

## ZUM EINSATZ DES GeWiss-KORPUS IM RAHMEN EINER MEHRSPRACHIGKEITSDIDAKTISCHEN VERMITTLUNG DER GESPROCHENEN WISSENSCHAFTSSPRACHE

Katharina Salzmann  
Freie Universität Bozen

### Abstract

Im vorliegenden Beitrag soll der Frage nachgegangen werden, wie das GeWiss Korpus, ein Vergleichskorpus zur gesprochenen Wissenschaftssprache, im Rahmen des mehrsprachig angelegten, universitären DaF-Unterrichts eingesetzt werden kann, um zum Auf- und Ausbau von meta- und zwischensprachlicher Bewusstheit beizutragen und um Fähigkeiten des kompetenten Sprachvergleichs zu fördern. Zu diesem Zweck sollen zwei exemplarische Korpusanalysen durchgeführt werden, anhand derer ein typisches Nomen der Wissenschaftssprache (dt. *Begriff / Konzept*, it. *concetto*, engl. *concept*) sowie ein Diskursmarker (dt. *entschuldigung*, it. *scusate*, engl. *sorry*) unter die Lupe genommen werden, um anschließend jeweils das spezifische Lernpotential einer solchen Korpusrecherche herauszuarbeiten.

**Keywords:** Korpora in DaF; gesprochene Wissenschaftssprache; Mehrsprachigkeitsdidaktik; Sprachbewusstheit

### Abstract

This paper aims at illustrating in which way the GeWiss corpus, a comparable corpus of spoken academic language, can be employed in a multilingual approach to teaching German as a foreign language and how its use can contribute to fostering and enhancing metalinguistic and crosslinguistic awareness as well as skills of competent language comparison among university students. For this purpose, two exemplary corpus analyses will be carried out, aimed at examining a typical noun of academic language (ger. *Begriff / Konzept*, it. *concetto*, engl. *concept*) as well as a discourse marker (ger. *entschuldigung*, it. *scusate*, engl. *sorry*), in order to subsequently highlight the specific learning potential of such a corpus search.

**Keywords:** corpora in teaching GFL; spoken academic language; multilingual didactics; language awareness

## 1. Einleitende Bemerkungen: Korpora im mehrsprachig angelegten DaF-Unterricht

Im Rahmen der Korpuslinguistik werden Korpora als für einen bestimmten Zweck zusammengestellte, meist digitalisierte Sammlungen schriftlicher oder gesprochener Äußerungen definiert, die neben den Sprachdaten auch Metadaten und linguistische Annotationen enthalten können (Lemnitzer / Zinsmeister 2015: 13). Anfänglich wurden Korpora vorwiegend für sprachwissenschaftliche Untersuchungen genutzt, in den letzten Jahren haben sich zudem immer mehr Forschende für einen verstärkten Einsatz von Korpora in der Fremdsprachenvermittlung, u.a. auch in der DaF-Didaktik, ausgesprochen (u.a. Lüdeling / Walter 2009; Flinz 2020; Flinz / Hufeisen 2021). Es gibt einige Gründe, die dafür sprechen, Korpora im Fremdsprachenunterricht einzusetzen (Lüdeling / Walter 2019: 3-13): Für Lehrende erleichtern sie die Erstellung von Unterrichtsmaterialien mit authentischen Daten und bieten einen Orientierungspunkt für die Korrektur, für Lernende ermöglichen sie ein datengesteuertes Lernen. Bei vorwiegend qualitativen Analysen können die Lernenden anhand von Konkordanzen, d.h. einem Suchausdruck innerhalb seines unmittelbaren Kontextes, selbstständig Gebrauchsregeln ableiten, was zum Ausbau von Wortschatz- und Grammatikkenntnissen beiträgt und das Bewusstsein für den tatsächlichen Sprachgebrauch schärft. Bei stärker quantitativ orientierten Analysen kann durch

den Einsatz von Korpora im Sprachunterricht außerdem das Bewusstsein für das Verhältnis von Norm und Varianz sowie für Unterschiede zwischen verschiedenen Varietäten (diaphasisch, diamesisch, diatopisch etc.) gefördert werden, indem z.B. Frequenzlisten erstellt und untersucht werden, ausgehend von einem Vergleich fachsprachlicher und allgemeinsprachlicher Korpora ein bestimmtes Fachvokabular ermittelt wird oder Kollokationen, d.h. typische Kombinationen von Wörtern (s. dazu Wallner 2014: 155 sowie Kapitel 3.1), analysiert werden. Hat man Korpora in verschiedenen Sprachen zur Verfügung, können die Textsammlungen zudem als Mittel des authentischen Sprachvergleichs dienen. Von Bedeutung ist dabei in jedem Fall, dass die Lernenden ein spezifisches Korpustraining erhalten, um schrittweise eine Korpusbenutzungskompetenz, eine sogenannte *corpus literacy* (Mukherjee 2002: 179-180), aufbauen zu können.

Aufgrund ihrer spezifischen pädagogischen Potenziale – insbesondere Förderung der Sensibilität für sprachliche Strukturen und deren Gebrauch, Möglichkeit des Daten- und Sprachvergleichs – eignen sich Korpora besonders gut für einen mehrsprachig angelegten DaF-Unterricht auf universitärer Ebene, in dessen Rahmen die Sprachbewusstheit und die Fähigkeit zum kompetenten Sprachvergleich gefördert werden sollen.

Die Sprachbewusstheit wird im *Dynamischen Modell der Mehrsprachigkeit* (Jessner 2008) als Kernelement der mehrsprachigen Kompetenz betrachtet. Die mehrsprachige Kompetenz, d.h. all jene Eigenschaften, die einen mehrsprachigen Sprecher bzw. eine mehrsprachige Sprecherin von einem bzw. einer einsprachigen unterscheiden, wird als dynamische Interaktion zwischen den einzelnen Sprachsystemen (L1, L2, L3 usw.), der zwischensprachlichen Interaktion (CLI = *cross-linguistic interaction*) und dem M(*ultilingualism*)-Faktor (Sprachbewusstheit und Sprachlernstrategien) beschrieben. Bei der Sprachbewusstheit (engl. *language awareness*) können zwei Dimensionen unterschieden werden (Jessner / Hufeisen 2019: 74): 1. die metasprachliche Bewusstheit (*metalinguistic awareness*), d.h. die Fähigkeit, abstrakt über Sprache nachdenken zu können; 2. die zwischensprachliche Bewusstheit (*cross-linguistic awareness*), d.h. das implizite und explizite Bewusstsein über Verbindungen zwischen den Sprachsystemen. Letztere wird insbesondere durch den Sprachvergleich, d.h. die Gegenüberstellung und das Herstellen von Bezügen zwischen zwei oder mehreren Sprachen, gefördert. Als didaktisches Kernelement der mehrsprachigkeitsdidaktischen Ansätze soll der Sprachvergleich den Lernprozess erleichtern (Lernstrategie) und dazu beitragen, die Sprachkompetenzen zu festigen (s. Salzmann 2020: 111).

Ziel dieses Beitrags ist es, anhand ausgewählter Beispielanalysen die Einsatzmöglichkeiten des GeWiss-Korpus im Rahmen einer mehrsprachig (Deutsch-Italienisch-Englisch) angelegten Vermittlung der Wissenschaftssprache aufzuzeigen und dabei zu veranschaulichen, wie der Einsatz des Korpus zur Förderung der meta- und zwischensprachlichen Bewusstheit der Studierenden sowie zu einer erhöhten interlingualen Vergleichsfähigkeit beitragen kann. Der Beitrag gliedert sich wie folgt: In Kapitel 2 wird das GeWiss-Korpus vorgestellt, indem näher auf das Korpusdesign, die Korpus-tools und die Besonderheiten der gesprochenen Wissenschaftssprache als Varietät eingegangen wird. In Kapitel 3 werden mithilfe der Konkordanzsuche zwei exemplarische Korpusanalysen – die erste zu einem typischen Nomen der Wissenschaftssprache, die zweite zu einem Diskursmarker als gesprochensprachliches Phänomen – durchgeführt, anhand derer das inhärente Lernpotential für die Studierenden herausgearbeitet wird. Abgerundet wird der Beitrag durch einige Schlussfolgerungen und einen kurzen Ausblick, in dem weitere Einsatzmöglichkeiten des Korpus aufgezeigt werden (Kapitel 4).

