

## Qualitative differences in Mandarin tone imitation between Mandarin and English speakers

Wei Zhang<sup>1</sup>, Meghan Clayards<sup>1,2</sup>, Morgan Sonderegger<sup>1</sup>

<sup>1</sup>*Department of Linguistics, McGill University (Canada)*, <sup>2</sup>*School of Communication Sciences & Disorders, McGill University (Canada)*

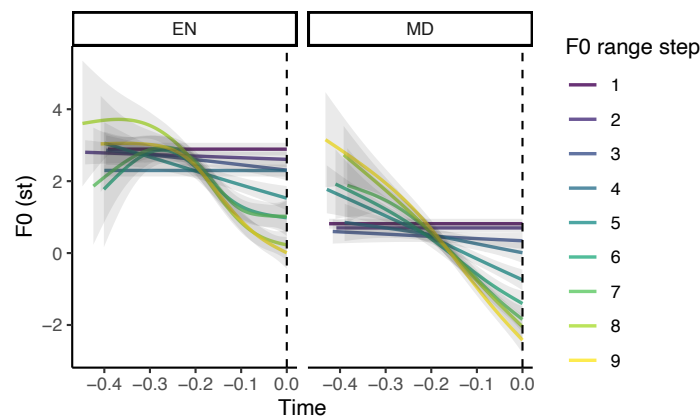
**Background.** Phonetic imitation offers a window on how L1 phonology mediates between perception and production. For instance, in the imitation of linearly-varied sounds between /b/ and /p/ along a VOT continuum, imitator-produced VOTs fell into clusters, with the centers aligning with the prototypical /b/ or /p/ in the imitators' native phonology [1]. Such nonlinearity in imitation related to phonological contrasts has also been found in the imitation of F0 contours between Mandarin flat and falling tones [2]. While such phonological mediation has been observed in native imitations, less is known about potential mediation in non-native speakers. English, lacking lexical tones, mainly use F0 contours for prosodic purposes. Thus, it may show a linear pattern when imitating F0 contours, or a non-linear pattern, but one differing from that of native speakers. This study investigates the linearity of English speakers' imitation of the Mandarin flat-falling (i.e., T1-T4) tonal continuum, comparing it to Mandarin speakers' imitation.

**Method.** We re-synthesized the F0 contour of a naturally produced syllable to create a tonal continuum containing 9 steps: the falling range increased from 0 Hz to 80 Hz, with a step size of 10 Hz (semitone scale: 0-6.1 st, step size = 0.7 st). Following [2], the syllable 'ba' was used to carry the tonal continuum. A pilot categorization study confirmed that Mandarin speakers labeled the endpoints as T1 and T4. Seventeen Mandarin speakers and 17 English speakers naïve to Mandarin were recruited remotely (EN: mean age 27, 9M 8F; MD: mean age 26, 9M 8F). Participants first read aloud a short passage in their native language which was used to normalize F0. Participants were then instructed to imitate the stimuli as closely as possible. There were 5 repetitions for each stimulus.

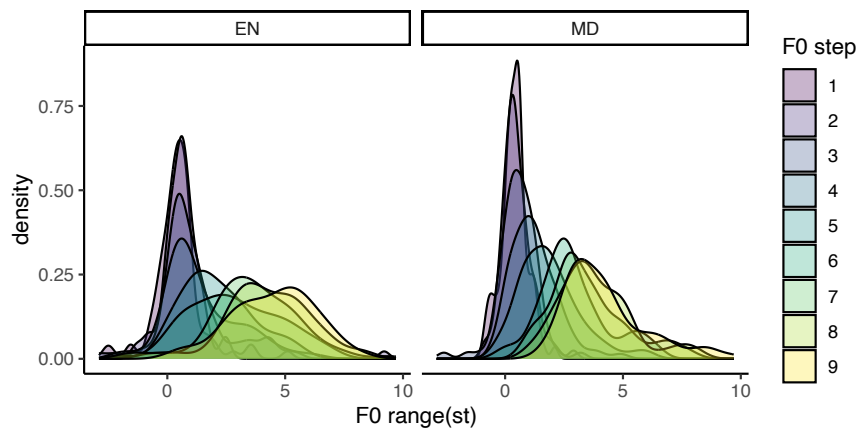
**Results.** Fig.1 shows the imitated F0 contours of each stimulus in the tonal continuum by Mandarin and English speakers. As the F0 has been normalized by speaker, the F0 height can be compared directly across groups. Compared to native Mandarin speakers, English speakers in general used a higher pitch height (relative to their mean F0 in the reading passage, i.e., 0 st in Fig. 1). Fig. 2 displays the density distribution of the imitated F0 range for each stimulus. Results of Bayesian mixture models reveal that both groups exhibit a bi-modal distribution of the F0 range: one for the flat category and the other for the falling category. Furthermore, the English group shows a wider variation for the second mode than the Mandarin group ( $\beta = -0.3$ , 95% CI=[-0.55, -0.04]). The results suggest that both languages are subjected to the mediation of a contrast, but the phonological nature of the contrast may differ between the groups. In particular, since this is a lexical contrast for the Mandarin group, the larger F0 ranges within the falling category may be interpreted as a hyper-articulated falling tone [3]. This hyper-articulation may leave other traces, such as expended vowel space and higher intensity [4]. For English speakers, however, no such predictions are assumed.

