Wuxi Speakers' Production and Perception of Coda Nasals in Mandarin

Lei Wang, Jie Cui, Ying Chen[*]

School of Foreign Studies, Nanjing University of Science and Technology, China

`laylaw@foxmail.com, jiecui@njust.edu.cn, ychen@njust.edu.cn`[*]

Abstract

Wuxi natives speak a dialect of Wu, which has only one coda nasal /n/ but allows allophones depending on the pre-nasal vowel [1], whereas in Mandarin, there are two coda nasals—alveolar /n/ and velar /ŋ/. Two perception experiments were conducted to investigate Wuxi speakers' perception and production of coda nasals in their second language (L2) Mandarin. First, two groups of Wuxi native speakers, age around 20 and 50, produced monosyllabic words with nasal coda in Mandarin and their production was used as the stimuli for native Mandarin speakers to identify. Second, the same Wuxi speakers participated in an identification task to judge the place of articulation of the nasal coda in monosyllabic words in standard Mandarin. The results of the first experiment indicate that young Wuxi speakers' Mandarin production was identified with higher accuracy by native Mandarin speakers than older Wuxi speakers', suggesting the young speakers produced more nativelike Mandarin than the older speakers. The results of the second experiment reveal that young Wuxi speakers identified coda nasals in Mandarin more accurately than older Wuxi speakers did, suggesting Wuxi speakers' production of Mandarin coda nasals is associated with their perception.

Index Terms: coda nasal, perception, production, age, vowel

1. Introduction

L2 learners tend to perceptually relate phonemes in their L2s to the closest sounds in their L1s [2]. The Speech Learning Model (SLM) claims L2 learners' perception influences their production. Similarly, the Perceptual Assimilation Model (PAM) [2] proposes that information perceived in speech is correlated to certain properties of speech production. Specifically, phonetic segments in L2 can be produced in a nativelike fashion only if they are perceived in a nativelike fashion [3]. Flege et al. [4] examined the production of English consonants /ð/ and /θ/ by native Italian speakers and found that Italian learners tended to perceive word-initial /ð/ as /d/. They argued that the inaccurate production of English /ð/ and /θ/ arose from the kind of discriminative failure [3]. Furthermore, SLM also proposes that age of Learning (AOL) influences L2 speech perception and production [3, 4, 5, 6]. That is, the likelihood of crosslinguistic phonetic differences being discerned decreases with AOL [3, 4, 5, 6, 7]. For example, Yamada [8] and Flege et al. [9] found that the earlier native speakers of Japanese were exposed to English and the longer they had been immersed in English, the more accurately they perceived the contrast between /l/ and /ɹ/. Similar findings were obtained for Japanese learners of English in MacKain et al.'s study [10]. Native Japanese speakers have difficulty producing and perceiving the distinction between English phoneme /l/ and /ɹ/ [10]. This is attributed to the fact that Japanese has only one liquid, while English has two. Besides AOL, other factors that may also affect perception and production of L2 speech,

including relative amount of L1/L2 use [6, 11] and the quantity and quality of input from native L2 speakers [11, 12]. More experienced listeners discern certain nonnative L2 contrasts better than less experienced listeners [10, 13, 14, 15, 16].

Wu is a Chinese language spoken in the south to Yangtze River. Wu speakers learn Mandarin as L2 and use it as an official language in daily life. The young generation of Wu speakers start learning Mandarin earlier and use more Mandarin in daily life than older speakers. In Mandarin, there are two coda nasals—alveolar /n/ and velar /ŋ/, whereas there is only one coda nasal /n/ in Wuxi Wu though allophones are allowed depending on the pre-nasal vowel [1, 17]. Previous studies of L2 speech led us to posit that Wuxi speakers may perceive and produce both nasals in Mandarin as the alveolar /n/ of their L1 Wu dialect. The production of closed syllables with nasal endings involves a process of coarticulation with preceding vowels [18]. Coarticulatory cues from the preceding vowels have been found to strongly affect the perception of place of articulation of postvocalic nasals [19]. According to Zee [18] and Chen et al. [20], the high front vowel /i/ resulted in more misidentification of coda nasals than mid vowel /ə/ and low vowel /a/. Given its potential effects in perceiving the coda nasal, vowel type is taken into consideration in the current study.

