THE PHONETIC CHANGE OF TONE SANDHI IN SHANGHAI CHINESE

Yasunori TAKAHASHI

National Institute for Japanese Language and Linguistics
ytakahashi@ninjal.ac.jp; tufs.y.takahashi@gmail.com

ABSTRACT

In Shanghai Chinese, “(broad) tone sandhi” occurs in a polysyllabic expression (i.e., word or phrase) in which the overall pitch of the tonal domain is determined by the tone of the initial syllable, and the F0 contour constantly falls from the second or third syllable. Although it is assumed that the phonetic realization of the pitch-fall in non-initial syllables differs across generations, this assumption is not supported by any objective data because the falling pattern in the newest variety (New Shanghai) has not been investigated in detail. To test this assumption, this study examines the phonetic realization of pitch-fall in New Shanghai. The results show that the rate of fall is very fast at first but rapidly decreases as time goes on, which is clearly different from the falling pattern in the older variety (Middle Shanghai). These results support the idea that the (default) low tone in non-initial syllables has changed from a “weak” target in Middle Shanghai to a “strong” one in New Shanghai.

Keywords: Shanghai Chinese, phonetics, tone sandhi, phonetic change

1. INTRODUCTION

Shanghai Chinese is a regional Chinese dialect spoken in Shanghai. According to Xu et al. [3], it has three varieties: Old, Middle, and New Shanghai, which are spoken by the old, middle, and young generations, respectively. The present study focuses on the latter two varieties.

Middle and New Shanghai have five tones in citation forms: high falling (T1: yinping), high rising (T2: yinu), low rising (T3: yangqu), short high (T4: yinru), and short low rising (T5: yangru). In a polysyllabic expression (i.e., word or phrase), the overall pitch shape is determined by the tone of the initial syllable. This tonal change is called “(broad) tone sandhi” (Chen [1], Xu et al. [3], among others).

There are several studies investigating the phonetic realization of Shanghai tone sandhi (Chen [1], Zee and Maddieson [7], Zhu [8], among others.). For example, Chen [1] comprehensively examined the phonetic realization of tone sandhi in Middle Shanghai (see Figure 1) and found that (1) the F0 contour of the tonal domain, particularly the first two syllables, is significantly influenced by the tone of the initial syllable, and (2) starting from the third syllable, all F0 contours are clearly falling.

In addition, Chen [1] reports that although gradual F0 lowering was the general pattern in tri- and quadri-syllabic expressions, as shown in Figures 1b-c, her three youngest Middle Shanghai speakers (all over age 50) sometimes produced a sharp F0 fall over the third syllable in quadri-syllabic expressions. She suggests that these speakers may be influenced by the younger generation because younger speakers (underage 30) may produce this sharply falling F0.

This would further indicate that a reinterpretation of the low tonal target over the third syllable may have occurred in Shanghai Chinese. According to Chen [1], older speakers regard it as “weak”, comparable to the neutral tone in Standard Chinese (cf. Chen and Xu [2]), while the younger generation of New Shanghai speakers may reinterpret it as “strong”. However, these assumptions are not supported by any objective data because F0 trajectories of pitch-fall in New Shanghai have not been investigated in detail.

The aim of this study is to compare the falling pattern of tone sandhi in New Shanghai with that in Middle Shanghai by examining the phonetic realization of the fall in the newest variety.

2. METHODS

2.1. Speech materials

In the present study, the targets of the investigation were restricted to tri- and quadri-syllabic noun phrases whose initial tones are T1-T4. Phrases that have T5 in the initial position were excluded because they have a different pattern of phonetic realization (i.e., a continuous pitch rise: Xu et al. [3]).
A frame sentence with a slot (1) was employed in order to elicit F0 contours of tri- and quadri-syllabic phrases. Within the frame sentence, the target phrases (shown in Table 1) were inserted in X. Each target phrase can form a tonal domain by itself.

1. Frame sentence of the experiment

I saw the fourth son at X.

