

Meike Meliss

Die LeGeDe-Ressource: korpusbasierte lexikografische Einblicke und anwendungsorientierte Ausblicke

Der Beitrag stellt die lexikografische Online-Ressource LeGeDe, den ersten korpusbasierten Prototypen für Besonderheiten der Lexik des Deutschen in der Interaktion vor. Dabei werden sowohl die Herausforderungen an das innovative Projekt thematisiert als auch Möglichkeiten für einen anwendungsorientierten Nutzen im DaF- und DaZ-Bereich diskutiert und als Ausblick Desiderata für die weitere Beschäftigung mit der lexikografischen Kodifizierung gesprochen Sprachlicher Spezifika des Deutschen aufgezeigt.

Schlüsselwörter: korpusbasierte Lexikografie; Lexik des gesprochenen Deutsch; DaF; DaZ

The LeGeDe resource: corpus-based lexicographic insights and application-oriented perspectives

This article introduces the online lexicographic resource LeGeDe, the first corpus-based prototype for special features of German vocabulary in interaction. The article discusses the challenges of this innovative project as well as possibilities for an application-oriented use in the field of German as a Foreign and Second language. Finally, an outlook points out desiderata for further work on the lexicographical codification of linguistic specifics of spoken German.

Keywords: corpus-based lexicography; lexis of spoken German; German as a Foreign Language; German as a Second Language

1 Einleitung

Die LeGeDe-Ressource¹ versteht sich als multimedialer Prototyp für ein korpusbasiertes Wörterbuch, das sich die lexikografische Darstellung der Besonderheiten des standardnahen, gemeinsprachlich gesprochenen Deutsch in der Interaktion zum Ziel gemacht hat.² Obwohl die Erstellung dieses Prototypen primär der Auslotung von Möglichkeiten und Grenzen der Sprachdokumentation gesprochen Sprachlicher lexikalischer Phänomene dienen soll und somit zunächst als mögliche Zielgruppe

Experten und Expertinnen anvisiert, existieren durchaus auch zahlreiche Anknüpfungspunkte für einen anwendungsorientierten Nutzen. So kann die lexikografische Ressource u. a. eine Grundlage für Maßnahmen zur Verbesserung eines an der Sprachrealität der Gegenwart orientierten Unterrichts von Deutsch als Fremd- bzw. Zweitsprache leisten. Daher werden im Laufe dieses Beitrages konkret für den DaF- und DaZ-Bereich verschiedene Verknüpfungen hergestellt.

Um die Spezifika gesprochen Sprachlicher Lexik des Deutschen auf der Basis des am Leibniz-Institut für Deutsche Sprache (=IDS) erstellten Forschungs- und Lehrkorpus zum Gesprochenen Deutsch (=Folk)³ identifizieren, analysieren und für die lexikografische An-

¹ Die Ressource wurde im September 2019 freigeschaltet und kann kostenlos über folgenden Link konsultiert werden: <https://www.owid.de/legede/>.

² Es handelt sich um ein Drittmittelprojekt der Förderlinie „Innovative Vorhaben“, welches im Leibniz-Wettbewerb eingeworben werden konnte (=LeGeDe-Projekt). An dieser Stelle möchte ich den Antragsteller(inne)n und dem LeGeDe-Team für die Zusammenarbeit danken. Für weitere Informationen zu dem Projekt „Lexik des gesprochenen Deutsch“ wird auf die Projektwebseite verwiesen: <http://www1.ids-mannheim.de/lexik/lexik-des-gesprochenen-deutsch.html>.

³ FOLK wird seit 2008 am IDS aufgebaut und versteht sich als erstes großes Gesprächskorpus des Deutschen. Parallel dazu wird die Datenbank gesprochenes Deutsch (=DGD), über die FOLK recherchierbar ist, permanent weiterentwickelt und mit innovativen korpus technologischen Funktionalitäten ausgestattet (Kupietz/Schmidt 2015; Schmidt 2018).

wendung aufbereiten zu können, wurden im LeGeDe-Projekt quantitative und qualitative Analysemethoden genutzt und innovative Beschreibungsformate für hochgradig kontextualisierte lexikalische Korpusdaten in audio-elektronischer Form, ebenso wie neuartige lexikografische Angabetypen, die auf die Funktion lexikalischer Einheiten in interaktionalen Kontexten Bezug nehmen, konzipiert.

Ziel des Beitrages ist, die LeGeDe-Ressource vorzustellen (Abschnitt 4) und zuvor sowohl auf den Hintergrund (Abschnitt 2) als auch auf die konzeptionellen Überlegungen (Abschnitt 3) des LeGeDe-Projektes einzugehen. In Abschnitt 5 werden als Fazit v. a. der innovative Aspekt und die anwendungsorientierten Möglichkeiten für die Sprachdidaktik aufgezeichnet. Als Ausblick werden verschiedene Bereiche angedeutet, in denen die lexikografische Arbeit an der LeGeDe-Ressource ausbaufähig ist.

2 Projekthintergrund

Die Entwicklung der LeGeDe-Ressource knüpft an die Tatsache an, dass sich die Divergenzen zwischen dem gesprochenen und dem geschriebenen Deutsch im lexikalischen Bereich sowohl im Bestand als auch in Verbindung mit der Formseite, der Bedeutung und der Verwendung der lexikalischen Elemente manifestieren können (vgl. Deppermann et al. 2017, Fiehler 2016, Imo 2007, Schwitalla 2012). Dennoch existieren zurzeit so gut wie keine korpusbasierten lexikografischen Projekte zur Lexik des gesprochenen Deutsch. Sowohl die Ergebnisse zweier projektbegleitender Umfragen zu den Erwartungen und Anforderungen an eine lexikografische Ressource für Spezifika des gesprochenen Deutsch (vgl. Meliss/Möhrs/Ribeiro Silveira 2018, 2019) als auch entsprechende lexikografische Untersuchungen bestätigen, dass die lexikografische Kodifizierung der gesprochenen Sprache unzureichend ist und besonders die Erfassung der interaktionalen Merkmale in den aktuellen Wörterbüchern bisher nicht zufriedenstellend berücksichtigt wurde (vgl. Meliss 2016: 195; Meliss/Möhrs 2018: 88 f.; Meliss/Möhrs/Ribeiro Silveira 2019: 94–95; Moon 1998: 352; Trap-Jensen 2004: 311). Trotz einiger aktuel-

ler Fortschritte in der Erstellung gesprochen-sprachlicher Korpora für unterschiedliche Sprachen und Sprachvarietäten (vgl. u. a. Barcala et al. 2018; Fandrych/Meißner/Wallner 2017; Verdonik/Sepesy Maučec 2017) sind die Erfahrungen mit gesprochen-sprachlichen Korpora in der Lexikografie bis jetzt eher selten. Zu beobachten ist andererseits, dass der Informationsbedarf zu typisch gesprochen-sprachlicher Lexik in unterschiedlichen Anwendungsbereichen, wie z. B. in Unterricht und Lehre im letzten Jahrzehnt gestiegen ist (vgl. Handwerker et al. 2016; Imo/Moraldo 2015; Sieberg 2013).⁴ So weisen u. a. Fandrych et al. 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 (Fandrych et al. 2018: 3 f.). Dies bestätigen auch die Ergebnisse der im LeGeDe-Projekt durchgeführten empirischen Erwartungsstudien (Meliss/Möhrs/Ribeiro Silveira 2018: 119–120; 2019: 103).

