

Zur Motivation einer neuen Sprachressource für funktional orientierte kontrastive Grammatiken: Das CoMPaRS-Projekt

Beata Trawiński und Piotr Bański

Während die bisherige kontrastive Grammatikschreibung weitgehend auf einzelsprachlichen Referenzgrammatiken und Wörterbüchern sowie der Introspektion der AutorInnen basierte, bedienen sich die neueren kontrastiven Forschungsvorhaben immer häufiger korpuslinguistischer Methoden (Chesterman 1998, Granger et al. 2003, Granger 2010, James 1980, Johansson 1999, 2007). Dabei kommen insbesondere bi- und multilinguale Parallelkorpora zum Einsatz, die es mittlerweile für zahlreiche Sprachpaare einschließlich des Deutschen gibt. Obwohl diese Ressourcen bereits einen soliden empirischen Ausgangspunkt für die sprachübergreifende Forschung bieten, sind sie oft unzureichend, da die in diesen Ressourcen vorhandene linguistische Information zu spärlich (die großen multilingualen Parallelkorpora wie OPUS, EuroParl und JRC-Acquis sind lediglich tokenisiert, satzsegmentiert und satzaligiert) oder sprachspezifisch ist (so bei InterCorp und ParaSol). Darüber hinaus ermöglichen die existierenden multilingualen Parallelkorpora ausschließlich form-basierte Abfragen. Bedeutungs- bzw. funktionsbezogene Recherchen der folgenden Art sind derzeit in keinem Korpus möglich: Wie kann EXISTENZ / POSSESSION / TRANSITIVITÄT / PERFEKTIVITÄT / EINORDNUNG IN ZEIT UND RAUM etc. in den Sprachen L1 ... Ln ausgedrückt werden? Dies sind jedoch zentrale Fragestellungen einer typologisch orientierten kontrastiven Grammatik, die funktional-semantische Konzepte (*funktionale Domänen* im Sinne von Givón 1984; vgl. auch Zifonun 2001, 2009 und Kutscher 2014) als Vergleichsbasis verwendet – wie sie zurzeit im GDE-V-Projekt (Grammatik des Deutschen im europäischen Vergleich – Verbgrammatik) am IDS Mannheim erarbeitet wird.

In diesem Beitrag wird auf die Beschränkungen existierender Parallelkorpora eingegangen und das Konzept einer funktional orientierten Datenbank von Parallelsequenzen des Deutschen und anderer europäischen Sprachen präsentiert. Die Sprachressource wird im Rahmen des CoMPaRS-Projekts (**C**ollection of **M**ultilingual **P**arallel **S**equences) als Teil von GDE-V am IDS realisiert. Die Kernidee von CoMPaRS ist, dass alle (aus den existierenden Parallelkorpora extrahierten) Daten neben den sprachspezifischen morphosyntaktischen Annotationen nach sprachübergreifenden funktionalen Informationen annotiert und auf mehreren Ebenen (auch ebenenübergreifend) miteinander verlinkt werden. CoMPaRS wird in TEI P5 (TEI Consortium, 2016) implementiert und sowohl als monolinguale als auch als multilinguale Sprachressource modelliert. Abbildung 1 illustriert das Datenmodell von CoMPaRS und Abbildung 2 dessen intendierte Infrastruktur.

