Successive Addition Boundary Tone in Chinese Disgust Intonation

Aijun Li[1], *Qiang Fang*[1], *Yuan Jia*[1], *Jianwu Dang*[2,3]

[1]Institute of Linguistics, Chinese Academy of Social Sciences, Beijing, China
[2]Tianjin University, Tianjian, China
[3]Japan Advanced Institute of Science and Technology, Japan
liaj@cass.org.cn; fangqiang@cass.org.cn; jdang@jaist.ac.jp

Abstract

Based on two Mandarin speakers' data, we found a kind of successive addition boundary tone in disgust intonation when the tones of the final boundary syllables are tone 1, tone 2 or tone 3. In this case, the disgust boundary tone is composed by two components, i.e., the lexical tone which is employed to express lexical meaning, and the expressive tone which is realized as a falling tone to express disgust emotion. Further, the acoustic features of the boundary tones in disgust emotion were analyzed, and the study proposed that tone1 has a longer successive addition tone which almost equals to the length of its lexical tone. In regard with tone 2 and tone 3, the successive tones are relatively shorter which account for 36% and 58% of the proceeding lexical tones. The mean slopes of the additive falling tones range from -53st/s to -33st/s.

Index Terms: Boundary tone, Successive addition boundary tone, Disgust emotion, Intonation, Mandarin Chinese

1. Introduction

Chao distinguished at least two types of additive patterns of tone and intonation: *simultaneous addition* and *successive addition* [1, 2]. In previous study, based on the analysis of the monosyllabic intonation of seven emotions, we observed that the successive addition tones in some emotional expressions as suggested by Chao [3], such as 'Disgust and Angry' have a kind of falling successive addition tones and 'Happy and Surprise' have a rising successive boundary tones. We suggested that boundary tones of 'Disgust and Angry' can be described as 'L-f%' and 'H-f%', 'Happy and Surprise' as 'H-r%' respectively.

Figure 1 displays the monosyllabic intonations of two speakers in 'Neutral and Disgust' emotions. In comparison with the 'Neutral' F_0 pattern, the 'Disgust' F_0 curves exhibits a falling successive addition tone. This addition parts can be obviously observed when the tone of the final syllable has a final H tone, i.e., tone1(HH), tone2(LH) and tone3 (LLH, where 'H' is a floating tone here [4,5]). This additive part does not belong to the lexical tone of the final syllable whereas it is produced to express the speakers' emotions. In the present study, we differentiate these two parts of the boundary tone into two components: the lexical tone and expressive tone.

We focused on 'Disgust' expression by examining the intonation patterns of disyllabic components with four boundary tones and those of longer utterances as well. Further, the acoustic features of the boundary tones were measured, including the durations of the lexical tone and the additive tone, and the falling slope of the additive tone.

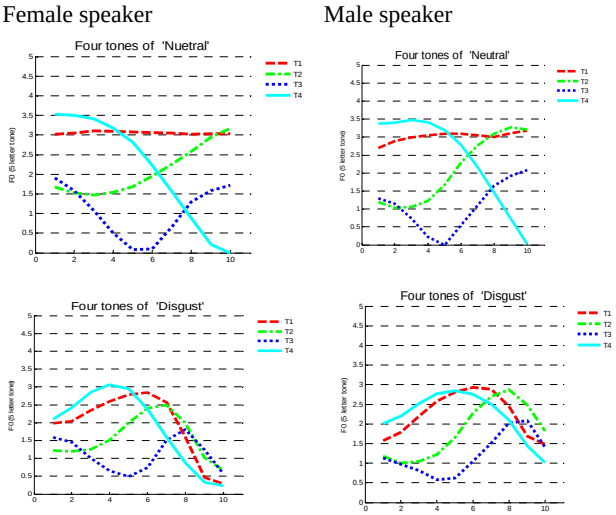


Figure 1, Intonation of monosyllabic constituents with four tones in 'Neutral' (upper) and 'Disgust' emotions (lower). The left column is F_0 contour of the female speaker while the right column is the male speaker. Tonal duration of each syllable is normalized through the selection of ten points for vowel part (tonal part).

