

Inferring errors and intended meanings with a generative model of language production in aphasia

Thomas Clark

Massachusetts Institute of Technology, Cambridge, Massachusetts, United States

Edward Gibson

Massachusetts Institute of Technology, Cambridge, Massachusetts, United States

Roger Levy

Massachusetts Institute of Technology, Cambridge, Massachusetts, United States

Abstract

We propose a generative modeling framework of impaired language production and an inference framework that models rational comprehension of impaired language. Given a task (e.g. picture-description), we approximate the prior distribution over intended sentences using a language model trained on unimpaired speakers' utterances. We define a generative model of operations (e.g., semantic and phonological errors, retracing, filled pauses) that intervene on the intended sentence to yield an utterance. The model is implemented in the Gen probabilistic programming language, with data from AphasiaBank's 'Window' picture-description task. Given observed utterances, a particle filter estimates posterior probabilities for latent variables (e.g. the speaker's intended sentence or sequence of errors). Our framework models comprehension as inference on a generative model of production, and provides a way to quantify incremental processing difficulty for impaired language in a way that combines a language model prior with explicit reasoning about errors.