

Objective evaluation of perceptual quality in the
dysarthrias: an engineering perspective
V. Berisha, J. Liss, A. Wisler, M. Tu

Rationale and Purpose:

The trained ear of the speech-language pathologist is the gold standard assessment tool for clinical practice in motor speech disorders. However, subjective judgments are vulnerable to bias, which poses a serious, but underappreciated, threat to clinical practice. We propose a computational solution, which uses machine-learning algorithms that are trained to model expert perceptual ratings.

Methods:

We present a suite of objective outcome measures based on machine learning and statistical signal processing. These models learn the relationship between speech acoustics and subjective ratings of abnormal nasality, prosody, articulatory precision, vocal quality and severity. After calibration, the models are able to reliably predict the rating of a new speaker on a 7-point scale.

The models operate by extracting a series of speech features that capture irregular prosody, voicing, and imprecise articulation. These features are then combined in a decision engine to produce ratings of a speaker along the 5 perceptual dimensions.

The models are evaluated on an extensive set of dysarthric speech and subjective ratings; evaluation metrics include the correlation coefficient and the mean absolute error between the predicted ratings and averaged subjective ratings.

Results and Discussion:

Results reveal correlations ranging between 0.75 – 0.85 between the objective ratings and the averaged expert evaluations and mean absolute errors of ~ 0.8 (on a 7-point scale). Research from others reveals that, for most perceptual dimensions, subjective evaluations from two evaluators fall within 1 of each other a 7-point scale. This compares favorably with the reliability of the model.