Die besagten Beobachtungen sind Ausgangspunkte für die konzeptionelle Entwicklung der LeGeDe-Ressource und führen u. a. zu folgenden zentralen theoretischen, methodologischen und anwendungsorientierten Aspekten, die sich bei der lexikografischen Umsetzung ergaben:

- Entwicklung von quantitativen und qualitativen Verfahren zur Identifizierung von interaktionstypischen, gesprochen-sprachlichen lexikalischen Elementen und deren spezifischen Merkmalen im Vergleich zur Lexik der geschriebenen Sprache (vgl. Meliss/Möhrs 2019, 2020)
- Erstellung einer Stichwortkandidatenliste und Auswahl geeigneter Stichwörter für die Ressource (vgl. Meliss et al. 2018)

⁴ Im Gemeinsamen europäischen Referenzrahmen für Sprachen wird u. a. zum Beurteilungsraster zur mündlichen Kommunikation und dem Parameter „Interaktion“ für Niveau C1 explizit darauf hingewiesen, dass die Lernenden die Fähigkeit besitzen sollten, aus einem Repertoire von Diskursmitteln eine geeignete Wendung auszuwählen, um z. B. eine Äußerung angemessen einzuleiten etc. (vgl. Europarat 2001: 37). Siehe dazu auch ausführliche Studien in Meliss/Möhrs (2018).

- Bestimmung der Besonderheiten und Divergenzen von mündlichem vs. schriftlichem Sprachgebrauch auf verschiedenen Ebenen (Form, Inhalt/Funktion, Situation etc.) im lexikalischen Bereich (vgl. Meliss/Möhrs 2018, 2019, 2020)
- Entwicklung neuartiger lexikografischer Angabeklassen, die u. a. auf die Funktion lexikalischer Einheiten in Interaktionskontexten Bezug nehmen (vgl. Meliss et al. 2019)
- Entwicklung innovativer lexikografischer Beschreibungsformate in multimedialer Form für kontextualisierte lexikalische Daten (vgl. Meliss et al. 2019)
- Entwicklung weiterer korpuslinguistischer Methoden und Tools zur Abfrage, Analyse und Strukturierung von korpusbasierten Daten (vgl. Möhrs/Meliss/Batinić 2017)

Die Beschäftigung mit den besagten Aspekten führte, neben der Entwicklung der lexikografischen LeGeDe-Ressource u. a. zu der projektbegleitenden Erstellung des Tools *Lexical Explorer*. Mit diesem Werkzeug ist es möglich, die quantitativen Daten von zwei Korpora aus der Datenbank für Gesprochenes Deutsch (= DGD) abzufragen⁵ und diese mit Hilfe von Häufigkeitstabellen bezüglich der Wortverteilung über Wortformen, Kookkurrenzen und Metadaten zu erzeugen.

3 Zur Konzeption der LeGeDe-Ressource

Die einzelnen Arbeitsschritte zur Entwicklung der LeGeDe-Ressource folgten grundsätzlich dem lexikografischen Prozess zur Erstellung von Internetwörterbüchern (Klosa/Tiberius 2016: 75–77). Die lexikografische Umsetzung stellte sich dabei neuen Herausforderungen, die in Verbindung mit einer Reihe von Fragestellungen u. a. zur Stichwortansetzung, zu möglichen Benutzungssituationen und zu angemessenen Darstellungsmöglichkeiten multimedialer Daten standen. Im Folgenden soll ein Einblick in die grundlegenden methodologischen Überlegungen gegeben wer-

den. Für genauere Information wird u. a. auf die Publikationen von Meliss/Möhrs (2019, 2020) und auf den lexikografischen Umtext der LeGeDe-Ressource „Über LeGeDe“ hingewiesen.

(i) Erstellung einer Stichwortkandidatenliste

Mit einer Methode, die auf dem Häufigkeitsklassenvergleich zwischen FOLK und Teilkorpora aus dem Deutschen Referenzkorpus (= DeReKo)⁶ beruht (Keibel 2008, 2009), konnte eine Stichwortkandidatenliste mit 322 Lemmata ermittelt werden, die verschiedene Phänomenbereiche unseres Interesses abdecken. In Meliss et al. (2018) wird diese Methode genau beschrieben und kann über den *Lexical Explorer* nachvollzogen werden. Auch das Problem der Vergleichbarkeit der besagten Korpora wird dort diskutiert. Die Stichwortkandidatenliste umfasst u. a. bestimmte Partikeln, Interjektionen und deiktische Ausdrücke. Diese für die gesprochene Sprache besonders relevanten Bereiche wurden zum großen Teil bereits in der Forschungsliteratur als typische Phänomene des Gesprochenen aufgelistet (u. a. Schwitalla 2012; Fiehler 2016) und teilweise in Einzelstudien ausführlich beschrieben. Als korpusbasierter Beweis für die Durchgängigkeit dieser Phänomene in der gesprochenen Sprache kann bereits die Liste der Top 25 Lemmata mit der höchsten Häufigkeitsklassendifferenz dienen, die über den *Lexical Explorer* abgerufen wurde (vgl. Meliss et al. 2018: 1013). Im Folgenden werden daher nur einige der Auffälligkeiten kommentiert.

(a) So weisen beispielsweise Partikeln wie *halt* und *mal* sowie die Interjektionspartikeln wie *ah*, *ach*, *oh*, *eh* im Vergleich zu den herangezogenen schriftsprachlichen Teilkorpora aus DeReKo zu Gunsten von FOLK eine hohe Häufigkeitsklassendifferenz auf, die zu der Annahme führt, dass diese Lemmata in FOLK spezifische Bedeutungen und /oder Funktionalitäten aufweisen. Die funktionale und interaktionale Beschreibung dieser Besonderheiten, die allgemein der gesprochenen Sprache in der

⁵ Diese sind FOLK und GeWiss (=Gesprochene Wissenschaftssprache; Fandrych et al. 2017).

⁶ Siehe dazu u. a. Kupietz/Lüngen/Kamocki/Witt 2018 und Kupietz/Schmidt 2015.

Interaktion zugeschrieben werden können, ist jedoch in den gängigen Wörterbüchern bis jetzt nur unzureichend bis gar nicht abgedeckt.

(b) Bei der Untersuchung der Adverbien in der Gesamtliste der 322 ermittelten Stichwortkandidaten konnten auch verschiedene temporale und lokale Adverbien (*nachher, jetzt, da, hier*) mit einer Häufigkeitsklassendifferenz zu Gunsten von FOLK identifiziert werden. Ihr deiktischer Gebrauch in der gesprochenen Sprache wird in den lexikografischen Werken ebenfalls kaum berücksichtigt.

(c) Einige Lemmata wie *gut, klar, ehrlich*, die ebenfalls als quantitativ hoch divergierend zu Gunsten von FOLK erkannt wurden, sind auch Gegenstand des Interesses. Aus den Belegen aus FOLK geht hervor, dass sie ein breites Funktionsspektrum aufweisen, da sie neben der kategorialen Zuordnung zu der Klasse der Adjektive u. a. auch als Diskurspartikel fungieren können.

