Revisiting the Nasal Coda Merger in Taiwanese Mandarin: The Merging Pattern in the /iN/ context in northern Taiwan

Chang-Lo Chiang 15493636

University of Amsterdam

Master's Thesis: General Linguistics

Supervisor: Silke Hamann

July 19th , 2026

Abstract

This study investigates the nasal coda merger in Taiwanese Mandarin, focusing on the demographic subgroup of young native speakers from northern Taiwan. While traditional measurements are centered around the acoustic correlates of the preceding vowel, this study examines the acoustic correlates of the nasal codas to identify the merging pattern in the /iN/ context. The results indicate that the nasal codas in the /iN/ context merge into a single category that is acoustically similar to the canonical alveolar [n] of /an/, rejecting the hypothesis of the velar pattern. High token exclusion rates for specific lexical items further reveal advanced consonantal reduction and vowel nasalization among northern speakers.

1. Introduction

Out of the three nasal consonants /m, n, ŋ/ in the phonetic inventory of Mandarin, only the alveolar nasal /n/ and the velar nasal /ŋ/ are allowed to appear in the coda positions (Dong, 2024). A merger of these two nasal codas has been reported in several Mandarin varieties, including Shanghai Mandarin and Taiwanese Mandarin (Faytak et al., 2020). In these varieties, the place contrast between the nasals is neutralized, while the contrast is maintained in other non-merging varieties such as Standard Mandarin (Faytak et al., 2020). Among the varieties that exhibit this merger, Taiwanese Mandarin has attracted considerable attention. Dating back to Kubler (1985), studies spanning decades have extensively documented the phenomenon (among many others, Fon et al., 2011; Hsu & Tse, 2007; Lai, 2009; Yang, 2007). Previous studies have reported varying degrees of merger among Taiwanese Mandarin speakers, raising questions about the current status of the merger and the factors that influence it. Addressing these questions requires consideration of the linguistic features of Taiwanese Mandarin, which may potentially shape the realization of the nasal codas. Taiwanese Mandarin (TM), often referred to as Taiwan Mandarin in the literature, is a variety of Mandarin spoken in Taiwan. Despite TM being mutually intelligible with Standard

Mandarin (SM) spoken in China, the two varieties differ in phonetics, phonology, syntax, lexicon, and pragmatics (Chen & Herring, 2024). Since they are historically derived from the same origin, TM and SM retain largely similar phonological systems (Torgerson, 2005). However, TM has developed distinctive features through contact with Taiwanese Hokkien, another tonal Sinitic language spoken in Taiwan (Lee-Kim & Chou, 2022). The observation that TM speakers produce tones in a lower register, resulting in lower tone heights and more conservative tonal contours, has been attributed to the influence of Taiwanese Hokkien, which is noted for its low tonal register (Fon & Chiang, 1999; Torgerson, 2005). In addition, due to the influence of Taiwanese Hokkien whose phonetic inventory lacks retroflex segments, TM speakers are reported to frequently deretroflex sibilants, reducing the contrast between the alveolar /ts, ts^h, s/ and the retroflex /$\widehat{ʈʂ}$, $\widehat{ʈʂ^h}$, ʂ/ sibilants (Lee-Kim & Chou, 2022). Another frequently discussed feature in the literature is the merger of nasal codas, which is the focus of the current study.

1.1 Previous Research On The Merger

Studies investigating the merger in TM have shown that it is vowel sensitive, which means the merging patterns vary across vowel contexts. The nasals only form minimal pairs when the preceding vowel is the high vowel /i/, the mid vowel /ə/, or the low vowel /a/ (Dong, 2024), and the merger has consistently been reported to be absent in the /aN/ context (Chiu et al., 2025; Chuang, 2017; Hsu & Tse, 2007; Lai, 2009; Yang, 2007), where N stands for the nasal codas. The absence of the merger in the /aN/ context is not arbitrary, as low vowels are more likely to accommodate the articulatory requirements of adjacent segments (Faytak et al., 2020). Acoustic studies suggest that the place contrast between the two nasals is largely neutralized in the /iN/ and /əN/ contexts. One frequently examined acoustic correlate is the vowel formants, particularly the second formant (F2). Since F2 reflects the tongue frontness/backness during the production of the vowel, the F2 transitions from the vowel to the nasal are expected to diverge between the alveolar /n/ and the velar /ŋ/ (Hsu & Tse, 2007). Focusing on the F2 values at the boundary between the vowel and the nasal

coda, Lai (2009) analyzed production data from 20 native TM speakers and observed no significant differences between the two nasals in the /iN/ and /əN/ contexts, whereas a significant difference was observed in the /aN/ context. In the extensive corpus study conducted by Chuang (2017), the F2 values were measured at the start, the midpoint, and the end of the vowel. The results showed that the F2 values at the end of the vowel were statistically indistinguishable in the /əN/ context, whereas the F2 value of /iŋ/ was approximately 120 Hz lower than that of /in/, and a slightly greater difference of approximately 170 Hz was observed in the /aN/ context. However, the opposite pattern was observed in a recent study conducted by Chiu et al. (2025) that also measured the F2 values at several time points across the vowel. They observed a small but consistent difference in the /əN/ context, whereas no such difference was observed in the /iN/ context, although the details of the statistical analyses regarding this measure were not elaborated in their study. Similar to the acoustic studies, ultrasound studies have provided extensive evidence for the merger. A series of ultrasound studies reported varying degrees of gestural overlap across vowel contexts. It is further confirmed that the merger is absent in the /aN/ context, where tongue postures for /an/ and /aŋ/ remain distinct and contrastive (Chiu & Lu, 2021; Chiu et al., 2019; Chiu et al., 2025). In the /iN/ context, the tongue postures are reported to be completely merged, with the tongue tip lowered and the tongue body retracted and bunched further back in the oral cavity, resembling the tongue posture for the velar nasal /ŋ/ (Chiu & Lu, 2021; Chiu et al., 2019; Chiu et al., 2025). In contrast, the tongue posture merger in the /əN/ context appears to be inconsistent and speaker dependent. While some speakers maintained distinct tongue postures between the two nasal codas, many exhibited overlapping tongue postures. Among speakers with merged tongue postures, the merging patterns varied. Some positioned their tongues with an elevated tongue tip and contact at the alveolar ridge, resembling the posture for the alveolar /n/, whereas others positioned their tongues with a lowered tongue tip and retracted and bunched tongue body, resembling the posture for the velar /ŋ/ (Chiu & Lu, 2021; Chiu et al., 2019; Chiu et al., 2025). Despite showing inconsistent results, the acoustic studies and ultrasound studies together have provided converging evidence that

the place contrast between the two nasals has been substantially reduced in TM. The evidence suggests that the merger is more advanced in the /iN/ and /əN/ contexts, while remaining absent in the /aN/ context.

