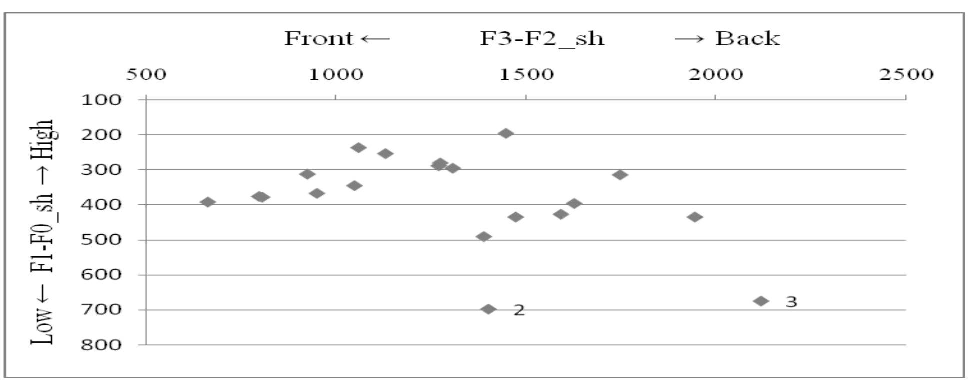
Figures 4-20 to 4-24 show the vowel distribution over the second half of the duration respectively in *pine/coin/gown/pain/cone* across speakers. The x-axis represents mean F3-F2_sh (the greater the F3-F2, the further back the glide); and the y-axis represents mean F1-F0_sh (the greater the F1-F0, the lower the glide). Each dot represents a glide position produced by each speaker (20 dots or speakers in total).

*Figure 4-20 Vowel plots over the second half of the duration for **pine** across speakers (Unt: Hz)*



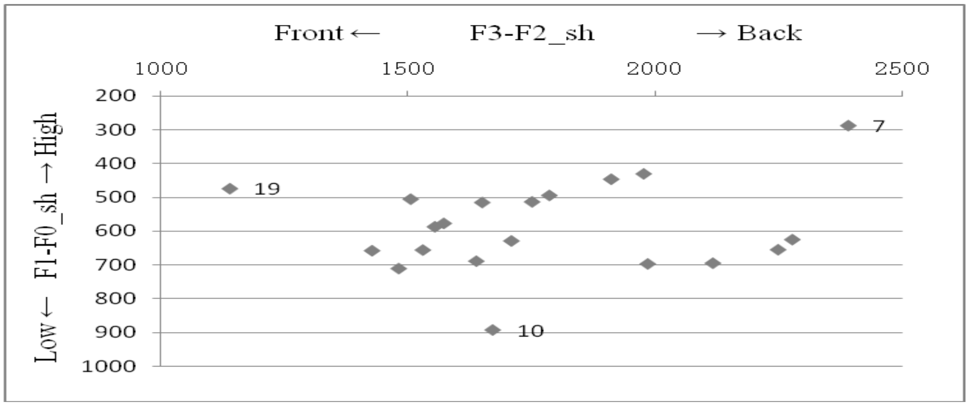
In Figure 4-20, most dots fall within 400-600Hz in F1-F0 and 700-1200Hz in F3-F2. Speaker 14 produced a particularly low glide, and Speaker 7 produced a relatively high, back glide in *pine*.

*Figure 4-21 Vowel plots over the second half of the duration for **coin** across speakers (Unt: Hz)*



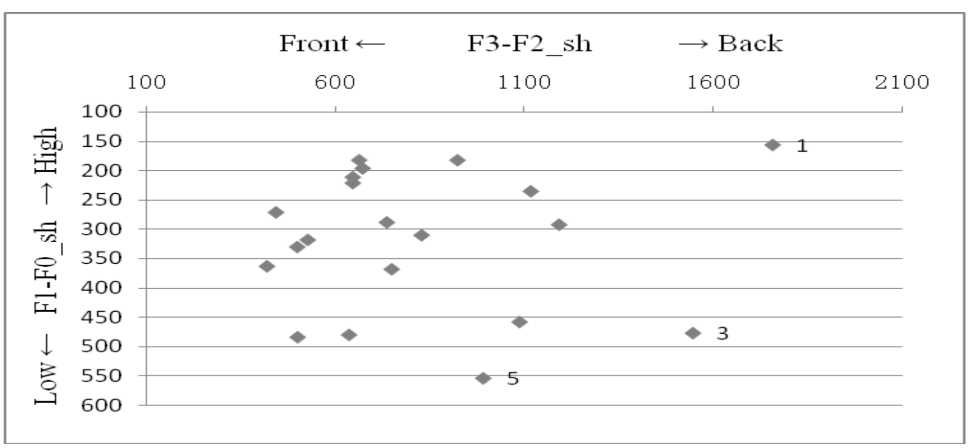
In Figure 4-21, most dots fall within 200-450Hz in F1-F0 and 750-1800Hz in F3-F2. Speaker 2/3 each produced a particularly low glide in *coin*.

Figure 4-22 Vowel plots over the second half of the duration for ***gown*** *across speakers (Unt: Hz)*



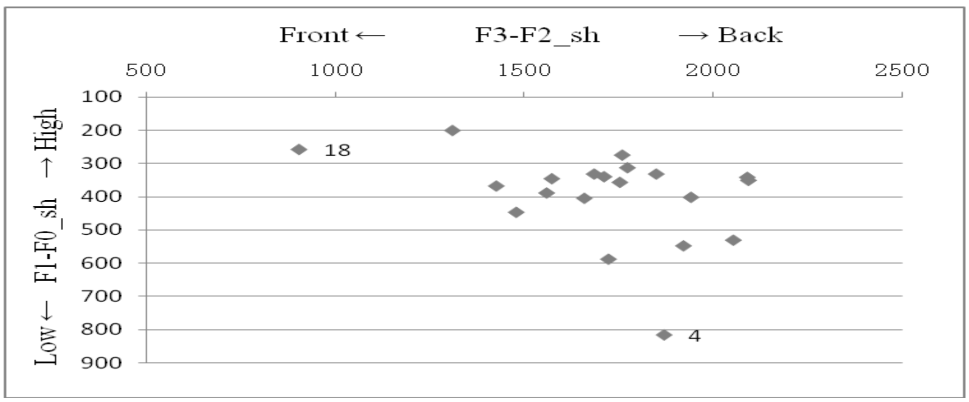
In Figure 4-22, most dots fall within 500-700Hz in F1-F0 and 1400-1800Hz in F3-F2. Speaker 19/10/7 respectively produced a particularly front, low, and back glide in *gown*.

Figure 4-23 Vowel plots over the second half of the duration for ***pain*** *across speakers (Unt: Hz)*



In Figure 4-23, most dots fall within 150-400Hz in F1-F0 and 400-1200Hz in F3-F2. Speaker 1 produced a relatively high, back glide, and Speaker 3 produced a relatively low, back glide in *pain*.

