Differences in the vocal space of deceptive and non-deceptive speech

Alessandro De Luca[1], Sarah Lim[1], Volker Dellwo[1]

[1]Institute of Computational Linguistics, University of Zurich

alessandro.deluca@uzh.ch

Abstract (max 300 words)

Deception is a common behaviour in not only humans, but also other taxa (Griffin & Ristau, 2013; Whiten & Byrne, 1988). Although deception can include a wide variety of different behaviours, here we only focus on lying by deliberately not telling the truth in human speech communication. We start from the assumption that humans have no interest in being identifiable when lying and hypothesise that a loss of acoustic congruency between utterances within the speaker may affect listeners' memory to what was being said and to the person saying it (Campeanu et al., 2013, 2015). We thus predict that voices should be less recognizable in situations in which speakers do not tell the truth. In a discrimination task with speech from the CSC Deceptive Speech database (Columbia University et al., 2013), human listeners decided between two voice samples of unfamiliar speakers whether they were produced by one or by two speakers either when speech was produced under deceptive or non-deceptive conditions. Results revealed a significant loss of discrimination performance under deceptive compared to truthful speech. To understand what the acoustic correlates of this effect were, we compared the area occupied by the two speech conditions in a latent feature space obtained by UMAP dimension reduction (McInnes et al., 2020) of a 39-dimensional MFCC, D, and D^2 feature set, extracted from utterances of the CSC Deceptive Speech database. Based on Latinus et al. (2013) and Dellwo et al. (2019) we hypothesise that the lower discrimination performance under deceptive speech may rise from speakers being less distinctive when lying by moving towards an average voice and thus occupying a smaller area in the vocal space when being deceptive. Preliminary results showed that instead deceptive speech occupied a slightly larger area than truthful speech in the latent feature space. To confirm this, we are currently working on replicating this analysis using different dimensionality reduction methods and feature sets.

References:

Campeanu, S., Craik, F. I. M., & Alain, C. (2013). Voice Congruency Facilitates Word Recognition. *PLOS ONE*, *8*(3), e58778. https://doi.org/10.1371/JOURNAL.PONE.0058778

Campeanu, S., Craik, F. I. M., & Alain, C. (2015). Speaker's voice as a memory cue. *International Journal of Psychophysiology*, *95*(2), 167–174. https://doi.org/10.1016/J.IJPSYCHO.2014.08.988

Columbia University, SRI International, & University of Colorado Boulder. (2013). *CSC Deceptive Speech*. https://doi.org/10.35111/q500-9a28

Dellwo, V., Pellegrino, E., He, L., & Kathiresan, T. (2019). The dynamics of indexical information in speech: Can recognizability be controlled by the speaker? *AUC PHILOLOGICA*, *2019*(2), 57–75. https://doi.org/10.14712/24646830.2019.18

Griffin, D. R. (Donald R., & Ristau, C. A. (2013). *Truth and Deception In Animal Communication*. 147–172. https://doi.org/10.4324/9780203761700-12

Latinus, M., McAleer, P., Bestelmeyer, P. E. G., & Belin, P. (2013). Norm-Based Coding of Voice Identity in Human Auditory Cortex. *Current Biology*, *23*(12), 1075–1080. https://doi.org/10.1016/J.CUB.2013.04.055

McInnes, L., Healy, J., & Melville, J. (2020). *UMAP: Uniform Manifold Approximation and Projection for Dimension Reduction*.

Whiten, A., & Byrne, R. W. (1988). Tactical deception in primates. *Behavioral and Brain Sciences*, *11*(2), 233–244. https://doi.org/10.1017/S0140525X00049682