JOURNAL ARTICLE

Nasal coda neutralization in Shanghai Mandarin: Articulatory and perceptual evidence

Matthew Faytak, Suyuan Liu and Megha Sundara
Department of Linguistics, University of California, Los Angeles, US
Corresponding author: Matthew Faytak (faytak@ucla.edu)

Shanghai Mandarin is reported to neutralize /n/ and /ŋ/ after non-low vowels, a change also reported for other varieties of Chinese. However, the place of articulation of the resulting nasal is unclear. We used production and perception experiments to determine the contexts which condition neutralization and the place of articulation of the resulting nasal. Analysis of ultrasound data using a dimensionality reduction technique revealed that neutralization in production was dependent on the height of the preceding vowel: The highest vowel /i/ conditioned a consistent, non-structure-preserving neutralization to a fronted velar or palatal nasal; the mid vowel /ə/ conditioned an inconsistent, structure-preserving neutralization to [n]; and no neutralization occurred after the low vowel /a/. In contrast, Standard Mandarin controls produced the contrast in all vowel contexts. A follow-up perception experiment confirmed that, consistent with complete neutralization, Shanghai Mandarin listeners failed to distinguish alveolar and velar nasals following /i/. More surprisingly, so did Standard Mandarin listeners, suggesting a near merger of the nasal codas following /i/ in Standard Mandarin. Our results provide support for a role for substantive factors in neutralization-driven sound changes.

Keywords: Nasals; Shanghai Mandarin; Standard Mandarin; ultrasound; neutralization

1. Introduction

In Chinese languages, the syllable coda is generally restricted to a small set of segments, including the nasals /m/, /n/, and /ŋ/. Most varieties have lost coda /m/; many varieties such as Standard Mandarin canonically still distinguish /n/ and /ŋ/ (Chen, 1973; Zee, 1985; Duanmu, 2007). In other modern Chinese varieties, mergers or positional neutralizations of /n/ and /ŋ/ are frequently observed. Interference from an L1 that lacks the relevant place contrast is the most commonly proposed cause for these mergers of nasal coda place in varieties of Chinese commonly spoken as an L2 (Lai, 2008; Chen & Guion-Anderson, 2011; Luo, 2015; Wu, Sloos, & van de Weijer, 2016; Wang, Cui, & Chen, 2018). However, loss of the nasal coda place contrast in modern Chinese, and the specific places of articulation which are favored in the merged or neutralized codas, can also be attributed to various substantive factors.

More specifically, nasal consonants are thought to be relatively vulnerable to sound change because of their low perceptual salience (Chen, 1973; Zee, 1981; Zee, 1985). Place contrasts in nasal consonants are more difficult to distinguish than in oral stops in both onset (Narayan, 2008) and coda position (Kawahara & Garvey, 2014), and are subject to positional neutralization more frequently than similar oral segments in the world's languages (Ohala, 1990; Steriade, 1994; Lombardi, 2001). This is likely because nasals typically have low intensity, and their nasal antiformants obscure the formant frequencies inherent to their places of articulation (Kurowski & Blumstein, 1984, 1993; Narayan, 2008).

Detecting nasal place contrasts is particularly challenging in the context of the vowel [i] and other front vowels. For instance, English listeners, who have a three-way place distinction for coda nasals, have been shown to misperceive coda velar nasals as alveolar after front vowels [i] and [e] (Zee, 1981). English listeners even confuse onset [m] and [n] more often when followed by [i] than other vowels (Kurowski & Blumstein, 1984).

The [i] vowel context may be especially conducive to place neutralization because of its biomechanical properties. Palatal segments, including front vowels and especially [i], have a relatively high *coarticulatory resistance*, making them insensitive to coarticulation with adjacent segments and more likely to impose their articulatory demands on those segments. Low vowels such as [a], on the other hand, have a lower coarticulatory resistance and are freer to vary contextually to accommodate adjacent segments' articulatory requirements (Stone & Vatikiotis-Bateson, 1995; Recasens & Espinosa, 2009; Recasens, 2012; Chen, Chang, & Iskarous, 2015; Recasens & Rodríguez, 2016). The articulatory requirements of producing non-low vowels may as such influence the course of nasal merger or neutralization, primarily in determining the outcome of merger: Articulatorily demanding vowels will tend to exert a larger influence on the non-contrastive place of nasals which effectively lack their own place targets, consistent with models of coarticulation and underspecification in phonetics (Keating, 1988a; Keating, 1988b; Bakovic, 2000).

In this paper we consider data from two regional accents of Standard Mandarin, the language variety known in its standardized forms as *Pǔtónghuà* 普通话 in Mainland China, *Guóyǔ* 国语 in Taiwan, and *Huáyǔ* 华语 in Singapore. We used production and perception experiments to investigate the extent and outcome of nasal coda neutralization in Shanghai Mandarin, the regional dialect of Standard Mandarin spoken in and around the city of Shanghai, and the factors that contribute to it.

In Experiment 1, we present ultrasound imaging data to evaluate the articulation of nasal codas in four vowel contexts by both native northern Standard Mandarin speakers and bilingual Shanghai Mandarin speakers. In Experiment 2, we present results from a perception experiment with the same groups focused primarily on discrimination of coda nasals following the high vowel /i/, the context in which neutralization is most often observed. Finally, we discuss the implications of the results from the production and perception experiments for the characterization of non-contrastive nasal place in the syllable coda in Chinese languages, and more generally for characterizing the outcomes of neutralization and contrast reduction processes in language change. In the rest of the introduction, we give an overview of the production of nasal codas in *Standard Mandarin* and past studies on nasal coda neutralization in varieties of Mandarin, with a focus on *Taiwan Mandarin* and *Shanghai Mandarin*, leading up to our proposed experiments.

1.1. Nasal codas in Standard Mandarin

The canonical variety of Standard Mandarin spoken in northern mainland China is described as having two nasal codas, alveolar /n/ and velar /ŋ/. The lingual constriction of both nasals is occasionally described as reduced or absent, mainly after the low vowel /a/; velum lowering and the resulting nasalization persist in these cases (Xu, 1993; Chen, 2000; Duanmu, 2007, p. 24). Not all combinations of the five monophthongal vowel phonemes /i y u ə a/ and the two coda nasals occur; the minimally different combinations of vowels and nasal codas that do occur are shown in **Table 1**.

Despite having a phonemic contrast between alveolar and velar nasal codas, Standard Mandarin listeners show some difficulty distinguishing the nasal place contrast after non-low vowels, particularly /i/. Error rates for coda nasal place identification after /i/ in Standard Mandarin can be as high as 20% (Mou, 2006, p. 125); this is much higher than the error rate of 10% expected for native contrasts (Polka, 1991). Furthermore, in gating

Table 1: Standard Mandarin rhymes containing nasal codas, after Duanmu (2007), with example characters. Pairs of rhymes examined in the present study are bolded.

Vowel	Coda /n/		Coda /ŋ/	
/i/	**/in/**	音 [in˥] 'sound'	**/iŋ/**	英 [iŋ˥] 'heroic'
/y/	/yn/	晕 [yn˥] 'faint'		
/u/			/uŋ/	工 [kuŋ˥] 'work'
/ə/	**/ən/**	根 [kən˥] 'root'	**/əŋ/**	羹 [kəŋ˥] 'custard'
	/uən/ [un ~ wən]	温 [wən˥] 'Wenzhou'	/uəŋ/	翁 [wəŋ˥] 'old man'
/a/	**/an/**	肝 [kan˥] 'liver'	**/aŋ/**	钢 [kaŋ˥] 'steel'
	/uan/	官 [kwan˥] 'official'	**/uaŋ/**	光 [kwaŋ˥] 'ray'
	/ian/ [jɛn]	烟 [jɛn˥] 'smoke'	/iaŋ/	秧 [jaŋ˥] 'sprout'

tasks, native Mandarin listeners can predict the place of articulation of an upcoming nasal 80 ms through /ə/ or /a/, but not /i/ (Mou, 2006). Following /i/, nasal tokens are increasingly judged to be /iŋ/ at each successive gate closer to the nasal coda past the vowel midpoint (Mou, 2006, p. 125). These results suggest that Mandarin listeners are unable to distinguish the sub-phonemic coarticulatory fronting of velar nasals in the context of /i/ from the retracted tongue position typical for native phonemic alveolar /n/ in this vowel context.

Acoustic investigations of Standard Mandarin provide a possible explanation for why nasal place perception is worse after /i/. In the non-high vowels /a/ and /ə/, a difference in monophthongal vowel quality co-occurs with nasal place difference: These vowels are more fronted before coda /n/ compared to /ŋ/, a pattern sometimes referred to as *Rhyme Harmony* (Duanmu, 2007). Rhyme Harmony results in a 250–500 Hz F2 frequency difference between the vowels in Standard Mandarin /an/ versus /aŋ/ (Shih, 1995; Chen, 2000; Mou, 2006), which likely contributes to the stability of the /n/ – /ŋ/ contrast after /a/. Similarly, the F2 frequency of the vowel in /əŋ/ is also 250–500 Hz lower than that of /ən/ (Chen, 2000; Mou, 2006), and the /n/ – /ŋ/ contrast after /ə/ is generally described as stable.

In contrast, the high vowel /i/ interacts differently and less consistently with nasal codas: The F2 of /i/ is not obviously affected by the place of the following nasal (Mou, 2006, p. 117). Some speakers of Standard Mandarin employ a strategy of schwa insertion in the nasal rhyme /iŋ/, producing it as [iə̯ŋ] (Mou, 2006; Duanmu, 2007). The inserted schwa has an F1 of about 500 Hz, which is higher than the F1 value of /i/ at roughly 300 Hz (Mou, 2006, p. 128). While this vowel quality difference may allow Standard Mandarin listeners to distinguish alveolar and velar nasal codas, it is not clear that schwa insertion is consistent across speakers. Duanmu describes schwa insertion as mandatory (2007, p. 58; see **Table 1**), but in an investigation by Mou (2006) three out of five speakers of Standard Mandarin failed to exhibit schwa insertion, producing [iŋ] rather than [iə̯ŋ] for /iŋ/. Thus, co-occurring vowel quality differences observed in /a/ and /ə/, but not /i/, are likely to facilitate successful discrimination of coda nasals.

Overall, both Rhyme Harmony and schwa insertion enhance the place contrast in coda nasals (Keyser & Stevens, 2006) when they follow the non-high vowels. Because not all speakers use schwa insertion as a strategy to differentiate /in/ and /iŋ/, it is likely more difficult for listeners to predict the upcoming coda nasal based on the vowel quality of /i/. Based on these results, some researchers have proposed a reduction in the /n/ – /ŋ/ contrast after /i/ in the mainland Standard Mandarin nasal inventory (Chen, 2000; Mou, 2006). The regional Standard Mandarin varieties considered in the following sections show contrast reduction not only in this context but also after other non-low vowels.

1.2. Coda nasal neutralization in Taiwan Mandarin

From the previous section we can see that there is some evidence that the nasal place contrast in Standard Mandarin is neutralized after /i/. In this section, we consider prior research on nasal codas in Taiwan Mandarin, a regional variety of Standard Mandarin that has been widely investigated. Like Standard Mandarin, Taiwan Mandarin exhibits the aforementioned tendency for F2 to be raised in the vowels /ə/ and /a/ when they are followed by an alveolar nasal [n] (Lin, 2002, p. 7; Lai, 2008, p. 165). However, some crucial differences exist between Taiwan Mandarin and mainland Standard Mandarin. Specifically, there is a consensus that speakers of Taiwan Mandarin typically merge the nasal codas following /i/ *and* merge variably after /ə/, but never after /a/ (Chen, 1991; Lin, 2002; Hsu & Tse, 2007; Fon, Hung, Huang, & Hsu, 2011; Chiu, Lu, Weng, Jin, Weng, & Wang, 2019).

Taiwan Mandarin speakers often also speak *Taiwanese Min Nan*, also known as *Hokkien*, a Min variety not mutually intelligible with Standard Mandarin. At least some of the specifics of the neutralization pattern in Taiwan Mandarin have been attributed to interference from Taiwanese Min Nan. Although Taiwanese Min Nan, unlike Standard Mandarin, has a three-way nasal place distinction in codas between /m/, /n/, and /ŋ/ (Chen & Guion-Anderson, 2011), Southern Taiwanese Min Nan allows /in/ but not /iŋ/. Seemingly as a consequence, Southern Taiwanese Mandarin speakers are often described as neutralizing the nasal codas to [n] following /i/, a pattern not observed in speakers of Northern Taiwanese Mandarin (Fon et al., 2011).

What remains controversial is the extent and outcome of the neutralization. Based on transcription data, there is near-total agreement in the literature that Taiwan Mandarin speakers neutralize the two coda nasals to alveolar [n] after /ə/ (Chen, 1991; Lin, 2002; Hsu & Tse, 2007; Fon et al., 2011). However, in a recent ultrasound study of nasal codas, Chiu, Lu, et al. (2019) report inter-speaker variation even in the context of /ə/. Four of nine participants in Chiu, Lu, et al.'s ultrasound study did not merge the nasals after /ə/; of the remaining five, three produced a merged [n] in stimuli containing both /ən/ and /əŋ/, but another two produced a merged [ŋ] in the same environment.

