A ‘pivot’ model for extracting formant measurements based on vowel trajectory dynamics

Formant measurements are commonly extracted at fixed fractions across a vowel's duration (e.g. the 1/2 point for a monophthong and the 1/3 and 2/3 points for a diphthong). This approach tacitly relies on the convenience assumption that a speaker always maximally approximates the intended acoustic target at roughly the same point across a vowel’s duration. The present paper proposes an alternate method whereby every formant point sampled within a vowel is considered as a possible 'pivot' (i.e. turning point), with monophthongs modeled as having one pivot and diphthongs modeled as having two pivots. The optimal pivot for the vowel is then determined by fitting regression lines to the formant trajectory and comparing the goodness-of-fit of these lines to the raw formant data. When applied to a corpus of an American English dialect, the resulting measurements were found to be significantly correlated with previous methods. This suggests that the aforementioned convenience assumption is unnecessary and that the proposed model, which is more faithful to our understanding of articulatory dynamics, is a viable alternative. Moreover, rather than being assumed a priori, the location of the measurement can be treated as an empirical question in its own right.