

MASARYKOVA UNIVERZITA

Filozofická fakulta

Ústav germanistiky, nordistiky a nederlandistiky

MAGISTERSKÁ DIPLOMOVÁ PRÁCE

2020

Bc. Tomáš Mach

MASARYKOVA UNIVERZITA

Filozofická fakulta

Ústav germanistiky, nordistiky a nederlandistiky

Učitelství německého jazyka a literatury pro střední školy

Bc. Tomáš Mach

**Analyse von Kollokationen ausgewählter
Abstrakta in der studentischen
Wissenschaftssprache**

Magisterská diplomová práce

Vedoucí práce: Mgr. Vlastimil Brom, Ph.D.

2020

Hiermit erkläre ich, dass ich die vorliegende Diplomarbeit selbstständig ausgearbeitet und nur die angeführte Literatur verwendet habe.

Brno, 13. April 2020

.....

Bc. Tomáš Mach

Danksagung

Besonders möchte ich mich an dieser Stelle bei meinem Betreuer, Mgr. Vlastimil Brom, Ph.D., bedanken, dass er mir in der Themenwahl besondere Freiheit gelassen hat, zugleich aber immer bereit war, mich kompetent und freundlich zu beraten. Mit seinen ausführlichen Kommentaren hat er maßgeblich zu der endgültigen Fassung dieser Arbeit beigetragen.

Besonderer Dank gebührt auch meinen Eltern, die mich während des ganzen Studiums unterstützt haben und immer für mich da waren.

Danke auch an meine Freunde, Eva und Peter, für die Ermunterung und die zahlreichen langen Nachtgespräche, nicht nur über die Sprachwissenschaft. Beide haben die letzten sechs Jahre in Brünn für mich unvergesslich gemacht.

Eine besondere Erwähnung verdient Herr James Thomas, der sich glücklicherweise auch andere Fragen als die auf der Seite 180 seines Buches (2017) gestellt hatte, von denen ich mich inspirieren lassen konnte.

Inhaltsverzeichnis

Abbildungsverzeichnis	5
Tabellenverzeichnis	7
Abkürzungsverzeichnis.....	8
1. Einleitung	9
2. Forschungsstand.....	11
2.1 Wissenschaftssprache.....	11
2.2 Kollokationen	12
2.2.1 Interpretationen des Kollokationsbegriffs	13
2.2.2 Kollokationen lexikographisch-didaktisch perspektiviert.....	14
2.3 Satzbezogene nominale Verweise (SNVs).....	20
2.3.1 Allgemeine Begriffsbestimmung	20
2.3.2 Probleme mit der Operationalisierung	22
3. Korpus	25
3.1 Korpusdesign.....	25
3.2 DESTWIKO (v0.2)	26
3.2.1 Vorüberlegungen zum Design.....	26
3.2.2 Erhebungsprozess und Technikalien	28
4. Methodische Vorgehensweise.....	33
4.1 Prozess.....	33
4.2 Exkurs: Dispersion	34
4.3 Instrumente.....	38
4.3.1 Word-Sketches	38
4.3.2 Visualisierungen.....	41
4.4 Weitere Aspekte des Profilaufbaus	47

5. Wortprofile	50
5.1 <i>Fall</i>	50
5.2 <i>Frage</i>	54
5.3 <i>Grund</i>	58
5.4 <i>Unterschied</i>	62
5.5 <i>Möglichkeit</i>	67
5.6 <i>Problem</i>	70
5.7 <i>Ergebnis</i>	74
5.8 <i>Ziel</i>	80
5.9 <i>Meinung</i>	85
5.10 <i>Aussage</i>	88
5.11 <i>Tatsache</i>	91
5.12 <i>Vorstellung</i>	96
5.13 <i>Aufgabe</i>	99
5.14 <i>Tendenz</i>	102
5.15 <i>Versuch</i>	106
6. Didaktische Implikationen	110
6.1 <i>Produkt</i>	110
6.2 <i>Prozess</i>	111
6.2.1 <i>Exkurs: Korpora im Fremdsprachenunterricht</i>	111
6.2.2 <i>Praktische Umsetzung</i>	114
6.2.3 <i>Fragen zur Korpusarbeit</i>	116
7. Zusammenfassung und Ausblick	127
Literaturverzeichnis	129

Abbildungsverzeichnis

Abbildung 1: Die Unterscheidung – Kollokationen, Kombinationen und Phrasemen (nach Hollós, 2010)	17
Abbildung 2: Wörterbucheintrag von Abend nach Quasthoff (2011).....	18
Abbildung 3: Wörterbucheintrag von Abend aus Buhofer et al. (2014).....	18
Abbildung 4: Wörterbucheintrag von Abend aus Hollós (2014a)	19
Abbildung 5: Histogramm zur Verteilung der Arbeiten über die Abschlussjahre	28
Abbildung 6: Box-Plot-Repräsentation der Länge von L1- und L2-Arbeiten	32
Abbildung 7: Darstellung der <i>DP</i> -Werte	37
Abbildung 8: Beispiel einer Gramrel aus Sketch Engine.....	39
Abbildung 9: Beispiel einer Sketch-diff-Gramrel zum Vergleich von Lexemen	42
Abbildung 10: Beispiel einer Sketch-diff-Gramrel zum Vergleich von Subkorpora.....	42
Abbildung 11: Neue, skalaartige Visualisierung einer Sketch-diff-Gramrel	43
Abbildung 12: Kreisdiagramm mit mehreren Gramrels.....	43
Abbildung 13: Mögliche Einstellungen der Kreisdiagramme.....	44
Abbildung 14: Kreisdiagramm mit Wortkreisen und ohne	44
Abbildung 15: Kreisdiagramm mit „score circles“	45
Abbildung 16: Kreisdiagramm mit variabler Textgröße	46
Abbildung 17: Kreisdiagramm mit einer ausgeglichenen Winkelgröße	46
Abbildung 18: <i>Fall L1</i>	50
Abbildung 19: <i>Fall L2</i>	50
Abbildung 20: <i>Frage L1</i>	54
Abbildung 21: <i>Frage L2</i>	54
Abbildung 22: <i>Grund L1</i>	58
Abbildung 23: <i>Grund L2</i>	58
Abbildung 24: <i>Unterschied L1</i>	62
Abbildung 25: <i>Grund L2</i>	62
Abbildung 26: <i>Möglichkeit L1</i>	67
Abbildung 27: <i>Möglichkeit L2</i>	67
Abbildung 28: <i>Problem L1</i>	70
Abbildung 29: <i>Problem L2</i>	70
Abbildung 30: <i>Ergebnis L2</i>	74

Abbildung 31: <i>Ergebnis</i> L1	74
Abbildung 32: <i>Ziel</i> L1	80
Abbildung 33: <i>Ziel</i> L2	80
Abbildung 34: <i>Meinung</i> L2	85
Abbildung 35: <i>Meinung</i> L1	85
Abbildung 36: <i>Aussage</i> L1	88
Abbildung 37: <i>Aussage</i> L2	88
Abbildung 38: <i>Tatsache</i> L1	91
Abbildung 39: <i>Tatsache</i> L2	91
Abbildung 40: <i>Tatsache</i> (diplomarbeiten24.de).....	92
Abbildung 41: <i>Tatsache</i> (German Web 2013)	94
Abbildung 42: <i>Vorstellung</i> L1	96
Abbildung 43: <i>Vorstellung</i> L2	96
Abbildung 44: <i>Aufgabe</i> L1	99
Abbildung 45: <i>Aufgabe</i> L2	99
Abbildung 46: <i>Tendenz</i> L2	102
Abbildung 47: <i>Tendenz</i> L1	102
Abbildung 49: <i>Versuch</i> L1	106
Abbildung 48: <i>Versuch</i> L2	106

Tabellenverzeichnis

Tabelle 1: Kookkurrenzen von SNVs in DESTWIKO (v0.1).....	21
Tabelle 2:Übersicht der Metadaten in DESTWIKO (v0.2).....	30
Tabelle 3: Angaben über die Verteilung der Subkorpora.....	31
Tabelle 4: Verteilung der 15 Abstrakta im L1- und L2-Subkorpus	36
Tabelle 5: <i>Fall</i> – statistische Übersicht	51
Tabelle 6: <i>Frage</i> – statistische Übersicht	55
Tabelle 7: <i>Grund</i> – statistische Übersicht	59
Tabelle 8: Vergleich von <i>aufgrund</i> und auf <i>Grund</i>	59
Tabelle 9: <i>Unterschied</i> – statistische Übersicht	63
Tabelle 10: <i>Möglichkeit</i> – statistische Übersicht.....	68
Tabelle 11: <i>Problem</i> – statistische Übersicht	71
Tabelle 12: Verteilung von <i>Problematik</i> im Korpus	71
Tabelle 13: <i>Ergebnis</i> – statistische Übersicht	75
Tabelle 14: <i>Anzahl</i> von Arbeiten mit Ergebnis	75
Tabelle 15: <i>Ziel</i> – statistische Übersicht	81
Tabelle 16: <i>Meinung</i> – statistische Übersicht	86
Tabelle 17: <i>Aussage</i> – statistische Übersicht	89
Tabelle 18: <i>Tatsache</i> – statistische Übersicht	92
Tabelle 19: Vergleich zwischen Gramrels in DESTWIKO (v0.2) [L2] und German Web 2013 [diplomarbeiten.de]	93
Tabelle 20: <i>Vorstellung</i> – statistische Übersicht	97
Tabelle 21: <i>Aufgabe</i> – statistische Übersicht	100
Tabelle 22: <i>Tendenz</i> – statistische Übersicht	103
Tabelle 23: <i>Versuch</i> – statistische Übersicht.....	107

Abkürzungsverzeichnis

AWS	Alltägliche Wissenschaftssprache (nach Ehlich)
CI	Konfidenzintervall
DDL	data-driven learning
<i>DP</i>	Deviation of Proportions (nach Gries)
GER	Gemeinsamer europäischer Referenzrahmen für Sprachen
L1	Muttersprache
L2	Fremdsprache
M	Mittelwert (arithmetisch)
<i>p</i>	Signifikanzwert
pM	per Million
SD	Standardabweichung
UWV	usuelle Wortverbindungen

1. Einleitung

Some famous and influential linguists have simply denied the relevance of corpus evidence to linguistic theory. Others have in recent years treated corpora as ‘fish ponds’ in which to angle for fish that will fit independently conceived hypotheses and theories. Fish that don’t fit the theory are thrown back into the pond. This book is not like that. It aims to find out what sort of fish are in the pond.

(Hanks, 2013, S. 7)

Die Fähigkeit, wissenschaftliche Texte in einem entsprechenden Stil verfassen zu können, ist ein nicht wegzudenkender Bestandteil der Kompetenzen jedes (erfolgreichen) Studierenden. Da der Einfluss des Deutschen im wissenschaftlichen Bereich sich innerhalb von letzten Jahrzehnten zugunsten des Englischen verminderte (Ehlich, 2000), ist Deutsch als Wissenschaftssprache auch proportional weniger erforscht worden (Graefen, 2009). Des Weiteren sind nicht nur empirische Studien zum Thema Mangelware, sondern auch aus diesen Studien eventuell hervorgehende didaktische Materialien. Das Nichtvorhandensein von diesen mag sich verständlicherweise in einer geringeren Bereitschaft von akademischen Lehrkräften äußern, Deutsch als Wissenschaftssprache zum Thema einer Lehrveranstaltung zu machen. Erschwerend kommt hinzu, dass die Wissenschaftssprache oftmals nur eine von mehreren, zu bewältigenden Aspekten des praktischen Teils des philologischen Studiums darstellt. Dies liegt vermutlich daran, dass üblicherweise auch Bewerber zum Studium aufgenommen werden, die sprachlich zwar relativ selbstständig, jedoch noch nicht kompetent sind (ca. B1 gemäß GER).¹ Zugleich behaupten die Autorinnen des praxisbezogenen Lehr- und Arbeitsbuch *Wissenschaftssprache Deutsch*, dass die Voraussetzung für die Arbeit mit dem Buch das C1 Niveau sei, denn es werde eine souveräne Beherrschung auf C2-Niveau angestrebt (Graefen & Moll, 2011, S. 14). Angesichts dessen liegt auf der Hand, dass den Bedürfnissen der tschechischen Lernenden/Studierenden kaum Rechnung getragen wird.² Um eine solche Schwachstelle zu beheben, stehen einem viele Wege zur Verfügung, die in der Literatur von allgemeinen Unterrichts- und Lernstrategien bis hin zu konkreten Entwürfen von Aufgabentypen reichen. Der Anspruch der vorliegenden Arbeit ist somit ein praktisches Beispiel von einem dieser Wege zu liefern und dabei konkrete Instrumente mit ihren Möglichkeiten und Grenzen vorzustellen, die sowohl von Studierenden als auch Lehrkräften in Anspruch genommen werden können.

Noch genauer ist diese Arbeit zum Teil als eine Fortsetzung meiner Bachelor Facharbeit konzipiert, die später in einer bearbeiteten Fassung publiziert wurde (Mach, im Druck). Von der Methodologie her war diese korpusbasiert, quantitativ ausgerichtet und widmete sich dem Gebrauch von nominalen Verweisen (siehe Abschnitt 2.3) in der studentischen Wissenschaftssprache, wobei das verfolgte Ziel in einem Vergleich zwischen L1- und L2-Studierenden bestand. Obgleich ich mir in der vorliegenden Arbeit eine Analyse des gleichen

¹ Dass dies gängige Praxis ist, sieht man z. B. an den Einstiegsvoraussetzungen auf den Webseiten der jeweiligen germanistischen Institute in Tschechien. Alle vier, die ich mir stichprobenweise angesehen habe, erwarten das Niveau B1 von den zukünftigen Bachelor-Studierenden. (FF UPOL, 2020; FF UJEP, 2020; FF CUNI, 2020; FF MU, 2020)

² Dies soll keineswegs als Kritik des Buches gedeutet werden. Vielmehr handelt es sich bloß um einen Hinweis auf die bestehende Kluft zwischen den Anforderungen und dem tatsächlichen Können.

Phänomens vorgenommen habe, unterscheidet sich ihre Herangehensweise erheblich. Die in dieser Arbeit erbrachte Leistung fußt mithin auf drei Säulen, aus denen sich die folgenden Ziele ergeben:

- I. Vergrößerung des zwecks der Facharbeit erstellten Korpus (DESTWIKO)
- II. Vorstellung einiger Visualisierungsmöglichkeiten von Kollokationen
- III. didaktisch motivierte Analyse des kollokationellen Umfelds von ausgewählten Abstrakta

Wo also früher quantitativ verglichen wurde, wird jetzt meistens qualitativ vorgegangen, wobei auch auf die Didaktisierung Wert gelegt wird. In Worten von Hanks wollen wir hier feststellen, welche Arten von Fischen im Teich zu finden sind.

Die erklärte Zielsetzung spiegelt sich in der Struktur dieser Arbeit wider. Zunächst werden ausgewählte theoretische Aspekte wie Wissenschaftssprache, Kollokationen, und nominale Verweisung, kurz beleuchtet. Daran schließt sich die Beschreibung der neuen Version des für die Analyse verwendeten Korpus an. In diesem Zusammenhang wird auf die neuen Visualisierungsmöglichkeiten von Kollokationen im Korpusmanager Sketch Engine eingegangen, wobei ich der Frage nachgehe, welche für diese Aufgabe am besten geeignet sind. Ausgehend davon werden danach die Ergebnisse in Form von Kollokationsprofilen und konkreten Gebrauchsbeispielen ausgewählter Verweise präsentiert. Zum Schluss werden konkrete Didaktisierungsmöglichkeiten des behandelten Sprachmaterials vorgestellt, die vornehmlich auf die Vertiefung der Schreibkompetenzen von nichtmuttersprachlichen Universitätsstudierenden abzielen.

2. Forschungsstand

2.1 Wissenschaftssprache

Unter *Wissenschaftssprache* wird hier Ehlichs (1993; 1999) etwas engere Konzeption der sog. Alltäglichen Wissenschaftssprache (AWS) verstanden. Eingeführt wurde der Begriff vor allem als gewisser Gegenpol zur *Fachsprache*. Die Notwendigkeit einer Unterscheidung zwischen den beiden Begriffen scheint für uns besonders vor dem Hintergrund der Didaktik einzuleuchten; so beklagt Graefen den Verzicht auf das Erstere und die Favorisierung des Letzteren (2009):

Gesichert wurde der Erwerb der grundlegenden fachlichen Terminologie, nicht aber die Vertrautheit mit der „Alltäglichen Wissenschaftssprache“ (s.u.), die im Deutschen ein breit ausgebautes Repertoire an fachübergreifend genutzten lexikalischen und syntagmatischen sprachlichen Mitteln umfasst. (S. 106)

Ebenso entscheidend für die Entwicklung der AWS war die etwas abstraktere, doch noch unmissverständlichere Einsicht Ehlichs (1999), nämlich dass „unser wissenschaftliches Arbeiten zentral gebunden ist an jene alltägliche Sprache, die das unumgängliche Verständigungsmittel der ganzen Sprachkommunität ist“ (S. 7). Aus den beiden Zitaten, die in ähnlicher Weise Bezug auf das bereits Erworbene und damit potenziell breit Anwendbare nehmen, lässt sich wohl das Objekt der AWS-Forschung unproblematisch ableiten. Es handelt sich um eine ganze Bandbreite von Mitteln, von völlig gefestigten Syntagmen (*aus diesem Grund*) über Konstruktionen mit Leerstellen (*Was X betrifft,...*) bis hin zu relativ freien Einheiten (*ersehen, Analyse*) und ihren Kollokationen (*eine Frage erörtern*), deren Zugehörigkeit zur AWS eher eine Frage der stilistischen Angemessenheit bzw. Usualität in der gegebenen Textsorte (z. B. eine wissenschaftliche Abhandlung, Masterarbeit u. Ä.) ist. Gemeinsam soll all diesen sein, dass sie „allgemein akademisch relevant sind“ (Graefen & Moll, 2011, S. 14).

Im Folgenden wird eine kleine Auswahl von Studien präsentiert, die sich mit solchen Einheiten auseinandersetzen, und die entweder wegen ihrem Fokus auf fremdsprachliche Daten oder eben auf fachübergreifende Gemeinsamkeiten von besonderer Relevanz sind. Dabei wird das Augenmerk hauptsächlich auf die Methode gerichtet, im Gegensatz zu den Ergebnissen. Graefens (1997) oft zitiertes Buch schuf das theoretische Fundament für weitere Analysen der Wissenschaftssprache, indem da womöglich ihre gängigste Textart, der wissenschaftliche Artikel, ins Visier genommen und diese hinsichtlich der Textorganisation einer korpusbasierten Untersuchung unterzogen wurde. Ebenfalls auf einer soliden empirischen Grundlage beruht die Arbeit von Völz (2015), deren Gegenstand „textorganisierende Ausdrücke“ darstellen. Vorwiegend besticht die Studie durch ihre Komplexität; es liegen sowohl Befunde qualitativer Einzelfallstudien, die aus gezielter Arbeit mit DaF-Studierenden entstanden, als auch ein quantitativ gewonnener Anhang im Umfang von 128 Seiten, der eine funktionale Gliederung der textorganisierenden Mittel bietet. Des Weiteren werden auch die Verfasser nach ihrer akademischen Laufbahn unterschieden. Die Autorin arbeitet dementsprechend mit vier Kategorien, spricht Professoren, promovierte Wissenschaftler, Doktoranden/wissenschaftliche Mitarbeiter, und Lehrer/Lektoren. Eine weitere, rezente Monografie, die sich des Themas Wissenschaftssprache vergleichend annimmt, ist die von Brommer (2018). Die Autorin bedient sich des Konzepts Musterhaftigkeit, untermauert die Analyse mit einer minutiösen

Beschreibung des methodischen Vorgehens, und erstellt ihr eigenes, relativ umfangreiches Spezialkorpus publizierter wissenschaftlicher Aufsätze aus zwei Disziplinen (germanistische Linguistik, Medizin), um in der Auswertung signifikante N-Gramm-förmige sprachliche Muster offenlegen zu können.

Von Belang für den Kontext der tschechischen Germanistik sind weiterhin ohne Zweifel die Dissertationen von Goldhahn (2016) und Hotařová (2019). In der ersteren werden allgemeine Unterschiede zwischen tschechischen und deutschen wissenschaftlichen Artikeln kontrastiv thematisiert. Neben Äußerlichkeiten und anderen Aspekten nicht-sprachlicher Natur, wie z. B. der Gebrauch von Fußnoten, Abbildungen usw., behandelt die Autorin, wenn auch nur qualitativ, auch die Thematik der Textkommentierungen. Diese wurden in weitere Kategorien nach ihrer Funktion unterteilt und auf dieser Basis jeweils kontrastiv verglichen. Da die eigentliche Sprache nicht im Mittelpunkt der Untersuchung war und an das Thema aus einer eher funktionalen Perspektive angegangen wurde (was wiederum zu Recht die Überlegungen zur Größe des erhobenen Textkorpus beeinflusst haben müsste), erfährt man einerseits bspw. vieles zur Modalität und Personalität in den Kommentierungen, andererseits reicht das Korpus leider nicht aus, um die festgestellten Tendenzen an einer breiteren Reihe von Beispielen zu demonstrieren, geschweige denn sie statistisch zu untermauern. Die Autorin der zweiten Dissertation geht dagegen andersherum vor. Hotařová nimmt nämlich eine Palette von korpusanalytischen Instrumenten in Anspruch, um N-Gramme aus einem zusammengestellten Vergleichskorpus von Zeitschriftenartikeln zu extrahieren. Die erhaltenen N-Gramme für die beiden Sprachen wurden in vier Fallstudien (abstrakte Substantivbasen, Hedges, das „Autoren-wir“ als Unpersönlichkeitsausdruck, Argumentationsformeln) vergleichend weiter analysiert und gewähren somit einen kontrastiven Einblick in die Besonderheiten der tschechischen und der deutschen Wissenschaftssprache.

Abschließend möchte ich nun noch den Stellenwert der bereits erwähnten Arbeiten für meinen Beitrag erläutern, denn sie dienen auf die eine oder andere Weise als Ausgangspunkt für die Überlegungen, die z. T. dieser Magisterarbeit zugrunde liegen. Dies spiegelt sich bereits in der Akzentuierung von und der Notwendigkeit der Einbeziehung von empirischen Daten, wie schon 1997 Graefen mit zwangsläufig eingeschränkten Möglichkeiten verfuhr. Weiterhin ähnelt sich z. B. der Versuch, ein didaktisch anwendbares Material zu produzieren, dem Teilergebnis von Völz' Studie. Die Gemeinsamkeiten mit der Arbeit Hotařová's beruhen hingegen auf der Ähnlichkeit der analysierten Einheiten in ihrer ersten Fallstudie (Abstrakta). Überdies ermöglichten mir die schon vorgelegten Ergebnisse der N-Gram-Analyse, auf eine andere Methode der Datengewinnung zurückzugreifen, wodurch sich der Fokus von syntagmatischen Ketten auf Kollokationen verschob.

2.2 Kollokationen

Sei es Ketten, Kookkurrenzen oder sogar beides auf einmal – usualisierte Einheiten spielen eine essenzielle Rolle in Texten aller Arten. Daher erfreuen sie sich zunehmender Aufmerksamkeit in der Fremdsprachendidaktik (Reder, 2011), wo sie als unentbehrlich für den natürlichen Sprachgebrauch gelten (Reder, 2006; Nation, 2013). Da Kollokation in der Sprachwissenschaft derzeit als eine längst etablierte Begrifflichkeit gilt, wird sie im Folgenden nur grob skizziert, denn vielmehr als um eine präzise, allgemeingültige Definition geht es mir um deren praktische Handhabung in der Analyse, insbesondere im Hinblick auf die Einschränkungen, die das Korpus und der Korpusmanager (siehe Kapitel 3) mit sich bringen. Nach diesem

terminologischen Exkurs wende ich mich der Frage der lexikographischen Erfassung von Kollokationen für fremdsprachendidaktische Zwecke zu.

2.2.1 Interpretationen des Kollokationsbegriffs

Gemeinsames Auftreten von zwei (oder mehreren) Lexemen wurde in der deutschen Sprachwissenschaft so früh wie in den dreißiger Jahren des letzten Jahrhunderts von Porzig (1934) thematisiert.³ Sein Aufsatz wird zwar in fast jeder Studie zum Thema Kollokationen zitiert, um Kollokationen im heutigen Sinne handelt es sich bei Porzigs „wesenhaften Bedeutungsbeziehungen“ dennoch nicht; man wäre stattdessen geneigt, von semantischer Restrangiertheit zu sprechen.⁴ Von der Semantik als eines der bestimmenden Merkmale von Kollokationen wird teilweise auch in zeitgenössischen Werken ausgegangen, gleichzeitig wird sie aber zugunsten der Häufigkeit verdrängt, so Bußmanns Bestimmung: „charakteristische, häufig auftretende Wortverbindungen, deren gemeinsames Vorkommen auf einer Regelmäßigkeit gegenseitiger Erwartbarkeit beruht, also primär semantisch (nicht grammatisch) begründet ist“ (Bußmann, 2008, S. 391). Gleichwohl betont Firth (1957), dem auch die Einführung von Kollokation im sprachwissenschaftlichen Sinne zugesprochen wird, dass „[m]eaning by collocation is an abstraction at the syntagmatic level and is not directly concerned with the conceptual or idea approach to the meaning of words“ (S. 196). Darüber hinaus wird schon bei ihm die Häufigkeit als eine der bestimmenden Eigenschaften von Kollokationen angesehen.⁵ Im Einklang mit der Idee, dass Kollokationen einen gewissen Grad an Usualität besitzen, steht auch die Anschauungsweise von Lutzeier (2007) und Schlaefer (2009), die sie als „[m]it eingegrenzter Variation auftretende Verkettung von Wortformen, wobei mindestens eine Wortform erst durch die jeweilige Umgebung inhaltsmäßig bestimmt wird“ (S. 160) bzw. als „konventionelle, nicht idiomatisierte syntagmaförmige Verbindung wortarverschiedener Lexeme“ (S. 187) definieren. Da sich aus den beiden Definitionen erschließen lässt, dass Kollokationen als Syntagmen/Ketten zu betrachten sind, führe ich zu guter Letzt der Ausgewogenheit halber noch eine etwas widersprüchliche Erklärung von Sinclair (1991), die sich aus dem britischen Kontextualismus herleitet: „[c]ollocation is the occurrence of two or more words within a short space of each other in a text... Collocation in its purest sense recognises only the lexical co-occurrence of words.“ (S. 171) Lemnitzer & Zinsmeister merken dazu noch an, es sei damit „keine normative Bewertung hinsichtlich der Korrektheit oder Grammatikalität dieser Wortverbindung verbunden“ (S. 30). Xiao (2015) verankert seine Sichtweise ebenfalls in der kontextualistischen Tradition und zieht diesbezüglich eine äußerst wichtige Grenze zwischen der sogenannten „collocation-via-concordance“ und „collocation-via-significance“, auf die ich ausgangs des nächsten Abschnitts noch zurückkomme:

As noted earlier, the terms like “habitual” and “customary” as used by Firth (1957) are vague and impressionistic; they rely on the simple frequency counts of co-occurrence (Krishnamurthy 2002). This approach to collocation analysis, which is also adopted by neo-Firthian linguists, is labeled as “collocation-via-concordance” as opposed to “collocation-via-significance” in McEnery and Hardie (2012: 126–127). This latter

³ Für eine umfassende Entwicklungsbeschreibung des Kollokationsbegriff im Deutschen verweise ich z. B. auf Reder (2006).

⁴ Deswegen sinkt immerhin nicht die Relevanz des Aufsatzes, zumal alle aufgeführten Beispiele auch heutzutage als Kollokationen einzustufen wären, beispielshalber *Baum + fällen* bzw. *Hund + bellen*.

⁵ Zugegebenermaßen war sie jedoch in dem Beitrag kaum thematisiert.

approach depends on more rigorous inferential statistical tests than simple frequency counts and is now extensively used in collocation analysis. (S. 108)

Im Anschluss an die obige knappe Übersicht gilt es nun die Vor- und Nachteile der Definitionen in Bezug auf die hier vorgenommenen Ziele und verwendeten Instrumente abzuwägen. Es werden in der Analyse ausschließlich Substantive und deren Kollokationen behandelt. Damit diese zumindest einen minimalen Anspruch darauf erheben können, didaktisch wertvoll zu sein, sollen sie, wie Völz' (2015) Beispielsätze und Hotařová's (2019) N-Gramme, eine bestimmte Funktion erfüllen. Eine Alternative dazu wäre es, eben die Kombinierbarkeit der verschiedenen Lexeme in den Vordergrund zu rücken. Beide Ansätze versprechen unterschiedliche Ergebnisse, deren Nützlichkeit für Studierende nicht notwendigerweise allein von der Gewinnungsmethode abhängig ist. Nehmen wir z. B. die folgenden Trigramme aus Hotařová's Dissertation: *im Gegensatz dazu, aus diesem Grunde, eine entscheidende Rolle, positiven Einfluss auf*. Während die ersten zwei relativ gefestigt sind und fast keine Variation einräumen (möglich sind je nach Ko(n)text z. B. *im Gegensatz zu, im Kontrast dazu, bzw. aus diesem Grund, aus dem Grund*), sind die letzten zwei in dieser Hinsicht das pure Gegenteil davon. Vielen Nomina können ja die Eigenschaften *entscheidend*, bzw. *positiv* zugeschrieben werden, ebenso können den Substantiven *Einfluss* und *Rolle* eine Mehrzahl von unterschiedlichen Adjektiven attribuiert werden. Es ist diese Variation, die uns hier interessiert. Wie bereits angedeutet, schlägt sich deren Absenz im Wesentlichen darin nieder, dass man meistens textorganisierende Ausdrücke auffindet, wozu sich eben die N-Gramm-Analyse besonders gut zu eignen scheint. Wenn man aber nicht organisieren, sondern bilden will, tut man es mithilfe von Sätzen, deren Satzglieder – den syntaktischen Regeln entsprechend – kaum in direkter Reihenfolge zueinanderstehen. Diese Feststellung hat zur Folge, dass ich mich hier der Sinclairschen Auffassung über die Nichtlinearität von Kollokationen anschließe. Folglich werden also Instrumente benutzt, die Zugriff auf Daten in einem bestimmten Bereich um das Lexem herum ermöglichen. Obgleich die Datenmenge im Korpus zur detaillierten Erkundigung des jeweiligen Lexemumfelds genügen mag, wird sie gewiss nicht zu einer statistischen Auswertung von dessen Kollokaten ausreichen (zumindest nicht bei allen Substantiven). Daher müssen wir uns hier mit „collocation-via-concordance“ zufriedengeben.

2.2.2 Kollokationen lexikographisch-didaktisch perspektiviert

Auch wenn sich Sprachwissenschaftler mittlerweile auf eine Parametrisierung von Kollokationen einigen könnten, würde der Weg zu ihrer erfolgreichen Didaktisierung wohl sowieso nicht weniger dornig ausfallen. Einerseits macht die Absenz eines Konsensus die Situation um ein Vielfaches komplizierter, andererseits ermöglicht sie im Rahmen des Dreiecks Kollokationstheorie – Kollokationslexikographie – Kollokationsdidaktik (Đurčo P. , 2016) eine Vielfalt von Strategien zu erproben und diese auf konkrete Zielgruppen von Lernenden abzustimmen.

