Predicting Listener Perception of Simulated Laryngeal Vocal Tremor Using A Novel Measure of Pitch Modulation Strength

R. Lester, B. Story

The acoustical bases for the perception of vocal tremor (VT) are currently unclear. This is problematic for assessing VT and conducting objective studies of treatment outcomes for individuals with VT. Previous studies have investigated the modulation rate and extent of the fundamental frequency (F0) and amplitude, the mean amplitude in frequency bands, and the coefficient of variation of amplitude in frequency bands without adequately accounting for listeners' perceptions. The purpose of the current study was to test a novel acoustical measure of VT that could better explain listeners' perception of the severity of VT. This research was carried out using a computational model of speech production that allowed for precise control and modulation of the vocal fold and vocal tract configurations. Healthy adults participated in a perceptual study involving pair-comparisons of the magnitude of "shakiness" with simulated samples of laryngeal VT that differed by either the F0, degree of vocal fold adduction, or vocal tract shape. An autocorrelation function was computed with the simulated samples as a measure of pitch modulation strength. Initial analyses revealed that the samples had a periodic structure that repeated at the tremor rate, with varying bandwidths centered around the most prominent tremor rate. Samples with a narrower temporal bandwidth were perceived as shakier than signals with a wider temporal bandwidth; that is, samples with a well-defined and prominent tremor rate were perceived as sounding shakier. For each contrast (F0, degree of vocal fold adduction, vocal tract shape), the participants' perceptual judgments will be modeled using a binary logistic stepwise regression analysis with the potential predictor variable representing the temporal bandwidth. The findings from this study are expected to be useful in identifying an acoustical measure that can better quantify the severity of VT as it relates to listener perception.