

ZUM KONZEPT EINER TYPOLOGISCH ORIENTIERTEN DATENBANK VON PARALLELSEQUENZEN DES DEUTSCHEN UND ANDERER EUROPÄISCHER SPRACHEN

Beata Trawiński

Institut für Deutsche Sprache, Mannheim

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 korpuslinguistischen Methoden (Chesterman 1998, Granger et al. 2003, Granger 2010, James 1980, Johansson 1999, 2007). Dabei kommen insbesondere bi- und multilinguale Vergleichs- und Parallelkorpora zum Einsatz, die es mittlerweile für zahlreiche Sprachpaaren einschließlich Deutsch gibt. Obwohl diese Ressourcen bereits einen soliden empirischen Ausgangspunkt für die sprachübergreifende Forschung bieten, sind sie aufgrund der Spärlichkeit an vorhandener linguistischer Information (die großen multilingualen Parallelkorpora wie OPUS, EuroParl und JRC-Acquis sind lediglich tokenisiert, satzsegmentiert und satzaligniert) oder dessen sprachspezifischen Charakters (InterCorp) oft unzureichend. Darüber hinaus ermöglichen die existierenden multilingualen Parallelkorpora ausschließlich form-basierte Abfragen wie beispielsweise:

- Was sind die Äquivalente von Lemma / Wortform / Phrase W in Sprache L_1 in den Sprachen $L_2 \dots L_n$?
- Kann Lemma / Wortform / Phrase W_1 in Sprache L_1 durch Lemma / Wortform / Phrase W_2 in Sprache L_2 ausgedrückt werden?
- Wie können Ketten von grammatischen Kategorien $K_1 \dots K_n$ in Sprache L_1 in Sprache L_2 ausgedrückt werden?
- Wie können Ausdrücke mit den grammatischen Merkmalen $M_1 \dots M_n$ in Sprache L_1 in Sprache L_2 ausgedrückt werden?

Bedeutungs- bzw. funktionsbezogene Anfragen der folgenden Art sind derzeit in keinem Korpus möglich:

- Wie kann EXISTENZ / POSSESSION / TRANSITIVITÄT / PERFEKTIVITÄT / DIE EINORDNUNG XY IN ZEIT UND RAUM etc. in Sprachen $L_1 \dots L_n$ 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 i.E.) als Vergleichsbasis verwendet und wie sie zurzeit im GDE-V-Projekt am IDS er-

arbeitet wird. Konzepte dieser Art sind auch für die Fremdsprachendidaktik von großem Gewinn, wie im Vortrag deutlich werden wird.

Im Vortrag wird auf die Beschränkungen existierender Parallelkorpora eingegangen und das Konzept einer typologisch orientierten Datenbank von Parallelsequenzen des Deutschen und anderer europäischen Sprachen präsentiert. Die Datenbank soll einerseits als empirische Grundlage der GDE-V-Grammatik fungieren und andererseits dem Benutzer (u.a. Germanisten, Allgemeine Sprachwissenschaftler, Fremdsprachendidaktiker) einen besseren Zugang zu abstrakten grammatischen Inhalten und zur Sprache selbst durch zahlreiche, natürlich vorkommende Sprachbeispiele ermöglichen.

Die Kernidee der geplanten Datenbank ist, dass alle Daten neben der sprachspezifischen morphosyntaktisch-semantischen Annotationen nach sprachübergreifenden funktionalen Domänen und Varianzparametern annotiert und auf mehreren Ebenen (auch ebenenübergreifend) aligniert sind. Die Annotation der Daten soll in XML nach den TEI 5-Richtlinien, möglicherweise mit einigen Erweiterungen, erfolgen. Die zugrundeliegende Datenstruktur soll die Durchsichtung der erfassten Parallelsequenzen nach unterschiedlichen Kriterien, und insbesondere nach funktionalen Domänen ermöglichen.

- Chesterman, A. 1998. *Contrastive Functional Analysis*. Amsterdam/Philadelphia: John Benjamins Publishing Company.
- Givón, Talmy 1984. *Syntax. A functional-typological introduction*. Bd. 1. Amsterdam, Philadelphia [Pa.]: J. Benjamins.
- Granger, S. / Lerot, J. / Petch-Tyson, S. (Hg.) 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*.
- Kutscher, S. (i.E.). *Entwurf einer Makrostruktur zum Bereich der Sachverhaltsversprachlichung im Rahmen des Projekts Grammatik des Deutschen im europäischen Vergleich (GDE-V)*. (Ms. Institut für Deutsche Sprache). Erscheint als Arbeitspapier der Abteilung Grammatik.
- James, C. 1980. *Contrastive Analysis*. London: Longman.
- Johansson, S. 1999. Corpora and contrastive studies. In: P. Pietilä & O-P. Salo (Hs.). *Multiple Languages – Multiple Perspectives*. AFinLA Yearbook 1999 / No. 57: 116-125.
- Johansson, S. 2007. *Seeing through multilingual corpora. On the use of corpora in contrastive studies*. Amsterdam: Benjamins.
- Zifonun, Gisela 2001. *Grammatik des Deutschen im europäischen Vergleich*. In: *Studia linguistica* 20: S. 171–186.
- Zifonun, Gisela 2009. Zur pragmatischen Fundierung des Konzepts der funktionalen Domäne im Sprachvergleich. In: Rita Brdar Szabó, Elisabeth Knipf-Komlósi und Attila Péteri (Hg.). *An der Grenze zwischen Grammatik und Pragmatik*. Frankfurt, M, Berlin, Bern, Bruxelles, New York, NY, Oxford, Wien: Lang (Deutsche Sprachwissenschaft international, Bd. 3), S. 243–252.