2. Materials

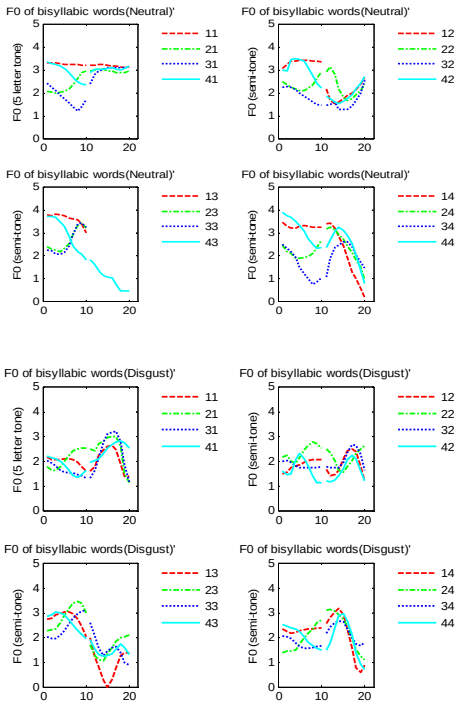
The data employed in this study were selected from the emotional speech corpus EmotionCASS [3]. A set of one hundred and eleven sentences with various sentence length (from one to fourteen syllables), different types (narrative or interrogative), and syntactic structures were recorded. The monosyllabic utterances cover the combinations of four tones and all the vowels. The disyllabic utterances contain sixteen tonal combinations. A male and a female professional actors were recruited to produce the sentences in seven kinds of emotions, i.e., Disgust(D), Sad(S), Angry(A), Happy(H), Surprise(SU), Fear(F) and Neutral(N). The sampling rate and resolution are 16 KHz and 16 bits.

All the F_0 and duration data were extracted by Praat

(http://www.fon.hum.uva.nl/praat/) and manually modified to ensure the accuracy of data. In the present paper, 'Neutral and Disgust' utterances are mainly concerned.

3. 'Disgust' intonation of longer utterances

Female speaker:



Male speaker:

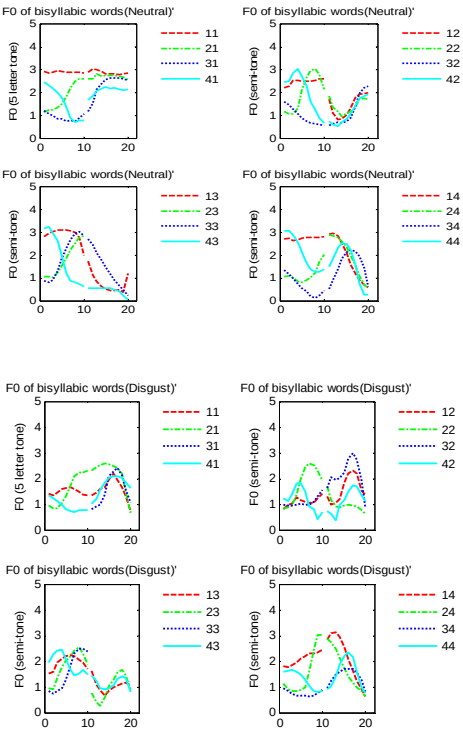


Figure 2, Intonation of disyllabic sentences in 'Neutral and Disgust' emotions with boundary tones in four tonal categories. The upper panel is the contour of female speaker and lower panel is for the male speaker.

Figure 2 shows the intonations of the disyllabic constituents in both 'Neutral and Disgust' emotions with 4*4=16 tonal combinations from two speakers. Tonal duration of each syllable is normalized into ten points (the same processing in plotting F_0 curves in Figures 3~5). From figure 2, it can be observed that (i) for the 'Neutral' intonation, the F_0 of boundary tones keep identical with their citation forms (as shown in figure , while for the 'Disgust' intonations, for those syllables tone1, tone2 and tone3 as the boundary tone, they show an additive final tailing. And, they show similar ways in the manifestation of the monosyllabic utterances, specifically, these additive tones are added after their lexical tone; (ii) examination of the intonations with tonal combinations of HH+HH / HH+HL /HL+HH / HL+HL (tone 1[1]+ tone 1 / tone1+tone4 / tone 4+tone1 / tone4+tone4), we found that the 'Neutral' intonations have an obvious tendency of declination, while the corresponding 'Disgust' intonations exhibit an opposite patterns. It is caused by the lowering and compressing of the F_0 on the first syllable which indicates that the final syllables could be more prominent than the first syllables in the expression of 'Disgust' emotion; (iii) 'L-f%' and 'L%' are employed to describe the 'Disgust and Neutral' boundary tones as for monosyllabic constituents; (iv) The final addition of the falling boundary tone accounts for a considerable length of the whole boundary tone.