Figure 4-24 Vowel plots over the second half of the duration for ***cone*** *across speakers (Unt: Hz)*

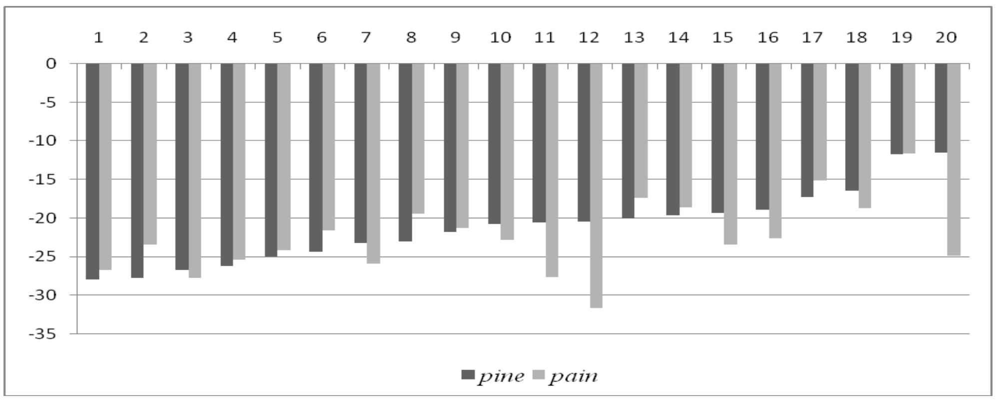


In Figure 4-24, most dots fall within 300-400Hz in F1-F0 and 1500-1800Hz in F3-F2. Speaker 4 produced a particularly low glide, and Speaker 18 produced a particularly front glide in *cone*.

Figure 4-20 – 4-24 together show that in *pine/coin/gown*, dots spread widely along the front-back dimension, and that in *pain/cone*, dots tend to cluster together. The difference in the distributional patterns suggests that vowels in *pain*/*cone* have fewer variations than in *pine/coin/gown*, probably due to the fact that these 2 words have a minimum vowel contrast in height/backness (i.e., they have a monophthong-like vowel).

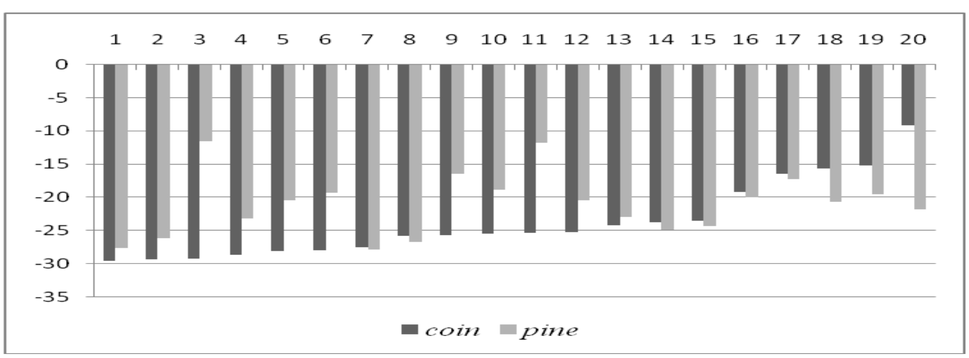
In Figure 4-25, 4-26, and 4-27, the x-axis represents 20 speakers and the y-axis represents mean ΔdB (the lower the ΔdB, the more fronted the nasal place). Dark/grey bars in each figure represent mean ΔdBs for 2 words in contrast.

*Figure 4-25 Mean ΔdB for **pine/pain** across speakers (Unit: dB)*



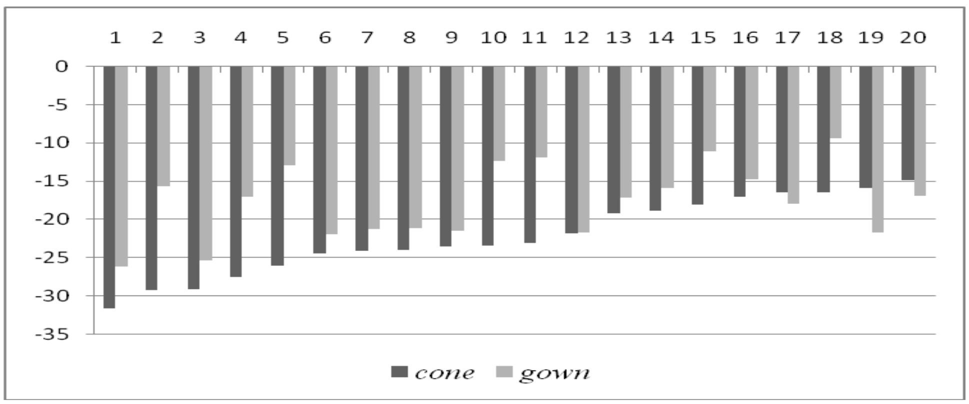
In Figure 4-25, five grey bars (for Speaker 11/12/15/16/20) are noticeably higher than the corresponding dark bars, indicating that in these 5 speakers' production, the nasal place in *pain* is probably less backed than in *pine* because there is a greater ΔdB decrease in *pain* than in *pine*.

*Figure 4-26 Mean ΔdB for **coin/pine** across speakers (Unit: dB)*



In Figure 4-26, three grey bars (for Speaker 18/19/20) are noticeably higher than the corresponding dark bars, indicating that in these 3 speakers' production, the nasal place in *pine* is probably less backed than in *coin* because there is a greater ΔdB decrease in *pine* than in *coin*.

Figure 4-27 Mean ΔdB for ***gown/cone*** *across speakers (Unit: dB)*



In Figure 4-27, two grey bars (for Speaker 19/20) are noticeably higher than the corresponding dark bars, indicating that in these 2 speakers' production, the nasal place in *gown* is probably less backed than in *cone* because there is a greater ΔdB decrease in *gown* than in *cone*.

Figure 4-25, 4-26, and 4-27 together indicate that in most speakers' production, the nasal place in the 5 English words becomes backward in the following order, *coin* -> *pine/pain/cone* -> *gown*, because mean ΔdB mostly falls between -15dB and -30dB for *coin*, between -15dB and -25dB for *pine/pain/cone*, and between -10dB and -25dB for *gown*. The greatest decrease in mean ΔdB for *coin* and the least decrease in mean ΔdB for *gown* suggest that the nasal place in *coin* and *gown* is respectively more fronted and more backed than the remaining words.

Chapter 5

DISCUSSION

5.1 VN production

The acoustic results from the Mandarin participants' production of the English VN and the corresponding Mandarin VN rimes suggest that Mandarin speakers rely on different acoustic cues to distinguish nasal place in their L1 and L2 production. Table 5-1 summarizes the significant cues used to differentiate English *sin/sing* and *son/song* and Mandarin *xìn/xìng* and *sàn/sàng*.

