



# Germanistik im Umbruch – Linguistik, Übersetzung und DaF

Irene Doval/Elsa Liste Lamas (Hg.)

Irene Doval/ Elsa Liste Lamas (Hg.)  
Germanistik im Umbruch – Linguistik, Übersetzung und DaF



Irene Doval / Elsa Liste Lamas (Hg.)

Germanistik im Umbruch –  
Linguistik, Übersetzung und DaF

Umschlagabbildung © paffy – stock.adobe.com

ISBN 978-3-7329-0327-6

ISBN E-Book 978-3-7329-9393-2

ISSN 2510-3792

© Frank & Timme GmbH Verlag für wissenschaftliche Literatur  
Berlin 2019. Alle Rechte vorbehalten.

Das Werk einschließlich aller Teile ist urheberrechtlich geschützt.  
Jede Verwertung außerhalb der engen Grenzen des Urheberrechtsgesetzes ist ohne Zustimmung des Verlags unzulässig und strafbar.  
Das gilt insbesondere für Vervielfältigungen, Übersetzungen, Mikroverfilmungen und die Einspeicherung und Verarbeitung in elektronischen Systemen.

Herstellung durch Frank & Timme GmbH,  
Wittelsbacherstraße 27a, 10707 Berlin.

Printed in Germany.

Gedruckt auf säurefreiem, alterungsbeständigem Papier.

[www.frank-timme.de](http://www.frank-timme.de)



## Inhalt

Vorwort | 7

### Teil I: Lexikografie, Lexikologie, Semantik und Syntax des Deutschen

*Meike Meliss / Christine Möhrs*

Herausforderungen eines innovativen lexikografischen Projekts zu Besonderheiten des gesprochenen Deutsch in der Interaktion | 13

*Sabine Geck*

Abwertende personengruppenbezogene Bezeichnungen (*slurs*): Semantische Verschiebungen im Deutschen als Zeichen sozialer und kultureller Veränderungen | 27

*Ulrike Oster*

Vier Jahrhunderte *Wut*. Entwicklung eines vielschichtigen Emotionsworts | 37

*Ewa Maria Majewska*

Besonderheiten des medizinischen Vokabulars fremden Ursprungs und sein Transfer in den deutschsprachigen Fachzeitschriften | 47

*Herbert J. Holzinger*

*In Zukunft vs. in der Zukunft*. Ein korpusbasierter Vergleich von Verwendungsweisen, Musterbildung und Festigkeit | 57

*Nely M. Iglesias Iglesias / Alejandro Alonso Santos*

*Viel mehr als nur zu zweit!* Korpusgesteuerte Beschreibung des Musters [für + ADJ + *Stunden zu zweit*] | 69

*Ireneusz Gaworski*

Zur emotionalisierenden Wirkung der Verb-Erst-Nachfeldbesetzung in der Sprache der protestantischen Leichenpredigten | 77

### Teil II: Kontrastive Linguistik und Übersetzung

*Christiane Nord*

Fremdheit oder Befremdlichkeit? Anredeformen in dem Roman *La sombra del viento* von Carlos Ruiz Zafón und seinen Übersetzungen ins Deutsche, Italienische und Englische | 91

*Ana Mansilla Pérez*

Schematisierung im Deutschen und im Spanischen: Das idiomatische Satzmuster  
[*Du kannst mich* + INF] aus konstruktionsgrammatischer Sicht | 105

*Irene Szumlakowski Morodo*

Redundanz in der Beschreibung von Fortbewegung in der vertikalen Achse  
im Deutschen und im Spanischen | 115

*Ewa Wojacek*

Neue sprachliche Abkürzungen als eine der Folgen der Welt- und Sprachdigitalisierung  
am Beispiel des Polnischen, Deutschen und Schwedischen | 127

### **Teil III: DaF-Didaktik**

*Irene Doval / Elsa Liste Lamas*

Vorschläge zum Einsatz von Parallelkorpora im DaF-Unterricht | 139

*Roswitha Althoff*

Typische Sprachstrukturen im DaF-Unterricht: Adverbien, trennbare Verbpartikeln und  
Richtungsadverbien | 151

*Begoña Velasco Arranz*

Ausarbeitung eines Tools zur Analyse und Bewertung von Software für das Erlernen der  
deutschen Sprache in heterogenen Gruppen von Lernenden | 161

*Mireia Calvet Creizet*

Die Sprachen in der Sprache und über die Sprache: korpusgestützte Förderung von  
fachgerechten Diskursfähigkeiten am Beispiel der deutschen Sprachgeschichte | 167

*Marta Panadés Guerrero*

*Sich beschweren* als expressiver Sprechakt in professionellen Umgebungen: emotionale  
Werte im Kommunikationsstil didaktisierter Beschwerden Spanisch/Deutsch | 175

*María Mar Soliño Pazó / Amador García Tercero*

Kiezdeutsch und seine Einführung in den DaF-Unterricht | 183

## Vorwort

Vom 13. bis 16. September 2017 fand an der Philologischen Fakultät der Universität Santiago de Compostela die Konferenz *Umbrüche gestalten: Germanistik in bewegter Zeit* statt, die von der Forschungsgruppe der Universität Santiago de Compostela *LitLinAl* (Literatura y Lingüística Alemana, *GI-1954*) organisiert wurde.<sup>1</sup> Im Rahmen der Tagung wurden unter anderem folgende Fragen aufgeworfen: *Wie beeinflussen die Veränderungen der Gegenwart die Studien der deutschen Sprache und Literatur? Wie können sich die germanistische Sprach- und Literaturwissenschaft einer sich kontinuierlich wandelnden und von konstantem Austausch geprägten globalen Welt anpassen? Welche Möglichkeiten bieten die Digital Humanities dem Germanistik-Studium? Welche Veränderungen und Chancen zeitigt die Digitalisierung für die Übersetzungswissenschaft und die Sprachdidaktik?*

