

ZUMAL: ZIELGRUPPENSPEZIFISCHE GESPRÄCHSAUSWAHL AUS KORPORA GESPROCHENER SPRACHE

Christian Fandrych, Universität Leipzig
Cordula Meißner, Universität Innsbruck
Matthias Schwendemann, Universität Leipzig
Franziska Wallner, Universität Leipzig

Abstract

ZuMal ist ein im Rahmen des Projekts *ZuMult* entwickeltes Werkzeug, welches eine Vorauswahl an Sprechereignissen aus den beiden Korpora *FOLK* und *GeWiss* nach sprachdidaktisch relevanten und schwierigkeitsbezogenen Kriterien ermöglicht. Hierfür werden zum einen Kriterien genutzt, die auf die Metadaten der Sprechereignisse zurückgehen (so etwa die Art des Gesprächs, bspw. Telefongespräch, Gespräch beim Arbeitsamt, Tischgespräch; die Gesprächsthemen, die Sprachregion und die Dauer). Zum anderen wird auf Informationen zurückgegriffen, die automatisiert aus den Daten gewonnen werden können und aus der Perspektive der Vermittlung des Deutschen als Fremd- und Zweitsprache mit besonderen Herausforderungen für Lerner:innen mit Deutsch als L2 assoziiert werden. Dazu gehören u.a. der Wortschatz und dessen Zugehörigkeit zu den Niveaustufen des gemeinsamen Europäischen Referenzrahmens, die Standardnähe/-ferne, die Sprechgeschwindigkeit, die Anzahl an Überlappungen verschiedener Sprecher:innen sowie das Vorkommen ausgewählter Mündlichkeitsphänomene und Wortarten. Im Beitrag sollen die in *ZuMal* bereitgestellten Auswahlparameter näher beschrieben werden. Im Fokus stehen dabei insbesondere die sprachdidaktischen Konzepte, die den mit Schwierigkeit assoziierten Parametern zugrunde gelegt wurden. Anhand exemplarischer Auswertungen wird zudem gezeigt, welche Ausprägungen die aktuell in *FOLK* und *GeWiss* enthaltenen Daten im Hinblick auf die Auswahlparameter zeigen.

Keywords: gesprochene Sprache; mündliche Korpora; FOLK; GeWiss; Korpora in DaF/DaZ; Merkmalsauswahl von Sprechereignissen

Abstract

ZuMal is a tool developed within the framework of the *ZuMult* project which enables a pre-selection of speech events from the two corpora *FOLK* and *GeWiss* according to language teaching criteria as well as criteria related to linguistic difficulty. The criteria are drawn from the metadata of the speech events (such as the type of conversation, e.g. telephone conversation, conversation at the employment office, kitchen table conversation; the topics of conversation, the language region and the duration of the speech event). In addition, information is used which can be obtained in an automated way from the data and which, from the perspective of teaching German as a foreign and second language, is associated with particular challenges for learners of German as an L2. These include, among others, vocabulary and its relation to the levels of the Common European Framework of Reference for Languages, closeness/remoteness to standard language, articulation rate, the number of overlaps of different speakers as well as the occurrence of selected oral phenomena and word classes. The article focuses on the parameters on which the selection options provided by *ZuMal* are based, in particular the didactic concepts underlying the parameters associated with linguistic difficulty. By means of exemplary evaluations, it will also be shown which characteristics the data currently contained in *FOLK* and *GeWiss* show with regard to the selection parameters.

Keywords: spoken language; oral corpora; FOLK; GeWiss; corpora in GFL/GSL; feature selection of speech events

1. Einleitung

Die Relevanz von möglichst realen, nicht simulierten oder eigens für didaktische Zwecke erstellten Gesprächen und Texten für den Sprachunterricht ist sehr groß: Nimmt man das Ziel der sprachlichen

Handlungsfähigkeit in einer zielsprachigen Umgebung ernst, so muss man die Lernenden auch mit den entsprechenden sprachlichen Handlungsmustern und den damit verbundenen sprachlichen Formen und Routinen vertraut machen. Dies gilt umso mehr für Lehr-Lernsituationen fernab vom deutschen Sprachraum. Auch wenn Digitalisierung, Medialisierung und weltweite kommunikative Vernetzung inzwischen deutlich mehr Möglichkeiten für den direkten Kontakt mit der deutschen Sprache geschaffen haben (auch in ihrer mündlichen Form), und trotz des wachsenden Angebots an spezifischen Sprachressourcen (online-Wörterbücher, Grammatik-Nachschlagerressourcen, Korpora der geschriebenen und gesprochenen Sprache) bleibt das Problem bestehen, gezielt kommunikative Ereignisse zu finden, die sich für eine konkrete Unterrichtssequenz und für bestimmte Lernziele besonders gut eignen. Insbesondere Korpora, also systematisch nach bestimmten Kriterien erhobene und aufbereitete Sprachdatenbanken, stellen für den Sprachunterricht potenziell wertvolle Ressourcen dar. Für das Deutsche gilt jedoch, dass wichtige und größere Korpora meist mit bestimmten (linguistischen) Forschungsinteressen geschaffen und auch für diese Ziele aufbereitet und nutzbar gemacht wurden (vgl. etwa Kupietz / Schmidt 2018; Deppermann et al. 2023), wodurch sie nur bedingt für die direkte Nutzung in sprachdidaktischen Kontexten verwendbar sind¹. Dies liegt zum einen am Design der jeweiligen Korpora, das viele aus sprachdidaktischer Sicht relevante Suchfunktionen nicht vorsieht, zum anderen aber auch an den für ihre Nutzung vorausgesetzten korpuslinguistischen Kompetenzen, die für viele Gruppen von Nutzer:innen eine relativ große Hürde darstellen (vgl. Fandrych et al. 2016).

Aus dieser Erkenntnis heraus wurde im Projekt *ZuMult (Zugänge zu multimodalen Korpora gesprochener Sprache)*² mit *ZuMal (Zugang zu Merkmalsauswahl von Gesprächen)*³ ein Werkzeug entwickelt, das es erlaubt, mündliche Interaktionen in größeren Korpora des Deutschen aus sprachdidaktischer Perspektive gezielt auszuwählen. *ZuMal* wurde so angelegt, dass es auch auf andere gesprochensprachliche Korpora anwendbar ist. Gegenwärtig ist es allerdings auf die Recherche in zwei größeren Korpora der gesprochenen Sprache beschränkt, die beide über die Datenbank für Gesprochenes Deutsch (DGD)⁴ des Leibniz-Instituts für Deutsche Sprache (kurz: IDS) abrufbar sind: Das *Forschungs- und Lehrkorpus Gesprochenes Deutsch (FOLK)*, das 400 Interaktionen mit einer Gesamtlänge von ca. 336 Stunden und 3,2 Millionen Token umfasst (Stand Januar 2023) sowie das Korpus *Gesprochene Wissenschaftssprache kontrastiv (GeWiss)*, das ausgewählte Gattungen aus der Hochschulkommunikation (im Bereich Germanistik / Deutsch als Fremdsprache sowie verschiedener weiterer Philologien) beinhaltet, mit 436 Interaktionen im Umfang von 146 Aufnahmestunden und ca. 1,2 Millionen Token⁵.

Für die Auswahloptionen, die im Tool *ZuMal* angelegt wurden, wurde nun zunächst ermittelt, welche der in den Daten bereits angelegten Informationen für sprachdidaktische Zwecke besonders relevant sein könnten⁶. Hierfür kamen einerseits die bei der Erarbeitung der Korpora miterhobenen Metadaten in Frage – also Angaben über die Art und den Kontext der Interaktionen, ihre Dauer, die behandelten Themen sowie die Sprecher:innen. Zum anderen wurden aber auch weniger direkt ab-

¹ Eine Ausnahme stellt die Plattform Gesprochenes Deutsch dar, die an der Universität Münster entwickelt wurde, vgl. <https://dafdaz.sprache-interaktion.de/> (15.02.2023). Sie ist allerdings in Umfang und Aufbereitung nicht mit den großen Korpora, etwa des IDS, vergleichbar.

² Siehe dazu <https://zumult.org/>. (15.02.2023).

³ Verfügbar unter <https://zumult.ids-mannheim.de/ProtoZumult/prototype/dist/zuMal.jsp> (15.02.2023).

⁴ Siehe https://dgd.ids-mannheim.de/dgd/pragdb.dgd_extern.welcome (15.02.2023). Das *GeWiss*-Korpus ist dort mit dem Kürzel *GWSS* abrufbar.

⁵ *GeWiss* ist mehrsprachig, die deutschsprachigen Daten enthalten 276 Interaktionen mit 742.332 Token und 92 Aufnahmestunden, vgl. ausführlicher Fandrych / Wallner (2023).

⁶ Eine Neuannotation war im Rahmen der Projektförderung durch die DFG – der wir an dieser Stelle herzlich danken – aufgrund der Richtlinien der Förderlinie *Wissenschaftliche Literaturversorgungs- und Informationssysteme (LIS)* nicht möglich, daher musste das Projekt sich darauf beschränken, die in den Datensätzen bereits angelegten Informationen zu sichten, auszuwählen und geeignete Wege der Operationalisierung und technischen Umsetzung zu entwickeln.

rufbare Informationen, die in den Datensätzen enthalten waren, auf ihre Relevanz für sprachdidaktische Zwecke hin gesichtet. Dabei standen zwei Aspekte im Vordergrund: Zum einen die sprachliche Schwierigkeit, denn es ist für die Sprachvermittlung natürlich wichtig, einigermaßen passgenaues Material für eine bestimmte Lernengruppe und einen bestimmten Lernkontext auswählen zu können; zum anderen aber auch die Spezifik der Mündlichkeit, die ja nach wie vor in der Sprachdidaktik häufig nicht ausreichend bedacht wird, nicht zuletzt aus Mangel an geeigneten Sprachbeispielen.

In der Folge werden diese Auswahlmöglichkeiten und die ihnen zugrunde liegenden Kriterien, Konzepte sowie die Art ihrer Operationalisierung in *ZuMal* der Reihe nach vorgestellt. Nach einer kurzen überblicksartigen Darstellung der *ZuMal*-Oberfläche (Kap. 2) folgt zunächst eine Übersicht über die metadatenbezogenen Auswahl- und Filtermöglichkeiten (Kap. 3). Im Anschluss werden sodann die schwierigkeitsbezogenen Auswahloptionen dargestellt (Kap. 4), danach die spezifisch auf die Mündlichkeit sowie auf die Frequenz von Wortarten bezogenen Filtermöglichkeiten (Kap. 5). Den Abschluss bildet ein kurzer Abschnitt, der auf das Zusammenspiel und die Kombinationsmöglichkeiten der verschiedenen Auswahloptionen eingeht (Kap. 6).

2. Die *ZuMal*-Oberfläche

Die *ZuMal*-Oberfläche wurde von Josip Batinić als Vue.js-Anwendung entwickelt, die mit den im Projekt *ZuMult* entwickelten Backend-Komponenten kommuniziert (vgl. Schmidt et al. in diesem Heft). Mit Hilfe von *ZuMal* kann gezielt nach geeigneten Sprechereignissen aus dem *FOLK* und dem *GeWiss*-Korpus gesucht werden. Hierfür wurden verschiedene Filteroptionen implementiert, die es ermöglichen, auf der Basis verschiedener Merkmale eine sprachdidaktisch orientierte Auswahl der Sprechereignisse zu treffen. Die Filter befinden sich auf der linken Seite der *ZuMal*-Oberfläche und sind in drei Blöcke untergliedert. Der erste Block beruht auf den Metadaten der Sprecherereignisse (darunter Gesprächstyp, Region, Dauer u.a., vgl. Kap. 3). Der zweite Block umfasst schwierigkeitsbezogene Parameter (darunter Niveaustufenzugehörigkeit des Wortschatzes und Sprechgeschwindigkeit, vgl. Kap. 4). Der dritte Block beinhaltet Wortarten und Mündlichkeitsphänomene (vgl. Kap. 5). Je nach Merkmal umfassen die Filter weitere Subkategorien und Schieberegler, mit denen sich die jeweilige Ausprägung der Merkmale bestimmen lässt. Die Filter können sowohl einzeln als auch in Kombination miteinander genutzt werden. Die Ergebnisse der Auswahl werden auf der rechten oberen Seite der *ZuMal*-Oberfläche in einem Streudiagramm visualisiert. Auf der x- und auf der y-Achse können dabei jeweils die auf die Dauer, die Schwierigkeit, die Wortarten und die auf Mündlichkeitsphänomene bezogenen Parameter eingestellt werden. Die Visualisierung der Ergebnismenge im Streudiagramm passt sich jeweils dynamisch den in den Merkmalsfiltern ausgewählten Einstellungen an, sodass nur die Sprechereignisse im Diagramm angezeigt werden, die den ausgewählten Merkmalsausprägungen entsprechen. Wird keine Anpassung der Filter vorgenommen, sind alle Sprechereignisse innerhalb des jeweils ausgewählten Korpus in dem Streudiagramm zu sehen (vgl. Abb. 1). Außerdem wird die durch die Anwendung der Filter vorgenommene Auswahl an Sprechereignissen in einer Tabelle unterhalb des Streudiagramms angezeigt. Ebenso wie das Streudiagramm passt sich die Anzeige der in den Merkmalsfiltern getroffenen Auswahl dynamisch an. In der voreingestellten Ansicht werden in der Tabelle die ID des Sprechereignisses im ausgewählten Korpus, die Art des Gespräches, die Dauer, die Normalisierungsrate (vgl. Kap. 4) und der Link zum Werkzeug *ZuViel* aufgelistet. Mit Hilfe dieses Links kann dann jedes einzelne Sprechereignis im Transkriptbrowser *ZuViel* (vgl. Schmidt / Schwendemann / Wallner in dieser Ausgabe) aufgerufen werden. Es ist aber auch möglich, die Tabellenansicht an individuelle Bedürfnisse anzupassen und für jeden der verfügbaren Filter die entsprechenden Spalten hinzuzufügen oder auch abzuwählen. Darüber hinaus lassen sich die Ergebnisse in dieser Ansicht auch sortieren, etwa um sich einen Überblick zu verschaffen,

welche Sprechereignisse besonders deutlich von einer (an der Schriftsprache orientierten) Standardvariante abweichen (ermittelt über „Normalisierungsrate“, vgl. Kap. 4).

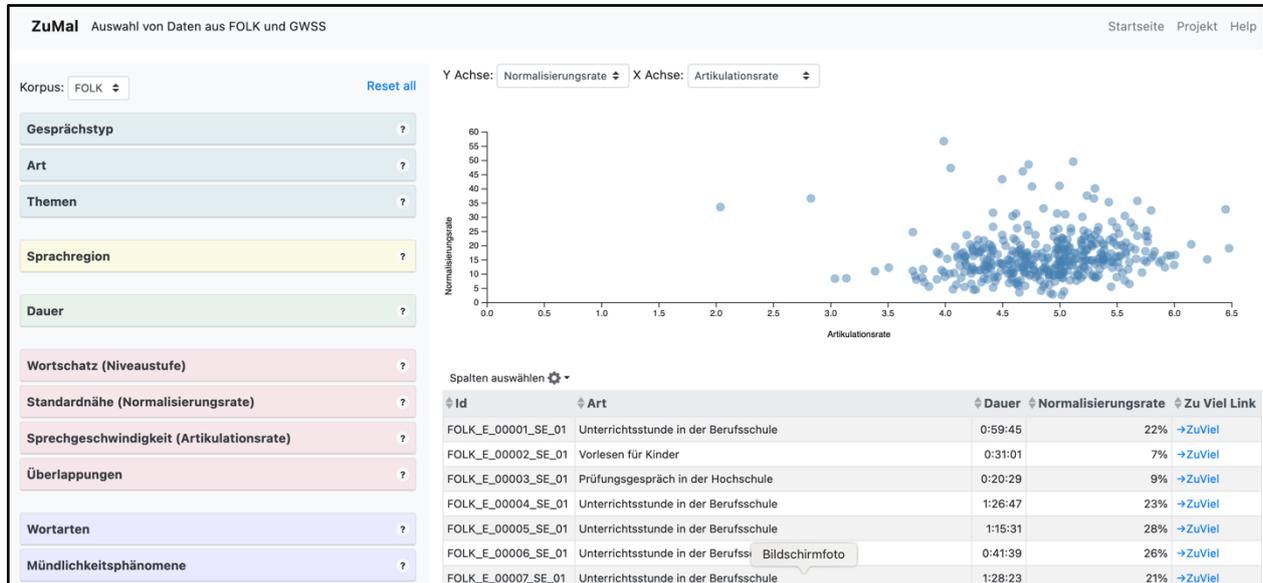


Abbildung 1
ZuMal-Oberfläche⁷

Im Folgenden werden nun die Auswahloptionen, die in der linken Spalte überblicksartig sichtbar sind, der Reihe nach dargestellt sowie die ihnen zugrunde liegenden Überlegungen transparent gemacht.

