Lexik des gesprochenen Deutsch

Projektbeschreibung

Das Projekt „Lexik des gesprochenen Deutsch“ (LeGeDe) ist ein Drittmittelprojekt im Leibniz-Wettbewerb 2016 (Förderlinie: Innovative Vorhaben) und wurde zwischen 2016 und 2019 am IDS als Kooperationsprojekt der Abteilungen Lexik und Pragmatik durchgeführt. In beiden Abteilungen sind in den letzten Jahren die Kompetenzen gereift, die es zum ersten Mal zulassen, ein korpusbasiertes Wörterbuch des gesprochenen Deutsch zu erstellen, das in dieser Form sowohl für das Deutsche als auch für andere Sprachen ein Novum darstellt.

Folgende Rahmenbedingungen konnten das Vorhaben ermöglichen:

(i) Mit dem am IDS erstellten „Forschungs- und Lehrkorpus Gesprochenes Deutsch“ (FOLK) existiert eine geeignete Datengrundlage, die für die Analysen genutzt werden konnte.

(ii) Das LeGeDe-Projekt konnte über Projektkooperationen und aktiven Austausch mit der Abteilung Pragmatik auf die vorhandenen Erfahrungen in der semantisch-funktionalen Analyse gesprochener Sprache in der Interaktion zurückgreifen.

(iii) In den Arbeitsbereichen „Computerlexikographie“ und „Empirische Methoden“ konnte über die letzten Jahre eine ausgewiesene Expertise in der Konzeption und Realisierung komplexer multimedialer Internetwörterbücher in Verbindung mit der Entwicklung entsprechender empirischer Methoden aufgebaut werden. Die LeGeDe-Ressource knüpft an diese Erfahrungen an.

Datengrundlage und Auswahl der Stichwörter

Die Datengrundlage der Wörterbuchartikel im LeGeDe-Prototyp ist FOLK (ca. 250h Aufnahmen/2,4 Mio. Tokens; Stand: September 2019), das am IDS als das erste große Gesprächskorpus des Deutschen entwickelt wurde und stetig weiter ausgebaut wird. FOLK ist über die „Datenbank für Gesprochenes Deutsch“ (DGD) abrufbar und beinhaltet Gesprächsaufnahmen und Transkripte im deutschsprachigen Raum aus unterschiedlichen privaten, institutionellen und öffentlichen Kontexten und bietet vielfältige Optionen der Erschließung mündlicher Daten nach linguistischen und interaktionalen Merkmalen.

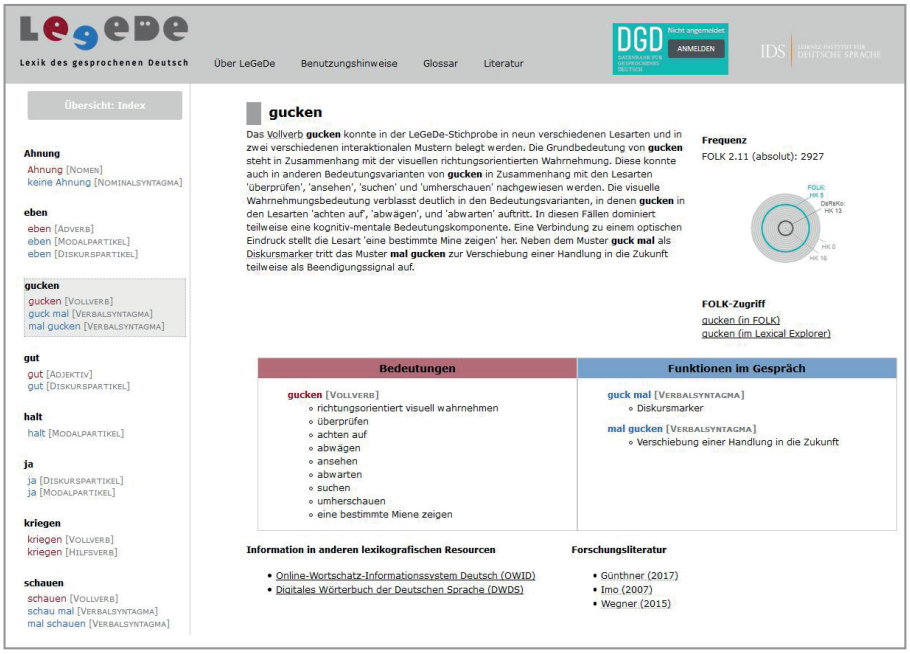
Die Auswahl der Stichwortkandidaten für die LeGeDe-Ressource orientierte sich an einem quantitativen Vergleich zwischen FOLK und Teilen aus dem „Deutschen Referenzkorpus“ (DeReKo), das den schriftsprachlichen Gebrauch der deutschen Sprache abbildet. Neben den so ermittelten Einwortlemmata konnten über manuelle, qualitative Analysen zudem mehrteilige Stichwortkandidaten angesetzt werden (z.B. *keine Ahnung*, *guck mal*), die als Einheit eine Funktion in der Interaktion besitzen.

Projektergebnisse

Mit der Erstellung des LeGeDe-Prototyps entstand sowohl aus der Sicht der Forschung zur gesprochenen Sprache als auch aus lexikografischer Perspektive eine völlig neuartige Form der Sprachbeschreibung und ihrer Darstellung. In den Wortartikeln werden sowohl lexikalische Eigenschaften als auch interaktionale Verwendungsweisen der Stichwörter in authentischen Gesprächen beschrieben (vgl. Abb. unten). Transkriptausschnitte mit Audios belegen die Analysen.