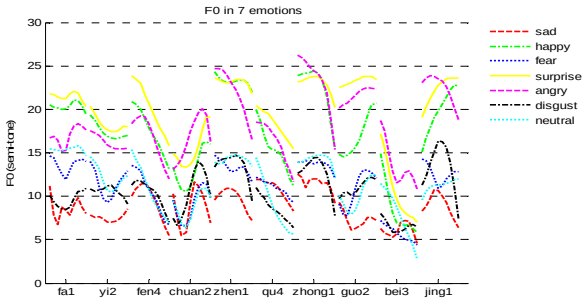


Figure 3. Intonation pattern of the sentence /fa1 yi2 feng4 chuang2 zhen1 qu4 zhong1 guo2 bei3 jing1/(Send a fax to Beijing, China.) with a high level boundary tone in seven emotions, the black dotted line is for the 'Disgust' emotion (Male speaker), light blue line for 'Neutral' emotion.

For longer utterances with final tone1, tone 2, and tone 3, the similar boundary patterns can be observed as in monosyllabic and disyllabic intonations. As depicted in figure 3, the sentence has ten syllables with a high level boundary tone (jing1). We may clearly see that an additive falling tone occurs after the final syllable /jing1/ in the 'Disgust' intonation. In comparison with the 'Neutral' boundary tone, the F_0 value of the 'Disgust' boundary tone is higher, which indicates a more prominent 'Disgust' boundary.

[1] In Standard Chinese, four tones is represented into tonal target H or L, specifically, it is HH for tone1, LH for tone2, LL for tone3 and HL for tone4.

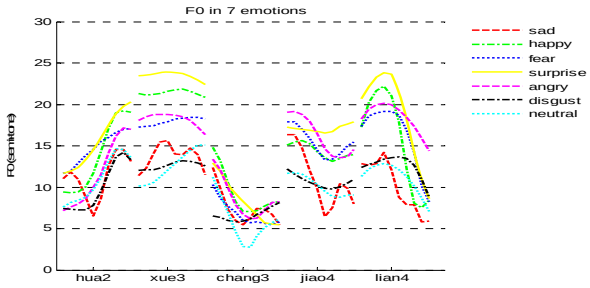


Figure 4. Intonation pattern of sentence /hua2 xue3 chang3 jiao4 lian4/(a skiing coach in ski field.)with a boundary tone 4, dark line is used to express 'Disgust' and light blue is for 'Neutral', from male speaker.

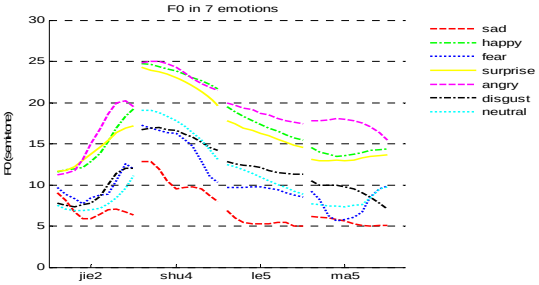


Figure 5. Intonation pattern of sentence /jie2 shu4 le ma?/ (Has it finished?) with an unstressed boundary tone (neutral tone), dark line for 'Disgust' and light blue line for 'Neutral', from male speaker.

For those boundary tones with tone4(HL) and neutral/unstressed tones, as shown in Figure 4 and Figure 5, the 'Disgust' falling boundary tone keep falling as its neutral counterparts, however, it also shows a slightly 'peak delay', even in the interrogative sentence in Figure 5, the 'neutral tone' counterparts has a rising tone, while the 'Disgust' intonation still has a falling boundary tone.

4. Acoustic analyses on the 'Disgust' boundary tone

As mentioned in the previous sections, the 'Disgust' boundary tones in T1~T3 (short for Tone1~Tone3) comprise two components, one contributes to the lexical tone of the boundary syllable and the other is the additive falling tone contributing to expression of the 'Disgust' emotion. In this section, we will make an acoustic measurement on the two components for all the utterances with the final boundary tones in tone 1, tone 2 and tone 3 (leaving tone 4 and neutral tones for future discussion). For male speaker, the amount of utterances measured with boundary T1~T3 is 21, 19, and 17 respectively. In regard with female speaker, the amount is 22, 17, and 19.

Figure 6 depicts a schematic representation of boundary tone 2 with a successive falling addition tone; the rising part represents the lexical tone while the successive falling part represents the 'Disgust' component. The acoustic features of the male and the female speakers' 'Disgust' boundary tones were analyzed according to this structure.

