UNCOVERING SYLLABLE-FINAL NASAL MERGING IN TAIWAN MANDARIN: AN ULTRASONOGRAPHIC INVESTIGATION OF TONGUE POSTURES AND DEGREES OF NASALIZATION

Chenhao Chiu[1], Yu-An Lu[2], Yining Weng[1], Shao-Jie Jin[2], Wei-Chen Weng[2], Tzu-Hsuan Yang[2]

[1] Graduate Institute of Linguistics, National Taiwan university, [2] Department of Foreign Languages and Literatures, National Chiao Tung University, Taiwan
chenhaochiu@ntu.edu.tw, yuanlu@nctu.edu.tw, wynjdws@gmail.com, courtney0419003@gmail.com, chris86112299@gmail.com, elsieyang05@gmail.com

ABSTRACT

Syllable-final nasals /n/ and /ŋ/ in Taiwan Mandarin have been reported to be undergoing merging. Perceptual studies have reported that the alleged merging is context-sensitive and the merging directions are vowel-dependent. These findings have been mostly attributed to dialectal and social factors. The current study uses ultrasonography to capture postures of the entire tongue during the production of syllable-final nasals. The results, though confirming that the merging directions of syllable-final nasals are vowel-dependent, are best accounted for by the biomechanics of the tongue, as supported by computational 3D model simulations. Furthermore, for some speakers, although nasals were merged in tongue postures, the degrees of nasalization of the preceding vowel were contrastive, suggesting that the merging process is incomplete. The effect of nasalization on different syllabic positions and potential influences from the exposure to Taiwan Southern Min are also discussed.

Keywords: nasal merging, nasalization, ultrasound, Taiwan Mandarin

1. INTRODUCTION

While Mandarin syllables allow nasal /n, ŋ/ codas, these syllable-final nasals have been reported to be undergoing merging in Taiwan Mandarin [4, 7, 8]. Previous studies have shown that /in/ is more likely to be merged to [iŋ], whereas /əŋ/ is more likely to be merged to [ən]. Moreover, merging only occurs in /i/N and /ə/N contexts but not in /a/N context. Acoustically, the syllable-final /n, ŋ/ affects the nasalization of the preceding vowels to different degrees [2, 11]. Specifically, pre-/ŋ/ vowels are more strongly nasalized than pre-/n/ vowels [11]. However, it remains unclear whether the alleged syllable-final nasal merging has an articulatory grounding (e.g., overlapped tongue postures) and whether the articulatory merged gestures are reflected in the degree of nasalization. The current study tackles these issues by asking two research questions: (1) Is there any articulatory evidence of syllable-final nasal merging in Taiwan Mandarin? If so, can the merging be accounted for by the biomechanics of the tongue? (2) Are there any differences regarding the degree of nasalization of the pre-nasal vowels in relation to the articulatorily merged nasals? To uncover the articulatory gestures of syllable-final nasals, we used ultrasonography to examine the overall posture of the whole tongue and a 3D tongue model to simulate tongue postures in different vowel contexts. The acoustics of the pre-nasal vowels were also measured and analyzed to answer the second research question.

2. METHODS

2.1. Participants

Data were collected and analyzed from nine participants (6F, 3M) with a mean age of 23.2. All participants were native speakers of Taiwan Mandarin, and none reported any auditory or visual disabilities. The study was conducted in accordance with ethical guidelines approved by National Chiao Tung University, Taiwan. All participants were compensated monetarily for their time.

2.2. Apparatus

The ultrasonography recording was conducted using a portable ultrasound machine (CGM OPUS5100) with a transvaginal electronic curved array probe (CLA 651). Participants sat upright wearing the ultrasound stabilization headset by Articulate Instruments [1]. The transducer was fixed at 90 degrees for every participant and was adjusted along the midsagittal tongue contour. A Samson C01U hypercardioid condenser microphone was placed directly facing the participant's mouth approximately 20 cm away. Acoustic and ultrasound data were recorded simultaneously by a USB 3.0 powered capture card (ExtremeCap U3), saved as .mp3 and .mp4, respectively. Acoustic signals were sampled at 48000 Hz, and the frame rate for the ultrasound videos was set at 40 fps.

