Acoustic Correlates of Mandarin Nasal Codas and Their Contribution to Perceptual Saliency[*]

Yi-hsiu Lai

National University of Kaohsiung

Two experiments were designed to investigate two issues: (a) acoustic correlates in the place of articulation in Mandarin syllable-final nasals, and (b) the contribution of acoustic cues to perceptual saliency for the alveolar-velar nasal contrast. In Experiment 1, an acoustic analysis identified phonetic features adopted by L1 and L2 speakers of Mandarin in Taiwan to distinguish the [n]-[ŋ] pair. Results revealed that spectral differences, found in Mainland Mandarin, were not significant in Taiwan Mandarin, which displayed certain nasal mergers. In Experiment 2, recordings of these speakers were rated by experienced teachers of Mandarin who maintained clear contrasts in nasal codas. Multiple regression analyses on acoustic measures and the perceptual rating indicated that extent of nasalization was the primary factor affecting native listeners' perception of the [n]-[ŋ] contrast.

Keywords: Mandarin Chinese nasal coda, formant transition, nasalization

1. Introduction

Over the past decades, segmental disorders have gradually attracted attention and gained much emphasis in the field of clinical phonetics. A large number of studies have established the important role of vowels in intelligibility of communication (Chen, Robb, Gilbert, and Lerman 2001, Flege, Bohn, and Jang 1997, Luo 2002, Ma 1995, Wang 1997). Little attention has been directed to phonetic inaccuracies of syllable-final nasals. A few acoustic reports of English nasals assessed the relative importance of murmurs and transitions as nasal place cues, but controversy has remained (Chen 1991b, Hajek 1997, Li 1999). Some studies indicated that nasal place of articulation was cued primarily by transitions (Malécot 1956, Recasen 1983). Others (Kurowski and Blumstein 1984, Repp 1986) showed that murmur made a significant contribution to place of articulation. A current perspective on acoustic correlates to nasal consonants is that both murmur and transitions together determine the nasal place of articulation (Harrington 1994, Mou 2006, Ohde 1994). Variable formant transitions, F2 in particular, when combined with murmurs, have been claimed most optimal for different places of articulation in English nasals.

Codas in the Mandarin syllable structure only allow nasals [n] and [ŋ]. Despite this simple structure, universal instability of nasal codas has also been found in

---

Mandarin Chinese spoken in Taiwan. Speakers of Mandarin in Taiwan are specified with an accent of Taiwan Mandarin,[1] genetically related to Putonghua, spoken in Beijing in Mainland China[2] (Lin and Yan 1991, Mou 2006). In Taiwan, the syllable-final distinction is frequently dropped (Hsu and Tse 2007, Tse 1992, Yueh 1992). In terms of merging directions, bifurcated conclusions have been made. Some researchers (Kubler 1985, Tse 1992) argued that alveolarization was the predominant trend of syllable-final nasal merger in [i] and [ə]. Others (Chen 1991a, Hsu and Tse 2007, Ing 1985) claimed that the syllable-final nasal following [i] was more likely to be velarized. These reports were somewhat limited partly because detailed discussion on acoustic information in the vowel region and in the murmur region was lacking, and partly because the possible relationship between acoustic correlates and perceptual saliency was neglected.

Inspired by previous literature, the present researcher aims to address two issues in the current study. The first is to examine to what extent murmur and vowel make a contribution to the nasal place of articulation distinction for L1 and L2 speakers of Mandarin in Taiwan. Whether these speakers follow the acoustic patterns of English nasals (cf. Harrington 1994, Kurowski and Blumstein 1984, Mou 2006, Ohde 1994, Repp 1986), or those of Mainland Mandarin (cf. Lin and Yan 1991, Mou 2006), is also discussed. The second issue is to investigate perceptual saliency for the alveolar-velar nasal distinction and then to correlate the acoustic cues adopted by speakers with the perceptual judgment rated by listeners. To what extent the syllable-final distinction is dropped and in which direction Mandarin nasal codas merge are examined in terms of acoustic measures and of perceptual saliency. In sum, this research investigates acoustic phonetics of Mandarin syllable-final nasal disorders and contributes to theories of Mandarin syllable-final nasal production and perception saliency.

2. Literature review

The present research intends to examine acoustic cues in syllable-final nasal contrasts and to interpret the possible merging of nasal codas from both acoustic and perceptual dimensions. Previous studies concerning acoustic correlates to place of

---

[1] Taiwan Mandarin, as Hsu and Tse (2007:2) define, refers to the "Mandarin natively spoken by people in Taiwan, particularly young people." With the constant contact with local languages, Taiwan Mandarin has developed its own linguistic system and has become distinct from Putonghua, which is mainly modeled after Beijing Mandarin. One reviewer indicated that although Taiwan Mandarin is generically related to Mandarin in China, it does not follow that the standard dialect in Taiwan is the version spoken in Beijing. The author would like to thank the anonymous reviewer for this suggestion.

[2] The term "Putonghua", as one reviewer suggested, referred to the accent of Mandarin spoken in Beijing in Mandarin China.

nasal articulation and to syllable-final nasal mergers are reviewed in this section.

2.1 Acoustic correlates to nasal distinction

Research has shown that place characteristics of nasals are difficult to produce and perceive (House 1957, Hura, Lindblom and Diehl 1992, Malécot 1956, Mohr and Wang 1968, Narayan 2008). Few acoustic reports on nasals have pointed out certain acoustic properties for nasal place contrasts. Formant transition, for example, signals the tongue movement from the target vowel to an adjacent nasal. Formant transitions in the vowel-nasal boundary especially reflect prominent coarticulated properties (Lin and Yan 1991, Lin 2002, Mou 2006, Ohde 1994), shown in the changing vocal tract shape under the influence of both the target vowel and the following nasal. It is generally assumed that the tongue body becomes more fronted when a vowel precedes an alveolar nasal, and more backed when the vowel takes precedence over a velar nasal. The change in vowel acoustics can be quantitatively tracked by the second formant frequency, F2, which is roughly equivalent to the tongue advancement (Ladefoged 2001a, 2001b, Pickett 1999). The higher F2 is, the more advanced the vowel is. Hence, a fronted vowel has higher F2, and a backed vowel possesses lower F2.

