Production and Perception of Nasal Coda Mergence in Hohhot Dialect: From a Prosodic Point of View

Ting WANG
Hongwei DING
Tongji University
2012-05-29
Introduction

Method

Analysis, Results and discussion

Conclusion
Introduction
Introduction

- Hohhot is the capital city of the Inner Mongolian Autonomous Region, which is located in north-central China.

- The dialect in Hohhot belongs to Zhangjiakou-Hohhot dialect, which is one of eight main branches of Jin dialect, one of the principal varieties of Chinese [1].

Introduction

- Notable features of the Hohhot dialect include a special intonation for yes-no questions, which is characterized by a prolonged contour at the end of the sentence;
- notable aspiration of [p], [t], and [k] sounds;
- absence of the [ʒ], [tʃ], and [ʃ] sounds, which are respectively changed into [z], [c] and [s];
**Introduction**

- Quansheng Qiao conducted a textual research on the inheritance and evolvement of the nasal initials of syllables of Jin Dialect from the Tang and the Five Dynasties to today [1].

- Zibing Liu discussed how Hohhot native speakers should learn the pronunciation of [n] and [ŋ] [2].

- Xiang Wang analyzed the negative transfer between Hohhot dialect and Mandarin Chinese [3].

• Investigating the prosodic features of nasal coda in Hohhot dialect.

• Finding out the acoustic and perceptive evidences of nasal coda mergence from \([n]\) to \([ŋ]\) in Hohhot dialect.

• Investigating if this nasal coda mergence is an inherent speech perception mode in the Hohhot dialect speaker’s brain.

• Investigating the correlation of production and perception of nasal codas in Hohhot dialect and the tonal aspect of perception.
Method
**Method**

- **Subjects**

<table>
<thead>
<tr>
<th>Group</th>
<th>Subjects</th>
<th>Age</th>
<th>Education background</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Experimental</strong></td>
<td>Male 1(EM1)</td>
<td>51</td>
<td>Senior high school</td>
</tr>
<tr>
<td></td>
<td>Male 2(EM2)</td>
<td>60</td>
<td>Junior high school</td>
</tr>
<tr>
<td></td>
<td>Female1(EF1)</td>
<td>45</td>
<td>Junior high school</td>
</tr>
<tr>
<td></td>
<td>Female2(EF2)</td>
<td>50</td>
<td>Senior high school</td>
</tr>
<tr>
<td><strong>Contrast</strong></td>
<td>Male 1(CM1)</td>
<td>24</td>
<td>University</td>
</tr>
<tr>
<td></td>
<td>Female 1(CF1)</td>
<td>24</td>
<td>University</td>
</tr>
</tbody>
</table>
Method

- Material and recordings

- Ten designed pairs of word are served as the speech text materials.
- Each pair of words only differs from each other with the nasal coda in the same place.
- In each pair, word 1 has the nasal coda [n] while word 2 has [ŋ].
- Four tones are involved in the syllables containing the nasal coda in the twenty words. They are HH (T1), LH (T2), MH (T3) and HL (T4).

<table>
<thead>
<tr>
<th>Pair</th>
<th>Word 1</th>
<th>Word 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pair 1</td>
<td>p’i$n^2$ f’an (频繁)</td>
<td>p’i$\emptyset$ f’an (平凡)</td>
</tr>
<tr>
<td></td>
<td>frequently</td>
<td>ordinary</td>
</tr>
<tr>
<td>Pair 2</td>
<td>s$\emptyset$n$^1$ s$i$ (绅士)</td>
<td>s$\emptyset$ŋ$^1$ s$i$ (声势)</td>
</tr>
<tr>
<td></td>
<td>gentleman</td>
<td>blockbuster</td>
</tr>
</tbody>
</table>

Table : Examples of word pairs
Method

• Listener groups

☐ Two groups of listeners are recruited for the perception tests in this study.

☐ The first group (LG.1) consists of ten Hohhot dialect listeners (five males and five females). Four of them are the subjects in the experimental group. They can only speak Hohhot dialect.

☐ In the second group (LG.2), ten speakers who are competent in standard Mandarin Chinese participate in the experiments.

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>LG.1</strong></td>
<td>Hohhot dialect listeners</td>
<td>10</td>
</tr>
<tr>
<td><strong>LG.2</strong></td>
<td>Mandarin Chinese listeners</td>
<td>10</td>
</tr>
</tbody>
</table>
Acoustic Analysis
Acoustic Analysis

• three important parameters of prosody:

  duration
  fundamental frequency
  intensity
• In the spectrogram of the speech sound involving a nasal coda, there appears a lower strong force frequency, which shows as a dark bar in the lower frequency area on the spectrogram.

• For nasal coda [ŋ], the nasal bar in the spectrogram appears longer and darker than the nasal coda [n]. That is because compared with [ŋ], the nasalization of [n] is in a lesser degree with a smaller resonant intensity which leads to a weaker intensity.

