Evidential value of long-term laryngeal voice quality acoustics

香 港 大 學
THE UNIVERSITY OF HONG KONG

THE HONG KONG POLYTECHNIC UNIVERSITY
香港理工大學

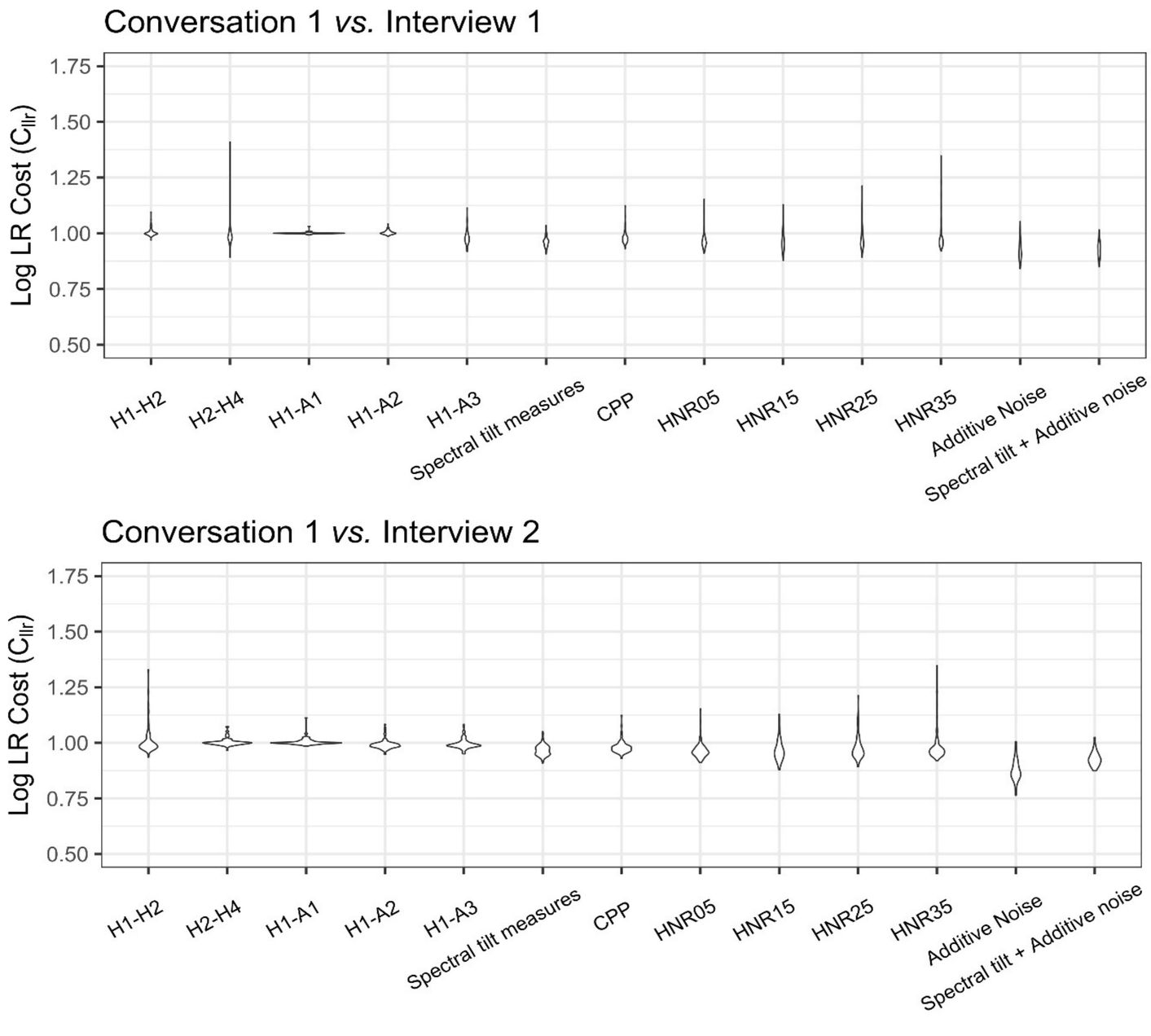
Ricky Chan[1]; Bruce Wang[2]

