Speech Prosody 2026 May 26-29, Philadelphia, Pennsylvania, USA

Trinity College Dublin
Coláiste na Tríonóide, Baile Átha Cliath
The University of Dublin

# Mandarin vowels: Intrinsic properties and interaction with tone

Lorna McGregor-Smith & Irena Yanushevskaya
lmcgrego@tcd.ie yanushei@tcd.ie

## Background

- The source-filter theory of speech production (Fant, 1960) proposes that the vocal fold vibration (source) and vocal tract shaping (filter) operate independently, although interaction is widely documented (Lieberman, 1963; Stevens, 1989; Garellek, 2022).
- Vowel articulation can influence phonatory behaviour and voice quality (Garellek, 2022; Ladefoged & Lindau, 1989). Close vowels are associated with higher fundamental frequency ($f_0$) and more regular phonation than open vowels (Ewan & Ohala, 1979; Whalen & Levitt, 1995; MacCallum et al., 2011).

Mandarin Vowel and Tone systems

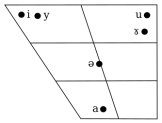



Figure1. *The six monophthongs of Mandarin* (Lee & Zee 2003).

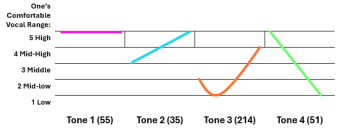



Figure 2. *Pitch contours of Mandarin Tones on Chao's* (1968) *scale.*

- Lexical tone in Mandarin is realised through $f_0$ modulation and is linked to different phonation types, particularly creaky voice in Tone 3 (Esling, 2005; Kuang, 2013; 2017; Davison, 1991). Previous study of Mandarin (Shu et. al, 2024) reported vowel-related differences in jitter, shimmer, HNR, and CPP, but were limited to Tone 1 and did not examine vowel-tone interactions.

## Discussion

- Significant vowel–tone interactions indicate dynamic interaction between the vocal tract filter and laryngeal source.
- Non-open vowels show higher $f_0$ than /a/, while /a/ exhibited higher jitter and shimmer and lower HNR.
- Close vowels are associated with more regular vocal fold oscillation and open vowels with less stable phonation.
- Significant vowel-tone interactions indicate that articulatory configuration influences how tonal targets are realised.
- Tone 1 shows the most stable phonation, and Tone 3 shows the greatest perturbation and lowest periodicity, consistent with descriptions of creaky phonation.
- Tone-related instability is weaker in close vowels but stronger in close-mid back vowels /ɤ/, /o/, suggesting interaction between articulatory configuration and tonal realisation.
- The findings support models of speech production in which source and filter are dynamically interconnected.

## Research questions

- Do Mandarin vowels of differing quality (height, backness, and lip rounding) exhibit systematically different intrinsic voice and pitch properties?
- Do different tones co-occur with characteristic patterns in perturbation or spectral/cepstral measures?
- How do vowel quality and tone interact in their effects on phonatory measures, and what does this reveal about the degree of interdependence between the voice source and the vocal tract filter?

## Methodology

- Participants: 12 native Mandarin speakers (10 F, 2 M), aged 18–30, with no reported voice disorders or smoking history. Only F data reported here.
- Six monophthongs /i/, /y/, /u/, /ɤ/, /o/, /a/ in monosyllabic words under four tones.

| Target Vowel IPA (Pinyin) | Target Tone | Word | Translation |
|---|---|---|---|
| i (i) | 1 | pī / 批 | critics |
| y (ū) | 1 | yǔ / 淤 | bruise |
| u (u) | 1 | kū / 哭 | to cry |
| ɤ (e) | 1 | kē / 科 | discipline of study |
| o (o) | 1 | bō / 波 | wave/storm |
| a (a) | 1 | bā / 巴 | to wish |

Table 1. *Partial wordlist used to elicit vowels.*

- Elicited both in citation form and in a carrier phrase:

| | |
|---|---|
| 再把_说一遍 | Simplified Chinese Characters |
| Zài bǎ _ shuō yībiàn | Pinyin |
| Say _ again | English Translation |

- Three repetitions; 1,728 tokens in total.
- Recorded in a semi-anechoic chamber using Audacity® (Audacity Team, 2024) and a Neumann TLM 103 microphone, with a sampling rate of 44.1 kHz, saved in .wav format.
- $f_0$, jitter, shimmer, HNR, CPPS, and formant frequencies (F1, F2) extracted in Praat (Boersma & Weenink, 2024).
- Linear mixed-effects models in R (R Core Team, 2025) fitted to each acoustic parameter, with vowel quality, lexical tone and their interactions as fixed effects and participant and word as random intercepts.