## 2. Das GeWiss-Korpus: ein Spezialkorpus der gesprochenen Sprache

Umfangreiche Korpora zur gesprochenen Sprache stehen erst seit relativ kurzer Zeit zur Verfügung. Die *Datenbank für gesprochenes Deutsch* (DGD)<sup>1</sup>, die am Leibniz-Institut für Deutsche Sprache (IDS, Mannheim) zusammengestellt wurde, umfasst zahlreiche Korpora unterschiedlicher Typologie und unterschiedlichen Umfangs; besonders erwähnenswert für die Zwecke der vorliegenden Arbeit sind dabei das *Forschungs- und Lehrkorpus Gesprochenes Deutsch* (FOLK)<sup>2</sup> sowie das Korpus *Gesprochene Wissenschaftssprache kontrastiv* (GeWiss)<sup>3</sup>. Bei beiden Korpora handelt es sich um Gesprächskorpora, d.h. sie enthalten aufgezeichnete Gespräche authentischer Interaktionen, die weitestgehend aus spontanen Äußerungen bestehen (Schmidt 2018: 210). Während es sich bei FOLK um ein einsprachig-deutsches, allgemeinsprachliches Referenzkorpus handelt<sup>4</sup>, liegt mit GeWiss ein Spezialkorpus aus der Domäne der Wissenschaftskommunikation vor. Insbesondere handelt es sich um ein mehrsprachiges Vergleichskorpus<sup>5</sup> zur gesprochenen Wissenschaftssprache, das nicht nur als Grundlage für sprachwissenschaftliche Untersuchungen dient, sondern auch eine wichtige Rolle für die Vermittlung des Deutschen als Wissenschaftssprache spielt (Fandrych / Meißner / Slavcheva 2014)<sup>6</sup>. Das GeWiss-Korpus umfasst Daten in drei bzw. vier verschiedenen Sprachen, und zwar Deutsch, Englisch, Polnisch und Italienisch<sup>7</sup>, und enthält unterschiedliche Gattungen der Mündlichkeit, nämlich Vorträge von Experten und Expertinnen, Studierendenreferate und mündliche Prüfungen aus sprachphilologischen Fächern. Zusätzlich wurden zu allen aufgenommenen Ereignissen soziolinguistische Metadaten zu den Sprechenden erhoben. Aufgrund seiner Zusammenstellung ermöglicht das Korpus unterschiedliche Vergleichsdimensionen (s. Abb. 1): 1. Gebrauch des Deutschen als fremder Wissenschaftssprache in verschiedenen akademischen Kontexten (L2-Daten von wissenschaftlich Tätigen mit nicht-deutscher Erstsprache); 2. Deutsch als Wissenschaftssprache im Vergleich zu Englisch, Polnisch, Italienisch als Wissenschaftssprache (kontrastive L1-Daten von wissenschaftlich Tätigen, die jeweils in ihrer Erstsprache vortragen).

<sup>1</sup> Die Datenbank ist zugänglich unter: [https://dgd.ids-mannheim.de/dgd/pragdb.dgd\\_extern.welcome](https://dgd.ids-mannheim.de/dgd/pragdb.dgd_extern.welcome) (01.07.2020).

<sup>2</sup> Abrufbar unter: <http://agd.ids-mannheim.de/folk.shtml> (01.07.2020).

<sup>3</sup> Abrufbar unter: <https://gewiss.uni-leipzig.de/index.php?id=home> (01.07.2020).

<sup>4</sup> FOLK umfasst als „nationales Gesprächskorpus“ 259 Gespräche im Umfang von 202 Stunden (ca. 2 Millionen Tokens), die verschiedene Gesprächstypen aus den Bereichen privater, institutioneller und öffentlicher Kommunikation enthalten (Schmidt 2018: 216).

<sup>5</sup> GeWiss ist ein Vergleichskorpus, da es sich im Unterschied zu einem Parallelkorpus um unabhängig voneinander entstandene Texte und keine Übersetzungen handelt (zur Korpus Typologie s. Lemnitzer / Zinsmeister 2015: 137-142).

<sup>6</sup> Weitere Datensammlungen, die speziell auf die Bedürfnisse von Lehrenden und Lernenden des Deutschen zugeschnitten sind, sind die Datenbank *Gesprochenes Deutsch für die Auslandsgermanistik* sowie die Plattform *Gesprochenes Deutsch – authentische Alltagsinteraktionen für die Forschung und Praxis im Bereich DaF und DaZ* (Imo / Weidner 2018: 238-241), abrufbar unter <https://daad-gda.sprache-interaktion.de/> bzw. <https://dafdaz.sprache-interaktion.de/> (01.07.2020). Anhand dieser Korpora können die Charakteristika der spontanen gesprochenen Sprache vermittelt und die mündlichen Kommunikationsfähigkeiten der Studierenden im Hinblick auf unterschiedliche Situations- und Kommunikationszwecke gefördert werden (Bachmann-Stein 2013: 51-52).

<sup>7</sup> Das italienische Teilkorpus, das ausschließlich wissenschaftliche Vorträge enthält, wurde von der Verfasserin im Rahmen ihrer Dissertation gesammelt und aufbereitet (s. dazu ausführlich Salzmann 2017).

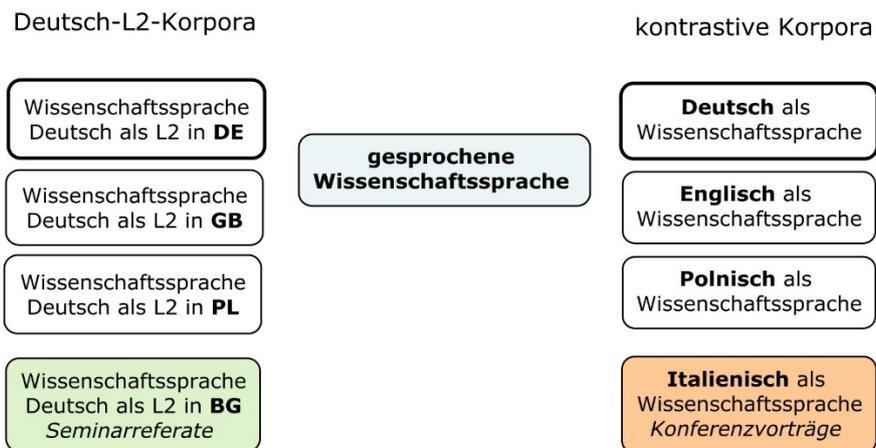


Abbildung 1  
Vergleichsdimensionen des GeWiss-Korpus<sup>8</sup>

Das GeWiss-Korpus ist nach Registrierung öffentlich zugänglich. Es besitzt unterschiedliche Suchfunktionen, und zwar die Recherche in den Volltexten, d.h. in den mit den Audiodateien alignierten Transkripten, sowie die Recherche nach Konkordanzen (KWIC-Format = *keywords in context*). Diese Funktion ermöglicht die Recherche eines Suchbegriffs im Korpus, wobei man auswählen kann, welche Teilkorpora durchsucht werden sollen. Der Suchausdruck wird dann innerhalb seines unmittelbaren linken und rechten Kontextes angezeigt (Abb. 2):

EV_DE_098	IT_0235	soziale praxis °h ansieht °h ma n sich auch die	frage	stellen muss wer gewissermaße n diese praxis praktiziert (1.2) also
EV_DE_098	IT_0235	solche prozessorientierung eine größere aufmerksam keit (0.5) für die	frage	wie und von wem diskursive wir klichkeit hergestellt wird °h geg enüber

Abbildung 2  
Konkordanzen zum Suchausdruck *Frage*<sup>9</sup>

Wie aus den beiden Beispielen für das Nomen *Frage* ersichtlich ist, kann man mithilfe der Konkordanzsuche die Verwendungsweise und syntaktische Einbettung des Suchausdrucks erfragen, wie hier z.B. die Kollokation *eine Frage stellen*<sup>10</sup> sowie das Gebrauchsmuster *die Frage* + indirekter Fragesatz (*die Frage, wie und von wem diskursive Wirklichkeit hergestellt wird*), wobei mit *Gebrauchsmuster* das Auftreten in typischen Ausdrucksroutinen lexikogrammatischer Natur (s. dazu Meißner 2020: 65 sowie Kapitel 3.1) gemeint ist.

<sup>8</sup> Aus: [https://gewiss.uni-leipzig.de/index.php?id=corpus\\_design](https://gewiss.uni-leipzig.de/index.php?id=corpus_design) (01.07.2020).

<sup>9</sup> Aus: <https://gewiss.uni-leipzig.de/index.php?id=search> (01.07.2020).

<sup>10</sup> *Eine Frage stellen* wird in der Literatur auch spezifischer als Funktionsverbgefüge (FVG) bezeichnet. Dabei handelt es sich um eine Verbindung eines lexikalisch verblassten Funktionsverbs mit einem inhaltsstärkeren abstrakten Nomen. Aus: <https://grammis.ids-mannheim.de/terminologie/85> (14.06.2021). Funktionsverbgefüge werden von Linguisten wie Burger (2015: 57-58) in die Gruppe der Kollokationen aufgenommen.

