THE PERCEPTUAL CUES OF TONES IN STANDARD CHINESE

HUA-CAN LIN

Institute of Linguistics
Chinese Academy of Social Sciences, Beijing, China

ABSTRACT

The synthesized speech of /shi/ /tuo/ and /ai/ were utilized to investigate the perceptual cues for tones.

The result of this experiment indicated that the four tones can be generated alone by F0 pattern with the possibility of about 95%, whereas the four tones can not be distinguished by amplitude contours alone. It also showed that the effect of duration on the naturalness of tone-3 and tone-4 is greater than that on the rate of identification of tone-1 and tone-2.

INTRODUCTION

In 1924, Liu Pu discovered the important role of F0 in Chinese tone (1). It was found that the F0 curve in syllable not only has a "tone-section", but also generally has a "onset-curving section" and "end-falling section" (2). Chuang et al. made the F0 analysis and identification test for colloquial Standard Chinese (3).

Howie demonstrated the primacy of F0 in the identification of the four tones (4). Wang talked about the role of F0 and amplitude in the four tones (5). Lin and Wang discovered that the judgement of tone category of the first syllable in bisyllabic word is often influenced by pitch of the second syllable and duration of the first one (6).

This experiment tried to investigate the role of F0, amplitude and duration in the four tones by varying these parameters in the synthesized speech.

THE PHYSICAL MANIFESTATION OF TONES

We made an acoustical analysis of 198 monosyllables consisted of 90 different Initial and Final Combinations with tones spoken by two speakers (m and f) in accordance with condition 1. Table 1 shows that the tone generally has its own peculiar F0 pattern.

Although the durations of the four tones did not show a regular relative relation, comparatively speaking, the duration of tone-3 was in most of the case the longest.

Four different types of amplitude contours could roughly be drawn from the amplitude curves in 276 monosyllables, namely: mid-hump, back-hump, two-hump and front-hump. It can be seen that the amplitude contours in tone-3 spoken by m were all two-hump, but those spoken by f were two-hump only in 60% of the cases. The peak of intensity in tone-3 showed in most of the case the lowest.

THE PERCEPTUAL EXPERIMENT OF TONES

The syllables of /shi/ /tuo/ and /ai/ were synthesized by a synthetic system (7) under five conditions shown in the lefth column of tables given below. All the speech sounds were randomized to make it impossible for the 14 subjects to predicate under which condition the speech sound was synthesized while he or she heard it. The average rate of identification of tone by subjects (14) was displayed in percentage in the right column of each table. The figures in parentheses in the tables represented the percentage of the speech sounds in good timbre judged by the subjects.

The data of the parameters in condition one roughly corresponded to the physical manifestation of tones. A consonant /tuo/ synthesized by condition one was played in fig. 2. Table 1 showed that the rate of correct identification of tone was 91.1%, and the speech sounds in good timbre amounted to 70.7%.

In condition two, the amplitude contours were only varied, i.e., the amplitude contour of low-falling-rising of F0 varied to mid-hump from two-hump, but other parameters were the same as those in condition one. A consonant /tuo/ in this condition was displayed in fig. 3. Table 2 showed that the rate of correct identification of tone was 97.6%, the speech sounds in good timbre amounted to 67.1%.

In condition three, F0 patterns were all mid-level, and durations were the same as those in condition one, but the amplitude contours had the four different types of mid-hump, back-hump, two-hump and front-hump. A consonant /tuo/ in this condition was displayed in fig. 4. Table 3 showed that the subjects (14) identified the speech sounds as tone-1 about 100%.

No one identified them as the other tones, namely, no one identified the speech sounds with amplitude contours of two-hump or back-hump and with mid-level of F0 as tone-3 or tone-4. This result indicated that the four tones cannot be distinguished by amplitude contours alone.
While the F0 patterns and the amplitude contours in condition four and five remained the same as those in condition one, the durations in condition four and five were different from those in condition one. In condition four, the durations of four different sounds were the same as those of tone-3 in condition one. In condition five, the durations of four different sounds were regulated as the same as those of tone-3 in condition one. These two results indicated that the effect of duration on the naturalness of tone-3 and tone-4 was greater than that on the rate of identification of tone-3 and tone-4.

The rate of identification of tone-3 decreased about 22% from those in condition one.

CONCLUSION

We may conclude that the four tones can be generated by F0 pattern alone with the possibility of about 95% ; The effect of duration on the naturalness of tone-3 and tone-4 is greater than that on the rate of identification of tone-3 and tone-4 ; The four tones can not be distinguished by amplitude contour alone.

REFERENCES

(1) Liu Fu, An experimental record of Chinese tone, 1924, Changhai Qingyi bookstore.