Dass dies jedoch meistens der Realität fernbleibt und immer noch eher als Desiderat aufzufassen ist, dokumentiert Đurčo (2016) am Beispiel pädagogischer Dokumentation. Von 1401 Lexemen waren nur 62 (4,4%) Kollokationen. Die Erklärung scheint die übergreifende Konzeption zu betreffen:

Die nationalen und internationalen Curricula in offiziellen Dokumenten zum Sprachenlernen enthalten jedoch keine durchdachten Konzepte dieser Problematik und

in den Wortschatzlisten der lexikalischen Minima kommen die Kollokationen nur sporadisch und offensichtlich zufällig vor. (S. 153)

Falls Kollokationen doch schon präsent sind, sind Probleme nicht ausgeschlossen; ganz im Gegenteil. Neben Aktualität, Angemessenheit, Vollständigkeit der Form und didaktischer Relevanz beklagt Ďurčo beispielsweise auch

die Reduktion einer reicheren Auswahl der Kollokate, z. B. neben *Anregungen bekommen* fehlen nicht nur „elementare“ Kollokationen *Anregungen geben, finden* aber auch semantisch und stilistisch relevantere spezifische Kollokationen, wie *Anregungen aufgreifen, verarbeiten, empfangen, holen, vorbringen, entgegennehmen, aufnehmen, weitergeben, sammeln, liefern, folgen, vortragen, umsetzen...* (S. 160)

Vor diesem Hintergrund liegt es allerdings nahe, dass Lehrbücher sowie Kollokationswörterbücher immer ein Kompromat und somit ein Kompromiss darstellen. Wie denn solche Kompromisse zustande kommen und was dabei abzuwägen gehört, erörtere ich im folgenden Abschnitt. Bevor ich zu den Einzelheiten komme, möchte ich nun noch verdeutlichen, dass ich mich ausschließlich mit herausgegebenen Produkten der korpusbasierten Forschung beschäftige.⁶

Aus dem strukturellen Blickwinkel haben wir oben zwar die rein syntagmatisch definierten Einheiten abge sondert; da war jedoch weder von Lexikografie noch von Didaktik die Rede. So eröffnet sich teilweise wieder die Diskussion darüber, welche Einheiten didaktisch am nützlichsten sind. Kathrin Steyer (2000) beispielsweise plädiert für das Konzept der usuellen Wortverbindungen (UWV). Folgendes ist besonders kennzeichnend für den Ansatz:

Korpusgesteuerte qualitative Sprachbeschreibung ist primär pragmatisch ausgerichtet. Der Topos, dass Bedeutung Gebrauch ist, manifestiert sich in diesem methodischen Paradigma auf besonders ausgeprägte Weise. Funktionales wird aus der Nische subjektiver Perspektiven auf einzelne sprachliche Akte herausgeholt und durch die Möglichkeit, massenhafte Gebrauchsmuster zu erkennen, »kollektiviert« und damit besser generalisierbar. (Steyer, 2009, S. 120)

Analog zu Hotařovás Dissertation werden Syntagmen gesucht, die danach anhand ihrer pragmatischen Leistung gruppiert werden. „UWV sind somit »funktionale Chunks«. UWV werden nun unter den beiden zentralen Bestimmungsmerkmalen »lexikalische Spezifiziertheit« und »funktionale Gebundenheit« [] betrachtet“ (S. 129). Praktisch wirkt sich die funktionale Gebundenheit so aus, dass strukturelle Merkmale wie Wortart und Valenz als sekundär zu betrachten sind. So sind laut Steyer beim Lemma *Grund* vier Gruppen von Adjektiven zu finden: zwei Sets von Qualifikationsadjektive, positive (*triftig, plausibel, verständlich*) und negative (*unerfindlich, unerklärlich, fadenscheinig*), Benennung (*familiär, persönlich, humanitär* usw.) und Quantifikation (*zweierlei*). Ähnlich sehen die UWV mit *Idee* aus; nehmen wir z. B. das Set Verbreitung (*~ aufgreifen, ~ umsetzen, zu einer ~ anregen, auf eine ~ bringen, jmdm. eine ~ nahe bringen* usw.). Der unbestrittene Vorteil der UWV besteht meiner Ansicht nach in dem gelungenen Versuch, syntagmatische Festigkeit und paradigmatische Variation in diesen Syntagmen unter einen Hut zu bringen; vgl. Steyer (2008):

⁶ Für einen Vergleich von Kollokationswörterbücher einschließlich einiger Webseiten siehe Hanzlíčková (2016).

Kollokationsmuster fokussieren die konzeptuelle Austauschbarkeit lexikalischer Vertreter in einem Paradigma. Wortverbindungsmuster beziehen sich auf ihre systematische Fixiertheit im ganzen Syntagma. Usuelle Wortverbindungen bilden das zu Grunde liegende Schema ab, mit der entsprechenden Funktion, die es in der Kommunikation realisiert... (S. 198)

Eine weitere hilfreiche Gliederung innerhalb der UWV erfolgt nach dem Potenzial der lexikalischen Austauschbarkeit. Steyer (2009, S. 132) unterscheidet zwischen

- Vorkommen mit minimaler lexikalischer Varianz (*auf eine Idee kommen*)
- Vorkommen mit eingeschränkter lexikalischer Varianz (*eine Idee ist geboren*)
- Vorkommen mit systematischer lexikalischer Varianz (*gute/zündende/schöne Idee*)

Wie jedes Konzept, das seine Existenz Kompromissen verdankt, gibt es auch an UWV einiges zu beanstanden. Das erste Problem mag einem gewissermaßen haarspalterisch vorkommen, denn es betrifft die semantische Zuordnung von Kategorien. Diese wirkt sich aber auf die Konsistenz der Ergebnisdarstellung aus. Bleiben wir beim Lemma *Idee*; man kann Ideen *haben, vergessen, umsetzen*, Ideen können *überzeugen, begeistern, überraschen, funktionieren, auf etwas basieren* usw. Es lassen sich für diese Verben nur schwer gemeinsame semantische Merkmale finden. Unter diesem Gesichtspunkt erscheint es auch nicht sonderlich praktisch, für fast jedes Lexem ein semantisches Netz zur Beheimatung aller relevanten Kollokate zu entwerfen. Das zweite Problem ist auf die Implikationen des folgenden Zitat Steyers (2008) zurückzuführen:

Betrachtet man die KWICs, wird deutlich, dass sogar innerhalb ein- und desselben Kotextclusters - trotz identischer lexikalischer Affinität - sehr verschiedene Funktionsbedeutungen zu lokalisieren sind. (S. 133)

Diese Feststellung läuft faktisch darauf hinaus, dass in einem Wörterbucheintrag fast gleiche Strukturen mehrmals aufgeführt wären, und zwar nur deswegen, weil sie pragmatisch anders sind. Durch diese Optik, so Steyer, könnte man die UWV *auf eine/die Idee kommen* als entweder Entstehung, Optionsangabe, Zurückweisung oder Frage interpretieren. Das Pragmatische übermitteln zu wollen, ist einerseits nachvollziehbar, andererseits könnte es ebenso gut zu Verwirrungen führen. Außerdem würde man viele (längere) Beispiele anführen müssen, um den pragmatischen Wert der jeweiligen UWV zu illustrieren. Am Ende des Tages würden den Lernenden wieder Syntagmen serviert.

Ein ähnlicher Ansatz im Hinblick auf den Untersuchungsgegenstand – das Typische – ist der von Zita Hollós. Sie beruft sich auf Feilkes (2004, zit. n. Hollós, 2010) Drei-Ebenen-Modell, in dem

die breite Basis der Kollokationen im weitesten Sinne das Fundament für die Gesamtheit der mehr oder weniger festen Wortverbindungen bildet, statt dass - wie das in der Phraseologieforschung bisher der Fall war und ist - das Irreguläre und Selten-Vorkommende fokussiert wird. (S. 86)

Demnach wird zwischen Kollokationen, Kombinationen und Phrasemen unterschieden (siehe Abb. 1):

Kombinationen bilden dabei eine zwischen Kategorie zwischen nicht- oder schwach idiomatischen Phrasemen, wobei sie im Vergleich zu den Kollokationen einen höheren Grad an Festigkeit und Idiomatizität aufweisen (S. 95)

Die Kategorien sind nicht genau festgelegt und es kann daher zu Überschneidungen kommen. Im Unterschied zu Steyers UWV sind sie aber nicht pragmatisch bestimmt.

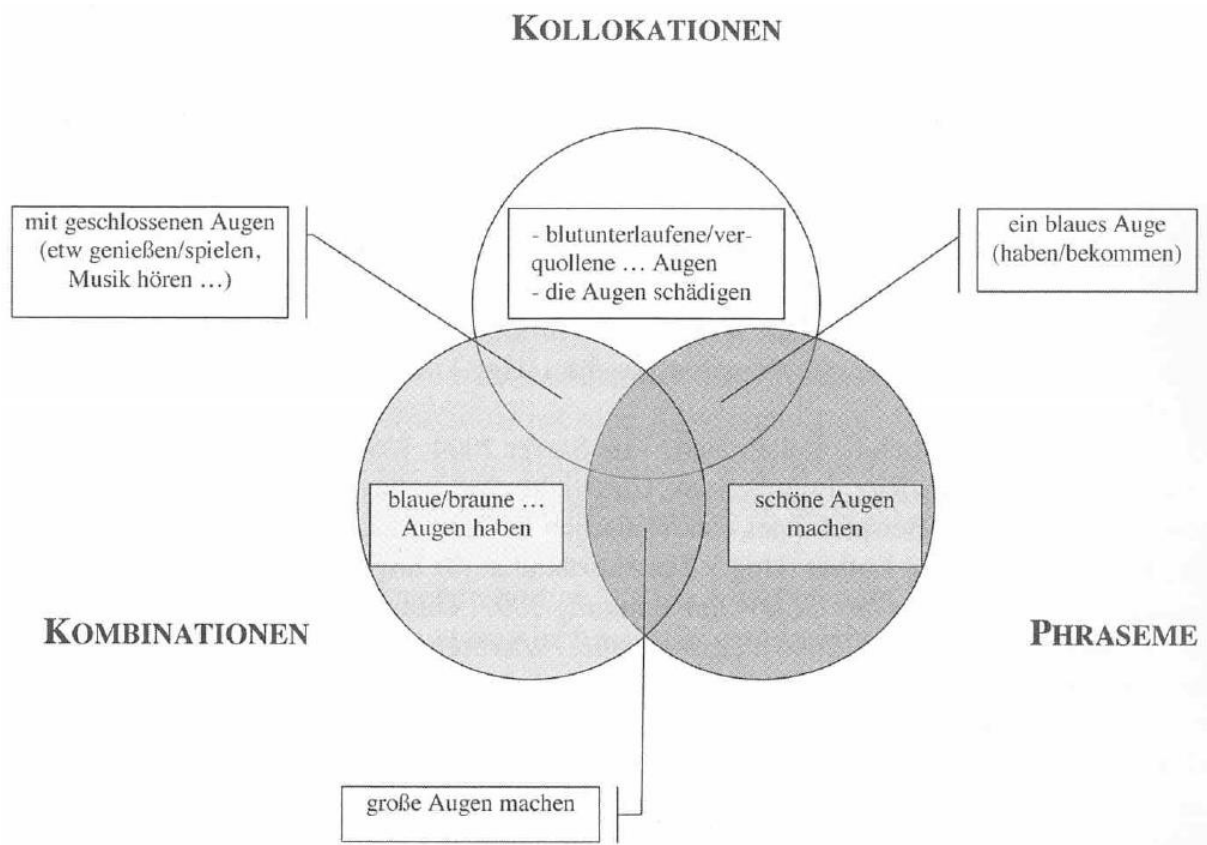


Abbildung 1: Die Unterscheidung von Kollokationen, Kombinationen und Phrasemen (nach Hollós, 2010)

Bislang habe ich ansatzweise einige lexikografisch-relevante Einheiten aufgezeichnet, die während der Erarbeitung von Wörterbüchern in Frage kommen. Angesichts dessen werde ich an dieser Stelle drei verschiedene Beispiele von Wörterbüchern kurz ansprechen.

Als erstes nehmen wir *Wörterbuch der Kollokationen im Deutschen* (Quasthoff, 2011). Es handelt sich um ein Kollokationswörterbuch schlechthin, denn mit der Ausnahme von den eigentlichen Kollokaten, ihrer Wortart und – streckenweise – Valenz wird auf andere Angaben verzichtet. Zu der inneren Struktur eines Eintrags merkt Quasthoff Folgendes an:

Die Kollokationen sind durch fette Punkte in Kollokationsgruppen eingeteilt. Die Wörter in einer solchen Kollokationsgruppe sollten syntaktisch und semantisch ähnlich sein. Innerhalb jeder Kollokationsgruppe sind die Wörter alphabetisch angeordnet. (S. X)

Wie aus dem Eintrag zum Wort *Abend* zu ersehen ist, dürfte die Bedeutungsähnlichkeit hin und wieder auf den ersten Blick nicht ganz erkennbar sein.

Abend der

V: NOM. beginnen · ausklingen · enden · misslingen · AKK. ankündigen · ansetzen · planen · eröffnen · inszenieren · moderieren · präsentieren · erleben · verbringen · abrunden · beenden · absagen · jmdm. verkürzen · vermiesen
A: außergewöhnlich · beeindruckend · denkwürdig · ereignisreich · feierlich · festlich · geschichtsträchtig · interessant · amüsant · beschaulich · beschwingt ·

Abbildung 2: Wörterbucheintrag von *Abend* nach Quasthoff (2011)

Das zweite Beispiel, *Feste Wortverbindungen des Deutschen: Kollokationenwörterbuch für den Alltag* (Buhofer, Dräger, Meier, & Roth, 2014), ist im Prinzip eine benutzerfreundliche Alternative. Im Gegensatz zu Quasthoffs Wörterbuch enthält das Nachschlagewerk Beispielsätze und mehr Informationen zum Gebrauch, sei es Konstruktionen oder stilistische Hinweise. Des Weiteren wirkt die grafische Darstellung besser durchdacht und die beachtliche Anzahl von Querverweisen erleichtert die Orientierung maßgeblich.

Abend m

ADJEKTIVE/ADVERBIEN

■ **schöner** | **guter** | **gemütlicher** A. *Wir verbrachten einen gemütlichen Abend bei einer Flasche Wein.* | **gelungener** | **angenehmer** | **netter** A. ■ **festlicher** | **geselliger** | **fröhlicher** | **bunter** A. ◊ ›Abendveranstaltung mit abwechslungsreichem Programm◁ *Drei Freundinnen organisierten einen bunten Abend, an welchem wir Spiele machten und Geschichten erzählten.* ■ **der frühe** | **späte** A. | bis **spät** in den A. ◊ ■ **der heutige** | **gestrige** A. *Der gestrige Abend wird mir für immer in Erinnerung bleiben.* ■ **der ganze** A. | **ruhiger** A.

VERBEN

■ **verbringen** einen A. mit jmdm./etw. **verbringen** | einen A. irgendwo **verbringen** *Ich verbrachte den ganzen Abend auf dem Flughafen und wartete auf meine Freundin.* ■ zu A. **essen** ◊ *Er lädt sie ein, bei ihm zu Hause zu Abend zu essen.* | jmdm. einen schönen/angenehmen/... A. **wünschen**

Abbildung 3: Wörterbucheintrag von *Abend* aus Buhofer et al. (2014)

Drittens ist *KOLLEX: deutsch-ungarisches KOLLOkationsLEXikon* (Hollós, 2014a) zu erwähnen. Es geht um ein bilinguales Wörterbuch, in dem deutsche Kollokationen und ihre ungarischen Äquivalente enthalten sind. Die Autorin weist darauf hin, dass KOLLEX „kein reines Kollokationswörterbuch [ist], sondern eine Kombination von Kollokations- und Valenzwörterbuch, also ein syntagmatisches Wörterbuch“ (Hollós, 2014b, S. 117). Diese Konzeption erfordert Vereinigung sämtlicher Aspekte, die den integrativen Kollokationsbegriff (Hollós, 2004, zit. n. Hollós, 2014b) ausmachen:

1. Fremdsprachendidaktik
2. Semantik
3. Korpuslinguistik
4. Syntax (Valenz) – Morphosyntax (Gebrauchspräferenzen und -restriktionen)

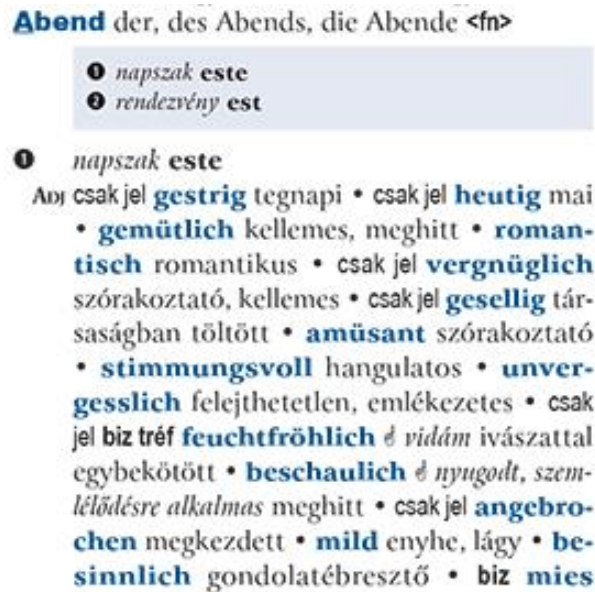


Abbildung 4: Wörterbucheintrag von Abend aus Hollós (2014a)

In Anbetracht des korpusbasierten Ansatzes soll man sich vor Augen führen, zu welchen Arten von relevanten Daten man nur mithilfe von Korpora ohne zusätzliche Analysen gelangen kann. In dieser Hinsicht erwähnt Quasthoff (2009):

1. Häufigkeitsangaben zu Wörtern, evtl. in verschiedenen, z. B. zeitlich geordneten Korpora: Diese eignen sich zur Stichwortauswahl für ein Wörterbuch oder zur Beurteilung der Gebrauchsentwicklung.
2. Nutzung von Beispielsätzen: Die Vorkommen einzelner Wörter lassen sich für die Auswahl von Belegstellen verwenden, für die Erstellung von Bedeutungsbeschreibungen sowie für die quantitative Beurteilung verschiedener Wortbedeutungen.
3. Wortkookurrenzen: Statistisch auffällig oft gemeinsam auftretende Wörter können durch verschiedenartige Relationen miteinander verknüpft sein. Sie lassen sich beispielsweise benutzen bei der Untersuchung von Kollokationen oder der Klassifikation von Wörtern nach Sachgebieten. (S. 161)

Das Erste wird in dieser Arbeit nur für die Lemmata benutzt werden, um die allgemeinen Gebrauchsunterschiede zwischen der L1- und L2-Gruppe von Studierenden. Das Zweite ist der Mittelpunkt der Untersuchung. Beispielsätze verhelfen der Kontextualisierung, veranschaulichen kontextuelle Einschränkungen und gewährleisten, dass der Leser attestierten Verwendungsbeispielen ausgesetzt wird. Laut Kilgarriff, Husák, McAdam, Rundell, & Rychlý (2008)⁷ lassen sich gute Beispielsätze anhand folgender Eigenschaften charakterisieren: Satzlänge, Häufigkeiten von Wörtern im Beispielsatz, Absenz von pronominalen Anaphern, die

⁷ Die Studie sowie die Kriterien darin orientieren sich auf eine *automatische* Identifizierung wertvoller Beispielsätze.

illustrierte Kollokation als Hauptsatz, Sätze mit den sog. „third collocates“⁸ und Sätze, die zuerst den Kontext bieten und erst dann kommt die Kollokation. Der dritte Punkt, Wortkookkurrenzen im Sinne von einem gemeinsamen Auftreten, wo statistische Signifikanz nachgewiesen ist, käme wegen der beschränkten Datenmenge eher sporadisch zum Tragen. Dass nicht jede angeführte Instanz gemeinsamen Vorkommens sich im Falle von SNVs in DESTWIKO (v0.2) pauschal als signifikant bezeichnen lässt, heißt nicht, es werden keine Kookkurrenzen ausgewertet werden. Bei solcher Auswertung geht es nach Steyer (2009) vorrangig um folgende Fragen:

1. Welche Binnenstruktur weist ein Kookkurrenzcluster auf?
2. Welche typischen Variationen und Gebrauchsrestriktionen lassen sich erkennen (u.a. Festigkeitsgrad/variable lexikalische Ausfüllungen)?
3. Gibt es andere Kookkurrenzcluster, die vergleichbare Merkmale und Gebrauchseigenschaften aufweisen? (S. 127)

Die Antworten auf 1 und 2 sollte man bis zu einem gewissen Grad mithilfe der Beispielsätze eruieren können. Die letzte Frage steht freilich nicht im Vordergrund der Untersuchung, denn sie ist z. T. pragmatisch motiviert. Wollte man sie strukturell bzw. semantisch interpretieren, würde es wohl reichen, auf authentische Gebrauchsbeispiele des Lexems bzw. der ganzen Kombination Blick zu werfen. Am besten würde man allerdings das Korpus benutzen, denn dieses bietet deutlich mehr Sätze per Kollokationspaar. So erhöht sich auch wesentlich die Wahrscheinlichkeit, dass z. B. die besagten Gebrauchsrestriktionen zutage treten.

Zusammengefasst lässt sich feststellen, dass für die Behandlung der Kollokationen als eines lexikographischen Phänomens unterschiedliche Aspekte maßgeblich sind, darunter auch die Perspektive der Didaktik. Ich habe an drei Beispielen zu zeigen versucht, wie sich diese Aspekte auf die finale Form von Wörterbüchern auswirken können und welche Bedeutung ihnen in diesem Projekt beigemessen wird.

2.3 Satzbezogene nominale Verweise (SNVs)

2.3.1 Allgemeine Begriffsbestimmung

An dieser Stelle werden zunächst einige Merkmale der nominalen Verweisung beleuchtet. Eine detaillierte Zusammenfassung der einschlägigen Literatur zum Thema lexikalische Verweisung ist im Artikel von Mach (im Druck) zu finden, der gerade eben diesem Thema gewidmet war und der sich mithin als schlüssig für die Auswahl des zu analysierenden Materials erweist.

SNVs sind aus der Suche nach einer lexikalischen Einheit entstanden, die im Sinne von Halliday & Hasan (1976) zur Textkohäsion beiträgt, und in der zugleich gewisse semantische, strukturelle und funktionale Eigenschaften enthalten sind. Das wichtigste Erkennungsmerkmal ist die sog. temporäre Hypostasierung, d. h. “potential to create context-dependent concepts and to reify the information to which they are linked” (Schmid, 1998). Ähnliche Konzepte wurden schon von Koeppel (1993) fürs Deutsche, und Ivanič (1991), Conte (1996) fürs Englische postuliert. In den folgenden Sätzen werden zum Beispiel die unterstrichenen Teile hypostasiert; ihre Inhalte werden für eine beschränkte Zeit zum Inhalt der zwei fettgedruckten Nomina (SNVs).

⁸ „words that occurred with high salience in sentences containing the node and primary collocate“ (Kilgarriff et al., 2008, S. 426)

(1) *Coninx wertete eher negativ Frischs **Versuche**, Schriftsteller zu werden oder als Journalist zu arbeiten. (0003T)⁹*

(2) *Das **Ziel** dieser Analyse war vor allem die angewandten Sprachformen in diesen Texten zu bestimmen, und somit die These überprüfen, dass die Standardsprache primär in ihrer schriftlichen Form und der Dialekt in seiner mündlichen Form gebraucht wird. (0001T)*

Mithilfe von SNVs können also Autoren Texte kohäsiver, kohärenter und kürzer gestalten. Wie den Beispielen 3-4 zu entnehmen ist, kann man dies sowohl anaphorisch als auch kataphorisch realisieren:

(3) *Außerdem gibt es bis dato in der Fachliteratur noch keine vergleichbare Forschung dazu. Das ist ein weiterer **Grund**, warum ich mich der Beantwortung dieser Fragen annehme. (0002D)*

(4) *Aufgrund der **Tatsache**, dass sie aber auch eine große Gemeinsamkeit haben, nämlich dass beide Systeme auf dem Phänomen der Kongruenz basieren, wird in vielen zeitgenössischen Publikationen argumentiert, dass keine Unterscheidung notwendig ist. (0002D)*

Aus der semantischen Perspektive handelt es sich bei den SNVs eindeutig um Abstrakta. Diese werden dann nach ihrer Bedeutung in weitere Kategorien aufgeteilt. In größeren Werken wie Schmid (2000) und später auch Flowerdew & Forest (2015) wird meistens zwischen sechs semantischen Klassen unterschieden: faktual, linguistisch, modal, mental, eventiv, und zirkumstantiell. Aufgrund des Fokus auf Kollokationen werden diese Differenzierungen in der vorliegenden Arbeit nicht eigens berücksichtigt.

Strukturell gesehen treten SNVs – wie aus der folgenden Tabelle ersichtlich – in einer relativ beschränkten Menge von Konstruktionen auf. Es werden Kollokate von getaggten SNVs aus der älteren Version des Korpus (DESTWIKO [v0.1]) angezeigt. Geordnet wurden sie nach ihrem logDice Wert (Rychlý, 2008).

	Lemma	Kookkurrenzen	LogDice
1	<i>dass</i>	765	9,45
2	<i>diese</i>	889	9,14
3	<i>ob</i>	98	8,94
4	<i>meine</i>	98	8,68
5	<i>nach</i>	213	8,35
6	<i>darin</i>	46	8,33
7	<i>dafür</i>	48	8,27
8	<i>bestehen</i>	52	8,15
9	<i>stellen</i>	66	7,86
10	<i>liegen</i>	42	7,76
11	<i>aus</i>	167	7,71
12	<i>sein</i>	586	7,51

Tabelle 1: Kookkurrenzen von SNVs in DESTWIKO (v0.1)

⁹ Das Sigel besteht aus der Dokumentnummer und der Angabe zur Muttersprache des Verfassers (D – Deutsch; T – Tschechisch).

Die anaphorisch verwendeten SNVs zeichnen sich dadurch aus, dass sie meistens durch definite Pronomina und seltener auch Artikel determiniert werden, denn sie verweisen in der Regel auf den Inhalt des vorhergehenden Satzes, der hiermit schon als bestimmt gilt. So wird der Inhalt zum Thema des Folgenden Satzes gemacht, weswegen sich die Mehrheit von Konstruktionen [dies.*] + [SNV] im Korpus jeweils am Anfang eines Satzes befindet. Der Beitrag eines SNV zur Kohäsion wird gut an dem folgenden Beispiel mit *Haltung* veranschaulicht:

(5) *Zusammenfassend kann gesagt werden, dass die oben genannten Forscher Köpcke und Zubin (1982), Corbett (1991), Bittner (2006) u.a. davon ausgehen, dass es in allen Sprachen, auch in jenen, die ein auf den ersten Blick höchst komplexes und undurchsichtiges Genusssystem haben, ein Regelsystem gibt, von welchem somit auch Fremdsprachlernende Gebrauch machen können. Auch ich schließe mich **dieser Haltung** an (vgl. Schwarze 2008: 88). 1.3.2 Typen der Genuszuweisung nach Corbett Corbett (1991) untersuchte über 200 Sprachen, um deren Regelwerk herauszuarbeiten. (0002D)*

Mit den kataphorisch gebrauchten SNVs ist die Situation ein wenig komplizierter, denn sie können entweder einem Infinitiv-(6-7) oder einem *dass*-Satz (8-9) bevorstehen. Darüber hinaus kann, wie die obige Tabelle zeigt, zwischen dem SNV und dem Nebensatz ein Kopula-(7) oder Pseudokopulaverb (9) stehen.

(6) *Die **Idee**, auf Dialekt zu verzichten, um die Hochsprache besser sprechen zu können, ist jedoch undenkbar, weil der Dialekt zu einem wesentlichen Bestandteil des Nationalbewusstseins der Deutschschweizer geworden ist. (0001T)*

(7) ***Aufgabe** dieser Arbeit war es zu untersuchen, ob es nach wie vor lohnend erscheint, Literaturgeschichten zu schreiben und zu verlegen, oder ob diese mittlerweile Relikte aus vergangenen Zeiten sind, die lediglich aus Prestigegründen weitergeführt werden. (0021D)*

(8) *Sie hatten den **Eindruck**, dass Lernende aufgrund der Undurchsichtigkeit des Genussystems vermehrt auf ihrer Muttersprache zurückgriffen. (0002D)*

(9) *In diesen Systemen wird die Klassenzugehörigkeit zwar wie bei den Deklinationsklassen an-hand von Affixen aufgezeigt, der **Unterschied** liegt jedoch darin, dass die Affixe in Genus- und Nominalklassensystemen nicht nur am Nomen selbst (obwohl das nicht zwingend der Fall sein muss), sondern vor allem an den mit ihnen verbundenen Wörtern, sowohl innerhalb wie auch außerhalb der Nominalphrase, erscheinen. (0002D)*

2.3.2 Probleme mit der Operationalisierung

Die oben skizzenhaft zusammengefasste Operationalisierung von SNVs in Mach (im Druck) war durch den Versuch motiviert, die drei möglichen Beschreibungsebenen in ein theoretisches Konstrukt zu integrieren. Um die Instanzen dieses Konstrukts zwecks der Analyse leichter auffindbar zu machen, wurden alle Substantive in DESTWIKO (v0.1), die folgenden Bedingungen entsprachen, mit Tags versehen:

- I. Ist das Nomen ein Typ (/abstrakt)?
- II. Wird das Nomen von einem anderen Textteil in seiner Bedeutung erweitert? / Fasst das Nomen einen Textteil zusammen?
- III. Ist dieser Textteil ein Satz bzw. könnte er ihn bilden?

- IV. Könnte dieser Textteil so umschrieben werden, dass er als Argument des Nomens entweder in Form eines *dass*- oder Infinitivsatzes fungiert? (S. 11)

Im Nachhinein betrachtet ist der Einbezug von allen drei Perspektiven zwar methodologisch inklusiv, indem keiner zu viel Gewicht beigemessen wird, er bringt aber weitere Probleme mit sich. Jede der drei Ebenen enthält nämlich distinktive Kriterien und benötigt somit eine Reihe von zusätzlichen Entscheidungen. Ein Beispiel hierfür wäre die Frage, ob nominale Gruppen im Genitiv als SNVs zu zählen sind (vgl. 10 und 11).

(10)

- a. *Das **Problem** der mangelnden Überprüfbarkeit der Dialektkompetenz, wie von BUCHELI BERGER/GLASER/SEILER (2012: 100) angeführt, ist außerdem zu hinterfragen.* (0016D)
- b. *Neben diesen Bemühungen stand stets die **Frage der Definition**, beziehungsweise der Eingrenzung des Kunstmärchens im Raum.* (0029D)

(11)

- a. *Dabei entsteht das **Problem**, dass es sich bei den Informanten meistens um keine ausgebildeten Linguisten, die auch Phonetik beherrschen, handelt.* (0215D)
- b. *Sprachliche Handlung erfolgt über sprachliche Mittel, es bleibt nur die **Frage**, was in dieser Definition unter 'sprachliche Mittel' verstanden wird.* (0115D)

Vergleichbar problematisch sind z. B. der Status von Satzinhalten, auf die jedoch nicht explizit verwiesen (12–14) wird oder Substantive, die einerseits abstrakt sind, verfügen andererseits nicht über die Fähigkeit, propositionale Inhalte in sich zu tragen (*Analyse, Modell, Ansatz, Ebene*). Angesichts dessen mag die eindimensionale Operationalisierung etwa konservativ erscheinen; es ist nichtsdestotrotz festzustellen, dass die identifizierten Substantive einen bedeutenden Teil der AWS darstellen. Wollte man außerdem weniger Ausschlusskriterien haben, und nur eine der drei Ebenen einbeziehen, wäre dies im Falle der Strukturebene mit gewöhnlichen Korpusinstrumenten durchführbar. Bräuchte man beispielsweise alle Substantive mit einem nachfolgenden *dass*-Satz, könnte man dies mithilfe einer einfachen Abfrage tun und die entsprechende Liste erstellen.

