Die Integration und Auswertung von Daten neuzugewanderter Schüler:innen in NaLeKo

Anja Binanzer[1], Hagen Hirschmann[2] & Miriam Langlotz[3]
[1]*Technische Universität Dresden,* [2]*Humboldt-Universität zu Berlin,*
[3]*Universität Kassel*
anja.binanzer@tu-dresden.de, hagen.hirschmann@hu-berlin.de,
m.langlotz@uni-kassel.de

Der Spracherwerb bei neuzugewanderten Schüler:innen ist ein bislang vernachlässigtes Forschungsgebiet (vgl. u.a. Goltsev & von Dewitz 2023, 268). In unserem Vortrag widmen wir uns der Frage, welche spezifischen Spracherwerbsherausforderungen bei neuzugewanderten Schüler:innen (sL2) im Vergleich zu Schüler:innen mit Deutsch als Erstsprache (L1) und als frühkindlich erworbene Zweitsprache (fL2) gelten. Wir nähern uns dieser Frage mit einem korpusbasierten Ansatz, indem wir Vergleichsdaten von sL2-Schüler:innen im Lernerkorpus NaLeKo integrieren und diese kontrastiv zu den bereits vorhandenen Lernergruppen (L1 und fL2) auf lernersprachliche Varianten auswerten.

NaLeKo (NArratives LErnerKOrpus, Hirschmann, Binanzer & Langlotz 2023) ist ein frei nutzbares Lernerkorpus mit schriftsprachlichen Erzähltexten von ein- und mehrsprachig in Deutschland aufgewachsenen Schüler:innen. Die bisher zugänglichen Textdaten, vorrangig erstellt, um die Entwicklung von narrativen Fähigkeiten von Schüler:innen zu untersuchen (vgl. Binanzer & Langlotz 2019, Binanzer, Hirschmann & Langlotz 2022), werden seit 2023 mit Vergleichsdaten neuzugewanderter Schüler:innen angereichert. Bis heute liegen Datensätze von 19 sL2-Schüler:innen (Klassenstufen 5 bis 11) mit den Erstsprachen Ukrainisch und Russisch vor; weitere Datenerhebungen dieser Art sind in Vorbereitung. Die Daten umfassen wie die meisten anderen in NaLeKo verfügbaren Datensätze pro Proband:in je zwei narrative Texte, C-Tests zur Erfassung des Sprachstands in der Zielsprache Deutsch sowie sprachbiografische Hintergrundinformationen zu den Schüler:innen (Fragebogen). Die Korpusannotationen umfassen neben standardmäßigen Verarbeitungen (Wortartentagging und Lemmatisierung) die Ausweisung von lernersprachlichen Varianten (Normabweichungen) durch die Formulierung von Zielhypothesen und die Eintragung von Abweichungskürzeln (vgl. Beh, Drewnowska-Vargáné, Kappel et al., 32–55) sowie die multidimensionale Annotation von Junktoren und die Annotation von Teilsätzen, Verbtypen und Tempuskonstruktionen.

Im Vortrag stellen wir zunächst die Vorgehensweise bei den Datenerhebungen und der Datenannotation dar und erläutern die Forschungsmöglichkeiten, die aus dieser Arbeit resultieren. Im zweiten Teil präsentieren wir eine Spracherwerbsstudie zum Vergleich lernersprachlicher Variation der drei Subgruppen L1-, fL2 und sL2.

References: Die Referenzen sind unter https://hu.berlin/AG14-NaLeKo-lit zu finden.