The outcome of nasal coda merger is especially unclear following the high vowel /i/. In perception experiments using an identification task, the nasal resulting from merger is sometimes identified as alveolar. For instance, the three annotators in Yang (2010) report perceiving an alveolar nasal after /i/. Note that these results are consistent with the perceptual bias reported for alveolar nasals in English listeners (Zee, 1981). In contrast are perceptual results from Lin (2002), who reports a bias toward perceiving a velar nasal /ŋ/ after the high front vowel /i/, and a bias toward perceiving an alveolar nasal /n/ after /ə/. The articulatory outcome of the nasal coda merger after /i/ is also similarly unclear. Fon et al. (2011) suggest that speakers from southern Taiwan merge coda nasals after /i/ to alveolar nasals, whereas speakers from northern Taiwan merge coda nasals in the same context to velar nasals. In contrast, Chiu, Lu, et al. (2019), who report the only articulatory data collected to date on the merger, suggest instead that /i/ conditions merger to /ŋ/.

Overall, Taiwan Mandarin listeners variably merge /n/ and /ŋ/, but the resulting nasal's place of articulation is unclear. Perceptual identification experiments alternately suggest that the outcome of neutralization is either alveolar or velar, both of which are *structure-preserving* outcomes in the sense that they are identified with one of the segments that participates in the neutralization. However, in all these perceptual experiments, listeners were forced to choose between alveolar and velar labels. Chiu, Lu, et al. also do not explicitly compare merged and unmerged nasals to identify the former's place of articulation. Thus, we cannot rule out the possibility that the place of the neutralized nasals in Taiwan Mandarin is neither alveolar nor velar, but intermediate between the two, and not structure-preserving in the aforementioned sense. This could account for listener groups identifying the nasals

sometimes as alveolar or velar in identification tasks. This brings us to the case of nasal coda merger in Shanghai Mandarin.

1.3. Coda nasals in Shanghai Mandarin and Shanghainese

Shanghai Mandarin is a Standard Mandarin variety spoken by residents of the Shanghai area which is often a second language learned after Shanghainese during schooling or acquired simultaneously with Shanghainese at home. *Shanghainese* is the mutually unintelligible local variety of Wu Chinese spoken in the Shanghai area. Unlike most mainland Standard Mandarin varieties, but like Taiwan Mandarin, Shanghai Mandarin is also reported to neutralize /n/ and /ŋ/ after the non-low vowels /i/ and /ə/ (Luo, 2015; Guan, 2019). In the limited research on Shanghai Mandarin, the neutralized post-/i/ nasals are also described as alveolar [n] (Guan, 2019; Luo, 2015). Both of these perceptual studies use an identification task; recall that based on results from such tasks, Taiwan Mandarin listeners have been reported to have either an alveolar or velar response bias when it comes to discriminating /n/ and /ŋ/ after /i/ (Mou, 2006, p. 125; Lin, 2002).

Coda nasal merger in Shanghai Mandarin has been explicitly attributed to the influence of Shanghainese (Luo, 2015), which lacks a coda nasal place contrast. Effects of the first language (L1) on incomplete acquisition of sound contrasts in a second language are ubiquitous; here we summarize findings that are specific to nasal codas. Increased experience with Chinese varieties which lack a coda nasal contrast has been shown to reduce listeners' perceptual sensitivity to the coda /n/-/ŋ/ contrast in English (Wu et al., 2016) and Standard Chinese (Wang et al., 2018). Likewise, Burmese speakers who learned Mandarin as an L2 are less sensitive in discriminating Mandarin coda nasals (Lai, 2008), in keeping with Burmese having only non-contrastive coda nasality like Shanghainese (Green, 2005). Given the absence of a coda nasal place contrast in Shanghainese, it is then not particularly surprising that Shanghai Mandarin speakers have difficulty distinguishing some codas differing in nasal place. What cannot be explained by an L1 interference account is why mergers in Shanghai Mandarin are reported in some but not all vowel contexts.

An L1 interference account is also limited in its ability to predict the place of articulation of the neutralized nasals in Shanghai Mandarin. A logical possibility is that the neutralized nasals in Shanghai Mandarin have the same place of articulation as nasals in similar rhymes in Shanghainese. However, other than after the low vowels /a/ and /ɑ/, where lingual closure during nasal codas is largely agreed to be incomplete or inconsistent, descriptions of Shanghainese offer multiple candidates for the place of articulation of the nasal, sometimes varying by vowel context (**Table 2**). There is general agreement that the back non-high vowel /o/ is followed by a velar nasal (Sherard, 1972; Shen, 1981; Qian 2003, p. 56). But there is significant disagreement on how the nasal coda is articulated after the remaining non-low, non-back vowels. Some impressionistic accounts state that alveolar [n] occurs after /i/, /y/, and /ə/ (Shen, 1981; Xu & Tang, 1988), and Ping (2005) provides electropalatographic evidence that at least some tokens of the Shanghainese coda nasal in these environments are alveolar [n] (21). Qian (2003, p. 56) posits a palatal

Table 2: Comparative phonetic transcription of selected Shanghainese rhymes containing a non-contrastive nasal coda or nasalization of the nuclear vowel, here denoted as /N/.

	/aN/	/ɑN/	/əN/	/ɪN/	/ʏN/
Sherard (1972, p. 72, 74)	ã ~ ãŋ	ɔ̃ ~ ɔ̃ŋ	ə̃ŋ	ĩŋ	iʊ̃ŋ
Xu and Tang (1988, p. 8)	ã	ɑ̃	ən	in	yn
Qian (2003, p. 56)	ã̠ⁿ		əɲ	iɲ	yɲ
Chen and Gussenhoven (2015)	ɐŋ	ɑŋ	əŋ	ɪŋ	ʏŋ

nasal coda [ɲ] for this group of rhymes, and still other analyses suggest a velar nasal [ŋ] (Sherard, 1972; Chen & Gussenhoven, 2015). Note that such variation in phonetic description of non-contrastive features, owing to the perceptual biases of transcribers, is a well-known limitation when relying solely on auditory transcription (Kerswill & Wright, 1990; Pouplier & Hardcastle, 2005; Munson, Edwards, Schellinger, Beckman, & Meyer, 2010). Regardless, such variability limits the utility of an L1 interference account. That is, these descriptions provide us with viable candidates for the merged place of articulation of nasals in Shanghai Mandarin under an L1 interference account, but no way to arbitrate between them.

1.4. The current experiments

In this paper we examined the extent of coda nasal place neutralization in Shanghai Mandarin, its outcome, and the vowel contexts which encourage it. We were motivated to investigate place neutralization because it has been studied less than neutralization of other features. The extant phonetic research on neutralization is heavily focused on voicing neutralization in coda obstruents (see Warner, Jongman, Sereno, & Kemps, 2004; Kleber, John, & Harrington, 2010; Roettger, Winter, Grawunder, Kirby, & Grice, 2014; Nicenboim, Roettger, & Vasishth, 2018) and flapping in North American English (see Herd, Jongman, & Sereno, 2010; Braver, 2014). These are usually described as structure-preserving processes in phonological accounts: That is, neutralization results in a sound structure present outside of the neutralizing environment, for instance a voiceless stop in the case of voicing neutralization. Instrumental phonetic studies of voicing neutralization demonstrate, however, that phonetic traces of the intended forms typically remain, making neutralization incomplete (Fourakis & Iverson, 1984; Port & O'Dell, 1985; Port & Crawford, 1989; Warner et al., 2004; Kleber et al., 2010; Roettger et al., 2014). Place neutralization, along with most other types of neutralization, has not been subject to this level of scrutiny (see Yu, 2011). It remains to be seen whether the phonetics of place neutralization are similar to those of voicing neutralization in generating novel phonetic structures from the participating segments.

Examining place neutralization conditioned by segmental context also provides an opportunity to probe contributions of perceptual and biomechanical biases to the process of contrast reduction. If outcomes are modulated by segmental context, as may be the case in Shanghai Mandarin, neutralization can be assessed with respect to substantive properties of this context; in the case of Shanghai Mandarin, for instance, neutralization may prove to be mediated by perceptual or articulatory properties of the preceding vowels. Unconditioned merger in place of articulation (Bukmaier, Harrington, & Kleber, 2014; Chang & Shih, 2015; Jannedy & Weirich, 2017; Chiu, Wei, Noguchi, & Yamane, 2019) does not afford this opportunity because it applies regardless of segmental context.

In our experiments, we first used ultrasound tongue imaging to examine nasal coda neutralization in production by Shanghai Mandarin and Standard Mandarin control speakers. We then played stimuli uttered during the production experiment to a separate group of Shanghai Mandarin and Standard Mandarin speaking controls. In this perception experiment, we used an AXB task instead of an identification task as has been done previously. Recall that forcing subjects to choose between alveolar or velar labels is not useful in cases where the place of articulation of the resulting merger may be neither alveolar nor velar. We used the perception results not to identify the place of the resulting neutralized nasals, but to probe the extent of neutralization in Standard versus Shanghai Mandarin. Comparisons of this sort are typically not made within the same study, even though studies of both Standard Mandarin (Mou, 2006) and regionally accented Mandarin (Yang, 2010; Chiu et al, 2019; Luo, 2015; Guan, 2019) suggest similar patterns of neutralization.

2. Ultrasound study

We first used ultrasound tongue imaging as a means of directly determining whether there was neutralization in production and to identify the non-contrastive place of articulation of the neutralized nasal. Given the rarity of direct articulatory imaging studies of place neutralization (but see Ramsammy, 2013; Chiu et al, 2019), the existing evidence for neutralization *to a particular place of articulation* is weak, likely because of the documented perceptual limitations in identifying nasal place of articulation in coda position. We additionally aimed to determine the extent of inter-speaker variation in production, which seems essential given conflicting reports in descriptions to date. All procedures described below were approved by UCLA's General Institutional Review Board.

2.1. Participants

Five self-identified native speakers of Standard Mandarin (3 male, 2 female, ages 21–23, mean age = 22) and fifteen self-identified bilingual speakers of Shanghai Mandarin and Shanghainese (4 male, 11 female, ages 19–25, mean age = 21.87) were recruited as participants in the ultrasound experiment. All participants were students at the University of California, Los Angeles and thus also spoke English. Mandarin control subjects were confirmed to have no ability in Shanghainese and never to have lived in a predominantly Shanghainese-speaking region. All bilingual participants except for speaker 9 acquired Shanghainese and Mandarin simultaneously from birth; Speaker 9 acquired Mandarin at age five. All participants were offered $15 as compensation.

2.2. Materials

Stimuli consisted of 16 pairs of Mandarin disyllabic words (32 words in total) differing only in whether they ended with an alveolar nasal coda /n/ or velar nasal coda /ŋ/. Each pair of stimuli was roughly frequency-matched based on occurrence in the SUBTLEX corpus (Cai & Brysbaert, 2010). The stimuli included four preceding vowel contexts for the nasal codas at issue: /i/, /ə/, /a/, and the diphthong /ua/ (**Table 3**, see Table A1 in the Appendix for full stimuli). Stimuli were presented to both participant groups using the same simplified Chinese characters.

2.3. Procedure and equipment

Stimuli consisted of 16 pairs of Mandarin disyllabic words (32 words in total) differing only in whether their second syllable ends in an alveolar or velar nasal. Simultaneous ultrasound tongue imaging and audio were recorded in a sound booth in the UCLA Phonetics Laboratory. Ultrasound images were recorded using a Telemed Micro device outfitted with a Telemed MC4-2R20S-3 convex probe recording at a rate of 82 frames per second. The probe's full field of view of 104° and an imaging depth of 90mm were used in all recordings. The probe was stabilized under each speaker's chin with an Articulate Instruments UltraFit headset (Spreafico, Pucher, & Matosova, 2018). Audio was recorded using a Røde smartLav+ omnidirectional lavalier microphone attached to the headset approximately five centimeters to the right of the speaker's mouth (sampling rate 44.1 kHz). Audio recordings were digitized

Table 3: Sample stimuli.

Vowel	/n/ coda		/ŋ/ coda	
/i/	押金 'deposit'	ia˥.t͡ɕin˥	压惊 'help sb. recover'	ia˥.t͡ɕiŋ˥
/ə/	清真 'Islamic'	t͡ɕʰiŋ˥.t͡ʂən˥	清蒸 'steamed'	t͡ɕʰiŋ˥.t͡ʂəŋ˥
/a/	青山 'green hills'	t͡ɕʰiŋ˥.ʂan˥	轻伤 'minor wound'	t͡ɕʰiŋ˥.ʂaŋ˥
/ua/	机关 'mechanism'	t͡ɕi˥.kuan˥	激光 'laser'	t͡ɕi˥.kuaŋ˥

using a Focusrite Scarlett 2i2 audio interface, which was also configured to receive the synchronization pulse train emitted by the ultrasound device as the time-aligned second channel of a stereo file.

Stimuli were presented as simplified Chinese characters. To prevent incorrect readings of the stimuli, subjects were briefly familiarized with the wordlist and frame sentence before recording. The experiment was presented using OpenSesame experimental software (Mathôt, Schreij, & Theeuwes, 2012). The stimuli were presented in a different pseudo-random order in ten blocks, for a total of 320 tokens for each speaker. Stimuli were embedded in a frame sentence, which was also presented in simplified Chinese characters (我知道怎么读 __ *Wǒ zhīdào zěnme dú* __ 'I know how to read __.'); participants in each of the two groups were instructed to produce the resulting full sentence with either a Shanghai Mandarin or Standard Mandarin reading respectively. All testing was carried out by the second author, a Standard Mandarin speaker who does not speak Shanghainese. The stimuli were located in utterance-final position in order to avoid influence on the production of the nasal codas of nasal place assimilation to, or resyllabification into, a following syllable. The experiment was self-paced: Participants started and ended the ultrasound and audio recording process using keypresses. The entire recording session lasted approximately 40 minutes.