- (12) *Alle Ehefrauen widerlegen, dass sie Schaad gewaltsam behandelte, was zu seinem Gunsten spricht. Eine von Ehefrauen behauptet sogar, dass er sie nicht kränken könnte. Eher würde er sich selbst verletzen. Ähnliche **Idee** erscheint auch in Montauk. Frisch wendet Gewalt nie an. Auch in der Ehe mit Rosalinde machte er keine Ausnahme....* (0003T)

– impliziter, kontextueller Verweis; weiter kompliziert durch die Entfernung von seinem Inhalt und durch das Wort *ähnlich*

- (13) *Für didaktische Zwecke ist es also notwendig, die konsonantischen Endungen dieser Nomina feiner aufzugliedern und auf ihre Regelmäßigkeiten bezüglich des Genus zu untersuchen. Bull stellte sich 1965 dieser **Herausforderung** und prüfte die häufigsten konsonantischen Endungen.* (0002D)

– durch *dieser* und die Wiederholung des Satzinhaltes expliziter gemacht

- (14) *Die Durchführung im Klassenzimmer barg zwar den **Nachteil**, dass der Erhebungsleiter immer vor Ort sein musste, um die Fragebögen auszugeben und wieder einzusammeln und die Befragten eventuell beeinflusste,* (0002D)

– expliziter Verweis

Im Lichte der obigen Ausführungen erscheint die Option, das SNV-Tagging auf die zweite Version des Korpus auszudehnen, als überflüssig, und zwar aus folgenden Gründen. Erstens würde eine Analyse einer größeren Menge von Substantiven wegen des damit verbundenen Aufwands über den Rahmen dieser Arbeit hinausgehen. Zweitens sind die 93 verschiedenen SNVs aus der vorherigen Untersuchung für eine Studie dieser Art vollkommen ausreichend. Die eigentlichen untersuchten Substantive werden also der angeführten Facharbeit entnommen, die jetzt berücksichtigte Datenbasis für die Auswertung des Gebrauchs der jeweiligen Lexeme ist hier jedoch umfangreicher. In diesem Zusammenhang muss auch hervorgehoben werden, dass die Lexeme hier zuerst ohne jegliche Rücksicht auf ihren Status als SNVs studiert werden. Dieser wird zwar *post hoc* berücksichtigt, wirkt sich aber gar nicht auf die Auswahl der Lexeme oder Beispiele.

3. Korpus

3.1 Korpusdesign

Zu Beginn dieses Kapitels wird hier auf einige allgemeine Kriterien des Korpusaufbaus eingegangen. Zu den wichtigsten gehört zweifelsohne die Repräsentativität, die McEnery & Hardie (2012) als ein nicht erreichbares Desiderat hinstellen. Dies ist insbesondere der Fall für Korpora, die ein breites Spektrum der Sprache erfassen sollen, z. B. gesprochene Sprache, Sprache im Zeitraum 1980-1989, Sprache der Nichtmuttersprachler usw. Für Spezialkorpora mit einem engeren Fokus erweist sich das Kriterium Repräsentativität als verhältnismäßig realistischer; wenn man beispielsweise den Mediendiskurs zu der Migrationskrise in deutschen Onlinemagazinen und -zeitungen untersucht, werden durch solche Fragestellung an sich konkrete Quellen, Textsorten und Zeiträume vorausgesetzt und andere hingegen ausgeschlossen. Man würde demnach weder mit Fernsehreportagen oder wissenschaftlichen Monografien noch mit Zeitungsartikeln von 2000 bis 2010 arbeiten. Auch wenn derartige Limitationen hinsichtlich der Zielsetzung vorliegen, entfällt das Kriterium nicht gänzlich. Entweder sind die Quellen in ihrer ganzen Bandbreite (alle politischen Richtungen; Blogs, Artikel, Reportagen usw.) einzuschließen, oder, falls sie eben nicht vorhanden sind, muss man die Interpretation der Endergebnisse dementsprechend modifizieren. Sie dürften etwa nicht zu dem Ausmaß verallgemeinert werden, wie es bei einem balancierten und repräsentativeren Korpus zulässig wäre, und die Aussagekraft der Auswahlkorpora wird entsprechend geringer.

Neben Repräsentativität sollte sich man in seinem Bestreben nach einem brauchbaren Korpus auch Gedanken über seine Größe machen. Obgleich die Korpusgröße mit Repräsentativität manchmal einhergeht, ist sie bei Weitem nicht das einzige Kriterium, mithilfe dessen man Repräsentativität gewährleisten kann. Äußerst ausschlaggebend sollen dabei zwei Faktoren sein, erstens die theoretische Bestimmung der Grundgesamtheit, in anderen Worten was alles gehört der untersuchten Population an, und zweitens die zur Stichprobenerhebung herangezogenen Methoden. (Biber, 1993, S. 243) Wenn wir aber die Korpusgröße unabhängig von Repräsentativität betrachten würden, würde es ans Lächerliche grenzen, zu behaupten, es bestünde kein Unterschied zwischen einem Korpus mit 1 Mio. Wörter und einem, das 100-mal größer ist. In dieser Hinsicht vgl. z. B. Flowerdew (2012): “[i]n general, researchers are unanimous in arguing that in principle the more text, the better” (S. 4). Präziser ausgedrückt sei die Korpusgröße laut McEnery & Wilson (2001)

ultimately dependent on the distribution of linguistic features, that is...frequent items are stable in their distributions and hence small samples are adequate for these. Rarer items on the other hand show more variation in their distributions and consequently require larger samples if they are to be fully represented in the corpus... (S. 80)

Es ist mithin der Forschungszweck, der immer zu berücksichtigen ist. So wird eine syntaktische Analyse in der Regel deutlich weniger Material beanspruchen als eine, der sich lexikalischen Besonderheiten widmet.

Bei der Beurteilung von Repräsentativität und Größe kommen auch verschiedene Varietätsaspekte zum Tragen. Davies (2015) erwähnt in diesem Zusammenhang die Gebrauchsunterschiede zwischen einzelnen Textsorten (diaphasisch), historischen Epochen (diachronisch), Dialekten und Gesellschaftsschichten (diastratisch). Konkretere und für diese

Arbeit relevante Faktoren liefert Goldhahn (2010) in ihrem Beitrag zum Design wissenschaftssprachlicher Korpora für kontrastive Analysen. Wichtig seien die Wahl der Disziplinen, Sprachen, Textsorte(n), Autoren, Größe, Zeit, sowie des Inhalts und Mediums. Es ist fraglich, inwiefern solche Auflistung den Spezifika der Erstellung von gerade den wissenschaftssprachlichen Korpora Rechnung trägt denn die Kriterien würden bei anderen Arten von Korpora höchstwahrscheinlich identisch aussehen. Die einzigen Ausnahmen, die für diese Arbeit nützlich sind und zugleich von dem Allgemeinen abheben, sind Disziplin und Schreiberfahrung. Goldhahn spricht in diesem Zusammenhang von einer „nahezu unüberschaubare[n] Vielfalt an Disziplinen, die eine weitere Einschränkung nötig machen“ (S. 230). Zweitens will sie für ihr eigenes (Deutsch-Tschechisch) Korpus „Autoren auswählen, die durch zahlreiche Publikationen (mind. 20) schon eine gewissen Routine und Stilsicherheit im wissenschaftlichen Schreiben entwickelt haben“ (S. 232).

Die zwei Kriterien werden hier folgendermaßen reflektiert: das Erste umgeht man zum Teil, indem die *allgemeine* Wissenschaftssprache zum Forschungsgegenstand wird, denn sie soll theoretisch disziplinübergreifend fungieren. Gibt man sich jedoch mit dieser Behauptung zufrieden, beraubt er sich der Möglichkeit, das Verhalten einiger AWS-Phänomene zwischen Disziplinen zu vergleichen. Zum Beispiel der Gebrauch von Zahlen, Maßeinheiten und dadurch die Art und Weise, wie man über statistische Angaben und Analysen berichtet, dürfte sich von Fachgebiet zu Fachgebiet unterscheiden, obwohl Zahlen und Maßeinheiten in vielen zu finden sind. Was die Stilsicherheit/Schreiberfahrung anbelangt, kommt es darauf an, ob man hinsichtlich der Stichprobe Homogenität oder Repräsentativität bevorzugt. Im Korpus von Arbeiten erfahrener Wissenschaftler erhält man zwar womöglich eine gewisse Homogenität, es stellt sich darauf immerhin die Frage, inwiefern von einer Wissenschaftssprache zu reden ist, wenn ein erheblicher Anteil der so schreibenden Population ausgeklammert wird. Umgekehrt müsste man sich zugunsten der Repräsentativität wiederum mit der Aussagekraft von einer mehr oder weniger heterogenen Population beschäftigen.

3.2 DESTWIKO (v0.2)

3.2.1 Vorüberlegungen zum Design

An die vorgenannten, eher allgemeingültigen Aspekte des Korpusdesigns knüpfe ich hier mit ihrer Bedeutung für das Korpus der deutschen studentischen Wissenschaftssprache kontrastiv¹⁰ (DESTWIKO) sowie einer Charakterisierung der aktuellen Version des Korpus an.

Es stehen bereits Korpora zur Verfügung, die der Erforschung von Lernaltersprache wie Falko (Reznicek, et al., 2012) oder L2-Wissenschaftssprache wie KoLaS (Andresen & Knorr, 2017) bzw. GeWiss (Gräfe, et al., 2013) dienen. Ihre Datenbasis stellen vorwiegend Studierende dar, die sich in Deutschland typischerweise für eine begrenzte Zeit aufhalten (Erasmus, Studium u. Ä.). Dies wirkt sich auf die Heterogenität der Daten aus, und hat überdies Limitationen bezüglich der Korpusgröße zur Folge. Textsammlungen, die zwecks einer der größeren Untersuchungen (siehe Forschungsstand) erhoben wurden, sind indessen geringfügig größer und deutlich homogener, in der Regel aber schwer zugänglich und oft nicht auf Fremdsprachen

¹⁰ Im Gegensatz zum üblichen Gebrauch des Wortes wird hier unter *kontrastiv* nicht „zwei (bzw. mehr) Sprachen vergleichend“ verstanden. Eher wird es im Rahmen des Spracherwerbs als „mögliche Kontraste zwischen L1- und L2-Sprachbenutzer (Mutter- und Fremdsprache) erfassend“ (Lüdeling & Walter, 2010, S. 319) aufgefasst.

orientiert. Aus dem Bedürfnis, einige dieser Defizite¹¹ im Hinblick auf die Ziele auszugleichen, resultierte 2017 die Erstellung der ersten Version von DESTWIKO.

Bei diesem Korpus war (und ist) es das Ziel, ein grobes Bild von der deutschen studentischen geschriebenen Wissenschaftssprache liefern zu können, die von entweder tschechischen Nichtmuttersprachlern oder deutschen/österreichischen Muttersprachlern produziert wird. Aufgrund ihrer Länge und Zugänglichkeit wurden elektronische Versionen von Abschlussarbeiten als Medium und Textsorte gewählt. In einer Abschlussarbeit bietet sich oft die Gelegenheit dar, eine Funktion (Ziele vorstellen, Argumente kontrastieren, eine Stellung nehmen usw.) mehrmals in Anspruch zu nehmen. Dies ist bei einem kürzeren Aufsatz nicht immer der Fall. Daraufhin liegt automatisch ein Text vor, der auf den ersten Blick wenig lexikalische Variation aufweist. Außerdem ist anzunehmen, dass die Motivation beim Schreiben einer Abschlussarbeit generell höher liegt und die Texte demgemäß eine relativ zuverlässige Einschätzung der Schreibfertigkeiten bieten. Aus diesen Gründen kennzeichnet sich die Textsorte m. E. durch einen hohen Grad an Repräsentativität für die studentische Wissenschaftssprache.¹² Die aktuelle Version des Korpus umfasst momentan nur Masterarbeiten, langfristig wird aber auch mit Bachelor- und Doktorarbeiten gerechnet.

Was die Größe der Sammlung angeht, kam es zu einer wesentlichen Erweiterung. Die erste, für die Zwecke der Facharbeit entstandene Version des Korpus setzte sich aus 54 Texten im Umfang von 1,1 Mio. Wortformen zusammen. DESTWIKO (v0.2) beinhaltet dagegen 280 Texte, die ca. 5,6 Mio.¹³ Wörtern entsprechen. Die vorgenommene Vergrößerung der Textbasis lag hauptsächlich darin begründet, dass Analysen von lexikalischen Phänomenen mehr Daten bedürfen. Wie jedoch schon am Ende des Abschnitts 2.2 erwähnt, wird die Erweiterung kaum dazu beitragen, dass Kollokationen Signifikanztesten unterzogen werden können. Dafür hätte sie vermutlich nicht im Umfang von Millionen, sondern mehreren Dutzend Millionen stattfinden müssen.

Abschließend wenden wir uns nun den von Goldhahn erwähnten Bereichen (siehe oben) des Korpusdesign, die noch nicht angesprochen worden sind, d. h. Autoren, Zeitraum und Sprachvarietät. Die Repräsentativität von Idiolekten ist durch die Gestaltungsweise der Sammlung sichergestellt, nämlich je eine Arbeit pro Verfasser. Der Zeitraum, in dem die Masterarbeiten eingereicht wurden, erstreckt sich über vierzehn Jahre und reicht von 2006 bis 2019. Mit der Ausnahme von 2019 sind im Korpus für jedes Jahr immer wenigstens 10 Arbeiten zu finden. Die genaue Verteilung ist im Histogramm unten (Abb. 5) angezeigt. Diatopisch gesehen stammen die Daten aus zwei digitalen Universitätsarchiven, erstens an der Masaryk Universität (MU) in Brünn und zweitens an der Universität Wien (UW). Es ist diesbezüglich unwahrscheinlich, dass die (Brünner) nichtmuttersprachliche Stichprobe lediglich durch die Tatsache beeinflusst werden könnte, dass die Arbeiten genau an dieser Universität angefertigt und eingereicht worden sind, weswegen anzunehmen ist, dass keine signifikanten Unterschiede z. B. zu Arbeiten aus der Karls-Universität zu verzeichnen wären. Andererseits muss in Kauf genommen werden, dass in den Dokumenten aus dem Wiener Archiv vorwiegend österreichisches Standarddeutsch repräsentiert ist.¹⁴

¹¹ Im Kontext der Zielsetzung dieser Arbeit.

¹² Obwohl sie selbstverständlich nur eine der Textsorten in der (studentischen) Wissenschaftssprache darstellt.

¹³ Einschließlich der ersten Version.

¹⁴ Es sei dahingestellt, inwiefern sich bundesdeutsches und österreichisches Deutsch in der Wissenschaftssprache unterscheiden.

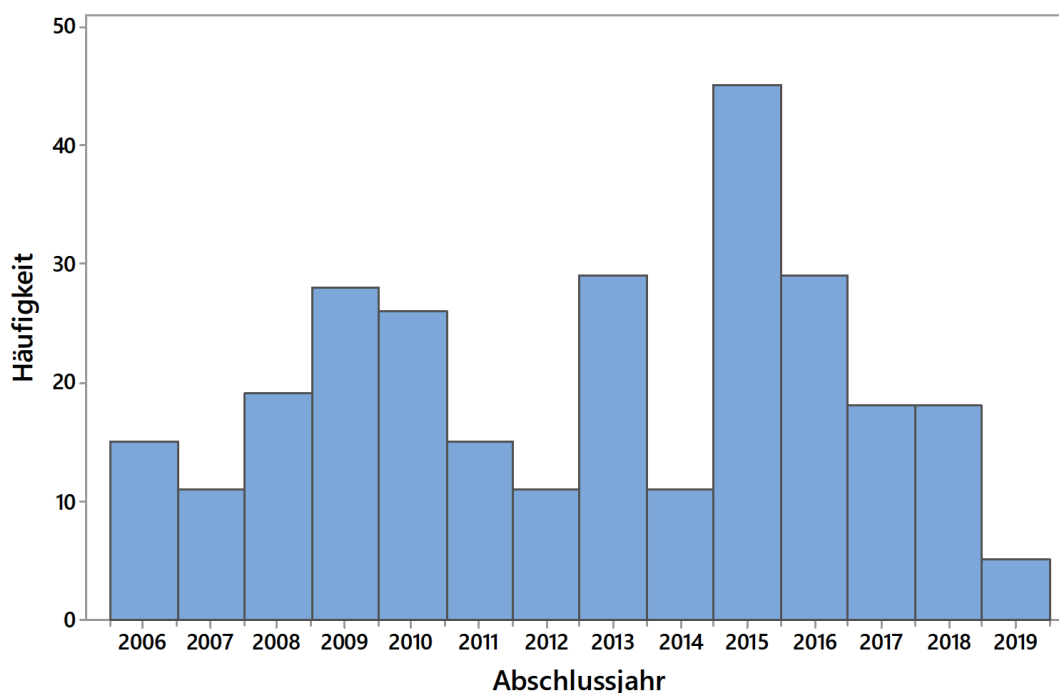


Abbildung 5: Histogramm zur Verteilung der Arbeiten über die Abschlussjahre

Das wichtigste Kriterium für die Ergebnisinterpretation ist letztlich die Wahl der Disziplinen bzw. des Inhalts. An dieser Stelle ist es hervorzuheben, dass DESTWIKO unabhängig von Version ein monodisziplinäres Korpus ist, denn es umfasst nur germanistische Werke. Gleichzeitig wurde die Abgrenzung in Bezug auf die Teildisziplinen¹⁵ beibehalten und auch in DESTWIKO (v0.2) wird zwischen sprachwissenschaftlich und literaturwissenschaftlich orientierten Arbeiten unterschieden.¹⁶ Was auf der anderen Seite undokumentiert bleibt sind einzelne Teilbereiche (z. B. Syntax, Textlinguistik, Phraseologie usw.), die sich aus den Forschungsinteressen der Betreuer bzw. Institute, an denen sie tätig sind, erschließen lassen. Da die Wiener Masterarbeiten mehreren Instituten entstammen, sind die Themen in dieser Gruppe relativ breitgefächert, zumindest verglichen mit der Varianz in dem zwangsläufig thematisch homogeneren Brunn.

3.2.2 Erhebungsprozess und Teknikalien¹⁷

Die Texte wurden aus den zwei Archiven in zwei Dateiformaten (doc[x], PDF) heruntergeladen. Die überwiegende Mehrheit des MU-Archivs waren doc(x)-Dateien während alle Dokumente im UW-Archiv nur als PDFs verfügbar waren. Da an einer PDF-Datei keine Änderungen vorgenommen werden können, mussten all diese Dokumente mithilfe von Microsoft Word ins doc(x)-Format umgewandelt werden. In DESTWIKO (v0.1) wurde zu demselben Zweck das iLovePDF-Online-Tool benutzt, weswegen sich die nachfolgende Datenvorbereitung als einigermaßen anspruchsvoller erwies. Mit MS Word hat sich die Prozedur etwas vereinheitlicht, indem der Einfluss der ursprünglichen Formatierung auf die

¹⁵ Diese werden nachfolgend nur *Disziplinen* genannt.

¹⁶ Didaktisch orientierte Arbeiten enthält das Korpus nach wie vor noch nicht.

¹⁷ Der Prozess ist für beide Versionen des Korpus identisch. In DESTWIKO (v0.1) wurden dennoch SNVs manuell annotiert.

neue Word-Datei verringert wurde. Infolgedessen sind die Änderungsschritte in den Dokumenten leichter zu dokumentieren.

Nach der Konversion wurden jeweils die ersten Seiten einer Arbeit gelöscht, nämlich das Titelblatt und das Inhaltsverzeichnis. Jedes Dokument beginnt also mit einer als fließender Text gestalteten einleitenden/einführenden Sektion. Weiter gelöscht wurden die am Ende der Arbeiten befindlichen Literatur- und Quellenverzeichnisse, Lebensläufe (im Falle von UW) und, falls vorhanden, auch Anhänge in jeglicher Form. Entfernt wurden darüber hinaus alle Fußnoten.¹⁸ Es spielte dabei keine Rolle, ob diese meistens als beiläufige Anmerkungen oder Quellenangabe dienten. Dasselbe galt für alle Bilder, Graphen und Tabellen. Diese Vorgehensweise spiegelt sich in der übergeordneten Zielsetzung wider, die übernommene Sprache, sei es nun Sekundärquellen oder konkrete Beispiele aus den Primärtexten, zu eliminieren. Auf dieser Grundlage wurde beschlossen, dass auch längere (mehr als drei Zeilen) Blockzitate sowie Textbeispiele ebenfalls zu beseitigen sind. Dies war auch der Fall für „Anhäufungen“ von kürzeren Beispielen, die unmittelbar oder fast unmittelbar (typischerweise mit einem oder zwei Sätzen inzwisch) aufeinanderfolgen und deren Länge insgesamt die drei Zeilen überschreitet. Viele von den Zitaten wurden nicht ganz gelöscht, sondern beträchtlich verkürzt; im Falle von Poesie normalerweise zu einer Strophe, ansonsten zu drei Zeilen. Die Verkürzung erklärt sich dadurch, dass es mitunter von Nutzen schien, die textuelle Integrierung von Zitaten zumindest streckenweise beizubehalten. Aus der Tatsache, dass Zitate zwar bei Weitem nicht willkürlich, jedoch auch nicht ganz deckungsgleich gehandhabt wurden, gehen folgende zwei Bezugspunkte hervor. Erstens berufe ich mich in dieser Hinsicht auf das Bestreben nach gewisser „Stummschaltung der nicht-Verfasserstimme“ als oberstes Ziel, denn es ist letzten Endes eben die Sprache der Verfasser, die hier analysiert wird. Zweitens ist darauf hinzuweisen, dass gerade deswegen das Korpus gar nicht zu Analysen geeignet ist, die sich auf den Umgang mit Zitaten und der übernommenen Sprache fokussieren. Wollte man dies dennoch tun und zugleich zwischen den Stimmen differenzieren können, hätten alle Zitate wohl schon bei der Erstellung aufgefunden und getaggt werden müssen.¹⁹

Nachdem die Daten entsprechend vorbereitet worden waren, wurde jede Datei mithilfe eines Tags mit Metadaten markiert, die der Tabelle 2 zu entnehmen sind. Dieser Schritt ermöglichte eine automatische Zuweisung zu den angezeigten Kategorien, die sich in der Software leicht abrufen und durchsuchen lassen. Danach wurden die Dateien in den Online-Korpusmanager Sketch Engine (Kilgarriff, et al., 2014) hochgeladen. Vor der endgültigen Kompilierung eines Korpus wurden die voreingestellten Optionen der Software angenommen und das Korpus wurde lemmatisiert und mit dem deutschen RFTagger PoS-getaggt.

¹⁸ So vermeidet man die Unterbrechung des Textzusammenhangs durch falsch integrierte Fußnoten, die als eine der Annotationsschwierigkeiten von Brommer (2018) dokumentiert wird.

¹⁹ Persönlich fand ich solche Prozedur zu anspruchsvoll und aufgrund der Vielfalt von möglichen Formaten der Zitationen ohne die notwendigen computerlinguistischen Kenntnisse nicht durchführbar. Daneben ist die Leistungsfähigkeit des Korpus nicht verringert, denn Zitationen, wenn nicht absichtlich studiert, sind für die meisten Elemente der Wissenschaftssprache von relativ geringer Bedeutung. Es handelt sich praktisch um Bezug nehmende (einleitende und zusammenfassende) Formulierungen, die nichtsdestoweniger im Korpus enthalten sind, nur deren Kontext fehlt.

Universität	Geschlecht	Studiengang	Abschlussjahr	Disziplin	Muttersprache	Note – gesamt ^b			
						Betreuer	Opponent	Betreuer	Opponent
Universität Wien	männlich	Bachelor ^a	2006	Linguistik	Deutsch	A	A	A	A
MU Brno	weiblich	Magister Doktor ^a	2007	Literatur	Tschechisch	B	B	B	B
			2008			C	C	C	C
			2009			D	D	D	D
			2010			E	E	E	E
			2011			F	F	F	F
			2012						
			2013						
			2014						
			2015						
			2016						
			2017						
			2018						
			2019						

^a vorgesehen

^b nur auf L2-Arbeiten angewandt

Tabelle 2: Übersicht der Metadaten in DESTWIKO (v0.2)

Einer der wichtigsten Vorteile des Korpus, der schon in Planungsphasen in den Mittelpunkt der Aufmerksamkeit gestellt wurde, war seine Ausgewogenheit im Hinblick auf die zwei Basen, muttersprachliche und nichtmuttersprachliche Verfasser. Aus diesen zwei Gruppen wurden zwei Subkorpora erstellt, denn wenn sie nur anhand von den Metadaten voneinander getrennt werden, würde dies die Funktionalität einiger Analyseinstrumente limitieren. Zusätzlich gab es noch ein anderes an die Variable Disziplin gebundenes quantitatives Kriterium – die sprach- und literaturwissenschaftlich ausgerichteten Arbeiten sollen im Verhältnis zueinander möglichst gleichmäßig verteilt sein, es sei denn diese Verteilung beeinträchtigt die Ausgewohnheit der L1- und L2-Stichprobe. Es sollen überdies in jedem Quadranten (Sprache x Disziplin) mindestens 50 Dokumente sein, deren Anzahl über die Disziplinen in der gegebenen Mutterprache hinweg gleich sein sollte. Eine detaillierte Übersicht der absoluten und prozentuellen Verhältnisse zwischen den Kategorien findet sich in der Tabelle 3.

		L1	L2	gesamt
		(n=100)	(n=180)	
Tokens	Linguistik (n=140)	1 574 039	1 625 104	3 199 143
	Literatur (n=140)	1 835 962	1 812 052	3 648 014
	gesamt	3 410 001	3 437 156	6 847 157
	Textlänge (<i>M</i>)	27 956	15 815	
	(<i>SD</i>)	8 890	4712	
prozentuelle Anteile	Linguistik (n=140)	23 %	23,7 %	46,7 %
	Literatur (n=140)	26,8 %	26,5 %	53,3 %
	gesamt	49,8 %	50,2 %	100 %

Tabelle 3: Angaben über die Verteilung der Subkorpora

Die fast doppelte Anzahl von L2-Arbeiten und die Mittelwerte deutet auf gravierende Unterschiede angesichts der Textlänge hin.²⁰ Dasselbe verdeutlichen auch die Box-Plots unten. Bemerkenswert sind dabei die Ausreißer in der L2-Gruppe. Eine Arbeit, die wegen ihrer Länge besonders ins Auge fällt, umfasst nach dem oben beschriebenen Vorbereitungsverfahren insgesamt 3728 Wörter. Grundsätzlich betrifft solche Senkung der Wortanzahl hauptsächlich die Arbeiten aus dem Bereich Phraseologie. Vor allem im explikativen Teil dieser Arbeiten geben die Verfasser viele Beispiele an, die anschließend kategorisiert werden, wodurch die oben angedeuteten Anhäufungen entstehen. Obwohl ihr Gebrauch an sich gewiss nicht zu kritisieren ist, wird es dennoch problematisch, wenn man auf sie als primäre Textgestaltungsstrategie zurückgreift, denn eine Abschlussarbeit bedarf komplexer Ideen, die sich in der Regel nicht mithilfe bloßer Auflistungen, sondern eher mit Fließtext effektiv übermitteln lassen. In vielen Fällen mangelt es zudem an Kommentaren, die die angeführten Beispiele bzw. Kategorisierungen in die Textmenge integrieren würden. In Anlehnung an das Verarbeitungsverfahren oben hatte die Verwendung solcher Anhäufungen zur Folge, dass maßgebliche Anteile so strukturierter Texte gelöscht wurden.

²⁰ Diese Unterschiede erwiesen sich als signifikant, Wilcoxon-Rangsummentest ($W = 21\,216,5; p < 0,001$).

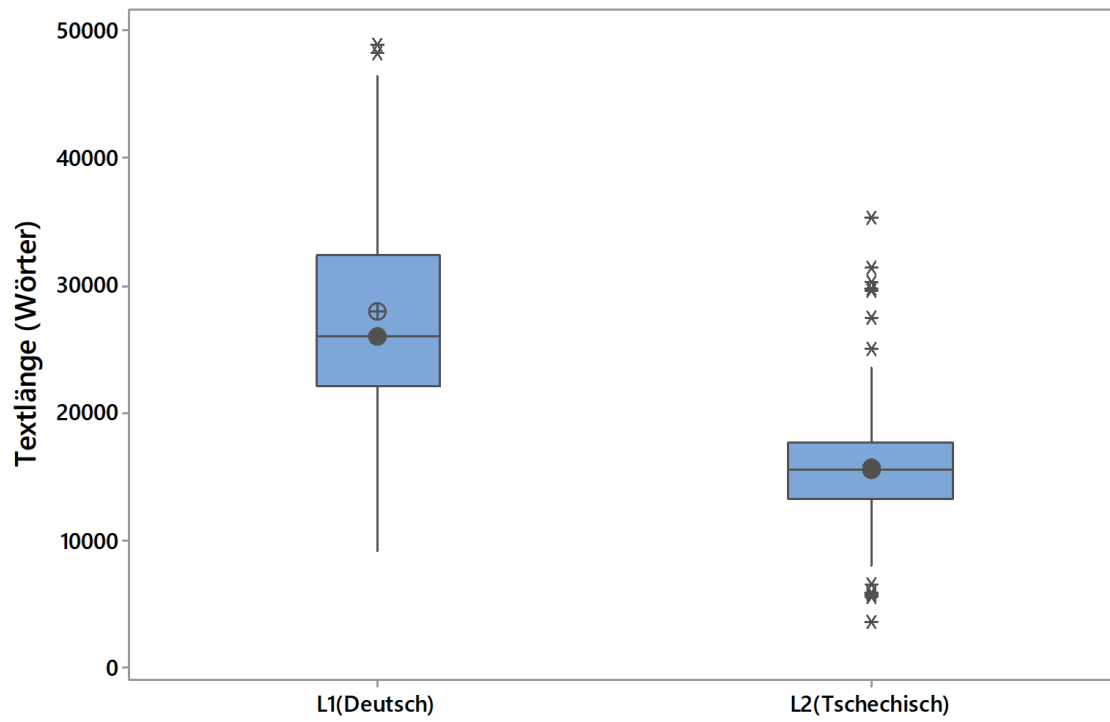


Abbildung 6: Box-Plot-Repräsentation der Länge von L1- und L2-Arbeiten

4. Methodische Vorgehensweise

Schon zu Beginn dieser Studie kündigte ich das Ziel an, das kollokationelle Umfeld von ausgewählten nominalen Verweisen zu analysieren. Der Hauptbeweggrund für die Analyse war das Potenzial ihrer eventuellen didaktischen Anwendung. Das Ziel solcher didaktisch-gesteuerten Herangehensweise hier ist es, plausible Verwendungsbeispiele zu zeigen statt unbedingt Unterschiede/Tendenzen zu berechnen.²¹ Diese Zielsetzung lässt sich gut mit den Limitationen hinsichtlich des zu analysierenden Materials vereinbaren, die ich im Abschnitt 3.2.1 vorbrachte, als ich auf die Korpusgröße zu sprechen kam.

Aus dem bisher Gesagten schälen sich ansatzweise die Umrisse der Endprodukte heraus. Diese werden die Form von lexikographischen Wortprofilen annehmen und werden anhand von selektierten Beispielsätzen aufgebaut. Im Folgenden werden sie näher erläutert. Nach den theoretisch-methodischen Ausführungen knüpfe ich mit der Beschreibung des Zustandekommens der Profile an, wobei vor allem die zur Datengewinnung und Datenvisualisierung angesetzten Instrumente berücksichtigt werden. Anschließend werden konkrete Aspekte und Sonderfälle bei der Erstellung von Wortprofilen thematisiert.

4.1 Prozess

Eingangs sind 15 Abstrakta aus der Liste in meiner Facharbeit ausgewählt worden. Die Auswahl hatte sich mit der Häufigkeit als Hauptkriterium vollzogen, indem 50 SNVs aus der Frequenzliste aus DESTWIKO (v0.1) herausgenommen wurden, und ihre Häufigkeiten diesmal in DESTWIKO (v0.2) berechnet wurden. Daraus entfernte ich das Substantiv *Gefühl*, dessen Kollokationen aufgrund ihrer Deskriptivität und Kontext- bzw. Themagebundenheit der AWS meistens nicht angehörten. Aus dem Rest sind die häufigsten 15 Lexeme zum Thema dieser Magisterarbeit geworden. Diesen Bedingungen, nämlich Häufigkeit und Zugehörigkeit zur AWS, unterliegen auch die Abstrakta in einer Studie von Tomášková (2016), die ihre Auswahl ebenfalls dadurch rechtfertigt, „dass es sich um zentrale Wörter in der Wissenschafts-sprache handelt, die auch bei Graefen (2011) im Lehrbuch zur Wissenschafts-sprache ... häufig vorkommen“ (S. 132).

