Syntactic Identification of Occurrences of MWEs in Text using a Lexicon with Dependency Structures

Eduard Bejček, Pavel Straňák, Pavel Pecina
Charles University in Prague, Institute of Formal and Applied Linguistics

Hypothesis: If two deep syntactic subtrees look the same, they represent the same multiword expression.
Schema of the experiment

Syntactic

Identification of Occurrences of MWEs in Text using a Lexicon with Dependency Structures

- Basic Form: ...
- Lemmatised Form: ...
- Gloss: ...
- Source: John_Doe
- Synonyms: ...
- PDT22_FREQ: 2
- Tree Structure:

- Basic Form: ...
- Lemmatised Form: ...
- Gloss: ...
- Source: Jane_Smith
- Synonyms: ...
- PDT22_FREQ: 1
- Tree Structure: