Discourse subordination licenses D-linking:
An experimental approach to rhetorical relation effects

We use evidence from an experiment on Amazon Mechanical Turk to show that subordinating rhetorical relations license the dependencies of English D-linked *wh*-phrases. This result demonstrates a regularity between rhetorical relations and information structural categories, and it provides a methodology for isolating the connection between the two. Additionally, some implications arise for the corpus annotation of rhetorical relations.

D-linked *wh*-phrases (e.g. *which store*) have been described as salient alternative sets (Pesetsky 1987, 2000; Comorovski 1996). D-linking has also been characterized in terms of previous mention (Grohmann 1998) and topicality (Polinsky 2001; Grewendorf 2012; a.o.). In (1a–d), the set of *stores* is mentioned in (c–d), and the definite article in (c, italics) indicates that the set is saliently shared by the speaker and addressee. However, D-linking of *store* is infelicitous in (e), so salience and previous mention are not sufficient.

1. a.) Speaker A: Randy bought a lot of vegetables$_i$ yesterday, didn't he?
   b.) Speaker B: Yes, he did.
   c.) Speaker B: First, he bought some onions and peppers$_i$ from *the cheapest store$_j$ in town*.
   d.) Speaker B: Then, he bought some interesting vegetables$_i$ from an Asian market$_j$.
   e.) Speaker A: #After that, **which store$_j$** did he go to?
   e'.) Speaker A: After that, **which vegetables$_i$** did he buy?

What does differentiate between (e) and (e') is *discourse subordination*, defined in Asher & Lascarides' (2003) Segmented Discourse Representation Theory as asymmetrical rhetorical relations where one sentence explains, elaborates, or expands upon another (cf. also Grosz & Sidner 1986; Asher & Vieu 2005). Sentences (1c–e) all *elaborate* on (a), a subordinating relation. *Which vegetables* is felicitous, because its antecedent is linked by that subordinating relation. But there is no subordinating relation linking *which store* to its antecedents.

In an Amazon Mechanical Turk experiment, 50 participants listened to 17 contexts like (1) and judged their naturalness on a sliding scale. D-linked *wh*-phrases in subordinating and non-subordinating relations with their antecedents were compared, along with non-D-linked *wh*-phrases like *what store*. Sentences like (1e) were rated significantly worse than others like (1e'; median rating: 52.25/100 vs. 80.5/100). In a linear mixed effects model including random effects for speaker, context, and syntactic argument, the presence of discourse subordination is found to raise median ratings by $24/100 \pm 8.3$ ($p < 0.05$).

We propose a methodology to test further links between rhetorical relations and information structure. Participants on Amazon Mechanical Turk can reliably detect differences in discourse structures via audio prompts, provided enough context is given. First, cue words like *After that…* are necessary to disambiguate rhetorical relations (here, Narration). Second, rhetorical relations should be chosen that provide the strongest contribution to the context (cf. Matsui 1995; Asher & Lascarides' 2003 Maximize Discourse Coherence). Third, the *preceding discourse* should be made explicit *for all speakers*. In (1a–b), both speakers explicitly arrange the conversation around *vegetables*. If Speaker B's agreement in (1b) is omitted, *which stores* is rated more highly. We hypothesize that this enables accommodation, such that participants interpret the discourse as being arranged around *stores* by Speaker B.

If our analysis is taken to be correct, the presence of a D-linked *wh*-phrase can be used to detect a subordinating relation leading to its antecedent. However, the sentence containing the antecedent may be implicit and accommodated. We assert that rhetorical relations can only be reliably annotated in *well-defined segments* of conversational discourse, where all speakers explicitly agree on the prior discourse structure.

Asher, N. & A. Lascarides. 2003. *Logics of Conversation*. Oxford UP.

Asher, N. & L. Vieu. 2005. "Subordinating and coordinating discourse relations." *Lingua* 115:591–610.

Comorovski, I. 1996. *Interrogative phrases and the syntax-semantics interface*. Springer.

Grewendorf, G. 2012. "Wh-movement as topic movement." In, L. Brugé, et al (eds.), *Functional Heads: The Cartography of Syntactic Structures*, Volume 7. Oxford UP.

Grohmann, K. K. 1998. "Syntactic inquiries into discourse restrictions on multiple interrogatives." *Groninger Arbeiten zur germanistischen Linguistik* 42:1–60.

Grosz, B. & C. Sidner. 1986. "Attention, intentions, and the structure of discourse." *Computational Linguistics* 12(3):175–204.

Matsui, T. 1995. *Bridging and Relevance*. Ph.D. thesis, University of London.

Pesetsky, D. 1987. "Wh-in-situ: Movement and unselective binding." In E. Reuland & A. ter Meulen (eds.), *The representation of (in)definiteness*. MIT Press: 98–129.

Pesetsky, D. 2000. *Phrasal movement and its kin*. MIT Press.

Polinsky, M. 2001. "Information structure and syntax: Topic, discourse-linking, and agreement." *Third Workshop on Spoken and Written Texts*. Austin: Texas Linguistic Forum.