Nach der Phase der Datensammlung und -transkription wurde das GeWiss-Korpus digital aufbereitet: Dazu wurde das Teilkorpus der deutschsprachigen Vorträge (ca. 87 000 Token<sup>11</sup>) zuerst von der Transkription in eine orthographisch normalisierte Version überführt und danach der Wortarten-Annotation (POS-Tagging = *parts of speech*) und der Lemmatisierung unterzogen (Fandrych 2017)<sup>12</sup>. In den Lemmalisten<sup>13</sup> werden verschiedene Wortformen zu einer Grundform (Lemma) zusammengeführt, so wie sie im Wörterbuch zu finden wäre. So werden z.B. Formen wie gesprochen sprachlich *haste* in den schriftsprachlichen Standard *hast du* überführt und unter dem Verb *haben* lemmatisiert. Die vollständige Lemmaliste kann alphabetisch oder nach der Frequenz geordnet werden, zusätzlich sind die Lemmalisten nach Wortarten (Adjektiv, Präposition, Nomen, Verb, Diskursmarker<sup>14</sup>) durchsuchbar. Die Liste führt zu jedem Lemma die Häufigkeit seines Vorkommens auf. Wie aus Abb. 3 ersichtlich ist, enthält die Lemmaliste in der Wortart Nomen insgesamt 12 612 Belege, für jedes Lemma wird die Anzahl der Vorkommnisse separat angegeben. So finden sich für das häufigste Nomen *Beispiel* 227 Belege, für die zweithäufigsten Nomina *Sprache* und *Frage* jeweils 158 Belege. Ist ein Lemma der Häufigkeitsklasse 5 bzw. 6 zugeordnet, bedeutet dies, dass das häufigste Lemma der Gesamtliste (*die*) etwa 25 bzw. 26 mal so häufig auftritt.

## Lemmalisten (DEU\_L1\_EV)

### alle Lemmata

Adjektive - Präpositionen - Nomina - Verben - Diskursmarker

### alphabetisch - Frequenz

\*Häufigkeitsklasse: Das häufigste Lemma (*die*) ist etwa  $2^n$  mal so häufig wie ein Lemma der Klasse n.

SUMME	12612	
		Klasse*
Beispiel	227	5
Sprache	158	6
Frage	158	6
Raum	147	6
Text	142	6

Abbildung 3  
Lemmaliste geordnet nach Nomina<sup>15</sup>

<sup>11</sup> Ein Token oder Worttoken bezeichnet das Vorkommen eines Wortes an einer bestimmten Stelle im Korpus (Lemnitzer / Zinsmeister 2015: 198).

<sup>12</sup> Zu den verschiedenen Möglichkeiten linguistischer Annotation von Korpora s. Lemnitzer / Zinsmeister (2015: 57-89).

<sup>13</sup> Die Listen sind abrufbar unter: [https://gewiss.uni-leipzig.de/index.php?id=lemma\\_lists](https://gewiss.uni-leipzig.de/index.php?id=lemma_lists) (01.07.2020).

<sup>14</sup> Die Kategorie des Diskursmarkers zählt nicht zu den klassischen Wortarten, wird aber seit Ende der 1990er Jahre von den Vertretern der deutschen Gesprächsforschung immer häufiger verwendet (s. u.a. Auer / Günthner 2005; Imo 2012). S. auch Kapitel 3.2.

<sup>15</sup> Aus: [https://gewiss.uni-leipzig.de/index.php?id=lemma\\_lists](https://gewiss.uni-leipzig.de/index.php?id=lemma_lists) (01.07.2020).

Außerdem wurden im deutschen Teilkorpus im Zuge der digitalen Aufbereitung Annotationen pragmatischer Phänomene vorgenommen, u.a. zu Diskurskommentierungen, Zitaten und Verweisen (Fandrych 2017: 22-26).

Hinsichtlich der Einordnung der gesprochenen Wissenschaftssprache auf dem Varietätenspektrum<sup>16</sup> kann festgehalten werden, dass diese Merkmale zweier Subsysteme der deutschen Sprache aufweist, einerseits der Wissenschaftssprache (z.B. *einer Frage nachgehen*), die auf der Ebene der diaphasischen Variation (Alltagssprache-Fachsprache-Wissenschaftssprache) anzuordnen ist, und andererseits der gesprochenen Sprache (z.B. *tschuldigung*, s. auch 3.2), welche die diamesische Variation, d.h. das Kontinuum zwischen Schriftdeutsch und gesprochenem Deutsch, betrifft. Unter Bezugnahme auf das Modell von Koch / Oesterreicher (2007) zeigt sich, dass die gesprochene Wissenschaftssprache Charakteristika unterschiedlichen Formalitätsgrades besitzt, d.h. formale wissenschaftssprachliche Strukturen (WS), die zum Pol der konzeptionellen Schriftlichkeit hintendieren, sowie relativ informale gesprochensprachliche Strukturen (GS), die eher am Pol der konzeptionellen Mündlichkeit anzusiedeln sind. In Abb. 4 wird die Position der gesprochenen Wissenschaftssprache mittig auf dem Kontinuum zwischen konzeptioneller Mündlichkeit (Sprache der Nähe) und konzeptioneller Schriftlichkeit (Sprache der Distanz) angenommen, wobei sich die gesprochene Wissenschaftssprache – je nach konkreter Realisierung – auch stärker am Pol der konzeptionellen Mündlichkeit bzw. an jenem der Schriftlichkeit ausrichten kann.

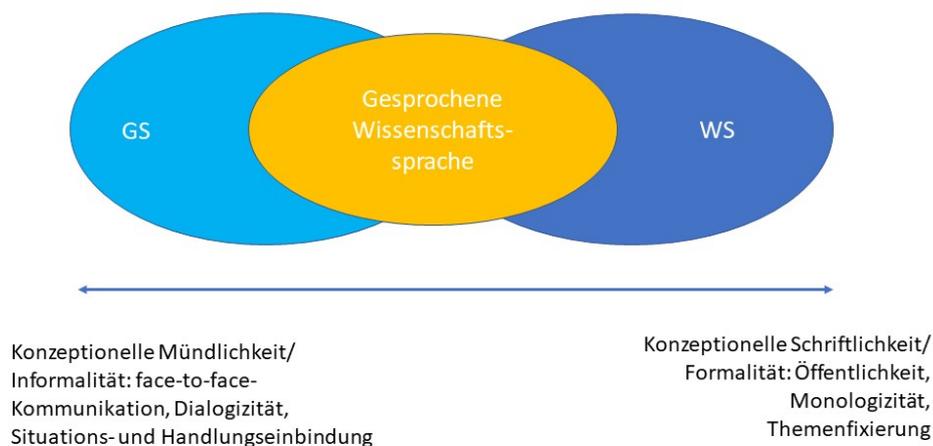


Abbildung 4  
Gesprochene Wissenschaftssprache auf der Achse der  
konzeptionellen Mündlichkeit / Schriftlichkeit (Koch / Oesterreicher 2007)

Zum Zweck der wissenschaftssprachdidaktischen Vermittlung gilt es zu beachten, dass im Unterricht sowohl ausgewählte Merkmale der Wissenschaftssprache als auch relevante Charakteristika der gesprochenen Sprache Berücksichtigung finden müssen, wenn man die Studierenden zu einer kompetenten Verwendung der gesprochenen Wissenschaftssprache befähigen will (Salzmann 2019).

<sup>16</sup> Unter *Varietät* verstehe ich in diesem Beitrag ein Subsystem, das sich teilweise der Mittel der deutschen Standardsprache bedient, sich aber teilweise von den anderen Varietäten durch die Verwendung bestimmter grammatischer und lexikalischer Mittel unterscheidet, die durch die jeweilige Dimension bedingt sind (Foschi Albert 2017: 57).

### 3. Exemplarische Korpusanalyse

In diesem Kapitel sollen auf der Grundlage des Teilkorpus der Experten- und Expertinnenvorträge zwei exemplarische Korpusanalysen durchgeführt werden, eine zu einem Phänomen der Wissenschaftssprache und eine zu einem Phänomen der gesprochenen Sprache. Nach der Auswertung der Daten werde ich jeweils darstellen, wie durch den Einsatz des GeWiss-Korpus eine Bewusstmachung grundlegender Eigenschaften der Wissenschaftssprache (ein typisches Nomen der allgemeinen Wissenschaftssprache) sowie der gesprochenen Sprache (ein Diskursmarker) im Sprachvergleich Deutsch-Italienisch-Englisch erfolgen kann.

#### 3.1 Bedeutungen und Kollokationen eines Nomens der alltäglichen Wissenschaftssprache am Beispiel des Suchausdrucks dt. *Begriff* und *Konzept* / it. *concetto* / engl. *concept*

Für die erste Analyse wurde ein typisches Nomen der alltäglichen oder allgemeinen Wissenschaftssprache (=AWS) ausgewählt, nämlich dt. *Begriff* und das sinnverwandte *Konzept* sowie die Entsprechungen it. *concetto* und engl. *concept*. Der Begriff der AWS geht auf Ehlich (1993: 33) zurück und meint jene fachübergreifenden Ausdrucksmittel, die der Versprachlichung wissenschaftsmethodologischer Inhalte dienen und durch wissenschaftliche Handlungen (u.a. Benennen, Vergleichen, Verweisen) entstehende Zusammenhänge darstellen.

