Danish Acquisition of Mandarin Chinese Tones

Marjoleine Sloos\textsuperscript{1}, Mengzhu Yan\textsuperscript{2}, Jie Liang\textsuperscript{3}, Chun Zhang\textsuperscript{4}

\textsuperscript{1}Interacting Mind Centre, Aarhus University, \textsuperscript{2}Linguistics, Aarhus University, \textsuperscript{3}Tongji University, \textsuperscript{4}Chinese Studies, Aarhus University
Prosody in L2 acquisition

Intonation is the most important cue for native speakers to recognize L2 accent (Wenk, 1985)

But intonation—and prosody in general—is rarely explicitedly taught (Wenk, 1985)

However, L2 learners seem to pick it up to a certain extent

- German vowel reduction is largely acquired by learners with Chinese, Italian, or Japanese as their native language (Gut 2003)
- Rhythm of Chinese is largely acquired by Danish learners (Sloos et al 2015-OCOCOSDA)
L2 acquisition of tone

Lexical tones are taught in the classroom
- usually in isolation or disyllabic words

Not explicitly taught
- Tones in connected speech
- Neutral tones
- Pitch range
Tones in connected speech

**Carry-over effects**

Fig. 2. Effects of preceding tone on the $F_0$ contour of the following tone in Mandarin. In each panel, the tone of the second syllable is held constant (H in panel a and R in panel b, while the tone of the first syllable varies among H, R, L and F. The vertical lines indicate the syllable boundaries (at the onsets of initial nasals) (Xu, 1997).

**Anticipatory effects**

Fig. 3. Effects of following tone on the $F_0$ contour of the preceding tone in Mandarin. In each panel, the tone of the first syllable is held constant (H in panel a and R in panel b), while the tone of the second syllable varies among H, R, L and F. The vertical lines indicate the syllable boundaries (at the onsets of initial nasals) (Xu, 1997).
Tones in connected speech

Tonal contours are fluently connected with each other.

Pitch and pitch contours are targeted and approximated, i.e. the target is not totally attained. Xu & Wang (2001)
Pitch Target Approximation Model (Xu & Wang 2001)

- T1: a high target
- T3: a low target
- T2: the rising contour as target
- T4: the falling contour as target
Pitch Target Approximation Model (Xu & Wang 2001)

Fig. 1. A schematic illustration of hypothetical pitch targets (dashed lines) and their surface realization (solid curved line). The three vertical lines represent the boundaries of the two consecutive pitch target-carrying units. The level dashed line on the right of the figure represents a static pitch target [low]. The oblique dashed line on the left represents a dynamic pitch target [rise]. In both cases, the targets are asymptotically approximated.
Pitch Target Approximation Model (Xu & Wang 2001)

The target peak of T1 and target valley of T3 occur close to the offset of the host syllable.

The onset of the rise of T2 and the onset of the fall of T4 are located in the centre of the host syllable.

The target peak of T2 and target valley of T4 occur after the host syllable.

The rise of T2 is slow in the beginning, the peak velocity is located towards the end of the syllable.
L2 acquisition of Chinese prosody

General impression among instructors
- word level pitch contours are not problematic for second language learners
- the problem in tone acquisition lies in the implementation of pitch contours on the sentence level (Yang, 2011:2)

Research on the L2 perception of Chinese tones is abundant

Production has mainly been investigated on the word level

Only limited research on L2 tone production in connected speech
American English-accented Chinese

- Yang (2013) showed that native speakers usually don’t attain the extreme low and high targets of the tones in connected speech.

- Rising tones were produced as falling tones in almost 50% of the cases and that more than half of the speakers produced falling tones as rising tones (Jian 2015).

- L2 learners often confused the high level tone with the falling tone and the low dipping with the rising tone (Yang 2015).

- Lacking: a detailed analysis of the form of the pitch contours of tones in foreign-accented connected speech as compared to native Standard Mandarin.
Questions

Do Danish learners of Chinese acquire Tones in connected speech? Neutral tones? Pitch range?
Design

Reading task

• Carrier sentence: Wǒ X shì Y
  • where X is a kinship term with T0 on the second syllable
  • Y is either yīshēng or lǎoshī
• eight sentences with the same structure, in which all tones, including T0, occur several times.
• alternated with other 81 sentences of the same length (6 syllables) but with different structures
我伯伯是医生。  
Wǒ bóbo shì yīshēng.  
My uncle is a doctor.

我弟弟是老师。  
Wǒ dìdi shì lǎoshī.  
My younger brother is a teacher.

我哥哥是老师。  
Wǒ gēge shì lǎoshī.  
My elder brother is a teacher.

我婶婶是老师。  
Wǒ shěnshen shì lǎoshī.  
My aunt is a teacher.

