

57. Jahrestagung des Leibniz-Instituts für Deutsche Sprache vom 9. bis 11. März 2021
(als Online-Konferenz)

SPRACHE IN POLITIK UND GESELLSCHAFT

Perspektiven und Zugänge

METHODENMESSE

Dienstag, 09. März 2021, 09:40 Uhr

Das Archiv für Gesprochenes Deutsch und das Forschungs- und Lehrkorpus für Gesprochenes Deutsch – Audiovisuelle Dokumentation von Sprachgebrauch in Gesellschaft und Politik

Silke Reineke / Thomas Schmidt

Im Archiv für Gesprochenes Deutsch (AGD) werden Korpora und Sammlungen archiviert, die mündlichen Sprachgebrauch in verschiedensten gesellschaftlichen Kontexten über einen Zeitraum von den 1950er Jahren bis heute dokumentieren. Darunter zählen etwa das Korpus "Belgische TV-Debatten" (2012), ein Korpus der Schlichtungsgespräche zu Stuttgart 21 (2010) das Korpus "Ausbildung im Bergbau" (1980er), das Korpus "Jugendkommunikation" (1998 bis 2000) oder die Bonner Längsschnittstudie des Alterns (1960er bis 1980er).

Viele Daten aus dem AGD werden online über die Korpus-Plattform "Datenbank für Gesprochenes Deutsch" (dgd.ids-mannheim.de) zur Verfügung gestellt und sind so für Online-Browsing und Online-Recherchen zu unterschiedlichen Forschungszwecken zugänglich. Aktuell bietet das AGD 40 für eine Nachnutzung aufbereitete Korpora an, davon sind knapp 4.000 Stunden Audio-Aufnahmen und 200 Stunden Videoaufnahmen sowie Transkripte im Umfang von über 17 Millionen Tokens in der DGD zugänglich.

Zu den online verfügbaren Korpora gehört auch das „Forschungs- und Lehrkorpus Gesprochenes Deutsch“ (FOLK), das Audio- und Videoaufnahmen natürlicher Interaktionen aus unterschiedlichsten Bereichen des gesellschaftlichen Lebens und somit auch aus politischer Kommunikation (z.B. öffentliche Diskussionen mit Bürgerbeteiligung, eine Bundestagsdebatte) enthält. Das Korpus adressiert primär Forschende und Lehrende aus den Bereichen der Gesprächsforschung und Korpuslinguistik, die Daten können jedoch für viele weitere Ansätze sinnvoll genutzt werden. Die Daten in FOLK werden nach zeitgemäßen Standards erschlossen (transkribiert, annotiert und dokumentiert). Annotationen der Daten auf vier Ebenen (Transkription, Normalisierung, Lemmatisierung, Part-of-Speech-Tagging) ermöglichen gezielte Recherchen auf den insgesamt 2.7 Millionen transkribierten Wörtern.

Auf der Projektmesse sollen die Projekte AGD und FOLK einerseits mit ihren Ressourcen und andererseits als Kooperationspartner für Datenerhebungen und Archivierung vorgestellt werden.

Literatur:

Kaiser, Julia (2018): Zur Stratifikation des FOLK-Korpus: Konzeption und Strategien. In: Gesprächsforschung 19, 515—552.

Deppermann, Arnulf/Hartung, Martin (2011): Was gehört in ein nationales Gesprächskorpus? Kriterien, Probleme und Prioritäten der Stratifikation des "Forschungs- und Lehrkorpus Gesprochenes Deutsch" (FOLK) am Institut für Deutsche Sprache (Mannheim). In: Felder et al. (Hg.): Korpuspragmatik. Berlin: de Gruyter, 414—450.

Stift, Ulf-Michael/Schmidt, Thomas (2014): Mündliche Korpora am IDS: Vom Deutschen Spracharchiv zur Datenbank für Gesprochenes Deutsch. In: Institut für Deutsche Sprache (Hrsg.): Ansichten und Einsichten. 50 Jahre Institut für Deutsche Sprache. Redaktion: Melanie Steine, Franz Josef Berens. S. 360—375. Mannheim: Institut für Deutsche Sprache.