Originalveröffentlichung in: Piosik, Michał/Taborek, Janusz/Woźnicka, Marta (Hrsg.): Korpora in der Lexikographie und Phraseologie: Stand und Perspektiven – Berlin, Boston: de Gruyter, 2021. S. 141-173. (Lexicographica. Series Maior 160)
DOI: *https://doi.org/10.1515/9783110716955-008*

Meike Meliss

# Informationsspektrum und Angabeklassen in der lexikografischen Ressource zum gesprochenen Deutsch: LeGeDe

**Abstract:** The main aim of this contribution is to present the range of lexicographic information from LeGeDe, an electronic prototype for lexical and interactional features of spoken German. The focus lies on the detailed description of the different lexicographical information classes using illustrative examples and figures from the resource. In addition to highlighting the lexicographic microstructure and providing an overview of the outer texts and the multimedia information offer, the contribution also presents detailed background data on the conception of the LeGeDe resource. Innovative aspects and possible applications are outlined and forward-looking desiderata are offered.

**Keywords:** online lexicography, spoken German in interaction, corpus linguistics, microstructure, lexicographic information classes

**Schlagwörter:** Onlinelexikografie, Lexik des gesprochenen Deutsch in der Interaktion, Korpuslinguistik, Mikrostruktur, lexikografische Angabeklassen

## 1 Einleitung

Die **LeGeDe-Ressource** versteht sich als Prototyp für die lexikografische Darstellung der Besonderheiten des standardnahen, gemeinsprachlich gesprochenen Deutsch in der Interaktion, so wie es in natürlichen Gesprächen im privaten und institutionellen Kontext genutzt wird[1]. Die Entwicklung dieses multimedialen Prototypen bietet erstmals die Möglichkeit der lexikografischen Sprachdokumentation korpusbasierter lexikalischer Besonderheiten des gesprochenen standardnahen Deutsch und spricht als mögliche Zielgruppen primär Gesprächsforschende, LexikologInnen, SprachwissenschaftlerInnen und Sprachlehrende an[2].

1 Der vorliegende lexikografische Prototyp wurde in dem Forschungsprojekt „Lexik des gesprochenen Deutsch" (=LeGeDe) zwischen 2016 und 2019 am Leibniz-Institut für Deutsche Sprache in Mannheim konzipiert und erstellt. Die Autorin war in der besagten Zeit am IDS Koleiterin des Projekts.

2 In Ergänzung zu den konzeptionellen Arbeiten im LeGeDe-Projekt wurden auch verschiedene Arten von empirischen Studien zu Erwartungshaltungen an eine neuartige lexikografische Ressource zu Besonderheiten der gesprochensprachlichen Lexik durchgeführt. Viele der Ergebnisse

**Meike Meliss,** Universität Santiago de Compostela, Spanien, meike.meliss@usc.es

https://doi.org/10.1515/9783110716955-008

*Publikationsserver des Leibniz-Instituts für Deutsche Sprache*
*URN: http://nbn-resolving.de/urn:nbn:de:bsz:mh39-104275*

Sowohl aus der Sicht der Gesprächsforschung als auch aus lexikologischer Perspektive war es notwendig, eine neuartige Form der lexikografischen Sprachbeschreibung und –darbietung zu entwickeln, für die bislang kaum Vorbilder existierten. Für die Erfüllung der Projektziele mussten daher im Laufe der Projektzeit innovative lexikografische Beschreibungsformate für stark kontextualisierte lexikalische Daten in audioelektronischer Form, ebenso wie neuartige lexikografische Angabetypen, die auf die Funktion lexikalischer Einheiten in interaktionalen Kontexten Bezug nehmen, entwickelt werden. Für die erstellten Wörterbuchartikel wurde somit sowohl für die formale als auch für die inhaltliche und interaktional-funktionale Ebene ein umfangreiches Spektrum von Angabeklassen als Artikelsegmente eingesetzt (Engelberg/Lemnitzer 2009: 156 ff., Wiegand 1989b: 427).

Seit September 2019 kann die LeGeDe-Ressource öffentlich frei konsultiert werden[3] und zeichnet sich u. a. durch folgende Merkmale aus (vgl. Engelberg/Lemnitzer 2009, Engelberg/Storrer 2016):

- elektronisch/digital, internetbasiert
- neu konzipiert
- Ausbauwörterbuch: dynamisch
- nicht automatisch sondern redaktionell ausgearbeitetes Informationsangebot
- ohne Nutzerbeteiligung
- einsprachig
- gegenwartssprachlich/standardsprachlich
- medial: gesprochen/interaktional
- korpusbasiert
- multimedial
- multimodular
- semasiologisch: form-bedeutungs-/funktionsorientiert
- mit internem und externem Verlinkungsangebot

Ziel des vorliegenden Beitrages ist es, die einzelnen Angaben als lexikografische Artikelsegmente detailliert zu beschreiben und damit hauptsächlich die lexikografische Mikrostruktur der elektronischen LeGeDe-Ressource vorzustellen (Engelberg/Lemnitzer 2009: 166 ff., Wiegand 1989a, 1989b). So befassen sich daher Abschnitt 2 zunächst mit der Präsentation des LeGeDe-Projekts und Abschnitt 3 mit der ausführlichen Information zu der Konzeption der LeGeDe-Ressource. Abschnitt 4 präsentiert das Informationsspektrum der Ressource mit illustrativen Beispielen und

konnten in dem lexikografischen Prozess berücksichtigt und in der Ressource entsprechend umgesetzt werden (vgl. Meliss/Möhrs/Ribeiro Silveira 2018b, 2019a).

**3** Die LeGeDe-Ressource ist in das IDS-Portal OWID$^{plus}$ eingebettet und kann unter www.owid.de/legede/ konsultiert werden.

Abbildungen. Abschließend werden die innovativen Aspekte und Anwendungsmöglichkeiten skizziert und ausblickende Desiderata angeboten.

## 2 Das LeGeDe-Projekt

Das LeGeDe-Projekt basiert auf folgenden Hauptannahmen und Beobachtungen, die kurz dargestellt werden. Einerseits existieren Unterschiede auf verschiedenen sprachlichen Ebenen zwischen dem gesprochenen im Vergleich zum geschriebenen Deutsch. Die Divergenzen bezüglich der Lexik können sich sowohl im Bestand als auch in Verbindung mit deren Form, Bedeutung und Verwendung manifestieren (vgl. Deppermann/Proske/Zeschel Hrsg. 2017, Fiehler 2016, Imo 2007, Schwitalla 2012). Obwohl eine Vielzahl von interaktionslinguistischen Einzelstudien vorliegen, steckt die korpusbasierte Lexikografie des Gesprochenen jedoch noch in den Kinderschuhen (Meliss 2016: 195; Eichinger 2017: 283). Sowohl die Ergebnisse der besagten projektbezogenen Umfragen zu den Erwartungshaltungen an eine lexikografische Ressource für Spezifika des gesprochenen Deutsch (vgl. Meliss/Möhrs/Ribeiro Silveira 2018b, 2019a) als auch entsprechende stichprobenhaft durchgeführte lexikografische Untersuchungen bestätigen, dass die lexikografische Kodifizierung der Merkmale der gesprochenen Sprache in der Interaktion in den aktuellen Wörterbüchern bisher nicht zufriedenstellend berücksichtigt wird (vgl. Meliss 2016, 2021, Meliss/Möhrs 2017, 2018, 2019, 2020, Moon 1998, Siepmann 2015, Trap-Jensen 2004)[4]. Trotz einiger jüngster Fortschritte in der Erstellung von gesprochensprachlichen Korpora zu verschiedenen Sprachen und Sprachvarietäten (vgl. Barcala et al. 2018, Fandrych/Meißner/Slavcheva Hrsg. 2014, Fandrych/Meißner/Wallner Hrsg. 2017, Schmidt 2014a, 2014b, 2017, Verdonik/Sepesy Maučec 2017) sind die Erfahrungen mit ihrer Nutzung in der Lexikografie bis jetzt eher selten. Daher kann sich die LeGeDe-Ressource kaum auf bestehende Modelle stützen, die z. B. der Erstellung einer geeigneten Stichwortliste oder der multimedialen Darstellung als Orientierungshilfe hätte dienen können. Es ist aber andererseits auch zu beobachten, dass der Informationsbedarf zu Inhalt, Form und Funktion typisch gesprochensprachlicher Lexik in den letzten Jahren 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/Bäuerle/Sieberg Hrsg. 2016, Imo/Moraldo Hrsg. 2015, Moraldo/Missaglia Hrsg. 2013, Reeg/Gallo/Moraldo Hrsg. 2012, Sieberg 2013). So wird z. B. im „Gemeinsamen europäischen Referenzrahmen für

4 Lediglich zum Dänischen kann hier ein kleines lexikografisches Projekt zu Interjektionen genannt werden (Hansen/Hansen 2012).

Sprachen“ (=GeR) zum Beurteilungsraster zur mündlichen Kommunikation und dem Parameter „Interaktion“ u. a. für Niveau C1 explizit darauf hingewiesen, dass der Lernende „[. . .] aus einem ohne weiteres verfügbaren Repertoire von Diskursmitteln eine geeignete Wendung auswählen [kann], um 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/North/Coste 2001: 37). Sowohl Meliss/Möhrs (2018: 81 ff.) als auch Fandrych/Meißner/Wallner (2018: 3 ff.) weisen darauf hin, dass Korpora der gesprochenen Sprache von großem Interesse für viele Forschungs- und Anwendungsszenarien sind und gewinnbringend für die Sprachdidaktik eingesetzt werden können, wenn den Lehrenden und Lernenden ein benutzergerechter Zugang zu dem Material angeboten wird. Schließlich zeigen auch die Ergebnisse der im LeGeDe-Projekt durchgeführten Studien zu den Erwartungen künftiger NutzerInnen an die Ressource, dass sowohl bei den L1- als auch bei den L2-Sprechenden des Deutschen zu jeweils über 70 % Bedarf an einem WB zu Spezifika des gesprochenen Deutsch vorhanden ist (Meliss/Möhrs/Ribeiro Silveira 2018b: 124 f.). Diese Beobachtung bestätigt die grundsätzliche Annahme zum ansteigenden Bedarf an einer Ressource zu Spezifika des Gesprochenen sowohl in der Forschung als auch in der Lehre.

Das Hauptziel des LeGeDe-Projekts ist die Erstellung einer korpusbasierten elektronischen Ressource des gesprochenen Deutsch auf der Grundlage von korpusbasierten Untersuchungen zu den Besonderheiten von mündlichem Sprachgebrauch in der Interaktion im Bereich des Lexikons[5]. Die besagten Annahmen sind Ausgangspunkte für die konzeptionellen Überlegungen zur Entwicklung der Ressource im LeGeDe-Projekt, für die quantitative und qualitative Methoden miteinander verknüpft wurden. Auf diese Weise konnten die Spezifika gesprochensprachlicher Lexik des Deutschen auf der Datengrundlage des am Leibniz-Institut für Deutsche Sprache erstellten Forschungs- und Lehrkorpus Gesprochenes Deutsch (=FOLK)[6] (Schmidt 2014a, 2014b, 2017) analysiert und für die lexikografische Anwendung aufbereitet werden. Folgende zentrale theoretische, methodologische und anwendungsorientierte Aspekte haben sich bei der Beschäftigung mit dem Thema in der Projektarbeit ergeben:

**5** Es handelt sich um ein Drittmittelprojekt der Förderlinie „Innovative Vorhaben“, welches im Leibniz Wettbewerb eingeworben werden konnte. Die Kooperation zwischen den Abteilungen Pragmatik und Lexik des Leibniz-Instituts für Deutsche Sprache in Mannheim ermöglichte während der dreijährigen Projektlaufzeit eine Verbindung der entsprechenden Fachkompetenzen, die für die Erstellung einer korpusbasierten lexikografischen Ressource des gesprochenen Deutsch in der Interaktion absolut notwendig sind. Für weitere allgemeine Informationen zu dem Projekt „Lexik des gesprochenen Deutsch“ verweisen wir auf die LeGeDe-Projektwebseite: <http://www1.ids-mannheim.de/lexik/lexik-des-gesprochenen-deutsch.html>.

**6** FOLK ist in das Archiv für Gesprochenes Deutsch integriert und kann über die Datenbank für Gesprochenes Deutsch (=DGD2) abgerufen werden.