3. Metadatenbezogener Zugang

Um den in *ZuMal* angelegten metadatenbezogenen Zugang zu Interaktionen und sprachlichen Phänomenen einordnen zu können, muss zunächst ein Blick auf die Art und Gliederung der mit den beiden Korpora *FOLK* und *GeWiss* verbundenen Metadaten geworfen werden, die derzeit mithilfe von *ZuMal* durchsuchbar sind. Sie sind die Grundlage für verschiedene Auswahl- und Filtermöglichkeiten, die auch für sprachdidaktische Zwecke von großer Bedeutung sind.

Um Korpora gesprochener Sprache gezielt nach verschiedenen relevanten Merkmalen der enthaltenen Daten durchsuchen zu können, ist eine möglichst konsistente Kategorisierung der Gespräche in verschiedenster Hinsicht notwendig. So ist etwa von hoher Relevanz, in welchem Kontext bzw. in welcher Interaktionsdomäne ein Gespräch stattfindet (privat, im Kontext einer Institution, öffentlich; in raum-zeitlicher Ko-Präsenz der Teilnehmenden oder vermittelt über ein Medium), wie die Sprecherkonstellation ist (wie viele Personen nehmen an dem Gespräch teil, welche Funktion haben sie in dem Gespräch, wie vertraut sind sie untereinander) und wie das Gespräch gestaltet ist (hat es einen handlungsbegleitenden Charakter oder nicht, welche Themen werden behandelt). Diese Merkmale können noch deutlich weiter ausdifferenziert werden – bei der Interaktionsdomäne kann nach bestimmten Arten von institutionellen Kontexten weiter differenziert werden (Bildungsinstitutionen, Gesundheitswesen, Wirtschaftsbetriebe, Ämter etc.), bei der Personenkonstellation können soziode-

⁷ Vgl. <https://zumult.ids-mannheim.de/ProtoZumult/prototype/dist/zuMal.jsp> (15.02.2023).

mografische Merkmale der Sprecher:innen wie Alter, Herkunft, Bildungshintergrund, Sprachkompetenzen, beruflicher Hintergrund und vieles andere mehr wichtige Parameter darstellen, und die Zahl der Themen und der Themenfokussierung ist im Prinzip unbegrenzt und nicht leicht zu klassifizieren.

Die genannten Merkmale sind für verschiedene Typen von Korpora unterschiedlich wichtig. Strebt man, wie das *Forschungs- und Lehrkorpus Gesprochenes Deutsch (FOLK)*, das am Leibniz-Institut für Deutsche Sprache angesiedelt ist, eine möglichst große Ausgewogenheit der in einem Korpus enthaltenen Daten in Bezug auf interaktionale, medial mündliche Sprache in Deutschland an (vgl. Deppermann / Hartung 2012: 418), dann stellt sich die Frage nach der Art von Merkmalen sowie ihrer Kategorisierung und Verteilung in sehr grundlegender Weise. Bei spezifischeren Korpora, wie es das Korpus *Gesprochene Wissenschaftssprache kontrastiv (GeWiss)* mit seiner begrenzten Zahl an kommunikativen Ereignissen in einer klar definierten kommunikativen Domäne darstellt, ist schon von vornherein eine relativ strikte Vorauswahl getroffen worden, die das Spektrum der Gesprächsmerkmale deutlich einschränkt. Während bei *FOLK* eine möglichst breite Abdeckung verschiedenster als besonders relevant erkannter Merkmale im Vordergrund steht, um Aussagen über das Spektrum der Mündlichkeit in unterschiedlichsten regionalen, sozialen und situativen Kontexten sowie bezüglich verschiedener Merkmale der Sprechenden machen zu können (vgl. Deppermann / Hartung 2012: 418-419; Kaiser 2018; Reineke / Deppermann / Schmidt 2023: 71-75), steht bei *GeWiss* die intra- und interlinguale Vergleichbarkeit gesprochener Sprache in relativ genau definierten Gattungen von Sprechenden mit einem verhältnismäßig ähnlichen Bildungshintergrund im Vordergrund (vgl. Fandrych / Meißner / Slavcheva 2012; Fandrych / Wallner 2023).

Die bei der Korpuserstellung als relevant angesehenen Merkmale werden dabei mithilfe von Metadaten erhoben und mit den Gesprächsereignissen verknüpft. So werden die Gesprächsereignisse gezielt nach bestimmten Eigenschaften durchsuchbar. Die Bestimmung, Auswahl und Zuordnung der Merkmale sowie die Abdeckung und Kombination der Merkmale ist eine sowohl theoretisch wie praktisch sehr herausfordernde Aufgabe, insbesondere, wenn – wie bei *FOLK* – das langfristige Ziel der Aufbau eines Referenzkorpus ist (vgl. Deppermann / Hartung 2012: 438-441; Kaiser 2018: 516-520). Neben der Auswahl der Parameter und Merkmale stellt auch deren Heterogenität und Hierarchisierung ein theoretisches wie praktisches Problem dar. Bei *FOLK* wird zwischen „primären Parametern“, welche das Gesprächsereignis betreffen, und „sekundären Parametern“, die sich auf Merkmale der Sprechenden beziehen, unterschieden (vgl. Kaiser 2018: 543-546; Reineke / Deppermann / Schmidt 2023: 74-75). Bei den gesprächsbezogenen primären Parametern wird zunächst nach vier grundlegenden *Interaktionsdomänen* unterschieden, nämlich privat, institutionell, öffentlich sowie „Sonstiges“⁸. Weitere Untergliederungen betreffen dann die verschiedenen *Lebensbereiche*, in denen die Gespräche stattfinden (etwa Bildung, Verwaltung, Medizin, Vereinsleben bei den institutionellen Gesprächen; Politik, Wissenschaft, Wirtschaft und Unterhaltung bei den öffentlichen – massenmedialen – mündlichen Ereignissen) sowie die Arten von *Aktivitäten*, die evtl. während der Gespräche ausgeführt werden (z.B. Renovieren, Kochen etc. im privaten Bereich; Fahrtschulstunde, Meeting etc. im institutionellen Bereich; Mediation oder Panel-Diskussion in der öffentlichen Interaktionsdomäne; vgl. Kaiser 2018: 543). Bei den sekundären Parametern wird nach Geschlecht, Alter und Bildung der Interaktanten differenziert sowie nach dem Ort der Sprachaufnahme, der als Indikator für regionale Spracheinflüsse der Sprechenden angesehen wird⁹.

Das Spezialkorpus *GeWiss* hingegen beschränkt sich auf einen konkreten Ausschnitt aus der Hochschulkommunikation: Vorträge und Prüfungsgespräche in ausgewählten philologisch-angewandten Fächern in vier verschiedenen Sprachen (Deutsch, Englisch, Polnisch, Italienisch) und an

⁸ Hierunter fallen vor allem Interviews und Maptask-Interaktionen, vgl. Kaiser (2018: 421).

⁹ Die Aufnahmeregion ist natürlich kein verlässlicher Indikator für die Art der regionalen Sprachprägung der Sprechenden; mittelfristig strebt *FOLK* an, stattdessen die „prägendste Aufenthaltsregion“ der Sprechenden als aussagekräftigeren Indikator zu nutzen, dies ist allerdings noch nicht umsetzbar, vgl. Kaiser (2018: 540-542).

verschiedenen Standorten (in Deutschland, Großbritannien, Polen, Bulgarien, Italien und Finnland). Neben der Erforschung von sprach- und gattungsvergleichenden Fragen dient *GeWiss* auch der Untersuchung von unterschiedlichen Stufen der wissenschaftlichen Sozialisation und kommunikativen Kompetenz. So wurden sowohl Vorträge von Expert:innen als auch von Studierenden erhoben, daneben auch gezielt ähnlich viele kommunikative Ereignisse mit Sprechenden, die Deutsch als L1, sowie solchen, die Deutsch als L2 sprechen¹⁰ (vgl. ausführlicher Fandrych / Meißner / Slavcheva 2012; Fandrych / Wallner 2023).

In *ZuMal* wurden nun auf der Basis der vielfältigen erhobenen Metadaten bestimmte, für die Recherche aus der Perspektive der (Fremd-)Sprachenvermittlung besonders relevante Auswahlfunktionen geschaffen, die es erlauben, gezielt nach unterschiedlichen Typen von Interaktionen zu suchen und diese für die weitere Arbeit auszuwählen. Wir stellen die metadatenbasierten Filterfunktionen im Folgenden kurz anhand der beiden Korpora *FOLK* und *GeWiss* dar und illustrieren dies mit einigen Beispielen. Wenn man *ZuMal* öffnet, findet man auf der linken Seite zunächst sechs zentrale Auswahloptionen:

The screenshot shows the 'ZuMal' interface with the title 'Auswahl von Daten aus FOLK und GWSS'. At the top left, there is a dropdown menu labeled 'Korpus:' with 'FOLK' selected. To the right of this menu is a blue 'Reset all' button. Below the menu are six filter categories, each in a light blue box with a question mark icon on the right: 'Gesprächstyp', 'Art', 'Themen', 'Sprachregion', and 'Dauer'. The 'Sprachregion' box is highlighted in yellow, and the 'Dauer' box is highlighted in light green.

Abbildung 2
Auswahloptionen nach Metadaten in *ZuMal*

Im Fenster oben links kann man zwischen den beiden Korpora *FOLK* und *GeWiss* wählen. Wählt man *FOLK* (wie in Abbildung 2), erhält man die Kategorien „Gesprächstyp“, „Art“, „Themen“ sowie „Sprachregion“ und „Dauer“. Hinter dem Reiter „Gesprächstyp“ verbergen sich die oben bereits genannten vier *Interaktionsdomänen* sowie die ihnen zugeordneten *Lebensbereiche*, aus denen man nun auswählen kann:

¹⁰ Daneben gibt es einige Sprechereignisse aus dem britischen Kontext mit Sprechenden, für die Englisch die L2 darstellt – diese Ausprägung wurde nicht systematisch erhoben.

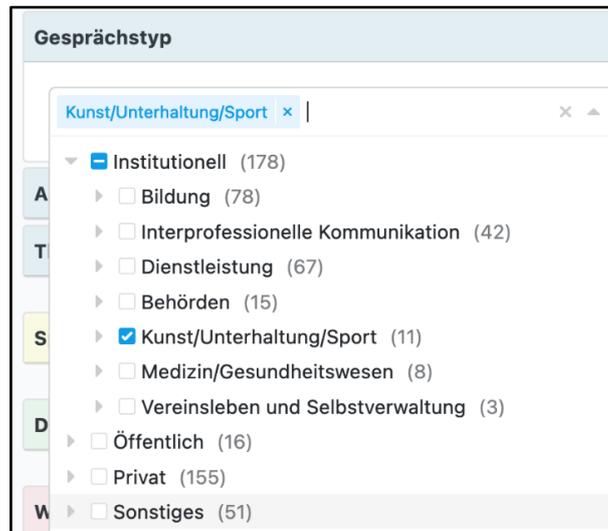


Abbildung 3
Auswahl der Gesprächstypen bei FOLK

In Abbildung 3 wurde aus der *Interaktionsdomäne* „institutionell“ der *Lebensbereich* „Kunst/Unterhaltung/Sport“ ausgewählt – in Klammern sieht man hier bereits, wie viele Interaktionen dem jeweiligen Bereich zugeordnet sind. Ein Klick auf den Pfeil links neben „Kunst/Unterhaltung/Sport“ öffnet weitere Unterkategorien (siehe das Szenario für die Auswahl eines institutionellen Gesprächs unten). Sehr vielfältig sind die Auswahloptionen, die sich hinter dem Reiter „Art“ verbergen:

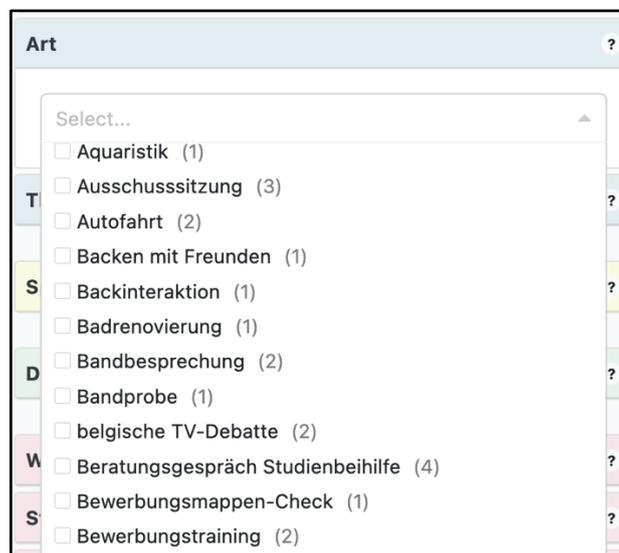


Abbildung 4
Auswahl der Gesprächsart bei FOLK

Wie aus dem Screenshot deutlich wird, sind die verschiedenen Interaktionen nach Kurzcharakterisierungen gegliedert, die wichtige weitere Merkmale benennen (und teils auf primären bzw. sekundären Parametern basieren): „Backen mit Freunden“ gibt einen Hinweis auf ein privates Gespräch mit vertrauten Personen, das eine nicht-verbale Aktivität begleitet, „Beratungsgespräch Studienbeihilfe“ auf ein institutionelles Gespräch im Bildungskontext. Mit der Aufstellung ist keine gattungstheoretisch fundierte Kategorisierung verbunden (diese wäre ohnehin kaum zu erreichen, vgl. die Diskussion

dazu bei Deppermann / Hartung 2012: 427-430), vielmehr stand bei der Erhebung der Interaktionen im Vordergrund, ein möglichst breites Spektrum von vielfältigen Interaktionen zu erheben; die im Auswahlfilter „Gesprächsart“ genutzten Kurzbezeichnungen haben somit den Charakter einer Hilfestellung bei der Suche und Orientierung zu den aufgenommenen Interaktionen.

Die vierte Auswahloption, „Themen“, eröffnet eine sehr große Palette von Themenbegriffen; diese wurden bei der Erhebung und Transkription der Daten in nicht systematisierter Form den Gesprächen zugeordnet (wobei einzelne Gespräche auch mehrere Themen enthalten können). Die verschiedenen Auswahlen beeinflussen sich gegenseitig, insofern ist der Einstieg in die Suche prinzipiell beliebig – mit einer Ausnahme: Ein Einstieg über den Auswahlfilter „Themen“ ist nicht möglich, da die Zahl der Themenschlagwörter hierfür zu groß und daher unübersichtlich zu handhaben ist. Für die Themenauswahl muss daher zunächst eine Vorauswahl bei den Filtern „Gesprächstyp“ oder „Art“ vorgenommen werden, um die Ergebnismenge zu reduzieren.

Unter „Sprachregion“ lässt sich die Region des Aufnahmeorts einer Interaktion auswählen (s.o.). Hierfür wurden im *FOLK*-Projekt die verschiedenen Dialekt- und Regionalsprachenräume in Deutschland auf der Basis dialektologischer Forschungsergebnisse in sechs Großregionen unterteilt (vgl. Kaiser 2018: 540-542): Nordwest, Nordost, Mittelwest, Mittelost, Südwest, Südost. Da das Korpus auch Telefongespräche enthält, kommt es auch zu Kombinationen der Regionen, daneben gibt es einige wenige Interaktionen, die außerhalb des zusammenhängenden deutschen Sprachraums aufgenommen wurden (z.B. auf Urlaubsreisen). *FOLK* versteht sich nicht als Korpus, das die dialektale Vielfalt des deutschen Sprachraums abbildet, es sollen vielmehr „*standardnahe bis gemäßigt dialektale Varianten* des gesprochenen Deutsch“ erhoben werden (Deppermann / Hartung 2012: 421, Hervorhebung im Original), dennoch finden sich Interaktionen, die relativ deutlich dialektal gefärbt sind (vgl. Kap. 5).

