

Christian Fandrych/Franziska Wallner (Leipzig)

# Das GeWiss-Korpus: Neue Forschungs- und Vermittlungsperspektiven zur mündlichen Hochschulkommunikation

**Abstract:** Das Korpus GeWiss (Gesprochene Wissenschaftssprache kontrastiv: Deutsch im Vergleich zum Englischen und Polnischen) bietet vielfältige Möglichkeiten zur Erforschung und Vermittlung der mündlichen Hochschulkommunikation. Mit den im Projekt ZuMult entwickelten Zugangswegen zu Korpora der gesprochenen Sprache eröffnen sich für einen deutlich größeren Personenkreis umfassende Nutzungsmöglichkeiten, die sowohl für sprachdidaktische Kontexte als auch für Forschungszwecke relevant sind. In diesem Beitrag wird eine Auswahl der in ZuMult geschaffenen Werkzeuge im Hinblick auf ihr Potenzial zur Arbeit mit den GeWiss-Daten vorgestellt. Im Anschluss wird anhand von expliziten sprachlichen Positionierungsmustern aufgezeigt, wie diese Korpustools für eine sprachdidaktisch orientierte empirische Untersuchung zu den Spezifika mündlicher Wissenschaftskommunikation genutzt werden können.

## 1 Einleitung

Die mündliche Hochschulkommunikation ist für die studienbezogene Sprachvermittlung von hoher Relevanz, denn der Studienerfolg hängt in wesentlichen Teilen auch von (rezeptiven wie produktiven) studienbezogenen mündlichen Kompetenzen ab (vgl. etwa zu den komplexen Kompetenzen in mündlichen Prüfungen Rahn 2022). Dennoch ist zum einen die empirische Grundlage zur Erforschung der mündlichen Wissenschaftskommunikation noch relativ dünn, zum anderen sind die vorliegenden Korpusressourcen noch nicht ausreichend für die Vermittlungskontexte aufbereitet. Mit GeWiss (Gesprochene Wissenschaftssprache kontrastiv, <https://gewiss.uni-leipzig.de/>, Stand: 5.7.2022) steht seit einiger Zeit ein öffentlich nutzbares Korpus zur Verfügung, das es erlaubt, verschiedene sprachdidaktisch relevante Fragestellungen zu untersuchen. Gleichwohl waren die bisher entwickelten Recherchemöglichkeiten noch relativ eingeschränkt und erlaubten viele aus forschungspraktischer sowie sprachdidaktischer Sicht relevante Such- und Nutzungsmöglichkeiten nicht.

Mit dem Projekt ZuMult (vgl. <https://zumult.org/>, Stand: 5.7.2022) wurden neue Zugangswege zu Korpora der gesprochenen Sprache geschaffen, die eine

<https://doi.org/10.1515/9783111085708-007>

Vielfalt an neuen Forschungs-, Nutzungs- und Anwendungsmöglichkeiten ermöglichen. Das Potenzial dieser neuen Zugangswege soll im vorliegenden Beitrag exemplarisch anhand des GeWiss-Korpus aufgezeigt werden. Dazu wird zunächst das GeWiss-Korpus bezüglich seiner Anlage, Datenstruktur und den bereits erfolgten Annotationen kurz vorgestellt (2). Im Folgenden werden dann in zwei Schritten die neuen Nutzungs- und Forschungsmöglichkeiten dargestellt, die sich durch die ZuMult-Werkzeuge ergeben: In 3.1 werden die mit den neuen Tools ermöglichten sprachdidaktischen Anwendungsszenarien beschrieben, in 3.2 wird anhand der expliziten argumentativen Positionierungen aufgezeigt, wie mithilfe der ZuMult-Werkzeuge didaktisch motivierte Forschungsfragen bearbeitet werden können. Den Abschluss bilden Überlegungen zu forschungsbezogenen und didaktischen Perspektiven, die sich durch solche, auf der Basis von Nutzerstudien erarbeitete Weiterentwicklungen von Korpustools ergeben.

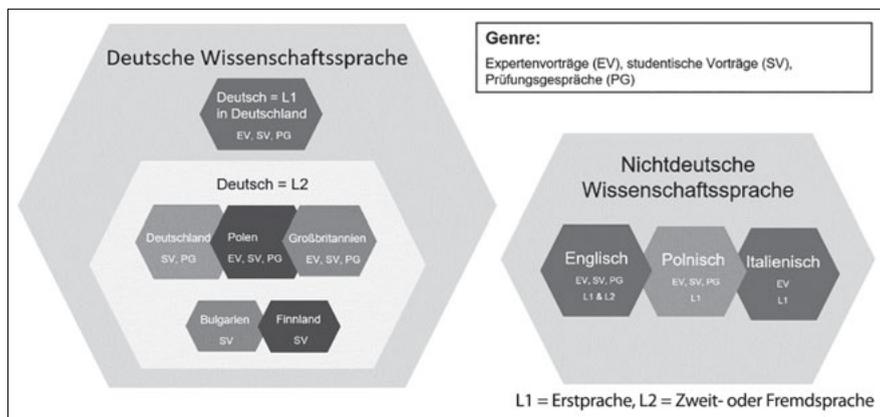
## 2 GeWiss – ein Korpus der gesprochenen Wissenschaftssprache

GeWiss ist ein Vergleichskorpus der gesprochenen Wissenschaftssprache, das Prüfungsgespräche, studentische Referate und Expertenvorträge aus philologischen Fächern in deutscher, englischer, italienischer und polnischer Sprache bereitstellt. Die Daten stammen aus unterschiedlichen akademischen Kontexten (darunter Bulgarien, Deutschland, Großbritannien, Polen und Italien)<sup>1</sup> und wurden in authentischen Kommunikationssituationen erhoben. Es handelt sich demnach um „natürliche“, d. h. nicht zum Zweck der Korpuserstellung elizitierte Sprachdaten. Neben L1-Daten enthält das Korpus auch L2-Produktionen für die Sprachen Deutsch und Englisch.<sup>2</sup> Abbildung 1 visualisiert die Zusammensetzung des GeWiss-Korpus.

---

<sup>1</sup> Zu einem späteren Zeitpunkt (2018) kamen auch in Finnland erhobene Daten dazu, diese sind jedoch ausschließlich über die Datenbank für Gesprochenes Deutsch zugänglich.

<sup>2</sup> Eine detaillierte Beschreibung der deutschsprachigen L2-Produktionen findet sich in Fandrych/Wallner (2022).



**Abb. 1:** Zusammensetzung des GeWiss-Korpus

Insgesamt enthält das Korpus 1.205.306 Token und 146 Aufnahmestunden. Die deutschsprachigen Daten bilden mit 742.332 Token und rund 92 Aufnahmestunden den Großteil des GeWiss-Korpus. Tabelle 1 gibt einen Überblick über die Anzahl der im gesamten sowie im deutschsprachigen GeWiss-Korpus vertretenen Genres und die jeweils zugehörigen Tokenzahlen:

**Tab. 1:** Überblick über die im GeWiss-Korpus vertretenen Genres

Genre	GeWiss gesamt		GeWiss deutsch	
	Sprechereignisse	Token	Sprechereignisse	Token
Expertenvorträge	76	374.380	33	166.639
Studentische Vorträge	137	328.533	106	240.935
Prüfungsgespräche	223	502.393	137	334.800
<b>gesamt</b>	<b>436</b>	<b>1.205.306</b>	<b>276</b>	<b>742.374</b>

Das GeWiss-Korpus beinhaltet die Audioaufnahmen zu den einzelnen Sprechereignissen sowie die dazugehörigen aussprachenahen Transkriptionen. Zu jedem Transkript ist zudem eine orthografisch normalisierte Fassung verfügbar. Darüber hinaus erfolgten weitere korpuslinguistische Aufbereitungsschritte, darunter das POS-Tagging für das gesamte Korpus sowie die Annotation von Sprachwechseln (Reershemius/Lange 2014), Diskurskommentierungen (Fandrych 2014), Verweisen und Zitaten (Sadowski 2017) in ausgewählten Teilkorpora.

Die GeWiss-Daten waren zunächst über das GeWiss-Portal (<https://gewiss.uni-leipzig.de/>, Stand: 5.7.2022) zugänglich, über ein CLARIN-D-Kurationsprojekt wurden sie ab 2017 zudem in die Datenbank für Gesprochenes Deutsch (DGD) (<https://dgd.ids-mannheim.de/>, Stand: 5.7.2022) integriert. Diese beiden Schnittstellen bieten jeweils unterschiedliche Nutzungsmöglichkeiten, sind aber auch mit verschiedenen Einschränkungen verbunden. So sind bspw. über das GeWiss-Portal nur die aussprachenahen Transkriptionen zugänglich. Auf die orthografisch normalisierte Fassung sowie auf das POS-Tagging kann nur über die DGD zugegriffen werden. Die annotierten Sprachwechsel, die Diskurskommentierungen sowie die Verweise und Zitate können wiederum nur über das GeWiss-Portal abgerufen werden (vgl. hierzu auch Frick/Helmer/Wallner *angen.*).

Die Zugangs- und Nutzungsmöglichkeiten für GeWiss sind, wie das hier nur sehr kurz angedeutet werden kann, also je nach Plattform unterschiedlich und dadurch auch komplex. Es kommt hinzu, dass viele Nutzer/-innen mit wichtigen korpusbezogenen Such- und Recherchefunktionen nur wenig vertraut sind, aber spezifische Anwendungsinteressen haben, die mithilfe der bestehenden Funktionen der Plattformen nicht oder nur sehr aufwendig erfolgreich bedient werden können. Dies hat auch eine Studie zur Nutzung von Korpora der gesprochenen Sprache (darunter auch die DGD und das GeWiss-Portal) bestätigt (Fandrych et al. 2016). Als eine besonders zentrale Nutzer/-innengruppe haben sich dabei Akteur/-innen aus dem Kontext der Sprachvermittlung erwiesen. Diese verfügen häufig über keine ausgeprägte korpuslinguistische Expertise, haben aber ein starkes Interesse daran, authentische Korpusdaten der gesprochenen Sprache für sprachdidaktische Zwecke zu nutzen. Hierfür sind sprachdidaktisch einschlägige Parameter als Suchkriterien relevant, etwa sprachliche Schwierigkeit, Standardnähe bzw. -ferne, Wortschatzabdeckung etc., um so die zielgenaue Auswahl von geeigneten Sprechereignissen für die Materialentwicklung oder auch den konkreten Sprachunterricht zu ermöglichen. Zudem wurde der Wunsch nach Werkzeugen geäußert, die eine umfassendere Aufbereitung der Transkripte für Unterrichtszwecke ermöglichen.

Die Ergebnisse der Nutzer/-innenstudie, die bereits erwähnten Einschränkungen der bisherigen Zugriffsmöglichkeiten auf die GeWiss-Daten sowie weitere Begrenzungen bisheriger Schnittstellen für Korpora der gesprochenen Sprache gaben den Anstoß für das Projekt ZuMult (Zugänge zu multimodalen Korpora gesprochener Sprache) – einem von der DFG geförderten Projekt am Leibniz-Institut für Deutsche Sprache (IDS) Mannheim, am Herder-Institut der Universität Leipzig sowie am Hamburger Zentrum für Sprachkorpora (HZSK).<sup>3</sup> Im Rah-

---

<sup>3</sup> Ausführliche Informationen zum Projekt finden sich unter <https://zumult.org/> (Stand: 5.7.2022).

men des Projekts wurden Zugangswege zu Korpora der gesprochenen Sprache entwickelt, die einerseits an den Nutzungsbedürfnissen von Sprachdidaktiker/-innen ausgerichtet sind, um somit für diese Zielgruppe einen bedarfsgerechten Zugriff auf authentische gesprochensprachliche Daten zu ermöglichen. Andererseits bieten sie aber auch einen über die bisherigen Nutzungsoptionen deutlich hinausgehenden Nutzungsumfang für die korpusorientierte Arbeit insgesamt und ermöglichen somit eine umfassendere Erforschung der gesprochenen Sprache.

Im folgenden Kapitel werden drei zentrale Werkzeuge vorgestellt, die im Projekt ZuMult erstellt wurden, und im Hinblick auf ihr Potenzial zur Vermittlung und Erforschung der mündlichen Hochschulkommunikation anhand der GeWiss-Daten beschrieben.

