The sociophonetics of gender in three Chinese varieties

Mary E. Beckman       Fangfang Li       Ya-ting Shih

Columbus, OH, 2018-03-09 — NACCL-30 (Chinese Linguistics: from the field, from the lab, and from the armchair)
Three varieties of Chinese, spoken in these two cities

Map sources: [Songyuan] User Joowwww / Wikipedia commons / Public domain; [Kaohsiung] User ASDFGHJ (work by Luuva) / Wikipedia commons / CC BY-SA 3.0
The Songyuan (Putonghua) corpus

<table>
<thead>
<tr>
<th>Researcher</th>
<th>Fangfang Li (李芳芳), Psychology, University of Lethbridge</th>
</tr>
</thead>
<tbody>
<tr>
<td>Participants</td>
<td>Young adult speakers of Dongbei Mandarin (10 men, 10 women), recorded in 2006-2007 as part of a study of phonological development.</td>
</tr>
<tr>
<td>Materials</td>
<td>Initial lingual obstruents before a variety of vowels in words such as /ti^51 tʰu^25/ (地图), /se^55 ja^35/ (塞牙), /ta^51 cən̩^51/, /cən̩^35/ (熊), and /su^55 pao^55/ (书包)</td>
</tr>
</tbody>
</table>
The Kaohsiung (Taiwanese & Guoyu) corpus

Researcher  Ya-ting Shih (施雅婷), Teaching Chinese as a Second Language, Chung Yuan Christian University

Participants  Adult speakers of the southern Taiwan varieties of Min Nan and Mandarin (23 young, 21 middle-aged, 20 elderly), recorded in 2011 as part of a study of phonological development in a context of cross-generational language shift.


Materials  Initial fricatives before a variety of vowels, in words such as Taiwanese /sjɔ te/ (熱茶) and Guoyu /ɕon^35/ (熊).
Elicitation: picture-prompted word repetition task

[Image of a bear]

target: [ɕoŋ\(^{35}\)] 熊
Previous work on gender effects in Chinese (1)

**lexical variables**


**morphological variables**


**phonological variables**

Previous work on gender effects in Chinese (2)

Current study builds on previous acoustic studies of gender effects in several varieties of Chinese.

**Fronting of alveolopalatal sibilants in women’s speech**


**Lengthening of aspiration in women’s /pʰ, tʰ, kʰ/**

Gender effects in phonetics — ubiquitous and yet diverse

Diversity (culture-specificity) due to a multiplicity of bases?

sexual dimorphism effects
lowering of the characteristic frequencies of some sounds in men’s speech relative to women’s speech (or vice-versa):
  ▪ Culture-specific exaggerations of physiological differences?

language socialization effects
enhancement of cues to a phonemic contrast in women’s speech relative to men’s speech:
  ▪ Typically dominant role of women in language socialization?

sound change effects
differentiation of men and women in only one age group:
  ▪ Role of women as leading agents in sound changes in progress?
Measuring vocal tract length (VTL) from midsagittal magnetic resonance imaging (MRI) shows a more than 40% increase between (left) MRI for a 4-year-old boy with VTL = 11.28 cm and (right) MRI for a 54-year-old man with VTL = 15.87 cm (Vorperian & Kent, 2007, Fig. 1).
The loci of adult gender differences

Measuring lip thickness (left) and posterior cavity length (right) as two major components of differences in overall VTL between men and women (Vorperian et al., 2009, Figs. 7 & 4).
Longer pharynx (larger head & lowered larynx) in men

Source: Vorperian et al. (2009) Figure 4.
Thicker lips (longer lip tube) in men

Source: Vorperian et al. (2009) Figure 7.
Articulation of the Mandarin sibilants

Relevant dimensions and their relationships

- Front cavity size: a. /s/ < b. /ɕ/ < c. /ʂ/
- Constriction length: a. /s/ < b. /ɕ/ > c. /ʂ/
- Back cavity size: a. /s/ > b. /ɕ/ < c. /ʂ/

Representative images of (a) dental /s/ vs (b) palatoalveolar [ɕ] vs. (c) retroflex [ʂ] for one of four speakers of “Northern varieties of Chinese” in an MRI study of languages with 1, 2, or 3 sibilant fricatives (Toda & Honda, 2003, Figure 3).
Measures of front and back cavity size in /\$/

Spectrum with centroid value:
- 3880 Hz

Spectrogram with onset F2 value:
- 1976 Hz
Measures of front and back cavity size in /ʃ/.

Spectrum with centroid value at 3880 Hz.