Nasal murmur offered additional information which may enhance the perception of nasal place. In nasal murmur, resonance characteristics of nasal consonants are clearly manifested. Several researchers (cf. Cheng 1972, Chung 1990, Zhang 1996) have claimed the anticipatory effect, arguing that nasalization in the velar [ŋ] is greater than that in the alveolar [n]. In other words, the vowel preceding [ŋ] is expected to be more nasalized than the vowel preceding [n]. To measure the extent of nasalization, Chen (1995, 1997, 2000) conducted a series of acoustic studies and suggested A1-P0 (the difference in amplitudes of the first formant frequency and the first nasal pole) for the low vowel or A1-P1 (the difference in amplitudes of the first formant frequency and the second nasal pole) for the high vowel. The smaller the measurements of A1-P0 or A1-P1, the more nasalized the vowel becomes.

Previous acoustic reports have assessed relative importance of murmurs and transitions as nasal place cues, but certain controversy has remained. Some studies on natural and synthetic speech have indicated that murmurs offered predominant information about manners of articulation and that nasal place of articulation was cued primarily by transitions (Malécot 1956, Recasen 1983). Though transitions were more important than murmurs, Recasen (1983) showed that murmurs contributed significantly to the [n]-[ŋ] distinction. Contribution to places of articulation made by murmurs was also reported in other studies (Kurowski and Blumstein 1984, Repp

1986). Kurowski and Blumstein (1984), for instance, regarded nasal murmurs as effective as transitions in cueing place of articulation. A current perspective on acoustic cues to nasal consonants holds that nasal place of articulation is determined by both the murmurs and transitions together (Harrington 1994, Mou 2006, Ohde 1994). It has been widely documented (Kurowski and Blumstein 1984, Ohde and Haley 1992, Ohde and Ochs 1992, Ohde and Perry 1994, Repp 1986) that in the high vowel [i] both murmur and transition cues are required to have accurate identification of place of articulation. In summary, variable formant transitions, F2 in particular, when combined with murmurs, have been claimed most optimal for different places of articulation.

Temporal cues in the vowel region and in the murmur region may serve as a third possible acoustic correlate. So far, few studies have examined the contribution of temporal cues in syllable-final nasal distinction. As Lin (2002) suggests, acoustic differences in dimensions of vowel duration and nasal duration are worthy of investigation.

2.2 Nasal codas in Taiwan Mandarin

It has been reported that historically nasal endings in Mandarin Chinese are susceptible to change and to undergo merging from Old Chinese, Middle Chinese, to Modern Chinese (Chen 1991b, Li 1999). Instability of nasal codas has also been recently found in Mandarin spoken in Taiwan.

Previous research on syllable-final nasals in Taiwan Mandarin, however, has been bifurcated regarding the merging directions. Some researchers (Kubler 1985, Tse 1992, Yueh 1992) argued that alveolarization was the predominant trend of syllable-final nasal merger in [i] and [ə]. Tse (1992), for example, examined production and perception of Mandarin syllable-final [n] and [ŋ] by young speakers of Mandarin in Taiwan. In production, factors of vowel types and nasal types were significant. The accuracy hierarchy of production was [aN] > [iN] > [əN][3] for three vowel types and [alveolar nasal] > [velar nasal] for two nasal types. Significant interaction of these two effects was found. The finding that final [n] was more accurately produced than final [ŋ] was in agreement with Yueh's (1992) sociolinguistic observation of an on-going merger of [-ŋ] > [-n] in Mandarin spoken in Taiwan.

Others (Chen 1991a, 1991b, Hsu and Tse 2007, Ing 1985, Lin 2002) claimed that the syllable-final nasal, when following [i], was more likely to be velarized. Both Chen (1991b) and Hsu and Tse (2007) collected data from speakers of different age

---

[3] In Tse (1992), the capital N was used as a cover term for nasals, including both the alveolar nasal and the velar nasal.

groups in Taipei, the major city in northern Taiwan. A tendency for *-in* to merge into *-ing* and for *-eng* to merge into *-en* was revealed in their findings. Discrepancies about the leading merger have remained. Chen (1991a) indicated that [in] to [iŋ] was in the leading position, while Hsu and Tse (2007) suggested that [əŋ] to [ən] has been losing its distinction for decades.

Certain limitations, however, existed in previous reports. Chen (1991b), for instance, exclusively investigated nasal endings in two vowel contexts (i.e. [i] and [ə]). Tse's (1992) research, though involving three target vowels (i.e. [i], [ə] and [a]), judged whether a syllable ended in [n] or [ŋ] on the sole basis of the experimenter's ears. In a most recent study, Hsu and Tse (2007) referred to spectrogram in their analysis, but the realization of the syllable-final nasals was determined mainly by the velar pinch, an acoustic cue for the velar nasal [ŋ] with the coming together of falling F3 and rising F2. But, one might question whether the perceptual judgment or the velar pinch[4] truly revealed the nasal performance of Mandarin speakers in Taiwan.

As previously reviewed, several acoustic correlates may contribute to the distinction in nasal place of articulation, such as formant transition in the vowel-nasal boundary, the extent of nasalization, or temporal cues. Lin and Yan (1991) examined Mainland Mandarin spoken in Beijing and significant differences between the F2 values were observed at the end-point for non-high vowels to distinguish the alveolar-velar nasal codas. Mou (2006) recruited Mandarin speakers in America, and compared nasal codas in English and in Mainland Mandarin. Mainland Mandarin low and mid vowels (i.e. [a], [ə]) shifted in F2, while the high vowel [i] did not. These studies, however, excluded other acoustic correlates, such as nasalization, vowel duration and nasal duration. What role these acoustic cues play in Mandarin syllable-final nasals, how these acoustic cues influence native listeners' judgment, and which acoustic cue is the determining factor for speakers' possible nasal mergers have remained unknown. Additionally, whether acoustic patterns in Taiwan Mandarin correspond to those in Mainland Mandarin (Lin and Yan 1991, Mou 2006) is worthy of investigation.

Inspired by the literature, the researcher aims to conduct two experiments – one on acoustic measures in Mandarin nasal codas and the other on perceptual rating. Two issues are addressed in the current study. The first is to examine the contribution of murmurs and vowels in the nasal place of articulation distinction for L1 and L2 Mandarin speakers in Taiwan. Whether these speakers follow the acoustic patterns of

---

[4] In the current investigation, several subjects' productions of the velar nasal were perfectly rated as [ŋ] by the rater, but no velar pinch was shown in their spectrograms. The spectrogram of [aŋ] with a falling F2 at the vowel-nasal boundary in Figure 1 (the right panel) is an example that somewhat contradicted to the criteria adopted in Hsu and Tse (2007). Hence, velar pinch might not be a reliable criterion in the analysis of Mandarin alveolar-velar distinction in nasal codas.