### 3 ZuMult: Neue Nutzungsmöglichkeiten von GeWiss

Wie bereits angesprochen, wurden im Rahmen des Projekts ZuMult verschiedene Tools entwickelt, die einen bedarfsgerechten Zugang zu Korpora der gesprochenen Sprache ermöglichen.<sup>4</sup> Anhand der Tools ZuMal (Zugang zu Merkmalsauswahl von Gesprächen), ZuViel (Zugang zu Visualisierungselementen für Transkripte) und ZuRecht (Zugang zur Recherche in Transkripten) sollen im Folgenden zunächst Nutzungsmöglichkeiten der GeWiss-Daten aus sprachdidaktischer Perspektive beschrieben werden (3.1). Daran anschließend wird am Beispiel der Positionierungen als einer zentralen wissenschaftssprachlichen Handlung gezeigt, wie vermittlungsrelevante Forschungsfragen zur mündlichen Hochschulkommunikation mit Hilfe der ZuMult-Tools angegangen werden können (3.2).

#### 3.1 Sprachdidaktische Anwendungsperspektiven

Aus sprachdidaktischer Perspektive bieten Korpora der gesprochenen Sprache ein großes Potenzial. Sie ermöglichen den Lernenden einen Zugang zu authentischen mündlichen Interaktionsdaten, können für die Förderung des Hörverste-

---

<sup>4</sup> Bei diesen Tools handelt es sich um Prototypen. Sie sind über die DGD sowie unter <http://zumult.ids-mannheim.de/ProtoZumult/index.jsp> (Stand: 5.7.2022) zugänglich. Für die Nutzung ist eine Registrierung bei der DGD erforderlich.

hens eingesetzt werden und bieten vielfältige Möglichkeiten zur Erarbeitung von genrespezifischen sprachlichen Mitteln und Formulierungs- bzw. Textroutinen (vgl. Feilke 2012). Dabei ist ein eher gesteuerter Einsatz ebenso denkbar wie eine eigenständige Suche nach typischen Formulierungen und Versprachlichungsmustern durch die Lernenden selbst (vgl. hierzu ausführlicher Dietz 2021; Fandrych/Meißner/Wallner 2018, 2021; Fandrych/Schwendemann/Wallner 2021; Meißner/Wallner 2022). Im Kontext der Vermittlung mündlicher Hochschulkommunikation ist eine solche Arbeit mit authentischen Sprachdaten besonders relevant, da erst auf dieser Grundlage eine erfolgreiche Aneignung einer adäquaten mündlichen Kommunikationsfähigkeit im akademischen Kontext ermöglicht werden kann. Das GeWiss-Korpus bietet hierfür eine ideale Grundlage, da mit den enthaltenen Genres (Vorträge und Prüfungsgespräche) zwei zentrale und für den Bildungserfolg hoch relevante Handlungsfelder der mündlichen Hochschulkommunikation abgedeckt werden.

Mit dem Tool ZuMal wurde ein Werkzeug geschaffen, das eine bedarfsgerechte Auswahl von Sprechereignissen aus dem GeWiss-Korpus ermöglicht. Dabei können verschiedene metadatenbezogene Filter wie bspw. Sprache, Genre, Aufnahmeort oder Dauer des Sprechereignisses eingesetzt werden. Zudem ist es möglich, gezielt nach Sprechereignissen zu suchen, die überwiegend Sprecher/-innen mit Deutsch als Erst- und/oder Fremd- bzw. Zweitsprache enthalten. Es eröffnen sich dadurch vielfältige Vergleichsdimensionen, die im studienvorbereitenden und -begleitenden DaFZ-Unterricht aufgegriffen werden können. Hierzu zählt etwa der Vergleich von Prüfungsgesprächen oder Vortragsweisen in verschiedenen akademischen Kontexten, von Eröffnungssequenzen in Vorträgen in verschiedenen Sprachen oder auch der Grad der Mündlichkeit bei Sprecher/-innen mit Deutsch als Erst- vs. Fremdsprache. Daneben können – sowohl in Kombination als auch gesondert von den metadatenbezogenen Filtern – schwierigkeitsbezogene Filter genutzt werden. So ist es bspw. möglich, für eine noch wenig fortgeschrittene Gruppe von Lernenden nach studentischen Vorträgen zu suchen, deren Wortschatz zu mindestens 80% der Niveaustufe A2 des Gemeinsamen europäischen Referenzrahmen (GER) zugeordnet werden kann und die auf der Ebene der sprachlichen Realisierung der einzelnen Wörter möglichst wenige Abweichungen von einem angenommenen schriftsprachlichen Standard aufweisen. Alternativ ist es aber auch denkbar, für eine weiter fortgeschrittene Gruppe gezielt nach herausfordernden Vorträgen oder Prüfungsgesprächen zu suchen, die bspw. viele Abweichungen vom schriftsprachlichen Standard aufweisen, eher schnell gesprochenen sind oder einen hohen Anteil an Phänomenen der gesprochenen Sprache wie bspw. Klitisierungen enthalten. Darüber hinaus ist auch eine wortartbezogene Filterung der Daten möglich. So können bspw. Vorträge mit hohem und niedrigem Anteil an Nomen ermittelt werden, um sie dann im Sprach-

unterricht im Rahmen der Thematisierung von Nominalstil in der gesprochenen Wissenschaftssprache miteinander zu vergleichen.

Die in ZuMal ermittelten Sprechereignisse können sodann nach dem Auswahlprozess im Transkriptbrowser ZuViel aufgerufen werden. Von dort aus können sie vollständig oder in Sequenzen abgespielt werden, wobei die Möglichkeit besteht, die Abspielgeschwindigkeit an die Bedürfnisse der Nutzer/-innen anzupassen. Die Transkripte zu den Sprechereignissen lassen sich sowohl in aussprachenaher Transkription als auch in orthografisch normalisierter Fassung anzeigen, was insbesondere für einen binnendifferenzierenden Einsatz von Transkripten im Sprachunterricht von Vorteil ist. Leistungsstarke Lernende könnten dabei die aussprachenahere Transkription erhalten, während schwächere Lernende mit der orthografisch normalisierten Fassung arbeiten würden. Ergänzend zum Transkript wird eine Lemmaliste bereitgestellt, die einen Überblick über den im Transkript enthaltenen Wortschatz gibt. Die Lemmaliste kann zusätzlich mit einer Referenzwortschatzliste abgeglichen werden. So lassen sich bspw. diejenigen Wörter eines Transkripts schnell identifizieren, die einer Lernendengruppe auf B1 vermutlich noch nicht bekannt sind, weil sie bspw. nicht zu den häufigsten 5000 Wörtern des Deutschen zählen, einer Bezugsgröße, die mit dem Niveau B2/C1 assoziiert wird (Tschirner 2019). Zudem werden in ZuViel eine Reihe von Markierungs- und Downloadoptionen angeboten, die eine umfassende didaktische Aufbereitung ermöglichen. So können bspw. sämtliche vom orthografischen Standard abweichend realisierte Token hervorgehoben werden, um so die Aufmerksamkeit der Lernenden gezielt auf diese Phänomene zu lenken. Auch Wortarten wie etwa Diskursmarker oder Modalpartikeln sowie die Niveaustufenzugehörigkeit des Wortschatzes lassen sich im Transkript markieren. Zusätzlich bietet der Density Navigator eine kompakte Übersicht über die zeitliche Anordnung der Gesprächsbeiträge der verschiedenen Sprecher/-innen. Dies ermöglicht es bspw., eher monologische Sequenzen in Vorträgen und Prüfungsgesprächen von stärker interaktiven Phasen zu unterscheiden. Auf diese Weise lassen sich auch potenzielle Eröffnungs- und Beendigungsphasen sowie Sequenzen mit einem hohen Aufkommen an Turn-Taking-Aktivitäten identifizieren und somit für die Vermittlung bzw. Aneignung wissenschaftssprachlicher Handlungsfähigkeit nutzbar machen.<sup>5</sup>

---

<sup>5</sup> Die Auswahl von Sprechereignissen mit ZuMal sowie Nutzungsoptionen mit ZuViel werden anhand von Daten aus dem Forschungs- und Lehrkorpus (FOLK) in Fandrych/Schwendemann/Wallner (2021) sowie in Meißner/Wallner (2022) ausführlich dargestellt. Eine detaillierte Beschreibung der einzelnen Komponenten von ZuMal und ZuViel findet sich zudem in den Hilfsdokumenten der beiden Tools.

Bei ZuRecht handelt es sich um eine Benutzeroberfläche, die mit Hilfe komplexer CQP-Suchanfragen<sup>6</sup> einen umfassenden Zugriff auf Transkriptionen gesprochener Sprache gestattet. Dabei sind sowohl Konkordanzsuchen als auch gezielte Suchen nach Sprechereignissen über die Wortschatzsuche möglich. Bezogen auf das GeWiss-Korpus eröffnen sich damit vielfältige für didaktische Kontexte relevante Rechercheoptionen, die über die bisherigen Zugangswege nicht umsetzbar waren. Neben der Suche nach einzelnen Wortformen und -verbindungen unter Einbezug der tokenbasierten Annotationen (wie Lemma, orthografische Normalisierung und Wortart) ist es mit Hilfe der Konkordanzsuche möglich, metadaten- und positionsbezogene Bedingungen direkt in die Suchanfrage zu integrieren. Dabei können auch bislang nicht zugängliche Bedingungen einbezogen werden wie bspw. Sprechgeschwindigkeit, Wiederholungen sowie sprachbiografische Informationen für eine Auswahl von Sprechereignissen, die überwiegend L1- bzw. L2-Sprecher/-innen enthalten. So kann etwa mit der folgenden Abfrage gezielt nach TagQuestions in Expertenvorträgen mit Deutsch als Erstsprache gesucht werden (Abb. 2):

---

<sup>6</sup> CQP ist ursprünglich die Abfragesprache des Corpus Query Processors – einer linguistischen Suchmaschine, die an der Universität Stuttgart als Teil der IMS Open Corpus Workbench (CWB) entwickelt wurde (vgl. <https://cwb.sourceforge.io/>, Stand: 5.7.2022). Eine detaillierte Beschreibung der Suchanfragesprache findet sich in ZuRecht unter dem Fragezeichen-Button neben dem CQP-Suchanfragefeld.

**<[pos="SEQU" within <e\_se\_art="Expertenvortrag"/>] within <e\_se\_sprachen="Deutsch \ (L1)"/>**

The screenshot shows the ZuRecht search interface. The search query is entered in the search bar: `<[pos="SEQU" within <e_se_art="Expertenvortrag"/>] within <e_se_sprachen="Deutsch \ (L1)"/>`. The interface includes a sidebar for selecting corpora (DH, FOLK, GWSS, MEND, DNAM, BETV, HMAT), a search bar with a search button, and a results section. The results section shows a total of 37 hits and a list of search results with columns for ID, text snippet, and actions (DGD, ZuViel).

Ergebnisse	
für die Suchanfrage <code>&lt;[pos="SEQU" within &lt;e_se_art="Expertenvortrag"/&gt;] within &lt;e_se_sprachen="Deutsch \ (L1)"/&gt;</code> (in GWSS)	
[Treffer gruppieren] [Metadatenansicht] [Download der KWIC]	
Insgesamt: 37	
[Erste Seite] [Vorherige Seite] [1] [2] [3] [4] [Nächste Seite] [Letzte Seite]	
1	GWSS_E_00027_SE_01_T_01 OR_0236 ... touristen oh (1,07) anlocken richtig hm hm wenn ... DGD ZuViel
2	GWSS_E_00029_SE_01_T_01 LF_0245 ... sofa gucken oder ne auf den distractor ... DGD ZuViel
3	GWSS_E_00029_SE_01_T_01 LF_0245 ... blinzeln zum beispiel ne dann hab ich ... DGD ZuViel
4	GWSS_E_00029_SE_01_T_01 LF_0245 ... wer hier also ne ah dann dann ... DGD ZuViel

**Abb. 2:** Suchanfrage nach TagQuestions in Expertenvorträgen mit Deutsch als Erstsprache

Außerdem ist es mit ZuRecht möglich, nach pragmatischen Annotationen wie Sprachwechseln, Diskurskommentierungen sowie Verweisen und Zitaten zu recherchieren. Diese lassen sich auch in Kombination mit tokenbasierten Annotationen abfragen. Denkbar wäre hier etwa die Suche nach Diskurskommentierungen, die eine Modalpartikel enthalten, mit der Suchanfrage

**<DK/> containing [pos="PTKMA" ]** (Abb. 3)

oder auch die Abfrage aller Verweise und Zitate, die die Präposition *nach* enthalten, mit der Suchanfrage

**<VZ/> containing [pos="APPR" & word="nach" ]** (Abb. 4).

<sup>7</sup> Die Suchanfragen können mit Hilfe eines Query Builders erstellt werden. Es handelt sich dabei um eine visuelle Komponente zur schrittweisen Erstellung der Suchanfragen über die grafische Benutzeroberfläche. Nutzer/-innen ohne Kenntnisse der CQP-Suchanfragesprache erhalten hier ein Auswalmenü mit Eingabeoptionen zur Vervollständigung der eigenen Suchanfrage.

