



Handbuch zum Korpus

März 2014

Karen Gräfe, Daisy Lange,

Magda Sieradz, Cordula Meißner, Adriana Slavcheva

Inhalt

Verzeichnis der Abkürzungen	3
1. Konzeption des GeWiss-Korpus.....	4
2. Daten im GeWiss-Korpus	6
2.1 Audioaufnahmen	6
2.2 Zusatzmaterialien (n.v.)	7
2.3 Metadaten	7
2.3.1 Das Metadatenset des GeWiss-Korpus.....	7
2.3.2 Erläuterungen zu den Genres im deutschen, britischen, polnischen, italienischen und bulgarischen akademischen Kontext.....	10
2.3.2.1 Expertenvorträge (EV).....	10
2.3.2.2 Studentische Vorträge (SV).....	12
2.3.2.3 Prüfungsgespräche (PG).....	15
2.4 Transkriptionen.....	18
2.4.1 Transkriptionssoftware und Darstellung	18
2.4.1.1 Spurentypen.....	19
2.4.1.2 Spurennamen	20
2.4.2 Transkriptionskonventionen.....	21
2.4.2.1 GAT 2-Minimaltranskript.....	21
2.4.2.2 Anpassungen	22
2.5 Annotationen.....	24
3. Zugriffsmöglichkeiten auf das GeWiss-Korpus.....	26
3.1 Registrieren und Anmelden im GeWiss-Korpus	26
3.2 Anmelden (Login) im GeWiss-Korpus.....	27
3.3 Zugang zu den Volltexten (Browse)	28
3.3.1 Darstellung und Zugriff.....	28
3.3.2 Metadaten einer Kommunikation einsehen	30
3.3.3 Volltext einer Kommunikation einsehen	31
3.4 Zugang zur Konkordanzsuche (Search)	33
3.4.1 Konkordanzen erstellen.....	33
3.4.2 Effiziente Suchen formulieren – Reguläre Ausdrücke nutzen	35
3.4.3 Metadaten hinzufügen	36
3.4.4 Filter-Optionen	36
3.4.5 Konkordanzen speichern.....	37
3.5 Webservices	37
4. Literatur.....	41
Anhang 1	43
Anhang 2.....	52
Anhang 3.....	54
Anhang 4.....	56

Verzeichnis der Abkürzungen

DEU = Deutsch

ENG = Englisch

EV = Expertenvortrag

EXAKT = EXMARaLDA Analyse- und Konkordanztool

EXMARaLDA = Extensible Markup Language for Discourse Annotation

GAT = Gesprächsanalytisches Transkriptionssystem

ISO = International Organization for Standardization

k.A. = keine Angabe

L1 = Erstsprache

L2 = Zweitsprache; *hier*: alle erworbenen Sprachen nach der Erstsprache

n/a = keine Angabe

n.v. = nicht verfügbar

OHP = Overheadprojektor

PG = Prüfungsgespräch

POL = Polnisch

SV = Studentischer Vortrag

TDN = TestDaF-Niveaustufen

1. Konzeption des GeWiss-Korpus

Das GeWiss-Korpus entstand im Rahmen des von der VolkswagenStiftung geförderten internationalen Forschungsprojektes GeWiss – Gesprochene Wissenschaftssprache kontrastiv: Deutsch im Vergleich zum Englischen und Polnischen. Vor dem Hintergrund einer steigenden Mobilität von Wissenschaftler(inne)n, Studierenden und Dozent(inn)en und dem damit verbundenen Qualifizierungsbedarf war es ein Hauptziel des Projekts, eine empirische Grundlage für die Beschreibung der wissenschaftssprachlichen Voraussetzungen im Bereich der Mündlichkeit zu schaffen, die Nicht-Muttersprachler(inne)n für einen Forschungs-/ Lehr- oder Studienaufenthalt in einem anderen europäischen Land benötigen. Im GeWiss-Projekt schlossen sich Partner aus Deutschland, Polen und Großbritannien zusammen, um gemeinsam ein Vergleichskorpus zur gesprochenen Wissenschaftssprache zu erstellen.

Das GeWiss-Korpus ist ein Vergleichskorpus zur gesprochenen Wissenschaftssprache. Es enthält Aufnahmen von akademischen Genres in den Sprachen Deutsch, Englisch, Polnisch und Italienisch. GeWiss umfasst dabei sowohl ein monologisches Genre (den wissenschaftlichen Vortrag) als auch dialogische Daten (Prüfungsgespräche sowie die sich an die Vorträge anschließenden Diskussionen). Die Aufnahmen dieser Genres in den drei Sprachen wurden in den jeweiligen akademischen Kontexten Deutschlands, Großbritanniens und Polens gemacht. Die Aufnahmen stammen aus philologischen Fächern und umfassen die Themenbereiche Linguistik, Literatur/ Kultur oder Didaktik.

Die dem GeWiss-Korpus zugrunde liegende Perspektive der Sprachvermittlung kommt in zwei Dimensionen zum Tragen. Einerseits enthält das Korpus neben Daten von Muttersprachler(inne)n auch Aufnahmen von Nicht-Muttersprachler(inne)n des Deutschen und Englischen, welche sowohl im deutschen als auch im britischen, polnischen und bulgarischen akademischen Kontext erhoben wurden. Zum anderen wurden Aufnahmen von Sprecher(inne)n mit unterschiedlichem akademischem Professionalisierungsgrad einbezogen. So enthält das Subkorpus mit studentischen Vorträgen Daten von Noviz(inn)en, das Subkorpus mit wissenschaftlichen Konferenzvorträgen Daten von Expert(inn)en des Wissenschaftsbetriebs. Die folgende Tabelle stellt den Aufbau des GeWiss-Korpus zusammenfassend dar.¹

Akademischer Kontext	Sprache	Deutsch		Britisch		Polnisch		Total	
		Deutsch L1	Deutsch L2	Englisch L1	Englisch L2	Deutsch L2	Polnisch L1		Deutsch L2
Expertenvortrag		10:03 h	-	5:06 h	2:47 h	5:21 h	4:53 h	4:54 h	33:04 h
Studentischer Vortrag		8:01 h	8:20 h	2:25 h	2:36 h	5:02 h	4:55 h	4:58 h	36:17 h
Prüfungsgespräch		12:07 h	9:01 h	3:32 h	6:57 h	10:22 h	9:59 h	10:02 h	62:00 h
Total		30:11 h	17:21 h	11:03 h	12:20 h	20:45 h	19:47 h	19:54 h	131:21 h

Tab. 1 Kernkorpusgröße in Stunden

¹ Zum Hintergrund und den Zielen des GeWiss-Projekts vgl. ausführlich Fandrych/Meißner/Slavcheva (2012).

Akademischer Kontext	Bulgarisch	Italienisch	Total
Sprache	Deutsch	Italienisch	
Genre	L2	L1	
Expertenvortrag		10:17 h	10:17 h
Studentischer Vortrag	05:22 h		05:22 h
Total	05:22 h	10:17 h	15: 39 h

Tab. 2 Größe der assoziierten Korpora (seit Oktober 2013)

Das Korpus umfasst somit die folgenden Vergleichsdimensionen:

- Wissenschaftssprache (Deutsch, Englisch, Polnisch, Italienisch)
- Sprachkompetenz (L1, L2)
- Grad akademischer Professionalisierung (Novizen, Experten)
- Institutioneller akademischer Kontext (deutsch, britisch, polnisch, italienisch und bulgarisch)

Im Folgenden wird nun erläutert, wie diese Konzeption umgesetzt und für den/die Nutzer(in) im Korpus zugänglich ist.

2. Daten im GeWiss-Korpus

Das GeWiss-Korpus enthält vier verschiedene Arten von Daten. Dies sind zum einen die Primärdaten in Form von

- (1) Audioaufnahmen ([2.1](#)) und
- (2) den dazugehörigen Zusatzmaterialien (n.v.) ([2.2](#)).

Zum anderen gehören zum GeWiss-Korpus solche Daten, die die Primärdaten beschreiben bzw. zur Nutzung aufbereiten:

- (3) Metadaten ([2.3](#)) und
- (4) Transkriptionen ([2.4](#)) mit Annotationen ([2.5](#)).

2.1 Audioaufnahmen

Alle Kommunikationen im GeWiss-Korpus liegen als Audiodateien vor und werden den Korpusnutzer(inne)n entsprechend zur Verfügung gestellt.

Zur Erhebung der Audiodaten wurde an den Standorten Leipzig, Wrocław und Pisa der Olympus LS-10 Linear PCM Recorder eingesetzt. Platziert wurde dieser in den meisten Fällen direkt vor oder zwischen den Kommunikationsteilnehmer(inne)n. Die Daten des Standorts Birmingham wurden mit dem Marantz PMD 660 Recorder aufgezeichnet.

Die Audioaufnahmen wurden standardgemäß mit folgenden Parametern als wav-Dateien gespeichert und archiviert:

- Abtastrate: 48 kHz
- Bit-Tiefe: 16 bit
- Stereo

Aufnahmen, die bereits vor Projektbeginn 2009 erhoben wurden, können von diesen Parametern abweichen. Die genauen Parameter jeder Aufnahme sind in den Metadaten der entsprechenden Kommunikation festgehalten.

Die Audiodateien wurden mit der freien Software Audacity bearbeitet, bevor sie zur Transkription freigegeben wurden. Die Bearbeitung umfasste insbesondere den Schnitt und die Maskierung der Daten. Anfangs- und Endpunkt wurden dabei präzise auf den Beginn und das Ende des eigentlichen Gesprächsereignisses festgelegt.

Die Maskierung erfolgte durch Verrauschen sensibler Stellen in der Datei. Maskiert wurden sämtliche Namen der am Gesprächsereignis beteiligten Sprecher(innen). Außerdem verrauscht wurden Namen von Institutionen sowie Informationen, die eine Identifizierung erleichtert hätten, z.B. Ortsnamen, Projektnamen, Konferenz- und Sektionsnamen, Verweise auf eigene Publikationen etc. Vereinzelt wurden in Expertenvorträgen längere Passagen, in denen Sprecher(innen) durch die Sektionsleitung vorgestellt wurden, in Stille umgewandelt, um eine permanente Verrauschung zu vermeiden. Alle Maskierungen wurden in den Transkripten durch entsprechende Metainformationen beschrieben, z.B. ((stadtname)) für „Magdeburg“. Personennamen wurden durch Pseudonyme ersetzt.

2.2 Zusatzmaterialien (n.v.)

Hinweis: Die Zusatzmaterialien sind im GeWiss-Korpus aufgrund urheberrechtlicher Beschränkungen nicht verfügbar (n.v.).

Neben den Audioaufnahmen wurden für zahlreiche Kommunikationen auch sogenannte Zusatzmaterialien erhoben, wie z.B. Handouts oder Vortragsfolien aus Power-Point-Präsentationen. Die Prüfungsgespräche des Korpus liefen nur selten unter Zuhilfenahme zusätzlicher Materialien ab. Vereinzelt, vor allem im deutschen akademischen Kontext, wird vom Prüfling ein Thesenpapier mitgebracht, das in geraffter Form die Prüfungsthemen inhaltlich skizziert inkl. eines Literaturverzeichnisses (vgl. auch die Erläuterungen zu den Genres unter Abschnitt [2.3.2](#)). Hier sehen Sie ein Thesenpapier aus einer Lehramtsprüfung:

Sie können in den Metadaten „Verwendete Medien“ und „Zusatzmaterial“ nachlesen, ob Medien in der Kommunikation verwendet wurden.

2.3 Metadaten

2.3.1 Das Metadatenset des GeWiss-Korpus

Die Metadaten geben Auskunft über den Inhalt, die Herkunft und die beteiligten Personen des Kommunikationsereignisses. Die folgenden zwei Übersichten zeigen das Metadatenset für die Aufnahmesituation (Tab.3) sowie für die Sprecher(innen) (Tab.4):

Kategorie	Beispiel	Kommentare
Projektname	GeWiss	
Teilkorpus	Deutschland	Projektstandort, zu dessen Subkorpus die Aufnahme erstellt wurde; nicht zwingend Ort der Aufnahme
Muttersprachliche Kommunikation	gemischt	ja = die Kommunikation findet in der Muttersprache der Hauptsprecher(innen) statt; nein = die Kommunikation findet nicht in der Muttersprache der Hauptsprecher(innen) statt; gemischt = für manche der Hauptsprecher(innen) ist die Sprache, in der die Kommunikation stattfindet, ihre Muttersprache, für andere dagegen eine Fremdsprache, z.B. ist in einigen Prüfungsgesprächen im Teilkorpus Deutschland Deutsch für den/die Prüfer(in) Muttersprache, für den Prüfling jedoch Fremdsprache
Genre	studentischer Vortrag	Expertenvortrag / studentischer Vortrag / Prüfungsgespräch

Kategorie	Beispiel	Kommentare
Kurzbezeichnung	Grammatik, EuroComGerm	Schlüsselwörter, die den Inhalt der Aufnahme beschreiben
Art des Vortrags / Art des Prüfungsgesprächs	Gruppenvortrag im Master-Studiengang	
Zusatzmaterial	Handout, Präsentation	Information zu vorliegenden Zusatzmaterialien (n.v.), die bei der Kommunikation eine Rolle gespielt haben
Zusammenfassung		kurze Zusammenfassung der Aufnahme
Ort		
Land	Deutschland	Ort der Aufnahme
Jahr	2010	Jahr der Aufnahme
Dauer	1 Stunde 15 Minuten	Dauer der Aufnahme
Institution	Universität	Einrichtung, in der die Kommunikation stattfand
Raum	Seminarraum	Beschreibung des Raums, in dem die Kommunikation stattfand
Sprache		
Basissprache	deu	die Basissprache der Interaktion; Angabe des ISO Language Codes gemäß 639-3
Grad der Mündlichkeit	frei gesprochen	Beschreibung als frei gesprochen / zum Teil abgelesen / vollständig abgelesen / scheint vorformuliert und auswendig gelernt zu sein
Wechsel in andere Sprache(n)	eng	Weitere Sprachen, die in der Kommunikation benutzt wurden, Angabe des ISO Language Code gemäß 639-3
Setting		
Anzahl der Teilnehmer	3 Vortragende, 1 Seminarleiter, ca. 22 Zuhörer	
Verwendete Medien	Handout, Präsentation	Alle Medien, die die Sprecher(innen) in ihrem Vortrag unterstützend verwendet haben, unabhängig davon, ob sie dem Forschendenteam vorliegen oder nicht

Kategorie	Beispiel	Kommentare
Beziehung der Sprecher zueinander und zum Publikum	für die Vortragenden sind die Zuhörer Kommilitonen und somit bekannt	Beschreibung der kommunikativ relevanten Beziehung der Sprecher(innen) zu den anderen an der Kommunikation beteiligten Personen

Tab. 3 Metadaten zur Aufnahmesituation im Gewiss-Korpus

Tabelle 4 zeigt, welche Metadaten zu den Sprecher(inne)n im Korpus erhoben wurden. Sprecherbezogene Daten wurden ausschließlich von Vortragenden, Seminarleiter(inne)n (bei SV), Prüflingen und Prüfenden erhoben, den Hauptsprecher(inne)n der Gesprächsereignisse.

Besondere Aufmerksamkeit wurde der Metadaten-Rubrik *Sprache* gewidmet. Grundsätzlich wurde hier zwischen zwei Typen von *Sprache* unterschieden:

- Erstsprache – die Sprache, in der die schulische Sozialisation (Amtssprache des Herkunftslandes) des/der Sprecher(s/in) erfolgt ist
- L2 – dieser Typ wird zur Angabe aller weiteren Sprachen verwendet, unabhängig von der Reihenfolge, in der sie gelernt wurden oder dem erreichten Niveau.

Es ist durchaus möglich, dass pro Sprecher(in) mehrere Sprachen vom Typ L2 aufgeführt sind.

Kategorie	Beispiel	Kommentare
Sprecherkürzel	MA_0743	
Name	Modini Alama	Pseudonym der Sprecherin
Alter	21	
Geschlecht	weiblich	
Rolle	Vortragender	Spezifikation der Rollen, die der Sprecherin im Korpus zukommen
Bildungshintergrund		Zur Angabe der Stationen des Bildungsweges
Schulbildung	Deutschland, 13 Jahre	
Auslandsstudium	k.A.	Auslandsaufenthalte zu Studienzwecken, bspw. Studium oder Promotion im Ausland, Auslandssemester oder -jahr an einer ausländischen Universität etc.
Auslandsaufenthalt	Lettland, 4 Jahre	Berufsbezogene Auslandsaufenthalte, Sprachkurse etc.
Sprachen		Zur Angabe der Sprachkompetenzen der Sprecherin
Erstsprache	fra	Die Sprache, in der die schulische Sozialisation der Sprecherin erfolgt ist, Angabe des Language Codes gemäß ISO 639-3
L2	deu - TestDaF, TDN 4	Alle weiteren Sprachen, unabhängig von der Reihenfolge, in der sie gelernt wurden und dem erreichten Niveau, Angabe

		des Language Codes gemäß ISO 639-3; bei L2 Deutsch zusätzlich Informationen zum Sprachstand der Sprecherin in der L2
--	--	---

Tab. 4 Metadaten set zu den Sprecher(inne)n im GeWiss-Korpus

Wie Sie auf die Metadaten zugreifen können und wo im Korpus sie zu finden sind, ist im Abschnitt [3. Zugriffsmöglichkeiten](#) beschrieben.

2.3.2 Erläuterungen zu den Genres im deutschen, britischen, polnischen, italienischen und bulgarischen akademischen Kontext

Die drei im GeWiss-Korpus enthaltenen Genres Expertenvortrag, Studentischer Vortrag und Prüfungsgespräch haben an den einzelnen Standorten - Deutschland, Großbritannien, Polen, Italien und Bulgarien - ganz spezifische Ausprägungen. Daher werden sie hier für jeden Standort einzeln charakterisiert.