English nasals, as previously reported (cf. Harrington 1994, Kurowski and Blumstein 1984, Mou 2006, Ohde 1994, Repp 1986), or confirm the acoustic patterns of nasals in Mainland Mandarin, as revealed in Lin and Yan (1991) and Mou (2006), will also be addressed. The second is to investigate perceptual saliency for the alveolar-velar nasal distinction. The results would indicate acoustic-perceptual salience in Mandarin syllable-final nasal place contrasts and provide helpful suggestions to language learners and to speech language pathologists in the clinical decision-making process.

3. Experiment 1: Acoustic measures in Mandarin nasal codas

An acoustic experiment on Mandarin nasal codas was carried out. It aimed to examine the major acoustic patterns in Taiwan Mandarin, as compared to those in Mainland Mandarin (Lin and Yan 1991, Mou 2006). To what extent the acoustic correlates contributed to place cues in nasals and to what extent linguistic variants occurred in Mandarin syllable-final nasals are discussed.

3.1 Participants

Twenty subjects without any reported speech or hearing defects participated in the first experiment. Among these subjects, ten were L1 speakers of Mandarin in Taiwan (the L1 group) and ten were Chinese as a second language learners from Burma (the L2 group). Both the L1 and L2 speakers of Mandarin were currently college students in southern Taiwan and their ages ranged from twenty to twenty-three years old. These L2 speakers of Mandarin had studied Chinese for one to two years and their Chinese proficiency was at the low-intermediate level. They were chosen because their native language – Burmese – enjoys a more marked nasal system (i.e. [m], [n], [ɲ] and [ŋ]). It would thus be interesting to examine whether the L2 learners endowed with a more marked L1 system would have no difficulty producing less marked L2 segments.

Speakers in the L1 group were young bilinguals with Mandarin and Taiwan Min as their mother tongues. The major rationale for choosing these speakers was to examine whether the nasal merger in Taiwan Mandarin (Hsu and Tse 2007, Tse 1992, Yueh 1992) existed in this age group. To what extent and in which acoustic correlates they had the alveolar-velar nasal mergers would be investigated.

3.2 Materials

Participants in the current study articulated the alveolar-velar nasal pairs ([-n] vs.

[-ŋ]) in three vowel contexts ([i], [ə], [a]) with the rising tone (i.e. *ying* /iŋ/ 'camp', *yin* /in/ 'silver', *chen* /tʂʰən/ 'Chen' (family name), *cheng* /tʂʰəŋ/ 'Cheng' (family name), *pan* /pʰan/ 'dish', *pang* /pʰaŋ/ 'side'). These target words were embedded in two-character phrases (i.e. *jun-ying* 軍營, *jun-yin* 軍銀, *xing-chen* 姓陳, *xing-cheng* 姓程, *i-pan* 一盤, *i-pang* 一旁) in a paragraph to ensure that the speakers uttered the target sounds in a natural way (Appendix A).[5]

3.3 Procedures

Nasal production of each subject was recorded individually in a quiet room free from ambient noise. The sound recorder in Praat (Boersma and Weenink 1999-2000) with sampling rate 22050, along with a SONY ECM-MS907 microphone, was used. These subjects were invited to read aloud the given paragraph in the most spontaneous way. The instruction was orally given to the subjects. They were told that their voices were being recorded and then directly saved in a wave file in Praat (Boersma and Weenink 1999-2000). The researcher monitored the recording session, and did not give the participants any hint or correction in the hope that they would articulate the segments in the most natural way.

3.4 Data analysis

Tape-recorded readings were later transformed into wave files in Praat. Several acoustic correlates in the vowel region (i.e. F2 value at the vowel-nasal boundary, duration of target vowel) and in the nasal murmur region (i.e. extent of nasalization and duration of nasal murmurs) were analyzed. Temporal cues were easily calculated in the spectrograms by defining the vowel region and the nasal murmur, as specified in Figure 1. The vowel region was usually characterized by visible periodic waves and vertical pulses in spectrograms (Fry 2001, Ladefoged 2001a, 2001b). Murmurs with lower or weaker formants were symbolic of the nasal region, which started with some abrupt changes in the waveform from the nucleus vowel. Additionally, the audible differences in Praat screens could offer an explicit way to distinguish the vowel region from the nasal region.

Four major acoustic cues of nasals, including the formant transitions at the vowel-nasal boundary (in Hz), degrees of nasalization (in dB), vowel duration (in msec) and nasal duration (in msec), were investigated. According to the coarticulated

---

[5] One reviewer suggested that the sentential contexts, such as an intonational phrase boundary, be taken into consideration in a future study. In the current study, the researcher tried to elicit nasal production from a two-character phrase as one unit, which might minimize the effect of sentential contexts. The author would like to thank the anonymous reviewer for this suggestion.

properties, it was expected that vowels would become more fronted next to [n], and more backed next to [ŋ], as manifested in F2 in the vowel region. It was then assumed that speakers might produce with higher F2 in the [Vn] tokens, and lower F2 in the [Vŋ] tokens.
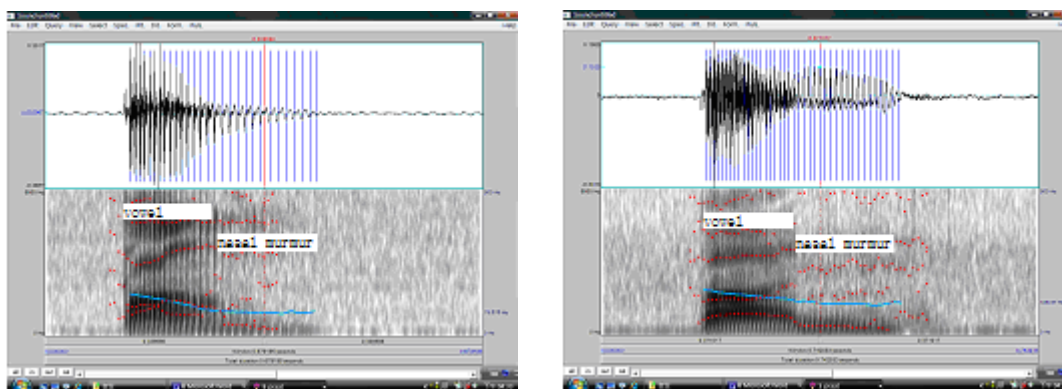

Figure 1. Spectrograms of [an] (the left panel) and [aŋ] (the right panel)

As for degrees of nasalization in the nasal region, the current measure was based on earlier studies of acoustic correlates for nasalization (Chen 1995, 1997, 2000). It was reported that an acoustic correlate A1-P0 (the first formant amplitude-the nasal peak below the first formant) or A1-P1 (the first formant amplitude-the nasal peak between the first two formants) helped quantify nasalized vowels (Figure 2).[6] A1-P0 was statistically proven to show significant difference between nasal and oral vowels, especially in non-high vowels (i.e. [ə] and [a] in the current study). A1-P1 difference, on the contrary, was specifically evident in non-low vowels (i.e. [i] in the current study). An inverse relationship was found between acoustic correlates and nasalization. The acoustic assumption regarding nasalization was that greater anticipatory effect was in the vowel preceding [ŋ] (Cheng 1972, Chung 1990, Zhang 1996); that is, greater extent of nasalization was anticipated in the [-ŋ] context than in the [-n] context.

