

Prosodic properties of contrastive information in spontaneous productions

Shari R. Speer and Kiwako Ito

Ohio State University

contact: speer@ling.osu.edu

Speakers frequently use pitch accents (localized pitch prominence assigned at the phrasal level) to express the contrastive status of words (cf. Pierrehumbert & Hirshberg, 1986; Krahenr & Swerts, 2001; Watson, Tanenhaus & Gunlogson, 2008). In spontaneous dialogue, the phonetic realization of pitch accent depends on multiple factors, including overall pitch range, utterance length and rhythmic structure, the location of the accented word within phrasal and discourse structures, and individual differences in speech style. The relationship between the physical properties of pitch accents and their phonological, semantic and pragmatic representations is difficult to observe in laboratory experiments, where artificial tasks often restrict and may even alter the development of discourse context. In this study, we examined the prosodic properties of contrastive referential expressions elicited in an interactive conversational task. Participants gave instructions on how to decorate holiday trees with a given set of ornaments, where the sequence and location of ornaments were prompted via photographs displayed on a computer screen. No speaking instructions or example utterances were provided. This method loosely constrained the discourse structure of the dialogue as well as the range of referential expressions produced by speakers, facilitating the examination of turn taking and the potential contrastive intent behind each utterance.

Each speaker participant gave instructions to our confederate decorator and they together completed four trees. The instructor sat in a soundproofed booth and was presented with a series of photo images, each showing a single ornament and a tree labeled at the intended hanging location. The instructor told the decorator sitting outside the booth what ornament to pick and where to place it. Each ornament label combined a color adjective and an object noun (e.g., green drum), and trial order was used to create adjective contrasts (e.g. green drum, blue drum) and noun contrasts (e.g. green drum, green ball). The conversations between the instructor and the decorator were fully transcribed and separately ToBI-annotated by annotators blind to experimental condition. ToBI annotation employed an ALT (alternative) tier for the recording of prosodically ambiguous productions (Veilleux, Shattuck-Hufnagle & Brugos, 2006). Duration and F0 values were extracted for words in the target [Adj+Noun] phrases and for their stressed syllables.

Transcription and ToBI annotation of data from ten speakers yielded 580 contrastive utterances, where target noun phrases produced consecutively in the discourse contained a repeated noun or adjective. Results indicate that contrastive adjectives were produced most frequently with L+H* (57% of trials), but also with H* (33%). Nouns following a contrastive adjective were most frequently unaccented (56%), many showing vowel reduction (34%). However, nouns produced after a contrastive adjective sometimes carried pitch accents (44%; a variety of pitch accent types). Contrastive nouns were produced more frequently with !H* or H* (55%) than with L+H* (25%). Adjectives preceding a contrastive noun were highly likely to carry a pitch accent (91%; 74% were H*). This was so despite the fact that the majority of these adjectives were the penultimate (immediately pre-nuclear) words in their utterances, and nuclear accents were found on the utterance-final word in the vast majority of contrastive noun productions (97%). Measures of the stressed vowels' relative F0 height revealed that

unambiguous L+H* accents were higher than unambiguous H* and ambiguous L+H*/H* (where an alternative choice of a pitch accent was annotated) in both the adjective and the noun locations. Duration measures did not show such consistent patterns for either adjectives or nouns. Taken together, the present results indicate that during spontaneous dialogue, speakers express contrastive information primarily by the use of F0 changes, yet that the distribution of L+H* and H* on contrastive terms is strongly influenced by word location in the utterance.

References:

Pierrehumbert, J. & Hirschberg, J. (1986). The meaning of intonational contours in the interpretation of discourse. *Intentions in Communication*, P.R. Cohen, J.L. Morgan, Eds. MIT Press.

Krahmer, E. & Swerts, M. (2001). On the alleged existence of contrastive accents. *Speech Communication*, pp. 391-405.

Veilleux, N., Shattuck-Hufnagel, S., Brugos, A. (2006). Transcribing Prosodic Structure of Spoken Utterances with ToBI, section 2.12. MIT OpenCourseWare, <http://anita.simmons.edu/~tobi/tutorial.html>

Watson, D., Tanenhaus, M., Gunlogson, C. (2008). Interpreting pitch accents in online comprehension: H* vs. L+H*. *Cognitive Science*, 32, 1232-1244.