Table 1: Target phrases of the experiment

<table>
<thead>
<tr>
<th>Initial tone</th>
<th>Tri-syllabic</th>
<th>Quadri-syllabic</th>
</tr>
</thead>
<tbody>
<tr>
<td>T1 (falling)</td>
<td>‘AB Road’</td>
<td>‘AB Univ.’</td>
</tr>
<tr>
<td></td>
<td>[su. tsv. lu]</td>
<td>[su. tsv. da. oʔ]</td>
</tr>
<tr>
<td></td>
<td>‘Suzhou Road’</td>
<td>‘Suzhou Univ.’</td>
</tr>
<tr>
<td></td>
<td>‘Wenzhou Road’</td>
<td>‘Wenzhou Univ.’</td>
</tr>
<tr>
<td>T2 (high rising)</td>
<td>广州大学</td>
<td>[ku. tsv. da. oʔ]</td>
</tr>
<tr>
<td></td>
<td>‘Guangzhou Road’</td>
<td>[ku. tsv. da. oʔ]</td>
</tr>
<tr>
<td></td>
<td>[kur. tsv. lu]</td>
<td>‘Guizhou Univ.’</td>
</tr>
<tr>
<td></td>
<td>‘Guizhou Road’</td>
<td>‘Guizhou Univ.’</td>
</tr>
<tr>
<td>T3 (low rising)</td>
<td>兰州大学</td>
<td>[li. tsv. da. oʔ]</td>
</tr>
<tr>
<td></td>
<td>‘Lanzhou Road’</td>
<td>[li. tsv. da. oʔ]</td>
</tr>
<tr>
<td></td>
<td>[hi. tsv. lu]</td>
<td>‘Wenzhou Univ.’</td>
</tr>
<tr>
<td></td>
<td>‘Zhaozhou Road’</td>
<td>‘Wenzhou Univ.’</td>
</tr>
<tr>
<td>T4 (short high)</td>
<td>福州大学</td>
<td>[fo. tsv. da. oʔ]</td>
</tr>
<tr>
<td></td>
<td>‘Fuzhou Road’</td>
<td>[fo. tsv. da. oʔ]</td>
</tr>
<tr>
<td></td>
<td>[tsv. lu]</td>
<td>‘Dezhou Road’</td>
</tr>
<tr>
<td></td>
<td>[tsv. da. oʔ]</td>
<td>‘Dezhou Univ.’</td>
</tr>
</tbody>
</table>

2.2. Subjects

Five native speakers of New Shanghai, two males and three females, participated. All speakers were born between 1983 and 1988 and had lived in the urban area (the so-called “shiqia (市區)” of Shanghai since childhood. All speakers were studying at various universities in Tokyo at the time of the experiment.

2.3. Procedures

Recording was conducted in a sound-treated booth in the Phonetics Laboratory at Tokyo University of Foreign Studies. The target sentences were presented in Chinese at random. The speakers read them five times at their normal speech rate. Recording was made using a Marantz PMD 660 at a 44.1 kHz sampling rate.

2.4. Measurements and normalization

The recorded materials were analyzed using the Praat software package. Nine F0 points were measured at 10% to 90% of the vowel duration of the target syllables by using the ProsodyPro script (ver 3.0: Xu [4]). In quadri-syllabic phrases, however, it was difficult to separate the third syllable [da] from the fourth syllable [oʔ] because they overlapped in the waveform and spectrogram (see Figure 2). Therefore, nine F0 points in the combined third and fourth syllables were measured, and the author assumed that the F0 values of points from 10 to 50% of the vowel duration correspond to the third syllable and those from 50 to 90% correspond to the fourth syllable.

Figure 2: Waveform and spectrogram of [fi. tsv. da. oʔ] ‘Yangzhou Univ.’ produced by F-SY (The portion within the black outline represents the third and fourth syllables [da. oʔ])

Each speaker’s measured F0 values were transformed into an LZ-score (Zhu [8, 9]). This score is calculated as the difference between the logarithmic F0 (log X) and the logarithmic mean F0 (log X̄) of each subject, divided by the logarithmic standard deviation (log SD) of the overall F0 of the same speaker.

2) LZ-score

\[ LZ = (\log X - \log X̄) / \log SD \]

In addition, the F0 contours were time-normalized. All times for the points measured at 10 to 90% points of the vowel duration of the target syllables were measured using the ProsodyPro script (Xu [4]), and the average times for all speakers were calculated.

3. RESULTS

Figure 3 (shown on the next page) illustrates the speaker-normalized pitch contours of the target phrases (hereafter, the initial syllable is used to refer to each tone sandhi category; e.g., T1 sandhi). The origin of the time axis (x-axis) corresponds to the starting point of the pitch-fall: the 10% point of the vowel duration in the second syllable in T1 sandhi (Figure 3a) and in the third syllable in T2-T4 sandhi (Figures 3b-d). The pitch contours of tri- and quadri-syllabic phrases are shown with dotted and solid lines, respectively. In a quadri-syllabic phrase, a square indicates the boundary between the third and fourth syllables.

Figure 3: Pitch contours of New Shanghai tone sandhi in tri-syllabic (dotted line) and quadri-syllabic (solid line) phrases
Tables 2 and 3 show the mean magnitude of fall of each syllable in tri- and quadri-syllabic T1 phrases, respectively. This score is calculated as the difference between the F0 values of the initial and final points of F0 measurement, which takes a negative value when a rise occurs.

In Tables 2 and 3, the magnitudes of the falls in the second syllable are far greater than those in the following syllables, and all subjects have negative values in the fourth syllable, which means that rises occurred there. To evaluate the significance of the differences in magnitude in Tables 2 and 3, two statistical analyses were conducted for each subject: a paired t-test between the second and third syllables in tri-syllabic phrases, and a repeated measures ANOVA with fall magnitude as the dependent variable and Syllable (2nd, 3rd, and 4th σ) as independent variables in quadri-syllabic phrases. The results of the t-tests revealed significant differences between the fall magnitudes of the second and third syllables in all subjects [F-SY: t(9) = 16.1, p < .0001, F-HC: t(9) = 14.6, p < .0001, F-ZZ: t(9) = 14.2, p < .0001, M-WJ: t(9) = 7.7, p < .0001, M-XL: t(9) = 6.0, p < .0001]. The results of the ANOVA also revealed significant effects of Syllable on the fall magnitude in all subjects [F-SY: F(2, 27) = 131.6, p < .0001, F-HC: F(2, 27) = 89.3, p < .0001, F-ZZ: F(2, 27) = 214.3, p < .0001, M-WJ: F(2, 27) = 279.5, p < .0001, M-XL: F(2, 27) = 55.2, p < .0001]. Tukey post-hoc tests showed that in all subjects, the fall magnitude in each syllable was significantly different from that in every other syllable.