Zuletzt ermöglicht es der Filter „Dauer“, mithilfe eines Schiebereglers eine Auswahl nach der Länge der Interaktionen zu treffen:

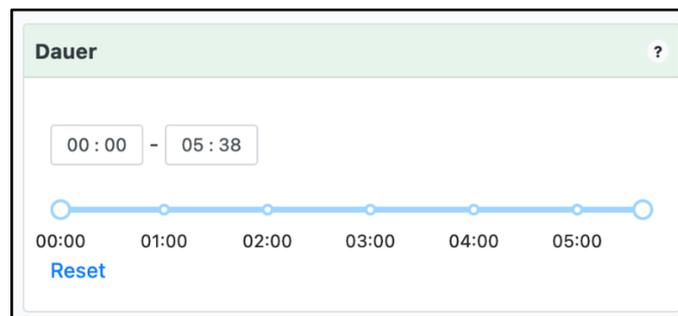


Abbildung 5
Auswahl der Gesprächsdauer

Zu allen Filtern gibt es Kurzerklärungen, die über das jeweilige Fragezeichen oben rechts im Hauptfenster der Auswahloption aufrufbar sind. Direkt zugeordnet ist dort auch jeweils ein Link, der zu einer „Handreichung zur Arbeit mit ZuMal“ führt¹¹.

Filtert man über die genannten Tools nach bestimmten Kriterien, wird gleichzeitig im rechten Bereich überblickartig visualisiert, welche Interaktionen in der Auswahl verbleiben (Kap. 2): In der rechten oberen Hälfte wird dies über ein Streudiagramm angezeigt, das nach verschiedenen sprachdidaktischen und schwierigkeitsbezogenen Parametern eingestellt werden kann (siehe dazu genauer Kap. 4), in der rechten unteren Hälfte durch eine Liste der verbleibenden Interaktionsereignisse, die ebenfalls modifiziert werden kann (siehe Kap. 2).

¹¹ Auch direkt abrufbar unter <https://zumult.ids-mannheim.de/ProtoZumult/doc/Handreichung-ZuMal.html> (15.02.2023).

Anhand von zwei möglichen Szenarien soll nun kurz gezeigt werden, wie man für einen bestimmten Sprachvermittlungskontext die besprochenen Filter einsetzen kann. Sucht man eine informelle private Interaktion, wählt man bei Gesprächstyp „privat“. Eine weitere Spezifizierung ist in diesem Filter nicht vorgesehen. Diese kann allerdings über die Auswahl „Art“ vorgenommen werden: Hier wird eine ganze Reihe von privaten Gesprächen gelistet. Als einen typisch informellen Gesprächsanlass könnte man etwa „Kaffeetrinken“ wählen. Klickt man dies an, verbleiben zwei Interaktionen (was man in der *ZuMal*-Ansicht auch in der Liste der Interaktionen auf der rechten Seite sieht). Der Auswahl-Filter „Themen“ zeigt, dass die Liste der mit den beiden Kaffeetrinken-Interaktionen verbundenen Themen verhältnismäßig groß ist, was bei einer solchen Interaktion nicht verwundert. Die Auswahloption „Sprachregion“ zeigt, dass beide Gespräche im Raum „Nordost“ aufgenommen wurden. Die Dauer ist verhältnismäßig lang (knapp eine Stunde bei einem, eineinhalb Stunden bei dem anderen Gespräch).

Als zweites Beispiel soll die Wahl einer institutionellen Interaktion dienen. Hier kann man nach der Auswahl „institutionell“ beim Filter „Gesprächstyp“ weiter differenzieren, indem man etwa „Bildung“ auswählt. Dies führt zu 78 verbleibenden Interaktionen (vgl. Abbildung 6):

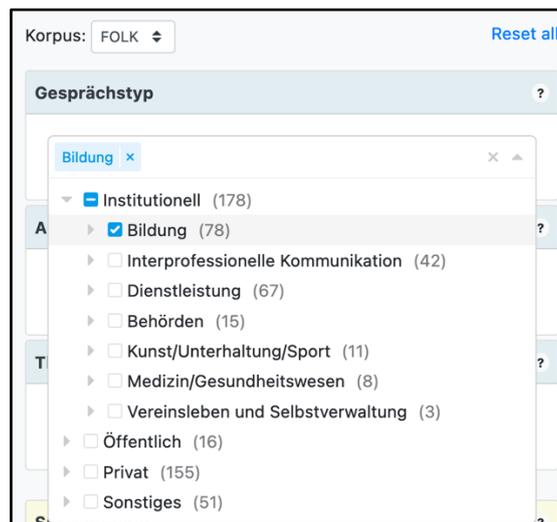


Abbildung 6
Auswahl von „Bildung“ innerhalb des Gesprächstyps „institutionell“

Durch das Klicken auf den Pfeil links neben „Bildung“ erhält man ein weiteres Auswahlmenü, das die verschiedenen Unterkategorien von Bildungsinteraktionen in *FOLK* auflistet. Interessiert man sich für Feedback-Gespräche, markiert man diese Option und erhält eine Auswahl von drei Interaktionen:

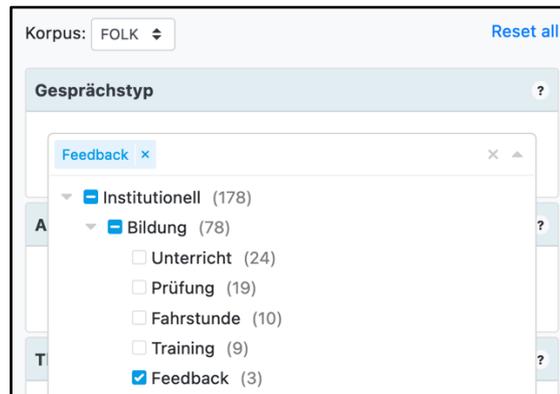


Abbildung 7
Auswahl von Feedback-Gesprächen als Unterkategorie von „Bildung“

Über den Filter „Art“ kann man sodann sehen, dass die drei mündlichen Feedback-Interaktionen aus einem Feedbackgespräch unter Lehrkräften und zwei Unterrichtshospitationen bestehen. Der Sprachregions-Filter zeigt zudem an, dass das Gespräch unter Lehrkräften in der Region „Mittelwest“, die Unterrichtshospitationsgespräche im Raum „Nordwest“ stattfanden. Auch bezüglich der Dauer unterscheiden sich die beiden Unterrichtshospitationsgespräche vom Lehrkräftegespräch, was über den Filter „Dauer“, aber auch über die Liste der ausgewählten Interaktionen in der Übersicht rechts deutlich wird.

Je nach Interesse kann die Suche nach mündlichen Sprachereignissen aber auch mit der Auswahl einer Sprachregion und / oder eines gewünschten Zeitrahmens („Dauer“) beginnen und dann sukzessive verfeinert werden, auch in Kombination mit den Parametern für Schwierigkeit und Mündlichkeitsphänomene (Kap. 4 und 5).

Bei *GeWiss* ist der Interaktionstyp, wie bereits geschildert, vorab festgelegt. Anders als bei *FOLK* allerdings sind verschiedene Sprachen im Korpus enthalten, weshalb statt des dort angebotenen Auswahlmenüs „Gesprächstyp“ der Filter „Sprachen“ vorgesehen ist. Hier kann man nicht nur nach der verwendeten Sprache, sondern auch nach dem Status als L1 oder L2 filtern. In einer Reihe von Interaktionen partizipieren sowohl L1- als auch L2-Sprechende des Deutschen bzw. Englischen – auch dies wird bei der Auswahl separat angezeigt. Die Gesprächsart besteht im Wesentlichen aus den drei genannten Gattungen Expertenvortrag, studentischer Vortrag und Prüfungsgespräch¹². Anders als bei *FOLK* wurden die ausgewählten Gattungen von Beginn an projektintern in gemeinsamen Workshops möglichst konzise definiert und es wurde während des Korpusaufbaus auf Vergleichbarkeit und Konsistenz der Gattungen im Gesamtkorpus geachtet¹³. Auch im *GeWiss*-Korpus ist die Liste der behandelten Themen sehr groß, sodass die Themenauswahl erst erfolgen kann, wenn die Zahl der Interaktionen bereits durch das Setzen anderer Filter eingeschränkt wurde. Auch hier besteht die Möglichkeit, nach der Dauer der Interaktionen weiter zu filtern; das Auswahlmenü „Sprachregion“ zeigt beim *GeWiss*-Korpus allerdings nur die Aufnahmeländer als Ganze an – eine weitere regionale Zuordnung erfolgt an dieser Stelle nicht.

¹² Im finnischen Kontext wurden einige Diskussionen, die im Anschluss an studentische Vorträge geführt wurden, separat aufgezeichnet; diese werden unter „Gesprächsart“ separat ausgewiesen.

¹³ Aufgrund der Vielzahl an Faktoren, die die konkrete Ausprägung der in *GeWiss* enthaltenen drei Gattungen beeinflussen, sind diese dennoch als relativ breite Kategorien zu verstehen – Faktoren, wie die genaue institutionelle und hochschuldidaktische Einbettung von studentischen Vorträgen, die Zahl der Zuhörenden und der situationelle Rahmen von Expertenvorträgen, der Gegenstand, Status und Ablauf von mündlichen Prüfungen können sich im Einzelfall unterscheiden. Hierzu geben zwar die zu den jeweiligen Ereignissen erhobenen Metadaten näheren Aufschluss, diese sind aber nicht systematisch über Auswahlfilter durchsuchbar, vgl. auch Fandrych / Wallner (2022).

Interessiert man sich etwa für deutschsprachige Prüfungsgespräche im universitären Kontext, kann man sich nach der Wahl der Sprache (z. B. „Deutsch als L1“) und der Gesprächsart („Prüfungsgespräch“) über das Auswahlmenü „Themen“ anzeigen lassen, welche Inhalte in den verbleibenden 18 Interaktionen behandelt wurden. Wählt man als Thema „Standard und Variation“, verbleiben drei Prüfungsgespräche, die alle eine Dauer von ca. einer Stunde aufweisen. Es muss dabei berücksichtigt werden, dass das gewählte Thema hier – und in vielen weiteren Fällen – nur *eines* von mehreren behandelten Themen ist und man die genaueren Passagen innerhalb des Gesprächs erst mithilfe einer Durchsicht der Transkripte oder über eine Suche nach bestimmten Schlüsselwörtern identifizieren kann.

Für Sprachvermittlungszwecke ist es nun von besonders großer Relevanz, dass man die in diesem Abschnitt beschriebenen metadatenbasierten Auswahloptionen mit verschiedenen weiteren Filterfunktionen kombinieren kann, die sich auf Aspekte der sprachlichen Schwierigkeit und auf bestimmte Mündlichkeitsphänomene beziehen. Diese werden in den folgenden beiden Abschnitten näher dargelegt.

4. Filter zur Auswahl nach schwierigkeitsbezogenen Parametern

Im Folgenden werden die schwierigkeitsbezogenen Filter vorgestellt. Diese umfassen Auswahlfilter zum in den Sprechereignissen vorhandenen Wortschatz, zur Standardnähe (bzw. der Normalisierungsrate), zur Sprechgeschwindigkeit (bzw. zur Artikulationsrate) der am Gespräch beteiligten Sprechenden und zu Überlappungen, also gleichzeitig gesprochenen Sequenzen dieser Sprechenden.

4.1 Zur Operationalisierung von Schwierigkeit im Rahmen des Projektes *ZuMult*

Die ‚Schwierigkeit‘ von Sprechereignissen ist ein schwer fassbares Konstrukt und von vielen verschiedenen Faktoren abhängig. Grundsätzlich sind unterschiedliche Operationalisierungen von Schwierigkeit denkbar¹⁴. Um eine Übertragbarkeit auf künftige Korpuserweiterungen zu gewährleisten, war es im Rahmen des Projekts *ZuMult* erforderlich, in den Korpusdaten Merkmale zu identifizieren, die mit Schwierigkeitskonzepten in Verbindung stehen und sich zugleich mit Hilfe automatisierter Verfahren ermitteln lassen.

Schwierigkeit wurde bei der Konzeption der Auswahlfilter vor allem als lernerseitige Schwierigkeit bei der Rezeption von Sprechereignissen konzeptualisiert. Die an dieser Stelle vorgestellten Auswahlfilter zielen daher darauf, die Rezeption der Sprechereignisse für die Lernenden entweder zu erleichtern oder anspruchsvoller zu gestalten. So könnte eine Erleichterung der Rezeption bspw. erfolgen, indem Sprechereignisse ausgewählt werden, die über einen potenziell weitgehend bekannten Wortschatz oder eine größere Standardnähe verfügen. Eine anspruchsvollere Gestaltung der Rezeption könnte wiederum vorgenommen werden, indem sehr dialogische Sprechereignisse ausgewählt werden, bei denen sich Redesequenzen oft überschneiden und zusätzlich vielleicht eine Vielzahl an Mündlichkeitsphänomenen enthalten sind. Zugrunde gelegt wird hierbei die Beobachtung, dass Merkmale der gesprochenen Sprache, wie die Sprechgeschwindigkeit oder Überlappungen verschiedener Sprechender Personen, für Lernende mit höherem kognitiven Verarbeitungsaufwand bei der Dekodierung einhergehen (vgl. Housen et al. 2019: 4). Sie sind demnach für die Lernenden unter

¹⁴ So könnte die Schwierigkeit von Sprechereignissen etwa mit Hilfe von Expert:innen oder auch durch experimentelle Erhebungen mit verschiedenen Nutzer:innen / Lernenden bestimmt werden und im Rahmen eines mehrstufigen Ratings annotiert werden. Ein solches Vorgehen ist jedoch mit einem erheblichen personellen Aufwand verbunden und daher für stetig wachsende Korpusressourcen nicht praktikabel.

Umständen schwieriger zu erfassen. Schwierigkeit (*difficulty*) wird bei Housen et al. (2019) als eine Dimension von Komplexität (*complexity*) verstanden. In diesem Kontext wird zudem immer wieder die Unterscheidung zwischen relativer und absoluter Komplexität herausgestellt, wobei ‚Schwierigkeit‘ der relativen Komplexität zuzuordnen ist und auch als nutzerbezogene Komplexität bezeichnet wird (vgl. Housen et al. 2019: 4)¹⁵.

Schwierigkeit ist in diesem Zusammenhang immer nur individuell bestimmbar, da sie sich aus der Beziehung der Sprachnutzenden zu den jeweiligen Texten bzw. Sprechereignissen ergibt (vgl. Bulté / Housen 2012: 23). Bei der kognitiven Verarbeitbarkeit der Sprechereignisse spielen hinsichtlich der Schwierigkeit zusätzlich Einflussfaktoren wie die generelle Sprachlerneignung, das Arbeitsgedächtnis, der erstsprachliche Hintergrund oder auch das Sprachniveau in der Fremd- bzw. Zweitsprache Deutsch eine entscheidende Rolle (vgl. Housen / Simoens 2016: 167; Housen et al. 2019: 10). Vor diesem Hintergrund sind wir bei der Konzeption der Filter für *ZuMal* davon ausgegangen, dass diese einerseits von Lehrkräften dazu verwendet werden können, um möglichst passgenaue Sprechereignisse für spezifische Unterrichtskontexte auszuwählen und andererseits von Lernenden dazu genutzt werden können, um Sprechereignisse auszuwählen, die sowohl den eigenen Interessen als auch dem jeweiligen Lernstand entsprechen und so eine gewinnbringende und lernförderliche Auseinandersetzung mit den Transkripten ermöglichen.

Der Auswahlfilter „Wortschatz (Niveaustufe)“

Der erste der im Projekt entwickelten schwierigkeitsbezogenen Auswahlfilter ist der Filter „Wortschatz (Niveaustufe)“. Dieser Filter operationalisiert die Textdeckung eines Sprechereignisses mit Hilfe unterschiedlicher Wortschatzlisten. Die Textdeckung eines Textes bezieht sich auf die Anzahl an Wörtern in einem Text oder einem Sprechereignis, die Lernenden mit Deutsch als Fremd- oder Zweitsprache, aber auch erstsprachlichen Lesenden oder Hörenden, mindestens bekannt sein müssen, damit ein Text angemessen verstanden werden kann (vgl. Vilkaitė-Lozdienė / Schmitt 2020: 80)¹⁶. In *ZuMal* ist die Operationalisierung der Textdeckung bzw. des Wortschatzfilters einerseits eng an den Wortschatz einzelner GER-Niveaustufen angelehnt, da diese auf der Grundlage von verschiedenen Wortschatzlisten des Goethe-Instituts berechnet werden kann. Zu diesem Zweck stehen die A1¹⁷, A2¹⁸- und die B1-Wortschatzliste¹⁹ zur Verfügung. Andererseits lässt sich die Textdeckung frequenzbezogen anhand der ‚Herder-Listen‘ ausgeben, die auf Grundlage des Frequenzwörterbuchs von Tschirner / Möhring (2019) berechnet wurden. Diese ermöglichen einen Abgleich mit den 1000, 2000, 3000, 4000 und 5000 häufigsten Wörtern des Deutschen. Die Frequenz eines Wortes wird dabei als Kriterium für die Relevanz eines Wortes für die Lernenden gesehen. Gleichzeitig wird ein indirekter Zusammenhang zwischen dem für die GER-Niveaustufen konzipierten Wortschatz und den jeweiligen Wortschatzlisten und den frequenzbezogenen Wortschatzlisten angenommen. Dieser wird in der Regel über die Wortschatzbreite operationalisiert. Tschirner / Hacking / Rubio (2018) arbeiten etwa für die Niveaustufe A1 eine Wortschatzbreite von 837 Wörtern, für A2 eine Wortschatzbreite von 1640 und für B1 eine Wortschatzbreite von 3246 Wörtern heraus (vgl. Tschirner / Hacking / Rubio 2018: 70). Dies legt für die Filter in *ZuMal* also vor allem Zusammenhänge zwischen den Filtern für

¹⁵ Absolute Komplexität bezieht sich im Gegensatz dazu auf der Sprache selbst inhärente Konstruktionen und Strukturen (vgl. für ausführliche Diskussionen zur absoluten Komplexität und zur oft operationalisierten linguistischen Komplexität Housen et al. 2019: 5; Pallotti 2015).

