

Technologies for Storing and Accessing Information about Syntactic and Prosodic Structures: the Potential Utility of CorpusSearch and the like.

Yelena Fainleib, University of Massachusetts Amherst
yfainlei@linguist.umass.edu

The talk will explore the options for storing and accessing syntactic and prosodic data using software for searching inside parsed corpora. Two tools that are used for these purposes will be covered: CorpusSearch (UPenn) and TregEx (Stanford University). Both programs allow searching sentences which are encoded in the format of labeled brackets, for tokens with certain types of syntactic structures. Therefore, it is possible that these tools could potentially provide quick and efficient access to syntactic and prosodic data collected from multiple participants. The talk will go over the format in which the input data must be encoded, the types of queries that the two programs allow and also their specific features.