(d) Außerdem konnten mit der verfolgten korpusbasierten Methode zur Erstellung einer Stichwortkandidatenliste zahlreiche verbale Lemmata identifiziert werden, die verschiedenen semantischen Untergruppen unseres Interesses zugeordnet werden können:

- Verben der Wahrnehmung (z. B.: *gucken, schauen, sehen, hören*);
- mentale Verben und Verben der Kognition (z. B.: *wissen, glauben, denken, meinen, überlegen, verstehen, kennen*);
- Verben der Fortbewegung (z. B.: *gehen*);
- Verben der Kommunikation (z. B.: *reden, sagen, fragen*);
- Verben der Emotion (z. B.: *mögen, gefallen*);
- Modalverben (z. B.: *können, müssen*);

Interaktional ausgerichtete Studien zeigen, dass die Verben dieser semantischen (Sub-) Klassen ein wichtiges Potenzial in interaktionalen Kontexten besitzen und entsprechend der verschiedenen speziellen Bedeutungen, Funktionen und formalen Besonderheiten u. a. den oben beschriebenen Phänomenbereichen zugeschrieben werden können (vgl. Imo 2007; Deppermann et al. 2017).

(ii) Auswahl der zu bearbeitenden Stichwörter

Die Erkennung von Einwortlemmata war der erste Schritt für die Erstellung einer Stichwortkandidatenliste für die Darstellung des Wortschatzes des gesprochenen Deutsch in der Interaktion. Darüber hinaus können diese Einwortausdrücke die lexikalische Grundlage von Mehrwortlemmata bzw. musterhaften Verwendungsweisen bilden (z. B.: *ich weiß nicht, guck mal, mal schauen, keine Ahnung*), die als lexikalische Elemente in einer erweiterten Stichwortliste der Ressource ebenfalls berücksichtigt werden (vgl. Meliss 2020b).

Die quantitative Auswahl der behandelten Stichwörter aus der umfangreichen Stichwortkandidatenliste einer jeden lexikografischen Ressource bezieht sich auf den anvisierten Umfang der Ressource und seine statische vs. dynamische Konzeption (Atkins/Rundell 2008: 160 f.; Klosa 2013: 518). Da die LeGeDe-Ressource zunächst als lexikografischer Prototyp geplant wurde, erfolgte die Auswahl der behandelten Stichwörter aus der Stichwortkandidatenliste nach Kriterien in Zusammenhang mit dem Anspruch, eine Bandbreite von unterschiedlichen Phänomenbereichen anzubieten und dabei gleichzeitig auch die Möglichkeit zu besitzen, Ergebnisse aus wissenschaftlichen Teilstudien berücksichtigen zu können (vgl. Blühdorn et al. 2017; Deppermann et al. 2017; Günthner 2018; Imo 2007). Aus formal-inhaltlicher und interaktionaler Perspektive interessierten daher u. a. Verwendungsformen mit Distributionsbeschränkungen in der Verbalmorphologie wie z. B. bei *gucken* (*guck*: Imperativ) und *wissen* (du weißt/weißt du → *weißte*: Verschmelzung bzw. Komplementierungsmuster), sowie auch zahlreiche Verwendungsmuster u. a. in Form von Verbal- oder Nominalsyntaxmen (z. B. *guck mal, mal schauen, was weiß ich, keine Ahnung*), die im interaktionalen Kontext eine spezifische Funktion besitzen (vgl. Meliss 2020b; Möhrs 2020). Die Verwendung von Teilsynonymen wie z. B. von *kriegen* vs. *bekommen/erhalten* oder von *gucken* bzw. *schauen/sehen* ist für die LeGeDe-Ressource v. a. in Verbindung mit ihrem unterschiedlichen Kombinationspotenzial von Interesse. Die für die gesprochene Sprache typischen Diskurspartikeln und Modalpartikeln bilden ebenfalls einen interessanten und

umfangreichen Bereich des Gegenstandsbe-
reiches. Aus den zahlreichen Möglichkeiten
wurden für den LeGeDe-Prototypen die mul-
tifunktionalen Stichwörter *gut* und *schön* ei-
nerseits und *eben*, *halt* und *ja* andererseits aus-
gewählt.

Für die Stichwortauswahl der ersten Ar-
beitsphase konnte somit an die Ergebnisse der
projektbegleitenden Onlineumfrage (=OU)
angeknüpft werden, die sowohl für die L1- als
auch für die L2-Nutzer besonders den Wunsch
nach Stichwörtern, die – im Vergleich zur
geschriebenen Sprache – eine distinktive Be-
deutung und Funktion und/oder eine spezi-
fische Kombinatorik aufweisen, deutlich un-
terstreichen (Meliss/Möhrs/Riberiro Silveira
2018: 125 f.; 2019: 103 f.).

(iii) Datenanalyse und Informations- strukturierung

Verschiedene quantitative und qualitative Ver-
fahren zur Bedeutungsdisambiguierung und
zur korpusbasierten Entwicklung von Wort-
profilen, die als methodologische Ansätze in
der korpusbasierten Lexikologie und Lexiko-
grafie gelten (vgl. Engelberg 2015), wurden
mit der Analyse sprachlicher Einheiten und
der Beschreibung von kommunikativen Funk-
tionen aus Sicht der Interaktionslinguistik
(vgl. Deppermann 2007) vereint. Datengrund-
lage für die Analyse war jeweils eine Zufalls-
stichprobe aus FOLK, aus der die ersten 100
gültigen Belege händisch kodiert wurden. Das
Kodierschema wurde in Zusammenhang mit
qualitativen lexikalisch-semantischen und in-
teraktionslinguistischen Einzelbeleganalysen
erstellt. In diesen ersten qualitativen Studien
wurden die Gesprächssequenzen zu entspre-
chenden Belegen nach Form, Bedeutung,
Kombinatorik, Sequenz, Prosodie und Funk-
tion in Verbindung mit metasprachlicher In-
formation analysiert und kodiert, die Ergeb-
nisse strukturiert und hinsichtlich ihrer
Relevanz für die LeGeDe-Ressource interpre-
tiert und gebündelt. Zu den Korpusbelegen
stehen in allen Fällen Metadaten und Audio-
aufnahmen zur Verfügung, die über die Da-
tenbank parallel zu den Transkripten abgeru-
fen werden können und die Dateninterpretation
jederzeit unterstützen.

(iv) Lexikografische Umsetzung

Für die lexikografische Umsetzung mussten
teilweise neue Angabeklassen konzipiert wer-
den, die den Spezifika der Lexik in der münd-
lichen Interaktion besonders entsprechen. In
Abschnitt 4 wird auf dieses Informationsan-
gebot gesondert eingegangen.

4 Informationsangebot der LeGeDe- Ressource

4.1 Allgemeines

Die LeGeDe-Ressource bietet ein umfangrei-
ches multimediales Informationsspektrum für
jedes Stichwort an, welches durch verschie-
dene Wörterbuchumtexte („Über LeGeDe“,
Benutzungshinweise, Glossar, Literatur) um-
rahmt wird. Die lexikografische Struktur ist
dementsprechend komplex. Der alphabetische
Zugriff auf die einzelnen Stichwörter oder die
spezifischen Lemmata erfolgt entweder über
die Indexseite oder die seitliche Menüleiste.