The reduction of the place contrast increases the relevance of several compensatory cues to the merger, and previous studies have suggested that some of these cues maintain contrastive in the merging /iN/ and /əN/ contexts. One such compensatory cue that is frequently discussed in the literature is the degree of nasalization of the preceding vowel. The degree of nasalization is obtained by measuring the A1-P0 for the non-high vowel /ə/, defined as the amplitude difference between F1 and the first nasal peak below F1, and measuring the A1-P1 for the high vowel /i/, defined as the amplitude difference between F1 and the nasal peak between F1 and F2 (Chiu & Lu, 2021). Regarding the degree of nasalization of the vowel, studies have presented mixed results. In the /iN/ context, Lai (2009) reported stronger vowel nasalization in the alveolar /in/ than in the velar /iŋ/, whereas other (Chuang, 2017; Chiu & Lu, 2021; Chiu et al., 2025) reported the opposite pattern, with a stronger vowel nasalization in /iŋ/. In the /əN/ context, Lai (2009) and Chiu & Lu (2021) reported stronger vowel nasalization in /ən/ than in /əŋ/, whereas Chuang (2017) reported the opposite pattern, with a stronger vowel nasalization in /əŋ/. Similarly, Chiu et al. (2025) also reported a slightly stronger vowel nasalization in /əŋ/ than in /ən/, although the difference was not statistically significant. The vowel duration and nasal duration have also been frequently examined as potential contrastive cues. Regarding the vowel duration, Lai (2009) observed no significant differences across vowel contexts, whereas Chuang (2017) reported longer vowels in /in/ than in /iŋ/ and longer vowels in /əŋ/ than in /ən/. In terms of the nasal duration, Lai (2009) observed no significant differences across vowel contexts. Chuang took a different approach by examining the ratio of nasal duration to syllable duration and reported a greater ratio for the velar /ŋ/ across vowel contexts, although it was not significant in the /əN/ context. In comparison, Chiu et al. (2025) examined the ratio of nasal duration to rime duration and reported no significant differences

between the nasals in both /iN/ and /əN/ contexts. Although these studies have demonstrated that some acoustic cues remain distinct in the merging /iN/ and /əN/ contexts, these acoustic distinctions are not well perceived by TM listeners. A follow-up study to the ultrasound study conducted by Chiu et al. (2025) investigated the listeners' accuracy of identifying the two nasal codas based on the compensatory cues produced by speakers with the merger. The study reported ceiling accuracy of 98% success in identifying the nasals in the non-merging /aN/ context, compared with the significantly lower accuracies of 60% in the /iN/ context and 71% in the /əN/ context. Moreover, an asymmetry was observed between the /iN/ and /əN/ contexts. In the /iN/ context, accuracy of successfully identifying the alveolar /n/ was 7% higher than that of the velar /ŋ/, whereas in the /əN/ context, the accuracy of successfully identifying /n/ was 5% lower than that of /ŋ/. Taken together, these findings indicate that, despite the place contrast being reduced, many TM speakers continue to preserve measurable acoustic distinctions between the two nasal codas. Nevertheless, the reported mismatch between production and perception suggests that the merger is incomplete and is rather closer to a near-merger.

1.2 Observed Issues in the Literature

While previous studies primarily focused on providing evidence for the merger, some of them have gone further to propose specific merging patterns in the /iN/ and /əN/ contexts. However, the merging patterns remain unclear since studies have reported inconsistent and contradicting findings. Theoretically, there are three possibilities for the merging patterns: the two nasals merge into the alveolar [n], the two nasals merge into the velar [ŋ], and the two nasals merge into a third distinct category. In the /iN/ context, Yang (2007) proposes the alveolar pattern of /iŋ/ being more likely to be realized as [in], whereas others (Chiu & Lu, 2021; Chiu et al., 2025; Hsu & Tse, 2007) propose the opposite velar pattern of /in/ being more likely to be realized as [iŋ]. The same discrepancy exists in the /əN/ context, with most studies reporting the alveolar pattern of /əŋ/ merging into [ən] (Chiu et al., 2025; Hsu & Tse, 2007; Yang, 2007), where Chiu & Lu (2021) reported both patterns,

suggesting that the merging pattern might be speaker dependent in this vowel context. Based on the available literature, no study has proposed that the nasals merge into a third distinct category in any vowel context. An overview of the proposed merging patterns is presented in Table 1. A notable asymmetry could be observed in that the suggestions for the velar pattern in the /iN/ context and the alveolar pattern in the /əN/ context are preferred over the alternative patterns. These discrepancies might be a consequence of differences in research focus and methodology. As discussed earlier, some studies employed acoustic analyses that examined various acoustic correlates of the preceding vowel and the nasal coda, whereas others used ultrasound imaging to examine the tongue postures. Given the observation that the merger is incomplete and is rather closer to a near-merger, these differing focuses and methodologies could lead to substantially different interpretations and conclusions. One potentially problematic methodology is the use of velar pinch as the primary criterion to identify the merging patterns in the study conducted by Hsu & Tse (2007). Velar pinch is defined as a rising F2 and a falling F3 at the end of the vowel that serves as a cue for the realization of a following velar segment (Hsu & Tse, 2007). In principle, the presence of a velar pinch should be a reliable indicator of the velar nasal /ŋ/. However, Chiu et al. (2025) observed a falling F2 at the end of the vowel in /aŋ/ in the non-merging /aN/ context, indicating the absence of a velar pinch in the canonical realization of the velar nasal in TM. This finding suggests that velar pinch might not be a reliable indicator of the velar nasal in TM, implying that any conclusions based primarily on this criterion should be interpreted with caution. Another recurring methodological limitation across the acoustic studies is the underdetermination of the acoustic measurements. Most studies reviewed above examined the acoustic correlates of the vowel, including the formants, duration, and degree of nasalization. However, the interpretations of these measures are not straightforward due to the absence of an established baseline. Consider F2 transitions as an example. In the canonical realizations of both /in/ and /iŋ/, F2 is expected to fall at the end of the vowel due to the high-front vowel /i/ being articulated further front than the alveolar nasal, with the F2 falling approximately 300 Hz more in [iŋ] than [in] (Faytak et al., 2020). Since

vowel formants have been reported to be statistically identical between the nasals in the merging /iN/ context in TM (Chiu et al., 2025), an observed F2 fall in the merging realization does not unambiguously indicate whether it reflects the smaller fall of [in] or the greater fall of [iŋ]. Due to the absence of historical data on the canonical non-merging realizations in the /iN/ and /əN/ contexts in TM, this methodological challenge remains to be resolved.

Table 1

The Merging Patterns Proposed Across Studies

Vowel context	/n, ŋ/ → [n]	/n, ŋ/ → [ŋ]
/iN/	Yang (2007)	Chiu & Lu (2021) Chiu et al. (2025) Hsu & Tse (2007)
/əN/	Chiu & Lu (2021) Chiu et al. (2025) Yang (2007)	Chiu & Lu (2021)

Previous studies have suggested that the merger is incomplete, implying that the merger is an ongoing and dynamic process that subjects to new emerging patterns. The discrepancies in the findings across previous research might also reflect limited knowledge regarding demographic variations. All studies reviewed so far have primarily focused on the general population of TM speakers instead of specific demographic subgroups. However, social factors such as age, gender, and region have been reported to influence the merger. The merger has been proposed as an innovation of younger generations (Tse, 1992) and has been reported to be more prevalent among

female speakers than among male speakers (Fon et al., 2011). In the study conducted by Fon et al. (2011), a regional difference was reported in that the alveolar merging pattern in the /iN/ context (i.e., /in, iŋ/ to [in]) was only attested among speakers from southern Taiwan who were heavily influenced by Taiwanese Hokkien. However, a recent study (Lee-Kim et al., 2024) observed the same alveolar merging pattern among young speakers from northern Taiwan. They reported approximately 4% of the production data from female participants and approximately 7% from male participants from northern Taiwan exhibited the alveolar pattern, whereas approximately 70% of the data across genders exhibited the velar pattern. The merging pattern in the /iN/ context among speakers from northern Taiwan remains unclear, since Fon et al. did not specify the criteria used to identify the merging patterns, and Lee-Kim et al. (2024) relied on the judgements of trained phoneticians rather than objective acoustic measures.

