Acoustic distance effect on the perception of sibilants mergers between retroflexes and alveolars in Taiwan Mandarin (SaSC)

Baichen Du
Department of Linguistics, University of Hong Kong | baichen@connect.hku.hk

Overview
- A production and perception experiment investigated aspects of the Mandarin sibilant merger.
- CoG lower center of gravity (CoG) in production of all sibilants than previously reported.
- Merger-in-progress of retroflexes and alveolars, conditioned by social factors.
- A link between production and perception: Lower center of gravity in production than previous reports.
- Implications on the cause of the merger and sound change: Listenability of environment on perception.
- Social effects of gender and region.
- Future research may study the cause and exact structure of the merger, e.g., category shifting or expanding.

Background
Sibilants in Mandarin
- A three-way contrast among sibilants sounds in Mandarin, distinguished by place and manner of articulation, and aspiration1.
- Alveolars: /tɕ/, /tʂl/, /tʃl/, /ʃl/.
- Retroflexes: /tʃ/, /tɕ/, /ʃr/, /ʃl/.
- Palatalvelars: /n/, /l/.
- Components of sibilants:
  - Affricates: closure, burst, frication, aspiration (production)
  - Frications: frication.
- Canonical production: higher frequency frication for alveolars and lower for retroflexes1,4,5.
- CoG for Alveolars: 9000-10000Hz.
- CoG for Retroflexes: 4000-5000Hz.
- Merger-in-progress: from retroflexes to alveolars1,4,6.
- Loss of retroflex: Fronted place of articulation.
- Acoustic correlate: Increased CoG for retroflex.
- Asymmetry & Unidirectionality

Speech production and perception link
- Bidirectional influence between production repertoire and perceptual space.
- Explained by the Exemplar Model6.
- Ample evidence: speech accommodation/adaptation7,8
  - Second language acquisition9.
- Complexity of the link: “Filtering” of tokens entering production repertoire, causing a mismatch between production and perception7.

Objectives
- The current study investigates if the link applies to the merger.
- Research Question: Is there a production and perception link among speakers in the sibilant merger? Specifically, do merged speakers also shift their perceptual boundary between retroflexes and alveolars?
- Hypothesis: Merged speakers have a perceptual boundary shifted towards higher frequency, i.e., the merger is relatively symmetric between production and perception.
- Null hypothesis: Merged speakers do not shift the boundary, suggesting an asymmetric and complex relationship between production and perception.

Analysis
- Annotation and segmentation of recordings and stimuli in Praat.
- Measurement of four spectral moments of frications:
  - Center of gravity (CoG)
  - Standard deviation
  - Skewness
  - Kurtosis.
- Kolmogorov-Smirnov test on CoG was performed to classify subjects into unmerged and merged speakers.
- Perception results and questionnaire were analyzed using R.
- Tests on the degree of merger predicted by social factors and perceptual boundary predicted by merger, step, and onsets.

Results
1. CoG of sibilants is lower than previous reports found.

Table 1. Production results.

<table>
<thead>
<tr>
<th>Par</th>
<th>Retrorhexis</th>
<th>Alveolar</th>
<th>Difference</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>/p/</td>
<td>3156</td>
<td>6657</td>
<td>3501</td>
<td>420</td>
</tr>
<tr>
<td>/b/</td>
<td>2649</td>
<td>5998</td>
<td>3349</td>
<td>540</td>
</tr>
<tr>
<td>/g/</td>
<td>2413</td>
<td>5232</td>
<td>2734</td>
<td>418</td>
</tr>
</tbody>
</table>

2. Female, northeastern and southeastern speakers lead the merger.

Figure 7. Perceptual boundary of merged and unmerged speakers.

Table 2. Logistic Regression on step, onsets, speaker status, and interactions.

<table>
<thead>
<tr>
<th>Estimate</th>
<th>Std. Error</th>
<th>z value</th>
<th>p (2 df)</th>
</tr>
</thead>
<tbody>
<tr>
<td>/nym/</td>
<td>0.5301</td>
<td>0.0055</td>
<td>97.07</td>
</tr>
<tr>
<td>/nym/</td>
<td>0.5301</td>
<td>0.0055</td>
<td>97.07</td>
</tr>
<tr>
<td>/nym/</td>
<td>0.5301</td>
<td>0.0055</td>
<td>97.07</td>
</tr>
<tr>
<td>/nym/</td>
<td>0.5301</td>
<td>0.0055</td>
<td>97.07</td>
</tr>
<tr>
<td>/nym/</td>
<td>0.5301</td>
<td>0.0055</td>
<td>97.07</td>
</tr>
<tr>
<td>/nym/</td>
<td>0.5301</td>
<td>0.0055</td>
<td>97.07</td>
</tr>
</tbody>
</table>


Conclusion & Future Directions
1. There is indeed a merger in the participants we tested, led by female speakers, consistent with previous sound change studies.
2. Compared to unmerged speakers, merged speakers are more likely to perceive more retroflexes along the same continua.
3. The merger might be conditioned by onsets or vowels, or cue weighting differences in perception among the population.
4. Future research should investigate if the merger is articulatorily or perceptually motivated, if there are any cue weighting differences among the population, if the merger is caused by category shifting or expanding, and if the two categories collapsed in production or perception first.

Selected References

**Table 2:** Logistic Regression on step, onsets, speaker status, and interactions.

**Figure 7:** Perceptual boundary of merged and unmerged speakers.

This study was approved by Human Research Ethics Committee at the University of Hong Kong (HKE). It was made possible by the Transdisciplinary Undergraduate Research Initiative Grant from the Common Core of Research from the University of Hong Kong. I would like to express my special thanks to Dr. Jonathan Havenshill and Dr. Youngah Do at the Language Development Lab of HKU for offering feedback and comments on the project, and Dr. Alexandra Pfiffer and Prof. Keith Johnson at UC Berkeley Linguistics for offering suggestions on the poster. I also thank all the subjects who participated in this study.