Ziel der Tagung war es, ein Diskussionsforum zur Untersuchung der vielfältigen Wechselwirkungen zwischen interdisziplinärem Wissen, Digitalisierung, Globalisierung und zunehmender Mobilität im Bereich der Germanistik, der Übersetzungswissenschaft und der Sprachdidaktik sowie den jeweils involvierten Akteuren und Institutionen zu schaffen. Über 80 WissenschaftlerInnen der In- und Auslandsgermanistik nahmen unsere Einladung nach Santiago de Compostela wahr und schufen durch eine Vielzahl an Vorträgen, gestützt auf unterschiedliche theoretische und methodologische Ansätze, eine fruchtbare Diskussionsplattform. Ergänzt und begleitet wurde dieser Austausch durch die Ausstellung *Spaniens Himmel – Die österreichischen Freiwilligen in den Internationalen Brigaden* sowie durch den Workshop *Einführung in die digitale Textedition*.<sup>2</sup>

Aus der Konferenz sind zwei Tagungsbände mit dem Titel *Germanistik im Umbruch* hervorgegangen. Der erste ist der deutschsprachigen Literatur gewidmet, während im Fokus des zweiten die germanistische Linguistik, die Übersetzungswissenschaft und die DaF-Didaktik stehen. Es ist dies der vorliegende Band, welcher eine Auswahl der Tagungsbeiträge aus jenen drei germanistischen Teildisziplinen enthält.

Die hier versammelten 17 Arbeiten illustrieren verschiedene thematische und methodologische Umbrüche und setzen sich mit hochaktuellen, zum Teil wenig erforschten Fragen der deutschen Linguistik, der Übersetzungswissenschaft und der DaF-Didaktik auseinander. Eine Reihe von Beiträgen zeigt neue Möglichkeiten der Anwendung von Sprachkorpora für lexikologische und syntaktische Fragestellungen, aber auch für die Entwicklung didaktischer Konzepte und Ressourcen auf. Andere Arbeiten beziehen gesprochene Varietäten des Deutschen oder neue Formen der Kommunikation mit ein oder behandeln konkrete Fragen der interkulturellen Pragmatik. Auch die kritische Auseinandersetzung mit Lehrcurricula und -inhalten ist Gegenstand einer weiteren Reihe von Beiträgen.

Der vorliegende Band gliedert sich in drei Teile: Teil I umfasst Beiträge, die sich mit lexikografischen, lexikologischen, semantischen und syntaktischen Fragen des Deutschen auseinandersetzen. Teil II enthält Aufsätze, die Fragen der kontrastiven Linguistik und der Übersetzungswissenschaft gewidmet sind. Teil III schließlich beinhaltet Arbeiten zur DaF-Didaktik.

---

<sup>1</sup> Unser Dank gilt der Galicischen Landesregierung (2014-PG083, 2017-PG023), der Nationalen Forschungsagentur (AEI) des Spanischen Ministeriums für Wirtschaft und Universitäten (Projekte FEM2015-71909-REDT, FFI2013-42571-P, FFI2017-85938 und FFI2013-45769-P), dem Österreichischen Kulturforum sowie dem Vize-Rektorat für Forschung der Universität Santiago de Compostela, durch deren finanzielle Unterstützung die Organisation der Konferenz ermöglicht wurde.

<sup>2</sup> An dieser Stelle sei Georg Pichler, dem Kurator der Ausstellung, herzlich gedankt. Weiterhin gilt unser Dank Erich Hackl für seine Lesung im Rahmen ebendieser sowie dem Workshop-Leiter Valentín Carrera.

Die in Teil I versammelten Beiträge behandeln ein breites Spektrum von lexikografischen, lexikologisch-semantischen sowie syntaktischen Fragen.

Im ersten Beitrag zeigen *Meike Meliss* und *Christine Möhrs* neue Perspektiven für lexikografische Arbeiten auf und beschreiben die methodologischen sowie wissenschaftlichen Herausforderungen des Projekts *LeGeDe (Lexik des Gesprochenen Deutsch)*, in dem es um die Erstellung eines korpusbasierten Wörterbuches zur Lexik des gesprochenen Deutsch geht. Hierbei geben die Autorinnen Einblick in die Mikrostruktur des Wörterbuches und stellen ein im Rahmen des Projektes entwickeltes Tool für die Gewinnung quantitativer Informationen vor.

*Sabine Geck* setzt sich in ihrem Beitrag mit drei sogenannten *slurs*, also Begriffen, die zur Herabwürdigung von Menschen aufgrund ihrer Gruppenzugehörigkeit verwendet werden, auseinander. Im Rahmen ihrer semantischen Analyse illustriert die Autorin dabei drei Arten von semantischer Verschiebung, nämlich die Aneignung (*schwul*), die Ersetzung (*Zigeuner* vs. *Sintil/Roma*) und die Tabuisierung (*Jude*). Zudem schlägt Geck ein Modell für die semantische Analyse von *slurs* anhand der *Natural Semantic Metalanguage* vor.

Ebenfalls semantische Fragestellungen spricht *Ulrike Oster* in ihrem Beitrag zur diachronen Bedeutungsentwicklung des Emotionswortes *Wut* an. Im Zuge ihrer korpusbasierten Methode begrenzt sich die Autorin dabei nicht auf eine rein quantitative Untersuchung, sondern analysiert komplementär hierzu die entsprechenden Kookkurrenzen auch qualitativ. Hierdurch gelingt es ihr, Bedeutungsaspekte zu identifizieren, die sonst unerkannt geblieben wären.

Basierend auf der Analyse medizinischer Fachschriften liefert die Arbeit von *Ewa Majewska* einen Einblick in die morphologischen und orthographischen Hauptmerkmale des deutschen medizinischen Fachvokabulars fremden Ursprungs. Im Mittelpunkt stehen dabei vor allem der griechische, englische und französische Einfluss, der anhand einer Reihe von Termini aus Anatomie und Klinik sowie weiteren medizinischen Feldern analysiert wird.

Die darauffolgenden zwei korpuslinguistisch basierten Arbeiten richten ihr Augenmerk auf usuelle Wortverbindungen und sind damit an der Schnittstelle zwischen Lexikon und Grammatik zu verorten. Zunächst stellt *Herbert Holzinger* in seinem Beitrag Strukturen des Typs [Präposition + Substantiv] und [Präposition + Artikel + Substantiv] gegenüber. Konkret untersucht er die Konstruktionen *in Zukunft* und *in der Zukunft* und identifiziert Unterschiede und Gemeinsamkeiten, indem er Verwendung, Musterbildung und interne Festigkeit beider Konstruktionen vergleicht. Schließlich zeigen *Nely Iglesias* und *Alejandro Alonso*, wie ausgehend von der Verbindung *für* + Zeiteinheit die Identifikation des usuellen Wortbindungsmusters [*für* + ADJ + *Stunden zu zweit*] gelingt. Zusätzlich zur konkreten Gebrauchsanalyse des Musters liefern die Autoren einen Vorschlag für dessen lexikografische Behandlung.

