

Personality, Working Memory Capacity and Expert Manual Annotation of Spontaneous Speech

Anton Stepikhov, Saint Petersburg State University

Manual annotation of sentence boundaries in spontaneous speech is usually considered the gold standard of annotation. Yet even for expert annotators the sentence boundary placement is typically ambiguous. This ambiguity manifests itself in boundary placement in different positions in the text as well as in the different number of boundaries; i. e. in varying average sentence lengths. The reasons for such ambiguity have rarely been extensively explored. Nonetheless an understanding of them could reveal the interaction between the expert's annotation decisions and their individual characteristics such as memory capacity or personality traits.

The presentation reports the results of experiments aimed at revealing a possible relationship between expert manual annotation of spontaneous speech (German and Russian) and annotator's individuality. Mixed linear regression was used to evaluate whether variability in sentence length can be partially explained by the annotator's personality traits, working memory capacity and processing speed. The analysis enabled us to determine a factor of ambiguous segmentation affecting inter-annotator agreement in different languages.