

# ASR Systems as Models of Phonetic Category Perception in Adults

**Thomas Schatz**

LSCP (ENS/EHESS/CNRS), Département d'Etudes Cognitives, Ecole Normale Supérieure, PSL Research University, France

**Francis Bach**

SIERRA Project-Team (ENS/INRIA/CNRS), Département d'Informatique, Ecole Normale Supérieure, PSL Research University, France

**Emmanuel Dupoux**

LSCP (ENS/EHESS/CNRS), Département d'Etudes Cognitives, Ecole Normale Supérieure, PSL Research University, France

**Abstract:** Adult speech perception is tuned to efficiently process native phonetic categories, causing difficulties with certain non-native categories. For example, Japanese has no equivalent of the distinction between American English /r/ and /l/ and native speakers of Japanese have a hard time discriminating between these two sounds. Here, we ask whether standard Automatic Speech Recognition (ASR) systems trained on large corpora of continuous speech can make correct quantitative predictions regarding such non-native phonetic category perception effects. By training an ASR system on language L1 and evaluating it on language L2, we obtain predictions for a native L1 speaker tested on L2 phonetic contrasts. Using a variety of L1 and L2, we show that ASR models correctly predict several well-documented effects. Beyond the immediate results, our evaluation methodology, based on a machine version of ABX discrimination tasks, opens the possibility of a more systematic investigation of computational models of phonetic category perception.