

## Annotating Corpora with Information Structure

### Abstract:

In this course, participants will get acquainted with contemporary information structure theory as well as state-of-the-art methods for creating manually annotated text and speech corpora. The course is aimed at students and researchers of both theoretical and computational linguistics.

We first provide an introduction to classical and contemporary interpretations of information structural notions including focus, aboutness topic, contrastive topic, givenness, anaphora and information status, as well as relevant underlying frameworks and concepts like Alternative Semantics, discourse trees and Questions under Discussion.

Turning from theory to practice, we will introduce the participants to the use of the EXMARaLDA annotation tool. In a practical exercise, participants will analyse texts with respect to referential information status (anaphora) as well as lexical information status (semantic relations). Furthermore, we will discuss ways of identifying focus and topic constituents in corpus data.

At the end of the class, participants will have gone through a detailed information structural analysis of an English text or piece of spoken language, and will have acquired methods for the analysis of their own preferred type of linguistic data.

### Instructors:

- Kordula De Kuthy
  - *Email:* [kdk@sfs.uni-tuebingen.de](mailto:kdk@sfs.uni-tuebingen.de)
- Arndt Riester
  - *Email:* [arndt.riester@ims.uni-stuttgart.de](mailto:arndt.riester@ims.uni-stuttgart.de)

### Course meets:

- Week II: 11:00–12:30, Room N9

### Coursewebpage and Syllabus:

- <http://www.sfs.uni-tuebingen.de/kdk/essli14>
- <http://www.sfs.uni-tuebingen.de/kdk/essli14/syllabus.pdf>

### Software and Corpora:

- EXMARaLDA: The annotation platform and all necessary documentation can be downloaded from the following webpage:
  - [http://www.exmaralda.org/en\\_downloads.html](http://www.exmaralda.org/en_downloads.html)

- Corpora: We will annotate parts from an interview with Edward Snowden broadcasted by the German TV station ARD:
  - video of the Snowden interview, YouTube link to the interview
  - transcription of the full Snowden interview
  - parts of the Snowden interview in EMARaLDA format
  - parts of the Snowden interview in simple text format

## Sessions:

- Monday: **Introduction to Information Structure**
  - Literature: Krifka (2007)
  - Slides: intro-slides.pdf
- Tuesday: **Focus, Givenness and Information Status**
  - Literature: Schwarzschild (1999), Baumann & Riester (2012)
  - Slides: givenness-infostatus.pdf
- Wednesday: **Annotating Information Status**
  - Literature: Riester & Baumann (2014)
  - Slides: lexical-infostatus.pdf
- Thursday: **Focus, Topic and Questions under Discussion**
  - Literature: Rooth (1992), Büring (2008), Wagner (2006), Büring (2003)
  - Slides: rooth-wagner-buering.pdf, qud-slides.pdf
- Friday: **Annotating Focus**

## References

- Baumann, S. & A. Riester (2012). Referential and lexical givenness: Semantic, prosodic and cognitive aspects. In G. Elordieta & P. Prieto (eds.), *Prosody and Meaning*, Berlin: Mouton de Gruyter, vol. 25 of *Interface Explorations*.
- Büring, D. (2008). What's New (and What's Given) in the Theory of Focus? In S. Berson, A. Bratkievich, D. Bruhn, R. E. Amy Campbell, A. Giovine, L. Newbold, M. P.-B. Marilola Perez & R. Rhomieux (eds.), *Proceedings of the Thirty-Fourth Annual Meeting of the Berkeley Linguistics Society*. Berkeley, CA: Berkeley Linguistics Society, pp. 403–424.
- Büring, D. (2003). On D-trees, beans, and B-accent. *Linguistics and Philosophy* 26(5), 511–545.
- Krifka, M. (2007). Basic Notions of Information Structure. In C. Fery, G. Fanselow & M. Krifka (eds.), *The notions of information structure*, Potsdam: Universitätsverlag Potsdam, vol. 6 of *Interdisciplinary Studies on Information Structure (ISIS)*.
- Riester, A. & S. Baumann (2014). *RefLex Scheme – Annotation Guidelines*.
- Rooth, M. (1992). A Theory of Focus Interpretation. *Natural Language Semantics* 1, 75–116.

Schwarzschild, R. (1999). GIVENness, Avoid F and other Constraints on the Placement of Focus. *Natural Language Semantics* 7(2), 141–177.

Wagner, M. (2006). Givenness and Locality. In *Proceedings of SALT XVI*. Ithaca: CLC Publications, vol. 16.

Last update: August 21, 2014