Lehrende können anhand der so ermittelten Belegstellen die sprachliche Ausgestaltung dieser wissenschaftssprachlichen Handlungen für ihre Lernenden illustrieren.

The screenshot shows the ZuRecht search interface. The search query is: `<DK/> containing [pos="PTKMA"]`. The results are for the search query `<DK/> containing [pos="PTKMA"]` in the GWSS corpus. There are 82 results in total. The first four results are shown:

Rank	Document ID	Text Snippet	Source	Actions
1	GWSS_E_00027_SE_01_T_01	OR_0236 ... und kollegen (0.49) oh (0.59) sie sehen schon (0.35) der titel is etwas lang geraten ( ) oh im vergleich ...	DGD ZuViel	
2	GWSS_E_00027_SE_01_T_01	OR_0236 ... gesehen ham (0.43) aber ( ) spätestens seit dem vortrag vorhin von herrn von herrn lichtenberg 'h oh (0.32) wissen wir ja dass raumkonstruktionen oh rich ...	DGD ZuViel	
3	GWSS_E_00027_SE_01_T_01	OR_0236 ... der dialektgeschichte ( ) erklären dazu hamma ja ( ) in den vorträgen bisher 'h auch schon ne ganze ( ) menge gehört das sin eben ...	DGD ZuViel	
4	GWSS_E_00027_SE_01_T_01	OR_0236 ... sich ( ) geschaffen hat 'h ich komm mal noch kurz noch ( ) drauf zurück 'h ich will jetzt ...	DGD ZuViel	

Abb. 3: Suchanfrage nach Diskurskommentierungen, die eine Modalpartikel enthalten

The screenshot shows the ZuRecht search interface. The search query is: `<VZ/> containing [pos="APPR" & word="nach"]`. The results are for the search query `<VZ/> containing [pos="APPR" & word="nach"]` in the GWSS corpus. There are 15 results in total. The first six results are shown:

Rank	Document ID	Text Snippet	Source	Actions
1	GWSS_E_00051_SE_01_T_01	OF_0216 genau (3.83) und ähm (0.81) nach schüller (0.23) teilt sie deutunglernen ...	DGD ZuViel	
2	GWSS_E_00051_SE_01_T_01	OF_0216 ... deutunglernen ( ) es gibt nach schüller implizites und explizites ...	DGD ZuViel	
3	GWSS_E_00051_SE_01_T_01	OF_0216 passiert fast automatisch (0.59) un dann gibt es explizites deutunglernen was nach sch nach schüller ( ) besser ist oder effektiver (0.65) äh und ( ) solche ...	DGD ZuViel	
4	GWSS_E_00055_SE_01_T_01	AGA_0214 der sprachlichen regeln (0.54) nach xxx xxx habe ich hier ein zitat von potapov (1.07) äh ( ) der sprechrhythmus ...	DGD ZuViel	
5	GWSS_E_00058_SE_01_T_01	CN_0205 leben in deutschland nach angaben des bundesamtes für migration und flüchtlinge cirka sieben komma ...	DGD ZuViel	
6	GWSS_E_00060_SE_01_T_01	LP1_0220 erst mal ähm (0.34) wir haben hier eine (0.27) uns an einer gliederung orientiert ( ) nach reich ( ) zwotausendacht (0.59) ähm ( ) der das ...	DGD ZuViel	

Abb. 4: Suchanfrage nach Verweisen und Zitaten, die die Präposition *nach* enthalten

Mithilfe der Wortschatzsuche können darüber hinaus in ZuRecht auch Sprechereignisse gesucht werden, die einen bestimmten Wortschatz enthalten. Hierfür stehen drei bereits vorgefertigte, auf dem Übungswortschatz „Sage und Schreibe“ (Fandrych/Tallowitz 2019) basierende Wortschatzlisten zu den Themenbereichen „Essen“, „Haus und Wohnung“ und „Schule und Ausbildung“ zur Verfügung. Daneben besteht die Möglichkeit, auch eigene Listen im Dateiformat txt hochzuladen, die den Bedürfnissen und Lernzielen einer Lerngruppe entsprechen. So ist es bspw. denkbar, eine Liste mit fachübergreifenden Verben aus dem GeSIG-Inventar<sup>8</sup> hochzuladen, um Sprechereignisse zu ermitteln, in denen diese besonders häufig genutzt werden. Die txt-Dateien können neben Wortschatzsammlungen auch einzelne Elemente enthalten, wie bspw. die Wortform *ich*. Es lassen sich so Sprechereignisse identifizieren, anhand derer sich der Ich-Gebrauch in Vorträgen gut veranschaulichen lässt. Alternativ kann die Wortschatzsuche auch dazu genutzt werden, nach Sprechereignissen mit einem hohen Anteil pragmatischer Annotationen zu suchen. Enthält die hochzuladende txt-Datei bspw. eine Suchanfrage nach Verweisen und Zitaten (<VZ/>), können diejenigen Sprechereignisse gefunden werden, in denen besonders viele Verweise und Zitate vorkommen. Die so ermittelten Sprechereignisse können wiederum im Transkriptbrowser ZuViel aufgerufen werden. Der Vorteil besteht darin, dass die jeweils bei der Wortschatzsuche fokussierten Elemente in ZuViel rot umrandet werden und damit für didaktische Zwecke bereits markiert sind (Abb. 5).<sup>9</sup>

---

<sup>8</sup> Das GeSIG-Inventar ist eine Wortschatzliste, die den Wortschatz umfasst, der in den Geisteswissenschaften fachübergreifend verwendet wird. Sie ist unter <https://www.esv.info/t/gesig/aktualisierung.html> (Stand: 5.7.2022) frei verfügbar.

<sup>9</sup> Weitere Beispiele für sprachdidaktische Nutzungsmöglichkeiten von ZuRecht werden in Frick/Helmer/Wallner (angen.) vorgestellt.

0057 nn ((räuspert sich))

0058 nn (0.32)

0059 LE\_0201 ja in unserem zweiten teil geht es um das thema was is literarizität (0.66) ((schmatzt)) unser zweiter punkt \*h

0060 nn ((räuspert sich))

0061 nn ((hustet))

0062 LE\_0201 äh wir orientieren uns an der definition von roman jacobson (0.92) äh demnach list literarizität ein merkmal (0.36) das an eine bestimmte einstellung des rezipienten beziehungsweise des produzenten eines textes gebunden ist (0.57) an die einstellung zitat jacobson (0.88) auf die nachricht die botschaft (.) als solche (0.33) ((schmatzt)) die zentrierung auf die nachricht (.) um ihrer selbst willen (0.52) ((schmatzt)) \*h die nachricht das sprachliche zeichen das wort beziehungsweise der text den text in diesem poetischen sinne zu fokussieren (0.46) heißt nach jacobson (0.46) die aufmerksamkei (0.28) von den äquivalenzen bei der auswahl der worte weg (0.76) und auf die äquivalenzen bei der verketzung (.) der worte hinzulenken

0063 (1.51)

0064 LE\_0201 ((schmalzt)) \*h in der formulierung von terry leagleton (0.82) wir reihen wörter aneinander (0.53) die semantisch (0.29) rhythmisch (0.36) phonetisch (0.26) oder auf eine andere weise äquivalent sind

0065 (1.2)

0066 LE\_0201 \*h jacobson (0.3) wo hamma s zitat nee das hab ich noch nich (0.28) xxx nich (.) pardon \*h jacobson (0.28) ähm (.) selbst illustriert seine these mit dem folgenden alltagsbeispiel das sie bestimmt schon mal (0.27) gehört haben a girl used to talk about \*h the horrible harry (0.4) ((schmatzt)) \*h why horrible (0.48) because i hate him (0.74) but why not dreadful terrible frightful disgusting (0.67) i don t know why (0.27) but horrible (0.27) fits him better

Abb. 5: Verweise und Zitate in einem Ausschnitt aus einem Expertenvortrag (GWSS\_E\_00021\_SE\_01) in ZuViel

### 3.2 Forschungsperspektiven: Das Beispiel argumentativer Positionierungen

Im Folgenden wollen wir anhand von Konstruktionen zur expliziten argumentativen Positionierung aufzeigen, wie das GeWiss-Korpus unter Nutzung einiger der im Rahmen von ZuMult entwickelten Tools für sprachdidaktisch ausgerichtete Forschung und in der Folge auch für die Anwendung in der Sprachausbildung verwendet werden kann. Hierzu erfolgt zunächst eine kurze theoretische Einführung zu argumentativen Positionierungen (3.2.1). Im nächsten Schritt werden exemplarische Analysen zum Vorkommen dieser wissenschaftssprachlichen Handlungen in Vorträgen aus dem GeWiss-Korpus einerseits und in wissenschaftlichen Texten (Dissertationen) andererseits vorgestellt (3.2.2 und 3.2.3). Im

Fokus stehen dabei die einzelnen Schritte zur Ermittlung und Analyse der Belegstellen sowie die Vorteile und Potenziale, welche die in ZuMult entwickelten Tools für eine solche Analyse bieten. Schließlich werden erste, auf den vorgestellten Ergebnissen der Analyse beruhende didaktische Schlussfolgerungen formuliert (3.2.4) und aufgezeigt, wie die Fachcommunity auf die ermittelten Belegstellen für argumentative Positionierungen zugreifen kann (3.2.5).

### 3.2.1 Argumentative Positionierungen als Untersuchungsgegenstand der mündlichen Hochschulkommunikation

Die mündliche Wissenschaftskommunikation ist bisher noch vergleichsweise wenig untersucht worden;<sup>10</sup> insofern überrascht es nicht, dass sie in der (Fremd-) Sprachendidaktik eine sehr untergeordnete Rolle spielt. Dies liegt auch an der mangelnden empirischen (Korpus-)Basis. Diese Situation hat dazu geführt, dass sich didaktische Publikationen zum Wissenschaftsdeutschen fast ausschließlich an den Gattungen, Konventionen und sprachlichen Mitteln der schriftlichen Wissenschaftskommunikation orientiert haben (vgl. etwa Steinhoff 2007; Graefen/Moll 2011; Gätje/Rezat/Steinhoff 2012; Moll/Thielmann 2017; Emmrich 2019). So wird das Spezifische, das die mündliche Wissenschaftskommunikation auszeichnet, übersehen und es kann leicht zu einer unhinterfragten Übertragung von Stilkonventionen und Schreibmustern auf mündliche Gattungen kommen.<sup>11</sup>

Das GeWiss-Korpus ermöglicht es u. a., ausgewählte mündliche Gattungen wie den wissenschaftlichen Vortrag oder das studentische Referat mit funktional ähnlichen schriftlichen Textsorten zu vergleichen. Eine interessante Frage ist dabei, in welcher Weise – häufig musterhaft realisierte – sprachliche Handlungen mit ähnlichen Funktionen in den verschiedenen Gattungen realisiert werden. Dass es hier Unterschiede gibt, lassen Analysen im englischsprachigen Kontext vermuten, die zeigen, dass sprachliche Handlungen, die im weitesten Sinne eigene Bewertungen und Positionierungen ausdrücken (*stance*), in mündlichen Gattungen sprachlich wesentlich expliziter realisiert werden als in geschriebe-

<sup>10</sup> Zu den Ausnahmen zählen etwa die Monografien von Meer (1998); Hanna (2003); Hohenstein (2006); Brinkschulte (2015); Rahn (2022), Beiträge aus dem Kontext der Projekte EuroWiss (Redder/Heller/Thielmann (Hg.) 2014; Redder 2019) sowie GeWiss (vgl. etwa die Beiträge in Fandrych/Meißner/Wallner (Hg.) 2017), daneben vereinzelte weitere Studien (vgl. etwa Baßler 2007; Ventola 2007; Günthner/Zhu 2014).