Das Informationsspektrum für jedes Stich-
wort besteht aus einem Überblicksartikel (vgl.
Abschnitt 4.2), einem Modul 1 zu lexikali-
scher Information in Verbindung mit den un-
terschiedlichen Lesarten, die in der LeGe-
De-Stichprobe identifiziert werden konnten
(vgl. Abschnitt 4.3), und einem Modul 2 mit
relevanter Information zu dem Funktions-
spektrum der Lemmata in der Interaktion (vgl.
Abschnitt 4.4). Verschiedene Schnittstellen
zwischen beiden Modulen, d. h. zwischen der
Bedeutungs- und Funktionsbeschreibung wer-
den durch eine interne Verknüpfung explizit
gemacht.

4.2 Überblicksartikel

Der Überblicksartikel mit allgemeiner Infor-
mation zu jedem Stichwort stellt neben der In-
formation zu dem Einwortlemma auch eine
direkte Verbindung zu den unterschiedlichen
mehnteiligen, musterhaften Verbindungen her
und bietet zunächst in beschreibender Form in
einem Überblickstext allgemeine bedeutungs-
und funktionsorientierte Informationen an
(vgl. *Abb. 1*). Eine deutliche modulare Zwei-
teilung des Informationsangebotes ermöglicht

gucken

Das Vollverb **gucken** konnte in der LeGeDe-Stichprobe in neun verschiedenen Lesarten und in zwei verschiedenen interaktionalen Mustern belegt werden. Die Grundbedeutung von **gucken** steht in Zusammenhang mit der visuellen richtungsorientierten Wahrnehmung. Diese konnte auch in anderen Bedeutungsvarianten von **gucken** in Zusammenhang mit den Lesarten 'überprüfen', 'ansehen', 'suchen' und 'umherschauen' nachgewiesen werden. Die visuelle Wahrnehmungsbedeutung verblasst deutlich in den Bedeutungsvarianten, in denen **gucken** in den Lesarten 'achten auf', 'abwägen', und 'abwarten' auftritt. In diesen Fällen dominiert teilweise eine kognitiv-mentale Bedeutungskomponente. Eine Verbindung zu einem optischen Eindruck stellt die Lesart 'eine bestimmte Miene zeigen' her. Neben dem Muster **guck mal** als Diskursmarker tritt das Muster **mal gucken** zur Verschiebung einer Handlung in die Zukunft teilweise als Beendigungssignal auf.

Frequenz
FOLK 2.11 (absolut): 2927

FOLK-Zugriff
gucken (in FOLK)
gucken (im Lexical Explorer)

Bedeutungen	Funktionen im Gespräch
gucken [VOLLVERB] <ul style="list-style-type: none"> ○ richtungsorientiert visuell wahrnehmen ○ überprüfen ○achten auf ○abwägen ○ansehen ○abwarten ○suchen ○umherschauen ○eine bestimmte Miene zeigen 	guck mal [VERBALSYNTAGMA] <ul style="list-style-type: none"> ○ Diskursmarker mal gucken [VERBALSYNTAGMA] <ul style="list-style-type: none"> ○ Verschiebung einer Handlung in die Zukunft

Information in anderen lexikografischen Ressourcen

- Online-Wortschatz-Informationssystem Deutsch (OWID)
- Digitales Wörterbuch der Deutschen Sprache (DWDS)

Forschungsliteratur

- Günthner (2017)
- Imo (2007)
- Wegner (2015)

Abb. 1: LeGeDe-Überblicksartikel: Beispiel: Stichwort *gucken*

einerseits die Präsentation der lexikalisch-semantischen Information zu den entsprechenden Lesarten eines Lexems (=Modul 1) und andererseits die der funktionsspezifischen interaktional ausgerichteten Information (=Modul 2). In beiden Modulen erfolgt der Zugriff über die verschiedenen Lemmazeichengestaltungsangaben (=LZGAs). Diese LZGAs werden formal und kategorial voneinander unterschieden (z. B. Modul 1: *gucken* = ‚Vollverb‘; Modul 2: *guck mal* = ‚Verbalsyntagma‘). Daneben werden lesartübergreifende Daten zur Frequenz des jeweiligen Stichwortes bezüglich der errechneten korpusbasierten Häufigkeitsklassendifferenz zwischen den Daten der Korpora DeReKo und FOLK im Vergleich in visualisierter Form angeboten und der direkte Zugriff auf FOLK und den *Lexical Explorer* bereitgestellt.

Weitere ressourcenexterne lexikografische Informationen zu den einzelnen Stichwörtern werden über die Verlinkung zu verschiedenen anderen elektronischen Wörterbüchern angeboten (z. B. zu DWDS und OWID). Ein Kurzhinweis auf relevante Forschungsliteratur ermöglicht die Verknüpfung zu den ausführlichen Literaturangaben, die für die einzelnen

Stichwörter der LeGeDe-Ressource relevant sind und die als Umtextangebot vorliegt.

4.3 Lexikografische Angaben in Modul 1

Die Bedeutungsinformation von Modul 1 wird neben relevanten lesartübergreifenden formalen Informationen hauptsächlich durch lesartspezifische Information zur Kombinatorik ergänzt (Strukturmuster, feste Verbindungen/Kollokationen). Ein weiteres optionales Informationsangebot ermöglicht Kommentare zu Besonderheiten.

Die formal-kategoriale Information der Einwortlemmata von Modul 1 erfolgt über grammatische Kategorien (z. B.: Vollverb: *gucken*, Hilfsverb: *kriegen*, Adjektiv: *gut*, Adverb: *eben*, Nomen: *Ahnung*). Weitere optionale formale Informationen beziehen sich auf die interne Morphemstruktur (z. B. *Ahnung*: ahn- (Verbalstamm) – Derivationsuffix: -ung). Neben der gelisteten Information zu möglichen Formvarianzen auf graphematischer (z. B. *gucken*: <kucken>) und/oder phonetischer Ebene (z. B.: *gucken*: [kucken]) steht außerdem ein optionaler Textblock für Auffälligkeiten

ten zur Verfügung. Zur Forschungsliteratur finden die Benutzer/-innen an dieser Stelle Kurzverweise.

Die einzelnen Lemmata weisen im Allgemeinen unterschiedliche Lesarten auf. So konnten z. B. in der LeGeDe-Stichprobe für das Verb *gucken* neun Lesarten identifiziert werden, während die gängigen Wörterbücher nicht mehr als drei unterschiedliche Lesarten verzeichnen. Die einzelnen Lesarten werden durch eine Kurzzettikettierung zur Bedeutungsdisambiguierung semantisch gekennzeichnet (z. B. ‚richtungsorientiert visuell wahrnehmen‘ und ‚überprüfen‘ für *gucken*) und ihre Bedeutung durch Bedeutungsparaphrasen beschrieben:

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‘.