A methodological limitation is observed in the literature, due to the absence of an established baseline for identifying the merging patterns through measurements of the acoustic correlates of the vowels. Consequently, inconsistent and contradicting findings have left the merging patterns unclear, particularly regarding the reported regional differences. These issues together motivate the focus of the current study, which investigates the merging pattern in the /iN/ context among young TM speakers from northern Taiwan and whether the pattern could be identified through measurements of the acoustic correlates of the nasal coda instead of the preceding vowel. Based on previous findings, it is hypothesized that young TM speakers from northern Taiwan will exhibit the merger in the /iN/ context, but not in the /aN/ context. Furthermore, the merging pattern in the /iN/ context is expected to be the velar pattern (i.e., /in, iŋ/ to [iŋ]) due to this pattern being reported most frequently in the literature.

2. Methods

2.1 Participants

13 native speakers of Taiwanese Mandarin (TM) residing in the Netherlands were initially recruited for the experiment. 10 of the participants were female and three were male. At the time of participation, all participants were under the age of 31 (M = 26.8). A background interview was conducted with each participant to verify that their place of origin and linguistic profile aligned with the focus of this study. All participants originated from northern Taiwan, which is defined as the administrative regions of Taipei, New Taipei, Keelung, Taoyuan, Hsinchu city/county, and Yilan. 12 participants had not lived outside of northern Taiwan for more than six months before the age of 18 (eight lived in Taipei, two in New Taipei, one in Taipei and New Taipei, and one in Taipei and Taoyuan). One participant was excluded from the study since she had spent the first eight years of her life in Shanghai, China, and this experience was suspected of having influenced her acquisition of TM during early childhood. This exclusion resulted in a sample of 12 participants. Information of the participants' linguistic profiles was collected regarding their native language(s), additional language(s) learned, and self-reported proficiency in those language(s). All participants identified TM as their only native language. All participants identified Taiwanese Hokkien (TH) as a second language, due to exposure through family members who are native speakers, which is common for Taiwanese people. Any degree of proficiency in TH is not expected to have an influence on the merger in TM for the participants, since previous research has reported the merger to be an innovation of younger generations who are not heavily influenced by TH (Chiu & Lu, 2021). All participants also identified English as a second language, which is introduced in the Taiwanese education system starting at the age of six. Any degree of proficiency in English is also not expected to have an influence on the merger, since the same educational background applies to the participants from previous research, and no effect of this factor has been reported. All participants reported no hearing impairments and no history of language disorders. The study was approved by

the Ethics Committee of the Faculty of Humanities of the University of Amsterdam, and participants received monetary compensation in accordance with the committee's guidelines.

2.2 Procedure and Stimuli

The experiment was conducted in the speech lab of the University of Amsterdam. Participants were seated in a sound-treated room and recorded using a Neumann TLM103 cardioid condenser microphone, into which the participants were instructed to speak. A self-paced reading task was employed. Participants were instructed to read a list of TM sentences in the most comfortable and natural manner, approximating their speech in a spontaneous conversation while avoiding any elision. Each sentence was read three times before participants proceeded to the next item. All sentences consisted of nine syllables, and the target stimuli were embedded in either the sentence-initial or the sentence-medial position. The target stimuli were disyllabic words, with the nasal codas occurring in the second syllable. Although no effect of tone on the merger has been reported in the literature, all target syllables carried the first tone (i.e., the high-level tone) to minimize potential confounding effects from this factor. The target syllables had the structure CVN, where C was any consonant, V was either the high vowel /i/ or the low vowel /a/, and N was either the alveolar nasal /n/ or the velar nasal /ŋ/. For each vowel and nasal combination, three lexical items were included on the list, resulting in 36 tokens recorded per participant (2 vowels × 2 nasals × 3 lexical items × 3 repetitions). The sentences containing target stimuli and filler sentences were presented to all participants in the same randomized order. On average, it took approximately seven minutes for participants to complete the reading task. The complete list of target stimuli is provided in Appendix A.

2.3 Measurements and Data Preparation

As noted in the introduction, the current study focused on the acoustic correlates of the nasal codas instead of the vowels in order to investigate whether these measurements could assist the

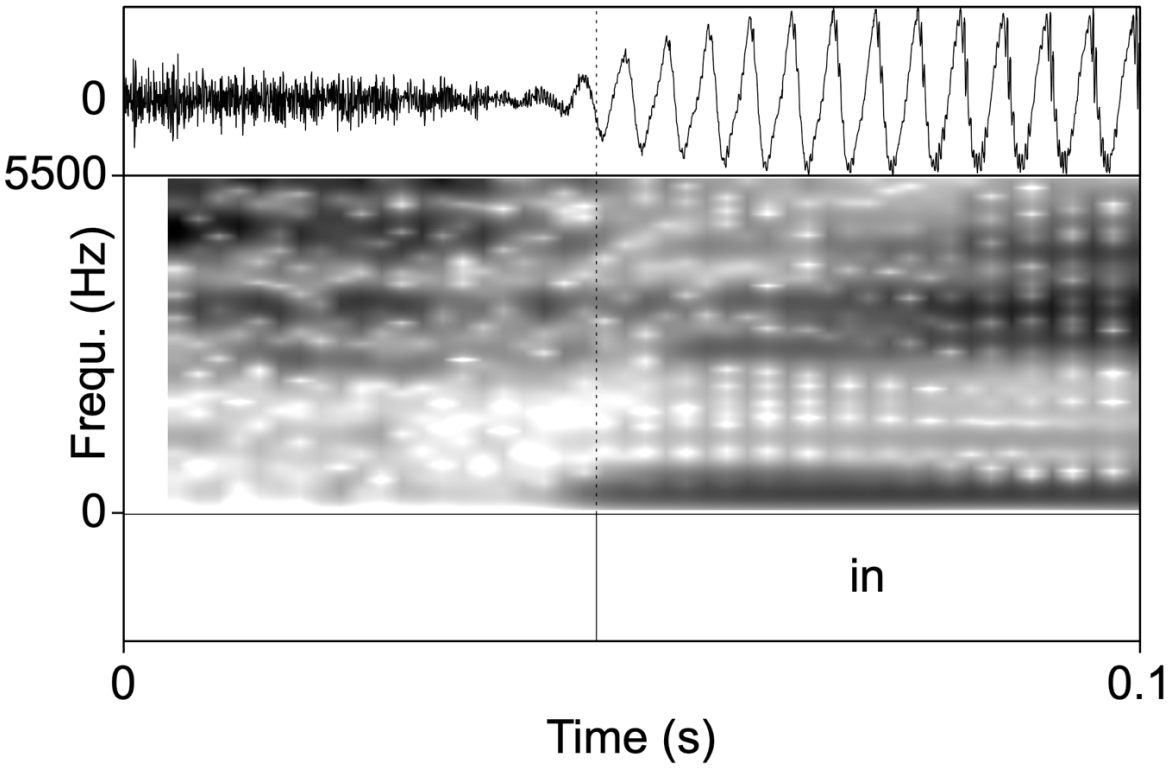
identification of the merging pattern in the /iN/ context. The acoustic correlates measured were the F1 and F2 values of the nasal, nasal duration, and the ratio of nasal duration to rime duration (nasal/rime ratio). In addition, the F2 values at the end of the vowel were also measured as to verify the merger in the /iN/ tokens compared to the non-merging /aN/ control tokens.