- Entwicklung von quantitativen und qualitativen Verfahren zur Identifizierung von interaktionstypischen, gesprochensprachlichen lexikalischen Elementen und deren spezifischen Merkmalen (vgl. Meliss/Möhrs 2017)
- Erstellung einer Stichwortkandidatenliste und Auswahl geeigneter Stichwörter für die Ressource (vgl. Meliss et al. 2018a, Meliss 2020b)
- Entwicklung von korpusbasierten quantitativen und qualitativen Verfahren für die lexikalische und interaktional ausgerichtete Datenanalyse und Beschreibung auf verschiedenen sprachlichen Ebenen (vgl. Meliss et al. 2018a)
- Bestimmung der lexikalischen Besonderheiten von mündlichem Sprachgebrauch auf verschiedenen Ebenen (Form, Inhalt/Funktion, Situation etc.) (vgl. Meliss/Möhrs 2017, 2018, 2019, 2020)
- Entwicklung sowohl neuartiger lexikografischer Angabeklassen, die u. a. auf die Funktion lexikalischer Einheiten in Interaktionskontexten Bezug nehmen, als auch von innovativen lexikografischen Beschreibungsformaten in multimedialer Form (vgl. Meliss et al. 2019b)
- Entwicklung weiterer korpuslinguistischer Methoden und Tools zur Abfrage und Strukturierung von automatisch generierten korpusbasierten Daten (vgl. Lemmenmeier-Batinić 2020, Möhrs/Meliss/Batinić 2017)

# 3 Hinweise zur Konzeption der LeGeDe-Ressouce

Die einzelnen Phasen und Arbeitsschritte zur Entwicklung der LeGeDe-Ressource folgen dem lexikografischen Prozess zur Erstellung von Internetwörterbüchern, wie es u. a. in Klosa (2013a, 2013b) und Klosa/Tiberius (2016: 75 f.) beschrieben wird. Die lexikografische Umsetzung stellt sich dabei neuen Herausforderungen, die in Verbindung mit einer Reihe von unterschiedlichen Fragestellungen stehen. Für den vorliegenden Beitrag sind v. a. folgende Fragen von zentralem Interesse:

- Welche Zielgruppen und Benutzungssituationen wären für ein Wörterbuch zum Gesprochenen denkbar?
- Wie können Informationen zu Form, Inhalt und Funktion von lexikalischen Elementen des gesprochenen in der Interaktion auf der Mikrostrukturebene der geplanten Ressource angemessen multimedial beschrieben und dargestellt werden?
- Wie können korpusbasierte, mündliche Sprachdaten (z. B. Audiodateien, Transkripte) gewinnbringend in die Ressource integriert werden?

Zu anderen relevanten lexikografischen Themenbereichen in Verbindung mit der Erstellung der LeGeDe-Ressource, wie z. B. zu der Stichwortansetzung soll auf den Beitrag von Möhrs (2021) und auf weitere projektrelevante Publikationen verwiesen werden (Meliss/Möhrs 2017, 2019, 2020; Meliss et al. 2018a; 2019b; Möhrs/Meliss/Batinić 2017).

## 3.1 Zielgruppe

Die Frage nach der anvisierten Zielgruppe spielt eine fundamentale Rolle für die Art und Weise der lexikografischen Umsetzung der erhobenen Daten. Da die LeGeDe-Ressource zunächst als Wissensspeicher und Wortschatzdokumentation fungieren soll, stehen die Daten v. a. einem wissenschaftlich interessierten Kreis zur Verfügung. Die besagte Ressource schafft aber auch eine Grundlage für Maßnahmen zur Verbesserung eines an der Sprachrealität der Gegenwart orientierten Unterrichts von Deutsch als Fremd- bzw. Zweitsprache und kann außerdem die Fundierung eines sprachreflexiven L1-Unterrichts unterstützen. So ist u. a. aus den Ergebnissen der projektbegleitenden empirischen Befragungen hervorgegangen, dass Benutzer und Benutzerinnen in Lernsituationen (v. a. in Produktionssituationen) von der LeGeDe-Ressource profitieren könnten, wenn Experten (Wissenschaftler, Lehrende etc.) eine entsprechende Mittlerposition einnehmen (Meliss/Möhrs/Riberiro Silveira 2018b: 132; 2019a: 116).

## 3.2 Datenanalyse und Datenstrukturierung

Unterschiedliche quantitative und qualitative Verfahren zur lexikalischen Bedeutungsdisambiguierung und zur Entwicklung von Wortprofilen, die als methodologische Ansätze in der korpusbasierten Lexikologie und Lexikografie gelten (vgl. u. a. Engelberg 2015, 2018, 2019; Engelberg et al. 2011), werden mit der interaktionslinguistischen Analyse sprachlicher Einheiten und der Beschreibung ihrer Formen und Funktionen in der Interaktion (vgl. u. a. Deppermann 2007, Couper-Kuhlen/Selting 2018) vereint und für den lexikografischen Prozess genutzt.

Über das im Projekt entwickelte Werkzeug ‚Lexical Explorer' wird ermöglicht, die quantitativen Daten von FOLK zu durchsuchen und abzufragen. Unterschiedliche Informationen, wie u. a. automatisch generierte Frequenzdaten zum Formenbestand, zu Kollokationen, Kookkurrenzen etc. werden nicht nur bei der ersten korpusbasierten Annäherung an die Daten durch entsprechende Hypothesenbildung sondern auch bei der Interpretation der Analyseergebnisse unterstützend genutzt. Außerdem bietet ein Häufigkeitsklassenvergleich mit Teilkorpora aus DeReKo (=Deutsches Referenzkorpus) einen quantitativen Vergleich zwischen den Daten beider Korpora an (vgl. Meliss et al. 2018a, 2019b).

Auf der Grundlage von qualitativen lexikalisch-semantischen und interaktionslinguistischen Einzelbeleganalysen wurden verschiedene Kodierschemata entwickelt, die jeweils die Besonderheiten der unterschiedlichen grammatischen Kategorien (Verb, Nomen, Diskurspartikel etc.) gerecht werden. Die Datengrundlage für die Analyse eines Lemmas ist jeweils eine Zufallsstichprobe aus FOLK, aus der die ersten 100 gültigen Belege auf der Basis von qualitativen Einzelanalysen kodiert wurden (= 100-Stichprobe). Neben Metadaten zum Treffer und zum Transkript, die automatisch extrahiert vorliegen, wurden formale (Person, Numerus, Modus etc.), inhaltliche (Bedeutung,

Bedeutungsumschreibung etc.), kombinatorische (Strukturmuster, Kookkurrenzen etc.), morphosyntaktische (Realisierungsmöglichkeiten der Aktanten etc.), funktionale, sequenzbezogene sowie prosodische Aspekte betrachtet und festgehalten. Da sich schon in der 100-Stichprobe die Notwendigkeit ergeben hat, einerseits die Ebene der Bedeutung und andererseits die funktionale Ebene der ausgewählten Stichwörter mit teilweise unterschiedlichen Analyseparametern und Methoden zu untersuchen und zu strukturieren, wurden für die weiteren Untersuchungen zwei teilweise unterschiedliche Verfahren angesetzt, die sich dann auch direkt auf eine unterschiedliche lexikografische Darstellung der Ergebnisse ausgewirkt haben. Im Zentrum von Modul 1 der LeGeDe-Ressource steht die semantisch und syntaktisch motivierte Disambiguierung der einzelnen Lesarten eines Stichwortes, während im Zentrum von Modul 2 die Beschreibung der interaktionalen Funktionen steht. Entsprechend wurden die weiteren Analyse-, Strukturierungs- und Beschreibungsverfahren den Anforderungen des jeweiligen Moduls angepasst und sollen im Folgenden detailliert dargestellt werden. Die Verteilung des Informationsangebotes auf zwei unterschiedliche Module ist folglich damit zu begründen, dass sich bestimmte Kategorien, wie z. B. die Diskurs- und Modalpartikeln und verfestigte interaktionale Einheiten adäquater mit interaktionslinguistischen als mit lexikalisch-semantischen Ansätzen beschreiben lassen.

### 3.2.1 Modul 1

Für das Informationsangebot in Modul 1 wurden die Kodierergebnisse der besagten 100-Stichprobe in ihren jeweiligen Gesprächssequenzen nach Form, Bedeutung und Kombinatorik (Strukturmuster, Kollokationen, feste Wendungen, interaktionale Einheiten etc.) in Verbindung mit dem Kontext und den Metadaten analysiert und auf der Grundlage einer semantisch-syntaktischen Disambiguierung einzelne Lesarten identifiziert. Als Orientierungshilfe für die jeweilige Lesartendisambiguierung dienten dafür entsprechende lexikografische Informationen ausgewählter Wörterbücher (z. B. E-VALBU, LGWB-DaF, Duden-online, DWDS). Durch den Vergleich der Daten aus der Stichprobe mit denen aus gängigen lexikografischen Werken, die hauptsächlich auf der geschriebenen Sprache basieren, wurden verschiedene Ziele verfolgt, die im Folgenden kurz dargestellt werden.

a. Mögliche, besonders interessante gesprochensprachliche Phänomene bieten solche Kandidaten, die sich als Lesarten identifiziren lassen, die nur in den FOLK-Daten belegt werden konnten, aber nicht in den konsultierten Wörterbüchern kodifiziert sind (so konnten z. B. für *gucken* in FOLK wesentlich mehr Lesarten belegt werden als in den konsultierten Wörterbüchern). Diese Fälle erlauben den Schluss, dass es sich zumindest um Lesarten handelt, die in der gesprochenen Interaktion – auf der Grundlage der LeGeDe-Stichprobe – als üblich gelten.

b. Die Identifizierung von Lesarten, die nicht in den FOLK-Daten belegt werden konnten, aber in den konsultierten Wörterbüchern kodifiziert sind, erlaubt den Schluss, dass es sich zumindest um Lesarten handelt, die in der gesprochenen Interaktion – auf der Grundlage der LeGeDe-Stichprobe – eher selten sind. Da sie nicht in der LeGeDe-Stichprobe belegt werden konnten, wurden sie für die LeGeDe-Ressource ausgeschlossen. Ein Verlinkungsangebot zu einigen elektronischen Wörterbüchern der deutschen Gegenwartssprache ermöglicht aber ein weiterführendes Informationsangebot.
c. Spezifische formale und morphosyntaktische Eigenschaften im Gesprochenen (z. B. Präferenz für Modus Imperativ bei *gucken* in der Bedeutung: „richtungsorientiert visuell wahrnehmen“) ermöglichen die Identifikation von Distributionsbeschränkungen.
d. Mögliche Fälle mit typisch gesprochensprachlichen kombinatorischen Merkmalen sind solche mit unterschiedlichen kombinatorischen Eigenschaften, die explizit in den FOLK-Daten der LeGeDe-Stichprobe belegt werden konnten, aber in den konsultierten Wörterbüchern nicht oder nur unzureichend explizit kodifiziert bzw. thematisiert sind. Diese Fälle umfassen verschiedene Bereiche u. a. in Verbindung mit folgenden Phänomenen:
   - Information zu den Strukturmustern und nicht realisierten Aktanten: z. B. konnte für *kriegen* in der Bedeutung „bekommen“ die Rolle „Sender“ nicht belegt werden;
   - Information zu reduzierten Strukturmustern und der ergänzenden Verwendung von deiktischen Elementen: z. B. bei *gucken* und *schauen* in der Bedeutung: „richtungsorientiert visuell wahrnehmen“;
   - Information zu typischen Verbindungen / Kollokationen: So zeigt z. B. *wissen* in der Bedeutung „informiert sein“ eine sehr häufige Verbindung mit den Partikeln *ja* und *aber* auf. Das Adjektiv *gut* in der Bedeutung von „positiv bewertet“ tritt häufig mit *echt* und das Adverb *eben* mit *gerade* auf.
   - Information zu festen Wendungen (z. B. *keine Ahnung haben*, *gucken wo jemand bleibt*, *frei kriegen*, *Besuch kriegen*, *die Kurve kriegen*) oder Routineformeln (z. B. *guten Appetit*, *danke schön*);

### 3.2.2 Modul 2

Die Analysen für das Informationsangebot in Modul 2 basieren ebenfalls auf der besagten 100-Stichprobe und liefern Hinweise auf relevante Phänomene und v. a. auf die Frage, welche Belege Kandidaten für eine interaktional ausgerichtete Beschreibung sein könnten. Dabei standen im Zentrum des Interesses u. a. Einheiten mit einer spezifischen kommunikativen Funktion im Gespräch wie z. B. die verschiedenen Muster mit den Verben *wissen* (*ich weiß nicht*, *weißt du*, *ich weiß*, *wer weiß*, *was weiß ich*), *gucken* (*guck mal*, *mal gucken*) und *schauen* (*schau mal*, *mal schauen*)

oder die polyfunktionalen Stichwörter *eben*, *gut*, *ja* und *schön*, die u. a. auch den Kategorien ‚Modalpartikel' und ‚Diskurspartikel' zuzuordnen sind. In den Fällen, in denen die 100-Stichprobe zwar Hinweise auf verfestigte Muster gibt, diese jedoch in nicht ausreichender Anzahl vorliegen, um belastbare Aussagen treffen zu können, wurden separate, gezielte Abfragen in FOLK generiert. Diese Belege wurden nach einer Zufallsauswahl extrahiert und ergänzend zur Analyse hinzugezogen. Für die Analysen für Modul 2 wurde außerdem das grundlegende Ausgangskodierschema um weitere relevante formale und funktionale Parameter ergänzt. Auf der Grundlage der Kodierung wurden anschließend die ermittelten typischen, rekurrenten und verfestigten Form-Funktions-Zusammenhänge unter Berücksichtigung der Sequenz- und Interaktionskontexte beschrieben und für ein lexikografisches Informationsangebot entsprechend aufbereitet. Da sich die Analysen beider Module zunächst auf die jeweilige LeGeDe-Ausgangsstichprobe beziehen, wurden nur die Phänomene (Bedeutungen, kommunikative Muster, Funktionen) beschrieben, die hier belegt werden konnten. Solche Phänomene, die ggf. Gegenstand von Beschreibungen in anderen wissenschaftlichen Einzelstudien und/oder lexikografischen Werken sind, aber nicht in der LeGeDe-Stichprobe dokumentiert werden konnten, sind daher nicht Teil des LeGeDe-Informationsangebotes.

# 4 Das Informationsangebot der LeGeDe-Ressource

Für jedes Stichwort bietet die LeGeDe-Ressource ein umfangreiches Informationsspektrum an (vgl. Abschnitt 4.2), welches durch verschiedene Außentexte (vgl. Abschnitt 4.1) „umrahmt" wird. Die lexikografische Struktur (vgl. Engelberg/Lemnitzer [4]2009, Wiegand 1989a, 1989b) und das Vernetzungsangebot (vgl. Engelberg/Müller-Spitzer/Schmidt 2016) sind dementsprechend komplex und vielfältig (vgl. Meliss/Möhrs 2020). Außerdem konnten verschiedene Wünsche aus den Umfragen bezüglich der multimedialen Präsentationsmodalitäten berücksichtigt werden (vgl. Abschnitt 4.3)[7].