¹⁶ Dieser Zusammenhang zwischen Textdeckung und Textverständnis wird in der Threshold-Hypothese zusammengefasst (vgl. Milton 2009). Dieser Hypothese liegt der Befund zugrunde, dass eine höhere Textdeckung auch mit größerem Textverstehen korreliert, dass aber gleichzeitig unterschiedliche Schwellen erreicht werden können bzw. müssen, um unterschiedliche Niveaus von Textverständnis zu ermöglichen.

¹⁷ Vgl. https://www.goethe.de/pro/relaunch/prf/de/A1_SD1_Wortliste_02.pdf (15.02.2023).

¹⁸ Vgl. https://www.goethe.de/pro/relaunch/prf/de/Goethe-Zertifikat_A2_Wortliste.pdf (15.02.2023).

¹⁹ Vgl. https://www.goethe.de/pro/relaunch/prf/de/Goethe-Zertifikat_B1_Wortliste.pdf (15.02.2023).

die Goethe-Wortschätze A1, A2 und B1 und den Listen der 3000 häufigsten Wörter nahe, wobei für einen Text mit B1-Wortschatz die Liste der 3000 häufigsten Wörter eine angemessene Schwelle darzustellen scheint.

Über die Filterfunktion kann für diese acht Listen die gewünschte Textdeckung ausgewählt werden. Mit Hilfe der Lemmata wird dann ein Abgleich der im Sprechereignis enthaltenen Wörter mit der jeweils ausgewählten Wortschatzliste durchgeführt. Die zur Verfügung stehenden Wortschatzlisten ermöglichen so eine wortschatzbezogene Einschätzung des ausgewählten Sprechereignisses. Ob allerdings eher die Wortschatzlisten des Goethe-Instituts oder die frequenzbezogenen Wortschatzlisten zur Einschätzung der lernerseitigen Schwierigkeit eines Sprechereignisses herangezogen werden sollten, hängt entscheidend davon ab, wie die von den Lernenden besuchten Kurse, die verwendeten Lehr- und Lernmaterialien und die zugrundeliegenden Curricula gestaltet sind.

Für das Lesen hat sich in den letzten Jahrzehnten eine Textdeckung von zwischen 95 % und 98 % als Schwellenwert etabliert, der ein weitgehendes Textverständnis sicherstellt (vgl. Hacking / Tschirner 2017: 503; Tschirner / Hacking / Rubio 2018: 60). Prinzipiell kann jedoch davon ausgegangen werden, dass für gesprochene Sprache eine niedrigere Wortschatzdeckung zum Verstehen eines kommunikativen Ereignisses notwendig ist als in geschriebenen Texten (vgl. Read 2004: 149). Van Zeeland / Schmitt (2013) gehen für das Englische etwa davon aus, dass 90-95 % Wortschatzdeckung für das Verständnis eines Hörtextes ausreichen, gleichzeitig gehen sie davon aus, dass für eine solche Deckung in Sprechereignissen der Alltagssprache etwa 2000-3000 Wörter bekannt sein müssten. Es muss allerdings darauf hingewiesen werden, dass die Schätzungen darüber, wie viele Wörter zum Erreichen der Schwellenwerte bekannt sein müssen, stark differieren und zum Teil deutlich höhere Werte angenommen werden. In diesem Zusammenhang ist es zudem wichtig zu erwähnen, dass Studien zur Wortschatzdeckung in mündlichen Sprechereignissen für das Deutsche noch weitgehend ausstehen.

Der Auswahlfilter „Standardnähe (Normalisierungsrate)“

Der Auswahlfilter „Standardnähe (Normalisierungsrate)“ zielt, wie die Filter „Sprechgeschwindigkeit“ und „Überlappungen“, in besonderer Weise auf grundlegende Eigenschaften gesprochener Sprache ab, die einen Einfluss auf die relative Komplexität bzw. auf die lernerseitig wahrgenommene Schwierigkeit eines Sprechereignisses haben können²⁰. Gesprochene Sprache unterscheidet sich in ihrer Realisierung in vielerlei Hinsicht von geschriebener Sprache. Dies zeigt sich etwa in *FOLK* vor allem durch dialektal gesprochene Formen, aber auch durch typische Phänomene der Mündlichkeit wie Häsitationen, Korrekturen oder Klitisierungen. Im *GeWiss*-Korpus kommen zu diesen Phänomenen potenziell noch lernersprachliche Strukturen hinzu, die durch eine grundsätzliche Variation (vgl. Fandrych / Wallner 2022; Wisniewski / Lüdeling / Czinglar 2022) gekennzeichnet sind. Der Filter „Standardnähe (Normalisierungsrate)“ wird operationalisiert, indem die Anzahl der Token berechnet wird, die aufgrund der aussprachenahen Transkription der Gespräche von der orthografischen Standardschreibung differieren. Als Grundlage der aussprachenahen schriftlichen Wiedergabe der mündlichen Daten dient dabei die Minimaltranskript-Version von cGAT2 (vgl. Schmidt / Schütte / Winterscheid 2015); diese wurde in der Folge auf einer weiteren Transkriptspur in eine orthografisch bereinigte („normalisierte“) Version übertragen (vgl. zur Normalisierung von gesprochenen Daten Winterscheid et al. 2019). Eine hohe Normalisierungsrate zeigt daher an, dass viele Token zwischen diesen beiden Spuren differieren und es sich damit um ein Sprechereignis handelt, das als „standardferner“ beschrieben werden kann.

Diesem Filter liegt – der Logik dieser Operationalisierung folgend – die Annahme zugrunde, dass Transkripte, die durch eine größere „Standardnähe“ (und damit eine geringere Normalisierungsrate) gekennzeichnet sind, grundsätzlich zugänglicher für Lernende des Deutschen als Zweit- und

²⁰ Der Filter „Standardnähe (Normalisierungsrate)“ steht in *ZuMal* nur für deutschsprachige Daten zur Verfügung.

Fremdsprache sind, da bis heute in unterrichtlichen Kontexten nach wie vor eher am schriftsprachlichen Standard orientierte Strukturen des Deutschen vermittelt werden. Zu berücksichtigen ist, dass die Normalisierungsrate ein Wert ist, der sich auf das gesamte Sprechereignis bezieht und die Werte innerhalb eines Sprechereignisses sowie bei einem Vergleich der Sprecher:innen untereinander durchaus variieren können.

Der Auswahlfilter „Sprechgeschwindigkeit (Artikulationsrate)“

Dem Filter „Sprechgeschwindigkeit (Artikulationsrate)“ liegt die Überlegung zugrunde, dass Sprechereignisse für Lernende mit einem höheren Dekodierungsaufwand einhergehen, wenn in diesen schneller gesprochen wird. Dieser Filter unterstützt Lehrende und Lernende demnach dabei, schneller und langsamer gesprochene Sprechereignisse für didaktische Zwecke auszuwählen. Die Sprechgeschwindigkeit wird dabei über die Rate der pro Sekunde artikulierten Silben operationalisiert. Zur Berechnung der Artikulationsrate werden die aussprachenahen Transkriptionen in die phonemische Umschrift SAMPA (vgl. Wells 1997) transponiert. Mithilfe der in dieser Umschrift markierten Silbengrenzen lässt sich für jedes Sprechereignis die Gesamtzahl an Silben berechnen, die dann durch die Gesamtzahl an Sekunden dividiert wird. Eine solche Operationalisierung bildet folglich nicht ab, ob es innerhalb eines Gespräches auch Phasen gibt, in denen deutlich langsamer gesprochen wird. Dies lässt sich allerdings über den Density Viewer in *ZuViel* nachvollziehen (vgl. Schmidt / Schwendemann / Wallner in dieser Ausgabe).

Der Auswahlfilter „Überlappungen“

Ähnlich wie der Filter „Sprechgeschwindigkeit (Artikulationsrate)“ wurde der Filter Überlappungen auf Basis der Annahme konzipiert, dass Gesprächssequenzen, in denen mehrere Personen gleichzeitig, d.h. überlappend, sprechen, von Lernenden eine größere Rezeptionsleistung verlangen. Die Auszählung überlappend gesprochener Sequenzen erfolgt direkt anhand der aussprachenahen Transkription, in der diese Sequenzen festgehalten werden. Wie für den Filter „Sprechgeschwindigkeit (Artikulationsrate)“ werden jeweils globale Werte für ein Sprechereignis angegeben. Das bedeutet, dass die Auswahl eines Sprechereignisses mit einer hohen Rate an Überlappungen nicht bedeutet, dass es in diesem Sprechereignis nicht auch Sequenzen gibt, in denen es kaum zu Überlappungen kommt.

Für den Kontext DaF/DaZ ist an dieser Stelle grundsätzlich zu konstatieren, dass die systematische Analyse von konkreten Faktoren, die lernerseitig zu größeren Schwierigkeiten bei der Rezeption von Sprechereignissen führen könnten, nur selten in den Forschungsfokus rückt, auch weil diese Faktoren aufgrund ihrer gegenseitigen Abhängigkeiten und dynamischen Beziehungen schwierig zu operationalisieren sind. Hier bestehen zudem noch dringende forschungsbezogene Desiderata²¹. Die vorgestellten Filter aus *ZuMal* bilden dennoch einen Versuch, solche schwierigkeitsbezogenen Faktoren beim Einsatz von authentischen Sprechereignissen systematisch in Unterrichtskontexte einzubringen und mitzudenken.

4.2 Die Filter in Anwendung: Spektren und Beispiele

Im Folgenden wird gezeigt, zu welchen Ergebnissen die Anwendung der schwierigkeitsbezogenen Filter bezogen auf die aktuelle Version 2.19 der DGD für das *FOLK*-Korpus führt. Neben den derzeit beobachtbaren Spektren der Merkmalsausprägungen wird auch jeweils angegeben, welche Werte die

²¹ Gerade für den Zusammenhang zwischen linguistischen Komplexitätsmaßen, die die Textkomplexität als Ganze beeinflussen, und Lesekompetenzen wurden in jüngster Zeit immer wieder Zusammenhänge herausgearbeitet. Forschungen, die solche Zusammenhänge hinsichtlich multimedialer Hör-Seh-Texte bzw. der Sprechereignisse in Korpora der gesprochenen Sprache wie *FOLK* oder *GeWiss* systematisch untersuchen, stehen bis heute jedoch weitgehend aus.

mittleren 50 % aller Sprechereignisse bezüglich der einzelnen Merkmale aufweisen²². Die Darstellung beruht dabei auf Informationen, die der tabellarischen Ergebnisansicht in *ZuMal* entnommen wurden. Für die Charakterisierung und Einordnung der Ergebnisse werden Interaktionsdomänen und Gesprächstypen angegeben. Ergänzend dazu werden zusätzlich auch einige ausgewählte Befunde aufgeführt, die auf Auswertungen mit dem Tool *ZuRecht* (vgl. auch Frick / Helmer / Wallner in dieser Ausgabe) beruhen.

Spektren des Auswahlfilters „Wortschatz (Niveaustufe)“

Für die Nutzung dieses Filters muss zunächst eine der unter 4.1 genannten Wortschatzlisten ausgewählt werden. Nach der Auswahl einer Liste wird unterhalb der Liste das Spektrum der Wortschatzdeckung angezeigt – also der minimale und der maximale Wert der Wortschatzdeckung, den die ausgewählte Liste bei den Sprechereignissen in *FOLK* erreicht. Alternativ kann in der tabellarischen Ergebnisansicht die gewünschte Wortschatzliste als Spalte hinzugefügt werden. Die Sprechereignisse können daraufhin in dieser Spalte nach Wortschatzdeckung sortiert werden. Abbildung 8 zeigt beide Optionen am Beispiel der Goethe-Zertifikatswortschatzliste der Niveaustufe A2:

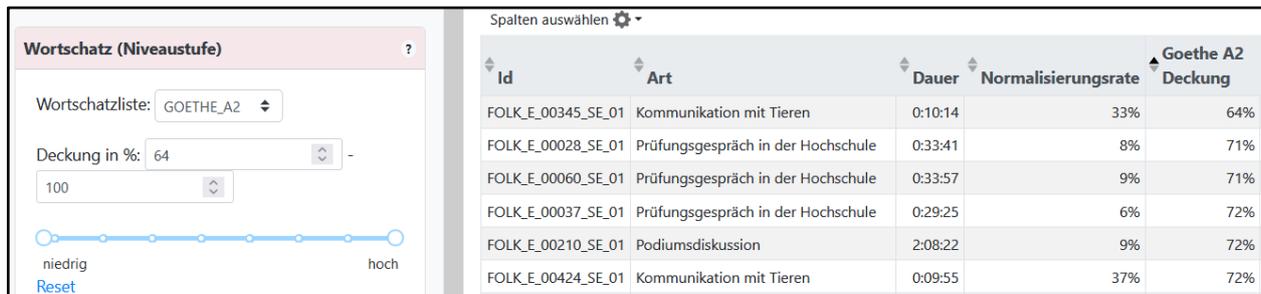


Abbildung 8
Auswahlfilter „Wortschatz (Niveaustufe)“

Bei der Anwendung der verschiedenen Goethe-Zertifikatswortschatzlisten auf die Sprechereignisse in *FOLK* lassen sich folgende Spektren bezüglich der Wortschatzdeckung beobachten: 55 % - 97 % (A1), 64 % - 100 % (A2) und 69 % - 100 % (B1). Dabei zeigen die mittleren 50 % der Daten eine Wortschatzdeckung zwischen 75 % und 79 % (A1), 81 % und 85 % (A2) bzw. zwischen 88 % und 91 % (B1). Bezogen auf die Niveaustufen A1 und A2 liegt damit die Hälfte der Daten unterhalb der – nach Van Zeeland / Schmitt 2013 – für ein ausreichendes Hörverständnis erforderlichen Wortschatzdeckung von 90 %. Soll von vornherein eine Einschränkung der Auswahl der Sprechereignisse – etwa mit einer Wortschatzdeckung von 90 % – vorgenommen werden, so ist dies mit Hilfe des Schiebereglers unterhalb der Anzeige der Deckungswerte (vgl. Abbildung 8) möglich. Aktuell zeigen für das Niveau A1 lediglich drei Datensätze eine Wortschatzdeckung von mindestens 90 %. Dabei handelt es sich um sehr kurze (weniger als eine Minute andauernde) Gespräche im Polizeirevier. Für das Niveau A2 erreichen neun Gespräche eine Wortschatzdeckung von mindestens 90 %. Neben fünf Gesprächen im Polizeirevier handelt es sich hierbei um zwei Tandemtreffen und zwei Verkaufsgespräche. Für das Niveau B1 ist die Auswahl deutlich größer: Hier erreichen 212 Gespräche eine Wortschatzdeckung von mindestens 90 %. Abgesehen von den Gesprächen, die bereits für A2 Deckungswerte über 90 % aufweisen, zeigen eine Maptask-Interaktion, zwei ethnografische Interviews, ein Meeting in einem Wirtschaftsunternehmen sowie ein Gespräch aus der Make-up-Artist-Ausbildung

²² Gemeint sind die Sprechereignisse, die innerhalb des Interquartilsabstands liegen und sich um den Mittelwert gruppieren. Es handelt sich dabei um diejenigen Sprechereignisse, die bezüglich des jeweiligen Merkmals besonders nah beieinanderliegen und damit als besonders typisch angesehen werden können.

besonders hohe Werte von 94 %. Im Hinblick auf die Domänenzugehörigkeit fällt auf, dass die Gespräche mit einer Wortschatzdeckung von mindestens 90 % hauptsächlich der privaten Interaktionsdomäne mit insgesamt 100 Gesprächen und der institutionellen Interaktionsdomäne mit 81 Gesprächen zuordnen lassen.