2.3.2.1 Expertenvorträge (EV)

Deutschsprachige EV im deutschen akademischen Kontext

Die Expertenvorträge im Teilkorpus Deutschland sind Konferenzvorträge (vgl. Metadatum „Art des Vortrags“). Sie richten sich also an interessierte Fachkolleg(inn)en. Aufgenommen wurden die Vorträge auf internationalen Konferenzen in Deutschland, auf denen die Konferenzsprache Deutsch war. Sie umfassen eine Dauer von 40-50 Minuten inklusive Diskussion von meist 10-15 Minuten. Genauere Informationen zum Inhalt des Vortrags sowie zur Länge der Diskussionsphase finden sich im Metadatum „Zusammenfassung“. Vor dem eigentlichen Vortrag wird der/die Vortragende i.d.R. durch die Sektionsleitung vorgestellt, da die Zuhörer(innen) typischerweise „fremdes Fachpublikum“ sind. Dies erfährt man aus dem Metadatum „Beziehung der Sprecher zueinander und zum Publikum“. Der Applaus der Hörer (Klopfen) beendet den Vortrag und ist Signal zur Eröffnung der Diskussionsphase, die von der Sektionsleitung moderiert wird und in der Fragen zum Vortrag gestellt werden können. Inwieweit die Vorträge frei gesprochen oder abgelesen sind, geht aus dem Metadatum „Grad der Mündlichkeit“ hervor. Ob und welche Medien (wie z.B. Power-Point-Präsentation, Handout) genutzt werden, ist aus dem Metadatum „verwendete Medien“ sowie „Zusatzmaterial“ ersichtlich (vgl. auch Abschnitt [2.2](#)).

Deutschsprachige EV im britischen akademischen Kontext

Da es sich als problematisch erwiesen hat, britische Germanist(inn)en in Großbritannien für die Aufnahmen zu gewinnen, die dort auf Deutsch vortragen, wurde auf eine internationale Konferenz mit Deutsch als Konferenzsprache ausgewichen. Es handelt sich dabei um halbstündige Vorträge, bei denen es vorrangig um Wissensvermittlung unter Forscher(inne)n geht. Das Publikum ist also ein internationales Konferenzpublikum von Germanist(inn)en. Es ging bei den aufgezeichneten Vorträgen um literaturwissenschaftliche und linguistische Themen. In einigen Fällen wurden Power-Point-Präsentationen eingesetzt, in anderen gab es ein Handout.

Englischsprachige EV im britischen akademischen Kontext

Für diesen Teil des Korpus wurden zwei verschiedene Formen von Vorträgen aufgezeichnet. Zum einen gibt es Vorträge von eingeladenen Gastsprecher(inne)n, die Teil einer Vortragsreihe im Department „English Language“ sind. Die Dauer dieser beträgt im Durchschnitt etwa 45 Minuten. Es geht bei den Vorträgen sowohl um den Austausch als auch um die Wissensvermittlung unter Forschenden. Die Vorträge sind öffentlich und wenden sich an Wissenschaftler(innen) und interessierte Studierende. Die Sprecher(innen) sind bekannte Sprachwissenschaftler(innen), die zu unterschiedlichen Themen der allgemeinen und anglistischen Sprachwissenschaften vortragen. Einige verwenden Power-Point-Präsentationen, einige Handouts. Zum anderen wurden Vorträge aufgezeichnet, die anlässlich einer Postgraduiertenkonferenz gehalten wurden. Die Sprecher(innen) sind Doktorand(inn)en der britischen Germanistik, die über zumeist literaturwissenschaftliche Themen auf Englisch vortragen. Das Publikum besteht überwiegend aus anderen Doktorand(inn)en. Die Vortragsdauer beträgt dabei eine halbe Stunde mit anschließender Diskussion.

Deutschsprachige EV im polnischen akademischen Kontext

Die deutschsprachigen Expertenvorträge im Teilkorpus Polen sind allesamt Konferenzvorträge (vgl. Metadatum „Art des Vortrags“). Sie richten sich also an interessierte Fachkolleg(inn)en. Aufgenommen wurden die Vorträge auf zwei internationalen Konferenzen in Polen, auf denen die Konferenzsprache Deutsch war. Sie umfassen eine Dauer von 20 bis 30 Minuten inklusive Diskussion von 10 bis 25 Minuten. Genauere Informationen zum Inhalt des Vortrags sowie zur Länge der Diskussionsphase finden sich im Metadatum „Zusammenfassung“. Die Vorstellung des/der Vortragenden an sich kommt im polnischen Teilkorpus nicht vor. Gewöhnlich wird lediglich der Name des/der Vortragenden vor Beginn des Vortrags von der Sektionsleitung bekannt gegeben. Die Zuhörer(innen) sind für den/die Vortragende(n) fremdes Publikum, einige darunter sind jedoch bekannt. Dies erfährt man aus dem Metadatum „Beziehung der Sprecher zueinander und zum Publikum“. Der Applaus der Hörer(innen) (Klopfen) beendet den Vortrag und ist Signal zur Eröffnung der Diskussionsphase, die von der Sektionsleitung moderiert wird und in der Fragen zum Vortrag gestellt werden können. Inwieweit die Vorträge frei gesprochen oder abgelesen sind, geht aus dem Metadatum „Grad der Mündlichkeit“ hervor. Ob und welche Medien (wie z.B. Power-Point-Präsentation, Handout) genutzt werden, ist aus dem Metadatum „Verwendete Medien“ sowie „Zusatzmaterial“ ersichtlich.

Polnischsprachige EV im polnischen akademischen Kontext

Die polnischsprachigen Expertenvorträge im Teilkorpus Polen sind Konferenzvorträge (vgl. Metadatum „Art des Vortrags“), gehalten auf zwei Konferenzen zum Thema Sprache und Kultur. Genauere Informationen zum Inhalt des Vortrags finden sich im Metadatum „Zusammenfassung“. Sie richten sich also in erster Linie an interessierte Fachkolleg(inn)en. Aufgenommen wurden die Vorträge auf Konferenzen in Polen, auf denen die Konferenzsprache Polnisch war. Sie umfassen eine Dauer von 12 bis 26 Minuten. Aus datenschutzrechtlichen und organisatorischen Gründen wurde die anschließende Gesamtdiskussion nicht aufgezeichnet. Die Diskussion schließt sich auf den polnischen Konferenzen jeweils an einen Themenblock an, der mehrere Vorträge umfasst. Es wird nicht unmittelbar nach jedem Einzelvortrag diskutiert. Die Vorstellung des/der Vortragenden an sich kommt im polnischen Teilkorpus nicht vor. Gewöhnlich wird lediglich der Name des/der Vortragenden vor Beginn des Vortrags von der Sektionsleitung bekannt gegeben. Die Zuhörer(innen) sind für den/die Vortragende(n) fremdes Publikum, einige darunter sind jedoch bekannt. Dies erfährt man aus dem Metadatum „Beziehung der

Sprecher zueinander und zum Publikum“. Der Applaus der Hörer(innen) (Klatschen) beendet den Vortrag. Inwieweit die Vorträge frei gesprochen oder abgelesen sind, geht aus dem Metadatum „Grad der Mündlichkeit“ hervor. Ob und welche Medien (wie z.B. Power-Point-Präsentation, Handout) genutzt werden, ist aus dem Metadatum „verwendete Medien“ sowie „Zusatzmaterial“ ersichtlich.

Italienischsprachige EV im italienischen akademischen Kontext

Die Expertenvorträge im Teilkorpus Italien wurden zum Teil bei einem eintägigen Seminar, zum Teil bei zwei größeren, mehrtägigen Konferenzen aufgezeichnet. Die Vortragsthemen reichen von linguistischen zu sprachdidaktischen Fragestellungen. Das Hauptziel der Vorträge besteht darin, den anwesenden Kollegen neue Forschungsergebnisse vorzustellen und diese gemeinsam zu diskutieren. Die Vorträge dauern durchschnittlich 20-30 Minuten. Vor Vortragsbeginn erfolgt in der Regel eine knappe Vorstellung der Vortragenden. Meist wird lediglich der Name des Vortragenden und der Titel des Vortrags genannt, auf die Forscherbiographie des Referenten wird nicht eingegangen. Die anschließende Diskussion gestaltet sich mitunter heterogen: In einigen Fällen folgt dem Vortrag unmittelbar eine Diskussion, in anderen Fällen findet die Diskussion nach zwei zusammenhängenden oder sogar nach einem längeren Block von Vorträgen statt. Alle Vorträge fanden in relativ kleinen Hörsälen statt, bei einer Zuhörerzahl von ca. 40-60 Personen. Bei den Konferenzen kannten sich die Vortragenden zum Großteil untereinander, da es sich um klar begrenzte Forschungsbereiche und Forschergruppen handelte. Die Referenten verwenden entweder eine Power-Point-Präsentation oder ein Handout.

2.3.2.2 Studentische Vorträge (SV)

Deutschsprachige SV im deutschen akademischen Kontext

Die Studentischen Vorträge sind Teil von Bachelor-, Master- oder Magisterseminaren und können das Format von Einzel- oder Gruppenvorträgen haben (vgl. Metadatum „Art des Vortrags“). Sie dienen dazu, wissenschaftliche Literatur zu einem Thema zu referieren bzw. über eigene Projekte wissenschaftlich zu berichten. In der Regel dauern Studentische Vorträge etwa 20 Minuten. In den modularisierten Studiengängen können Vorträge als Vorleistung zur Modulprüfung fungieren und später zu einer Hausarbeit ausgearbeitet werden, wodurch sie gewissermaßen Vorstufen zu einer Prüfungsleistung darstellen. Im Magisterstudiengang ist die Ausarbeitung zu einer Hausarbeit nicht Pflicht, kann aber als eine Möglichkeit wahrgenommen werden, den Leistungsschein zu erwerben.

Die meisten Vortragenden haben Handouts für die Zuhörer(innen) vorbereitet, darüber hinaus nutzen sie beim Sprechen eigene vorbereitete Power-Point-Präsentationen (vgl. Metadaten „Zusatzmaterial“ und „verwendete Medien“ sowie Abschnitt [2.2](#)). Ein Vortrag ist immer auch sprachlich vorbereitet und/oder eingeübt. Die Einstufung im Metadatum „Grad der Mündlichkeit“ gibt an, wie sehr der Vortragende an der Textvorlage aus seiner Vorbereitung bleibt oder ob frei gesprochen wird.

Hat der/die Vortragende Student(in) seine/ihre Ausführungen beendet, bedankt er/sie sich bei den Hörer(inne)n und bekommt Applaus von seinen Kommiliton(inn)en in Form von Klopfen. Damit wird die Diskussionsphase eröffnet, die die Möglichkeit für Rückfragen und Ergänzungen bietet. Die Dauer der Diskussionsphase, wenn es sie gibt, ist sehr unterschiedlich und reicht von 2 bis 26 Minuten. In der Mehrheit der Datensätze ist sie jedoch nicht länger als 10 Minuten. Diese Zeitspanne ist im Metadatum „Zusammenfassung“ mit angegeben und kann für eine bessere Durchsuchbarkeit genutzt werden.

Deutschsprachige SV im britischen akademischen Kontext

Die meisten Module, die britische Studierende der Germanistik absolvieren, enthalten ein Vortragselement, das zumeist mit zehn oder zwanzig Prozent in die Gesamtbewertung einfließt. Sie sind also ein Pflichtteil des Moduls. Die Studierenden werden über die jeweiligen Bewertungskriterien am Anfang des akademischen Jahres durch eine Modulbeschreibung, die Teil des studentischen Handbuchs ist, informiert.

Je nach Modul können diese Vorträge unterschiedlich gestaltet sein. Es gibt Gruppen- oder Einzelvorträge, Vorträge, die auf Postern oder Power-Point-Präsentationen beruhen, oder auch solche, die in eine Art Planspiel integriert werden, wie zum Beispiel eine „mock conference“. Die Zuhörer(innen) sind dabei in den meisten Fällen die anderen Modulteilnehmer(innen). Manchmal werden die Vorträge aber auch öffentlich gemacht, wenn, zum Beispiel, die Magisterstudent(inn)en die Themen ihrer Dissertationen vorstellen. Die Vorträge können je nach Modul unterschiedlich lang sein (zwischen 5 und 32 Minuten) und sich auch in ihrem Ablauf unterschiedlich gestalten, je nach Vorgabe in den Modulbeschreibungen. Die Lehrenden haben bei der Gestaltung ihrer Module und damit auch der Prüfungs- und Vortragsgestaltung relativ freie Hand. Neue Modulbeschreibungen werden vom „Teaching Committee“ der jeweiligen Fakultät geprüft.

Englischsprachige SV im britischen akademischen Kontext

Die meisten Module, die britische Studierende absolvieren, enthalten ein Vortragselement, das zumeist mit zehn oder zwanzig Prozent in die Gesamtbewertung einfließt. Sie sind also ein Pflichtteil des jeweiligen Moduls. Die Studierenden werden über die jeweiligen Bewertungskriterien am Anfang des akademischen Jahres durch eine Modulbeschreibung, die Teil des studentischen Handbuchs ist, informiert.

Je nach Modul können diese Vorträge unterschiedlich gestaltet sein. Es gibt Gruppen- oder Einzelvorträge, Vorträge, die auf Postern oder Power-Point-Präsentationen beruhen, oder auch solche, die in eine Art Planspiel integriert werden, wie zum Beispiel eine „mock conference“. Die Zuhörer(innen) sind dabei in den meisten Fällen die anderen Modulteilnehmer(innen). Die Referate können je nach Modul unterschiedlich lang sein und sich auch in ihrem Ablauf unterschiedlich gestalten, je nach Vorgabe in den Modulbeschreibungen. Die Lehrenden haben bei der Gestaltung ihrer Module und damit auch der Prüfungs- und Vortragsgestaltung relativ freie Hand. Neue Modulbeschreibungen werden vom „Teaching Committee“ der jeweiligen Fakultät geprüft. Die Dauer der Referate ist ebenfalls modulabhängig unterschiedlich und bewegt sich zwischen minimal fünf und maximal zwanzig Minuten.

Deutschsprachige SV im polnischen akademischen Kontext

Die meisten studentischen Vorträge sind Konferenzvorträge (vgl. Metadatum „Art des Vortrags“). Sie richten sich an ein internationales, den Vortragenden fremdes Konferenzpublikum von Linguist(inn)en. Aufgenommen wurden die Vorträge auf zwei internationalen Konferenzen in Polen, auf denen die Konferenzsprache Deutsch war. Sie umfassen eine Dauer von 20 bis 40 Minuten inklusive Diskussion von meist 5-15 Minuten. Genauere Informationen zum Inhalt des Vortrags sowie zur Länge der Diskussionsphase finden sich im Metadatum „Zusammenfassung“. Die Vorstellung des/der Vortragenden an sich kommt im polnischen Teilkorpus nicht vor. Gewöhnlich wird lediglich sein/ihr Name vor Beginn des Vortrags von der Sektionsleitung bekannt gegeben. Die Zuhörer(innen) sind für den/die Vortragende(n) fremdes Publikum, einige darunter sind jedoch bekannt. Dies erfährt man aus dem Metadatum „Beziehung der Sprecher zueinander und zum Publikum“. Der Applaus der Zuhörendenschaft (Klopfen) beendet

den Vortrag und ist Signal zur Eröffnung der Diskussionsphase, die von der Sektionsleitung moderiert wird und in der Fragen zum Vortrag gestellt werden können. Inwieweit die Vorträge frei gesprochen oder abgelesen sind, geht aus dem Metadatum „Grad der Mündlichkeit“ hervor. Ob und welche Medien (wie z.B. Power-Point-Präsentation, Handout) genutzt werden, ist aus dem Metadatum „Verwendete Medien“ sowie „Zusatzmaterial“ ersichtlich.

Einen weiteren Teil bilden Vorträge, die im Rahmen eines studentischen Kolloquiums gehalten wurden (vgl. Metadatum „Art des Vortrags“), in dem auf Deutsch vorgetragen und diskutiert wurde. Sie richten sich also an interessierte Kommiliton(inn)en. Jeder Vortrag dauert von 5 bis 20 Minuten inklusive Diskussion von meist 3-5 Minuten, die durch den/die Kolloquiumsleiter(in) moderiert wird. Genauere Informationen zum Inhalt des Vortrags sowie zur Länge der Diskussionsphase finden sich im Metadatum „Zusammenfassung“. Ähnlich wie bei den Konferenzvorträgen kommt die Vorstellung des/der Vortragenden als solche nicht vor. Es wird lediglich sein/ihr Name vor Beginn des Vortrags von der Sektionsleitung bekannt gegeben. Die Zuhörer(innen) sind für den/die Vortragende/n bekannt, vgl. „Beziehung der Sprecher zueinander und zum Publikum“. Der Applaus der Hörer(innen) (Klopfen) beendet den Vortrag und ist Signal zur Eröffnung der Diskussionsphase. Inwieweit die Vorträge frei gesprochen oder abgelesen sind, geht wiederum aus dem „Grad der Mündlichkeit“ hervor. Ob und welche Medien (wie z.B. Power-Point-Präsentation, Handout) genutzt werden, ist aus dem Metadatum „Verwendete Medien“ sowie „Zusatzmaterial“ ersichtlich.