## 4.1 Außentexte

Die Außentexte stellen ein zusätzliches Informationsangebot zur Verfügung, welches sich an unterschiedliche Nutzer wendet und die verschiedensten Möglichkeiten einer internetbasierten Ressource ausnutzt (vgl. Klosa/Gouws 2015). Der Informationstext „Über LeGeDe" stellt in ausführlicher Form die theoretischen und methodologischen

7 Siehe dazu auch das Informationsangebot in der LeGeDe-Ressource.

Grundlagen des LeGeDe-Projekts dar und bietet detaillierte Informationen zu den einzelnen lexikografischen Angabeklassen an. Dieser Text richtet sich wegen seiner Ausführlichkeit und Komplexität besonders an ein interessiertes Fachkollegium, das neben dem Interesse an der Ressource an sich, Interesse an metalexikografischen und methodologischen Fragestellungen hat. Die „Benutzungshinweise“ sind hingegen so konzipiert, dass sie den Wörterbuchbenutzenden auf schnelle, anschauliche und direkte Art und Weise den Zugang zu der Ressource und den einzelnen Strukturelementen durch eine „Guided Tour“ visuell näherbringen. Das „Glossar“ verfolgt das Ziel, bestimmte linguistische Fachbegriffe, die in der Ressource eine zentrale Rolle einnehmen, durch Kurzdefinitionen zu erklären bzw. über ein Verlinkungsangebot zu terminologischen Datenbanken zu verweisen. Im Vordergrund steht dabei hauptsächlich das Ziel, den für die Erstellung der LeGeDe-Ressource verfolgten terminologischen Ansatz anzudeuten bzw. zu skizzieren. Der Zugriff auf den entsprechenden Glossareintrag ist dabei entweder direkt über einen konkreten Terminus in der Ressource möglich oder über das alphabetisch angeordnete komplette Glossar. Der Außentext „Literatur“ bietet einen Einblick in die relevante Forschungsliteratur. Der Zugriff ist ebenfalls entweder über die Kurzverweise in den einzelnen Wörterbuchartikeln oder über die komplette Liste möglich.

## 4.2 Lexikografisches Informationsspektrum

Der Wunsch der zukünftigen Nutzenden einer lexikografischen Ressource zu Spezifika des Gesprochenen nach einem Informationsangebot, welches in speziellem Maße die Besonderheiten des gesprochenen Deutsch hervorhebt, wird u. a. in der Frage *Welche Informationen sollten in einem Onlinewörterbuch des gesprochenen Deutsch angeboten werden?* unserer Onlineumfrage (vgl. Abb. 1) deutlich geäußert (vgl. Meliss/Möhrs/Riberiro Silveira 2018b, 2019a). Sowohl durch die korpusbasierte Methode zur Stichwortkandidatenauswahl (vgl. Meliss et al. 2018a) als auch durch die korpusbasierte Analysemethode konnten quantitative und qualitative Daten zu Besonderheiten des Gesprochenen aus FOLK ermittelt und bei der lexikografischen Umsetzung an verschiedenen Stellen besonders explizit dargestellt, kommentiert bzw. visualisiert werden.

Das Informationsspektrum für jedes Stichwort wird in einem **Überblicksartikel** (vgl. Abschnitt 4.2.1), einem **Modul 1** zu lexikalisch-semantischer Information in Verbindung mit den unterschiedlichen Lesarten der Lemmata, die in der LeGeDe-Stichprobe identifiziert werden konnten (vgl. Abschnitt 4.2.2) und in einem **Modul 2** mit relevanter Information zu dem Funktionsspektrum der Lemmata (vgl. Abschnitt 4.2.3) dargestellt. Für beide Module wurden für die weitere Beschreibung der einzelnen Lemmata spezifische lexikografische Angabeklassen genutzt bzw. neu entwickelt, die sich neben dem klassischen Angebot auch durch völlig neuartige lexikografische Informationen und Formate auszeichnen. Verschiedene inhaltlich-funktionale

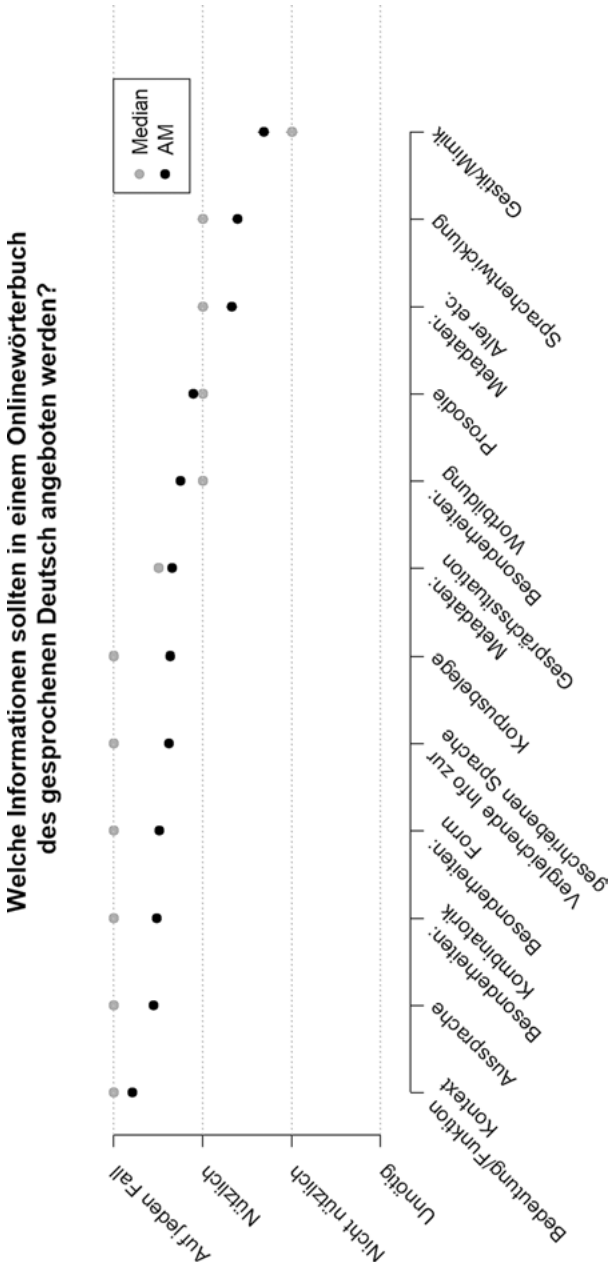



**Abb. 1:** Erwartungen an das Informationsangebot (Online-Umfrage).

Schnittstellen zwischen beiden Modulen werden durch interne Verknüpfungen explizit gemacht.

### 4.2.1 Überblicksartikel zu jedem Stichwort

Für jedes Stichwort, welches für die LeGeDe-Ressource aus der Stichwortkandidatenliste ausgewählt wurde (vgl. Möhrs 2021), steht zunächst ein Überblicksartikel, der aus unterschiedlichen Teilen besteht, zur Verfügung. Der deskriptive Überblickstext bietet allgemeine, zusammenfassende bedeutungs- und funktionsorientierte Informationen an (vgl. Abb. 2 zu WISSEN[8]). Die deutliche modulare Zweiteilung des Informationsangebotes ermöglicht einerseits die Präsentation der lexikalisch-semantischen Information, die an den jeweiligen Lesarten eines Lexems orientiert ist (=Modul 1: Bedeutungen) und andererseits die Darbietung der funktionsspezifischen, interaktional ausgerichteten Information (=Modul 2: Funktionen im Gespräch). Neben den Einwortlemmata (z. B. *wissen*) können auch unterschiedliche mehrteilige, teilweise musterhafte Einheiten als komplexe, mehrteilige Lemmata (z. B. *ich weiß nicht*) auftreten. Nicht für jedes Stichwort ist allerdings die modulare Zweiteilung, wie sie für WISSEN in Abb. 2 vorgestellt wird, sinnvoll. So wird für das Stichwort KRIEGEN nur Information in Modul 1 bereitgestellt (*kriegen* als Voll- und Hilfsverb). Für Partikeln wie HALT und JA werden hingegen nur Informationen in Modul 2 (*halt*: Modalpartikel; *ja*: Diskurspartikel, Modalpartikel) angeboten. In beiden Modulen erfolgt der Zugriff über die verschiedenen Lemmazeichengestaltangaben (=LZGAs). Diese werden formal und kategorial voneinander unterschieden (*wisse*n als ‚Vollverb' und als ‚Verbalsyntagma'). So sind z. B. die LZGAs zu dem Stichwort WISSEN folgende:

Modul 1: *wissen* als Vollverb

Modul 2: *wissen* als Nukleus eines Verbalsyntagmas:

- *ich weiß nicht*
- *weißt du*
- *ich weiß*
- *wer weiß*
- *was weiß ich*

Jeder LZGA werden entsprechende Etikettierungen zugeordnet. In Modul 1 handelt es sich um lesartdisambiguierende (z. B. *wissen*: ‚informiert sein', ‚kennen' etc.) und in Modul 2 um funktionsdifferenzierende Label (‚Unsicherheitsmarker', ‚Diskursmarker', ‚etcetera-Formel' etc.).

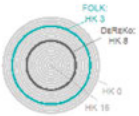
8 Alle Stichwörter, die noch keine formal-kategoriale Zuordnung erfahren haben, werden in diesem Beitrag durch Kapitälchen ausgezeichnet. Es handelt sich zunächst nur um eine abstrakte Einheit, die typographisch von den formgleichen Lemmata differenziert werden soll.

## wissen

Das mentale Verb **wissen** tritt in der LeGeDe-Stichprobe in verschiedenen Bedeutungsvarianten auf, die dem Bereich der Epistemik (Wissen und Nicht-Wissen) zuzuordnen sind. Neben der Grundbedeutung 'informiert sein' konnte **wissen** vereinzelt in den Bedeutungsvarianten 'kennen' und 'Information besitzen' dokumentiert werden. Zudem ließen sich in der LeGeDe-Stichprobe interaktionale Einheiten mit **wissen** belegen (**ich weiß nicht**, **weißt du**, **ich weiß**, **wer weiß**, **was weiß ich**), die vielfältige Funktionen in der Interaktion übernehmen und ebenfalls im Bereich der Epistemik liegen. Diese Einheiten dienen einerseits dazu, Wissen des Sprechenden hinsichtlich eines Sachverhalts anzuzeigen und zu beanspruchen sowie Unsicherheit (nicht-sicheres Wissen) zu kennzeichnen. Andererseits werden sie dazu genutzt, um die Aufmerksamkeit des Gegenübers zu lenken und einen Wissenstransfer erwartbar zu machen, die Aufmerksamkeit und das Verständnis des Gegenübers zu sichern, auf die Unvorhersagbarkeit einer Situation zu verweisen und eine Äußerung bzw. einen Teil einer Äußerung als potenziell unangemessen, fortführbar oder beispielhaft zu markieren.

**Frequenz**

FOLK 2.11 (absolut): 7246




**FOLK-Zugriff**

wissen (in FOLK)
wissen (im Lexical Explorer)

| Bedeutungen | Funktionen im Gespräch |
|---|---|
| **wissen** [VOLLVERB]<br>∘ informiert sein<br>∘ kennen<br>∘ Information besitzen | **ich weiß nicht** [VERBALSYNTAGMA]<br>∘ Unsicherheitsmarker<br>∘ Markierung potenzieller Unangemessenheit<br>**weißt du** [VERBALSYNTAGMA]<br>∘ Diskursmarker<br>∘ Vergewisserungssignal<br>**ich weiß** [VERBALSYNTAGMA]<br>∘ Anzeige von Wissen<br>∘ Beanspruchung von Wissen<br>**wer weiß** [VERBALSYNTAGMA]<br>∘ Verweis auf Unvorhersagbarkeit<br>**was weiß ich** [VERBALSYNTAGMA]<br>∘ etcetera-Formel<br>∘ Markierung eines Beispiels<br>∘ Unsicherheitsmarker |

**Information in anderen lexikografischen Resourcen**

- Online-Wortschatz-Informationssystem Deutsch (OWID)
- Elektronisches Wörterbuch zur Verbvalenz (eValbu)
- Digitales Wörterbuch der Deutschen Sprache (DWDS)

**Forschungsliteratur**

- Bergmann (2017)
- Erman (2001)
- Günthner (2017)
- Hagemann (2009)
- Helmer/Deppermann/Reineke (2017)
- Helmer/Deppermann (2017)
- Helmer/Reineke/Deppermann (2016)
- Imo 2007
- König/Stoltenburg (2013)
- König (2017)
- Reineke (2016)
- Zeschel (2017)

**Abb. 2:** Überblicksartikel zu WISSEN (Screenshot).

In dem Überblicksartikel werden außerdem sowohl modulübergreifende Informationen zu automatisch errechneten, korpusbasierten Häufigkeitsklassen des jeweiligen Stichwortes für die Daten der zu vergleichenden Korpora (DEREKO vs. FOLK) in visualisierter Form angeboten als auch der Zugriff auf FOLK und den Lexical Explorer bereitgestellt. Die kreisförmige Abbildung zu WISSEN zeigt die Häufigkeitsklasse (HK) des Lemmas in FOLK und in DeReKo. Vergleicht man die beiden eingefärbten Kreisringe, erhält man eine Vorstellung des Frequenzunterschiedes eines Lemmas in den beiden Korpora (je größer der Kreis im Vergleich, desto häufiger das Lemma im

Verhältnis). Das Lemma WISSEN ist in FOLK der HK3 und in DEREKO der HK8 zugeordnet. Die Differenz beträgt daher 5 zu Gunsten von FOLK.[9]

Weitere lexikografische Informationen zu den einzelnen Stichwörtern werden über die externe Verlinkung zu verschiedenen anderen Wörterbüchern mit online-Zugriff angeboten (z. B. DWDS). Ein Kurzhinweis auf relevante Forschungsliteratur ermöglicht die Verknüpfung zu den ausführlichen Literaturangaben, die für die einzelnen Stichwörter der LeGeDe-Ressource relevant sind und als Umtextangebot vorliegt. Zur Verständnissicherung werden außerdem ausgewählte Fachtermini in den Textpassagen mit einem Glossar, welches ebenfalls als Umtext vorliegt, verbunden. So kann man z. B. in dem Übersichtstext zu WISSEN zu dem Begriff *Epistemik* genauere Information im Glossar erhalten. Wie komplex die Vernetzungsstruktur der LeGeDe-Ressource ist, wird schon an dem präsentierten Überblicksartikel zu WISSEN deutlich.