The recordings were manually annotated in Praat (Boersma & Weenink, 2025) to identify the boundaries of the vowel and nasal segments in the target syllables. Praat displays both the waveforms and spectrograms of imported audio files, and both were taken into consideration simultaneously when identifying the segment boundaries. Because the onsets of the target syllables included aspirated segments, the aspiration of these segments could persist into the vowel and interfere with the waveforms, resulting in irregular patterns. An example of such interference is presented in Figure 1, where the waveforms at the start of the vowel appears jagged and irregular. Thus, the start of the vowel was defined as the time point at which the waveforms began to exhibit stable, periodic patterns without noticeable irregularities, while clear formant structures became visible in the spectrogram. Similarly, the boundary between the vowel and the nasal was defined as the time point at which the waveforms ceased to exhibit periodic patterns of the vowel and began to exhibit periodic patterns of the nasal, while the formant structures of the vowel in the spectrogram diminished to the point of no longer being visible. Analogous to the vowel being affected by the onset, the nasal coda could also be affected by the onset of the following syllable. Consequently, the end of the nasal was defined as the time point at which the waveforms ceased to exhibit stable, periodic patterns without noticeable irregularities, while clear formant structures were no longer visible in the spectrogram.

During annotation, several tokens were excluded due to that the vowel and nasal segment could not be reliably demarcated according to the criteria described above. Two tokens of /in/ (one from P01, and one from P04) and one token of /iŋ/ (from P11) were excluded due to the vowel formants not

Figure 1

An Example of the Aspiration in the Onset Interfering with Vowel Soundwaves



being sufficiently visible to identify the boundaries. An example of these excluded tokens is presented in Figure 2. Eight tokens of /in/ (one from P02, one from P03, two from P04, three from P06, and one from P10) and one token of /an/ (from P05) were excluded due to the vowel formants remaining visible until the end of the syllable, which resulted in the vowel and nasal being indistinguishable in both the waveforms and the spectrogram. In these tokens, the nasal coda appeared to be reduced from an independent segment to the complete or partial nasalization of the vowel. An example of these excluded tokens is presented in Figure 3. One /an/ token (from P02) was excluded due to the substantial overlap between the nasal coda and the following velar fricative /h/, making the two segments indistinguishable in both the waveforms and the spectrogram. One /in/ token (from P11) was excluded due to the periodic patterns of the vowel formants diminishing abruptly and reappearing shortly after with the same patterns. This was suspected to be a result of a recording equipment malfunction. In total, 11 /in/ tokens, one /iŋ/ token, and two /an/ tokens were

Figure 2

An Example of an Excluded Token Due to the Vowel Formants not Being Visible Enough in the Spectrogram

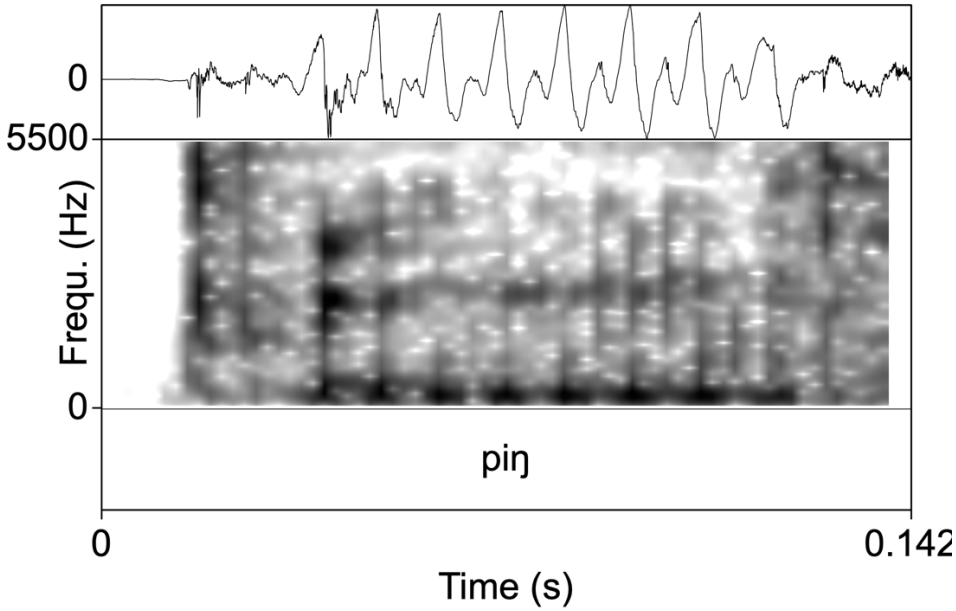
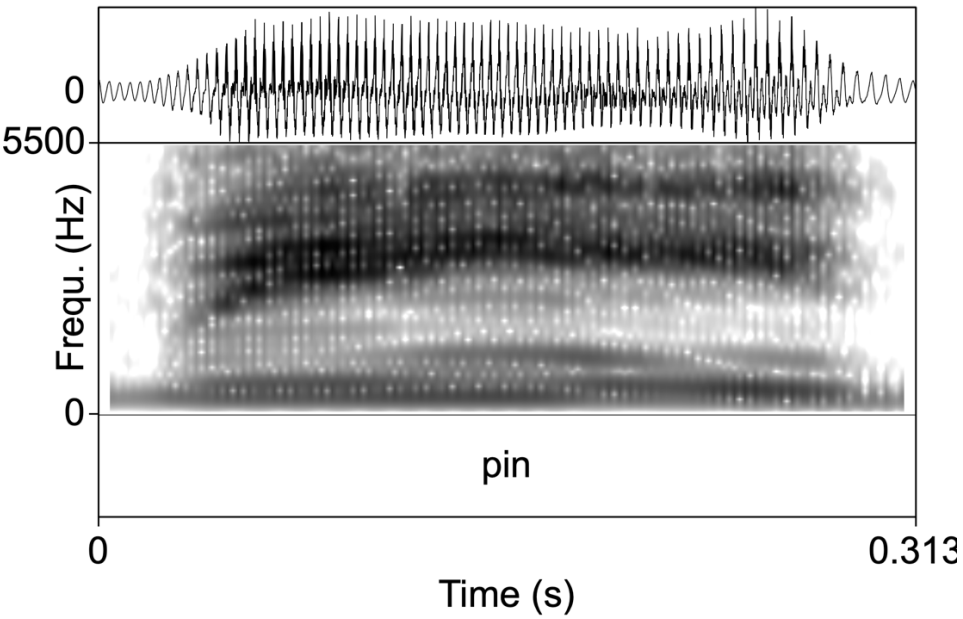


Figure 3

An Example of an Excluded Token Due to the Vowel Formants Lasting All the Way to the End of the Syllable



excluded, leading to an exclusion rate of 2.5%. The numbers of excluded tokens per lexical item are presented in Table 2.

Table 2

Numbers of Excluded Tokens Per Lexical Items

Lexical item	Number
/pin/	8
/tɕʰin/	3
/piŋ/	1
/tan/	2

Measurements of the vowel and nasal segments were obtained using a Praat script. The formant ceiling was set to 5,500 Hz for female participants and 5,000 Hz for male participants. The script employed the formant tracker function built in Praat to extract the F2 values at the end of the vowel (V-F2) and the F1 and F2 values of the nasal (N-F1 and N-F2). The V-F2 value was averaged over the duration of the end fourth of the vowel, and the N-F1 and N-F2 values were averaged over the entire duration of the nasal. However, it is important to point out that the formant tracker in Praat posed a substantial challenge for the current study. A close inspection of the waveforms and spectrograms would reveal that the nasal formants were often visible and stable during the vowel, indicating some extent of vowel nasalization. This was particularly evident when the measurement was taken at the end of the vowel, when nasal formants inevitably emerged as a result of coarticulation. Due to nasal formants typically having low frequencies, Praat could potentially mistake the nasal formants during the vowel as the F1 or F2 of the vowel. Thus, the F2 values

extracted could represent a mixture of the true F2 of the vowel and the nasal formants during the vowel. This resulted in a systematic variation and a wide range of F2 values across participants in both the /iN/ and /aN/ tokens. This issue reflects a broader challenge in the acoustic analysis of vowels followed by nasals, where coarticulatory nasalization in the vowel could complicate formant tracking. Given the absence of a reliable method for solving this issue within the scope of this study, the obtained measurements were not further adjusted. Therefore, the observed variation in the F2 values should be regarded as an inherent limitation of the dataset, and the results reported in the following section should be interpreted with caution.