Die Funktion der Nachfeldbesetzung in Verb-Erst-Strukturen im späten Frühneuhochdeutschen und frühen Neuhochdeutschen ist Gegenstand der Arbeit von *Ireneusz Gaworski*. Der Autor analysiert Imperativsätze und rhetorische Fragesätze aus protestantischen Leichenpredigten des 16. bis 18. Jahrhunderts und zeigt auf, wie der Verb-Erst-Nachfeldbesetzung in derartigen illokutionären Kontexten eine besondere, emotionalisierende Funktion zukommt.

Die Arbeiten in Teil II setzen sich aus mehrsprachiger und/oder kontrastiver Perspektive mit Übersetzungsstrategien, Konstruktionen und Wortbildungsmechanismen auseinander.

*Christiane Nord* befasst sich in ihrem Beitrag mit der Übersetzung von Anredeformen in literarischen Texten. Anhand eines Korpus von Dialog-Turns aus einem spanischen Roman und deren Übersetzungen ins Deutsche, Italienische und Englische analysiert sie, wie die Strategien der Verfremdung, der Adaptation und der Neutralisierung im Zuge der Übersetzung einer Reihe von Anredeformen sprachlich kodiert werden. Dabei werden Unterschiede zwischen den drei Übersetzungen deutlich, die Auswirkungen auf die Rezeption des Werkes mit sich bringen.

Gegenstand des Beitrags von *Ana Mansilla* ist die Phrasemkonstruktion [*Du kannst mich* + INF] und ihre Entsprechungen im Spanischen. Die Autorin konzentriert sich dabei auf die Verwendung der Konstruktion zum Ausdruck von Sprechakten des Ablehnens, Zurückweisens und Kritisierens und kontrastiert u. a. Füller, Satztypen und kommunikative Funktionen, die in deutschen und spanischen Belegen zu identifizieren sind.

Der Fokus des Beitrages von *Irene Szumlakowski* liegt auf Fortbewegungsbeschreibungen der vertikalen Achse im Deutschen und Spanischen. Basierend auf einer Analyse von Belegen aus einem Parallelkorpus geht sie darauf ein, wie spanische vektorielle Verben im Deutschen wiedergegeben werden und inwieweit Fortbewegungsbeschreibungen mit doppelter Richtungsbestimmung in beiden Sprachen als redundant zu interpretieren sind.

Sprachliche Abkürzungen im Polnischen, Deutschen und Schwedischen stehen im Zentrum des Beitrags von *Ewa Wojacek*. Die Autorin ermittelt verschiedene Abkürzungsfunktionen und stellt eine detaillierte Abkürzungstypologie vor, welche durch Beispiele aus den drei Sprachen illustriert wird. Überdies analysiert sie für die Textsorten E-Mail und SMS typische Abkürzungen von Anrede-, Gruß- und Abschiedsformeln sowie von emotiven Gefühlslagen und geographischen Orten.

Der dritte und letzte Teil des vorliegenden Bandes umfasst Arbeiten, die sich mit der Didaktik des Deutschen und der Germanistik im spanischen Kontext befassen.

*Irene Doval* und *Elsa Liste* erörtern die Einsatzmöglichkeiten von Parallelkorpora im fortgeschrittenen DaF-Unterricht anhand eines Parallelkorpus Deutsch/Spanisch. Neben allgemeinen Überlegungen zum Einsatz von Parallelkorpora im Fremdsprachenunterricht stellen die beiden Autorinnen die Hauptmerkmale des Korpus PaGeS vor und verdeutlichen sein Einsatzpotenzial exemplarisch anhand zweier Aufgaben zur deutschen Grammatik und zum deutschen Wortschatz.

*Roswitha Althoff* geht in ihrem Beitrag der Frage nach, wie und wann deutsche Adverbien, trennbare Verbpartikeln und Richtungsadverbien im DaF-Unterricht für Spanischmuttersprachler eingeführt werden sollen. Basierend auf der Analyse der Frequenz solcher adverbialer Strukturen in einem Korpus und deren Behandlung in DaF-Lehrbüchern plädiert die Autorin für einen früheren und expliziten Einsatz im Unterricht.

*Begoña Arranz* stellt in ihrem Aufsatz die Entwicklung eines Tools für die Bewertung digitaler DaF-Inhalte und Materialien vor, das sich besonders für heterogene Lernergruppen eignet. Sie diskutiert dabei das Konzept der Heterogenität im DaF-Unterricht und zeigt, wie das geplante Tool die Adäquatheit von Inhalten und Materialien anhand von vier Kriterienblöcken und eines Punktesystems bewerten soll.

Lehre und Studium der deutschen Sprachgeschichte im Kontext der Bologna-Hochschulreform stehen im Mittelpunkt des Beitrages von *Mireia Calvet*. Die Autorin erörtert mehrere Umbrüche im Zuge der Reform und zeigt anhand eines konkreten Didaktisierungsbeispiels zur Arbeit mit diachronen Daten, wie der Umgang mit der beschriebenen Umbruchsituation gelingen kann.

Ausgehend von Gemeinsamkeiten und Unterschieden zwischen deutschem und spanischem Wirtschaftskommunikationsstil setzt sich *Marta Panadés* mit der didaktischen Behandlung des illokutionären Aktes der Beschwerde in Lehrwerken für Wirtschaftsdeutsch und -spanisch auseinander. Dabei analysiert sie aus kontrastivem und intrakulturellem Blickwinkel die Struktur von Beschwerden ebenso wie die kommunikativen Strategien und sprachlichen Mittel, die für diese Sprechakte sprachspezifisch und sprachübergreifend typisch sind.