Two perception experiments were conducted to examine Wuxi speakers' perception and production of coda nasals in their L2 Mandarin. Based on the previous literature, both speakers' age and vowel quality were included as exploratory factors. Three research questions will be investigated: (1) Do native Mandarin speakers perceive young Wuxi speakers' and older Wuxi speakers' Mandarin production of coda nasals differently? (2) Do the young Wuxi speakers and the older Wuxi speakers perceive Mandarin coda nasals differently? (3) Is Wuxi speakers' Mandarin production of coda nasals correlated to their perception?

2. Experiment 1: Mandarin speakers' perception of Wuxi speakers' Mandarin

2.1. Participants

Twelve native Wuxi Wu speakers were recruited for two age groups: young (mean age 23.3 yrs), and older (mean age 49.2 yrs). Each group had three males and three females. All participants are bilingual speakers of Wuxi Wu and Mandarin. Twenty native Mandarin participants (mean age 24.1 yrs) who were born and raised in northern China were recruited from Nanjing University of Science and Technology. None of these Mandarin participants speak other Mandarin varieties or other Chinese languages. All participants are familiar with the romanization convention for /n/ and /ŋ/, which are noted as 'n' and 'ng' respectively in Pinyin script.

2.2. Stimuli

Six monosyllabic words, all bearing the high-level tone in Mandarin, served as the stimuli for this experiment: [pʰin], [pʰiŋ], [pʰən], [pʰəŋ], [pʰan], [pʰaŋ]. The stimulus words were embedded in a carrier sentence: "x, *zhe ge zi du* x (x, this character is read as x)". The twelve Wuxi speakers were requested to read these stimulus sentences with five repetitions in a random order in Mandarin. The first "x" in the carrier sentence, was extracted as the target word for the perception experiment. The first, third and fifth repetitions of each Wuxi speakers' production were chosen as the perceptual stimuli. This procedure yielded 216 tokens (6 words * 3 times * 12 participants) for the perception experiment.

2.3. Procedure

The stimuli were recorded in mono channel at a 44,100 sampling rate by a Marantz PMD661 professional recorder and a Shure SM10A-CN head-worn microphone. Recordings took place in a quiet room and the sound files were then digitized on an SD card to save on a personal computer. The intensity of the stimulus words was RMS normalized to 70 dB. The perception experiment was conducted in the sound- attenuated booth of Language Cognition and Speech Science Lab at Nanjing University of Science and Technology. Stimuli were presented in randomized order using the Praat ExperimentMFC program with Sony MDR-7506 professional studio headphones. Mandarin speakers are asked to participate in a forced-choice identification task to judge the nasal type (either 'n' or 'ng') by clicking the corresponding button on computer screen.

2.4. Results

A logistic regression model was applied to the data using the statistics software R (version 3.4.3). Participants' response (two levels: accurate, inaccurate) to /n/ or /ŋ/ was set as the dependent variable. Vowel type (three levels: /a/, /ə/, /i/) and age (two levels: young, older) were set as independent variables. The first level of the response, i.e., the 'accurate' response, was set as the reference level. Both vowel type and speakers' age emerged as significant factors that correlate with the accuracy of nasal identification and no interaction between vowel and age was found (see Table 1).

Table 1: *Logistic regression results of nasal identification accuracy in Experiment 1.*

	b	*SE*	*z*	*p*
(Intercept)	2.262	0.103	22.030	0.000 ***
Vowel	-0.483	0.022	-21.958	0.000 ***
Age	0.518	0.070	7.400	0.000 ***

The mean percent accuracy of nasal identification by vowel type and age group is illustrated in Figure 1. Mandarin speakers performed well in the perception of both the young (90%) and the older (89%) Wuxi speakers' Mandarin nasals preceded by the vowel /a/ than by the vowel /i/ (both 48%). Collapsing vowel types, Mandarin listeners more accurately identified the young Wuxi speakers' production of Mandarin nasals (72%) than the older speakers' (62%).
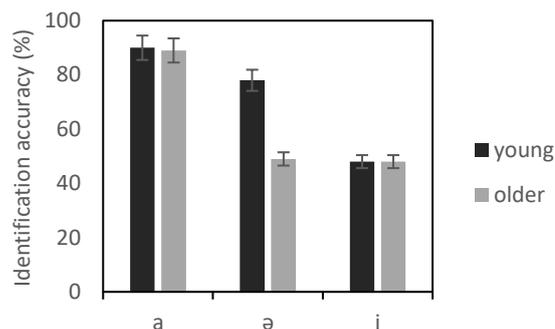
Figure 1: *Percent accuracy of coda nasal identification by vowel and age in Experiment 1.*

Further prediction analyses were applied to examine the probability of accurate response by each vowel type and age group, in order to find out which type of vowel and age group are more likely to lead to accurate perception. The probability that the accurate response was calculated based on the logistic regression model.