2.4 Statistical Analysis

All statistical analyses were conducted in R (R Core Team, 2025). Several linear mixed-effects regression models were fitted to examine the proposed acoustic correlates. The fixed effects included nasal place, vowel context, and the interaction between these two factors. The categorical predictors were sum coded, with /ŋ/ coded as -0.5 and /n/ coded as +0.5, and with /a/ coded as -0.5 and /i/ coded as +0.5. Repetition was included as a covariate and was centered by subtracting the mean value (i.e., 2) from each observation. Random intercepts were included for participants and lexical items to account for multiple observations per participant per lexical item.

The first model examined the V-F2 values as a sanity check. As discussed in the introduction, a difference between /an/ and /aŋ/ tokens was expected, whereas the same was not expected for /in/ and /iŋ/ tokens. Subsequently, separate models were fitted for each proposed nasal correlate to identify which correlate maintained a contrast between the two nasals in the /aN/ context. For the correlate that maintained contrastive, two follow-up analyses were conducted. The first compared /iN/ tokens with /an/ tokens, and the second compared /iN/ tokens with /aŋ/ tokens. These analyses were designed to determine whether the potentially merged nasal in the /iN/ tokens was acoustically distinct from the alveolar /n/, the velar /ŋ/, or both.

3. Results

A total of 421 tokens produced by 12 participants were analyzed. The mean values for each measurement are presented in Table 3 and Table 4, separated by gender. The results of the /iN/ tokens will be described first, followed by the results of the /aN/ tokens. For both females and males, the mean V-F2 values showed minimal differences of less than 30 Hz between the two nasals in the /iN/ tokens. This was expected, as previous research has reported identical vowel formants in the /iN/ context. Similarly, the mean N-F2 values showed minimal differences of less than approximately 50 Hz in the /iN/ tokens for both genders. The mean N-F1 value was 61.46 Hz lower in the /in/ tokens than in the /iŋ/ tokens for females, and the same pattern was observed for males, despite a smaller difference of 26.15 Hz. The mean nasal duration showed a minimal difference of less than 1 ms in the /iN/ tokens for females, while the mean duration was 7.83 ms longer in the /in/ tokens than in the /iŋ/ tokens for males. In contrast, the mean nasal/rime ratio showed minimal differences of less than 0.02 in the /iN/ tokens for females, while the mean ratio was approximately 0.04 greater in the /iŋ/ tokens than in the /in/ tokens for males.

For both females and males, the mean V-F2 value was approximately 400 Hz higher in the /an/ tokens than in the /aŋ/ tokens. This was expected, since previous research has reported a rising F2 toward the end of the vowel in /an/ and a falling F2 in /aŋ/. The mean N-F1 value was 148.3 Hz lower in the /an/ tokens than in the /aŋ/ tokens for females, and the same pattern was observed for males, with a greater difference of 200.87 Hz. While the mean N-F2 value was 168.9 Hz higher in the /an/ tokens than in the /aŋ/ tokens for females, the same was not observed for males, whose mean N-F2 values showed a minimal difference of approximately 70 Hz. For both females and males, the mean nasal duration showed minimal differences of less than 3 ms in the /aN/ tokens, and the mean nasal/rime ratio also showed minimal differences and less than 0.01. Overall, females exhibited longer mean nasal duration and greater nasal/rime ratio than males in the /iN/ tokens, whereas this pattern was not observed in the /aN/ tokens.

Table 3

The Means of the Acoustic Correlates from Female Participants

Acoustic correlate	/i/		/a/	
	/n/	/ŋ/	/n/	/ŋ/
V-F2	1772.38	1800.75	1737.37	1320.91
N-F1	351.42	412.88	463.52	611.82
N-F2	1601.71	1551.61	1550.59	1381.69
Duration	71.57	70.93	57.03	56.91
Ratio	0.4345	0.4504	0.3357	0.3431

Note. The formant values were in Hz, and the duration was in ms.

The model examining V-F2 values showed significant main effects of nasal place and vowel context, as well as a significant interaction effect between these two factors. The mean V-F2 value of the /n/ tokens was significantly higher than that of the /ŋ/ tokens, $b = 183.61$, $SE = 43.36$, $t(31.38) = 4.23$, $p < 0.001$, 95% CI [101.54, 265.78], which is expected due to the correlation between F2 values and tongue backness. The mean V-F2 value of the /i/ tokens was significantly higher than that of the /a/ tokens, $b = 235.95$, $SE = 43.36$, $t(31.39) = 5.44$, $p < 0.001$, 95% CI [153.87, 318.11], which is expected due to /i/ inherently having a higher F2 value than /a/. The interaction effect indicates that the effect of nasal place was significantly greater in the /aN/ context than in the /iN/ context, $b = 441.17$, $SE = 86.73$, $t(31.39) = 5.09$, $p < 0.001$, 95% CI [276.82,

Table 4

The Means of the Acoustic Correlates from Male Participants

Acoustic correlate	/i/		/a/	
	/n/	/ŋ/	/n/	/ŋ/
V-F2	1535.52	1555.03	1526.69	1173.39
N-F1	420.89	447.04	443.39	644.26
N-F2	1650.35	1626.98	1521.57	1587.53
Duration	65.56	57.73	54.16	56.84
Ratio	0.3589	0.3954	0.3587	0.3599

Note. The formant values were in Hz, and the duration was in ms.

605.29]. This result is consistent with previous research that the two nasals remain more distinct in the non-merging /aN/ context than in the merging /iN/ context. Although the model does not directly indicate whether the nasals were acoustically identical between /in/ and /iŋ/ tokens, the significant interaction effect provides evidence that the merger is more advanced in the /iN/ context.

The next series of models examined the nasal correlates. The model that examined N-F1values showed significant main effects of nasal place and vowel context, with no interaction effect. The mean N-F1 value of the /ŋ/ tokens was significantly higher than that of the /n/ tokens, b = 103.66,

SE = 41.04, *t*(31.28) = 2.53, *p* = 0.017, 95% CI [26.37, 181.01]. This effect of nasal place indicates that the two nasal categories differ reliably in N-F1 values, suggesting that this correlate is sufficiently informative to be included in the follow-up analyses that aimed at identifying the merging pattern of /iN/ tokens. The mean N-F1 values of the /a/ tokens was significantly greater than that of the /i/ tokens, *b* = 141.97, *SE* = 41.04, *t*(31.38) = 3.46, *p* = 0.002, 95% CI [64.67, 219.32]. This effect of vowel context indicates that the preceding vowel could potentially influence the acoustic properties of the following nasal segment. The model that examined N-F2 values showed neither a main effect nor an interaction effect. This finding suggests that N-F2 values do not reliably distinguish the two nasal categories and is therefore not sufficiently informative to be included in the follow-up analyses. The model that examined the nasal duration showed a significant main effect of vowel context. The nasal duration of the /i/ tokens was significantly longer than that of the /a/ tokens, *b* = 11.59, *SE* = 3.65, *t*(31.29) = 3.173, *p* = 0.003, 95% CI [4.68, 18.50]. This effect of vowel context indicates that the nasal duration differs across vowel contexts, which might be a consequence of different vowel duration. However, since no effect of nasal place nor an interaction effect was observed, the nasal duration is not sufficiently informative to be included in the follow-up analyses. The model that examined the nasal/rime ratio showed a significant main effect of vowel context. The nasal/rime ratio of the /i/ tokens was significantly greater than that of the /a/ tokens, *b* = 0.0793, *SE* = 0.0179, *t*(31.36) = 4.43, *p* < 0.001, 95% CI [0.0455, 0.113]. This effect of vowel context indicates that the proportion of the rime occupied by the nasal segment differs across vowel contexts, potentially reflecting the differences in nasal duration reported earlier. However, no effect of nasal place nor interaction was observed, suggesting that the nasal/rime ratio does not reliably distinguish the two nasal categories. Consequently, it was not included in the follow-up analyses.

