

## Examples

- (1) a \*When this song I heard last week, I remembered my first love  
 b When last week I heard this song, I remembered my first love  
 c *Quand cette chanson je l'ai entendue...* (French)  
 when this song I it-have-1sg heard-PART-FSG
- (2) a. (%)\*John regrets that this book Mary read. (Maki *et al*, 1999: 3, (2c))  
 b. I regret that in those days I didn't realize the importance of classical languages.  
 c. *Mi dispiace che questo problema gli studenti non l'abbiano potuto risolvere.*  
 me displeases that this problem the students *non* it have been able to solve
- (3) a. John resents that *this book* Mary read from cover to cover, and not that the other (his favorite) she didn't even open. (compare to (2a), which is not easily read contrastively)  
 b. It's that *this book* Mary read that John resents.
- (4) *János AZT felejtette el, hogy MARI tegnap kit választott.*  
 John Expl forgot Prt Comp Mary yesterday whom chose  
 "What John forgot is whom MARY chose yesterday."

## References

- Arsenijević, B. (2009), Clausal complementation as relativization. *Lingua* 119: 39–50.
- Bhatt, R. & Pancheva, R. (2006), Conditionals. In *The Blackwell Companion to Syntax* (M. Everaert, H. van Riemsdijk, eds.). Vol 1: 638-687.
- Bianchi, V. & Frascarelli, M. (2009), Is topic a root phenomenon? <http://ling.auf.net/lingBuzz/000954>
- De Cuba, C. & Ürögdi, B. (2009), Eliminating Factivity from Syntax: Sentential complements in Hungarian. In *Approaches to Hungarian*. Amsterdam and New York: John Benjamins.
- Demirdache, H. & Uribe-Etxebarria, M. (2004), The syntax of time adverbs. In *The syntax of time* (J. Guéron, J. Lecarme, eds.), Cambridge, Mass.: MIT Press, 143-180.
- Geis, M. (1970), *Adverbial subordinate clauses in English*. Ph.D.diss., Cambridge, Mass.: MIT.
- Haegeman, L. (2007), Operator movement and topicalization in adverbial clauses. *Folia Linguistica* 18: 485-502.
- Haegeman, L. (2010), The internal syntax of adverbial clauses. In *Exploring the left periphery. Lingua thematic issue*, (K. Grohmann, I. Tsimpli, eds.), 120: 628-648.
- Hegarty, M. (1992), Familiar complements and their complementizers: On some determinants of A'-locality. Unpublished manuscript: University of Pennsylvania.
- Hooper, J. & Thompson, S. (1973), On the applicability of root transformations. *Linguistic Inquiry* 4: 465-97.
- Kiparsky P. & Kiparsky, C. (1970), Fact. In *Progress in Linguistics*, (M. Bierwisch, K.E. Heidolph, eds). The Hague: Mouton. 143-73
- Kiss, K. É. (2002), *The Syntax of Hungarian*. Cambridge University Press.
- Larson, R. (1990), Extraction and multiple selection in PP. *The Linguistic Review* 7: 169-182.
- Melvold, J. (1991), Factivity and definiteness. In *More Papers on Wh-Movement, MIT Working Papers in Linguistics*. No. 15 (L. Cheng, H. Demirdache, eds.), MIT: Cambridge, Mass, 97-117
- Munsat, S. (1986), *Wh-complementizers*. *Linguistics and Philosophy* 9: 191-217.
- Rizzi, L. (2004), Locality and left periphery. In *Structures and Beyond. The Cartography of Syntactic Structures*, vol. 3 (A. Belletti, ed.), Oxford University Press. Oxford.
- Starke, M. (2001), *Move dissolves into Merge: a theory of locality*. Ph.D. diss. University of Geneva. [http://theoling.auf.net/papers/starke\\_michal/](http://theoling.auf.net/papers/starke_michal/)

### On interplay of information structure, anaphoric links and discourse relations

Hajicova, Eva  
 (Charles University, Prague)

There is an abundant evidence of research carried out in the domain of information structure of the sentence, on anaphoric and coreferential relations and on most different aspects of discourse structure. However, there is no systematic account of the interplay of these aspects. At the same time, the current existence of large text corpora available for linguistic investigations invites linguists to try and examine the interaction and cooperation of the above mentioned aspects on a corpus annotated systematically with regard to the syntactic structure of sentences, to their information structure as well as to the (basic) anaphoric/coreferential links and (at least) to some aspects of discursive relations.

