

## DEUTSCH IN EUROPA

Sprachpolitisch, grammatisch, methodisch

### METHODENMESSE

Mittwoch, 11. März 2020, 16:00 Uhr

#### **CoMPaRS: EINE SAMMLUNG VON MULTILINGUALEN PARALLELSEQUENZEN DES DEUTSCHEN UND ANDERER EUROPÄISCHER SPRACHEN**

Beata Trawiński, Susan Schlotthauer und Piotr Bański (IDS)

Seit der empirischen Wende bedienen sich nicht nur einzelsprachliche, sondern auch sprachübergreifende Forschungsvorhaben immer häufiger korpuslinguistischer Methoden. Dabei kommen insbesondere bi- und multilinguale Parallel- und Vergleichskorpora zum Einsatz, die es mittlerweile für zahlreiche Sprachen einschließlich des Deutschen gibt. Dieser Beitrag präsentiert die neue multilinguale Ressource CoMPaRS (Collection of Multilingual Parallel Sequences), die im Rahmen des am IDS laufenden Projekts Grammatik des Deutschen im europäischen Vergleich (GDE) erstellt wird. CoMPaRS ist als eine funktional-semantisch orientierte Datenbank von Parallelsequenzen des Deutschen und anderer europäischer Sprachen (darunter Englisch, Französisch, Italienisch, Polnisch, Ungarisch) zu verstehen, die primär aus den existierenden Parallelkorpora extrahierte Daten und darüber hinaus konstruierte, introspektionsbasierte Daten enthält. Die Kernidee von CoMPaRS ist, dass alle Daten neben den sprachspezifischen und universellen (im Sinne von Universal Dependencies) morphosyntaktischen Annotationen auch nach sprachübergreifenden funktional-semantischen Informationen auf der neudefinierten Annotationsebene Funktionale Domänen annotiert und auf mehreren Ebenen (auch ebenenübergreifend) miteinander verlinkt werden. Das für diese Zwecke entwickelte Tagset reflektiert die Ergebnisse der laufenden Forschung im GDE-Projekt. CoMPaRS wird in TEI P5 XML kodiert und sowohl als monolinguale als auch als multilinguale Sprachressource modelliert.

#### **Literatur:**

- Bański, Piotr (erscheint demnächst): The logical architecture of CoMPaRS and its XML implementation. In: Beata Trawiński & Angelika Wöllstein (Hrsg.) (in Vorbereitung): Pilotstudien im Projekt GDE-V (Arbeitstitel). Berlin/Boston: De Gruyter (= Konvergenz und Divergenz).
- Bański, Piotr/Kamocki, Paweł/Trawiński, Beata (2017): Legal canvas for a patchwork of multilingual quotations. The case of CoMPaRS. Corpus Linguistics 2017 Conference, University of Birmingham, 25-28 July 2017. <http://www.birmingham.ac.uk/Documents/college-artslaw/corpus/conference-archives/2017/general/paper49.pdf>
- Granger, Sylviane/Lerot, Jacques/Petch-Tyson, Stephanie (Hrsg.) (2003): Corpus-based approaches to contrastive linguistics and translation studies. Amsterdam/Atlanta: Rodopi.
- Gunkel, Lutz/Murelli, Adriano/Schlotthauer, Susan/Wiese, Bernd/Zifonun, Gisela (2017): Grammatik des Deutschen im europäischen Vergleich. Das Nominal. Unter Mitarbeit von Christine Günther und Ursula Hoberg. 2 Teilbände. Berlin/Boston: De Gruyter (= Schriften des Instituts für Deutsche Sprache 14).
- Trawiński, Beata (2016): Linguistic data in contrastive studies. Addressing the need for a multilingual parallel resource annotated with semantic-functional information. In: María José Domínguez Vázquez und Silvia Kutscher (Hrsg.): Interacción entre gramática, didáctica y lexicografía. Estudios contrastivos y multicontrastivos. Berlin/Boston: De Gruyter. S. 85–98.