Articulation and perception of Mandarin coda nasals by Shanghainese-Mandarin bilinguals

Suyuan Liu and Matthew Faytak
liusuyuan@ucla.edu; faytak@ucla.edu

LSA 94th Annual Meeting
Jan. 3, 2020

Overview

Mandarin nasal codas /n/ and /ŋ/ are prone to merger after non-low vowels

- Previous literature is overly focused on language contact as cause and is generally based on auditory impression

Present study: another look at articulation and perception of **Mandarin nasal codas** by two groups, **Shanghai Mandarin** speakers and control northern Mandarin speakers

- Speakers vary between [n] and [ŋ] for some lexemes
- After /i/, merger to place which is neither [n] nor [ŋ]
- Poor discrimination performance; perceptual bias towards [ŋ]

Outline

Background

Ultrasound Study: characterize **articulatory** properties of merged nasal codas

- Linear discriminant analysis (LDA) method
- Visualizing typical nasal coda tongue shapes

Perception Study: **discrimination** of the merged nasal codas after /i/

- Stimulus selection using LDA
- AXB task

Discussion

Background

Shanghai Mandarin nasal codas

Regional dialect of Standard Mandarin spoken in Shanghai

- Not to be confused with Shanghainese (Wu dialect)
- Speakers usually bilingual: L1 or co-L1 Shanghainese

Merges its two **nasal codas** /n/, /ŋ/ after non-low vowels[1]

- /an/–/aŋ/ remain distinct
- /ən/–/əŋ/ and /in/–/iŋ/ merge

This pattern of nasal merger is also seen in other dialects of Standard Mandarin, such as Taiwan Mandarin[2]

---

[1] Guan, 2019; Luo, 2015.

[2] Chiu et al., 2019; Wu, Sloos, and van de Weijer, 2016.

Non-contrastive place

At issue here: **non-contrastive place** that nasals settle on in Shanghai Mandarin

- All described as [n], but using only auditory coding[3]
- **Perception** of coda nasal place after [i], [e] has known bias toward [n][4]

In Mandarin dialects where place has been investigated using ultrasound, e.g. Taiwan Mandarin[5]; place is known to **vary according to the preceding vowel**

- After mid vowel /ə/, merge to [n]
- After high vowel /i/, merge to [ŋ]

]

[3] Guan, 2019; Luo, 2015.

[4] Zee, 1981.

[5] Chiu et al., 2019.

Influence of L1 transfer?

Merger *and the merged coda's place* are often attributed to contact with the local language

…regardless of the contact language and its nasal contrasts!

» Shanghainese (contact with Shanghai Mandarin) lacks nasal coda contrasts[6]

» Southern Min (contact with Taiwan Mandarin) contrasts coda /m, n, ŋ/[7]

We find this account unlikely

» Involvement of coarticulatory pressures?

» Perception?

[6] Luo, 2015.

[7] Y. Chen and Guion-Anderson, 2011; Chiu et al., 2019.

Ultrasound study

Study objectives

We find it more likely that **biomechanical factors** dictate place of non-merged nasal

- **Ultrasound study** of Shanghai Mandarin to determine:
  - Place of merged nasal
  - Which vocalic contexts encourage merger

Recording method

Synchronized **ultrasound video** and **audio** recorded in UCLA Phonetics Lab

- UltraFit stabilization headset used[8]

[8] Spreafico, Pucher, and Matosova, 2018.

Participants

Two groups recruited on the UCLA campus

Shanghai Mandarin speakers (n=15)

» L1 or co-L1: Shanghainese
» Expected to exhibit mergers

Mandarin control speakers (n=5)

» L1: standard northern/Beijing Mandarin
» Report no experience with Shanghai Mandarin or Shanghainese
» Not expected to exhibit mergers

Stimuli

Frequency-matched[9] minimal pairs differing only in final nasal

	n coda	**ŋ** coda
a-	隐含 in˨˩˦.ha**n**˧˥ ‘imply’	引航 in˨˩˦.ha**ŋ**˧˥ ‘pilot’
ə-	清真 tɕʰiŋ˥.tʂə**n**˥ ‘Islamic’	清蒸 tɕʰiŋ˥.tʂə**ŋ**˥ ‘steamed’
i-	山林 ʂan˥.li**n**˧˥ ‘mountain forest’	山陵 ʂan˥.li**ŋ**˧˥ ‘lofty mountains’

- Presented in random order in a frame sentence
- Utterance-final position to avoid place assimilation

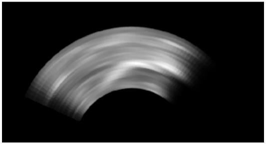
[9] Cai and Brysbaert, 2010.