Durch die Analyse der Konkordanzen im deutschen, italienischen und englischen Teilkorpus soll der Frage nachgegangen werden, welche spezifischen Bedeutungen das ausgewählte Nomen besitzt und welche typischen Kollokationen und syntaktischen Muster es jeweils eingeht. Die Kollokationen der Wissenschaftssprache, insbesondere die Verb-Nomen-Verbindungen, sind zwar auf der Ebene des individuellen Lemmas einzelsprachspezifisch geprägt – weshalb sie häufig einen Problem- bereich für Fremdsprachenlernende darstellen –, weisen aber gleichzeitig einige charakteristische sprachenübergreifende Merkmale auf, die sich für eine mehrsprachigkeitsdidaktische Vermittlung anbieten (s. dazu Meißner 2020)<sup>17</sup>. In Anlehnung an Wallner (2014: 55) sind Kollokationen binäre, nicht bzw. schwach idiomatische Wortverbindungen, zwischen deren Bestandteilen entweder eine attributive (Adjektiv + Nomen) oder eine auf Satzgliedfunktionen beruhende syntaktische Relation (Nomen + Verb) besteht und die durch ein präferiertes Miteinandervorkommen, verbunden mit mehr oder weniger stark ausgeprägten Einschränkungen hinsichtlich der Substituierbarkeit der Bestandteile durch bedeutungsähnliche Ausdrücke, gekennzeichnet sind<sup>18</sup>. Eine Kollokation besteht dabei aus einer Basis, die den Kollokator bestimmt: So werden zur Kollokationsbasis *Begriff* und *Konzept* im Lehrbuch *Wissenschaftssprache Deutsch* von Graefen / Moll (2011: 127) u.a. die folgenden Verbindungen angegeben: *einen Begriff / ein Konzept entwickeln, prägen, einführen, präzisieren, verwenden, ein weiter / enger Begriff*. Außerdem sind für ein Nomen wie *Begriff* bestimmte Gebrauchsmuster belegt, darunter verstehe ich das usuelle Vorkommen eines Ausdrucks in einem bestimmten syntaktischen Kontext, insbesondere in Bezug auf syntaktische Funktionen (Subjekt, Objekt), z.B. *ein Begriff (S) setzt sich durch*, sowie auf die Verbindung mit anderen Phrasen, z.B. NP + PP mit *von*: *ein weiter Begriff von Wissenschaft* (zu Gebrauchsmustern von wissenschaftssprachlichen Verben s. Meißner 2020: 65). Umschrieben wird der Ausdruck *Begriff / Konzept* in seiner wissenschaftssprachlichen Verwendung

<sup>17</sup> So sind die Verben vieler Wissenschaftssprachen u.a. durch Polysemie, Vagheit, bestimmte Gebrauchsmuster und eine formseitige Unauffälligkeit gekennzeichnet (Meißner 2020: 64-67). Zur Zusammensetzung des stark figurativ geprägten Verbwortschatzes der deutschen Wissenschaftssprache s. Meißner (2020: 63).

<sup>18</sup> Zu Kollokationen in der Korpuslinguistik und zur Abgrenzung von Kookkurrenzen s. Lemnitzer / Zinsmeister (2015: 177-179). Für eine vertiefte Auseinandersetzung mit dem Kollokationsbegriff s. Belica / Perkuhn (2015).

von den beiden Autorinnen als „Inhalt einer wortgebundenen Vorstellung“, „komplexe Bedeutung sprachlicher Ausdrücke, [...] in einem erweiterten Sinn, ähnlich einer Theorie“ (Graefen / Moll 2011: 127).

### 3.1.1 Datenauswertung

Mithilfe der Funktion der Konkordanzanalyse könnten die Studierenden zunächst aufgefordert werden, die Frequenzen der Suchausdrücke zu ermitteln: Die Recherche ergibt dabei, dass für dt. *Begriff* 25 Belege vorliegen, für dt. *Konzept* 20, für it. *concetto* ebenfalls 20 Belege und für engl. *concept* 10.

Nach der quantitativen Analyse sollen die Bedeutungen der Suchausdrücke näher unter die Lupe genommen werden. Analysiert man das Lexem *Begriff* anhand ausgewählter Verwendungsbeispiele (1-3), ergeben sich verschiedene Bedeutungsnuancen:

- (1) würd ich auch plädieren für den begriff laienlinguistik h° äh der also von h° ja im deutschsprachigen raum h° von gerd antos h° äh propagiert eingeführt wurde (EV\_DE\_093)<sup>19</sup>
- (2) und ich werde äh heute wenn ich mich auf das der pronomens beziehe werd ich h° den begriff des äh de pronomens äh benutzen (EV\_DE\_100)
- (3) mittlerweile äh wird zumindest in annoncen und auch offiziellen texten intern der begriff des stamm(.)rollenprofessors (nicht mehr) verwendet (EV\_DE\_104)

In den Beispielen (1) und (2) geht es jeweils um einen Fachbegriff (Laienlinguistik, D-Pronomen) und die damit verbundene Vorstellung, während man in (3) den Eindruck hat, dass der Ausdruck *Begriff* weniger in der Bedeutung des dahinterstehenden Konzepts, sondern primär in der Bedeutung ‚Bezeichnung‘ verwendet wird, was linguistisch gesehen einer umgangssprachlichen Verwendung entspricht<sup>20</sup>. Der Vortragenden geht es nämlich darum, dass der Ausdruck *Stammrollenprofessor* an der mehrsprachigen Universität Bozen für das it. *professore di ruolo* nicht mehr verwendet wird.

Was die Wortverbindungen betrifft, können mögliche Kollokationen zunächst intuitiv aus dem Kontext des Suchausdrucks ermittelt und bei Bedarf durch andere Quellen abgesichert werden. Da viele Kollokationen im GeWiss-Korpus nur ein bis zwei Mal belegt sind, kann es nicht darum gehen, statistisch signifikante Häufigkeiten zu bestimmen; vielmehr sollen zu rein didaktischen Zwecken einige ‚gute‘, typische Verbindungen ermittelt werden.

Zu dem Suchausdruck *Begriff* finden sich im Korpus Adjektiv-Nomen-Kollokationen, z.B. *ein allgemeiner / umfassender / zentraler / schwammiger Begriff*, Nomen-Verb-Kollokationen, z.B. *einen Begriff benutzen, einen Begriff einführen* sowie das Gebrauchsmuster Nomen + NP im Nominativ bzw. Genitiv, z.B. *Begriff des D-Pronomens, der Begriff Raum, jener der Transliteration, der Begriff Volkslinguistik / Laienlinguistik*. Interessant ist zudem, dass die Wortverbindung in Beispiel (1) reformuliert wird: Der Sprecher wählt zuerst das Verb *propagieren*, das wohl kaum als usuelle Verbindung gesehen werden kann, und ersetzt dieses in der Folge durch das Verb *einführen*, welches hingegen eine typische Kollokation mit dem Nomen *Begriff* eingeht<sup>21</sup>.

Bei den Korpusbelegen zu dem Suchausdruck *Konzept* zeigt sich, dass das Lexem im weiten Sinne in der Bedeutung ‚Ansatz, Theorie‘ belegt ist (Bsp. 4). Auffallend ist auch hier die

<sup>19</sup> Die Transkripte wurden anhand des Minimaltranskripts GAT2 verschriftlicht (Selting et al. 2009). Für die verwendeten Transkriptionskonventionen s. Anhang.

<sup>20</sup> So weisen auch Graefen / Moll (2011: 127) darauf hin, dass das Wort *Begriff* oft in einer zweiten Bedeutung, nämlich im Sinne von ‚Ausdruck‘, benutzt wird, die es den Autorinnen zufolge eher zu vermeiden gilt.

<sup>21</sup> Im DWDS wird das Verb *einführen* unter den typischen Verbindungen zu dem Lexem *Begriff* angeführt, *propagieren* hingegen nicht (s. <https://www.dwds.de/wb/Begriff>, 01.07.2020).

Reformulierung des Verbs (*nehmen oder entwickeln*), eingeleitet durch den Häsitationsmarker *öh*, wobei wie in (1) das erste Verb nicht durch ein präferiertes Vorkommen mit der Basis *Konzept* gekennzeichnet ist, während das zweite Verb als guter Kollokator betrachtet werden kann<sup>22</sup>. In Beispiel (5) hat *Konzept* die Bedeutung ‚Ansatz‘, in (6) kann *Konzept* als ‚Vorstellung‘ oder ‚Idee‘ umschrieben werden, d.h. in diesem Fall die Vorstellung von *English as a lingua franca* (= ELF) als Varietät. In (7) geht es hingegen um die Bedeutungsseite im Gegensatz zur Ausdrucksseite, d.h. um die Frage, wie Bedeutungen gemeinsam mit den Wörtern gespeichert werden.

- (4) und es ist wichtig dass wir (.) in unseren (.) herangehenweisen ein interdisziplinäres h° konzept nur h° öh (0.4) nehmen (.) oder entwickeln (.) können (EV\_DE\_095)
- (5) zu verweisen ist in diesem zusammenhang (0.3) auch auf das von gudrun belke (0.3) h° entwickelte konzept zur spracharbeit mit kindern (EV\_DE\_005)
- (6) [...] dass bei der elfkonferenz in wien im im mai h° ähm (.) man sich verabschiedet hat von dem konzept elf als varietät aufzufassen (EV\_DE\_093)
- (7) öhm (.) wie wird ein konzept mehrsprachig gespeichert, das konzept is (1.2) wasser (EV\_DE\_094)

Zu dem Suchwort *Konzept* sind u.a. folgende Verbindungen belegt: Nomen-Verb-Kollokationen: *ein Konzept entwickeln*, das Kompositum *Interkomprehensionskonzept* sowie weitere Verbindungen mit PP bzw. Konjunkionalphrase, u.a. *Konzept zum transferbasierten Tertiärsprachenerwerb*, *Konzept wie Sprachnähe*.

