Generating Questions under Discussion

Kordula De Kuthy
Universität Tübingen
kdk@sfs.uni-tuebingen.de

Tobias Kolditz
Universität Tübingen

Detmar Meurers
Universität Tübingen
dm@sfs.uni-tuebingen.de

Tag
Datum
Zeit
Raum

Questions under Discussion (QUD) (Roberts, 2012) have emerged as a way to spell out the hinge between the properties of the sentence and the nature of the discourse in which the sentence can function. Riester, Brunetti & De Kuthy (2017) spell out a discourse annotation approach based on explicit pragmatic principles for determining a QUD for every assertion in a text. We propose to partially automate the annotation by generating all potentially relevant questions for a given sentence. From this set, the annotator can then select the contextually appropriate QUD.

In computational linguistics, question generation has been tackled in several applied contexts for English. To explore the possibilities and challenges of question generation for German, we piloted the development and implementation of a transformation-based question generation system (Kolditz, 2015). We showed that the majority of generated questions are of high quality for the limited set of targeted question types, potential answer phrases, and syntactic environments in which they can occur. We are currently extending the empirical reach to cover a broader range of question types and syntactic configurations for which questions need to be expressible.

The nature and formulation of the questions we need to be able to generate for a sentence crucially is constrained by the nature of the immediate context in which the sentence occurs. Following the Q-Givenness principle (Riester et al., 2017), QUDs essentially consist of Given material. Generating questions for a text sentence by sentence makes it possible to model the constantly evolving set of given material, using approximations of the formal pragmatic notion of givenness (Ziai, De Kuthy & Meurers, 2016).

References: • Kolditz, T. (2015): *Generating questions for German text.* Master thesis. University of Tübingen. • Riester, A., L. Brunetti & K. De Kuthy (2017): Annotation guidelines for questions under discussion and information structure. In E. Adamou, K. Haude, & M. Vanhove (eds.), *Information structure in lesser-described languages: Studies in prosody and syntax.* John Benjamins. • Roberts, C. (2012): Information structure in discourse: Towards an integrated formal theory of pragmatics. *Semantics and Pragmatics* 5(6). 1–69. • Ziai, R., K. De Kuthy & D. Meurers (2016): Approximating Givenness in Content Assessment through Distributional Semantics. In *Proceedings of *SEM*, 209–218. ACL.