Nach der Festlegung der zu untersuchenden Substantive wurden mithilfe der Word-sketch-Funktion die Kollokationsprofile ermittelt, die anschließend visualisiert wurden (mehr dazu in 4.3). Es sind jeweils zwei Visualisierungen (L1 und L2) pro Substantiv generiert worden, damit sich der Leser einen schnellen Überblick über eventuelle Unterschiede und Ähnlichkeiten zwischen den beiden Subkorpora verschaffen kann. Im Gegensatz dazu sind die Kollokationen enthaltenden Beispielsätze aufgrund ihres vergleichsweise höheren Musterhaftigkeitsgrad ausschließlich dem muttersprachlichen Subkorpus entnommen worden. Dass auch von muttersprachlichen Studierenden verfasste Texte in gewissem Maße unterschiedlich von Fachaufsätzen und Monografien erfahrener Wissenschaftler sind, wird hier beiseitegelassen, denn dies würde das Heranziehen eines entsprechenden Vergleichskorpus erfordern. Schließlich geht es ja nicht um eine Nachahmung des professionellen Wissenschaftsstils, sondern um eine Präsentation wohlgeformter Beispielsätze mit Kollokationen.

²¹ Es sei dazu angemerkt, dass die Verwendung von Datensammlungen mit Werken unerfahrener Verfasser für didaktische Zwecke zwar mit einigen Schwierigkeiten behaftet ist, zugleich ist sie aber keineswegs präzedenzlos (vgl. Römer, 2012).

Da es bei der automatischen Identifikation von Kollokationskandidaten oft zu Ungenauigkeiten kommt²², vor allem bei kleineren Korpora, unterscheiden sich hin und wieder die in den erstellten Diagrammen abgebildeten Kookkurrenzen von denjenigen, die sich in den Beispielsätzen vorfinden. Die Notwendigkeit manueller Überprüfung von Kollokationen unterstreicht in ihrer Arbeit sowohl Tomášková (S. 131) als auch Steyer (2008): „[u]m zu den konstitutiven Bildungsmustern zu gelangen, muss man iterativ vorgehen: Ermittelte rekurrente Syntagmen werden wiederum daraufhin analysiert, in welchen Umgebungsmustern sie eingebettet sind“ (S. 198). Zum Schluss wurde, wie bereits mehrmals angedeutet, zu jedem Kollokationspaar ein Beispielsatz gefunden (siehe 4.3 für die Auswahlkriterien).

Da uns zwei Subkorpora zur Verfügung stehen, bietet es sich an, Unterschiede und Ähnlichkeiten dazwischen entdecken zu wollen. Vergleiche von Relativfrequenzen häufig auftretender Wortkombinationen, sprich auf der Wortebene statt (Sub)korpusebene, anhand von einem Referenzkorpus aufzustellen, wie es z. B. Tomášková (2016) tut, halte ich allerdings für methodologisch unüberlegt, denn dies kann insbesondere im Falle einer lexikalisch-motivierten Untersuchung von Kleinkorpora²³ Probleme mit der Streuung von Daten bereiten.²⁴ Konsequenterweise könnte man vermutlich über die idiosynkratischen Schreibgewohnheiten der einzelnen Autoren sprechen, nicht jedoch verallgemeinernd über die Wissenschaftssprache als Ganzes, wie eingeschränkt das Konzept in der jeweiligen Studie auch immer sein mag. Zudem muss man in solchen Fällen eine der Limitationen des korpusbasierten Ansatzes im Hinterkopf behalten, und zwar, dass Korpora keine Beweise für das Fehlen liefern können (Hunston, 2002, S. 22) – die Extrapolation von mangelnden Daten mittels Relativhäufigkeiten ist somit ebenfalls als unzulänglich zu betrachten.

Vor diesem Hintergrund ist festzuhalten, dass die Anzahl der einzelnen Kollokate in DESTWIKO (v0.2) zu einem L1-L2-Vergleich auf der Lexemebene erwartungsgemäß auch aus methodischen Gründen bei Weitem nicht ausreicht. Daraus folgt, dass die einzigen in den Profilen angeführten quantitativen Vergleiche nur die Abstrakta selbst ins Visier nehmen. Konkret wird mit einem Chi-Quadrat-Test ($\alpha = 0,05$) auf Unabhängigkeit von der Muttersprache und der Disziplin getestet. In anderen Worten wird danach gefragt, ob sich die Anzahl jeweils eines der 15 gewählten Abstrakta in den sprachwissenschaftlichen und den literarwissenschaftlichen Arbeiten in Bezug auf die Muttersprache signifikant unterscheidet. Die Testergebnisse zusammen mit der Gruppenverteilung sind dem Wortprofil jeweils in einer Tabelle beigelegt. Neben diesem Test und den Häufigkeitsangaben ist auch der Index der Dispersion eingeschlossen worden, mit der ich mich im nächsten Abschnitt auseinandersetze.

4.2 Exkurs: Dispersion

Wie die Übersichtstabelle 4 veranschaulicht, sind die Gesamthäufigkeiten im L1- und L2-Subkorpus in vielen Fällen sehr ähnlich. Dies kann zwar zu der Annahme verleiten, dass die Subkorpora in dieser Hinsicht gleichartig sind, aber die Häufigkeiten an sich lassen diesen Schluss nicht zu. Es ist gerade in dieser Situation, wo die Nützlichkeit, oder sogar

²² Vgl. jedoch Kilgariff, Kovář, Krek, Srdanović, & Tiberius (2010) für die Genauigkeitsauswertung der benutzten Funktion Word sketch.

²³ Die Korpusgröße in Tomáškovás (2016) Studie beläuft sich auf 844 767 Wörtern aus insgesamt 13 Dokumenten.

²⁴ Siehe z. B. die Grafik auf S. 131 (Tomášková, 2016), wo mindestens 3 von 7 Belegen der Kollokation *einen Versuch unternehmen* in demselben Dokument zu finden sind.

Unentbehrlichkeit, von Dispersion zutage tritt; so z. B. Gries (im Druck) in seinem Plädoyer für das Konzept:

just like trying to summarize the distribution of any numeric variable using only a mean can be treacherous (especially when the numeric variable is not normally distributed), so is trying to summarize the overall 'behavior' (or the co-occurrence preferences or the keyness) of a word x on the basis of just its frequency/frequencies because ... words with identical frequencies can exhibit very different distributional behaviors. (S. 2)

Es gibt viele Dispersionsmaße, wovon eines der meist benutzten das Juillands D ist. Allerdings haben Biber, Reppen, Schnur, & Ghanem (2016) in ihrer Studie demonstriert, worauf schon früher hingedeutet wurde, und zwar, dass bei einem vierteiligen Korpus Juillands D falsch-positive, d. h. von einer uniformen Verteilung zeugende, Ergebnisse produzieren kann. Empfohlen wird stattdessen von Gries (2008) eingeführtes Maß – DP (Deviation of Proportions), bzw. DP_{norm} ²⁵. Außer seiner Robustheit im Hinblick auf die Teileanzahl, bringt das Maß weitere Vorteile mit sich, nämlich die relative Einfachheit seiner Berechnung, und eine klare, intuitive Spannweite (von 0 bis 1 für DP_{norm}).

Für ein S -Wörter großes Korpus mit n Teilen, deren Größen (Wörteranzahl) s_{1-n} betragen, und die gesuchte Erscheinung mit der Gesamtfrequenz von F und Frequenzen f_{1-n} in den jeweiligen Korpusteilen sieht die Formel folgendermaßen aus:

$$DP = \frac{\sum \left| \left(\frac{s_{1-n}}{S} - \frac{f_{1-n}}{F} \right) \right|}{2}$$

Es handelt sich eigentlich um die Summe der Differenzen zwischen erwarteten und beobachteten Anteilen geteilt durch zwei. Dazu gibt es eine Normalisierung, die in Anbetracht der Extremwerte etwas genauer ausfällt:

$$DP_{norm} = \frac{DP}{1 - \left(\frac{s_x}{S} \right)_{min}}$$

Es ist jedoch darauf hinzuweisen, dass ihre Bedeutung mit zunehmender Anzahl an Korpusteilen sinkt und keine entscheidenden Unterschiede dementsprechend erwartet werden können.

²⁵ Die Berechnung der normalisierten Variante von DP wurde später von Lijffijt & Gries (2012) korrigiert und unterscheidet sich daher von der ursprünglichen Formel in Gries (2008). Mit der korrigierten Version von DP_{norm} der Formel wird auch hier gerechnet.

	<i>N</i>			Mittelwert ^a			<i>SD</i>			Median ^a		
	Gesamt	L1	L2	Gesamt	L1	L2	Gesamt	L1	L2	Gesamt	L1	L2
<i>Fall</i>	4675	2276	2399	16,75	22,88	13,35	16,38	17,32	14,81	13	17,5	9
<i>Frage</i>	4290	2440	1850	15,33	24,41	10,28	15,67	20,66	8,67	11	20	8
<i>Grund</i>	2982	1603	1379	10,66	16,06	7,64	9,18	11,63	5,60	8	13,5	7
<i>Unterschied</i>	2508	1284	1224	9,00	12,88	6,85	9,39	12,27	6,42	6	10,5	5
<i>Möglichkeit</i>	1979	1015	964	7,07	10,16	5,36	6,30	7,74	4,52	5	8	4
<i>Problem</i>	1810	798	1012	6,50	8,07	5,62	7,45	8,98	6,31	5	5	4
<i>Ergebnis</i>	1748	1122	626	6,30	11,39	3,48	12,00	17,15	6,25	2	5	2
<i>Ziel</i>	1720	866	854	6,15	8,67	4,74	6,65	8,91	4,41	4	6,5	3,5
<i>Meinung</i>	1677	564	1113	6,00	5,67	6,18	7,14	7,01	7,23	4	4	4
<i>Aussage</i>	1439	794	645	5,15	7,96	3,58	6,87	8,91	4,76	3	5	2
<i>Tatsache</i>	1361	582	779	4,94	5,84	4,44	5,32	5,57	5,12	3	5	2,5
<i>Vorstellung</i>	1084	537	547	3,87	5,37	3,04	5,22	6,79	3,87	2	4	2
<i>Aufgabe</i>	1048	443	605	3,74	4,43	3,36	4,71	4,57	4,76	2	3	2
<i>Tendenz</i>	868	409	459	3,11	4,13	2,55	4,94	6,35	3,85	1	2	1
<i>Versuch</i>	815	464	351	2,91	4,64	1,95	4,18	6,03	2,12	2	3	1

^a per Arbeit

Tabelle 4: Verteilung der 15 Abstrakta im L1- und L2-Subkorpus

Neben dem Wunsch, den Ergebnisinterpretationen mehr Aussagekraft zu verleihen, lag dem Einsatz von Dispersionsmaßen auch die Hypothese zugrunde, es mögen – trotz vergleichbarer Gesamtergebnisse pro Subkorpus – aufgrund der mutmaßlich höheren Variabilität in Sprachniveaus und Schreibkompetenzen zwischen einzelnen L2-Studierenden doch gewisse Unterschiede vorliegen. Kurzum wurde angenommen, dass obwohl die L2-Studierenden mit den Muttersprachlichen in absoluten Zahlen zwar gleichziehen können, weisen sie zugleich als Gruppe eine weniger homogene Verteilung auf. Nun zur Berechnung von DP_{norm} in unserem Korpus. Der DP -Wert jedes der 15 untersuchten Substantive wurde für das Korpus als Ganzes und für die beiden Subkorpora (L1, L2) ermittelt (siehe Abb. 7).

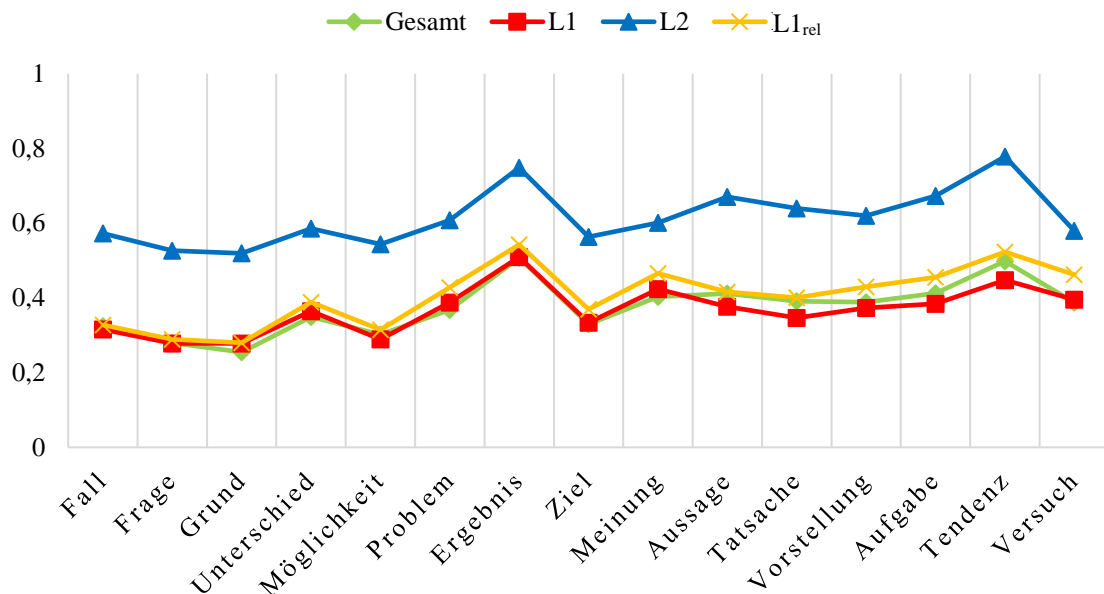


Abbildung 7: Darstellung der DP-Werte

Die Werte wurden mittels eines gepaarten t-Tests verglichen. Wie die Graphik schon nahelegt, zeigen die Testergebnisse einen signifikanten Unterschied zwischen den zwei Stichproben: $t(15) = 23,28$; $p < 0,001$; 95 % CI [0,2253; 0,2710]. Man dürfte allerdings einwenden, dass die Differenz schlichtweg als Konsequenz der unterschiedlichen Textlänge (siehe 3.2.2) zu deuten ist – weniger Text könnte sehr wohl weniger Material bedeuten. Obwohl DP/DP_{norm} relativ sind, kann die beschriebene Eventualität trotzdem eintreten. Vor allem betrifft sie die weniger häufigen Substantive

Um den Einfluss dieser Variable auszufiltern führte ich eine Modifikation von L1 ein, deren DP -Werte in der Graphik als $L1_{rel}$ verzeichnet sind. Sie beruht auf der folgenden Formel, wo a_x für eines der 15 Substantive (a) in einer konkreten Abschlussarbeit (x) steht:

$$f_{rel}(a_x) = \left[\frac{med_{L2}(s_{n1-n180})}{med_{L1}(s_{n1-n100})} \times f(a_x) \right]$$

Man errechnet aus den Medianwerten den Koeffizienten, mit dem die Frequenz eines bestimmten Substantivs in einer bestimmten Arbeit multipliziert wird. Für $a_1 = Fall$ haben wir also

$$f_{rel}(a_x) = \left[\frac{15525,5}{25995,5} \times 63 \right]$$

bzw.

$$f_{rel}(a_x) = \lfloor 0,597237983 \times 63 \rfloor$$

Das Produkt, hier 37,62599296, wird letztendlich auf die nächstniedrigere ganze Zahl abgerundet.

$$f_{rel}(a_x) = 37$$

Der Abrundungsschritt ist ausgesprochen wichtig, denn für Texte, wo ein Substantiv nur einmalig vertreten ist, würde die übliche Rundung (1-4 nach unten, 5-9 nach oben) keine²⁶ Nullwerte ergeben. Da eben diese die Dispersion in L2 beeinflussen mögen, sollen sie auch in L1_{rel} repliziert sein. Durch diesen Prozess wird die ganze Matrix der absoluten Frequenzen in L1 relativ zu den absoluten Frequenzen in L2 proportional verkleinert.²⁷ Auch mit der Textlänge isoliert, erwiesen sich die Unterschiede zwischen L2 und L1_{norm} als signifikant; $t(15) = 19,55$; $p < 0,001$; 95 % CI [0,1860; 0,2319]. Wir können folglich die Eingangshypothese für grundsätzlich bestätigt erklären. Im Gegensatz zu der L2-, werden in der L1-Teilmenge die 15 Substantive konsistenter benutzt. Die Unterschiede zwischen den L2-Verfassern in dieser Hinsicht sind einer Fülle von möglichen Faktoren zuzuschreiben, darunter z. B. Sprachkenntnisse, eine textsortenbezogene Unsicherheit, oder eine andere Informations- bzw. Textgestaltungsstrategien, die weniger SNVs/Abstrakta/Kohäsionsmittel beanspruchen.²⁸ Alternativ dürften die L1-Verfasser die analysierten Wörter übernutzen. Letzteres sehe ich angesichts der absoluten Häufigkeiten pro Arbeit (Tab. 4) als eher unwahrscheinlich, obwohl nur ein Vergleich mit erfahrenen Verfassern eine definitive Antwort liefern könnte.

4.3 Instrumente

4.3.1 Word-Sketches

Die Funktion Word sketch, der Sketch Engine sein Name verdankt und durch die sich die Software von anderen korpusanalytischen Instrumenten abgrenzt, bringt mit sich einen deutlichen Vorteil für die Kollokationsanalyse. Im Gegensatz zu den gängigen Instrumenten, die zur Ermittlung von Kollokaten alle Wörter im Umfeld von ± 3 bzw. ± 5 Tokens vom Stichwort mithilfe verschiedener statistischer Assoziationsmaße auswerten. In einem Word-Sketch werden dagegen nicht nur statistische Informationen zu den Kookkurrenzen zusammengetragen, sondern auch zu ihrer grammatischen Gebrauchsweise:

Rather than looking at an arbitrary window of text around the headword, we look, in turn, for each grammatical relation that the word participates in. The word sketch then provides one list of collocates for each grammatical relation the word participates in. (Kilgariff, Rychlý, Smrž, & Tugwell, 2004, S. 107)

Die Abbildung 8 zeigt ein Beispiel der Darlegung einer mehrerer grammatischen Beziehungen, der sog. gramrels.

²⁶ Es sei denn natürlich die ursprüngliche Zahl beträgt auch 0.

²⁷ Umgekehrt vorzufahren (L2 vergrößern) würde in der Erfindung von Daten resultieren, was gegen gute Praxis verstößt.

²⁸ In den L2-Arbeiten werden beispielsweise oft (direkt übernommene) Auflistungen verwendet, die beinahe definitionsgemäß kohäsiv sind und demnach keiner anderen kohäsionsleistenden Elemente bedürfen.

verbs with "Dokument" as accusative object			
vorlegen	143	8,34	...
stimmen habe für dieses Dokument gestimmt	141	7,86	...
geben	56	4,44	...
erhalten Dokumente erhalten	52	6,49	...
lesen	48	8,8	...
begrüßen Ich habe dieses Dokument begrüßt	39	6,35	...
erarbeiten Dokument erarbeitet	37	8,04	...
veröffentlichen	35	8,42	...
unterstützen Ich habe dieses Dokument unterstützt , da	35	5,22	...
annehmen Dokument anzunehmen	33	6,67	...

Abbildung 8: Beispiel einer Gramrel aus Sketch Engine

Außer den einzelnen Kollokaten enthalten die Gramrels auch Informationen zur Häufigkeit der Kombination (die zweite Spalte) und Assoziation (die dritte Spalte) in Form von logDice (Rychlý, 2008). Dabei hat der Benutzer auch die Wahl, das sog. „longest-commonest match“ (Kilgarriff, Baisa, Rychlý, & Jakubíček, 2015) unter den Kollokaten angezeigt zu bekommen. Kilgarriff et al. definieren es als

a straightforward extension of [] two-word collocations into multiword expressions [which] ... are also very useful when it is not clear from what texts a collocation pair comes, e.g. <flame_n, object-of, put_v>, <love_v, object, neighbor_n>, etc. The longest-commonest match is therefore also a representative expression for collocational pairs. (S. 397)

Jede grammatische Relation wird durch Corpus Query Language (CQL) in einer Sketch-Grammatik bestimmt. In DESTWIKO (v0.2) wird Jakubíček's vereinfachte Version der deutschen Sketch-Grammatik von Ďurčo & Ďurčo (2012) verwendet. Für Substantive erbringt diese Version sogar 16 Gramrels. Es liegt mithin auf der Hand, dass einige davon relevanter sind als andere. Eine engere Auswahl für diese Arbeit wurde auf Basis des Kollokationswörterbuchs von Buhofer et al. (2014) getroffen. Nämlich wurden insgesamt nur 5 Relationen berücksichtigt:

1. Attribute in Form von Adjektiven/Partizipien
2. Verben mit dem untersuchten Substantiv als Subjekt
3. Verben mit dem untersuchten Substantiv als Akkusativobjekt

4. Verben mit dem untersuchten Substantiv als Dativobjekt²⁹
5. Genitivattribute (vereinzelt)

Zum Abschluss wird die Bestimmung der fünf Relationen wegen Unterschieden zwischen Sketch-Grammatiken der Vollständigkeit halber näher erläutert.

Attribute in Form von Adjektiven/Partizipien

```
*DUAL
=modifiers of "%w"/nouns modified by "%w"
*UNIMAP modifier/modifies
      2:[adj_] [gap_nN]{0,4} 1:"N.*"
```

Jedes Adjektiv gefolgt von 0 bis 4 Tokens, die keine Substantiva sind, mit einem Substantiv gleich danach.

Verben mit dem untersuchten Substantiv als Subjekt

```
*DUAL
=subjects of "%w"/verbs with "%w" as subject
*UNIMAP subject/subject_of
      no_prep 2:"N.*Nom.*" [gap_nV]{0,10} 1:"VFIN.Full.*"
      1:"VFIN.Full.*" [gap_nVAp]{0,10} 2:"N.*Nom.*"
```

Jedes finite Vollverb gefolgt von 0 bis 10 Tokens, die weder Verben noch Adpositionen sind, mit einem Substantiv im Nominativ gleich danach.

Verben mit dem untersuchten Substantiv als Akkusativobjekt

```
*DUAL
=accusative objects of "%w"/verbs with "%w" as accusative
object
*UNIMAP object/object_of
      no_prep 2:"N.*Acc.*" [gap_nV]{0,10} 1:"V.*Full.*"
      1:"V.*Full.*" [gap_nVAp]{0,10} 2:"N.*Acc.*"
```

Jedes finite Vollverb gefolgt von 0 bis 10 Tokens, die weder Verben noch Adpositionen sind, mit einem Substantiv im Akkusativ gleich danach.

Verben mit dem untersuchten Substantiv als Dativobjekt

```
*DUAL
=dative objects of "%w"/verbs with "%w" as dative object
*UNIMAP object/object_of
      no_prep 2:"N.*Dat.*" [gap_nV]{0,10} 1:"V.*Full.*"
      1:"V.*Full.*" [gap_nVAp]{0,10} 2:"N.*Dat.*"
```

²⁹ Im Einklang mit der Gramrel-Klassifizierung in Sketch Engine werden in dieser Arbeit als Dativ- und Akkusativobjekte auch Präpositionalobjekte im entsprechenden Kasus betrachtet (*führen zu, eingehen auf* usw.).

Jedes finite Vollverb gefolgt von 0 bis 10 Tokens, die weder Verben noch Adpositionen sind, mit einem Substantiv im Dativ gleich danach.

Genitivattribute

```
*DUAL
=genitive objects of "%w"/nouns with "%w" as genitive
object
*UNIMAP modifier
  1:"N.*" [gap_nVAp]{0,4} 2:"N.*Gen.*"
  2:"N.*Gen.*" [gap_nVAp]{0,4} 1:"N.*"
```

Jedes Substantiv im Genitiv, dem eine Kette von einem Substantiv gefolgt von 0 bis 4 Tokens, die weder Verben noch Adpositionen sind, vorangestellt ist.

4.3.2 Visualisierungen

Als Komponente der neuen Benutzeroberfläche sind in Sketch Engine verschiedene Visualisierungsmöglichkeiten eingearbeitet worden, durch die sich die Aussagekraft einer Analyse erhöhen kann. In Anlehnung an Chen, Hårdle, & Unwin (2008, S. 4-5, zit. n. Bubenhofer, 2018, S. 36) ist im Wesentlichen zwischen Präsentations- und Explorationsgrafiken zu differenzieren. Die ersteren dienen einer präziseren, zusammenfassenden Darlegung von Ergebnissen und komplementieren somit die textuelle Beschreibung. Explorationsgrafiken verwendet man dagegen, um zu Ergebnissen zu gelangen. Als solche sind sie durch einfache und übersichtliche Gestaltung gekennzeichnet, damit sich Schlussfolgerungen schneller ziehen lassen. Neben dieser Unterscheidung kann man nicht unterlassen zu erwähnen, was diese zwei umgekehrt verbindet; in dieser Hinsicht ist die folgende Anmerkung Bubenhofers (2018) von besonderer Bedeutung: „[m]ethoden der Visualisierung, egal ob für explorative Zwecke oder für die Präsentation, sind nicht einfach ein zusätzliches Element der Analyse, sondern prägen ganz grundsätzlich unsere Perspektive auf die Daten“ (S. 54).

Die ältere Version der Sketch-Engine-Benutzeroberfläche verfügte über fast keine Visualisierungsmöglichkeiten. Würde man sich die Word-Sketches wegdenken, die auch an sich optische Darstellung einer Art ermöglichen, beschränkte sich Datenvisualisierung von Kollokationen auf die am Rande einer Gramrel befindlichen Farbtöne zwischen grün und rot, die die Präferenz für Kookkurrenz markieren. Diese eignen sich besonders gut zu Vergleichen zwischen einzelnen Lexemen (Abb. 9) bzw. Subkorpora (Abb. 10) und sind nach wie vor Teil der Word-Sketch-Differenz-Funktion.

verbs with "Problem/Schwierigkeit" as accusative object			
lösen	2806	21	...
darstellen	505	8	...
angehen	590	7	...
ansprechen	360	7	...
geben	1947	309	...
bewältigen	219	33	...
verursachen	272	52	...
auftreten	98	33	...
bereiten	221	171	...
überwinden	82	105	...
stoßen	40	50	...
geraten raten	0	30	...

Abbildung 9: Beispiel einer Sketch-diff-Gramrel zum Vergleich von Lexemen

modifiers of "Grund"			
genannt	12	0	...
offensichtlich	5	0	...
tatsächlich	6	0	...
möglich	26	8	...
verschieden	24	13	...
finanziell	6	9	...
praktisch	4	8	...
politisch	6	15	...
ökonomisch	0	6	...
rhythmisch	0	6	...
wahrscheinlich	0	14	...
sprachökonomisch	0	10	...

Abbildung 10: Beispiel einer Sketch-diff-Gramrel zum Vergleich von Subkorpora

Für die Zwecke dieser Arbeit kamen die zusammen mit der zweiten Version herausgebrachten Visualisierungen in Frage, die auf der Masterarbeit von Kocincová (2015) basieren.³⁰ Entweder die Farbenskala, die fürs Vergleichen bestimmt ist (Abb. 11), oder der Gramrel-Kreis (Abb. 12).

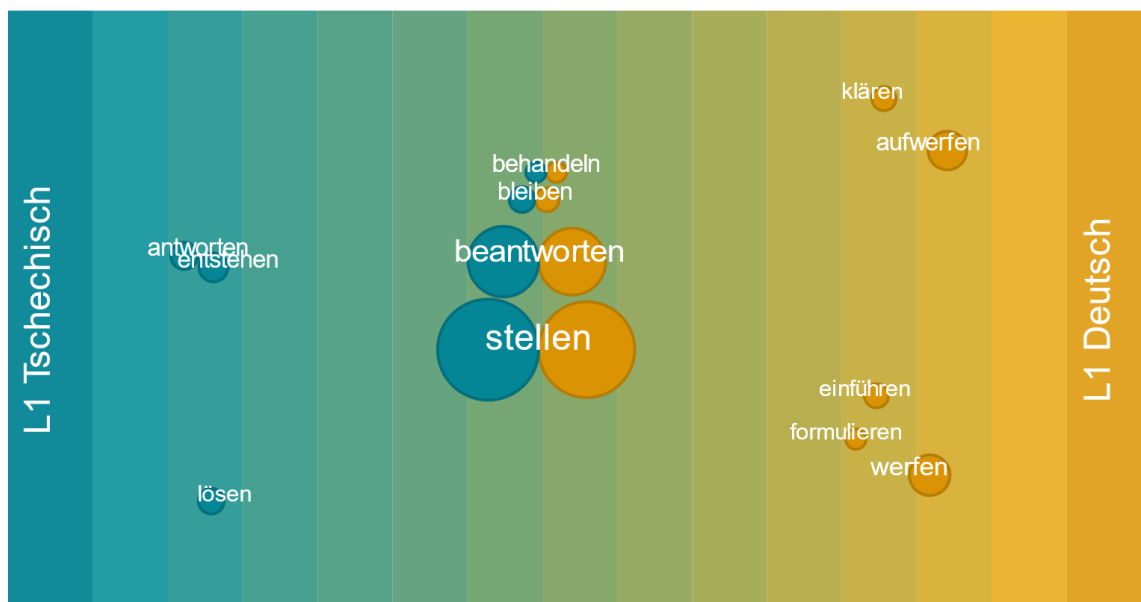


Abbildung 11: Neue, skalaartige Visualisierung einer Sketch-diff-Gramrel

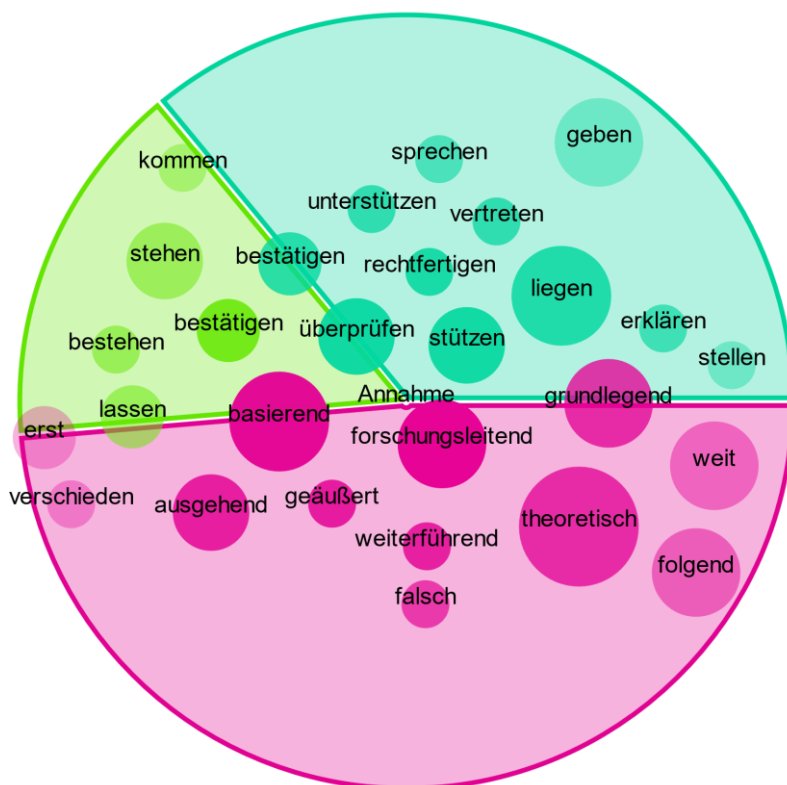


Abbildung 12: Kreisdiagramm mit mehreren Gramrels

³⁰ Siehe Kocincová, Baisa, Jakubíček, & Kovář (2015) für die Details der Implementierung in Sketch Engine.