Was die Bedeutungen des italienischen Lexems *concetto* betrifft, zeigt sich, dass dieses wie in (8) einen linguistischen Fachbegriff (hier *topic*) und die dahinterstehende Bedeutung meinen kann, auch in (9) und (10) geht es um Fachbegriffe der Sprachwissenschaft (*errore, variabile sociolinguistica*) und deren Inhaltsseite, in Beispiel (11) meint *concetto* hingegen die semantische Vorstellung, dass der Hund ins Wasser fällt.

- (8) dunque (.) perché noi assumiamo un concetto di topic nel senso di elemento dislocato (EV\_IT\_016)
- (9) in questa citazione [...] che diceva in realtà il concetto di errore presuppone un parlante che parli una lingua non sua (EV\_IT\_006)
- (10) il modello laboviano in altre parole che è fondato sul concetto di variabile sociolinguistica (0.4) si concentra sui fatti di variazione (EV\_IT\_005)
- (11) di nuovo ripete il concetto che h° il cane cade nell'acqua (EV\_IT\_009)

Was die Verwendungsweise des Suchausdrucks *concetto* betrifft, sind im Korpus u.a. folgende Verbindungen belegt: Verb-Nomen-Kollokationen wie *partire da un concetto, illustrare un concetto, assumere un concetto*, Nomen-Nomen-Kollokationen wie *concetto cardine* sowie das Gebrauchsmuster Nomen + PP: *il concetto di familiarità debole alla roberts 2003, il concetto di topic / di errore / di marcatezza*.

Untersucht man die Konkordanzen der englischen Entsprechung *concept*, zeigt sich, dass auch diese im Sinne eines linguistischen Fachbegriffs (hier *corpus*) und der damit verbundenen Bedeutung verwendet werden kann (Bsp. 12) oder als Vorstellung, insbesondere als Selbstbild (Bsp. 13).

- (12) erm (0.5) but they don't seem to understand the underlying concept either i had (0.3) second year students ask me what is corpus they didn't understand the concept (EV\_UK\_012)
- (13) and he says these changes challenge our self concept (0.2) as language teachers (EV\_UK\_012)

<sup>22</sup> Dies wird durch die typischen Verbindungen zu *Konzept* im DWDS bestätigt (<https://www.dwds.de/wb/Konzept>, 01.07.2020).

Zu dem Suchausdruck *concept* liegen u.a. folgende Verbindungen vor: Verb-Nomen-Kollokationen wie *understand a concept, come across a concept, be / become familiar with a concept, develop a concept* sowie das Gebrauchsmuster Nomen + PP: *the whole concept of self directed learning and learner centredness, the concept of portfolios and portfolio assessment, the concept of web quests*.

### 3.1.2 Zusammenfassung

Lässt man die Studierenden im universitären DaF-Studium eine solche korpusbasierte Analyse durchführen, könnte dies dazu beitragen, die metasprachliche Bewusstheit zu fördern. Dabei kann den Studierenden das Konzept der Polysemie näher gebracht werden, indem unterschiedliche Bedeutungen einzelner Lexeme herausgearbeitet werden.

Zusammenfassend treten die Suchausdrücke in den folgenden Bedeutungen auf:

- *Begriff*: Bedeutungsseite eines Fachterminus = Fachbegriff; Bezeichnung (ugsprl.)
- *Konzept*: Theorie, Ansatz; allgem. Bedeutungsseite (Semantik), Vorstellung / Idee
- *concetto*: Bedeutungsseite eines Fachterminus = Fachbegriff; Vorstellung
- *concept*: Bedeutungsseite eines Fachterminus = Fachbegriff; Bild

Im Hinblick auf die Förderung der zwischensprachlichen Bewusstheit kann den Studierenden gezeigt werden, dass keine 1:1 Relationen zwischen den sprachlichen Ausdrücken dt. *Begriff*, it. *concetto*, engl. *concept* vorliegen. Im Sprachvergleich scheint den Lexemen it. *concetto* und engl. *concept* eher das dt. *Begriff* zu entsprechen, mit der Hauptbedeutung ‚inhaltliche Vorstellung / Bedeutungsseite eines Fachausdrucks bzw. eines oder mehrerer alltagssprachlicher Ausdrücke‘; der Ausdruck *Konzept* wird häufig etwas weiter im Sinne eines wissenschaftlichen Ansatzes verwendet.

Außerdem lernen die Studierenden, dass Kollokationen in der alltäglichen Wissenschaftssprache (AWS) in allen drei Sprachen eine wichtige Rolle spielen. Unter Anleitung des Dozenten oder der Dozentin kann herausgearbeitet werden, welche Kollokationen (Adjektiv-Nomen, Nomen-Verb) von besonderer Relevanz sind, z.B. dt. *ein umfassender / zentraler Begriff, einen Begriff verwenden; ein Konzept entwickeln*; it. *illustrare / assumere un concetto, concetto cardine*; engl. *understand a concept, develop a concept, be familiar with a concept*<sup>23</sup>. Neben diesen Kollokationen kann auch auf die Gebrauchsmuster der Suchausdrücke eingegangen werden, z.B. *Begriff (des / der) + Fachterminus*, das (Fachterminus-) *Konzept, concetto di + Fachterminus, concept of + Fachterminus*. Im Deutschen findet sich dabei das Muster Nomen + NP im Genitiv bzw. ein Nominalkompositum, im Italienischen und Englischen wird auf die Struktur Nomen + PP mit *di* bzw. *of* zurückgegriffen. Außerdem können die Studierenden darauf hingewiesen werden, dass in der spontanen mündlichen Interaktion auch bei geübten Muttersprachlern und Muttersprachlerinnen Unsicherheiten bei der Wahl der Verben und Reparaturen bei Kollokationen auftreten können.

### 3.2 Die Kategorie der Diskursmarker am Beispiel des Suchausdrucks dt. *(en)tschuldigung / it. scusate / engl. Sorry*

Für die zweite exemplarische Korpusanalyse wurde eine Kategorie der gesprochenen Sprache ausgewählt, die als Varietät ebenfalls ihre Besonderheiten aufweist. Dabei handelt es sich um die Kategorie

<sup>23</sup> Diese Belege decken sich mit den in Meißner (2020: 74-77, 79) angeführten Kollokatoren der Nomina dt. *Begriff* und engl. *concept*.

der Diskursmarker. Diese werden zusammenfassend umschrieben als syntaktisch nicht oder nur schwach angebundene Einheiten, die häufig prosodisch abgetrennt, morphologisch und lexikalisch heterogen sind, wenig deskriptive Bedeutung besitzen (*semantic bleaching*) und dem Bezugsausdruck vor- oder nachgestellt werden können (Blühdorn / Foolen / Loureda 2017: 17-25). Zwar handelt es sich bei Diskursmarkern um keine grammatisch klar abgrenzbare Kategorie, aber aufgrund der großen Relevanz dieser Ausdrücke in spontanen gesprochensprachlichen Daten – nicht zuletzt aus sprachdidaktischer und sprachvergleichender Perspektive – erscheint es durchwegs sinnvoll, Diskursmarker in Korpora der gesprochenen Sprache zu annotieren und zu untersuchen (Fandrych 2017: 21). Als funktionale Kategorie sind Diskursmarker gesprächsorganisierende Einheiten, die der Initiierung von Redebeiträgen, zur Einleitung von Reparatursequenzen, als Gliederungssignale usw. dienen. In den anderen beiden Sprachen liegen mit it. *segnale discorsivo*<sup>24</sup> (Bazzanella 2011) und engl. *discourse marker*<sup>25</sup> vergleichbare Kategorien vor, die anhand ähnlicher Kriterien definiert werden.

Für die exemplarische Korpusanalyse wurde das deutsche Lexem (*en*)*entschuldigung*<sup>26</sup> und seine Entsprechungen it. *scusate* und engl. *sorry* gewählt. Im Sprachvergleich zeigt sich, dass die ausgewählten Lexeme unterschiedlichen Wortarten angehören – ein für Diskursmarker typischer Umstand: *Entschuldigung* ist ein Nomen, *scusate* ein Verb im Imperativ der 2. Person Plural und *sorry* ein Adjektiv.

Die Bedeutung von *entschuldigung* wird im DWDS<sup>27</sup> folgendermaßen umschrieben: ‚um Nachsicht, Verständnis bitten‘ bzw. ‚für etwas Nachsicht und Verständnis zeigen, etwas verzeihen‘; zusätzlich wird darauf hingewiesen, dass die Ausdrücke *entschuldigen Sie* und *entschuldigung* als Höflichkeitsformel dienen.

