

The phonation types in Fuzhou Chinese

Changhe Chen

Department of Linguistics and Modern Languages, The Chinese University of Hong Kong, China

<https://doi.org/10.36505/ExLing-2020/11/0018/000433>

Abstract

This pilot study provides an acoustic description of the phonation types in Fuzhou Chinese. Speech samples from 5 native speakers show that tones [21 242 24] are breathy, while [44 32 4] are mainly modal and [51] is modal-breathy. Acoustic measure HNR35 can distinguish these phonation types, while H1*-A1* can only differentiate [24] and the end of [51] from other tones.

Keywords: Fuzhou Chinese, phonation type, voice quality, tone-vowel interaction

Introduction

Fuzhou Chinese is a variety of Eastern Min spoken in southeastern China. It has seven citation tones and is known for its tone-vowel interaction. For instance, [i] (only carrying citation tones [44 51 32 4]²) - [ei]/[ɛi] ([21 242 24]), [y] - [øy]/[ɛy], [e] - [a], [ə] - [ɔ], [o] - [ɔ], [u] - [ɔu], [øy] - [ɔy] and so on. In tone sandhi position, the lower or diphthongal vowels will become the higher or monophthongal ones. Donohue (2017) reports that tones [21 242 24] are creaky and explains such interaction from the perspective of phonetic enhancements. However, Esposito (2006) analyses recordings of a male speaker (aged 27 in 2004) and shows that tone [21] is produced with breathiness which can be distinguished from modal voice by measures CPP, H1-H2, H1-A1, H1-A2 and so on. In view of scanty treatises and inconsistent findings, this pilot study aims at providing an acoustic description of seven citations tones in terms of their phonation types.

Methods

Speech samples from 5 native speakers (3 females) aged above 60 were recorded³. They read a word list for five times at the normal speech rate. From that list, the vocalic parts in [ts^hə⁴⁴], [e⁵¹], [tɔy³²], [ei²¹], [tɔy²⁴²], [ɔ²⁴], [o²⁴]⁴ are analysed regarding values of HNR35 (Harmonic-to-Noise Ratio 0-3500 Hz) and H1*-A1* (the amplitude difference between the first harmonic and the harmonic nearest to the first formant) obtained by using VoiceSauce (Shue et al., 2011). The second measure is corrected for formant frequencies and bandwidths in VoiceSauce, so data from different vowels can be compared. Ten time points are evenly sampled throughout the whole duration excluding the onset and end⁵.

Results and Conclusions

HNR measures noise in signal, and both creakiness and breathiness can lead to lower values (Garellek, 2012). Thus, prior to quantitative analysis, the spectrograms and waveforms are scrutinized. Some samples are presented below.