Polnischsprachige SV im polnischen akademischen Kontext

Die Studentischen Vorträge sind Konferenzvorträge (vgl. Metadatum „Art des Vortrags“), gehalten auf zwei Konferenzen zum Thema Linguistik und Didaktik. Genauere Informationen zum Inhalt des Vortrags finden sich im Metadatum „Zusammenfassung“. Die Vorträge richten sich an interessierte Kommiliton(inn)en. Aufgenommen wurden die Vorträge auf Konferenzen in Polen, auf denen die Konferenzsprache Polnisch war. Sie umfassen eine Dauer von 12 bis 24 Minuten. Aus datenschutzrechtlichen und organisatorischen Gründen wurde die anschließende Gesamtdiskussion nicht aufgezeichnet. Die Diskussion schließt sich auf den polnischen Konferenzen jeweils an einen Themenblock an, der mehrere Vorträge umfasst. Es wird nicht unmittelbar nach jedem Einzelvortrag diskutiert. Die Vorstellung der Vortragenden an sich kommt im polnischen Teilkorpus nicht vor. Es wird lediglich sein/ihr Name vor Beginn des Vortrags von der Sektionsleitung bekannt gegeben. Die Zuhörenden sind für den/die Vortragende(n) fremdes Publikum, einige darunter sind jedoch bekannt. Dies erfährt man aus dem Metadatum „Beziehung der Sprecher zueinander und zum Publikum“. Der Applaus der Hörer(innen) (Klatschen) beendet den Vortrag. Inwieweit die Vorträge frei gesprochen oder abgelesen sind, geht aus dem Metadatum „Grad der Mündlichkeit“ hervor. Ob und welche Medien (wie z.B. Power-Point-Präsentation, Handout) genutzt werden, ist aus dem Metadatum „verwendete Medien“ sowie „Zusatzmaterial“ ersichtlich.

Deutschsprachige SV im bulgarischen akademischen Kontext

Die studentischen Vorträge in der bulgarischen Germanistik können Bestandteil von regulären Bachelor- oder Masterseminaren sein, müssen es aber nicht. Den Dozent(innen) ist es freigestellt, ob sie in ihren Veranstaltungen studentische Vorträge einsetzen oder nicht. Die Entscheidung, ob und in welchem Maße die Referate eine Prüfungsleistung darstellen, obliegt ebenfalls den Seminarleiter(inne)n, ist im Falle dessen allerdings in der jeweiligen Seminar- bzw. Modulbeschreibung festzuhalten. Sowohl in den Bachelor- als auch in den Masterseminaren kann das Halten eines Vortrags Voraussetzung für die Zulassung zur Prüfung sein, würde dann

selbst aber nicht bewertet werden. Lediglich in den Masterseminaren kann der studentische Vortrag auch Vorleistung zu einer Hausarbeit sein.

Die studentischen Vorträge des bulgarischen Teilkorpus wurden im Rahmen der Veranstaltung „Techniken wissenschaftlichen Arbeitens“ (Bachelor-Studiengang, 6. Semester) gehalten. Ziel der Vorträge war es jeweils eine linguistische Teildisziplin, ihren Gegenstand, die wichtigsten Fragestellungen, Methoden und Perspektiven vorzustellen. Es handelt sich um Gruppenvorträge von jeweils 20-30 min. und einer anschließenden Diskussionsphase von nicht mehr als 10 min. Dauer. Die Vorträge wurden nicht benotet, ihr Halten war aber Voraussetzung für die Zulassung zur Prüfung. Den Studierenden wurde freigestellt, welche Medien sie einsetzen: In der Regel wurde Gebrauch von einer Powerpoint-Präsentation, gelegentlich auch nur von Handouts gemacht. Die meisten studentischen Vorträge wurden zudem protokolliert.

2.3.2.3 Prüfungsgespräche (PG)

Deutschsprachige PG im deutschen akademischen Kontext

In einem Prüfungsgespräch wird unter Leitung einer Prüfungskommission ein Gespräch entwickelt, um das Wissen und das wissenschaftliche Argumentieren des Prüflings zu testen. Die Prüfungsleistung wird benotet und fließt in die Gesamtnote des Studienabschlusses ein. Bei Nicht-Bestehen kann und muss die Prüfung wiederholt werden. Wiederholt nicht bestandene Prüfungen können den Studienabschluss gefährden.

Aus rechtlichen Gründen müssen der Prüfungskommission mindestens zwei Personen angehören. Dies sind entweder zwei gleichberechtigte Prüfer(innen) oder ein(e) Prüfer(in) und ein(e) Beisitzer(in) (vgl. Metadatum „Anzahl der Teilnehmer“). Beisitzer(innen) sind für organisatorische Dinge und Prüfungsunterlagen verantwortlich und greifen sonst in das Prüfungsgeschehen nicht ein. Bei Lehramtsprüfungen nimmt obligatorisch ein(e) Vertreter(in) der Schulaufsichtsbehörde am Prüfungsgespräch teil. Die Prüfenden sind dem Prüfling i.d.R. aus Seminaren, Vorlesungen oder sonstigen universitären Zusammenhängen bekannt. Aus dem Metadatum „Beziehung der Sprecher zueinander“ erfährt man, ob und inwieweit Prüfer(in) und Beisitzer(in) dem Prüfling bekannt sind.

Folgende Prüfungsarten gibt es im GeWiss-Teilkorpus Deutschland (vgl. Metadatum „Art des Prüfungsgesprächs“):

- Bachelor-Modulprüfung
- Master-Modulprüfung
- Magisterzwischenprüfung
- Magisterabschlussprüfung
- Lehramtsprüfung

Die Magisterzwischenprüfung findet nach der Hälfte der Studienzeit statt, die Magisterabschlussprüfung sowie die Lehramtsprüfung am Ende der Studienzeit. In den modularisierten Studiengängen wird dagegen studienbegleitend am Ende der Studienmodule geprüft (Bachelor-Modul- und Master-Modulprüfung). Die Themenbreite ist eng verbunden mit der Art der Prüfung. Eine Modulprüfung prüft das Wissen, das in diesem speziellen Modul behandelt wurde und ist somit enger gefasst und kürzer. Eine Studienabschlussprüfung prüft breiteres Wissen mehrerer Semester und mehrerer Lehrveranstaltungen und ist inhaltlich und zeitlich deutlich umfangreicher. Es wird auch überprüft, wie gut der/die Student(in) in der Lage ist, zu wissenschaftlichen Positionen Stellung zu beziehen.

Der Prüfling bereitet sich eigenständig mit Hilfe von wissenschaftlicher Literatur, Skripten, Readern etc. auf die Prüfung vor. Ergänzend dazu können Kolloquien oder Konsultationen besucht werden, um die Prüfungsmodalitäten abzusprechen und ggf. die Themenauswahl und -eingrenzung festzulegen. In manchen Fällen muss der Prüfling zu den abgesprochenen Themen sogenannte Thesenpapiere ausarbeiten, die bei der Prüfung thematisiert werden. In anderen Fällen bringt der/die Prüfende Prüfungsmaterialien ins Gespräch ein, etwa zu analysierende Text- oder Audiobeispiele, oder es werden gar keine zusätzlichen Materialien verwendet. Diese Informationen sind in den Metadaten unter „verwendete Medien“ vermerkt (vgl. auch Abschnitt 2.2).

Die Dauer der Prüfungen bewegt sich zwischen 15 (Master-Modulprüfung) und 55 Minuten (Magisterabschlussprüfung). Die Reihenfolge der Prüfungsthemen kann sich nach der Gliederung des Thesenpapiers richten, es kommt aber auch vor, dass dem Prüfling angeboten wird, die Reihenfolge zu Beginn des Prüfungsgesprächs selbst festzulegen.

Deutschsprachige PG im britischen akademischen Kontext

Alle aufgenommenen Prüfungsgespräche sind 30-minütige Viva-voce-Examen (Viva), d.h. es handelt sich um eine Verteidigung der schriftlichen Abschlussarbeit (final year research dissertation) für den BA in Modern Languages (German). Diese final year research dissertation ist ein Modul, das mit 20 credits bewertet wird (von 120 im Jahr zu absolvierenden); 60 Prozent der Bewertung liegen dabei auf der schriftlichen Arbeit und 40 Prozent auf der Verteidigung (Viva) dieser in einer mündlichen Prüfung. Die Studierenden werden über diese Bewertungskriterien am Anfang des akademischen Jahres durch eine Modulbeschreibung, die Teil des studentischen Handbuchs ist, informiert.

Das Thema der Arbeit wird von den Studierenden in Absprache mit den Lehrenden am Anfang des akademischen Jahres selbst gewählt. Die Arbeit wird nach 24 Unterrichtswochen eingereicht und von Erst- und Zweitgutachtern bewertet, in einigen Fällen wird auch ein(e) externe(r) Gutachter(in) (external examiner) herangezogen.

Grundlage der Vivas sind die schriftlichen Arbeiten. Zwei Hochschullehrende, von denen eine(r) der Erstgutachter(in) ist, stellen Fragen, die die Arbeit direkt betreffen, manchmal aber auch über sie hinausgehen. Abhängig von dem/der Prüfenden können einige Studierende ihre Arbeit zu Beginn der Prüfung in einer kurzen, etwa sechsminütigen Präsentation zusammenfassen. Einige benutzen dazu Folien.

Englischsprachige PG im britischen akademischen Kontext

Die mündlichen Prüfungen, die für das GeWiss-Projekt aufgezeichnet wurden, sind Teil eines MA-Moduls, in dem es um eine eigenständige Forschungsleistung (dissertation) geht. Es wird für verschiedene Magisterstudiengänge in der School of Languages and Social Sciences angeboten. Bevor die Studierenden mit der schriftlichen Ausarbeitung ihrer dissertations beginnen können, müssen sie ihr Thema und das Konzept ihrer geplanten Arbeit vorstellen. Die Vorstellung (bewertet entweder mit „Pass“ oder „Fail“ von dem/der Betreuer(in) der Arbeit) findet vor Publikum statt. Anwesend sind der/die Betreuer(in) der geplanten Arbeit und andere MA-Studierende. Die Vorstellungen sind öffentlich, d.h. alle Lehrenden und Postgraduierten der Fakultät werden eingeladen, wobei nicht alle dieser Einladung Folge leisten. In dieser Form der mündlichen Prüfung gibt es also nicht die Rollenverteilung Prüfer(in) – Prüfling, sondern es gibt einen Chair, also eine(n) Moderator(in), den Prüfling und mehrere Diskutant(inn)en, zu denen auch der/die Betreuer(in) der Arbeit gehören kann. Die Vorstellungen sollen mindestens

zehn Minuten und maximal eine halbe Stunde lang sein. Es ist den Studierenden selbst überlassen, ob sie Handouts verteilen wollen oder eine Power-Point-Präsentation als Grundlage ihrer Vorstellung vorbereiten. Im Anschluss an den Vortrag dürfen alle Anwesenden im Publikum Fragen stellen oder Anmerkungen machen.

Weitere Prüfungen wurden im Rahmen des sogenannten Preessional English Programme aufgezeichnet. Dieses Programm richtet sich speziell an Studierende, die ihr Englisch verbessern wollen, da sie nicht mit Englisch aufgewachsen oder mit Englisch erzogen worden sind. Hier wurden kurze (10 bis 15 minütige) Einzel- oder Gruppenprüfungen aufgezeichnet, die je nach Modul zehn oder zwanzig Prozent der Gesamtbewertung ausmachen.

Die Studierenden werden über die Form und den Ablauf der Prüfung durch die Modulbeschreibung informiert, die sie zu Beginn ihres Studiums in Form eines „Student handbook“ bekommen.

Deutschsprachige PG im polnischen akademischen Kontext

In einem Prüfungsgespräch wird unter Leitung eines Prüfers ein Gespräch entwickelt, um das Wissen und das wissenschaftliche Argumentieren des Prüflings zu testen. Die Prüfungsleistung wird benotet und fließt in die Gesamtnote des Studienabschlusses ein. Bei Nicht-Bestehen kann und muss die Prüfung wiederholt werden. Wiederholt nicht bestandene Prüfungen können den Studienabschluss gefährden.

Der/die Prüfer(in) ist den Prüflingen i.d.R. aus Seminaren, Vorlesungen oder sonstigen universitären Zusammenhängen bekannt. Aus dem Metadatum „Beziehung der Sprecher zueinander“ erfährt man, ob und inwieweit Prüfende den Prüflingen bekannt sind.

Im polnischen Teilkorpus befinden sich drei Arten von Bachelor-Semesterprüfungen (vgl. Metadatum „Art des Prüfungsgesprächs“):

Die erste Gruppe bilden die sprachwissenschaftlichen Prüfungsgespräche, bei denen die Prüflinge einzeln geprüft werden, indem sie eine Frage aus einer Themenliste lösen. Jeder Prüfling bekommt Zeit für die Vorbereitung, während der/die andere Kandidat(in) geprüft wird. Die Prüflinge bereiten sich also im selben Raum vor und hören die Kandidat(inn)en vor ihnen. Nachdem ein Prüfling seine Prüfung beendet hat, verlässt er den Raum. Dies ist Zeichen für den/die nächste(n) einzutreten und seine/ihre Frage zu lösen. In der Regel sind 2-3 Prüflinge im Raum anwesend. Geprüft wird von einem Hochschullehrer zu den Vorlesungsinhalten (vgl. „Anzahl der Teilnehmer“). Die Prüfungsgespräche dauern zwischen 3 und 9 Minuten. Außer Prüfungsfragen werden keine weiteren Zusatzmaterialien verwendet (vgl. „Verwendete Medien“).

Die zweite Gruppe bilden kontrastiv (deutsch-polnisch) orientierte Prüfungsgespräche, bei denen zwei bzw. drei Prüflinge geprüft werden, indem sie die Fragen nacheinander beantworten (vgl. Metadaten „Zusammenfassung“ und „Anzahl der Teilnehmer“). Zuerst kommt die sog. Buchfrage, die sich auf die „Deutsch-polnische kontrastive Grammatik“ (Engel et al. 1999) bezieht. Die Frage wird von den Prüflingen vor der Prüfung gewählt und vorbereitet. Der vergleichenden Charakteristik eines ausgewählten Sprachphänomens folgt die zweite Fragerunde mit der sog. „weichen Frage“. Die Frage wird aus einer Themenliste gelöst (vgl. „Zusatzmaterial“). Geprüft wird ebenfalls von einem Hochschullehrer zu den Vorlesungsinhalten. Die Dauer der Prüfungen ist sehr unterschiedlich und reicht von 14 bis 30 Minuten bei zwei Prüflingen

oder von 29 bis 38 Minuten bei drei. Außer den Prüfungsfragen werden keine weiteren Zusatzmaterialien verwendet.

Die dritte Gruppe machen die literaturwissenschaftlichen Prüfungsgespräche aus, bei denen drei oder vier Prüflinge geprüft werden, indem sie die Fragen nacheinander beantworten (vgl. „Zusammenfassung“ und „Anzahl der Teilnehmer“). Geprüft wird wiederum von einem Hochschullehrer zu den Seminar- und Vorlesungsinhalten. Die Prüfungsgespräche dauern von 27 bis 33 Minuten, vgl. „Dauer“. In einigen Fällen kommt es vor, dass den Prüflingen während des Prüfungsgesprächs deutschsprachige literarische Texte vorliegen, die vorgelesen und interpretiert werden. Diese werden vom Prüfer bereitgestellt, vgl. „Verwendete Medien“.

Polnischsprachige PG im polnischen akademischen Kontext

In einem Prüfungsgespräch wird meist unter Leitung eines Prüfers ggf. einer Prüfungskommission ein Gespräch entwickelt, um das Wissen und das wissenschaftliche Argumentieren des Prüflings zu testen. Die Prüfungsleistung wird benotet und fließt in die Gesamtnote des Studienabschlusses ein. Bei Nicht-Bestehen kann und muss die Prüfung wiederholt werden. Wiederholt nicht bestandene Prüfungen können den Studienabschluss gefährden.

Die im Teilkorpus Polen vorkommenden Bachelor- und Magister-Semesterprüfungen (vgl. Metadatum „Art des Prüfungsgesprächs“) umfassen Prüfungsgespräche zur polnischen Literatur und Kultur, sowie zur Didaktik des Polnischen im Fremdsprachenunterricht (vgl. „Zusammenfassung“). Die Prüflinge werden einzeln oder gruppenweise geprüft, indem sie gewöhnlich 2-3 Fragen direkt von dem/den Prüfer(n) bekommen (vgl. „Anzahl der Teilnehmer“). Die Prüfungen umfassen eine Dauer von 10 bis 26 Minuten bei einem Prüfling und von 35 Minuten bis 2,5 Stunden (inklusive Zeit für Vorbereitung, was mit aufgezeichnet ist) bei 4-5 Prüflingen. In der Aufnahme hört man nicht, was die Prüflinge erarbeiten. Sie bereiten sich still vor. Geprüft wird von einem oder ggf. mehreren Hochschullehrenden zu den Seminar- und Vorlesungsinhalten. Die Prüfenden sind den Prüflingen i.d.R. aus Seminaren, Vorlesungen oder sonstigen universitären Zusammenhängen bekannt. Aus dem Metadatum „Beziehung der Sprecher zueinander“ erfährt man, ob und inwieweit die Prüfenden den Prüflingen bekannt sind. Außer Prüfungsfragen werden keine weiteren Zusatzmaterialien verwendet (vgl. „Verwendete Medien“).

2.4 Transkriptionen

2.4.1 Transkriptionssoftware und Darstellung

Das erhobene Audiomaterial wurde mit dem EXMARALDA-Partitur-Editor (www.exmaralda.org) in Partiturschreibweise transkribiert. Sprechereignisse finden sich damit sequentiell und simultan in einzelnen Segmenten/Events auf unterschiedlichen Transkriptionsspuren untereinander repräsentiert. Die Segmentierung erfolgte nicht auf Grundlage eines spezifischen pragmatischen Ansatzes, sondern nach praktischen Gesichtspunkten individuell durch die Transkribierenden. Festlegungen im Hinblick auf gesonderte Segmentierungen wurden lediglich für folgende Phänomene getroffen:

- Pausen über eine Länge von 0.2 sek wurden in gesonderten Segmenten notiert, um ihre Dauer konkret bestimmen zu können.

- Eine Segmentierung erfolgte insbesondere für Simultanpassagen, um parallel geäußerte Ereignisse präzise abzubilden.
- Segmentierungen erfolgten auch bei zusätzlichen Hinweisen oder Annotationen in korrespondierenden Spuren zu konkreten Äußerungen in der Verbalspur, z.B. Wechsel oder Kommentare.