Bei der Anwendung der häufigkeitsbezogenen Herder-Listen zeigen sich folgende Spektren in der Textdeckung: 77 % - 97 % (Herder 1000), 61 % - 97 % (Herder 2000), 63 % - 100 % (Herder 3000), 64 % - 100 % (Herder 4000) und 65 % - 100 % (Herder 5000). Die mittleren 50 % der Datensätze liegt zwischen 80 % und 84 % (Herder 1000), 83 % und 87 % (Herder 2000), 85 % - 88 % (Herder 3000), 86 % und 90 % (Herder 4000) und 87 % und 90 % (Herder 5000). Die Hälfte der Gesprächsdaten erreicht damit erst für die 4000 häufigsten Wörter eine Wortschatzdeckung von mindestens 90 %. Bezüglich der häufigsten 1000 Wörter zeigt lediglich eines der sehr kurzen Gespräche im Polizeirevier eine Wortschatzdeckung von über 90 %. Bezüglich der 2000 häufigsten Wörter sind es bereits zwölf Gespräche, darunter erneut Gespräche im Polizeirevier, Tandemtreffen sowie Verkaufs- und Telefongespräche. Im Hinblick auf die 3000 häufigsten Wörter erreichen wiederum 43 Gespräche eine Wortschatzdeckung von mindestens 90 %, bezüglich der 4000 häufigsten Wörter sind es schon 109 Gespräche und bezüglich der 5000 häufigsten Wörter sind es 262 Gespräche. Hinsichtlich der häufigsten 5000 Wörter gehören die Gespräche mit einer Wortschatzdeckung von mindestens 90 % überwiegend der privaten Interaktionsdomäne (mit 71 Gesprächen) an. 47 Gespräche lassen sich der institutionellen Interaktionsdomäne zuordnen.

Insgesamt wird deutlich, dass die Gesprächsdaten im *FOLK* vor allem ab der Niveaustufe B1 geeignet zu sein scheinen. Das heißt aber nicht, dass die Daten nicht auf niedrigeren Niveaustufen eingesetzt werden können. Da es sich bei der Wortschatzdeckung um einen globalen Wert handelt, der sich auf das Gesamtgespräch bezieht, ist durchaus denkbar, dass einzelne Sequenzen höhere (ggf. aber auch niedrigere) Deckungswerte aufweisen und damit auch für niedrigere Niveaustufen geeignet sein können. Zudem ist es auch möglich, den nicht durch eine Wortliste gedeckten Wortschatz vorzuentlasten (vgl. hierzu Schmidt / Schwendemann / Wallner in dieser Ausgabe).

Spektren des Auswahlfilters „Standardnähe (Normalisierungsrate)“

Die Spektren des Auswahlfilters „Standardnähe (Normalisierungsrate)“ lassen sich analog zur Wortschatzdeckung über den Filter sowie über die tabellarische Ergebnisansicht anzeigen (vgl. Abbildung 9)²³.



Abbildung 9
Auswahlfilter „Standardnähe (Normalisierungsrate)“

Bezüglich der Normalisierungsrate zeigen die Gesprächsdaten in *FOLK* ein sehr breites Spektrum. Dieses liegt zwischen 3 % und 57 %. Das heißt, dass bei einigen Sprechereignissen lediglich 3 %

²³ Zu beachten ist dabei, dass es sich in der tabellarischen Ansicht um Rundungswerte handelt und sich ggf. dadurch der niedrigste sowie der maximale Wert von den unterhalb der Filter angezeigten Werten geringfügig unterscheiden können.

aller Token abweichend von der Standardorthografie transkribiert wurden und somit deutlich von der Standardlautung abweichen, während dies bei einem Sprechereignis bei 57 % aller Token der Fall ist. Die mittleren 50 % der Daten besitzen eine Normalisierungsrate zwischen 11 % und 19 %. Besonders niedrige Normalisierungsraten lassen sich bei öffentlichen Gesprächsdaten beobachten (so etwa bei einer Plenarsitzung im Bundestag), bei mehreren Ausschusssitzungen sowie bei einer Podiumsdiskussion mit jeweils 3 %. Es ist davon auszugehen, dass es sich hierbei um überwiegend geplante, teilweise auch medial inszenierte Sprechereignisse handelt. Die höchste Normalisierungsrate, die sich in den aktuell verfügbaren Gesprächsdaten aus der öffentlichen Interaktionsdomäne beobachten lässt, liegt bei 12 %. Dabei handelt es sich um zwei öffentliche Schlichtungsgespräche. Besonders hohe Normalisierungsraten besitzen wiederum zwei Fokusgruppen-Interviews mit 57 % bzw. 49 %, in denen jeweils Schweizerdeutsch gesprochen wird. Höhere Normalisierungsraten lassen sich zudem bei Gesprächen aus der Interaktionsdomäne „privat“ beobachten (bspw. ein Telefongespräch mit 48 % und ein Tischgespräch mit 47 %). Aber auch institutionelle Gespräche können hohe Normalisierungsraten aufweisen (so etwa ein Meeting in einer sozialen Einrichtung mit 46 % und ein Gespräch beim Friseur mit 41 %). Vielfach zeigen zudem Gesprächsdaten mit höheren Normalisierungsraten eine stärkere dialektale Prägung (vgl. den Ausschnitt aus FOLK_E_00319_SE_01 mit 41 % Normalisierungsrate in Abbildung 10).

0002	MG	so hh° mogst du was trinka
0003		(0.75)
0004	EZ	i hätt gern ein cappuccino bitte
0005	MG	(.) äh i kann da vui mui noi dua (.) ne cappuccino mach[a mer koin]
0006	EZ	[des bass]t au
0007	MG	roicht des a[u]
0008	EZ	[j]a ja

Abbildung 10
Ausschnitt aus einem Gespräch beim Friseur (FOLK_E_00319_SE_01)

Die Gespräche mit durchschnittlichen Normalisierungsraten und hier insbesondere die mittleren 50 % der Daten, die zwischen 11 % und 19 % liegen, enthalten meist umgangssprachliche Phänomene, die regionenübergreifend vorkommen (vgl. den Ausschnitt aus einem Meeting einer sozialen Einrichtung mit 14 % Normalisierungsrate in Abbildung 11).

0014	SF	((schmatzt)) °h (.) habt ihr ne tagesordnung (.) vorhin schon festgele[gt]
0015	AM	[nee]wir ham no nich gesammelt wir ham erst ma mit der beratung angefangen (.) damit das (.) [erledigt]is un wir dann hier mit dem orgakram schnell weitermachen (.) °hh [was habt ihr denn für themen

Abbildung 11
Ausschnitt aus einem Meeting in einer sozialen Einrichtung (FOLK_E_00254_SE_01)

Grundsätzlich ist allerdings zu berücksichtigen, dass anhand der Normalisierungsrate nicht erkennbar ist, inwieweit es sich um dialektale oder eher umgangssprachliche Phänomene handelt.

Zu den häufigsten übergreifend beobachtbaren Normalisierungsfällen gehören Reduktionen wie bspw. *is*, welches 43.179 Mal zu *ist* normalisiert wurde und in 99 % der Sprechereignisse und bei 78 % der dokumentierten Sprecher:innen zu beobachten ist. Weitere übergreifend beobachtbare Reduktionen sind *nich* (25.195 Mal zu *nicht* normalisiert, beobachtbar in 94 % der Sprechereignisse und bei 86 % der dokumentierten Sprecher:innen) und *s* (19.867 Mal zu *es* normalisiert, beobachtbar in 98 % der Sprechereignisse und bei 96 % der dokumentierten Sprecher:innen). Daneben betreffen die Normalisierungen auch gesprochensprachliche Ersatz- oder auch Schnellsprechformen, wie bspw. *nee*, welches 9.507 Mal zu *nein* normalisiert wurde und in 93 % aller Sprechereignisse und bei 77 % aller dokumentierten Sprecher:innen vorkommt. Weitere Beispiele sind *ham* (8.921 Mal normalisiert zu *haben* und in 93 % der Sprechereignisse sowie bei 79 % der dokumentierten Sprecher:innen beobachtbar) sowie *nix* (2.369 Mal normalisiert zu *nichts* und in 71 % der Sprechereignisse sowie bei 49 % der dokumentierten Sprecher:innen beobachtbar).

Überdies handelt es sich bei den Normalisierungsfällen oft auch um Klitisierungen. Zu den häufigsten übergreifend vorkommenden Klitisierungen im *FOLK* zählt aktuell *son* bzw. *so_n*, welches in 3.006 Fällen zu *so ein* normalisiert wurde und in 78 % der Sprechereignisse sowie bei 54 % der Sprecher:innen beobachtet werden kann. Weitere häufige und verbreitete Klitisierungen sind *gibt_s* (1.942 Mal normalisiert zu *gibt es*, beobachtbar in 71 % der Sprechereignisse und bei 51 % der dokumentierten Sprecher:innen) und *geht_s* (1.134 Mal normalisiert zu *geht es*, beobachtbar in 72 % der Sprechereignisse und bei 42 % der dokumentierten Sprecher:innen).

Spektren des Auswahlfilters „Sprechgeschwindigkeit (Artikulationsrate)“

Auch die Spektren zur Sprechgeschwindigkeit können über den Filter sowie über die tabellarische Ergebnisansicht angezeigt werden (vgl. Abbildung 12).

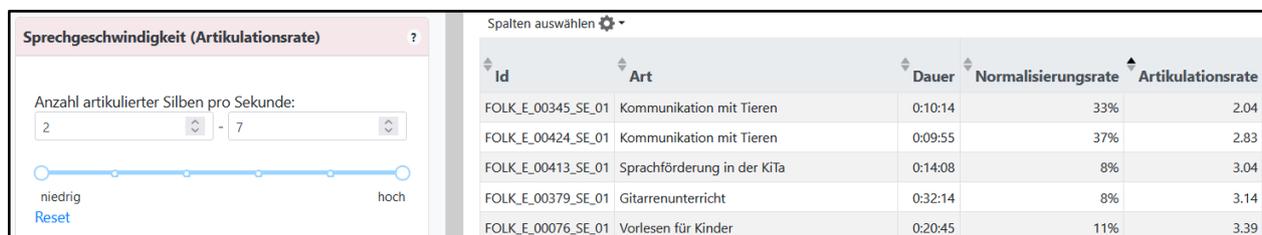


Abbildung 12
Auswahlfilter „Sprechgeschwindigkeit (Artikulationsrate)“

Bei der Betrachtung der Artikulationsrate zeigen die Gesprächsdaten in *FOLK* aktuell ein Spektrum zwischen 2,04 Silben pro Sekunde und 6,48 Silben pro Sekunde. Die mittleren 50 % der Daten liegen zwischen 4,5 und 5,25 Silben pro Sekunde. Zu den Sprechereignissen mit besonders niedrigen Artikulationsraten gehören Kommunikationen mit Tieren (bspw. *FOLK_E_00345_SE_01* mit 2,04 Silben pro Sekunde) und verschiedene Interaktionen mit Kindern (darunter eine Sprachförderung in der KiTa mit 3,04 Silben pro Sekunde, Gitarrenunterricht mit 3,14 Silben pro Sekunde und mehrere Vorlese-Interaktionen mit Kindern (z. B. *FOLK_E_00076_SE_01* mit 3,39 Silben pro Sekunde). Besonders hohe Artikulationsraten zeigen u. a. ein Training in einer Hilfsorganisation mit 6,48 Silben pro Sekunde, ein Feedbackgespräch unter Lehrkräften mit 6,45 Silben pro Sekunde sowie ein Gespräch beim Friseur mit 6,29 Silben pro Sekunde. Bezüglich der Zugehörigkeit zu den Interaktionsdomänen ist keine Tendenz erkennbar.

Insgesamt ist auch bezüglich der Artikulationsrate zu beachten, dass diese erheblichen Schwankungen unterliegt. So wird bei der Betrachtung einzelner Sprecherbeiträge deutlich, dass in fast allen Sprechereignissen und bei nahezu von allen dokumentierten Sprecher:innen sowohl eher

langsame Sprecherbeiträge (mit 2,0 bis 3,0 Silben pro Sekunde) als auch solche mit schnell gesprochene Sequenzen (mit 6,0 oder mehr Silben pro Sekunde) vorkommen. Ein erster explorativer Vergleich macht deutlich, dass es sich bei den langsameren Sprecherbeiträgen in den meisten Fällen um der Verständnissignalisierung dienende Einzelwörter (wie *ja*, *hmhm* und *hm*) handelt, die häufig auch überlappend gebraucht werden. Bei den Sprecherbeiträgen mit hoher Artikulationsrate überwiegen zwar auch die Einzelwörter, doch unterscheidet sich deren Frequenz nur geringfügig von Sequenzen, die mehrere Token umfassen.

Spektren des Auswahlfilters „Überlappungen“

Die Spektren bezüglich der Anzahl der Überlappungen sind ebenfalls über den Filter sowie über die tabellarische Ergebnisansicht aufrufbar (vgl. Abbildung 13).

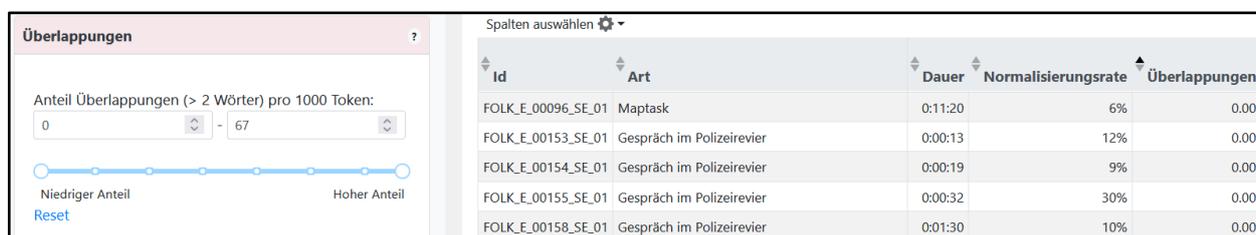


Abbildung 13
Auswahlfilter „Überlappungen“

Ein Blick in die Daten zeigt, dass die Anzahl der Überlappungen in *FOLK* erheblich differiert. So gibt es in einigen der Sprechereignisse keine Überlappungen. Darunter fallen bspw. mehrere Gespräche im Polizeirevier (vgl. auch Abbildung 13), ein Verkaufsgespräch in der Apotheke und ein Verkaufsgespräch im Supermarkt. Nur wenige Überlappungen sind in zwei Ausschusssitzungen, einer Podiumsdiskussion sowie in zwei Interviews zu beobachten. Die höchste Anzahl an Überlappungen zeigt wiederum ein Tischgespräch mit durchschnittlich 66,94 Überlappungen pro 1000 Token. Zudem gehören zu den Sprechereignissen mit besonders vielen Überlappungen zwei Spielinteraktionen, eine Kommunikation beim Kochen, ein WG-Casting, eine Teambesprechung und ein Mädelsabend.

Bei der Mehrheit der Sprechereignisse in *FOLK* sind zwischen 3,46 und 11,98 Überlappungen pro 1000 Token zu beobachten. Hinsichtlich der Interaktionsdomäne zeigt sich jedoch keine Tendenz. Auffällig ist allerdings, dass es sich bei den Überlappungen häufig um kurze Sequenzen handelt, mit denen Sprecher:innen Verständnis rückmelden bzw. Zustimmung oder Erstaunen signalisieren (*hm hm*, *ach so*, *ja ja*, *ah ja*, *ja genau*) oder die eigene Turnübernahme initialisieren (*na ja*, *na dann*, *ja aber*, *ja ich*).

5. Filter zur Auswahl nach Wortarten und Mündlichkeitsphänomenen

Ein weiterer Filterbereich erlaubt die Auswahl von Sprechereignissen nach der relativen Vorkommenshäufigkeit bestimmter linguistischer Kategorien. Es handelt sich dabei einerseits um Kategorien, die traditionelle Wortarten abbilden²⁴, andererseits um solche, die für die POS-Annotation gesprochener Sprache und für Spezifika der Mündlichkeit entwickelt wurden (vgl. Westpfahl et al. 2017). Im Filter werden sechs auf Inhaltswortarten bezogene Kategorien verfügbar gemacht: Nomen (NN), Eigennamen (NE), Verben (V), Adjektive (ADJ), Adverbien (ADV) sowie in Distanzstellung verwendete trennbare Verben (vermittelt über das Tag PTKVZ für die abgetrennte Partikel). Daneben

²⁴ Im Fall der getrennt gebrauchten Verbpartikel (PTKVZ) handelt es sich um Kategorien unterhalb der Wortart.

umfasst der Bereich sechs Kategorien, die mündlichkeitsbezogene Phänomene abbilden: Häsitationen (NGHES), die Gruppe der Interjektionen, Responsive bzw. Rezeptionssignale (NGIRR), Modalpartikeln (PTKMA), Diskursmarker (SEDM), Tag Questions (SEQU) und Klitisierungen (CLITIC).

