TONGUE CONTOURS FOR THE JAPANESE MORAIC NASAL BY SPEAKERS OF STANDARD CHINESE

Ai Mizoguchi[1, 2], Maho Morimoto[3, 4], Li Weiyu[3], Takayuki Arai[3]

[1]Maebashi Institute of Technology, [2]NINJAL, [3]Sophia University, [4]JSPS
aimizoguchi@maebashi-it.ac.jp, maho.morimoto.jp@gmail.com, w-li-5x4@eagle.sophia.ac.jp, arai@sophia.ac.jp

ABSTRACT

The Japanese moraic nasal /N/ has been transcribed as a uvular nasal [ɴ] in the utterance final position, although variability in the place of articulation by speaker has been reported. For L2 learners whose L1 is English, /N/ is often produced as [n], avoiding an unfamiliar place of articulation. The current study investigated the tongue contours for /N/ by native speakers of Standard Chinese learning Japanese using ultrasound to analyze whether the same avoidance strategy is applied for learners with different L1s.The results showed that two of four speakers used [n] for /N/, while one speaker created a tongue blade constriction and one showed a tongue back gesture. It was unclear whether proficiency level affected their articulation because even an advanced learner showed /n/ substitution. English and Chinese speakers seem to have different acquisition processes and strategies for unfamiliar phonemes.

Keywords: Japanese moraic nasal, Standard Chinese, L2, articulation, ultrasound.

1. INTRODUCTION

A Japanese coda nasal, also called a moraic nasal, /N/, which forms a mora on its own, has often been transcribed as a uvular nasal, [ɴ], in the utterance final position [1–3]. Recent articulatory studies have revealed the variability in its place of articulation by speakers from the alveolar ridge to the uvula [4–6], although most speakers produce /N/ as a velar or uvular nasal.

By contrast, for learners of Japanese, it has been observed that /N/ is produced as an alveolar nasal, [n], by native speakers of English [7]. This is probably because there is no uvular sound in the English phoneme inventory and thus native English speakers use a peripheral variation of /N/, which is used in their L1. As all three speakers including an advanced learner have been found to use [n] for /N/ in [7], this seems to be a robust characteristic for learners of Japanese.

To investigate whether this strategy of avoiding an unfamiliar place of articulation is universal across languages, the current study observed the tongue contours for /N/ uttered by native speakers of Standard Chinese, which also has no uvular sound.

Standard Chinese has two coda nasals, /n/ and /ŋ/. The coda /n/ is often produced with a weaker constriction than the onset /n/, resulting in an incomplete closure at the alveolar ridge [8, 9]. The current study first investigated the tongue contours for the Chinese coda nasals and then compared them with the tongue contours for the Japanese moraic nasal uttered by the same speakers. It was hypothesized that Standard Chinese speakers substitute /n/ for the Japanese /N/, as English speakers do.

2. METHODS

2.1. Participants

Six native Standard Chinese speakers (3 males and 3 females) who had been learning Japanese participated in the experiment. Of these, four participants were included in the current analysis to compare learners of different proficiency levels. Two were advanced learners (ACF02 and ACM02), while the other two self-reported as basic-level learners (ACF03 and ACM03). The advanced learners had lived in Japan more than three years and reached the highest level (N1) in the Japanese-Language Proficiency Test. The basic-level learners had lived in Japan for fewer than four months. Two speakers were from Beijing (ACF03 and ACM03), one speaker was from Dalian (ACF02), and one speaker was from Inner Mongolia (ACM02). A speaker ID prefix 'ACF' was used for female speakers and 'ACM' was used for male speakers. In addition, three native speakers of Japanese (2 females and 1 male) participated in the experiment for comparison purposes, although the articulatory analysis was not conducted for the current study.

The participants were recruited by the authors and were compensated for their participation in the study. They consented to the experimental procedure approved by the ethics committee of Sophia University. They were in their 20s, had no history of speech/hearing impairment, and had corrected-to-normal vision. Language background information about the participants was collected using a questionnaire.