2.4.1.1 Spurentypen

In den GeWiss-Transkripten finden sich **Verbal-**, **Kommentar-** und **Annotationsspuren**. Sie sind jeweils für alle am Gesprächsereignis Teilnehmenden und ggf. für während des Vortrags oder Prüfungsgesprächs eingesetzte Audio- oder Videobeispiele angelegt worden.

Verbalspur:

Verbalspuren werden mit der Kategorie ‚v‘ für ‚verbal‘ angegeben. In ihnen wurden alle Ereignisse transkribiert, die mündlich geäußert wurden oder als Metainformation für verbal Geäußertes stehen, z.B. für maskierte Stellen ((universitätsname)), ((stadtname)) etc.

Kommentarspur:

Kommentarspuren werden mit der Kategorie ‚k‘ für ‚Kommentare‘ angegeben.

Zu jeder Verbalspur wurde auch eine Kommentarspur angelegt. Hierin sind in erster Linie non-verbale Ereignisse von Seiten der Sprecher(innen) oder der Zuhörenden vermerkt (z.B. *blättert*, *Klopfen*, *klickt* usw.), die unmittelbar Einfluss auf das Kommunikationsereignis haben.

Notiert wurden in dieser Spur außerdem:

- Originalschreibweisen von Wortbildungsprodukten mit Abkürzungen, z.B. *[DDR-Geschichte, FDJler]*
- Hochlautung von Dialektalismen bzw. Regionalismen
- die orthographisch korrekte Schreibweise von in der Verbalspur aufgeführten Klitisierungen, bei denen die Wortgrenzen beider an der Verschmelzung beteiligten Wörter nicht mehr zu erkennen sind, also von jenen Klitisierungen, die ohne Unterstrich in der Verbalspur notiert werden, z.B. [können wir], [haben wir] für *kömma*, *hamma* (genauer dazu vgl. Abschnitt [2.4.2.2 Anpassungen](#))
- die orthographisch korrekte Schreibweise von in der Verbalspur aufgeführten Wörtern, in denen wortinterne Prozesse wie Tilgung, Assimilation und Reduktion so stark auftreten, dass deren Realisierung kaum der Originallautung gleichkommt, z.B. *mitnanner* [miteinander], *anners* [anders], *wern* [werden].

Alle Wörter und Passagen, die aus der Verbalspur unmittelbar in die Kommentarspur übertragen wurden, wurden dabei in eckigen Klammern aufgeführt.

In den englischen Transkripten finden sich in der Kommentarspur außerdem zusätzliche Hinweise auf Stotterer, False Starts, Versprecher und Abbrüche.

In den polnischen Transkripten treten in der nn-Kommentarspur zudem noch Hinweise auf Hintergrundgeräusche und nonverbale Handlungen in Form des Kürzels „zza“ auf.

Annotationsspur:

In den GeWiss-Transkripten finden sich drei Typen von Annotationsspuren:

- Annotationsspuren der Kategorie ‚a‘, in der Fälle von Code-switching mit „Wechsel“ kenntlich gemacht wurden
- Annotationsspuren der Kategorie ‚t‘ für deren sinngemäße Übersetzung dieser Phänomene
- Annotationsspuren der Kategorie ‚DK‘ für die pragmatisch-funktionale Annotation von Diskurskommentierungen

Zu weiteren Informationen zur Annotation der GeWiss-Daten s. [2.5 Annotation](#).

The screenshot shows the GeWiss interface for a transcript titled 'EV_DE_093'. The main content area displays a transcript with various annotations. A specific annotation 'english Wechsel English' is circled in orange. The transcript includes the following text and annotations:

[226] NR_0242 [v] detailergebnisse ah (.) vorstellen ah das war also h ah (.) es wir war also auch öhm

[227] NR_0242 [v] die fragen waren ah <lachend> eng lisch gestellt und mussten english und auch ich

[228] nn [v] ((räuspert sich)) ((lachen))

[229] NR_0242 [v] musste diese ah ah interviews dann die ((unverständlich)) interviews <lachend> (in) in

[230] nn [v] ((hustet))

NR_0242 [v] meinem english als lingua franca> english stellen sich diese diese ah diesem

NR_0242 [a] Wechsel

NR_0242 [t] English

[231] NR_0242 [v] zwang hab ich mir selbst unterzogen weil ich also wei ich ah (i had to) talk about this

NR_0242 [a] Wechsel

NR_0242 [t] ich musste darüber auch auf

[232] NR_0242 [v] also in other conferences <lachend> where i (li to stb) talk: in english- das war der grund

NR_0242 [a] anderen Konferenzen sprechen wo ich (...) auf English sprechen (musste)

Abb. 1: Beispiel einer Volltext-Partitur mit Wechsel

2.4.1.2 Spurennamen

Im Transkript wurden für die verschiedenen Sprecher jeweils unterschiedliche Spurennamen vergeben.

a) Spuren für die Hauptsprecher(innen) der Kommunikationen (von denen sprecherbezogene Daten erhoben wurden)

Die Spuren der Hauptsprecher(innen) jedes Gesprächsereignisses (d.h. der Vortragenden, Seminarleitenden sowie von Prüflingen und Prüfenden) wurden mit den entsprechenden Sprecherkürzeln benannt, z.B. MA_0743. Sprecherbezogene Daten wurden ausschließlich von den Hauptsprecher(inne)n der Kommunikation erhoben.

b) Spuren für weitere Sprecher(innen) (von denen keine sprecherbezogenen Daten erhoben wurden)

Spuren jeder weiteren Person, die neben den Hauptsprecher(inne)n mit einem Redebeitrag am Gesprächsereignis teilnimmt, wurden mit einer anderen als der in a) aufgeführten Spurenzeichnung versehen.

Vortrag:

Im Vortrag betrifft dies:

- Diskutant(inn)en, die sich in ihrem Redebeitrag unmittelbar auf den Inhalt eines Vortrags beziehen, Fragen dazu stellen etc.: DIS 1, DIS 2 usw.
- Personen, die Vortragende vorstellen und die Moderation der Diskussion übernehmen: MOD 1, MOD 2.

Prüfungsgespräch:

In Prüfungsgesprächen können solche Sprecher(innen) beispielsweise Beisitzer(innen) sein, die das Gesprächsereignis protokollieren und dieses nicht direkt mit Fragen oder längeren Redebeiträgen beeinflussen. Diese werden im Spurennamen mit BS aufgeführt.

c) Gesonderte Spuren

nn-Spur:

Ereignisse, die keinem/keiner konkreten Sprecher(in) zugeordnet werden können, werden in der nn-Spur notiert. Zumeist handelt es sich hierbei um Äußerungen von den an der Kommunikation beteiligten Zuhörenden. Darüber hinaus wurden insbesondere in dialogischen Kommunikationen auch Pausen in der nn-Spur vermerkt, um Interpretationen hinsichtlich ihrer pragmatischen Funktionen - durch die Zuordnung zu einzelnen Sprecher(inne)n - zu vermeiden.

Spur für Beispiele:

Für Audio- oder Videobeispiele in Vorträgen oder Prüfungsgesprächen wurde eine BSP-Spur angelegt. Wurden lediglich Wörter oder kurze Sätze gesprochen (z.B. Lehrwerk-Übungen zur Aussprache) wurden diese Beiträge in eine einzige Beispielspur integriert. Handelte es sich um umfangreichere Dialoge oder Redebeispiele, wurde für die verschiedenen Sprecher(innen) innerhalb dieses Hörbeispiels jeweils eine entsprechende Spur (BSP 1, BSP 2, usw.) angelegt.

2.4.2 Transkriptionskonventionen

2.4.2.1 GAT 2-Minimaltranskript

Alle GeWiss-Transkriptionen sind in Anlehnung an die GAT 2-Konventionen des Minimaltranskripts (Selting et al. 2009) erstellt worden. Dieses stellt die Grundausstufe von GAT 2 dar und umfasst Konventionen sowohl zur Notation des Wortlauts von Redebeiträgen (darunter die präzise Darstellung von z.B. Tilgungen, Klitisierungen, Regionalismen, Komposita, Abkürzungen, Zahlen usw.) als auch zur Notation von Verzögerungs- und Rezeptionssignalen, Pausen, Atmen, Lachen, nicht- oder schwerverständlichen Passagen und von nonverbalen Handlungen und Ereignissen.

Eine Liste der von der Standardsprache abweichenden Realisierungen in den GeWiss-Transkripten kann im [Anhang 2](#) nachgeschlagen werden.

2.4.2.2 Anpassungen²

Das GAT 2 ist primär für die Transkription deutschsprachiger Daten entwickelt worden, was eine Anpassung einzelner Konventionen an die Kontrastsprachen Englisch und Polnisch erforderlich machte. Auch die Transkription deutscher L2-Daten erforderte vereinzelt Ergänzungen der GAT 2-Konventionen. Im Folgenden werden lediglich die von GAT 2 abweichenden Transkriptionsregeln festgehalten, die projektintern beschlossen wurden.

Deutschsprachige Daten

Buchstabierungen

Buchstabierungen und Akronyme wurden, sofern sie phonetisch ungebunden sind, als Wörter ausgeschrieben und je nach Vokallänge mit Doppelpunkt als Zeichen der Dehnung notiert (z.B. KDW = ka: de: we:, DDR = de: de: er). Das obligatorische Leerzeichen zwischen den einzelnen Buchstaben ist aus Gründen der adäquaten Darstellung der silbischen Struktur und aktuellen phonetischen Realisierung auch bei Wortbildungsprodukten mit Akronymen gesetzt worden (z.B. FDJler = eff de: jott ler, CDU-Abgeordneter = ce: de: u: abgeordneter).

Die silbische Darstellung der Buchstaben C, J, Q, V und X wurde wie folgt transkribiert, z.B. ce:, jott, qu:, vau, ix.

Klitisierungen

Klitisierungserscheinungen wurden in zwei Gruppen klassifiziert und ihre Notation dahingehend angepasst:

- i)** Klitisierungen, bei denen *eines* der beteiligten Wörter in seiner ursprünglichen Form verändert wurde, bei denen der klitische Vokal getilgt wurde und das Klitikon dadurch i.d.R. seinen Silbenstatus verlor (vgl. u.a. Nübling 1992, 15), wurden mit Unterstrich notiert: z.B. *hab_s*, *gibt_s* usw.

Für die Transkription derartiger Phänomene wurde in GeWiss gemeinhin ein konservatives Vorgehen beschlossen, d.h. ausschließlich deutlich erkennbare Klitisierungen wurden als solche markiert. Nicht als Klitisierungen notiert wurden:

- Elemente, bei denen – wenn auch schwachtonig bzw. reduziert – die ursprüngliche Lautung wahrgenommen wird. Indikator dafür waren z.B. der Vokalneueinsatz, hörbare Pausen o.ä.
 - Bei Proklitika die Verschmelzung des unbestimmten Artikels mit dem rechten Nachbarn (z.B. *n_auto)
 - Verschmelzungen von Präposition und Artikel, die bereits grammatikalisiert bzw. kodifiziert sind, wie *im*, *aufs* usw.³
- ii)** Klitisierungen, bei denen der Einfluss weiterer phonetisch-phonologischer Prozesse so stark war, dass nicht nur *ein* inkorporiertes Element identifiziert werden konnte, sondern die Wortgrenzen *beider* an der Verschmelzung beteiligter Wörter nicht mehr erkennbar waren, in der Regel also auch die Basis betroffen war (vgl. Nübling 1992, 15), wurden in einem Wort und ohne Unterstrich notiert: z.B. *haste*, *hamma*, *kömma*, *simma*. Fälle dieser Art wurden darüber hinaus in der Kommentarspur in ihrer Standardform notiert.

² Für eine ausführliche Darstellung der verschiedenen Anpassungen von GAT 2 im Projekt GeWiss vgl. Lange et al. (i.Dr.)

³ Zugrunde gelegt wurde dabei DUDEN, Bd. 1 (2009)

Eine Liste aller deutschen im Korpus enthaltenen Klitisierungen finden Sie in [Anhang 2](#). Diese Liste kann als Unterstützung für mögliche Konkordanzsuchanfragen (vgl. auch [3.4.2 Effizientes Suchen](#)) nach derartigen Phänomenen fungieren.

Deutschsprachige L2-Daten

Spezifische Anpassungen für deutschsprachige Daten von L2-Sprecher(inne)n waren in Bezug auf folgende Phänomene nötig:

Diakritika:

Diakritika in Fremdwörtern oder fremdsprachigen Passagen wurden der Originalorthographie entsprechend verschriftlicht. Lediglich Apostrophe wurden durch Unterstriche ersetzt.

Aussprachebesonderheiten

Vereinzelt auftauchende und prägnante Aussprachebesonderheiten wurden im Transkript an entsprechender Stelle nach Gehörtem notiert.

Fremdsprachliche Elemente und Äußerungen sind, sofern sie standardgemäß artikuliert wurden, grundsätzlich in ihrer jeweiligen standardsprachlich kodifizierten Orthographie notiert. In nicht eindeutigen Fällen, etwa bei Sprecher(inne)n mit deutlich hervortretendem fremdem Akzent, wurde auf die GAT 2-Konvention für unverständliche Passagen zurückgegriffen und es wurden Alternativen angegeben, z.B. (*direkt/direct*), (*charakter/character*), (*repräsentativ/representative*).

Englischsprachige Daten

Die Anpassungen der GAT 2-Konventionen für das Englische folgen zum Großteil den von Selting et al. (2009) vorgeschlagenen und den von Couper-Kuhlen/Barth-Weingarten (2011) adaptierten Konventionen. Projektinterne Festlegungen betreffen folgende Phänomene:

Tilgungen und Kurzformen:

Zum Zwecke der besseren Korpusdurchsuchbarkeit wurde im Englischen die Vielzahl an möglichen Schreibweisen insbesondere für Tilgungen und Kurzformen auf meist eine einzige reduziert, z.B. wurden *cos*, *cuZ*, *cus*, *cause* für *because* einheitlich durch *cause* verschriftlicht.

Klitisierungen

Apostrophe in englischen Klitisierungen wurden durch Unterstriche substituiert: *i_m*, *we_ll*, *don_t*. Weniger konventionalisierte Formen, die starken Assimilations- und Reduktionsprozessen unterliegen, wurden mit einem Wort notiert und auf eine Schreibung vereinheitlicht, z.B. *wanna* (*want to*), *gimme* (*give me*) usw. Die orthographisch korrekte Schreibweise wurde in der Kommentarspur wiedergegeben.

Komposita

Für die Entscheidung über Getrennt- oder Zusammenschreibung von Komposita wurde die Schreibweise des Oxford English Dictionary zugrunde gelegt. Diese standardsprachlich kodifizierte Orthographie wurde zudem in der Kommentarspur festgehalten.

Buchstabierungen

Buchstaben und Abkürzungen wurden nicht, wie im Deutschen, in silbischer Schreibweise als Wörter notiert, sondern als Buchstaben, die mit einem Leerzeichen voneinander abgesetzt sind, z.B. *m a*. In der Kommentarspur findet sich jeweils der Hinweis auf die Originalschreibweise, deren Realisierung (hier wiederum in silbischer Darstellung) und ggf. die "Aufschlüsselung" des Akronymes, *m a [em ay; Master of Arts]*.

Rezeptions- und Verzögerungssignale

Rezeptions- und Verzögerungssignale wurden im Sinne der vereinfachten Korpusdurchsuchbarkeit auf wenige mögliche Notationsvarianten reduziert: *er, erm, um; hm_hm, hm, yeah, no*. Stark davon abweichende Formen wurden in ihrer realisierten Form dargestellt: *yep, nope*.

Polnischsprachige Daten

Abweichungen und Anpassungen der GAT 2-Konventionen für die Verschriftlichung polnischsprachiger Daten betreffen lediglich:

Normierte Schreibung in der Verbalspur

Da das gesprochene Polnisch nur geringfügige Abweichungen vom Schriftpolnischen zeigt, sind Anpassungen im Bereich der für das Deutsche und Englische diskutierten Phänomene (wie Klitisierungen, Reduktionen usw.) nicht notwendig. Seltene Fälle von Aussprachebesonderheiten oder -abweichungen wurden deshalb in der Kommentarspur aufgeführt.

Diakritika

Diakritika sind im Polnischen unerlässlich und werden aus diesem Grund in den Transkripten entsprechend notiert.

Verzögerungssignale

Alle Verzögerungssignale werden ungeachtet ihrer Länge mit *yyy* notiert.

Rezeptionssignale

Sämtliche Hörersignale und Rückmeldungen werden mit *mmm* verschriftlicht, ungeachtet ihrer affirmativen oder negierenden Funktion.

Nonverbales:

Nonverbale Handlungen und Hintergrundgeräusche werden mit *zza* in der nn-Kommentarspur vermerkt.

2.5 Annotationen

Die GeWiss-Transkripte enthalten zum einen Annotationen von Sprachwechselphänomenen im engeren Sinne. Diese definieren wir nach Myers-Scotton (2005) und Matras (2009) als Wechsel zwischen Sprachen in einer Kommunikation, die auf Äußerungs- oder Wortebene gesche-

hen können.⁴ Sie werden in einer Annotationsspur [a] mit „Wechsel“ kenntlich gemacht und in einer zweiten Annotationsspur [t] sinngemäß übersetzt (vgl. auch Abb. 2).

Sprachwechselphänomene finden sich im GeWiss-Korpus vor allem in den nichtmuttersprachlichen deutschen Daten. Annotiert wurden bei diesen L2-Äußerungen insbesondere:

- Wechsel in andere Sprachen resultierend aus Wortfindungsschwierigkeiten (eingesetzt als kommunikative Strategie)
- Wechsel in der Funktion des Metalanguaging (vgl. Maschler 1994, Beispiele vgl. Reershemius/Lange i.Dr.)
- Der Einsatz von Diskursmarkern aus anderen Sprachen, meist der Erstsprache
- Unklare Fälle bei akzentgeprägter Aussprache gleicher oder ähnlicher Wörter in zwei oder mehr Sprachen (ggf. Fälle pragmatischer Dominanz anderer Sprachen): (*document/dokument*), (*corpus/korpus*), (*representative/repräsentativ*) usw.