Der Band schließt mit dem Beitrag von *Mar Soliño* und *Amador García*, in dem die Autoren zeigen, wie das sogenannte Kiezdeutsch fruchtbar in den DaF-Unterricht einfließen und

## Vorwort

ihn nachhaltig bereichern kann. Hierfür werden zunächst grundlegende soziologische und sprachliche Merkmale des Kiezdeutschen diskutiert, ehe Soliño und García mit Nachdruck auf Vorteile und Wichtigkeit der Einbeziehung von Jugendsprachvarietäten in den Fremdsprachenunterricht eingehen. Illustriert wird dieses Plädoyer anhand einer didaktischen Sequenz.

Noch einmal möchten wir uns an dieser Stelle bei allen AutorInnen sowohl für ihre Beiträge im vorliegenden Band als auch während der Konferenz in Santiago de Compostela herzlich bedanken. Dieser Dank gilt darüberhinaus auch allen weiteren KonferenzteilnehmerInnen, deren Diskussionsfreudigkeit einen erheblichen Beitrag zum Gelingen der Konferenz geleistet hat.

Die Herausgeberinnen  
Irene Doval und Elsa Liste Lamas

## **Teil I**

# **Lexikografie, Lexikologie, Semantik und Syntax des Deutschen**



# Herausforderungen eines innovativen lexikografischen Projekts zu Besonderheiten des gesprochenen Deutsch in der Interaktion

MEIKE MELISS / CHRISTINE MÖHRS

*Universidade de Santiago de Compostela / Leibniz-Institut für Deutsche Sprache*

## Abstract

The article presents the scientific and methodological challenges for the development of an innovative, corpus-based lexicographic resource for the lexicon of spoken German in interaction and points out new ways for lexicographical work. In addition to general project information on the starting points, the data basis, the methods, objectives and the concrete subject area, selected results of two project-related empirical studies on expectations on a lexicographic resource of spoken German are outlined. For corpus-based quantitative information, the possibilities of a tool developed in the framework of the project are presented. Furthermore, an insight into the conceptual and methodological considerations with regard to the microstructure of the lexicographical resource is provided.

**Keywords:** lexicology, interaction, spoken German, lexicography, corpus-based

## Zusammenfassung

Der Beitrag stellt die wissenschaftlichen und methodologischen Herausforderungen für die Erstellung einer innovativen, korpusbasierten lexikografischen Ressource zur Lexik des gesprochenen Deutsch in der Interaktion vor und zeigt neue Wege für lexikografische Arbeiten auf. Neben allgemeinen Projektinformationen zu den Ausgangspunkten, der Datengrundlage, den Methoden, Zielen und dem konkreten Gegenstandsbereich werden ausgewählte Ergebnisse von zwei projektbezogenen empirischen Studien zu Erwartungshaltungen an eine lexikografische Ressource des gesprochenen Deutsch präsentiert. Für korpusbasierte quantitative Informationen werden die Möglichkeiten eines Tools, welches im Rahmen des Projekts entwickelt wurde, aufgezeigt. Außerdem wird ein Einblick in die konzeptionellen und methodologischen Überlegungen zur Mikrostruktur der geplanten Ressource gegeben.

**Schlüsselwörter:** Lexikologie, Interaktion, gesprochenes Deutsch, Lexikografie, korpusbasiert

## 1. Einleitung

Ziel des Beitrages<sup>1</sup> ist es, die wissenschaftlichen und methodologischen Herausforderungen für die Erstellung einer korpusbasierten lexikografischen Ressource zur Lexik des gesprochenen Deutsch im Rahmen des Projekts LeGeDe<sup>2</sup> vorzustellen. Zunächst werden allgemeine Projektinformationen zu den Ausgangspunkten, zum Gegenstandsbereich, zu den Methoden und Zielen sowie zu der Datengrundlage skizziert. Die Ausgangspunkte sollen außerdem in Zusammenhang mit ausgewählten Ergebnissen aus zwei projektbezogenen empirischen Studien zu Erwartungshaltungen an eine lexikografische Ressource des gesprochenen Deutsch betrachtet werden (vgl. Abschnitte 2 und 3).<sup>3</sup> Schließlich wird von einigen konzeptionellen und methodologischen Überlegungen sowie von zukunftsgerichteten Betrachtungen zur lexikografischen

---

<sup>1</sup> Wir danken an dieser Stelle Annette Klosa-Kückelhaus für die Durchsicht des Manuskriptes und ihre vielen wertvollen Hinweise.

<sup>2</sup> Das Forschungsprojekt „Lexik des gesprochenen Deutsch“ (LeGeDe) ist ein drittmittelgefördertes Projekt der Leibniz-Gemeinschaft, wird als Kooperationsprojekt der Abteilungen Pragmatik und Lexik am Leibniz-Institut für Deutsche Sprache (IDS) in Mannheim umgesetzt und hat eine Laufzeit von drei Jahren (2016-2019). Eine detaillierte Projektbeschreibung findet sich in Meliss/Möhres (2017).

<sup>3</sup> Informationen zu den Rahmenbedingungen beider Umfragen sind auf den LeGeDe-Projektseiten zusammengestellt, außerdem können dort auch die Fragebögen abgerufen werden: <<http://www.ids-mannheim.de/lexik/lexik->

Umsetzung berichtet (vgl. Abschnitte 4 und 5). Auch diese Überlegungen werden in Teilen mit Ergebnissen aus den oben genannten empirischen Studien in Verbindung gesetzt.

## 2. Ausgangspunkte und Hauptannahmen

Das LeGeDe-Projekt verfolgt das Ziel, eine korpusbasierte elektronische Ressource, die lexikalische Besonderheiten des gesprochenen Deutsch in der Interaktion zum Gegenstand hat, zu entwickeln.<sup>4</sup> Im Rahmen der Projektarbeit sollen mit unterschiedlichen korpusbasierten Methoden v. a. diejenigen Spezifika der gesprochensprachlichen Lexik in der Interaktion identifiziert, analysiert und beschrieben werden, die in der bisherigen lexikografischen Kodifizierung und Forschung eher vernachlässigt wurden und werden. Die Entwicklung neuartiger lexikografischer Beschreibungsformate und Angabetypen, die Formen, Bedeutungen und Funktionen von lexikalischen Einheiten in interaktionalen Kontexten besonders in den Blick nehmen, ist dabei ein wichtiger Teil der Konzeption des geplanten Wörterbuches (= WB<sup>5</sup>). Der anvisierte lexikografische Prototyp soll als mögliche Zielgruppe v. a. Gesprächsforscher, Lexikologen, Sprachwissenschaftler und Sprachlehrende ansprechen (vgl. dazu Ergebnisse aus den Umfragen Meliss/Möhrs/Ribeiro Silveira 2018: 124f. und 132; Meliss/Möhrs/Ribeiro Silveira 2019: 116).