2.4. Analysis

Audio files were segmented with the Montreal Forced Aligner (McAuliffe, Socolof, Mihuc, Wagner, & Sonderegger, 2017). The resulting segment boundaries were inspected in Praat (Boersma & Weenink, 2019) and hand-corrected as required. For the 40 tokens recorded for each vowel-nasal combination, spectrograms were inspected to confirm the presence of a lingual closure, signaled by lowered intensity during the force-aligned /n/ or /ŋ/ interval. Tokens whose nasal codas did not demonstrate lingual closure were discarded so as not to present a confound to the analysis of lingual closure place. Nasal codas in the /i/ and /ə/ contexts demonstrated a relatively consistent lingual closure, and most tokens were retained (average number retained per speaker: 38.95 /in/, 38.5 /iŋ/, 37.35 /ən/, 37 /əŋ/). Tokens following the vowel /a/ were more frequently excluded (average retained: 22.9 /an/, 26.6 /aŋ/), consistent with prior data on both Shanghainese and Standard Mandarin demonstrating a tendency toward reduced or absent lingual constriction after low vowels (Xu, 1993; Chen, 2000; Duanmu, 2007, p. 24). Tokens following /ua/ were retained somewhat more often (32.3 /uan/, 35.4 /uaŋ/). Out of the 320 utterances per speaker collected, an average of 269 tokens per speaker remained for analysis.

In order to evaluate the tongue shapes each speaker used to produce the two nasal codas, a single representative frame at the midpoint of each stimulus-final nasal coda token was selected based on the force-aligned segment boundaries and the synchronization signal emitted by the ultrasound device. After applying a series of filtering operations to reduce speckle noise in the ultrasound signal, which improves the signal-to-noise ratio for the analysis (Carignan, 2014), an arc-shaped region of interest was selected for each frame (**Figure 1a–b**) and the set of midpoint frames for each speaker was submitted to a separate principal components analysis (PCA), which takes each frame's pixels as its numerical input.

The resulting principal components (PCs) capture patterns of covariation among the pixels in the set of frames. These patterns can be displayed as so-called eigentongues (Hueber et al., 2007), maps relating the new dimensions of variation to scores assigned to individual frames. The eigentongue shown in **Figure 1c**, for instance, captures much of the covariation in pixel values across one speaker's /n/ and /ŋ/ frames: PC1 scores positively correlate with brighter pixels in the red area and negatively correlate with brighter pixels
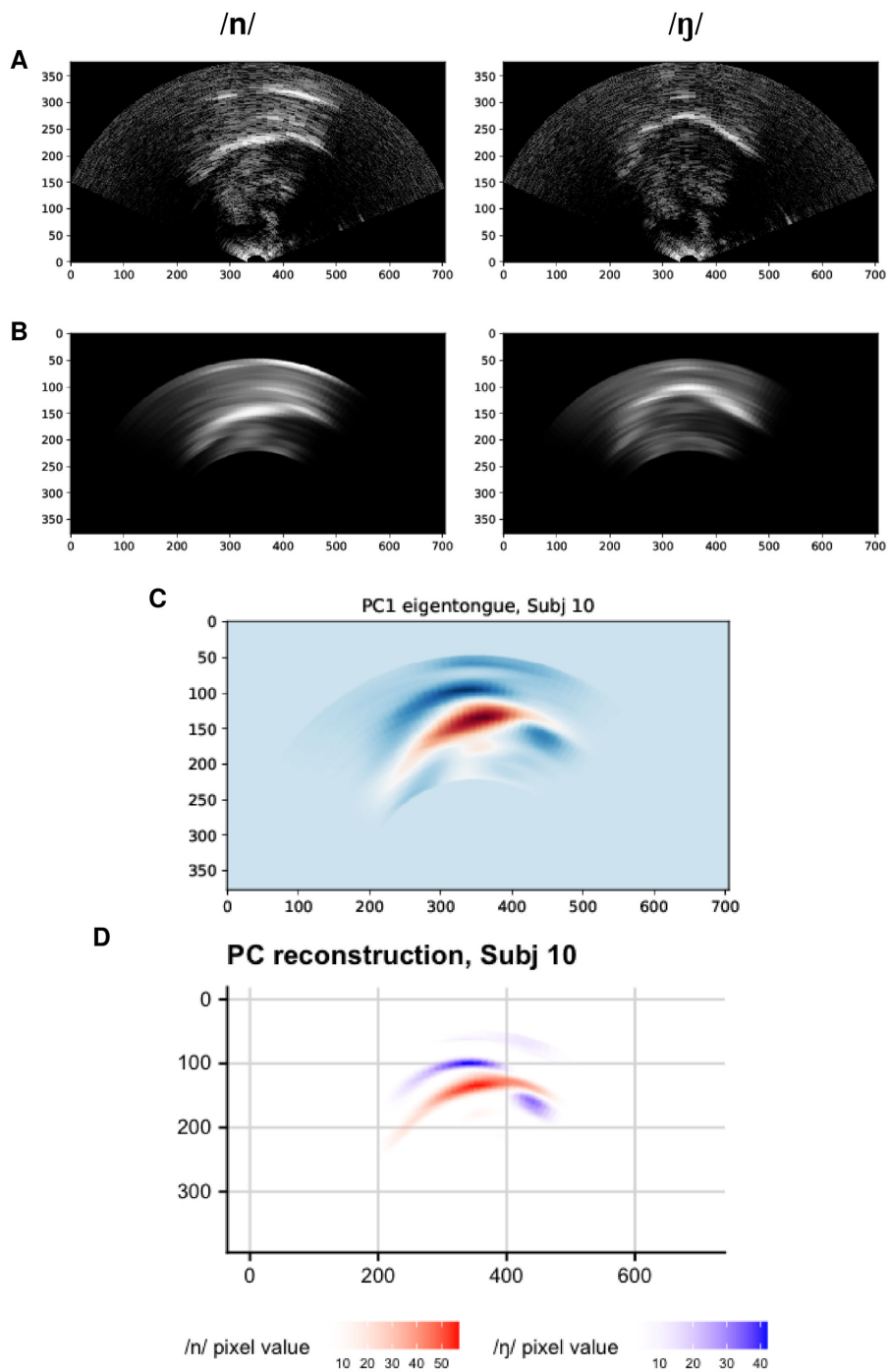
Figure 1: Sample data for speaker 10; anterior is right. **(a)** Raw tokens of /n/ (left) and /ŋ/ (right); **(b)** the same filtered and trimmed to region of interest. **(c)** PC1 eigentongue for full image set, showing much of the variation between pixel values for /n/ frames (high PC1 scores, red) and /ŋ/ frames (low PC1 scores, deep blue). **(d)** Reconstruction of typical tongue position in /n/ and /ŋ/ frames from eigentongues for PCs 1–10 (see Section 2.5).

in the deep blue area; these red and blue pixel clusters tend to have brighter values during /n/ frames and /ŋ/ frames, respectively (**Figure 1d**).

The first ten PCs were retained as a lower-dimensional representation of tongue shape for an individual speaker, explaining an average of 80.12% of variation (standard deviation = 2.57%). For each speaker, the scores for these PCs and the target of each token, i.e., velar /ŋ/ or alveolar /n/, were submitted to a linear discriminant analysis (LDA). The resulting

linear discriminant (LD) score, when normalized to a [0,1] range for all speakers, may be taken as an index of how distinctly /n/- or /ŋ/-like the coda nasal in each token is: The LDA was structured such that /n/ was consistently near 0, and /ŋ/ was consistently near 1, for all speakers. We refer readers to Hueber et al. (2007) and Hoole and Pouplier (2017) for more information on the application of principal components analysis to image data, and to Mielke, Carignan, and Thomas (2017), Strycharczuk and Sebregts (2018), and Kochetov, Faytak, and Nara (2019) for applications of similar dimensionality reduction procedures to ultrasound data.

Statistical comparisons across vowel context and/or across speaker groups were done using linear mixed-effects models using lme4 (Bates, Mächler, Bolker, & Walker, 2014) in R v3.5.3 (R Core Team, 2019). These models maximize statistical power despite an unequal number of observations under each condition; they also allowed us to evaluate the contribution of the independent variables while controlling for baseline differences across speakers and items using the highest random effects structure with intercepts and slopes that converged. We then present reconstructions based on individual speakers' first ten PCs' eigentongues to verify the specific tongue shapes employed in the coda nasal contrast.

2.5. Ultrasound results

To determine whether there was a merger in the production of Shanghai Mandarin /n/ and /ŋ/, we compared the classification accuracy of each speaker's LDA. The classification outcome for each item submitted to the LDA, on the basis of the linear discriminant function for each speaker's data, was compared to the actual category membership, and thus used as a measure of predictability of the coda nasals in production given their expected pronunciation as velar or nasal.

Classification accuracies for speaker-specific LDAs, pooled for all speakers (**Figure 2**), were in keeping with previous findings that the frequency of merger varies across vowel contexts. Accurate classification of nasal codas following /a/ and /ua/ was at or near ceiling for all speakers, suggesting very distinct means of production across both groups; this is in line with previous findings that the contrast is maintained after low vowels. Accurate classification of nasal codas following the non-low vowels /i/ and /ə/ was also at ceiling for control speakers, demonstrating that a contrast is maintained in production by Standard Mandarin speakers across all four vowel environments tested.

Classification accuracy for bilingual Shanghai Mandarin speakers was considerably lower for rhymes containing one of the non-low vowels /i/ or /ə/, and accuracy for the rhymes /əŋ/, /in/, and /iŋ/ was often at or below chance. To explore further, a mixed-effects linear regression was run to predict LDA accuracy, with fixed effects of speaker group, vowel,
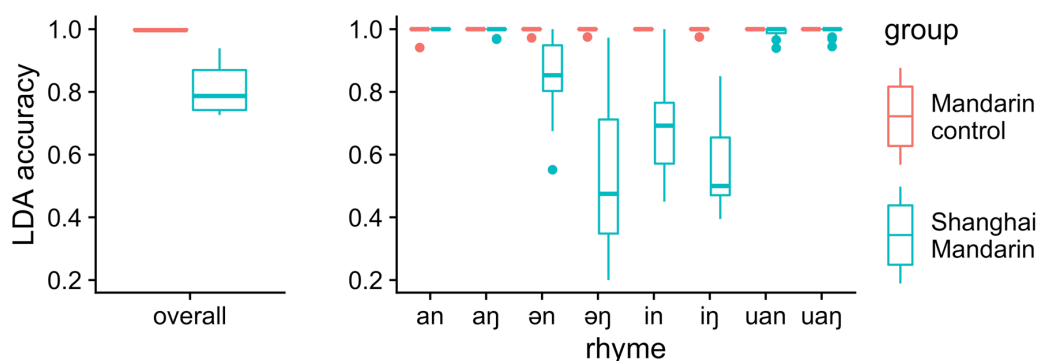

Figure 2: LDA classification performance on each speaker's basis data for nasal codas overall (left) and by rhyme (right).

coda nasal, as well as their interactions and random intercepts for subjects. There was a significant positive intercept, close to 1, indicating that LDA accuracy was overall quite high ($\beta = 0.99$; SE $= 0.05$; $t(133.3) = 19.8$, $p < 0.0001$). No main effects of group, vowel, or coda nasal on LDA accuracy reached significance, but there was a significant three-way interaction between speaker group, vowel, and coda nasal [$\chi^2(10) = 102$, $p < 0.0001$]. This interaction was driven by the Shanghai Mandarin group's less distinct production of nasal codas only in the context of non-low vowels. Post-hoc comparisons confirmed that compared to Standard Mandarin speakers, Shanghai Mandarin speakers were less accurate at producing the velar nasal in the /ə/ vowel context [$t(133) = 8$; $p < 0.0001$], and both alveolar [$t(133) = 5.2$; $p = 0.0001$] and velar nasals [$t(133) = 7.7$; $p < 0.0001$] in the /i/ vowel context. In sum, the LDA accuracy data show that alveolar and velar nasals produced by the Shanghai Mandarin group were less distinguishable than those produced by the Standard Mandarin group in the /ə/ and /i/ vowel context.

The LDA performance data considered so far only reveal patterns in *distinctiveness* or *similarity* of articulations used for the coda nasals. To appreciate the actual tongue configurations employed, we first examined the overall distribution of LD scores by segment and vowel context. We use the preceding low nuclei /a/ and /ua/ as a point of reference since they were consistently well discriminated, suggesting distinct articulatory configurations. Since lenition disproportionately resulted in discarded tokens of the rhymes /(u)an/ and /(u)aŋ/, we pool the /a/ and /ua/ data in the analysis that follows to even out the sample sizes. We view this as appropriate since inspection of the LD scores revealed no difference in coda nasal articulation conditioned by preceding /a/ and /ua/ nuclei. Figures comparing LD score distributions for /a/ and /ua/ are given in the appendix (Figures A1–A2).