Nicht annotiert werden demnach:

- Zitationen und Beispielnennungen in anderen als der Basissprache
- Borrowings⁵, z.B. call for papers
- Eigennamen (*lycée français*, *business administration*) und fach- bzw. wissenschaftsprachliche Termini (*critical period*, *code-switching*)

Darüber hinaus wurde mit dem zweiten Release der GeWiss-Korpora eine weitere Annotationsebene für eines der deutschsprachigen Teilkorpora verfügbar gemacht. Die Expertenvorträge aus dem deutschen akademischen Kontext wurden pragmatisch nach Diskurskommentierungen (vgl. Fandrych erscheint) annotiert. Detaillierte Informationen zum Inhalt als auch dem methodischen Vorgehen finden Sie bei Baur et al. (2014) (Annotationsdokumentation unter <https://gewiss.uni-leipzig.de/help>).

⁴ Ausführlicher zur Annotation von Sprachwechseln in den GeWiss-Daten finden sich bei Reershemius/Lange (i.Dr.).

⁵ Zu einer Definition vgl. auch Reershemius/Lange (i.Dr.), Matras (2009)

3. Zugriffsmöglichkeiten auf das GeWiss-Korpus

Das GeWiss-Korpus steht Ihnen unter der Internetadresse <https://gewiss.uni-leipzig.de> zur Verfügung. Die Seiten sind für den Webbrowser Mozilla Firefox ab Version 18.0.2. optimiert.

Die GeWiss-Seiten haben folgenden prinzipiellen Aufbau:

Navigationsleiste	Menüleiste	Spracheinstellung
	Anzeigefeld	

In der **Menüleiste** können Sie über den Eintrag **Korpusrecherche** in den Unterpunkten **Volltexte** oder **Konkordanzen** auf die GeWiss-Daten zugreifen.

In der **Spracheinstellung** können Sie zwischen den Sprachversionen *Deutsch*, *English* und *Polski* wählen. Dadurch verändert sich die in der Anzeige verwendete Sprache und in manchen Kontexten die Ihnen zur Verfügung stehende Datenauswahl (vgl. Abschnitt [3.3 Volltexte](#)).

Die Einträge in der **Navigationsleiste** ändern sich in Abhängigkeit zu den in der Menüleiste ausgewählten Einträgen.

Im **Anzeigefeld** werden Ihnen die in der Menüleiste ausgewählten Daten angezeigt.

Bevor Sie alle Funktionalitäten wie zum Beispiel die Volltextansicht oder Konkordanzsuche nutzen können, müssen Sie sich auf der oben genannten Seite registrieren und anmelden (Login).

3.1 Registrieren und Anmelden im GeWiss-Korpus

Um auf Volltexte zugreifen (Browse) oder über das Konkordanztool in den Daten recherchieren (Search) zu können, müssen Sie sich auf der Startseite <https://gewiss.uni-leipzig.de> registrieren. Klicken Sie dazu in der Navigationsleiste auf Registrierung.



Abb. 3 Startseite des GeWiss-Korpus

Tragen Sie in der anschließend erscheinenden Registrierungsmaske Ihre Daten ein. Nach Zustimmung zu den Nutzungsbedingungen und Kenntnisnahme der Datenschutzerklärung können Sie sich durch Anklicken der Schaltfläche **Senden** registrieren. Nach dem Senden Ihrer Daten erhalten Sie an die von Ihnen angegebene E-Mail-Adresse eine Kontakt-Mail. Nach Bestätigung dieser Mail werden Ihnen die Zugangsdaten an Ihre E-Mail-Adresse zugeschickt.

Hinweis! Beachten Sie bei Ihrer Registrierung und der Nutzung der Korpusdaten die Nutzungsbedingungen (vgl. Abschnitt [4. Rechtliche Fragen](#))

3.2 Anmelden (Login) im GeWiss-Korpus

Mit dem Ihnen per Mail zugeschickten Benutzernamen und Passwort können Sie sich auf der Seite <https://gewiss.uni-leipzig.de> im GeWiss-Korpus anmelden.

Geben Sie dazu in der Navigationsleiste links Ihren Benutzernamen im Eingabefeld Username und Ihr Passwort im darunter liegenden Feld ●●●●● ein. Bestätigen Sie Ihre Eingabe durch Anklicken der Schaltfläche **Login**.

Falls Sie Ihr Passwort vergessen haben sollten, können Sie durch Anklicken von **Passwort vergessen?** ein neues Passwort anfordern.

Nach einer erfolgreichen Anmeldung wird oberhalb der Sprachauswahl nach dem Eintrag **Eingeloggt als:** Ihr Benutzername angezeigt. Über die nebenstehende Schaltfläche **Logout** können Sie sich wieder abmelden.

3.3 Zugang zu den Volltexten (Browse)

3.3.1 Darstellung und Zugriff

Nach dem Login können Sie die Volltexte aller im GeWiss-Korpus enthaltenen Gesprächsergebnisse als Partitur mit allen relevanten Metadaten sowie den dazu gehörigen Audiodateien unter

Menüleiste: Korpusrecherche > Volltexte

abrufen.

Hier steht Ihnen eine Korpusübersicht der einzelnen Teilkorpora (Abb. 4) zur Verfügung. Sie sind geordnet nach den drei Basissprachen:

- **Deutsch**
- **Englisch**
- **Polnisch**
- **Italienisch**

mit der jeweiligen Unterebene, die die Sprachkompetenz sowie den akademischen Kontext angibt, etwa:

- **L1, akademischer Kontext: Deutschland.**

Anklickbar ist schließlich die Ebene der *Genres(-Schaltflächen)*:

- **PG**
- **EV**
- **SV**

Je nachdem, mit welcher Basissprache Sie arbeiten möchten und für welchen akademischen Kontext Sie sich interessieren, erhalten Sie durch Anklicken der farbigen Genre-Schaltflächen die für die einzelnen Genres verfügbaren Volltexte.

Achtung! *In Abhängigkeit von der auf der Startseite gewählten Spracheinstellung für die Homepage sind jeweils nur die in dieser Sprache verfügbaren Daten zugänglich. Unter der polnischen Spracheinstellung sind nur Daten aus dem polnischen akademischen Kontext mit Metadaten in polnischer Sprache zugänglich; unter der englischen Spracheinstellung sind nur Daten aus dem englischen akademischen Kontext mit Metadaten in englischer Sprache zugänglich. Unter der Spracheinstellung Deutsch sind alle Daten aus allen akademischen Kontexten zugänglich mit Metadaten in deutscher Sprache.*

Beispiel:

1. Um Prüfungsgespräche ausländischer Studierender an einer deutschen Universität einzusehen, klicken Sie auf die grüne Genre-Schaltfläche **PG** unter Basissprache: **Deutsch**, unter **L2, akademischer Kontext: Deutschland**.
2. Möchten Sie hingegen Prüfungsgespräche englischer Germanistikstudierender an einer britischen Universität einsehen, klicken Sie auf die grüne Genre-Schaltfläche **PG** unter Basissprache: **Deutsch**, unter **L2**, aber **akademischer Kontext: Großbritannien**.

Deutsch		
L1, akademischer Kontext: Deutschland		
PG	EV	SV
L2, akademischer Kontext: Deutschland		
PG		SV
L2, akademischer Kontext: Großbritannien		
PG	EV	SV
L2, akademischer Kontext: Polen		
PG	EV	SV
L2, akademischer Kontext: Bulgarien		
		SV
Englisch		
L1, akademischer Kontext: Großbritannien		
PG	EV	SV
L2, akademischer Kontext: Großbritannien		
PG	EV	SV
Polnisch		
L1, akademischer Kontext: Polen		
PG	EV	SV
Italienisch		
L1, akademischer Kontext: Italien		
	EV	

Abb. 4 Korpusübersicht

Für eine gezielte Auswahl der Subkorpora können Sie die Tabelle des [Anhangs 3](#) verwenden.

Durch Anklicken einer Genre-Schaltfläche (**PG**, **EV** oder **SV**) werden im linken Teil des folgenden Anzeigefelds eine Liste aller in diesem Teilkorpus vorhandenen Kommunikationen und im rechten Teil eine Liste aller Sprecher(innen) dargestellt (Abb. 5).

Die Kommunikationen werden mit ihrem Kommunikationskürzel nach dem Muster Genre_akademischer Kontext_laufende Nummer aufgelistet. So bezeichnet z.B. das Kürzel PG_DE_046 das Teilkorpus Prüfungsgespräch im akademischen Kontext Deutschland.

Die Sprecher(innen) werden ebenfalls mit einem (Pseudonymen-)Kürzel dargestellt. Dieses Kürzel folgt dem Muster Initialen_laufende Nummer - Vor- und Zuname.

Zu Ihrer besseren Orientierung werden Ihnen unterhalb der Menüleiste die Umgebungsvariablen für dieses Genre in der Form Basissprache / Sprachkompetenz / Genre angezeigt.

The screenshot shows the GeWiss web interface. At the top, there is a navigation bar with 'Über GeWiss', 'Aktivitäten', 'Korpusrecherche', 'Publikationen', and 'Kontakt'. Below this, the current context is 'Deutschland / Deutsch L1 / PG'. The main content area is divided into two columns: '19 Kommunikationen' and '33 Sprecher'. The 'Kommunikationen' column lists items like PG_DE_046, PG_DE_051, PG_DE_054, PG_DE_080, PG_DE_088, and PG_DE_108, each with a 'Partitur' link. The 'Sprecher' column lists speakers like AP_0241 - Agnes Peukert, AK_0267 - Anke Krippe, AT_0343 - Anke Traum, BR_0272 - Barbara Rabe, BR_0259 - Bärbel Ringel, and CR_0257 - Carola Reinke. A sidebar on the left shows 'Navigation' and 'Korpusübersicht'. The top right corner indicates the user is logged in as 'mustermann' with a 'Logout' button. Language options 'Deutsch', 'English', and 'Polski' are also visible.

Abb. 5 Liste der Kommunikationen und Sprecher(innen) für das Genre

Um alle Metadaten zu einer Kommunikation anzuzeigen, klicken Sie in der Liste der Kommunikationen auf das Kommunikationskürzel. Den Volltext der Kommunikation können Sie sich durch Anklicken des Eintrags „Partitur“ neben dem Kommunikationskürzel anzeigen lassen. Um die Metadaten zu einem/einer Sprecher(in) in diesem Genre aufzurufen, klicken Sie in der Sprecher(innen)liste auf das entsprechende Kürzel.

3.3.2 Metadaten einer Kommunikation einsehen

Menüleiste: Korpusrecherche > Volltexte

Anzeigefeld: Genre-Schaltfläche > Kommunikationskürzel

Klicken Sie auf das **Kommunikationskürzel** der gewünschten Kommunikation, um sich die Metadaten zu dieser Kommunikation anzeigen zu lassen. Im folgenden Anzeigefenster (Abb. 6) werden in der linken Bildschirmhälfte die Metadaten angezeigt, die das kommunikative Ereignis beschreiben. Rechts daneben finden Sie die Metadaten zu den beteiligten Sprecher(inne)n.

Achtung! Ein Ausdruck der Metadaten über die Druckfunktion des Browsers ist nicht möglich. Klicken Sie zum Ausdrucken der Metadaten im **Navigationsmenü** links auf **Drucken**.

The screenshot shows the GeWiss web application interface. At the top, there is a navigation bar with 'GeWiss' and 'Gesprochene Wissenschaftsgespräche'. Below it, a menu includes 'Über GeWiss', 'Aktivitäten', 'Korpusrecherche', 'Publikationen', and 'Kontakt'. The main content area is titled 'Deutschland / Deutsch L1 / PG' and is split into two columns. The left column, labeled '1 Kommunikation', displays metadata for communication 'PG_DE_046'. The right column, labeled '2 Sprecher', displays metadata for two speakers: 'DM_002 - Denise Müller' and 'JS_002 - Jara Schmitzchen'. A sidebar on the left contains a 'Navigation' menu with options like 'Korpusübersicht', 'Liste der Kommunikationen', and 'Drucken'. A top navigation bar shows 'Eingelogg als: mustermann' and 'Logout'.

Abb. 6 Metadaten-Übersicht zur Kommunikation PG_DE_046

3.3.3 Volltext einer Kommunikation einsehen

Menüleiste: Korpusrecherche > Volltexte
Anzeigefeld: Genre-Schaltfläche > Partitur

Zur Anzeige des Volltextes einer Kommunikation klicken Sie in der Kommunikationen-Liste auf **Partitur** rechts neben dem Kürzel der gewünschten Kommunikation.

Sie können die Partitur der Kommunikation auch direkt aus den Metadaten der entsprechenden Kommunikation aufrufen. Klicken Sie dazu in diesen Metadaten oben rechts auf den Link **Partitur**.

Im folgenden Anzeigefeld sehen Sie links die Kürzel der beteiligten Sprecher(innen) und rechts die Transkription ihrer Äußerungen. Während des Lesens der Partitur können Sie Audiosegmente von 15 Sekunden mithören. Klicken Sie dazu in der Partitur auf das weiße Dreieck über dem gewünschten Segment (Abb. 7).

Hinweis: Die Audiosegmente werden auf dem nativen Format des jeweils verwendeten Browsers abgespielt, also entweder im wav- oder mp3-Format. Als Player wird das HTML5-Audio-Element verwendet.

Abb. 7 Partitur eines deutschsprachigen Prüfungsgesprächs

Im **Navigationsmenü** links neben der Partitur finden Sie die Rubriken

- **Navigation,**
- **Spurauswahl,**
- **Rollen.**

Unter **Navigation** können Sie zur Gesamtübersicht des Korpus bzw. zur Liste der für die gewählte Rubrik verfügbaren Kommunikationen wechseln. Außerdem haben Sie hier die Möglichkeit, sich die Partitur als PDF-Datei herunterzuladen und auszudrucken.

Über die **Spurauswahl** können Sie Spuren, die angezeigt werden sollen, separat an- und abwählen (vgl. zu den Spuren Abschnitt [2.4.1.1 Spurentypen](#) und [2.4.1.2 Spurennamen](#)). Voreingestellt ist die Anzeige aller vorhandenen Spuren.

Unter **Rollen** wird Ihnen angezeigt, welche(r) Sprecher(in) welche Rolle in der Kommunikation hat.

Achtung! Ein Ausdruck der Partituren über die Druckfunktion des Browsers ist nicht möglich. Nutzen Sie zum Ausdrucken der Partitur die zugehörige PDF-

Datei, die Sie im Navigationsmenü links in der Rubrik Navigation über den Link Download PDF herunterladen können.

3.4 Zugang zur Konkordanzsuche (Search)

Mit diesem Recherchetool können Sie:

- Konkordanzen (= eine Liste aller Vorkommen eines Suchwortes im Kontext) aus der Verbal- und Annotationsspur erstellen,
- Metadaten hinzufügen,
- die Suchergebnisse filtern.

Von den Treffern aus abrufen können Sie außerdem:

- die Metadatenübersicht und
- das Volltranskript zu den jeweiligen Kommunikationen.

Nach dem Login können Sie konkrete Suchanfragen durchführen unter

Menüleiste: Korpusrecherche > Konkordanzen

3.4.1 Konkordanzen erstellen

Sie sehen in Abb. 8 von links nach rechts

- das Eingabefeld für das Suchwort,
- ein Drop-down-Menü zur Spurenauswahl und
- ein Drop-down-Menü zur Auswahl des zu durchsuchenden Teilkorpus.

Geben Sie ein Suchwort ein und wählen Sie das zu durchsuchende Korpus aus:

<input type="text"/>	Verbalspur ▾	DEU_L1_SV ▾
----------------------	--------------	-------------

Abb. 8 Eingabemaske für die Konkordanzsuche

Wählen Sie für eine Suchanfrage die zu durchsuchende Spur aus (vgl. Abschnitt [2.4.1.1 Spurentypen](#) und [2.4.1.2 Spurennamen](#)). Ihnen stehen sowohl die Verbal- als auch die Annotationsspur zur Auswahl.

Während Sie in der Verbalspur nach beliebigen Wörtern und Konstruktionen suchen können, stehen Ihnen für die Suche in der Annotationsspur sowohl die Option der Suche nach Sprachwechseln als auch die Suche nach Diskurskommentierungen zur Verfügung. Beim Klick auf die entsprechende Annotationsspur erscheinen die jeweiligen Kategorien automatisch an Stelle des Suchfeldes.

Für die Auswahl des zu durchsuchenden Teilkorpus orientieren Sie sich gern wiederum an der Korpusübersicht in [Anhang 3](#).

Bestätigen Sie nach Eingabe Ihres Suchworts (vgl. auch [3.4.2 Effizientes Suchen](#)) Ihre Suchanfrage entweder mit **Enter** oder durch Klick auf den **Suchen**-Button auf der rechten Bildschirmseite.



Der Klammereintrag zeigt Ihnen die Anzahl der in der Konkordanzdarstellung aufgeführten Kontextwörter.