### 3.2.1 Datenauswertung

Das Lexem *entschuldigung* wurde in den Lemmalisten von GeWiss als Diskursmarker kategorisiert<sup>28</sup> (zur Korpusannotation s. Kapitel 2) und befindet sich an neunter Stelle nach *also*<sup>29</sup> (über 900 Treffer, Häufigkeitsklasse 3), *ja* (Klasse 5), *so* (Klasse 7), *gut* (Klasse 8), *na*, *okay*, *weil* und *genau*, und gehört gemeinsam mit letzteren der Klasse 9 an. Die Frequenzanalyse im Teilkorpus der Experten- und Expertinnenvorträge hat ergeben, dass für *entschuldigung* 9 Treffer vorliegen, für *tschuldigung* 2 Treffer (d.h. insgesamt 11), für das italienische *scusate* finden sich 19 und für das englische *sorry* 8 Belege<sup>30</sup>. Dabei kann der Dozent oder die Dozentin die Studierenden darauf aufmerksam machen, dass das

<sup>24</sup> “I segnali discorsivi (detti anche *marcatori di discorso*) sono elementi linguistici (parole, espressioni, frasi), di natura tipicamente pragmatica, diffusi in specie nella lingua parlata, che, a partire dal significato originario, assumono ulteriori funzioni nel discorso a seconda del contesto: sottolineano la strutturazione del testo, connettono elementi nella frase e tra le frasi, esplicitano la posizione dell’enunciato nella dimensione interpersonale, evidenziano processi cognitivi in atto (pragmatica).” (Bazzanella 2011, verfügbar unter: [http://www.treccani.it/enciclopedia/segnali-discorsivi\\_\(Enciclopedia-dell%27Italiano\)/](http://www.treccani.it/enciclopedia/segnali-discorsivi_(Enciclopedia-dell%27Italiano)/), 01.07.2020).

<sup>25</sup> “Discourse markers are words or phrases like *anyway*, *right*, *okay*, *as I say*, *to begin with*. We use them to connect, organise and manage what we say or write or to express attitude. [...] Discourse markers do not always have meanings that you will find in your dictionary. However, they do have certain functions, and some discourse markers, such as *well*, can have a number of functions.” Aus: Cambridge Dictionary. Der Ausdruck (*I'm*) *sorry* wird hier ebenfalls aufgelistet, und zwar mit der Funktion, eine Einstellung oder einen Standpunkt darzulegen.

<sup>26</sup> Bei *tschuldigung* handelt es sich um eine für die gesprochene Sprache typische phonetische Reduktion.

<sup>27</sup> <https://www.dwds.de/wb/Entschuldigung> (01.07.2020).

<sup>28</sup> Zur Annotation von Diskursmarkern in FOLK s. Westpfahl (2017).

<sup>29</sup> Zu den Diskursmarkern *also* und *so* in wissenschaftlichen Vorträgen s. die Studie von Wallner (2017).

<sup>30</sup> Außerdem liegen im deutschen Teilkorpus jeweils ein Treffer für *sorry* und *pardon* vor, die jedoch hier aus Gründen der Übersichtlichkeit nicht berücksichtigt werden. Mögliche alternative Formen wie dt. *entschuldigen Sie*, it. *scusa / scusi* und engl. *excuse me*, *pardon* sind nicht belegt.

ausgewählte Lexem in allen drei Teilkorpora auch in syntaktisch integrierten Formen auftritt, die jedoch nicht als Diskursmarker betrachtet werden sollten: dt. *dafür bitte ich um Entschuldigung*, it. *scusate la cacofonia verbale* und engl. *i'm really sorry and i'd really like to stay*.

Für quantitativ zuverlässige Aussagen eignen sich die Lemmata aufgrund der relativ niedrigen Häufigkeit nicht, sie können jedoch für qualitative Analysen herangezogen werden. Mithilfe der Konkordanzsuche könnten die Studierenden der Frage nachgehen, in welchen pragmatischen Funktionen diese Diskursmarker in wissenschaftlichen Vorträgen auftreten und wie man ihre Bedeutung paraphrasieren kann<sup>31</sup>.

Für (*en*)*tschuldigung* können folgende Funktionen unterschieden werden: Erstens erfolgt damit ein Verweis auf die Gesprächssituation, insbesondere auf Probleme mit der Technik, der Power-Point-Präsentation (Bilder, Audio) und Ähnliches. Im folgenden Beispiel zeigt der Vortragende Bilder von Schweizer Abstimmungsplakaten. Aus dem Transkript ist herauszulesen, dass zunächst vermutlich nicht die richtige Folie zu sehen ist, weshalb er mit der Interjektion *hups* und dem Diskursmarker *tschuldigung* eine Reparatur einleitet.

- (14) und zwar (.) spiele mit dem schweizer kreuz h° das erste beispiel hier (0.3) hups (0.8) tschuldigung das äh (.) ist aus den fünfziger jahren (EV\_DE\_106)

Zweitens kann ein Sprecher mit dem Diskursmarker (*en*)*tschuldigung* das Rederecht beanspruchen. In (15) möchte die Vortragende im Rahmen der Diskussion eine Rückfrage an den Diskutanten stellen und leitet diese mit *tschuldigung* ein:

- (15) tschuldigung auf welchen typ beziehen sie sich jetzt grade okay (EV\_DE\_097)

Als dritte Funktion kann die Reparatur von Versprechern beschrieben werden. Allgemein unterscheidet man bei Reparaturen zwischen dem Reparandum (Problemquelle), der Reparatur-Initiierung und der Reparatur-Operation. Reparaturen können von demselben Sprecher oder von einer anderen Sprecherin initiiert bzw. durchgeführt werden (Selbst- bzw. Fremdreparatur)<sup>32</sup>. In Beispiel (16) äußert die Sprecherin zunächst das Wort *Dimensionsbeschreibungen*, verbessert dieses anschließend in *Beschreibungsdimensionen* und fügt den Diskursmarker *entschuldigung* hinzu (Selbstreparatur). In (17) möchte die Moderatorin die nächste Vortragende vorstellen und irrt sich dabei im Nachnamen. Die Vortragende korrigiert sie, woraufhin die Moderatorin den Nachnamen korrekt wiederholt und anschließend zwei Mal *entschuldigung* hinzufügt (Fremdreparatur).

- (16) die die erkenntnisse h° anderer (.) dimensions(.)beschreibungen oder beschreibungsdimensionen entschuldigung h° wie der lexik (.) wie der syntax (.) und d der stilistik (EV\_DE\_095)

- (17) MOD: und nach dem motto ladies first sabrina pickert  
nn: ((lacht))  
SP: peckert  
MOD: peckert entschuldigung entschuldigung (da) so kommt es wenn (ich) äh wenn man ganz kurzfristig nur noch äh h° surft im netz  
SP: kein Problem ((lacht))

(EV\_DE\_005)

Auch das italienische *scusate* wird häufig zur Reparatur von Versprechern eingesetzt, wie in den folgenden Beispielen, wo jeweils ein Reparandum (*della quale, undici, funzione grammaticale*) und eine

<sup>31</sup> Für die Analyse der pragmatischen Funktionen reicht der unmittelbare Kontext der Konkordanzen oft nicht aus. Die Suchmaske von GeWiss ermöglicht es aber, durch einen Klick von den Konkordanzen direkt zu dem Volltext zu gelangen.

<sup>32</sup> Zu Reparaturen in wissenschaftlichen Vorträgen s. ausführlich Salzmann (2017: 247-260).

Reparatur-Operation (*del quale, dieci, funzione argomentale*) vorliegen, in deren Mitte (bzw. im Anschluss danach) sich der Diskursmarker *scusate* befindet, der jeweils als Reparatur-Initiierung fungiert (Bsp. 18–20).

- (18) cioè mi sfugge il criterio fonetico (.) o fonologico sulla base della quale h° del quale *scusate* (.) ehm (0.4) affermiamo questo (EV\_IT\_017)  
 (19) eh in per esempio eh (0.6) undici eh *scusate* in dieci (0.5) siamo ancora in dieci in dieci bi: (EV\_IT\_017)  
 (20) attraverso eh l'uso del clitico eh ah viene abbinata una eh funzione grammaticale *scusate* una funzione eh argomentale (.) alla funzione eh discorsiva (EV\_IT\_020)

Auch mit *scusate* kann gelegentlich ein Verweis auf die Gesprächssituation, insbesondere auf eine bestimmte Stelle in der Power-Point-Präsentation und Ähnliches erfolgen. Beispiel (21) stammt aus einer Diskussion, bei der die Vortragende zu den zuvor gezeigten Beispielen zurückkehren möchte, um ihre Aussagen zu belegen:

- (21) e infatti negli esempi adesso che eh *scusate* ritorno indietro e lo ritrovo h° (0.7) hai sia ha macchiato la camicia ha lavorato è venuto (EV\_IT\_015)

Für das englische *sorry* ist die Funktion der Beanspruchung des Rederechts in der Diskussion belegt. In dem folgenden Ausschnitt erfolgt eine Überschneidung der (schwer verständlichen) Äußerungen der Sprecher EM und DIS10, danach ergreift die DiskutantIn mithilfe des Diskursmarkers *sorry* das alleinige Rederecht.