Spectrogram with onset F2 value at 1976 Hz.
Measures in /s/ (left) vs /ʃ/ (right)

centroid=8659 Hz

1837 Hz

centroid=4973 Hz

2436 Hz
Measures in /s/ (left) vs /ɕ/ (right)

centroid=8659 Hz

centroid=4973 Hz

dental /s/  

alveolopalatal /ɕ/
The Songyuan (PTH) sibilant fricative space

all 10 male talkers

all 10 female talkers

F2 at vowel onset (Hz)

centroid frequency over middle 40 ms of fricative (Hz)
The Songyuan (PTH) sibilant fricative space

---

The diagram shows scatter plots for the F2 at vowel onset (Hz) and centroid frequency over middle 40 ms of fricative (Hz) for all 10 male talkers and all 10 female talkers. The data is represented by different colors: [s] (brown), [ɕ] (red), and [ʃ] (blue).

- **F2 at vowel onset (Hz)**:
  - Key: [s] (brown), [ɕ] (red), [ʃ] (blue)
  - axis: 1000 to 2500

- **Centroid frequency over middle 40 ms of fricative (Hz)**:
  - Key: [s] (brown), [ɕ] (red), [ʃ] (blue)
  - axis: 4000 to 8000

The labels on the diagram indicate:

- **14m**: alveolopalatal /ɕ/
- **03f**: “fronted” /ɕ/
Teasing apart between- vs. within-speaker variation

all 10 male talkers

all 10 female talkers

median F2 at vowel onset (Hz)

median centroid frequency over middle 40 ms of fricative (Hz)

[ʂ] [ɕ] [s]

onsetF2

women

men
Men with most extreme centroid values in /ʂ/

subject 14m

subject 20m

F2 at vowel onset (Hz)

centroid frequency over middle 40 ms of fricative (Hz)

14m: alveolopalatal /ɕ/

20m: “fronted” /ɕ/
Men with most extreme centroid values in /ɕ/

- **F2 at vowel onset (Hz)**
- **Centroid frequency over middle 40 ms of fricative (Hz)**
- **Subject 14m**
- **Subject 20m**
- [s] represented by gold dots
- [ɕ] represented by red diamonds
- [s] represented by blue squares
All of the centroid values in /ɛ/
Comparing centroid values in /c/ and /s/ (men)

men's /c/ productions

men's /s/ productions

centroid frequency at middle 40 ms (Hz)
Comparing centroid values in */ð*/ and */s*/ (women)

women's */ð*/ productions

women's */s*/ productions

centroid frequency at middle 40 ms (Hz)
Primary school and included provisions for establishing teacher-training courses in pronunciation as well as pedagogy at government-funded centers such as Jílín Teachers' College (see Cheng 1979; Yin and Baldauf 1990; Chen 1999: Chapter 2). Although there was a hiatus when the "Cultural Revolution from 1966 to 1976 dramatically ruined the education system" (Wang 2012: 60), the Ministry of Education then implemented a wide-scale curricular reform in the 1980s, which included a heavy re-investment in local teacher training and local documentation efforts such as the Sūn et al. survey. As a result of this reinvestment, Yin and Baldauf (1990: 285) could state a decade later, "Most students graduating from primary schools can speak Putonghua fluently, and they will continue to speak Putonghua if they go to secondary schools or universities for further study."

Note that in non-Mandarin regions, speaking Pǔtōnghuà meant the spread of societal bilingualism into rural areas and the emergence (or reinforcement) of regional urban L2 Mandarin accents, which could take on the same social meanings that speaking the regional language has always had. For example, Shànghǎi-Fig. 6: Distribution of pronunciations of the anterior sibilant fricative in the Sūn, Lù, and Lǐ (1986: 45) dialect survey map superimposed on a map of administrative divisions in Jílín Province (https://en.wikipedia.org/wiki/File:Jilin_prfc_map.png). Zone 4 is the division of the provincial capital, Jílín City, and zone 7 is the Sōngyuán City administrative division. The metropolitan area proper is the Níngjiāng district, the part at the northern edge of zone 7 that is at the center of the '[ʂ] only' region on the dialect survey map.

The Songyuan sibilant space

- Young women (at least university-educated women) have fronted /ɕ/ relative to men.
- What is the relationship to the long-standing “feminine accent” of the Beijing vernacular?
- Larger sociolinguistic context: (1) Influx of ethnic Han began only in early 1900s. (2) In the generation born after 1976, the standard Putonghua contrast between /s/ and /ʂ/ has been solidified and there are no longer the hypercorrections of [s] for /ʂ/ observed among men of the previous generation (who came of age between 1966 and 1976).
- Some young men are also exaggerating the dental quality of the standard /s/, perhaps to enhance the contrast with /ʂ/ (or perhaps to conform to a more traditionally northern and less southern-oriented “yuppie” norm — cf. Zhang, 2005).
Glottal physiology (Grays [1918 ed.], Titze [1989])

Fig. 960.—Muscles of the larynx, seen from above.