Table 5-1 Significant acoustic cues used to differentiate English ***sin/sing*** *and* ***son/song*** *and Mandarin* ***xìn/xìng*** *and* ***sàn/sàng***

cues	*sin vs. sing*	*xìn vs. xìng*	*son vs. song*	*sàn vs. sàng*
F3-F2/ F1-F0	n.s.*	front vs. central	high- vs. low-mid central vs. back	low-mid vs. low
N2	n.s.	large vs. small	n.s.	large vs. small
ΔdB	n.s.	large vs. small	n.s.	large vs. small
D	n.s.	n.s.	small vs. large	n.s.
N_D%	n.s.	n.s.	large vs. small	n.s.

*not significant.

Table 5-1 shows that English nasal place is distinguished mainly by durational measurements, D and N_D%, rather than nasal place measurements, N2 and

ΔdB. Specifically, D for *son* is significantly smaller than for *song*, and N_D% for *son* is significantly larger than for *song*. The difference in D and N_D% between *son* and *song* is due to the inherent vowel height difference; that is, the high-mid vowel in *son* is shorter and therefore has a lower degree of nasalization and a longer nasal murmur than the low-mid vowel in *song*. Since the difference in D and N_D% is mainly a result of the vowel difference in the two words, the nasal place in *son/song* can be considered to have an extrinsic difference in duration.

As for *sin/sing*, although they are not distinctively produced in terms of any acoustic parameters measured in this study, there are observable differences between *sin* and *sing* in Figure 4-3, 4-4, 4-5, and 4-6: the two durational measurements, D and N_D%, are respectively smaller and larger for *sin* than for *sing*, N2 for *sin* is a little higher than for *sing*, and ΔdB for *sin* is a little larger than for *sing*. The subtle differences in D, N_D%, N2, and ΔdB between *sin* and *sing* suggest that the nasal place in *sin* is less backed than in *sing*. Figure 4-3 also shows that the difference in D between *sin* and *sing* is mainly due to the durational difference of the nasal; that is, /n/ in *sin* is a little shorter than /ŋ/ in *sing*. Since alveolar /n/ is inherently shorter than velar /ŋ/, the nasal place in *sin/sing* can be considered to have an intrinsic difference in duration.

Now the observation that Mandarin speakers often confuse /n/ and /ŋ/ can be explained based on the above results. Since the durational difference in the nasal place, intrinsic or extrinsic, are easily lost in casual settings, without additional cues such as vowel quality change over the duration to enhance the nasal place perception, /n, ŋ/ produced by Mandarin speakers can sound very similar, especially in the high (close) vowel context (e.g., *sin/sing*) where the vowel-nasal coupling effect is inherently weak.

Unlike the English nasal place, the Mandarin nasal place is distinguished by N2 and ΔdB rather than D and N_D%; also, the vowel quality change over the duration helps to cue the nasal place. In the L1 production of *xìn/xìng*, for example, nasal place covaries with vowel height/backness; that is, the more backed the nasal place, the more backed and lower the preceding vowel. The robust nasal place cues in N2 and ΔdB, enhanced by the additional vowel quality cue, render the Mandarin nasal place maximally differentiated along the front-back dimension. Note that the combined vowel quality and nasal place cues in Mandarin /ŋ/ provide a phonetic basis for Hansen's (2001) observation that /ŋ/ in Mandarin is more prominent than in English.

Also recall that in Flege's (1995) study, the English speakers used the combined vowel length and stop voicing contrasts to produce final /t, d/ distinctively, but the Mandarin speakers attended only to the stop voicing contrast. The L2 production of /n, ŋ/ again indicates native speakers' tendency to signal a phonological contrast using multiple phonetic cues or "ensembles of gestures and gestural components that have mutually enhancing auditory effects" (Blicher, Diehl, & Cohen, 1990, p47). The use of combined cues to make a phonological distinction by native speakers but not by L2 speakers also indicates that non-native speakers do not seem to be able to integrate multiple cues into their L2 production.

Last, it is not surprising that duration is not a good cue for distinguishing Mandarin nasal place, given that Mandarin is a syllable-timed tone language. In a syllable-timed language, to stabilize syllable length and duration is crucial to maintain the rhythmic regularity of the language. Also in a tone language, vowel quantity (duration) is closely linked to the tonal difference as well as segmental difference (Lin & Wang, 2008). According to Lin's (2002) tone dominance theory, the short duration of Mandarin VN rimes is determined by the inherent short fourth tone. Therefore, the durational contrast is very likely reserved for rhythmic/tonal purposes in Mandarin. However, the Mandarin participants relied on durational cues in their English nasal place distinction. Why did they not rely on the same cues used in their L1 production to produce L2 nasals? The reason may be attributed to the Mandarin participants' awareness of the rhythmic difference between Mandarin and English.

As shown in the acoustic results, the Mandarin VN rimes tend to have a higher degree of vowel-nasal coarticulation than the English counterparts, which may be related to their relatively fixed and short duration. In order to maintain a duration, for example, more overlap for Vŋ than for Vn is necessary because /ŋ/ is longer than /n/. In English, on the other hand, syllable duration can be varied. As a result, the vowel in Mandarin Vŋ but not in English Vŋ is subject to a strong nasal coda influence and undergoes a significant quality change, and this is actually what happened in the Mandarin Vŋ production. Therefore, the vowel-nasal coarticulation effect in English Vŋ production is not or at least not perceived by Mandarin speakers as strong as in Mandarin. If Mandarin speakers assume that English does not encourage a strong vowel-nasal coarticulation due to flexible syllable duration/length, they will not be likely to use the vowel quality change to cue nasal place in English. Without using the vowel quality changing cue, the speakers can rely only on other cues such as intrinsic duration in the close vowel context and extrinsic duration in the open vowel context to distinguish nasal place in their English production.

As for what is an indicator of a strong vowel-nasal co-articulation effect, opinions differ. Clumeck (1976), for example, considers the early lowering of the velum such an indicator. He claimed that English also exhibits a strong vowel-nasal coarticulation effect because vowel nasalization begins as early as the first vowel starts (i.e., the velum is lowered in advance). However, Solé (1992, 1995) considered the early lowering of the velum not as a result of the strong coarticulation effect but as a phonological process to turn an oral into a nasal vowel when it precedes a nasal. Unlike the two views, this study considers a strong vowel-nasal co-articulation effect as having a high degree of gestural overlap rather than the degree of temporal overlap (e.g., the timing of velum lowering). In other words, a strong co-articulation effect means that a vowel changes drastically in terms of height and/or backness under the nasal coda influence.

5.2 VGn production

The acoustical and statistical results from the Mandarin speakers' production of the English VGn and the corresponding VG rimes reveal that both V and G in a VGn rime can undergo a significant change either in height, in backness, or in both. Generally, V and G in a VGn rime tend to be more centralized than in a VG rime. For example, the low vowel /a/ in *pine* is higher than in *pie*, the high vowel /e/ in *pain* is lower than in *pay*. Also, the high, front glide /j/ in *pine* is lower and more backed than *pie*, and the high, back glide /w/ in *gown* is lower and more fronted than in *cow*.