- (22) EM: [whereas you know whereas with]  
 DIS10: [(i just think you know in terms of)]  
 DIS10: (0.2) *sorry* in [term]s of staffing (0.3) i mean how do you pay teachers  
 EM: [yeah] (EV\_UK\_012)

Auch in (23) signalisiert *sorry* die Übernahme des Rederechts und somit die Beendigung der Diskussion durch den Moderator.

- (23) MOD: okay i think (0.1) i'm afraid we're going to have to stop there (0.1) erm (0.6) thank you very much indeed for a (0.1) very stimulating thought provoking talk  
 nn: ((xxx))  
 MOD: [before <<laughing>] yes yes *sorry* but i mustn't forget h° the next talk here  
 nn: [((laughter))] (EV\_UK\_012)

Außerdem kann *sorry* auch eine Nachfrage bei Nicht-Verstehen darstellen, wie in (24). Durch *sorry* erfolgt hier eine Fremd-Initiierung der Reparatur, die Reparatur-Operation wird im Anschluss von dem Diskutanten durchgeführt.

- (24) DIS1: so is personality not something that varies then (0.3)  
 MB: *sorry*  
 DIS1: (0.2) do we not have multiple personalities tha[t (.) one might be (...)]  
 MB: [h° we do it's just (...)] (EV\_UK\_016)

### 3.2.2 Zusammenfassung

In Bezug auf die metasprachliche Bewusstheit könnten die DaF-Studierenden durch die vorliegende Korpusanalyse die Kategorie der Diskursmarker kennenlernen und erfahren, dass diese bestimmte gesprächsorganisierende Funktionen erfüllt, durch Multifunktionalität charakterisiert ist und als sprachenübergreifendes Phänomen betrachtet werden kann.

Im Hinblick auf den Ausbau der zwischensprachlichen Bewusstheit sollen die Studierenden lernen, Bezüge zwischen den Diskursmarkern der drei Sprachen herzustellen, indem sie erkennen, dass diese in ähnlichen, wenn auch nicht identischen Funktionen belegt sind. Zusammenfassend geht aus der Analyse hervor, dass *(en)tschuldigung* als Verweis auf die Gesprächssituation (Probleme mit Audio, Power-Point) in der paraphrasierten Bedeutung ‚tut mir leid‘, als Reparatur von Versprechern in der Bedeutung ‚ich korrigiere mich‘ sowie als Beanspruchung des Rederechts in der Bedeutung ‚darf ich etwas fragen / sagen‘ dienen kann. Zudem zeigt sich, dass der Diskursmarker *entschuldigung* sowohl vorgestellt als auch nachgestellt (nach der Reparatur) auftreten kann; er ist in monologischen, aber auch in dialogischen Phasen belegt. Die italienische Entsprechung *scusate* dient vorwiegend der Reparatur von Versprechern in der Bedeutung ‚ich korrigiere mich‘, marginal handelt es sich auch um einen Verweis auf die Gesprächssituation in der Bedeutung ‚tut mir leid‘. In Bezug auf die Stellung zeigt sich, dass *scusate* meist vorgestellt wird zur Einleitung der Reparatur und primär in monologischen Phasen belegt ist. Die englische Entsprechung *sorry* dient der Beanspruchung bzw. Übernahme des Rederechts und kann paraphrasiert werden als ‚darf ich sprechen‘; außerdem fungiert *sorry* als Nachfrage bzw. Initiierung einer Reparatur in der Bedeutung ‚wie bitte‘. *Sorry* tritt vorgestellt auf und steht in engem Zusammenhang mit dem Sprecherwechsel in dialogischen Phasen.

## 4. Schlussfolgerungen und Ausblick

In diesem Beitrag wurde veranschaulicht, wie das GeWiss-Korpus zum Zweck der Vermittlung der deutschen Wissenschaftssprache in Bezug zur italienischen und englischen Wissenschaftssprache eingesetzt werden kann. Dabei wurde gezeigt, auf welche Weise die Arbeit mit dem Korpus im universitären DaF-Unterricht die meta- und zwischensprachliche Bewusstheit der Studierenden fördern, den interlingualen Transfer sprachenübergreifender Aspekte anregen und zum Ausbau von Fähigkeiten des kompetenten Sprachvergleichs beitragen kann. Lässt man die Studierenden im Unterricht eine solche Korpusrecherche durchführen, kann eine Bewusstmachung grundlegender Eigenschaften der gesprochenen Wissenschaftssprache – in diesem Fall das Auftreten eines ausgewählten Nomens der AWS in typischen Kollokationen und dessen Polysemie sowie die funktionale Kategorie des Diskursmarkers mit spezifischen gesprächsorganisatorischen Funktionen – erfolgen. Außerdem kann der Sprachvergleich auf effiziente Weise trainiert werden und die Lernenden der fremden deutschen Wissenschaftssprache können für vorhandene Gemeinsamkeiten und Unterschiede sensibilisiert werden, die sie dann beim weiteren Ausbau der Wissenschaftssprachkompetenz gezielt berücksichtigen können. Dies stellt eines der Hauptziele des mehrsprachig orientierten DaF-Unterrichts dar und erfüllt vor allem im Studium der Fremdsprachenphilologien eine zentrale Koordinierungsfunktion.

Mögliche zukünftige Anwendungsgebiete für eine korpusbasierte, vergleichende Vermittlung der gesprochenen Wissenschaftssprache sind dabei der universitäre DaF-Unterricht in den Fremdsprachenphilologien der nicht-deutschsprachigen Länder sowie eher sprachpraktisch orientierte Kurse für Studierende oder Wissenschaftler unterschiedlicher Disziplinen, die sich (auch) in den deutschsprachigen Ländern ab dem Sprachniveau C1 mit der deutschen Wissenschaftssprache beschäftigen. Im Rahmen dieser Lehrveranstaltungen könnte anhand empirischer Untersuchungen der Frage nachgegangen werden, welche Herausforderungen der Einsatz des GeWiss-Korpus und die konkrete

Datenanalyse für universitäre DaF-Studierende mit sich bringen, auf welche Aspekte beim Ausbau der *corpus literacy* besonders geachtet werden muss und wie Studierende das Korpus im Sinne eines *data-driven learning* direkt für die Produktion und Rezeption eigener Texte, wie z.B. wissenschaftliche Präsentationen und Seminare, verwenden können.

## Literatur und Ressourcen

Auer, Peter / Günthner, Susanne (2005): Die Entstehung von Diskursmarkern im Deutschen – ein Fall von Grammatikalisierung? In: Leuschner, Torsten / Mortelmans, Tanja / De Groot, Sarah (Hrsg.): *Grammatikalisierung im Deutschen*. Berlin: De Gruyter, 335-362.

Bachmann-Stein, Andrea (2013): Authentische gesprochene Sprache im DaF-Unterricht: Pro und Contra. In: Moraldo, Sandro / Missaglia, Federica (Hrsg.): *Gesprochene Sprache im DaF-Unterricht. Grundlagen – Ansätze – Praxis*. Heidelberg: Winter, 39-58.

Bazzanella, Carla (2011): Segnali discorsivi. In: *Enciclopedia dell'italiano*. [http://www.treccani.it/enciclopedia/segnali-discorsivi\\_\(Enciclopedia-dell%27Italiano\)/](http://www.treccani.it/enciclopedia/segnali-discorsivi_(Enciclopedia-dell%27Italiano)/) (01.07.2020).

Belica, Cyril / Perkuhn, Rainer (2015): Feste Wortgruppen/Phraseologie I: Kollokationen und syntagmatische Muster. In: Haß, Ulrike / Storjohann, Petra (Hrsg.): *Handbuch Wort und Wortschatz*. Berlin / Boston: De Gruyter, 201-225.

Blühdorn, Hardarik / Foolen, Ad / Loureda, Óscar (2017): Diskursmarker: Begriffsgeschichte – Theorie – Beschreibung. Ein bibliographischer Überblick. In: Blühdorn, Hardarik et al. (Hrsg.): *Diskursmarker im Deutschen. Reflexionen und Analysen*. Göttingen: Verlag für Gesprächsforschung, 7-47.

Cambridge Dictionary: <https://dictionary.cambridge.org/it/> (14.06.2021).

Datenbank für gesprochenes Deutsch (DGD): [https://dgd.ids-mannheim.de/dgd/pragdb.dgd\\_extern.welcome](https://dgd.ids-mannheim.de/dgd/pragdb.dgd_extern.welcome) (14.06.2021).

Digitales Informationssystem der deutschen Sprache (DWDS): <https://www.dwds.de/> (14.06.2021).

Ehlich, Konrad (1993): Deutsch als fremde Wissenschaftssprache. In: *Jahrbuch Deutsch als Fremdsprache* 19, 13-41.

