

Towards Multilingual Constraint-Based Grammars

Previous grammars developed within constraint-based frameworks such as Head-driven Phrase Structure Grammar or Lexical Functional Grammar were concerned with individual natural languages such as English, Polish or Japanese. This paper presents an approach that allows us to deal with linguistic expressions in a multilingual context.

We assume that (i) every linguistic expression bears information concerning the language it belongs to and that (ii) a linguistic expression can belong to exactly one language. A language is understood here as a system of signs (in the sense of Saussurean signs) bearing the same language labels. A multilingual grammar is expected to license a set of signs with different language labels. Signs bearing the same language labels form subsets associated with the individual languages. Besides the information about the language a sign belongs to, every sign incorporates information about its phonological, morphosyntactic, semantic, discourse, and phrase-structural properties. Pieces of this information might be identical across languages. We assume that these identities between pieces of information embedded in distinct signs bearing different language labels reflect typological peculiarities.

This proposal is implemented within the formal framework of Head-driven Phrase Structure Grammar, which uses typed feature structures to represent (idealizations of) natural language expressions. The model of a multilingual grammar contains thus feature structures (admitted by the grammar) associated with expressions of all the languages covered by the grammar. This approach allows us to account for universal invariants and cross-linguistic variations model-theoretically.