In der *KWIC-Tabelle* (*keyword in context*) erscheinen die Treffer Ihrer Suchanfrage.

also Verbalspur DEU_L1_SV

1 2 3 ... 28 >>

Suchen (10) Kontext - Kontext +
Erweiterte Suche Export

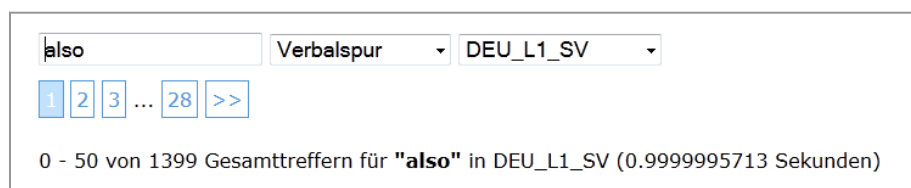
0 - 50 von 1399 Gesamttreffern für "also" in DEU_L1_SV (0.9999995713 Sekunden)

Kommunikation	Sprecher	Linker Kontext	Treffer	Rechter Kontext
SV_DE_036	AT_0213	das ist nicht als ob der irgendwann (.) ja (.)	also	[[fragend] hm] ja (1.0)
SV_DE_036	AT_0213	(0.5) aber der ist zu (1.1) gar nicht mal dreißig	also	sehr (.) er ist schon in unserem alter das ist nicht
SV_DE_077	DB_0221	s (0.6) hat eigentlich keine funktion in den (unterlagen) (0.8)	also	beschreiben (.) (werden) der verfahren (0.2) auf der basis der
SV_DE_077	DB_0221	ja dieses synoptische beschreibung bei bei ehlich und und andern (.)	also	von (comoer) und (.) (0.2) (schülers) (1.1) warum wollen
SV_DE_122	SP_0202	nich °hh is ähm ((schmatzt)) (0.6) warum speziell s (.)	also	ich finde den endetext eigentlich gut oder denke dass da da
SV_DE_122	SP_0202	welche aspekte (.) wollen oder können sie daran auffächern (0.6)	also	da w (0.8) das is für mich noch ein fragezeichen
SV_DE_064	SK_0232	das is (.) doch ziemlich einfach (.) (könnte sein) (.)	also	das is (.) auch für leute die sonst gar äh
SV_DE_122	SP_0202	sogar als äh datei geben (0.3) da gab_s n lehrerpa	also	n unterrichtsprojekt an österreichischen schulen (0.2) das heißt zwischen welten

Abb. 9 Konkordanz des Suchwortes *also*

Der Ansicht in Abb. 9 können folgende Informationen entnommen werden:

Über der Tabelle finden Sie den Hinweis auf die Gesamtanzahl der Treffer zu Ihrem Suchwort in dem jeweiligen Teilkorpus. Ihnen werden pro Seite 50 Suchergebnisse angezeigt. Über die Seitennavigation unter dem Suchfeld können Sie beliebig zwischen ihnen wechseln.



Hinweis: Die Teilkorpora sind unterschiedlich groß. Quantitative Anhaltspunkte zu den einzelnen Teilkorpora können Sie dem Dokument „Statistische Daten“ entnehmen. Beachten Sie bitte, dass die Angabe der Gesamttreffer auf 10.000 beschränkt ist. Für Suchwörter, die mit der Gesamtanzahl von 10.000 Treffern angegeben werden, kann somit von einer Gesamttrefferanzahl ausgegangen werden, die darüber liegt.

In der Konkordanztabelle erscheinen folgende Spalten:

- Kommunikation – Hier wird Ihnen das Kürzel der Kommunikation angegeben, aus dem das Genre (EV, SV, PG), der Standort (DE, GB, PL, IT, BG) und die Nummer der Aufnahme hervorgehen. Durch Klick auf die Kürzel gelangen Sie zum Volltranskript, in dem Sie das Suchwort, wenn nötig, in einem größeren Kontext einsehen sowie die korrespondierenden Audioteile abspielen können.
- Sprecher – Hier wird Ihnen das Kürzel des Sprechers/der Sprecherin angegeben, der/die das Suchwort äußert. Durch Klick auf das Kürzel gelangen Sie zur Metadatenübersicht dieser Kommunikation, der Sie zusätzliche Informationen zu Sprecher/in und Kommunikation entnehmen können.
- Linker/Rechter Kontext – Links und rechts neben dem Suchwort finden Sie den Kontext, in dem das Suchwort auftrat. Sie können den Kontext beliebig vergrößern und verkleinern, indem Sie rechts über der Tabelle auf *Kontext+* oder *Kontext-* und anschließend wieder auf *Suchen* klicken.



Im Feld *Suchen* links daneben ändert sich dabei die Anzahl der Kontextwörter, die in Klammern angegeben ist.

Wenn Sie auf die Spaltenüberschriften klicken, können Sie sich die Tabelle sortieren lassen. Beachten Sie bitte, dass die Sortierfunktion für eine sehr hohe Ergebnisanzahl (>8000) nicht immer zuverlässig anwendbar ist.

3.4.2 Effiziente Suchen formulieren – Reguläre Ausdrücke nutzen

Reguläre Ausdrücke sind Platzhalterzeichen. Manchmal werden sie auch logische Operatoren genannt. Sie beschreiben ein bestimmtes Textmuster in einer abgekürzten oder unspezifizierten Weise (vgl. Lemnitzer/Zinsmeister 2006: 90-92).

Im GeWiss-Suchtool stehen Ihnen folgende reguläre Ausdrücke zu einer effizienten Suche zur Verfügung:

| für ODER (z.B. also|so)

* für BELIEBIG LANGE WORTTEILE vor oder nach der angegebenen Buchstabenkombination (z.B. *stellen für Ergebnisse wie herstellen, bereitstellen, gegenüberstellen usw.; stell* für Ergebnisse wie stellen, stellte, Stellenangebot, *arbeit* für verarbeiten, herausgearbeiteten usw.)

Nutzen Sie für die Suche nach gesprochensprachlich bedingten Reduzierungen oder Klitisierungen auch die Tabelle in [Anhang 2](#).

3.4.3 Metadaten hinzufügen

Mit einem Klick auf **Erweiterte Suche** können Sie Metadaten zu ihrer Suchanfrage hinzufügen.



Es erscheint eine Liste aller Metadaten unter den Kategorien *Kommunikation* und *Sprecher*. Metadaten können durch Klick auf **Hinzufügen** an- bzw. durch wiederholtes Klicken (**Entfernen**) abgewählt werden.

Kommunikation	Auswahl	Filtertyp	
Teilkorpus	Hinzufügen	kein	Deutschland
Muttersprachliche Kommunikation	Hinzufügen	kein	ja
Genre	Hinzufügen	kein	Expertenvortrag
Art	Hinzufügen		
Kurzbezeichnung	Hinzufügen		
Zusammenfassung	Hinzufügen		
Jahr	Hinzufügen		
Land	Hinzufügen		
Institution	Hinzufügen		
Dauer	Hinzufügen		
Raum	Hinzufügen	kein	Computerpool
Basissprache	Hinzufügen	kein	deu
Grad der Mündlichkeit	Hinzufügen	kein	frei gesprochen
Wechsel in andere Sprache(n)	Hinzufügen	kein	eng
Verwendete Medien	Hinzufügen	kein	Präsentation
Anzahl der Teilnehmer	Hinzufügen		
Beziehung der Sprecher zueinander	Hinzufügen		
Sprecher	Auswahl	Filtertyp	
Alter	Hinzufügen		
Geschlecht	Hinzufügen	kein	weiblich
Rollen	Hinzufügen	kein	Vortragender
Erstsprache	Hinzufügen	kein	ara

Abb. 10 Metadatenübersicht unter Erweiterte Suche

Wenn Sie dann auf **Aktualisieren** klicken, kehren Sie zurück zur Konkordanzansicht, wo die ausgewählten Metadaten als zusätzliche Spalten angezeigt werden.

Für einige, insbesondere die Text umfangreicheren Metadateneinträge, erhalten Sie einen Link mit dem Titel Metadatum, der Sie zum ausführlichen Metadatenformular (vgl. [Abschnitt 3.3.2](#)) leitet.

3.4.4 Filter-Optionen

Über **Erweiterte Suche** erhalten Sie auch Möglichkeiten zum Filtern Ihrer Suchergebnisse. Für die meisten Metadaten stehen Ihnen hier über Drop-down-Menüs anwählbare Optionen zur Anzeige **aller** Ergebnisse **mit** (diesem Kriterium) oder **aller** Ergebnisse **ohne** (dieses Kriterium) zur Verfügung. Im Drop-down-Menü rechts daneben finden Sie die entsprechenden zur Auswahl stehenden Werte für das jeweilige Metadatum.

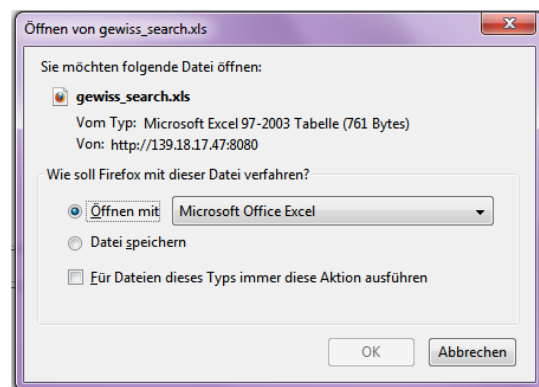
Grad der Mündlichkeit	Hinzufügen	kein	frei gesprochen
Wechsel in andere Sprache(n)	Hinzufügen	kein alle mit alle außer	eng
Verwendete Medien	Hinzufügen	kein	Präsentation

Abb. 11 Auswahl Filteroptionen

Übernehmen Sie Ihre Auswahl für die Konkordanzdarstellung wiederum mit dem Klick auf *Aktualisieren*.

3.4.5 Konkordanzen speichern

Sie können Konkordanzen als xls-Datei über die **Export**-Funktion lokal auf Ihrem Rechner speichern und diese mit Tabellenkalkulationsprogrammen wie z.B. Excel öffnen und bearbeiten.



3.5 Webservices

Für die Arbeit mit dem GeWiss-Korpus steht darüber hinaus eine Reihe von Funktionen in Form von RESTful Java Webservices zur Verfügung, welche teilweise auch im GeWiss-Portal Anwendung finden. Sie erlauben die Ausgabe von Auflistungen und Metadaten zu den vorhandenen Teilkorpora, Kommunikationen, Sprechern und Transkripten. Darüber hinaus stehen Funktionen wie die Suche nach Worten und Wortverbindungen in den Sprachdaten sowie die Anzeige und das Filtern dieser nach bestimmten Metadaten zur Verfügung. Die Webservices unterstützen i.d.R. die Ausgabe der Ergebnisse in XML für die weitergehende Verarbeitung sowie in HTML für die Anzeige in Browsern, was bspw. die Darstellung von Konkordanzsuchergebnissen im KWIC-Format in Form einer HTML-Tabelle ermöglicht. Eine ausführliche Dokumentation der Funktionalitäten und Anwendungsmöglichkeiten der einzelnen Webservices ist in der entsprechenden Dokumentation unter <https://gewiss.uni-leipzig.de/help> einzusehen.

Nutzungsbedingungen

Die folgenden Nutzungsbedingungen gelten für die Nutzung des GeWiss-Korpus. Zugriff und (Weiter-)Nutzung des GeWiss-Korpus werden Ihnen nur gewährt, wenn Sie diese Nutzungsbedingungen durch Auswahl des dafür vorgesehenen Zustimmungsfeldes anerkennen.

1. Persönliche Angaben

Bei der Registrierung werden Sie zur Eingabe von Vorname, Nachname und einer gültigen E-Mail-Adresse aufgefordert. Bei Eingabe von Vorname, Nachname und E-Mail-Adresse sind Sie zu wahrheitsgemäßen Angaben verpflichtet. Die Verwendung von Fantasienamen ist ebenso unzulässig wie die Verwendung ungültiger E-Mail-Adressen. Ungültige E-Mail-Adressen werden gelöscht und die betreffenden Nutzer zu einer Neueingabe aufgefordert.

2. Schutzrechte

Sie werden darauf hingewiesen, dass die im GeWiss-Korpus enthaltenen Informationen, Texte, Bilder, Software, Videos, Grafiken, Audios und andere Materialien in der Regel durch Urheber-, Marken- oder Patentrechte geschützt sind.

3. Nutzung der Informationen

Das Ihnen auf das GeWiss-Korpus gewährte Zugriffs- und Nutzungsrecht ist nicht übertragbar und steht nur Ihnen selbst zu. Sie sind berechtigt, die im GeWiss-Korpus zugänglichen Informationen in dem angebotenen Umfang für private, nicht kommerzielle Zwecke zu nutzen, insbesondere zu lesen, herunterzuladen oder auszudrucken.

Die im GeWiss-Korpus bereitgestellten Daten dürfen nur für Forschung, Lehre und Studium genutzt werden. Die angebotenen Daten können im Rahmen des Zitatrechtes wiedergegeben werden. Dabei ist als Quelle "GeWiss - Gesprochene Wissenschaftssprache, Herder-Institut, Universität Leipzig" und die im GeWiss-Korpus verzeichnete Kennung des zitierten Dokuments anzugeben.

Für andere Zwecke ist eine Nutzung der im GeWiss-Korpus hinterlegten Informationen nicht gestattet. Insbesondere dürfen die Informationen nicht an Dritte weitergegeben, publiziert oder in anderer Form Dritten oder der Öffentlichkeit zugänglich gemacht werden. Die Informationen dürfen daher auch nicht in externen Postings an Newsgruppen, Mailinglisten, elektronischen schwarzen Brettern u. ä. verwendet werden.

Für den Fall, dass Sie die im GeWiss-Korpus enthaltenen Informationen in einem nach den vorstehenden Absätzen nicht gestatteten Umfang nutzen wollen (bspw. um Ausschnitte in Lehrmaterialien oder zu Testzwecken zu verwenden), wenden Sie sich bitte an den Betreiber des GeWiss-Korpus unter gewiss-korpus@uni-leipzig.de.

Der Betreiber des GeWiss-Korpus ist derzeit das Herder-Institut, Universität Leipzig.

Bei der Publikation von Forschungsergebnissen, die auf der Grundlage des GeWiss-Korpus entstanden sind, bittet der Betreiber um eine Benachrichtigung an gewiss-korpus@uni-leipzig.de.

4. Standort und Änderung dieser Nutzungsbedingungen

Die Nutzungsbedingungen können unter dem Punkt "Nutzungsbedingungen" auf den Einstiegsseiten des GeWiss-Korpus eingesehen werden. Der Betreiber des GeWiss-Korpus ist berechtigt, die Nutzungsbedingungen jederzeit zu ändern. Jede(r) Nutzer(in) wird nach jeder Änderung der Nutzungsbedingungen automatisch aufgefordert, den geänderten Bedingungen durch die Auswahl des dafür vorgesehenen Betätigungsfeldes ausdrücklich zuzustimmen. Stimmt der/die Nutzer(in) den geänderten Nutzungsbedingungen nicht zu, wird der Account des/der Nutzer(s/in) gelöscht.

5. Registrierung (Anmeldung) und Zugangskennung

Mit der Nutzung des GeWiss-Korpus unter diesen Bedingungen versichern Sie, dass

- Sie Ihre Kennung und Ihr Passwort niemandem mitteilen werden,
- Ihre Registrierungsdaten richtig und vollständig sind,

Sie erkennen an, dass

- Sie für jede Nutzung des Angebots des GeWiss-Korpus unter Verwendung Ihrer Kennung und Ihres Passwortes verantwortlich sind,
- Sie dafür Sorge zu tragen haben, dass Ihre Kennung und Ihr Passwort vor unautorisiertem Gebrauch geschützt sind,
- Sie verpflichtet sind, die Administration unverzüglich zu benachrichtigen, falls Sie eine missbräuchliche Verwendung Ihrer Zugangsdaten bemerken oder vermuten.

6. Ausschluss von der Nutzung

Der Betreiber des GeWiss-Korpus kann unter Beachtung der hochschulrechtlichen Bestimmungen den Zugang zum GeWiss-Korpus zeitweise in Teilen oder gänzlich sperren oder Ihre Nutzungsberechtigung (Zugriff mit Zugangskennung und Passwort) dauerhaft beenden, wenn das GeWiss-Korpus mit Ihrer Kennung und Ihrem Passwort, sei es durch Sie oder durch Dritte, unter Verletzung dieser Nutzungsbedingungen oder in anderer Weise zum Nachteil des Betreibers des GeWiss-Korpus genutzt wird. Entstehen der Universität durch die missbräuchliche Nutzung Schäden, kommen zudem Schadensersatzansprüche in Betracht.

7. Gewährleistung und Produkthaftung

Die Inhalte des GeWiss-Korpus werden auf der Basis "as-is" und "as-available" angeboten. Eine Gewährleistung wird nicht übernommen.

Sie stimmen ausdrücklich zu, dass der Gebrauch des GeWiss-Korpus sowie dessen Inhalte auf eigenes Risiko erfolgt. Der Betreiber des GeWiss-Korpus übernimmt keine Gewähr, dass die Programmfunktionen den Anforderungen der Nutzer(innen) genügen oder in der von ihnen getroffenen Auswahl zusammenarbeiten. Weder der Betreiber des GeWiss-Korpus noch die Hersteller oder Entwickler haften für Schäden, die aufgrund der Benutzung des GeWiss-Korpus oder dessen Inhalte entstehen.

Der/die Nutzer(in) stellt den Betreiber des GeWiss-Korpus von sämtlichen Ansprüchen Dritter frei, die infolge einer unzulässigen Nutzung, u.a. der verwendeten Audioaufnahmen, Folien und weiteren verfügbaren Zusatzmaterialien durch ihn/sie entstehen oder infolge einer Pflichtverletzung des/der Nutzer(s/in) gegen den Betreiber des GeWiss-Korpus geltend gemacht werden.

Ferner ist der/die Nutzer(in) verpflichtet, für darüber hinaus entstehende Schäden inklusive der Kosten für eine eventuell erforderliche Rechtsverfolgung und -verteidigung aufzukommen. Diese Verpflichtung besteht auch bei einem Missbrauch der Internet-Portal-Registrierung durch Dritte, sofern ein Verschulden des/der Nutzer(s/in) hierfür vorliegt.

GeWiss-Datenschutzerklärung

Das personalisierte Webangebot des GeWiss-Korpus erfordert eine einmalige Benutzerregistrierung. Dabei werden Ihre im Registrierungsformular eingegebenen persönlichen Daten gespeichert. Diese Daten sind für die korrekte Funktionsweise des GeWiss-Korpus erforderlich.

Alle Zugriffe auf den Server des GeWiss-Korpus werden registriert. Aus Sicherheitsgründen werden Datum, Rechneradresse und gelesenes Dokument für einen Monat gespeichert. Diese Daten werden zudem für statistische Zwecke ausgewertet.