[1]Speech, Language and Cognition Laboratory, School of English University of Hong Kong

[2]Department of Chinese and Bilingual Studies, Hong Kong Polytechnic University

Results

- VQ parameters generally performed poorly
  - Mostly: C_{llr}~1 & EER > 40%
- Improved when combining the five spectral tilt or the five additive noise parameters
- BUT: using both spectral tilt & additive noise parameters → worse performance → conflicting information for distinguishing speakers?
- Non-contemporaneous recordings → higher C_{llr}
- Stable system performance (low *SD* in C_{llr} & EER values)



Background

- Purported usefulness of voice quality (VQ) analysis in FVC[1]
  - BUT: limited empirical validation
- Few FVC studies compared
1. match vs. mismatch in speech style
2. contemporaneous vs. non-contemporaneous recordings
- This study: laryngeal VQ + non-contemporaneous recordings; Bayesian likelihood ratio

Methods

- 75 male speakers aged 18-45
- Australian English
  - Sydney/New South Wales[3]
- Tasks: casual telephone conversation (CNV) & mock police interview (INT)
- Two recordings per speaker
  - ~ 2-week interval
- 33s of vocalic material per recording
- Spectral tilt and additive noise acoustic parameters
- MVKD + logistic regression
- 25 speakers: training, test, reference
- 100 replications[4]

Discussion

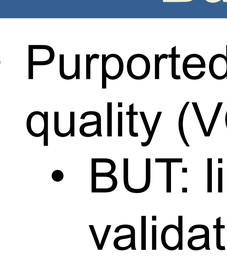
- Speech style mismatch + non-contemporaneous recordings: typical in FVC casework

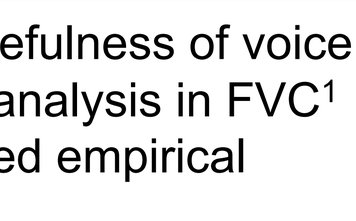
→ VQ parameters: limited evidential value
  - At best be combined with other features (e.g. MFCC)[2]

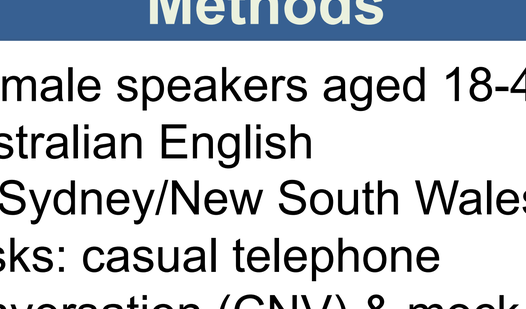
- VQ: not phonologically contrastive; no socially conditioned variation (males)

→Little room for between-speaker variation?

- Methodological issues:
  - MVKD—suitable?
  - Large age range
  - Not enough speech material for VQ analysis?

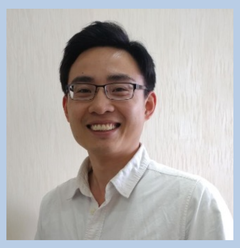






Ricky Chan
rickykwc@hku.hk

Bruce Wang
bruce.wang@alumni.york.ac.uk



References

1. Gold, E., & French, P. (2011). International practices in forensic speaker comparison. *International Journal of Speech Language and the Law*, *18*(2).
2. Hughes, V., Cardoso, A., Foulkes, P., French, J. P., Harrison, P. & Gully, A. (2019). Forensic voice comparison using long-term acoustic measures of voice quality. *Proceedings of the 19th International Congress of Phonetic Sciences (ICPhS)*. Melbourne, Australia.
3. Morrison, G. S., Zhang, C., Enzinger, E., Ochoa, F., Bleach, D., Johnson, M., Folkes, B. K., De Souza, S., Cummins, N., & Chow, D. (2015). Forensic database of voice recordings of 500+ Australian English speakers.
4. Wang, B. X., Hughes, V., & Foulkes, P. (2019). The effect of speaker sampling in likelihood ratio based forensic voice comparison. *International Journal of Speech Language and the Law*, *26*(1), 97–120