### 4.2.2 Lexikografische Angaben in Modul 1

Im Zentrum von Modul 1 steht die spezifische Information zu den einzelnen Lesarten und deren Kombinatorik, die durch relevante lesartübergreifende Informationen ergänzt wird. Ein weiteres optionales Informationsangebot ermöglicht Kommentare zu Besonderheiten.

#### 4.2.2.1 Lesartübergreifende Information zu den einzelnen Lemmata

(i) Die bearbeiteten Lemmata von Modul 1 treten in einfacher Form, d. h. zunächst als Einwortlemmata auf. Die formal-kategoriale Information erfolgt über folgende grammatische Kategorien:

- Vollverb: *gucken*, *kriegen*, *schauen*, *wissen* [. . .]
- Hilfsverb: *kriegen* [. . .]
- Adjektiv: *gut*, *schön* [. . .]
- Adverb: *eben* [. . .]
- Nomen: *Ahnung* [. . .]

9 Die Darstellung ist zusammengesetzt aus 17 konzentrischen Kreisen, also 16 Kreisringen, die jeweils für eine HK stehen. Das häufigste Wort eines Korpus definiert die HK 0, die in der Abbildung durch den äußersten Kreisring dargestellt ist. Der innerste Kreisring steht für die HK 16 ($1/2^{16}$ so häufig, wie das häufigste Wort), da diese die seltensten Wörter in FOLK repräsentiert. Ein Vergleich mit niedrigeren HK in DEREKO (bis HK 32) ist für unsere Zwecke nicht sinnvoll, daher wurden für die Visulisierung für DEREKO, parallel zu FOLK auch nur 16 HK berücksichtigt. Der Ring, der die HK visualisiert, in der sich das Lemma in FOLK wiederfindet, ist jeweils türkis eingefärbt. Der Ring, der die HK visualisiert, in der das Lemma in unserem Vergleichskorpus DEREKO zu finden ist, ist hingegen grau eingefärbt (vgl. Meliss et al. 2019b).

(ii) Weitere optionale formale Informationen beziehen sich auf die Morphemstruktur bei komplexen lexikalischen Formen (z. B.: *Ahnung*: ahn- (Verbalstamm) – Derivationssuffix: -ung).

(iii) Für ein Informationsangebot zu möglicher Formvarianz u. a. auf phonischer Ebene steht ein optionaler Textblock zur Verfügung (vgl. Bsp. 1)

(1) *kriegen* [Vollverb]
Anstelle der Realisierung eines auslautenden ‚k-Lautes' (Auslautverhärtung) lässt sich eine deutliche Tendenz zu der Realisierung eines ‚ich-Lautes' feststellen [kriecht] [. . .].

(iv) Zur Forschungsliteratur können Kurzverweise, die speziell für diesen Informationsblock relevant sind, aufgeführt werden. Sie verlinken mittels eines Pop-up-Fensters zum Umtext „Literatur".

#### 4.2.2.2 Lesartenspezifische Information

(i) Die einzelnen LZGAs weisen im Allgemeinen unterschiedliche Lesarten auf. So konnten z. B. in der LeGeDe-Stichprobe für *gucken* neun Lesarten und verschiedene Muster (u. a. *guck mal*, *mal gucken, müssen wir mal gucken*) identifiziert werden (vgl. Abb. 3). Gerade zu dem Stichwort GUCKEN konnte festgestellt werden, dass die Diskrepanz zwischen der bisherigen lexikografischen Kodifizierung und der Korpusbefunde sehr groß ist, denn in den gängigen Wörterbüchern konnten nicht mehr als drei Lesarten zu *gucken* dokumentiert werden (Meliss/Möhrs 2017, 2018). Die einzelnen Lesarten werden durch eine **Kurzettikettierung** zur semantischen Disambiguierung gekennzeichnet und ihre Bedeutung durch kurze **Bedeutungsparaphrasen** umschrieben.

(ii) Für alle Lesarten werden illustrative Audiobelege mit einem entsprechenden Transkriptausschnitt angeboten. Diese sind in einen ausführlichen **Belegblock** integriert, zu dem neben einem aussagekräftigen Belegtitel auch eine kurze Beschreibung des Kontextes und optional ein Analysekommentar gehört. Ein Audiosymbol lädt den Benutzenden explizit ein, sich die Datei anzuhören (vgl. Abb. 4).

(iii) Zusätzlich gibt es die Möglichkeit, in einem kurzen Kommentar auf formale Auffälligkeiten hinzuweisen. In diesem optionalen Textfeld können unterschiedliche Aspekte, wie u. a. das auffällige kombinatorische Verhalten in Verbindung mit Modalverben, der Gebrauch von bestimmten Modi oder die Verknüpfung mit bestimmten Partikeln oder deiktischen Ausdrücken kommentiert werden (vgl. Bsp. 2–3).

(2) *gucken* in der Bedeutung von ‚abwarten':
In dieser Lesart ist der Gebrauch der 1. Person Plural Indikativ bzw. des Imperativs in Verbindung mit der Modalpartikel *mal* auffällig. Dies ist als Aufforderung (teilweise auch modalisiert durch das Modalverb *müssen*) an alle Interagierenden zu verstehen, zunächst zu überlegen, nicht zu handeln, sondern abzuwarten.

**richtungsorientiert visuell wahrnehmen** ▶

Mit der Grundbedeutung von **gucken** wird ausgedrückt, dass jemand etwas durch einen visuellen Eindruck wahrnimmt, indem er oder sie den Blick dabei bewusst auf etwas oder jemanden oder auch auf mentale Bezugsgrößen (z.B. Probleme) richtet.

**überprüfen** ▶

Mit **gucken** wird ausgedrückt, dass jemand etwas durch eigene visuelle Wahrnehmung in Verbindung mit einer Recherche 'überprüft' bzw. 'feststellt' oder 'beobachtet'.

**achten auf** ▶

Bei **gucken** in der Bedeutung von 'achten auf' entweder im Sinne von 'kontrollieren', was der Fall ist, bzw. was passieren wird oder im Sinne von 'aufpassen', dass etwas (nicht) passiert oder eingehalten wird, tritt die visuelle Wahrnehmungsbedeutung zugunsten der kognitiven Tätigkeit einer bewussten Aufmerksamkeitslenkung in den Hintergrund.

**abwägen** ▶

Bei **gucken** als 'abwägen' tritt die Grundbedeutung der visuellen Wahrnehmung in den Hintergrund. Im Vordergrund steht eine mental-kognitive Bedeutung im Sinne eines zukunftsorientierten Abwägens.

**ansehen** ▶

Bei **gucken** in der Bedeutung 'sich etwas ansehen/anschauen' richtet sich die visuelle Wahrnehmung darauf, sich etwas (z.B. einen Film) bewusst in einem entsprechenden Medium anzuschauen.

**abwarten** ▶

Bei **gucken** zum Ausdruck von einem zukunftsorientierten Abwarten in Bezug auf das ungewisse Eintreten eines Sachverhaltes oder der Durchführung einer zukünftigen Handlung verblasst die visuelle Wahrnehmungsbedeutung deutlich.

**suchen** ▶

In Verbindung mit einer visuellen Wahrnehmung kann **gucken** in der Bedeutung von 'suchen', d.h. nach einem Gegenstand oder einer Information gezielt 'Ausschau halten', auftreten.

**umherschauen** ▶

Bei **gucken** mit der Bedeutung 'umherschauen' mit dem Ziel, sich etwas beiläufig anzuschauen, wird die nicht zielgerichtete, visuelle Wahrnehmung begleitet durch eine Art des Herumstöberns.

**eine bestimmte Miene zeigen** ▶

Mit **gucken** wird ausgedrückt, dass jemand sich nonverbal, d.h. durch eine bestimmte Miene bzw. einen bestimmten Gesichtsausdruck ausdrückt und dadurch einen bestimmten optischen Eindruck erweckt.

**Abb. 3:** Etikettierungen und Bedeutungsparaphrasen zu *gucken* (Modul 1) (Screenshot).

[2] **gucken** als Aufforderung zu einer richtungsorientierten visuellen Wahrnehmung

0586 **TZ** sieht cool aus °h (.) da gu_mal (.) da (.) genau vor deim kopf jetzt
0587 (0.83)
0588 **PZ** kann ich nich ich muss jetz kurz hier gucken °h (.) damit ich nix anmale
0589 (1.52)
0590 PZ äh
0591 (0.31)
0592 PZ ups was
0593 (0.42)
0594 **TZ** da jetz gu_mal da vor deim kopf da sieht man es (.) °h da is ungleichmä[ßig]
0595 **PZ** [ouh] (.) des da
0596 **TZ** nein
0597 (0.42)
0598 **PZ** ((schmatzt)) °h (.) da (.) ja
0599 **TZ** (.) ja
0600 (0.39)
0601 **PZ** ja da ka_ma ja auch mit ner größeren rolle noch [des is nur]
0602 **TZ** [nein aber wei]sch des da is des is
0603 **PZ** is besser wenn man_s gleich wegmacht ebe[n]

Privat: Gespräch beim Renovieren - FOLK_E_00217_T_03

**Belegkontext**

Zwei Schwestern, PZ und TZ, streichen zusammen das Zimmer der einen Schwester neu.

**Analyse**

Die Sprecherin (TZ) fordert ihre Schwester (PZ) dazu auf, etwas visuell wahrzunehmen, indem sie direkt auf das, was visuell wahrgenommen werden soll, verweist bzw. hindeutet. Diese Zeigegeste wird auch durch den Gebrauch deiktischer Elemente, wie hier dem Lokaladverb **da** deutlich.

**Abb. 4:** Belegblock zu *gucken* in der Bedeutung ‚richtungsorientiert visuell wahrnehmen' (Modul 1) (Screenshot).

(3) *schauen* in der Bedeutung von ‚überprüfen':
Diese Bedeutungsvariante von *schauen* zeigt eine deutliche Verlagerung des Fokus von der visuellen Wahrnehmung auf den Aspekt des Überprüfens bzw. Feststellens auf. In der Hälfte der Belege wird außerdem durch die Kombination mit den Modalpartikeln *mal* und *halt* eine gewisse Beiläufigkeit bzw. Unverbindlichkeit durch den Sprechenden ausgedrückt.

(iv) Die relevante Information zu der **Kombinatorik** in Verbindung mit den einzelnen Lesarten wird durch die Beschreibung der Strukturmuster, der festen Verbindungen sowie der Kollokationen und der interaktionalen Einheiten explizit gemacht.