2.2. Speech materials

The speech materials consisted of 29 Chinese words and 15 Japanese words including nasal consonants preceded and followed by a low-vowel /a/ with and without oral consonants designed for a broad analysis. For the current investigation, four Chinese words and one Japanese word were analyzed. The target phonemes were the Chinese onset /t/ in /an.ta/ (安打 [ān.tǎ] 'a hit'), onset /n/ in /an.na/ (安娜 [ān.nà] 'Anna'), coda /n/ in /an.an/ (暗暗 [àn.àn] 'secretly'), coda /ŋ/ in /aŋ.aŋ/ (昂昂 [áŋ.áŋ] 'brave-looking'), and the Japanese coda /N/ in /kaN.aN/ (勘案 ([kaɴ.aɴ] 'consideration'). The Chinese onset /t/ and onset /n/ were used for comparison purposes. The participants read aloud the words displayed on the screen one at a time. The Chinese words were written in Chinese characters and the Pinyin and Japanese words were written in Japanese characters (Kanji and Hiragana). The words were randomized and each word was repeated 10 times. The Chinese trials followed the Japanese trials after a short break.

2.3. Data collection

The audio and ultrasound recordings took place in a soundproof room. The audio signal was digitally recorded monaurally at 22,050 Hz using a RODE-NT2-A microphone through an audio interface (Focusrite, Scarlett Solo 2nd Gen.), connected to a laptop computer.

Real-time mid-sagittal images of the oral cavity were recorded with an ultrasound system (MicrUs, EXT-1H) using a microconvex probe (MC10-5R10S-3). The probe was placed under the participant's chin with an UltraFit Headset [10] stabilizing the relative position of the probe and head. The video frame rate was 113 fps. The tongue images were recorded while the speakers uttered the words and the palate images were recorded while they swallowed water. The audio and ultrasound video were recorded and synchronized using AAA software [11]. Nine tokens were excluded from the analysis due to a synchronization error.

2.4. Analysis

The midpoint of each target phoneme from the audio recording was measured using Praat software [12]. The tongue contours at the acoustic midpoint and palate edge were manually traced using GetContours [13] and interpolated to 100 points on an x/y coordinate for each frame. The origin of the ultrasound fan shape was set to zero. If the tongue images at the acoustic midpoint were unclear, then an image up to three frames before the midpoint was used. One token was not included in the analysis due to poor image quality. The number of analyzed tokens was 190 (5 target phonemes × 10 repetitions × 4 speakers − 10 excluded tokens).

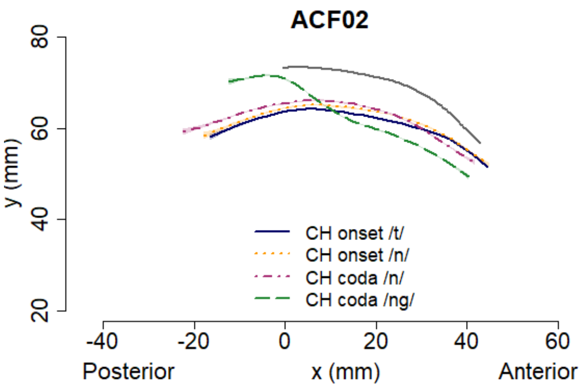
The predicted tongue contours from the repetitions were plotted by speaker for each phoneme using a generalized additive model (GAM) [14]. The contours were trimmed at the observed maximum and minimal points on the x-axis.