## Results

Vowel categories were clearly separated in F1–F2 space.

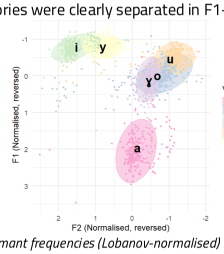



Figure 3. *Formant frequencies (Lobanov-normalised) of Mandarin Chinese vowels in this study.*

**Effects of vowel quality**

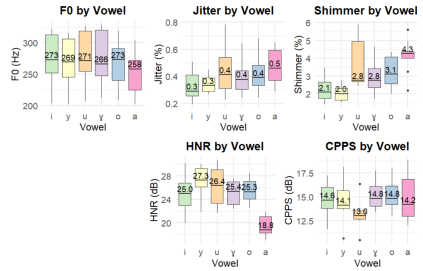
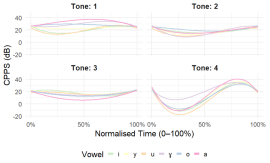



Figure 4. *Acoustic measures (median, IQR, min and max values) of female voice by Vowel (Tone 1 data only).*

- Close vowels /i/, /y/, /u/ showed higher $f_0$, higher HNR, and lower jitter and shimmer than /a/ (=more regular phonation).
- Vowel /u/ showed significantly lower CPPS than other vowels.

**Dynamic measures (qualitative analysis)**




- Tones 1, 2, and 3 CPPS are relatively stable across vowel duration.
- Tone 4 shows a rising CPPS trajectory corresponding with increasing $f_0$ across the vowel.

Figure 6. *Mean contour plots of CPPS (dB) by tone.*

**Effects of vowel quality and tone**

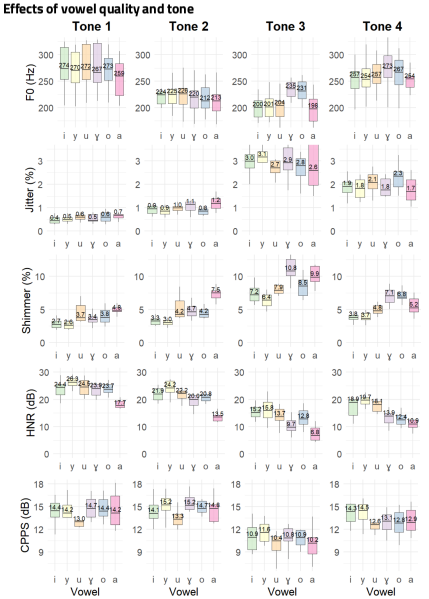



Figure 5. *Acoustic measures (median, IQR, min and max values) of female voice by vowel and tone.*

- The vowel /u/ showed significantly lower CPPS values than other vowels across tones.
- Tones 2, 3, and 4 showed lower $f_0$ and reduced periodicity compared to Tone 1. (Tone 1 = stable phonation patterns)
- Tone 3 showed the greatest phonatory instability (higher jitter/shimmer, lower HNR and CPPS).
- Vowel–tone interactions observed:
  - Close vowels showed reduced tone-related instability.
  - Close-mid back vowels /ɤ/, /o/ showed greater instability under certain tones.
- Tone-related instability was weaker for close vowels but stronger for close-mid back vowels /ɤ/, /o/.

## Conclusions

**Vowel quality and lexical tone jointly shape phonation**
Phonatory behaviour in Mandarin depends on the interaction between vowel articulation and tonal context.

**Close vowels are associated with more stable phonation**
Close vowels (/i/, /u/, /y/) showed higher $f_0$ and HNR and lower jitter and shimmer than the open vowel /a/.

**Tone 3 produces the greatest phonatory instability**
Significant vowel-tone interactions indicated that articulatory configuration influences how tonal targets are realised.