<sup>11</sup> Bisher liegen sehr wenige didaktische Materialien vor, die explizit – und auf empirischer Basis – die Mündlichkeit thematisieren, vgl. aber Ylönen (2006) und Lange/Rahn (2017). Für eine neuere Übersicht vgl. auch Ylönen (2018).

nen Gattungen: „In general, stance is overtly marked to a greater extent in the spoken registers than the written registers“ (Biber 2006, S. 103).<sup>12</sup> Zu den hierfür einschlägigen Ausdrücken zählt Biber auch *stance complement clauses* wie *I believe / think ... that ...* (ebd.), deren deutsche Äquivalente im Weiteren im Mittelpunkt unserer Analyse stehen sollen. Für das Deutsche legen erste Analysen nahe, dass Sprecher- und Hörerbezug sowie sprecherseitige modalisierende Einordnungen des Gesagten auch in eher formellen, vorgeplanten kommunikativen Ereignissen wie wissenschaftlichen Vorträgen durchaus mit gewisser Frequenz auftreten.<sup>13</sup> Dabei scheinen sich auch spezifisch gesprochensprachliche konstruktionsartige Muster zu etablieren, die sich von vergleichbaren schriftlichen Mustern deutlich unterscheiden (vgl. Meißner 2017).

Wir wollen im Folgenden exemplarisch korpusbasiert die Verwendung von expliziten argumentativen Positionierungsstrukturen analysieren, mit denen die eigene Meinung bzw. Überzeugung in einem wissenschaftlichen Kontext verbalisiert wird.<sup>14</sup> Hierzu liegen empirische Studien zu schriftlichen argumentativen Gattungen vor, die versuchen, den Erwerb und die Verwendung solcher Konstruktionen über die schulische bis hin zur universitären Schreibpraxis nachzuzeichnen (vgl. zusammenfassend Gätje/Rezat/Steinhoff 2012). Sie zeigen, dass verbale Konstruktionen des Typs *ich finde + NP + Adj.* (*ich finde Schule spannend*) sowie *ich finde + Nebensatz/Hauptsatz* (*ich finde, dass Schule spannend ist / Schule ist spannend*) im Primarschulbereich vorherrschen, aber im Laufe der Schreibsozialisation zunehmend von nominalen Mustern des Typs *meines Erachtens, meiner Meinung/Auffassung nach* abgelöst werden. Dies wird mit der Beanspruchung von „intersubjektiver Geltung“ sowie domänenspezifischer Kontextualisierung begründet (Gätje/Rezat/Steinhoff 2012, S. 147): Der nominale Stil solcher Konstruktionen nimmt insgesamt das kommunikative

---

**12** Biber untersucht allerdings eine breite Palette von kommunikativen Ereignissen, von Seminarsgesprächen, studienorganisatorischen und -beratenden Gesprächen über Lehrmaterial bis hin zu Studienunterlagen.

**13** Dies wurde auf der Basis des GeWiss-Korpus etwa für Phänomene wie Diskursmarker, Modalpartikeln, Modaladverbien, agensorientierten Stil und Modalverberwendung bei Metakommentierungen gezeigt, vgl. dazu u. a. Fandrych (2014); Slavcheva/Meißner (2014); Wallner (2017). Man kann diese Befunde auch als Ausdruck von stärkerem sprecherseitigem *involvement* lesen, vgl. Barbieri (2015).

**14** Damit wird nur ein kleiner Ausschnitt von kommunikativen Positionierungspraktiken in den Blick genommen. Vgl. allgemein zu Ansätzen der Beschreibung von sozialen Positionierungen Deppermann (2015), zu Formen der sozialesemiotischen Positionierung in wissenschaftlichen Vorträgen im chinesisch-deutschen Vergleich Günthner/Zhu (2014).

Gewicht der Positionierung und den Bezug auf die Sprecher/-innen bzw. Autor/-innen zurück.

Wie allerdings in Fandrych (2021) explorativ gezeigt wurde, scheinen diese Befunde für mündliche Gattungen wie den wissenschaftlichen Vortrag nicht in gleicher Weise zu gelten: Hier sind Positionierungskonstruktionen mit verbalen Konstruktionen des Typs *ich denke, ich glaube, ich finde, ich meine, ich halte ... für* (vgl. die Belege 1–5), die von Gätje/Rezat/Steinhoff (2012) dem Schreibentwicklungsstand des schulischen Primarbereichs zugeordnet werden, durchaus einschlägig. Wie die folgenden Belege zeigen, treten sie häufig an argumentativen Schlüsselstellen auf; zudem sind sie deutlich frequenter als entsprechende nominale Varianten:<sup>15</sup>

- (1) dies widerspricht eindeutig der aussage sternefelds dass es im deutschen kein do support gäbe (0.5) und **ich denke** die indizien (0.4) sind eindeutig dass es (0.7) diesen (0.4) sehr wohl gibt (0.4)  
(GWSS\_E\_00032\_SE\_01\_T\_01)
- (2) und dass das langsame lesen °h äh (.) deshalb auch (al) texte bekommen muss die es wert sind langsam gelesen zu werden **ich finde** auch diese position von weinrich übrigens eine absolut unterschätzte position  
(GWSS\_E\_00021\_SE\_01\_T\_01)
- (3) **ich glaube** (.) **nicht** dass diese einschätzung grundsätzlich (0.2) falsch is aber **ich glaube** dass äh sozusagen (.) sie nur bedingt auf die (.) stilistik (0.5) zutrifft (GWSS\_E\_00028\_SE\_01\_T\_01)
- (4) **ich meine** dass °h unterschiede im korrektheits und anderen teilen im meiner untersuchung zu aspekten der mehrsprachigkeit °h nicht von vornherein auf laienexpertinnen [...] reduzierbar sind  
(GWSS\_E\_00024\_SE\_01\_T\_01)
- (5) wir haben versucht mit dieser modellbildung (0.7) [...] einfach be zu betonen dass **wir** die akteure **für** aktive protagonisten **halten** °hh (.) deren handeln nicht nicht einfach nur vom diskurs geprägt ist sondern die durch ihr handeln umgekehrt auch den diskurs prägen  
(GWSS\_E\_00028\_SE\_01\_T\_01)

---

<sup>15</sup> Gezählt wurden dabei sowohl Fälle, in denen die Belege im Vortrag selbst auftraten, als auch solche, in denen sie in der anschließenden Diskussion vorkamen. Die hier aufgeführten Belege sind nach den Konventionen des GAT2-Minimaltranskripts verschriftlicht worden, vgl. Selting et al. (2009).

Es scheint also so zu sein, dass die für die Schreibentwicklung konstatierten Konventionen zur Realisierung von expliziten argumentativen Positionierungskonstruktionen in der mündlichen Gattung wissenschaftlicher Vortrag nicht in gleicher Weise gelten, obwohl sowohl der wissenschaftliche Artikel bzw. die Monografie als auch der wissenschaftliche Vortrag als relativ formale, argumentativ eristisch angelegte und wissenschaftlichen Standards verpflichtete Gattungen angesehen werden können.<sup>16</sup>

Im Folgenden soll dies anhand von verschiedenen Positionierungskonstruktionen in Vorträgen und in vergleichbaren wissenschaftlichen Texten genauer untersucht werden. Hierfür wird die Häufigkeit von Positionierungskonstruktionen in den unterschiedlichen Textsorten miteinander verglichen, wobei bezüglich der Vorträge auch berücksichtigt wird, ob für die Sprecher/-innen Deutsch die Erstsprache (L1) oder Zweit- bzw. Fremdsprache (L2) darstellt. Damit sollte überprüft werden, ob sich anhand der untersuchten Phänomene möglicherweise Hinweise auf unterschiedliche Vortragskonventionen auffinden lassen. Zudem wird überprüft, inwieweit die Positionierungskonstruktionen in den einzelnen Textsorten verbal oder nominal realisiert werden und inwieweit es sich bei den Positionierungen um Eigen- und Fremdpositionierungen handelt.<sup>17</sup>

### 3.2.2 Korpusgrundlagen und Vorgehen

Als Datengrundlage für die Untersuchung wurden Expertenvorträge aus GeWiss und ein Teilkorpus mit Dissertationen aus dem Projekt GeSIG<sup>18</sup> herangezogen. Bezüglich der GeWiss-Daten wurden zwei Teilkorpora gebildet, um L1- und L2-Sprecher/-innen des Deutschen separat untersuchen zu können (GeWiss EV\_L1, GeWiss EV\_L2). Es wurden dabei lediglich die Sprechanteile der Vortragenden selbst berücksichtigt, nicht die Sprechanteile von Moderator/-innen sowie die

---

**16** Argumentative Positionierungen sind selbstverständlich nicht auf die hier untersuchten Ausdrücke des Meinens, Denkens und Glaubens beschränkt; eine umfassendere Untersuchung müsste auch andere Formen der Positionierung mit einbeziehen, wie sie etwa im *stance*-Begriff gefasst werden. Zu *stance* in der Wissenschaftskommunikation vgl. Gray/Biber (2012); zur Begründung eines soziolinguistisch-diskursorientierten *stance*-Begriffs vgl. Du Bois (2007).

**17** In den im Folgenden vorgestellten Analysen wird lediglich eine Auswahl an möglichen Untersuchungsperspektiven eingenommen, um das Potenzial der in ZuMult geschaffenen Tools für die Erforschung der mündlichen Hochschulkommunikation aufzuzeigen.

**18** Bei GeSIG handelt es sich um ein am Herder-Institut der Universität Leipzig erstelltes Korpus geisteswissenschaftlicher Dissertationen für Forschungszwecke, vgl. ausführlicher Meißner/Wallner (2019). Das Korpus ist nicht öffentlich zugänglich.

Diskussionsbeiträge von Zuhörenden, da für letztere keine Metadaten zur L1 vorliegen.

Eine Innovation des Tools ZuRecht besteht darin, dass für solche spezifisch zusammengestellten Teilkorpora auch Gesamttokenzahlen ermittelt werden können. Folgende Suchanfragen wurden hierfür genutzt:

Deutschsprachige Expertenvorträge mit Deutsch als L1 (ausschließlich Sprecherrolle Vortragende/r):

**((<word/> within <e\_se\_art="Expertenvortrag"/>) within <e\_se\_sprachen="Deutsch \ (L1)" />) within <ses\_rolle\_s="Vortragender.\*"/>**

Deutschsprachige Expertenvorträge mit Deutsch als L2 (ausschließlich Sprecherrolle Vortragende/r):

**((<word/> within <e\_se\_art="Expertenvortrag"/>) within <e\_se\_sprachen="Deutsch \ (L2)" />) within <ses\_rolle\_s="Vortragender.\*"/>**

Über die Metadatenansicht kann zudem die Anzahl der Vortragenden sowie die jeweils von den Vortragenden produzierten Token abgerufen werden (Abb. 6).

The screenshot shows the ZuRecht web interface. The search query is: `((<word/> within <e_se_art="Expertenvortrag"/>) within <e_se_sprachen="Deutsch \ (L1)" />) within <ses_rolle_s="Vortragender.*"/>`. The results table shows 7 entries with columns for document ID, OR ID, and text snippets. A 'Metadatenansicht' popup window is open, showing a table of metadata for 18 results, including 'S: Sprecher ID' and 'Treffler, Insgesamt: 75016'.