Persönliche Daten werden darüber hinaus nicht gespeichert.

Eine Weitergabe an Dritte findet nicht statt.

Es gelten die Datenschutzrichtlinien des Freistaates Sachsen.

Bei der Registrierung zur Korpusrecherche müssen diese Erklärungen beachtet und ihnen zustimmt werden.

4. Literatur

Couper-Kuhlen, Elizabeth / Barth-Weingarten, Dagmar (2011). A system for transcribing talk-in-interaction: GAT 2. *Gesprächsforschung - Online-Zeitschrift zur verbalen Interaktion* [Online]; 12, 1–51. Abrufbar unter: <http://www.gespraechsforschung-ozs.de/heft2011/px-gat2-englisch.pdf> [Stand: 22.2.2013]

Engel, Ulrich / Rytel-Kuc, Danuta / Cirko, Lestaw / Dębski, Antoni / Gaca, Alicja / Jurasz, Alina / Kątny, Andrzej / Mecner, Paweł / Prokop, Izabela / Sadziński, Roman / Schatte, Christoph / Schatte, Czesława / Tomiczek, Eugeniusz / Weiss, Daniel (1999/2000). *Deutsch-polnische kontrastive Grammatik I-II*. Heidelberg / Warszawa: Julius Groos / Wydawnictwo Naukowe PWN.

Fandrych, Christian (erscheint): Metakomentierungen in wissenschaftlichen Vorträgen, in: Fandrych, Christian / Meißner, Cordula / Slavcheva, Adriana (Hg.): *Gesprochene Wissenschaftssprache: Korpusmethodische Fragen und empirische Analysen*. Heidelberg: Synchron-Verlag (= Wissenschaftskommunikation).

Fandrych, Christian / Meißner, Cordula / Slavcheva, Adriana (Hg.) (i.Dr.). *Gesprochene Wissenschaftssprache: Korpusmethodische Fragen und empirische Analysen*. Heidelberg: Synchron-Verlag. (= Wissenschaftskommunikation).

Fandrych, Christian / Meißner, Cordula / Slavcheva, Adriana (2012). The GeWiss Corpus: Comparing Spoken Academic German, English and Polish, in: Schmidt, Thomas / Wörner, Kai (Hg.): *Multilingual corpora and multilingual corpus analysis*. Amsterdam: Benjamins. (= Hamburg Studies in Multilingualism).

Lange, Daisy / Slavcheva, Adriana / Rogozińska, Marta / Morton, Ralph (i.Dr.). GAT 2 als Transkriptionssystem für multilinguale Sprachdaten? Zur Adaption der Notationskonventionen im Rahmen des Projekts GeWiss, in: Fandrych, Christian / Meißner, Cordula / Slavcheva, Adriana (Hg.). *Gesprochene Wissenschaftssprache: Korpusmethodische Fragen und empirische Analysen*. Heidelberg: Synchron-Verlag. (= Wissenschaftskommunikation).

Lemnitzer, Lothar / Zinsmeister, Heike (2006). *Korpuslinguistik*. Tübingen: Narr.

Maschler, Yael (1994). Metalanguaging and discourse markers in bilingual conversation. *Language in society*; 23, 325-366.

Matras, Yaron (2009). *Language Contact*. Cambridge: Cambridge University Press.

Myers-Scotton, Carol (2005). *Multiple voices: an introduction to bilingualism*. Oxford: Blackwell-Wiley.

Nübling, Damaris (1992). *Klitika im Deutschen: Schriftsprache, Umgangssprache, alemannische Dialekte*. Tübingen: Narr. (= Script Oralica; 42).

Reershemius, Gertrud / Lange, Daisy (i.Dr.). Sprachkontakt in der mündlichen Wissenschaftskommunikation, in: Fandrych, Christian / Meißner, Cordula / Slavcheva, Adriana (Hg.). *Gesprochene Wissenschaftssprache: Korpusmethodische Fragen und empirische Analysen*. Heidelberg: Synchron-Verlag. (= Wissenschaftskommunikation).

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 / Mertzlufft, Christine / Meyer, Christian / Morek, Miriam / Oberzaucher, Frank / Peters, Jörg / Quasthoff, Uta / Schütte, Wilfried / Stuckenbrock, Anja / Uhmann, Susanne (2009). Gesprächsanalytisches Transkriptionssystem 2 (GAT 2). *Gesprächsforschung - Online-Zeitschrift zur verbalen Interaktion* [Online]; 10, 353–402. Abrufbar unter: <http://www.gespraechsforschung-ozs.de/heft2009/px-gat2.pdf> [Stand: 22.2.2013]

Anhang 1

Dieser Anhang enthält drei Tabellen, die alle im Kernkorpus (+) enthaltenen Kommunikationen mit ihren Themen verzeichnen. Die Tabelle wird demnächst um die Angaben der assoziierten Korpora ergänzt.

Expertenvorträge

Kommunikation	Thema
EV_DE_004	Multimediales Lernen
EV_DE_005	Literarizität im FSU
EV_DE_092	Kindliche Schulwegbeschreibungen
EV_DE_093	Multilingual Awareness
EV_DE_094	EuroCom
EV_DE_095	Intertextualität, Fachsprache
EV_DE_096	Dialekt, Prestige usw.
EV_DE_097	SOV-Struktur, L1-Transfer
EV_DE_098	Textlinguistik, Stilistik, Diskurslinguistik
EV_DE_100	Demonstrativ- und Personalpronomen „der“
EV_DE_101	Interkulturelle Kompetenz
EV_DE_104	Mehrsprachigkeit # interkulturell, Bozen
EV_DE_105	„tun“-Periphrase
EV_DE_106	Politische Plakate
EV_UK_001	GermanC-Korpus
EV_UK_002	Standardisierungsprozess
EV_UK_003	Verben und Adverbien
EV_UK_004	Raoul Schrott
EV_UK_005	Interkulturelles Theater
EV_UK_006	Modalverben
EV_UK_007	NS-Vergangenheitsbewältigung
EV_UK_008	Zelter im Briefwechsel mit Goethe
EV_UK_009	Starke Verben: Englisch und Russisch
EV_UK_010	Small Talk
EV_UK_011	DDR
EV_UK_012	Learning Technology
EV_UK_013	Knowledge Transfer
EV_UK_014	musical self reflexivity
EV_UK_015	Literary dreams
EV_UK_016	Idiolects
EV_UK_017	authorical reframing
EV_UK_018	Conflicts
EV_PL_002	Operatorpartikeln

EV_PL_003	Abtönungspartikeln
EV_PL_004	Partikeln
EV_PL_009	Metasprachliche Determinanten des Erbgutes
EV_PL_011	Phraseologismen
EV_PL_013	Sprachtheorie
EV_PL_019	Nominativ
EV_PL_020	Kausalsätze
EV_PL_021	Sprache und Wissen
EV_PL_022	Niederländische Aussprache
EV_PL_023	Übersetzung
EV_PL_027	Kultur des Unternehmensmanagements
EV_PL_028	Aktionärenbrief
EV_PL_029	Ablehnungsakt in der Politik
EV_PL_032	Vorhersehbarkeit der Sprechakte
EV_PL_033	Entwertungsakt
EV_PL_038	Sportliche Berichte
EV_PL_040	Pressemeldungen zum Ehejubiläum
EV_PL_044	Gattung der Globalisierung gegenüber
EV_PL_048	Antrag
EV_PL_049	Sprechakte und Sprechsorten
EV_PL_050	Sprechakte, Akte der Gebärdensprache
EV_PL_056	Aufsätze der Lizeumschüler
EV_PL_058	Der sprachliche Humor
EV_PL_068	SMS-Nachrichten
EV_PL_075	Deaf studies

Studentische Vorträge

Kommunikation	Thema
SV_DE_010	Versprecher
SV_DE_014	Perzeptionsmodelle
SV_DE_016	Sprechrhythmus
SV_DE_021	Phraseologismen, Wörterbücher
SV_DE_022	Phraseologismen, Lernerwörterbücher
SV_DE_023	Interjektionen, Wörterbücher
SV_DE_033	Literarizität, Deutungslernen
SV_DE_035	Grammatik, EuroComGerm
SV_DE_036	Indirekte Rede
SV_DE_057	E-Learning
SV_DE_059	Hybrides Lernen, Lehrerrolle
SV_DE_060	Medienkompetenz
SV_DE_064	Autorenwerkzeug Hot Potatoes

SV_DE_065	Leipzig, Wandel
SV_DE_068	Wohnen, Deutschland
SV_DE_069	Generationen
SV_DE_070	Identitäten
SV_DE_075	Sprachtendenzen, tun-Periphrase
SV_DE_077	DaZ, Sprachstandsdiagnosen
SV_DE_079	DaZ, Herkunftsunterricht
SV_DE_121	Lehrwerkanalyse
SV_DE_122	Literatur, Fremdsprachenunterricht
SV_UK_001	Deutsche Firmen im chinesischen Markt
SV_UK_002	Opel in der Krise
SV_UK_003	Der Nationalsozialismus im Film
SV_UK_004	Englischer Einfluss
SV_UK_005	Aldi auf dem Britischen Markt
SV_UK_006	Rechtsextreme Gewalt
SV_UK_007	Islam in Deutschland
SV_UK_008	Vergangenheitsbewältigung
SV_UK_009	NS-Zeit
SV_UK_010	Soziale Gesetzgebung, Bismark
SV_UK_011	Kindermärchen
SV_UK_012	Dada
SV_UK_013	Theaterpädagogik
SV_UK_014	Polens EU-Beitritt
SV_UK_015	Hans Christian Andersen
SV_UK_016	Mangas
SV_UK_017	Spätmittelalter in Deutschland
SV_UK_018	Weine aus Deutschland
SV_UK_019	Rendezvous
SV_UK_020	Urlaub
SV_UK_021	Chameleon
SV_UK_022	Germany
SV_UK_023	Shared resouces
SV_UK_024	Expert witness
SV_UK_025	interlingual identifiers
SV_UK_026	IT Metaphors
SV_UK_027	Lobbying
SV_UK_028	political Metaphors
SV_UK_029	IT Metaphors
SV_UK_030	Spiegel
SV_UK_031	Übersetzungs-Organisationen
SV_UK_032	Tausendundeine Nacht
SV_UK_033	Literary translations

SV_UK_034	Family Guy
SV_UK_035	Igbo
SV_PL_001	Attributive Adjektiv- und Partizipphrasen
SV_PL_002	Duden-Grammatik
SV_PL_003	Wysockie Obcasy - Zeitschrift, Werbung, Frau
SV_PL_004	Der sportliche Kommentar in Medien
SV_PL_005	Aggression in der Sprache
SV_PL_006	„Ein“
SV_PL_007	Werbung
SV_PL_008	Gefahren für die Entwicklung der Polnischen Sprache
SV_PL_009	Defizite in der Grammatikschreibung
SV_PL_010	Schönheit
SV_PL_011	Begrüßung
SV_PL_012	Sprachlicher Relativismus
SV_PL_013	Geburtsikone als Verschmelzung
SV_PL_014	Feminativa
SV_PL_015	Feminismus
SV_PL_016	Zeit und Prozess
SV_PL_017	Überredung
SV_PL_018	In vitro
SV_PL_019	Konstruktion am-Progressiv
SV_PL_020	Branna
SV_PL_021	Lipowa - Laznie
SV_PL_022	Jüdisches Theater
SV_PL_023	Evangelisches Stift Tübingen
SV_PL_024	Philipp Gotthard von Schaffgotsch
SV_PL_025	Freiwaldau
SV_PL_026	Vulgarismen
SV_PL_027	„Die Reise“ Friedrich Nicolai
SV_PL_028	Patriotismus
SV_PL_030	„Wysockie obcasy“ - Zeitschriftenanalyse
SV_PL_031	Sprachphilosophie
SV_PL_032	Sprache der Kletterer
SV_PL_033	Bürgerjournalismus

Prüfungsgespräche

Kommunikation	Thema
PG_DE_046	Phonetik, Produktionsmodelle
PG_DE_047	Phonetik, Lauterwerb
PG_DE_048	Phonetik

PG_DE_050	Phonetik, Sprechrhythmus
PG_DE_051	Phraseologie, trennbare Verben, Zweitspracherwerb
PG_DE_054	Geschichte der Sprachwissenschaft
PG_DE_055	Lesen, Erzählen, Sprachlernberatung, DaF/nE, GER
PG_DE_080	Morphologie
PG_DE_081	Morphologie
PG_DE_084	Morphologie
PG_DE_085	Morphologie
PG_DE_087	Morphologie
PG_DE_088	Morphologie
PG_DE_089	Morphologie
PG_DE_090	Morphologie
PG_DE_091	Morphologie
PG_DE_108	Regionale Intonation, Spracherwerb
PG_DE_109	Code-Switching, Sprache und Rundfunk
PG_DE_110	Literarisches Lernen, Leseförderung
PG_DE_111	Literarisches Lernen
PG_DE_112	Literarisches Lernen: Drama, Grammatikunterricht
PG_DE_114	Metapherntheorie, Typologie/Syntax, Realismus
PG_DE_115	Literarisches Lernen: Didaktik epischer Texte
PG_DE_116	Rechtsschreibung, Rechtsschreiberwerb, Textsorten
PG_DE_117	Lyrik, Rechtsschreibung
PG_DE_118	Sprachdidaktik, Lesesozialisation, Literaturdidaktik
PG_DE_120	Mündliche Kommunikation, Literaturdidaktik
PG_DE_123	Sprechen und Denken, Phonetik
PG_DE_124	Perzeptionsmodelle, Phonetik
PG_DE_125	Alter, Aussprache, Critical Period Hypothesis
PG_DE_127	Bedeutungswandel, Phraseodidaktik, kontrastive Lexikologie, Lexikographie, Textsorten
PG_DE_131	Nachkriegs-, Väterliteratur
PG_DE_132	Luther, Restaurationszeit, Mörike, Lasker-Schüler, Standard, Variation
PG_DE_135	Gedächtnisroman, Restauration, Kasus, Verständlichkeit
PG_DE_136	Mariendichtung, Passionsspiel, Weimarer Republik, Wortarten, Standard-Variation
PG_DE_137	Jüdische AutorInnen, Orient-Mittelalter, Standard-Variation, Wortarten
PG_DE_140	Erinnerungskultur
PG_DE_141	Zweitspracherwerb, Kontrastivhypothese, Identitätshypothese, Transfer, Informationsstruktur, Basic variety
PG_DE_142	Zweitspracherwerb, Informationsstruktur, Basic variety
PG_DE_146	Mittelhochdeutsch
PG_UK_001	Fußball
PG_UK_002	Helmut Kohl
PG_UK_003	Wählerprofile in Baden-Württemberg
PG_UK_004	Weimarer Dreieck
PG_UK_005	Marketing und Wiedervereinigung
PG_UK_006	Finanzierung von Fußball
PG_UK_007	Frauen in Führungspositionen

PG_UK_008	Frauen in der DDR
PG_UK_009	Der Euro
PG_UK_010	Die deutsche Minderheit in Belgien
PG_UK_011	Talentmanagement
PG_UK_012	Rechtsradikalismus in Österreich
PG_UK_013	Deutsch-Polnische Beziehungen
PG_UK_014	Fremdsprachenunterricht in Deutschland
PG_UK_015	Schwyzerdütsch
PG_UK_016	Wirtschaftswunder
PG_UK_017	Luftfahrtindustrie
PG_UK_018	Autoindustrie
PG_UK_019	Weißerose
PG_UK_020	Emil Nolde
PG_UK_021	Maßnahmen zur Integration der türkischen Minderheit nach PISA
PG_UK_022	Multikulturalität und Ausländerfeindlichkeit in Österreich
PG_UK_023	Tierschutz in Deutschland
PG_UK_024	Das Thema Jugend und Alkohol in der deutschen Presse
PG_UK_025	Englisch als globale Sprache
PG_UK_026	Alte und neue Frauenbilder
PG_UK_027	Stalingrad
PG_UK_028	Herausforderungen für die deutsche Automobilindustrie
PG_UK_029	Investitionen der deutschen Wirtschaft in China
PG_UK_030	Bio als Verkaufsfaktor
PG_UK_031	Flachbildfernseher
PG_UK_032	Course books
PG_UK_033	Authorship analysis
PG_UK_034	cognitive stilistics
PG_UK_035	Syntactic ordering
PG_UK_036	Polish Please
PG_UK_037	Near synonyms
PG_UK_038	Interlingual identifiers
PG_UK_039	New horizon college English
PG_UK_040	Experienced life in Birmingham
PG_UK_041	ESL classrooms
PG_UK_042	National Geographic
PG_UK_043	Geschlechterrollen
PG_UK_044	Chinese ELT coursebook
PG_UK_045	Forensic analysis
PG_UK_046	Authorship analysis
PG_UK_048	Mother tongue
PG_UK_049	Gender
PG_UK_050	Headway
PG_UK_051	Kultur und Sprachen
PG_UK_052	Geschlechterdiskriminierung
PG_UK_053	Interference
PG_UK_054	Kritisches Denken
PG_UK_055	Teacher
PG_UK_056	Sexismus
PG_UK_057	Authorship attribution
PG_UK_058	Data driven learning