---

[6] The length of the window is 1.880 seconds. This is for a presentation window, not for an analysis window. For an analysis window, the length of the window would be shorter to fetch the data.
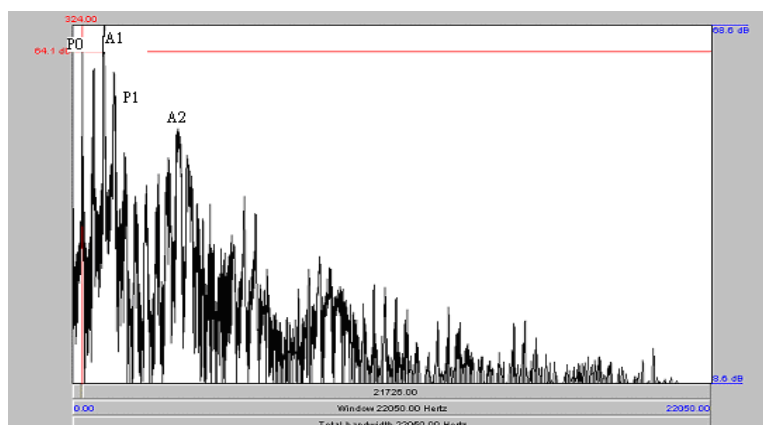
Figure 2. P0, P1, A1 and A2 in the spectra of /an/

Additionally, the role of temporal cues in distinguishing the alveolar-velar nasal pairs, though neglected in earlier research reports (Lin 2002), was addressed in the current study. All together, nasal production of L1 and L2 speakers of Mandarin in Taiwan was analyzed from four acoustic dimensions, inclusive of formant transition, temporal cues in the vowel region, nasalization (A1-P0 or A1-P1), and temporal cues in the nasal murmur.

After all the acoustic data were tabulated, statistical analyses were performed. ANOVA was used to examine the variances in the acoustic cues articulated by L1 and L2 speakers of Mandarin Chinese. Three independent variables were Language Background (L1, L2), Nasal Type ([n], [ŋ]) and Vowel Type ([i], [ə], [a]). Four acoustic cues (i.e. F2, nasalization, vowel duration, nasal duration) served as the dependent variables to observe subjects' nasal coda production. To what extent murmur and vowel contributed to the syllable-final nasal distinction for these speakers is discussed in the following section.

3.5 Results and discussion on acoustic measures

In this section, several acoustic factors in the vowel region and in the murmur region are compared to demonstrate the way in which L1 and L2 speakers of Mandarin differentiate the syllable-final nasal contrast. Statistical results are presented in terms of formant transitions in vowel (F2), degrees of nasalization, vowel duration and nasal duration.

The second formant frequency (F2) as produced by L1 and L2 speakers in relation to three factors (Language Background, Nasal Type and Vowel Type) was examined by three-way repeated measures ANOVA. The main effect of Vowel Type was significant [$F(2, 36) = 41.068$, $p<.01$]. However, main effects of Language Background and Nasal Type were not significant [for Language Background, $F(1, 18)$

= 3.175, p>.05; for Nasal Type, F(1, 18) = 0.209, p>.05]. Only the interaction between Language Background and Vowel Type was significant (p<.05).

Disregarding the other two insignificant factors, a post-hoc analysis was conducted to examine a simple effect of vowel types.[7] The finding of the post-hoc Tukey's analysis on preceding vowels was similar to that in Hsu and Tse (2007). The syllable-final nasals preceded by the vowel [a] were significantly least likely to converge (p <.05). No significant differences between the syllable-final nasals preceded by [i] and [ə] were found. This is not a surprising result for Mandarin vowel [a] has different allophonic realizations before two nasals. This would make it easier to distinguish Mandarin nasals in the [a] context.[8]

It was clear that F2 was only significant in differentiating different vowel types (i.e. [i], [ə] and [a]), not in the nasal types (i.e. [n] and [ŋ]). Both L1 and L2 speakers shared the same pattern in that they did not refer to F2 to distinguish the syllable-final nasal codas. This finding was obviously different from that in Lin and Yan (1991) and Mou (2006), which revealed that speakers of Mainland Mandarin significantly adopted F2 in the alveolar-velar nasal distinction. Their tongue body became more fronted with higher F2 when a vowel preceded an alveolar nasal, and more backed with lower F2 when the vowel took precedence over a velar nasal. By contrast, L1 and L2 speakers of Mandarin in Taiwan failed to distinguish the alveolar-velar nasal contrast in terms of F2 at the vowel-nasal boundary.

Table 1. Statistical results of ANOVA of nasalization

Factors		Mean	F
Language Background	L1	4.657	13.192**
	L2	6.715	
Nasal Type	[n]	6.360	5.661*
	[ŋ]	5.012	
Vowel Type	[i]	6.433	2.161
	[ə]	5.633	
	[a]	4.993	

Note: **$p< .01$; *$p< .05$; Nasals are surface forms.

Table 1 summarizes the statistical results of degree of nasalization as articulated by participants concerning three factors (Language Background, Nasal Type and Vowel Type); main effects of Language Background and Nasal Type were significant

---

[7] The mean F2 in the conditions across vowel types and nasal types is presented in Appendix B.
[8] This is only a conjecture. As one reviewer suggested, future studies should measure the acoustic attributes of the different allophones to validate this suggestion.

[for Language Background, F(1, 18) = 13.192, p<.01; for Nasal Type, F(1, 18) = 5.661, p<.05]. Also, the interaction between Language Background and Nasal Type was significant (p<.05). However, the effect of Vowel Type was not significant [F(2, 36) = 2.161, p>.05].

