A Linguistics Annotation System

Tools to Facilitate Annotating Linguistic Corpora

The final three screenshots show an example of how a custom-designed application can facilitate the task of building an annotated corpus from a linguistics text. The components demonstrated here enable several strategies (which can be combined) for describing parsing structures and the logical forms ascribed to language samples:

S-Expressions Representing linguistic units as semantic and syntactic transformations triggered by words assigned to "functional" (lexical or Part of Speech) types.

Dependency Grammar Representing phrase structures via inter-word syntactic relationships.

Link Grammar Representing linguistic structure via connectors internal to each word-sense — inter-word links are activated when each word in the pair has a connector compatible with the other word's connector. Intuitively, connectors represent how one word's meaning or grammatical contribution can be "completed" by the phenomenon of its linking to another word.