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
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/–/æŋ/ 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

<table>
<thead>
<tr>
<th>n coda</th>
<th>η coda</th>
</tr>
</thead>
<tbody>
<tr>
<td>a- 隐含 in˨˩.haŋ¹</td>
<td>引航 in˨˩.haŋ¹</td>
</tr>
<tr>
<td>‘imply’</td>
<td>‘pilot’</td>
</tr>
<tr>
<td>ə- 清真 tʂʰiŋ˥.tʂəŋ¹</td>
<td>清蒸 tʂʰiŋ˥.tʂəŋ¹</td>
</tr>
<tr>
<td>‘Islamic’</td>
<td>‘steamed’</td>
</tr>
<tr>
<td>i- 山林 ʂan˥.liŋ¹</td>
<td>山陵 ʂan˥.liŋ¹</td>
</tr>
<tr>
<td>‘mountain forest’</td>
<td>‘lofty mountains’</td>
</tr>
</tbody>
</table>

» 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.

Liu and Faytak

Mandarin Nasal Codas
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.

Liu and Faytak

Mandarin Nasal Codas
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.
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

Liu and Faytak
Mandarin Nasal Codas

LSA 94
Reconstructing frames

Middle of LD (not near 0 or 1): not easily classified as /n/ or /ɳ/, 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, [ɳ]**

**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]](image1.png) ![Reconstruction](image2.png)

Liu and Faytak
Mandarin Nasal Codas
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]](image1)

![Reconstruction](image2)
Reconstructing frames

Middle of LD (not near 0 or 1): not easily classified as /n/ or /η/, 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
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
Reconstructions, control speaker

Tongue position for each nasal is consistent regardless of vowel

-μ
Reconstructed AN, speaker 05

-η
Reconstructed ANG, speaker 05

-α
Reconstructed EN, speaker 05

-η
Reconstructed ENG, speaker 05

-ə
Reconstructed IN, speaker 05

-η
Reconstructed ING, speaker 05

-ï
Reconstructed IN, speaker 05

-η
Reconstructed ING, speaker 05
Reconstructions, Shanghai Mandarin speaker

Mergers affect codas in red

-\(\text{-n}\)

Reconstructed AN, speaker 06

\(\text{a-}\)

Reconstructed EN, speaker 06

\(\text{ø-}\)

Reconstructed IN, speaker 06

\(\text{i-}\)

Reconstructed ING, speaker 06

-\(\text{-ŋ}\)

Reconstructed ANG, speaker 06
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/

Liu and Faytak

Mandarin Nasal Codas

LSA 94
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” on¹²

» After /i/: [ŋ] 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/–/in\(\text{n}^j\)/ contrast\(^{14}\)

» Perception of nasal codas is also influenced by preceding vowels in language-specific fashion\(^{15}\)

\(^{13}\text{Kawahara and Garvey, 2014.}\)

\(^{14}\text{M. Y. Chen, 2000; Mou, 2006.}\)

\(^{15}\text{Y. Chen and Guion-Anderson, 2011; Zee, 1981.}\)
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?\(^\text{16}\)

\(^{16}\text{Zee, 1981.}\)
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]

![Graph showing normalized LD values for /n/ and /ŋ/ productions]
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
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

» \( \text{resp} \sim \text{stim type*listener group} + (1 | \text{word}) + (1 | \text{talker}) + (1 | \text{listener}) \)
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 [ŋ] 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/-/inj/ 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


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.


## Appendix: Stimuli

<table>
<thead>
<tr>
<th>Vowel Context</th>
<th>/n/</th>
<th>/ŋ/</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Chinese Char</strong></td>
<td><strong>English</strong></td>
<td><strong>Chinese Char</strong></td>
</tr>
<tr>
<td>山林 shānlín</td>
<td>‘mountain forest’</td>
<td>山陵 shānlíng</td>
</tr>
<tr>
<td>全民 quánmín</td>
<td>‘all the people’</td>
<td>全名 quánmíng</td>
</tr>
<tr>
<td>押金 yānjīn</td>
<td>‘deposit’</td>
<td>压惊 yājīng</td>
</tr>
<tr>
<td>风琴 fēngqín</td>
<td>‘organ (instrument)’</td>
<td>风情 fēngqíng</td>
</tr>
<tr>
<td>青山 qīngshān</td>
<td>‘green hills’</td>
<td>轻伤 qīngshāng</td>
</tr>
<tr>
<td>出产 chūchǎn</td>
<td>‘yield’</td>
<td>出厂 chūchǎng</td>
</tr>
<tr>
<td>隐含 yǐnhán</td>
<td>‘imply’</td>
<td>引航 yǐnháng</td>
</tr>
<tr>
<td>造反 zàofǎn</td>
<td>‘rise in rebellion’</td>
<td>造访 zàofǎng</td>
</tr>
</tbody>
</table>
## Appendix: Stimuli

<table>
<thead>
<tr>
<th>Vowel Context</th>
<th>/n/</th>
<th>/ŋ/</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Chinese Char Pinyin</strong></td>
<td><strong>English</strong></td>
<td><strong>Chinese Char Pinyin</strong></td>
</tr>
<tr>
<td><strong>/ə/</strong></td>
<td>人参 rénshēn</td>
<td>'ginseng'</td>
</tr>
<tr>
<td></td>
<td>解闷 jiěmèn</td>
<td>'amuse'</td>
</tr>
<tr>
<td></td>
<td>清真 qīngzhēn</td>
<td>'Islamic'</td>
</tr>
<tr>
<td></td>
<td>水深 shuǐshēn</td>
<td>'depth (of waterway)'</td>
</tr>
<tr>
<td><strong>/ua/</strong></td>
<td>机关 jīguān</td>
<td>'mechanism'</td>
</tr>
<tr>
<td></td>
<td>经传 jīngzhuàn</td>
<td>'classics'</td>
</tr>
<tr>
<td></td>
<td>高官 gāoguān</td>
<td>'manager'</td>
</tr>
<tr>
<td></td>
<td>旁观 pángguān</td>
<td>'look on'</td>
</tr>
</tbody>
</table>
Appendix: individual LD values in /əN/ context

Normalized LD

Speaker ID

S

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20

M

5
6
7
8
9
10
11
12
13
14
15
16
17
18

/n/ 0.0 0.2 0.4 0.6 0.8 1.0 /ŋ/

Normalized LD

Speaker ID

coda

vowel

M

S

n
a
ŋ
e
Appendix: individual LD values in /iN/ context