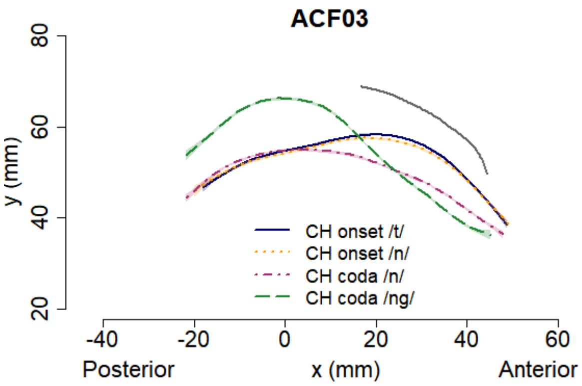
3. RESULTS

3.1. Tongue contours for the Chinese phonemes

Fig. 1 shows the predicted tongue contours with 95% confidence intervals for the Chinese phonemes: the onset /t/, onset /n/, coda /n/, and coda /ŋ/ by speaker. All the speakers showed a tongue front gesture for the coda /n/ and a tongue back gesture for the coda /ŋ/.

For three of the speakers (ACF02, ACF03, and ACM03), the onset /n/ showed a significant difference around the tongue tip from the coda /n/, but no significant difference from the onset /t/. As it has been reported [8, 9], the constrictions for the coda /n/ seemed to be weaker than those for the onset /n/ for these three speakers.





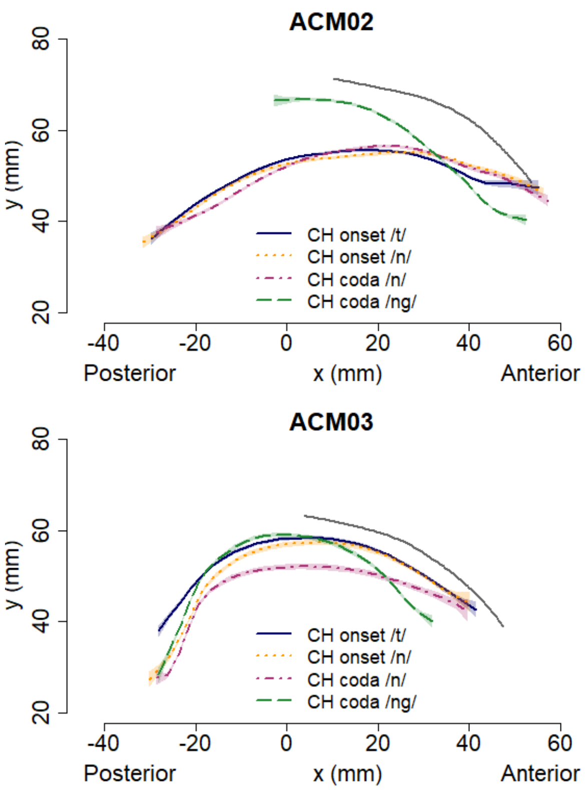
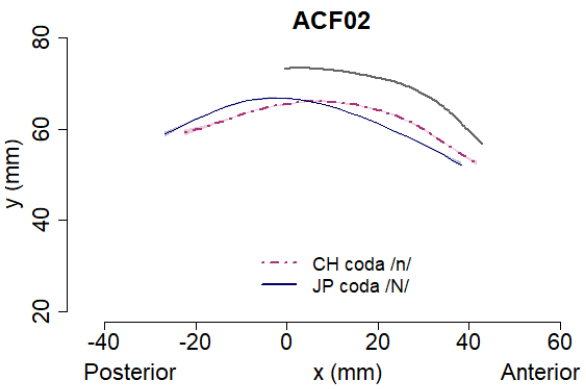


Figure 1: The GAM contours with 95% confidence intervals for the four Chinese phonemes: the onset /t/, onset /n/, coda /n/, and coda /ŋ/ by speaker. The gray line above the tongue contours shows the palate contour.

3.2. Tongue contours for the Japanese moraic nasal

Fig. 2 shows the predicted tongue contours with 95% confidence intervals for the Chinese coda /n/ and Japanese coda /N/ by speaker.

