

On the edge: Acoustic cues to layered prosodic domains

Prosodic structure encodes the grouping of words into hierarchically layered prosodic constituents, including the prosodic word, intermediate phrase (ip) and intonational phrase (IP) (Ladd 1996). Acoustic cues to prosodic boundaries are observed in the lengthening of segments in the preboundary syllable rhyme, with greater effects of lengthening at successively higher levels of prosodic domains (Wightman et al. 1992). A second dimension of prosodic structure is the encoding of prominence, which also gives rise to lengthening effects in the prominent syllable (stressed or accented) (Turk & Sawusch 1997). Given two distinct sources of lengthening, the question arises whether lengthening on its own can serve as a cue to either prosodic context, a question which is not addressed in any prior work. Additional cues to prominence are found in F0 (Pierrehumbert 1980) and intensity (Kochanski et al. 2005), but these same features also give rise to distinctive patterns in pre-boundary position. This paper investigates the phonetic encoding of prosodic structure through a study of the acoustic correlates of prosodic boundary and their interaction with accent at three levels of prosodic structure: Word, ip, and IP. Guided by earlier evidence that boundary cues are local (Wightman et al. 1992), evidence for acoustic effects of prosodic boundary is considered in measures of duration, F0 intensity local to the domain-final rhyme.

Acoustic data is collected in read speech from the Radio News corpus (Ostendorf et al. 1995) from words in three contexts of prosodic phrase structure: phrase-medial, intermediate-phrase final, and intonational-phrase final positions. The corpus consists of about 10,000 words produced by 5 speakers, and is prosodically labeled using the ToBI transcription system. F0 measurements are normalized (z-transform) based on F0 within the IP. Intensity and duration are normalized for each phone based on all tokens in the corpus. All measurements are taken from the word-final syllable nucleus in the relevant prosodic context.

Separate ANOVAs were conducted for each acoustic measure with the independent factors of Boundary (Wd, ip, IP) and Accent (Accented, Unaccented). Results show a main effect of prosodic boundary on normalized duration, and post-hoc analysis reveals a three-way contrast, with Wd < ip < IP. There is no interaction effect (Boundary \times Accent), and mean values show that there is an additive effect of lengthening such that within each boundary condition, accented vowels are longer than unaccented vowels. For F0 and intensity measures, there are main effects of Boundary, and an interaction between Boundary and Accent: the effect of Accent on F0 and intensity masks that of Boundary. For example, F0 at the beginning of the nucleus and the maximum intensity over the nucleus serves to differentiate intermediate and intonational phrases only under the condition that pitch accent is not present on the pre-boundary word.

These findings provide strong evidence for prosodic theory, showing acoustic correlates of a 3-way distinction in boundary level. The evidence supports a model of prosody encoding in which cues to prosodic boundaries are local to the edges of prosodic domains. (Word Count 498)