Metadaten Descriptor (Wählen Sie):	S: Sprecher ID	Treffler, Insgesamt: 75016
Sortiert nach (Wählen Sie):	Treffler (abs) absteigend	
Insgesamt: 18 S: Sprecher ID Treffler, Insgesamt: 75016		
1	GWSS_S_00047	6485 KWAC öffnen
2	GWSS_S_00046	6312 KWAC öffnen
3	GWSS_S_00048	5776 KWAC öffnen
4	GWSS_S_00051	5657 KWAC öffnen

Abb. 6: Ermittlung der Gesamttokenzahlen für ein Teilkorpus mit Hilfe von ZuRecht

Das Teilkorpus aus GeSIG umfasst insgesamt 11 Dissertationen aus dem Fachbereich Germanistik.<sup>19</sup> Die Ermittlung der Tokenzahlen erfolgte hier mithilfe der Sketch Engine.<sup>20</sup> Tabelle 2 gibt einen Überblick über die untersuchten Teilkorpora:

Tab. 2: Übersicht über die Teilkorpora

GeWiss EV_L1	GeWiss EV_L2	GeSIG
75.016 token	60.434 token	1.025.691 token
18 Vortragende	19 Vortragende	11 Autor/-innen

Zur Ermittlung potenzieller Kandidaten für Positionierungskonstruktionen wurden in den Teilkorpora alle Belege mit den Lemmata *denken, finden, glauben, halten ... für, meinen* sowie *Auffassung, Erachten, Meinung, Überzeugung* extrahiert. Für GeSIG fand auch die Schreibabkürzung *m. E.* Berücksichtigung.

Die Belegsuche in den GeWiss-Teilkorpora erfolgte ebenfalls mit ZuRecht, was eine sprecherbezogene, metadatensensible Lemmasuche gestattet. Hierfür wurde die folgende Suchanfrage genutzt:

```
(([lemma="(denken|finden|glauben|halten|meinen|Auffassung|Erachten|Meinung|Überzeugung)" within <e_se_art="Expertenvortrag"/>) within <e_se_sprachen="Deutsch (L1)"/>) within <ses_rolle_s="Vortragender.*"/>21
```

Über den Button „Treffer gruppieren“ (Abb. 6) kann eine Übersicht über die Verteilung der Treffer auf die einzelnen Lemmata aufgerufen werden (Abb. 7):

<sup>19</sup> Die Zuordnung zur „Germanistik“ erfolgte nach der Fachbereichsbestimmung des statistischen Bundesamtes, die Arbeiten stammen aus literatur- und sprachwissenschaftlichen Arbeiten zum Deutschen, dem Fach Deutsch als Fremdsprache, der Niederlandistik sowie der Skandinavistik (vgl. auch Meißner/Wallner 2019).

<sup>20</sup> [www.sketchengine.eu/](http://www.sketchengine.eu/) (Stand: 5.7.2022).

<sup>21</sup> Die Anfrage bezieht sich auf das GeWiss EV\_L1, für die Suche in GeWiss EV\_L2 muss „L1“ durch „L2“ ersetzt werden.

**Ergebnisse**  
für die Suchanfrage ((lemma=" (denken|finden|glauben|halten|meinen|auffassung|Erachten|Meinung|Überzeugung)') within <e\_se\_art="Expertenvortrag"/>) within <e\_se\_sprachen="Deutsch |(L1)"/>) within <ses\_rolle\_s="Vortragender."/> (in GWSS)

Gruppirt nach (Wählen Sie):

Sortiert nach (Wählen Sie):

Insgesamt: 8	Lemma	Treffer, Insgesamt: 341	
1	finden	94	KWIC öffnen
2	denken	83	KWIC öffnen
3	meinen	64	KWIC öffnen
4	glauben	57	KWIC öffnen
5	Meinung	25	KWIC öffnen
6	halten	7	KWIC öffnen
7	Auffassung	7	KWIC öffnen
8	Erachten	4	KWIC öffnen

**Ergebnisse**  
für die Suchanfrage ((lemma=" (denken|finden|glauben|halten|meinen|auffassung|Erachten|Meinung|Überzeugung)') within <e\_se\_art="Expertenvortrag"/>) within <e\_se\_sprachen="Deutsch |(L1)"/>) within <ses\_rolle\_s="Vortragender."/> (in GWSS)

Gruppirt nach (Wählen Sie):

Sortiert nach (Wählen Sie):

Insgesamt: 9	Lemma	Treffer, Insgesamt: 231	
1	finden	84	KWIC öffnen
2	glauben	52	KWIC öffnen
3	meinen	35	KWIC öffnen
4	denken	29	KWIC öffnen
5	Meinung	13	KWIC öffnen
6	halten	7	KWIC öffnen
7	Erachten	6	KWIC öffnen
8	Auffassung	4	KWIC öffnen
9	Überzeugung	1	KWIC öffnen

**Abb. 7:** Überblick über die Anzahl potenzieller Belegstellen für Positionierungen in GeWiss EV\_L1 und in GeWiss EV\_L2 sortiert nach den einzelnen Lemmata

Für die weitere Analyse wurden die Belege anschließend über den Button „Download KWIC“ exportiert. ZuRecht ermöglicht hier eine An- bzw. Abwahl von Metadaten sowie die Bestimmung der Kontextgröße. Zudem wird mit den heruntergeladenen Belegen über eine Beleg-URL eine Verknüpfung mit der DGD sowie mit dem Transkriptbrowser ZuViel bereitgestellt. Dies erlaubt bei der weiteren Analyse (bspw. in Excel) einen direkten Zugriff auf die Belegstellen und ihre Kontexte inklusive der jeweiligen Audios. Zudem können mit Hilfe dieser Beleg-URLs auch Analyseergebnisse veröffentlicht und replizierbar gemacht werden (vgl. auch 3.2.4).

Die schriftsprachlichen Belege aus dem GeSIG-Teilkorpus wurden mithilfe von Sketch Engine über eine wortartsensible Lemma-Suche ermittelt und als csv-Datei exportiert. Die Gesamtzahl der Belege mit den oben genannten Suchlemmata betrug 2374 (GeWiss EV\_L1: 341, GeWiss EV\_L2: 231, GeSIG: 1802).

Im nächsten Schritt wurden die Belegstellen in Excel manuell ausgewertet, um Verwendungsweisen auszuschließen, die nicht der argumentativen Positionierung dienen.<sup>22</sup> Hierunter fallen zum einen andere Bedeutungen (Polysemie; etwa *eine Belegstelle finden* oder *an eine transzendente Macht glauben*). Zum anderen aber wurden auch stark grammatikalisierte Verwendungen herausgefiltert, bei denen eine Diskursmarkerfunktion (realisiert durch *ich mein*, *ich denk*,

<sup>22</sup> An dieser Stelle danken wir ganz herzlich unserer wissenschaftlichen Hilfskraft Laura-Jane Schmengler für die Unterstützung bei der Aufbereitung und Analyse der Belegstellen.

*ich find*; vgl. Beleg (6)) oder eine deutlich modalisierende Funktion vorliegt (vgl. Beleg (7)):

- (6) zudem ham i elemente keinen eigenen semantischen gehalt <sup>9</sup>h und ich denke mal diese auflistung wird jetzt für sie ein wenig langweilig (GWSS\_E\_00032\_SE\_01\_T\_01)
- (7) doyé hat hier (0.33) neun bereiche sind es glaub ich (0.35) aufgeschrieben (GWSS\_E\_00025\_SE\_01\_T\_01)

Diese letzte Unterscheidung ist eine Tendenzentscheidung, die u. a. aufgrund von Substitutionstests (Ersetzung durch einen anderen Meinungs Ausdruck, mögliche Umwandlung in eine Matrixkonstruktion des Typs *ich bin der Meinung, dass ...*) sowie durch die Analyse der Intonationsverhältnisse in Verbindung mit syntaktischen Merkmalen (z. B. syntaktische Selbstständigkeit) und der morphologischen Verhältnisse getroffen wurde (z. B. reduziertes parenthetisches *glaub* als modalisierender Ausdruck).<sup>23</sup> Für diese qualitative Auswertung wurde bei Bedarf immer wieder auf die Belegstellen und die Audios im Transkriptbrowser ZuViel zurückgegriffen.

Die verbleibenden, als explizite argumentative Positionierungshandlungen kategorisierten Belege wurden zusätzlich daraufhin analysiert, ob sie der Eigen- oder Fremdpositionierung dienen, da sich hier interessante Tendenzen zeigten (siehe unten). Im Rahmen der vorliegenden Studie konnte keine weitere Analyse der realisierten syntaktisch-lexikalischen Muster bzw. Konstruktionstypen oder der argumentativen Funktion im Vortrag bzw. Text durchgeführt werden. Dies wäre aber nicht nur aus linguistischer, sondern auch aus didaktischer Sicht äußerst lohnenswert, da eine Durchsicht der Belege zeigt, dass sie stark musterhaft geprägt sind und nicht selten an rhetorisch und argumentativ herausgehobener Stelle auftreten (vgl. Fandrych 2021). All dies prädestiniert sie dazu, im Rahmen eines „konstruktionsdidaktischen Ansatzes“ (vgl. dazu ausführlich Amorocho Duran/Pfeiffer *angen.*) im studienbezogenen Sprachunterricht thematisiert zu werden.<sup>24</sup> Dies könnte ein wichtiger Baustein zur Ergänzung bestehen-

**23** Näheres zur Unterscheidung von grammatikalisierten und nicht-grammatikalisierten Formen der hier interessierenden Verben findet sich u. a. bei Günthner/Imo (2003); Auer/Günthner (2005); Imo (2016) und, spezieller für wissenschaftliche Vorträge, bei Fandrych (2021). Zu betonen ist, dass auch diskursmarkerähnliche und modalisierende Verwendungen zu *stance*-Ausdrücken gerechnet werden können – allerdings sollten in der vorliegenden Studie die meinungs-betonenden verbalen Positionierungshandlungen im Vordergrund der Untersuchung stehen, die sich auch mit nominalen Konstruktionen des Typs *meiner Meinung nach* vergleichen lassen.

**24** Zur Nutzung der Konstruktionsgrammatik im Sprachunterricht auch Herbst (2019).

der, stark schriftorientierter didaktischer Materialien zur sprachlichen Studierfähigkeit sein.

### 3.2.3 Auswertung und Analyse

Bei der Analyse wurden insgesamt 377 Belege für explizite Positionierungen in den drei Korpora identifiziert. Eine Betrachtung der relativen Frequenzen lässt erkennen, dass die Intuition, dass Positionierungen dieser Art in Vorträgen deutlich häufiger auftreten als in wissenschaftlichen Texten, zutrifft (Tab. 3):

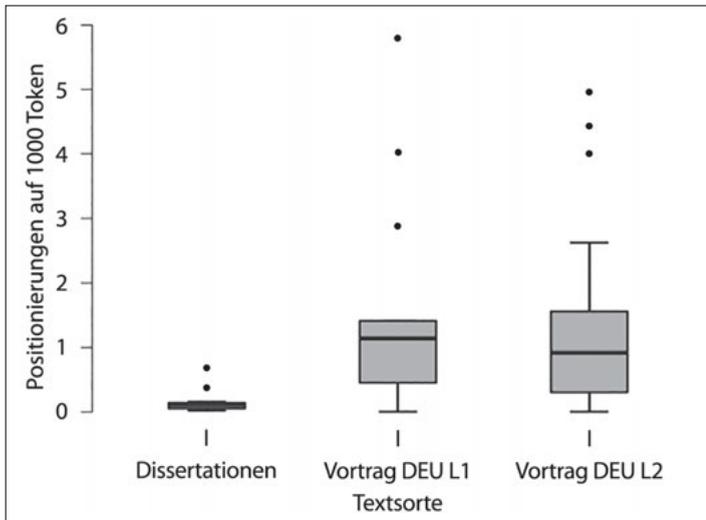
**Tab. 3:** Überblick über Anzahl der Belege für explizite Positionierungen

	GeWiss EV_L1	GeWiss EV_L2	GeSIG
Beleganzahl	120	89	168
<b>Beleganzahl/1000 Token</b>	<b>1,60</b>	<b>1,75</b>	<b>0,16</b>

Im Durchschnitt enthalten die einzelnen Vorträge mit Deutsch als L1 rund 1,5 explizite Positionierungen auf 1000 Token; bei den Vorträgen mit Deutsch als L2 fällt der Wert etwas geringer aus. Dabei ist anzumerken, dass in den L1-Vorträgen 17 von 18 Sprecher/-innen Positionierungen realisiert haben. Unter den 19 L2-Vorträgen gibt es wiederum drei, in denen keine Positionierungen vorkommen. Insgesamt fällt auf, dass die Anzahl der jeweils produzierten Positionierungen sehr unterschiedlich ausfällt, was sich auch in einer recht hohen Standardabweichung niederschlägt. Auch in den einzelnen Dissertationen ist eine unterschiedliche Anzahl an Positionierungen zu beobachten. Durchschnittlich fanden sich dort 0,163 Positionierungen auf 1000 Token (Tab. 4 und Abb. 8).

