



P1-2

The relationship between redundancy, prosodic prominence and acoustic cues in spontaneous speech

Matthew Aylett

Rhetorical Systems Ltd. and Department of Linguistics University of Edinburgh, U.K.

matthewa@cogsci.ed.ac.uk

In spontaneous speech, there are wide variations in the realisation of the acoustic cues for the same phoneme articulated by the same speaker. This paper explores two related factors which influence the variation of both duration and spectral cues, prosodic structure and redundancy.

The constraint of producing robust communication while efficiently expending articulatory effort leads to an inverse relationship between language redundancy and care of articulation. More careful articulation leads to more distinct acoustic cues. Thus this inverse relationship improves communication robustness by spreading information more evenly across the speech signal, yielding a smoother signal redundancy profile. We argue that prosodic prominence (as in degree of vowel reduction, lexical stress, and pitch accent) is a linguistic means of achieving smooth signal redundancy, as prosodic prominence increases care of articulation and coincides with unpredictable sections of speech.

The results of linear regressions carried out between measures of redundancy, a spectral and duration measure of care of articulation and prosodic structure in a large corpus of spontaneous speech confirm: 1) a strong relationship between prosodic prominence and care of articulation, and 2) an inverse relationship between language redundancy and care of articulation. The corpus consisted of fifteen hours of spontaneous speech from 64 different speakers.

In addition, when variation in duration due to the proximity of prosodic boundaries is controlled for, language redundancy can predict up to 65% of the variance in raw syllabic duration. This is comparable to 64% predicted by prosodic prominence (accent, lexical stress and vowel type). Moreover most (62%) of this predictive power is shared.

For a subset of material, a care of articulation measurement based on the spectral characteristics of the syllables' vowel supported these results.

The implications of this research are that a perceptual model should take into account the statistical variation of acoustic cues within a language, and that the importance of these cues should relate inversely to language redundancy.