The post-hoc Tukey's analysis on syllable-final nasals indicated that L1 speakers produced mean values of nasalization in the alveolar nasal [n] significantly larger than those in the velar nasal [ŋ] (p<.05) in all three vowel types ([in]=5.55; [iŋ]=3.93; [ən]=6.23; [əŋ]=4.49; [an]=4.95; [aŋ]=2.79). As measurements of A1-P0 or A1-P1 (Figure 2) were inversely correlated with the extent of nasalization, it could be inferred that the vowel preceding [ŋ] was more nasalized than that preceding [n]. For all three vowels, L1 speakers of Mandarin nasalized to a lesser extent in the VN context than in the VNG context.

L2 speakers, by contrast, expressed different nasalized patterns from L1 speakers. Except for the [an]-[aŋ] pair ([an]=6.77; [aŋ]=5.46), the remaining two pairs (i.e. [in]-[iŋ] and [ən]-[əŋ]) were in disagreement with the anticipatory effect of nasalization in Mandarin (Cheng 1972, Chung 1990, Zhang 1996). The vowels [i] and [ə] preceding [ŋ] were less nasalized than those preceding [n]. Unlike the L1 pattern, these L2 speakers of Mandarin nasalized to a greater extent in the VN context than in the VNG context. To sum up, these L2 speakers, when compared to L1 speakers, articulated the Mandarin nasal pairs with less significant distinction, violating the acoustic assumptions of nasalization in Mandarin.

Table 2. Statistical results of ANOVA of vowel duration (in msec)

Factors		Mean	F
Language Background	L1	147	0.716
	L2	141	
Nasal Type	[n]	147	1.131
	[ŋ]	140	
Vowel Type	[i]	153	4.159*
	[ə]	148	
	[a]	131	

Note: *p< .05

In Table 2, duration of the target vowel as articulated by L1 and L2 speakers in relation to three factors (Language Background, Nasal Type and Vowel Type) was assessed by three-way repeated measures ANOVA. The main effect of Vowel Type was significant [F(2, 36) = 4.159, p<.05]. Disregarding the other two insignificant factors, a post-hoc analysis was conducted to examine a simple effect of vowel types.

As the finding in the second formant frequency, the syllable-final nasals following [a] remained significantly distinctive (p<.05) in terms of vowel duration. Vowel duration was significantly different for the three vowels, but the vowel [a] was usually produced longer than the [i] vowel (Hillenbrand et al. 1995, Ladefoged 2001a, 2001b). It was thus interesting to find the vowel tendency (Table 2) exactly opposite the universal tendency. This finding might be a result of different research methods adopted in different studies. Previous reports focused mostly on the vowel token in isolation, but in the current study the vowel duration was fetched in the VN or VNG contexts. This might be a reflection of idiosyncratic production for participants in the current study. However, the main effects of Language Background and Nasal Type were not significant [for Language Background, F(1, 18) = 0.716, p>.05; for Nasal Type, F(1, 18) = 1.131, p>.05].

Table 3. Statistical results of ANOVA of nasal duration (in msec)

Factors		Mean	F
Language Background	L1	143	0.465
	L2	137	
Nasal Type	[n]	145	1.344
	[ŋ]	135	
Vowel Type	[i]	159	16.068**
	[ə]	155	
	[a]	107	

Note: **$p< .01$

Table 3 presents the statistical results of nasal murmur duration as uttered by subjects regarding three factors (Language Background, Nasal Type and Vowel Type). The main effect of Vowel Type was significant [F(2, 36) = 16.068, p<.01]. However, main effects of Language Background and Nasal Type were not significant [for Language Background, F(1, 18) = 0.465, p>.05; for Nasal Type, F(1, 18) = 1.344, p>.05]. A post-hoc analysis was conducted to examine a simple effect of vowel types without concerning the other two insignificant factors. Again, the syllable-final nasals following [a] remained significantly distinctive (p<.05) in terms of nasal duration. Similar to the finding in Table 2, no significant difference in duration for the alveolar-velar pair across all three vowel types was found in Table 3.

4. Experiment 2: Perceptual rating

An experiment on perceptual rating was designed to determine how native

speakers of Mandarin with the alveolar-velar contrast categorized talkers with possible nasal mergers and to identify the acoustic cues that listeners used for perceptual rating. To what extent acoustic correlates interacted with perceptual saliency was addressed in the second experiment.

4.1 Raters

Two male evaluators,[9] also teachers of Mandarin for more than ten years in Taiwan, took part in the perception rating. One rater was forty-five years old, and the other was forty-three years old. The main reasons why the researcher invited teachers of Mandarin were two-fold. First, the current study followed some previous perceptual experiments (cf. Ohde 1994, Ohde and Sharf 1981) to carefully select evaluators who represented standard production to make the perceptual judgments. Secondly, these teachers of Mandarin produced clear alveolar-velar contrasts in all three vowel types (i.e. [i], [ə], and [a]) and were more sensitive to perceptual differences in Mandarin syllable-final nasals. More accurate rating in the perception experiment would hopefully be generated.

4.2 Criteria for the rating and data analysis

Each nasal production, collected in Experiment 1, was independently rated by two male experienced teachers of Mandarin based on a four-point rating scale (1=good [n], 2=not so good [n], 3=not so good [ŋ], 4=good [ŋ]). An ambiguous production was operationally defined as one with an average perceptual rating for a listener between 1.8 and 2.2 for [n], and between 2.8 and 3.2 for [ŋ]. Criteria and the scale for ratings were adapted from Ohde (1994).[10]

Following these teachers' rating, the researcher, also a native speaker of Mandarin, independently checked the tokens to make sure that all syllables were rated and none of them was missed. Cronbach's alpha for nasal codas was .94 *(SD = 0.05), ranging from 0.88 to 0.99, which indicated high inter-rater reliability.

Logistic regression between overall human ratings with acoustic cues (i.e. formant transition in vowel, extent of nasalization, vowel duration, and murmur duration) were then computed for each VN syllable and each VNG syllable. Through Multiple regression analysis, relative importance of each predictive variable to enhance

---

[9] This study was limited in that only two raters participated in the perceptual rating and that they were teachers of Mandarin. One reviewer suggested that in future studies, ordinary native speakers of Mandarin should be recruited as raters to check whether there are significant differences between their rating results and to ensure the reliability of the perceptual rating.

[10] Ohde (1994) examined the perceptual cues to the [m]-[n] distinction in CV syllables with the rating scale: 1=good [m], 2=not so good [m], 3=not so good [n], 4=good [n].

perceptual saliency for the alveolar-velar nasal distinction was compared. The most salient perceptual cues for Mandarin nasal contrast and the most efficient predictive variables were then identified. Results from the perceptual ratings might further reveal whether acoustic correlates for place of nasal articulation were embedded more in the nasal murmur region than in the vowel region, or vice versa.