2.3. Procedures and stimuli

The experiment employed a design of self-paced sentence reading. The stimuli were embedded in a sentence-final position with the sentences ranging from seven to nine syllables, following Hsu and Tse [7] and Fon et al. [4]. The target syllable had the structure of (C)VN, where V was one of the three /i, ə, a/ vowels and N was either an alveolar nasal /n/ or a velar nasal /ŋ/. The experiment began with 4 practice sentences that differed from the stimulus sentences, followed by 84 test trials (= 3 vowels × 2 nasal codas × 14 tokens) in one block. The test trials were repeated three times by blocks (84 × 3 = 252). All experimental stimuli were randomized and delivered using E-prime.

2.4. Data preparation and analyses

The vowel and nasal in the target syllables were first labelled in Praat. To determine the degree of nasalization of the pre-nasal vowels, the differences between *a1* and *p0*, and between *a1* and *p1* were measured [2]. *A1* represents the amplitudes in dB of the first formant. *P0* represents the amplitude in dB of the nasal peak at low frequencies, while *p1* represents the amplitude in dB of the nasal peak above the first formants. *A1 – p0* is typically negatively correlated with the degree of nasalization for non-high vowels, whereas *a1 – p1* is negatively correlated with that for high vowels [2]. That is, the lower the *a1 – p0* and *a1 – p1*, the stronger the nasalization. The values of *a1 – p0* and *a1 – p1* were obtained using the Nasality Automeasure Script Package developed at the University of Colorado's Phonetics Lab [10].

The labeled vowel and nasal segments with textgrids were imported into ELAN (http://tla.mpi.nl/tools/tla-tools/elan/) for video annotation and image capturing. Still images of tongue postures were captured from the midpoint of the labelled nasal interval. Next, tongue postures were first manually traced and then optimized using EdgeTrak. The results were first generated in the Cartesian coordinate system and later converted into polar coordinates, which were fit through smoothing spline analysis of variance (SS ANOVA) with 95% confidence interval (dashed-lined ribbons in the following figures) around the predicted fit (solid lines in the following figures) [3]. The SS ANOVA not only visualizes the optimal fits for tongue postures but also provides statistically explanatory power: any white space between two pairs of ribbons indicates a statistical difference at 95% confidence interval. To fit predicted tongue contours, a virtual origin was used as the polar origin for the SS ANOVA results following the method introduced in Heyne and Derrick [6].

2.5. 3D model simulation

To examine if the merging directions across different vowel contexts can be accounted for by the biomechanics of the tongue gesture (i.e., less movement efforts), biomechanical simulations were conducted to compare different nasal postures and merging directions. A 3D finite-element method (FEM) tongue model developed within the ArtiSynth simulation toolkit [9] was used. The tongue model implements a number of intrinsic and extrinsic tongue muscles to simulate tongue postures, including genioglossus (posterior, medial, anterior; GGp, GGm, GGa), hyoglossus (HG), inferior longitudinal (IL), superior longitudinal (SL), and styloid (STY). Tongue postures for /in, iŋ, ən, əŋ, an, aŋ/ were achieved by activating various muscles as outlined in Table 1. Muscle activations were manually set to achieve canonical postures for the six syllable combinations that are under examination. For each posture, the tongue muscles were set to simulate a coarticulatory event with a nasal target superposed onto the tongue posture of the preceding vowel [5].