Using mixed effects regression, we modeled the LD value as a function of speaker group (Standard or Shanghai Mandarin), intended nasal (alveolar or velar), and vowel context (/a/, /ə/, and /i/). We also included random intercepts for speaker and item. There was a significant three-way interaction between speaker group, intended nasal, and vowel context, as expected based on the LDA accuracy scores [$\chi^2(7) = 2643$, $p < 0.0001$]. Planned comparisons confirmed that Standard Mandarin control speakers used distinct tongue configurations for intended /n/ and /ŋ/ in all vowel contexts [/a/: z-ratio $= -30.783$; $p < 0.0001$; /ə/: z-ratio $= -23.677$; $p < 0.0001$; /i/: z-ratio $= -22.433$; $p < 0.0001$]. Shanghai Mandarin speakers also used distinct tongue configurations after /a/ and /ə/, but not after /i/ [/a/: z-ratio $= -28.112$; $p < 0.0001$; /ə/: z-ratio $= -8.011$; $p < 0.0001$; /i/: z-ratio $= -2.899$; $p = 0.1410$]. Below, we discuss the distributional differences in tongue configurations for intended alveolar and velar nasals in more detail.

The distribution of linear discriminant (LD) values by segment and speaker group is shown in **Figure 3** for the /(u)a/, /ə/, and /i/ vowel contexts. For the Standard Mandarin control group, regardless of preceding vowel context, alveolar nasals almost exclusively exhibited an LD value lower than 0.4 (98.5% of /n/ tokens), and velar nasals an LD value higher than 0.6 (99.7% of /ŋ/ tokens). We will henceforth refer to these as canonical ranges for the alveolar and velar nasals. The probability density distributions in the LD value for both nasals were compact and well-separated. The Shanghai Mandarin group had similar but less compact distributions for its coda nasals when preceded by /a/ (**Figure 3**, top right), with most alveolar nasals preceded by /a/ scoring lower than 0.4 (97.3% of /n/ tokens) and most velar nasals scoring higher than 0.6 (96.7% of /ŋ/ tokens). However, the distribution of LD values varied substantially from speaker to speaker for both of the Shanghai Mandarin group's coda nasals when preceded by /i/ or /ə/, with the overall result that many fewer canonical nasals were produced (58.5% of /n/ tokens below 0.4 and 32.1% of /ŋ/ tokens above 0.6).
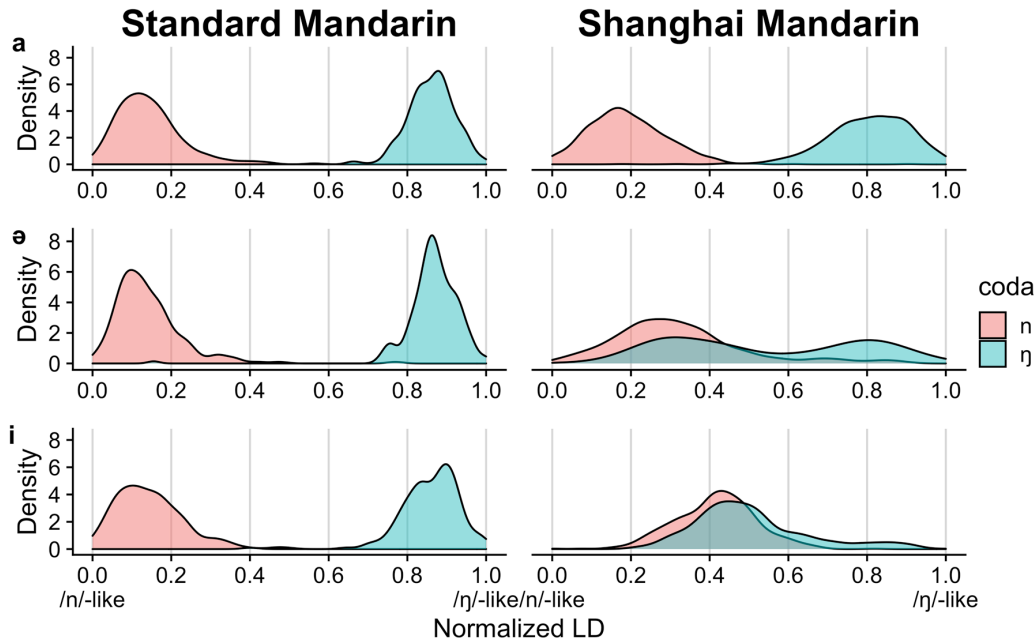
Figure 3: Distribution of LD values for nasal codas in the /(u)a/, /ə/, and /i/ contexts for the Standard Mandarin (left) and Shanghai Mandarin (right) groups.

Following /ə/, Shanghai Mandarin speakers' LD values (**Figure 3**, middle right) showed evidence of inconsistent coda nasal neutralization to /n/, a result mainly driven by inter-speaker variation between three major patterns shown in **Figure 4** (center). Two speakers showed control-like separation of /n/ and /ŋ/ after /ə/ (speakers 12 and 13). Five speakers had a near-complete overlap of the coda nasals after /ə/ (speakers 1, 3, 4, 6, and 19). The remaining speakers exhibited a more complex, bimodal distribution of the coda nasal in /əŋ/, which alternated between canonical /n/-like and canonical /ŋ/-like variants (i.e., speakers 15, 16, 17). Note that this pattern cannot be predicted by age of acquisition: All aforementioned speakers learned Standard Mandarin and Shanghainese simultaneously at home.

At the group level, this yielded a roughly unimodal, canonical /n/-like distribution for the /ən/ rhyme, with occasional /ŋ/-like productions contributing to a long tail approaching the /ŋ/ end of the LD scale. For /əŋ/, which had /n/-like characteristics much more often and in more speakers, there was an overall bimodal distribution containing /n/-like and /ŋ/-like peaks. The /ŋ/-like tokens were partly produced by non-merging speakers, but many were produced by speakers who variably produce the coda in /əŋ/ as /n/-like or /ŋ/-like. This distribution was driven in part by a bimodal distribution for all /əŋ/ items with the exception of 解梦 /t͡ɕiɛ↓.məŋ\/ 'dream reading' in which speakers nearly always produce a canonical coda [ŋ].

Following /i/, LD values for Shanghai Mandarin speakers' nasal codas suggest a more complete merger (**Figure 3**, bottom right). Most speakers did not distinguish between /n/ and /ŋ/ in production after /i/, as shown by the predominant pattern for individual speakers (**Figure 4**, right): a distinct cluster in the middle range of the LD value, flanked on either side by the canonical velar and alveolar nasals preceded by /(u)a/. This single non-canonical variant was typically centered around an LD value of about 0.4. While for many speakers these two distributions overlapped very closely (i.e., speakers 1, 13, and 19), others show more token-to-token variability in production, particularly for /iŋ/. Speakers 2 and 20, for instance, produced tokens of /iŋ/ ranging across nearly the entire LD range. This group generally exhibited two clear modes for both /n/ and /ŋ/ which
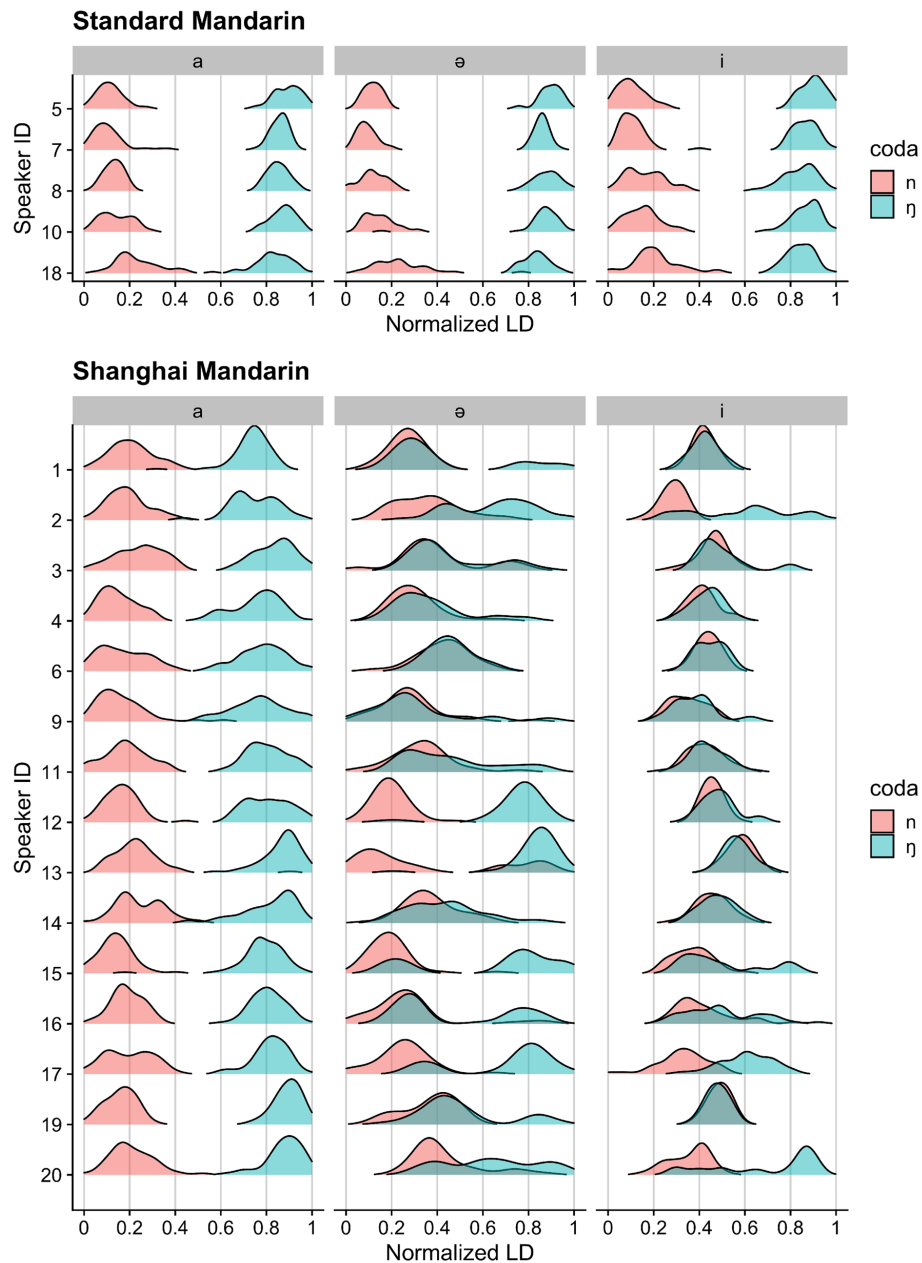
Standard Mandarin

Shanghai Mandarin

Figure 4: Distribution of LD values for nasal codas in the /(u)a/, /ə/, and /i/ contexts for each individual in the Standard Mandarin (top) and Shanghai Mandarin (bottom) groups. Trivially low probability density values (below 0.01) are not plotted for readability.

fell just outside of the 0.4–0.6 range for non-canonical variants, but which nonetheless differed substantially from the modes for post-/a/ vowels. A small group of speakers (i.e., Speakers 2, 17, and 20) distinguished the two post-/i/ coda nasals to some extent, but there was considerable overlap between the categories, and some speakers (i.e., Speakers 15 and 16) tended toward a bimodal distribution for /iŋ/ with a small subset of tokens that were produced like canonical /ŋ/. This pattern does not appear to be driven by lexical factors as is the case for /əŋ/ above, since it affects all /iŋ/ lexemes. Rather, unclear mediating factors keep some /iŋ/ tokens distinct.

All in all, these data show that most Shanghai Mandarin speakers produced a single non-canonical variant after /i/. The remainder of this section is given over to identifying these non-canonical variants, the identity of which cannot be deduced from LD values alone:

The LD encodes only similarity to the two maximally distinct articulations at either end, with the middle of the LD scale not necessarily representing an interpolation between these two variants. Anything not particularly similar to either maximally distinct nasal stop will fall in between; in this case, such segments could be nasal stops with full closure at a mid-palatal location, but could also potentially be reduced coda nasals which do not achieve full closure.

We used the PC scores which serve as the basis data for the LDA to determine the specific tongue shapes used for the nasal coda of each rhyme. In doing so, we also aim to rule out a contribution of coda nasal lenition or deletion to intermediate LD values, though efforts were taken to remove tokens of lenited nasals from inclusion in the analysis. Reconstructed typical ultrasound frames were created using the average value of each of the first ten PC scores from the PCA described in Section 2.4. The ten eigentongues, or maps of covariance in the original signal's pixels described by each PC, obtained from the same analyses were multiplied by the resulting ten average PC scores and summed to obtain a graphical representation of the typical production for both nasal codas when preceded by the rhymes /i, ə, a, ua/. We used tongue reconstruction for this analysis because it is fully automated; in contrast, the more typical semi-automated process of tongue surface contour extraction and spline fitting requires manual annotator correction for accurate extraction of tongue surfaces from noisy ultrasound data. Such interventions are time-consuming and have been reported to decrease replicability (see Hoole & Pouplier, 2017; Roettger, 2019).

