



Research Week 2020

Improving Machine Learning Models of Paraphasia Classification Using Crowdsourced Semantic Features

Katy McKinney-Bock, Ph.D., Linying Li, Gerasimos Fergadiotis, and Steven Bedrick
mckinnka@ohsu.edu
OHSU

Keywords

Natural Language Processing, Machine Learning, Aphasiology, Language Assessment

Abstract

Confrontation naming tasks such as the Philadelphia Naming Test (PNT) are essential tools for the assessment and treatment of people with aphasia. One challenge to clinical use of the PNT is the complexity of its scoring guidelines. We build on work towards automatic classification of paraphasias by (1) demonstrating that machine learning models substantially improve performance on the paraphasia classification task, (2) providing insight into the behavior of the highest-performing model, and (3) reporting results on experiments with contextual semantic and item-level information. Our best model improves 8.1 F1 points over a manually-constructed decision tree classifier. For development of the semantic features, we used crowdsourcing methods to construct a dataset of linguistic contexts from the visually presented items on the PNT assessment.