Das LeGeDe-Projekt basiert auf folgenden vier Hauptannahmen und Beobachtungen: (i) Es existieren Unterschiede in der Lexik des gesprochenen im Vergleich zum geschriebenen Deutsch auf unterschiedlichen Ebenen (vgl. Deppermann et al. Hg. 2017; Fiehler 2016; Imo 2007; Schwitalla <sup>4</sup>2012). (ii) Die lexikografische Kodifizierung der interaktionstypischen Besonderheiten der gesprochensprachlichen Lexik des Deutschen ist unzureichend (vgl. u. a. Meliss 2016; Meliss/Möhrs 2018; Meliss/Möhrs/Ribeiro Silveira 2019; Moon 1998; Trap-Jensen 2004). Der Vergleich des Informationsangebotes aus verschiedenen lexikografischen Standardwerken<sup>6</sup> lässt die oben genannten Schlüsse zu, sodass es gilt, die bestehenden Desiderata für diesen Gegenstandsbereich aufzudecken und neue korpusbasierte Methoden für die Erstellung einer lexikografischen Ressource zu erarbeiten, die die vorhandenen Lücken schließen möchten. (iii) Es ist außerdem zu beobachten, dass in den letzten Jahren der Informationsbedarf zu typisch gesprochensprachlicher Lexik allgemein und in unterschiedlichen Anwendungsbereichen, wie z. B. in Unterricht und Lehre (speziell im Sekundarbereich und in den Bereichen Deutsch als Fremd- und/oder Zweitsprache) sowie im Verlagswesen in Verbindung mit der Erstellung von geeigneten Unterrichtsmaterialien gestiegen ist (vgl. Handwerker et al. Hg. 2016; Imo/Moraldo Hg. 2015; Moraldo/Missaglia Hg. 2013; Reeg et al. Hg. 2012; Sieberg 2013). Auch im „Gemeinsamen europäischen Referenzrahmen für Sprachen“ (= GeR) wird u. a. zum Beurteilungsraster zur mündlichen Kommunikation und dem Parameter „Interaktion“ für Niveau C1 explizit darauf hingewiesen, dass der Lernende „[...] aus einem ohne weiteres verfügbaren Repertoire von Diskursmitteln eine geeignete Wendung auswählen [kann], um

---

des-gesprochenen-deutsch/projektbeschreibung/empirische-forschung.html>. Ergebnisse zu den Studien allgemein finden sich in Meliss/Möhrs/Ribeiro Silveira (2018), mit L2-Perspektive in Meliss/Möhrs/Ribeiro Silveira (2019) sowie in Meliss/Möhrs (2018).

<sup>4</sup> Am IDS sind in den letzten Jahren die Kompetenzen gereift, die es ermöglichen, ein solches empirisch fundiertes Wörterbuch des gesprochenen Deutsch zu erstellen. Es liegen hinreichend große Korpora des gesprochenen Deutsch, Expertise in der Konzeption und Realisierung komplexer multimedialer Internetwörterbücher und Erfahrungen in der lexikologischen und semantischen Analyse gesprochener Sprache in der Interaktion vor.

<sup>5</sup> Im Folgenden wird der Ausdruck „Wörterbuch“ (in allen Flexionsformen) mit „WB“ abgekürzt.

<sup>6</sup> Wir verweisen hier exemplarisch auf die lexikografischen Einträge zu *gucken* und *okay* in LGWB-DaF, DWDS und Duden-online.

seine/ihre Äußerung angemessen einzuleiten, wenn er/sie das Wort ergreifen oder behalten will, oder um die eigenen Beiträge geschickt mit denen anderer Personen zu verbinden“ (Trim et al. 2001: 37). Die Lernenden können den formulierten Ansprüchen des GeR und der Lehrpläne nur schwerlich gerecht werden, wenn keine Ressourcen vorliegen, in denen die genannten gesprochen sprachlich typischen Phänomene der Lexik konsultiert werden können (vgl. Melliss/Möhrs 2018). Die Ergebnisse der im LeGeDe-Projekt durchgeführten Studien zu der Erwartung künftiger Nutzer an die geplante Ressource zeigen, dass sowohl bei den L1-SprecherInnen als auch bei den L2-SprecherInnen des Deutschen zu jeweils über 70 % Bedarf an einem WB zu Spezifika des gesprochenen Deutsch vorhanden ist. Diese Beobachtung bestätigt die grundsätzliche Annahme zum ansteigenden Bedarf. (iv) Es existieren zurzeit kaum korpusbasierte lexikografische Projekte zur Lexik der gesprochenen Sprache. Lediglich zum Dänischen wurde ein kleines Projekt zu Interjektionen (vgl. Hansen/Hansen 2012) durchgeführt und es liegt ein Paper zum gesprochenen Slowenisch in Verbindung mit lexikografischen Überlegungen vor (vgl. Verdonik/Sepesy Maučec 2017).

Diese vier Grundannahmen sind Ausgangspunkte für die konzeptionellen Überlegungen der geplanten lexikografischen Ressource im LeGeDe-Projekt. Daneben wurden gleich zu Beginn des Projekts die zwei oben bereits erwähnten empirischen Studien zu Erwartungshaltungen an eine lexikografische Ressource mittels zweier Befragungen<sup>7</sup> durchgeführt. Sie fokussieren neben Informationen zu soziodemografischen Daten und zur Nutzung von Online-Ressourcen allgemein hauptsächlich Fragen zur Einschätzung der Relevanz der Spezifika der gesprochenen Lexik in unterschiedlichen Situationen sowie zu Vorstellungen und Erwartungen an eine zukünftige Ressource zur Lexik des gesprochenen Deutsch.

