Voice parade parameters:

Investigating the effect of parade size and voice sample duration on earwitness identification accuracy

Alice Paver¹, Harriet M. J. Smith², Nikolas Pautz², Kirsty McDougall¹, Katrin Mueller-Johnson³, Francis Nolan¹¹Theoretical & Applied Linguistics Section, University of Cambridge; ²Department of Psychology, Nottingham Trent University; ³Centre for Criminology, University of Oxford

1. Background

Voice parades enable collection of evidence by asking an earwitness to identify a perpetrator if present in a voice lineup.

Current procedure (Home Office guidelines, 2003¹):

- 9 speakers (8 foils + suspect)
 - Foils with same accent
 - 60s collage of speech of innocuous content
- Issue: process is challenging, time consuming and expensive.

Research question: can process be safely streamlined

- by reducing speech sample length
- by reducing parade size

... without compromising accuracy?

2. Methodology

Participants heard 1 of 6 different target speakers:

- 3 x SSBE (DyViS²)
- 3 x Yorkshire (1 YorViS³, 2 WYRED⁴)

They were exposed to the voice for 60s & later, after a 5 min filler task, tried to identify the speaker in a lineup of foils from the same corpus.

Experimental design

Target-present & target-absent lineup

- Mimic situation of either guilty or innocent suspect.

15s, 30s or 60s speech sample length

- Testing effect of sample length on accuracy

9 person (exp. 1) and 6 person (exp. 2) lineup size

- Testing effect of parade size on accuracy

3. Results

In both experiments (9 & 6 person), there was no statistically significant difference between accuracy levels for the different sample durations (Fig. 1).

Fig 1. Voice identification accuracy rates for 15s, 30s and 60s sample lengths, and 6 and 9 person parades, both target-present and target-absent

Performance overall was low in both experiments, but better in target-present condition compared to target-absent. The main effect of target presence was statistically significant, and there was no interaction effect between parade size (6, 9) and sample length (15s, 30s, 60s).

Fig 2. Mean accuracy rates for each target speaker in the 6 and 9 person parades, both target-present and target-absent.

Performance varied noticeably by target speaker (Fig. 2), even within the same accent group. This is especially true for DyViS (highlighted on figure).

4. Discussion

- These results suggest that the current¹ voice parade procedure can be adapted.

- Since there is no meaningful difference in performance between 15s, 30s and 60s, sample length could safely be reduced.

→ This would significantly impact time and money spent constructing voice parades.

- Smaller parade sizes increase the risk of innocent suspects being chosen by chance.

→ No improvement in performance here for smaller parades to offset this risk, so we don't recommend adapting the procedure by changing parade size.

5. References

1. Home Office. (2003). Home Office circular 057/2003: Advice on the use of voice identification parades. Retrieved from <https://bit.ly/3v13jJE>
2. Nolan, F., McDougall, K., de Jong, G., & Hudson, T. (2009). The DyViS database: Style-controlled recordings of 100 homogeneous speakers for forensic research. *International Journal of Speech, Language, and the Law*, 16(1), 31-57.
3. McDougall, K., Duckworth, M., & Hudson, T. (2015). Individual and group variation in disfluency features: A cross-accent investigation. *ICPhS 2015. Proc. 18th International Congress of Phonetic Sciences*.
4. Gold, E., Ross, S., & Earnshaw, K. (2018). The 'West Yorkshire Regional English Database': Investigations into the generalizability of reference populations for forensic speaker comparison casework. *Proc. Annual Conference of the International Speech Communication Association, INTERSPEECH, September 2018*, 2748-2752.

ESRC Grant Reference: ES/S015965/1