**Tab. 4:** Verteilung der Positionierungen/1000 Token auf die einzelnen Vorträge bzw. Dissertationen

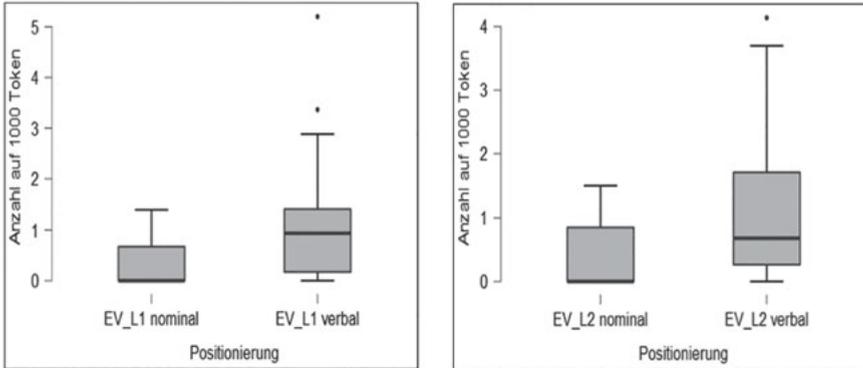
	Positionierungen auf 1000 Token		
	GeWiss EV_L1	GeWiss EV_L2	GeSIG
Mittelwert	1,546	1,367	0,163
Median	1,140	0,916	0,106
Standardabweichung	1,599	1,543	0,199



**Abb. 8:** Anzahl der Positionierungen/1000 Token in den einzelnen Korpora

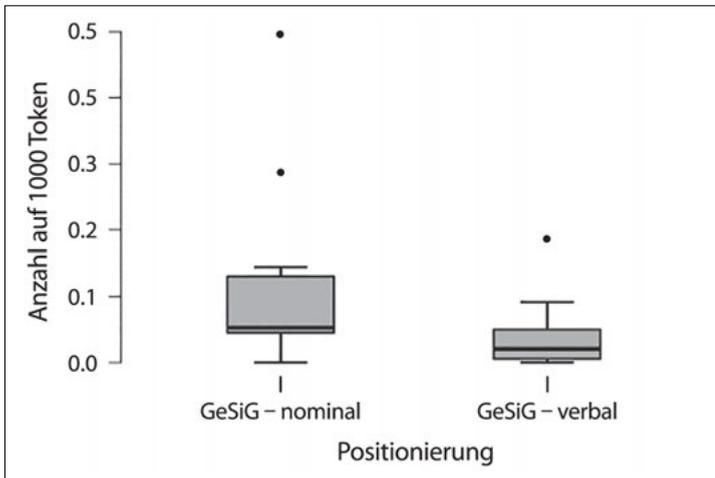
Die größere Streuung der Daten in den Vortragskorpora im Vergleich zu den Dissertationen ist aus didaktischer Perspektive interessant, da dies darauf hindeuten könnte, dass sich Vorträge durch höhere stilistisch-kommunikative Varianz auszeichnen als wissenschaftliche Texte – allerdings müsste diese Hypothese anhand von weiteren Merkmalen und größeren Korpora überprüft werden. Eine erste vorläufige statistische Analyse der Daten (Kruskal-Wallis-Test) zeigt jedoch bereits, dass sich die Korpora signifikant und mit großer Effektstärke voneinander unterscheiden ( $H(2) = 11,231$ ;  $p = 0,004$ ;  $\eta^2 = 0,205$ ). Auch die paarweisen Vergleiche zwischen den Dissertationen und den L1- bzw. L2-Vorträgen (Dunn-Bonferroni-Test mit angepassten p-Werten (Bonferroni-Korrektur)) ergeben jeweils signifikante Unterschiede mit hoher Effektstärke (GeSIG vs. GeWiss EV\_L1:  $z = -3,217$ ,  $p = 0,002$ , Cohen's  $d = -1,085$ ; GeSIG vs. GeWiss EV\_L2:  $z = -2,748$ ,  $p = 0,009$ , Cohen's  $d = -0,974$ ). Die L1- und L2-Vorträge unterscheiden sich bei schwacher Effektstärke nicht signifikant ( $z = 0,578$ ,  $p = 0,845$ , Cohen's  $d = 0,114$ ).

Eine wichtige Eingangsfrage war, wie sich verbale und nominale Realisierungen von Positionierungen verteilen, ob also verbale Realisierungen des Typs *ich meine/denke/glaube ...* in Vorträgen auch statistisch gesehen signifikant häufiger auftreten als entsprechende nominale Positionierungen des Typs *meiner Meinung nach*. Die folgende Abbildung zeigt die Ergebnisse der Gegenüberstellung dieser beiden Typen in den Vorträgen mit Deutsch als L1 und Deutsch als L2:



**Abb. 9:** Nominal vs. verbal realisierte Positionierungen in L1 und L2-Expertenvorträgen

Es bestätigt sich, dass verbal realisierte Positionierungen in Vorträgen deutlich häufiger auftreten als nominale Positionierungen. Für die L1-Vorträge ist das Ergebnis mit hoher Effektstärke statistisch signifikant ( $H(1) = 6,081$ ;  $p = 0,014$ ;  $\eta^2 = 0,15878$ ), für die L2-Vorträge ist der Unterschied bei mittelgradiger Effektstärke knapp nicht mehr statistisch signifikant ( $H(1) = 3,823$ ;  $p = 0,051$ ;  $\eta^2 = 0,094$ ). Betrachtet man im Vergleich die wissenschaftlichen Texte, so zeigt sich hier eine umgekehrte Tendenz (vgl. Abb. 10):



**Abb. 10:** Nominal vs. verbal realisierte Positionierungen in deutschen wissenschaftlichen Texten (GeSiG)

Die ohnehin deutlich weniger frequenten expliziten Positionierungen werden in den wissenschaftlichen Texten offensichtlich eher nominal formuliert. Allerdings ist der Unterschied zwischen nominal und verbal realisierten Positionierungen bei mittlerer Effektstärke statistisch nicht signifikant ( $H(1) = 3,281$ ;  $p = 0,070$ ;  $\eta^2 = 0,11405$ ).

Ein Blick auf die Frage, ob mithilfe der hier interessierenden Konstruktionen die eigene Position (Eigenpositionierung) oder die Position anderer Wissenschaftler/-innen verdeutlicht wird, hilft, die Unterschiede zwischen wissenschaftlichen Vorträgen und schriftlichen Texten in noch differenzierterer Weise zu verstehen. Tabelle 5 zeigt zunächst die Verteilung der verbalen Muster auf Eigen- und Fremdpositionierung in den Teilkorpora.

**Tab. 5:** Verteilung der verbalen Realisierungen auf Eigen- und Fremdpositionierungen

	GeWiss EV_L1	GeWiss EV_L2	GeSIG
Eigenpositionierung	92	62	1
Fremdpositionierung	6	7	41

Es zeigt sich, dass in Vorträgen die verbalen Realisierungen überwiegend zur Hervorhebung der eigenen wissenschaftlichen Position genutzt werden, während dies in wissenschaftlichen Texten praktisch nicht der Fall ist (nur eine Realisierung in einem Korpus von über 1 Mio Token). Interessant ist auch die Beobachtung, dass korpusübergreifend *meinen* für die Fremdpositionierung genutzt wird: In den beiden Vortragskorpora zusammengenommen tritt es in dieser Funktion in 10 der 13 Belege auf, im GeSIG-Korpus (wissenschaftliche Texte) finden sich 31 Fremdpositionierungsbelege mit *meinen* (meist formelhaft durch *X meint*), daneben 9 Belege mit *halten ... für (X hält ... für)*. Es scheint hier also eine lexikalische Spezialisierung zu geben.

Wir können die Schlussfolgerung ziehen, dass verbale Positionierungen in wissenschaftlichen Texten – im Gegensatz zu Vorträgen – erstens weniger frequent sind als nominale Positionierungen, und wo sie auftreten, sehr formelhaft der Fremdpositionierung dienen, wobei *meinen* das dominant auftretende Lexem ist.

Betrachtet man die nominalen Realisierungen bezüglich der Eigen- und Fremdpositionierung, so wird deutlich, dass hier das Verhältnis in den wissenschaftlichen Texten ausgeglichener ist (vgl. Tab. 6):

**Tab. 6:** Verteilung der nominalen Realisierungen auf Eigen- und Fremdpositionierungen

	GeWiss EV_L1	GeWiss EV_L2	GeSIG
Eigenpositionierung	18	17	56
Fremdpositionierung	4	3	70

Hier zeigt sich sowohl in den wissenschaftlichen Texten als auch in den Vorträgen eine starke Formelhaftigkeit. Eigenpositionierungen in wissenschaftlichen Texten wurden in 46 von 56 Fällen mit *meines Erachtens* bzw. *m. E.* verbalisiert, während Fremdpositionierungen v. a. mit den Lexemen *Auffassung* (39 Belege) bzw. *Meinung* (24 Belege) formuliert wurden, wobei unterschiedliche Muster auftraten (z. B. *X ist der Auffassung/Meinung; seiner Auffassung/Meinung nach; nach X' Auffassung/Meinung* etc.). In den Vorträgen wird für die Eigenpositionierung ebenfalls oft formelhaft *meines Erachtens* (11 Belege) genutzt, auch das Lexem *Meinung* ist bei der Eigenpositionierung relativ häufig vertreten (21 Belege). Für die Fremdpositionierung wird v. a. *Meinung* genutzt (6 Belege), nur einmal *Auffassung*.

Die bei den verbalen Mustern festgestellte lexikalische Spezialisierung von *meinen* auf Fremdpositionierungen findet hier ihr Pendant mit *Meinung*, zusätzlich wird nominal auch noch *Auffassung* relativ häufig für Fremdpositionierungen genutzt.

Was die Analyse der Daten sowohl für die verbalen als auch für die nominalen Positionierungen angeht, so zeigen sich zwischen den Expertenvorträgen von L1-Sprecher/-innen und L2-Sprecher/-innen auf den ersten Blick wenige Unterschiede. Eine feinkörnigere Analyse der lexikalischen Varianz sowie der auftretenden Formulierungsmuster wäre nötig, um evtl. bestehende Unterschiede aufzudecken. Auch wäre es interessant zu untersuchen, inwiefern sich Unterschiede zwischen Vortragenden aus dem britischen und dem polnischen Kontext ergeben – diese Daten sind ebenfalls über ZuRecht abrufbar, im Rahmen des vorliegenden Beitrags konnte eine feinkörnigere Untersuchung allerdings nicht durchgeführt werden.

### 3.2.4 Didaktische Schlussfolgerungen

Insgesamt bestätigt unsere Untersuchung Gätje/Rezat/Steinhoffs (2012) Beobachtung, dass in der geschriebenen Wissenschaftssprache argumentative Positionierungsprozeduren selten verbal ausgedrückt werden – mit einer wichtigen Einschränkung: Ganz offenbar können sie (v. a. *meinen* und *halten für*) durchaus

zur Fremdpositionierung verwendet werden. Ein domänenspezifischer Vorteil der Verwendung nominaler Varianten in der Wissenschaftskommunikation könnte darin bestehen, dass mit *Auffassungen, Meinungen, Überzeugungen* (aber nicht *Erachten*) ohne Bezug auf eine/-n Urheber/-in auf allgemein oder in einer bestimmten Domäne akzeptierte Konzepte, Annahmen und Theorien Bezug genommen werden kann. Diese Übergänge sieht man etwa bei *Auffassung* sehr deutlich: „*X's Auffassung*“ bzw. „*X ist der Auffassung, dass ...*“ steht an einem Ende eines Kontinuums, dessen anderes Ende von Formulierungen wie „*die traditionelle Auffassung*“ / „*die Auffassung zu X hat sich gewandelt*“, „*allgemein herrscht die Auffassung, dass ...*“ gebildet wird.