我爸爸是医生。  
Wǒ bàba shì yīshēng.  
My father is a doctor.

我爷爷是老师。  
Wǒ yéye shì lǎoshī.  
My grandpa is a teacher.

我妈妈是医生。  
Wǒ māma shì yīshēng.  
My mother is a doctor.

我姐姐是医生。  
Wǒ jiějie shì yīshēng.  
My elder sister is a doctor.
Analysis

ProsodyPro script (Xu, 2013)
- Measures the F0 at 10 equidistant points of the rhyme
- Pitch range
- Pitch values were speaker-normalized by Herz-to-semitone conversion
  - (formula: $12 \times \log_2(\text{Hz})$ according to Patel (2006))

For each speaker
- mean pitch
- pitch range
- the standard deviation (as a measure for the variability of pitch movements)
Subjects

Twelve undergraduate students enrolled in the Chinese Studies program at Aarhus University in Denmark.

- Danish native BA student
- 4 males, 8 females
- had spent one semester in Beijing as part of their studies
- about to graduate

Also one female and one male native speaker of Standard Mandarin

- to compare the pitch contours of Danish- accented with native Chinese
Procedure

Microsoft PowerPoint presentation

• Each sentence was presented on a single slide in both Chinese characters and in (the alphabetical) pinyin orthography (tones were indicated with standard diacritics, e.g. ㄇㄢ, ㄇㄚ, ㄇㄢˇ, ㄇㄚˋ, ㄇㄚ, ma)

• The experiment was self-paced

• Duration 15-20 minutes on average

• Recorded with a Tascam DR05 V2 recorder in the linguistics lab at Aarhus University
<table>
<thead>
<tr>
<th></th>
<th>Mean Pitch (Hz)</th>
<th>Maximum Pitch range (st)</th>
<th>Relative Standard Deviation (in %)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Native Standard Mandarin</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>217.9</td>
<td>32</td>
<td>55.4</td>
</tr>
<tr>
<td>Male</td>
<td>125.7</td>
<td>24</td>
<td>34.5</td>
</tr>
<tr>
<td><strong>Native Danish</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>?</td>
<td>5-10</td>
<td>0.5-5.2</td>
</tr>
<tr>
<td><strong>Danish-accented</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>189.4</td>
<td>27</td>
<td>40.4</td>
</tr>
<tr>
<td>Male</td>
<td>138.4</td>
<td>25</td>
<td>26.4</td>
</tr>
</tbody>
</table>

Pitch

Similar finding

- the pitch range of native speakers of Chinese was 1.5 times as large as the pitch range of English
- but when English speakers spoke Chinese, their pitch range increased, although not as much as that of the native speakers
- Chen (1974)
Female Native Speaker

Wǒ bóbo shì yīshēng.
Wǒ dìdi shì lǎoshī.
Wǒ gēge shì lǎoshī.
Wǒ shěnshen shì lǎoshī.
Wǒ bàba shì yīshēng.
Wǒ yéye shì lǎoshī.
Wǒ māma shì yīshēng.
Wǒ jiējie shì yīshēng.

Native Male Speaker

Wǒ bóbo shì yīshēng.
Wǒ dìdi shì lǎoshī.
Wǒ gēge shì lǎoshī.
Wǒ shěnshen shì lǎoshī.
Wǒ bàba shì yīshēng.
Wǒ yéye shì lǎoshī.
Wǒ māma shì yīshēng.
Wǒ jiējie shì yīshēng.
Female Native Speaker

Wǒ bóbo shì yīshēng.
Wǒ dìdi shì lǎoshī.
Wǒ gēge shì lǎoshī.
Wǒ shěnshēn shì lǎoshī.
Wǒ bàba shì yīshēng.
Wǒ yéye shì lǎoshī.
Wǒ māma shì yīshēng.
Wǒ jiějie shì yīshēng.

Native Male Speaker

Wǒ bóbo shì yīshēng.
Wǒ dìdi shì lǎoshī.
Wǒ gēge shì lǎoshī.
Wǒ shěnshēn shì lǎoshī.
Wǒ bàba shì yīshēng.
Wǒ yéye shì lǎoshī.
Wǒ māma shì yīshēng.
Wǒ jiějie shì yīshēng.
Female Native Speaker

Wǒ bóbo shì yīshēng.
Wǒ dìdi shì lǎoshī.
Wǒ gēge shì lǎoshī.
Wǒ shěnshen shì lǎoshī.
Wǒ bàba shì yīshēng.
Wǒ yéye shì lǎoshī.
Wǒ māma shì yīshēng.
Wǒ jiējie shì yīshēng.