Da es bei der Zielsetzung der Arbeit nicht primär um Vergleiche geht, wurden die Gramrel-Kreise als zweckmäßiger erachtet. Des Weiteren würde man bei der Untersuchung von drei Relationen in zwei Subkorpora drei skalaartige Graphen benötigen (3 x 1 eine Relation), jedoch nur zwei Gramrel-Kreise (2 x 1 Subkorpus).

Sketch Engine erlaubt es dem Benutzer, viele Parameter der Diagramme individuell anzupassen, weshalb sich ihre Form von Arbeit zu Arbeit unterscheiden kann. Daher werden an dieser Stelle die wichtigsten parametrisierbaren Einstellungen erörtert. Ich folge dabei der Einstellungsliste (Abb. 13) von oben nach unten.

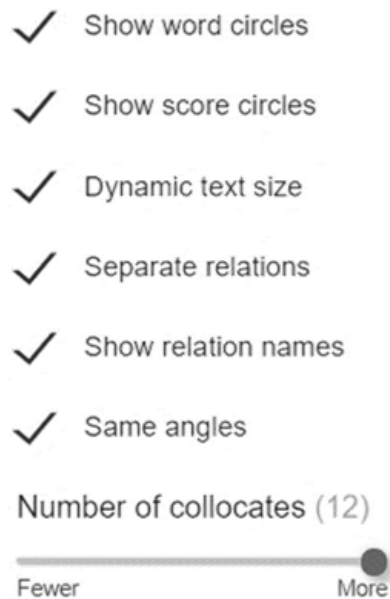


Abbildung 13: Mögliche Einstellungen der Kreisdiagramme

Als erstes besteht die Möglichkeit, Kreise um die einzelnen Kollokate herum anzuzeigen.

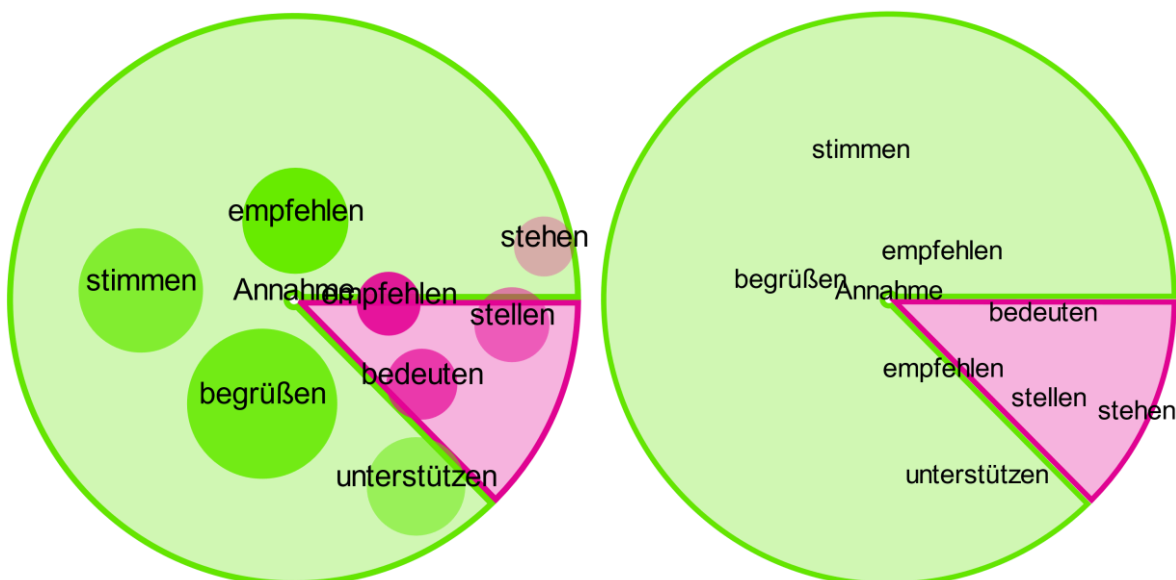


Abbildung 14: Kreisdiagramm mit Wortkreisen und ohne

Diese Option war bei der Erstellung der Wortprofilen aus zwei Gründen eingeschaltet. Erstens kommt infolge anderer Restriktionen (siehe den Rest dieses Abschnitts) zeitweise zu

Fehlplatzierungen von Kollokaten in andere Ausschnitte. Mit Kreisen gleicher Farbe wie die entsprechende Gramrel lassen sich alle Lexeme auf den ersten Blick leicht auseinanderhalten. Zweitens geben die Kreise Auskunft über die Häufigkeit des Kovorkommens. Dabei soll man sich nach dem Flächeninhalt (nicht Durchmesser) orientieren.

Die Diagramme können auch durch sog. „score circles“ bereichert werden (Abb. 15). Diese dienen einer besseren Einschätzung der Entfernung vom Mittelpunkt des ganzen Kreises. Je näher am Mittelpunkt ein Kollokat angesiedelt ist, desto höher der logDice-Wert der Kombination. Dass die Ringe in den Wortprofilen keine Anwendung gefunden haben, liegt daran, dass erstens die angezeigte Entfernung gleich ist, unabhängig davon, ob die „score circles“ angehakt sind. Zweitens ist die Entfernung immer relativ zu dem Minimal- und Maximal-logDice-Wert der gegebenen Kollokate, d. h. sie taugt nicht zum Vergleichen, zumal der Maßstab derzeit nirgendwo zu finden ist. Es ist auch darauf hinzuweisen, dass der logDice-Wert, sprich die Entfernung von der Mitte, der Platzierung in die richtige Gramrel gegenüber priorisiert ist, was unter bestimmten Bedingungen zur optischen Fehlplatzierung führen kann (siehe *bedeuten* in Abb. 15).

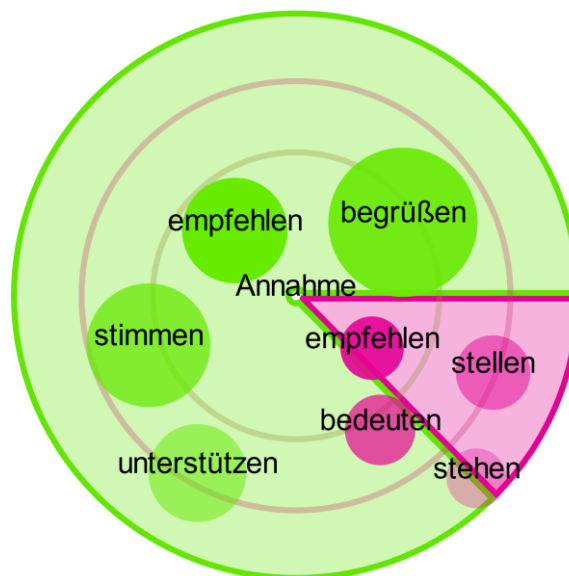


Abbildung 15: Kreisdiagramm mit „score circles“

Eine andere Option, die dynamische Textgröße (Abb. 16), ähnelt sich der Kreisgröße, indem beide die Häufigkeit des gemeinsamen Vorkommens darstellen. Das Ziel der Diagramme ist hier Übersichtlichkeit und dazu trägt die unterschiedliche Textgröße nicht besonders bei, vor allem wenn mehr Kollokate abgebildet werden. Alles in allem wäre sie überflüssig.

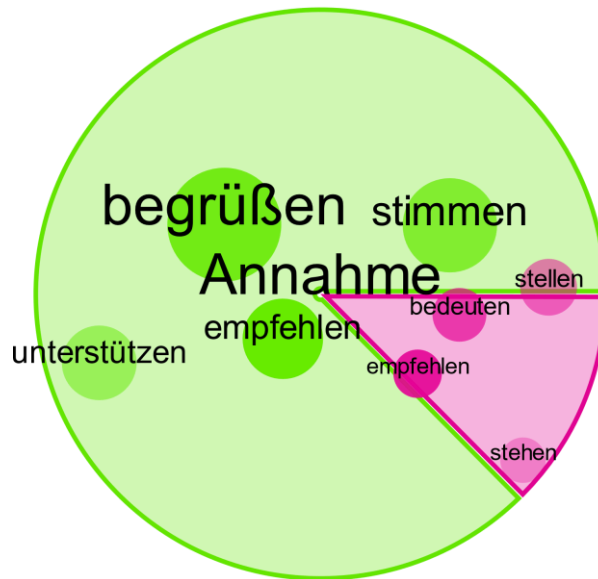


Abbildung 16: Kreisdiagramm mit variabler Textgröße

Die Menüpunkte „Separate relations“ und „Show relation names“ sind ebenfalls nicht benutzt worden. „Separate relations“ löst die Kreisform aus und zeigt alle möglichen Gramrels mit ihren Attributen, was die Entstehung eines höchstens unübersichtlichen Clusters zur Folge hat. Der zweite Punkt versieht die Ausschnitte mit Bezeichnungen der Gramrels. Da sich die Bezeichnung und die Kollokate oft in die Quere kommen, werden hier die Ausschnittbezeichnungen stattdessen unter den Abbildungen angeführt.

Als letztes hat man die Möglichkeit, die Ausschnittgrößen auszugleichen (siehe Abb. 17). Wenn man sie nicht aktiviert, repräsentieren die Winkel die Frequenz sämtlicher Kombinationen, die einer konkreten angezeigten Gramrel angehören, relativ zu anderen angezeigten Gramrels. Die gleiche Größe bedeutet einerseits, dass in jeder Relation eine gleiche Anzahl von Kollokationen abgebildet werden kann, andererseits kann solch eine Darstellung geringfügig irreführend wirken. Es kann somit der Eindruck erweckt werden, dass der Gebrauch über die Gramrels gleichmäßig verteilt wird. Da dies fast nie der Fall ist, werden die unterschiedlichen Größen beibehalten.

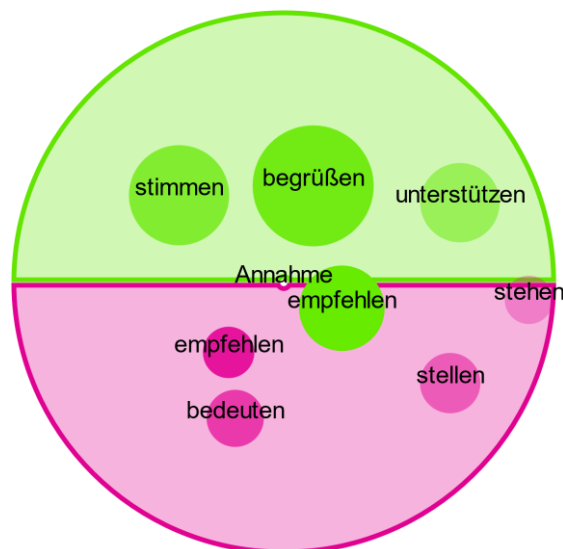


Abbildung 17: Kreisdiagramm mit einer ausgeglichenen Winkelgröße

Zum Schluss muss eine Anzahl von gezeigten Kollokationen pro Gramrel festgelegt werden. In dieser Arbeit war diese Zahl auf das Maximum (=12) gesetzt, womit jedoch folgendes Problem einhergeht. Wie man aus den obigen Abbildungen ersehen kann, ist die voreingestellte Größe der Gramrel-Kreise zu klein, um max. 12*(Anzahl von Gramrels) Kollokate aufzunehmen. Merkwürdigerweise lässt sich die Diagrammgröße in der aktuellen Version nicht einstellen. Die einzige Lösung ist, die Oberfläche im Internetbrowser zu verkleinern und die Seite zu aktualisieren.

4.4 Weitere Aspekte des Profilaufbaus

In diesem Abschnitt werden einige Aspekte/Probleme der Wortprofilerstellung thematisiert, die während ihrer Vorbereitung auftauchten. Diese werden in Form einer Liste gefasster Entschlüsse dargelegt.

1. Anzahl der Gramrels

Die Anzahl der Gramrels pro Substantiv richtet sich nach der Gesamthäufigkeit der Kollokate in einer Gramrel. Die Häufigkeit ist der obigen Beschränkung auf fünf Gramrels untergeordnet, d. h. auch wenn es theoretisch eine Gramrel gäbe, die außer der gewählten fünf läge und zugleich häufig vertreten wäre, würde dies nicht in deren Inklusion resultieren. Gleichzeitig sollen die meisten Kollokate zu dem Substantiv tatsächlich in der Beziehung stehen, die die Gramrel angibt. Dabei wird die bereits erwähnte manuelle Überprüfung von Kollokationen vorausgesetzt. Überprüft werden alle Sätze aus den fünf Gramrels, indem die syntaktische Beziehung zwischen der Kombination und die lexikografische Natur (Kollokation, Kolligation, Phrasem) dieser Beziehung kontrolliert werden. Demnach werden diejenigen, die der syntaktischen Beziehung der Gramrel nicht entsprechen oder Kolligationen sind, ausgelassen.

2. Kollokatenhäufigkeit

Ein weiterer, mit der Gramrelanzahl einhergehender Aspekt betrifft die Häufigkeit der einzelnen Kombinationen. Diese ist auf das Minimum von 3 im konkreten Subkorpus gelegt worden. Damit sich die vorgelegten Beispielsätze nicht als arbeit- bzw. autorenspezifisch auslegen ließen, müssen die 3+ Belege unterschiedlichen Arbeiten entstammen. Kollokationen, die nur in einer Arbeit nachweisbar sind, wurden zwar in den Diagrammen angezeigt, im Wortprofil weiterhin aber nicht berücksichtigt.

3. Fachsprachliche Kollokationen

Fachsprachliche Kollokationen wie *polare Frage* oder *illokutionäres Ziel* findet man in der Auflistung ebenfalls nicht, denn sie sind der AWS generell nicht zuzurechnen. Außerdem sind sie typischerweise nicht genug verteilt, viel zu arbeitsspezifisch und daher nicht sonderlich nützlich. Dies gilt auch für Attribute, die höhere Deskriptivität aufweisen (*finanzieller Grund literarischer Versuch*).

4. Beispielsätze

Es wurde, wenn möglich, Wert darauf gelegt, dass den im Abschnitt 2.2.2 zitierten Regeln über gut gewählte Beispielsätze Rechnung getragen wird. Die Kriterien Satzlänge und Absenz von pronominalen Anaphern waren im Allgemeinen nicht schwer zu erfüllen. Dabei soll man auf die Tatsache Rücksicht nehmen, dass Sätze in der Wissenschaftssprache länger sein können als in der (geschriebenen) Alltagskommunikation. An das Kollokation-im-Hauptsatz-Kriterium habe ich mich fast überhaupt nicht gehalten, obwohl nach einer kursorischen Durchsicht im

Nachhinein evident ist, dass die meisten – oder zumindest sehr viele – Kollokationen sich tatsächlich in den Hauptsätzen vorfinden. Das besondere Augenmerk galt den „third collocates“; praktisch hieß dies, dass Sätze, die auch andere, für die AWS typische Erscheinungen beinhalteten, sei es Kollokationen oder andere AWS-Phänomene, einen Vorrang vor denjenigen hatten, die es eben nicht taten. Letztens wurden Sätze, die zugleich den Kontext bieten, zwar intuitiv präferiert, jedoch konnte dieser Aspekt keinesfalls systematisch berücksichtigt werden.

5. Passivumschreibungen

Verben, die in der sog. Mittelkonstruktion (Eisenberg, 2013, S. 124) auftreten, wie (*sich*) *zeigen* und (*sich*) *ergeben* werden in der Subjekt-Gramrel aufgeführt, denn streng genommen handelt es sich bei den untersuchten Substantiven um grammatische Subjekte dieser Verben. Zudem tendieren die betreffenden Verben, in dieser Konstruktion in der Regel stark grammatikalisiert zu sein. Falls das gleiche Verb auch in der Akkusativ-Gramrel vorkommt, wird sie auch da mit einem entsprechenden Beispielsatz angegeben. Die Priorität liegt in beiden Fällen auf der Überprüfung der Sätze, sodass sie der richtigen Gramrel zugerechnet werden können. Analog zu den Medialkonstruktionen tauchen Schwierigkeiten mit Sätzen mit dem Verb *lassen* auf. Das Verb ordnet Sketch Engine der Subjekt-Gramrel zu und auch wenn jene Zuordnung prinzipiell zutrifft, ist sie nicht von praktischer Relevanz, denn es handelt sich mehr oder weniger um eine kolligationelle Tendenz, nämlich zum Gebrauch des Passivs (oder seiner Alternativen). Da diese aber schon an sich ein Aspekt der grammatischen Spezifika der AWS darstellt, braucht man sie nicht als Kollokat bei jedem Substantiv aufs Neue vorzubringen.

6. Andere Konstruktionen

Außer Kollokationen sind im Umfeld der Nomina auch andere Konstruktionen festzustellen. und einige davon, wie z. B. *etwas_[DAT] auf den Grund gehen*, sind als Teile der Gramrels ans Licht gekommen. Es wäre nicht sinnvoll, solche Erscheinungen lediglich durch ihren verbalen Kern zu definieren. Angesichts dessen wird in dem Profil, hier *Grund*, nicht nur das Verb (*gehen*), sondern die ganze Phrase angeführt. Andere, wie feste Verbindungen mit keiner Variation oder Kolligationen, bedurfte einer Erstellung von N-Grammen. Dabei hat man mit Lemmata als Suchattributen gearbeitet und den Umfang der N-Gramme auf 3 bis 5 festgelegt. Im Falle, dass eine usuelle Konstruktion vorhanden war, fiel diese, versehen mit einer Häufigkeitsangabe, in das entsprechende Wortprofil unter die Kategorie „Andere Konstruktionen“. Sollten zwei oder mehr N-Gramme fast identisch sein (z. B. *von* + [DAT] vs. [GEN]), wurden beide Varianten in Form ihres zugrundeliegenden Musters angegeben³¹, ihre Frequenzen jedoch addiert.

7. Kommentare

Jedes Wortprofil wird um einen Kommentar ergänzt, dessen Aufgabe es ist, die statistischen Angaben zu kontextualisieren und in den meisten Fällen die Funktion des Wortes innerhalb des Textes zu beleuchten. Einige Kommentare bilden „Ministudien“, wo einer zusätzlichen Frage im Kontext des analysierten Substantivs nachgegangen wird. Thematisiert werden in diesen zum Beispiel Phorik, Schreibweise, Synonyme, stilistische Angemessenheit u. Ä.

³¹ In Anlehnung an die Prozedur von Brommer (2018, S. 140).

8. Statistische Angaben

Dem Kommentar geht im jeden Wortprofil eine Tabelle mit statistischen Angaben voraus. Sie gibt Auskunft über die Verteilung des Lexems in den Kategorien Muttersprache und Disziplin, bzw. ihren Überschneidungen. Im Falle der Muttersprache wird auch der *DP*-Wert angeführt. Außerdem werden die Ergebnisse des Chi-Quadrat-Tests (siehe 4.1) angezeigt. Eine andere Einheit, die jedoch wiederholt in den Kommentaren auftritt ist das relative Risiko (RR). RR gibt die Effektgröße des errechneten Unterschieds an und seine Berechnung wurde mittels des Online-Instruments Calc (Cvrček, 2020) durchgeführt. Wenn es angegeben wird, ist das Ergebnis des damit einhergehenden Testverfahrens (der Chi-Quadrat-Test oder der Exakte-Fisher-Test) immer signifikant. Für alle hier durchgeführten Tests wird das Signifikanzniveau von $\alpha = 0,05$ verwendet.

5. Wortprofile

5.1 Fall

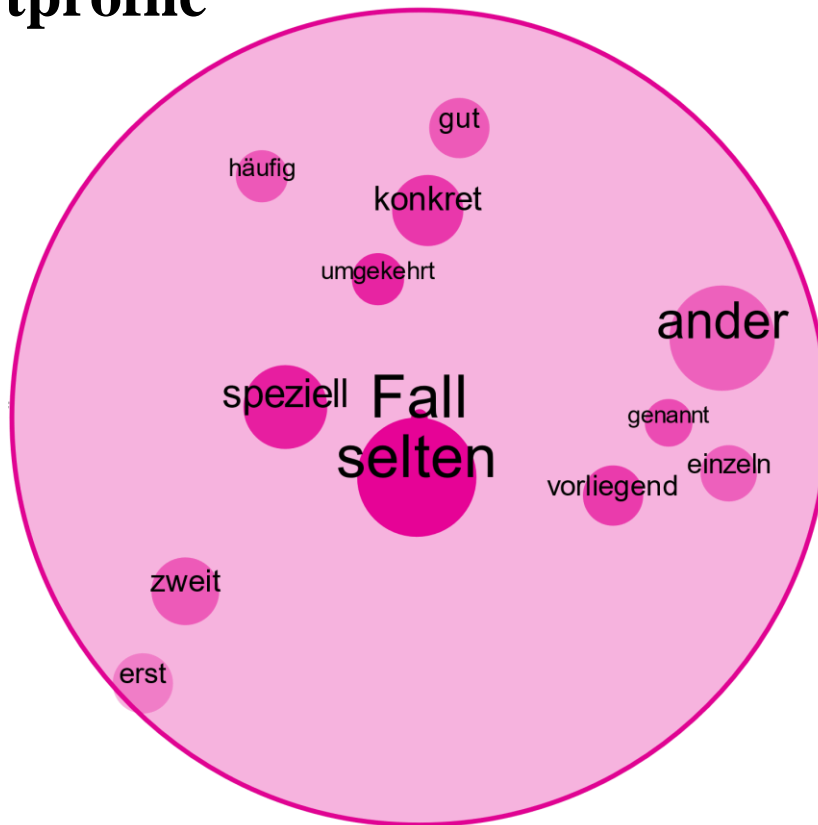


Abbildung 18: Fall L1 **ATTRIBUT**

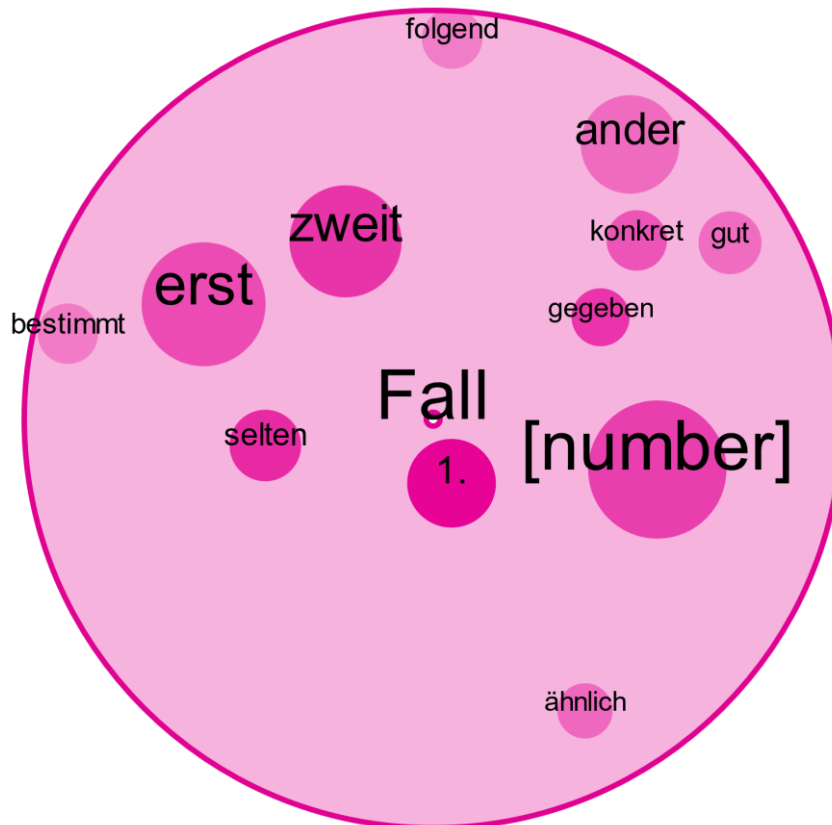


Abbildung 19: Fall L2 **ATTRIBUT**

<i>Fall</i>	Linguistik	Literatur	Σ	DP_{norm}
L1	1434	842	2276	0,32
L2	1557	842	2399	0,57
Σ	2991	1684	4675	
X^2	(1, $N = 4675$) = 1,823, $p = 0,177$			

Tabelle 5: *Fall* – statistische Übersicht

Kommentar

Obwohl über die Muttersprache und Disziplin hinweg keine Gebrauchstendenzen sichtbar sind, ändert sich das Bild schlagartig, wenn wir die zwei Variablen isoliert betrachten. Während die Zahlen innerhalb der Gruppe Muttersprache fast perfekt ausgewogen sind, lässt sich bei dem Substantiv eine eindeutige Präferenz für linguistische Arbeiten feststellen (RR = 1,37; CI [1,341; 1,399]). Um diese nahezu zweifache Überzahl zu erklären, ziehen wir semantische und strukturelle Kriterien heran. Was die Semantik anbelangt hat *Fall* eine zirkumstantielle Bedeutung; Schmid (2000, S. 289) schreibt der englischen Entsprechung *case* überdies sogar eine konditionale Bedeutung zu. Strukturell gesehen treten 74,3 % (Linguistik) und 64,3 % (Literatur) des Substantivs in der folgenden Konstruktion auf: [lemma = "in"] []{0,4} [lemma = "Fall"].

In den sprachwissenschaftlichen Abschlussarbeiten gibt es generell mehr metasprachliche Äußerungen; darunter zählen sich auch Bezugnahmen auf Beispiele (= Fälle). Oft werden auch komplexere Taxonomien und Klassifikationen angesprochen, auf die man sich nachher beruft. Damit hängt auch das Bedürfnis zusammen, Vergleiche zwischen Beispielen und Klassifikationen anzustellen, wodurch sich das erhöhte Präsenz von numerischen Attributen substantiiert. Typisch sind diese eben für die linguistischen Arbeiten, in denen *Fall* als Genitivattribut anderer Merkmale, wie das Prozentzeichen, *Hälfte*, und *Prozent* festzustellen ist und die im literarischen Subkorpus so gut wie keine Verwendung finden. Dort werden als *Fälle* üblicherweise konkrete literarische Werke (5) bzw. Autoren (6) oder Gestalten (7) dieser Werke benannt. Andere für das literaturwissenschaftliche Schreiben charakteristische Funktionen mit *Fall* waren z. B. die Handlungsnacherzählung (8) oder Deutung des Textes (9).

- (15) *Im Falle des Romans Imperium soll sich um ein Plagiat, wo nur die Namen der Autoren und die Titel abgewandelt werden handeln.* (0045T)
- (16) *Vor allem im Falle von Kleist ist solch ein Ausschlussverfahren unbedingt nötig.* (0004D)
- (17) *Die Auswirkungen zeigen sich in Stellas Fall sowohl auf psychischer als auch auf körperlicher Ebene.* (0152D)
- (18) *In diesem Fall bekommt Seppe zwei magische Dinge und eine findet er selbst.* (0023D)
- (19) *Ob die typische Einschränkung des Blickwinkels bereits ausreicht, um einen Ich-Erzähler als unzuverlässig zu kategorisieren, sei dahingestellt; zu bedenken ist allerdings, dass in vielen dieser Fälle die "Diagnose unzuverlässiges Erzählen" als Naturalisierungsstrategie für das Textverständnis gar nicht nötig ist.* (0087D)

In all diesen Beispielen fungiert *Fall* als ein Isolierungs- bzw. Hervorhebungswort – ein oder mehrere Fälle werden fokussiert oder betont. Auf der Textebene dürfte man behaupten, dass *Fall* in diesen Konstruktionen der Wiederaufnahme dient, wodurch der (oft rhematische) Inhalt des vorhergehenden Satzes/Textabschnitts zum Thema gemacht werden kann.

Aufgrund der starken Präferenzen für das Auftreten in der besagten Präpositionalgruppe beschränken sich die graphischen Darstellungen sowie die folgenden Beispielsätze nur auf Attribute, denn bei jeder anderen Gruppe handelte es sich um bloße Kookkurrenzen, deren Großteil syntaktisch falsch eingestuft war.