Of the four proposed nasal correlates, only N-F1 showed a significant nasal place effect, making it the only informative correlate included in the follow-up analyses. Furthermore, the results of V-F2

reported earlier support the interpretation that the two nasals were merging in the /iN/ tokens. Therefore, the two nasals in the /iN/ tokens could be regarded as one unidentified nasal category. To further examine this category, two separate models were fitted to compare the N-F1 of the unidentified nasal with the nasals in the /an/ and /aŋ/ tokens, respectively. The first model compared the /iN/ tokens with the /an/ tokens, which showed no effect of vowel context. In contrast, the second model that compared the /iN/ tokens with the /aŋ/ tokens showed a main effect of vowel context. The mean N-F1 value of the /iN/ tokens was significantly lower than that of the /aŋ/ tokens, $b = 221.6$, $SE = 39.11$, $t(24.11) = 5.67$, $p < 0.001$, 95% CI [145.82, 297.48]. This pattern was already apparent from the descriptive statistics presented in Table 3 and 4, where the mean N-F1 value of the /iN/ tokens closely resembled that of the /an/ tokens but greatly differed from that of the /aŋ/ tokens. The presence of the main effect of vowel context in the /iN/ vs. /aŋ/ model indicates that the N-F1 value of the unidentified nasal in the /iN/ tokens was statistically different from the velar nasal in the /aŋ/ tokens. Conversely, the absence of a main effect in the /iN/ vs. /an/ model should be interpreted with caution. A non-significant result could only indicate that the unidentified nasal in the /iN/ tokens did not significantly differ from the alveolar nasal in the /an/ tokens, whereas it did not indicate whether the two were statistically identical to each other. Nevertheless, the main effect observed in the /iN/ vs. /aŋ/ model is sufficient to conclude that the results reported here contradict with previous research, which has suggested that the merged nasal in the /iN/ context is more similar to the velar [ŋ]. As a consequence, the hypothesis proposed in this study, which is based on these earlier suggestions, is not supported.

4. Discussion

The primary objective of this study was to revisit the nasal coda merger in Taiwanese Mandarin (TM) from an approach that has received relatively little attention in the literature, focusing on the acoustic correlates of the nasals rather than relying exclusively on the correlates of the vowels. In addition, this study targeted a specific demographic subgroup, the young TM speakers from

northern Taiwan, to provide empirical evidence to resolve the contradicting merging patterns reported in the literature. The analysis of V-F2 confirmed the previously established vowel sensitive nature of the merger. Consistent with decades of research spanning from Kubler (1985) to the recent study by Chiu et al. (2025), the place contrast between the alveolar /n/ and the velar /ŋ/ remains robust and acoustically distinct in the /aN/ context. For both female and male participants, the V-F2 values of the /an/ tokens were approximately 400 Hz higher than those of the /aŋ/ tokens, reflecting the expected divergence in tongue postures between a fronted posture for the alveolar and a retracted posture for the velar. Conversely, in the /iN/ context, the V-F2 difference between the two nasals was minimized, exhibiting a mean difference of less than 30 Hz for both genders. This finding provides a successful sanity check, confirming that the recruited participants indeed exhibit a substantial neutralization of the acoustic cues traditionally used to define the place contrast in the /iN/ context. In addition, the investigation of acoustic correlates of the nasal codas revealed that not all acoustic correlates perform equally well as indicators of nasal place. Among the four acoustic correlates of the nasal examined, namely the N-F1, N-F2, nasal duration, and nasal/rime ratio, only N-F1 reliably distinguished the alveolar /n/ and the velar /ŋ/ in the /aN/ context. For both female and male participants, the N-F1 values were significantly lower in the alveolar /an/ tokens than in the velar /aŋ/ tokens. Neither N-F2, nasal duration, nor the nasal/rime ratio consistently distinguished the two nasals. This finding establishes N-F1 as an acoustic correlate that serves as a viable indicator of nasal place in the merger in TM. Furthermore, in the follow-up analyses where the validated correlate N-F1 was examined to identify the category of the merged nasal in the /iN/ context, the results revealed that the merged nasal significantly differed from the velar [ŋ] and did not significantly differ from the alveolar [n]. This contrast was apparent in the group means where the mean N-F1 values of the /iN/ tokens for both female and male participants aligned closer to the mean N-F1 of the /an/ tokens (463.5z Hz) than to the higher mean of the /aŋ/ tokens (611.82 Hz). This finding contradicts with the prediction derived from previous studies that the category of the merged nasal is acoustically closer to the velar /ŋ/ than the alveolar /n/ in the /iN/ context.

Consequently, the hypothesis that young TM speakers from northern Taiwan favor the velar pattern in the /iN/ context (i.e., /in, iŋ/ to [iŋ]) is rejected.

The most important debates in the literature of the nasal coda merger in TM concerns the nature of the merger, and the findings of the present study has theoretical implications regarding this aspect. The present findings complement the view that the merger is highly sensitive across vowel contexts. The V-F2 results exhibited the previously established asymmetry between the /iN/ and /aN/ contexts. While the nasals remain distinct in the /aN/ tokens, the /iN/ tokens showed substantially reduced differences. This observed pattern is consistent with the findings reported by many previous studies (Chiu et al., 2025; Chuang, 2017; Hsu & Tse, 2007; Lai, 2009; Yang, 2007) and supports the idea that the high vowel /i/ is less likely to accommodate the articulatory requirements of adjacent segments, compared to the low vowel /a/ (Faytak et al., 2020). In addition, although previous studies generally agree that extensive neutralization is exhibited in the /iN/ context, there is less agreement on whether the merger is complete or rather a near-merger, which is characterized by a mismatch where the contrast in production persists despite that it is not perceived by listeners. While the previous perceptual study by Chiu et al. (2025) reported that listeners exhibit low accuracy in identifying the nasals in the /iN/ context, the production data in the current study reveals that TM speakers are not producing a random or intermediate category of the nasal in the /iN/ context. TM speakers are instead systematically producing a merged nasal target whose acoustic features closely matches the canonical realization of an alveolar [n]. Despite the low accuracy in identifying the nasals among the listeners, the fact that this consistency exists in production of the nasal suggests that the use of traditional acoustic correlates of the preceding vowel has led to the overlook of the stable realization of the alveolar nasal /n/ in the /iN/ context in the literature. Furthermore, there is little agreement in the literature when it comes to the merging patterns. Theoretically, three potential possibilities exist for the merger: the two nasals merge into the alveolar /n/, the two nasals merge into the velar /ŋ/, and the two nasals merge into a third

distinct category. By demonstrating that the two nasal codas in the /iN/ tokens merge into a single, unidentified category that is not statistically different from the alveolar [n] in the /aN/ context, this study provides empirical support for the alveolar merging pattern. This finding directly challenges the prominent view in the literature adopted by Hsu & Tse (2007), Chiu & Lu (2021), and Chiu et al. (2025), which preferred the velar merging pattern in the /iN/ context. Additionally, the finding that the N-F1 values of the merged nasal in the /iN/ context consistently track the alveolar baseline suggests that the phonetic realization of the merger in TM is highly patterned rather than random. Take together, these implications reframe the nature of the merger from a random sound change to a stable process conditioned by vowels, indicating that looking beyond traditional acoustic correlates of the vowel is necessary to fully capture the phonological reality of TM.