Table 1: Muscle activation (%) for simulated tongue postures across /i, ə, a/N contexts

Muscles	/in/	/iŋ/	/ən/	/əŋ/	/an/	/aŋ/
GGp	15	15	20	–	–	–
GGm	20	–	5	–	–	–
GGa	–	–	10	10	5	5
HG	–	–	5	5	5	5
IL	95	95	35	35	35	–
SL	25	25	20	20	20	–
STY	–	60	–	60	–	60

The relative strain of the tongue was visualized as RGB images in ArtiSynth, with blue representing a low level of strain, red representing a high level of strain, and green representing in between. To evaluate overall strain of the tongue, we calculated the mean RGB values individually using ImageJ. The ratio of the mean red value over the mean blue value and the ratio of the mean green value over the mean blue value were calculated to determine the degree of overall strain. The larger the ratio, the more the strain induced.

3. RESULTS

3.1. Acoustic results

Standardized *a1 – p0* and *a1 – p1* measurements were reported for /i, ə/ and for /a/, respectively, at the end of the pre-nasal vowels. As mentioned above, the lower the values, the stronger the nasalization. The values with an errorflag produced using the Nasality Automeasure Script Package were excluded (∼20%). Two-tailed independent *t*-tests showed stronger nasalization for the vowels preceding the underlying /ŋ/ than preceding the underlying /n/ in the /a/N context (p = .04, Fig. 1).

Figure 1: Standardized *a1 – p0* and *a1 – p1* measured at the end of the pre-nasal vowels. The lower the values, the stronger the nasalization.

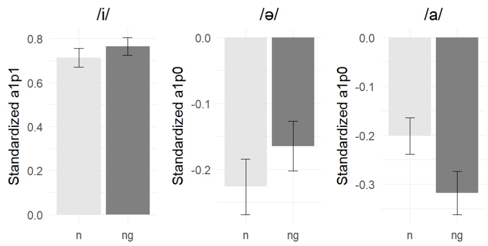
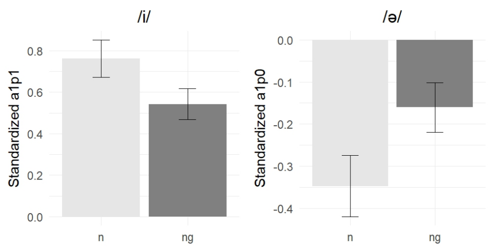


Figure 2: Standardized *a1 – p0* and *a1 – p1* from selected participants. The lower the values, the stronger the nasalization.

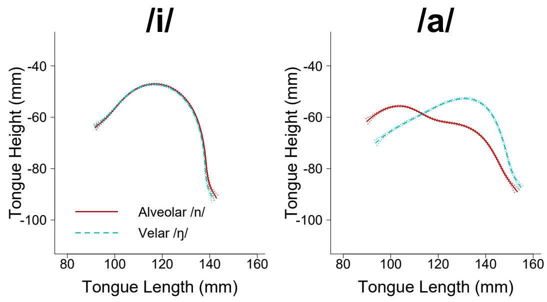


Apart from the articulatorily and acoustically distinction between /an/ and /aŋ/, the degrees of nasalization between pre-/n/ and pre-/ŋ/ vowels were not significantly different (/i/ context, p =.39; /ə/ context, p =.27). Upon further analysis, we found three speakers for each of the /i, ə/ vowels who merged /n/ and /ŋ/ articulatorily yet still had distinct contrast in the degree of nasalization at a marginal level (Fig. 2, p = .06 for /i/ and p = .046 for /ə/). The results suggest that for some speakers, the merging of /n/ and /ŋ/ is incomplete.

3.2. Ultrasound results

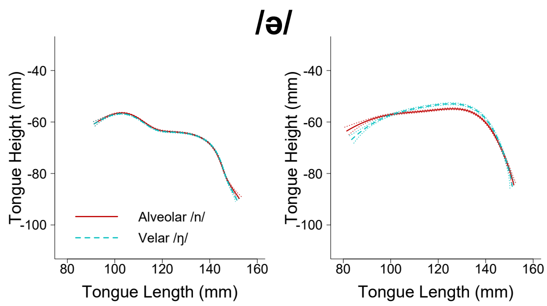
All nine participants showed complete merging in tongue postures for the /i/N contexts and no merging for the /a/N context. In the /ə/N contexts, some speakers showed nasal merging while others didn't. SS ANOVA results of the tongue postures across different vowels are illustrated in Fig. 3 - 4.

