

Cue phrases in Spoken Language: Discourse Pragmatics at the Forefront

Catherine Lai, Johanna Moore

School of Informatics, University of Edinburgh
Edinburgh, EH8 9AB, United Kingdom
clai@inf.ed.ac.uk, jmoore@ed.ac.uk

Being able to model discourse structure is crucial for developing robust spoken language understanding and generation applications. While previous computational studies have tended to focus on newswire style text corpora (Carlson et al., 2003; Prasad et al., 2008; Reese et al., 2007), it is easy to identify discourse related linguistic resources that are extremely frequent in spontaneous speech but rare in written corpora. A prominent example of this is the use of cue phrases such as ‘well’, ‘so’, ‘okay’, and ‘I mean’, (Schiffrin, 1987). These *pragmatic* markers are often excluded from studies on the grounds that they do not describe *semantic* discourse relations (Redeker, 1990; Prasad et al., 2008; Taboada and Gómez-González, 2012). Nevertheless, cue phrases give us vital information about how parts of a discourse relate to one another.

The importance of pragmatic markers comes to the forefront when we examine spoken language, particularly dialogue. They provide cues for how to process utterances in an ongoing discourse. The following discussion on a US political party leadership contest demonstrates a number of these uses.¹ Here, ‘well’ signals that a speaker cannot directly answer a question and more information needs to be elicited to provide a satisfactory answer. ‘I mean’ qualifies a previous assertion while ‘so’ relates utterances to a previous parts of the discourse.

- (1) A: Did you have a sense coming in that the anti-Boehner revolt was picking up steam? Were the offices of the Daily Color cheering with every ‘yo ho’ shout out on the house floor?
- B: Well, you know, first of all you have to remember a couple things.
- B: One, journalists, even center right journalists, are very skeptical and cynical.
So, I think everybody was skeptical that anything was going to happen.
- B: On the other hand, you’re rooting for the story. I mean, everybody is sort of rooting, not that you have a stake in it, but wouldn’t it be great if it came close, right?
that would be fun to write about
- B: So, i think you sort of have these conflicting, you know, instincts or urges.

The frequency of these markers in spoken language is indicative of the need to build and alter discourse structures on the fly in this modality. In this vein, (Tonelli et al., 2010) report the need to extend the PDTB sense taxonomy in or-

der to account for an increased frequency of the pragmatic use of discourse connectives.

We argue that cue phrases are best understood in the Question Under Discussion (QUD) framework (Roberts, 1996; Ginzburg, 2012). Here, discourse constituents are related through context induced questions (setting the discourse topic) rather than a closed set of relations. This framework is process driven and allows us to model cue phrases as signalling potential updates to the common ground via the stack (equivalently a tree) of questions. A benefit of modelling discourse using a question driven approach is that the question structure allows us to make the link to sentential Information Structure (Steedman, 2000). This, in turn, has proven valuable for spoken dialogue generation, particularly in terms of prosody (White et al., 2010).

Coherence in the QUD framework is defined very locally: an utterance is coherent if it directly answers a (possibly implicit) relevant question. This results in an infinite range of potential questions. To make such a structure tractable for computational applications, we would like to see whether questions can be mapped to broad classes, how well these classes fit classic taxonomies of relations (cf. Jasinskaja (2010)), and how they relate to non-verbal aspects of speech such as prosody. Similarly, we would like to see how the process driven QUD model applies to spoken monologues. A barrier to using supervised machine learning techniques to investigate this is the lack of speech data annotated for both QUD/information structure and discourse relations, as both are notoriously difficult annotation tasks (Stent, 2002; Calhoun, 2007).

To make some headway on this we propose to use unsupervised topic modelling techniques, e.g. LDA (Blei, 2012), to investigate the connection between pragmatic markers, subordination, and coordination (Asher and Vieu, 2005). Initially, we very broadly describe subordinate utterances as specifying details about the topic raised in the previous utterance (adding a subquestion), while coordinate utterances introducing new topics (adding a sister question). We propose to approximate discourse structure by inducing topic vectors for utterances and then inducing tree structures based on topic similarity. Besides helping us to understand how topic modelling relates to discourse structure, we hope that using these automatically derived structures as a first pass will facilitate fine-grained discourse annotation. More generally, we hope that examination of pragmatic and ideational discourse markers with respect to this sort of representation will lead us to a better understanding of how to model spoken language.

¹<http://bloggingheads.tv/videos/32893>

References

- Asher, N. and Vieu, L. (2005). Subordinating and coordinating discourse relations. *Lingua*, 115(4):591–610.
- Blei, D. M. (2012). Probabilistic Topic Models. *Commun. ACM*, 55(4):77–84.
- Calhoun, S. (2007). *Information structure and the prosodic structure of English: A probabilistic relationship*. PhD thesis, University of Edinburgh.
- Carlson, L., Marcu, D., and Okurowski, M. E. (2003). Building a Discourse-Tagged Corpus in the Framework of Rhetorical Structure Theory. In van Kuppevelt, J. and Smith, R. W., editors, *Current and New Directions in Discourse and Dialogue*, number 22 in Text, Speech and Language Technology, pages 85–112. Springer Netherlands.
- Ginzburg, J. (2012). *The Interactive Stance: Meaning for Conversation*. Oxford University Press.
- Jasinskaja, E. (2010). Modelling Discourse Relations by Topics and Implicatures: The Elaboration Default. In Kühnlein, P., Benz, A., and Sidner, C., editors, *Constraints in Discourse 2*, pages 61–79. John Benjamins.
- Prasad, R., Dinesh, N., Lee, A., Miltsakaki, E., Robaldo, L., Joshi, A., and Webber, B. (2008). The Penn Discourse Treebank 2.0. In *Proceedings of the 6th International Conference on Language Resources and Evaluation (LREC 2008)*, pages 2961–2968.
- Redeker, G. (1990). Ideational and pragmatic markers of discourse structure. *Journal of Pragmatics*, 14(3):367–381.
- Reese, B., Denis, P., Asher, N., Baldridge, J., and Hunter, J. (2007). Reference Manual for the Analysis and Annotation of Rhetorical Structure (Version 1.0). Technical report, Technical report. Austin: University of Texas, Departments of Linguistics and Philosophy.
- Roberts, C. (1996). Information structure in discourse: Towards an integrated formal theory of pragmatics. *Working Papers in Linguistics-Ohio State University Department of Linguistics*, pages 91–136.
- Schiffrin, D. (1987). *Discourse markers*. Studies in interactional sociolinguistics. Cambridge University Press.
- Steedman, M. (2000). Information structure and the syntax-phonology interface. *Linguistic Inquiry*, 31(4):649–689.
- Stent, A. J. (2002). A conversation acts model for generating spoken dialogue contributions. *Computer Speech & Language*, 16(34):313–352.
- Taboada, M. and Gómez-González, M. D. I. A. (2012). Discourse markers and coherence relations: Comparison across markers, languages and modalities. *Linguistics and the Human Sciences*, 6(1-3).
- Tonelli, S., Riccardi, G., Prasad, R., and Joshi, A. K. (2010). Annotation of Discourse Relations for Conversational Spoken Dialogs. In *LREC*.
- White, M., Clark, R. A. J., and Moore, J. D. (2010). Generating Tailored, Comparative Descriptions with Contextually Appropriate Intonation. *Comput. Linguist.*, 36(2):159–201.