PG_UK_059	Bible translation
PG_UK_060	Online news
PG_UK_061	Herr der Ringe
PG_UK_062	Gesundheit
PG_UK_063	Hound
PG_UK_064	Harry Potter
PG_UK_065	Galaxy
PG_UK_066	Kasachstan
PG_UK_067	Embassies
PG_UK_068	Disney
PG_UK_069	3D TV
PG_UK_070	3D technology
PG_UK_071	Innovation
PG_UK_072	Photoshop
PG_UK_073	Microsoft Surface
PG_UK_074	Spotify
PG_UK_075	Pen
PG_UK_076	Xbox
PG_UK_077	Maglev
PG_UK_078	Igoogle
PG_UK_079	Computer Viruses
PG_UK_080	Natal
PG_UK_081	Windows 7
PG_UK_082	Broadband internet
PG_UK_083	dvd
PG_UK_084	Robotics
PG_UK_085	Gaming
PG_UK_086	Sony
PG_PL_001	„Wesele“ von Stanisław Wyspiański
PG_PL_002	„Lalka“ von Bolesław Prus
PG_PL_003	Rustikalismus, Lob der Natur und des Primitiven in der Literatur, „Na wspak“ von Huysmans
PG_PL_004	Frauenbild in der Literatur, Dramaturgie von Maeterlinck
PG_PL_005	Naturalismus, Dichtung von Leopold Staff
PG_PL_006	Dichtung der zweiten Hälfte des XIX. Jahrhunderts, „Bez dogmatu“ von Henryk Sienkiewicz
PG_PL_007	Der modernistische Künstler, „Marta“ von Eliza Orzeszkowa
PG_PL_008	„Nad Niemnem“ von Eliza Orzeszkowa, Modernismus
PG_PL_009	Der philosophische Pessimismus, „Wyzwolenie“ von Stanisław Wyspiański
PG_PL_010	Kulturelles Leben im XIX. Jahrhundert, Symbolismus
PG_PL_012	Positivismus und Junges Polen
PG_PL_013	Polnische Literatur des 20. Jahrhunderts
PG_PL_016	Kohärenz
PG_PL_017	Sprachliche Höflichkeit
PG_PL_018	Textsorte
PG_PL_019	Kohäsion
PG_PL_020	Wortbildung
PG_PL_021	Sprache als Zeichensystem

PG_PL_022	Grices Konversationsmaximen
PG_PL_023	Sprechakte
PG_PL_024	Wortbildungsverfahren
PG_PL_025	Flexion
PG_PL_026	Unterschiede zwischen gesprochenen und geschriebenen Texten
PG_PL_027	Intentionalität
PG_PL_028	Wortbildungsverfahren
PG_PL_029	Wortbildung
PG_PL_030	Textverarbeitung
PG_PL_031	Bühlers Organonmodell
PG_PL_032	Grices Konversationsmaximen
PG_PL_034	Terminus Sprache
PG_PL_035	Kohäsion
PG_PL_036	Textverarbeitung
PG_PL_037	Struktur des Textes
PG_PL_038	Semantische Relationen
PG_PL_039	Terminus Sprache
PG_PL_040	Lexik
PG_PL_041	System
PG_PL_042	Textverarbeitung
PG_PL_043	Semantik
PG_PL_044	Bühlers Organonmodell
PG_PL_045	Kommunikationsmodell von Shannon und Weaver
PG_PL_047	Zeicheneinteilung
PG_PL_050	Wortbildung
PG_PL_051	Semantik
PG_PL_052	Zeicheneinteilung
PG_PL_053	Informativität und Situativität
PG_PL_054	Textsorte
PG_PL_055	Terminus Sprache
PG_PL_056	Bühlers Organonmodell
PG_PL_057	Zeicheneinteilung
PG_PL_058	Intentionalität
PG_PL_059	Intentionalität
PG_PL_061	Semantik
PG_PL_062	Informativität und Situativität
PG_PL_063	Wortbildungsverfahren
PG_PL_064	Text, Äußerung, Diskurs
PG_PL_065	Struktur des Textes
PG_PL_066	Struktur des Textes
PG_PL_067	Intentionalität
PG_PL_068	Bühlers Organonmodell
PG_PL_069	Kohärenz
PG_PL_070	Sprache als Zeichensystem
PG_PL_071	Informativität und Situativität
PG_PL_072	Sprachliche Höflichkeit
PG_PL_075	Sprache als Zeichensystem
PG_PL_076	Intentionalität
PG_PL_077	Unterschiede zwischen gesprochenen und geschriebenen Texten
PG_PL_078	Wortbildung

PG_PL_079	Textverarbeitung
PG_PL_080	Textlinguistik
PG_PL_081	Textverarbeitung
PG_PL_082	Wortbildung, Apposition, Linguistik, Aspiration und Affrikaten, Klassifizierung der Sprachen, Quantität vs. Qualität
PG_PL_086	Wortbildung, Satelliten, Areale Einordnung des Deutschen und des Polnischen, Konsonanten, Klassifizierung der Sprachen, Genus verbi
PG_PL_090	Attributsätze, Dativergänzung, Wortklassen, Interferenz, Alternationen, Linguistik, Deklination, Interferenz, Alternationen
PG_PL_110	Steigerung, Apposition, Linguistik, Diphthonge, Interferenz, Konsonanten
PG_PL_114	Wortbildung, Apposition, Alternation, Klassifizierung der Sprachen, Pidginsprachen, Kreolsprachen, Steigerung, Einordnung der Sprachen, die Liquidä
PG_PL_120	Schreibzeichen, Sprechakte, Wortbildung, Genetische Klassifizierung der Sprachen, Liquidä, Linguistik, Affrikaten, Kontakt, Transferenz, Transfer, Interferenz, Integration, Alternation
PG_PL_211	Hölderlin, "Nibelungenlied", Eichendorff, Schiller
PG_PL_212	Aufklärung, Sturm und Drang, Goethe
PG_PL_213	Gedichte, die Sturm und Drang-Periode, Eichendorff
PG_PL_214	Angelus Silesius, E.T.A. Hoffmann, Weimarer Klassik, „Faust“, „Maria Stuart“
PG_PL_215	Polnisch als Fremdsprache
PG_PL_216	Polnisch als Fremdsprache
PG_PL_217	Die populäre Literatur, Der ritterliche Roman, Die Bearbeitung, Die polnische Literatur
PG_PL_218	Kultursupermarkt, Literaturreichtungen, Die negativen Helden, Die populäre Literatur und die Hochliteratur
PG_PL_219	Populäre Literatur, Sensationelle Literatur, Einfluss der ausländischen Literatur auf die polnische Fantasy-Literatur
PG_PL_220	Literaturtheorie
PG_PL_221	Kultur, Soziogenese der Begriffe <i>Kultur</i> und <i>Zivilisation</i> in Deutschland nach Elias
PG_PL_222	Literaturtheorie
PG_PL_223	Gegenwartspolnisch
PG_PL_224	Polnisch als Fremdsprache
PG_PL_225	Polnisch als Fremdsprache
PG_PL_226	Polnisch als Fremdsprache
PG_PL_227	Polnisch als Fremdsprache

Anhang 2

Im Folgenden finden Sie eine alphabetische Liste der von der Standardsprache abweichenden Realisierungen (Klitisierungen) aus den deutschsprachigen Transkripten.

Standardsprache	Transkribierte Realisierungen
also einem/einen	also_m, also m, also_n, also n
andere(n)	anre, annern
anders	anners
auf der	auffa
auf ein, auf einen	auf_n, auf n, auf nen
aus dem	aus_m, aus m
brauche es	brauch_s
die/der es	die_s, der_s
eigentlich	einklich
einem	eim
es gibt	s gibt, s_gibt
es hatten	s_hatten, s hatten
es ist	s is, s ist, s_is, s_ist
es kommt	s kommt, s_kommt
es wird	s wird, s_wird
finde(n) es	find_s, finden_s
fünfundneunzig	fümmunneunzig
gab es	gab_s
gehen wir	gehma
geht es	geht_s
gibt es	gibt_s
habe es	hab_s
haben sie	hamse, haben se, ham sie
haben wir	hamwa, ham wir, hamma, haben wa
hast es	hast_s
hätten wir	hättma
hinter dem	hinterm
ich es	ich_s
ihr es	ihr_s
in den, in einen	in nen, innen
ist denn	is_n
ist ein	is n, is_n
ist eine	is ne, is_ne
ist es	isses
ist so	isso
kann man	kamma, kamman, kann ma
kommen wir	kommen wa, komma
können sie es	könnses

Standardsprache	Transkribierte Realisierungen
können wir	kömma
macht es	macht_s
man es	man_s
mit dem/einem/den/einen	mit_m, mi_m, mit_n
miteinander	mitnanner
nochmal	nomma
sag es	sag_s
schon ein	schon n
sehen es	sehen_s
sie es	ses, sie_s
sie es denn	sesn
sind es	sind_s
sind wir	simma
so ein	so n, so_n
so eine	so ne; so_ne
so noch nicht	so no nich
solche	sone
war es	war_s
waren es	waren_s
weil es	weil_s
wenn man	wenn ma, wemma, wemman
werden	wern
wie es	wie_s
wo es	wo_s
wollen es	wollen_s
wollten es	wollten_s

Anhang 3

Im Folgenden finden Sie eine Tabelle mit den unter Volltranskripten und für die Konkordanzsuche zur Verfügung stehenden Subkorpora des Kernkorpus. Die Tabelle wird demnächst um die Angaben der assoziierten Korpora ergänzt.

	Kürzel ⁶	Erläuterung					
1	DEU_L1_PG	deutschsprachige	Prüfungsgespräche	mit	Muttersprachlern	im deutschen	akademischen Kontext
2	DEU_L1_EV	deutschsprachige	Expertenvorträge	mit	Muttersprachlern	im deutschen	akademischen Kontext
3	DEU_L1_SV	deutschsprachige	Studentische Vorträge	mit	Muttersprachlern ⁷	im deutschen	akademischen Kontext
4	DEU_L2_UK_PG	deutschsprachige	Prüfungsgespräche	mit	Nicht-Muttersprachlern	im britischen	akademischen Kontext
5	DEU_L2_UK_EV	deutschsprachige	Expertenvorträge	mit	Nicht-Muttersprachlern	im britischen	akademischen Kontext
6	DEU_L2_UK_SV	deutschsprachige	Studentische Vorträge	mit	Nicht-Muttersprachlern	im britischen	akademischen Kontext
7	DEU_L2_PL_PG	deutschsprachige	Prüfungsgespräche	mit	Nicht-Muttersprachlern	im polnischen	akademischen Kontext
8	DEU_L2_PL_EV	deutschsprachige	Expertenvorträge	mit	Nicht-Muttersprachlern	im polnischen	akademischen Kontext
9	DEU_L2_PL_SV	deutschsprachige	Studentische Vorträge	mit	Nicht-Muttersprachlern	im polnischen	akademischen Kontext
10	DEU_L2_D_PG	deutschsprachige	Prüfungsgespräche	mit	Nicht-Muttersprachlern	im deutschen	akademischen Kontext
11	DEU_L2_D_SV	deutschsprachige	Studentische Vorträge	mit	Nicht-Muttersprachlern	im deutschen	akademischen Kontext
12	ENG_L1_PG	englischsprachige	Prüfungsgespräche	mit	Muttersprachlern	im britischen	akademischen Kontext
13	ENG_L2_PG	englischsprachige	Prüfungsgespräche	mit	Nicht-Muttersprachlern	im britischen	akademischen Kontext
14	ENG_L1_EV	englischsprachige	Expertenvorträge	mit	Muttersprachlern	im britischen	akademischen Kontext
15	ENG_L2_EV	englischsprachige	Expertenvorträge	mit	Nicht-Muttersprachlern	im britischen	akademischen Kontext
16	ENG_L1_SV	englischsprachige	Studentische Vorträge	mit	Muttersprachlern	im britischen	akademischen Kontext
17	ENG_L2_SV	englischsprachige	Studentische Vorträge	mit	Nicht-Muttersprachlern	im britischen	akademischen Kontext
18	POL_PG	polnischsprachige	Prüfungsgespräche	mit	Muttersprachlern	im polnischen	akademischen Kontext
19	POL_EV	polnischsprachige	Expertenvorträge	mit	Muttersprachlern	im polnischen	akademischen Kontext
20	POL_SV	polnischsprachige	Studentische Vorträge	mit	Muttersprachlern	im polnischen	akademischen Kontext

⁶ L1 und L2 in den Kürzeln bezieht sich jeweils auf die „Muttersprachlichkeit“ der Kommunikation, d.h. im Vortrag > Vortragende, im Prüfungsgespräch > Prüflinge und Prüfer. Welche Sprachkompetenz weitere an der Kommunikation beteiligte Sprecher(innen) wie Seminarleitende haben, ist in den Metadaten verzeichnet.

⁷ In einigen Teilkorpora kommen jeweils dieselben sogenannten gemischten Kommunikationen vor: Bei den SV halten den Vortrag beispielsweise nicht-muttersprachliche (Einteilung bei L2) und muttersprachliche (Einteilung bei L1) Studierende als Gruppe.

Für die Konkordanzsuche stehen Ihnen zusätzlich folgende Teilkorpora zur Auswahl:

	Kürzel⁸	Erläuterung					
21	ENG_L1	alle englischsprachigen Kommunikationen	mit	Muttersprachlern	im	britischen	akademischen Kontext
22	ENG_L2	alle englischsprachigen Kommunikationen	mit	Nicht-Muttersprachlern	im	britischen	akademischen Kontext
23	DEU_L1	alle deutschsprachigen Kommunikationen	mit	Muttersprachlern	im	deutschen	akademischen Kontext
24	DEU_L2	alle deutschsprachigen Kommunikationen	mit	Nicht-Muttersprachlern	in	allen	akademischen Kontexten
25	ENG	alle englischsprachigen Kommunikationen	mit	Muttersprachlern und Nicht-Muttersprachlern	im	britischen	akademischen Kontext
26	DEU	alle deutschsprachigen Kommunikationen	mit	Muttersprachlern und Nicht-Muttersprachlern	in	allen	akademischen Kontexten
27	POL	alle polnischsprachigen Kommunikationen	mit	Muttersprachlern	im	polnischen	akademischen Kontext

⁸ L1 und L2 in den Kürzeln bezieht sich jeweils auf die „Muttersprachlichkeit“ der Kommunikation, d.h. im Vortrag > Vortragende, im Prüfungsgespräch > Prüflinge und Prüfer. Welche Sprachkompetenz weitere an der Kommunikation beteiligte Sprecher(innen) wie Seminarleitende haben, ist in den Metadaten verzeichnet.

Anhang 4

Dieser Anhang enthält eine Tabelle aller standardisierten Ausdrücke, wie sie zur Beschreibung der Metadaten angewandt wurden. Die Tabelle wird demnächst um die Angaben der assoziierten Korpora ergänzt.

Rubrik	verwendete Ausdrücke	Erläuterungen
Rubrik Communication		
Muttersprachliche Kommunikation	Ja	
	nein	
	gemischt	
Genre	Expertenvortrag	
	Studentischer Vortrag	
	Prüfungsgespräch	
Art des Vortrags	Konferenzvortrag	<i>bei Expertenvorträgen</i>
Art des Vortrags	Einzel-/Gruppenvortrag im Bachelor-Studiengang	<i>bei Studentischen Vorträgen</i>
	Einzel-/Gruppenvortrag im Master-Studiengang	
	Einzel-/Gruppenvortrag im Magisterstudiengang	
Art des Prüfungsgesprächs	Bachelor-Prüfung	
	Bachelor-Modulprüfung	
	Master-Prüfung	
	Master-Modulprüfung	
	Magisterzwischenprüfung	
	Magisterabschlussprüfung	
	Lehramtsprüfung	
Zusatzmaterial	Handout	<i>alle Arten von Handreichungen an das Publikum, bspw. Foliendrucke, extra gestaltete Zusammenfassungen, Beispiel-, Zitat- oder Arbeitsblätter etc.</i>
	Präsentation	<i>PowerPoint-Folien</i>
	Skript	<i>schriftliche Vorformulierungen oder Stichworte der Vortragenden selbst für den Vortrag</i>
	Audiobeispiele	
	Videobeispiele	
	Thesenpapier	<i>schriftliche Grundlage für eine mündlich</i>

Rubrik	verwendete Ausdrücke	Erläuterungen
		Prüfung, Stichworte zu einem Themenkomplex, die der Prüfling anhand Literaturrecherche vorbereitet und die in der Prüfung aufgegriffen werden können
	OHP-Folien	Folien, die mit einem Overheadprojektor an die Wand projiziert werden
	fehlt	wenn nichts vorliegt
Die Rubrik Location		
Raum	Seminarraum	
	Hörsaal	
	Konferenzraum	wenn Konferenzvorträge nicht in universitären Gebäuden stattfanden
	Büro	z.B. bei Prüfungsgesprächen
Die Rubrik Languages		
Basissprache	deu	die Basissprache der Interaktion; Angabe des ISO Language Codes gemäß 639-3 (http://www-01.sil.org/iso639-3/codes.asp)
	eng	
	pol	
Grad der Mündlich-keit	frei gesprochen	
	zum Teil abgelesen	
	vollständig abgelesen	
	scheint vorformuliert und auswendig gelernt zu sein	
Wechsel in andere Sprache(n)	eng	Weitere Sprachen, die in der Kommunikation benutzt wurde; Angabe des ISO Language Codes gemäß 639-3
Die Rubrik Setting		
Verwendete Medien	Präsentation	
	Handout	
	Skript	
	Thesenpapier	
	OHP-Folien	
	Audiobeispiele	
	Videobeispiele	
	keine	
Die Rubrik Speaker		
Rollen	Vortragender	Studierender oder Experte, der einen Vortrag hält

Rubrik	verwendete Ausdrücke	Erläuterungen
	Seminarleiter	<i>Dozent, der ein Seminar leitet und einen studentischen Vortrag anmoderiert und die anschließende Diskussion leitet</i>
	Prüfling	<i>Studierender, der geprüft wird</i>
	Prüfer	<i>Person, die prüft, d.h. in der Prüfung Prüfungsfragen stellt. In einer Prüfung können auch mehrere Prüfer anwesend sein, wenn bspw. ein Prüfling in der Abschlussprüfung zu verschiedenen Themen von den entsprechenden Prüfern befragt wird</i>
Erstsprache		Angabe des ISO Language Codes gemäß 639-3 (http://www-01.sil.org/iso639-3/codes.asp)
L2		Angabe des ISO Language Codes gemäß 639-3 (http://www-01.sil.org/iso639-3/codes.asp)