ATTRIBUT + Fall

selten, anderer, speziell, konkret, erste/zweite, vorliegend, beste, einzeln, umgekehrt, häufig, genannt, bestimmt, untersucht, besondere

*Er wird vielmehr vom Sprecher selbst verursacht, in **seltenen Fällen** auch von jemand anderem. (0147)*

*So könnte man auch die Behauptung aufstellen, dass es sich um eine Variante des Phraseologismus auf der Lauer liegen handelt. In diesem wie auch in genügend **anderen Fällen** scheinen die Grenzen zwischen Variation und Modifikation zu verschwimmen. (0033)*

*Zudem sieht er einen weiteren Grund dafür, "daß Grabbes doch einigermaßen umfängliches Werk trotz intensiver Diskussionen und trotz der Forschung, die ja gerade in den letzten zwanzig Jahren darauf aus ist, das Territorium mit dem Pflug neu zu umbrechen, in der Sekundärliteratur kaum vorhanden ist", in der Tatsache, dass in diesem **speziellen Fall** die Biographie des Autors interessanter ist als sein Werk. (0100)*

*Stellt die eigene Sprache jedoch eine andere Varietät als die örtliche Dachsprache dar, ist die Beherrschung jener für die Teilnahme am gesellschaftlichen Leben oft Voraussetzung. Im **konkreten Fall** von Mehrsprachigkeit spricht man von Bilingualismus. (0026)*

*Im **ersten Fall** wird ein Adressat direkt angesprochen, im **zweiten Fall** nicht. (0147)*

*Der Fragebogen stellt im **vorliegenden Fall** deshalb eine geeignete Erhebungsmethode dar, da auf diese Weise eine relativ große Datenmenge generiert werden konnte. (0025)*

*Eine genaue wissenschaftliche Klärung von SH-Sets bedarf allerdings umfassenderer Untersuchungen, im **besten Fall** empirischer. (0115)*

*Bei einer Vokabularstudie mit mehreren Kindern zwischen 5 und 6 Jahren kam heraus, dass unter den älteren im Durchschnitt 10% und in **einzelnen Fällen** bis zu fast 20% neue lexikale Zusammensetzungsformen verwendeten. (0126)*

*Demgegenüber wurde lediglich von einem englischen Broca Aphasiker berichtet, bei dem der **umgekehrte Fall**, nämlich ein selektives Defizit der irregulären Flexion, festgestellt werden konnte (siehe Cholin et al. 2007). (0032)*

*Es besteht kein Zweifel, dass Gewalt - in den **häufigsten Fällen** körperliche, nicht verbale Gewalt - Kriege, Revolutionen wie auch Aufstände bestimmt. (0004)*

*Während in den **genannten Fällen** die einzelnen Sprachen gemischt oder verändert werden, existieren jedoch auch Sprachkontaktsituationen, in denen zwei unterschiedliche Sprachen oder Varietäten autonom nebeneinander bestehen und sich in ihren Funktionen ergänzen. (0026)*

*In **bestimmten Fällen** können auch Substantive und Adjektive über eine Valenz verfügen, die Aktanten des Adjektivs können obligatorisch oder fakultativ sein, während die des Subjekts in der Regel fakultativ sind. (0136)*

*Im Suffix <-ion> wird das offene ə, das nicht mit Nasalierung einhergeht, von den Gewährspersonen am häufigsten realisiert; es wird in 14 von 25 **untersuchten Fällen** (etwa 56 Prozent) angewendet. (0035)*

*Aufgrund ihrer konstanten Position in Wörtern sind Phonästhemer ein **besonderer Fall** im Lautsymbolismus der mehr morphemischen Eigenschaften besitzt als etwa reine Lautikonizität oder Ideophone. (0130)*

ANDERE KONSTRUKTIONEN

in diesem Fall (328)

*Deshalb kann **in diesem Fall** nicht von Kongruenz gesprochen werden. (0002)*

der Fall SEIN (220)

*Es kann auch **der Fall sein**, dass Menschen eine Heterotopie aufsuchen, da sie glauben, dass in ihr die perfekte Ordnung herrscht. (0008)*

*Zudem stellt der Umgang mit der Sprache bei den Dadaisten an sich einen völlig anderen dar, als es bei den Beat-, Pop- und UndergroundautorInnen **der Fall ist**. (0019)*

im Falle [GEN]/VON (219)

*Ebenfalls erstaunlich ist, dass dieser Einfluss **im Falle der** Deutsch-lernenden auf beiden Kompetenzstufen gleich bleibt und **im Falle der** Spanischler-nenden im Fortgeschrittenenniveau sogar ansteigt. (0002)*

*Vor allem **im Falle von** Kleist ist solch ein Ausschlussverfahren unbedingt nötig. (0004)*

in den meisten Fällen (111)

*Diese Biographien verlaufen **in den meisten Fällen** zur Gänze chronologisch. (0029)*

auf jeden Fall (60)

*Deshalb ist Mühlbergers Literaturgeschichte nur mit Vorsicht zu verwenden, **auf jeden Fall** muss zusätzliche Literatur ergänzend gelesen werden. (0099)*

5.2 Frage

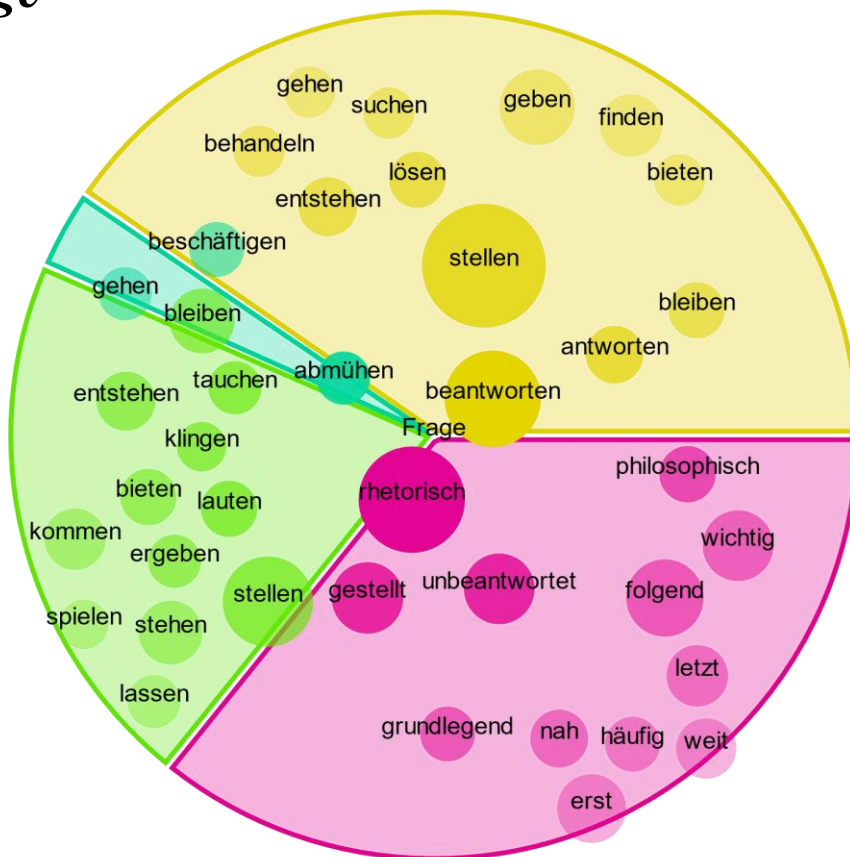


Abbildung 20: Frage L1 **VERB+AKK** **VERB+DAT** **VERB+SUBJEKT** **ATTRIBUT**

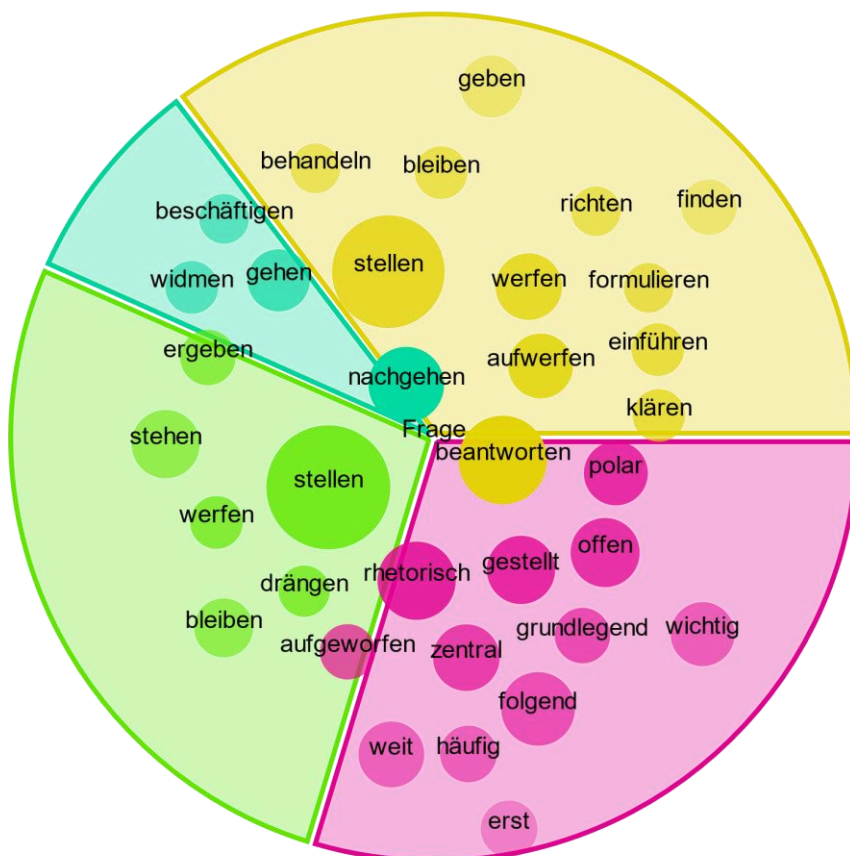


Abbildung 21: Frage L2 **VERB+AKK** **VERB+DAT** **VERB+SUBJEKT** **ATTRIBUT**

<i>Frage</i>	Linguistik	Literatur	Σ	DP_{norm}
L1	1375	1065	2440	0,28
L2	963	887	1850	0,53
Σ	2338	1952	4290	
χ^2	(1, $N = 4290$) = 7,84, $p = 0,005$			

Tabelle 6: *Frage* – statistische Übersicht

Kommentar

Ein Blick auf die Visualisierungen macht deutlich, dass *Frage* von den beiden Gruppen syntaktisch sehr ähnlich gebraucht wird. Wenn wir aber die Disziplin miteinbeziehen und uns die statistischen Ergebnisse ansehen, zeichnet sich ein vollkommen anderes Bild. Der angesetzte Chi-Quadrat-Test kann zwar keine Kausalitäten offenbaren, es lassen sich nichtsdestoweniger folgende grobe Schlüsse ziehen: in der L2-Teilmenge wird das Substantiv weniger benutzt; die linguistischen Arbeiten scheinen etwas mehr Instanzen von *Frage* zu enthalten; der größte Gebrauchsunterschied besteht zwischen L1-Linguistik und L2-Literatur. Die Dispersionszahlen legen die Vermutung nahe, dass die Wahrnehmung von *Frage* in der L2-Gruppe sich möglicherweise geringfügig weniger auf ihre Funktion als SNV ausdehnt, sondern nur auf ihre lexikalische Bedeutung beschränkt ist. Dies scheint im Einklang mit den Daten aus DESTWIKO (v0.1) zu sein, wo sich die Anzahl von manuell getaggtten, als SNV fungierenden Erscheinungen von *Frage* auf 2,55 per 10^4 (L1) und 1,47 per 10^4 Wörter (L2) belief.

Obwohl *Frage* von der Semantik her zu linguistischen SNVs zu zählen ist (Schmid, 2000), überschneidet sie sich auch mit der Bedeutung der faktualen SNVs wie *Problem*. Somit könnte man im Hinblick auf den hypostasierten Inhalt vielleicht von Problemen, Hypothesen, Annahmen usw. reden, die nur der Form nach in einen Fragesatz umgewandelt sind. Von den drei genannten abstrakten Konzepten dient *Frage* mit nur 16 % anaphorischen Gebrauchs (Mach, im Druck) jedoch am häufigsten zur allerersten Einführung des jeweiligen Inhalts.

VERB + *Frage* als Akkusativobjekt

stellen, beantworten, aufwerfen, klären, formulieren, behandeln

*Im Falle der vorliegenden Studie werden allerdings keine expliziten **Fragen gestellt**.* (0002)

*Allerdings kann die aktuelle Untersuchung diese **Frage nicht abschließend beantworten**.* (0024)

*Diese Feststellung **wirft die Frage nach einer möglichen weiteren Kategorie auf**.* (0147)

*Um diese für vorliegende Studie hoch relevante **Frage zu klären**, soll auf die Erkenntnisse von PENZINGER (1985 und 1994) eingegangen werden.* (0025)

*"Innerhalb all jener möglichen Vorgehensweisen für Leitfadeninterviews habe ich mich für jene Arbeit dafür entschieden, die **Fragen im Vorfeld zu formulieren** und an die Interview-Personen vorab zu schicken, sodass sie sich bereits vor dem Interview mit der Thematik auseinandersetzen konnten.* (0118)

*Der zweite empirische Teil **behandelt die Frage**, mit welchen Themen die Klientinnen und Klienten ihre Redebeiträge einleiten.* (0129)

VERB + Frage als Dativobjekt

nachgehen, (auf den Grund) gehen, widmen, beschäftigen

*Es wird dabei der **Frage nachgegangen**, von welchen Bedingungen die Verwendung einer dialektalen/dialektnahen oder eben einer (eher) standardsprachlichen Varietät abhängt. (0025)*

*Um diesen **Fragen auf den Grund gehen** zu können soll zunächst eine Annäherung über den Staatsbegriff erfolgen, der sich nicht nur als sehr umfangreich und somit schwer abgrenzbar zeigt sondern vielmehr auch hinsichtlich der disziplinären Zugänge zu diesem Begriff eine starke Varianz aufweist. (0109)*

*Ein weiterer Schwerpunkt **widmet sich** der **Frage** nach der Verortung der Figuren - infolgedessen soll betrachtet werden, ob sich die Protagonist_innen über fixe Orte und Räumlichkeiten definieren oder ob sie sich an einem Third Space, sozusagen einem Ort im 'Dazwischen' positionieren. (0159)*

*Eher am Rande gehe ich darauf ein, welche Funktion dieses Amerikabild haben könnte und wie es vermittelt wird, da sich **mit** diesen **Fragen** schon zahlreiche andere Arbeiten zur Genüge **beschäftigt** haben. (0153)*

VERB + Frage als Subjekt

(sich)stellen, stehen, offenbleiben, (sich) ergeben, (sich) aufdrängen

*Es **stellt sich** die **Frage**, ob die Antwort ihn wirklich interessiert, oder ob es wirklich seine Widersprüchlichkeit und Angst sind, die ihn davor zurückschrecken lassen, eine befriedigende Antwort einzufordern. (0012)*

*Bei den >>Who-dunit<<-Werken **steht** immer die **Frage** "Wer war der Mörder /die Mörderin? (0110)*

*So **bleibt** etwa die **Frage** nach dem Tatmotiv **offen** und der Fokus wird auf deren Beantwortung gerichtet. (0111)*

*Daraus **ergibt sich** die **Frage**, inwiefern es einen common sense von Studierenden gibt, der sich vom common sense von Expert*innen unterscheidet. (0141)*

*An dieser Stelle **drängt sich** endlich die **Frage auf**, ob die dem Erzähler bzw. der fokalisierten Person "diagnostizierte" psychische Labilität zwingend mit unreliable narration einhergeht - was verneint werden kann. (0087)*

ATTRIBUT + Frage

folgend, gestellt, zentral, weiter/erst/zweit/letzt, wichtig, aufgeworfen, grundlegend/grundsätzlich, einzeln, relevant, wesentlich, diskutiert, weiterführend

*Aufgrund dieser Überlegungen beinhaltete der Fragebogen auch **folgende** drei **Fragen**: Wie sprechen Sie normalerweise mit Kindern, die schon sprechen können? (0025)*

*Somit kann ich meine eingangs **gestellte Frage** mit "ja" beantworten. (0129)*

*Eine **zentrale Frage** hierbei ist, auf welchen Ebenen die Untersuchungen Suchys auf die Exilthematik Bezug nehmen. (0020)*

Eine **weitere Frage**, die sich in diesem Zusammenhang aufmacht, ist, ob es eine "österreichische" Phraseologie bzw. eine österreichische Varietät der standarddeutschen Phraseologie gibt. (0033)

Der Titel des Romans, *Das Märchen vom letzten Gedanken*, wirft eine **wichtige Frage** auf, will man das Werk als historischen Roman analysieren. (0106)

Kapitel 7 stellt schließlich die *Conclusio* dar und es wird auf die in der Einleitung **aufgeworfenen Fragen** eingegangen. (0032)

Die **grundlegendste Frage** nach einer Wissensordnung wäre aber jene danach, was Wissen, was Nicht-Wissen und was Glauben ist. (0086)

Bei der Beantwortung der **einzelnen Fragen** werden jeweils Level von 0 (Nichterfüllung), 1 (teilweise Erfüllung) und 2 (vollständige Erfüllung) vergeben. (0120)

Um diese für vorliegende Studie hoch **relevante Frage** zu klären, soll auf die Erkenntnisse von PENZINGER (1985 und 1994) eingegangen werden. (0025)

Die zweite für die Identität **wesentliche Frage**, die sich die Protagonisten stellen müssen, ist die nach der Herkunft. (0028)

Eine in dieser Hinsicht oft eher implizit aufgeworfene und daher wenig **diskutierte Frage** ist, ob auch kontrastive Topiks eigene Intonationsphrasen bilden und daher eine eigene Endkontur aufweisen. (0031)

Im nächsten Kapitel werden die Ergebnisse der Arbeit zusammengefasst und mögliche **weiterführende Fragen** formuliert. 5 Zusammenfassung 5 Zusammenfassung Die Arbeit setzt sich aus einem theoretischen und einem empirischen Teil in Form einer Korpusanalyse zusammen. (0114)

5.3 Grund

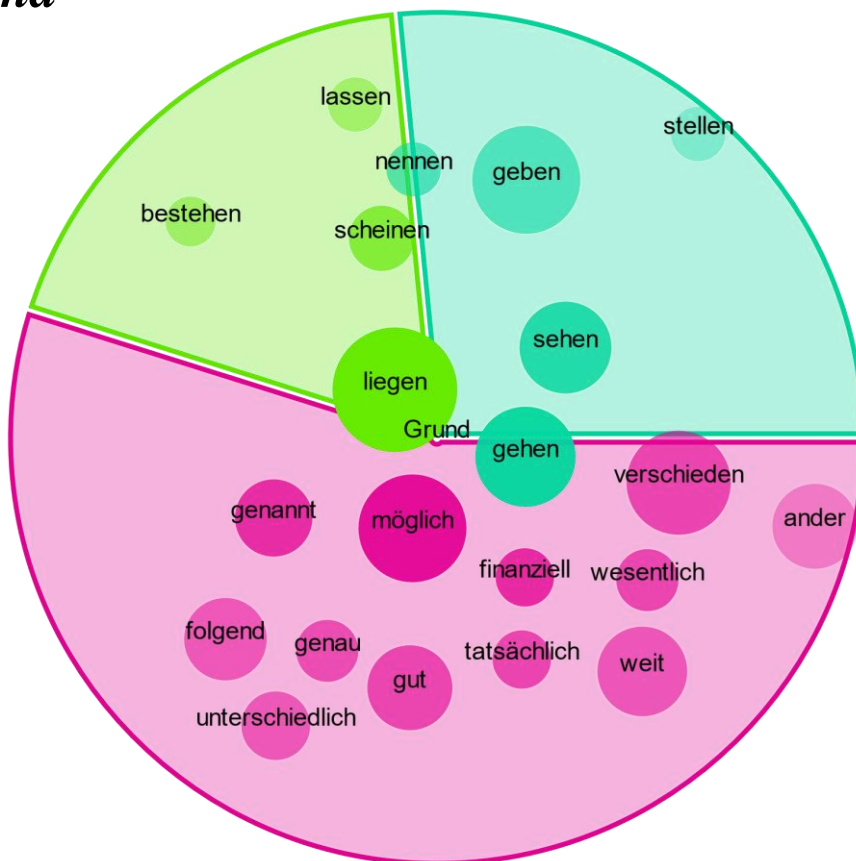


Abbildung 22: Grund L1 VERB+AKK VERB+SUBJEKT ATTRIBUT

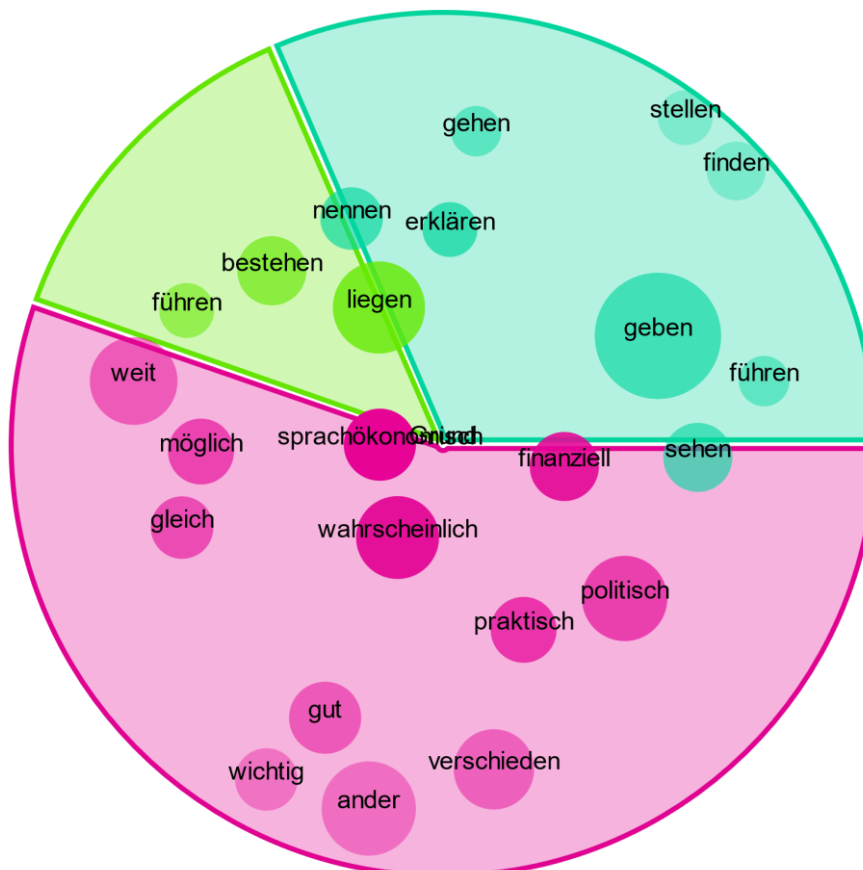


Abbildung 23: Grund L2 VERB+AKK VERB+SUBJEKT ATTRIBUT

<i>Grund</i>	Linguistik	Literatur	Σ	DP_{norm}
L1	802	801	1603	0,28
L2	676	703	1379	0,52
Σ	1478	1504	2982	
X^2	(1, $N = 2982$) = 0,303, $p = 0,582$			

Tabelle 7: *Grund* – statistische Übersicht

Kommentar

Wie die Tabelle zeigt, ist keine Interaktion zwischen der Muttersprache und Disziplin nachzuweisen. Der einzige Unterschied liegt zwischen L1 und L2 ($X^2 = 15,11$; $p < 0,001$) obwohl die Effektstärke relativ gering ist (RR = 1,071; CI [1,036; 1,107]).

Außer den unten angeführten Kollokaten, treten mit dem Substantiv *Grund* häufig Verben *scheinen, sich lassen...* (*feststellen, nachweisen*) auf:

(20) *Die Gründe dafür **scheinen** die hohe Komplexität der zu beschreibenden Positionierungshandlung, aber vor allem auch der nicht typische Zielzustand des bewegten Objekts zu sein.* (0116D)

(21) *Diese Gründe **lassen sich** auch bei mir als Möglichkeit nachweisen.* (0114D)

Grund kommt oft am Satzanfang vor, insbesondere in der Präpositionalgruppe mit *aus*, wo es sich auf den vorhergehenden Satz zurückbezieht. Ein weiterer interessanter Punkt betrifft die Schreibweise von *aufgrund* bzw. *auf Grund*. Obwohl beide Formen als richtig beurteilt werden (Dudenredaktion, 2009, S. 898), begegnet man der grammatikalisierten Verschmelzung heutzutage viel öfter. Die folgende Tabelle veranschaulicht die Gebrauchsunterschiede zwischen den österreichischen und den tschechischen Studierenden.

		<i>aufgrund</i>	<i>auf Grund</i>
L1	Häufigkeit ^a	1653 (241,41 pM)	122 (17,82 pM)
	Arbeiten	99 (99 %)	21 (21 %)
	med/Arbeit ^b	14	1
L2	Häufigkeit	657 (95,95 pM)	216 (31,55 pM)
	Arbeiten	133 (47,5 %)	71 (25,3 %)
	med/Arbeit ^b	4	2

^a rel. Freq. per Subkorpus berechnet

^b Arbeiten mit 0 Treffer nicht einbezogen

Tabelle 8: Vergleich von *aufgrund* und *auf Grund*

Schon auf den ersten Blick wird deutlich, dass die L1-Gruppe für die Präferenz der synthetischen Präposition stärker sensibilisiert ist. Was an den Ergebnissen zudem noch auffällt sind die absoluten Zahlen sowie die Verteilung des Ausdrucks über die Arbeiten; unabhängig von der Schreibung haben ihn nämlich alle L1-Verfasser benutzt im Vergleich zu nur 52,1 % der Nichtmuttersprachler.

VERB + Grund als Akkusativobjekt

geben, auf den Grund gehen, sehen, nennen, darstellen

Es gibt mehrere Gründe, wieso es mehr als 50 Jahre später noch möglich und auch durchaus gerechtfertigt war, von einem Affekt der Intellektuellen gegen den Sport zu sprechen. (0089)

Der Frage, um welche der drei Varianten, wie sie in Kapitel 5.1. dargestellt wurden, es sich dabei genau handelt, wird nun in folgendem Absatz auf den Grund gegangen. (0101)

Den Grund dafür sieht sie in einem methodischen Problem, da noch kein systematisches Verfahren zur Erhebung von Aussagen über Textsorten etabliert sei, das relevante Resultate bringen würde (vgl. Adamzik 2008: 147). (0123)

Einen Grund dafür sieht der Schriftsteller in der Verkommenheit moralischer Werte, sowie der Schnelllebigkeit seiner Generation: ... (0155)

Generell werden in der Fachliteratur viele verschiedene Gründe genannt, weshalb nicht-satzwertige Konstituenten ins Nachfeld verschoben werden (können). (0017)

Er nennt drei Gründe: Erstens bezieht sich das Schichtenmodell, welches von HOFFMANN begründet wurde und lange als Grundlage für die Fachsprachenforschung galt, in erster Linie auf lexikalische Merkmale. (0148)

Die Traumata der in Skandinavien gelandeten Flüchtlingskinder stellten einen weiteren Grund für den Willen zur Veränderung bei den skandinavischen Kinderliteratur-Schaffenden dar. (0090)

VERB + Grund als Subjekt

liegen, bestehen

Ein Grund für die Instrumentalisierung des Verlags für ideologische Zwecke lag Hall zufolge darin, dass der Verlag "zu 60 % im Besitz der Gemeinde war und vom Kulturamt verwaltet wurde". (0013)

Die Kinder seien laut ihm nicht in der Lage, den komplexen Inhalten zu folgen. Ein weiterer Grund dafür liegt in der Verwendung der Umgangssprache. (0029)

Der Grund für diesen Willen besteht darin, dass sie keine Seele besitzen und diese nur durch die Ehe mit einem Menschen erlangen können. (0029)

Beim Schreiben von Aufsätzen später in der Schule erweist sie sich auch als gut, sodass alles in einem kein Grund zur Annahme einer ernsthaften Sprachstörung besteht. (0126)

ATTRIBUT + Grund

möglich, verschieden, weiter, folgend, genannt, unterschiedlich, wesentlich, tatsächlich

Ein möglicher Grund könnte laut FLEISCHER/KASPER/LENZ (2012: 32) die klare Stigmatisierung solcher Konstruktionen sein. (0016)

Generell werden in der Fachliteratur viele verschiedene Gründe genannt, weshalb nicht-satzwertige Konstituenten ins Nachfeld verschoben werden (können). (0017)

Ein **weiterer Grund**, warum hier mit Hoffmanns Terminus gearbeitet wird, ist seine Abgrenzung vom Begriff des Raums als Metapher. (0157)

Rechtspopulismus ist aus **folgenden** beiden **Gründen** eine bloß modernisierte und salonfähigere Politik-Form rechtsextremer Politik: Rechtspopulistinnen und Rechtspopulisten appellieren an Verteidigung und Bewahrung der ethnisch-kulturellen Identität. (0146)

Aus **genannten Gründen** habe ich mich entschieden, die Figuren der Erzählungen zunächst in drei große Gruppen zu untergliedern und erst dann auf etwaige nähere Differenzierungen einzugehen. (0153)

Diese Kategorisierungsversuche erweisen sich aus **unterschiedlichen Gründen** als problematisch. (0159)

Die Märchen sollten es den Kindern aber auch ermöglichen, aus dem Alltag zu fliehen. Das war ein **wesentlicher Grund**, auch ausländische Märchen zu übersetzen. (0095)

Als diskussionswürdig stellt sich nicht nur eine einheitliche Terminologie dar, sondern auch die empirische Absicherung, da es oftmals nicht so einfach ist, die **tatsächlichen Gründe** für die vorgenommene Ausklammerung festzustellen. (0017)

ANDERE KONSTRUKTIONEN

der Grund (da)für (147)

Der Grund dafür könnte sein, dass sie sich ja noch immerhin in ihrem Element befunden haben und nicht am Land. (0029)

Der Grund für die relativ freie Positionierbarkeit von Epitheta und Vokativen (besonders in Kontrast auf Hanging Topics) wird von De Vries eben in der Verankerung an den Diskurs vermutet. (0133)

aus [ADJ|ART] GRUND (101)

In Kapitel 4.1.2 wurde bereits darauf eingegangen, dass die Abwesenheit des Explorators **aus verschiedensten Gründen** auch als Nachteil gesehen werden kann - allen voran werden ein Kontrollverlust sowie das Versäumen unmittelbarer Reaktionen genannt. (0016)

5.4 Unterschied

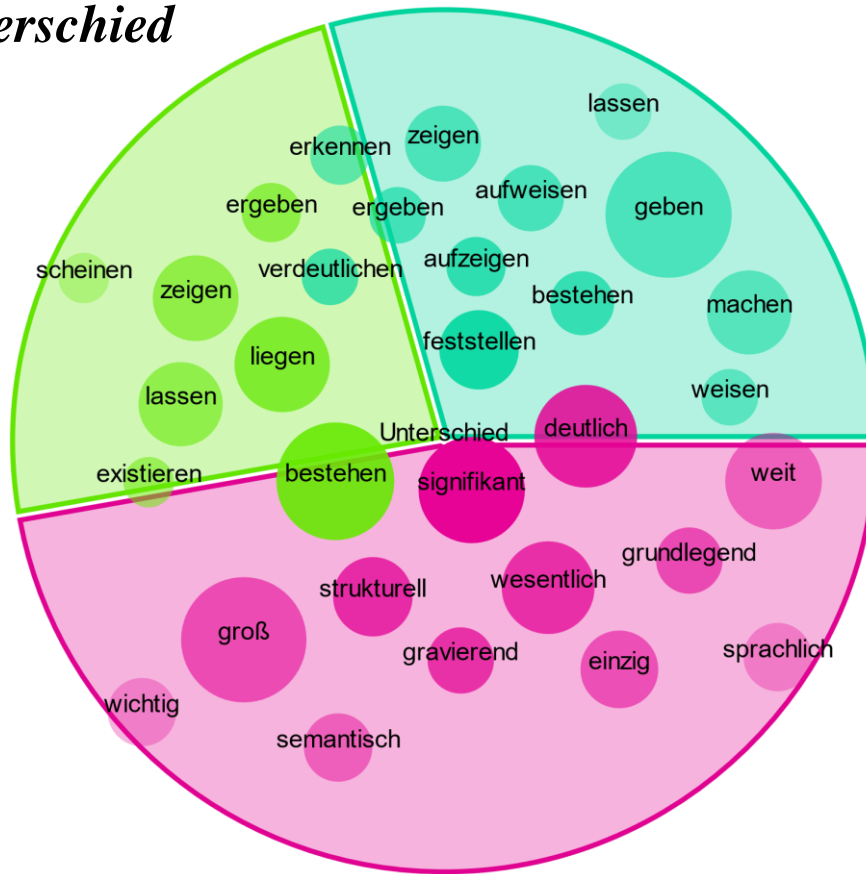


Abbildung 24: *Unterschied* L1 VERB+AKK VERB+SUBJEKT ATTRIBUT

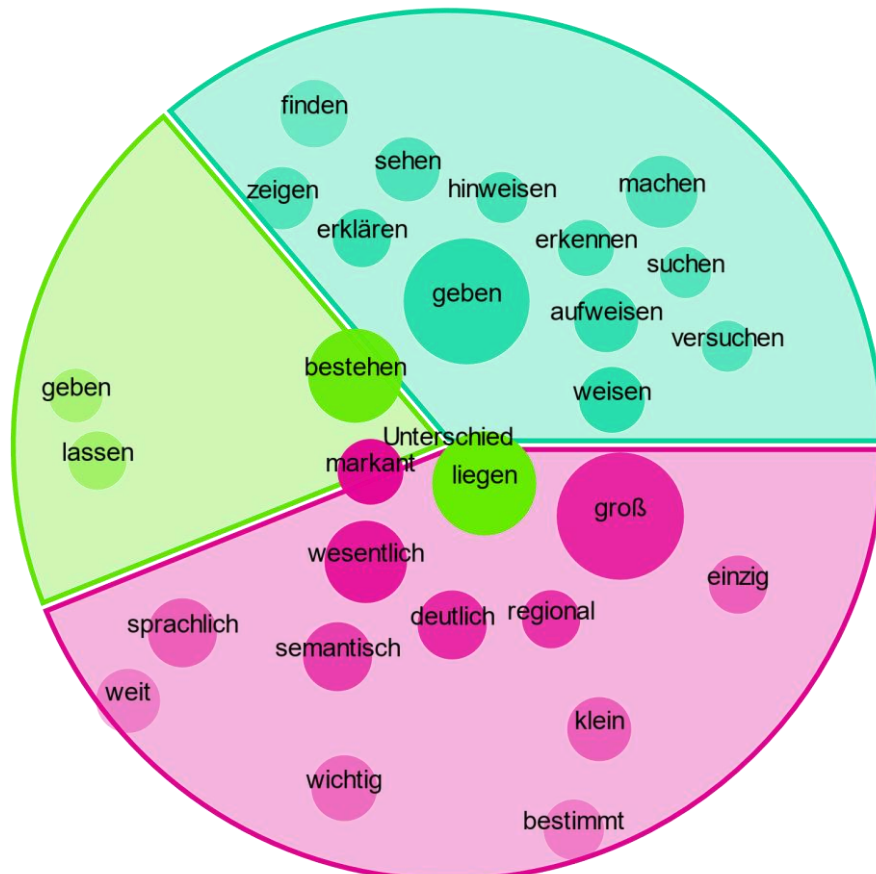


Abbildung 25: *Grund* L2 VERB+AKK VERB+SUBJEKT ATTRIBUT

<i>Unterschied</i>	Linguistik	Literatur	Σ	DP_{norm}
L1	942	342	1284	0,36
L2	803	421	1224	0,59
Σ	1745	763	2508	
X^2	(1, $N = 2508$) = 17,827, $p < 0,001$			

Tabelle 9: *Unterschied* – statistische Übersicht

Kommentar

Den Zahlen in der Tab. 9 nach könnte man *Unterschied* als ein disziplinspezifisch gebrauchtes Substantiv betrachten. Demzufolge scheinen Unterschiede in sprachwissenschaftlich orientierten Abschlussarbeiten mehr angesprochen zu werden (RR = 1,489; CI [1,451; 1,529]). Man könnte zwar behaupten, dass auch Literatur kontrastiver Analyse unterzogen werden kann, wobei die aufgedeckten Unterschiede versprachlicht werden müssen, diese sind tendenziell jedoch eher abstrakter Natur und daher komplizierter, um als Inhalt eines einzigen Abstraktums erfasst zu werden. Zudem ziehen sie sich in der Regel durch die ganze Arbeit, mitunter auch als das leitende Thema. In der Sprachwissenschaft sind die Unterschiede oft quantifiziert und daher etwas greifbarer – sie müssen folglich direkt als *Unterschiede* beschrieben werden, so sind zumindest die Gepflogenheiten in der Disziplin.