Fandrych, Christian (2017): Gesprochene-Sprache-Forschung und Korpuserschließung am Beispiel von *GeWiss digital*. In: Fandrych, Christian / Meißner, Cordula / Wallner, Franziska (Hrsg.): *Gesprochene Wissenschaftssprache – digital. Verfahren zur Annotation und Analyse mündlicher Korpora*. Tübingen: Stauffenburg, 13-30.

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

Flinz, Carolina (2020): Vergleichbare Spezialkorpora für den Tourismus: eine Chance für den Fachsprachenunterricht. In: Hepp, Marianne / Salzmann, Katharina (Hrsg.): *Sprachvergleich in der mehrsprachig orientierten DaF-Didaktik. Theorie und Praxis*. Roma: Studi Germanici, 133-151.

- Flinz, Carolina / Hufeisen, Britta (2021): *Korpora in DaF und DaZ: Theorie und Praxis*. (Themenheft). Zeitschrift für Interkulturellen Fremdsprachenunterricht. Didaktik und Methodik im Bereich Deutsch als Fremdsprache. <https://tujournals.ulb.tu-darmstadt.de/index.php/zif/> (14.06.2021).
- Forschungs- und Lehrkorpus Gesprochenes Deutsch (FOLK): <http://agd.ids-mannheim.de/folk.shtml> (14.06.2021).
- Foschi Albert, Marina (2017): Grammatica del tedesco scritto e parlato. In: Costa, Marcella / Foschi Albert, Marina (Hrsg.): *Grammatica del tedesco parlato. Con un saggio introduttivo di Reinhard Fiehler*. Pisa: Pisa University Press, 51-67.
- Gesprochenes Deutsch – authentische Alltagsinteraktionen für die Forschung und Praxis im Bereich DaF und DaZ: <https://dafdaz.sprache-interaktion.de/> (14.06.2021).
- Gesprochenes Deutsch für die Auslandsgermanistik: <https://daad-gda.sprache-interaktion.de/> (14.06.2021).
- Gesprochene Wissenschaftssprache kontrastiv (GeWiss): <https://gewiss.uni-leipzig.de/index.php?id=home> (14.06.2021).
- Graefen, Gabriele / Moll, Melanie (2011): *Wissenschaftssprache Deutsch: lesen – verstehen – schreiben. Ein Lehr- und Arbeitsbuch*. Frankfurt am Main: Lang.
- Grammis Grammatisches Informationssystem: <https://grammis.ids-mannheim.de/> (14.06.2021).
- Imo, Wolfgang (2012): Wortart Diskursmarker? Nicht-flektierende Wortarten. In: Rothstein, Björn (Hrsg.): *Nicht-flektierende Wortarten*. Berlin: De Gruyter, 48-88.
- Imo, Wolfgang / Weidner, Beate (2018): Mündliche Korpora im DaF- und DaZ-Unterricht. In: Kupietz, Marc / Schmidt, Thomas (Hrsg.): *Korpuslinguistik*. Berlin / Boston: de Gruyter, 231-252.
- Jessner, Ulrike (2008): A DST Model of Multilingualism and the Role of Metalinguistic Awareness. In: *The Modern Language Journal* 92: 2, 270–283.
- Jessner, Ulrike / Hufeisen, Britta (2019): The Psycholinguistics of Multiple Language Learning and Teaching. In: Singleton, David / Aronin, Larissa (Hrsg.): *Twelve Lectures on Multilingualism*. Bristol: Multilingual Matters, 65-100.
- Koch, Peter / Oesterreicher, Wulf (2007): Schriftlichkeit und kommunikative Distanz. In: *Zeitschrift für germanistische Linguistik* 35, 346-375.
- Lemnitzer, Lothar / Zinsmeister, Heike (2015<sup>3</sup>): *Korpuslinguistik. Eine Einführung*. Tübingen: Narr Francke Attempto.
- Lüdeling, Anke / Walter, Maik (2009): *Korpuslinguistik für Deutsch als Fremdsprache: Sprachvermittlung und Spracherwerbsforschung*, 1-37. <https://www.linguistik.hu-berlin.de/de/institut/professuren/korpuslinguistik/mitarbeiter-innen/anke/pdf/LuedelingWalterDaF.pdf> (01.07.2020). Stark erweiterte Fassung von Lüdeling, Anke / Walter, Maik: Korpuslinguistik. In: Fandrych, Christian et al. (Hrsg.) (2010): *Deutsch als Fremd- und Zweitsprache. Ein internationales Handbuch*, Berlin: Mouton de Gruyter, 315–322.
- Meißner, Cordula (2020): Der Verbwortschatz der deutschen Bildungs- und Wissenschaftssprache als Lerngegenstand: Möglichkeiten einer mehrsprachigkeitsdidaktischen Vermittlung. In: Hepp, Marianne / Salzmann,

- Katharina (Hrsg.): *Sprachvergleich in der mehrsprachig orientierten DaF-Didaktik. Theorie und Praxis*. Roma: Studi Germanici, 59-83.
- Mukherjee, Joybrato (2002): *Korpuslinguistik und Englischunterricht. Eine Einführung*. Berlin: Peter Lang.
- Salzmann, Katharina (2017): *Expansionen in der deutschen und italienischen Wissenschaftssprache. Kontrastive Korpusanalyse und sprachdidaktische Überlegungen*. Berlin: Erich Schmidt Verlag.
- Salzmann, Katharina (2019): Die Wissenschaftssprachkomparatistik als Ansatz zur Förderung der akademischen Mehrsprachigkeit: Ein Unterrichtsmodul aus der italienischen DaF-Didaktik. In: *Bulletin suisse de linguistique appliquée* 109, Themenheft *Wissenschaftssprache(n) kontrastiv*, hrsg. v. Fandrych, Christian / Schmidlin, Regula, 189-206.
- Salzmann, Katharina (2020): Phraseodidaktik mehrsprachig. Ein Unterrichtsmodul zur metaphorischen Verwendung der Farbadjektive im Sprachvergleich. In: Hepp, Marianne / Salzmann, Katharina (Hrsg.): *Sprachvergleich in der mehrsprachig orientierten DaF-Didaktik. Theorie und Praxis*. Roma: Studi Germanici, 111-131.
- Schmidt, Thomas (2018): Gesprächskorpora. Aktuelle Herausforderungen für einen speziellen Korpusstyp. In: Kupietz, Marc / Schmidt, Thomas (Hrsg.): *Korpuslinguistik*. Berlin / Boston: de Gruyter, 209-230.
- Selting, Margret et al. (2009): Gesprächsanalytisches Transkriptionssystem 2 (GAT 2). *Gesprächsforschung – Online-Zeitschrift zur verbalen Interaktion* 10, 353-402. <http://www.gespraechsforschung-ozs.de/heft2009/px-gat2.pdf> (01.07.2020).
- Wallner, Franziska (2014): *Kollokationen in Wissenschaftssprachen: Zur lernerlexikographischen Relevanz ihrer wissenschaftssprachlichen Gebrauchsspezifika*. Tübingen: Stauffenburg.
- Wallner, Franziska (2017): Diskursmarker funktional: Eine quantitativ-qualitative Beschreibung annotierter Diskursmarker im GeWiss-Korpus. In: Fandrych, Christian / Meißner, Cordula / Wallner, Franziska (Hrsg.): *Gesprochene Wissenschaftssprache – digital. Verfahren zur Annotation und Analyse mündlicher Korpora*. Tübingen: Stauffenburg, 107-122.
- Westpfahl, Swantje (2017): Diskursmarker aus korpuslinguistischer Sicht – POS-Annotation von Diskursmarkern in FOLK. In: Blühdorn, Hardarik et al. (Hrsg.): *Diskursmarker im Deutschen. Reflexionen und Analysen*. Göttingen: Verlag für Gesprächsforschung, 285-310.

## Anhang

Transkriptionskonventionen GAT 2 Minimaltranskript (gekürzt nach Selting et al. 2009)

[ ]	Überlappungen
h°	Ein- bzw. Ausatmen
(.)	Mikropause
(1.0)	gemessene Pause von ca. 1.0 Sek.
(solche)	vermuteter Wortlaut
((xxx xxx))	unverständlich
<<lachend>	> Lachpartikeln in der Rede
äh, ähm	Verzögerungssignale

**Kurzbio: Katharina Salzmann** ist Dozentin für Deutsche Sprachwissenschaft an der Freien Universität Bozen, Italien. 2015 Abschluss des Doktorats an der Universität Pisa, 2016-2018 Forschungsstipendiatin am Istituto Italiano di Studi Germanici (Rom) im Rahmen eines Projekts zu Mehrsprachigkeit und Tertiärsprachendidaktik. Ihre Forschungsinteressen liegen in den Bereichen mehrsprachig orientierte DaF-Didaktik, kindlicher Mehrsprachenerwerb, Wissenschaftssprache und gesprochene Sprache.

**Anschrift:**

Katharina Salzmann  
Freie Universität Bozen  
Fakultät für Bildungswissenschaften  
Regensburger Allee 16  
39042 Brixen-Bressanone  
Italien  
[katharina.salzmann@unibz.it](mailto:katharina.salzmann@unibz.it)



Lizenz: CC BY 4.0 International - Creative Commons, Namensnennung.