Die Strukturmuster werden zunächst durch eine abstrahierte, formelhafte Information angeboten. Die einzelnen Aktanten, aus denen die Strukturmuster zusammengesetzt sind, werden bezüglich ihrer semantischen Rolle, ihrer syntaktischen Funktion und den morphosyntaktischen Realisierungsmöglichkeiten genauer erklärt und Besonderheiten angegeben (vgl. Bsp. 4 zu *schauen*). Kurze Belegblöcke illustrieren die jeweiligen kombinatorischen Möglichkeiten (vgl. Abb. 5)

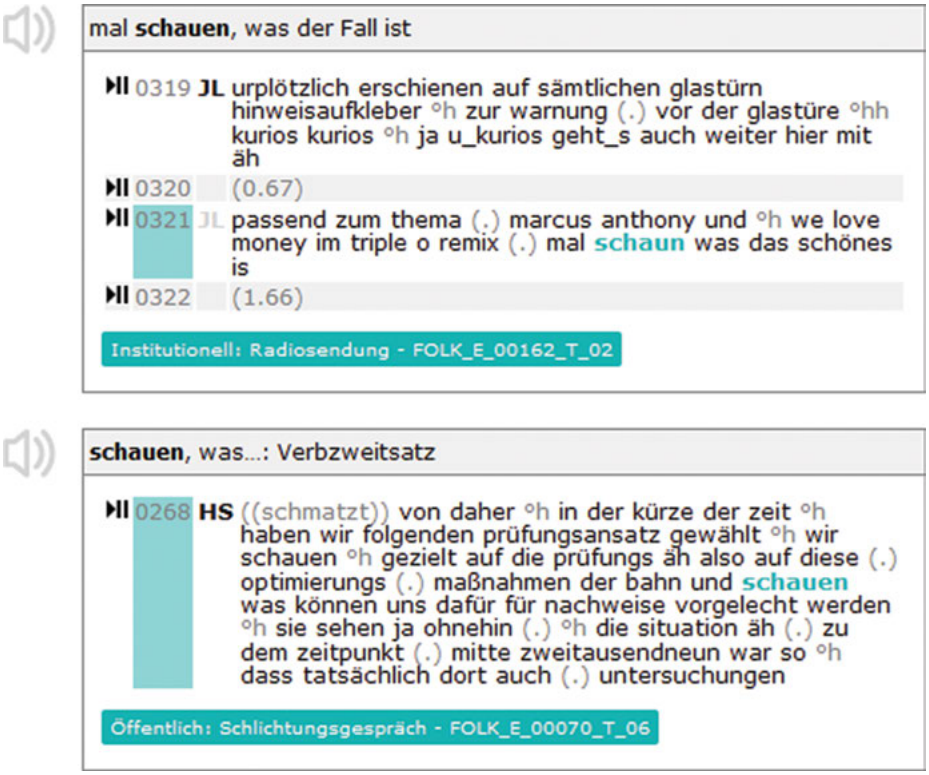



**Abb. 5:** Ausschnitt: Kurzbelegblöcke zu *schauen* in der Bedeutung ‚überprüfen' (Modul 1) (Screenshot).

(4) *schauen* in der Bedeutung ‚überprüfen':
**jemand *schaut* (was der Fall ist)** [irgendwo]

*jemand:*
Person, die etwas überprüft/feststellt
syntaktische Funktion: Subjekt

*was der Fall ist:*
Sachverhalt, der an einem bestimmten Ort überprüft bzw. festgestellt wird
syntaktische Funktion: satzförmiges Komplement: fakultativ

Realisierungsform:
- indirekte Entscheidungsfrage eingeleitet mit *ob*
- indirekte Ergänzungsfrage eingeleitet durch ein *w-Element*
- abhängiger, uneingeleiteter Verbzweitsatz

Besonderheiten:
- die indirekten Ergänzungsfragen weisen teilweise eine Verbzweitposition auf
- wenn das satzförmige Komplement nicht realisiert wird, steht das Verbalsyntagma (z. B.: *ich schau mal*) meist am Ende einer Sequenz und bildet einen thematischen Abschluss mit teilweise spezifischen interaktionalen Funktionalitäten.

*irgendwo:*
Ort der visuellen Wahrnehmung zur Überprüfung bzw. Feststellung von etwas
syntaktische Funktion: Situativsupplement: lokal

Realisierungsform:
- meistens in Form von deiktisch gebrauchten Lokaladverbien wie *hier*, *da*, *dort*

Unter dem weiten Oberbegriff „Feste Wendungen/Kollokationen" werden in der LeGeDe-Ressource verschiedene Arten von mehr oder weniger festen lexikalischen Einheiten und (usuelle) Verbindungen, ohne eine weitere Spezifizierung bzw. terminologische Präzisierung vorzunehmen, zusammengefasst. So stehen hier Kollokationen, feste (idiomatische) Wendungen, Routineformeln, Sprichwörter etc. nebeneinander (vgl. Bsp. 5–8) und werden auch durch entsprechende Belegblöcke illustriert.

(5) *wissen* in der Bedeutung von ‚informiert sein':
**Bescheid wissen**: Mit dieser Wendung geben Sprechende zu verstehen, dass sie selber oder eine andere Person tatsächlich über einen nicht näher benannten Sachverhalt informiert sind oder stellen dies in Frage.

(6) *gut* in der Bedeutung ‚positiv bewertend':
Vor das Adjektiv *gut* treten in der LeGeDe-Stichprobe weitere Adjektive bzw. Adverbien, die die Bedeutung von *gut* entweder relativieren (z. B. *ganz*) oder intensivieren (z. B. *sehr* und *echt*).

(7) *gut* in der Bedeutung ‚positiv bewertend':
**guten Appetit:** Die Routineformel *guten Appetit* wird von Sprechenden typischerweise vor Beginn einer Mahlzeit als eine Art Wunschformel geäußert, womit auch der Beginn des gemeinsamen Essens (als Vorgang) markiert wird. Im Transkriptausschnitt zeigt sich außerdem eine typische Paarsequenz, die bei der Verwendung dieser Formel in der Interaktion zu beobachten ist [. . .]

(8) *schauen* in der Bedeutung ‚richtungsorientiertes visuelles Wahrnehmen':
**Einem geschenkten Gaul schaut man nicht ins Maul**:Dieses Sprichwort bedeutet, dass man, wenn man etwas geschenkt bekommt, damit zufrieden sein und das Geschenk nicht kritisch überprüfen oder hinterfragen sollte.

Durch die Auflistung in Modul 1 von interaktionalen Einheiten, die in der LeGeDe-Stichprobe in direkter inhaltlicher Verbindung mit einzelnen Lesarten des entsprechenden Stichwortes belegt werden konnten, wird die Einrichtung einer direkten Bedeutungs-Funktions-Schnittstelle zu interaktionalen Funktionalitäten ermöglicht, die in Modul 2 genauer beschrieben werden (vgl. Bsp. 9–11).

(9) *Ahnung* in der Bedeutung ‚Wissen':
*keine Ahnung:* Anzeige von Nicht-Wissen: Das Muster *keine Ahnung* fungiert als interaktionale Einheit, mit der Sprechende anzeigen, dass sie kein Wissen über einen Sachverhalt besitzen.

(10) *wissen* in der Bedeutung ‚informiert sein':
*ich weiß nicht:* Unsicherheitsmarker: Mit *ich weiß nicht* drücken Sprechende aus, dass sie sich hinsichtlich der Wahrheit, Genauigkeit, Gültigkeit bzw. Angemessenheit einer Aussage unsicher sind und markieren die Vagheit dieser Angabe.
*ich weiß:* Beanspruchung von Wissen: Das Muster *ich weiß* fungiert als Anzeige von Wissen bei Reaktion auf eine Äußerung oder als Beanspruchung von Wissen. Sprechende bekennen sich mit diesem Muster dazu, informiert zu sein.

(11) *schön* in der Bedeutung ‚passend':
*schön:* Diskurspartikel: Bewertung: Zu der Verwendung als Adjektiv mit der Bedeutung ‚passend' tritt in der LeGeDe-Stichprobe auch die Verwendung von *schön* als Diskurspartikel auf. Dabei nehmen Sprechende mit *schön* Bezug auf eine vorausgegangene Äußerung und bewerten den Inhalt dieser Äußerung, die Äußerung an sich oder eine Handlung als positiv. Sie drücken so auf inhaltlicher Ebene ihre persönliche Haltung dem jeweiligen Bewertungsgegenstand gegenüber aus.

In einem weiteren optionalen Textfeld ist es außerdem möglich, auf Auffälligkeiten in Bezug auf ausgewählte Metadaten und ihre Frequenz hinzuweisen (vgl. Bsp. 12).

(12) *schauen* in der Bedeutung ‚richtungsorientiert visuell wahrnehmen':
Etwa ein Drittel aller Belege der LeGeDe-Stichprobe zu *schauen* konnten dieser Lesart zugeordnet werden. Dabei können mehr als die Hälfte dieser Belege privaten Interaktionssituationen zugerechnet werden (Alltagsgespräche: Spieleinteraktion, beim Kochen, beim Renovieren, beim Spazierengehen etc.). In diesem Kontext kommt der situationsbezogene imperativische Gebrauch in Verbindung mit einer lokalen Deixis besonders deutlich zum Ausdruck. In öffentlichen Interaktionssituationen konnte diese Bedeutung von *schauen* hingegen kaum belegt werden.

### 4.2.3 Lexikografische Angaben in Modul 2

Im Zentrum von Modul 2 stehen neben allgemeiner funktionsübergreifender Information v. a. die funktionsspezifischen Besonderheiten, die durch Informationen zur Syntax/Sequenz und Prosodie bereichert werden. Auch für dieses Modul stehen ausführliche **Belegblöcke** zur Verfügung, die die Transkriptausschnitte[10] in Verbindung mit der entsprechenden Audiodatei anbieten.

#### 4.2.3.1 Funktionsübergreifende Information

(i) Gegenstandsbereich von Modul 2 sind sowohl Einwortlemmata (z. B. *eben*[11], *gut, ja*) als auch Mehrwortlemmata (z. B. *keine Ahnung*[12], *weiß nicht*, *mal gucken*) (vgl. Meliss 2020b). Die formal-kategoriale Information erfolgt über folgende Kategorien:

- Nominalsyntagma: *keine Ahnung* [. . .]
- Verbalsyntagma: *ich weiß nicht*, *weißt du, mal gucken* [. . .]
- Diskurspartikel: *ja, gut* [. . .]
- Modalpartikel: *eben, halt, ja* [. . .]

(ii) Bei komplexen Formen werden weitere genauere formale Informationen bezüglich der beteiligten Elemente aufgeführt (vgl. Bsp. 13 + 14):

(13) *keine Ahnung*: Determinativum *keine* und Substantiv *Ahnung*

(14) *weißt du:* Phrase aus der 2. P. Sg. des Verbs *wissen* (weißt) und dem Personalpronomen *du*

---

**10** Die entsprechenden Belege wurden nach GAT2-Basistranskript nachtranskribiert. Zu den Transkriptionskonventionen GAT2 vgl. Selting et al. 2009.

**11** Zu *eben* siehe auch Möhrs/Torres (2020).

**12** Zu *keine Ahnung* siehe auch Möhrs (2020).

(iii) Die Information zu möglichen formalen Varianten erfolgt u. a. auf der lautlichen (vgl. Bsp. 15, 16) und der kompositionellen Ebene (vgl. Bsp. 17):

(15) *eben*: [ebent] – Epithese eines stimmlosen [t]

(16) *wissen*: [weißt du], [weißte], [weiß_u], [weiß_du] – teils Klitisierung und phonetische Reduktion

(17) *(ich) weiß nicht* / *weiß (ich) nicht*

(iv) Eine Auflistung von in der LeGeDe-Stichprobe belegten Möglichkeiten zur Kombinatorik wird meistens mit einem Kommentar begleitet (vgl. Bsp. 18–20)

(18) *eben*: (vgl. auch Abb. 6)
dann eben, halt eben, ja eben, eben auch, eben dann, eben nicht
Manche Partikelkombinationen können eine Funktionsverstärkung mit sich bringen bzw. ihre eigene spezifische Funktion eröffnen (z. B. Verweis auf gemeinsames Wissen durch *ja*).

(19) *halt*:
dann halt, halt eben, halt dann, halt jetzt
Die meisten kombinierten Elemente werden rechts von der Partikel realisiert. In der linken Peripherie findet sich in der LeGeDe-Stichprobe lediglich die Partikel *dann*.

(20) *ja*:
*ja klar, ja natürlich, ja aber;*
Durch die Kombination mit weiteren Partikeln wie *klar* und *natürlich* kann der Grad der Zustimmung in der Funktion ‚Zustimmungssignal' erhöht werden. Die Kombination *ja aber* ist eine verfestigte Konstruktion mit der eigenen interaktionalen Funktion ‚Einräumung'.

(v) Kurzverweise auf relevante Literaturangaben, die speziell für diesen Informationsblock, d. h. für die Information zu der interaktionalen Funktion, relevant sind, werden ebenfalls funktionsübergreifend angeboten (vgl. Abb. 6 zu *eben*).

(vi) In einem optionalen Textblock können außerdem funktionsübergreifende Informationen zu Metadaten angeführt werden (vgl. Abb. 6 zu *eben*).

**eben** [MODALPARTIKEL]

| Form | Varianz | Kombinatorik | Forschungsliteratur |
|---|---|---|---|
| --- | eben<br>[ebent]<br>Epithese eines stimmlosen [t] | dann **eben**, halt **eben**, ja **eben**, **eben** auch, **eben** dann, **eben** nicht<br>*Manche Partikelkombinationen können eine Funktionsverstärkung mit sich bringen bzw. ihre eigene spezifische Funktion eröffnen (z.B. Verweis auf gemeinsames Wissen durch **ja**).* | • Blühdorn (2019)<br>• Helbig (1990)<br>• Helbig/Helbig (1995)<br>• Müller (2016)<br>• Thurmair (1989)<br>• Torres Cajo (2019)<br>• Weydt (1969) |

Auffälligkeiten:

In der LeGeDe-Stichprobe wird **eben** auffällig häufig in institutioneller bzw. öffentlicher sowie in handlungsorientierter Alltagsinteraktion verwendet (Prüfungs- und Schlichtungsgespräche, Interviews, Lehr-Lern-Kontexte und Spieleinteraktionen).

**Abb. 6:** Funktionsübergreifende Information zu *eben* als Modalpartikel (Modul 2) (Screenshot).

#### 4.2.3.2 Funktionsspezifische Informationen

(i) Jede einzelne Funktion, die einem ein- oder mehrteiligen Lemma zugeschrieben werden kann, wird entsprechend etikettiert. So weist z. B. *was weiß ich* als komplexe Mehrworteinheit ein breites funktionales Spektrum auf. Eine Funktionskurzbeschreibung soll dazu beitragen, die einzelnen Möglichkeiten voneinander abzugrenzen (vgl. Abb. 7).

**etcetera-Formel** ▸

Mit **was weiß ich** zeigen Sprechende an, dass eine Aufzählung potenziell fortführbar ist, schließen diese Aufzählung mit **was weiß ich** jedoch zugleich ab und strukturieren somit ihren Redebeitrag.

**Markierung eines Beispiels** ▸

Mit **was weiß ich** zeigen Sprechende an, dass das folgende Beispiel ein ungefähr passendes aus einer Menge potenzieller Beispiele darstellt, das zur Illustration der Aussage ausreicht.

**Unsicherheitsmarker** ▸

Mit **was weiß ich** zeigen Sprechende an, dass sie sich hinsichtlich der Wahrheit/Genauigkeit einer Aussage unsicher sind und markieren die Vagheit dieser Angabe.