The above VG differences between VGn and VG rimes suggest that Mandarin speakers' production of English *pine/coin/gown/pain/cone* exhibits a strong vowel-nasal coarticulation effect, similar to that in their Mandarin Vŋ production; however, the similarity in the vowel-nasal coarticulation effect between the English VGn production and the Mandarin Vŋ production does not necessarily suggest L1 transfer, because the controlled 2-second interval between two tokens in the word-reading test may also induce a similar effect. Given the time restriction, for example, long VGn rimes have to be produced in a more condensed way than short VN rimes, resulting in a strong "phonetic process of compression" effect (Hajek, 1997, p.129), and its acoustic consequence is to co-articulate VG and /n/ with more overlap and thus to cause VG undergo a significant quality change, similar to the strong vowel-nasal backness assimilation effect in the Mandarin Vŋ production. The strong coarticulation effect is also manifested in mean D (the total vowel and nasal murmur duration). Note that English VN rimes produced by the Mandarin speakers have different mean Ds depending on the vowel context and the nasal place, but the mean Ds for *pine/coin/gown/pain/cone* are about the same (around 0.54s), and so are for the Mandarin VN rimes (around 0.34s). The durational contrasts between English VN and VGn rimes and between English and Mandarin VN rimes suggest that the production of VGn may have been constrained by the fixed time frame given in the test, just as the Mandarin VN production have been constrained by the fixed tone length.

Also as mentioned in 4.2.3, ΔdB for /n/ in *gown* is significantly smaller than in the rest of the VGn words but comparable to that in Mandarin *gòng*, suggesting that the nasal place in *gown* is at the back rather than at the presumed front (alveolar) region. Note that the assimilation hypothesis mentioned in Section 2.3 concerns only with the influence of a consonantal context on surrounding vowels, but the L2 production of *gown* reveals a reverse impact of the vowel quality change over the duration on the nasal place, indicating that the backness assimilation process between vowels and nasals is possibly more complicated than the assimilation hypothesis has suggested.

If the nasal place of *gown* being at the back to assimilate the back glide /w/ makes sense because of the backness assimilation effect, then the nasal place in *cone* is also expected to be assimilated towards the back as in *gown*, given that /ow/ in *cone* is a high back vowel and therefore subject to the same assimilation effect. However, ΔdB, N2, and D for *cone* all indicate that the nasal place in *cone* is more fronted than in *gown*. To explain the nasal place difference between *cone* and *gown*, it is necessary to examine further the vowel context difference between *cone* and *gown*.

Figure 4-15 has shown that the vowel movement over the duration is from back to front for *cone* but from front to back for *gown*. Two factors may cause the forward vowel movement in *cone*. First, vowel nasalization has a general effect to centralize a vowel so that a back vowel is expected to move forward. Second, the back vowel in *cone* has to move forward to anticipate the front nasal coda /n/. Due to coarticulation, the forward vowel movement will have an inertial (or carry-over) effect on the following nasal. Thus, the nasal place in *cone* is not likely to move to the back to assimilate the back vowel but to follow the vowel movement towards the front. The carry-over effect can be also seen in the production of *gown*. The diphthongal movement from central /a/ to back /w/ carries the nasal /n/ over to the back, so the nasal place in *gown* can not stay at the presumed front as in *cone*, but follows the vowel movement towards the back, even though articulators have to move forward in anticipation for /n/. Since the vowel-nasal assimilation in *cone*/*gown* is relevant to the direction of the vowel movement along the front-back dimension and the associated carry-over effect, it should be viewed as a dynamic phonetic process.

Although the vowel-nasal assimilation process can be captured through a static phonological rule (i.e., the backness assimilation rule) in Mandarin, the L1 production of *gàng*/*gòng* can further show that it is a result of the dynamic interaction between vowels and nasals. As shown in Figure 4-14 and 4-15, *gòng* and *cone* both start with a high, back vowel and move towards the front; *gàng* and *gown* both start with a low, non-back vowel and move towards the back. Despite the inherent nasal place difference, a parallel can be made between *gown* and *gàng* and between *cone* and *gòng*. In the L1 production of *gàng*, for example, the velar nasal place moving towards the further back to the uvular position can now be attributed to the carry-over effect caused by the backward vowel movement, similar to that in *gown*; that is, the nasal place in *gàng* can not just stay at the velar region if the central vowel /a/ keeps moving backward under the nasal coda influence. As a result, both the vowel and the nasal place in *gàng* arrive at a more backed region than their initial position.

In the L1 production of *gòng*, the nasal place is already at the back as anticipated. Even though the carry-over effect caused by the forward vowel movement in *gòng* may bring the nasal place to the further front, its influence does not override the anticipatory effect so that the nasal place in *gòng* is not significantly fronted, just as the nasal place in *cone* is not significantly backed due to the anticipatory effect.

The above discussion suggests that the backness assimilation rule is conditionally applied in both languages; specifically, the condition depends on the relative strength of the anticipatory as opposed to carry-over effect. For example, the carry-over effect is larger than the anticipatory effect in the L2 production of *gown* but smaller in the L1 production of *gòng.* However, why do the anticipatory and carry-over effects differ in the L1 and L2 production? Although the answer is not all that clear at this stage of the study, the connection between rhythm and coarticulation found in previous studies may help to shed lights on this question. For example, Magen (1984) found that in English vowel-to-vowel coarticulation, carry-over effects are larger than anticipatory effects. The vowel-to-vowel articulation refers to a type of coarticulation that centers around vowels (Öhman, 1967), and it is a characteristic of stress-timed languages such as English.

Last, the dynamic view of the vowel-nasal interaction provides an explanation for why Mandarin speakers tend to confuse *gown* with *gàng*. The vowel /a/ and the nasal /ŋ/ in *gàng* move to the far back over the duration due to a strong vowel-nasal assimilation effect. In the production of *gown*, the vowel movement from /a/ to /w/ is towards the back, but /w/ does not reach its presumable height due to the nasal coda influence. The backward vowel movement further causes the nasal place to move backward. Consequently, /aw/ in *gown* has a backward movement trajectory similar to /a/ in *gàng* and the nasal place in *gown* is also at the back similar to that in *gàng* (though is is more backed in *gàng* than in *gown*). The acoustic similarities between *gown* and *gàng* may cause perceptual confusion so that Mandarin speakers' production of *gown* can be heard as *gàng*.

From the acoustic point of view, Mandarin speakers do not intend to confuse *gown* with *gàng* because at least the nasal place is still different in the final production of the two words: [ŋ] in *gown* and possible [ɴ] in *gàng*. It is just that as the coarticulation proceeds, the vowel-nasal interaction causes the vowel and the nasal in the two words to go through different changes but reach a similar acoustic effect at the end. Therefore, the impressionistic explanation that the L2 production of *gown* as *gàng* is due to the L1 syllable structure and the same backness constraints is not accurate, because the glide /w/ in *gown* is not deleted as assumed in the beginning of this study.

