

58. Jahrestagung des Leibniz-Instituts für Deutsche Sprache vom 15. bis 17. März 2022
(als Online-Konferenz)

KORPORA IN DER GERMANISTISCHEN SPRACHWISSENSCHAFT – MÜNDLICH, SCHRIFTLICH, MULTIMEDIAL

METHODENMESSE

Mittwoch, 16. März 2022, 15:45 Uhr bis 17:45 Uhr

Erhebung, Aufbereitung und (kollaborative) Nutzung des Korpus „Deutsch in Österreich. Variation - Kontakt – Perzeption“

Jakob Bal, Johanna Fanta-Jende, Katharina Korecky-Kröll, Katharina Kranawetter,
Markus Pluschkovits, Rita Stiglbauer, Rebecca Stocker, Florian Tavernier &
Anja Wittibschlager

Der Spezialforschungsbereich (FWF F60) „Deutsch in Österreich“ (im Folgenden „DiÖ“, siehe für einen Überblick Lenz 2018 oder Budin et al 2018) gliedert sich in drei wissenschaftliche Task-Cluster („Variation“, „Kontakt“, „Perzeption“), die von einem technischen Task-Cluster („Forschungsplattform“) unterstützt werden. Auf Basis der kollaborativen Projektarbeit ist ein umfassendes Sprachkorpus entstanden, dessen Methoden und Tools im Rahmen der Tagung vorgestellt werden sollen. Im Fokus stehen dabei die linguistischen und technischen Zugänge zu Datenerhebung, -aufbereitung und -analyse sowie die Zugänglichkeit für projektexterne Anwendungen und die Übertragbarkeit auf zukünftige Forschungsvorhaben (alle Tools werden auf Open-Source-Basis entwickelt und stehen in den GitHub-Repositoryen des DiÖ-Projekts zur Verfügung, siehe DiÖ 2021).

Die linguistische Grundlage der Präsentation stellt ein umfassendes gesprochen sprachliches Korpus dar, das sich an den Kernthemen der Teilprojekte PP03 „Sprachrepertoires und Varietätenspektren“ und PP08 „Standardvarietäten aus Perspektive der perzeptiven Variationslinguistik“ orientiert. Zwischen 2016 und 2019 wurden ca. 570 Stunden Audiomaterial von rund 150 Personen unterschiedlichen Alters und Geschlechts aus ganz Österreich aufgezeichnet. Um möglichst viel über die inter- und intraindividuelle Variation in Österreich zu erfahren, wurden die Personen mit einem breiten Mix an Erhebungsmethoden konfrontiert: Kontrollierte Settings umfassten mündliche Übersetzungsaufgaben aus dem Dialekt in das Standarddeutsche und umgekehrt, multimodale Sprachproduktionsexperimente im Standard und im Dialekt sowie Bildbenennungs- und Vorleseaufgaben (mit Fokus auf dem standardnahen Pol). Freie Settings zielten auf die Spontansprache ab: Forscher*innen führten mit den Gewährspersonen soziolinguistische Interviews, um den standardnahen Spontansprachegebrauch sowie die Spracheinstellungen in einem formellen Kontext zu elizitieren. Demgegenüber sollten Freundesgespräche, in denen Gespräche der Gewährspersonen mit jeweils einer befreundeten Person aus demselben Ort erhoben wurden, dialektnahe spontansprachliche Register und Spracheinstellungen in einem informellen Setting abdecken.

Für den Umgang mit dieser heterogenen Datengrundlage wurden mehrere technische Tools im Rahmen des DiÖ-Projekts entwickelt. Die Transkriptionssoftware *SpeechHouse* (bisheriger Arbeitstitel *transcribe*, Graf et al. in Vorbereitung) ist ein plattformunabhängiges, multi-user taugliches und token-basiertes Transkriptionstool. Das Tool orientiert sich an den Prinzipien des User-Centered Designs (vgl. Chammas / Quaresma / Mont'Alvão 2015) und ermöglicht Nutzer*innen ein flüssiges, zeitaligniertes und gleichzeitiges Transkribieren von Sprachdaten.

Es unterstützt außerdem die Transkription auf verschiedenen Tiers, um Ansprüchen unterschiedlicher linguistischer Systemebenen (z. B. phonetische Transkription, lautorientierte Transkription für Syntax usw.) nachzukommen. Darauf folgt die linguistische Annotation in der projekteigenen Datenbank, die ebenfalls auf die unterschiedlichen Forschungsinteressen abzielt.

Für den anschließenden Schritt der Datenauswertung soll ein sich in Entwicklung befindliches Kartierungstool (Alpha-Fassung mit Hilfe des Vue.js Framework, vgl. You 2020) vorgestellt werden. Dieses dient einerseits den Linguist*innen des DiÖ-Projekts als Werkzeug zur Visualisierung des annotierten georeferenzierten Sprachmaterials im Raum und soll andererseits einen niederschweligen Zugang zum Korpus für ein interessiertes Lai*innenpublikum ermöglichen. Dabei steht der explorative Zugang im Vordergrund, der es Nutzer*innen erlauben soll, mit den Karten dynamisch zu interagieren und aktuelle Fragen der Variationslinguistik in Österreich auf Basis der regionalen Verortung und einer entsprechenden Verknüpfung zu den Audiobeispielen kennenzulernen.

Schließlich soll mit Abschluss des Projekts 2023 eine geplante Online-Forschungsplattform das gesamte Datenmaterial allen Interessierten öffentlich zugänglich gemacht werden. Die Präsentation bei der Methodenmesse bietet also einen Vorgeschmack auf die zu erwartenden umfassenden Möglichkeiten der Datenerfassung, -aufbereitung und -auswertung zum Deutschen in Österreich.

Literatur:

- Budin, Gerhard, Stephan Elspaß, Alexandra N. Lenz, Stefan Michael Newerkla & Arne Ziegler (2018). Der Spezialforschungsbereich „Deutsch in Österreich (DiÖ). Variation – Kontakt – Perzeption“. Zeitschrift für germanistische Linguistik 46(2). <https://doi-org.uaccess.univie.ac.at/10.1515/zgl-2018-0017> [Abruf 12. September 2021].
- Chammas, Adriana / Quaresma, Manuela / Mont'Alvão, Cláudia. 2015." A Closer Look on the User Centred Design". *Procedia Manufacturing* 2015/3, 5397–5404.
- DiÖ. 2021. Github Repositorien des Spezialforschungsbereich Deutsch in Österreich. O.O. <https://github.com/german-in-austria> (Abruf 12. September 2021).
- Graf, Arnold, Ludwig Maximilian Breuer, Tahel Singer & Markus Pluschkovits. in Vorbereitung. *Transcribe: a web-based linguistic transcription tool*. O.O.
- Lenz, Alexandra N. (2018). The Special Research Programme "German in Austria: Variation – Contact – Perception". *Sociolinguistica* 32(1), 269-277.
- You, Evan (2020): Vue.Js. The Progressive JavaScript Framwork. URL: <https://vuejs.org> [Abruf 1. Februar 2021].