



P3-17

**The ADS-to-IDS continuum:
Developing an index for the study of infant directed speech prosody**

Simone Ashby¹, Fred Cummins¹, and Deb Roy²

1: Department of Computer Science, University College Dublin

2: Cognitive Machines Group, MIT Media Laboratory

simone.ashby@ucd.ie, fred.cummins@ucd.ie, dkroy@media.mit.edu

Studies of infant directed speech (IDS) prosody have revealed highly modulated pitch contours as one of the primary salient features in distinguishing IDS from adult directed speech (ADS) prosody across a range of languages (Fernald and Simon, 1984; Garnica, 1977; Grieser and Kuhl, 1988; Stern et al., 1983). While serving a number of communicative intentions on the part of the caregiver, e.g. Attention-Bid, Prohibition, and Approval, there is evidence to suggest that the exaggerated pitch contours characteristic of IDS may serve an additional role of highlighting within the speech stream various cues to the linguistic structure of the ambient language (Jusczyk, 1992, 1997; Sundberg, 1998).

This paper is part of an open-ended investigation of those prosodic cues present in IDS that facilitate the segmentation and lexical acquisition processes that surface within the latter part of the infant's first year of life. We seek to develop an f_0 modulation index for English, compiled from data gathered on both IDS and ADS, as a means of capturing the degree of f_0 variation that occurs within utterances across the two speech styles. The purpose of the index is to establish an observable threshold marker that can provide a local characterization of the speech stream as relatively IDS- or ADS-like. Our goal is to use this index to subcategorize a corpus of IDS, isolating those utterances which clearly fall within the bounds of IDS or IDS-like speech from those that typify more ADS-like speech. If those prosodic cues which identify particularly IDS-like speech also serve to flag parts of the continuous audio stream to which the infant should pay particular attention, a case can be made for the utility of prosodic modification in IDS as facilitating an initial segmentation of the stream.

Preliminary results on index development and its application to a roughly 7,600-utterance IDS database will be presented.