4.3 Results and discussion on perceptual saliency

The issue in Experiment 2 was about how the acoustic cues produced by speakers contributed to the perceptual saliency for listeners. Multiple regression analysis was conducted to elaborate the possible interaction and the predictive power of acoustic measurements. It is valid to assess the predictive power of acoustic variables (i.e. F2 at the vowel-nasal boundary, A1-P1 or A1-P0, vowel duration, nasal murmur duration) on perceptual rating. The perceptual rating was the criterion variable, and four acoustic variables were the predictive variables. Statistical results of a step-wise multiple regression analysis are reported in Table 4.

Table 4. Statistical results of multiple regression analysis

Ranking for PR-VN	R	R^2	Adjusted R^2	F	F change	Beta
1. AP-VN	.408	.167	.167	5.59	5.59	-.415
2. ND-VN	.572	.327	.161	6.57	6.45	.401
Ranking for PR-VNG	R	R^2	Adjusted R^2	F	F change	Beta
1. AP-VNG	.534	.285	.285	11.16	11.17	-.534

Note: PR for "perceptual rating"; AP for "A1-P0 or A1-P1"; ND for "nasal murmur duration"

When all acoustic variables were entered as predictors, two of them (i.e. F2 and vowel duration) in the VN context became statistically insignificant and were excluded. An insignificant effect of F2 might be a possible result of insignificant inputs uttered by these speakers in terms of formant movements. For this reason, the raters were thus incapable of using spectral differences as a perceptual cue to the identity of Mandarin syllable-final nasals. Another possible explanation for the insignificant effect of spectral cues was that the listeners might allow a wide range of tolerance in formant transitions in nasal codas and did not attend much to spectral differences in categorizing Mandarin syllable-final nasals.

Two statistically significant predictors, including AP-VN (the extent of

nasalization) and ND-VN (the nasal murmur duration), were contained in the final model. This finding suggested that the raters primarily attended to these two acoustic measures in categorizing Mandarin nasal codas. As shown in the adjusted R^2, AP-VN (with 16.7% predictive power) was slightly more predictive than ND-VN (with 16.1% predictive power). Taken together, these two variables could predict 32.7% of the perceptual rating, as shown in the overall multiple determination coefficient (R^2 = .327). But, in the VNG context, there was only one statistically significant predictor AP-VNG (extent of nasalization). As revealed in the adjusted R^2 (R^2 = .285), AP-VNG could predict 28.5% of the perceptual rating.

The standardized regression coefficient (Beta) was -.415 for AP-VN and .401 for ND-VN in the VN context, and was -.534 for AP-VNG in the VNG context. The respective standardized equations for the PR-VN model and for the PR-VNG model are listed below.

$$PR\text{-}VN = (-.415 \text{ x AP-VN}) + (.401 \text{ x ND-VN})$$
$$PR\text{-}VNG = (-.534 \text{ x AP-VNG})$$

To conclude, two acoustic-phonetic attributes, inclusive of nasalization and murmur length, could reliably predict accented affiliation for L1 and L2 speakers of Mandarin in Taiwan. Hence, in Mandarin syllable-final nasals, acoustic correlates in the nasal region contributed more to perceptual saliency than those in the vowel region.

5. General discussion

The discussion addresses two issues in the current study. The first issue concerns the acoustic cues that L1 and L2 Mandarin speakers in Taiwan produced. It was found that these speakers differentiated the nasal types significantly in terms of nasalization. No significant contrasts in Mandarin nasal codas, however, were found in the vowel transition.

Additionally, the current investigation revealed that L2 speakers were less sensitive to the Mandarin syllable-final nasal contrast. They articulated the alveolar-velar nasal pairs with insignificant differences, especially from the perspective of nasalization. Since Burmese enjoys a more marked nasal system (i.e. [m], [n], [ɲ] and [ŋ]) than Mandarin, it is expected native speakers of Burmese should not have difficulty distinguishing less marked Mandarin nasal codas. Referring to current speech learning theories, Eckman (1977, 2004) has advocated the Markedness Differential Hypothesis (MDH). The areas of difficulty for a language learner can be

predicted based on a systematic comparison of native language, target language and the markedness relations stated in the universal grammar. For instance, some elements in L2, being different and more marked than L1, will be difficult for L2 learners, while some elements in L1, being different but not more marked, will be easier. Usually, unmarked structures are simpler than the corresponding marked ones within the areas of L1-L2 distinction. The interpretation of MDH, however, is somewhat limited in the current findings. MDH argues that learners, endowed with a more marked L1 system, are assumed to encounter less difficulty when learning a less marked L2 system. That is, Burmese speakers with more marked L1 nasal system were supposed to acquire the less marked Mandarin nasals with ease. But, this assumption was not overwhelmingly supported in the current investigation. Therefore, it could be argued that confusion in the Burmese group didn't result from the phonological system of their own native language (i.e. Burmese), but from merging nasals of the target language or from the low-intermediate L2 language proficiency. Also, it was possible that differences between these two languages could be the cause of different acoustic realizations in Mandarin nasals. Same phonemic categories did not necessarily guarantee the same phonetic realizations. Further studies on acoustic properties of Burmese nasals (cf. Dantsuji 1984, 1986, Gordon and Ladefoged 2001) should be conducted to substantiate this claim.

These acoustic patterns in Mandarin nasals produced by L1 and L2 speakers in Taiwan were extremely different from those in Mainland Mandarin, as reported in Lin and Yan (1991) and Mou (2006). In Mainland Mandarin, significant differences in F2 at the end-point for non-high vowels (i.e. [ə] and [a]) helped distinguish the alveolar-velar nasal codas. In Taiwan, speakers produced the [an]-[aŋ] pair with prominent differences in vowel formant transition. Mergers in the surface forms (i.e. [in]-[iŋ] and [ən]-[əŋ] pairs) with insignificant vowel formant transitions were thus representative of the modified nasals in Taiwan Mandarin. With no significant cues for the place of nasal articulation in formant movement, the subjects relied more on other acoustic cues for the alveolar-velar contrast, such as vowel duration and nasal murmur duration, as a compensational strategy in production.

