

**Deutsch-ungarischer Sprachvergleich**  
*Zwischenbericht vom Lernerkorpus Dulko*

30.10.2018, 11.00 Uhr  
Vortragssaal, IDS Mannheim

Dr. Andreas Nolda  
<http://andreas.nolda.org>  
Universität Szeged

Seit 2017 besteht zwischen der Abteilung Grammatik des IDS und dem Lehrstuhl für Germanistische Linguistik an der Universität Szeged eine von der Alexander-von-Humboldt-Stiftung geförderte Institutspartnerschaft namens „Deutsch-ungarischer Sprachvergleich: korpustechnologisch, funktional-semantisch und sprachdidaktisch (DeutUng)“ unter Einbeziehung des Programmbereichs Korpuslinguistik am IDS, des Lehrstuhls für Korpuslinguistik und Morphologie an der Humboldt-Universität zu Berlin sowie der Ungarischen Akademie der Wissenschaften. Im Szegeder Teilprojekt wird ein deutsch-ungarisches Lernerkorpus namens „Dulko“ mit kontrolliert erhobenen Texten ungarischer Deutschlerner erstellt. Diese Lernertexte werden im Dulko-Projekt unter anderem nach Wortarten, Lemmata, Zielhypothesen und Fehlerkategorien annotiert und sollen nach Abschluss der Annotation auf der AN-NIS-Plattform (<https://korpling.german.hu-berlin.de/falko-suche/>) für die Zweitspracherwerbsforschung verfügbar gemacht werden. Im Fokus des empirischen Interesses stehen dabei Interferenzfehler und andere Transferphänomene.

Das im Dulko-Projekt eingesetzte Annotationsverfahren orientiert sich an Standards der Falko-Lernerkorpora, geht aber in wesentlichen Punkten darüber hinaus. Dazu gehören eine flexible, verallgemeinerte Konzeption von Zielhypothesen sowie ein mehrdimensionales Tagset für eine explizite, grammatisch motivierte Fehlerannotation. Als Annotationstool verwendet das Dulko-Projekt den EXMARaLDA-Partitureditor (<http://www.exmaralda.org>) mit projektspezifischen Transformationsszenarien in Form von XSLT-Stylesheets, die an der Universität Szeged entwickelt und unter dem Namen „EXMARaLDA (Dulko)“ als Open Source veröffentlicht wurden (<https://bitbucket.org/nolda/exmaralda-dulko>). Mit Hilfe dieser Transformationsszenarien können (Lerner-)Texte automatisch tokenisiert und mit Wortarten, Lemmata und Satzspannen getaggt werden; außerdem unterstützt EXMARaLDA (Dulko) den Annotator bei der Formulierung von Zielhypothesen und der Annotation von Abweichungen mit Hilfe von Fehlerkategorien. Annotationsverfahren und Implementierung wurden auf der Methodenmesse der IDS-Jahrestagung 2018 einer breiteren Öffentlichkeit präsentiert (Hirschmann und Nolda im Ersch.) und im Mai 2018 mit dem Innovationspreis der Universität Szeged ausgezeichnet. Mittlerweile werden sie auch in Lernerkorpus-Projekten in Deutschland und weiteren Ländern eingesetzt.

Der Vortrag stellt das Dulko-Annotationsverfahren und die dafür implementierte Software am Beispiel ausgewählter Lernerdaten aus dem Dulko-Projekt vor und diskutiert diese Daten im Hinblick auf mögliche Transferphänomene.

### **Literatur**

Hirschmann, Hagen und Andreas Nolda (im Ersch.). Dulko – auf dem Weg zu einem deutsch-ungarischen Lernerkorpus. In *Jahrbuch des Instituts für Deutsche Sprache 2018*, Berlin: de Gruyter.