

KORPORA IN DER GERMANISTISCHEN SPRACHWISSENSCHAFT – MÜNDLICH, SCHRIFTLICH, MULTIMEDIAL

Donnerstag, 17. März 2022, 10:10 Uhr

Akustisches Signal, Mehrebenenannotation und Aufgabendesign. Flexible Korpusarchitektur als Voraussetzung für die Wiederverwendung gesprochener Korpora

Carolin Odebrecht, Malte Belz (HU Berlin)

Eine Korpusarchitektur muss – insbesondere für Korpora nicht-standardisierter Sprache – verschiedene epistemische Anknüpfungspunkte für unterschiedliche Forschungskontexte bieten, wenn sie die Weiterentwicklung und Wiederverwendung von Korpora ermöglichen soll. In unserem Vortrag diskutieren wir drei besonders relevante Korpusarchitekturkomponenten für gesprochene Korpora: die Bereitstellung eines akustischen Signals, die konzeptionell und technisch damit verbundene Annotation in einem Mehrebenenmodell und das Aufgabendesign. Zudem stellen wir methodische Überlegungen zur Normalisierung und zum Forschungsdatenmanagement vor, die für iterative Verbesserungen und für die Wiederverwendung relevant sind. Exemplarisch hierfür ziehen wir das gesprochene Lernerkorpus WroDiaCo heran (Wesolek et al. 2021), welches für eine studentische Abschlussarbeit erhoben (Version 1) und schon für eine andere linguistische Fragestellung zu Normalisierungsmechanismen wiederverwendet wurde (Version 2, vgl. Belz & Odebrecht [angenommen]). Die methodischen Überlegungen werden anhand einer Forschungsfrage zum Fremdsprachnakzent (Vokalrealisierung polnischer Deutschlerner*innen) illustriert.

Literatur:

Belz, Malte & Carolin Odebrecht. [angenommen]. Abschnittsweise Analyse sprachlicher Flüssigkeit in der Lernaltersprache: Das Ganze ist weniger informativ als seine Teile. Zeitschrift für germanistische Linguistik.

Wesolek, Sarah, Malte Belz & Christine Mooshammer. 2021. Wroclaw Dialogue Corpus (WroDiaCo): Version 2. Medien-Repositoryum. <<https://rs.cms.hu-berlin.de/phon>>.