

Same data, different results? An evaluation of the robustness of approaches used for establishing individual distinctiveness in mammalian acoustic cues

Kaja Wierucka^{1,2}, Judith M. Burkart^{1,2}

¹Department of Anthropology, University of Zurich

²The Swiss National Centre of Competence in Research (NCCR) Evolving Language

kaja.wierucka@uzh.ch

Abstract (max 300 words)

Complex, automated acoustic data analysis methods are becoming increasingly popular for the study of animal vocal communication. Establishing caller identity is a critical element of many behavioural studies and call classification models are used to demonstrate the ability of cues to convey individual identity information. This is often not only a basis for hypothesis testing through bioassays, but also a key factor in evaluating cognitive and communication abilities in many animal species. However, feature extraction methods as well as classifiers used vary greatly between studies making it impossible to accurately compare results across species, conduct comparative research and draw comprehensive conclusions of evolutionary significance. While previous studies have evaluated the accuracy with which models can predict caller identity, none have investigated the effect that data preparation and feature extraction have on the consistency of these results and whether certain combinations of processing and classification methods produce more reliable results. By incorporating data from multiple species (including both primates and non-primates), we evaluate the robustness and consistency of accuracy scores across different feature extraction methods (frequency measurements, mel-frequency cepstral coefficients, and highly comparative time-series analysis) and classifiers (discriminant function analysis, permutational multivariate analysis of variance, support vector machines, gaussian mixture models, neural networks and random forests). We also explore the effect of various data processing methods (balancing, outlier removal, normalisation and consequent sample size) on the robustness of obtained results. We identify trends that are generalisable across species and provide guidelines for the processing and analysis of mammalian vocalisations in relation to determining the individual distinctiveness of acoustic cues and establishing caller identity.

[Abstract Word count: 262]