The resulting reconstructions (**Figure 5**) show displacements of the tongue dorsum and blade (for /ŋ/ and /n/, respectively) consistent with lingual closure during the nasal codas. As expected from the LD score data, Standard Mandarin control speakers employed distinct tongue shapes which align well with the intended place of the coda nasal (**Figure 5**, top; see Figure A3 for additional speakers). Investigation of the typical tongue shapes used by
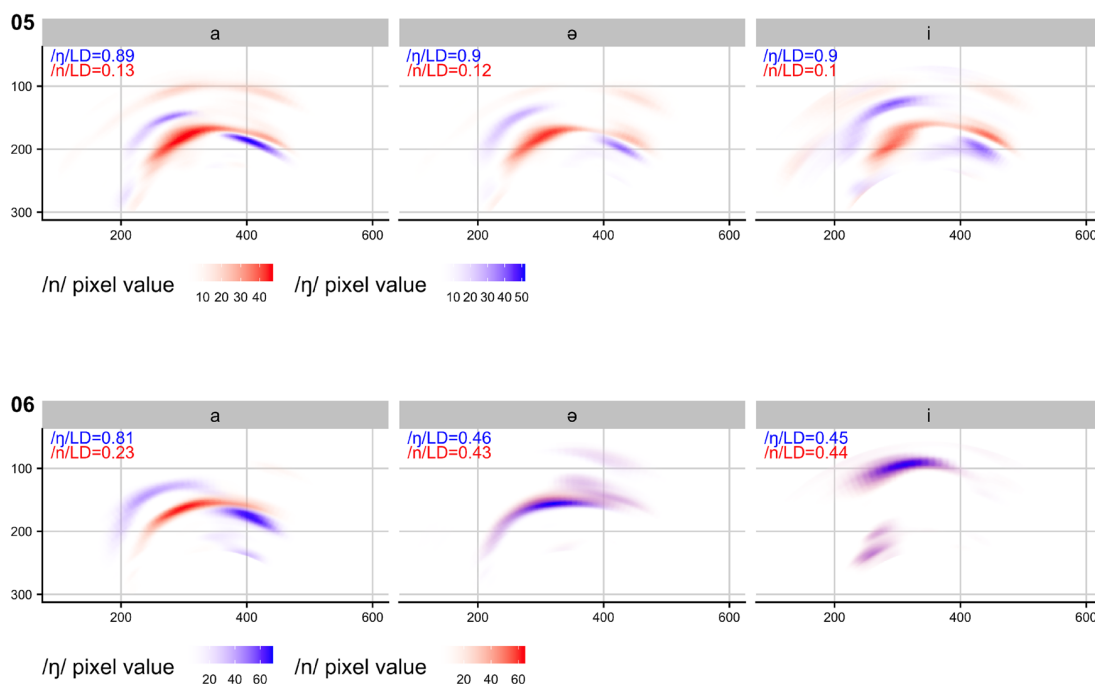

Figure 5: Mean reconstructed ultrasound frames at midpoint of coda nasals for a typical Standard Mandarin control subject (top) and Shanghai Mandarin subject (bottom). Anterior is right. Estimated pixel values below zero, which are artifacts of reconstruction, are flattened to zero. Note the overlap of estimated contours for the Shanghai Mandarin speaker's /n/ (red) and /ŋ/ (blue) in the /ə/ and /i/ context.

speakers in the Shanghai Mandarin group suggests that the non-canonical, merged nasal codas used by most of this group after /i/ were in fact articulated with tongue shapes intermediate in place between the alveolar and velar nasals (**Figure 5**, bottom; see Figure A4 for additional speakers). The mergers observed for Shanghai Mandarin participants in the LD data can be seen in the reconstructions: The two columns which contain tokens produced after /i/ and /ə/ (center, right) contain appreciably similar shapes, resembling /n/ in the case of the /ə/ context and a fronted velar in the /i/ context.

Darker shading in the reconstructed frames indicates that the tongue surface is more consistently present at that location. The speakers shown in **Figure 5** have a compact, unimodal distribution of productions for /n/ as well as /ŋ/, which is reflected by the clear reconstructions for each coda-vowel pair. Faint tongue contours which reflect variable production should not be confused with faint artifacts of reconstruction appearing at the edges of some frames (i.e., arcs at about y = 100 in **Figure 5**, top). Shanghai Mandarin speakers with more variable productions (i.e., multimodal distribution of LD value, particularly for /ŋ/) show fainter or more diffuse reconstructions; see Figure A4 in the appendix for examples.

2.6. Interim discussion

The ultrasound data presented above provide a nuanced position on coda nasal neutralization in Shanghai Mandarin. There are striking differences in both the outcome and consistency of neutralization among the vowel environments examined here. In keeping with previous literature, nasal stops before /a/ and /ua/ were articulated with two distinct tongue shapes, even by Shanghai Mandarin speakers. Nasals after /ə/ were articulated distinctly by Standard Mandarin control speakers and by some in the Shanghai Mandarin group. However, many in the Shanghai Mandarin group produced intended /ŋ/ in this environment as either canonical [ŋ] or canonical [n], with little in the way of intermediate articulations. This is evident from the bimodal distribution of velar nasals after /ə/ on the LD. This effect was asymmetrical: The LD value distribution of /ŋ/ was bimodal, typically merging to [n] but sporadically produced as [ŋ]; /n/ however, was largely unimodal and produced as [n]. These findings replicate and extend Chiu et al.'s (2019) findings of inconsistent merger after schwa in Taiwan Mandarin.

Nasals after /i/ presented the most consistent evidence of neutralization: Unlike post-/ə/ tokens, nasal productions by the Shanghai Mandarin group after /i/ were unimodally distributed and centered in the non-canonical middle of the LD scale. These non-canonical variants were produced by most Shanghai Mandarin speakers and, upon reconstruction, did not closely resemble either coda [n] or coda [ŋ] after low vowels, a context that does not promote neutralization. Our findings in the /i/ vowel context differ from Chiu, Lu, et al.'s (2019), who report velar nasals as an outcome of merger after /i/.

We argue that the merged non-canonical nasals are fronted velar or palatal nasals, fronted strongly enough as to be difficult to classify as canonical velars on the LD. This tongue configuration is in keeping with the occasional description of coda nasalization after /i/ in Shanghainese as palatal (Qian, 2003). If so, it may be due to transfer from Shanghainese to L2 Mandarin, since this degree of vowel-to-nasal coarticulation after /i/ is not observed in the control Standard Mandarin group. However, a more parsimonious account is that in Shanghai Mandarin this configuration is simply a result of coarticulation with the high vowel /i/. A similar explanation based on articulatory ease has been proposed for Taiwan Mandarin in Chiu, Lu, et al. (2019).

In the following sections, we made use of materials produced by speakers in the production study to conduct a follow-up perception study for nasal codas after /i/, in light of known perceptual difficulties in distinguishing the nasal codas in this context (i.e.,

Mou, 2006). We restricted the follow-up perception experiment described below to the /i/ context because after /ə/, any variation present for speakers resembled a substitution error of [n] for intended [ŋ], the rate of which varied among lexemes. Neutralization after /i/ was observed more consistently across all items, and thus likely creates more general perceptual difficulties.

3. Perception study

We conducted a perception experiment to determine the extent of neutralization between the alveolar and velar nasals in Shanghai Mandarin and Standard Mandarin. All study procedures described below were approved by the UCLA General Institutional Review Board.

As we saw from the ultrasound data, Shanghai Mandarin speakers had the most consistent neutralization of velar nasals after the vowel /i/. Thus, we used an AXB task to assess Shanghai Mandarin and Standard Mandarin listeners' ability to distinguish place of articulation of nasal codas in the context of the vowel /i/. We chose to use an AXB task instead of an identification task in order to minimize listeners' bias towards specific labels. In an AXB task, listeners simply determine whether the middle syllable is more like the first or the third one, without a need for explicit labels. The anchors, that is, the first and the third syllable, were always produced by Standard Mandarin speakers, and were confirmed to be canonical tokens of alveolar and velar nasals based on their LD values. To encourage phonemic categorization, we varied the speaker across each of the three syllables presented within a trial: Listeners had to ignore between-talker variability to successfully categorize the nasal codas.

If alveolar and velar nasal codas are completely merged in Shanghai Mandarin, we expected Shanghai Mandarin (and Standard Mandarin) listeners to be at chance at categorizing Shanghai Mandarin nasals. In that case, Shanghai Mandarin, but not Standard Mandarin, listeners were additionally expected to fail to categorize Standard Mandarin nasals as well. It is also possible that the two groups of speakers differ in whether they completely neutralize the distinction or signal it with cues other than the place of the nasal, such as the quality of the preceding vowel (Rhyme Harmony; see Section 1.1). In that case, we expected listeners in each group to be better at distinguishing nasals produced by their own group of speakers, but not the other.

3.1. Participants

Fourteen Standard Mandarin listeners (7 males, 7 females, mean age = 24.5) and 14 Shanghai Mandarin listeners (3 males, 11 females, mean age = 29.1) were recruited for the perception experiment. Listeners were compensated with course credit or cash payment ($12 USD). No participants reported a history of speech or hearing disorders. All participants filled out a questionnaire for their residential history and language background. None of the Standard Mandarin participants had experience with Shanghainese. All Shanghai Mandarin bilingual listeners self-rated their level of Shanghainese ability as native or native-like (at a minimum of 6 on a 7-point Likert scale). All listeners in the Shanghai Mandarin bilingual group acquired Shanghainese and Mandarin simultaneously in China except for listener 3, who acquired Mandarin at age 5, and listener 13, who acquired Mandarin at age 18 in the United States. The pattern of results reported later remains the same even when we exclude these two participants.

3.2. Stimuli

Stimuli were drawn from the audio recordings collected in the production experiment described above. Only nasal codas that followed the vowel /i/ were selected because they demonstrated the most consistent pattern where intended /ŋ/ fell in the intermediate

range described previously as non-canonical (0.4–0.6). Target syllables were extracted from the original recordings and trimmed to exclude the syllable onset, which varies from token to token (see **Table 3** and Table A1 in Appendix). Stimuli were normalized to an average amplitude of 70 dB.

Canonical and non-canonical /iŋ/ and canonical /in/ were selected for Shanghai Mandarin speakers based on LD value. Many Shanghai Mandarin speakers produced both canonical and non-canonical variants of the velar nasal after /i/, often varying in which variant was used across repetitions of a single word. In the remainder of this section, we use /N/ as a shorthand for non-canonical variants of velar nasals. Intended /iŋ/ with an LD value above 0.6 were treated as canonical /iŋ/, and intended /iŋ/ with an LD value between 0.4 and 0.6 were treated as non-canonical /iN/. Nearly all /in/ were canonical; canonical /in/ were selected with an LD value below 0.4. One canonical /in/ token produced by Standard Mandarin control speaker 08 and one canonical /iŋ/ produced by Standard Mandarin control speaker 18 served as anchor files in positions A and B in the AXB task. Twenty Standard Mandarin /in/ tokens and 20 Standard Mandarin /iŋ/ tokens (all canonical) produced by control speaker 7 were used as controls in experimental position X. Other test stimuli were drawn from four Shanghainese speakers (speakers 2, 16, 17, 20; 2 females, 2 males) who produced the full range of canonical and non-canonical nasal coda types described above. All three types of stimuli were drawn from each speaker's materials: 14 canonical /in/, 7 canonical /iŋ/, and 7 non-canonical /iN/. This led to a total of 112 (56 canonical /in/, 28 canonical /iŋ/, and 28 non-canonical /iN/) Shanghai Mandarin test stimuli.

The acoustic characteristics of the vowels preceding nasals in Standard Mandarin and Shanghai Mandarin are summarized in **Tables 4** and **5**, respectively. As shown in **Table 4**, Standard Mandarin alveolar and velar nasals could potentially be distinguished by the duration of the preceding vowel, because the force-aligned vowels were longer before alveolars compared to velars; furthermore, as expected, F3 at vowel midpoint was lower for velars compared to alveolars. Note that the vowel duration differences in the perception stimuli in Standard Mandarin were not consistent with schwa insertion before velars, which would be expected to increase rhyme duration. Before velars, F1 was also higher at the vowel endpoint compared to the vowel midpoint, the opposite of what might be expected with schwa insertion (Mou, 2006). **Table 5**, in contrast, shows that

Table 4: Summary of vowel duration and mean formant values (Min:Max) for perceptual stimuli produced by Standard Mandarin speakers. Results from t-tests comparing each potential cue in alveolar and velar stimuli are shown in the "Cue status" column.

Cue	Mean (min:max)		Cue status
	/n/	/ŋ/	
Vowel duration (ms)	192 (134:236)	128 (62:200)	$t(38) = 5.9, p < 0.001$
Formant frequencies at vowel midpoint			
F1 (Hz)	509 (406:564)	516 (451:589)	n.s.
F2 (Hz)	2381 (1083:2907)	2616 (1225:3011)	n.s.
F3 (Hz)	3498 (2554:3819)	3305 (2784:3821)	$t(38) = 2.14, p = 0.04$
Formant frequencies at vowel endpoint			
F1 (Hz)	474 (190:584)	503 (404:593)	n.s.
F2 (Hz)	1904 (530:2379)	2198 (1566:3108)	n.s.
F3 (Hz)	3226 (2241:3717)	3306 (2789:3871)	n.s.

Table 5: Summary of vowel duration and mean formant values (Min:Max) for perceptual stimuli produced by Shanghai Mandarin speakers. Results from t-tests comparing each potential cue in alveolar and either canonical or non-canonical velar stimuli are shown in the "Cue Status" column; no comparisons reach significance.

Cue	Mean (min:max)			Cue status
	/n/	Canonical /ŋ/	Non-canonical /N/	
Vowel duration (ms)	192 (134:236)	128 (62:200)	130 (75:198)	$t(38) = 5.9$, $p < 0.001$
Formant frequencies at vowel midpoint				
F1 (Hz)	509 (406:564)	516 (451:589)	515 (387:709)	n.s.
F2 (Hz)	2381 (1083:2907)	2616 (1225:3011)	2294 (954:2975)	n.s.
F3 (Hz)	3498 (2554:3819)	3305 (2784:3821)	3045 (2329:3728)	$t(38) = 2.14$, $p = 0.04$
Formant frequencies at vowel endpoint				
F1 (Hz)	474 (190:584)	503 (404:593)	486 (264:729)	n.s.
F2 (Hz)	1904 (530:2379)	2198 (1566:3108)	1999 (521:2985)	n.s.
F3 (Hz)	3226 (2241:3717)	3306 (2789:3871)	3022 (2290:3763)	n.s.

neither preceding vowel duration nor formant frequencies at vowel midpoint or endpoint distinguished upcoming alveolar or velar nasal place in Shanghai Mandarin stimuli.