Für alle Lesarten werden illustrative Audio-belege und ein entsprechender Transkriptausschnitt angeboten. Diese sind in einen Belegblock integriert, zu dem neben einem aussagekräftigen Belegtitel auch eine kurze Beschreibung des Kontextes und optional ein Analysekommentar gehören. Zusätzlich existiert die Möglichkeit, in einem kurzen optionalen Kommentar auf formale und/oder kombinatorische Auffälligkeiten hinzuweisen. Diese Information lautet z. B. für *schauen* in der Bedeutung von ‚richtungsorientiert visuell wahrnehmen‘ wie folgt:

Der häufige auffordernd imperativische Gebrauch in Verbindung mit der Modalpartikel *mal* ist auffällig. Die Sprechenden fordern die Interagierenden auf, ihren Blick in eine bestimmte Richtung zu wenden, um etwas visuell wahrzunehmen. Dabei werden für den aktuellen lokalen Bezug häufig deiktische Ausdrücke (hier, da) oder Demonstrativa (des/das) verwendet. Der Gebrauch des deontischen Infinitives ist ebenfalls belegt und kann mit Negation und in reduplizierter Form auftreten.

Die differenzierte Beschreibung der Kombinatorik in Verbindung mit den einzelnen Lesarten wird durch die ausführliche Information zu den (a) Strukturmustern, (b) den festen Verbindungen und Kollokationen und (c) den interaktionalen Einheiten abgedeckt. Zahlreiche Kurzbelegblöcke illustrieren die entsprechenden Phänomene.

(a) Die Strukturmuster werden zunächst durch eine abstrahierte Formel umschrieben. Die einzelnen Aktanten, aus denen die Strukturmuster zusammengesetzt sind, werden bezüglich ihrer semantischen Rollen, ihrer syntaktischen Funktion und den morphosyntaktischen Realisierungsmöglichkeiten genauer beschrieben und es wird auf Besonderheiten hingewiesen.

(b) Unter dem Oberbegriff ‚Feste Wendungen/Kollokationen‘ werden in der LeGeDe-Ressource verschiedene Arten von mehr oder weniger festen lexikalischen Einheiten zusammengefasst, ohne eine weitere Spezifizierung bzw. terminologische Präzisierung vorzunehmen. So stehen hier Kollokationen (z. B. *eben*: ‚vorhin‘ → gerade eben/eben gerade), feste (idiomatische) Wendungen (z. B. *wissen*: ‚informiert sein‘ → Bescheid wissen), Routineformeln (z. B. *gut*: ‚positiv bewertend‘ → guten Appetit), Sprichwörter (z. B. *schauen*: ‚richtungsorientiertes visuelles Wahrnehmen‘ → Einem geschenkten Gaul *schaut* man nicht ins Maul) etc. nebeneinander.

(c) Komplexe lexikalische Einheiten, die in der LeGeDe-Stichprobe als interaktional relevante Muster identifiziert und in direkter inhaltlicher Verbindung mit der einen oder anderen Lesart des lexikalischen Nukleus des Musters belegt werden konnten, werden aufgelistet und knapp in ihrer Funktion beschrieben. Die Einrichtung einer direkten Schnittstelle zwischen beiden Modulen ermöglicht hier die interne Verknüpfung zu den entsprechenden ausführlichen Funktionsbeschreibungen der Muster in Modul 2 (vgl. Information zu den interaktionalen Einheiten von *wissen* in der Bedeutung von ‚informiert sein‘: *ich weiß nicht*: Unsicherheitsmarker, *ich weiß*: Anzeige von Wissen, *ich weiß*: Beanspruchung von Wissen etc.).

In einem weiteren optionalen Textfeld wird auf auffällige Metadaten und ihre Frequenz hingewiesen. So lautet z. B. die Information zu

schauen in der Bedeutung von ‚richtungsorientiert visuell wahrnehmen‘ wie folgt:

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.

4.4 Lexikografische Angaben in Modul 2

In Modul 2 werden neben funktionsübergreifender Information (u. a. zu formalen, kategorialen und kombinatorischen Phänomenen) ausführliche funktionspezifische Informationen beschrieben, die durch Kommentare zur Syntax/Sequenz und Prosodie angereichert werden. Auch für dieses Modul stehen ausführliche Belegblöcke zur Verfügung, die die entsprechenden Transkriptausschnitte mit Verbindung zur Audiodatei anbieten.

Die formal-kategoriale Information der Einwort- und Mehrwortlemmata erfolgt in Modul 2 über verschiedene Kategorien (z. B. Nominalsyntaxma: *keine Ahnung*; Verbalsyntaxma: *ich weiß nicht, mal gucken*; Diskurspartikel: *eben, ja*; Modalpartikel: *eben, halt*). Zu komplexen Formen werden genauere formale Informationen bezüglich der beteiligten

Elemente angeboten (z. B.: *weißt du*: Phrase aus der 2. P. Sg. des Verbs *wissen* (weißt) und dem Personalpronomen *du*). Mögliche formale Varianten werden auf mehreren Ebenen kommentiert: so z. B. auf der lautlichen Ebene (*wissen*: [weißt du], [weißte], [weiß_u], [weiß_du] teils Klitisierung und phonetische Reduktion) oder der kompositionellen Ebene (*(ich) weiß nicht/weiß (ich) nicht*). Die in der LeGeDe-Stichprobe belegten Möglichkeiten zur Kombinatorik werden aufgelistet und mit einem optionalen Kommentar versehen. Kurzverweise zu weiterführender Literatur, die mit der Literaturliste als Umtext verlinkt sind, können ebenfalls aufgelistet werden. Zu allgemeinen Auffälligkeiten in Verbindung mit den Metadaten wurde ein optionales Kommentarfeld vorgesehen (z. B. *eben* als Modalpartikel).

Jede einzelne Funktion, die einem Lemma zugeschrieben werden kann, wird entsprechend etikettiert und durch eine Funktionskurzbeschreibung ergänzt. So weist z. B. *keine Ahnung* als komplexe Mehrworteinheit ein breites Funktionsspektrum auf (vgl. Abb. 2). Wenn noch eine Verknüpfung zu der Grundbedeutung des Stichwortes erkennbar ist, dann wird eine Verlinkung zu der entsprechenden Lesart von Modul 1 hergestellt (vgl. *keine Ahnung* als ‚Anzeige von Nicht-Wissen‘ und Verbindung zu *Ahnung* in der Bedeutung ‚Wissen‘).