The dynamic vowel-nasal coarticulation effect also provides a possible phonetic motivation for the glide-hardening process in Northern Italian described but not explained by Hajek (1997) (see Section 2.2). Since a diphthong vowel preceding alveolar /n/ has an acoustic effect similar to a monophthong vowel preceding velar /ŋ/, the similarity can cause perceptual confusion and eventually lead to the phonological change from a diphthong vowel preceding /n/ to a monophthong vowel preceding /ŋ/. Note that Italian is classified more as a syllable-timed language than as a stress-timed language (Fox, 2000), so it is plausible that Italian has a strong vowel-nasal coarticulation effect similar to that in Mandarin.

5.3 Theoretical implications

The Mandarin speakers' production of the English VN and VGn rimes has an implication for the role of prosody in L2 segmental production; that is, L2 segmental production may be shaped by the high levels of supra-segmental differences between L1 and L2. For example, differences between the L1 and L2 post-vocalic nasal production may be caused by the different degrees of the vowel-nasal coarticulation effect, and the difference in the degree of the co-articulation effect may be further caused by the rhythmic difference between the two languages. Figure 5.1 illustrates such a top-down influence from the supra-segmental to segmental level.

Figure 5-1 Top-down influence on L1 and L2 post-vocalic nasal production

Level of influence	Mandarin	English
Rhythm	Syllable-timed	Stress-timed
↓	↓	↓
Duration/syllable length	Generally fixed	Varied
↓	↓	↓
The co-articulation effect	Strong	Relatively weak
↓	↓	↓
Vowel quality change	Significant	Non-significant
Nasal place distinction	Distinctively contrasted in all vowel contexts	Not distinctively contrasted, especially in the close vowel context

Note that the top-down influence does not involve a direct L1 transfer to L2 at the segmental level. For example, the Mandarin speakers did not use the same nasal place cues used in Mandarin to produce English *sin/sing/son/song;* instead, they seem to have adopted the stress-timed rhythm of English and assumed a weak vowel-nasal co-articulation effect in English. Consequently, their English VN production is not distinctively contrasted in both vowel quality change and nasal place. Therefore, a higher level of supra-segmental difference (e.g., the rhythmic difference) can travel down the phonological hierarchy and indirectly link to the bottom-level of segmental difference.

Previous L2 theories such as the L1 transfer theory, Eckman's (1977) MDH, and Flege's (1995) SLM focus mainly on the segmental comparison between L1 and L2. In these theories, the relationship between L1 and L2 segments is described in terms of the universal markedness, similarities, and differences, because they are considered as major sources of influence on L2 production and acquisition, though James (1988) did mention the possibility of a high level of L1 influence on L2 phonology. This study goes further to postulate that phonological differences between L1 and L2 at the supra-segmental level, particularly, the rhythmic level, may determine differences between L1 and L2 productions at the segmental level.

Note that this postulation not only accord with the view that some top-down control is needed to regulate segmental duration (White & Mattys, 2007), but also has its phonetic basis. According to Fox (2000), prosodic features are specified by subglottal and laryngeal (i.e., the source) settings and segments are modified by supralaryngeal articulators (i.e., the filter), so prosodic features are more basic than segmental features. If prosody comes before segments, then it is not surprising that prosody can shape the timing relationship between adjacent segments. From the phonological point of view, however, prosody is mostly secondary to segments such as stress is an add-on feature of segments. Smith's (1995) study of prosodic patterns in the coordination of vowel and consonant gestures also claims that gestural organization may give rise to timing patterns; that is, prosody may be shaped by gestural timing.

In fact, Shattuck-Hufnagel (2006) succinctly termed the above phonetically based view as "Prosody First" and the opposing phonologically based view as "Prosody Last." Clearly, the L2 post-vocalic nasal production shown in this study supports the Prosody First view. Whether or not the role of prosody is dominant in L2 production still awaits for further research, but one thing is borne out from this study: coarticulation patterns are intimately associated with prosody patterns.

In addition, Aoyama's study of L2 perception discussed in Section 2.2 suggests that "the perceived relationship between L1 and L2 segments plays an important role in how L2 segments are perceived" (2003, p. 263). From the perspective of L2 production, this study suggests that the perceived relationship between L1 and L2 prosodic structures may dictate how L2 segments can be generally produced. Note that most of the Mandarin speakers in this study were intermediate or advanced ESL learners based on the fact that they had to meet certain English proficiency requirement to study at the university level, so they should be familiar with English rhythmic features and perceive Mandarin and English rhythms as different. Moreover, Lin and Wang's (2005) study of Chinese speakers' Mandarin and Canadian English productions confirms that advanced L2 speakers are able to achieve a native-like rhythm in their L2 production.

Another theoretical implication of this study is for the nature of coarticulation. The different vowel-nasal assimilation effects reflected in the L1 production of *gàng/gòng* and L2 production of *gown/cone* suggest that coarticulation is a dynamic process rather than static feature assimilation. If it is understood as merely feature spreading between adjacent segments such as backness assimilation, then the nasal place difference in *gown/cone* and *gàng/gòng* can not be explained, because based on feature assimilation, the nasal place in the four words would be all at the back, similar to one another. Therefore, this study delimits some general claims concerning the co-articulation effect such as the assimilation hypothesis and the Mandarin backness assimilation rule, suggesting that the coarticulation effect between adjacent segments is a dynamic process, so that how a feature is changed during the process is as important as which feature is changed.

To sum up, if the high level of prosodic difference between L1 and L2 can be considered to be a fundamental factor in shaping L2 post-vocalic nasal production in general, then the dynamic vowel-nasal coarticulation process can be considered to have a fine-tuning effect on L2 post-vocalic nasal production in particular.

Chapter 6

CONCLUSION

6.1 Answers to the research questions and evaluation of the research hypotheses

This study has considered how Mandarin speakers produce the two English nasal codas /n, ŋ/ in the preceding open versus close and monophthongal versus diphthongal vowel contexts. Also, the L1 production of post-vocalic nasals in similar vowel contexts has been used to contrast with the L2 production. The acoustic results have provided answers to the 3 research questions:

1) *How do vowel context and nasal place interact respectively in L1 and L2 production?*

In the L1 VN production, the nasal place tends to co-vary with the backness of the preceding vowel, whereas in the L2 production, the nasal place tends to covary with the syllable duration. Consequently, Mandarin /n, ŋ/ are more distinctively differentiated than English /n, ŋ/, which supports the first hypothesis that the actual nasal place in Mandarin speakers' production of English and Mandarin velar /ŋ/ is different.

2) *Can systematic similarities and differences be identified in the L1 and L2 production? If yes, what linguistic factors may come into play?*

Yes. Both the English VGn and Mandarin Vŋ production exhibit a significant vowel quality change. Also, the vowel quality change has a reverse effect on the nasal place change. However, the vowel quality change and the associated nasal place change are not evident in the English Vŋ production.