• For this reason, the nasal bar in the spectrogram can be regarded as an acoustic parameter to describe the nasal coda [n] and [ŋ].
In order to eliminate the effect of different segment number in the word and personal factors during recording, we define \( D \) (the ratio of nasal coda length (sec) to the length of the word involving this nasal coda) as:

\[
D = \begin{bmatrix}
\frac{d_{n11}}{d_{w1}}, & \frac{d_{n21}}{d_{w2}}, & \frac{d_{n12}}{d_{w3}}, & \frac{d_{n22}}{d_{w4}}, & \ldots, & \frac{d_{n1x}}{d_{w(y-1)}}, & \frac{d_{n2x}}{d_{wy}}
\end{bmatrix}
\]

Where \( d_n \) refers to the duration of the nasal coda,
- \( d_w \) refers to the duration of the corresponding word;
- \( x \) refers to the pair number (here \( x = 10 \)),
- \( y \) refers to the word number (here \( y = 20 \)),
- \( d_{n1x} \) means the duration of the nasal coda in the first word of pair \( x \),
- \( d_{n2x} \) is the duration of the nasal coda in the second word of pair \( x \).
Acoustic Analysis - Duration

Figure 2: Word 1 in a word pair produced by a male Mandarin Chinese speaker

Figure 3: Word 2 in a word pair produced by a male Mandarin Chinese speaker

Figure 4: Word 1 in a word pair produced by a male Hohhot dialect speaker

Figure 5: Word 2 in a word pair produced by a male Hohhot dialect speaker
Acoustic Analysis - Duration

Dcm1 and Dcm2 are obviously different. Dhm1 is much longer than Dcm1, while Dhm1 is almost the same as Dhm2.

Dcf1 is quite different from Dcf2. However Dhf1 is much longer than Dcf1 and almost the same as Dhf2 in all the word pairs.

Figure 6: Duration ratio for male speaker in Hohhot dialect and Mandarin Chinese

Figure 7: Duration ratio for female speaker in Hohhot dialect and Mandarin Chinese
### Acoustic Analysis - Duration

<table>
<thead>
<tr>
<th>Mean(s)</th>
<th>N(pair)</th>
<th>Std. Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dhm1</td>
<td>0.23206</td>
<td>0.026704</td>
</tr>
<tr>
<td>Dcm1</td>
<td>0.19404</td>
<td>0.047415</td>
</tr>
</tbody>
</table>

**Table 4:** Duration ratio in word 1 for **male speaker**

\[ P=0.028<0.05 \]

<table>
<thead>
<tr>
<th>Mean(s)</th>
<th>N(pair)</th>
<th>Std. Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dhm1</td>
<td>0.23206</td>
<td>0.026704</td>
</tr>
<tr>
<td>Dhm2</td>
<td>0.22641</td>
<td>0.024824</td>
</tr>
</tbody>
</table>

**Table 5:** Duration ratio in word 1 and word 2 for **male speaker**

\[ p=0.289>0.05 \]

<table>
<thead>
<tr>
<th>Mean(s)</th>
<th>N(pair)</th>
<th>Std. Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dhf1</td>
<td>0.23350</td>
<td>0.056717</td>
</tr>
<tr>
<td>Dcf1</td>
<td>0.18642</td>
<td>0.048968</td>
</tr>
</tbody>
</table>

**Table 6:** Duration ratio in word 1 for **female speaker**

\[ P=0.002<0.05 \]

<table>
<thead>
<tr>
<th>Mean(s)</th>
<th>N(pair)</th>
<th>Std. Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dhf1</td>
<td>0.23350</td>
<td>0.056717</td>
</tr>
<tr>
<td>Dhf2</td>
<td>0.23567</td>
<td>0.051768</td>
</tr>
</tbody>
</table>

**Table 7:** Duration ratio in word 1 and word 2 for **female speaker**

\[ p=0.091>0.05 \]
Acoustic Analysis -- F₀

- Due to the influence of the nasal coda, the nasal vowel has a higher fundamental frequency (F₀) and a lower first formant (F₁) [1]. Therefore, the nasal vowel before [n], compared with the same nasal vowel before [ŋ], has a lower F₀ but a higher F₁.

- Normalization

\[ F_{st} = 12 \times \log_2 \left( \frac{f}{f_r} \right) \]

Where \( F_{st} \) is the transformed F₀;
\( f \) is the original F₀;
\( f_r \) is the reference value, 55 Hz for male and 80 Hz for female.

**Acoustic Analysis-- F₀**

- $F_{0c1}$ stands for the average $F_0$ of the nasal vowel in the first word in each word pair produced by Mandarin Chinese speakers and $F_{0c2}$ in the second word.
- $F_{0h1}$ stands for the average $F_0$ of the nasal vowel in the first word in each word pair produced by Hohhot dialect speakers, and $F_{0h2}$ in the second word.

