Formant centralization ratio (FCR) as a sensitive metric of vowel undershoot: Comparison with vowel space area (VSA) in impaired versus healthy speakers

Shimon Sapir (University of Haifa), Monika Połczyńska (School of English, Adam Mickiewicz University, Poznań), Yishai Tobin (Ben Gurion University, Beer Sheva), Lorraine Ramig (National Center for Voice and Speech & University of Colorado, Boulder, Co.), Cynthia Fox (University of Colorado, Boulder, Co. & University of Arizona), Jennifer Spielman (National Center for Voice and Speech), Christopher Dromey (Brigham Young University, Provo, Utah), Shawn Nissen (Brigham Young University, Provo, Utah), & Nelson Roy (University of Utah, Salt Lake City)

Vowel space area (VSA) has been traditionally used as a metric of vowel hypoarticulation (as in dysarthric speech) and hyperarticulation (as in clear speech). The VSA is highly reliable and sensitive to vowel centralization when a single speaker is concerned (e.g., comparing clear vs. conversational speech, stressed vs. unstressed syllables, impaired vs. unimpaired articulatory movements). However, when it is used for group statistics, it is much less sensitive and reliable, due, to a large extent, to inter-speaker variability, and statistical "noise" produced by this variability. The purpose of this presentation is to introduce a new acoustic metric – formant centralization ratio, or FCR, which has been designed to reduce sensitivity to inter-speaker variability and enhance sensitivity to vowel centralization. We compared healthy speakers to speakers with the diagnosis of either Parkinson's disease (PD), traumatic brain injury (TBI) or muscle tension dysphonia (MTD). The speakers with PD and MTD and their matched controls spoke English whereas individuals with TBI and their matched controls spoke Polish. The VSA and FCR were constructed from 4 vowels (/i/, /u/, /a/, and /ae/) or from 3 vowels (/i/, /u/, and /a/). These vowels were extracted from words embedded in phrases. Hi Fi recordings and state of the art acoustic analysis programs were used to measure vowel formant frequencies. Our preliminary results indicate a greater sensitivity of the FCR to abnormal vowel articulation than the VSA, both in terms of statistical significance and effect size measures. The FCR was also more sensitive than the VSA to behavioral treatment effects, although both metrics showed sensitivity to treatment effect. The coefficient of variation (CV) was much smaller in the FCR than the VSA, and, unlike the VSA, the FCR was insensitive to gender effects. Both FCR and VSA seemed to perform better when the formant frequencies were first logarithmically transformed, and when 4 rather than 3 cardinal vowels were used to construct the metrics. In all, the FCR proved to be minimally sensitive to interspeaker variability and highly sensitive to vowel centralization associated with impaired vowel articulation. These findings have significant theoretical implications and useful research and clinical applications.