The similarity between the L1 Vŋ and L2 VGn production can be generally attributed to a strong vowel-nasal coarticulation effect. The strong vowel-nasal coarticulation effect is further related to rhythmic factors. For a Mandarin Vŋ rime, duration is relatively fixed for a given tone, so the vowel-nasal coarticulation effect is mainly reflected by the vowel quality (i.e., height/backness) change as opposed to quantity (i.e., duration) change. As for an English VGn rime, since its duration is relatively longer than a VN rime, the controlled 2-second interval condition in the word-reading test may have induced a strong co-articulation effect similar to that in the Mandarin Vŋ production.

The difference between the L1 and L2 Vŋ production, on the other hand, can be attributed to the different degrees of the vowel-nasal coarticulation effect. While the strong vowel-nasal coarticulation effect in the L1 Vŋ production is related to the syllable-timed nature of Mandarin, the weak vowel-nasal coarticulation effect in the L2 Vŋ production is related to the stress-timed nature of English, since English syllables can vary greatly in duration depending on the number of stresses and their locations. Although the English test words with the VN type of rime are all monosyllabic and the stress effect is not obvious, the length distinction among them is still evident enough to indicate that no strong coarticulation effect has occurred in the L2 production of these words.

The similarity and difference found in the L1 and L2 post-vocalic nasal production also supports the second hypothesis that English post-vocalic nasal production by Mandarin speakers is related to supra-segmental factors, or specifically, rhythmic factors.

3) *Is there any acoustic evidence that can be used to qualify/quantify previous phonological claims such as the backness assimilation rule?*

Yes. The backness assimilation effect in the L1 and L2 production is found to be a result of the dynamic interaction between vowels and nasals. Since the production of English *gown/cone* and Mandarin *gàng/gòng* reveals that the nasal place is related to but not correlated with the backness of the vowel itself, the backness assimilation rule is no longer accurate and can be refined, for example, as "nasal place depends on both vowel backness and the vowel backness change over the duration." Also, the third hypothesis that nasal place co-varies with vowel backness in Mandarin speakers' production of both

Mandarin and English post-vocalic nasals needs to be revised as "nasal place co-varies with vowel backness in Mandarin speakers’ production of Mandarin but not English post-vocalic nasals."

6.2 Limitations of the study

The major limitation of this study involves its methodology. First, since this study used standard computer programs (e.g., Praat scripts) to measure acoustic parameters automatically, some special speech data may not have been measured accurately. In a few cases, speakers' production of a word token was rather creaky at the end of articulation so that the related acoustic results looked abnormal and had to be manually adjusted based on the spectral display of the token. Second, the choice of the participants was limited, because it is hard to find a homogeneous group of Mandarin speakers in a university setting. When recruiting participants, this study generally looked for speakers with a similar English proficiency level and language background; however, some ESL students were still included and they may deviate from regular university students in their English production because of their relatively low English proficiency level. In addition, most of the participants speak different Mandarin dialects as well as Standard Mandarin, so their dialectal difference which has not been identified by this study may have been reflected in their English production. Last, the choice of the test words is also limited. For example, some contrastive word pairs such as *sin*/*xìn* and *cow/gown* do not have the identical onset. The voicing contrast in the syllable onset of *cow/gown*, for example, may have a different influence on the following vowel production. Thus, some of the results in this study may need to take the onset influence into account.

6.3 Future directions

The acoustic approach to L1 and L2 post-vocalic nasal production allows this study to describe the dynamic coarticulation process between nasal codas and the preceding vowels accurately and to explain the L2 production patterns adequately. Given the effectiveness of the acoustic approach, a possible direction for future research is to use a similar approach to study other L2 segmental production, especially those segments which are easily confused by L2 speakers, such as English *r/l* produced by Japanese speakers.

If future research continues to explore post-vocalic nasal production with a similar acoustic approach, then it can be conducted on a larger scale such as using more test words with various contexts, recruiting more participants, and adopting more advanced technical means. Better yet, it can compare English speakers' production of Mandarin and English post-vocalic nasals, in contrast with Mandarin speakers' production of English and Mandarin post-vocalic nasals.

Another possible future direction is to study L2 prosody related to nasals. Nasals as sonorant consonants may have a special status in L2 prosody. A text full of nasal sounds, for example, may have very different prosodic structure from a text with fewer nasals; a 4-second time interval instead of the 2-second time interval in the test word reading may change the vowel-nasal coarticulation pattern in the L2 VGn production. Also, Mandarin speakers' production of Japanese nasals may be different from their production of English nasals because Japanese is considered to be a mora-timed language and different from both Mandarin and English. Therefore, various acoustically based rhythmic studies can be carried out to help our understanding of the prosodic nature of nasal sounds.

To sum up, this study provides a systematic account of how post-vocalic /n, ŋ/ are differentiated both in L1 and L2 production and in different vowel contexts. Consequently, it is able to provide a relatively satisfactory explanation for the difference between the L1 and L2 production by appealing to prosody; that is, L2 post-vocalic nasal production is ultimately shaped by the speakers' interpretation of the rhythmic difference between L1 and L2. Also, by adopting the acoustically based research methodology, this study is able to discover the dynamic nature of vowel-nasal coarticulation and contributes to the refining of broad theoretical claims such as the backness assimilation rule.

References

Aoyama, K. (2003). Perception of syllable-initial and syllable-final nasals in English by Korean and Japanese speakers. *Second Language Research, 19 (3),* 251-265.

Ainsworth, W.A. (1981). Duration as a factor in the recognition of synthetic vowels. *Journal of Phonetics, 9,* 333-342.

Beddor, P.S. (1991). Predicting the structure of phonological systems. *Phonetica, 48*, 83-107.

Beddor, P.S. , Brasher, A., & Narayan, C. (2007). Applying perceptual methods to the study of phonetic variation and sound change. In M-J. Solé, P.S. Beddor, & M. Ohala (Eds.), *Experimental approaches to phonology* (pp. 127-143). New York: Oxford University Press.

Bladon, R. A.W., & Nolan, F. (1977). A video-fluorographic investigation of tip and blade alveolars in English. *Journal of Phonetics*, *5,* 185-193.

Blicher, D. L., Diehl, R. L., & Cohen, L. B. (1990). Effects of syllable duration on the perception of the Mandarin Tone 2/Tone 3 distinction: Evidence of auditory enhancement. *Journal of Phonetics, 18(1)*, 37-50.

Busà, M. G. (2007). Coarticulatory nasalization and phonological developments: Data from Italian and English nasal-fricative sequences. In

M-J. Solé, P.S. Beddor, & M. Ohala (Eds.), *Experimental approaches to phonology* (pp. 155-174). New York: Oxford University Press.

Cao, W. (2000). *Hanyu fayin yu jiuyin* (*Chinese pronunciation and correction*). Beijing: Beijing Daxue Chubanshe (Beijing University Press).

Chafcouloff, M., & Marchal, A. (1999). Velopharyngeal coarticulation. In W. J. Hardcastle, & N. Hewlett (Eds.), *Coarticulation: Theory, data and techniques* (pp. 69-79). Cambridge: Cambridge University Press.