Analysis of articulatory signal

Ultrasound image data is rich, but noisy and high-dimensional

- Each observation is tens of thousands of pixels
- Each pixel contains numerical data: for an 8-bit grayscale image, brightness between 0 (black) and 255 (white)

One solution: reduce dimensionality of data using **principal component analysis** (PCA), capturing important variation in a few new features

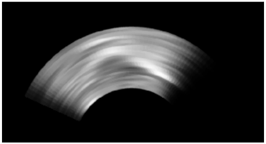


Analysis of articulatory signal

Ultrasound image data is rich, but noisy and high-dimensional

- Each observation is tens of thousands of pixels
- Each pixel contains numerical data: for an 8-bit grayscale image, brightness between 0 (black) and 255 (white)

One solution: reduce dimensionality of data using **principal component analysis** (PCA), capturing important variation in a few new features

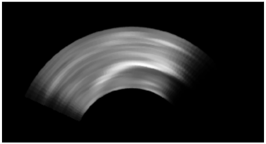


Analysis of articulatory signal

Ultrasound image data is rich, but noisy and high-dimensional

- Each observation is tens of thousands of pixels
- Each pixel contains numerical data: for an 8-bit grayscale image, brightness between 0 (black) and 255 (white)

One solution: reduce dimensionality of data using **principal component analysis** (PCA), capturing important variation in a few new features

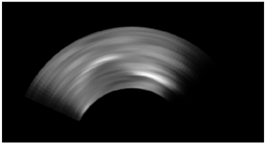


Analysis of articulatory signal

Ultrasound image data is rich, but noisy and high-dimensional

- Each observation is tens of thousands of pixels
- Each pixel contains numerical data: for an 8-bit grayscale image, brightness between 0 (black) and 255 (white)

One solution: reduce dimensionality of data using **principal component analysis** (PCA), capturing important variation in a few new features



Analysis of articulatory signal

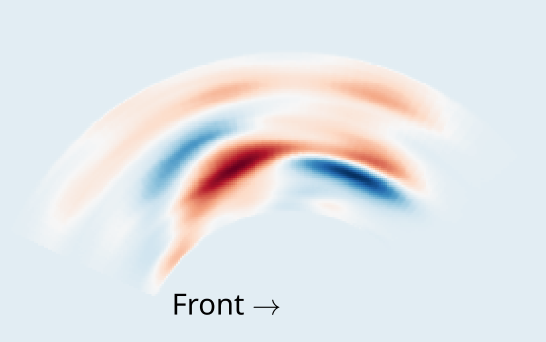
Ultrasound image data is rich, but noisy and high-dimensional

- Each observation is tens of thousands of pixels
- Each pixel contains numerical data: for an 8-bit grayscale image, brightness between 0 (black) and 255 (white)

One solution: reduce dimensionality of data using **principal component analysis** (PCA), capturing important variation in a few new features

Representing principal components

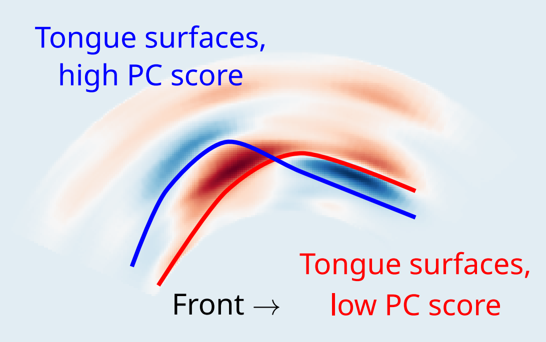
Map of **loadings** on that PC, or roughly **covariation in pixel intensity**[10]



[10] Hueber et al., 2007; Mielke, Carignan, and Thomas, 2017.

Representing principal components

Map of **loadings** on that PC, or roughly **covariation in pixel intensity**[10]



[10] Hueber et al., 2007; Mielke et al., 2017.

