Available online at www.sciencedirect.com

SciVerse ScienceDirect

Lingua 133 (2013) 101–126

Lingua

www.elsevier.com/locate/lingua

Prosodic correlates of discourse boundaries and hierarchy in discourse production

Joseph Tyler *

University of Michigan, 440 Lorch Hall, 611 Tappan Street, Ann Arbor, MI 48109-1220, USA

Received 3 August 2011; received in revised form 7 April 2013; accepted 10 April 2013
Available online 23 May 2013

Abstract

A well-formed discourse is more than just a series of well-formed sentences. While often left implicit, this structure to discourse is sometimes overtly cued. And though most attention in this area has focused on lexicalized cues like discourse markers, prosody can also convey information about the structure of discourse. This paper presents the results of a production study examining prosodic correlates of discourse structure in readings of a newspaper article. Prosodic measures of pause duration, pitch, intensity and speech rate were found to significantly correlate with discourse structural measures of boundary size, discourse coordination/subordination, and their interaction. This interaction effect shows that the effect of boundary size on an utterance's prosody often depends on whether that utterance is coordinated or subordinated, and vice versa. These results expand our understanding of how prosody correlates with discourse structure, setting the stage for follow-up perception studies of what prosodic variation listeners use in discourse interpretation.

Keywords: Discourse; Prosody; Intonation; Coordination; Subordination; Segmented discourse representation theory

1. Introduction

Language is clearly structured in many different ways. Established areas of linguistics have for decades studied the systematic organization of sounds (phonology) and parts of a sentence (syntax). Similarly, the sentences of a discourse are structured, and a well-formed discourse is more than just a series of well-formed sentences. One way to reveal this structure is to remove it, perhaps by re-ordering the sentences of a discourse. For instance, if you were to read the sentences of this paragraph from last to first, the resulting discourse would be quite hard to follow. Even the two possible orderings of two sentences can lead to different interpretations of the events narrated.

(1) John banged his head. He fell over.
(2) John fell over. He banged his head.

A natural interpretation of the discourse in (1) is that John's banging his head happened before his falling over, while a natural interpretation of (2) is that John first fell over and then banged his head. In addition to the temporal ordering contrast, these two discourses likely also have different causal relationships. In (1), the banging of his head seems likely to have caused John to fall over. In (2), John's falling over seems likely to have led to him to bang his head.

* Present address: P.O. Box 2713, Department of English Literature and Linguistics, Qatar University, Doha, Qatar. Tel.: +974 734 761 5549; fax: +974 734 936 3406.
E-mail addresses: jctyler@umich.edu, jctyler5@gmail.com.

0024-3841/$ – see front matter
http://dx.doi.org/10.1016/j.lingua.2013.04.005

While it seems clear there is structure in discourse, it is less clear exactly what that structure is. Sometimes aspects of discourse structure are explicitly cued, while other times a speaker leaves the structure implicit, leaving listeners to fill in the gaps with their own reasoning. Most work that has analyzed explicit cues to discourse structure has focused on *lexical* cues, e.g. discourse markers. If (1) was instead produced as (3), with the addition of the explicit marker of temporal succession *then*, the temporal relationship between the two sentences would be explicit.

(3) John banged his head. Then he fell over.

In (3), it is explicit that John banged his head and subsequently fell over. An alternative, though dispreferred, interpretation of (1) could have been that it described two separate, independent events with no information about when each happened. In this interpretation, (1) would describe two independent events that happened to John, banging his head and falling over. With the addition of the discourse marker *then* in (3), the temporal ordering is explicitly encoded and this alternative is ruled out. Thus, the addition of a lexical item like a discourse marker can make explicit how the sentences of a discourse are related.

One feature of discourse identified by many theorists (Grosz and Sidner, 1986; Hobbs, 1985; Mann and Thompson, 1988; Polanyi, 1988; Van Kuppevelt, 1995) is that it is hierarchically structured. Asher and Vieu (2005) discuss the intuitions motivating hierarchical structure in the context of Segmented Discourse Representation Theory (SDRT) (Asher and Lascarides, 2003). They mention paragraph structure as an orthographic manifestation of discourse hierarchy, where paragraph-initial sentences are in some sense higher-order than paragraph-medial sentences. A paragraph-medial sentence likely provides more detail about whatever was introduced by the paragraph-initial sentence. They also argue that temporal structure motivates a hierarchical conception of discourse. If one sentence introduces an event and a second sentence describes something occurring at the same time as that first event, the second is likely providing more detail about the first event. By contrast, if a second sentence describes an event at a different time, the two events likely have equal status.

Like most theories of discourse structure, SDRT analyzes the structure of discourse by segmenting the discourse, identifying relations that hold between segments, and constructing a hierarchy from the segments and relations. SDRT focuses on both semantic and pragmatic information for all stages of analysis (segmentation, relation identification, hierarchy). SDRT also provides an inventory of discourse relations (e.g. Elaboration, Background, Result) that are claimed to hold between the segments of a discourse. But most importantly here, SDRT builds hierarchy in discourse by classifying all discourse relations as either coordinating or subordinating. Coordinating relations link discourse segments at an equal hierarchical level while subordinating relations link a discourse segment with another segment one hierarchical level lower.

Rhetorical Structure Theory (RST) (Mann and Thompson, 1988), like SDRT, analyzes a discourse into segments, identifies relations between segments, and constructs the discourse into a hierarchical structure. RST also has a local hierarchical structure contrast in its nucleus-satellite distinction. In RST, all discourse segments are considered to be either a nucleus or a satellite. The distinction between the two is defined in terms of a segment's relative importance to the coherence of the discourse. One diagnostic test is that satellites can be deleted without harming the overall message of the discourse, while deleting a nucleus would disrupt the discourse's coherence. This test reveals one of RST's applications: automatic text summarization. If all satellites in a text were deleted, the result would be a stripped down summary of the discourse.

While RST's nuclearity principle has been compared to SDRT's coordinating/subordinating contrast (Danlos, 2010), there are points of contrast. In RST, nuclearity is a feature of a discourse segment. This means that every discourse segment is either a nucleus or a satellite. In SDRT, coordinating and subordinating relations are theorized to hold between discourse segments, but are not strictly features of the segments themselves. This means that any one segment in an SDRT analysis could be coordinated to one segment and subordinated to another. Another difference between RST's nuclearity and SDRT's coordinating/subordinating contrast is in terms of how an analyst identifies a segment's nuclearity or CoordSubord status. In RST, a central criterion for satellite status is that a discourse segment be expendable: if it can be deleted without harming the discourse's coherence, it is a satellite. In SDRT, the main point of contrast between coordination and subordination is in terms of the level of detail. So, RST and SDRT both supply theoretical constructs that account for local hierarchical contrasts, but the nature of those local hierarchical constructs is not exactly the same.

Another influential theory of discourse that analyzes discourse into segments, relations between segments, and hierarchy is the Grosz and Sidner model (1986). Unlike SDRT and RST, which focus on the propositional content of utterances as the basis of their analyses, the Grosz & Sidner model analyzes discourse using speaker purposes, goals and intentions. In this theory, a speaker may have one overall purpose to their discourse, e.g. to give directions on how to replace a car battery. Then, this overall purpose may be subdivided into a series of subgoals, e.g. how to identify the battery, how to remove the old battery, and how to install the new battery. Grosz & Sidner propose two structural relations that organize these discourse purposes into a hierarchical structure: dominance and satisfaction-precedence. The higher-order purpose of replacing a car battery is said to dominate the three subgoals. And since the removal of the old battery needs to be complete before the