

Modeling dysphonic voice quality perception using psychometric theory and auditory-processing models

Rahul Shrivastav, Ph.D., CCC-SLP

Assistant Professor, Department of Communication Science and Disorders
University of Florida, Gainesville.

Abstract

Despite years of research, there is little consensus on the optimum methods to describe and quantify dysphonic voice quality. Research to establish a relationship between the vocal acoustic signal and the perception of voice quality typically involves obtaining perceptual judgments from a group of listeners and comparing these to the vocal acoustic signal. Unfortunately, such research has shown highly inconsistent results. These inconsistencies have been hypothesized to have resulted from a number of factors, including individual differences in voice quality perception and the assumption that the vocal acoustic signal is linearly related to the perceived voice quality.

First, the hypothesis that there are large individual differences in how listeners perceive voice quality is based upon findings of poor “agreement” in perceptual ratings (Kreiman, Gerratt, Kempster, Erman, & Berke, 1993; Kreiman, Gerratt, Precoda, & Berke, 1992) as well as the failure of multidimensional scaling solutions to account for the variance in perceptual judgments of voice quality (Kreiman & Gerratt, 1996, , 2000; Kreiman, Gerratt, & Berke, 1994). However, research in our laboratory suggests that individual differences in the perception of voice quality are not as large as previously reported. It is hypothesized that the findings of previous research result from the use of experimental procedures that fail to adequately capture a listener’s percept of voice quality, resulting in excessive “noise” in perceptual data (Shrivastav, Sapienza, & Nandur, In Press).

Second, it is often assumed that acoustic and perceptual measures of voice quality are linearly related. However, speech perception is often influenced by multiple acoustic cues (for example, Repp, 1988) and acoustic-perceptual relationships are known to be non-linear. The use of an auditory processing model as a signal processing front-end can account for a part of the non-linear acoustic-perceptual relationship, thereby improving the fit between acoustic and perceptual measures of voice quality (Shrivastav, 2003; Shrivastav & Sapienza, 2003).

This lecture will describe the theoretical motivations behind the use of psychophysical theory and auditory processing models in voice quality research as well as present data obtained in our laboratory.

References:

- Kreiman, J., & Gerratt, B. R. (1996). The perceptual structure of pathologic voice quality. *Journal of the Acoustical Society of America*, 100(3), 1787-1795.
- Kreiman, J., & Gerratt, B. R. (2000). Sources of listener disagreement in voice quality assessment. *Journal of the Acoustical Society of America*, 108(4), 1867-1876.
- Kreiman, J., Gerratt, B. R., & Berke, G. S. (1994). The multidimensional nature of pathologic vocal quality. *Journal of the Acoustical Society of America*, 96(3), 1291-1302.
- Kreiman, J., Gerratt, B. R., Kempster, G. B., Erman, A., & Berke, G. S. (1993). Perceptual evaluation of voice quality: review, tutorial, and a framework for future research. *Journal of Speech & Hearing Research*, 36(1), 21-40.
- Kreiman, J., Gerratt, B. R., Precoda, K., & Berke, G. S. (1992). Individual differences in voice quality perception. *Journal of Speech & Hearing Research*, 35(3), 512-520.
- Repp, B. H. (1988). Integration and Segregation in Speech Perception. *Language & Speech*, 31(3), 239-271.
- Shrivastav, R. (2003). The Use of an Auditory Model in Predicting Perceptual Ratings of Breathy Voice Quality. *Journal of Voice*, 17(4), 502-512.
- Shrivastav, R., & Sapienza, C. (2003). Objective Measures of Breathy Voice Quality Obtained Using an Auditory Model. *Journal of the Acoustical Society of America*, 114(4), 2217-2224.
- Shrivastav, R., Sapienza, C., & Nandur, V. (In Press). Application of Psychometric Theory to the Measurement of Voice Quality Using Rating Scales. *Journal of Speech, Language, & Hearing Research*.