The result shows stimuli containing the high front vowel /i/ are less likely to be accurately identified. In other words, the nasal consonants preceded by /i/ tend to be misidentified (see Table 2). In terms of age, the young speakers' production of Mandarin coda nasals was more likely to be identified correctly by native speakers of Mandarin than the older speakers' (see Table 3).

Table 2: *Probability of accurate identification of coda nasals by vowel type in Experiment 1.*

Vowel	Probability
an	0.88
aŋ	0.82
ən	0.74
əŋ	0.64
in	0.53
iŋ	0.41

Table 3: *Probability of accurate response by age group in Experiment 1.*

Age	Probability
Young	0.72
Older	0.62

3. Experiment 2: Wuxi speakers' perception of native speakers' Mandarin

3.1. Participants

The same twelve Wuxi speakers in Experiment 1 participated in this perception experiment.

3.2. Stimuli

The stimuli were taken from part of the stimuli in Chen and Guion-Anderson [21], where four Beijing Mandarin speakers were recruited to produce the target words, [pʰin], [pʰiŋ], [pʰən], [pʰəŋ], [pʰan], [pʰaŋ], with the high-level tone in Mandarin with three repetitions in random order. A total of 72 tokens (6 words

* 3 times * 4 participants) were obtained. The experiment was composed of two trials. In the first trial, all stimuli were presented with no background noise. In the second trial, the stimuli were embedded in pink noise, whose frequency spectrum is closer to nasals than other types of noise, with -4dB SNR, in order to avoid ceiling effects.

3.3. Procedure

The same procedure as in the perception task in Experiment 1 was adopted. Experiment 1 was conducted before Experiment 2 because we hoped to obtain natural and normal production of Wuxi speakers' Mandarin. Thus, we recorded their Mandarin production first to edit as the stimuli of Experiment 1 and then requested them to do Experiment 2. By running Experiment 1 before Experiment 2, we avoided possible effects of unintentional perceptual training.

3.4. Results

Similar to Experiment 1, a logistic regression was applied to predict accurate identification. Participants' responses (two levels: accurate, inaccurate) were set as the dependent variable. Vowel type (three levels: /a/, /ə/, /i/), age (two levels: young, older), noise (two levels: with, without) were set as the independent variables. The results were presented in Table 4. Age and vowel were both found to contribute to the prediction but not noise. Again, no interaction between vowel and age was found.

Table 4: *Logistic regression results of nasal identification accuracy in Experiment 2.*

	b	*SE*	*z*	*p*
(Intercept)	1.189	0.158	7.508	0.000***
age	1.173	0.113	10.380	0.000***
noise	-0.169	0.110	-1.536	0.125
vowel	-0.243	0.033	-7.391	0.000***

Figure 2 shows the accuracy of Wuxi speakers' identification of Mandarin coda nasals. Both young and older Wuxi speakers scored higher in vowel /a/ (92% and 62% respectively) than in vowel /i/ (63% and 51% respectively). Collapsing vowel types, the young speakers obtained higher accuracy (81%) compared to the older speakers (58%).
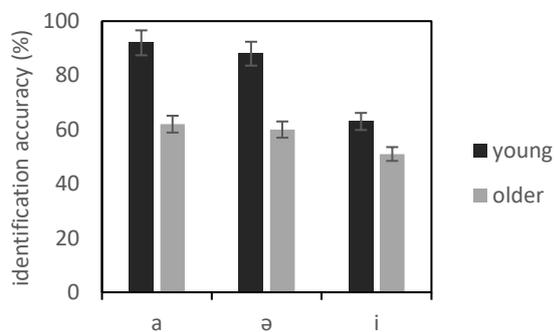
Figure 2: *Percent accuracy of coda nasal identification by vowel and age in Experiment 2.*

The same prediction function of logistic regression model as in Experiment 1 was applied to determine which type of vowel or age group was the most likely to be accurately identified. The results are consistent with the findings in Experiment 1. The high front vowel /i/ led to the most perception difficulties (see Table 5). In terms of age, the young Wuxi Wu speakers are more likely to get the accurate response (see Table 6).