(a) Sagittal View

(b) Horizontal Section
Glottal size differences affect voice quality (and VOT)

Spectrograms of initial 400 ms of 2 talkers’ productions of the word /tʰao⁵⁵ tsi/ (桃子), showing tag points for burst, voice onset, and end of following vowel (V end), with VOT = 69 ms vs 109 ms.
Glottal size differences affect voice quality (and VOT)

male 17 <tao1zi>

female 06 <tao1zi>

17m : pressed voice /tʰ ao⁵⁵/
06f : breathy voice /tʰ ao⁵⁵/
Measuring the aspiration contrast

Spectrograms of one talker’s productions of the words /tʰaʊ 55 tsi/ (桃子), with 69 ms VOT, vs /ta 51 ɕan 51/ (大象), with 9 ms VOT.
Glottal size differences affect voice quality (and VOT)

17m: pressed voice /tʰ ao^55/
17m: pressed voice /ta^51/
06f: breathy voice /tʰ ao^55/
06f: pressed voice /ta^51/
Voice onset time (VOT) in Songyuan (PTH) talkers

subject ID (ordered within gender by median VOT in aspirated stops)

aspirated, women
aspirated, men
Voice onset time (VOT) in Songyuan (PTH) talkers

Subject ID (ordered within gender by median VOT in aspirated stops)

Voice onset time (ms)

- aspirated, women
- aspirated, men
- unaspirated, women
- unaspirated, men
Voice onset time (VOT) in Songyuan (PTH) talkers

subject ID (ordered within gender by median VOT in aspirated stops)
VOT in relationship to a speaking rate measure

<table>
<thead>
<tr>
<th>Type of Stops</th>
<th>Median Voice Onset Time (ms)</th>
<th>Median Duration of Following Vowel (ms)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aspirated Stops</td>
<td>Men: 36, 45, 55, 67, 82</td>
<td>Women: 11, 13, 16, 20, 25</td>
</tr>
<tr>
<td></td>
<td>Unaspirated Stops</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Men: 145, 170, 200, 240</td>
<td>Women: 11, 13, 16, 20, 25</td>
</tr>
</tbody>
</table>
The Songyuan aspiration contrast

- Women have longer VOT values in aspirated stops relative to men.
- They also have slower average speaking rates, as gauged by the following vowel durations.
- However, the correlation between speaking rate and VOT in aspirated stops cannot be a direct, causal relationship, because ...
- Women have **shorter** VOT values in unaspirated stops, suggesting that ...
- Women enhance the contrast between aspirated and unaspirated stops relative to men as part of a larger ensemble of “clear speech” effects that include a slower speech rate.
Larger versus smaller fricative inventories

The two main Chinese varieties in contact in Taiwan have different tone, vowel, and consonant inventories.

The Guoyu fricative inventory

<table>
<thead>
<tr>
<th>place</th>
<th>IPA</th>
<th>example word</th>
<th>phonotactics</th>
</tr>
</thead>
<tbody>
<tr>
<td>labiodental</td>
<td>/f/</td>
<td>/fan\textsuperscript{51}/ (飯)</td>
<td></td>
</tr>
<tr>
<td>dental</td>
<td>/s/</td>
<td>/san\textsuperscript{55} ka/ (三個)</td>
<td>not before /i, y/</td>
</tr>
<tr>
<td>alveolopalatal</td>
<td>/ɕ/</td>
<td>/ɕa\textsuperscript{55} tsì/ (蝦子)</td>
<td>not before /u/</td>
</tr>
<tr>
<td>retroflex</td>
<td>/ʂ/</td>
<td>/ʂa\textsuperscript{55} fa\textsuperscript{56}/ (沙發)</td>
<td>not before /i, y/</td>
</tr>
<tr>
<td>velar</td>
<td>/χ/</td>
<td>/xai\textsuperscript{35} tsì/ (孩子)</td>
<td>not before /i, y/</td>
</tr>
</tbody>
</table>

The Taiwanese Min Nan fricative inventory

<table>
<thead>
<tr>
<th>place</th>
<th>IPA</th>
<th>example word</th>
<th>allophony</th>
</tr>
</thead>
<tbody>
<tr>
<td>alveolar</td>
<td>/s/</td>
<td>/se\textsuperscript{55} sā\textsuperscript{55}/ (洗衣)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>/si\textsuperscript{33} tsiĕn\textsuperscript{55}/ (時鐘) “palatalizes” to [ɕ] before /i, j/</td>
<td></td>
</tr>
<tr>
<td>glottal</td>
<td>/h/</td>
<td>/hei\textsuperscript{13} a\textsuperscript{53}/ (蝦子)</td>
<td></td>
</tr>
</tbody>
</table>
Previous work on the phonetics of Guoyu fricatives

phonetics of L2 and L1 Guoyu (from Peng, 1993)