Analysis

Midpoint frames from each nasal coda submitted to PCA

- PCs 1-10 retained (avg. 80% of variance explained)
- **Separate PCAs** for each speaker: PCA including all speakers might capture non-linguistic variation
  - Morphological variation (size, palate shape, etc.)
  - Ultrasound probe placement variation

PCs 1-10 submitted to a **linear discriminant analysis**, which helps interpret the components

- Learns dimension that **maximally separates** etymological /n/ and /ŋ/
- Yields **linear discriminant values** for each nasal coda token and **classification** as /n/ or /ŋ/ based on these values
- LD values are **normalized** to 0-1 range for all speakers
  - /n/ always low, /ŋ/ always high

Predictions

Because less consistent contrasts will be **less learnable by**, and **worse separated** on, the LDA:

- Shanghai group will have worse separability of /n/ and /ŋ/ than the control group
- Shanghai group will have lower rate of correct classification of /n/ and /ŋ/ than control group

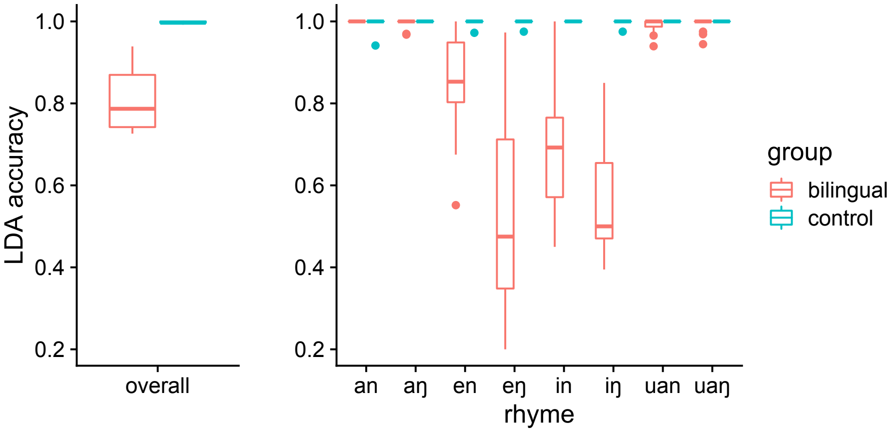
Expected place of coda nasals, based on Chiu et al:[11]

- Merger to [n] after /ə/ in Shanghai group
- Merger to [ŋ] after /i/ in Shanghai group
- No mergers after low vowels, and no mergers for control group

[11] Chiu et al., 2019.