5.1 Vermittlungsbezogene Relevanz

Die Möglichkeit, Gespräche im Korpus nach dem Vorkommen bestimmter Wortartenkategorien auszuwählen, kann in verschiedener Hinsicht für die Vermittlung des Deutschen als fremder Sprache relevant sein. Zunächst stellen Wortarten selbst einen Lern- bzw. Vermittlungsgegenstand dar, da an sie bestimmte grammatische Eigenschaften gebunden sind, die entsprechend realisiert werden müssen (z. B. bzgl. Flexion, Position in der Äußerung, prosodischer Merkmale, Rektions- und Kongruenzbeziehungen, vgl. Bryant 2021; Fandrych / Thurmair 2021). Datenbeispiele auswählen zu können, in denen eine bestimmte Wortartenkategorie häufig zu finden ist, kann daher aus Perspektive der Grammatikvermittlung hilfreich sein²⁵. Daneben sind mit den einzelnen Wortarten auch unterschiedliche Schwierigkeitsprofile verbunden, die gezielt thematisiert werden können. Beispielweise bilden Verben einen anspruchsvollen Lerngegenstand, da sie als relationale Wortart und semantisch-morphosyntaktische Schnittstelle des Satzes bzw. der Äußerung bei der Verarbeitung die größte Integrationsleistung erfordern (vgl. z.B. Peters 2020: 129-130; Behrens 1999). Für das Deutsche sind zudem mit Partikelverben aufgrund ihrer morphologischen und syntaktischen Trennbarkeit besondere Herausforderungen verbunden, insbesondere für Lernende, die in der L1 nicht über vergleichbare Strukturen verfügen (vgl. Boers 2020: 152-153; Thurmair 1991). Vor diesem Hintergrund kann daher etwa eine Auswahl von authentischen Sprechereignissen mit einer relativ hohen Frequenz von getrennt gebrauchten Verben relevant sein.

Diskursmarker, Modalpartikeln, Verständnissicherungssignale, gefüllte Pausen oder Klitisierungen stellen Phänomene dar, die v.a. für die gesprochene Sprache bzw. die interaktionale Kommunikation charakteristisch sind und somit Vermittlungsgegenstände für die Mündlichkeitsdidaktik bilden (vgl. Moraldo / Missaglia 2013; Imo / Moraldo 2015; Aguado 2021: 257). Am Beispiel von Klitisierungen, d. h. Verschmelzungen von Wortformen, lässt sich dies näher illustrieren. Während etwa im Bereich der Verschmelzungen von Präposition und Artikel auch schriftsprachlich einzelne Formen üblich sind (vgl. Axel-Tobler 2022: 820-821), ist das Phänomen in der gesprochenen Sprache häufiger und hinsichtlich der betroffenen Formen vielfältiger (vgl. Thurmair 2022: 553). So haben korpusbasierte Untersuchungen wie bspw. Lenort / Pohle / Sakhno (2022) gezeigt, dass schriftsprachlich seltener synthetisch verwendete Kombinationen (z.B. *fürs*, *übers*, *vorm*), im Mündlichen deutlich häufiger auftreten. In der gesprochenen Sprache sind Klitisierungen auch bei Pronomen (*biste* [bist du], *hamwa* [haben wir]) und anderen Wortformen (z. B. *hömma* [hör mal]) zu finden (vgl. Thurmair 2022: 552, 554). Empirische Untersuchungen deuten hier auf domänen- bzw. medialitätsspezifische Gebrauchspräferenzen hin. So betrachten Schwendemann / Wallner (2022) das Verschmelzungsverhalten sprachlicher Einheiten in der gesprochenen Sprache auf der Basis des *FOLK*-Korpus insgesamt (ohne Einschränkungen auf einen Klitisierungstyp) und zeigen, dass es einen deutlichen Unterschied in der Gebrauchshäufigkeit in Abhängigkeit von der Interaktionsdomäne gibt: In den Datensätzen der öffentlichen Domäne (vgl. Kap. 3) erscheinen signifikant weniger Klitisierungen als in Datensätzen der institutionellen, privaten oder sonstigen Interaktionsdomäne (vgl. Fandrych / Meißner / Wallner 2021: 14). Um authentischen mündlichen Sprachgebrauch in der von den Lernenden angestrebten

²⁵ In diesem Fall können über *ZuMal* Sprechereignisse ausgewählt werden, die im Vergleich zu allen im Korpus verfügbaren Datensätzen eine relativ hohe Frequenz des Phänomens aufweisen. Wenn ein so selektiertes Sprechereignis in *ZuViel* geöffnet wird, erlaubt die Suche nach der POS-Kategorie eine Markierung aller Vorkommen im Transkript, die dann direkt im Kontext betrachtet werden können.

Zieldomäne zu vermitteln und auch die Besonderheit etwa medial vermittelter Sprechsituationen (Radio, Podcast) im Gegensatz zu alltäglichen und beruflichen Sprechsituationen deutlich zu machen, kann daher eine Auswahl von Datensätzen nach der relativen Häufigkeit klitisierter Formen hilfreich sein.

Die genannten, Mündlichkeitsphänomene einschließenden Wortartenkategorien stellen zudem nicht nur aus grammatisch-lexikalischer bzw. interaktionaler Perspektive Vermittlungsgegenstände dar. Vielmehr bilden sie in ihrem quantitativ unterschiedlichen Vorkommen auch Merkmale von Registern und von verschiedenen mündlichen Gattungen bzw. Textsorten. So haben Arbeiten der korpusbasierten Registeranalyse bspw. gezeigt, dass verschiedene schriftliche und mündliche Kommunikationskontexte an der Hochschule spezifische Häufigkeitsprofile von Wortarten aufweisen (v. a. im Anteil von Nomen und Verben, vgl. Biber 2006: 47). Eine Auswahl von Sprechereignissen nach der Häufigkeit bestimmter Wortartenkategorien kann daher auch im Hinblick auf die Vermittlung register- bzw. textsortenbezogener Kompetenzen relevant sein (vgl. u.a. Efing 2014; Fandrych / Thurmair 2011), etwa um zu zeigen, für welche Gesprächsarten ein häufigeres Vorkommen der Kategorie typisch ist (z. B. die Häufigkeit von Nomen in Formen des öffentlichen Sprechens, wie es bspw. in Plenardebatten im Bundestag oder in Podiumsdiskussionen zu finden ist). Zum anderen kann für die jeweilige Gesprächsart der Einsatz eines Mittels differenziert betrachtet werden (z. B. die Funktionen von Rückversicherungssignalen als frequentes Phänomen bei einer Stadtführung, im Verkaufsgespräch oder einem privaten Telefongespräch). Anwendungen dieser Art gewinnen an Bedeutung angesichts der zunehmend domänenspezifischen Bedarfe des Fremdsprachenlernens (v. a. im Kontext der Berufsorientierung), die eine verstärkte Vermittlung von register- bzw. varietätenspezifischen Kompetenzen erfordern (vgl. Venohr 2021: 175-176; Kniffka / Riemer 2022: 132).

5.2 Operationalisierung

Die Basis für die Filtermöglichkeiten nach den genannten Wortartenkategorien bildet die automatische POS-Annotation der Korpusdaten. Diese liegen hierfür aussprachenah transkribiert in den cGAT-Konventionen vor (vgl. Schmidt / Schütte / Winterscheid 2015), die auf dem GAT2-Minimaltranskript (vgl. Selting et al. 2009) basieren. Sie werden zunächst in einem vorgelagerten Prozessschritt semi-automatisch orthografisch normalisiert (vgl. Schmidt 2016: 135-136)²⁶ und anschließend nach dem für gesprochene Sprache entwickelten Tagset (vgl. Westpfahl et al. 2017) annotiert. Dieses Tagset basiert auf dem Stuttgart-Tübingen-Tagset (vgl. Schmid 1995), das um Kategorien gesprochener Sprache erweitert wurde²⁷. Die Tags beziehen sich jeweils auf einzelne Token, Mehrworteinheiten werden also nicht abgebildet (vgl. Westpfahl et al. 2017: 6). Die Annotation liegt im Korpus für jedes transkribierte Token eines Sprechereignisses vor. Es kann somit automatisch der prozentuale Anteil einer Tagkategorie relativ zu allen Token des Sprechereignisses ermittelt werden. Für den Filterbereich Wortarten und Mündlichkeitsphänomene in *ZuMal* wurden hier die oben genannten Kategorien ausgewählt. Bei den Verben werden für den Filter alle als Voll-, Modal- und Hilfsverbformen getaggt Einheiten (VV..., VM... und VA...) zusammengefasst. Bei Adjektiven umfasst der Filter

²⁶ Die orthografische Normalisierung erfolgt zunächst in einem automatischen Schritt, der eine Fehlerquote von 20% aufweist und in einem zweiten Schritt manuell nachkorrigiert werden muss (vgl. Schmidt 2016: 136; vgl. auch Schmidt 2022: 458-460).

²⁷ Der Tagger wurde dazu anhand eines manuell annotierten Datensets (Goldstandard) trainiert. Die Annotation der Datensätze im *FOLK*-Korpus erfolgt automatisch durch den so trainierten Tagger. Dabei liegt die Fehlerquote der Wortartenannotation derzeit bei 5 % (vgl. Westpfahl / Schmidt 2016). Die didaktische Arbeit mit Markierungen von POS-Kategorien im Transkript, wie sie in *ZuViel* angeboten wird (vgl. Fußnote 25), erfordert daher Bewusstheit dafür, dass dieser Visualisierung eine automatische Datenaufbereitung zugrunde liegt, die auch fehlerhafte Tagzuweisungen umfassen kann.

als attributiv (ADJA) und als adverbial (ADJD) gebraucht getaggte Vorkommen. Zu Adverbien (ADV) zählen gemäß dem STTS nur die nicht-flektierbaren Modifizierer, also keine adverbial gebrauchten Adjektive (vgl. Schiller et al. 1999: 55-56). Die für den Filter ausgewählten Kategorien von Mündlichkeitsphänomenen haben folgende Grundlagen: Häsitiationssignale (NGHES) basieren auf den zur Form *äh* normalisierten Realisierungen von gefüllten Pausen (vgl. Westpfahl et al. 2017: 16). Die Kategorie NGIRR fasst Interjektionen, Responsive und Rezeptionssignale zusammen. NG in diesem Tag steht für die Gruppe ‚nicht grammatischer‘ Elemente, die syntaktisch nicht in den Satz integriert sind (vgl. Westpfahl et al. 2017: 15). Neben lexikalischen Formen (etwa *ach*, *oh* als Interjektion) werden durch ihre positionale Alleinstellung auch homonym zu anderen Wortarten in der Funktion der Interjektion oder des Responsivs auftretende Einheiten erfasst (z. B. *oh* (NGIRR) *gott* (NGIRR) *das ist so teuer* oder *ja* (NGIRR) *klar* (NGIRR)²⁸, vgl. Westpfahl et al. 2017: 15-16). Die Kategorie PTKMA erfasst Modal- bzw. Abtönungspartikeln (z.B. *das wäre aber* (PTKMA) *cool*), die aufgrund ihrer Stellungseigenschaften von Adverbien und Fokuspartikeln abgegrenzt werden (vgl. Westpfahl et al. 2017: 21). Diskursmarker (SEQU) werden zu ‚satzexternen Elementen‘ (SE) gezählt, die im Vor-Vorfeld stehen und eine Verstehensanweisung für die folgende Äußerung geben, die also im Unterschied zu NGIRR nicht am Ende eines Turns stehen können (z. B. *also* (SEDM) *da sprechen alle noch mehr oder weniger platt*, vgl. Westpfahl et al. 2017: 26-27). Ebenfalls zu den satzexternen Elementen zählen Rückversicherungssignale und Question-tags (SEQU), die häufig im rechten Außenfeld auftreten, aber auch im Vor-Vorfeld erscheinen können und an eine Struktur gebunden sind, deren Verstehen sie abfragen (z. B. *ich bin dran gell* (SEQU), vgl. Westpfahl et al. 2017: 27-28).

Neben den genannten direkt auf annotierten Tag-Kategorien basierenden Filteroptionen wurde mit den Klitisierungen (CLITIC) eine Kategorie hinzugefügt, der selbst keine Tag-Kategorie entspricht, die aber indirekt auf der Basis der vorliegenden Korpusdaten bestimmt werden kann. Um Klitisierungen auch für zukünftig ins Korpus aufgenommene Datensätze automatisch ermittelbar zu machen, wurde hierfür ein Verfahren gewählt, das Informationen ausnutzt, die in der Transkription oder in den Wortartentags enthalten sind. Hierzu wurden als Klitisierungen einerseits die in der Transkription als assimiliert (*kommt_s*) festgehaltenen Formen gezählt. Zum anderen wurden die Fälle ausgewertet, in denen einem transkribierten Token im Rahmen der orthografischen Normalisierung zunächst zwei normalisierte Token und darauf basierend zwei POS-Tags zugeordnet wurden (z.B. wurden dem Token „s“ in *wenn s net selwer tusch* (‚wenn du es nicht selber tust‘) die beiden normalisierten Formen *du* und *es* sowie entsprechend die beiden POS-Tags PPER und PPER zugeordnet). Durch manuelle Überprüfung, eine Bereinigung auf POS-Abfolgen, die klitisiert vorkommen können, und den Ausschluss von nur einmal auftretenden Kombinationen wurde so eine Liste an Klitisierungen ermittelt. Sie wurde für den Auswahlfilter in ZuMal an eine Perzentilgrenze relativ zum Korpus gekoppelt, die aus der Ermittlung der prozentualen Anteile von Klitisierungen jeweils die 5 % der seltensten Klitisierungen (d.h. alle unterhalb des 5. Perzentils) ausschließt²⁹.

5.3 Die Filter in Anwendung: Spektren und Beispiele

Im Folgenden soll für die Wahloptionen des Filterbereichs „Wortarten und Mündlichkeitsphänomene“ illustriert werden, welche Ergebnisse sich bei ihrer Anwendung zeigen und welche Spektren

²⁸ In Fällen wie *oh gott* und *ja klar* in den Beispielen ließen sich indirekt aus der Abfolge zweier als NGIRR getaggtter Token auch Mehrworteinheiten rekonstruieren.

²⁹ CLITIC kann derzeit, anders als alle anderen beschriebenen POS-Kategorien, nicht in *ZuViel* im Transkript gesucht und markiert angezeigt werden. Eine Umweglösung ist es, in *ZuViel* nach *pos=.* zu suchen. Damit erhält man Klitisierungen, die jedoch auch solche umfassen, die für die *ZuMal*-Auswahl ausgeschlossen wurden (z. B. *hmhm* oder *Stuttgart 21*).

die Werte bezogen auf die aktuelle Version 2.19 der DGD für das *FOLK*-Korpus annehmen. Veranschaulicht werden soll dabei, analog zu Kap. 4.2, jeweils ein Beispiel aus dem Bereich jener Sprechereignisse, die aktuell die höchsten relativen Anteile der jeweiligen Kategorien aufweisen. Es wurde hierfür in *ZuMal* in der Ergebnisliste die entsprechende Wortartenkategorie als Spalte dazugewählt und nach absteigender Größe der Werte sortiert. Das Spektrum der Werte ist zusätzlich auch im Auswahlmenü auf der linken Seite ersichtlich.

So reicht das Spektrum für Nomen (NN) von einem Anteil an 6 % bis zu 19 % der Token des Sprechereignisses. Hohe Werte mit über 16 % finden sich in öffentlichen Gesprächen des Lebensbereichs Politik (Plenarsitzung im Bundestag, Ausschusssitzung), in einer Podiumsdiskussion, aber auch einer Unterrichtsstunde im Wirtschaftsgymnasium sowie einer Führung im Schloss/Schlosspark. Der Blick in die Transkripte zeigt den jeweiligen Hintergrund für die häufige Verwendung der Nomen. So offenbart etwa die Ausschusssitzung FOLK_E_00388 einen an der konzeptionellen Schriftlichkeit (vgl. Koch / Oesterreicher 2008) orientierten und durch den fachlichen Gegenstandsbezug geprägten Stil bei den Redebeiträgen, der sich in einem gehäuften Gebrauch von Nomen wieder spiegelt (z. B. ... *drittens alle maßnahmen³⁰ zur entspannung der wohnungsmärkte und zur begrenzung der mietpreisdynamik* °h sind sicher zu begrüßen °h insbesondere müssen dringend zusätzliche °h und bezahlbare wohnungen für *einpersonenhaushalte* geschaffen werden ...).