3.3. Procedure

The performance of each listener was evaluated using an AXB discrimination task presented using MATLAB. Each participant heard three syllables in each AXB trial, each separated by 500 ms. They were informed that these syllables were produced by different talkers. The first and third syllables were always distinct: one a canonical /n/ and the other a canonical /ŋ/. Participants were asked to determine if the middle syllable was more similar to the first or third syllable by clicking buttons "choice 1" or "choice 2." They were asked to disregard the differences in speakers, vowels, and tones and instead to focus on the differences between the final consonants in each item. No feedback was provided. The experiment took place in a sound-treated room in the UCLA Phonetics Lab. Stimuli were presented through 3M PELTOR (HTB79A-02) closed-system over-ear headphones. The experiment began with ten practice trials to familiarize subjects with the task. In the test phase that followed, participants were presented 304 trials in one sitting. These trials included counterbalancing of 152 files by switching the order of anchor files in position A and B, which contained one canonical /n/ and one canonical /ŋ/ produced by two Standard Mandarin control speakers. The presentation of trials was randomized. The experiment took about 40 minutes.

3.4. Analysis

To determine whether Shanghai Mandarin and Standard Mandarin listeners were able to distinguish coda nasals, we used mixed-effects logistic regression (Bates et al., 2014) in R v3.5.3 (R Core Team, 2019). The model included the between-subjects variable, listener group (Standard Mandarin or Shanghai Mandarin), the within-subjects variable, stimulus type (Standard Mandarin, Shanghai Mandarin canonical, or Shanghai Mandarin non-canonical), and their interaction. Standard Mandarin alveolar nasals and Standard Mandarin listeners were coded as reference conditions. Finally, random intercepts for individual speakers of stimuli and individual listeners were also included. This was the maximal random effects structure with which the model converged. Fixed effects were evaluated by comparing the log likelihood of two models with and without the variable of

interest using a Likelihood Ratio test in R. Follow-up planned comparisons were conducted using emmeans (Lenth, 2020).

To determine how well Shanghai Mandarin and Standard Mandarin listeners were able to distinguish the different stimulus types, we used d-prime. Values were calculated separately for each listener group using R v3.5.3, for a total of six groupings by stimulus type (Standard Mandarin, Shanghai Mandarin canonical, and Shanghai Mandarin non-canonical) and listener type (Standard Mandarin, Shanghai Mandarin).

3.5. Perception results

The full logistic regression model had a significant negative intercept (β = –0.8; SE = 0.32; z-value = –2.5, p = 0.01), indicating that all listeners had a bias towards categorizing nasals as Standard Mandarin /ŋ/. Additionally, the interaction of listener group and stimulus type was significant [$\chi^2(4)$ = 33.98, p < 0.001].

The effects plot for the interaction term is presented in **Figure 6**. Planned comparisons confirmed that Standard Mandarin listeners were able to distinguish Standard Mandarin /n/ from Standard Mandarin /ŋ/ [z-ratio = –9.1; p < 0.0001], Shanghai Mandarin /n/ [z-ratio = –4.4; p = 0.0005] and Shanghai Mandarin canonical /ŋ/ [z-ratio = –4.5; p = 0.0003] and non-canonical /N/ [z-ratio = –4.3; p = 0.0007]. But they failed to distinguish between Standard Mandarin /ŋ/, Shanghai Mandarin /n/, and Shanghai Mandarin /N/. That is, Standard Mandarin listeners were only able to distinguish between Standard Mandarin /n/ and canonical /ŋ/, and categorized all Shanghai Mandarin nasals the same as Standard Mandarin /ŋ/.

In contrast, Shanghai Mandarin listeners failed to distinguish Standard Mandarin /n/ from any other stimulus type [Standard Mandarin /ŋ/: z-ratio = –2.2; p = 0.5; Shanghai Mandarin /n/: z-ratio = –2.8; p = 0.1; Shanghai Mandarin canonical /ŋ/: z-ratio = –2.2; p = 0.5; Shanghai Mandarin noncanonical /N/: z-ratio = –2.25; p = 0.4]. That is,
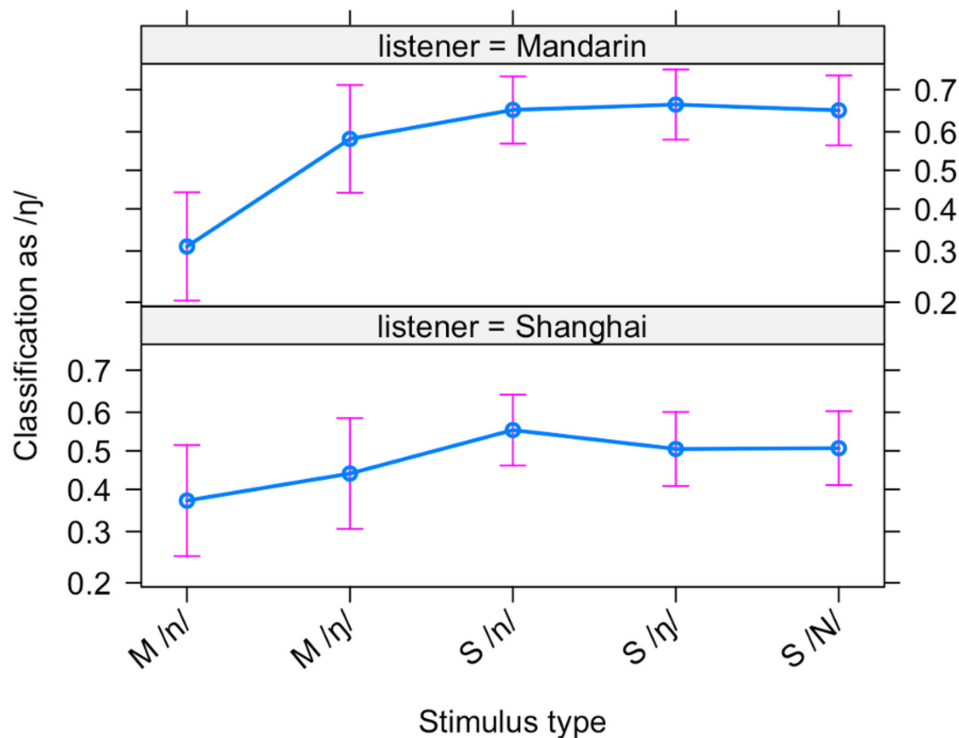

Figure 6: Effects plot for mixed-effects logistic regression. Standard Mandarin /n/ and the Standard Mandarin listener group served as reference levels.
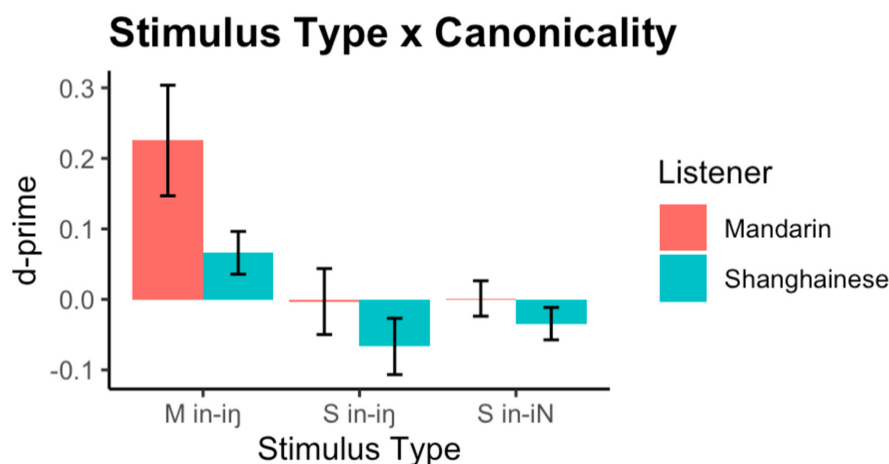
Figure 7: d-prime for AXB discrimination task by talker and stimulus type.

Shanghai Mandarin listeners failed to distinguish between /n/ and /ŋ/ produced by either Standard Mandarin or Shanghai Mandarin speakers.

Furthermore, we can see from their d-prime values that even Standard Mandarin listeners were not too successful in distinguishing Standard Mandarin /n/ and /ŋ/ (**Figure 7**). Native-like performance typically results in d-prime values of 3; Standard Mandarin listeners had average d-prime values of 0.2. As expected from the logistic regression results reported previously, d-prime values were close to zero for Shanghai Mandarin nasals; their d-prime value for the Standard Mandarin /n/-/ŋ/ contrast was only marginally higher.

In summary, we see from the perception results that Shanghai Mandarin listeners failed to distinguish both Standard Mandarin and Shanghai Mandarin /in/-/iŋ/, providing evidence for neutralization of coda nasal place after /i/ in Shanghai Mandarin. Standard Mandarin listeners also failed to distinguish Shanghai Mandarin /in/-/iŋ/, regardless of the similarity of the articulatory configuration to canonical productions. Additionally, although they were able to distinguish Standard Mandarin /in/-/iŋ/, the near zero d-prime value attests to a near merger in Standard Mandarin as well.

4. Discussion

In this study we gauged the extent of nasal coda neutralization in Shanghai Mandarin in both production and perception, and determined the outcome of the neutralization process in terms of similarity to the unmerged nasals produced in contexts that do not promote neutralization. Using ultrasound recordings, we showed that in Shanghai Mandarin the extent and outcome of neutralization is conditioned by the vowel context. The alveolar-velar distinction in nasal place was maintained after the low-vowel /a/ and /ua/, but not after /ə/ or /i/. Following the mid vowel /ə/, the two coda nasals neutralized to a nasal resembling /n/, but with a tendency for speakers to produce occasional /ŋ/ in a few words having etymological /ŋ/. Following the high vowel /i/, the two nasals typically neutralized to a fronted velar or palatal nasal which did not resemble unmerged realizations of /ŋ/ in low vowel environments. The perception experiment further confirmed that Shanghai Mandarin listeners were unable to distinguish /n/ and /ŋ/ following /i/. Thus, for speakers of Shanghai Mandarin, the nasal contrast between alveolar and velar nasal was completely neutralized in the context of /i/.

The pattern of neutralization observed in this study in Shanghai Mandarin differs from the pattern observed in Shanghainese: Recall that Shanghainese lacks a coda nasal place contrast altogether. Because reduction of the coda nasal contrast in Shanghai Mandarin is observed in some but not all vowel contexts, it is unlikely to be driven by L1 interference

alone as has been previously proposed. Coda nasal place neutralization in Shanghai Mandarin also does not always result in nasals which closely resemble either [n] or [ŋ], as claimed previously based on transcription data (Yang, 2010; Fon et al., 2011; Luo, 2015; Guan, 2019). In our ultrasound data, while speakers did produce a coda similar to alveolar [n] after the mid vowel /ə/, after /i/, speakers' nasals were neutralized to a fronted velar place which resembled the preceding /i/ more than any non-neutralized segment. In other words, the outcome of neutralization was structure-preserving (like one of the input segments, i.e., [n] or [ŋ]) after /ə/, but not after /i/.

These findings highlight how moving away from transcription, and towards the use of direct articulatory recordings, is critical to pinpointing the outcomes of contrast reduction. Transcription-based investigations of neutralization are likely to be affected by the perceptual biases of the annotators, including the influence of whatever orthographic representation is used to indicate the choices. Given the ease with which perceptual biases can warp interpretation of the data, and the popularity of identification tasks in this literature, non-structure-preserving outcomes in mergers and neutralizations of the sort studied here may even be undercounted in the relevant phonetic and phonological literature in the absence of direct articulatory recordings. We thus view direct articulatory recordings as a useful complement to transcription evidence to confirm the articulatory implementation of neutralized or merging contrasts and to add to our understanding of contrast reduction processes (see Yu, 2011).

Our ultrasound data show evidence for complete neutralization following /i/. Chiu, Lu, et al. (2019) describe a similar pattern in Taiwan Mandarin, but describe the neutralized place of articulation after /i/ in Taiwan Mandarin as velar, and thus structure-preserving. Chiu, Lu, et al., however, do not evaluate the degree of fronting of velar nasals after /i/ relative to those encountered in low vowel contexts, a comparison made here with the reconstruction of typical frames. We conjecture that such a reconstruction of tongue shape for nasals after /i/ in Taiwan Mandarin is likely to replicate the results reported here, and identify the neutralized nasal as a fronted velar or palatal. This midpalatal constriction location, and its likely ambiguous acoustic effect, may explain the highly variable conclusions for the /i/ context in the literature on Taiwan Mandarin. Comparison of tongue position during the articulation of fronted and non-fronted velars by speakers of Taiwan Mandarin will be necessary to test this prediction.

Based on three-dimensional simulations of tongue shape and the muscular activations required to achieve these shapes, Chiu, Lu, et al. (2019) argue that the place of articulation of coda nasals neutralized after /i/ is determined by biomechanical constraints. Specifically, they argue that producing an [iŋ] sequence involves less muscular strain than producing an [in] sequence, which promotes neutralization to [ŋ] after /i/. However, counter to what is actually observed in our and Chiu, Lu, et al.'s data, these simulations also predict neutralization to [ŋ] after /ə/, since muscular strain is lower for an [əŋ] sequence compared to an [ən] sequence as well. A further limitation of such an account is that effort reduction alone cannot explain differences in neutralization outcome observed after /i/ versus /ə/, with only the latter being structure-preserving.

