Mapping segmental phonology to the speech signal: what's the acoustic energy baseline?

Is the phonetic expression of phonological features defined primarily with respect to the articulatory or the auditory-acoustic domain? This might have been expected to be one of the most basic questions dividing different feature theories. However, the edge is taken off the issue by the widely held (if not necessarily justified) assumption that feature definitions framed within one domain translate in a fairly direct way to the other.

There is a much more fundamental question to be asked about how features are phonetically specified: what is the baseline of acoustic energy in relation to which the definitions are framed or translated? The mainstream assumption, rarely made explicit, is that the baseline is zero. This thinking mirrors (and was probably heavily influenced by) the way acoustic energy is depicted in a sound spectrogram in terms of a unidimensional scale of intensity. For example, the acoustic difference between the minus and plus values of [continuant] can be defined as the extent to which the energy produced by a segment deviates from silence: oral stops deviate less from this baseline that do fricatives and vowels.

This paper reviews work with a very different take on how features map to acoustics: the baseline is not zero but rather the energy associated with the carrier signal in speech. This approach is explicitly founded on the conception of speech as a schwa-like carrier wave modulated by acoustic events. The linguistic message is contained in the modulations. The carrier makes the message audible but is in itself linguistically void (although it carries non-linguistic information about the talker). Features only map to energy that is linguistically significant: they encode modulations but not the carrier.

According to this approach, the phonetic expression of each feature is defined in terms of how it perturbs the energy baseline set by the carrier signal. In changing the shape of the carrier, some features decrease this energy, while others add to it.

Amongst the various feature frameworks out there, some versions of Element Theory (though by no means all) come nearest to implementing this idea. For example, the element [edge] (which characterises non-continuant consonants) radically reduces the energy of the carrier. The element [noise] (which characterises fricatives and plosives) adds aperiodic energy, potentially replacing the periodicity of the carrier.

The paper summarises some of the conceptual and empirical advantages this approach brings to the analysis of a range of phonological phenomena: consonant phonotactics, intervocalic voicing, consonant lenition, vowel reduction, vowel epenthesis and vowel syncope.