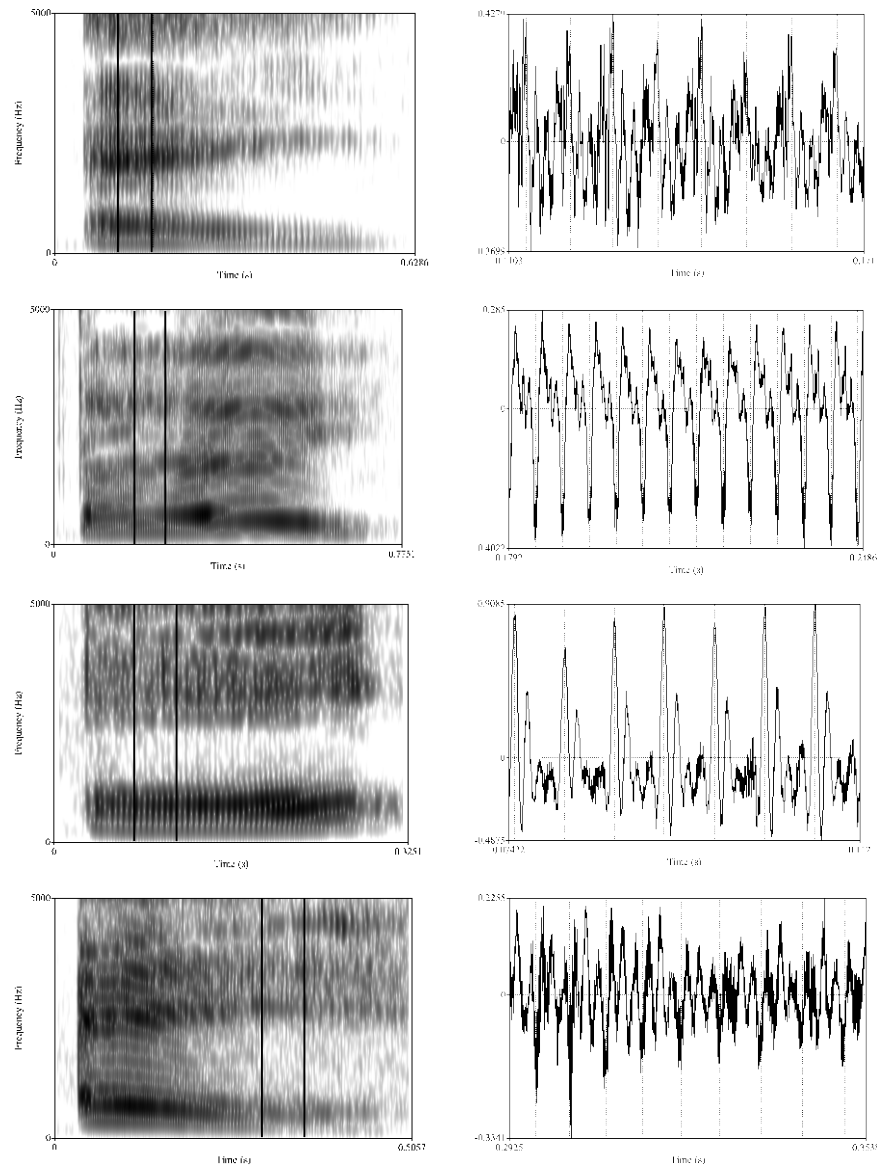


Figure 1. Samples of [ei²¹] (the first row, by Male 1), [ɔy²⁴²] (the second, by Female 3), [ɔ²²⁴] (the third, by Male 2) and [e⁵¹] (the last, by Female 2).

Tones [21 242 24] and the end of [51] are characterized by prickly waveforms and blurry formant structures, suggesting breathy voice. Most tokens of these tones show varying degrees of breathiness. Figure 2 presents the data of HNR35 and H1*-A1*.

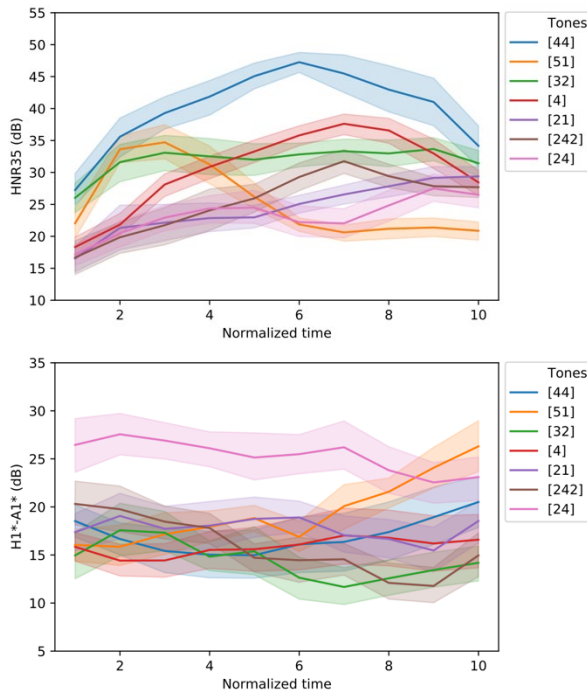


Figure 2. The averages of HNR35 and H1*-A1*, and 95% confidence intervals.