$F_{0c1}$ and $F_{0c2}$ appear differently. $F_{0h1}$ and $F_{0h2}$ are nearly the same, but larger than $F_{0c1}$.

*Figure 8: Average $F_0$ of the nasal vowel*
### Acoustic Analysis-- $F_0$

<table>
<thead>
<tr>
<th></th>
<th>Mean(st)</th>
<th>N(word)</th>
<th>Std. Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Foh1</td>
<td>17.679837</td>
<td>20</td>
<td>3.7556139</td>
</tr>
<tr>
<td>Foh2</td>
<td>17.493915</td>
<td>20</td>
<td>3.9285395</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Mean(st)</th>
<th>N(word)</th>
<th>Std. Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Foh1</td>
<td>17.679837</td>
<td>20</td>
<td>3.7556139</td>
</tr>
<tr>
<td>Foc1</td>
<td>15.905799</td>
<td>20</td>
<td>3.9300937</td>
</tr>
</tbody>
</table>

Table 8: Average $F_0$ in word 1 and word 2 for Hohhot speakers

Table 9: Average $F0$ in word 1 for all speakers

$p=0.111>0.05$  

$p=0.000<0.05$
• In Mandarin Chinese, the intensity of the nasal coda [ŋ] is larger than [n] due to the different degree of nasalization.
• Here we may wonder if it is the same in Hohhot dialect.

Figure 9: Average intensity of nasal codas (male Hohhot dialect speakers)

Figure 9: Average intensity of nasal codas (female Hohhot dialect speakers)

p=0.361>0.05

p=0.886>0.05
• Hohhot dialect speakers cannot produce the nasal coda [n] since nasal coda [n] has obvious difference from that in Mandarin Chinese but are the same with [ŋ] in Hohhot dialect. [n] changes to [ŋ] when it works as a coda.

• $n \rightarrow \eta / X_-$

• where X stands for the other segments in the word.

• This rule means that in Hohhot dialect, nasal [n] merges into [ŋ] when it appears as a coda.
Perceptual Analysis
### Table 3: Information of the perception tests

<table>
<thead>
<tr>
<th></th>
<th>Stimuli</th>
<th>Listener</th>
<th>Options</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Test.1</strong></td>
<td>Hohhot dialect</td>
<td>Mandarin Chinese speaker</td>
<td>Word1, Word2, Uncertainty</td>
</tr>
<tr>
<td><strong>Test.2</strong></td>
<td>Hohhot dialect</td>
<td>Hohhot dialect speaker</td>
<td>Word1, Word2, Uncertainty</td>
</tr>
<tr>
<td><strong>Test.3</strong></td>
<td>Mandarin Chinese</td>
<td>Hohhot dialect speaker</td>
<td>Word1, Word2, Uncertainty</td>
</tr>
</tbody>
</table>
Inherent speech perception mode in the Hohhot dialect speaker’s brain

Figure 11: Accuracy rate of Test.1 by LG.2

Figure 12: Accuracy rate of Test.2 by LG.1

Figure 13: Accuracy rate of Test.3 by LG.1

the nasal coda mergence is an inherent speech perception mode in the Hohhot dialect speaker’s brain
Correlation between production and perception

The accuracy of production is highly correlated with that of perception \((r=0.981)\) and the correlation is significant \((p=0.012 < 0.05)\)
Tonal aspect of perception

Figure 15: Accuracy of different tones

Among the correct choices by four subjects in test. 3, nasal codas with T1 and T4 are perceived much more accurately than the other two. Nasal codas with T2 is perceived worst.
Conclusion
Conclusion

- Hohhot dialect speakers may not produce the nasal coda [n].
- There appears no significant difference between [n] and [ŋ] in Hohhot dialect, and there exists a mergence from [n] to [ŋ] in Hohhot Dialect.

- Hohhot dialect speakers cannot discern the distinction between [n] and [ŋ] in both Hohhot dialect and the Mandarin Chinese.
- This mergence is an inherent speech perception mode in the Hohhot dialect speaker’s brain.

- For each individual subject, the accuracy of production of nasal codas is highly correlated with that of perception.
- Perception performances of nasal codas in Tone 1 and Tone 4 are much better than those in Tone 2 and Tone 3.
- Perceptual accuracy of nasal codas in Tone 2 is worst.
Acknowledgements

- Innovation Program of Shanghai Municipal Education Commission (No. 12ZS030)
- Cross-linguistic Comparison of Speech Prosody, and Error Analysis and Automatic Assessment of Second Language Prosody (2010JDXM024)

References

Thank you for your attention!

For more questions, please write to:

Ting WANG
2011ting_wang@tongji.edu.cn