### 3. Gegenstandsbereich und Korpusgrundlage

Der Gegenstandsbereich, mit dem sich das LeGeDe-Projekt beschäftigt, ist die Lexik des gesprochenen Deutsch in unterschiedlichen Interaktionstypen. Die im Fokus stehende Lexik ist durch die Merkmale „standardnah“ und „distinktiv“ ausgezeichnet, die eine Abgrenzung zu anderen Sprachvarietäten erlauben. Regionale, soziolektale, funktionale oder idiolektale Sprachvarietäten werden in den Analysen nicht behandelt. Im Zentrum des Interesses stehen besonders die distinktiven Merkmale im Vergleich zu der Lexik der geschriebenen Standardsprache. Die LeGeDe-Datengrundlage beruht auf dem Forschungs- und Lehrkorpus Gesprochenes Deutsch (= FOLK<sup>8</sup>). FOLK ist das größte Korpus zum gesprochenen Deutsch in der Interaktion und erweist sich in seiner Zusammensetzung als geeignete Datengrundlage für die geplante LeGeDe-Ressource. Für den korpusbasierten Vergleich mit der geschriebenen Sprache werden entsprechende Teilkorpora aus dem Deutschen Referenzkorpus (= DeReKo<sup>9</sup>) genutzt. FOLK ist integriert in das Archiv für Gesprochenes Deutsch und über die Datenbank für Gesprochenes Deutsch (= DGD)<sup>10</sup> abrufbar. Es beinhaltet Gesprächsaufnahmen aus dem deutschsprachigen Raum in unterschiedlichen privaten, institutionellen und öffentlichen Kontexten mit entsprechenden Transkripten und teilweise auch Videoaufnahmen.

---

<sup>7</sup> Wenn im weiteren Textverlauf auf die beiden Studien referiert wird, so wird für das durchgeführte Experteninterview die Abkürzung „EXPI“ und für die Onlineumfrage „OU“ verwendet.

<sup>8</sup> FOLK: <http://agd.ids-mannheim.de/folk.shtml>; vgl. auch Schmidt 2014a und 2016.

<sup>9</sup> DeReKo: <http://www.ids-mannheim.de/direktion/kl/projekte/korpora/releases.html>; vgl. Kupietz/Schmidt 2015; Institut für Deutsche Sprache 2017.

<sup>10</sup> DGD: <http://dgd.ids-mannheim.de>; vgl. auch Schmidt 2014b.

Eine der zentralen Forschungs- und Methodikfragen, mit denen sich das LeGeDe-Projekt beschäftigt, steht in Verbindung mit dem Vergleich zur Lexik der geschriebenen Sprache und damit in Bezug zur Erstellung einer Stichwortliste mit Kandidaten, die möglichst typische Phänomene des Gesprochenen aufweisen. Dazu wurde ein frequenzgesteuertes korpusbasiertes Verfahren entwickelt, bei dem die Häufigkeitsklassen der Lemmata aus FOLK und aus DEREKO ermittelt und diese dann in einem direkten Vergleich betrachtet werden (vgl. Meliss et al. 2018). Dieser Häufigkeitsklassenvergleich wurde zunächst intern genutzt, er ist ab September 2018 aber auch über das frei zugängliche Tool „Lexical Explorer“<sup>11</sup> nachvollziehbar. Dieses Tool wurde für die konkreten Ziele im Rahmen der LeGeDe-Projektarbeit entwickelt und ermöglicht u. a., verschiedene quantitative Daten aus FOLK abzufragen und mit Hilfe von Häufigkeitstabellen bezüglich der Wortverteilung über Wortformen, Kookkurrenzen und Metadaten zu erforschen.

Die Top-10 der möglichen Stichwortkandidaten (vgl. in Abb. 1 die Spalte „Lemma“) vermitteln einen ersten Eindruck von denjenigen Kandidaten, die für die Ressource relevant sind. Die hohe Häufigkeitsklassendifferenz bei den Interjektionen (*ah*, *ach*, *oh*), Abtönungspartikeln (*ja*, *halt*), Gesprächspartikeln (*okay*, *ja*, *na*) und Verben (*gucken*, *kriegen*) weist schon mit dieser vor allem quantitativen Perspektive auf Besonderheiten im Gesprochenen hin, die dann im Projekt durch qualitative Studien genauer analysiert werden.

■ Study corpus vs. DeReKo

Show/hide columns CSV Show 10 entries

Lemma	Study corpus freq	DeReKo freq	Study corpus HK	DeReKo HK	HK Diff	PoS
<b>okay</b>	7314	199942	4	14	10	NG
<b>ah</b>	4676	152521	4	14	10	NG
<b>ach</b>	4427	445211	4	13	9	NG
<b>ja</b>	73125	10506472	0	8	8	ADV/NG/PTK
<b>gucken</b>	2989	375327	5	13	8	V
<b>oh</b>	3732	313095	5	13	8	NG
<b>halt</b>	7072	802658	4	12	8	ADV/NG/PTK
<b>du</b>	28145	6206378	2	9	7	PRO
<b>kriegen</b>	2249	667783	5	12	7	V
<b>na</b>	3459	520673	5	12	7	NG

Filter Lemma Filter Study corpus Filter DeReKo Freq <9 <15 >1 Filter PoS

Showing 1 to 10 of 307 entries (filtered from 57,666 total entries)

Abb. 1: Screenshot aus dem „Lexical Explorer“: Top-10 Lemmata mit höchster Häufigkeitsklassendifferenz (HK Diff) zwischen FOLK (= Study corpus) und DEREKO<sup>12</sup>

<sup>11</sup> Der „Lexical Explorer“ kann über OWID<sup>plus</sup> abgerufen werden: [http://www.owid.de/lexex/index\\_de](http://www.owid.de/lexex/index_de)

<sup>12</sup> Die Berechnungen basieren auf dem Stand vom DGD Release 2.10 (23.05.2018).

