ACOUSTICS2008/3280
Evaluation of a simplified phonetic annotation scheme for disordered speech

Kathleen Murray[a], David Graff[b], Yolanda Heman-Ackah[a] and Robert Sataloff[a]
[a]Drexel University College of Medicine, Department of Otolaryngology, Head and Neck Surgery, 1721 Pine Street, Philadelphia, PA 19103, USA
[b]Linguistic Data Consortium, 3600 Market St. Suite 810, Philadelphia, PA 19104-2653, USA

Attempts to train a computer to mimic a vocal pathology expert's perception of perceivable voice problems have had limited success. A recent study successfully used a Cepstrum-based calculation (CPPs) to detect dysphonic speech, but it made significant false negative and false positive errors [1]. Appropriate training data could improve calculation accuracy, but logistical and legal issues from the medical domain render large amounts of detailed, consistent training data difficult to produce. A simple scheme has been developed based on labeling one point in time for each of a few vowels in a read statement. Calculations from [1] and [2], centered on neighborhoods of these vowels will augment the CPPs classifier. The annotation scheme will be evaluated with respect to its ability to assist a classifier to automatically predict perceptual labels [1] from data, and practicalities such as inter/intra-annotator agreement.
1. Heman-Ackah, Y.D., Heuer, R.J., Michael, D.D., Ostrowski, R., Horman, M., Baroody, M., Hillenbrand, J.M., and Sataloff, R.T. (2003). "Cepstral peak prominence: A more reliable measure of dysphonia," Annals of Otology, Rhinology, and Laryngology, 112, 324-333. 2. "Perceived Level of Noise by Mark VII and Decibels (E)" S. S. Stevens, J. Acoustical Soc. Am. v. 51(2, part 2) 1972. pp. 575-602.