The role of voice quality in voice memorability: an experiment

Alice Paver[1] *and Kirsty McDougall*[1]
[1]*Phonetics Laboratory, University of Cambridge, UK*
{aegp2|kem37}@cam.ac.uk

Research has shown that perceptually distinctive voices are easier for lay-listeners to recall in voice memory experiments (Babel et al., 2020; Foulkes & Barron, 2000; Mullennix et al., 2011; Skuk & Schweinberger, 2013; Stevenage et al., 2018). However, the relationship between distinctiveness and recognisability is not straightforward, and the phonetic characteristics underlying these effects remain poorly understood. Much of the existing work on voice memorability has focused on overall impressions of distinctiveness rather than the contribution of specific phonetic variables, yet other research shows that phonetic features such as f0, long-term formant measurements and voice quality (VQ) contribute to listener perceptions of voice similarity and distinctiveness (Paver & McDougall, in press).

Some studies have attempted to control for phonetic typicality; Sørensen (2012) observed that speakers with extreme f0 were more accurately identified in voice parades than speakers with middling f0. Conversely, Paver et al. (submitted) demonstrated that this was not the case, and illustrated that VQ may play a part in determining which voices are more memorable. VQ is of particular interest as it is highly speaker-specific and thus plays an important role in distinguishing between speakers. In Paver and McDougall (2024), it was shown that some VQ settings, e.g. lowered larynx, fronted tongue body, yield judgements of greater distinctiveness among pairs of speakers matched for that particular VQ. The hypothesis being tested in the present study is that speakers with such a VQ setting will also be more memorable within a line-up of matched-VQ speakers. The present study will therefore use stimuli controlled for VQ to explore whether VQ indeed affects listeners' ability to encode and recognise unfamiliar voices.

An online listening experiment is being conducted to explore listener identification accuracy for speech stimuli with different VQ. Stimuli have been constructed from the PASR dataset (Hughes et al., 2022). Four male British English speakers each produced six different VQ settings, creating 24 distinct voice recordings of the Rainbow Passage. For each recording, one target-exposure sample (~20s) and two short parade samples (each ~3s) has been created; these samples have been normalised for mean f0 (107Hz) and intensity (70dB). 240 listeners will be recruited online via Prolific and randomly assigned to one of 24 voice parades (one parade per target voice stimulus). After hearing a target-exposure stimulus, listeners will complete a distractor task before hearing a voice line-up consisting of eight stimuli matched to the exposure voice for VQ (two samples per speaker). For each sample in the line-up, listeners will indicate whether the voice belonged to the speaker they heard previously, creating both target-present and target-absent identification decisions. This design enables exploration of voice memory accuracy for each of the speaker × VQ combinations, allowing for an understanding of the impact of VQ on voice memorability across different speakers and ensuring that memorability effects are not confounded with individual speaker effects, as they may have been in previous studies. This presentation will discuss the implications of the results for the understanding of voice memory, particularly as it relates to earwitness behaviour.

References

Babel, M., McGuire, G., & Willis, C. (2020). The role of voice evaluation in voice recall. In B. Weiss, J. Trouvain, M. Barkat-Defradas, & J. J. Ohala (Eds), *Voice Attractiveness: Studies of Sexy, Likable and Charismatic Speakers* (pp. 101–124). Springer Singapore Pte. Limited. https://doi.org/10.1007/978981-15-6627-1_6

Foulkes, P., & Barron, A. (2000). Telephone speaker recognition amongst members of a close social network. *International Journal of Speech, Language and the Law*, *7*(2), 180–198. https://doi.org/10.1558/sll.2000.7.2.180

Hughes, V., Foulkes, P., Harrison, P., Wormald, J., Xu, C., van der Vloed, D., & Kelly, F. (2022). *Personspecific ASR* [Data set]. http://pasr.york.ac.uk/

Mullennix, J. W., Ross, A., Smith, C., Kuykendall, K., Conard, J., & Barb, S. M. (2011). Typicality effects on memory for voice: Implications for earwitness testimony. *Applied Cognitive Psychology*, *25*(1), 29–34. https://doi.org/10.1002/acp.1635

Paver, A., & McDougall, K. (2024). *The role of voice quality in perceived voice distinctiveness: An experiment*. International Association for Forensic Phonetics and Acoustics, Montréal, Canada.

Paver, A., & McDougall, K. (in press). Perceived voice similarity. In F. Nolan, K. McDougall, & T. Hudson (Eds), *The Oxford Handbook of Forensic Phonetics*. Oxford University Press.

Paver, A., McDougall, K., & Nolan, F. (submitted). *Voice memory, acoustic-phonetic typicality and perceptual distinctiveness: The role of f0 distribution in a population*.

Skuk, V. G., & Schweinberger, S. R. (2013). Gender differences in familiar voice identification. *Hearing Research*, *296*, 131–140. https://doi.org/10.1016/j.heares.2012.11.004

Sørensen, M. H. (2012). Voice line-ups: Speakers' F0 values influence the reliability of voice recognitions. *International Journal of Speech Language and the Law*, *19*(2), 145–158. https://doi.org/10.1558/ijsll.v19i2.145

Stevenage, S. V., Neil, G. J., Parsons, B., & Humphreys, A. (2018). A sound effect: Exploration of the distinctiveness advantage in voice recognition. *Applied Cognitive Psychology*, *32*(5), 526–536. https://doi.org/10.1002/acp.3424