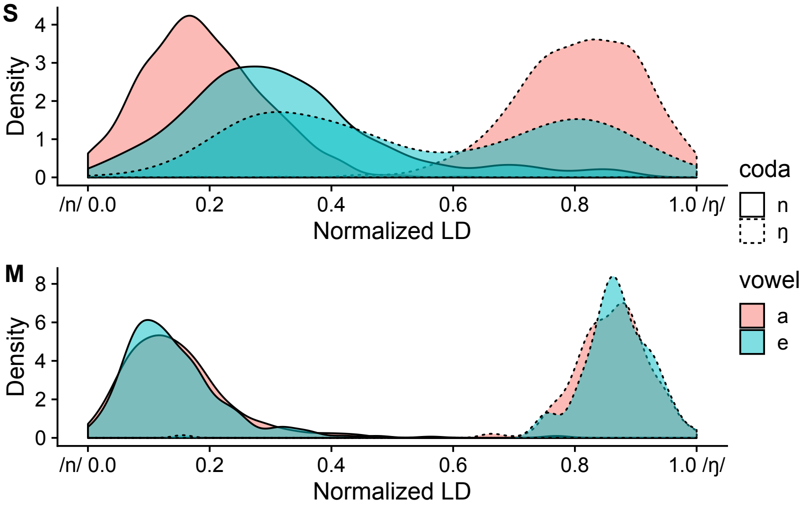
Classification accuracy

LDA performs worse on **Shanghai group**; particularly for /əŋ/ and /iŋ/ rhymes



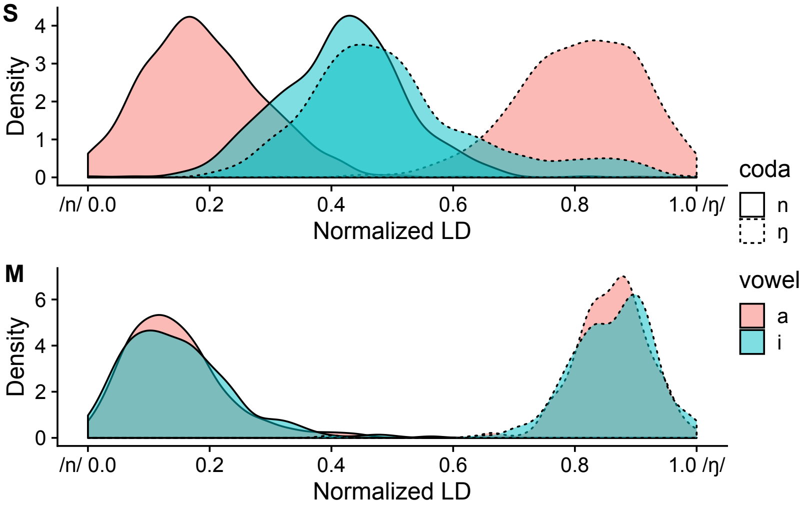
LD values, nasals after /ə/

/ən/ and /əŋ/ merge to (roughly) [ən]; /əŋ/ is **bimodal**



LD values, nasals after /i/

/iŋ/ and /in/ merged, in middle of LD; /iŋ/ is also bimodal



Reconstructing frames

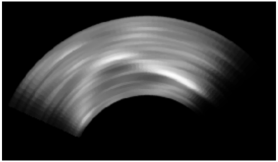
Middle of LD (not near 0 or 1): not easily classified as /n/ or /ŋ/, but not necessarily similar to either prototype

For segments which cluster mid-LD, we can **reconstruct** a **typical frame** for each nasal after each vowel height from the PCA

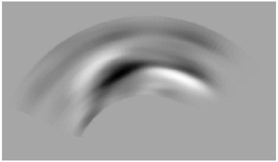
» **Multiply** each PC's loading by its average PC score (i.e. the nasal in /iŋ/)

» **Add** together the contributions of each PC

Basis frame, [ŋ]



Reconstruction



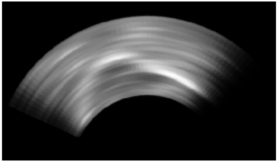
Reconstructing frames

Middle of LD (not near 0 or 1): not easily classified as /n/ or /ŋ/, but not necessarily similar to either prototype

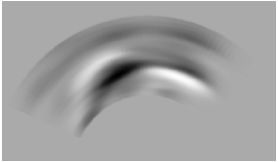
For segments which cluster mid-LD, we can **reconstruct** a **typical frame** for each nasal after each vowel height from the PCA

- **Multiply** each PC's loading by its average PC score (i.e. the nasal in /iŋ/)
- **Add** together the contributions of each PC

Basis frame, [ŋ]



Reconstruction



Reconstructing frames

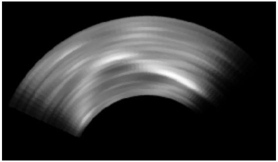
Middle of LD (not near 0 or 1): not easily classified as /n/ or /ŋ/, but not necessarily similar to either prototype

For segments which cluster mid-LD, we can **reconstruct** a **typical frame** for each nasal after each vowel height from the PCA

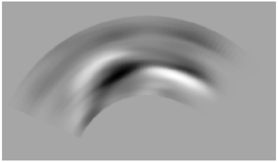
» **Multiply** each PC's loading by its average PC score (i.e. the nasal in /iŋ/)

» **Add** together the contributions of each PC

Basis frame, [ŋ]



Reconstruction



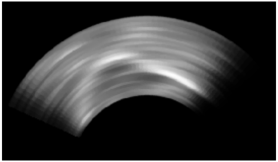
Reconstructing frames

Middle of LD (not near 0 or 1): not easily classified as /n/ or /ŋ/, but not necessarily similar to either prototype

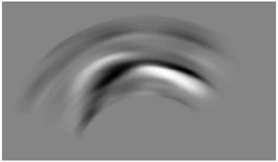
For segments which cluster mid-LD, we can **reconstruct** a **typical frame** for each nasal after each vowel height from the PCA

- **Multiply** each PC's loading by its average PC score (i.e. the nasal in /iŋ/)
- **Add** together the contributions of each PC

Basis frame, [ŋ]



Reconstruction



Reconstructing frames

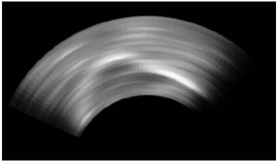
Middle of LD (not near 0 or 1): not easily classified as /n/ or /ŋ/, but not necessarily similar to either prototype

For segments which cluster mid-LD, we can **reconstruct** a **typical frame** for each nasal after each vowel height from the PCA

» **Multiply** each PC's loading by its average PC score (i.e. the nasal in /iŋ/)

» **Add** together the contributions of each PC

Basis frame, [ŋ]



Reconstruction

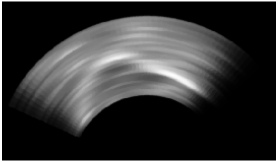
Reconstructing frames

Middle of LD (not near 0 or 1): not easily classified as /n/ or /ŋ/, but not necessarily similar to either prototype