In der gesprochenen Wissenschaftssprache hingegen dienen die verbalen Positionierungen offenbar der variabel einsetzbaren Akzentuierung der eigenen Position vor einem Publikum. In dem interaktiven Kontext eines Vortrags stehen sie für das *involvement* (Barbieri 2015) des Sprechers/der Sprecherin mit Gegenstand und Publikum, das sich durch die verbale Realisierung unter Nutzung der Sprecherdeixis deutlicher manifestiert als über nominale Ausdrücke. Daneben erlauben verbale Formulierungen eine nuancierte Modalisierung der eigenen Positionierung – von emphatischer Assertion oder Ablehnung einer Position bis hin zu einer epistemischen Abschwächung, die wohl auch der Präsenz anderer Wissenschaftler/-innen geschuldet ist. Nicht zuletzt eignen sich verbale Muster – wie etwas ausführlicher in Fandrych (2021) gezeigt – auch in besonderem Maße, um an Schlüsselstellen eine „eristische Reliefgebung“ zu erzeugen, also die eigene Position gegenüber anderen wissenschaftlichen Positionen rhetorisch deutlich hervorzuheben (häufig unter Nutzung weiterer modalisierender und intensivierender sprachlicher Mittel, vgl. dazu die Belege 1–5 oben).

Diese Befunde sind aus Sicht einer Didaktik der Wissenschaftskommunikation von großer Relevanz – rezeptiv, für das Erkennen argumentativer Muster in Vorlesungen, Vorträgen oder Seminardiskussionen, aber auch produktiv bei der Planung eigener Vorträge, bei denen Studierende häufig vor dem Problem stehen, eine eigene Position zu entwickeln und diese in angemessener Weise zu verbalisieren (vgl. dazu auch Grzella/Plum 2018). Für eine solche Didaktik von argumentativen Formulierungskonstruktionen müssten in Folgestudien neben den lexikalischen sowie syntaktisch-morphologischen und intonatorischen Realisierungsmustern auch die jeweiligen Funktionen in den Vorträgen noch genauer herausgearbeitet werden (z. B. emphatisch-rhetorische Verstärkung, epistemische Abschwächung; vorsichtige Vermutung etc.). Dies könnte sodann die Basis für eine konstruktionsdidaktische Nutzung (Amarocho Duran/Pfeiffer angehen.) darstellen, die aus unserer Sicht gerade für diesen Gegenstand lohnenswert ist.

### 3.2.5 Nachnutzung der Ergebnisse

Die im Rahmen der vorgestellten Studie ermittelten Belegstellen aus dem GeWiss-Korpus können einerseits in didaktischen Kontexten und andererseits auch in Folgestudien von der Wissenschaftsgemeinschaft nachgenutzt werden. Dies ist möglich, da beim Download der Ergebnisse der Konkordanzsuche in ZuRecht – wie unter 3.2.2 bereits angesprochen – für jede Belegstelle eine Beleg-URL bereitgestellt wird. Die Beleg-URLs für die hier ermittelten Positionierungen führen in den Transkriptbrowser ZuViel. Die jeweilige Belegstelle wird in ZuViel durch einen roten Rahmen hervorgehoben. Es ist vorgesehen, die Beleg-URLs zu den Positionierungen in GeWiss auf dem Blog zum Projekt ZuMult unter <https://zumult.org/> (Stand: 5.7.2022) zu veröffentlichen. Neben einer Linksammlung für die einzelnen Belegstellen sollen auch Sammel-URLs zur Verfügung gestellt werden. Diese ermöglichen es, alle innerhalb eines Vortrags vorkommenden Belegstellen mit einem Klick aufzurufen (Abb. 11):

Ref wordlist	None	0322 (1.69)
A1-Niveau	1	0323 LS_0237 das is alles öhm h*
A1-Außerungen	1	0324 (0.59)
A2	1	0325 LS_0237 ja sehr relativ also ich bin (.) *h nicht ganz einverstanden mit diesen (0.34) kompetenzniveaus die (0.66) herr meißner da (0.47) gefunden hat
ab	2	0326 nn ((unverständlich))
aber	37	0327 (0.51)
Abitur	1	0328 LS_0237 ich glaube da (.) kann man ah (.) durchaus inzwischen
Abiturient	1	0329 (0.57)
abschließen	1	0330 LS_0237 öh in einzelnen kompetenzen andere ergebnisse (0.29) vorweisen (0.43) nach vier jahren sprachenunterricht sagt er a zwei bei
Abschlussarbeit	1	0331 (0.45)
Abschlussklausur	1	0332 LS_0237 vielleicht allen schülern
Abstract	1	0333 (0.46)
abtesten	1	0334 LS_0237 ah im lesen und hören aber bei sprechen und schreiben setzt: er schon fragezeichen (0.4) und (.) was das lesen angeht (0.43) erreichen einige (0.27) beins also wie gesagt ich denke dass das (0.26) inzwischen (0.39) doch durch (0.35) auch innovationen im fremdsprachenunterricht sich n bisschen geändert hat *h was wir aber belegen können
ach	3	0335 nn *hhh hhh*
acht	1	
achten	1	
acqua	1	

**Abb. 11:** Hervorhebung von Belegstellen für Positionierungen in einem Expertenvortrag (GWSS\_E\_00025\_SE\_01) im Transkriptbrowser ZuViel

Mit der Bereitstellung der Beleg-URLs eröffnet sich auch die Möglichkeit für eine dezentrale Forschungszusammenarbeit bei der Erschließung von mündlichen Korpora wie GeWiss, indem Fragestellung, Begründung der Suchabfrage und

praktische Schritte zur Korpusuche dokumentiert werden und sodann zusammen mit den Suchabfrageergebnissen (einschließlich der gebildeten Kategorien und den jeweiligen Beleg-URLs) veröffentlicht werden. So werden Forschungsergebnisse langfristig überprüfbar und können die Basis für weiterführende Fragestellungen und Untersuchungen bilden.

## 4 Fazit und Perspektiven

Das GeWiss-Korpus umfasst naturalistische mündliche Sprachdaten aus der Wissenschaftskommunikation. Es wurde nach relativ einheitlichen Kriterien aufgebaut und erlaubt so vielfältige Untersuchungsmöglichkeiten und Vergleichsperspektiven der gesprochenen Wissenschaftssprache. Das Korpus umfasst monologische wie auch dialogische kommunikative Ereignisse aus verschiedenen akademischen Kontexten (Deutschland, Polen, Großbritannien, Bulgarien, Italien, Finnland). Es enthält L1- und L2-Daten von Studierenden und Expertinnen/Experten sowie deutsche, polnische, englische und italienische Teilkorpora. So ermöglicht GeWiss genrebezogene sprachvergleichende Untersuchungen und Vergleiche zwischen Sprecher/-innen mit unterschiedlichem Expertenstatus. Es enthält somit Lernerdaten – nämlich studentische Vorträge und Prüfungsgespräche – sowohl von L1- als auch von L2-Sprecher/-innen des Deutschen. So lassen sich auch (fortgeschrittene) lernersprachliche Phänomene untersuchen, wenn gleich es hier eine Reihe von konzeptuellen und methodischen Faktoren zu berücksichtigen gilt (vgl. ausführlicher dazu Fandrych/Wallner 2022). Mit ZuMult werden die Zugänge zum Korpus und die Forschungs- und Nutzungsmöglichkeiten von GeWiss deutlich verbessert. Insbesondere stehen für die didaktische Nutzung über ZuMal vielfältige neue Filtermöglichkeiten auch für sprachdidaktische Zwecke bereit, es können gezielt Sprechereignisse ausgewählt werden und dann über ZuViel nach verschiedenen didaktischen Vorgaben visualisiert oder auch mit Wortschatzlisten abgeglichen werden. ZuRecht gestattet die gezielte Recherche von sprachlichen Phänomenen für Forschungs- und Vermittlungszwecke, die letztlich eine Community-orientierte kooperativ-kritische Forschungsperspektive ermöglicht. Aus didaktischer Perspektive verhindert die rechtliche Einschränkung der Datennutzung von GeWiss sowie der anderen Korpora, die in der DGD enthalten sind, allerdings eine breitere Anwendung außerhalb des universitären Kontexts. Hier wäre für die Zukunft dringend nach Möglichkeiten einer Ausweitung der Nutzungsmöglichkeiten zu suchen.

## Literatur

- Amorocho Duran, Simone/Pfeiffer, Christian (angen.): Konstruktionsdidaktik – Grundzüge einer sprachdidaktischen Konzeption. In: Deutsch als Fremdsprache.
- Auer, Peter/Günthner, Susanne (2005): Die Entstehung von Diskursmarkern im Deutschen – ein Fall von Grammatikalisierung? In: Leuschner, Torsten/Mortelmans, Tanja/De Groot, Sarah (Hg.): Grammatikalisierung im Deutschen. (= Linguistik – Impulse & Tendenzen 9). Berlin/ New York: De Gruyter, S. 335–362.
- Auer, Peter/Baßler, Harald (Hg.) (2007): Reden und Schreiben in der Wissenschaft. Frankfurt a. M.: Campus.
- Barbieri, Frederica (2015): Involvement in university classroom discourse: register variation and interactivity. In: Applied Linguistics 36, 2, S. 151–173.
- Baßler, Harald (2007): Diskussionen nach Vorträgen bei wissenschaftlichen Tagungen. In: Auer/Baßler (Hg.), S. 133–155.
- Biber, Douglas (2006): Stance in spoken and written university registers. In: Journal of English for Academic Purposes 5, 2, S. 97–116.
- Brinkschulte, Melanie (2015): (Multi-)mediale Wissensübermittlung in Vorlesungen. Diskursanalytische Untersuchungen zur Wissensübermittlung am Beispiel der Wirtschaftswissenschaft. (= Wissenschaftskommunikation 11). Heidelberg: Synchron.
- Deppermann, Arnulf (2015): Positioning. In: De Fina, Anna/Georgakopoulou, Alexandra (Hg.): The handbook of narrative analysis. (= Blackwell Handbooks in Linguistics). Chichester: Wiley-Blackwell, S. 369–386.
- Dietz, Gunther (2021): Korpora gesprochener Sprache als Quelle für die Erstellung von Mikro-Hörübungen mit authentischen Hörmaterialien im Daz-/Daf-Unterricht. In: Korpora Deutsch als Fremdsprache 1/2021, S. 97–123. <https://doi.org/10.48694/tujournals-41> (Stand: 5.7.2022).
- Du Bois, John W. (2007): The stance triangle. In: Englebretson, Robert (Hg.): Stancetaking in discourse. Subjectivity, evaluation, interaction. (= Pragmatics & beyond, New Series 164). Amsterdam u. a.: Benjamins, S. 139–182.
- Emmrich, Volker (2019): Kontroversen darstellen: Kontrastieren und Positionieren. In: Steinseifer/Feilke/Lehnen (Hg.), S. 209–241.
- Fandrych, Christian (2014): Metakomentierungen in wissenschaftlichen Vorträgen. In: Fandrych/Meißner/Slavcheva (Hg.), S. 95–111.
- Fandrych, Christian (2021): *Ich denke, die Indizien sind eindeutig ...*: Positionierungshandlungen als spezifisch mündliche Phänomene in wissenschaftlichen Vorträgen. In: Günthner, Susanne/Schopf, Juliane/Weidner, Beate (Hg.): Gesprochene Sprache in der kommunikativen Praxis. Analysen authentischer Alltagssprache und ihr Einsatz im DaF-Unterricht. (= Stauffenburg Deutschdidaktik). Tübingen: Stauffenburg, S. 219–246.
- Fandrych, Christian/Tallowitz, Ulrike (2019): Sage und Schreibe: Übungswortschatz Grundstufe A1–B1 mit Lösungen. Neubearbeitung mit Audio-CD. Stuttgart: Klett.
- 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 (ZGL) 50, 1, S. 202–239.