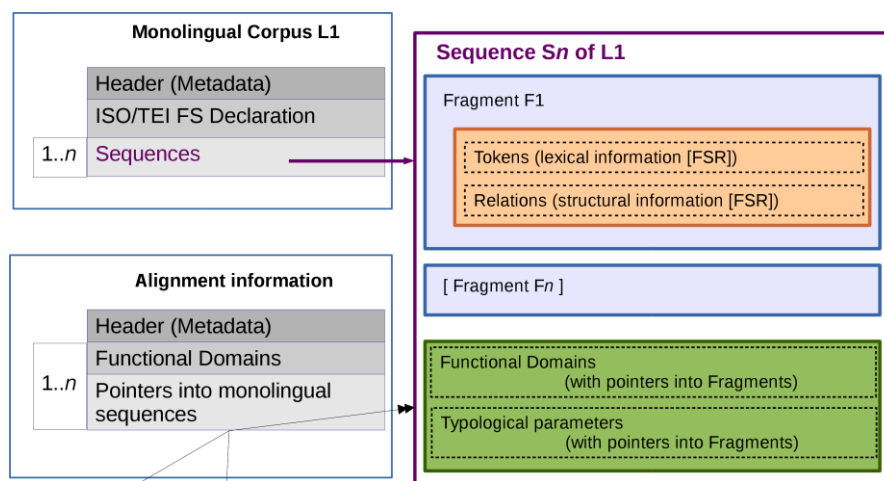


Abbildung 1 Datenmodell von CoMPaRS

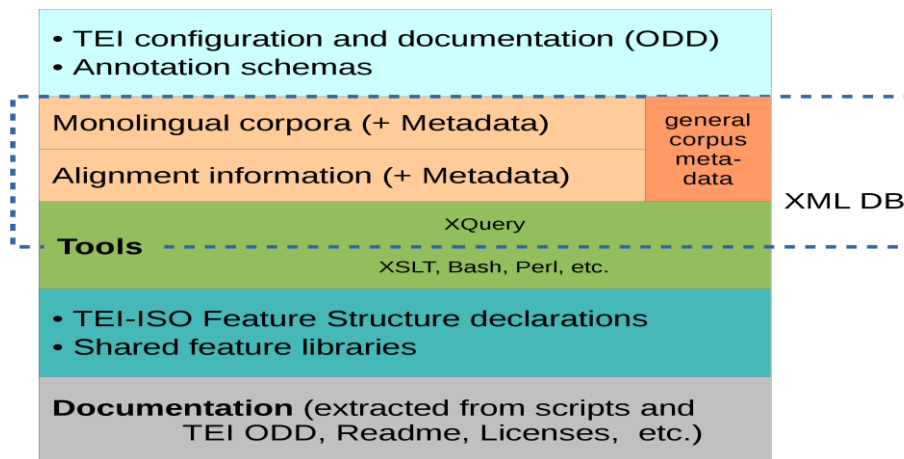


Abbildung 2 Intendierte Infrastruktur von CoMPaRS

Literatur:

- Chesterman, A. (1998): *Contrastive Functional Analysis*. Amsterdam/Philadelphia: John Benjamins Publishing Company.
- Givón, T. (1984): *Syntax. A functional-typological introduction*. Bd. 1. Amsterdam, Philadelphia (Pa.): Benjamins.
- Granger, S. / Lerot, J. / Petch-Tyson, S. (Hrsg.) (2003): *Corpus-based Approaches to Contrastive Linguistics and Translation Studies*. Amsterdam / Atlanta: Rodopi.
- Granger, S. (2010): *Comparable and translation corpora in cross-linguistic research. Design, analysis and applications*. In: *Journal of Shanghai Jiaotong University*, Bd. 2, S. 14-21.
- Kutscher, S. (2014): *Entwurf einer Makrostruktur zum Bereich der Sachverhaltsversprachlichung im Rahmen des Projekts Grammatik des Deutschen im europäischen Vergleich (GDE-V)*. Institut für Deutsche Sprache: Arbeitspapier der Abteilung Grammatik.
- James, C. (1980): *Contrastive Analysis*. London: Longman.
- Johansson, S. (1999): *Corpora and contrastive studies*. In: Pietilä, P. / Salo, O.-P. (Hrsg.), *Multiple Languages – Multiple Perspectives*. AFinLA Yearbook Nr. 57, S. 116-125.
- Johansson, S. (2007): *Seeing through multilingual corpora. On the use of corpora in contrastive studies*. Amsterdam: Benjamins.
- TEI Consortium (2016): *TEI P5: Guidelines for Electronic Text Encoding and Interchange, ver. 3.0.0*. <http://www.tei-c.org/Guidelines/P5/>. Last updated on 29th March 2016.
- Zifonun, G. (2001): *Grammatik des Deutschen im europäischen Vergleich*. In: *Studia linguistica* 20, S. 171–186.
- Zifonun, G. (2009): *Zur pragmatischen Fundierung des Konzepts der funktionalen Domäne im Sprachvergleich*. In: Brdar-Szabó, R./Knipf-Komlósi, E./Péteri, A. (Hrsg.): *An der Grenze zwischen Grammatik und Pragmatik*. Frankfurt/M., Berlin, Bern, Bruxelles, New York, Oxford, Wien: Lang (*Deutsche Sprachwissenschaft international*, Bd. 3), S. 243–252.