For HNR35, all trajectories start from lower positions, but only [44 51 32 4] rise sharply after the first or second point. Results of repeated measures ANOVA at the 3rd time point show a significant main effect of tones on the values of HNR35 ($F(6, 144)=43.4, p<0.001, \eta_p^2=0.644$), and post hoc tests (p adjusted by the Holm correction) reveal that the HNR35 values of [21 242 24] are significantly lower than those of other tones ($p<0.001$, except $p=0.005$ for the comparison between [24] and [4]) and no significant difference is found amongst [21 242 24] (all $p=1.000$). Tone [44] has the highest value ($p<0.001$, except $p=0.013$ for the comparison between [44] and [51]). The value of [51] is significantly higher than those of [4 21 242 24] (all $p<0.001$) but no significant difference is found for the comparison between [51] and [32] ($p=1.000$).

Tone [51] shows considerable falling after the 3rd time point and reaches its lowest point after the 6th. The same statistical analysis carried out at the 7th time point also show a significant main effect of tones on the values of HNR35 ($F(6, 144)=86.1, p<0.001, \eta_p^2=0.782$). Post hoc tests (p adjusted by the Holm

correction) reveal that [51 24] have the lowest HNR35 values ($p < 0.001$, except $p = 0.004$ for the comparison between [21] and [24]) and no significant difference is found between these two tones ($p = 0.496$).

For H1*-A1*, only [24] and the end of [51] stand out. Mean values over the ten time points are obtained for all tones except [51] (which shows large changes across time) and results of repeated measures ANOVA reveal a significant main effect of tones on the values of H1*-A1* ($F(5, 120) = 22.5$, $p < 0.001$, $\eta_p^2 = 0.484$) and post hoc tests (p adjusted by the Holm correction) find that only [24] with higher H1*-A1* values is significantly different from other tones (all $p < 0.001$). The same statistical analysis at the 9th time point also shows that there is a significant main effect of tones on the values of H1*-A1* ($F(6, 144) = 13.9$, $p < 0.001$, $\eta_p^2 = 0.367$), and [51] is significantly higher than all other tones (all $p < 0.001$ except $p = 0.040$ for the comparison between [44] and [51]) except [24] ($p = 1.000$).

Based on the features of waveforms and spectrograms as well as the data of HNR35 and H1*-A1*, it is safe to conclude that [21 24 24] are produced with breathy voice, while [51] is modal-breathy. Measure H1*-A1* can only differentiate [24] and the end of [51] from other tones.

Notes

1. Current affiliation: Department of Linguistics, The University of Hong Kong.
2. In this study, tone letters of Chao (1930) are used to represent tone height. ‘5’ is the highest and ‘1’ is the lowest.
3. Female 1, 2 and Male 1 are born in 1940s. The exact birth years of Female 3 and Male 2 were not noted down, but they are very likely in their sixties and seventies respectively. All recordings were made in 2018.
4. Male 1 produced [tøy³²] as [tui³²] (with modal voice). One token of [tøy²⁴²] by Female 2 is replaced by [tøy²⁴²] due to too strong breathiness (aperiodic signal).
5. The data of F0, F1 and F2 are checked against the spectrograms. In a few cases of incorrect tracking (48 out of 1750, about 2.74%), the nearby points closest to the original points are used.

References

- Chao, Y. R. 1930. A system of tone letters. *Le maître phonétique* 45, 24-27.
- Donohue, C. 2017. Tones and vowels in Fuzhou revisited. In *Segmental Structure and Tone*, 99-108. De Gruyter.
- Esposito, C. M. 2006. The effects of linguistic experience on the perception of phonation. Doctoral dissertation, University of California, Los Angeles.
- Garellek, M. 2012. The timing and sequencing of coarticulated non-modal phonation in English and White Hmong. *Journal of Phonetics* 40(1), 152-161.
- Shue, Y.-L., Keating, P., Vicenik, C., & Yu, K. 2011. VoiceSauce: A program for voice analysis. In *Proceedings of the 17th International Congress of Phonetic Sciences*, 1846-1849.