Table 5: *Probability of accurate identification of coda nasals by vowel type in Experiment 2.*

Vowel	Probability
an	0.80
aŋ	0.77
ən	0.72
əŋ	0.68
in	0.62
iŋ	0.60

Table 6: *Probability of accurate response by age group in Experiment 2.*

Age	Probability
Young	0.81
Older	0.58

4. Discussion

Experiment 1 revealed that young Wuxi speakers' Mandarin production was identified with higher accuracy by native Mandarin speakers than that of older Wuxi speakers' production. This result suggests that the young Wuxi speakers produced more nativelike Mandarin than the older speakers. It is consistent with the results of Flege et al. [12], where the experienced learners were found to produce and perceive English vowels more accurately than the relatively inexperienced learners. The differences in the participants' age differentiate their L2 Mandarin experience. Young Wuxi speakers started learning Mandarin in early childhood, while the majority of the older speakers acquired Mandarin at a later stage in life. Therefore, young Wuxi speakers' AOL of Mandarin is much earlier than the older speakers. Also, due to the countrywide popularity of Mandarin Chinese as the official language and the rapid development of Mandarin media in China since the late 1970s, young Wuxi speakers were exposed to higher quality and quantity of standard Mandarin than the older Wuxi speakers in childhood and in school. Moreover, young Wuxi speakers use more L2 Mandarin in their daily life. The amount of L2 use also correlates to the outcome of L2 speech learning [22].

In Experiment 2, the accuracy rate of Mandarin nasal identification by the older Wuxi speakers was around 50% (see Figure 2), close to chance. This finding confirms previous findings in the literature that L2 consonants that do not exist in the listeners' L1 phonological inventory tend to yield perception difficulties [10]. Given the mismatch of the nasal phonemes between Wuxi Wu (/n/) and Mandarin (/n/ and /ŋ/), it is not surprising that Wuxi Wu speakers have difficulty in discerning the contrast between /n/ and /ŋ/. Experiment 2 also showed that the young Wuxi speakers obtained higher accuracy of Mandarin nasal identification compared to the older speakers. This result is consistent with the result of Experiment

1 that native Mandarin speakers auditorily judged the young Wuxi speakers' production of Mandarin coda nasals better than those of the older speakers, suggesting a correlation between the perception and production in Mandarin of Wuxi speakers. This finding supports the proposal of PAM and SLM that perception predicts production in L2 speech learning. Therefore, we give positive answers to all the three research questions in the current study.

In addition, vowel quality was found to influence the perception of coda nasals in the current study as in the literature [20, 21, 23]. Specifically, coda nasals following the high front vowel /i/ were more likely to be misidentified by Wuxi speakers than those following the mid vowel /ə/ and the low vowel /a/ in Experiment 2. This was also reflected in their production based on the perception of native Mandarin speakers in Experiment 1. However, the background noise did not affect the accuracy of coda nasal identification in Experiment 2, suggesting a relatively stable status of the acquisition of coda nasals in L2 Mandarin by Wuxi speakers.

5. Conclusions

The current study investigated Wuxi speakers' production and perception of coda nasals in Mandarin and found that young Wuxi speakers produced and perceived Mandarin coda nasals more accurately than older Wuxi speakers. The high front vowel /i/ resulted in more coda nasals errors than the mid vowel /ə/ and the low vowel /a/ in both production and perception of Wuxi speakers' L2 Mandarin.

Future work will statistically examine the correlation between Wuxi speakers' production and perception of their L2 Mandarin to look into whether there are correlated on the level of individual speakers. Acoustic analysis of Wuxi speakers' Mandarin production will be conducted and compared with the native Mandarin speakers' production. Future work will also involve the analysis of Wuxi speakers' production of coda nasals in Wu in order to trace back the L1 effect on the L2 perception and production, especially the potential influence of the allophonic variations of coda nasals in Wu.

6. Acknowledgements

7. References

[1] N. R. Qian, *The Research of Contemporary Wu (Dangdai Wuyu Yanjiu)*. Shanghai: Shanghai Education Press, 1992.

[2] C. T. Best, A direct realistic view of cross-language speech perception. In W. Strange (Ed.), *Cross-Language Speech Perception* (pp. 171-204). Timonium, MD: York Press, 1995.

[3] J. E. Flege, Second language speech learning: Theory, findings and problems. In Speech Perception and Linguistic Experience. W. Strange (Eds), *Issues in Cross-Language Research* (pp. 233–272.) MD: York, Timonium, MD, 1995.