Chen, M. (1972). *Nasals and nasalization in Chinese: Explorations in phonological universals*, Ph.D. dissertation, University of California, Berkeley.

Chen, M. (1975). An aerial study of nasalization in Chinese. *Journal of Chinese Linguistics*. 3, 16-59.

Chen, M. Y. (1997). Acoustic correlates of English and French nasalized vowels. *Journal of the Acoustical Society of America,102 (4),* 2360-2370.

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

Chomsky, N., & Halle, M. (1968). *The sound pattern of English*. New York: Harper & Row.

Clumeck, H. (1976). Patterns of soft palate movements in six languages. *Journals of Phonetics*, *4 (4)*, 337-351.

Esling, J. H. (2005). There are no back vowels: The laryngeal articulator model. *Canadian Journal of Linguistics, 50 (1)*, 13-44.

Espy-Wilson, C.Y. (1994). A feature-based semivowel recognition system. *Journal of the Acoustic Society of America*, *96 (1)*, 65-72.

Fant, G. (2004). *Speech acoustics and phonetics: Selected writings*. Dordrecht, The Netherlands: Kluwer Academic Publishers.

Flege, J. E. (1995). Second language speech learning theory, findings, and problems. In W. Strange (Ed.). *Speech perception and linguistic experience: Issues in cross-language research* (pp. 233-277). Baltimore: York Press.

Fox, A. (2000). *Prosodic features and prosodic structure: The phonology of suprasegmentals*. New York: Oxford University Press.

Goldsmith, J. A. (1990). *Autosegmental and metrical phonology*. Oxford: Blackwell.

Greenberg, J. H. (2005). *Language universals: With special reference to feature hierarchies*. New York: Mouton de Gruyter.

Hajek, J. (1997). *Universals of sound change in nasalization*. Boston: Blackwell Publishers.

Hansen, J. G. (2001). Linguistic constraints on the acquisition of English syllable codas by native speakers of Mandarin Chinese. *Applied Linguistics,* 22 *(3)*, 338-365.

Hansen, J. G. (2006). *Acquiring a non-native phonology: Linguistic constraints and social barriers*. New York: Continuum.

Harrington, J. (1994). The contribution of the murmur and vowel to the place of articulation distinction in nasal consonants. *Journal of the Acoustic Society of America, 96 (1)*, 19-32.

Harrington, J., & Cassidy, S. (1999). *Techniques in speech acoustics*. Dordrecht : Boston: Kluwer Academic Publishers.

Hayward, K. (2000). *Experimental phonetics*. Harlow, England: Longman.

Ho, A. T. (1976). The acoustic variation of Mandarin tones. *Phonetica, 33,* 353-367.

James, A. (1988). *The acquisition of a second language phonology : A linguistic theory of developing sound structures*. Tübingen: G. Narr.

Jha, S. K. (1986). The nasal vowels in Maithili: An acoustic study. *Journal of Phonetics*, *14 (2)*, 223-230.

Kingston, J. (1991). Integrating articulations in the perception of vowel height. *Phonetica, 48*, 149-179.

Kurowski, K., & Blumstein, S. (1987). Acoustic properties for place of articulation in nasal consonants. *Journal of the Acoustic Society of America, 81(6)*, 1917-1927.

Kühnert, B., & Nolan, F. (1999). The origin of coarticulation. In W. J. Hardcastle, & N. Hewlett (Eds.), *Coarticulation: Theory, data and techniques* (pp. 7-30). Cambridge: Cambridge University Press.

Ladefoged, P., & Maddieson, I. (1996). *The sounds of the world's languages*. Cambridge: Blackwell.

Ladefoged, P. (2001). *Vowels and consonants: An introduction to the sounds of languages*. Malden, MA: Blackwell.

Lee, W-S., & Zee, E. (2003). Standard Chinese (Beijing). *Journal of the International Phonetic Association*, *33*, 109-112.

Lehiste, I. (1976). Suprasegmental features of speech. In N.J. Lass (Ed.), *Contemporary issues in experimental phonetics* (pp. 225-242). New York: Academic Press.

Lin, H. (2001a). *A grammar of Mandarin Chinese*. Muenchen: Lincom Europa.

Lin, H. (2001b). Stress and the distribution of the neutral tone in Mandarin. In D.B. Xu (Ed.). *Chinese Phonology in Generative Grammar* (pp. 139-161). San Diego: Academic Press.

Lin, H. (2002). Hanyu shengdiao de zhudao diwei. [Tone dominance in Mandarin Chinese]. Paper presented at the Conference in Celebration of the 50-Year Anniversary of the Journal of Chinese Linguistics (Zhongguo Yuwen). Nanchang, China, June 27-July 1, 2002.

Lin, H., & Wang, Q. (2005). Syllable-timedness in the ESL production of Mandarin Chinese speakers. Paper presented at the Conference Between Stress and Tone, Leiden, the Netherlands. June 2005.

Lin, H., & Wang, Q. (2008). Mandarin rhythm: An acoustic study. *Journal of Chinese Linguistics and Computing*, *17(3)*, 127-140.

Lin, Y. (2007). *The sounds of Chinese*. New York: Cambridge University Press.

Magen, H. S. (1984). The extent of vowel-to-vowel coarticulation in English. *Journal of Phonetics*, *25 (2)*, 187-206.

Manuel, S. (1999). Cross-language studies: Relating language-particular coarticulation patterns to other language-particular facts. In W. J. Hardcastle, & N. Hewlett (Eds.), *Coarticulation: Theory, data and techniques* (pp. 179-198). Cambridge: Cambridge University Press.

Moll, K., & Daniloff, R. (1971). Investigation of the timing of velar movements during speech. *Journal of the Acoustic Society of America, 50 (1)*, 678- 684.

O'Grady, W., & Archibald, J. (Eds.). (2000). *Contemporary linguistic analysis: An introduction.* Toronto: Pearson Education Canada.

Öhman, S. (1967). Coarticulation in VCV sequences. *Journal of the Acoustic Society of America, 39,* 151-168.

Pruthi, T., & Espy-Wilson, C. (2007). Acoustic parameters for the automatic detection of vowel nasalization. Proceedings of Interspeech 2007, Antwerp, Belgium, Aug 27-31. 1925-1928.

Qi, Y., & Fox, R. A. (1992). Analysis of nasal consonants using perceptual linear prediction. *Journal of the Acoustic Society of America, 91 (3),* 1718-1726.

Recasens, D. (1983). Place cues for nasal consonants with special reference to Catalan. *Journal of the Acoustic Society of America, 73 (4),* 1346-1353.

Recasens, D. (1988). Phonology, and speech production and perception: the case of nasal stops. In W.U. Dressler, H.C. Luschützky, O.E. Pfeiffer, & J.R. Rennison (eds.). *Phonologica 1988: Proceedings of the 6th international phonology meeting*. Cambridge: Cambridge University Press, 229-238.

Recasens, D. (1997). A model of lingual coarticulation based on articulatory constraints. *Journal of the Acoustic Society of America, 102 (1),* 544-561.

