# Grammatical Framework for implementing multilingual frames and constructions

Normunds Gruzitis[1,2], Dana Dannélls[3], Aarne Ranta[3,4], Francis M. Tyers[5]
[1]University of Latvia, [2]LETA
[3]University of Gothenburg, [4]Chalmers University of Technology
[5]UiT Norgga árktalaš universitehta
normunds.gruzitis@lumii.lv, dana.dannells@svenska.gu.se, aarne.ranta@cse.gu.se, francis.tyers@uit.no

Several approaches to Construction Grammar have been proposed. Remarkable examples include Sign-Based Construction Grammar that uses Head-Driven Phrase Structure Grammar as the underlying formalism, Fluid Construction Grammar, and Embodied Construction Grammar. For FrameNet-like resources of different languages, various annotation schemas are used as well.

We propose Grammatical Framework, GF (Ranta, 2004), as a unified formalism and a toolkit for implementing both computational frame-based grammars and computational construction grammars, allowing for seamless combination of both perspectives. We show that such grammars, as well as lexicons, can be extracted systematically and, thus, largely automatically from the existing semi-formal framenets and constructicons by extending the existing GF resource grammar library, RGL (Ranta, 2009). Moreover, we propose GF as a framework for implementing multilingual frame-based and construction grammars, currently testing our approach on English and Swedish (Gruzitis and Dannélls, 2015; Gruzitis et al., 2015), as well as Russian.

GF is characterized by its two-level approach to natural language representation. One level, the abstract syntax, accounts for the language-independent aspects, and the other level, the concrete syntax, accounts for the language-specific aspects. The same abstract syntax can be equipped with many concrete syntaxes – reversible mappings from abstract syntax trees to records (feature structures) and strings – making the grammar multilingual.

GF RGL currently supports 30 languages, all implementing the same abstract syntax. The built-in support for multilingual grammars has a great potential for implementing, unifying and interlinking frames and constructions of different languages, which, in turn, would be particularly beneficial for the use in machine translation, second-language learning and cross-lingual information extraction and summarization, involving controlled semantic parsing and natural language generation.

Our approach leverages FrameNet-annotated corpora to automatically extract a set of shared cross-lingual semantico-syntactic valence patterns. The implementation of the grammar and lexicon is supported by the design of FrameNet, providing a frame semantic abstraction layer, an interlingual semantic API, over the interlingual syntactic API already provided by GF RGL. The evaluation of the currently acquired bilingual grammar and lexicon shows the feasibility of the approach.

As a side result, we suggest a unified method for comparing and mapping semantic and syntactic valence patterns and lexical units across framenets. The systematic conversion also helps to debug and to improve the consistency of the original FrameNet and Constructicon resources. The major advantage is that language-dependent clause-level specifications to a large extent are hidden by the semantic frame-level API, making the potential application grammars more robust and flexible.

Regarding constructions, the resulting grammars (including the abstract syntax) are so far language-specific. A future task is to take this approach from the monolingual construction grammar to a multilingual one. Apart from direct mappings (where possible), this would require considering the links to FrameNet, as far as we deal with constructions with a referential meaning. The GF construction grammar and FrameNet grammar approaches are complementary to each other, and integrating them would be mutually beneficial.

**Keywords**: Grammatical Framework, FrameNet, Constructicon, Natural Language Generation

**Acknowledgements**. This work is supported in part by Swedish Research Council under grant agreement No. 2012-5746, Centre for Language Technology in Gothenburg, Latvian State Research Programme NexIT, and the H2020 project SUMMA under grant agreement No. 688139.

## References

Gruzitis N. and Dannélls D. A multilingual FrameNet-based grammar and lexicon for Controlled Natural Language. Journal of Language Resources and Evaluation, Springer, 2015 (online)

Gruzitis N., Dannélls D., Lyngfelt B., Ranta A. Formalising the Swedish constructicon in Grammatical Framework. Proceedings of the ACL Workshop on Grammar Engineering Across Frameworks, 2015

Ranta A. Grammatical Framework, a type-theoretical grammar formalism. Journal of Functional Programming, 14(2), 2004

Ranta A. The GF Resource Grammar Library. Linguistic Issues in Language Technology, 2(2), 2009