We suggest a different account in which the extent of structure preservation is modulated by the coarticulatory resistance of the conditioning environment, here the preceding segment. Structure preservation in sound changes due to misperception is thought to result from resolving ambiguous inputs based on forms more frequent in the language (Blevins, 2009; Kavitskaya, 2014). Our data suggest that this type of structure preservation has a demonstrable effect on sound change only in contexts with low coarticulatory resistance. Compared to /ə/, the higher vowel /i/ has a higher coarticulatory resistance, and thus more coarticulatory 'aggression' (Recasens & Espinosa, 2009; Chen et al. 2015; Recasens

& Rodríguez, 2016). Producing a lingual closure while minimizing the articulatory excursion from a preceding /i/ results in a non-structure preserving palatal nasal that neither resembles unmerged /n/ nor /ŋ/. This relationship between the coarticulatory resistance of the conditioning environment and the outcome of contrast reduction merits further investigation, given that it may constrain the outcomes of certain neutralizations.

It is also possible that the greater extent of neutralization in the /i/ compared to /ə/ context can be attributed to the lower perceptual salience of the nasal place contrast in the high vowel context. This is consistent with proposals that speakers tolerate articulatory simplifications only if they are 'perceptually inconspicuous' (Kohler, 1990; Kawahara & Garvey, 2014). Recall that in our perception experiment Shanghai Mandarin listeners failed to distinguish nasals following /i/, whether produced by Shanghai Mandarin speakers or Standard Mandarin speakers. Whether Shanghai Mandarin listeners are better at distinguishing nasals following /ə/, as such an account would predict, remains to be determined.

What was more surprising was that both Shanghai Mandarin listeners and Standard Mandarin listeners failed to correctly categorize 'canonical' productions by Shanghai Mandarin speakers. (Recall that at least some intended /in/ or /iŋ/ tokens produced by Shanghai Mandarin listeners had linear discriminant scores similar to canonical productions by Standard Mandarin speakers.) We think this might be because Standard Mandarin listeners attend to vowel dynamics in addition to, or in place of, the nasals themselves, since schwa insertion may cue the presence of a velar nasal coda after /i/ for some populations (see Section 1.1, Mou, 2006; Duanmu, 2007). Whether this is indeed the case remains to be confirmed in future research.

The perception and production experiments led to an additional surprising finding about Standard Mandarin. All speakers produced distinct coda /n/ and /ŋ/ in every vowel context. However, Standard Mandarin control listeners' discrimination of the Standard Mandarin /n/-/ŋ/ contrast after /i/ was close to chance. These findings are consistent with the perceptual difficulties with nasals in the context of /i/ documented in Mou (2006). More crucially, the data presented here suggest a *near merger* (see Yu, 2011) for alveolar and velar nasal codas following /i/ even for Standard Mandarin speakers.

So how can Standard Mandarin speakers still produce distinct alveolar and velar nasal codas in the context of /i/? It has been argued in other cases of incomplete neutralization that speakers might continue to distinguish between neutralized segments in articulation because of supporting orthographic differences (Fourakis & Iverson, 1984; Warner et al., 2004; Warner, Good, Jongman, & Sereno, 2006; Roettger et al., 2014). In our study, we displayed stimuli using simplified Chinese characters, a mostly ideographic script which does not consistently signal the place of articulation of segments. Nonetheless, it is possible that even the presentation of characters in the reading task elicited more canonical productions than are typical in spontaneous speech. Alternately, associations among characters and Hanyu Pinyin, a romanization which orthographically indicates the difference between the codas, may also have resulted in an exaggerated production difference between the alveolar and velar nasal codas. Finally, it is also possible that using minimal pairs to elicit nasals in our production experiment resulted in exaggerated articulatory differences between the alveolar and velar nasals (see Port & Crawford, 1989; Kharlamov, 2014). Further investigations of Standard Mandarin spontaneous speech are needed to confirm whether production differences between alveolar and velar nasal codas following /i/ are robust.

5. Conclusion

In this study we have presented a proof of concept for the usefulness of complementary articulatory methods in the study of contrast reduction. The improved characterization of the nasal place contrast for individual speakers in the three vowel contexts studied

using direct articulatory recordings allowed us to arrive at a more nuanced understanding of the effects of substantive factors on the process of positional neutralization, and on structure preservation in sound change more generally. Our combined ultrasound tongue imaging and perception results suggest that neutralization of nasal coda place in Shanghai Mandarin does not occur after low vowels for any speaker examined, but that the non-low vowels /i/ and /ə/ do condition neutralization. Additionally, neutralization after /i/ is not structure-preserving, resulting in a fronted velar or palatal nasal which does not occur elsewhere in Shanghai Mandarin. Unexpectedly, we also found that although Standard Mandarin control speakers produce the nasal place distinction after /i/, they cannot reliably distinguish it in perception, suggesting a near-merger in this context for this group. Whether neutralization of Shanghai Mandarin nasal codas following /i/ is due to biomechanical properties as we have posited remains to be confirmed.

Additional Files

An Open Science Framework project page (https://osf.io/rmpxz/) has been created to store the scan-converted, filtered ultrasound frames used for the articulatory analysis in compressed format. The project page also contains a Jupyter notebook which provides a means of viewing the ultrasound data and reproducing the dimensionality reduction analysis described in Section 2.4 for a reference speaker. We additionally provide the following files:

- **Appendix.** A PDF file containing the full list of stimuli and supplemental figures (linear discriminant scores for /ua/ and eigentongue reconstructions for each individual speaker). DOI: https://doi.org/10.5334/labphon.269.s1
- **Supplementary Material 1.** A ZIP file containing numerical data obtained in the ultrasound study (linear discriminant values and classification accuracies) for each speaker. DOI: https://doi.org/10.5334/labphon.269.s2
- **Supplementary Material 2.** A CSV file containing numerical data obtained in the perception study (discrimination results; d-primes). DOI: https://doi.org/10.5334/labphon.269.s3

Acknowledgements

Competing Interests

The authors have no competing interests to declare.

Author Contributions

Matthew Faytak designed the production experiment and analyzed the articulatory data, and Suyuan Liu and Megha Sundara designed the perception experiment. Matthew Faytak and Suyuan Liu collected the articulatory data, and Suyuan Liu collected perception data. All authors contributed to interpretation of data and writing the manuscript.

References

Bakovic, E. (2000). Nasal place assimilation in Spanish. *University of Pennsylvania Working Papers in Linguistics, 7*(1), article 2. https://repository.upenn.edu/pwpl/vol7/iss1/2

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2014). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67*(1), 1–48. DOI: https://doi.org/10.18637/jss.v067.i01

Blevins, J. (2009). Structure-preserving sound change: A look at unstressed vowel syncope in Austronesian. In A. Adelaar & A. Pawley, (Eds.), *Austronesian historical linguistics and culture history: a festschrift for Robert Blust* (pp. 39–55). Canberra: Pacific Linguistics.

Boersma, P., & Weenink, D. (2019). *Praat, a system for doing phonetics by computer*. Software. Version 6.0.50, retrieved 31 March 2019, from http://www.praat.org/

Braver, A. (2014). Imperceptible incomplete neutralization: Production, non-identifiability, and non-discriminability in American English flapping. *Lingua, 152*, 24–44. DOI: https://doi.org/10.1016/j.lingua.2014.09.004

Bukmaier, V., Harrington, J., & Kleber, F. (2014). An analysis of post-vocalic /s-ʃ/ neutralization in Augsburg German: Evidence for a gradient sound change. *Frontiers in Psychology, 5*, paper 828. DOI: https://doi.org/10.3389/fpsyg.2014.00828

Cai, Q., & Brysbaert, M. (2010). SUBTLEX-CH: Chinese word and character frequencies based on film subtitles. *PLoS ONE, 5*(6), paper e10729. DOI: https://doi.org/10.1371/journal.pone.0010729

Carignan, C. (2014). *TRACTUS (Temporally Resolved Articulatory Configuration Tracking of UltraSound)*. MATLAB software suite. Retrieved 15 October 2019, from http://christophercarignan.github.io/TRACTUS

Chang, Y., & Shih, C. (2015). Place contrast enhancement: The case of the alveolar and retroflex sibilant production in two dialects of Mandarin. *Journal of Phonetics, 50*, 52–66. DOI: https://doi.org/10.1016/j.wocn.2015.02.001

Chen, C. (1991). The nasal endings and retroflexed initials in Peking Mandarin: Instability and the trend of changes. *Journal of Chinese Linguistics, 19*(2), 139–71. Retrieved from https://www.jstor.org/stable/23756149

Chen, M. [Matthew] (1973). Across-dialectal comparison: A case study and some theoretical considerations. *Journal of Chinese Linguistics, 1*(1), 39–63. Retrieved from https://www.jstor.org/stable/23749776

Chen, M. [Marilyn] (2000). Acoustic analysis of simple vowels preceding a nasal in Standard Chinese. *Journal of Phonetics, 28*(1), 43–67. DOI: https://doi.org/10.1006/jpho.2000.0106

Chen, W., Chang, Y., & Iskarous, K. (2015). Vowel coarticulation: Landmark statistics measure vowel aggression. *Journal of the Acoustical Society of America, 138*(2), 1221–32. DOI: https://doi.org/10.1121/1.4928307

Chen, Y. [Ying] & Guion-Anderson, S. (2011). Perceptual confusability of word-final nasals in Southern Min and Mandarin: Implications for coda nasal mergers in Chinese. In *Proceedings of the 17th International Congress of Phonetic Sciences*, Hong Kong (pp. 464–7). Hong Kong: City University of Hong Kong. Retrieved from https://www.internationalphoneticassociation.org/icphs-proceedings/ICPhS2011/index.htm

Chen, Y. [Yiya] & Gussenhoven, C. (2015). Shanghai Chinese. *Journal of the International Phonetic Association, 45*(3), 321–37. DOI: https://doi.org/10.1017/S0025100315000043

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*, Melbourne (pp. 398–402). Canberra: Australasian Speech Science and Technology Association. Retrieved from https://www.internationalphoneticassociation.org/icphs-proceedings/ICPhS2019/index.php

Chiu, C., Wei, P., Noguchi, M., & Yamane, N. (2019). Sibilant fricative merging in Taiwan Mandarin: An investigation of tongue postures using ultrasound imaging. *Language and Speech*. December 2020. DOI: https://doi.org/10.1177/0023830919896386

Duanmu, S. (2007). *The Phonology of Standard Chinese*. Oxford: Oxford University Press.

Fon, J., Hung, J., Huang, Y., & Hsu, H. (2011). Dialectal variations on syllable-final nasal mergers in Taiwan Mandarin. *Language and Linguistics*, *12*(2), 273–311. Retrieved from https://www.AiritiLibrary.com/Publication/Index/1606822X-201102-201505270022-201505270022-273-311

Fourakis, M., & Iverson, G. (1984). On the 'incomplete neutralization' of German final obstruents. *Phonetica*, *41*(3), 140–149. DOI: https://doi.org/10.1159/000261720

Green, A. (2005). Word, foot, and syllable structure in Burmese. In J. Watkins (Ed.), *Studies in Burmese Linguistics* (pp. 1–25). Canberra: Pacific Linguistics.

Guan, Y. (2019). Nasal coda realization in speech production of Shanghai Mandarin. In *Proceedings of the 19th International Congress of Phonetic Sciences*, Melbourne (pp. 467–71). Canberra: Australasian Speech Science and Technology Association. Retrieved from https://www.internationalphoneticassociation.org/icphs-proceedings/ICPhS2019/index.php

Herd, W., Jongman, A., & Sereno, J. (2010). An acoustic and perceptual analysis of /t/ and /d/ flaps in American English. *Journal of Phonetics*, *38*(4), 504–516. DOI: https://doi.org/10.1016/j.wocn.2010.06.003

Hoole, P., & Pouplier, M. (2017). Öhman returns: New horizons in the collection and analysis of imaging data in speech production research. *Computer Speech and Language*, *45*, 253–277. DOI: https://doi.org/10.1016/j.csl.2017.03.002

Hsu, H., & Tse, J. (2007). Syllable-final nasal mergers in Taiwan Mandarin – Leveled but puzzling. *Concentric: Studies in Linguistics*, *33*(1), 1–18. DOI: https://doi.org/10.6241/concentric.ling.200701_33(1).0001

Hueber, T., Aversano, G., Cholle, G., Denby, B., Dreyfus, G., Oussar, Y., Roussel, P., & Stone, M. (2007). Eigentongue feature extraction for an ultrasound-based silent speech interface. In *Proceedings of the 2007 International Conference on Acoustics, Speech and Signal Processing*, Honolulu (pp. I1245–I1248). DOI: https://doi.org/10.1109/ICASSP.2007.366140

Jannedy, S., & Weirich, M. (2017). Spectral moments vs discrete cosine transformation coefficients: Evaluation of acoustic measures distinguishing two merging German fricatives. *Journal of the Acoustical Society of America*, *142*(1), 395–405. DOI: https://doi.org/10.1121/1.4991347

Kavitskaya, D. (2014). *Compensatory lengthening: Phonetics, phonology, diachrony*. New York: Routledge.