Follow up exploratory analyses aimed to look for differences in hyper-articulation in other acoustic dimensions. Fig. 3 visualizes the effect of stimulus F0 step on F1 and intensity for both groups. In Mandarin speakers' productions, it can be observed that more extreme F0 falling ranges have larger F1 and intensity values than the intermediate steps. GAM models revealed that F0 step significantly influences F1 and intensity of Mandarin speakers (F1: edf=3.47,  $p < 0.001$ ; Intensity: edf=2.19,  $p < 0.001$ ), but not English speakers (F1: edf=1,  $p = 0.06$ ; Intensity: edf=1,  $p = 0.71$ ). This result further confirms the qualitative difference in F0 range imitation between Mandarin speakers and English speakers.

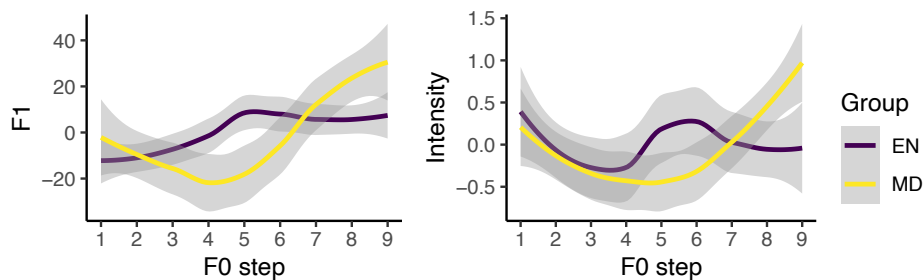
In conclusion, although English speakers do not have the phonology of lexical tones, both English and Mandarin groups imitate the F0 ranges as flat and falling categories. Their imitation behavior still differs qualitatively due to hyper-articulation by Mandarin but not English speakers, plausibly reflecting the different phonological status of these categories (tone vs. intonation) in the two languages.



**Fig. 1.** Smoothed F0 tracking of the imitated F0 contours, grouped by the steps in the continuum. The contours were aligned to the end of the target syllable, as indicated by the dashed line. Each line is smoothed using a generalized additive model fit to the empirical data for visualization purposes.



**Fig. 2.** Density distribution of the imitated F0 range by English and Mandarin speakers. Colors indicate the step indexes of the target stimuli.



**Fig. 3.** Effect of F0 step on the imitated F1 and intensity (both are centered) for the two groups of participants.

## References

- [1] Flege, J. E., & Eefting, W. (1988). Imitation of a VOT continuum by native speakers of English and Spanish: Evidence for phonetic category formation. *The Journal of the Acoustical Society of America*, 83(2), 729–740.
- [2] Zhang, W., Clayards, M., & Torreira, F. (2023). Phonological mediation effects in imitation of the Mandarin flat-falling tonal continua. *Journal of Phonetics*, 101, 101277.
- [3] Lindblom, B. (1990). Explaining phonetic variation: A sketch of the H&H theory. In *Speech production and speech modelling* (pp. 403–439). Springer.
- [4] De Jong, K. J. (1995). The supraglottal articulation of prominence in English: Linguistic stress as localized hyperarticulation. *The Journal of the Acoustical Society of America*, 97(1), 491–504.