L2 elderly  L2 middle-aged  L2 young  L1 Guoyu

percent of tokens

0 20 40 60 80 100

/\ as [h]  fricated /x/
Effect of language contact on Min Nan /h/ phonetics of Guoyu /x/ and Min Nan /h/ (from Peng, 1993)
Stage 1: F2 onset differs in [ɕ] vs [s] allophones of MN /s/
Stage 1: F2 onset differs in [ɕ] vs [s] allophones of MN /s/
Stage 1: GY /ɕ/ vs. /s/ differentiated only at vowel onset
Stage 3: GY /ʂ/ vs. /s/ differentiated at fricative center

9 male 20-40 year olds
14 female 20-40 year olds

median F2 at vowel onset (Hz)

median centroid at 20 ms window at frication center (kHz)

assimilation of Guoyu /s/ to Taiwanese /s/ in /ɕe⁵¹ ɕe/
Stage 3: GY /ɕ/ vs. /s/ differentiated at fricative center

9 male 20-40 year olds

14 female 20-40 year olds

median F2 at vowel onset (Hz)

median centroid at 20 ms window at frication center (kHz)
Stage 3: \([\varsigma]\) vs \([s]\) allophones look like Guoyu \(/\varsigma/\) vs \(/s/\)

- 9 male 20-40 year olds
- 14 female 20-40 year olds

\[/s/+V = \begin{array}{c}
[\varsigma i] \\
[\varsigma j c] \\
[\varsigma e] \\
[\varsigma e] \\
[\varsigma a] \\
[\varsigma a] \\
[\varsigma u] \\
[\varsigma u] \\
\end{array} \]

- median F2 at vowel onset (Hz)
- median centroid at 20 ms window at frication center (kHz)
Stage 2: GY production patterns differentiated by gender

- 8 male 41-65 year olds
- 13 female 41-65 year olds

- median F2 at vowel onset (Hz)
- median centroid at 20 ms window at frication center (kHz)
Stage 2: MN allophones also differentiated by gender

8 male 41-65 year olds

13 female 41-65 year olds

\[ /s/+V = \begin{align*}
[\text{ɪ}] & \quad \text{red diamond} \\
[\text{s}j\text{ɛ}] & \quad \text{red square} \\
[\text{ɛ}] & \quad \text{blue triangle} \\
[\text{s}ɛ] & \quad \text{blue square} \\
[\text{sa}] & \quad \text{blue circle} \\
[\text{su}] & \quad \text{brown square}
\end{align*} \]
The Kaohsiung sibilant space(s)

- Stage 1: The oldest speakers assimilate the GY contrast between dental /s/ vs alveolopalatal /ɕ/ to the L1 allophonic variation between an alveolar [s] and a “palatalized” [ɕ].
- Stage 3: Most of the youngest speakers have acquired the phonetics of the GY /s/ vs /ɕ/ contrast, which has influenced their L1 phonetics, to differentiate a more dental [s]allophone from a robustly alveolopalatal [ɕ] allophone for MN /s/.
- Stage 2: In the middle generation, men tend to produce the older assimilation patterns of their parents and women tend to produce the younger allophonic patterns of their children.
The Songyuan Putonghua sibilant space
Many young women have a fronted ([s̩]-like) /ɕ/, much like the long-standing “feminine accent” variant of Beijing.

The Songyuan Putonghua aspiration contrast
These young women have longer VOT values in /tʰ, kʰ/ and shorter VOT values in /t, k/, enhancing the aspiration contrast relative to most men’s productions, a “clear speech” effect?

The Kaohsiung Guoyu sibilant space
Middle-aged women are more advanced in the shift from heavily accented Guoyu sibilants to a Putonghua-like differentiation between /ɕ/ and /s/ (and even /ʂ/).

The Kaohsiung Southern Min sibilant allophones
Middle-aged women are more influenced by Guoyu contrasts in their differentiation of allophones of /s/ in Southern Min.
Acknowledgments

Funding sources

- OSU Center for Cognitive Science Interdisciplinary Summer Fellowships to Fangfang Li and to Ya-ting Shih
- Targeted Investment grant to Fangfang Li & Eunjong Kong
- Alumni Grant for Graduate Research and Scholarship to Ya-ting Shih
- NIDCD grant 02932 to Jan Edwards
- NSF grants BCS-0729306, BCS-0729140, & BCS-0729277 to Mary Beckman, Jan Edwards, & Benjamin Munson

Special thanks ...

- to Marjorie Chan and Li-mei Chen for help in designing the word lists and recruiting participants in Kaohsiung,
- to all of the participants in both Kaohsiung and Songyuan,
- and to you, our audience, for your kind attention.