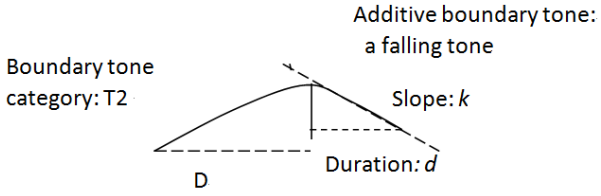


Fig.6 Schematic representation of an additive falling tone with duration'*d*' and slope '*k*'.

The features that we are concerned here are the duration of the additive falling tone 'd' (in second), the duration of the lexical part 'D'(in second), the duration ratio of the additive falling tone of the preceding lexical tone 'd/D', and the additive falling tone slope 'k' (-SemiTone/s.). 'D and d' were delimited by the peak F_0 of the boundary tone.

Figure 7 depicts the average duration features of 'D, d and D+d' for the two speakers. The female speaker has longer 'D' and 'd' than the male speaker in three conditions. Tone 3 has longer duration than the other two tones.

For male speaker, ANOVA analysis is conducted that 'D' and 'd' have significant difference across three tones (p=0.0 and 0.003 respectively), while the total length of Tone1~Tone3 have no significant difference (p=0.54). However, all these three parameters have significant difference for female speaker (p=0.034, 0.00, 0.005, respectively).

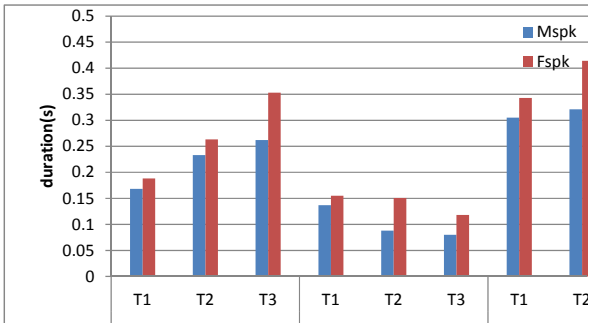


Figure 7. Mean durations: 'D, d and D+d' of the final boundary tone for three tonal categories of two speakers.

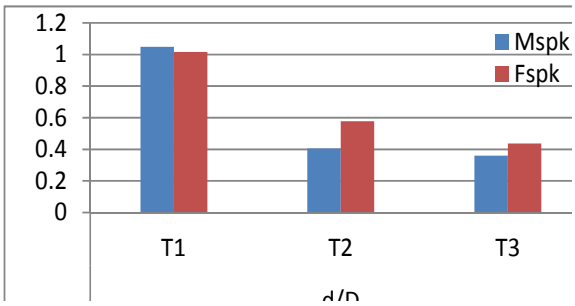


Figure 8. Mean duration ratios: 'd/D' of the final boundary tone for three tonal categories of two speakers.

Figure 8 depicts the Mean duration ratios of two speakers, the lexical component and the expressive component for boundary T1 are almost in equal length, while the emotional components shows shorter duration in T2 and T3, for male speaker: d/D≈

40% and 36% for T2 and T3 respectively, and for female speaker: d/D≈58% and 44% for T2 and T3 respectively. For both speakers, d/D of T2 and T3 are clustered in one group while T1 in another group.(alpha = 0.05).

For male speaker, the minimum ratio ('d/D') is 0.15 and maximum ratio is 3.6, the 95% Confidence Interval is from 0.25 to 1.38. For female speaker, the minimum ratio ('d/D') is 0.15 and maximum ratio is 3.59, the 95% Confidence Interval is from 0.20 to 1.33. Therefore, both speakers have consistent change space for 'd/D'.

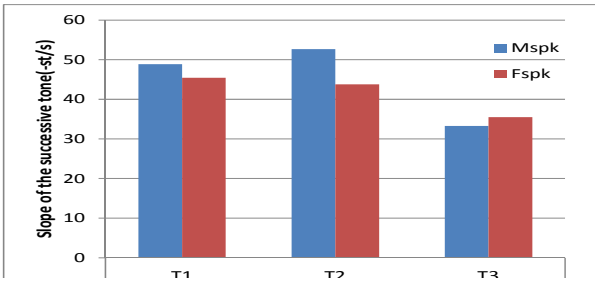


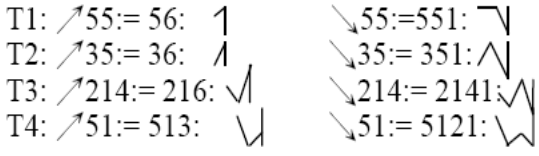
Figure 9. Average additive tone slopes (-st/s): 'k' of the final boundary tone for three tonal categories of two speakers, reference F_0=75Hz.