Eine genauere Untersuchung der erzielten Stichwortkandidaten verweist auf verschiedene Phänomene bzw. Phänomenbereiche, die sich auch durch Rückbezüge zu einschlägigen Studien – als besonders gesprochen sprachlich identifizieren lassen (vgl. Fiehler 2016, Schwitalla <sup>4</sup>2012):

- (i) Im Verbalbereich sind aus formal-inhaltlicher und interaktionaler Perspektive u. a. Verwendungsformen mit Distributionsbeschränkungen in der Verbalmorphologie [*ich dachte* (Person, Numerus und Tempus), *guck* (Modus: Imperativ), *meinste* (Verschmelzungen bzw. Komplementierungsmuster), *ich kann kein Deutsch* (absoluter Gebrauch von Modalverben)] und spezifische Bedeutungsvarianten, die häufig nur reduzierte morphologische Paradigmen aufweisen (*geht/gehen* mit einer spezifischen Semantik in der 3. Pers. Sg./Pl. z. B. in den Bedeutungen ‚in Bewegung sein‘, ‚einigermaßen akzeptabel sein‘ etc.) für die LeGeDe-Ressource von Interesse.
- (ii) Auch im Bereich der Entlehnungen – sowohl aus anderen Sprachen (Anglizismen: *cool, okay* etc.) als auch aus deutschen Sprachvarietäten (soziolektal: *geil, krass, Mama, wieso* etc.) – zeigen sich interessante, gesprochen sprachliche Phänomene, die u. a. in Zusammenhang mit ihrer Gebrauchsfrequenz, dem Vorkommen in bestimmten Interaktionstypen, den Sprechergruppen, der grammatischen Integration und der phonetischen Realisierung stehen.
- (iii) Im Bereich der Wortbildung lassen sich u. a. zahlreiche Affixe mit formalen Besonderheiten und/oder neuen Bedeutungsnuancen verbuchen:
  - rum-* (-*ärgern, -blödeln, -gammeln, -hängen, -laufen, -machen, -sitzen* etc.),
  - rein-* (-*bringen, -gehen, -kommen, -machen, -passen, -tun* etc.),
  - rauf-* (-*gehen, -schmier* etc.),
  - mega-* (-*crazy, -einfach, -gut, -krass, -lang, -schlecht, -schwer, -stressig, -viel* etc.),
  - super-* (-*cool, -easy, -gut, -lecker, -safe, -schön, -toll, -viel* etc.),
  - sau-* (-*blöd, -dumm, -geil, -gut, -teuer* etc.),
  - mäßig* (*alters-, freizeit-, hobby-, kosten-, vereins-, zahlen-* etc.).
- (iv) Die Untersuchung zu Gebrauch und Bedeutung von Passepartoutwörtern (machen, tun, Ding, Sache etc.) sowie von Ausdrücken und Formeln der Vagheit (*irgend-* [*-wann, -was, -wie, -wo*] etc.) stehen v. a. in Verbindung mit der Bedeutungskonstituierung in der Interaktion.
- (v) Die Verwendung von Teilsynonymen (*werfen/schmeißen, kriegen/bekommen/erhalten, gucken/schauen/sehen, warum/weshalb/wieso, darum/deshalb/deswegen* etc.) interessiert insbesondere aus der Perspektive von Stil und Register.
- (vi) Die für die gesprochene Sprache besonders typischen Interjektionen (*ach, ah, oh* etc.) sowie Diskurspartikeln (*einfach, genau, gut, okay, richtig, sicher* etc.) und Abtönungspartikeln (*eben, halt, mal* etc.) bilden ebenfalls einen interessanten und umfangreichen Bereich des anvisierten Gegenstandsbereiches.
- (vii) Schließlich sind von ganz zentralem Interesse auch zahlreiche Verwendungsweisen bzw. feste Verwendungsmuster in Verbindung mit einer Verbalform (*guck mal, wenn du meinst, sag ich mal, wenn's sein muss, was weiß ich, ich weiß nicht* etc.), einer Nominalform (*keine Ahnung, mein Gott, Gott sei Dank, in Ordnung* etc.) oder adjektivischen bzw. adverbialen Elementen (*alles klar, na gut, na schön* etc.), die im interaktionalen Kontext eine spezifische Funktion besitzen.

Der Gegenstandsbereich wurde auch in den durchgeführten empirischen Befragungen thematisiert. Diesbezüglich kann daher an die Ergebnisse der Frage „*Welche Art von Stichwörtern würden Sie in einem WB des gesprochenen Deutsch erwarten?*“ aus der OU angeknüpft werden, die besonders den Wunsch nach Stichwörtern, die eine spezifische Kombinatorik aufweisen, deutlich unterstreichen. Jeweils rund drei Viertel aller Befragten der OU erwarten in einem WB des gesprochenen Deutsch sowohl Stichwörter, die eine formelhafte Verwendung (79,5 %) besitzen, als auch Stichwörter mit einem – im Vergleich zur geschriebenen Sprache – anderen Kombinationspotenzial (74,6 %). Der Vergleich mit entsprechenden Antworten zu einer parallelen Frage aus dem EXPI zeigt ähnliche Ergebnisse. Lexikalische Einheiten, die in der gesprochenen Sprache ein anderes Kombinationspotenzial als in der geschriebenen Sprache aufweisen, stehen bei den ExpertInnen an oberster Stelle auf der Wunschliste von möglichen Stichwörtern (94,1 %). Dies beinhaltet – laut Äußerungen der interviewten ExpertInnen – Konstruktionen, lexikalische Ausdrücke, syntagmatische Kombinationen, Formeln, Chunks etc. und auch Mehrwortlemmata.

#### **4. Datengewinnung und lexikografische Umsetzung**

Die einzelnen Phasen und Arbeitsschritte zur Entwicklung der LeGeDe-Ressource folgen grundsätzlich dem lexikografischen Prozess zur Erstellung von Internetwörterbüchern, wie es in Klosa/Tiberius (2016: 75f.) beschrieben wird (vgl. dazu einige exemplarisch ausgewählte Daten zu Pilotstudien in Abschnitt 4.1). Die lexikografische Umsetzung stellt sich dabei neuen Herausforderungen, die in Verbindung mit einer Reihe von unterschiedlichen lexikografischen Fragestellungen stehen: Welche Stichwortansetzung kann sinnvollerweise für ein Wörterbuch zur gesprochenen Sprache verfolgt werden (z. B. Einwortlemmata, Mehrwortlemmata)? Welche Benutzungssituation wäre für ein Wörterbuch zum Gesprochenen denkbar? Wie können authentische mündliche Sprachdaten (z. B. Audiodateien, Transkripte) in die Ressource integriert werden? Wie können Informationen zu Form, Inhalt und Funktion von gesprochensprachlichen lexikalischen Elementen in der Mikrostruktur der geplanten Ressource (vgl. hierzu Abschnitt 4.2) angemessen dargestellt werden?