Hinsichtlich der attributiven Kollokationen des Substantivs lässt sich folgende disziplinbezogene Tendenz ausmachen: Attribute *signifikant*, *markant*, *strukturell*, *regional* und *deutlich* findet man im Korpus eher in den sprachwissenschaftlichen Texten. *Sozial* und *entscheidend* findet man in dem Korpus eher in der Literaturwissenschaft.

VERB + *Unterschied* als Akkusativobjekt

geben, machen, feststellen, zeigen, aufweisen, aufzeigen, erkennen, verdeutlichen, ergeben, erklären, beschreiben, sehen (in), darstellen

*Gibt es signifikante **Unterschiede** im quantitativen Gebrauch von Konnektoren zwischen Bachelorarbeiten und wissenschaftlichen Fachartikeln?* (0119)

*Für die Interpretation kann es durchaus einen **Unterschied machen**, ob der erinnerte Tod des Vaters, die Vergangenheit, oder die via eigener Dichtung prophezeite Zukunft, in welcher der Tod Nathanael aus Claras Augen ansieht, Nathanael in den Wahnsinn treibt.* (0088)

*Dies ist eines der Merkmale Grus Lebensweise, welches den groben **Unterschied** zu den gängigen Werwolfdarstellungen in der Literatur augenscheinlich **macht**.* (0098)

*Insgesamt ließen sich wieder bei 28 Lemmata **Unterschiede feststellen**.* (0131)

*RIESEL sprach von "nationalen Varianten" des Deutschen und **zeigte** die **Unterschiede** zwischen diesen hinsichtlich Lexik, Aussprache etc. auf.* (0117)

*Nationale Varietäten können laut CLYNE auf verschiedenen Ebenen **Unterschiede aufweisen**.* (0117)

*Das Ziel der Arbeit soll einerseits sein, die **Unterschiede** in der Wirkung der Synchronfassungen **aufzuzeigen**, welche Gründe es dafür gibt, und ob es einen anderen Weg hätte geben können.* (0120)

*Auch zwischen Esmeralda und Stella sind klassenspezifische **Unterschiede** zu **erkennen**.* (0152)

Ein einfaches Beispiel, das Christa Dürscheid zitiert, **verdeutlicht** den **Unterschied** beider Erklärungsansätze. (0138)

Dass sich innerhalb der beiden Fassungen einige **Unterschiede** in der Führungsgestaltung **ergeben**, soll hier explizit gemacht werden. (0109)

Kieweg (2003) **erklärt** den **Unterschied** zwischen den Begriffen Emotion, Stimmung und Affekt folgendermaßen: Emotionen sind interne, introspektiv wahrnehmbare Zustände, die einhergehen können mit physiologischen Korrelaten (wie Angstschweiß, Pulsanstieg usw.), sie sind in starkem Maße sozial beeinflusst. (0147)

KRETZENBACHER (vgl. 1998: 138) **beschreibt** auch einen **Unterschied** zwischen der deutsch- und englischsprachigen Wissenschaftskommunikation. (0148)

Den **Unterschied** zum öffentlichen Epitext **sieht** Genette nicht **darin**, dass der private Epitext nicht zur Veröffentlichung gedacht sei, sondern **darin**, dass sich der private Epitext grundsätzlich an eine konkrete Person und nicht zu allererst an die Öffentlichkeit richte. (0143)

Die Figur wird in Abhängigkeit zum Raum bewegt und ihr Endpunkt ist kein menschlicher Rezipient, was den größten **Unterschied** zu Besitzwechselhandlungen **darstellt**. (0116)

VERB + Unterschied als Subjekt

bestehen (in), liegen (in), sich zeigen, sich ergeben, existieren

Auffallend ist, daß kaum ein **Unterschied** zu den jeweils vergleichbaren Presstexten **besteht**. (0024)

Der **Unterschied** der MentorInnen zu jenen im Erziehungsroman **besteht** **darin**, dass sie mehr durch ihre konkrete Lebenserfahrung, die bildend wirkt, **bestehen** als durch pädagogische Ratschläge. (0012)

Ein weiterer, bedeutender **Unterschied** **liegt** **darin**, dass die Möglichkeit einer Kombination mehrerer Modalverben nur in der "Duden-Grammatik" aufscheint. (0030)

Betrachtet man nun im Vergleich die spanische Sprache, **zeigt** **sich** ein deutlicher **Unterschied**: Im Falle des Spanischen gab es ebenfalls, wie beim Englischen und Französischen, ein Sprachzentrum. (0117)

Bei der Analyse der Wortfelder **ergaben** **sich** keine bedeutenden **Unterschiede**, was darauf schließen lässt, dass zumindest die untersuchten Wortfelder keinen Einfluss auf die Verbreitung des gesamtösterreichischen Wortschatzes haben. (0024)

Dennoch **existieren** auch gravierende **Unterschiede**, vor allem in Bezug auf den Verlauf der Erzählungen und die Erzählweise. (0106)

ATTRIBUT + Unterschied

groß, signifikant, deutlich, weiter, wesentlich, grundlegend, gravierend, bedeutend, entscheidend, eindeutig, etwaig, klar

Ein weiterer **großer** **Unterschied** findet sich jedoch bereits in der Erarbeitung der Satzarten an sich. (0030)

Und genau hier offenbart sich ein **signifikanter Unterschied** zwischen Charly, Joe, Birgit, Monika und Fausers Harry Gelb: Gelb möchte aktiv an der Revolution und den Veränderungen teilhaben, selbst in das Geschehen eingreifen. (0019)

Foucault stellte dabei fest, dass es einen **deutlichen Unterschied** zwischen dem, "was sich praktisch alles machen lässt oder was sich theoretisch alles denken lässt, und dem, was Menschen tatsächlich sagen, machen und denken", gibt. (0102)

Weitere interessante Unterschiede treten bei den Affixen auf. (0102)

Jedoch verweist Göttische auch darauf, dass **wesentliche Unterschiede** in der Form der Sprachproblematisierung vorliegen. (0093)

Ein **grundlegender Unterschied** in Bezug auf die der Analyse zugrundeliegenden Belege ist, dass Steinhoff mit den gesamten Belegen arbeitet, diese Untersuchung aber ausschließlich mit bereinigten Belegen. (0141)

Der Frage, ob es **gravierende Unterschiede** in der Verwendung und Formulierung von Aufforderungen in diesen beiden Epochen gibt, wird im vierten Kapitel, wo die Ergebnisse der Untersuchung vorgestellt werden, noch ausführlicher nachgegangen. (0127)

Ein **bedeutender Unterschied** bestehe auch in der Faktenpräsentation bzw. in der Gestaltung der Autobiografie, die vom Autobiografen/der Autobiografin subjektiv intendiert sei und vom Biografen/von der Biografin als objektiv anzusehen ist. (0091)

In Bezug auf die Konzepte der Multi- und Interkulturalität weist der Begriff der Hybridität also einen **entscheidenden Unterschied** auf, da er nicht kulturelle Diversität in den Vordergrund rückt, sondern die kulturelle Differenz. (0159)

Auch bei den Bewertungsaufgaben zeigen sich **eindeutige Unterschiede** zwischen Dialekt- und Standarddurchgang (siehe Abb. 20). (0016)

Eine (kritische) Darstellung bisheriger Ergebnisse und Überlegungen in Bezug auf **etwaige Unterschiede** ist jedoch unerlässlich, um eine anschlussfähige Arbeit vorlegen zu können. (0141)

Im Vergleich zwischen den beiden Gruppen ergab sich nur bei den Bildbeschreibungen von Activities ein sehr **klarer Unterschied**: Während die erste Gruppe ausschließlich tun-Periphrasen bildete, bevorzugte es die Gruppe der am-Ablehner mehrheitlich, keine der untersuchten Progressivkonstruktionen zu verwenden. (0016)

ANDERE KONSTRUKTIONEN

DER/EIN Unterschied zwischen (148)

Auf die Signifikanz **der Unterschiede zwischen** regulären und irregulären Verben wird in dieser Studie jedoch nicht eingegangen. (0032)

Unterschied in DER (89)

Betrachtet man zunächst nur die **Unterschiede in den** Beispielen (8) und (9), so lässt sich schnell erkennen, dass in beiden Fällen eine NP in Funktion einer Akkusativergänzung vorliegt. (0017)

im Unterschied zu (85)

Im Unterschied zu seinen Brüdern wohnt Richard noch bei seinen Eltern, doch das ändert sich rasch nachdem er im April 2004 Kerstin kennenlernt. (0110)

5.5 Möglichkeit

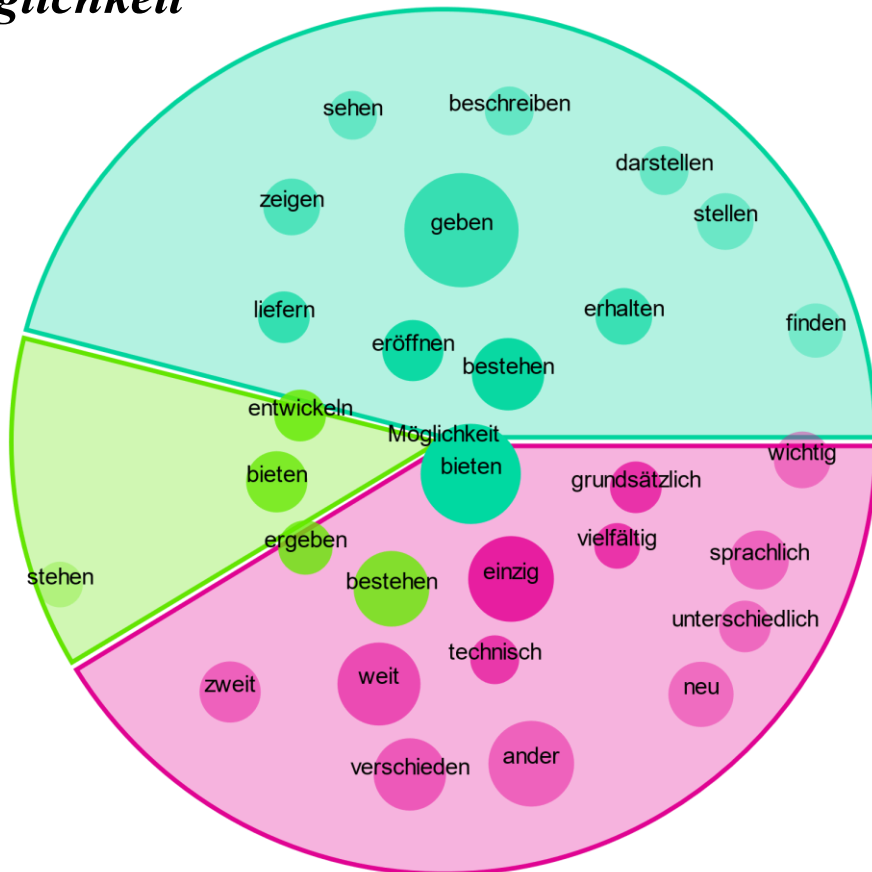


Abbildung 26: Möglichkeit L1 VERB+AKK VERB+SUBJEKT ATTRIBUT

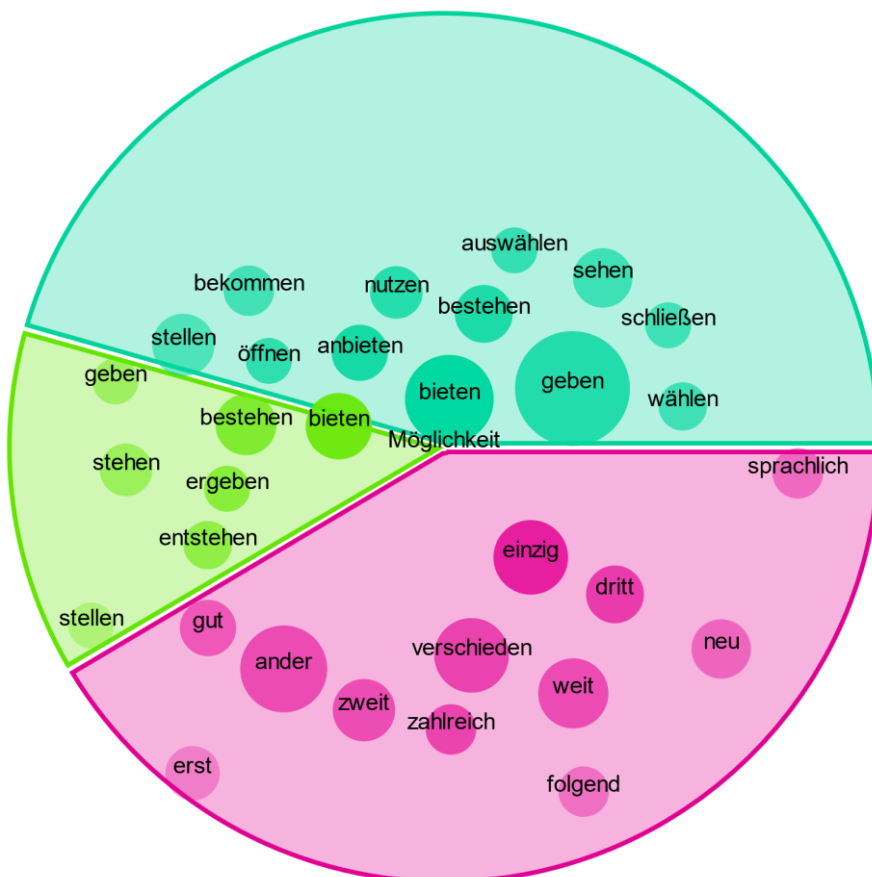


Abbildung 27: Möglichkeit L2 VERB+AKK VERB+SUBJEKT ATTRIBUT

<i>Möglichkeit</i>	Linguistik	Literatur	Σ	DP_{norm}
L1	538	477	1015	0,29
L2	557	407	964	0,54
Σ	1095	884	1979	
X^2	(1, $N = 1979$) = 4,561, $p = 0,033$			

Tabelle 10: *Möglichkeit* – statistische Übersicht

Kommentar

Bei *Möglichkeit* lässt sich eine relativ kleine Diskrepanz zwischen den Disziplinen festzustellen (RR = 1,184; CI [1,138; 1,232]) Betrachtet man zugleich beide Variablen, stellt sich heraus, dass in der L1-Gruppe die Unterschiede fast unauffällig sind (64 Treffer), zumindest verglichen mit der L2-Gruppe (150). Obgleich die kontrastiven Word-sketches zu Überprüfungszwecken geschaffen worden sind, hielt ich die Häufigkeiten der in Gramrels aussortierten Kollokationen für zu gering, um sie hier anzuführen.

Im Allgemeinen ist *Möglichkeit* in der Literatur üblicherweise entweder auf den Autor des analysierten Werks oder die einzelnen Figuren bei der Wiedergabe/Interpretation der Handlung bezogen. In der Sprachwissenschaft berufen sich die Verfasser tendenziell eher auf die angewandten Methoden/Ansätze; gegebenenfalls wird das Substantiv auch metasprachlich verwendet, d. h. es gibt Auskunft über die Leistung des jeweiligen sprachlichen Phänomens.

VERB + *Möglichkeit* als Akkusativobjekt

es gibt, bieten, erhalten, eröffnen, (auf)zeigen, darstellen, (sich) finden, liefern, sehen (in), eingehen auf, schaffen, anführen

Es gibt drei Möglichkeiten, in welcher Form diese Sätze im Y-Feld der Nominalphrasen auftauchen. (0017)

Die Forschung bietet also verschiedene Möglichkeiten zur Kategorisierung der Modelle, welche Gemeinsamkeiten und Unterschiede verdeutlichen. (0115)

So wird die Adoleszenzphase als Proberaum dargestellt, in dem sie die Möglichkeit erhalten, sich selbst zu erproben und mit ihrer Identität zu experimentieren. (0028)

Ihre Analyse eröffnet die Möglichkeit, zum einen z.B. räumliche Gebundenheit von Personen und zum anderen persönliche Zugangsbeschränkungen in Bezug auf einen literarischen Text und dessen Handlung zu erklären. (0027)

Cordula Meißner (2014) legt eine Systematisierung und Beschreibung zu figurativen Verben vor, und Franziska Wallner (2014) belegt in ihrer korpuslinguistischen Studie die abweichende Gebrauchspraxis von Kollokationen in der Wissenschaftssprache und zeigt eine Möglichkeit zur Aufbereitung von Kollokationen auf. (0119)

Diese Unterscheidung untereinander wurde auch als Sieg über das herrschende Regime wahrgenommen und stellte gleichzeitig eine Möglichkeit dar, ein gewisses Maß an Normalität zu konstruieren. (0152)

In Untersuchungsgruppe 1 finden sich durchaus unterschiedliche Möglichkeiten der Besetzung gerade dieser beiden Felder, die nun im Detail näher erläutert werden sollen. (0017)

Der dabei vorherrschende Diskurs **liefert die Möglichkeiten** der Rechtfertigung, verurteilt jedoch auch im Gegenzug illegitime und illegale Gewalt. (0152)

Karlheinz Valtl **sieht** hierin die **Möglichkeit** des Erziehungssystems bzw. seiner ausführenden Kräfte unterstützend auf die Jugendlichen und Kinder einzuwirken. (0118)

In "Das kunstseidene Mädchen" wird außerdem **auf die schier unbegrenzten Möglichkeiten** in einer Großstadt **eingegangen**. (0104)

Somit **schafft** Konversion zwar die **Möglichkeit** für idiomatische Veränderung der Verbglieder, beinhaltet sie aber nicht. (0113)

Weiters **führt** Fleischer drei **Möglichkeiten an**, wie sowohl usuelle als auch okkasionelle Varianten realisiert werden können. (0033)

VERB + Möglichkeit als Subjekt

bestehen, sich bieten, sich ergeben

Umgekehrt **besteht** die **Möglichkeit**, dass bestimmte Denotate Wertungen hervorrufen, die den Wortinhalt mit konstituieren. (0146)

Durch die Indoktrination der Jugendliteratur **bot sich** gleichsam die **Möglichkeit**, faschistische Werte voraussetzungslos zu propagieren. (0013)

Erst daraus **ergibt sich** die **Möglichkeit**, städtisches Material, wie Gesprächsfetzen, Werbungstexten, Zeitungsmeldungen, Lieder und Hinweise der Behörden, zusammenzufügen, und damit eine unglaubliche Informations- und Materialfülle zu schaffen, die den Leser zwar überfordern kann, gleichzeitig aber die Komplexität der Großstadt verdeutlicht. (0104)

ATTRIBUT + Möglichkeit

einzig, andere, weiter, verschieden, unterschiedlich, vielfältig

Dabei ist diese Art der Unterscheidung bei weitem nicht die **einzige Möglichkeit**, Sprachproduktionen zu kategorisieren. (0134)

Im Gegensatz zu den **anderen Möglichkeiten** gibt es hier sogar eine eigene grammatische Form, die gebildet werden kann. (0127)

Eine **weitere Möglichkeit** eines solchen Verbalkomplexes, die ebenso dessen Umfang beeinflusst, ist die Verknüpfung zweier Modalverben. (0030)

Handlungen, die im alltäglichen Leben von großer Relevanz sind, bedürfen notwendigerweise auch **verschiedener Möglichkeiten** der Versprachlichung. (0116)

In Untersuchungsgruppe 1 finden sich durchaus **unterschiedliche Möglichkeiten** der Besetzung gerade dieser beiden Felder, die nun im Detail näher erläutert werden sollen. (0017)

Die einfachen Regeln und die **vielfältigen Möglichkeiten** der Spielentfaltung (von der robusten Härte der einfachen Industriearbeiter bis zum verschlagenen, kreativen Spiel der Straßenfußballer) machten Fußball zu einem Sport, den man auch ohne große Vorkenntnisse oder teure Ausrüstung spielen konnte. (0089)

5.6 Problem

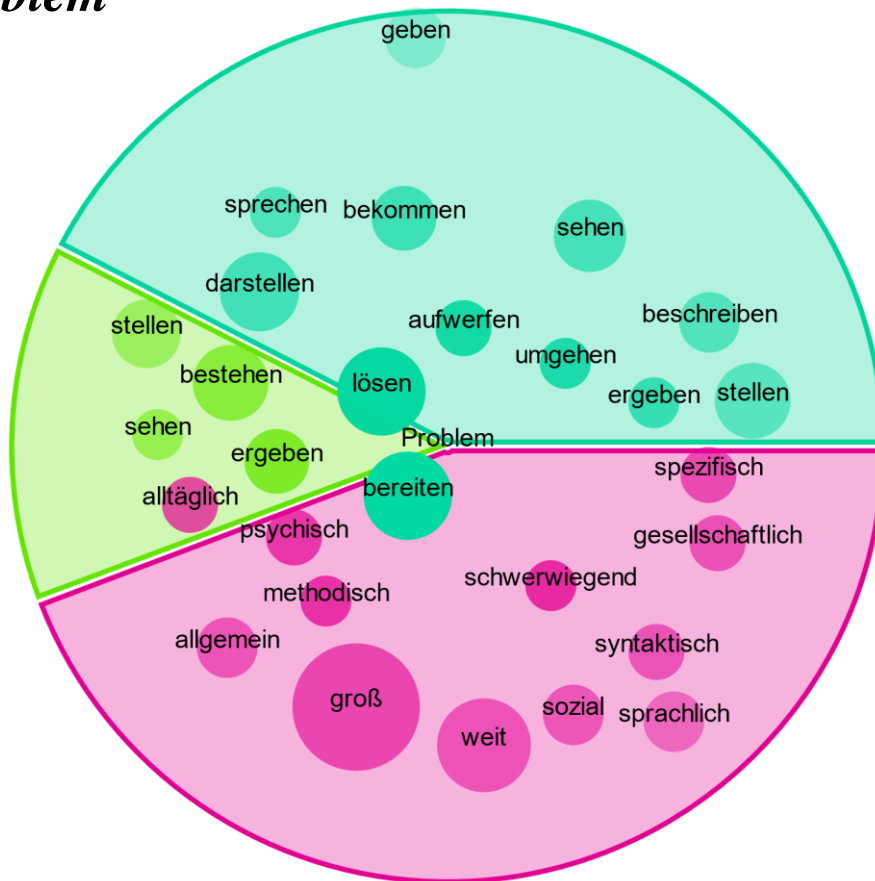


Abbildung 28: Problem L1 VERB+AKK VERB+SUBJEKT ATTRIBUT

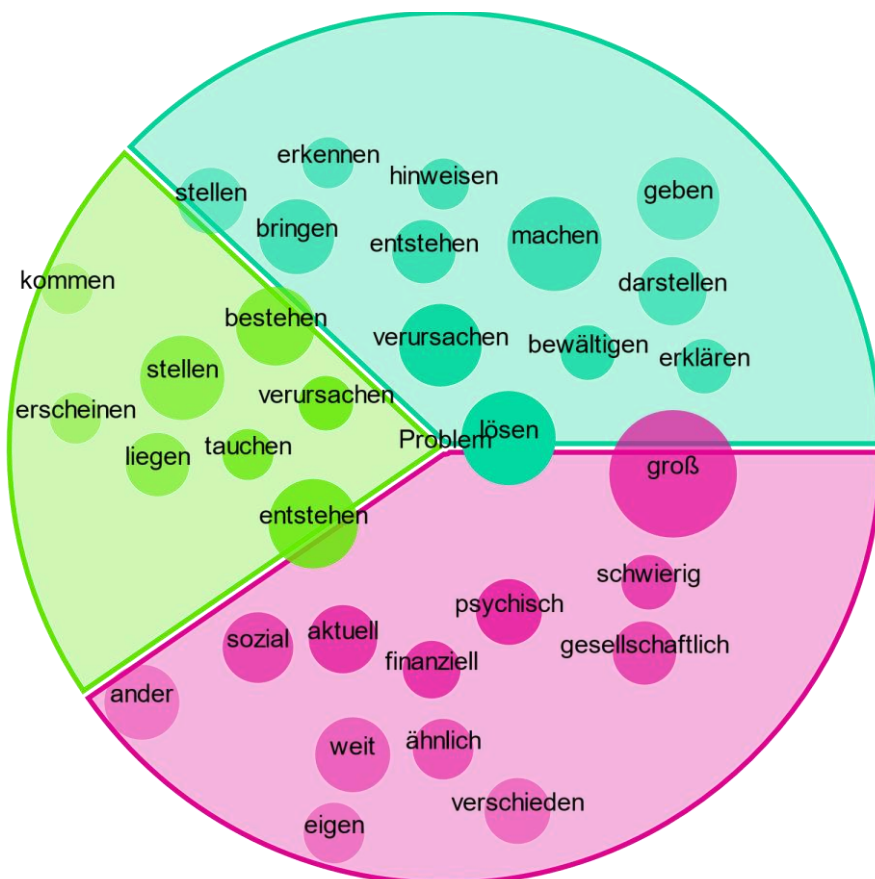


Abbildung 29: Problem L2 VERB+AKK VERB+SUBJEKT ATTRIBUT

<i>Problem</i>	Linguistik	Literatur	Σ	DP_{norm}
L1	521	277	798	0,39
L2	474	538	1012	0,61
Σ	995	815	1810	
X^2	(1, $N = 1810$) = 61,36, $p < 0,001$			

Tabelle 11: *Problem* – statistische Übersicht

Kommentar

Wenn man das Kreisdiagramm betrachtet, liegen viele Attribute vor, zu denen keine Beispielsätze im Profil gesammelt worden sind. Dies betrifft Adjektive wie z. B. *psychisch, gesellschaftlich, finanziell, syntaktisch, semantisch*, von denen man behaupten könnte, sie gehören der AWS nicht an. Allerdings soll gleichzeitig ihre Nützlichkeit nicht in Frage gestellt werden.

Wie das Ergebnis des Chi-Quadrat-Tests in der Tabelle zeigt, ergeben sich aus der Kombination von Sprache und Disziplin klar zu erkennende Gebrauchsunterschiede. In L1-Literatur wird kaum problematisiert, in der L2 jedoch sogar mehr als in der L2-Linguistik. Dies könnte daran liegen, dass die L1-Gruppe mehr Synonyme in Anspruch nimmt. Qualitative Unterschiede im Kollokationsbereich findet man nicht. Die einzige Ausnahme stellt wohl das Verb *bereiten*, das als Akkusativobjekt fast ausschließlich von der L1-Gruppe verwendet wird.

Aus Neugier habe ich auch das Profil der morphologisch verwandten *Problematik* abgerufen, wo die Verhältnisse geringfügig anders ausschauen:

	Linguistik	Literatur	Σ
L1	90	120	210
L2	248	131	379
Σ	338	251	589

Tabelle 12: Verteilung von *Problematik* im Korpus

Zum Teil ist dies an die Bedeutungsnuancen zurückzuführen. *Problem* steht für eine einzige, wie auch immer komplexe Schwierigkeit, wohingegen *Problematik* (bzw. *Problemstellung*) „umfangreicher, abstrakter, schwerer überschaubar [ist] und [] mehrere Probleme beinhalten oder mehrere ‚Ebenen‘ haben [kann]“ (Graefen & Moll, 2011, S. 43). Häufig wird das Wort auch eher im Sinne von *Themenfeld* benutzt, ohne dass dieses an sich automatisch Probleme birgt:

- (22) *Im theoretischen Teil wird die **Problematik** der Kollokationen mit Hilfe von Fachliteratur zusammengefasst.* (0010T)
- (23) *Auch eine Auseinandersetzung über die **Problematik** der Zuweisung in die Gattung des Kunst- und Volksmärchens soll in dieser Arbeit vorliegen.* (0029D)

VERB + *Problem* als Akkusativobjekt

bereiten, lösen, darstellen, sehen, beschreiben, es gibt, aufwerfen, umgehen, erkennen, ansprechen

*Im Rahmen einer Übersetzungskritik gibt es allerdings gewisse Punkte, die immer wieder **Probleme bereiten.*** (0120)

Eine umfangreichere Befragung, durch welche auch die Ergebnisse in den Unterteilungsgruppen stärker belegt werden würden, könnte dieses **Problem lösen**. (0144)

Die Zahlen legen allerdings offen, dass dies für die Lernenden so gut wie gar kein **Problem darstellt**, wenn die spanischen Nomina auf - o oder - a enden. (0002)

Oben haben wir gezeigt, dass es auch nicht- restrukturierende Infinitive geben kann, die kein unabhängiges tense aufweisen. Auf den ersten Blick **stellt** das ein **Problem dar**, immerhin sind wir bisher davon ausgegangen, dass nicht- restrukturierende Konstruktionen im Unterschied zu restrukturierenden funktionalen Kategorien projizieren und damit Satzwertigkeit aufweisen. (0140)

Andreas Bittner **sieht** das **Problem** des Grammatikunterrichts in der Schule ähnlich wie die bisher genannten Autoren und Autorinnen, er meint, dass vor allem die metasprachliche Kompetenz wichtig ist und gefördert werden muss. (0136)

Clahsen und Almazan (1998), Clahsen und Kollegen (2004) und Penke und Krause (2004) **sehen** diese **Probleme** als selektives Defizit der irregulären Flexion und aufgrund der intakten regulären Flexion gleichzeitig als Beweis für die Existenz eines Regelsystems und eines lexikalischen Systems. (0032)

Bachmair **beschreibt** hier die **Probleme**, denen BeraterInnen in einem "dysfunktionalen Arbeits- und Betriebsklima (vgl. Bachmair et al. 1994:53ff.) begegnen können. (0150)

Zusätzlich **gibt es** das methodologische **Problem**, dass nur positive Evidenz, aber keine negative Evidenz produziert werden kann, die Hypothesen widerlegen könnte. (0031)

Obwohl der Chat als gesprochene Sprache deklariert wird, haben die Logs rein schriftlichen Charakter, weshalb auch die Kategorisierung nach möglicher phonischer Realisierung schnell **Probleme aufwirft**. (0132)

Er **umgeht** das **Problem** mit den unterschiedlichen Phrasen, indem er annimmt, dass beide ursprünglich vom selben Typ waren (erweiterte VPs, immer noch ohne exakte Bestimmung) und in der spezifizierenden Phrase später redundantes Material phonologisch gelöscht wurde. (0133)

Auch Dürscheid **erkannte** dieses **Problem** und stellt die Frage nach der Einordnungsmöglichkeit der neuen Kommunikationsformen: Eingeordnet werden können nur Textsorten bzw. Diskursarten, nicht aber Kommunikationsformen. (0132)

Syntaxorientierte Modelle Constant (2012a) **spricht** zunächst die **Probleme** von Bürings CT-Bedeutung in Hinsicht auf CTmarkierte Fragen und Sätzen mit mehr als einer CT-Markierung an. (0031)

VERB + Problem als Subjekt

bestehen, sich ergeben

Das **Problem** der grammatikalischen Beschreibung der Elemente der AWS **besteht** darin, dass die Phänomene weder genau zum Aufgabengebiet der Lexik noch der klassischen Grammatik gezählt werden können. (0141)

*Auch der Prozess der Rückbildung kann für Immobilität nicht verantwortlich gemacht werden, da hierbei diverse **Probleme bestehen**.* (0113)

*Im Laufe der Auswahl **ergab sich das Problem**, nicht zweifelsfrei feststellen zu können, ob ein Element tatsächlich wissenschaftssprachlich ist.* (0141)