**Abb. 7:** Funktionsspezifische Information zu *was weiß ich*: Funktionskurzbeschreibung (Modul 2) (Screenshot).

Im Falle einer semantischen Verknüpfung zu einer der Lesarten des entsprechenden Einwortlemmas wird ein Querverweis hergestellt und über die Etikettierung auf die entsprechende Lesart, die in Modul 1 genauer beschrieben wird, verwiesen und verlinkt (vgl. Bsp. 21–23).

(21) *keine Ahnung:*
**Anzeige von Nicht-Wissen**
Mit *keine Ahnung* zeigen Sprechende an, dass sie kein Wissen über einen Sachverhalt besitzen. (vgl. dazu auch *Ahnung* in der Bedeutung ‚Wissen')

(22) *ich weiß nicht*
**Unsicherheitsmarker**
Mit *ich weiß nicht* zeigen Sprechende an, dass sie sich hinsichtlich der Wahrheit/Genauigkeit einer Aussage unsicher sind und markieren die Vagheit dieser Angabe. (vgl. dazu auch *wissen* in der Bedeutung ‚informiert sein')

(23) *mal schauen*
**Verschiebung einer Handlung in die Zukunft**
Mit *mal schauen* schieben Sprechende eine Handlung, Entscheidung, Einschätzung, Bewertung, Planung o. Ä. zeitlich auf. Damit dient *mal schauen* häufig als (vorläufiges) Beendigungssignal (vgl. dazu auch *schauen* in den Bedeutungen ‚abwarten' und ‚überprüfen').

(ii) Ein **Langbelegblock** mit entsprechendem Belegtitel, Belegkontext, funktional ausgerichteter Beleganalyse und einem Link zur Audiodatei dient zur Illustration und wird auch visuell deutlich durch ein Audiosymbol markiert, um den Benutzenden explizit darauf hinzuweisen, sich die Datei auch anzuhören (vgl. Abb. 8 + 9).

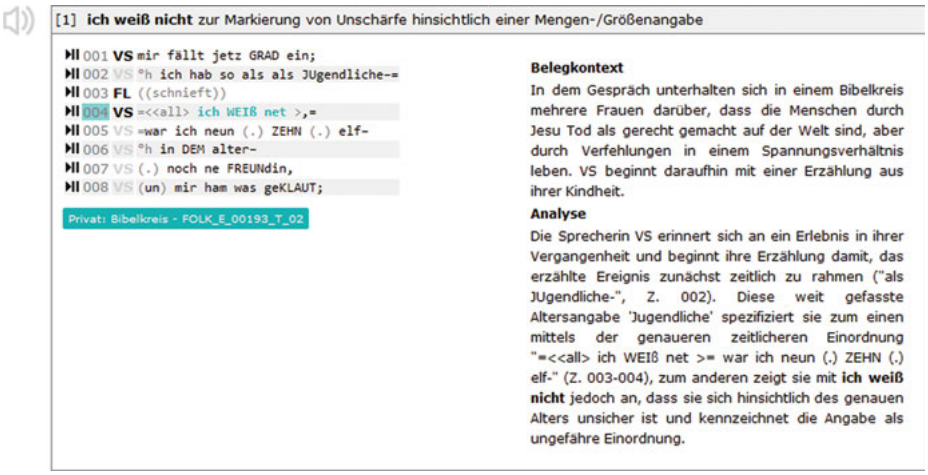



**Abb. 8:** Belegblock für *ich weiß nicht* als Unsicherheitsmarker (Modul 2) (Screenshot).

[1] **guck mal** zur Salienzerhöhung und Fokussierung

001 **FR** er !HAT! mich ja sogar selber gefagt;=
002 FR =ob ich das überhaupt SCHAFfe zeitlich,
003 FR un dann hab [ich gSAGT-=]
004 **EG** [ja, ]
005 **FR** =ja ich denke scho:n,
006 FR aber ich HAB schon viel zu tun;
007 FR °h un wenn ich dann sage ich mach es IMmer noch gerne abe:r;
008 FR (.) °h ähm ende FEbrua:r,
009 FR äh kann ich würd des WIRKlich machen?=
010 FR =weil des STIMMT,
011 FR ich hab dann auch nich mehr so VIEL zu tun,
012 FR °h (.) weil GUCK mal,=
013 FR =des proBLEM is jetz auch,
014 FR das hatt ich nämlich gar nich beDACHT,
015 FR ich bin jetz auch jeden tag bei der ((unverständlich)) gewesen,
016 FR einfach DAdurch dass ich ja-
017 FR °hh jetz für ne woche WEGfahre,
018 FR dann fällt (.) dieser ZEITraum weg;
019 FR (.) in dem ich (.) dann normalerweise immer dran ARbeite;

Privat: Telefongespräch - FOLK_E_00084_T_02

**Belegkontext**

Zwei Freundinnen sprechen am Telefon miteinander. FR erklärt ihrer Freundin die Gründe dafür, dass sie eine Ausarbeitung nicht rechtzeitig bei ihrem Dozenten abgeben kann.

**Analyse**

Zu Beginn des Ausschnitts rekonstruiert FR ein Gespräch mit Ihrem Dozenten, in dem sie ihm zugesichert hatte, dass sie ihre Arbeit rechtzeitig abgeben könnte (Z. 001-005). Sie räumt ein, dass sie die Arbeit Ende Februar tatsächlich angehen würde, da sie "dann auch nich mehr so VIEL zu tun" hat (Z. 011). Daraufhin liefert sie eine Erklärung für ihr derzeitiges Zeitproblem, das sie beim vorhergehenden Gespräch mit ihrem Dozenten nicht bedacht hatte: Sie ist täglich verhindert und fährt zudem für eine Woche in den Urlaub (Z. 012-019). Diese Rechtfertigung wird mit **guck mal** eingeleitet ("weil GUCK mal,= =des proBLEM is jetz auch, [...]", Z. 012ff.), wodurch die mentale Aufmerksamkeit des Gegenübers auf den Folgebeitrag fokussiert wird. Dadurch wird die Salienz des folgenden Arguments erhöht, es bekommt in der Rechtfertigungssequenz mehr Gewicht und Relevanz.

**Abb. 9:** Belegblock für *guck mal* als Diskursmarker (Modul 2) (Screenshot).

(iii) Die Funktionsabstrahierung (vgl. Abb. 10 + 11) fasst die jeweils beschriebene Funktion auf einer allgemeineren Ebene zusammen. Das verwendete Symbol, welches an dieser Stelle als Strukturanzeiger verwendet wird (Engelberg/Lemnitzer 2009: 155 f., Wiegand 1989b: 428) illustriert, dass es sich hierbei um zentrale Informationen handelt.

Mit **ich weiß nicht** als Unsicherheitsmarker markieren Sprechende, dass sie sich hinsichtlich des genauen Zutreffens einer oder mehrerer Angaben unsicher sind und sich nicht festlegen, die genannten Angaben jedoch ungefähr zutreffen und zum beabsichtigten Verständnis des Gesagten ausreichen. Der/die Sprechende erhebt mit **ich weiß nicht** somit keinen Anspruch auf die Richtigkeit seiner/ihrer Angabe und beansprucht keine epistemische Autorität. Meist wird damit gleichzeitig angezeigt, dass die Äußerung bzw. die genaue Angabe für das eigentliche, zentrale Gesprächsthema oder die Sprechhandlung nicht bzw. wenig relevant ist. Dies betrifft graduierbare Einheiten wie Mengen-, Zeit- oder Größenangaben oder Aussagen zu Sachverhalten, die durch den Gebrauch von **ich weiß nicht** als Schätzungen, ungefähr passende Angaben oder Vermutungen markiert werden. Dies wird teils durch Nennung mehrerer Möglichkeiten oder durch weitere Unschärfemarker (**ungefähr**, **oder so**) zusätzlich kenntlich gemacht.

**Abb. 10:** Funktionsabstrahierung: *ich weiß nicht* als Unsicherheitsmarker (Modul 2) (Screenshot).

**guck mal** wird von Sprechenden als Diskursmarker eingesetzt, um die gemeinsame Ausrichtung der Interagierenden zu koordinieren und die Relevanz einer folgenden Äußerung zu erhöhen. Es tritt häufig vor längeren Redebeiträgen eines/-r Sprechenden und/oder vor gewichtigen Argumenten auf. Aufgrund der syntaktischen Unabgeschlossenheit projiziert **guck mal** einen Folgebeitrag, was zur Rederechtssicherung des/der Sprechenden beiträgt. Die ursprüngliche Semantik von **guck mal** ('visuell wahrnehmen') ist stark verblasst, stattdessen wird die mentale Aufmerksamkeit aktiviert und gelenkt ('mental wahrnehmen').

**Abb. 11:** Funktionsabstrahierung: *guck mal* als Diskursmarker (Modul 2) (Screenshot).

(iv) Kurze Informationen zur **Syntax- und Sequenz-Realisierung** (vgl. Abb. 12 + 13) und zur **Prosodie** (vgl. Abb. 14 + 15) die, wie bei *ja* und *eben* als Modalpartikel vs. Diskurspartikel distinktiv sein können, werden ebenfalls geliefert und mit kurzen Transkriptausschnitten belegt.

**Syntax-Sequenz-Realisierung**

**ja** wird immer turnmedial, syntaktisch integriert und im Mittelfeld der Äußerung realisiert.

**ja** syntaktisch integriert, Mittelfeld

```
001 ST durch das TEXTverstehen,
002 ST also man man LIEST ja nen nen;
003 ST (.) literarischen TEXT?
004 ST und in diesem literarischen TEXT,
005 ST (.) wird einem immer °h n weltentwurf (.)
        VORgeschlagen;= ne?
```

Institutionell: Prüfungsgespräch in der Hochschule - FOLK_E_00033_T_01

**Abb. 12:** Syntax-Sequenz-Realisierung: *ja* als Modalpartikel zur Wissensaktualisierung (Modul 2) (Screenshot).

**Syntax-Sequenz-Realisierung**

**ja** kommt immer turninitial vor, ist syntaktisch desintegriert und bildet eine eigene TCU.

**ja** turninitial, syntaktisch desintegriert

```
001 MS ich komm mir immer vor wie_n SCHULmädchen wenn ich nix
       weiß,
002    (0.86)
003 AW ((Lachansatz))
004    (1.12)
005 MS des könnt ma jetz AUCH mal analysieren,
006 MS [°hh ]
007 SZ [hm:_HM- ]
008    (2.54)
009 AW °hhh (.) ja:-
010    (0.56)
011 AW also TEAMkonflikte ham mer keine,
```

Institutionell: Meeting in einer sozialen Einrichtung - FOLK_E_00024_T_03

**Abb. 13:** Syntax-Sequenz-Realisierung: *ja* als Diskurspartikel: Diskursmarker (Modul 2) (Screenshot).

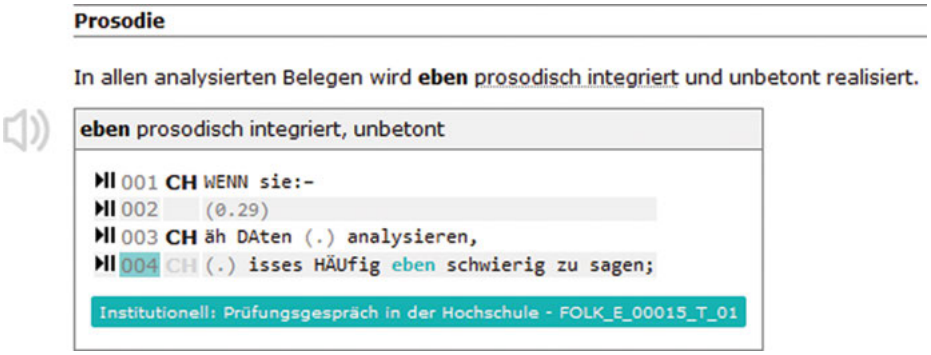



**Abb. 14:** Prosodie: *eben* als Modalpartikel zur Kohärenzherstellung (Modul 2) (Screenshot).

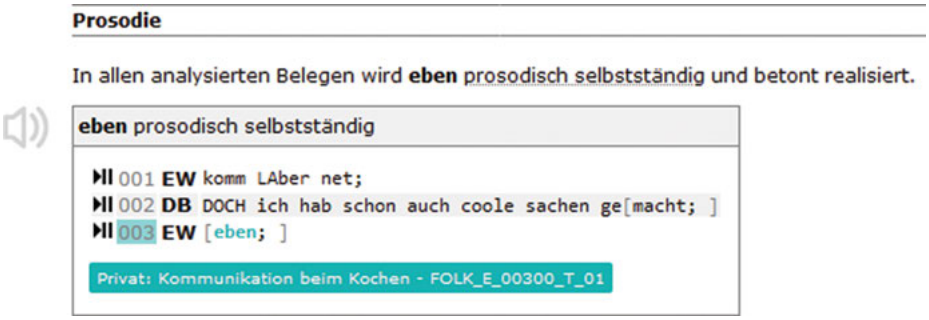



**Abb. 15:** Prosodie: *eben* als Diskurspartikel: Zurückforderung eines Wissensanspruch (Modul 2) (Screenshot).

(v) Informationen zu weiteren Auffälligkeiten, die hauptsächlich in Verbindung mit den Metadaten stehen, können in einem optionalen Textfeld bereitgestellt werden (vgl. Bsp. 24).

