The Roles of F0 Range/Slope and Duration in Cueing the Mandarin Rising or Falling Tones

Wei Zhang1 & Wentao Gu2,3

1Department of Linguistics, McGill University, Canada
2School of Communication Sciences and Disorders, McGill University, Canada
3School of Chinese Language and Literature, Nanjing Normal University, China

wei.zhang16@mail.mcgill.ca, wtgu@njnu.edu.cn

It has been well known that rising/falling pitch is employed to distinguish the rising (R) or falling (F) tones from the high-level (H) tone in Mandarin [1], but which F0 cue–F0 range or F0 slope–is primary or more critical to perception of these dynamic tones is still inconclusive. Since F0 range and F0 slope are closely related through the variable ‘duration’, the research question is equivalent to which F0 cue is associated with a perceptual boundary that is less dependent on duration. Also, duration itself may serve as a secondary cue due to intrinsic durational differences among isolated syllables of the four tones, specifically, the F tone is the shortest, while the R tone has a duration comparable to, or marginally longer than, the H tone [2].

To elucidate this issue, first of all, we took the H-F tonal contrast as the test case (since H-R and H-F are basically symmetric in F0), and recruited 30 native speakers of Mandarin (15F, 15M) to conduct two-alternative forced choice (2AFC) identification tests on two types of two-dimensional H-F tonal continua, one of which, as shown in Fig. 1A, varied along F0 range and duration (‘F0 range continuum’), while the other, as shown in Fig. 1B, varied along F0 slope and duration (‘F0 slope continuum’). The identification rates of the F tone for each continuum are shown in Fig. 2. Analysis with mixed-effects logistic models revealed a significant interaction between F0 slope and duration in the F0 range continuum, but not between F0 range and duration in the F0 range continuum. Moreover, at each duration step we calculated sharpness and position of the perceptual boundary, of which the ratios relative to the values at 100 ms, as illustrated in Fig. 4, are approximated by linear regression to indicate the rates of change with duration. Results suggest that F0 range is the primary cue as it results in a more robust (less duration-dependent) perceptual boundary than F0 slope. Meanwhile, position of the perceptual boundary in the F0 range continuum is not fully independent of but shifting towards the F tone mildly with duration, suggesting that duration (or equivalently, F0 slope) plays a secondary role in identifying the H-F tonal contrast.

There are two ways to interpret this supplementary effect. On the one hand, aside from the primary cue of F0 range, there might be a potential threshold in F0 slope to ensure an identifiable falling pitch. If so, the effect should apply almost equally to the H-F and H-R contrasts – thus the perceptual boundary of the H-R contrast will shift towards the R tone with a longer duration. On the other hand, this effect may be specific to the H-F contrast, mainly attributed to the effect of duration itself (like the role of duration in identifying vowels /u/-/i/ in English) – the F tone is inherently shorter than the H tone, resulting in a perceptual boundary closer to the F tone with a longer duration. If this is true, the effect will be missing or even reversed in the H-R contrast because the R tone has a duration comparable to, or marginally longer than, the H tone.

To further clarify which interpretation better accounts for the supplementary effect of duration in tone identification, we recruited another set of 30 native speakers of Mandarin (21F, 9M) to conduct 2AFC identification tests on an H-F and an H-R tonal continua which were both two-dimensional F0 range continua, with shared step sizes in F0 range and duration.

Figure 3 shows the identification rates of the F and R tones for the respective F0 range continua. Analysis with mixed-effects logistic models revealed significantly negative effects of duration in both continua, with a larger effect size in the H-F than in the H-R continuum. Moreover, we calculated the perceptual boundary position at each duration step, and the ratios relative to the values at 100 ms are illustrated in Fig. 5. The results of linear regression indicate that the boundary shifts mildly towards the F tone when duration is longer (with a significant correlation) in the H-F tonal continuum (consistent with the line in Fig. 4), but no significant correlation is observed in the H-R tonal continuum. This supports the second interpretation that the supplementary role of duration (or F0 slope) in tone identification is specific to the H-F tonal contrast due to the intrinsic shorter duration of the F tone, but not applicable to the H-R tonal contrast.

In summary, F0 range instead of F0 slope is the primary cue to identify the dynamic R/F tones from the level tone, while duration plays a secondary role in identifying the H-F instead of the H-R tonal contrast. This study will have broader implications for clarifying the roles of differing cues of tones in other languages.

References

Fig. 1: Identification rates of the F tone as a function of F0 range/slope at varying durations.

Fig. 2: Identification rates of the F tone as a function of F0 range/slope at varying durations.

Fig. 3: Identification rates of (A) the F tone and (B) the R tone as a function of F0 range/slope at varying durations.

Fig. 4: Variation of sharpness (in dots) and position (in triangles) of the perceptual boundary at all duration steps in the H-R tonal continua. The solid/dashed lines indicate the results of linear regression for the F0 range/slope continuum.

Fig. 5: Variation of perceptual boundary position at all duration steps in the H-F (in solid triangles) and the H-R (in hollow triangles) tonal continua. The solid lines indicate the results of linear regression.