Figure 3: Two representative SS ANOVA results of the nasals in the /i, a/N contexts. Tongue tip on the left of the figure.



The syllable-final alveolar /n/ and velar /ŋ/ were completely merged in the /i/N context (Fig. 3 left). In particular, the tongue for the alveolar nasal /n/ is bunched at the body but not at the tip close to the alveolar ridge. This retraction of the bunched tongue overlapped with the tongue posture for velar nasal /ŋ/. The merging direction is from /in/ to [iŋ] and thus is considered a process of velarization. As for the /a/N context, no overlapping tongue postures were observed across all participants. The two nasals contrasted each other in the positioning of both the tongue tip and tongue dorsum (Fig. 3 right).

Figure 4: Two representative SS ANOVA results of the nasals in the /ə/N context. Tongue tip on the left of the figure.

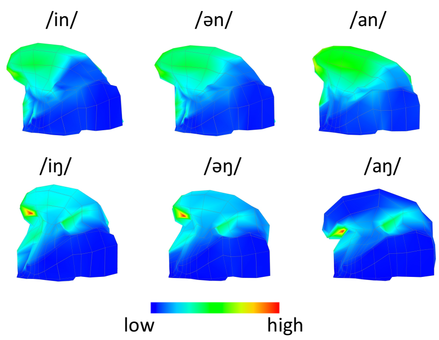


While velarization was found across all the participants in the /i/N context, differential nasal merging was observed in the /ə/N context. In this context, five speakers merged the tongue postures for both nasals (Fig. 4 left), while others showed non-overlapping tongue postures (Fig. 4 right). Among the five participants who demonstrated nasal merging in the /ə/N context, three positioned their tongues with an elevation at the tongue tip and contact at or close to the alveolar ridge (/əŋ/ → [ən]; Fig. 4 left). The two other participants demonstrated overlapped tongue postures with a lowered tongue tip and with more bunching and retraction in the tongue body (/ən/ → [əŋ], results similar to those shown in Fig. 4 right but with overlapped tongue posture contours). These results suggest that coronalization is the dominant merging pattern the /ə/N context. For those who did not show merged tongue postures in the /ə/N context, the tongue body was positioned fairly flat with some retraction close to the velar position (Fig. 4 right).

3.3. Simulation results

The simulation results of tongue postures across /i, ə, a/N contexts are illustrated in Fig. 5.

Figure 5: Midsagittal views of relative strain from low (blue) to high (red), for tongue postures in different contexts.



Compared with the tongue posture of /in/, the relative strain of the tongue for /iŋ/ was smaller. The majority of the strain occurred at the tongue tip and body for /in/, whereas more strain was evident at the tongue tip and dorsum, and a lesser extent at the tongue body, for /iŋ/ (Fig. 5 leftmost panel). In the posturing of /ən/, greater strain was observed in the tongue tip and body, but for that of /əŋ/, strain was present at the tongue tip and dorsum. Similarly, notable strain was observed in tongue tip and body for /an/ while more strain was observed in the tongue tip and dorsum towards the velum for /aŋ/. A summary of red/blue and green/blue ratios is presented in Table 2.

Table 2: The ratios of mean red over mean blue values and mean green over mean blue values across /i, ə, a/N contexts. Larger ratio indicates greater strain (i.e., more effortful).