Native Male Speaker

Wǒ bóbo shì yīshēng.
Wǒ dìdi shì lǎoshī.
Wǒ gēge shì lǎoshī.
Wǒ shěnshen shì lǎoshī.
Wǒ bàba shì yīshēng.
Wǒ yéye shì lǎoshī.
Wǒ māma shì yīshēng.
Wǒ jiējie shì yīshēng.
Danish-accented intonation

- Flat...
Danish-accented intonation

• Creaky....
Danish-accented intonation

- More than native carry-over effects...

M1

- Wǒ bóbo shì yīshēng.
- Wǒ dìdi shì lǎoshī.
- Wǒ gēge shì lǎoshī.
- Wǒ shěnshen shì lǎoshī.
- Wǒ bàba shì yīshēng.
- Wǒ yéye shì lǎoshī.
- Wǒ māma shì yīshēng.
- Wǒ jiējie shì yīshēng.
Danish-accented intonation

• Inconsistent
Danish-accented intonation

- No T3 sandhi
- L% tone
Patterns

- Initial T3 is largely correctly realized but sandhi is not applied.
- T4 in general is not as steep as in native production and the onset often too low.
- Final T1, although often correctly realized:
  - is sometimes subject to more-than-native carryover effects.
  - an incorrect L% boundary tone.
  - and replacement by a rising contour.
Patterns

• For the tonal distinction between T1 T2 T3 and T4
  – Distinction between T1 and T4 is almost never attained
  – A ‘default’ rising-falling contour is often realized for the sequences L-H-0 and L-F-0
  – Often, the sequence L-R-0 is similarly realized as this rising-falling contour
Default rising-falling contour

F6

Wǒ bóbo shì yīshēng.
Wǒ dìdi shì lǎoshī.
Wǒ gēge shì lǎoshī.
Wǒ shěnshen shì lǎoshī.
Wǒ bàba shì yīshēng.
Wǒ yéye shì lǎoshī.
Wǒ māma shì yīshēng.
Wǒ jiějie shì yīshēng.
Underlying mechanisms

**Errors**
- tone shift
  - T3 is replaced by T1 or vice versa

**Native language transfer**
- low pronunciation of sentence final T1
  - Although a high-level tone in sentence-final position is usually lowered, also in native speech, it is expected to be higher than a preceding T3. In our data, this is not always the case.
- some speakers more or less consistently used a global gradually declining sentence intonation with a low boundary tone
Underlying mechanisms

Inherent articulatory difficulty

- rising-falling contours appear like transition contours from a low tone to a high peak
- Initial T3 is often realized as a low tone, clearly distinct from the rising-falling contour → the subjects in general paid attention to tone realization, but failed to implement the subtle differentiation in timing of the turning point of the peak of T1, T2, T4 in connected speech, which can be observed in the native speaker’s productions.
Patterns

It requires high flexibility of the vocal chords to attain the required (native) velocity in pitch rises and falls

- Difficult to acquire for L2 learners
- The fact that the subjects had a faster-than-native realization of the syllables was certainly not an advantage for accurate tone implementation
- Extensive creaky voice of some speakers is of a disadvantage as well
- T4 turned out not to have such a steep contour as in native speech
- The standard deviation of F0 is lower than in native production
  - Compare: slower rises and falls for German-accented Chinese than for native Chinese speakers Ding et al. (2010)
Female Native Speaker

Wǒ bóbo shì yīshēng.
Wǒ dìdi shì lǎoshī.
Wǒ gēge shì lǎoshī.
Wǒ shěnshen shì lǎoshī.
Wǒ bàba shì yīshēng.
Wǒ yéye shì lǎoshī.
Wǒ māma shì yīshēng.
Wǒ jiějie shì yīshēng.

Native Male Speaker

Wǒ bóbo shì yīshēng.
Wǒ dìdi shì lǎoshī.
Wǒ gēge shì lǎoshī.
Wǒ shěnshen shì lǎoshī.
Wǒ bàba shì yīshēng.
Wǒ yéye shì lǎoshī.
Wǒ māma shì yīshēng.
Wǒ jiějie shì yīshēng.
Like in native production, T0 pronounced by Danish learners of Chinese seems less stable than the lexical tones.

T0 shows characteristics of

- as a transition between the contours of the adjacent syllables
- progressive assimilation
- stretch of the preceding contour
Conclusions

• The pitch range was largely acquired
• But pitch modulation, as expressed by the standard deviation of F0, was lower than in native speech
• The inability of the second language learners to use F0 with native-like flexibility
  – falling contours that are not steep enough
  – inaccurate timing of the turning point or peak dislocation in the continuously changing F0 in connected speech
• Moreover, the extensive use of creaky voice of a considerable part of the subjects could hinder flexible use of the F0
References

Mange tak!

谢谢