##### **4.1. Gewinnung, Kodierung, Analyse und Strukturierung der Daten**

Verschiedene Verfahren zur Bedeutungsdisambiguierung und zur korpusbasierten Entwicklung von Wortprofilen, die als methodologische Ansätze in der korpusbasierten Lexikologie und Lexikografie gelten, sollen mit der Analyse sprachlicher Einheiten und der Beschreibung von kommunikativen Funktionen aus Sicht der Interaktionslinguistik (vgl. u. a. Deppermann 2007) vereint werden. Unterschiedliche Informationen, wie u. a. automatisch generierte Daten (Frequenzdaten zum Formenbestand, zu Kollokationen, Kookkurrenzen etc.) sollen nicht nur bei der ersten korpusbasierten Annäherung an die Daten durch entsprechende Hypothesenbildung, sondern auch bei der Interpretation der Analyseergebnisse unterstützen. In ersten Studien wurden die Gesprächssequenzen zu entsprechenden Belegen nach Form, Bedeutung und Funktion in Verbindung mit metasprachlicher Information analysiert, die Ergebnisse strukturiert und hinsichtlich der Relevanz der Ressource interpretiert. Datengrundlage für die Analyse ist jeweils eine Zufallsstichprobe aus FOLK. Zu den Korpusbelegen stehen in allen Fällen Audioaufnahmen zur Verfügung, die über die Datenbank parallel zu den Transkripten abgerufen werden können. Neben Metadaten zum Treffer und zum Transkript, die automatisch extrahiert vorlie-

gen, können in manuellen Untersuchungen inhaltlich funktionale, syntaktisch-formale, sequenzrelevante sowie grammatische Aspekte im Abgleich mit den Metadaten betrachtet werden. Die Möglichkeiten dieses Verfahrens lassen sich exemplarisch an dem Verballemma *gucken* aufzeigen. Die erzielten Kodierungs- und Strukturierungsergebnisse von 100 Belegen einer Zufallsstichprobe ergaben für *gucken* zunächst mindestens neun verschiedene Bedeutungsvarianten, von denen nur zwei im LGWB-DaF und dem DWDS sowie drei in Duden-online verzeichnet sind. Die semantische Disambiguierung erfolgte durch ein Zusammenspiel von Kodierungen bezüglich der Form (Strukturmuster) und der spezifischen Bedeutungen (vgl. dazu Auszüge aus dem Bedeutungsspektrum des Lemmas *gucken* auf Basis der LeGeDe-Stichprobe in Meliss/Möhrs 2017: 47, Möhrs/Meliss/Batinić 2017: 293). Zusätzlich zu diesen Beobachtungen, die die inhaltliche Beschreibungsebene betreffen, finden sich in der Stichprobe spezielle, distributionell eingeschränkte und/oder modalisierende Verwendungsformen (Imperativ / Modalpartikel / Modalverb: *guck (mal)*; (*muss*) *mal gucken*). Eine auf diesen Beobachtungen aufbauende Analyse und Systematisierung der Ergebnisse zeigen weitere musterhafte Verwendungsmöglichkeiten von *gucken* mit unterschiedlichen interaktionalen Funktionalitäten, z. B. als Diskursmarker auf (vgl. dazu u. a. Günthner 2017, Wegner 2015, Proske 2017).

#### 4.2. Überlegungen zur Mikrostruktur der geplanten Ressource

Erste Überlegungen zur Entwicklung einer lexikografischen Mikrostruktur der LeGeDe-Ressource verfolgen ein breit gefächertes Informationsangebot. Jeder Wortartikel enthält bedeutungs- und funktionsübergreifende frequenzorientierte Informationen, die (semi-) automatisch und teilweise auch im Vergleich zu Daten aus DEREKO generiert werden, zum Beispiel zu Häufigkeiten und Häufigkeitsklassen. Weitere frequenzorientierte Angaben zum Formeninventar (vgl. Tab. 1), zur Kombinatorik (Bigramme, Trigramme und Kookkurrenzen; vgl. zu Kookkurrenzen Abb. 2) oder zu Metadaten (vgl. Abb. 3) können über den „Lexical Explorer“ abgerufen werden und sind Ausgangspunkt für weitere qualitative lexikografische Analysen. Ein Link aus dem Wortartikel zum Tool ermöglicht den NutzerInnen weitere explorative Untersuchungen. Entsprechende korpuslinguistische Werkzeuge (wie zum Beispiel eine Kookkurrenzanalyse) sind in bisherigen lexikografischen Projekten (vgl. z. B. das DWDS oder *lexiko*) zur automatischen Analyse und Strukturierung von Sprachdaten nicht neu. Allerdings wurden diese Verfahren bislang noch nicht auf Gesprächsdaten angewendet. Fehlende Satzgrenzen, Überlappungen, Pausen oder Abbrüche sind nur einige der Herausforderungen, die bei der Erarbeitung von diesen Werkzeugen auch für gesprochensprachliches Material zu bedenken sind. Frequenzorientierte Korpusabfragen sowie die Abfragen quantitativer Daten über den „Lexical Explorer“ werden im Folgenden an verschiedenen Beispiellemmata vorgestellt.

(i): Durch frequenzbasierte Daten zum Formeninventar in FOLK und DEREKO, die hier durch das Häufigkeitsklassenmaß (Frequenzklassen) und das entsprechende Maß für Frequenzklassendifferenzen angegeben werden, lassen sich u. a. Schlüsse zu gesprochensprachlichen Besonderheiten ziehen. So geben beispielsweise die Daten zur Frequenzklasse für *denke* (z. B. 1. Pers. Sg. Indikativ Aktiv Präsens) mit der HK-Klasse „6“ in FOLK und einer Häufigkeitsklassendifferenz zu Gunsten von FOLK im Vergleich zu DEREKO (HK-Differenz) von „3“ (vgl. Tab. 1) einen Hinweis darauf, dass die Verbindung *ich denke* besondere gesprochensprachliche Funktionen haben kann (z. B. als „gesprächsstrukturelles Gliederungssignal“, vgl. Zeschel 2017: 291).