Another issue concerned the interaction between acoustic cues and perceptual saliency. Among acoustic variables, acoustic cues in the nasal region (i.e. the extent of nasalization and the length of nasal murmur) were significantly correlated with the perceptual saliency and thus offered more predictive contribution than those in the vowel region. This finding disconfirmed the reports that murmurs provided predominant information about manners of articulation and that the nasal place of articulation was cued primarily by transitions (Malécot 1956). Instead, this was in agreement with the anticipatory effect in the alveolar-velar distinction (Cheng 1972,

Chung 1990, Zhang 1996) and the dominant contribution of murmurs in the [n]-[ŋ] distinction (Recasens 1983). Hence, this finding contributed to the importance of nasal murmur as nasal place cues in Mandarin nasal codas for speakers and listeners in Taiwan.

6. Conclusion

The current study examined the acoustic patterns of Mandarin nasal codas and its contribution to perceptual saliency. There were two major findings. Firstly, L1 and L2 speakers differentiated the nasal types significantly in terms of nasalization. Both groups adopted certain temporal cues as a compensational strategy to differentiate the alveolar nasal and the velar nasal. Compared with L1 speakers, L2 speakers produced Mandarin nasals in a significantly different way, especially in nasalization. Secondly, acoustic cues in the nasal region (i.e. the extent of nasalization and the nasal murmur duration) were significantly correlated with perceptual saliency and offered more predictive contribution than those in the vowel region.

Results in this acoustic investigation further offered prominent implications and contributions to some frequently-discussed issues. Firstly, acoustic properties in syllable-final nasals in Taiwan Mandarin were explicitly identified. Therefore, L2 learners and L1 speakers of Mandarin can be directed to different contributions of acoustic correlates in nasal coda distinction. Secondly, contribution of nasalization and nasal duration cannot be neglected in the Mandarin alveolar-velar distinction. Previous reports on nasal place cues in Mandarin have emphasized formant transitions in the vowel region, but paid little attention to the nasal region. To make clearer Mandarin alveolar-velar nasal distinction, the acoustic correlates in both the vowel region and the nasal region should be addressed. Thirdly, the findings will hopefully assist psycholinguistics in exploring the link between nasal production and its perceptual saliency. A hierarchy of priorities for teaching pronunciation to learners of Mandarin can be established following a better understanding of acoustic factors that weigh most heavily in a native speaker's reactions. In sum, results in this study will help sensitize both phoneticians and language professionals to the modified nasals of Mandarin produced by L1 speakers and L2 learners in Taiwan and provide useful implications for teachers and learners of Mandarin nasal codas.

References

Boersma, Paul, and Daivd Weenink. 1999-2000. Praat: Doing Phonetics by Computer

(Version 3.8-3.9) [Computer program]. Retrieved January 12, 1999 and October 18, 2000, from http://www.fon.hum.uva.nl/praat/.

Chen, Chung-yu. 1991a. Shengdiao de zhuanbian yu kuosan: Taipei butong nianlingqun de quyang [The tonal change and diffusion: From the sampling of different age groups in Taipei]. *Journal of the Chinese Language Teachers Association* 16:69-99.

Chen, Chung-yu. 1991b. The nasal endings and retroflexed initials in Peking Mandarin: Instability and the trend of changes. *Journal of Chinese Linguistics* 19:139-155.

Chen, Marilyn Y. 1995. Acoustic parameters of nasalized vowels in hearing-impaired and normal-hearing speakers. *Journal of the Acoustical Society of America* 98:2443-2453.

Chen, Marilyn Y. 1997. Acoustic correlates of English and French nasalized vowels. *Journal of the Acoustical Society of America* 102:2360-2370.

Chen, Marilyn Y. 2000. Acoustic analysis of simple vowels preceding a nasal in Standard Chinese. *Journal of Phonetics* 28:43-67.

Chen, Yang, Michael Robb, Harvey Gilbert, and Jay Lerman. 2001. Vowel production by Mandarin speakers of English. *Clinical Linguistics & Phonetics* 15:427-440.

Cheng, Chin-chuan. 1972. *A Synchronic Phonology of Mandarin Chinese*. The Hague: Mouton.

Chung, Raung-fu. 1990. Underspecification and Chinese phonology. Paper presented at the 6th Linguistic Society of America Annual Meeting, Ohio State University.

Dantsuji, Masatake. 1984. A study on voiceless nasals in Burmese. *Studia Phonologica* [*Study of Phonology*] 18:1-14.

Dantsuji, Masatake. 1986. Some acoustic observations on the distinction of place of articulation for voiceless nasals in Burmese. *Studia Phonologica* [*Study of Phonology*] 20:1-11.

Eckman, Fred. 1977. Markedness and the contrastive analysis hypothesis. *Language Learning* 27:315-330.

Eckman, Fred. 2004. From phonemic differences to constraint rankings: Research on second language phonology. *Studies in Second Language Acquisition* 26:513-549.

Flege, James-Emil, Ocke-Schwen Bohn, and Sunyoung Jang. 1997. Effects of experience on non-native speaker's production and perception of English vowels. *Journal of Phonetics* 25:437-470.

Fry, Dennis Butler. 2001. *The Physics of Speech*. Cambridge: Cambridge University Press.

Gordon, Matthew, and Peter Ladefoged. 2001. Phonation types: A cross-linguistic overview. *Journal of Phonetics* 29:383-406.

Hajek, John. 1997. *Universals of Sound Change in Nasalization.* Oxford: Blackwell Publishers.

Harrington, Jonathan. 1994. The contribution of the murmur and vowel to the place of articulation distinction in nasal consonants. *Journal of Acoustical Society of America* 96:19-32.

Hillenbrand, James, Laura Getty, Michael Clark, and Kimberlee Wheeler. 1995. Acoustic characteristics of American English vowels. *Journal of Acoustical Society of America* 97:3099-3111.

House, Arthur S. 1957. Analog studies of nasal consonants. *Journal of Speech and Hearing Disorders* 22:190-204.

Hsu, Hui-ju, and John Kwock-ping Tse. 2007. Syllable-final nasal mergers in Taiwan Mandarin: Leveled but puzzling. *Concentric: Studies in Linguistics* 33:1-18.

Hura, Susan L., Björn Lindblom, and Randy L. Diehl. 1992. On the role of perception in shaping phonological assimilation rules. *Language and Speech* 35:59-72.

Ing, R. O. 1985. Guoyu fayin zai Taiwan: Muqian qushi yu yiban cuowu zhi tantao [Mandarin pronunciation on Taiwan: An analysis of recent trends and errors]. *Proceedings of the First International Conference on the Teaching of Chinese as a Second Language*, 413-425. Taipei: Shijie Huawen.

Kubler, Cornelius. 1985. *The Development of Mandarin in Taiwan: A Case Study of Language Contact*. Taipei: Student Publishing.