- Fandrych, Christian/Meißner, Cordula/Wallner, Franziska (2018): Das Potenzial mündlicher Korpora für die Sprachdidaktik. Das Beispiel GeWiss. In: *Deutsch als Fremdsprache* 1/2018, S. 3–14.
- Fandrych, Christian/Meißner, Cordula/Wallner, Franziska (2021): Korpora Gesprochener Sprache und Deutsch als Fremd- und Zweitsprache: Eine chancenreiche Beziehung. In: *Korpora Deutsch als Fremdsprache* 2/2021, S. 5–30. <https://kordaf.tu-journals.ulb.tu-darmstadt.de/> (Stand: 5.7.2022).
- 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: *InfoDaF* 50, 6, S. 711–729.
- Fandrych, Christian/Frick, Elena/Hedeland, Hanna/Iliash, Anna/Jettko, Daniel/Meißner, Cordula/Schmidt, Thomas/Wallner, Franziska/Weigert, Kathrin/Westpfahl, Swantje (2016): User, who art thou? User profiling for oral corpus platforms. In: Calzolari, Nicoletta/Choukri, Khalid/Declerck, Thierry/Goggi, Sara/Grobelnik, Marko/Maegaard, Bente/Mariani, Joseph/Mazo, Helene/Moreno, Asuncion/Odijk, Jan/Piperidis, Stelios (Hg.): *Proceedings of the tenth international conference on language resources and evaluation (LREC 2016)*, Portorož, Slovenia. Paris: European Language Resources Association (ELRA), S. 280–287.
- Fandrych, Christian/Meißner, Cordula/Slavcheva, Adriana (Hg.) (2014): *Gesprochene Wissenschaftssprache: Korpusmethodische Fragen und empirische Analysen.* (= *Wissenschaftskommunikation* 9). Heidelberg: Synchron.
- Fandrych, Christian/Meißner, Cordula/Wallner, Franziska (Hg.) (2017): *Gesprochene Wissenschaftssprache – digital. Verfahren zur Annotation und Analyse mündlicher Korpora.* (= *Deutsch als Fremd- und Zweitsprache Schriften des Herder-Instituts (SHI)*). Tübingen: Stauffenburg.
- Feilke, Helmuth (2012): Was sind Textroutinen? Zur Theorie und Methodik des Forschungsfelds. In: Feilke/Lehnen (Hg.), S. 1–33.
- Feilke, Helmuth/Lehnen, Katrin (Hg.) (2012): *Schreib- und Textroutinen. Erwerb, Förderung und didaktisch-mediale Modellierung.* (= *FORUM ANGEWANDTE LINGUISTIK* 52). Frankfurt a. M. u. a.: Lang.
- Feilke, Helmuth/Lehnen, Katrin/Steinseifer, Martin (2019): Eristische Literalität – Theorie und Parameter einer Kompetenz. In: Steinseifer/Feilke/Lehnen (Hg.), S. 11–33.
- Frick, Elena/Helmer, Henrike/Wallner, Franziska (angen.): *ZuRecht: Neue Recherchemöglichkeiten in Korpora gesprochener Sprache für Gesprächsanalyse und Deutsch als Fremd- und Zweitsprache.* In: *KorDaF, Themenheft „Zugänge zu multimodalen Korpora gesprochener Sprache“.*
- Gätje, Olaf/Rezat, Sara/Steinhoff, Torsten (2012): Positionierung. Zur Entwicklung des Gebrauchs modalisierender Prozeduren in argumentativen Texten von Schülern und Studenten. In: Feilke/Lehnen (Hg.), S. 125–153.
- Graefen, Gabriele/Moll, Melanie (2011): *Wissenschaftssprache Deutsch: lesen – verstehen – schreiben. Ein Lehr- und Arbeitsbuch.* Frankfurt a. M.: Lang.
- Gray, Bethany/Biber, Douglas (2012): Current conceptions of stance. In: Hyland, Ken (Hg.): *Stance and voice in written academic genres.* Basingstoke: Palgrave Macmillan, S. 15–33.
- Grzella, Markus/Plum, Sabine (2018): Auf fremdem Terrain? Wissenschaftssprachliche Ausdrucksformen in studentischen Referaten. In: Albert, Georg/Diao-Klaeger, Sabine

- (Hg.): Mündlicher Sprachgebrauch. Zwischen Normorientierung und pragmatischen Spielräumen. (= Stauffenburg Linguistik 101). Tübingen: Stauffenburg, S. 25–42.
- Günthner, Susanne/Imo, Wolfgang (2003): Die Reanalyse von Matrixsätzen als Diskursmarker: *ich mein*-Konstruktionen im gesprochenen Deutsch. In: *Interaction and Linguistic Structures* 37, S. 1–31.
- Günthner, Susanne/Zhu, Qiang (2014): Wissenschaftsgattungen im Kulturvergleich – Analysen von Eröffnungssequenzen chinesischer und deutscher Konferenzvorträge. In: Meier, Simon/Rellstab, Daniel/Schiewer, Gesine (Hg.): *Dialog und (Inter-)Kulturalität. Theorien, Konzepte, empirische Befunde*. Tübingen: Narr, S. 175–196.
- Hanna, Ortrun (2003): Wissensvermittlung durch Sprache und Bild. Sprachliche Strukturen in der ingenieurwissenschaftlichen Hochschulkommunikation. (= *Arbeiten zur Sprachanalyse* 42). Frankfurt a. M. u. a.: Lang.
- Herbst, Thomas (2019): Über Kognition zur Konstruktion: zielorientiertes Lernen fremdsprachlicher Konstruktionen von links nach rechts. In: Erfurt, Jürgen/De Knop, Sabine (Hg.): *Konstruktionsgrammatik und Mehrsprachigkeit*. (= OBST 94). Duisburg: Universitätsverlag Rhein-Ruhr, S. 149–172.
- Hohenstein, Christiane (2006): Erklärendes Handeln im Wissenschaftlichen Vortrag. (= *Studien Deutsch* 36). München: ludicum.
- Imo, Wolfgang (2016): *Ich finde, mit Matrixsätzen kann man eine Menge machen... Von der Redeanführung über den Matrixsatz zum Diskursmarker*. In: Handwerker, Brigitte/Bäuerle, Rainer/Sieberg, Bernd (Hg.): *Gesprochene Fremdsprache Deutsch*. (= *Perspektiven Deutsch als Fremdsprache* 32). Baltmannsweiler: Schneider, S. 45–74.
- Lange, Daisy/Rahn, Stefan (2017): *Mündliche Wissenschaftssprache. Kommunizieren – Präsentieren – Diskutieren*. Lehr- und Arbeitsbuch. (= *Deutsch für das Studium*). Stuttgart: Ernst Klett Sprachen.
- Meer, Dorothee (1998): „Der Prüfer ist nicht der König“. Mündliche Abschlußprüfungen in der Hochschule. (= *Germanistische Linguistik* 20). Tübingen: Niemeyer.
- Meißner, Cordula (2017): Gute Kandidaten. Ein Ansatz zur automatischen Ermittlung von Belegen für sprachliche Handlungen auf der Basis manueller pragmatischer Annotation. In: Fandrych/Meißner/Wallner (Hg.), S. 169–218.
- Meißner, Cordula/Wallner, Franziska (2019): *Das gemeinsame sprachliche Inventar der Geisteswissenschaften. Lexikalische Grundlagen für die wissenschaftspropädeutische Sprachvermittlung*. (= *Studien Deutsch als Fremd- und Zweitsprache* 6). Berlin: ESV.
- Meißner, Cordula/Wallner, Franziska (2022): Korpora gesprochener Sprache als virtuelle Lernräume der Mündlichkeitsdidaktik: Affordanzen eines außerunterrichtlichen Sprachlernsettings. In: Feick, Diana/Rymarczyk, Jutta (Hg.): *Zur Digitalisierung von Lernorten – Fremdsprachenlernen im virtuellen Raum*. (= *Inquiries in language learning* 34). Bern: Lang, S. 215–239.
- Moll, Melanie/Thielmann, Winfried (2017): *Wissenschaftliches Deutsch. Wie es geht und worauf es dabei ankommt*. (= utb 4650). Konstanz: UVK- Verlagsgesellschaft.
- Rahn, Stefan (2022): *Universitäre Prüfungsgespräche mit deutschen und internationalen Studierenden. Eine diskursanalytische Studie aus der Perspektive von Deutsch als Fremdsprache*. (= *Deutsch als Fremd- und Zweitsprache Schriften des Herder-Instituts (SHI)*). Tübingen: Stauffenburg.
- Redder, Angelika (2019): *Diskursive und textuelle Eristik – Systematik und komparative Analysen*. In: Steinseifer/Feilke/Lehnen (Hg.), S. 35–64.

- Redder, Angelika/Heller, Dorothee/Thielmann, Winfried (Hg.) (2014): Eristische Strukturen in Vorlesungen und Seminaren deutscher und italienischer Universitäten. Analysen und Transkripte. Heidelberg: Synchron.
- Reershemius, Gertrud/Lange, Daisy (2014): Sprachkontakt in der mündlichen Wissenschaftskommunikation. In: Fandrych/Meißner/Slavcheva (Hg.), S. 57–74.
- Sadowski, Sabrina (2017): Die Annotation von Zitaten und Verweisen im GeWiss-Korpus. In: Fandrych/Meißner/Wallner (Hg.), S. 147–166.
- Selting, Margret/Auer, Peter/Barth-Weingarten, Dagmar/Bergmann, Jörg/Bergmann, Pia/Birkner, Karin/Couper-Kuhlen, Elizabeth/Deppermann, Arnulf/Gilles, Peter/Günthner, Susanne/Hartung, Martin/Kern, Friederike/Mertzluff, Christine/Meyer, Christian/Morek, Miriam/Oberzaucher, Frank/Peters, Jörg/Quasthoff, Uta/Schütte, Wilfried/Stukenbrock, Anja/Uhmann, Susanne (2009): Gesprächsanalytisches Transkriptionssystem 2 (GAT 2). In: Gesprächsforschung – Online-Zeitschrift zur verbalen Interaktion 10, S. 353–402. <http://www.gespraechsforschung-ozs.de/heft2009/px-gat2.pdf> (Stand: 4.10.2022).
- Slavcheva, Adriana/Meißner, Cordula (2014): *Also* und *so* in wissenschaftlichen Vorträgen. In: Fandrych/Meißner/Slavcheva (Hg.), S. 113–132.
- Steinhoff, Torsten (2007): Wissenschaftliche Textkompetenz. Sprachgebrauch und Schreibentwicklung in wissenschaftlichen Texten von Studenten und Experten. (= Reihe germanistische Linguistik 280). Tübingen: Niemeyer.
- Steinseifer, Martin/Feilke, Helmuth/Lehnen, Katrin (Hg.) (2019): Eristische Literalität. Wissenschaftlich streiten – wissenschaftlich schreiben. (= Wissenschaftskommunikation 13). Heidelberg: Synchron.
- Tschirner, Erwin (2019): Der rezeptive Wortschatzbedarf in Deutschen als Fremdsprache. In: Studer, Thomas/Thonhauser, Ingo/Peyer, Elisabeth (Hg.): IDT 2017. Brücken gestalten – mit deutsch verbinden: Menschen – Lebenswelten – Kulturen. Beiträge der XVI. Internationalen Tagung der Deutschlehrerinnen und Deutschlehrer. Fribourg/Freiburg, 31. Juli–4. August 2017. Bd. 1: Hauptvorträge. Berlin: ESV, S. 98–111.
- Ventola, Eija (2007): Konferenzvorträge: Sprechen englisch muttersprachige Konferenzteilnehmer wirklich anders? In: Auer/Baßler (Hg.), S. 115–132.
- Wallner, Franziska (2017): Diskursmarker funktional: Eine quantitativ-qualitative Beschreibung annotierter Diskursmarker im GeWiss-Korpus. In: Fandrych/Meißner/Wallner (Hg.), S. 110–125.
- Ylönen, Sabine (2006): Training wissenschaftlicher Kommunikation mit E-Materialien. Beispiel mündliche Hochschulprüfung. In: Ehlich, Konrad/Heller, Dorothee (Hg.): Die Wissenschaft und ihre Sprachen. (= Linguistic Insights – Studies in Language and Communication 52). Frankfurt a. M.: Lang, S. 115–146.
- Ylönen, Sabine (2018): Oral discourse in scientific research. In: Budin, Gerhard/Laurén, Christer/Humbley, John (Hg.): Language for special purposes. An international handbook. (= De Gruyter Reference). Berlin/Boston: De Gruyter, S. 364–380.