Human Error Rates for Speaker Recognition

Wade Shen\textsuperscript{1}*, Joseph P. Campbell\textsuperscript{1}*, Reva Schwartz\textsuperscript{2}

swade@l.mit.edu

It is commonly assumed that speaker identification by human listeners is an innate skill under certain conditions. As such, human listening tests have served as the benchmark for automatic recognition systems. In recent evaluations comparing human and machine performance on a speaker comparison task, error rates of naïve human listeners far exceed those of machines [special session on Human Assisted Speaker Recognition, IEEE ICASSP, Prague, 2011]. In this presentation, we quantify the performance of naïve listeners in a variety of challenging channel conditions and we compare these results against automatic systems and trained human listeners. The results of these experiments impact the admissibility of both forensic voice analysis and courtroom testimony by human listeners.

\textsuperscript{1} MIT Lincoln Laboratory
\textsuperscript{2} United States Secret Service

* This work is sponsored by the Department of Defense under Air Force contract FA8721-05-C0002. Opinions, interpretations, conclusions, and recommendations are those of the authors and are not necessarily endorsed by the United States Government.