ACF02 and ACF03 showed different tongue shapes between the Chinese coda /n/ and Japanese coda /N/, while the other two speakers showed similar tongue shapes for the two phonemes. ACF02 showed a tongue front gesture for the Chinese /n/ and a tongue back gesture for the Japanese /N/. ACF03, on the contrary, showed tongue front gestures for both the Chinese /n/ and the Japanese /N/, although the tongue blade positions differed.



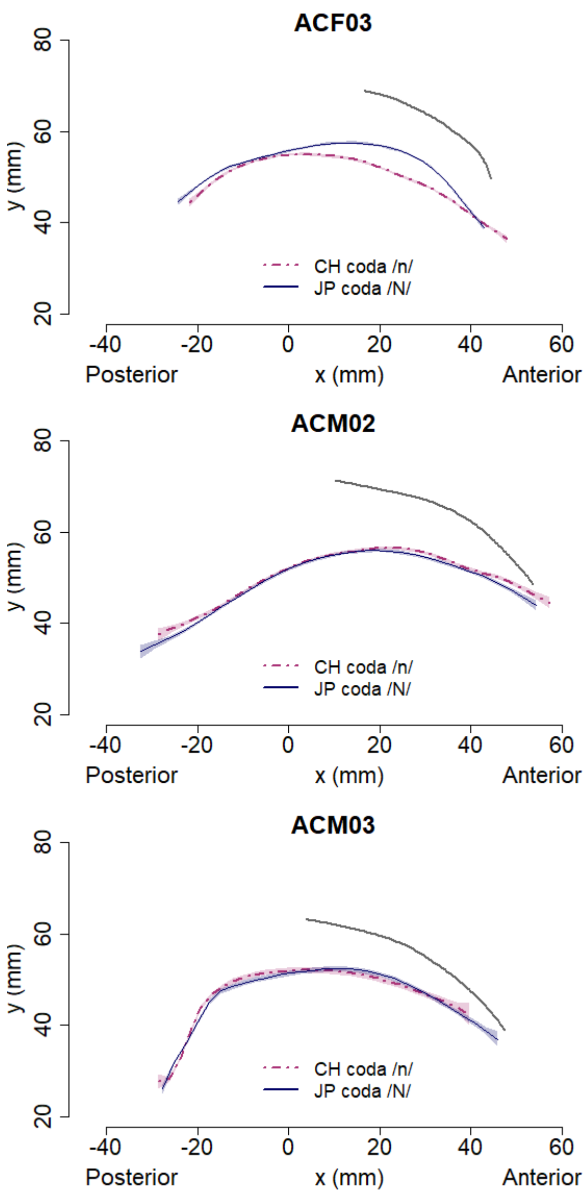


Figure 2: The GAM contours with 95% confidence intervals for the Chinese coda /n/ and Japanese coda /N/ by speaker. The gray line above the tongue contours shows the palate contour.

4. DISCUSSION

The tongue contours for the Chinese coda nasals /n/ and /ŋ/ and the Japanese moraic nasal /N/ by native speakers of Standard Chinese were investigated using ultrasound imaging. It was confirmed that the Standard Chinese speakers used the tongue front gesture for /n/ and tongue back gesture for /ŋ/. Moreover, the coda /n/ had a weaker constriction than the onset /n/ for three of the speakers, in line with [8, 9].

Two speakers (ACM02 and ACM03) showed a tongue front gesture for /N/ that was similar to their /n/ articulation. One of these two speakers was an advanced learner and the other was a basic-level learner. This /n/ substitution for /N/ has been observed in native speakers of English [7] regardless of their proficiency level. This suggests that, as seen in English speakers, the Japanese moraic nasal was not realized as a back nasal, as most Japanese speakers do, by some of the Standard Chinese speakers, including one advanced learner.

However, one basic-level learner (ACF03) showed a tongue blade gesture for /N/ dissimilar from /n/ and /ŋ/. This speaker appeared to notice some difference in /N/ but could not achieve a tongue back constriction. It is hard to conclude whether this speaker misperceived /N/ as a front nasal less anterior than /n/ or perceived /N/ as a back nasal but had difficulty to articulate as a uvular nasal.