For segments which cluster mid-LD, we can **reconstruct** a **typical frame** for each nasal after each vowel height from the PCA

- **Multiply** each PC's loading by its average PC score (i.e. the nasal in /iŋ/)
- **Add** together the contributions of each PC

Basis frame, [ŋ]



Reconstruction

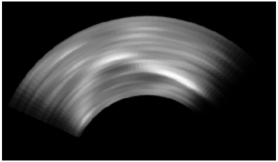
Reconstructing frames

Middle of LD (not near 0 or 1): not easily classified as /n/ or /ŋ/, but not necessarily similar to either prototype

For segments which cluster mid-LD, we can **reconstruct** a **typical frame** for each nasal after each vowel height from the PCA

- **Multiply** each PC's loading by its average PC score (i.e. the nasal in /iŋ/)
- **Add** together the contributions of each PC

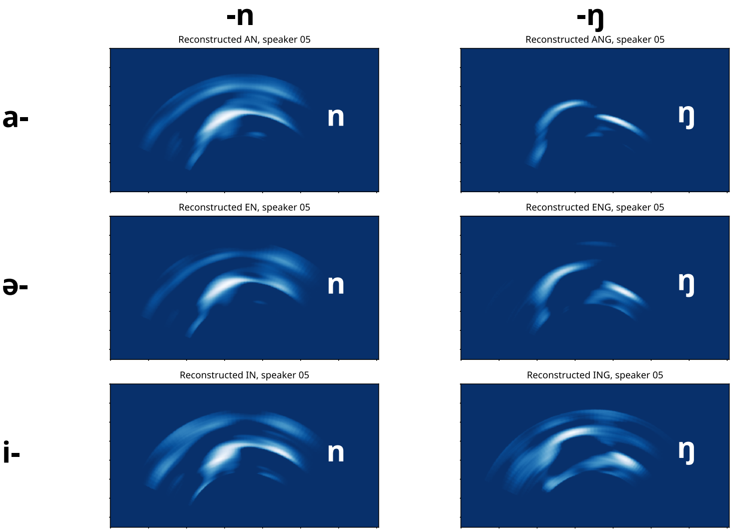
Basis frame, [ŋ]



Reconstruction

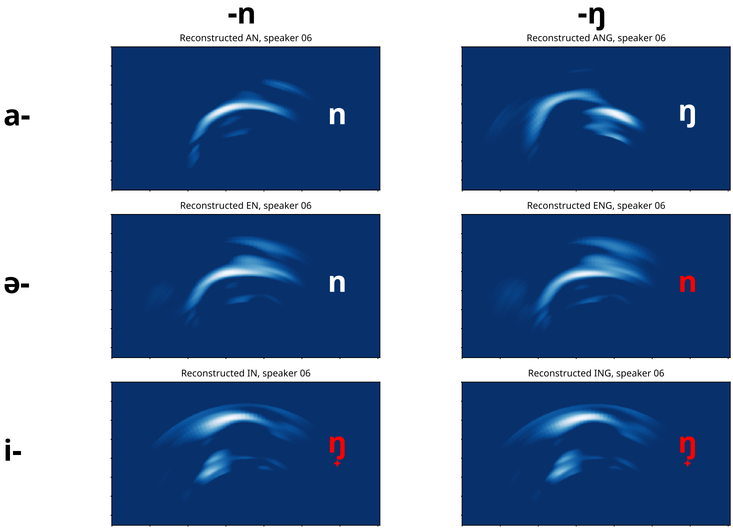
Reconstructions, control speaker

Tongue position for each nasal is consistent regardless of vowel



Reconstructions, Shanghai Mandarin speaker

Mergers affect codas in **red**



Discussion

Realization of merged nasal is **dependent on the preceding vowel**, consistent with other descriptions of dialectal Mandarin

- After /ə/, merger to [n]
- After /i/, merger to [ŋ̟] or perhaps [ɲ]

Realization of merged nasals is **less clear-cut** than usually depicted

- Neutralized segment after /i/ is **neither** [n] nor [ŋ]
- Some speakers vary between a canonical [ŋ] and a merged variant after non-low vowels, especially for intended /iŋ/ and /əŋ/

Discussion

Taiwan and Shanghai Mandarin have roughly the same merger patterns, but their situations involve different L1s

- Suggests that L1 transfer does *not* determine the place of the merged nasal that results

Biomechanical factors seem to determine the place that the merged, non-contrastive nasal "settles" on[12]

- After /i/: [ŋ̟] has maximally similar tongue position to [i]; requires least muscular strain
- After /ə/: producing [n] is not biomechanically easier than producing [ŋ], so other factors may be involved
- **Misperception**?