Bei Eigennamen (NE) zeigen die Datensätze in *FOLK* aktuell Anteile von 0 % bis 6 % der Token. Besonders hohe Werte liegen hier mit 5,63 % im Sprechereignis Vorlesen für Kinder (FOLK_E_0002) oder mit 4,8 % in einer Radiosendung (FOLK_E_00162) vor (z. B. in Programmankündigungen (*die cristel* (.) berichtet uns (.) über (.) onlinevorlesungen) oder in den Nachrichten (*die gewerkschaft verdi* rief heute zu warnstreiks an vielen deutschen flughäfen auf).

Bei Verben (V) reichen die Anteile von 2 % der Token des Sprechereignisses bis zu 15 %. Hohe Werte weisen mit 14,47 % z.B. ein Fahrschulgespräch (FOLK_E_00146) oder mit 14,43 % ein Training bei einer Hilfsorganisation (FOLK_E_000137) auf. In den Transkripten zeigt sich, dass die häufige Verbverwendung hier mit instruktionalen Sprachhandlungen verbunden ist (etwa durch die Lehrperson im Fahrschulgespräch z. B. ... *du musst dir ch das (0.23) du musst dir das vorher (0.49) merken (0.44) pass auf die kinder auf (0.22) und dann (.) entsprechend (.) überprüfen ob alles da is* °h (.) *das is jedes mal am anfang das gleiche (0.82) °h so wir fahren an der ampel (.) rechts weiter (4.24) lass die kupplung los ...* oder im Training der Hilfsorganisation durch die helfende Person an eine Patient:in z. B. *beruhigen s[ie sich] ... atmen ganz locker ... setzen sie sich mal hin*).

Adjektive (ADJA/ADJD) nehmen in den *FOLK*-Datensätzen Anteile von 1 % bis 12 % ein. Einen hohen Wert weist mit 7,38 % etwa ein Verkaufsgespräch im Gartencenter auf (FOLK_E_00211), in dem durch die Verkaufsperson u.a. Pflanzen beschrieben werden (z. B. ... *un gut winterh[art ja]... das is n japanischer ahorn ein schlitzahorn (0.4) °h ähm der bekommt (.) eine krone die so n bisschen dachförmig is* °h ähm wie so n schirm (0.26) je nachdem (.) nach flanze die wachsen alle n bisschen unterschiedlich auch eventuell en bisschen buschiger oder n bisschen höher auch ne [°hh] ähm wäre was so für den sonne bis halbschatten ...).

Adverbien (ADV) zeigen in den Sprechereignissen aktuell Anteile von 0 % bis 18 % der Token. Besonders hohe Werte finden sich z. B. mit über 17 % in den im Rahmen einer Experimentalsituation geführten Maptask-Gesprächen, bei denen die Proband:innen Wegbeschreibungen produzieren müssen. Daneben zeigt die Sportinteraktion FOLK_E_00376 mit 16,61 % einen hohen Wert. Hier finden Adverbien u.a. Verwendung, um räumliche Orientierungen und die Züge in einem Handballspiel wiederzugeben (z. B. JF: *dann habe ich dann halt von hinten geschossen ...* NR: *aber das +++++ [da war dass die] Nilly nicht von hinten schießen kann die kann von hinten nicht schießen ...* NC: *ja ich kann ja nicht rückwärts*)³¹.

³⁰ In den Beispielen sind jeweils die Vorkommen der thematisierten Wortart hervorgehoben.

³¹ Zur besseren Nachvollziehbarkeit wurde dieses Beispiel in der orthografisch normalisierten Fassung wiedergegeben.

In Distanzstellung gebrauchte Verben, deren Anteil im Filter über die Häufigkeit des Tags für die abgetrennte Verbpartikel (PTKVZ) erfasst wird, nehmen in den Sprechereignissen Anteile von 0 % bis 3 % ein. Den mit 2,76 % höchsten Wert zeigt die Kommunikation mit Tieren FOLK_E_00424, in der für die an ein Pferd gerichteten Äußerungen bzw. Anweisungen oft Partikelverben in Verberst- oder Verbzweitstellung verwendet werden (z. B. *charlie heb ma ma de kopf hoch ... komm her ... sodele h° jetz siekscht du wieder hübsch aus ge*).

Häsitationssignale (NGHES) weisen Anteile an den Token der Sprechereignisse von 0 % bis 11 % auf. Besonders ausgeprägt finden sie sich bspw. mit 7,88 % in der Ausschusssitzung FOLK_E_00389 oder mit 7,47 % im Prüfungsgespräch FOLK_E_00057. Hier können sie etwa auch in den Äußerungen des Prüfenden beobachtet werden, wo sie Planungsvorgänge deutlich machen (z. B. ... *das möcht ich auf jeden fall noch diskutiern °h [äh h°](.) also das ist (.) äh (.) ein (.) begriff den sie (0.67) aufführen wie s wie er vor allem vom (.) von peter auer ver äh vertreten wird °h äh also ...*).

Die Gruppe der Interjektionen, Responsive bzw. Rezeptionssignale (NGIRR) erreicht in den Sprechereignissen Anteile von 1 % bis 27 % der Token. Mit 22,03 % nimmt sie bspw. in einem Verkaufsgespräch im Supermarkt (FOLK_E_00367) einen hohen Anteil ein (z. B. Kundin CR: *ah okay gut* (6.61) Verkäuferin NN: *ich einmal bitte den ausweis sehen* CR: *hmhm* (8.41) NN: *danke schön* (17.03) *dreizehn euro dreiunsechzig +++ bitte* (0.37) *ham sie ne payback karte* CR: *nee* (2.59) NN: *danke schön*).

Modal- und Abtönungspartikeln (PTKMA) erreichen Anteile von 0 % bis 5 % der Token in den Sprechereignissen. Einen hohen Wert weist mit 5,01 % ein Telefongespräch (FOLK_E_00439) auf (hier z. B. in einer narrativen Sequenz HB: ... *nja und dann hab ich s ihm halt so nach bestem* (0.34) *°h wissen und gewissen geholfen hab halt* (0.24) *einfach so nach* (0.26) *logik oder* (0.87) *äh °hh orthografiefehlern halt geschaut und ...*).

Diskursmarker (SEDM) nehmen Anteile von 0 % bis 4 % der Token in den Sprechereignissen ein. Mit 3,71 % recht hoch liegt der Wert etwa im Prüfungsgespräch FOLK_E_00003. SEDM finden sich darin u. a. bei der Prüfungskandidatin, die damit ihre Antwort strukturiert (DM: *((schmatzt)) okay (.) also zum ersten das* (0.28) *leveltmodell oder levelt [...] wie auch immer es genannt wird °hhh ähm (.) also das modell wird ja sehr oft [rezi]piert einfach auch in einem psycholinguistischen einführungen (.) bei allen möglichen studien [°hh] und ähm °hh also dieses modell bietet einfach (.) also s is sehr detailliert (.) auch im bezug auf die (.) artikulation ...*)

Verständnissicherungssignale oder Tag Questions (SEQU) weisen Anteile von 0 % bis 3 % auf. Mit 3,16 % besitzt die Stadtführung FOLK_E_00311 hier einen hohen Wert. Im Transkript zeigt sich, dass der Stadtführer dieses Mittel in seinem Vortrag regelmäßig einsetzt (z. B. TO: ... *jetzt würd ich eigentlich ganz gern en bisschen was zur architektur des schlosses erzähl wir ham jetzt nur en bisschen das problem dass wir jetz nich besonders viel sehn aber ich deute das zumindestens mal an °h ne es gilt als das erste frühklassizistische gebäude berlins* (0.4) *ne das kann man an der fassade ganz schön sehn die is* (0.69) *im vergleich zu einem barockschloss noch verhältnismäßig stringent und einfach strukturiert °h ne und en zweites element was auch auf diesen frühklassizismus verweist °h das is das dach* (0.86) *ne und zwar ham wir s hier mit m sogenannten walmdach zu tun °h und barockschlösser die hatten flachdächer ne das kennt man vielleicht zum beispiel vom schloss versailles ne ...*)

Klitisierungen (CLITIC) betreffen 0 % bis 6 % der Token in den Sprechereignissen. Mit 5,71 % hat etwa die Schichtübergabe im Krankenhaus FOLK_E_00112 einen hohen Wert (z. B. *fang mer an mit den damen (.) mi m viererzimmer*).

Die gezeigten Beispiele verdeutlichen, wie mit Hilfe der Filter gezielt Gesprächsbeispiele aus dem FOLK-Korpus ausgewählt werden können, die natürlich bzw. authentisch vermehrte Vorkommen bestimmter didaktisch fokussierter Wortartenkategorien enthalten.

6. Didaktischer Ausblick: Die *ZuMal*-Filter im Zusammenspiel

In diesem Beitrag wurden die Filter und die Ergebnisse der Anwendung der Filter sowie die dann beobachtbaren Merkmalsausprägungen jeweils einzeln betrachtet³². Aus sprachdidaktischer Perspektive ist es jedoch sinnvoll, die Ausprägungen der Merkmale in Kombination miteinander zu betrachten, um Sprechereignisse zu identifizieren, die für spezifische Unterrichtssequenzen und für individuelle Lernziele besonders geeignet sind. In *ZuMal* ist dieses kombinatorische Vorgehen sehr leicht umzusetzen, da die ausgewählten Filter die Ergebnismenge nacheinander weiter filtern. Das bedeutet, dass der zweite Filter, der in *ZuMal* gesetzt wird, die Ergebnismenge, die durch den ersten gesetzten Filter gefunden wurde, weiter filtert. Ein weiterer gesetzter Filter reduziert die relevanten Ergebnisse dann entsprechend weiter. Ein kurzes Anwendungsszenario könnte dieses Zusammenspiel verschiedener Filter bei der Auswahl von Sprechereignissen verdeutlichen: Zunächst soll für eine Unterrichtssequenz zu gesprochener Sprache ein privates Sprechereignis (n=155) mit einer B1-Wortschatzdeckung von 90 % oder mehr ausgewählt werden. Diese Auswahl reduziert die potenziellen Sprechereignisse auf 100, die zunächst aus Wortschatzperspektive für einen Einsatz im Unterricht geeignet erscheinen. Zugleich weisen aber einige der Sprechereignisse eine sehr hohe Normalisierungsrate auf und passen aufgrund dieser Tatsache vielleicht weniger für eine Gruppe von Lernenden auf Niveau B1. In einem nächsten Schritt wird daher der Filter „Standardnähe (Normalisierungsrate)“ auf einen Bereich zwischen 2 % und 10 % eingestellt, um Sprechereignisse mit einer möglichst geringen Normalisierungsrate auszuwählen. Es bleiben noch 13 Sprechereignisse übrig. Nun sollen aus dieser Ergebnismenge Sprechereignisse gewählt werden, in denen möglichst wenig durcheinander gesprochen wird. Der Filter „Überlappungen“ wird daher auf einen Bereich zwischen 0 und 5 Überlappungen pro 1000 Token eingestellt. Es bleiben noch vier Telefongespräche. In einem letzten Schritt werden diese vier Sprechereignisse nach der Sprechgeschwindigkeit gefiltert. Dazu wird der Filter auf den Wertebereich von 2-4 artikulierte Silben pro Sekunde eingestellt. Es bleibt ein Telefongespräch übrig (FOLK_E_00395_SE_01), welches folglich eine Reihe von schwierigkeitsbezogenen Voraussetzungen erfüllt, um potenziell in einer Unterrichtssequenz auf dem Niveau B1 eingesetzt zu werden. Wenn nun der Filter „Themen“ angesteuert wird, kann zusätzlich ein Überblick über die in diesem Sprechereignis behandelten Themen gewonnen werden. Diese sind: Gartenarbeit, Gottesdienst, Herbstbeginn, Internetauftritt, Kaminofen, Kirchenchor und Zugvögel. Es ist nun möglich, das Transkript dieses Sprechereignisses über den *ZuViel*-Link aufzurufen und dieses dann mithilfe von *ZuViel* auf vielfältige Weise für den Einsatz im Unterricht aufzubereiten und besonders geeignete Sequenzen auszuwählen³³.

Eine solch mehrperspektivische Auswahl von Sprechereignissen aus Korpora der gesprochenen Sprache für unterrichtliche Kontexte war bisher nur unter großem Aufwand möglich; *ZuMal* bietet nun einen niedrighschwelligem und intuitiven Zugang zu authentischen Gesprächsdaten für Sprachdidaktiker:innen. *ZuMal* wurde hierzu in mehreren Durchläufen von Studierenden im Masterstudengang Deutsch als Fremd- und Zweitsprache am Herder-Institut der Universität Leipzig erprobt und an die Nutzer:innenbedürfnisse angepasst. Wir hoffen nun auf eine weitere Erprobung von *ZuMal* in (fremd- und zweit-)sprachdidaktischen Zusammenhängen und auf eine Erleichterung des Einsatzes von passgenauen und authentischen Sprechereignissen in Unterrichtskontexten. Wie genau die verschiedenen Filter und Filterkombinationen, und hier vor allem die schwierigkeitsbezogenen und die

³² Dabei ist zu berücksichtigen, dass sich die in diesem Beitrag berichteten Spektren auf die Sprechereignisse in *FOLK* Version 2.19 beziehen und sich natürlich mit jeder Korpuserweiterung verändern können.

³³ Weitere Beispiele für die Anwendung des Werkzeugs *ZuMal* finden sich in Fandrych / Schwendemann / Wallner (2021) sowie in Meißner / Wallner (2022).

Wortarten- und Mündlichkeitsfilter, die Interaktion von Lernenden mit ausgewählten Sprechereignissen beeinflussen, stellt dabei eines der großen Desiderata der Mündlichkeitsforschung im Fachzusammenhang DaF/DaZ dar.

Literatur und Ressourcen

Aguado, Karin (2021): Sprachliche Teilkompetenzen (1): Mündlichkeit. In: Altmayer, Claus / Biebighäuser, Katrin / Haberzettl, Stefanie / Heine, Antje (Hrsg.): *Handbuch Deutsch als Fremd- und Zweitsprache. Kontexte - Themen - Methoden*. Stuttgart: J.B. Metzler, 253-267.

Axel-Tober, Katrin (2022): Präposition. In: Wöllstein, Angelika / Dudenredaktion (Hrsg.): *Duden: Die Grammatik*. 10. völlig neu verfasste Aufl. Berlin: Dudenverlag, 810-821.

Behrens, Heike (1999): Was macht Verben zu einer besonderen Kategorie im Spracherwerb? In: Meibauer, Jörg / Rothweiler, Monika (Hrsg.): *Das Lexikon im Spracherwerb*. Tübingen: Francke, 32-50.

Biber, Douglas (2006): *University language. A corpus-based study of spoken and written registers*. Amsterdam: John Benjamins.

Boers, Frank (2020): Factors affecting the learning of multiword items. In: Webb, Stuart (Hrsg.): *The Routledge handbook of vocabulary studies*. London: Routledge, 143-157.

Bryant, Doreen (2021): Die deutsche Sprache aus der Lernendenperspektive. In: Altmayer, Claus / Biebighäuser, Katrin / Haberzettl, Stefanie / Heine, Antje (Hrsg.): *Handbuch Deutsch als Fremd- und Zweitsprache. Kontexte - Themen - Methoden*. Stuttgart: J.B. Metzler, 124-147.

Bulté, Bram / Housen, Alex (2012): Defining and operationalising L2 complexity. In: Housen, Alex / Kuiken, Folkert / Vedder, Ineke (Hrsg.): *Dimensions of L2 Performance and Proficiency*. Amsterdam: John Benjamins, 21-46. 10.1075/lllt.32.02bul.

Deppermann, Arnulf / Hartung, Martin (2012): Was gehört in ein nationales Gesprächskorpus? Kriterien, Probleme und Prioritäten der Stratifikation des „Forschungs- und Lehrkorpus Gesprochenes Deutsch“ (FOLK) am Institut für Deutsche Sprache (Mannheim). In: Felder, Ekkehard / Müller, Marcus / Vogel, Friedemann (Hrsg.): *Korpuspragmatik: Thematische Korpora als Basis diskurslinguistischer Analysen*. Berlin / Boston: De Gruyter, 415-450.

Deppermann, Arnulf / Fandrych, Christian / Kupietz, Marc / Schmidt, Thomas (Hrsg.) (2023): *Korpora in der germanistischen Sprachwissenschaft. Mündlich, schriftlich, multimedial*. Jahrbuch des Instituts für Deutsche Sprache 2022. Berlin / Boston: de Gruyter.