Ein Langbelegblock mit entsprechendem Belegtitel, Belegkontext, Beleganalyse und Link zur Audiodatei dient auch in Modul 2

<p>Anzeige von Nicht-Wissen ▶</p> <p>Mit keine Ahnung zeigen Sprechende an, dass sie kein Wissen über einen Sachverhalt besitzen (vgl. dazu auch Ahnung in der Bedeutung <u>Wissen</u>).</p>
<p>Unsicherheitsmarker ▶</p> <p>Mit keine Ahnung zeigen Sprechende an, dass sie sich hinsichtlich der Wahrheit/Genauigkeit einer Angabe unsicher sind und markieren die Vagheit dieser Angabe.</p>
<p>Markierung potenzieller Unangemessenheit ▶</p> <p>Mit keine Ahnung zeigen Sprechende an, dass sie sich hinsichtlich der Angemessenheit einer Äußerung unsicher sind.</p>
<p>Markierung eines Beispiels ▶</p> <p>Mit keine Ahnung zeigen Sprechende an, dass das auf keine Ahnung folgende Beispiel ein ungefähr passendes aus einer Menge möglicher Beispiele darstellt, das jedoch zur Illustration der Aussage ausreicht.</p>
<p>etcetera-Formel ▶</p> <p>Mit keine Ahnung zeigen Sprechende an, dass eine Aufzählung potenziell fortführbar ist, schließen diese Aufzählung mit keine Ahnung jedoch zugleich ab und strukturieren somit ihren Redebeitrag.</p>

Abb. 2: Modul 2: Etikettierung und Funktionskurzbeschreibung von *keine Ahnung*



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

- MI 001 FR er IHAT! mich ja sogar selber gefagt;=
- MI 002 FR =ob ich das überhaupt SCHAFfe zeitlich,
- MI 003 FR un dann hab [ich gSAGT=]
- MI 004 EG [ja,]
- MI 005 FR =ja ich denke scho:n,
- MI 006 FR aber ich HAB schon viel zu tun;
- MI 007 FR °h un wenn ich dann sage ich mach es IMMER noch gerne abe:r;
- MI 008 FR (.) °h ähm ende FEbrua:r,
- MI 009 FR äh kann ich würd des WIRKLich machen?=
MI 010 FR =weil des STIMMT,
- MI 011 FR ich hab dann auch nich mehr so VIEL zu tun,
- MI 012 FR °h (.) weil GUCK mal,=
MI 013 FR =des probleM is jetzt auch,
- MI 014 FR das hatt ich nämlich gar nich bedACHT,
- MI 015 FR ich bin jetzt auch jeden tag bei der (unverständlich) gewesen,
- MI 016 FR einfach DAdurch dass ich ja-
- MI 017 FR °hh jetzt für ne woche WEGfahre,
- MI 018 FR dann fällt (.) dieser ZEITraum weg;
- MI 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 nicht 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 jetzt 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. 3: Modul 2: Belegblock zu *guck mal*: Diskursmarker

zur Illustration (vgl. Belegblock von *guck mal* als Diskursmarker in Abb. 3).

Eine Funktionsabstrahierung bietet resümierende Information zu jeder Funktionsart eines Lemmas von Modul 2 an und lautet z. B. für *guck mal* als Diskursmarker wie folgt:

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‘).

Informationen sowohl zur Syntax- und Sequenz-Realisierung als auch zur Prosodie werden in der Ressource ebenfalls bereitgestellt und mit kurzen Transkriptausschnitten belegt. Dabei kann diese Information distinktiv sein, wie in dem Fall von *eben* als Modalpartikel (unbetont, prosodisch integriert) vs. *eben* als Diskurspartikel (betont, prosodisch selbstständig) belegt wird.

5 *Fazit und Ausblick*

5.1 *Innovative Aspekte*

Da das LeGeDe-Projekt auf keine Vorbilder zurückgreifen konnte, ist allein 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, eine innovative Leistung. Weitere konkrete innovative Aspekte der Ressource sind u. a. folgende: (i) Die Ressource basiert ausschließlich auf korpusbasierten Daten. (ii) Die korpusbasierten Daten wurden durch ein eigens entwickeltes methodologisches Vorgehen quantitativ ermittelt und qualitativ analysiert und strukturiert. (iii) Die Erstellung der Stichwortkandidatenliste erfolgte durch eine korpusbasierte Methode des Häufigkeitsklassenvergleiches zwischen zwei (Teil-)Korpora. (iv) Das Informationsangebot zu den lexikografischen Angabeklassen, welche klassische Angaben mit neuartigen Angabeklassen kombiniert, ist multimedial. (v) Das Angebot von authentischen Korpusbelegen erfolgt zunächst über ausgewählte Transkriptausschnitte, über die eine Schnittstelle zu den Audiodateien und

den ausführlichen Informationen zu den Metadaten der DGD hergestellt wird. (vi) Der multimediale Charakter der Ressource zeichnet sich 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* ermöglicht den Interessierten außerdem einen direkten Zugriff auf die der Ressource zugrundeliegenden Korpusdaten und eigene erweiterte quantitative Analysemöglichkeiten. (vii) Die völlige Neukonzipierung einer korpusbasierten lexikografischen Ressource zur Darstellung von gesprochensprachlichen Spezifika ermöglichte die konkrete Berücksichtigung von bestimmten, empirisch erhobenen Erwartungshaltungen zu Inhalt, Präsentationsmodalitäten und Funktionalitäten zukünftiger Benutzer/-innen an so eine neuartige Ressource.

5.2 Anwendungsmöglichkeiten im DaF- und DaZ-Bereich

Die Frage nach der anvisierten Zielgruppe und Benutzungssituation spielt eine fundamentale Rolle für die Art und Weise der lexikografischen Umsetzung der erhobenen Daten und ihrer Darstellung. Da die LeGeDe-Ressource zunächst als Wissensspeicher und Wortschatzdokumentation fungieren soll, stehen die Daten primär einem wissenschaftlich interessierten Benutzer/-innen-Kreis zur Verfügung und wurden entsprechend mit einem relativ hohen Komplexitätsgrad dargeboten. Aus den Ergebnissen der projektbegleitenden empirischen Befragungen ist aber auch hervorgegangen, dass Benutzer/-innen in Lernsituationen (v.a. in Produktionssituationen) von der LeGeDe-Ressource profitieren könnten, wenn Expert/-innen (Wissenschaftler/-innen, Lehrende etc.) eine entsprechende Mittlerposition einnehmen (Meliss/Möhrrs/Riberiro Silveira 2018: 132; 2019: 116). Ausgehend von den Daten, die die LeGeDe-Ressource bereitstellt, könnten daher aus anwendungsorientierter Perspektive heraus u. a. korpusbasierte Lernmaterialien zur konkreten Behandlung spezifischer Phänomene der Lexik in der mündlichen Interaktion für die Nutzung im DaF- und DaZ-

Unterricht entstehen. Diese Möglichkeiten gehen konform mit den Antworten der OU zu der Frage nach der möglichen künftigen Zielgruppe einer Ressource zum Gesprochenen, denn DaF- und DaZ-Lehrende und Lernende wurden sowohl von den L1- als auch von den L2-Befragten als Hauptzielgruppe gesehen (Meliss/Möhrrs/Riberiro Silveira 2019: 108 f.).

