Abstract for WPSI IV. Joseph Tyler, University of Michigan.

At least since Harris (1952), researchers have studied how language is organized above the level of a single sentence. A number of approaches have been proposed, e.g. using semantic and pragmatic information, assessments of social action, and inferences about speaker intentions. But how do we know if a theory captures how speakers mentally represent the structure of discourse? I argue that prosody can help us gain empirical traction on what aspects of discourse structure are in speakers' mental representations.

Therefore, this study asks, do speakers communicate discourse structure in their prosody? Breaking down this larger question, what specific features of discourse do speakers encode prosodically? To address these questions, it is necessary first to have a representation of the structure of a discourse and then code that structure in such a way as to make it possible to test for prosodic correlates. In general, discourse theories analyze discourse into segments, relations that link those segments, and some kind of hierarchical structure (Asher & Lascarides, 2003; Grosz & Sidner, 1986; Mann & Thompson, 1988). Segmented Discourse Representation Theory (SDRT), which has not, to my knowledge, been used previously in a study looking for prosodic correlates of discourse, is one theory that has all of these components (segmentation, relations, hierarchy). As a result, discourse as modeled with SDRT can be used to address the following three sub-questions that depend on these components. First, does prosody indicate whether some boundaries between parts of a discourse are small and others large? Second, does prosody indicate whether a segment of the discourse is higher or lower in the overall structure? And third, does prosody distinguish relations linking hierarchically equal segments (coordinating) from relations linking segments at different hierarchical levels (subordinating)? This study uses SDRT to address each of these questions experimentally.

The following text is an excerpt from a Wall Street Journal article, segmented according to SDRT (Reese, Denis, Asher, Baldridge, & Hunter, 2007).

40. But the mainstream civil-rights leadership generally avoided the rhetoric of "law and order,"
41. regarding it as a code for keeping blacks back.
42. Law and order didn't mean justice,
43. Mr. Jackson used to say,
44. but "just us."
45. In the past, many were hesitant to speak about crime in public
46. because "the larger community would talk about 'lock them up and throw the key away'
   and hide behind black leaders in doing it,"
47. explains Rep. Craig Washington,
48. the Houston Democrat who led the caucus hea
49. Now there is escalating discourse within the black community about what it can and
   must do to stop crime.
50. Just after the new year, Mr. Jackson held the first of several conferences focusing on just
   that.
51. "The premier civil-rights issue of this day is youth violence in general and black-on-
   black violence in particular,"
52. he has said.

Within a full SDRT analysis, the boundary between segments 41 and 42 is smaller than between 48 and 49. Would that different boundary size correlate with prosody? Furthermore, segments 40 and 49 are higher in the overall structure than all of the other segments in this excerpt; would this hierarchical height correlate with prosody? Finally, segments 40 and 49
are linked by a coordinating relation while segments 40 and 45 are linked by a subordinating relation; could this coordinating/subordinating contrast correlate with prosody?

To test these questions, this study investigates three specific prosodic measures for how they correlate with discourse structure: the commonly studied measures of pause duration and maximum pitch level (f0max), as well as the little studied measure of minimum pitch level (f0min). All of the questions above are investigated experimentally by having participants read aloud a news article, and then identifying correlations between the prosody of those readings and the article’s discourse structure.

Results showed that larger boundaries had significantly longer pause durations, while a discourse segment’s height in an overall discourse structure did not correlate with any of the prosodic measures. Results also indicated that relations linking hierarchically equal segments and relations linking segments at different hierarchical levels correlated differently with pause duration and f0min but not f0max. The first argument of a subordinating relation tends to have a longer preceding pause and higher f0min than the second argument. By contrast, the two arguments of a coordinating relation tend to have relatively similar pause duration and f0min. As a result, one way of indicating in speech that a discourse segment is subordinated, i.e. at a greater level of detail, appears to be to have a shorter preceding pause and lower f0min than the segment that it is subordinated to. Another notable finding is that f0min correlated differently with coordinating and subordinating relations. This finding is particularly important because prior discourse prosody research has tended to assume f0min does not vary and so has excluded it from its analysis (den Ouden, Noordman, & Terken, 2009; Hirschberg & Grosz, 1992; Pierrehumbert & Hirschberg, 1990).

This study’s results provide evidence that speakers communicate discourse structure in their prosody. Specifically, speakers use pause duration to indicate the size of a discourse boundary, and they use pause duration and f0min to indicate whether a relation links segments at the same hierarchical level (coordinating relation) or different hierarchical levels (subordinating relation). These correlations provide evidence for the claim that SDRT captures something about how speakers actually represent the structure of discourse.