Folgende Ergebnisse konnte das Projekt in seiner Laufzeit erzielen:

- die Durchführung verschiedener empirischer Studien zu Erwartungen, Anforderungen und Wünschen an eine lexikografische Ressource des gesprochenen Deutsch,
- die Bestimmung der Besonderheiten von mündlichem Sprachgebrauch im lexikalischen Bereich auf allen Ebenen (Form, Inhalt/Funktion, Situation etc.),
- die Entwicklung neuartiger lexikografischer Angabetypen, die u.a. auf die Funktion lexikalischer Einheiten in Interaktionskontexten Bezug nehmen, und ihrer Visualisierungsmöglichkeiten,
- die Entwicklung neuartiger Beschreibungsformate in multimedialer Form (Transkripte und Audiodateien) für hochgradig kontextualisierte lexikalische Daten sowie
- die Erstellung des Tools *Lexical Explorer*, mit dem Nutzende auf der Basis von Korpusdaten aus FOLK und GeWiss (Gesprochene Wissenschaftssprache) eigenständig Wortprofile erzeugen und diese mit Blick auf Wortformen, Kookkurrenzen und Metadaten quantitativ untersuchen können.



Im Ergebnis visiert der Prototyp hauptsächlich Benutzersituationen im wissenschaftlichen Bereich (Sprachdokumentation) an. Perspektivisch leistet das Projekt auch einen Beitrag zur Verbesserung eines an der Sprachrealität der Gegenwart orientierten Unterrichts von Deutsch als Fremd- bzw. Zweitsprache und unterstützt außerdem die Fundierung eines sprachreflexiven muttersprachlichen Unterrichts. Der LeGeDe-Prototyp erlaubt es, auf der Grundlage von Korpusevidenz zu einem besseren Verständnis der Besonderheiten des Wortschatzes des gesprochenen Deutsch in der Interaktion zu gelangen. Im Kontext der Lehre (z.B. DaF, DaZ und DaM) kann die Ressource als Fundament für die Erstellung von Unterrichtsmaterialien dienen.

Forschungsgegenstand

Moderne Wörterbücher des Deutschen gründen sich traditionellerweise eher auf das geschriebene Deutsch. Im Projekt „Lexik des gesprochenen Deutsch“ wurde zum ersten Mal ein Prototyp eines korpusbasierten Wörterbuchs des gesprochenen Deutsch konzipiert. Dazu wurden lexikalische Phänomene, die besonders typisch für das gesprochene Deutsch sind, auf der Basis des am Leibniz-Institut für Deutsche Sprache erstellten FOLK-Korpus beschrieben, analysiert und lexikografisch aufbereitet. Das Wörterbuch ist in die Plattform OWID[plus] integriert und unter https://www.owid.de/legede/ online verfügbar. Es erfasst die lexikalischen und funktionalen Besonderheiten des gesprochenen Deutsch in der Interaktion, wie es im privaten und institutionellen Kontext verwendet wird.

Projektleitung

Prof. Dr. Meike Meliss (IDS / Universidad de Santiago de Compostela), Dr. Christine Möhrs (IDS)

Beteiligte Projekte

- FOLK (Forschungs- und Lehrkorpus Gesprochenes Deutsch)
- Arbeitsbereiche „Computerlexikographie“ und „Empirische Methoden“

Beteiligte MitarbeiterInnen über die Projektlaufzeit

Katja Arens, Prof. Dr. Arnulf Deppermann, Prof. Dr. Stefan Engelberg, Dolores Lemmenmeier (geb. Batinić), Dr. Antje Nolting (geb. Töpel), Dr. Thomas Schmidt, Sarah Torres Cajo, Hilfskräfte, PraktikantInnen

Projekt-Kontakt

Dr. Christine Möhrs, Abteilung Lexik
Leibniz-Institut für Deutsche Sprache
legede@ids-mannheim.de
http://www1.ids-mannheim.de/lexik/lexik-des-gesprochenen-deutsch.html

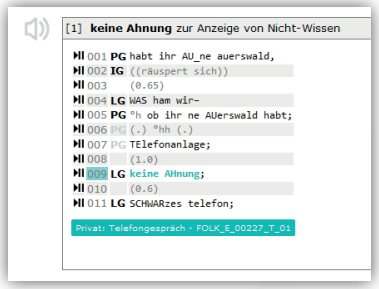
IDS LEIBNIZ-INSTITUT FÜR DEUTSCHE SPRACHE

R 5, 6-13 • D-68161 Mannheim
Tel.: +49 621 1581-0
Fax: +49 621 1581-200
info@ids-mannheim.de
www.ids-mannheim.de

www.facebook.com/ids.mannheim
www.twitter.com/ids_mannheim

Das Leibniz-Institut für Deutsche Sprache (IDS) ist die zentrale Einrichtung zur Erforschung und Dokumentation der deutschen Sprache in ihrem Gebrauch und in ihrer neueren Geschichte. Es gehört zu den über 90 außeruniversitären Forschungs- und Serviceeinrichtungen der Leibniz-Gemeinschaft.

IDS LEIBNIZ-INSTITUT FÜR DEUTSCHE SPRACHE



Abteilung Lexik
Leibniz-Institut für Deutsche Sprache

LEGEDE
LEXIK DES GESPROCHENEN DEUTSCH

Stand 09/2019

Leibniz Gemeinschaft