
Studying the prosodic properties of referring expressions in corpus resources: Obstacles and state of the art

Arndt Riester, Kerstin Eckart, Ina Rösiger, Antje Schweitzer,
Katrín Schweitzer, Sabrina Stehwien
Universität Stuttgart

{arndt.riester, kerstin.eckart, ina.roesiger, antje.schweitzer, katrin.schweitzer,
sabrina.stehwien}@ims.uni-stuttgart.de

Tag
Datum
Zeit
Raum

We provide a synopsis of long-standing efforts to create German spoken corpus resources – the DIRNDL corpus of radio news and the GRAIN corpus (part of the SFB732 silver standard collection, Eckart & Gärtner 2016) of radio interviews – that combine the annotation levels of information status (Riester & Baumann 2017), information structure (Riester et al. 2017), intonation – manual (DIRNDL) and automatic (GRAIN, Schweitzer 2010, Stehwien & Vu 2017) – and syntax. We are interested in the relation between prosody and referentiality beyond the widespread observation that new information is often, but not always, prosodically more prominent than given information. Major complications arise from three sources: (i) annotation quality (on all levels) is an issue which can strongly influence results. (ii) Syntactic complexity: new information tends to be expressed by longer phrases, and these may exhibit some form of tonal prominence for non-pragmatic reasons. Referring expressions can be nested inside each other. (iii) Information structure: focus typically receives sentence stress while backgrounded information is prosodically reduced. However, information can be backgrounded for reasons other than givenness, e.g. accommodation or parallelism.

References: • Eckart, K. & Gärtner, M. (2016): Creating silver standard annotations for a corpus of non-standard data. In: Proceedings of KONVENS, Bochum, 90–96. • Riester, A. & Baumann, S. (2017): The RefLex Scheme – Annotation Guidelines. SinSpeC 14, SFB 732, Universität Stuttgart. • Riester, A., Brunetti, L. & De Kuthy, K. (to appear) Annotation guidelines for Questions under Discussion and information structure. In: Adamou, E. et al. (eds.) Information Structure in Lesser-Described Languages: Studies in Syntax and Prosody. Benjamins. • Schweitzer, A. (2010): Production and Perception of Prosodic Events – Evidence from Corpus-based Experiments. PhD thesis. Universität Stuttgart. • Stehwien, S. & Vu, N. T. (2017): Prosodic event recognition using convolutional neural networks with context information. In: Proceedings of Interspeech, Stockholm, 2326–2330.