(24) *schön*: Diskurspartikel: Abschlussmarker
Die Verwendung von *schön* als Abschlussmarker tritt überwiegend in handlungsschematischen Interaktionen wie u. a. Unterrichtsstunden, Planungen oder Spielinteraktionen auf, in denen einzelne Themen und Handlungen systematisch nacheinander bearbeitet werden und die aufeinander aufbauen.

## 4.3 Multimediales Informationsangebot

Bei der Präsentation des Informationsangebotes in der LeGeDe-Ressource wurde versucht, möglichst viele der Modalitäten, die von den Befragten unserer Umfragen erwünscht wurden, aufzugreifen (Meliss/Möhrs/Ribeiro Silveira 2018b: 129 ff., 2019a: 113 ff.). Neben dem allgemeinen Wunsch nach der Inkorporierung von Audiodateien (vgl. Abb. 16) wurden ebenfalls auf Wunsch der Befragten Transkripte in den Belegblöcken und ein vielseitiges internes (u. a. zu den Umtexten: Glossar und Literatur) und externes Verlinkungsangebot (u. a. zu FOLK, dem Lexical Explorer und anderen Wörterbüchern) (vgl. Abb. 17) in die Ressource „eingebaut“.

**Abb. 16:** Visualisierung der Ergebnisse der OU zu möglichen Präsentationsmodalitäten.

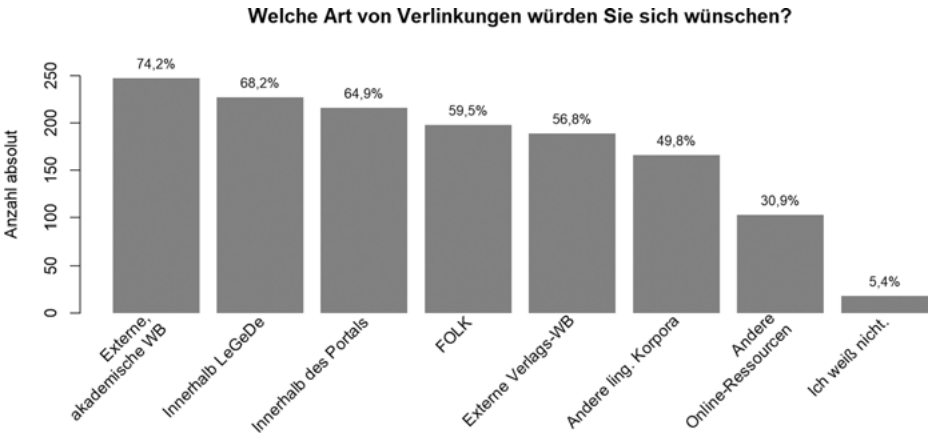



**Abb. 17:** Ergebnisse der OU zu Verlinkungsmöglichkeiten.

# 5 Innovation, Anwendung und Ausblick

Da das LeGeDe-Projekt auf keine Vorbilder zurückgreifen konnte, kann durchaus behauptet werden, dass die Tatsache, einen lexikografischen Prototypen zur Darstellung von Spezifika der gesprochensprachlichen Lexik des Deutschen auf der Datengrundlage eines Korpus des Gesprochenen in der Interaktion erstellt zu haben, äußerst innovativ ist. In Verbindung mit diesem Vorhaben gab es Anlass zu einer Reihe von methodologischen und technischen Herausforderungen. Dazu gehören u. a. die quantitative Ermittlung der relevanten Stichwortkandidaten durch ein eigens entwickeltes methodologisches Vorgehen (Meliss et al. 2018a, Möhrs 2021) sowie die qualitative Analyse und anschließende Strukturierung der Daten für die lexikografische Umsetzung. Das Spektrum an lexikografischen Angabeklassen, welches klassische Angaben mit neuartigen Angabeklassen kombiniert, die eigens für die Beschreibung bestimmter lexikalischer Phänomene in der Interaktion entwickelt wurden, ist multimodular. Damit wurde zum ersten Mal für die Darstellung lexikalischer Phänomene des Gesprochenen in der Interaktion ein Vorschlag für neuartige Angabeklassen entwickelt, der ermöglicht, die spezifischen Phänomene für lexikografische Zwecke adäquat zu strukturieren und zu beschreiben. Die Integration von authentischen Korpusbelegen erfolgt zunächst über ausgewählte Transkriptausschnitte, über die eine Schnittstelle zu den Audiodateien und zu den ausführlichen Informationen zu den Metadaten der DGD hergestellt wird. Damit ist die LeGeDe-Ressource eine der wenigen lexikografischen Ressourcen, die eine direkte Verknüpfung zu den entsprechenden Korpusdaten in Audioformat aufweist und gleichzeitig außerdem das Mitlesen der Transkriptausschnitte ermöglicht. In diesem Sinne zeichnet sich der multimediale Charakter der Ressource auch dadurch aus, dass für die Korpusdaten neben den Transkripten Audiodateien und teilweise auch entsprechende Videodateien über den Zugriff auf die DGD zur Verfügung stehen. Die Verknüpfung zu dem Tool ‚Lexical Explorer' stellt den Benutzenden außerdem einen direkten Zugriff auf die der Ressource zugrundeliegenden Korpusdaten und eigene erweiterte quantitative Analysemöglichkeiten bereit. Die völlige Neukonzipierung einer lexikografischen Ressource zur Darstellung von gesprochensprachlichen Spezifika ermöglichte in besonderem Maße die konkrete Berücksichtigung von bestimmten empirisch erhobenen Erwartungshaltungen der zukünftigen Benutzenden an so eine neuartige Ressource.

Ausgehend von den Daten, die die LeGeDe-Ressource bereitstellt, könnten in Zukunft aus anwendungsorientierter Perspektive heraus u. a. korpusbasierte Lernmaterialien zur konkreten Behandlung spezifischer Phänomene der Lexik in der mündlichen Interaktion für die Nutzung im DaF-, DaZ- und DaM-Unterricht entstehen. Diese Möglichkeit geht konform mit den Anworten der OU zu der künftigen Zielgruppe einer Ressource zum Gesprochenen, da DaF-Lehrende und Lernende sowohl von den L1- als auch von den L2-Befragten als Hauptzielgruppe gesehen wurden (Meliss/Möhrs/Riberiro Silveira 2019a: 108 f.). So kann man durchaus behaupten, dass die LeGeDe-Res-

source dank der unterschiedlichen Zugriffe auf die Korpusdaten und das strukturierte und selegierte, multimediale Informationsangebot neben ihrem primären Ziel der sprachwissenschaftlichen Dokumentation vielfältige Möglichkeiten für anwendungsorientierte Szenarien bietet. Das Glossar und die Benutzungshinweise, die mit vielen ausführlichen Informationen für die Benutzenden als Außentexte zur Verfügung stehen, leisten dazu u. a. eine optimale Hilfestellung. Bedingung für eine adäquate anwendungsorientierte Nutzung ist allerdings, dass die SprachdidaktikerInnen diese Nutzungsmöglichkeiten kennen (vgl. Fandrych/Meißner/Wallner 2018, Meliss/Möhrs 2018, Meliss 2021).

Die Bearbeitung derjenigen Phänomenbereiche, die in dem LeGeDe-Prototypen der ersten Arbeitsphase nicht aufgenommen werden konnten[13], könnten zusammen mit der quantitativen Erweiterung der Lemmaanzahl zu den schon behandelten Phänomenen ein breites zukünftiges lexikografisches Arbeitsfeld bieten. Auch die zusätzliche Entwicklung einer erweiterten Suchfunktion u. a. über kategoriale Parameter und kommunikative Funktionen würde besonders dem Wunsch der L2-Befragten der OU entgegenkommen (Meliss/Möhrs/Riberiro Silveira 2019a: 111 f.) und könnte somit gleichzeitig eine Verbindung zu Nachschlagehandlungen für Produktionszwecke im L2-Bereich herstellen. Neben diesen konkreten Desiderata für mögliche weiterführende Arbeiten wäre es wünschenswert, wenn die theoretischen, methodologischen und technischen Grundlagen, die zur Erstellung des LeGeDe-Prototypen entwickelt wurden, in zukünftigen korpusbasierten lexikografischen Projekten zum Gesprochenen in der Interaktion gewinnbringend genutzt und weiterentwickelt werden könnten (Meliss/Möhrs 2020: 29 ff.). Auf diese Weise kann der Kreis von lexikologischer und interaktionslinguistischer Forschung, lexikografischer Umsetzung und Anwendung durch ihre multiplen Wechselwirkungen sinnvoll geschlossen werden.

# Literaturangaben

## Forschungsliteratur

Barcala, Mario/Domínguez, Eva/Fernández, Alba/Rivas, Raquel/Santalla, Maria Paula/Vázquez, Victoria/Villapol, Rebeca (2018): El corpus ESLORA de español oral: diseño, desarrollo y explotación. In: *CHIMERA: Romance Corpora and Linguistic Studies* 5.2, 217–237.

Couper-Kuhlen, Elizabeth/Selting, Margret (2018): *Interactional Linguistics. Studying language in social interaction*. Cambridge etc.: Cambridge University Press.

**13** Weitere lexikalische Phänomenbereiche (wie z. B. der Bereich der Wortbildung oder der Bereich der Interjektionen) bilden ebenfalls höchst interessante Felder für die lexikografische Umsetzung von gesprochensprachlichen Spezifika (cfr. Meliss 2020a). Für weitere Ausbau- und Erweiterungsmöglichkeiten soll besonders auf die Ausführungen in Meliss/Möhrs (2020) hingewiesen werden.

Deppermann, Arnulf (2007): *Grammatik und Semantik aus gesprächsanalytischer Sicht*. Berlin/New York: de Gruyter.

Deppermann, Arnulf/Proske, Nadine/Zeschel, Arne (2017) (Hrsg.): *Verben im interaktiven Kontext. Bewegungsverben und mentale Verben im gesprochenen Deutsch*. Tübingen: Narr.

Eichinger, Ludwig M. (2017): Gesprochene Alltagssprache. In: Deutsche Akademie für Sprache und Dichtung/Union der deutschen Akademien der Wissenschaften (Hrsg.): *Vielfalt und Einheit der deutschen Sprache. Zweiter Bericht zur Lage der deutschen Sprache*. Tübingen: Stauffenburg, 283–331.

Engelberg, Stefan (2015): Gespaltene Stimuli bei Psych-Verben. Kombinatorische Mustersuchen in Korpora zur Ermittlung von Argumentstrukturverteilungen. In: Engelberg, Stefan, et al. (Hrsg.): *Argumentstruktur zwischen Valenz und Konstruktion*. Tübingen: Narr, 469–491.

Engelberg, Stefan (2018): The argument structure of psych-verbs: A quantitative corpus study on cognitive entrenchment. In: Boas, Hans/Ziem, Alexander (Hrsg.): *Constructional approaches to syntactic structures in German*. Boston/Berlin: de Gruyter, 47–84.

Engelberg, Stefan (2019): Argumentstrukturmuster. Ein elektronisches Handbuch zu verbalen Argumentstrukturen im Deutschen. In: Czicza, Dániel/Dekalo, Volodymyr/Diewald, Gabriele (Hrsg.): *Konstruktionsgrammatik VI. Varianz in der konstruktionalen Schematizität*. Tübingen: Stauffenburg, 13–38.

Engelberg, Stefan/König, Svenja/Proost, Kristel/Winkler, Edeltraud (2011): Argumentstrukturmuster als Konstruktionen? Identität – Verwandtschaft – Idiosynkrasien. In: Engelberg, Stefan/Holler, Anke/Proost, Kristel (Hrsg.): *Sprachliches Wissen zwischen Lexikon und Grammatik*. Berlin/New York: de Gruyter, 71–112.

Engelberg, Stefan/Lemnitzer, Lothar (2009): *Lexikographie und Wörterbuchbenutzung*. 4. überarbeitete und erweiterte Auflage. Tübingen: Stauffenburg.

Engelberg, Stefan/Müller-Spitzer, Carolin/Schmidt, Thomas (2016): Vernetzungs- und Zugriffsstrukturen. In: Klosa, Annette/Müller-Spitzer, Carolin (Hrsg.): *Internetlexikografie. Ein Kompendium*. Berlin/Boston: de Gruyter, 111–152.

Engelberg, Stefan/Storrer, Angelika (2016): Typologie von Internetwörterbüchern und -portalen. In: Klosa, Annette/Müller-Spitzer, Carolin (Hrsg.): *Internetlexikografie. Ein Kompendium*. Berli/Boston: de Gruyter, 31–63.

Fandrych, Christian/Meißner, Cordula/Slavcheva, Adriana (2014) (Hrsg.): *Gesprochene Wissenschaftssprache: Korpusmethodische Fragen und empirische Analysen*. Heidelberg: Synchron.

Fandrych, Christian/Meißner, Cordula/Wallner Franziska (2018): Das Potenzial mündlicher Korpora für die Sprachdidaktik: Das Beispiel GeWiss. In: *Deutsch als Fremdsprac*he 55.1, 3–13.

Fandrych, Christian/Meißner, Cordula/Wallner Franziska (2017) (Hrsg.): *Gesprochene Wissenschaftssprache –digital. Verfahren zur Annotation und Analyse mündlicher Korpora*. Tübingen: Stauffenburg.

Fiehler, Reinhard (2016): Gesprochene Sprache. In: Wöllstein, Angelika, (Hrsg.): *Duden – Die Grammatik. Unentbehrlich für richtiges Deutsch*. Berlin: Dudenverlag, 1181–1260.

Handwerker, Brigitte/Bäuerle, Rainer/Sieberg, Bernd (2016) (Hrsg.): *Gesprochene Fremdsprache Deutsch*. Baltmannsweiler: Schneider.

