Classifying ambiguous sentences and predicting their syntactic structure based on prosodic cues
S.J. Joustra

Abstract

Natural language processing machines often have trouble to identify the correct syntactic structure of sentences. In this thesis, prosodic annotation is used to help understand the structure of sentences and thereby disambiguate them. The research consists of two experiments, where the first has the aim to create a parser that combines a language corpus that is annotated both prosodically and syntactically, with a data-oriented syntax parser. This experiment successfully provides a toolkit to analyse sentence structure and classify for ambiguity. In the second experiment, the aim is to increase the accuracy of the best parse results, which is initially given by the data-oriented syntax parser. By analysing the location of prosodic symbols, adjoined by the syntactic parse trees, the result is that a certain amount of incorrect parse trees are successfully filtered. However, the numbers for improving the best parse tree selection only resulted in an approximate 1% increase in accuracy.