Our examination of the interplay of the above mentioned aspects of sentence and discourse structure is based on the material of the Prague Dependency Treebank, an annotated corpus of Czech texts. This corpus consists of 3165 documents containing 49431 sentences (with the total of 833195 word tokens); all the sentences are annotated on a morphological, surface and underlying (syntactico-semantic) layers including the information-structure of the sentence; in about 90% of them also basic anaphoric links ((pro)nominal coreference and associative links) is registered and an annotation of discursive relations is carried out for which an annotation manual has already been published..

The following aspects are taken into consideration (notation: X::=Y stands for “there is a relation between X and Y”)

(i) Tectogrammatical (underlying syntactic) trees (TGTS) in the PDT multilayer scheme of annotation capture underlying dependency syntactic relations, and thus are supposed to represent the literal meaning of sentence. Nodes are reconstructed in case of surface deletions of sentence elements. The elements “missing” in the surface shape of the sentence are indispensable if the sentences have to be understood in the appropriate (discourse) contexts. Therefore, the analysis of any intersentential relations is to be based on the underlying rather than superficial shape of the sentence.

(ii) Dependency relations::= Discursive relations

TGTSs of sentences capture also several types of relations (both explicit and implicit) that hold intrasententially between clauses as well as intersententially. The amount of discursive relations we get “for granted” is not negligible: in PDT there are 87916 clauses, i.e. approx. twice as many as the number of sentences, and the relations between them are already established. In a general perspective, it is an interesting research issue to look at the sets of intra- and inter-sentential relations and to study their commonalities and differences.

(iii) Information structure of the sentence::= Discursive relations

Information structure of the sentence is captured in the TGTSs by means of one of the values of a special attribute with each node, differentiating between a contextually bound contrastive element, a contextually bound non-contrastive element and a contextually non-bound element. One of the examples of an interaction between information structure and discourse is contrast: Also, the dynamics of discourse structure can be shown as depending, among other things, on the information structure of sentence in the texts.

(iv) Information structure of the sentence ::= Anaphoric relations

The contextually bound items, by their definition, are expected to be found in an anaphoric relation to some item(s) in the preceding co(n)text, however, an interesting research question is under which circumstances there is an anaphoric link to an unbound item, and, more generally, to look for these interrelations in case of bridging anaphora.

(v) Anaphoric relations ::= Discursive relations

Some anaphorical expressions also act as connective elements establishing a certain semantic discursive relation (e.g. *therefore*, *instead of*). Thus it is advantageous if these relations are studied in their cooperation.

The interactions of all the above mentioned aspects will be illustrated on a text fragment and it will be demonstrated how a multi-aspectual annotation of a text makes it possible to reach a more consistent and complex picture discourse.

## References

- Hajič, J. et al. (2006), *Prague Dependency Treebank 2.0*. Linguistic Data Consortium, Philadelphia.
- Hajičová, E. & Jínová, P. (in prep.), Why do we annotate discourse relations on syntactic trees rather than on linear texts. In (A. Lee, R. Prasad, A. Joshi, N. Dinesh, B. Webber, eds.), *Complexity of Dependecnies in Discourse*, In: *Proceedings of TLT 2006*, Prague, 79-90.
- Mladová, L (in prep.), Where coreference and discourse relations interact.
- Mladová, L, Zikánová, Š & Hajičová E. (2008), From Sentence to Discourse: Building an Annotation Scheme for Discourse Based on Prague Dependency Treebank. In *Proceedings of the 6th International Conference on Language Resources and Evaluation (LREC 2008)*. 2564–2570
- Nedoluzhko, A. et al. (2009), Extended Coreferential Relations and Bridging Anaphora in the Prague Dependency Treebank. *Proceedings of the 7th Discourse Anaphora and Anaphor Resolution Colloquium (DAARC 2009)*, Goa, India
- The PDTB-Group (2006), *The Penn Discourse Treebank 1.0. Annotation Manual. Technical Report IRCS-06-01*, Univ. of Pennsylvania.
- Webber B., Joshi A., Stone M. & Knott, A. (2003), Anaphora and Discourse Structure. *Computational Linguistics* 29(3)
- Xia, F., Bhatt, R., Rambow, O., Palmer M. and D. Misra Sharma, (2009), Towards a Multi-Representational Treebank, In *Proceedings of TLT 2009*, 159-170.