Hansen, Carsten/Hansen, Martin H. (2012): A dictionary of spoken Danish. In: Fjeld, Ruth Vatvedt/Torjusen, Julie Matilde (Hrsg.): *Proceedings of the 15th EURALEX International Congress. 7–11 August 2012. Oslo, Norway*. Oslo: Department of Linguistics and Scandinavian Studies, University of Oslo, 929–935.

Imo, Wolfgang (2007): *Construction Grammar und Gesprochene-Sprache-Forschung. Konstruktionen mit zehn matrixsatzfähigen Verben im gesprochenen Deutsch*. Tübingen: Niemeyer.

Imo, Wolfgang/Moraldo, Sandro M. (2015) (Hrsg.): *Interaktionale Sprache und ihre Didaktisierung im DaF-Unterricht*. Tübingen: Stauffenburg.

Klosa, Annette (2013a): The lexicographical process (with special focus on online dictionaries). In: Gouws, Rufus H., et al. (Hrsg.): *Dictionaries. An international encyclopedia of lexikography. Supplementary volume: Recent developments with focus on electronic and computational lexicography*. Berlin/Boston: de Gruyter, 517–524.

Klosa, Annette (2013b): Aktuelle Tendenzen in der deutschen Lexikographie der Gegenwart. In: Stickel, Gerhard/Váradi, Tamás (Hrsg.): *Lexical Challenges in a Multilingual Europe. Contributions to the Annual Conference 2012 of EFNIL in Budapest*. Frankfurt am Main: Lang, 75–93.

Klosa, Annette/Gouws, Rufus Hjalmar (2015): Outer features in e-dictionaries. In: *Lexicographica* 31, 142–172.

Klosa, Annette/Tiberius, Carole (2016): Der lexikografische Prozess. In: Klosa, Annette/Müller-Spitzer, Carolin (Hrsg.): *Internetlexikografie. Ein Kompendium*. Berlin/Boston: de Gruyter, 65–110.

Lemmenmeier-Batinić, Dolores (2020): Lexical Explorer: Extending access to the Database of Spoken German for user-specific purposes. In: *Corpora* 15.1, 55–76.

Meliss, Meike (2016): Gesprochene Sprache in DaF-Lernerwörterbüchern. In: Handwerker, Brigitte/Bäuerle, Rainer/Sieberg, Bernd (Hrsg.): *Gesprochene Fremdsprache Deutsch*. Baltmannsweiler: Schneider, 179–199.

Meliss, Meike (2021): Die LeGeDe-Ressource: korpusbasierte lexikografische Einblicke und anwendungsorientierte Ausblicke. In: *ZDaF* 1/2021, 3–15.

Meliss, Meike (2020a): Lexik des gesprochenen Deutsch in der Interaktion: Partikelverben und ihre Besonderheiten. In: García Adánez, Isabel/Gil Valdés, Mª Jesús/Sánchez Hernández, Paloma/Schilling Rodríguez, Mª Luisa/Szumlakowski Morodo, Irene (Hrsg.): *Das Leben in einem Rosa Licht sehen. Festschrift für Rosa Piñel/Homenaje a Rosa Piñel*. Bern: Peter Lang, 77–97.

Meliss, Meike (2020b): Musterhaftigkeit in der Lexik des gesprochenen Deutsch: eine handlungs- und funktionsorientierte Annäherung mit lexikografischer Anwendungsperspektive. In: Mellado Blanco, Carmen, et al. (Hrsg.): *Produktive Muster in der Phraseologie. Monolingual und kontrastiv*. Hamburg: Verlag Dr. Kovač, 181–202.

Meliss, Meike/Möhrs, Christine (2017): Die Entwicklung einer lexikografischen Ressource im Rahmen des Projektes LeGeDe. In: *Sprachreport* 4.17, 42–52.

Meliss, Meike/Möhrs, Christine (2018): Lexik in der spontanen, gesprochensprachlichen Interaktion: Eine anwendungsorientierte Annäherung aus der DaF-Perspektive. In: *GFL* 3/2018, 79–110.

Meliss, Meike/Möhrs, Christine (2019): Herausforderungen eines innovativen lexikografischen Projekts zu Besonderheiten des gesprochenen Deutsch in der Interaktion. In: Doval, Irene/Liste Lamas, Elsa (Hrsg.): *Germanistik im Umbruch – Linguistik, DaF und Übersetzung*. Berlin: Frank & Timme, 13–25.

Meliss, Meike/Möhrs, Christine (2020): Der LeGeDe-Prototyp: Zur Erstellung eines korpusbasierten Online-Wörterbuchs zu lexikalischen Besonderheiten im gesprochenen Deutsch. Aktueller Stand und Perspektiven. In: *Lexikos* 30, 1–41.

Meliss, Meike/Möhrs, Christine/Batinić, Dolores/Perkuhn, Rainer (2018a): Creating a List of Headwords for a Lexical Resource of Spoken German. In: Čibej, Jaka/Gorjanc, Vojko/Kosem, Iztok/Krek, Simon (Hrsg.): *Proceedings of the XVIII EURALEX International Congress. Lexicography in Global* Contexts, *17–21 July*, *Ljubljana*. Ljubljana: Znanstvena založba Filozofske fakultete Univerze v Ljubljani, 1009–1016.

Meliss, Meike/Möhrs, Christine/Ribeiro Silveira, Maria (2018b): Erwartungen an eine korpusbasierte lexikografische Ressource zur „Lexik des gesprochenen Deutsch in der Interaktion“: Ergebnisse aus zwei empirischen Studien. In: *Zeitschrift für Angewandte Linguistik* 68, 103–138.

Meliss, Meike/Möhrs, Christine/Ribeiro Silveira, Maria (2019a): Anforderungen und Erwartungen an eine lexikografische Ressource des gesprochenen Deutsch aus der L2-Lernerperspektive. In: *Lexicographica* 34, 89–121.

Meliss, Meike/Möhrs, Christine/Ribeiro Silveira, Maria/Schmidt, Thomas (2019b): A corpus-based lexical resource of spoken German in interaction. In: Kosem, Iztok/Zingano Kuhn, Tanara/Correia, Margarita/Ferreria, José Pedro/Jansen, Maarten/Pereira, Isabel/Kallas, Jelena/Jakubíček, Miloš/Krek, Simon/Tiberius, Carole (Hrsg.): *Electronic lexicography in the 21st century. Proceedings of the eLex 2019 conference*. 1–3 October 2019, Sintra, Portugal. Brno: Lexical Computing CZ, s.r.o, 783–804.

Möhrs, Christine (2020): "*Hast du eine Ahnung, . . . ?*" Eine lexikografische und korpusbasierte Untersuchung am Beispiel des Lexems *Ahnung*. In: Behr, Janina, et al. (Hrsg.): *Schnittstellen der Germanistik. Festschrift für Hans Bickes*. Berlin: Peter Lang, 185–214.

Möhrs, Christine (2021): Quantitative und qualitative Ansätze zu Stichwortkandidaten für die lexikografische Ressource zum gesprochenen Deutsch: LeGeDe. In: Piosik, Michał/Taborek, Janusz/Woźnicka, Marta (Hrsg.): *Korpora in der Lexikographie und Phraseologie – Stand und Perspektiven*. Berlin/Boston: de Gruyter, 175–195.

Möhrs, Christine/Meliss, Meike/Batinić, Dolores (2017): LeGeDe – towards a corpus-based lexical resource of spoken German. In: Kosem, Iztok, et al. (Hrsg.): *Electronic Lexicography in the 21st century. Proceedings of eLex 2017 Conference. 19–21 September 2017. Leiden, the Netherlands*. Brno: Lexical Computing CZ s.r.o., 281–298.

Möhrs, Christine/Torres Cajo, Sarah (2020): The microstructure of a lexicographical resource of spoken German: meanings and functions of the lemma *eben*. In: *Rasprave*: *Journal of the Institute of Croatian Language and Linguistics* 46/2, 921–939.

Moon, Rosamund (1998): On using spoken data in corpus lexicography. In: Fontenelle, Thierry, et al. (Hrsg.): *Proceedings of the 8th EURALEX International Congress. 4–8 August 1998. Liège, Belgium*. Liège: English and Dutch Departments, University of Liège, 347–355.

Moraldo, Sandro M./Missaglia, Federica (2013) (Hrsg.): *Gesprochene Sprache im DaF-Unterricht. Grundlagen – Ansätze – Praxis*. Heidelberg: Winter.

Reeg, Ulrike/Gallo, Pasquale/Moraldo, Sandro M. (2012) (Hrsg.): *Gesprochene Sprache im DaF-Unterricht. Zur Theorie und Praxis eines Lerngegenstandes*. Berlin: Waxmann.

Schmidt, Thomas (2014a): The Research and Teaching Corpus of Spoken German – FOLK. In: Calzolari, Nicoletta, et al. (Hrsg.): *Proceedings of the 9th International Conference on Language Resources and Evaluation. 26–31 May 2014. Reykjavik, Iceland*. Reykjavik: European Language Resources Association, 383–387.

Schmidt, Thomas (2014b): Gesprächskorpora und Gesprächsdatenbanken am Beispiel von FOLK und DGD. *In: Gesprächsforschung – Online-Zeitschrift zur verbalen Interaktion* 15, 196–233.

Schmidt, Thomas (2017): DGD –Die Datenbank für Gesprochenes Deutsch. Mündliche Korpora am Institut für Deutsche Sprache (IDS) in Mannheim. In: *Zeitschrift für Germanistische Linguistik* 45, 451–463.

Schwitalla, Johannes (2012): *Gesprochenes Deutsch. Eine Einführung*. 4. Aufl. Berlin: Schmidt.

Selting, Margret/Auer, Peter/Barth-Weingarten, Dagmar u. a. (2009): Gesprächsanalytisches Transkriptionssystem 2 (GAT2). In: *Gesprächsforschung – Online-Zeitschrift zur verbalen Interaktion* 10, 353–402.

Sieberg, Bernd (2013): *Sprechen lehren, lernen und verstehen. Stufenübergreifendes Studien- und Übungsbuch für den DaF-Bereich*. Tübingen: Julius Groos.

Siepmann, Dirk (2015): Dictionaries and spoken language: A corpus-based review of French dictionaries. In: *International Journal of Lexicography* 28, 139–168.

Trap-Jensen, Lars (2004): Spoken language in dictionaries: Does it really matter? In: Williams, Geoffrey/Vessier, Sandra (Hrsg.): *Proceedings of the 11th EURALEX International Congress. 6–10 July 2004. Lorient, France*. Lorient: Faculté des Lettres et des Sciences Humaines, Université de Bretagne Sud, 311–318.
Trim, John/North, Brian/Coste, Daniel (2001): *Gemeinsamer europäischer Referenzrahmen für Sprachen: lernen, lehren, beurteilen. Niveau A1, A2, B1, B2, C1, C2*. Berlin: Langenscheidt.
Verdonik, Darinka/Sepesy Maučec, Mirjam (2017): A speech corpus as a source of lexical information. In: *International Journal of Lexicography* 30, 143–166.
Wiegand, Herbert Ernst (1989a): Arten von Mikrostrukturen im allgemeinen einsprachigen Wörterbuch. In: *HSK* 1, 462–501.
Wiegand, Herbert Ernst (1989b): Der Begriff der Mikrostruktur: Geschichte, Probleme, Perspektiven. In: *HSK* 1, 409–462.

## (Online)Wörterbücher

Duden-online: Dudenredaktion (o.J.): Online unter: <https://www.duden.de/woerterbuch> [letzter Zugriff: 13.08.2019].
DWDS = Das Wortauskunftssystem zur deutschen Sprache in Geschichte und Gegenwart. [Online unter: <www.dwds.de>; letzter Zugriff: 13.08.2019].
E-VALBU: Elektronisches Valenzwörterbuch deutscher Verben. IDS-Mannheim. Online unter: <https://grammis.ids-mannheim.de/verbvalenz> [letzter Zugriff: 13.08.2019].
Götz, Dieter (Hrsg.) (2015): Langenscheidt Großwörterbuch Deutsch als Fremdsprache: das einsprachige Wörterbuch für alle, die Deutsch lernen. (Neubearbeitung) Unter Zusammenarbeit mit der Langenscheidt-Redaktion. München: Langenscheidt. (=LGWB-DaF)
LeGeDe = Lexik des gesprochenen Deutsch (2019). [Online unter: <www.owid.de/legede/>; letzter Zugriff: 13.11.2020].
Schumacher, Helmut/Kubczak, Jacqueline/Schmidt, Renate/de Ruiter, Vera (2004): VALBU. Valenzwörterbuch deutscher Verben. Tübingen: Narr (Studien zur Deutschen Sprache 31). = VALBU

## Sonstige Onlinequellen

DeReKo = Deutsches Referenzkorpus 2017-I (Release 08.03.2017). Mannheim: Institut für Deutsche Sprache. [Unter: <http://www1.ids-mannheim.de/kl/projekte/korpora.html>; letzter Zugriff: 13.08.2019].
DGD = Datenbank für gesprochenes Deutsch (2012). [Unter: <https://dgd.ids-mannheim.de/dgd/pragdb.dgd_extern.welcome>; letzter Zugriff: 13.08.2019].
FOLK = Forschungs- und Lehrkorpus Gesprochenes Deutsch (Release 2.11 vom 30.11.2018). Mannheim: Institut für Deutsche Sprache. [Unter: <http://agd.ids-mannheim.de/folk.shtml>; letzter Zugriff: 13.08.2019].
LeGeDe-Projektwebseite. [Unter: <http://www1.ids-mannheim.de/lexik/lexik-des-gesprochenen-deutsch/>; letzter Zugriff: 13.08.2019].