The findings of this study directly address the ongoing methodological limitations regarding the underdetermination of acoustic measurements in the merger. Due to the predominant reliance on the acoustic correlates of the preceding vowel, such as the F2 transition, vowel duration, and degree of nasalization, previous studies have struggled to establish an unambiguous baseline for identifying the merging patterns. The F2 at the end of the vowel is expected to fall in the canonical realization of both /in/ and /iŋ/, and the absence of a baseline has result in difficulties in determining an observed F2 fall in the data is which. By shifting the analyses to the acoustic correlates of the nasal codas themselves, the current study largely avoided this underdetermination. The data showed that while the N-F2, nasal duration, and nasal/rime ratio fail to reliably distinguish the two nasals, the N-F1 provides a stable, independent baseline for the place of articulation in the non-merging /aN/ context. Using N-F1 as an informative cue, this allowed the category of the merged nasal in the /iN/ tokens to be mapped directly against the canonical realizations in the /aN/ tokens, demonstrating that analyzing the nasals themselves provides a clarity that analyzing the vowels lack. In addition, this approach highlights the systemic limitations of automated formant tracking in software like Praat. Due to coarticulated nasalization, nasal formants that inherently have low frequencies

inevitably emerge during the vowel, leading the formant tracker to potentially mistake the nasal formants for the true F1 or F2 of the vowel. This creates a wide range of variation across participants, reinforcing the suggestion that traditional acoustic correlates of the vowel should be interpreted with caution and is better supplemented by analysis of the nasals.

The suggestion of the alveolar merging pattern among young TM speakers from northern Taiwan requires an evaluation of the demographic and regional differences reported in previous studies. The nasal coda merger has been characterized as an innovation of younger generations. Fon et al. (2011) suggested that the alveolar merging pattern in the /iN/ context is exclusive among speakers from southern Taiwan and attributed this regional difference to the heavy influence of Taiwanese Hokkien in that region. Conversely, it was suggested that speakers from northern Taiwan favor the velar pattern in the /iN/ context. The results of the current study contradict with this reported regional difference, demonstrating that young speakers in northern Taiwan also systematically exhibit the alveolar pattern in the /iN/ context. This finding provides empirical support for the recent observations made by Lee-Kim et al. (2024), in which trained phoneticians identified an emerging presence of the alveolar pattern among young speakers from northern Taiwan. Since the present results indicate that the place contrast in minimized regarding V-F2 and that the two nasals are merged into the alveolar /n/ regarding N-F1, it is evident that the alveolar merging pattern is no longer an exclusive pattern to southern Taiwan. However, whether this is due to a different methodology adopted in this study or due to the alveolar pattern diffusing into northern Taiwan remains unclear.

The contradictory findings across previous studies, in which some argue for the alveolar pattern in the /iN/ context and others for the velar pattern, could be traced directly to the reliance on problematic criteria and the absence of an established baseline. One key example is the use of velar pinch as the primary criterion to identify the merging patterns in the study by Hsu & Tse (2007).

Theoretically, the velar pinch should serve as a reliable cue for a following velar segment. However, results from Chiu et al. (2025) and from the present study in the /aN/ context show that the canonical realization of the /aŋ/ in TM exhibits a falling F2 at the end of the vowel. This indicates that velar pinch is absent even in the canonical realization of the velar nasal in TM, meaning that any previous study that identify the merged /iN/ tokens as the alveolar [in] due to the velar pinch not being observed was operating on an unreliable criterion. Similarly, contradictions rise due to differing research focuses and methodologies. Acoustic studies examining vowel formants at various time points frequently lead to contradicting patterns. Chuang (2017) observed a lower F2 for /iŋ/ compared to /in/, whereas Chiu et al. (2025) found a small difference in the /əN/ context and no significant difference in the /iN/ context. These discrepancies are further complicated when compared to data from ultrasound studies. Ultrasound studies show that in the /iN/ context, the tongue tip is lowered and the tongue body is retracted and bunched, closely resembling the tongue posture for the velar nasal [ŋ]. The findings of the current study provide an addition to this mismatch based on acoustic results. While ultrasound studies show a velarized tongue posture in the /iN/ context, the results of N-F1 show the acoustic features of the alveolar [n]. This indicates that the physically velarized tongue posture captured by ultrasound does not necessarily point to a canonical realization of the velar nasal [ŋ], which leads to unresolved contradicting conclusions between studies that measure tongue postures and studies that measure acoustic correlates.

The alveolar merging pattern among young TM speakers from northern Taiwan proposed by the current study has broader implications for language technology, sociolinguistic identity, and phonological description. Regarding computational linguistics, the neutralization of place contrast in the /iN/ context in TM presents a major challenge for automated speech recognition. Since Mandarin recognition technologies are more likely to be trained on Standard Mandarin than TM, in which the former maintains the contrast between the alveolar /n/ and the velar /ŋ/, the speech

recognition systems will likely struggle with high error rates when processing TM speech. Taking into consideration the empirical findings presented in this study that TM speakers systematically produce the alveolar nasal [n] in the /iN/ context could potentially help speech technology engineers adjust and improve the accuracy for TM. In terms of sociolinguistics, the currently attested alveolar merging pattern in the /iN/ context among young TM speakers from northern Taiwan reflects the development of TM being a distinct and independent variety of Mandarin. As noted by Chen & Herring (2024), TM exhibits distinct phonetic, phonological, and lexical features that differentiate it from Standard Mandarin, despite the two varieties being mutually intelligible. The attested stable alveolar pattern in the /iN/ context in northern Taiwan indicates that this pattern should no longer be viewed as a phenomenon caused by language contact, but a native phonological feature of TM instead. Finally, these findings emphasize again that the nasal coda merger is essentially vowel sensitive. The observation that the nasal place contrast remains intact and robust in the /aN/ context while being largely neutralized in the /iN/ context demonstrates that phonological features cannot be described globally across an entire phonetic inventory. Phonological features are strictly bound by the phonetic environments and coarticulatory constraints.