[12] Chiu et al., 2019.

Perception study

Perceptual factors

Mergers in production are typically preceded by mergers in perception

- articulatory merger only if perceptually inconspicuous[13]

Open question: whether Shanghai Mandarin listeners can tell the coda nasals apart

- Difficulty reported for Standard Mandarin listeners perceiving /in/–/iŋ/ contrast[14]
- Perception of nasal codas is also influenced by preceding vowels in language-specific fashion[15]

[13] Kawahara and Garvey, 2014.

[14] M. Y. Chen, 2000; Mou, 2006.

[15] Y. Chen and Guion-Anderson, 2011; Zee, 1981.

Study objectives

Can Mandarin-speaking listeners distinguish the coda nasals in perception?

- Shanghai Mandarin listeners expected to have difficulty distinguishing between [in], [iŋ]
- Control Mandarin listeners should not

If discrimination is poor, does bias towards [n] in perception drive it?[16]

[16] Zee, 1981.

Participants

Two groups recruited on the UCLA campus, defined as in ultrasound study (new speakers)

Shanghai Mandarin (SH) listeners (n=14)

- expected to perform worse in discriminating nasals

Mandarin control (MM) listeners (n=14)

- expected to perform better in discriminating nasals
- Not including one Mandarin speaker in our analysis because they had phonetic training

Method and stimuli

AXB discrimination task in sound-treated room at UCLA

- Non-orthographic presentation of choices (“Choice 1/2”)
- No feedback was provided
- Total of 304 test trials and 10 practice trials

Tokens of [in] or [iŋ] drawn from ultrasound study data

A and B anchor tokens, produced by control speakers:

- Canonical [in] (LD < 0.4) from MM
- Canonical [iŋ] (LD > 0.6) from MM

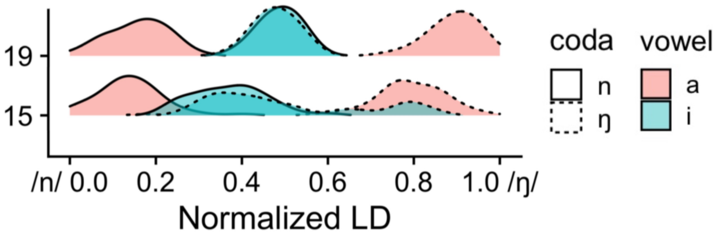
Method and stimuli

X tokens, [in] and [iŋ] produced by both groups:

- Canonical [in] or [iŋ] (LD < 0.4 or > 0.6) produced by one MM speaker and four SH speakers
- Non-canonical [iŋ] (LD 0.4–0.6) from the same SH speakers

Within-talker design is due to a quirk of the production data

- Many speakers **split** productions of /iŋ/ between canonical [iŋ] and non-canonical [iŋ̟]

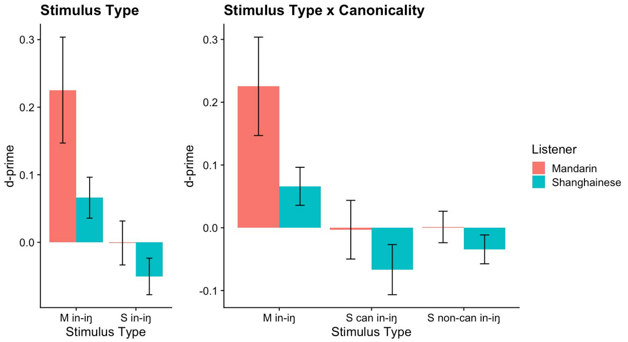


d-prime

Poor discrimination performance for all listener groups

» MM group outperforms SH group at distinguishing coda nasals produced by MM speakers

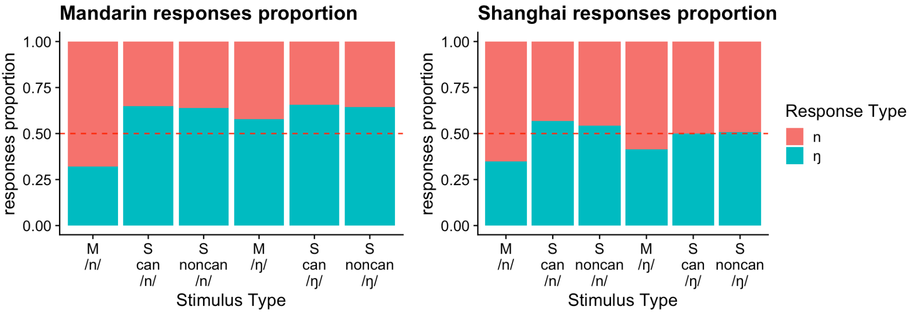
» All groups worse at distinguishing SH coda nasals, even canonical ones