5.3 Ausblick

Eine thematische Ausweitung auf andere Phänomenbereiche, die für die vorgelegte LeGeDe-Ressource in der ersten Phase nicht bearbeitet werden konnten, wäre wünschenswert. (i) Im Bereich der Entlehnungen – sowohl aus anderen Sprachen (Anglizismen: z. B. *okay*) als auch aus deutschen Sprachvarietäten (soziolektal: *geil, krass, Mama, wieso* etc.) – zeigen sich interessante, gesprochensprachliche Phänomene, die u. a. in Zusammenhang mit ihrer Gebrauchsfrequenz, dem Vorkommen in bestimmten Interaktionstypen, den Sprechergruppen, der grammatischen Integration und der phonetischen Realisierung stehen. (ii) Im Bereich der Wortbildung lassen sich u. a. zahlreiche Affixe mit formalen Besonderheiten und/oder neuen Funktions- und Bedeutungsnuancen verbuchen (vgl. Meliss 2020a). (iii) Die Untersuchung zu Gebrauch und Bedeutung von Passepartoutwörtern (z. B.: *machen, tun, Ding, Sache*) sowie von Ausdrücken der Vagheit (z. B.: *irgend-[-wann, -was, -wie, -wo]*) stehen v. a. in Verbindung mit der Bedeutungskonstituierung in der Interaktion. (iv) Die für die gesprochene Sprache in der Interaktion besonders typischen Interjektionen (z. B.: *ach, ah, oh*) sind weitere interessante Stichwortkandidaten für eine lexikografische Ressource zum Gesprochenen in der Interaktion.⁷ Die Bearbeitung der Phänomenbereiche, die in den LeGeDe-Prototypen der ersten Arbeitsphase nicht aufgenommen werden konnten, bildet zusammen mit der quantitativen Erweiterung der Lemmaanahl zu den schon

⁷ Zum Dänischen wurde ein korpusbasiertes lexikografisches Projekt zu Interjektionen (Hansen/Hansen 2012) durchgeführt.

behandelten Phänomenen ein breites zukünftiges lexikografisches Arbeitsfeld. Darüber hinaus würde die zukünftige Entwicklung einer erweiterten Suchfunktion u. a. über kategoriale Parameter und kommunikative Funktionen besonders dem Wunsch der L2-Befragten der OU entgegenkommen (Meliss/Möhrs/Ribeiro Silveira 2019: 111 f.) und könnte somit eine Verbindung zu Nachschlagehandlungen für Produktioszwecke im L2-Bereich herstellen.

Abschließend kann festgehalten werden, dass die LeGeDe-Ressource dank der unterschiedlichen Zugriffe auf die Korpusdaten und das strukturierte und seligierte multimediale Informationsangebot neben ihrem primären Ziel der sprachwissenschaftlichen Dokumentation vielfältige Möglichkeiten für anwen-

dungsorientierte Szenarien bietet. Das Glossar und die Benutzungshinweise, die als Umtex te zur Verfügung stehen, leisten dazu u. a. eine adäquate Hilfestellung. Bedingung dafür ist allerdings, dass die Sprachdidaktiker/-innen diese Nutzungsmöglichkeiten kennen. Daher soll für eine intensivere Beschäftigung im Studium und durch Weiterbildungsangebote im Bereich der anwendungsorientierten Korpuslinguistik und der Lexikografie plädiert werden.

Prof. Dr. Meike Meliss

Universidad de Santiago de Compostela

Facultad de Filología, Departamento de Filología

meike.meliss@usc.es

Literatur

- Atkins, Beryl T.S./Rundell, Michael (2008): *The Oxford guide to practical lexicography*. Oxford.
- Barcala, Mario/Domínguez, Eva/Fernández, Alba/Rivas, Raquel/Santalla, María Paula/Vázquez, Victoria/Villapol, Rebecca (2018): *El corpus ES-LORA de español oral: diseño, desarrollo y explotación*. In: CHIMERA: Romance Corpora and Linguistic Studies, 5/2, 217–237.
- Blühndorn, Hardarik/Depperman, Arnulf/Helmer, Henrike/Spranz-Fogasy, Thomas (Hg.) (2017): *Diskursmarker im Deutschen. Reflexionen und Analysen*. Göttingen.
- Deppermann, Arnulf (2007): *Grammatik und Semantik aus gesprächsanalytischer Sicht*. Berlin.
- Deppermann, Arnulf/Proske, Nadine/Zeschel, Arne (Hg.) (2017): *Verben im interaktiven Kontext. Bewegungsverben und mentale Verben im gesprochenen Deutsch*. Tübingen, 7–40.
- Engelberg, Stefan (2015): *Gespaltene Stimuli bei Psych-Verben. Kombinatorische Mustersuchen in Korpora zur Ermittlung von Argumentstrukturverteilungen*. In: Engelberg, Stefan/Meliss, Meike/Proost, Kristel/Winkler, Edeltraud (Hg.): *Argumentstruktur zwischen Valenz und Konstruktion*. Tübingen, 469–491.
- Europarat (2001): *Gemeinsamer europäischer Referenzrahmen für Sprachen: lernen, lehren, beurteilen*. Europarat. Rat für kulturelle Zusammenarbeit. Berlin.
- Fandrych, Christian/Meißner, Cordula/Wallner, Franziska (2018): *Das Potenzial mündlicher Korpora für die Sprachdidaktik. Das Beispiel Ge-Wiss. Deutsch als Fremdsprache 57/1, 3–13*.
- Fandrych, Christian/Meißner, Cordula/Wallner, Franziska (Hg.) (2017): *Gesprochene Wissenschaftssprache – digital. Verfahren zur Annotation und Analyse mündlicher Korpora*. Tübingen.
- Fiehler, Reinhard (2016): *Gesprochene Sprache*. In: Wöllstein, Angelika (Hg.): *Duden – Die Grammatik. Unentbehrlich für richtiges Deutsch*. Berlin, 1181–1260.
- Günthner, Susanne (2018): *Routinisierte Muster in der Interaktion*. In: Deppermann, Arnulf/Reineke, Silke (Hg.): *Sprache im kommunikativen, interaktiven und kulturellen Kontext*. Berlin/Boston, 29–50.
- Handwerker, Brigitte/Bäuerle, Rainer/Sieberg, Bernd (Hg.) (2016): *Gesprochene Fremdsprache Deutsch*. Baltmannsweiler.
- Hansen, Carsten/Hansen, Martin H. (2012): *A dictionary of spoken Danish*. In: Fjeld, Ruth V./Torjusen, Julie M. (Hg.): *Proceedings of the 15th EURALEX International Congress. 7–11 August 2012. Oslo, Norway*. Oslo, 929–935.
- Imo, Wolfgang (2007): *ConstructionGrammar und Gesprochene-Sprache-Forschung. Konstruktionen mit zehn matrixsatzfähigen Verben im gesprochenen Deutsch*. Tübingen.