Furthermore, one advanced learner (ACF02) showed a clear tongue contour difference between the Chinese coda /n/ and Japanese moraic /N/. This suggests that the Japanese /N/ could be acquired like native speakers of Japanese by Chinese speakers and the /n/ substitution is not robust, as seen in English speakers. However, it will be necessary to collect data from a sufficient number of speakers to draw a conclusion.

The reasons why /N/ is substituted by /n/ and not /ŋ/ remain unclear. One possible reason is orthography, as [7] pointed out. Since /N/ is often written with ‘n’ in the Roman alphabet, learners are already used to it when beginning to learn the Japanese writing system. Some Japanese language textbooks for Chinese learners also use ‘n’ for the utterance final /N/ to explain the sound. However, it seems less common to write Japanese in the Roman alphabet than in English settings. Rather, typing letters on computers might play a role because either ‘n’ or ‘nn’ is required to type /N/ in Hiragana.

Another possible reason for the /n/ substitution is the preceding vowel quality. In Chinese, /a/ in /an/ is a front vowel [a], whereas /a/ in /aŋ/ is a back vowel [ɑ] [9]. By contrast, /a/ in Japanese has been described as a centralized vowel [ä] [2, 3] and few descriptions have been reported in relation to the following consonant. If the difference in vowel qualities between /an/ and /aŋ/ are salient enough for Chinese speakers to distinguish the two nasals, vowel quality could be a cue to identify the Japanese /aN/, whose coda consonant is unfamiliar to them. In other words, if the Japanese /a/ in /aN/ is perceived as [a] by Chinese speakers, they would identify the sound as /an/ and thus produce /an/ for /aN/.

To test this hypothesis, we conducted a post-hoc acoustic analysis. Fig. 3 shows the F2 values for /a/ before /n/, /ŋ/, and /N/ by the two native speakers of Chinese (ACF02 and ACM02) and two Japanese speakers (1 female and 1 male). /an/ and /aŋ/ were produced in the speakers’ first languages.

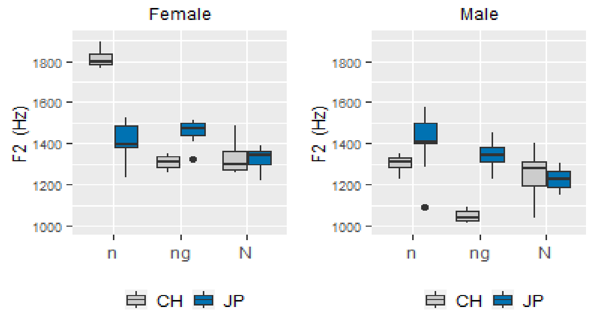


Figure 3: F2 values for /a/ in /an/, /aŋ/, and /aN/.

As shown in Fig. 3, the /a/ uttered by the Chinese speakers had higher F2 values before /n/ than /ŋ/. A linear mixed-effects model analysis [15] and a pairwise post-hoc test [16] revealed that the differences were significant (female: β=499.52, t=17.74, p<0.01; male: β=255.40, t=6.94, p<0.01). By contrast, for the /a/ uttered by the Japanese speakers, no significant differences between /an/ and /aŋ/ were found (female: β=−46.02, t=−1.63, p=1.00; male: β=59.60, t=1.62, p=1.00). However, there were significant differences between /aŋ/ and /aN/ (female: β=132.37, t=4.70, p<0.01; male: β=118.4, t=3.21, p=0.03). A difference between /an/ and /aN/ was significant for the male speaker (β=178.00, t=4.83, p<0.01) and that for the female speaker approached statistical significance (β=86.35, t=3.07, p=0.0508), but did not reach the conventional threshold of 0.05.