	Red/Blue	Green/Blue
/in/	0.61	0.84
/iŋ/	0.60	0.81
/ən/	0.61	0.87
/əŋ/	0.60	0.79
/an/	0.66	0.93
/aŋ/	0.57	0.71

4. DISCUSSION

The current study used ultrasonography to investigate the supposed syllable-final nasal merging in Taiwan Mandarin. The results showed that the nasal /n, ŋ/ merging found in perceptual and acoustic studies in previous studies have an articulatory grounding. In particular, more pervasive merging in terms of tongue postures was observed in the /i/N than in the /ə/N contexts, while no merging was observed in the /a/N context. The merging directions are also vowel-dependent; velarization was found in the /i/N context whereas coronalization was more dominant in the /ə/N context. Fon et al. [4] reported velarization as well as coronalization in the /i/N context with the coronalization considered having a Taiwan Southern Min origin. The across-the-board velarization in the /i/N context might have suggested a decrease influence on Taiwan Mandarin from Southern Min. These merging directions can be attributed to biomechanical reasons. When the tongue is already in the /i/ position moving into a nasal, posturing an alveolar nasal would induce higher strain to the tongue tip and body muscles compared to the strain that occurs in the tongue body muscles for /ŋ/ production. A costlier movement is disfavored, and thus the merging direction in the /i/N context undergoes a process of velarization, in which /in/ becomes [iŋ]. The evaluation of tongue biomechanics also accounts for the less pervasive merging in the /ə/N context. Five out of our nine participants showed oveeeerlapped tongue postures in this context; two of them demonstrated a velarized nasal, along with the four participants without merging who also showed more back tongue positions. The divergence of the two merging directions for the /ə/N context would suggest that the merging for this context is incomplete. While biomechanics of the tongue would favor velarization, there may be some other reasons for coronalized merging. Our results also demonstrate that different degrees of nasalization can still be maintained even when the tongue postures may merge across different conditions, further suggesting an incomplete merging (cf. [11]). That is, for some speakers, the tongue positions merge while nasalization remains distinct.

5. ACKNOWLEDGEMENTS

We would like to thank Dr. Bryan Gick and the ICPhS reviewers for their precious comments and suggestions. This work was supported by the Ministry of Science and Technology Grant (MOST 106-2410-H-009-031) awarded to Dr. Yu-an Lu.

6. REFERENCES

[1] Articulate Instruments Ltd, 2008. *Ultrasound Stabilisation Headset Users Manual: Revision 1.4.* Edinburgh, UK: Articulate Instruments Ltd.

[2] Chen, M. Y. 2000. Acoustic analysis of simple vowels preceding a nasal in Standard Chinese. *Journal of Phonetics* 28, 43–67.

[3] Davidson, L. 2006. Comparing tongue shapes from ultrasound imaging using smoothing spline analysis of variance. *Journal of the Acoustical Society of America* 120(1), 407–415.

[4] Fon, a. H. J.-M. a. H. Y.-H., J., Hsu, H.-J. 2011. Dialectal variations on syllable-final nasal mergers in Taiwan Mandarin. *Language and Linguistics* 12(2), 273–311.

[5] Gick, B., Stavness, I., Chiu, C. 2013. Coarticulation in a whole event model of speech production. *International Congress of Acoustics 2013 Proceedings*.

[6] Heyne, M., Derrick, D. 2015. Using a radial ultrasound probe's virtual origin to compute midsagittal smoothing splines in polar coordinates. *Journal of the Acoustical Society of America Express Letters* 138(6), EL509–EL514.

[7] Hsu, H.-J., Tse, J. K.-P. 2007. Syllable-final nasal nergers in Taiwan Mandarin-leveled but puzzling. *Concentric: Studies in Linguistics* 33(1), 1–18.

[8] Kubler, C. C. 1985. The influence of Southern Min on the Mandarin of Taiwan. *Anthropological Linguistics* 27(2), 156–176.

[9] Lloyd, J. E., Stavness, I., Fels, S. 2012. Artisynth: A fast interactive biomechanical modeling toolkit combining multibody and finite element simulation. In: Payan, Y., (ed), *Soft Tissue Biomechanical Modeling for Computer Assisted Surgery* volume 11. Springer 355–394.

[10] Styler, W. 2015. *On the Acoustical and Perceptual Features of Vowel Nasality*. PhD thesis University of Colorado at Boulder.

[11] Zhang, J. 2000. Non-contrastive features and categorical patterning in chinese diminutive suffixation: Max [F] or Ident [F]. *Phonology* 17(3), 427–478.