- Imo, Wolfgang/Moraldo, Sandro (Hg.) (2015): Interaktionale Sprache und ihre Didaktisierung im DaF-Unterricht. Tübingen.
- Keibel, Holger (2008, 2009): Mathematische Häufigkeitsmaße in der Korpuslinguistik: Eigenschaften und Verwendung. Mannheim. Unter: <http://www.ids-mannheim.de/kl/dokumente/freqMeasures.html> [24.06.2019].
- Klosa, Annette (2013): The lexicographical process (with special focus on online dictionaries). In: Gouws, Rufus/Heid, Ulrich/Schweickard, Herbert/Wiegand, Herbert Ernst (Hg.): Dictionaries. An international encyclopedia of lexicography. Supplementary volume: Recent developments with focus on electronic and computational lexicography. Berlin, 517–524.
- Klosa, Annette/Tiberius, Carole (2016): Der lexikografische Prozess. In: Klosa, Anette/Müller-Spitzer, Carolin (Hg.): Internetlexikografie. Ein Kompendium. Berlin, 65–110.
- Kupietz, Marc/Lüngen, Harald/Kamocki, Pawel/Witt, Andreas (2018): The German Reference Corpus DeReKo: New Developments – New Opportunities. In: Calzolari, Nicoletta/Choukri, Khalid/Cieri, Christopher/Declerck, Thierry/Goggi, Sara/Hasida, Koiti/Isahara, Hitoshi/Maegaard, Bente/Mariani, Joseph/Mazo, Hélène/Moreno, Asuncion/Odijk, Jan/Piperidis, Stelios/Tokunaga, Takenobu (Hg.): Proceedings of the Eleventh International Conference on Language Resources and Evaluation (LREC 2018). Miyazaki, 4353–4360.
- Kupietz, Marc/Schmidt, Thomas (2015): Schriftliche und mündliche Korpora am IDS als Grundlage für die empirische Forschung. In: Eichinger, Ludwig M. (Hg.): Sprachwissenschaft im Fokus. Positionsbestimmungen und Perspektiven. Berlin, 297–322.
- Meliss, Meike (2016): Gesprochene Sprache in DaF-Lernerwörterbüchern. In: Handwerker, Brigitte/Bäuerle, Rainer/Sieberg, Bernd (Hg.): Gesprochene Fremdsprache Deutsch. Baltmannsweiler, 179–199.
- Meliss, Meike (2020a): Lexik des gesprochenen Deutsch in der Interaktion: Partikelverben und ihre Besonderheiten. In: García Adánez, Isabel/Valdés, Jesús Gil/Hernández, Paloma Sánchez/Schilling Rodríguez, Luisa M./Szumlakowski Morodo, Irene (Hg.): Das Leben in einem Rosa Licht sehen. Festschrift für Rosa Piñel/Homenage a Rosa Piñel (= Perspektiven der germanistik und Komparatistik in Spanien, Bd. 16). Bern, 77–97.
- Meliss, Meike (2020b): Musterhaftigkeit in der Lexik des gesprochenen Deutsch: Eine handlungs- und funktionsorientierte Annäherung mit lexikografischer Anwendungsperspektive. In: Mellado Blanco, Carmen/Holzinger, Herbert/Mansilla, Ana/Iglesias, Nely M. (Hg.): Produktive Muster in der Phraseologie. (= Studia Phraseologica et Paroemiologica, Bd. 4) Hamburg, 181–202.
- Meliss, Meike/Möhrs, Christine (2018): Lexik in der spontanen, gesprochensprachlichen Interaktion: Eine anwendungsorientierte Annäherung aus der DaF-Perspektive. In: German as a foreign language 3, 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 (Hg.): Germanistik im Umbruch – Linguistik, DaF und Übersetzung. Berlin, 13–25.
- Meliss, Meike/Mörs, Christine/Batinic, Dolores/Perkuhn, Rainer (2018): Creating a List of Headwords for a Lexical Resource of Spoken German. In: Cibej, Jaka/Gorjanc, Vojko/Kosem, Iztok/Krek, Simon (Hg.): Proceedings of the XVIII EURALEX International Congress. Lexicography in Global Contexts, 17–21 July, Ljubljana. Ljubljana, 1009–1016.
- Meliss, Meike/Möhrs, Christine/Ribeiro Silveira, Maria (2018): 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 (2019): Anforderungen und Erwartungen an eine lexikografische Ressource des gesprochenen Deutsch aus der L2-Lernerperspektive. In: Lexicographica 34. Berlin/Boston, 2018, 89–121.
- Meliss, Meike/Möhrs, Christine/Ribeiro Silveira, Maria/Schmidt, Thomas (2019): A Corpus-Based Lexical Resource of Spoken German in Interaction. In: Kosem, Iztok/Zingano Kuhn, Tanara/Correia, Margarita/Ferreira, José Pedro/Jansen, Maarten/Pereira, Isabel/Kallas, Jelena/Jakubiček, Miloš/Krek, Simon/Tiberius, Carole (Hg.): Electronic lexicography in the 21st century. Proceedings of the eLex 2019 conference. 1–3 October 2019, Sintra, Portugal. Brno, 783–804.
- Möhrs, Christine (2020): „Hast du eine Ahnung, ...?“ Eine lexikografische und korpusbasierte Untersu-

- chung am Beispiel des Lexems Ahnung. In: Behr, Janina/Conrad, François/Kornmesser, Stephan/Tschernig, Kristin (Hg.): Schnittstellen der Germanistik. Festschrift für Hans Bickes. Berlin: Peter Lang, 185–214.
- Möhrs, Christine /Meliss, Meike/Batinić, Dolores (2017): LeGeDe – towards a corpus-based lexical resource of spoken German. In: Kosem, Itzok/Tiberius, Carole/Jakubiček, Miloš/Kallas, Jelena/Krek, Simon/Baisa, Vit (Hg.): Electronic Lexicography in the 21st century. Proceedings of eLex 2017 Conference. 19–21 September 2017. Leiden, the Netherlands. Brno, 281–298.
- Moon, Rosamund (1998): On using spoken data in corpus lexicography. In: Fontenelle, Thierry/Hilgsmann, Philippe/Michiels, Archibald/Moulin, André/Theissen, Siegfried (Hg.): Proceedings of the 8th EURALEX International Congress. 4–8 August 1998. Liège, Belgium. Liège, 347–355.
- Schmidt, Thomas (2018): Gesprächskorpora. In: Kupietz, Marc/Schmidt, Thomas (Hg.): Korpuslinguistik. Berlin/Boston, 209–230.
- Schwitalla, Johannes (2012): Gesprochenes Deutsch. Eine Einführung. 4. Aufl. Berlin.
- Sieberg, Bernd (2013): Sprechen lehren, lernen und verstehen. Stufenübergreifendes Studien- und Übungsbuch für den DaF-Bereich. Tübingen.
- Trap-Jensen, Lars (2004): Spoken language in dictionaries: Does it really matter? In: Williams, Geoffrey/Vessier, Sandra (Hg.): Proceedings of the 11th EURALEX International Congress. 6–10 July 2004. Lorient, France. Lorient, 311–318.
- Verdonik, Darinka/Sepesy Maučec, Mirjam (2017): A speech corpus as a source of lexical information. In: International Journal of Lexicography 30, 143–166.

Webseiten

- DGD. Datenbank für gesprochenes Deutsch (2012). Unter: https://dgd.ids-mannheim.de/dgd/pragdb.dgd_extern.welcome [04.07.2019].
- FOLK. Forschungs- und Lehrkorpus Gesprochenes Deutsch (Release 2.8 vom 06.04.2017). Unter: <http://agd.ids-mannheim.de/folk.shtml> [04.07.2019].
- LeGeDe. Lexik des gesprochenen Deutsch (2019). Unter: <https://www.owid.de/legede/article-main.jsp> [13.01.2020].
- DeReKo: Institut für Deutsche Sprache (2017). Deutsches Referenzkorpus/Archiv der Korpora geschriebener Gegenwartssprache 2017-I (Release 08.03.2017). Mannheim: Institut für Deutsche Sprache. Unter: <http://www.ids-mannheim.de/direktion/kl/projekte/korpora/releases.html>. [04.07.2019].
- Lexical Explorer (2018). Unter: <https://www.owid.de/lexex/index>. [13.01.2020].