Efing, Christian (2014): Berufssprache & Co.: Berufsrelevante Register in der Fremdsprache. Ein varietätenlinguistischer Zugang zum berufsbezogenen DaF-Unterricht. In: *Informationen Deutsch als Fremdsprache* 41: 4, 415-441. 10.1515/infodaf-2014-0403.

Fandrych, Christian / Thurmair, Maria (2011): Plädoyer für eine textsortenbezogene Sprachdidaktik. In: *Deutsch als Fremdsprache* 47: 2, 84-93.

Fandrych, Christian / Meißner, Cordula / Slavcheva, Adriana (2012): The GeWiss Corpus: Comparing Spoken Academic German, English and Polish. In: Schmidt, Thomas / Wörner, Kai (Hrsg.): *Multilingual corpora and multilingual corpus analysis*. Amsterdam: John Benjamins, 319-337.

Fandrych, Christian / Frick, Elena / Hedeland, Hanna / Iliash, Anna / Jettka, Daniel / Meißner, Cordula / Schmidt, Thomas / Wallner, Franziska / Weigert, Kathrin / Westpfahl, Swantje (2016): User, who art thou? User Profiling for Oral Corpus Platforms. In: *Proceedings of the Tenth International Conference on Language Resources and Evaluation*. Portorož: European Language Resources Association (ELRA), 280-287. http://www.lrec-conf.org/proceedings/lrec2016/pdf/210_Paper.pdf (15.02.2023).

Fandrych, Christian / Thurmair, Maria (2021): *Grammatik im Fach Deutsch als Fremd- und Zweitsprache. Grundlagen und Vermittlung*. 2. Auflage. Berlin: Erich Schmidt.

Fandrych, Christian / Meißner, Cordula / Wallner, Franziska (2021): Korpora gesprochener Sprache und Deutsch als Fremd- und Zweitsprache: Eine chancenreiche Beziehung. In: *Korpora Deutsch als Fremdsprache* 1: 2, 5-30. 10.48694/tujournals-76.

Fandrych, Christian / Schwendemann, Matthias / Wallner, Franziska (2021): „Ich brauch da dringend ein passendes Beispiel ...“: Sprachdidaktisch orientierte Zugriffsmöglichkeiten auf Korpora der gesprochenen Sprache aus dem Projekt ZuMult. In: *Informationen Deutsch als Fremdsprache* 48: 6, 711-729.

Fandrych, Christian / Wallner, Franziska (2022): Funktionale und stilistische Merkmale gesprochener fortgeschrittener Lerner:innensprache: Methodische und konzeptionelle Überlegungen am Beispiel von GeWiss. In: *Zeitschrift für germanistische Linguistik* 50: 1, 202-239.

Fandrych, Christian / Wallner, Franziska (2023): Das GeWiss-Korpus: Neue Forschungs- und Vermittlungsperspektiven zur mündlichen Hochschulkommunikation. In: Deppermann, Arnulf / Fandrych, Christian / Kupietz, Marc / Schmidt, Thomas (Hrsg.): *Korpora in der germanistischen Sprachwissenschaft: Mündlich, schriftlich, multimedial*. Berlin / Boston: De Gruyter, 129-160.

Hacking, Jane F. / Tschirner, Erwin (2017): The Contribution of Vocabulary Knowledge to Reading Proficiency. The Case of College Russian. In: *Foreign Language Annals* 50: 3, 500-518.

Housen, Alex, / Kuiken, Folkert / Vedder, Ineke (Hrsg.) (2012): *Dimensions of L2 Performance and Proficiency*. Amsterdam: John Benjamins. 10.1075/llt.32.

Housen, Alex / Simoens, Hannelore (2016): Introduction: Cognitive Perspectives on Difficulty and Complexity in L2 Acquisition. In: *Studies of Second Language Acquisition* 38: 2, 163-175. 10.1017/S0272263116000176.

Housen, Alex / De Clerq, Bastien / Kuiken, Folkert / Vedder, Ineke (2019): Multiple approaches to complexity in second language research. In: *Second Language Research* 35: 1, 3-21.

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

Kaiser, Julia (2018): Zur Stratifikation des FOLK-Korpus: Konzeption und Strategien. In: *Gesprächsforschung* 19, 515-552. <https://ids-pub.bsz-bw.de/frontdoor/index/index/start/0/rows/10/sortfield/score/sortorder/desc/searchtype/simple/query/Kaiser+stratifikation/docId/8668> (15.03.2023).

Koch, Peter / Oesterreicher, Wulf (2008): Mündlichkeit und Schriftlichkeit von Texten. In: Janich, Nina (Hrsg.): *Textlinguistik. 15 Einführungen*. Tübingen: Narr, 199-215.

Kniffka, Gabriele / Riemer, Claudia (2022): Methodisch-didaktische Prinzipien für die Sprachvermittlung und Sprachförderung DaF und DaZ – mehr Synergien als Unterschiede? In: *Deutsch als Fremdsprache* 59: 3, 131-141. 10.37307/j.2198-2430.2022.03.02.

- Kupietz, Marc / Schmidt, Thomas (Hrsg.) (2018): *Korpuslinguistik*. Berlin, Boston: De Gruyter. 10.1515/9783110538649.
- Lenort, Lisa / Pohle, Anna / Sakhno, Anna (2022): Zur Verschmelzung von Präposition und bestimmtem Artikel im gesprochenen Deutsch. In: *Deutsch als Fremdsprache* 59: 1, 27-37. 10.37307/j.2198-2430.2022.01.04.
- Meißner, Cordula / Wallner, Franziska (2022): Korpora gesprochener Sprache als virtuelle Lernräume der Mündlichkeitsdidaktik: Affordanzen eines außerunterrichtlichen Sprachlernsettings. In: Feick, Diana / Rymarczyk, Jutta (Hrsg.): *Zur Digitalisierung von Lernorten – Fremdsprachenlernen im virtuellen Raum*. Tagungsband zum 28. DGFF-Kongress an der Julius-Maximilians-Universität Würzburg 2019, 215-239.
- Milton, James (2009): *Measuring Second Language Vocabulary Acquisition*. Bristol: Multilingual Matters. doi.org/10.21832/9781847692092.
- Moraldo, Sandro M. / Missaglia, Federica (Hrsg.) (2013): *Gesprochene Sprache im DaF-Unterricht. Grundlagen, Ansätze, Praxis*. Heidelberg: Winter.
- Pallotti, Gabriele (2015): A simple view of linguistic complexity. In: *Second Language Research* 31: 1, 117-134. 10.1177/0267658314536435.
- Peters, Elke (2020): Factors affecting the learning of single word items. In: Webb, Stuart (Hrsg.): *The Routledge handbook of vocabulary studies*. London: Routledge, 125-142.
- Read, John (2004): Research in Teaching Vocabulary. In: *Annual Review of Applied Linguistics* 24, 146-161.
- Reineke, Silke / Deppermann, Arnulf / Schmidt, Thomas (2023): Das Forschungs- und Lehrkorpus für Gesprochenes Deutsch (FOLK). Zum Nutzen eines großen annotierten Korpus gesprochener Sprache für interaktionslinguistische Fragestellungen. In: Deppermann, Arnulf / Fandrych, Christian / Kupietz, Marc / Schmidt, Thomas (Hrsg.): *Korpora in der germanistischen Sprachwissenschaft. Mündlich, schriftlich, multimedial*. Jahrbuch der Jahrestagung des Leibniz-Instituts für Deutsche Sprache 2022. Berlin, Boston: De Gruyter, 71-102.
- Schiller, Anne / Teufel, Simone / Stöckert, Christine / Thielen, Christine (1999): Guidelines für das Tagging deutscher Textcorpora mit STTS (Kleines und großes Tagset). <http://www.sfs.uni-tuebingen.de/resources/stts-1999.pdf> (15.02.2023).
- Schmid, Helmut (1995): Improvements in Part-of-Speech Tagging with an Application to German. In: Proceedings of the ACL SIGDAT-Workshop. Dublin. <http://www.cis.uni-muenchen.de/~schmid/tools/TreeTagger/data/tree-tagger2.pdf> (15.02.2023).
- Schmidt, Thomas (2016): Construction and dissemination of a corpus of spoken interaction – tools and workflows in the FOLK project. In: *Journal for language technology and computational linguistics (JLCL)* 31: 1, 127-154.
- Schmidt, Thomas (2022): Werkzeuge für die Transkription gesprochener Sprache. In: Beißwenger, Michael / Lemnitzer, Lothar / Müller-Spitzer, Carolin (Hrsg.): *Forschen in der Linguistik. Eine Methodeneinführung für das Germanistik-Studium*. Paderborn: Wilhelm Fink, 451-460.
- Schmidt, Thomas / Schütte, Wilfried / Winterscheid, Jenny (2015): *cGAT. Konventionen für das computergestützte Transkribieren in Anlehnung an das Gesprächsanalytische Transkriptionssystem 2 (GAT2)*. Institut für Deutsche Sprache. Mannheim. <https://ids-pub.bsz-bw.de/frontdoor/index/index/docId/4616> (15.02.2023).

- Schwendemann, Matthias / Wallner, Franziska (2022): Mündlichkeitsphänomene in der gesprochenen Wissenschaftssprache: Korpuslinguistische Befunde und didaktische Perspektiven. Vortrag auf der 49. Jahrestagung des Fachverbandes Deutsch als Fremd- und Zweitsprache (FaDaF) 2022 in Kassel. https://zumult.org/wp-content/uploads/2022/09/FaDaF-2022_Schwendemann-Wallner.pdf (15.02.2023).
- Selting, Margret / Auer, Peter / Barth-Weingarten, Dagmar / Bergmann, Jörg / Bergmann, Pia / Birkner, Karin et al. (2009): Gesprächsanalytisches Transkriptionssystem 2 (GAT 2). In: *Gesprächsforschung: Online-Zeitschrift zur verbalen Interaktion* 10, 353-402. <https://ids-pub.bsz-bw.de/frontdoor/index/index/start/1/rows/10/sortfield/score/sortorder/desc/searchtype/simple/query/gat+2/docId/222> (15.02.2023).
- Tschirner, Erwin / Hacking, Jane F. / Rubio, Fernando (2018): The Relationship Between Reading Proficiency and Vocabulary Size. An Empirical Investigation. In: Ecke, Peter / Rott, Susanne (Hrsg.): *Understanding Vocabulary Learning and Teaching: Implications for Language Program Development*. Boston: Cengage, 58-77.
- Tschirner, Erwin / Möhring, Jupp (2019): *A Frequency Dictionary of German. Core vocabulary for learners*. 2. Auflage. London, New York: Routledge.
- Thurmair, Maria (1991): Warten auf das Verb. Die Gedächtnisrelevanz der Verbklammer im Deutschen. In: *Jahrbuch Deutsch als Fremdsprache* 17, 174-202. 10.5283/EPUB.25061.
- Thurmair, Maria (2022): Gesprochene Sprache. In: Wöllstein, Angelika / Dudenredaktion (Hrsg.): *Duden: Die Grammatik*. 10. völlig neu verfasste Auflage. Berlin: Dudenverlag, 288-315, 551-559, 886-899.
- Van Zeeland, Hilde / Schmitt, Norbert (2013): Lexical Coverage in L1 and L2 Listening Comprehension: The Same or Different from Reading Comprehension? In: *Applied Linguistics* 34, 457-479.
- Venohr, Elisabeth (2021): Varietäten- und Soziolinguistik in DaF/DaZ unter besonderer Berücksichtigung von Fachsprachen. In: Altmayer, Claus / Biebighäuser, Katrin / Habertzettl, Stefanie / Heine, Antje (Hrsg.): *Handbuch Deutsch als Fremd- und Zweitsprache. Kontexte - Themen - Methoden*. Stuttgart: J.B. Metzler, 163-179.
- Vilkaitė-Lozdienė, Laura / Schmitt, Norbert (2020): Frequency as a Guide for Vocabulary Usefulness. High-, Mid-, and Low-Frequency Words. In: Webb, Stuart (Hrsg.): *The Routledge Handbook of Vocabulary Studies*. London / New York: Routledge, 81-96.
- Wells, John (1997): SAMPA computer readable phonetic alphabet. In: Gibbon, Dafydd / Moore, Roger / Winkski, Richard (Hrsg.): *Handbook of Standards and Resources for Spoken Language Systems. Band 4: Spoken Language Reference Materials*. Berlin, Boston: De Gruyter Mouton, 60-107.
- Westpfahl, Swantje / Schmidt, Thomas (2016): FOLK-Gold — A gold standard for part-of-speech-tagging of spoken German. In: Calzolari, Nicoletta / Choukri, Khalid / Declerck, Thierry / Goggi, Sara / Grobelnik, Marko (Hrsg.): *Proceedings of the Tenth International Conference on Language Resources and Evaluation (LREC 2016), Portorož, Slovenia. Paris: European Language Resources Association (ELRA)*, 1493-1499. <https://ids-pub.bsz-bw.de/frontdoor/index/index/docId/5078> (15.02.2023).
- Winterscheid, Jenny / Deppermann, Arnulf / Schmidt, Thomas / Schütte, Wilfried / Schedl, Evi / Kaiser, Julia (2019): *Normalisieren mit OrthoNormal. Konventionen und Bedienungshinweise für die orthografische Normalisierung von Folker-Transkripten*. Version 1.1. Mannheim. https://ids-pub.bsz-bw.de/frontdoor/deliver/index/docId/9326/file/Winterscheid_etal._Normalisierungskonventionen_2019.pdf (15.02.2023).
- Wisniewski, Katrin / Lüdeling, Anke / Czinglar, Christine (2022): Zum Umgang mit Variation in der Lerner-sprachenanalyse. Perspektiven aus und für DaF / DaZ. In: *Deutsch als Fremdsprache* 59: 4, 195-206.

Westpfahl, Swantje / Schmidt, Thomas / Jonietz, Jasmin / Borlinghaus, Anton (2017): *STTS 2.0. Guidelines für die Annotation von POS -Tags für Transkripte gesprochener Sprache in Anlehnung an das Stuttgart Tübingen Tagset (STTS)*. <https://ids-pub.bsz-bw.de/frontdoor/index/index/docId/6063> (15.02.2023).

Biographische Notiz: Christian Fandrych ist Professor für Linguistik des Deutschen als Fremdsprache am Herder-Institut der Universität Leipzig. Schwerpunkte seiner Tätigkeit sind Wortbildung und Wortschatz des Deutschen, Grammatikvermittlung, Wissenschaftssprache, Text- und Gesprächslinguistik sowie Korpuslinguistik im Kontext des Deutschen als Fremd- und Zweitsprache.

Kontaktanschrift:

Christian Fandrych
Herder-Institut der Universität Leipzig
Beethovenstr. 15
04107 Leipzig
Deutschland
fandrych@uni-leipzig.de

Biographische Notiz: Cordula Meißner ist Assistenzprofessorin am Institut für Germanistik an der Universität Innsbruck. Ihre Forschungsschwerpunkte umfassen u. a. die gebrauchsbasierte Sprachbeschreibung im Kontext der bildungsbezogenen Linguistik, Korpuspragmatik und Lexikogrammatik geschriebener und gesprochener Sprache sowie die Sprachverwendung in Wissenschaft, Bildung und Beruf.

Kontaktanschrift:

Cordula Meißner
Institut für Germanistik
Universität Innsbruck
Innrain 52d
A-6020 Innsbruck
Österreich
cordula.meissner@uibk.ac.at

Biographische Notiz: Matthias Schwendemann ist wissenschaftlicher Mitarbeiter im Bereich Linguistik am Herder-Institut der Universität Leipzig. Seine Arbeitsschwerpunkte in Forschung und Lehre liegen in den Bereichen Lexikologie, Wissenschaftssprache und Erwerb und Entwicklung des Deutschen als Fremd- und Zweitsprache sowie der Analyse von Lerner Sprache.

Kontaktanschrift:

Matthias Schwendemann
Herder-Institut der Universität Leipzig
Beethovenstr. 15
04107 Leipzig
Deutschland
matthias.schwendemann@uni-leipzig.de

Biographische Notiz: Franziska Wallner ist wissenschaftliche Mitarbeiterin am Herder-Institut der Universität Leipzig. Ihre Forschungsschwerpunkte sind unter anderen das Deutsche als fremde Bildungs- und Wissenschaftssprache, die korpusbasierte Erforschung der gesprochenen Sprache, Mündlichkeitsdidaktik sowie die Nutzung von Korpora im Kontext von Deutsch als Fremd- und Zweitsprache.

Kontaktanschrift:

Franziska Wallner
Herder-Institut
Universität Leipzig
Beethovenstr. 15
04107 Leipzig
Deutschland
f.wallner@uni-leipzig.de