The Chinese female speaker, who was an advanced learner and could differentiate /an/ and /aN/ in articulation (Fig. 2, ACF02), had significantly different F2 values between them (β=478.47, t=16.99, p<0.01) and similar values for /aŋ/ and /aN/ (β=−21.05, t=−0.75, p=1.00). The Chinese male speaker, who was an advanced learner but could not differentiate /an/ and /aN/ well in articulation (Fig. 2, ACM02), had similar F2 values for them (β=48.3, t=1.31, p=1.00) and different F2 values between /aŋ/ and /aN/ (β=−207.1, t=−5.62, p<0.01).

The clear difference of the nucleus vowel quality between /an/ and /aŋ/ in Chinese implies that vowel quality could be a cue to perceive these syllables. However, a lack of systematic differences in the Japanese /a/ depending on the coda consonants could result in variability among the speakers to realize the Japanese /aN/, making it harder for learners to use such information.

5. CONCLUSION

/n/ substitution for /N/ by native speakers of Standard Chinese was observed in two of the four speakers. It is necessary to evaluate more speakers to understand their acquisition processes and strategies for unfamiliar phonemes compared with speakers of other languages.

6. ACKNOWLEDGMENTS

This study was supported by JSPS KAKENHI Grant Numbers JP19K13254, JP22J01381, JP23K12212, and Sophia University Special Grant for Academic Research (Research in Priority Areas).

7. REFERENCES

[1] Maddieson, I. 2011. Uvular consonants, In: Dryer, M. S., Haspelmath, M. (eds), *The World Atlas of Language Structures Online.* Max Planck Digital Library.

[2] Okada, H. 1999. Japanese. In: International Phonetic Association. (ed), *Handbook of the International Phonetic Association*, Cambridge University Press, 117–119.

[3] Vance, T. J. 2008. *The Sounds of Japanese*. Cambridge University Press.

[4] Mizoguchi, A., Tiede, M. K., Whalen, D. H. 2022. Inter-speaker variability of articulation for the Japanese moraic nasal : An ultrasound study. *Phonological Studies*, 25, 121–132.

[5] Yamane, N., Gick, B. 2010. Speaker-specific place of articulation: Idiosyncratic targets for Japanese coda nasal, *Can. Acoust.*, 38, 3, 136–137.

[6] Maekawa, K. 2021. Production of the utterance-final moraic nasal in Japanese: A real-time MRI study, *J. Int. Phon. Assoc.*, 1–24.

[7] Mizoguchi, A., Tiede, M. K., Whalen, D. H. 2019. Production of the Japanese moraic nasal /N/ by speakers of English: An ultrasound study. *Proc. 19th ICPhS* Melbourne, 3493–3497.

[8] Dow, F. D. M. 1972. *An Outline of Mandarin Phonetics*. Australian National University Press.

[9] Duanmu, S. 2007. *The Phonology of Standard Chinese*. Oxford University Press.

[10] Spreafico, L., Pucher, M., Matosova, A. 2018. UltraFit: A speaker-friendly headset for ultrasound recordings in speech science. *Proc. 19th Interspeech 2018* Hyderabad, 1517–1520.

[11] Articulate Instruments Ltd. 2022. Articulate Assistant Advanced (AAA).

[12] Boersma, P., Weenink, D. 2021. Praat: Doing phonetics by computer. Version 6.1.55, retrieved October 2021 from http://www.praat.org/.

[13] Tiede, M. K. 2022. GetContours. GitHub repository, https://github.com/mktiede/GetContours.

[14] Wood, S. N. 2006. *Generalized Additive Models: An Introduction with R*. chapman and hall/CRC.

[15] Bates, D., Mächler, M., Bolker, B., Walker, S. 2014. Fitting linear mixed-effects models using lme4, *arXiv Prepr. arXiv1406.5823*.

[16] Lenth, R., Singmann, H., Love, J., Buerkner, P., Herve, M. 2018. Emmeans: Estimated marginal means, aka least-squares means. *R package.*