Kurowski, Kathleen, and Sheila E. Blumstein. 1984. Perceptual integration of the murmur and formant transitions for place of articulation in nasal consonants. *Journal of Acoustical Society of America* 76:383-390.

Ladefoged, Peter. 2001a. *Vowels and Consonants.* Mass: Blackwell.

Ladefoged, Peter. 2001b. *A Course in Phonetics*. 4th ed. Orlando: Harcourt Brace College.

Li, Wen-chao. 1999. *A Diachronically-Motivated Segmental Phonology of Mandarin Chinese*. New York: Peter Lang Publishing Inc.

Lin, Chien-jer. 2002. Nasal endings of Taiwan Mandarin: Production, perception and linguistic change. Paper presented at the 35th International Conference on Sino-Tibetan Languages and Linguistics, Arizona State University.

Lin, Maocan, and Jingzhu Yan. 1991. Putonghua dai biwei lingshengmu yinjie zhong de xietong fayin [Coarticulation in the zero-initial syllable with nasal ending in Standard Chinese]. *Yuyin Yanjiu Baogao* [*Report of Phonetic Research*], 68-86. Beijing: Chinese Academy of Social Sciences Institute of Linguistics Phonetics Laboratory.

Luo, Tsun-lin. 2002. Production and Perception of Similar and New Vowels by Mandarin Speakers: An Experimental Study of English High Vowels. Doctoral

dissertation, National Taiwan Normal University.

Ma, Lin. 1995. Production of Five American English Front Vowels by Adult Mandarin Chinese Speakers. Doctoral dissertation, Texas A&M University.

Malécot, André. 1956. Acoustic cues for nasal consonants: An experimental study involving a tape-splicing technique. *Language* 32:274-284.

Mohr, B., and W. S.-Y. Wang. 1968. Perceptual distance and the specification of phonological features. *Phonetica* [*Phonetics*] 18:31-45.

Mou, Xiao-min. 2006. Nasal Codas in Standard Chinese: A Study in the Framework of the Distinctive Feature Theory. Doctoral dissertation, Massachusetts Institute of Technology.

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

Ohde, Ralph N. 1994. The development of the perception of cues to the [m]-[n] distinction in CV syllables. *Journal of Acoustical Society of America* 96:675-686.

Ohde, Ralph N., and Katarina Haley. 1992. The developmental role of interspeaker variability in the perception of the /m/-/n/ distinction of CV syllables. *Journal of Acoustical Society of America* 92:2463-2464.

Ohde, Ralph N., and Marleen Ochs. 1992. The developmental role of spectral relations in processing place of articulation of nasal consonants from short duration segments. *Journal of Acoustical Society of America* 91:2360-2361.

Ohde, Ralph N., and Anne Perry. 1994. The role of short-term and long-term auditory storage in processing spectral relations for adult and child speech. *Journal of Acoustical Society of America* 96:1303-1313.

Ohde, Ralph N., and Donald Sharf. 1981. Stop identification from vocalic transition plus vowel segments of CV and VC syllables: A follow-up study. *Journal of Acoustical Society of America* 69:297-300.

Pickett, J. M. 1999. *The Acoustics of Speech Communication: Fundamentals, Speech Perception Theory, and Technology*. Boston: Allyn and Bacon.

Recasen, Daniel. 1983. Place cues for nasal consonants with special reference to Catalan. *Journal of Acoustical Society of America* 73:1346-1353.

Repp, Bruno H. 1986. Perception of the [m]-[n] distinction in CV syllables. *Journal of Acoustical Society of America* 79:1987-1999.

Tse, Kwock-ping. 1992. Production and perception of syllable final [n] and [ŋ] in Mandarin Chinese: An experimental study. *Studies in English Literature and Linguistics* 18:143-156.

Wang, Xin-chun. 1997. The Acquisition of English Vowels by Mandarin ESL Learners: A Study of Production and Perception. MA thesis, Simon Fraser University.

Yueh, Li-chi. 1992. The Drift of the Velar Nasal Ending in Taiwan Mandarin: A Sociolinguistic Survey. MA thesis, Fu Jen Catholic University.

Zhang, Jie. 1996. Variation in degree of nasalization for Mandarin vowels with superimposed rhoticity. *Journal of Acoustical Society of America* 100:2687.

Department of Western Languages & Literature
National University of Kaohsiung
Kaohsiung, TAIWAN
Yi-hsiu Lai: laiyihsiu@yahoo.com.tw

Appendix A

Please read the following paragraph in Chinese.

又到了年度軍營中，軍官們盤點軍銀的時間。只見姓程的軍官端進一盤又一盤的銀兩，但是，姓陳的軍官卻只站在一旁，袖手旁觀。

Appendix B. Means, SD, Minimum and Maximum in F2 (in Hz)

Group	Sound	N	Mean	SD	Max	Min	t-value
TW	[in]	10	1895	279	1354	2339	-1.77
	[iŋ]	10	2136	239	1812	2534	
	[ən]	10	1590	136	1434	1923	0.77
	[əŋ]	10	1532	231	1139	1985	
	[an]	10	1569	168	1373	1946	1.98*
	[aŋ]	10	1244	256	1022	1910	
BU	[in]	10	1813	508	1227	2822	-2.77
	[iŋ]	10	2367	503	1543	3167	
	[ən]	10	1672	236	1404	2133	0.67
	[əŋ]	10	1630	322	1084	2230	
	[an]	10	1645	156	1227	2822	3.64*
	[aŋ]	10	1426	291	1063	1918	

Note: TW = Taiwanese speakers; BU = Burmese speakers; N = number; SD = standard deviation; Max = maximum; Min = minimum; *$p < .05$

華語鼻音結尾之聲學特質
與其感知清晰貢獻之研究

賴怡秀

國立高雄大學

　　本文藉由兩項實驗探討兩大議題：一、華語鼻音結尾發音部位之聲學特質；二、該聲學線索對舌根鼻音與舌尖鼻音（[-n]：[-ŋ]）感知清晰之貢獻程度。實驗一，聲學分析華語母語人士與以華語為第二語言學習者之華語鼻音結尾發音。結果顯示：共陣峰走勢在本研究受試者語料未達顯著差異，呈現舌根鼻音與舌尖鼻音某種程度混淆現象。實驗二，將受試者發音語料請具經驗之華語教師評聽。發音聲學數據與感知評聽結果經回歸分析後，發現鼻音鼻化程度是主要影響華語教師區辨舌根鼻音與舌尖鼻音之依據。

關鍵詞：華語鼻音結尾、共陣峰走勢、鼻化程度