In T2-T4 sandhi (Figures 3b-d), the pitch falls sharply in the third syllable but becomes nearly level in the fourth syllable. Table 4 shows the mean magnitude of fall of each syllable in T2-T4 phrases.

### Table 4: Mean fall magnitude of T2-T4 sandhi (in Hz; F: female, M: male)

<table>
<thead>
<tr>
<th>Subject</th>
<th>SY</th>
<th>HC</th>
<th>ZZ</th>
<th>WJ</th>
<th>XL</th>
</tr>
</thead>
<tbody>
<tr>
<td>3rd σ</td>
<td>50.8</td>
<td>69.3</td>
<td>31.6</td>
<td>31.8</td>
<td>23.9</td>
</tr>
<tr>
<td>4S</td>
<td>72.6</td>
<td>50.9</td>
<td>51.3</td>
<td>41.7</td>
<td>18.1</td>
</tr>
<tr>
<td>3rd σ</td>
<td>-3.6</td>
<td>-1.2</td>
<td>1.0</td>
<td>8.8</td>
<td>0.8</td>
</tr>
<tr>
<td>4th σ</td>
<td>-6.4</td>
<td>-1.2</td>
<td>1.0</td>
<td>8.8</td>
<td>0.8</td>
</tr>
</tbody>
</table>

In quadri-syllabic phrases, the magnitudes of the falls in the third syllable are far greater than those in the fourth syllable, which take negative values in two speakers (i.e., F-SY and F-HC). The results of a paired t-test between the fall magnitudes of the third and fourth syllables in quadri-syllabic phrases revealed that there were significant differences in all subjects [F-SY: t(29) = 25.8, p < .0001, F-HC: t(29) = 27.1, p < .0001, F-ZZ: t(29) = 25.8, p < .0001, M-WJ: t(29) = 15.6, p < .0001, M-XL: t(29) = 15.3, p < .0001].

### 4. DISCUSSION AND CONCLUSION

This study examined the phonetic realization of the pitch-fall pattern in New Shanghai. The results of the experiment show that, in all circumstances, the pitch contour first falls very sharply, then, the rate of fall rapidly decreases. In quadri-syllabic phrases, pitch-fall tends to stop in the final syllable, and in fact a pitch rise sometimes occurs there (especially in T1 sandhi).

The pitch-fall pattern in New Shanghai is clearly different from that in Middle Shanghai. As shown in Figure 1, Middle Shanghai generally has gradual pitch-fall in both tri- and quadri-syllabic phrases. In contrast, New Shanghai
has “asymptotic” pitch-fall: the rate of fall is very fast at first but rapidly decreases as time goes on. This indicates that, as Chen [1] suggests, a diachronic change has occurred in the phonetic realization of Shanghai tone sandhi. Chen [1] argues that the low tone on the third syllable has a “weak” target in Middle Shanghai, while it has become “strong” in New Shanghai. According to Chen and Xu [2], the implementation of a “weak” tonal target is slower and less robust than that of a “strong” (or lexical) one; therefore, in New Shanghai, the implementation of the low tone on the third syllable should be faster and more robust as compared to that in Middle Shanghai. This prediction is fully supported by the results of the experiment above.

This study also demonstrates that a phonetic value can change in a neutral element. In Shanghai Chinese, non-initial syllables can be considered “neutral” because their pitch values are determined by the tone of the initial syllable and, according to phonological studies like Zee [6], the pitch-fall from the third syllable is due to the insertion of a “default” low tone. Based on these assumptions, the results of the experiment above show that the phonetic value of the default low tone can change across generations, which indicates that the value of a “default” element may be different in each language or variety rather than universal (cf. Yip [5]).

Further studies are needed to understand how this diachronic change in pitch-fall developed in Shanghai Chinese. Chen [1] found that the youngest speakers of Middle Shanghai (who are over age 50) sometimes produced a sharp F0 fall similar to that in Figure 2. In contrast, the young speakers in this study (under age 30) consistently produced it, whereas gradual falling, which is the general pattern in Middle Shanghai, did not occur at all. It is still not clear whether Chen’s youngest speakers of Middle Shanghai were influenced by the younger generation, as she claimed, or whether they had begun changing the phonetic realization of pitch-fall little by little. To fully understand the process of this phonetic change, it is necessary to conduct a closer investigation of the phonetic realization of pitch-fall across generations of Shanghai Chinese speakers.

5. ACKNOWLEDGMENTS

The research reported in this paper was supported in part by a project fund from the National Institute for Japanese Language and Linguistics (NINJAL) and by a grant-in-aid from the Japan Society for the Promotion of Science (no. 25884094).

6. REFERENCES