ASR Systems as Models of Phonetic Category Perception

Submission ID   3000209

Submission Type   Poster

Topic   Cognitive Science

Status   Submitted

Submitter   Thomas Schatz

Affiliation   LSCP (ENS/EHESS/INRIA/CNRS)

SUBMISSION DETAILS

Presentation Type Either Poster or Oral Presentation

Presentation Abstract Summary Recent advances in machine learning raised the prospect of developing quantitative models of human perception that can handle realistic sensory input. In this paper, we focus on the question of phonetic category perception, i.e. the way we perceive basic speech sounds (roughly consonants and vowels), which is largely determined by the language(s) to which we were exposed as a child. For example, native speakers of Japanese have a hard time discriminating between American English /r/ and /l/, a phonetic contrast that has no equivalent in Japanese. We show that typical GMM-HMM Automatic Speech Recognition (ASR) systems trained on large corpora of continuous speech correctly predict several perceptual effects observed in humans. Our work illustrates the value of considering large-scale machine learning systems in the context of modeling human perception.

Paper Upload (PDF) CCN2017-3.pdf

Co-author Information

* Presenting Author

First Name	Last Name	Affiliation	E-mail
Thomas *	Schatz *	LSCP (ENS/EHESS/INRIA/CNRS)	thomas.schatz@laposte.net
Francis	Bach	SIERRA TEAM (ENS/INRIA/CNRS)	francis.bach@ens.fr
Emmanuel	Dupoux	LSCP (ENS/EHESS/INRIA/CNRS)	emmanuel.dupoux@gmail.com

Keywords

Keywords
perception

phonetic categories
modeling
ASR

phonetic categories

modeling

ASR