Kawahara, S., & Garvey, K. (2014). Nasal place assimilation and the perceptibility of place contrasts. *Open Linguistics*, *1*, 17–36. DOI: https://doi.org/10.2478/opli-2014-0002

Keating, P. (1988a). Underspecification in phonetics. *Phonology*, *5*(2), 275–92. DOI: https://doi.org/10.1017/S095267570000230X

Keating, P. (1988b). The window model of coarticulation: Articulatory evidence. *UCLA Working Papers in Phonetics*, *69*, 3–29.

Kerswill, P., & Wright, S. (1990). The validity of phonetic transcription: Limitations of a sociolinguistic research tool. *Language Variation and Change*, *2*(3), 255–75. DOI: https://doi.org/10.1017/S0954394500000363

Keyser, S., & Stevens, K. (2006). Enhancement and overlap in the speech chain. *Language*, *82*(1), 33–63. DOI: https://doi.org/10.1353/lan.2006.0051

Kharlamov, V. (2014). Incomplete neutralization of the voicing contrast in word-final obstruents in Russian: Phonological, lexical, and methodological influences. *Journal of Phonetics, 43*, 47–56. DOI: https://doi.org/10.1016/j.wocn.2014.02.002

Kleber, F., John, T., & Harrington, J. (2010). The implications for speech perception of incomplete neutralization of final devoicing in German. *Journal of Phonetics, 38*(2), 185–96. DOI: https://doi.org/10.1016/j.wocn.2009.10.001

Kochetov, A., Faytak, M., & Nara, K. (2019). Manner differences in the Punjabi dental-retroflex contrast: an ultrasound study of time-series data. In *Proceedings of the 19th International Congress of Phonetic Sciences*, Melbourne (pp. 2002–6). Canberra: Australasian Speech Science and Technology Association. Retrieved from https://www.internationalphoneticassociation.org/icphs-proceedings/ICPhS2019/index.php

Kohler, K. (1990). Segmental reduction in connected speech in German: Phonological facts and phonetic explanations. In W. J. Hardcastle & A. Marchal (Eds.), *Speech production and speech modelling*. Dordrecht: Kluwer Academic Publishers. DOI: https://doi.org/10.1007/978-94-009-2037-8_4

Kurowski, K., & Blumstein, S. (1984). Perceptual integration of the murmur and formant transitions for place of articulation in nasal consonants. *Journal of the Acoustical Society of America, 76*(2), 383–90. DOI: https://doi.org/10.1121/1.391139

Kurowski, K., & Blumstein, S. (1993). Acoustic properties for the perception of nasal consonants. In M. Huffman & R. Krakow (Eds.), *Nasals, Nasalization and the Velum*. San Diego: Academic Press. DOI: https://doi.org/10.1016/B978-0-12-360380-7.50012-0

Lai, Y. (2008). Production of Mandarin Chinese nasal coda by L1 and L2 speakers of Mandarin Chinese. *Journal of Chinese Language Teaching, 5*(1), 155–80. DOI: https://doi.org/10.6393/JCLT.200806.0155

Lenth, R. (2020). *emmeans: Estimated marginal means, aka least-squares means*. R package. Version 1.4.4. Retrieved 19 Feb. 2020, from https://CRAN.R-project.org/package=emmeans

Lin, C. (2002). Nasal endings of Taiwan Mandarin: Production, perception, and linguistic change [paper presentation]. *The 35th International Conference on Sino-Tibetan Languages and Linguistics*. Tempe, Arizona. Retrieved from http://www.u.arizona.edu/~clin/professional/papers/02nasal.pdf

Lombardi, L. (2001). Why Place and Voice are different: Constraint-specific alternations in Optimality Theory. In L. Lombardi (Ed.), *Segmental phonology in Optimality Theory: Constraints and Representations*. Cambridge: Cambridge University Press. DOI: https://doi.org/10.1017/CBO9780511570582

Luo, M. (2015). Perception and production of Mandarin nasal codas by Shanghainese speakers. In M. Sloos & J. van de Weijer (Eds.), *Proceedings of the First Workshop on Chinese Accents and Accented Chinese*. Shanghai. Retrieved from https://www.researchgate.net/profile/Marjoleine_Sloos/publication/299770827_Proceedings_of_the_first_workshop_Chinese_Accents_and_Accented_Chinese_CAAC2014/links/570517b308aef745f71730da/Proceedings-of-the-first-workshop-Chinese-Accents-and-Accented-Chinese-CAAC2014.pdf

Mathôt, S., Schreij, D., & Theeuwes, J. (2012). OpenSesame: An open-source, graphical experiment builder for the social sciences. *Behavior Research Methods, 44*(2), 314–24. DOI: https://doi.org/10.3758/s13428-011-0168-7

McAuliffe, M., Socolof, M., Mihuc, S., Wagner, M., & Sonderegger, M. (2017). Montreal Forced Aligner: Trainable text-speech alignment using Kaldi. *Proceedings of Interspeech 2017* (pp. 498–502). DOI: https://doi.org/10.21437/Interspeech.2017-1386

Mielke, J., Carignan, C., & Thomas, E. (2017). The articulatory dynamics of pre-velar and pre-nasal /æ/-raising in English: An ultrasound study. *Journal of the Acoustical Society of America, 142*(1), 332–49. DOI: https://doi.org/10.1121/1.4991348

Mou, X. (2006). Nasal codas in Standard Chinese: A study in the framework of the distinctive feature theory. Doctoral dissertation, Massachusetts Institute of Technology. Retrieved from http://hdl.handle.net/1721.1/35283

Munson, B., Edwards, J., Schellinger, S. K., Beckman, M. E., & Meyer, M. K. (2010). Deconstructing phonetic transcription: Covert contrast, perceptual bias, and an extraterrestrial view of Vox Humana. *Clinical Linguistics and Phonetics, 24*(4–5), 245–260. DOI: https://doi.org/10.3109/02699200903532524

Narayan, C. (2008). The acoustic-perceptual salience of nasal place contrasts. *Journal of Phonetics, 36*(1), 191–217. DOI: https://doi.org/10.1016/j.wocn.2007.10.001

Nicenboim, B., Roettger, T., & Vasishth, S. (2018). Using meta-analysis for evidence synthesis: The case of incomplete neutralization in German. *Journal of Phonetics, 70*, 39–55. DOI: https://doi.org/10.1016/j.wocn.2018.06.001

Ohala, J. (1990). The phonetics and phonology of aspects of assimilation. In J. Kingston & M. Beckman (Eds.), *Papers in Laboratory Phonology*. Cambridge: Cambridge University Press. DOI: https://doi.org/10.1017/CBO9780511627736.014

Ping, Y. (2005). 上海方言语音动态腭位研究 [An electropalatographical investigation of Shanghainese sounds]. Hong Kong: Hong Kong Wenhui Publishing House.

Polka, L. (1991). Cross-language speech perception in adults: Phonemic, phonetic, and acoustic contributions. *Journal of the Acoustical Society of America, 89*(6), 2961–77. DOI: https://doi.org/10.1121/1.400734

Port, R., & Crawford, P. (1989). Incomplete neutralization and pragmatics in German. *Journal of Phonetics, 17*(4), 257–282. DOI: https://doi.org/10.1016/S0095-4470(19)30444-9

Port, R., & O'Dell, M. (1985). Neutralization of syllable-final voicing in German. *Journal of Phonetics, 13*(4), 455–71. DOI: https://doi.org/10.1016/S0095-4470(19)30797-1

Pouplier, M., & Hardcastle, W. (2005). A re-evaluation of the nature of speech errors in normal and disordered speakers. *Phonetica, 62*(2–4), 227–43. DOI: https://doi.org/10.1159/000090100

Qian, N. (2003). 上海语言发展史 [History of linguistic developments in Shanghai]. Shanghai: 上海人民出版社 [Shanghai People's Press].

R Core Team. (2019). R: A language and environment for statistical computing. Software. R Foundation for Statistical Computing. Retrieved from http://www.R-project.org/

Ramsammy, M. (2013) Word-final nasal velarisation in Spanish. *Journal of Linguistics, 49*(1), 215–55. DOI: https://doi.org/10.1017/S0022226712000187

Recasens, D. (2012). A study of jaw coarticulatory resistance and aggressiveness for Catalan consonants and vowels. *Journal of the Acoustical Society of America, 132*(1), 412–20. DOI: https://doi.org/10.1121/1.4726048

Recasens, D., & Espinosa, A. (2009). An articulatory investigation of lingual coarticulatory resistance and aggressiveness for consonants and vowels in Catalan. *Journal of the Acoustical Society of America, 125*(4), 2288–98. DOI: https://doi.org/10.1121/1.3089222

Recasens, D., & Rodríguez, C. (2016). A study on coarticulatory resistance and aggressiveness for front lingual consonants and vowels using ultrasound. *Journal of Phonetics, 59*, 58–75. DOI: https://doi.org/10.1016/j.wocn.2016.09.002

Roettger, T. B. (2019). Researcher degrees of freedom in phonetic research. *Laboratory Phonology, 10*(1). DOI: https://doi.org/10.5334/labphon.147

Roettger, T., Winter, B., Grawunder, S., Kirby, J., & Grice, M. (2014). Assessing incomplete neutralization of final devoicing in German. *Journal of Phonetics*, *43*, 11–25. DOI: https://doi.org/10.1016/j.wocn.2014.01.002

Shen, T. (1981). 上海话老派新派的差别 [Differences between old and new Shanghainese]. 方言 [Dialect] 4, 275–83.

Sherard, M. (1972). Shanghai phonology. Doctoral dissertation, Cornell University.

Shih, C. (1995). Study of Vowel Variations for a Mandarin Speech Synthesizer. In *Proceedings of the Fourth European Conference on Speech Communication and Technology (EUROSPEECH '95)*, Madrid (pp. 1807–10). Retrieved from https://www.isca-speech.org/archive/archive_papers/eurospeech_1995/e95_1807.pdf

Spreafico, L., Pucher, M., & Matosova, A. (2018). UltraFit: A Speaker-friendly headset for ultrasound recordings in speech science. In *Proceedings of Interspeech 19*, 1517–20. DOI: https://doi.org/10.21437/Interspeech.2018-995

Steriade, D. (1994). Positional Neutralization and the Expression of Contrast. Unpublished manuscript, University of California, Los Angeles. Retrieved from http://lingphil.mit.edu/papers/steriade/contrastive-gesture.pdf

Stone, M., & Vatikiotis-Bateson, E. (1995). Trade-offs in tongue, jaw, and palate contributions to speech production. *Journal of Phonetics*, *23*(1–2), 81–100. DOI: https://doi.org/10.1016/S0095-4470(95)80034-4

Strycharczuk, P., & Sebregts, K. (2018). Articulatory dynamics of (de)gemination in Dutch. *Journal of Phonetics*, *68*, 138–49. DOI: https://doi.org/10.1016/j.wocn.2018.03.005

Wang, L., Cui, J., & Chen, Y. (2018). Wuxi Speakers' Production and Perception of Coda Nasals in Mandarin. In *Proceedings of Interspeech 19*, 2559–62. DOI: https://doi.org/10.21437/Interspeech.2018-2224

Warner, N., Good, E., Jongman, A., & Sereno, J. (2006). Orthographic vs. morphological incomplete neutralization effects. *Journal of Phonetics*, *34*(2), 285–93. DOI: https://doi.org/10.1016/j.wocn.2004.11.003

Warner, N., Jongman, A., Sereno, J., & Kemps, R. (2004). Incomplete neutralization and other sub-phonemic durational differences in production and perception: Evidence from Dutch. *Journal of Phonetics*, *32*(2), 251–76. DOI: https://doi.org/10.1016/S0095-4470(03)00032-9

Wu, M., Sloos, M., & van de Weijer, J. (2016). The perception of the English alveolar-velar nasal coda contrast by monolingual versus bilingual Chinese speakers. In Lee, T., Xie, L., Dang, J., Wang, H., Wei, J., Feng, H., Hou, Q., & Wei, Y. (Eds.), *Proceedings of the Tenth International Symposium on Chinese Spoken Language Processing*, Tianjin. DOI: https://doi.org/10.1109/ISCSLP.2016.7918401

Xu, B., & Tang, Z. (1988). 上海市区方言志 [A description of Shanghainese spoken in the urban districts of Shanghai City]. Shanghai: 上海教育出版社 [Shanghai Education Press].

Xu, D. (1993). Unexceptional Irregularities: Lexical Conditioning of Mandarin Nasal Deletion. *Diachronica*, *10*, 215–36. DOI: https://doi.org/10.1075/dia.10.2.04xu

Yang, J. (2010). Phonetic evidence for the nasal coda shift in Mandarin. *Taiwan Journal of Linguistics*, *8*(1), 29–56. DOI: https://doi.org/10.6519/TJL.2010.8(1).2

Yu, A. (2011). Mergers and neutralization. In van Oostendorp, M., Ewen, C., Hume, E., & Rice, K. (Eds.), *The Blackwell companion to phonology* (pp. 1892–1918). Oxford: Blackwell. DOI: https://doi.org/10.1002/9781444335262.wbctp0080

Zee, E. (1981). Effect of vowel quality on perception of post-vocalic nasal consonants in noise. *Journal of Phonetics*, *9*(1), 35–48. DOI: https://doi.org/10.1016/S0095-4470(19)30925-8

Zee, E. (1985). Sound change in syllable-final nasal consonants in Chinese. *Journal of Chinese Linguistics*, *13*, 291–330. Retrieved from https://www.jstor.org/stable/23767518