Raw response data

Both groups perform badly at the discrimination task

- SH listeners select more or less **randomly**
- MM listeners have a slight bias towards **/ŋ/** except for control /n/



Mixed-effect logistic regression

To model the impact of **listener group** and **stimulus type**

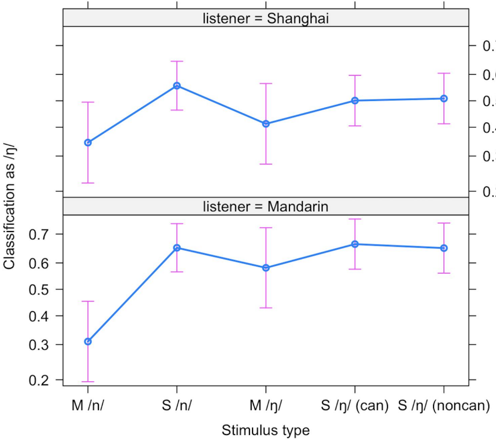
- Listener group
  - Control Mandarin or Shanghai Mandarin
- Stimulus type, five levels which incorporate:
  - Control Mandarin or Shanghai Mandarin speaker
  - Canonical or non-canonical /n/ or /ŋ/
  - **Mandarin control /n/** as baseline

Used maximal effects structure which converged

- resp ~ stim type*listener group + (1|word) + (1|talker) + (1|listener)

Fixed effects

- Effect of stimulus type is significant
- Main effect of listener group is not significant, but interactions between listener group and stimulus type are significant



Conclusion

Ultrasound findings

Data from the ultrasound study shows that

- Control Mandarin speakers produce nasal coda contrast
- Shanghai Mandarin speakers merge nasal codas after non-high vowels
  - To [ŋ̟] after /i/; to [n] after /ə/
  - But speakers occasionally produce canonical [n] and [ŋ]

Evidence against L1 transfer conditioning the place of the non-contrastive nasal

- Shanghai Mandarin shows the same pattern as Taiwan Mandarin despite different contact languages

Tying in perception findings

The coda nasals are apparently indistinguishable after /i/ for control and Shanghai Mandarin speakers

- Both SH *and MM* listeners poorly distinguish nasal coda place
- MM listeners exhibit velar bias, which is unexpected
- Still to test: nasal place contrast after /ə/

Possible **near merger**[17] for MM speakers, who have no difficulty producing an /in/–/iŋ/ contrast but cannot perceive it

- Merger in production is "perceptually inconspicuous"

[17] Yu, 2011.

Nasal coda merger as sound change

Findings allow for a more **nuanced** understanding of nasal coda merger in Chinese as a multi-step sound change

Some role for **L1 transfer** cannot be ruled out in the initial loss of the contrast

But following this, articulatory ease and perception may condition the resulting non-contrastive place in some contexts

- » Such as merger after /i/

While perception may play a greater role in other contexts

- » Such as merger after mid vowels such as /ə/ (to be tested)

Thanks

- **Megha Sundara**; **Henry Tehrani** for technical help
- **Canaan Breiss** for model convergence tips
- UCLA Dept. of Linguistics for experiment funds provided to first author

References I

Cai, Q., & Brysbaert, M. (2010). SUBTLEX-CH: Chinese word and character frequencies based on film subtitles. *PloS One*, *5*(6), e10729.

Chen, M. Y. (2000). Acoustic analysis of simple vowels preceding a nasal in standard chinese. *Journal of Phonetics*, *28*(1), 43–67.

Chen, Y., & Guion-Anderson, S. (2011). Perceptual Confusabiltiy of Word-final Nasals in Southern Min and Mandarin: Implications for Coda Nasal Mergers in Chinese. In *Proceedings of icphs 17* (pp. 464–467).

Chiu, C., Lu, Y.-A., Weng, Y., Jin, S.-J., Weng, W.-C., & Yang, T.-H. (2019). Uncovering syllable-final nasal merging in Taiwan Mandarin: An ultrasonographic investigation of tongue postures and degrees of nasalization. In *Proceedings of ICPhS 19*.

Guan, Y. (2019). Nasal coda realization in speech production of shanghai mandarin. In *Proceedings of ICPhS 19*.