[4] J. E. Flege, M. Munro, and I. MacKay, "Effects of age of second language learning on the production of English consonants," *Speech Communication*, vol. 16, no. 1, pp. 1-26, 1995.

[5] J. E. Flege, M. Munro, and L. Skelton, "Production of the word-final English /t/-/d/ contrast by native speakers of English, Mandarin and Spanish," *Journal of the Acoustical Society of America*, vol. 92, no. 1, pp. 128-143, 1992.

[6] J. E. Flege and I. R. A. MacKay, "Native Italian speakers' perception and production of English vowels," *Journal of Acoustical Society of America,* vol. 106, no. 5, pp. 2973-2987, 1999.

[7] R. Yamada, and Y. Tohkura, "Age effect on acquisition of non-native phonemes: perception of English /r/ and /l/ for native speakers of Japanese," *Proceedings of the 12th international congress of phonetic sciences, August 19-24, Aix-en-Provence,* France, 1991.

[8] R. Yamada, Age and acquisition of second language speech sounds: Perception of American English /ɹ/ and /l/ by native speakers of Japanese. In W. Strange (Eds), *Speech Perception and Linguistic Experience, Issues in Cross-Language Research*, (pp. 305–320). MD: York, Timonium, 1995.

[9] J. E. Flege, N. Takagi, and V. Mann, "Lexical familiarity and English-language experience affect Japanese adults' perception of /ɹ/ and /l/," *Journal of the Acoustical Society of America*, vol. 99, no. 2, pp. 1161-1173, 1996.

[10] K. S. MacKain, C. T. Best, and W. Strange, "Categorical perception of English /r/ and /l/ by Japanese bilinguals," *Applied Psycholinguistics*, vol. 2, no. 4, pp. 369-390, 1981.

[11] G. Jia, and D. Aaronson, "A longitudinal study of Chinese children and adolescents learning English in the United States," *Applied Psycholinguistics*, vol. 24, no. 1, pp. 131–161, 2003.

[12] J. E. Flege, and S. Liu, "The effect of experience on adults' acquisition of a second language*," Studies in Second Language Acquisition*, vol. 23, no. 4, pp. 527-552, 2000.

[13] C. T. Best and W. Strange, "Effects of phonological and phonetic factors on cross-language perception of approximants," *Journal of Phonetics*, vol. 20, no. 3, pp. 305-330, 1992.

[14] E. Frieda, A. Walley, J. E. Flege, and M. Sloane, "Adults' perception and production of the English vowel /i/," *Journal of Speech & Hearing Research*, vol. 43, no. 1, pp. 129-143, 2000.

[15] J. E. Flege, O-S. Bohn, and S. Jang, "The effect of experience on nonnative subjects' production and perception of English vowels," *Journal of Phonetics*, vol. 25, no. 4, pp. 437-470, 1997.

[16] J. E. Flege, and I. R. A. MacKay, "Perceiving vowels in a second language*," Studies in Second Language Acquisition,* vol. 26, no. 1, pp. 1-34, 2004.

[17] Y. R. Zhao, *The Research of Modern Wu (Xiandai Wuyu de Yanjiu).* Beijing: Commercial Press, 1928.

[18] E. Zee, "Effect of vowel quality on perception of post-vocalic nasal consonants in noise." *Journal of Phonetics*, vol. 9, pp. 35-48, 1981.

[19] P. Lageford, and I. Maddieson, *The Sounds of the World's Languages*. Oxford: Blackwell, 1996.

[20] Y. Chen, V. Kapatsinski, and S. Guion, "Acoustic cues of vowel quality to coda nasal perception in Southern Min," *Proceedings of the 13th Annual Conference of the International Speech Communication Association*, 2012, pp. 1412-1415.

[21] Y. Chen and S. Guion-Anderson, "Perceptual confusability of word-final nasals in Southern Min and Mandarin: implications for coda nasal mergers in Chinese," *Proceedings of the 17th International Congress of Phonetic Sciences*, 2011, pp. 464-467.

[22] Y. Chen, Y. Xu, and S. Guion-Anderson, "Prosodic realization of focus in bilingual production of Southern Min and Mandarin," *Phonetica*, vol. 71, no. 4, pp. 249-270, 2014.

[23] J. Cole and K. Iskarous, Effects of vowel context on consonant place identification: implications for a theory of phonologization. In E. Hume and K. Johnson (Eds), *The Role of Speech Perception in Phonology* (pp.103-122). New York: Academic Press, 2011.