While this study provides valuable insights by shifting the focus of analyses from the acoustic correlates of the vowel to those of the nasal, several critical limitations regarding the data and measurement must be acknowledged. One major limitation is the systematically higher exclusion rate of /in/ tokens, compared to the other types of excluded tokens. Out of all tokens, 11 /in/ tokens, one /iŋ/ token, and two /an/ tokens were excluded due to the difficulties in determining the boundary between the vowel and the nasal. This issue is heavily concentrated in the lexical item /pin/ (eight tokens excluded) and /tɕ^{h}in/ (three tokens excluded). In most of these excluded tokens, the vowel formants remained visible and stable all the way to the end of the syllable, which indicates that the nasal codas were reduced from an independent segment to the complete or partial nasalization of the vowel. Due to the strict segmentation criteria set that is crucial to measure the

acoustic correlates of the nasal, these tokens with a reduced nasal coda had to be excluded from the statistical analyses. Consequently, the present results are inherently biased toward tokens that maintained a relatively more independent nasal coda, indicating that the study might overlook the potential phenomenon of nasal coda deletion in TM. Another major limitation of the design of this study is the absence of the mid vowel /əN/ context in the stimuli. Previous studies have established that the merge is not only more advanced in the /iN/ context, but also more advanced in the /əN/ context, compared to the non-merging /aN/ context. By excluding the /əN/ context in the design and only including the /iN/ and /aN/ contexts, the current study was unable to evaluate whether the N-F1 correlate serves as an informative cue across all vowel contexts in the merger. Finally, the automated formant tracker built in Praat posed a substantial methodological challenge and introduces inherent limitations to the collected dataset. Close inspection of the waveforms and spectrograms of the data revealed that the nasal formants were often already visible and stable during the vowel due to coarticulated vowel nasalization. Since these nasal formants typically have low frequencies, the automated tracker frequently mistook them for the F1 or F2 of the vowel, which influenced the V-F2 measured at the end of the vowel where nasal formants had inevitably emerged. This limitation resulted in a wide range of variations across participants that could not be reliably adjusted within the scope of this study. Therefore, it is evident that the results and suggestions derived from the data with the systematic variations should be interpreted with caution. Future studies should address these limitations by preventing the use of stimuli in which the nasal codas are frequently deleted, including the missing /əN/ context in the stimuli, and utilizing other formant tracking mechanisms that avoid the formant tracking issue in Praat. In addition, including a perception task in the research design would reveal whether listeners could utilize the reliable nasal correlate N-F1 identified in this study.

References

Chen, Y.-A. J., & Herring, S. C. (2024). “What a standard Taiwan Mandarin accent”: Online metalinguistic commentary on linguistic performances of non-native Chinese speakers. *Language & Communication*, 99, 141–158. https://doi.org/10.1016/j.langcom.2024.10.001

Chuang, C.-T. (2017). Revisiting nasal coda merger in Taiwan Mandarin: A corpus study. *Concentric: Studies in Linguistics*, *43*(2), 1–27. https://doi.org/10.6241/concentric.ling.43.2.01

Chiu, C., & Lu, Y.-A. (2021). Articulatory evidence for the syllable-final nasal merging in Taiwan Mandarin. *Language and Speech*, *64*(4), Article 0023830920948084. https://doi.org/10.1177/0023830920948084

Chiu, C., Lu, Y.-A., Lin, K.-H., & Jin, S.-J. (2025). Prosody-modulated and vowel-dependent nasal merger in Taiwan Mandarin. *The Journal of the Acoustical Society of America*, *157*(2), 1523–1541. https://doi.org/10.1121/10.0035943

Chiu, C., Lu, Y., Weng, Y., Jin, S., Weng, W., & Yang, T. (2019). Uncovering syllable-final nasal merging in Taiwan Mandarin: Ultrasonographic investigation of tongue postures and degrees of nasalization. In *Proceedings of the 19th International Congress of Phonetic Sciences* (pp. 398-402). Melbourne, VIC; Canberra, ACT: Australasian Speech Science and Technology Association.

Dong, T. (2024). Standard Mandarin phonology. *Historical Phonology of Chinese* (1st ed., pp. 14–32). Routledge. https://doi.org/10.4324/9781003411444-2

Faytak, M., Liu, S., & Sundara, M. (2020). Nasal coda neutralization in Shanghai Mandarin: Articulatory and perceptual evidence. *Laboratory Phonology: Journal of the Association for Laboratory Phonology*, *11*(1), 23. https://doi.org/10.5334/labphon.269

Fon, J. & Chiang, W.-Y. (1999). What does Chao have to say about tones?: A case study of Taiwan Mandarin. *Journal of Chinese Linguistics,* 27(1), 13–37.

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

Hsu, H. J., & Tse, J. K. P. (2007). Syllable-final nasal mergers in Taiwan Mandarin–Leveled but puzzling. *Concentric: Studies in Linguistics*, *33*(1), 1-18.

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

Lai, Y. (2009). Acoustic correlates of Mandarin nasal codas and their contribution to perceptual saliency. *Concentric:Studies in Linguistics*, *35*(2), 143–166.

Lee-Kim, S. I., Chen, S. Y., Wade, L., & Tamminga, M (2024). Multiple nasal mergers in Taiwan Mandarin: a case of perception-production misalignment. [Poster presentation]. *Lab Phon 19: The 19th Laboratory Phonology Conference*.

Lee-Kim, S.-I., & Chou, Y.-C. I. (2022). Unmerging the sibilant merger among speakers of Taiwan Mandarin. *Laboratory Phonology*, *13*(1). https://doi.org/10.16995/labphon.6446

Torgerson, R. C., Jr. (2005). A comparison of Beijing and Taiwan Mandarin tone register: An acoustic analysis of three native speech styles. [Master's thesis, Brigham Young University]. *BYU ScholarsArchive*.

Tse, J. K.-P. (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.

Yang, J. H. (2007, September). The role of sound changes in the speech recognition System: A phonetic analysis of the final nasal shift in Mandarin. In *Proceedings of the 19th Conference on Computational Linguistics and Speech Processing* (pp. 245-256).

Appendix A

List of Target Stimuli

/in/

來賓已經全部就位了

/lái.pīn ǐ.tɕīŋ tɕʰɥén.pù tɕjòu.wèɪ lə/

‘The guests are all set.’

租金的價格一直在漲

/tsū.tɕīn tə tɕjà.kɤ́ ì.ʈʂì̯ tsài ʈʂǎ̯ŋ/

‘The rent keeps rising.’

生態入侵是當務之急

/ʂəŋ.tʰài ʐù.tɕʰīn ʂì̯ tāŋ.ù.ʈʂì̯.tɕí/

‘The invasive species is an urgent matter.’

/iŋ/

他已經溜冰好幾年了

/tʰā ǐ.tɕīŋ ljōu.pīŋ hǎu.tɕǐ njén lə/

‘He has been skating for years.’

這家店的水晶很便宜

/tʂɤ̌.tɕjā tjèn tə ʂwěi.tɕīŋ hən pʰjén.í/

‘The crystals in this shop are cheap.’

地球的衛星只有一顆

/tì.tɕʰjóu tə wèi.ɕīŋ ʈʂ͡ɻ̩̆.jǒu í kʰɤ̄/

‘Earth only has one satellite.’

/an/

每天下班都不想運動

/měɪ.tʰjɛ̄n ɕjà.pān tōu pù ɕjǎŋ ỳn.tòŋ/

‘Everyday after work I don’t want to exercise.’

買單好像要去櫃檯買

/mǎɪ.tān hǎu.ɕjàŋ jàu tɕʰỳ kwèi.tʰái mǎi/

‘It seems that you need to pay at the counter.’

那座海灘你一定要去

/nà.tswò hǎɪ.tʰān nǐ í.tìŋ jàu tɕʰỳ/

‘We haven’t been to that beach.’

/aŋ/

這個地方我們沒來過

/ʈʂ͡ɤ̀.kə tì.fāŋ wǒ.mən méɪ lái kwò/

‘You must go to that place.’

天冷喝熱湯非常舒服

/tʰjɛ̄n lə̌ŋ hɤ̄ rɤ̀.tʰāŋ fē̄i.ʈʂ͡ʰáŋ ʂū.fú/

‘It feels good to drink hot soup when it’s cold.’

今年外商好像賠很多

/tɕīn.njɛ́n wài.ʂāŋ hǎu.ɕjàŋ pʰéi hə̌n tūɔ/

‘The foreign companies are making a lot of losses this year.’