Remijisen, B. (2004). Praat scripts. Retrieved December, 2007, from http://www.ling.ed.ac.uk/~bert/praatscripts

Rosner, B. S., & Pickering, J. B. (1994). *Vowel perception and production*. Oxford: Oxford University Press.

Seitz, P.F., McCormick, M.M., Watson, I.M.C. & Bladon, R.A. (1990). Relational spectral features for place of articulation in nasal consonants. *Journal of the Acoustic Society of America, 87 (1),* 351-358.

Shattuck-Hufnagel, S. (2006). Prosody first or prosody last? Evidence from the phonetics of word-final /t/ in American English. Goldstein, L., Whalen, D.H., & Best, C. (Eds.). *Laboratory phonology 8*. 445-472.

Smith, C. L. (1995). Prosodic patterns in the coordination of vowel and consonant gestures. Connell, B., & Arvaniti, A. (Eds.). *Phonology and phonetic evidence: Papers in Laboratory phonology IV*. 205-222.

Solé, M.J. (1992). Phonetic and phonological processes: the cases of nasalization. *Language and Speech, 35,* 29-43.

Solé, M.J. (1995). Spatio-temporal patterns of velopharyngeal action in phonetic and phonologically nasalization. *Language and Speech, 38 (1),* 1-23.

Stevens, K. N. (1989). On the quantal nature of speech. *Journal of Phonetics, 17*, 3-45.

Sussman, H. (1990). Acoustic correlates of the front/back vowel distinction: A comparison of transition onset versus "steady state". *Journal of the Acoustic Society of America, 88 (1),* 87-96.

Van Son, R.J.J.H., & Pols, L.C.W. (1990). Formant frequencies of Dutch vowels in a text, read at normal and fast rate. *Journal of the Acoustic Society of America, 88 (4)*, 1683-1693.

White, L., & Mattys, S. L. (2007). Calibrating rhythm: First language and second language studies. *Journal of Phonetics*, *35 (4)*, 501-522.

Xu, S. (1999). *Pu tong hua yu yin chang shi (Common Standard Chinese pronunciation)*. Beijing: Yu wen chu ban she (Chinese Language Press).

Yamane-Tanaka, N. (2008). Place of moraic nasal in Japanese. Poster presentation at the Canadian Linguistic Association 2008 conference, Vancouver, BC, May 31- June 2, 2008.

Zee, E. (1981). Effect of vowel quality on perception of post-vocalic nasal consonants in noise. *Journal of Phonetics, 9 (1)*, 35-48.

Zee, E. (1985). Sound change in syllable-final nasal consonants in Chinese. *Journal of Chinese Linguistics*, *13*, 291-330.

Appendix 1 A sample questionnaire

Thank you for participating in my language survey about Mandarin speakers' English nasal production. There is no right or wrong answer. The data will be used to provide background information for my study. No identification information will be asked in this survey. The consent form will be detached from this survey.

Gender______ School year________ Major________

Age group: 25 and below, 25-30, 30-40, 40 and above.

1. Which part of China are you from and how many years have you stayed there?

2. Did you stay in another part of China more than one year? If yes, please specify the location and the length of your stay.

3. Do you speak a Chinese dialect other than Mandarin? If yes, please specify the dialect you speak.

4. How long have you stayed in Canada?

5. How many years of formal English education have you received in China?

6. Have you taken any ESL class outside China? If so, where and how long have you taken?

Appendix 2 A summary of the participants' background information

No.	Age	Region: Province	LOESL_China	LOESL	LOR_Canada
F1	< 25	N: Beijing	6	2	8
F2	25-30	N: Hebei	10	0	0.5
F3	< 25	N: Shandong	8	0.5	1
F4	< 25	N: Henan	6	0.75	0.75
F5	25-30	N: Shandong	17	0	7
F6	25-30	N: Beijing	11	0	2.5
F7	< 25	NE: Jilin	3	0	0.25
F8	< 25	N: Beijing	10	0	1.5
F9	< 25	N: Hebei	7	0.8	0.9
F10	< 25	NE: Heilongjiang	6	1	5
M1	30-40	N: Beijing	15	0	0.8
M2	25-30	N: Inner Mongolia	15	0	0.8
M3	30-40	N: Beijing	10	0	2
M4	25-30	N: Shandong	12	0.5	2
M5	< 25	N: Shandong	9	0.25	0.25
M6	25-30	NE: Liaoning	7	0.7	5
M7	25-30	N: Beijing	15	0	0.4
M8	25-30	N: Shandong	10	0	3
M9	< 25	N: Beijing	8	2	6
M10	25-30	N: Shijiazhuang	10	0	0.5

Note:

Column 1: F: Female, M: Male.

Column 3: N: North China, NE: Northeast China.

Column 4: LOESL_China (year): Length of formal ESL education received in China.

Column 5: LOESL(year): Length of formal English education received outside China.

Column 6: LOR (year): Length of Residence in Canada.

Appendix 3 A description of the Praat scripts

Script #	Description	Source of adaptation
1	It reads .wav and .TextGrid files in a directory all at once into the Praat window.	Remijsen, 2004
2	It extracts tokens from the initial .wav file of each speaker.	Remijsen, 2004
3	It creates a preliminary .TextGrid file for each token based on the intensity contour.	By the researcher
4	It measures V_D and N_D and calculates D and N_D%.	Remijsen, 2004
5	It measures mean F0/F1/F2/F3_fh and _sh and calculates mean F1-F0_fh and _sh and mean F3-F2_fh and _sh.	Remijsen, 2004
6	It measures N1/N2/N3 at the mid point of the nasal murmur duration.	Remijsen, 2004
7	It measures ΔdB between 0-525Hz and 525-1265Hz bands over the nasal murmur duration.	By the researcher

Appendix 4　A sample word-list

Reading task 1

Token#	**Words**	Rime with the underlined
1 2 3 4	**pie** (馅饼)	lie (撒谎)
1 2 3 4	**pine** (松树)	wine (酒)
1 2 3 4	**cow** (母牛)	now (现在)
1 2 3 4	**gown** (长袍)	down (下)
1 2 3 4	**coy** (隐蔽的)	toy (玩具)
1 2 3 4	**coin** (硬币)	
1 2 3 4	**sin** (罪恶)	win (赢)
1 2 3 4	**sing** (唱歌)	
1 2 3 4	**son** (儿子)	
1 2 3 4	**song** (歌曲)	
1 2 3 4	**go** (走)	
1 2 3 4	**cone** (蛋筒，圆锥形)	phone (电话)
1 2 3 4	**pay** (付钱)	
1 2 3 4	**pain** (痛苦)	

Reading task 2

Token#	**单词 (words)**	汉语拼音 (Chinese Pinyin)
1 2 3 4	共	gòng
1 2 3 4	杠	gàng
1 2 3 4	靠	kào
1 2 3 4	够	gòu
1 2 3 4	信	xìn
1 2 3 4	兴	xìng
1 2 3 4	散	sàn
1 2 3 4	丧	sàng