Hueber, T., Aversano, G., Cholle, G., Denby, B., Dreyfus, G., Oussar, Y., … Stone, M. (2007). Eigentongue feature extraction for an ultrasound-based silent speech interface. In *Icassp 2007* (Vol. 1, pp. I–1245). IEEE.

Kawahara, S., & Garvey, K. (2014). Nasal place assimilation and the perceptibility of place contrasts. *Open Linguistics*, *1*(1).

Luo, M. (2015). Perception and Production of Mandarin Nasal Codas by Shanghainese Speakers. (p. 16).

Mielke, J., Carignan, C., & Thomas, E. R. (2017). The articulatory dynamics of pre-velar and pre-nasal/æ/-raising in English: An ultrasound study. *JASA*, *142*(1), 332–349.

Mou, X. (2006). *Nasal codas in standard chinese: A study in the framework of the distinctive feature theory*. (Doctoral dissertation, Massachusetts Institute of Technology).

References II

Spreafico, L., Pucher, M., & Matosova, A. (2018). Ultrafit: A speaker-friendly headset for ultrasound recordings in speech science.

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 *2016 10th international symposium on chinese spoken language processing (iscslp)* (pp. 1–5). IEEE.

Yu, A. C. (2011). Mergers and neutralization. *The Blackwell companion to phonology*, 1–27.

Zee, E. (1981). Effect of vowel quality on perception of post-vocalic nasal consonants in noise. *JPhon*, *1*(9), 35–48.

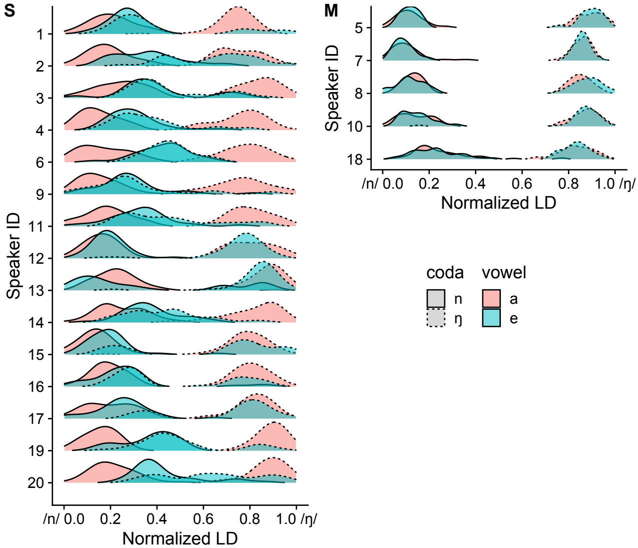
Appendix: Stimuli

Vowel Context	/n/		/ŋ/	
	Chinese Char Pinyin	English	Chinese Char Pinyin	English
/i/	山林 shānlín	’mountain forest’	山陵 shānlíng	’lofty mountains’
	全民 quánmín	’all the people’	全名 quánmíng	’full name’
	押金 yānjīn	’deposit’	压惊 yājīng	‘help sb. get over a shock’
	风琴 fēngqín	’organ (instrument)’	风情 fēngqíng	’amorous feelings’
/a/	青山 qīngshān	’green hills’	轻伤 qīngshāng	’minor wound’
	出产 chūchǎn	’yield’	出厂 chūchǎng	’(of products) dispatch from the factory’
	隐含 yǐnhán	’imply’	引航 yǐnháng	’pilot a ship’
	造反 zàofǎn	’rise in rebellion’	造访 zàofǎng	’pay a visit’

Appendix: Stimuli

Vowel Context	/n/		/ŋ/	
	Chinese Char Pinyin	English	Chinese Char Pinyin	English
/ə/	人参 rénshēn	’ginseng’	人声 rénshēng	’voice’
	解闷 jiěmèn	’amuse’	解梦 jiěmèng	’dream reading’
	清真 qīngzhēn	’Islamic’	清蒸 qīngzhēng	’steamed’
	水深 shuǐshēn	’depth (of waterway)’	水声 shuǐshēng	’water sounds’
/ua/	机关 jīguān	’mechanism’	激光 jīguāng	’laser’
	经传 jīngzhuàn	’classics’	精壮 jīngzhuàng	’strong’
	高官 gāoguān	’manager’	高光 gāoguāng	’highlight’
	旁观 pángguān	’look on’	膀胱 pángguāng	’bladder’

Appendix: individual LD values in /əN/ context



Appendix: individual LD values in /iN/ context