The average slopes of the additive falling tone are plotted in figure 9. The male speaker has slight steeper falling slopes than the female speaker in Tone 1 and Tone 2. However, Tone 1 and Tone 2 have steeper final falling than T3 for both speakers. While T1 and T2 are grouped in one cluster (average from -53~-43st/s) and T3 (average from -35~-33st/s) in another for both speakers (alpha = 0.05).

For male speaker, the minimum slope (k) is -110.83 st/s and maximum ratio is -5.77st/s, the 95% Confidence Interval is from -60st/s to -24st/s. For female speaker, the minimum slope (k) is -100.30st/s and maximum slope is -0.68st/s, the 95% Confidence Interval is from -50st/s to -28st/s.

5. Summary

Chao[3] gave some descriptions on successive addition tones in a 5 tone letter scale as followings: left column is for rising addition and right for falling addition.



This paper focused on the production aspect of successive falling addition boundary tone in 'Disgust' intonation. The results show that T1~T3 have the same falling successive addition patterns as suggested by Chao, but for our data, T4 and 'neutral tone' might have some variations, there is no successive addition falling after T4 and Neutral tone in 'Disgust' boundary tone. This suggests that the successive addition tone can have variations in expressing multifarious emotions or attitudes in additional to the patterns given by Chao.

This study discriminates the successive addition boundary tone in two components: lexical tone and expressive tone. The acoustic features of successive addition boundary tone were measured for a male and a female speaker. We found some slight differences between them. Specifically, the male speaker keeps the boundary tone in a constant length while the female speaker has longer boundary tone especially for T3. The additive falling tone of the male speaker is steeping than that of the female speaker. But they all have a longer expressive tone component for T1, that is, the two components are almost in equal length in T1, but have a shorter expressive tone component for T2 and T3. Moreover, T1 and T2 have a steeping slope than T3 for both speakers. The mean slopes of the additive falling tones range from -53st/s to -33st/s, the maximum and minimum slope shows a broad varying range from -100st/s to -0.69st/s.

Additionally, the 'Disgust' boundary tone sounds more prominent than the 'Neutral' boundary for both speakers. All these results indicate that speakers have similar patterns in general but have some personal characteristics as well. The statistic data have been used in our perceptual experiment to examine the interaction between the form and function of the successive boundary tone, and to explore how the emotional information is coded in parallel express pragmatic information in Chinese [6-10].

6. Acknowledgements

This research was funded by JSPS Ronpaku Program and NSFC Project with No. 60975081 and CASS innovation project.

7. References

[1] Chao, Y.R. 1933. Tone and intonation in Chinese. *Bulletin of the Institute of History and Philology* 4, 121- 134

[2] Chao Y. R, "A Grammar of Spoken Chinese", Berkeley, CA: University Of California Press, 1968.

[3] Li,A., Fang,Q. and Dang J., "Emotional Intonation in a Tone Language: Experimental Evidence From Chinese", ICPhS XVII, Hong Kong, 17-21, August 2011.

[4] Gussenhoven, C. 2004. The Phonology of Tone and Intonation. Cambridge: CUP.

[5] Wang, J. and Smith N., Studies in Chinese phonology, Walter de Gruyter.

[6] Xu, Y., "Speech Melody as Articulatorily Implemented Communicative Functions", Speech Communication, 46: 220-251, 2005.

[7] Ladd D. R., "Intonational Phonology", Cambridge, UK: Cambridge University Press, 1996.

[8] Liu, F. and Xu Y., "Parallel Encoding of Focus and Interrogative Meaning in Mandarin Chinese", Phonetica, 62, 70-87, 2005.

[9] Mueller-Liu, P., "Signaling Affect in Mandarin Chinese - the Role of Utterance-final Non-lexical Edge Tones", Proc. 5th of Speech Prosody, PS6-3-0048, 2006.

[10] Xu, Y., Kelly, A. and Smillie, C. Emotional expressions as communicative signals. Under review for S. Hancil and D. Hirst (eds.) *Prosody and Iconicity*, 2011.

[This paper was published on tal 2012]