*Daraus **ergeben sich Probleme** der Prädiktabilität, da sich Diminutive und Basen unterschiedlich verhalten, und das führt dazu, dass, obwohl produktive Klassen eher zur Einverleibung neuer Wörter dienen, sich auch eine Tendenz herauskristallisiert, auch bzw. mehr opakere unproduktive Klassen zu verwenden.* (0126)

ATTRIBUT + Problem

groß, weiter, allgemein, spezifisch, methodisch

*Gerade aber die Vielzahl der Kategorien stellt WissenschaftlerInnen vor ein **großes Problem**.* (0152)

*BOAS (vgl. 2009: 74) nennt ein **weiteres Problem**, das mit der relativ hohen Zahl an angeblich deutschsprechenden TexanerInnen einhergeht.* (0124)

*Weiters sollen die grundlegenden Merkmale von Parenthesen auf sämtlichen grammatiktheoretischen Ebenen diskutiert, und die **allgemeinen Probleme** der Erklärung von Parenthesen innerhalb syntaktischer Modelle erläutert werden.* (0133)

*Deren Probleme, Wünsche und Hoffnungen sind denn auch vor dem Hintergrund **spezifischer Probleme** des Bürgertums zu sehen, die infolge gesellschaftlicher Veränderungen auftraten.* (0013)

*Den Grund dafür sieht sie in einem **methodischen Problem**, da noch kein systematisches Verfahren zur Erhebung von Aussagen über Textsorten etabliert sei, das relevante Resultate bringen würde (vgl. Adamzik 2008: 147).* (0123)

5.7 Ergebnis

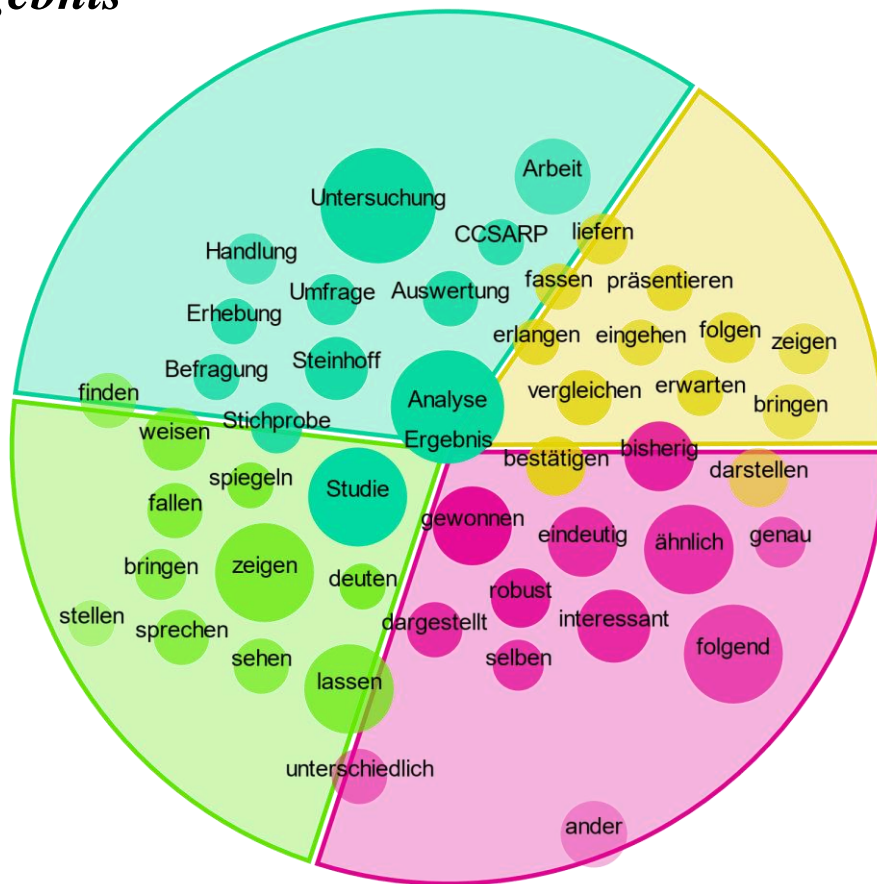


Abbildung 31: Ergebnis L1 **VERB+AKK** **VERB+SUBJEKT** **ADJ-ATTRIBUT** **GEN-ATTRIBUT**

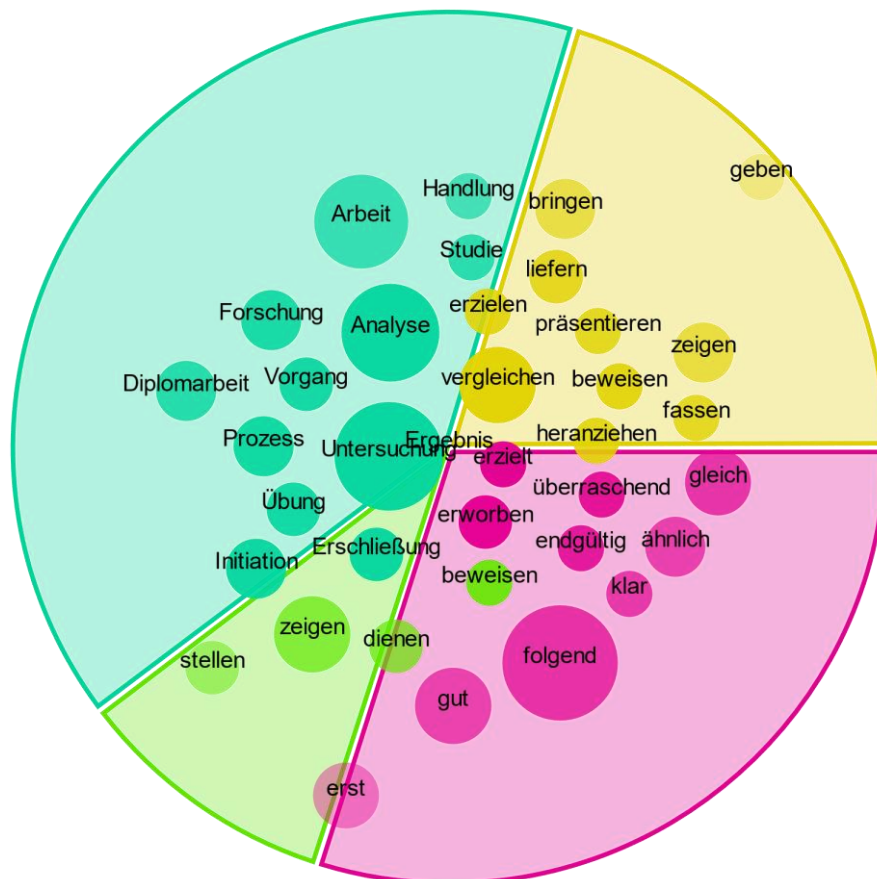


Abbildung 30: Ergebnis L2 **VERB+AKK** **VERB+SUBJEKT** **ADJ-ATTRIBUT** **GEN-ATTRIBUT**

<i>Ergebnis</i>	Linguistik	Literatur	Σ	DP_{norm}
L1	991	131	1122	0,51
L2	481	145	626	0,75
Σ	1472	276	1748	
X^2	(1, $N = 1748$) = 39,878, $p < 0,001$			

Tabelle 13: *Ergebnis* – statistische Übersicht

Kommentar

Für *Ergebnis* hat man auch die Genitivattribute miteinbezogen, denn sie repräsentieren sehr typische Beispiele der in der AWS gebrauchten Kombinationen. Nun zu der quantitativen Auswertung. Man würde in diesem Fall vermutlich auch ohne die induktive Statistik auskommen – die Kontraste sind auf schon bei flüchtiger Betrachtung der Frequenzen eindeutig. Das Substantiv kommt vorwiegend in linguistischen Arbeiten vor (RR = 1,803; CI [1,776; 1,84]), zugleich kommen zwei Drittel aller Treffer dem L1-Subkorpus zu (RR = 1,289; CI [1,245; 1,335]). Besonders gut schlägt sich die Kluft zwischen Literatur und Linguistik in dem DP -Wert nieder; die normalerweise niedrigere Dispersion von L2 ist bei *Ergebnis* radikal gestiegen und ist gewiss eben der relativen Absenz des Wortes in literaturwissenschaftlichen Arbeiten zuzuschreiben. DP wird aber deshalb weniger aussagekräftig, da es nur zwecks der Kontrastierung von L1 und L2 berechnet wurde und mithin die Disziplin ausblendet, die somit als Störfaktor wirken kann. Die folgende Tabelle präzisiert die Informationen zur Dispersion auf sehr einfache Art, indem sie die Anzahl der Arbeiten mit dem Wort *Ergebnis* in jedem der vier Korpusteile angibt.

	Linguistik		Literatur	
	abs.	rel./Subk.	abs.	rel./Subk.
L1	50	100%	39	78%
L2	74	82,2%	59	60%

Tabelle 14: Anzahl von Arbeiten mit *Ergebnis*

Ähnlich wie bei *Fall* scheint die methodische Herangehensweise in der Disziplin, mit den Gebrauchstendenzen einherzugehen – die Formulierung von Hypothesen und der Stellenwert von Produkten von Analysen sind mehr oder weniger eher die Domäne von der Linguistik als von der Literatur und beides setzt voraus, dass man auf Ergebnisse zu sprechen kommt. In der Sprachwissenschaft ist jedoch ein gewisser Zwang zu spüren, die Ergebnisse expliziter zu präsentieren. Hinsichtlich der Unterschiede zwischen dem L1- und L2-Subkorpus könnte man vielleicht versucht sein, zu schlussfolgern, die L2-Arbeiten sind weniger empirisch. Eine solche Behauptung könnte zwar richtig sein (und meiner Auffassung nach ist sie es auch³²), ihre Gültigkeit gleichwohl nur mit der Häufigkeit eines einzigen Lexems überprüfen würde m. E. den Bogen weit überspannen, denn die Erklärungen für ein so komplexes Phänomen wie Methodik können ja sehr vielfältig sein.

³² Vor allem im Verlauf der Erstellung von der ersten Version des Korpus mussten alle Arbeiten manuell getaggt und daher auch sorgfältig gelesen werden. Sie hatten immer ein eindeutig bestimmtes Ziel, das aber nicht notwendigerweise an Ergebnisse im traditionellen Sinne gebunden war.

VERB + Ergebnis als Akkusativobjekt

bestätigen, darstellen, vergleichen, zeigen, bringen, liefern, erlangen, erwarten, präsentiert, zusammenfassen, eingehen (auf), erhalten

*Die Ergebnisse zu diesen Konstruktionen spiegeln somit die bereits formulierten Erwartungen wider und **bestätigen** auch die **Ergebnisse** aus den freien Angaben. (0016)*

*Das Fazit versucht die prägnantesten und aussagekräftigsten **Ergebnisse** der Untersuchung **darzustellen**. (0116)*

*Ich habe ihre jeweiligen **Ergebnisse verglichen**, um eine größere Daten(=Text)grundlage zu haben und um zu sehen, welche Korpora aussagekräftiger sind, wo es Mängel gibt, und wo sich die Ergebnisse ergänzen. (0139)*

*Tabelle 17 im Anhang **zeigt** die zusammengefassten **Ergebnisse**. (0032)*

*Der versteckte Euphemismus von "bitca" erweist sich aufgrund der kulturellen Komponente daher als überraschend unübertragbar ins Deutsche, weswegen argumentiert werden kann, dass in Bezug auf die Wirksamkeit des Dialogs eine ganz andere, wenn auch nicht so ambitionierte, Übersetzungsstrategie eventuell ein besseres **Ergebnis gebracht** hätte. (0120)*

*Wie ist nun aber zu erklären, dass die Auswertung von Korpus II, der ja Dramen der Gegenwart beinhaltet, ein völlig anderes **Ergebnis liefert**? (0127)*

*Außerdem wurde das Ziel verfolgt, vergleichbare **Ergebnisse** zu **erlangen**, wofür sich ein standardisierter Fragebogen überaus eignet. (0002)*

*Dies wird dem Leser/der Leserin bereits ein erstes Bild darüber liefern, welche **Ergebnisse** in der korpusbasierten Analyse zu **erwarten** sein könnten bzw. auch wie die Methodik dahinter aussehen könnte. (0017)*

*Im Anschluss werden im Kapitel 11 die **Ergebnisse** der Erhebung **präsentiert**, wobei diese auch immer durch ein Resümee bzw. eine Interpretation ergänzt werden. (0025)*

*Ich berufe mich also nicht auf verschiedene Angaben der Sekundärliteratur und **fasse** deren **Ergebnisse zusammen**, auch weil es zu meiner speziellen Thematik sehr wenig an Literatur gibt. (0139)*

*Wir werden in der Conclusio noch näher **auf** die **Ergebnisse** der vorliegenden **Analyse eingehen**. (0120)*

*Parallel dazu wurde eine Internetfrequenzanalyse der Wörter durchgeführt, um noch genauere **Ergebnisse** zu **erhalten**. (0024)*

VERB + Ergebnis als Subjekt

(sich) zeigen, hinweisen auf, ausfallen, sprechen (für/gegen), aussehen, sich finden, hindeuten auf, widerspiegeln

*Ähnliche **Ergebnisse zeigten sich** bei der Befragung von MigrantInnen aus verschiedenen Ländern, die ich zusammen mit zwei StudienkollegInnen 2009 durchgeführt habe. (0150)*

Laut Cholin und Kollegen (2007) **zeigen die Ergebnisse**, dass der Aphasiker weniger korrekte Antworten bei jenen Verben produzierte, die eine besondere Gedächtnisleistung erforderten. (0032)

Die **Ergebnisse weisen** eindeutig **auf** einen Einfluss des common sense auf die sprachliche Entwicklung in den untersuchten Texten **hin**. (0141)

Im Falle der Nicht-Markierungen (Hypothese 3) **fielen** die Ergebnisse wie erwartet **aus**. (0002)

Diese **Ergebnisse sprechen dafür**, dass die Vorstellbarkeit durchaus ein Faktor ist, der sich auf die Verbflexion auswirken kann und deuten zudem auf die Wichtigkeit der Semantik bei diesem Prozess hin. (0032)

Allerdings **sprechen die Ergebnisse** von Tyler und Kollegen (2002b) **gegen** die alleinige Rolle der Phonologie bei der regulären Flexion, da selbst unter Berücksichtigung der phonologischen Komplexität längere Reaktionszeiten für gleich/ungleich Entscheidungen bei regulären Verben gemessen wurden. (0032)

Das **Ergebnis sieht** folgendermaßen **aus** (vgl. Anhang Abb. 159): Die 20 prozentuell sich am besten entwickelnden Bezeichnungen setzen sich zu beinahe gleichen Teilen aus österreichischen (in rot) und bundesdeutschen Ausdrücken (in blau) zusammen. (0036)

Die **Ergebnisse finden sich** in der nachfolgenden Tabelle: Für die Wörter aus den einzelnen Regionen sah die Verteilung für die Region A-west folgendermaßen aus: Mit dem dritten Schritt schloss die Untersuchung ab. (0024)

Das sohin erhaltene **Ergebnis deutet auf** die Verwendung der beiden Ausdrücke fast ausschließlich im Singular **hin** (vgl. Anhang Abb. 1 und 2). (0036)

Die **Ergebnisse** zu diesen Konstruktionen **spiegeln** somit die bereits formulierten Erwartungen **wider** und bestätigen auch die Ergebnisse aus den freien Angaben. (0016)

(Adjektiv)ATTRIBUT + Ergebnis

folgend, ähnlich, gewonnen, interessant, bisherig, eindeutig, dargestellt, aussagekräftig, überraschend, vorliegend, wichtig

Vergleicht man die Grammatiken der Neuzeit miteinander kommt man zu **folgendem Ergebnis**: Das starke Präteritum wird nur noch in der Wendung Er pflog der Ruhe. gebraucht (vgl. Duden 1998 u.a.). (0122)

Zu **ähnlichen Ergebnissen** kommt auch Richard Schrod, der die These aufstellt, dass sich der Grammatikunterricht seit dem Ende des Dritten Reiches stark verändert hat. (0136)

Im abschließenden Analyseteil werden die in der Untersuchung **gewonnenen Ergebnisse** ausführlich diskutiert. (0116)

An dieser Stelle sei noch einmal auf einige **interessante Ergebnisse** der Untersuchung verwiesen. (0033)

Daher wurde exemplarisch an einem Paar des Protokolls Nr. 10 eine qualitative Artikelanalyse durchgeführt, die einerseits **bisherige Ergebnisse** evaluiert, andererseits neue, ungenannte, aber nicht minder bedeutsame Aspekte beleuchten und damit zu einem konkludenten

Gesamtbild zum österreichischen Deutsch in den österreichischen Printmedien beitragen soll. (0036)

Im Grunde müsste man alle Erstglieder historisch auf diese Weise analysieren, um zu einem **eindeutigen Ergebnis** zu kommen. (0125)

Aufgrund der hier **dargestellten Ergebnisse** bezüglich der Relevanz der Satzklammer soll jener Aspekt auch nicht in die später vorgenommene Analyse der Nominalphrasen einfließen, da er wohl auch nicht zu einem engeren Komplexitätsbegriff beitragen kann. (0017)

Das Fazit versucht die prägnantesten und **aussagekräftigsten Ergebnisse** der Untersuchung darzustellen. (0116)

GILBERTS (vgl. BOAS 2009: 146) Untersuchung zu der Verteilung der Plosive [d] und [t] in New Braunfels liefern dabei ein **überraschendes Ergebnis**. (0124)

D.h., die Tatsache, dass das Verb einmal schwach war, hat sich in keinster Weise auf die hier **vorliegenden Ergebnisse** ausgewirkt. (0122)

Die Beantwortung der Fragen musste daher reflektiert und kritisch erfolgen, wodurch es unter anderem zu drei besonders **wichtigen Ergebnissen** kam: Raum in der Literatur ist nie völlig qualitätslos. (0107)

Ergebnis + (Genitiv)ATTRIBUT

Untersuchung, Analyse, Studie, Arbeit, Auswertung, Stichprobe, Erhebung, Befragung

Die **Ergebnisse der Untersuchung** werden schlussendlich Möglichkeiten eröffnen, komplexe Nominalphrasen zu beschreiben und ebenso zu generieren. (0017)

Die **Ergebnisse ihrer Analysen** zeichnen ein tristes Bild für die Stellung der Exilliteratur im Literaturgeschehen. (0020)

Die **Ergebnisse dieser Studie** gleichen sich zum Teil mit Daten, die aus Untersuchungen mit englischsprachigen Patienten hervorgehen (u.a. Cortese et al. 2006, Ullman et al. 1997). (0032)

In einem Abschlusskapitel sind einerseits die **Ergebnisse dieser Arbeit** zusammengefasst, andererseits Ideen für künftige Recherchen und wissenschaftliche Arbeiten aufgezeigt. (0030)

Hypothese 2 wurde durch die **Ergebnisse der Auswertung** von Korpus II widerlegt und kann in keinem Punkt bestätigt werden. (0127)

Beim **Ergebnis der Stichprobe** (vgl. Anhang Abb. 70) sei an die Vorbedingungen erinnert: Der hohe Wert von Filet in der Kategorie Nicht-Lebensmittel ist auf den Gebrauch des Ausdrucks in Konnex mit anderen als den oben spezifizierten Tieren zurückzuführen. (0036)

Im Anschluss werden im Kapitel 11 die **Ergebnisse der Erhebung** präsentiert, wobei diese auch immer durch ein Resümee bzw. eine Interpretation ergänzt werden. (0025)

Es wird zu beobachten sein, inwiefern die darin postulierten Thesen hinsichtlich der Unterschiede von Werte- und Moralvorstellungen auch in den **Ergebnissen der empirischen Befragung** Niederschlag finden. (0144)

VERB + Ergebnis als Dativobjekt

führen zu, kommen zu

*Die spontaneren Aufnahmen und das einmalige Zeigen der Kurzvideos scheinen zu authentischeren **Ergebnissen zu führen**. (0116)*

*Im anschließenden Kapitel wird analysiert, inwiefern und ob die hier beschriebenen Vorgangsweisen sowie die Erhebung mittels indirekter Methode **zu** entsprechend validen **Ergebnissen geführt** haben. (0016)*

***Zu ähnlichen Ergebnissen kommt** auch Richard Schrod, der die These aufstellt, dass sich der Grammatikunterricht seit dem Ende des Dritten Reiches stark verändert hat. (0136)*

*PANZIG muss also **zu dem Ergebnis kommen**, dass die Voraussetzungskette, wie er sie zuerst annahm, nicht aufrechterhalten werden kann. (0105)*

5.8 Ziel

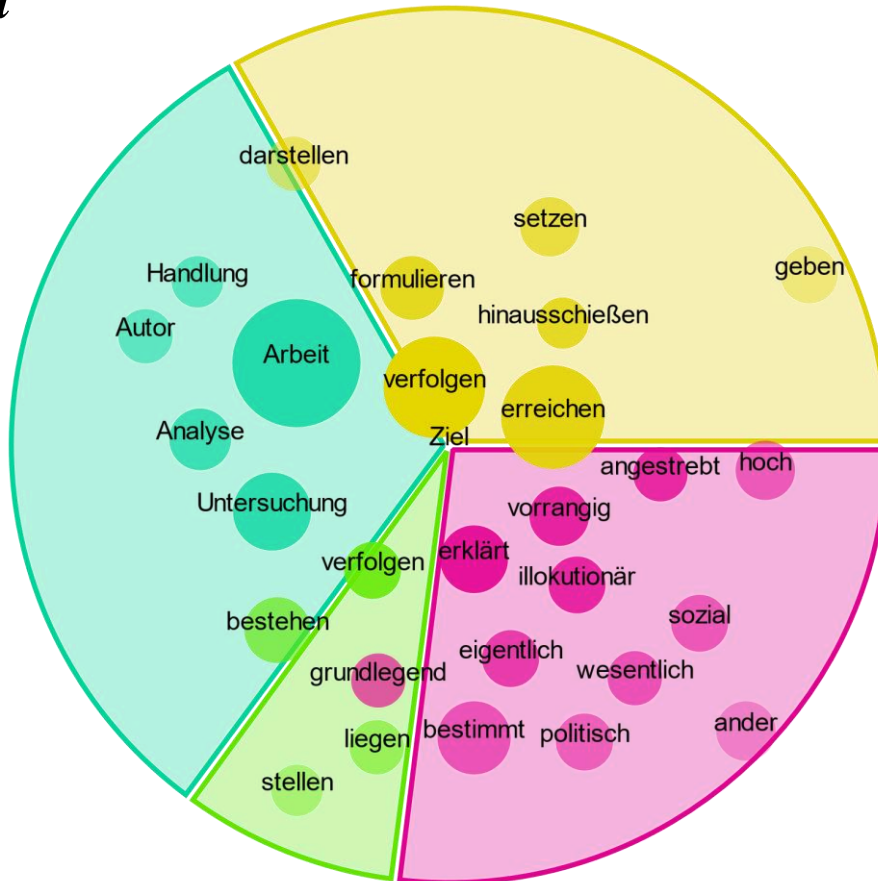


Abbildung 32: Ziel L1 VERB+AKK VERB+SUBJEKT ADJ-ATTRIBUT GEN-ATTRIBUT

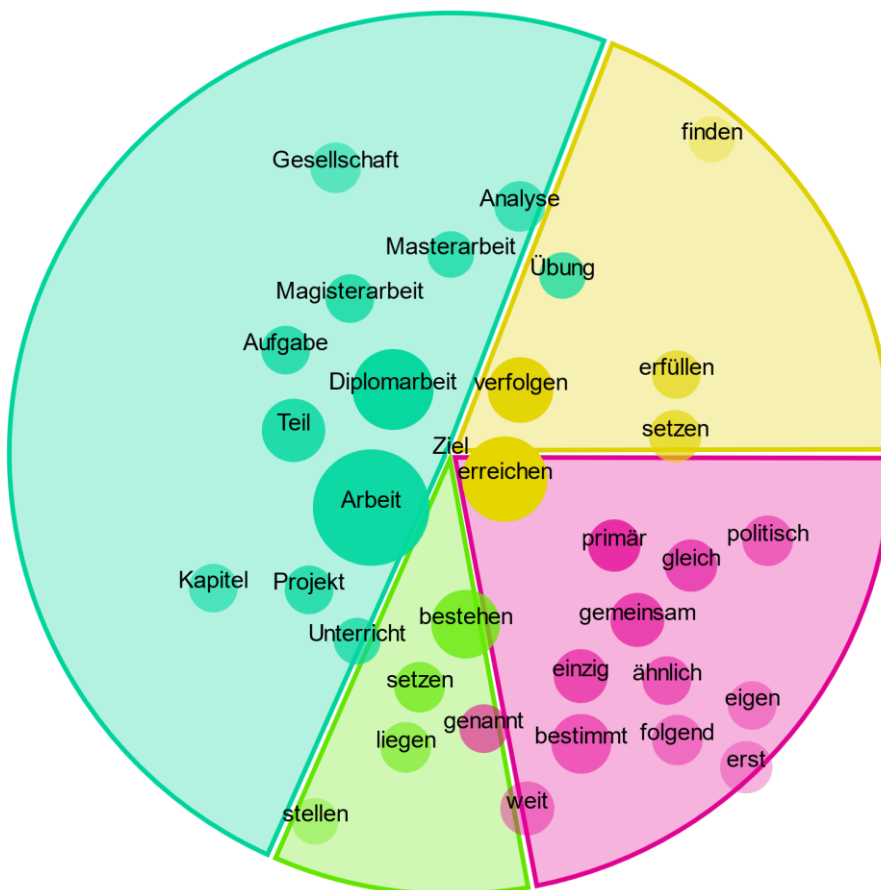


Abbildung 33: Ziel L2 VERB+AKK VERB+SUBJEKT ADJ-ATTRIBUT GEN-ATTRIBUT

<i>Ziel</i>	Linguistik	Literatur	Σ	DP_{norm}
L1	384	482	866	0,33
L2	448	406	854	0,56
Σ	832	888	1720	
X^2	(1, $N = 1720$) = 11,344, $p = 0,001$			

Tabelle 15: *Ziel* – statistische Übersicht

Kommentar

Obwohl der Chi-Quadrat-Test auf signifikante Differenzen hindeutet, sind die praktischen Unterschiede zwischen den zwei Korpusteilen nicht so groß (RR für Disziplin und Muttersprache beide nicht signifikant). Das Substantiv kommt in den meisten Dokumenten in der Einleitung viel zum Einsatz, um den Zweck der Arbeit vorzustellen, und manchmal auch im Fazit zur Zusammenfassung. Die Genitivattribute (*Arbeit, Untersuchung, Aufgabe* usw.) machen diese Funktionen besonders anschaulich.

In der Gruppe L2 gibt es unter den Verben mit *Ziel* als Akkusativobjekt auch das etwas seltsame Verb *hinausschießen*. Überraschenderweise sind alle vier Treffer im Korpus nicht einer Arbeit vorbehalten, sondern vier. In jenen Arbeiten handelt es sich in allen vier Fällen um Teil des eher umgangssprachlichen Ausdrucks *über das Ziel hinauschießen*.

- (24) *Zwar sind die Dadaisten ebenso wie andere europäische Avantgardebewegungen als erste Wegbereiter anzusehen, sie in diesem Rahmen unmittelbar voranzustellen, finde ich dennoch über das Ziel hinausgeschossen.* (0019D)
- (25) *Er entgegnete in seinem Buch "Die beliebtesten Rasenspiele" bereits 1893 der "Fußlümmelei" in ablehnender Form. Er spreche dem Autor Planck nicht den Ernst an der Sache ab, aber er habe... ..weit über das Ziel hinausgeschossen und einen Schlag ins Wasser geführt.* (0089D)

Als letzter Punkt ist noch die Korrelatkonstruktion kurz zu erwähnen. In dem umfangreichen Web-Korpus German Web 2013 (deTenTen13) ist *Ziel* das häufigste Substantiv mit einem *es*-Korrelat und einem nachgestellten Subjektsatz. Dass es sich bei Korrelaten um einen Zweifelsfall schlechthin handelt, bezeugen auch mitunter sogar widersprüchliche Äußerungen aus verschiedenen Grammatiken; Pittner & Berman (2015) sehen das Korrelat als „meist optional, jedoch insbesondere beim Subjektsatz **sehr häufig**“ (S. 131). Der erste Teil stimmt auch mit der Behauptung Mollicas (2010) überein, keine festen Tendenzen seien festzustellen. Im gleichen Atemzug fügt er jedoch hinzu, dass wenn „die *Nominalergänzung* vor der finiten Verbform [steht], erscheint das Korrelat **nicht sehr häufig**“ (S. 95; Hervorhebungen [Fettdruck] T.M.). Um mir einen flüchtigen Überblick zu verschaffen, habe ich mithilfe der folgenden CQL-Suchabfrage in German Web 2013 das Vorkommen der Konstruktionen mit und ohne Korrelat festgestellt.

```
[lemma = "Ziel"] [lemma = "sein"] [word = "es"] [tag != "V.*" & word != "zu"]{0,15} [word = "zu"] [tag = "V.*"] within <s/>
```

```
[lemma = "Ziel"] [lemma = "sein"] [word != "es" & tag != "ADJ.*"] [word != "zu|:" & tag != "V.*"] {0,15} [word = "zu"] [tag = "VINF.*"] within <s/>
```

Die Zahlen 195 466 (9,87 pM) bzw. 62 491 (3,15 pM) sprechen eher für die Tendenz von Pittner & Berman.³³

Das Ziel war (es), meinen Verdacht, dass Tschechen das *es*-Korrelat mit *Ziel* meiden, entweder zu bestätigen oder widerzulegen. So habe ich mit einer vereinfachten Version der obigen Korpusabfragen auch DESTWIKO durchsucht. Bei der Alternative mit dem Korrelat fallen 13 (1,9 pM) Treffer auf L1 und 11 (1,61 pM) Treffer auf L2.

(26) *Kleins Ziel ist es dabei, die klassifikatorische mit der funktionalen Textsortenanalyse zu verknüpfen.* (0142D)

(27) *Mein größtes Ziel ist es also, möglichst objektiv zu bleiben, sich nicht von der einseitig ausgerichteten Fachliteratur beeinflussen zu lassen und zu versuchen, die beiden Seiten zu verstehen.* (0220T)

Die Suche nach dem Gegenteil, wo das *es* unmittelbar nach *Ziel* verboten wird, ergab nur 2 (0,29 pM) Treffer für L1 und 13 (1,9 pM) Treffer für L2.

(28) *Seine selbst formulierten Ziele waren, das Schweigen zu brechen, und die Lücken der türkischen Geschichtsbücher zu füllen.* (0106D)

(29) *Ziel ist, eine größere Kompatibilität und Vergleichbarkeit der Hochschulsysteme vollständig zu verwirklichen und die internationale Wettbewerbsfähigkeit des europäischen Hochschulsystems zu verbessern.* (0145D)

(30) *Das Ziel ist, die Wortschatzliste um die Kollokationen zu ergänzen, die für das jeweilige Sprachniveau einschlägig sind.* (0248T)

Anhand der gewonnenen Ergebnisse lässt sich nicht plausibel zwischen den konkurrierenden Interpretationen entscheiden, nämlich dass die L2-Gruppe das Korrelat mit dem Substantiv nicht verwendet, oder dass die L1-Gruppe eben die Alternative ohne das Korrelat meidet.

VERB + Ziel als Akkusativobjekt

erreichen, verfolgen, (sich) setzen, es gibt, darstellen

Die Satire bedient sich nun, um ihr Ziel zu erreichen, verschiedener Mittel: Zu nennen wären hier als Beispiele der Witz, der Spott oder die Karikatur. (0100)

Die Leserschaft dieser Magazine verfolgte das Ziel, durch Umsetzung der präsentierten Trends ein Zugehörigkeitsgefühl hervorzurufen: "Die Biedermeiergesellschaft war noch nicht auf die Anonymität des bürgerlichen Individuums eingerichtet. (0092)

Deshalb hat sich diese Arbeit das Ziel gesetzt, dazu beizutragen, die Zugänge zu den Werken Drachs zu erweitern und die darin enthaltenen potentiellen Momente der Kritik an vermeintlich modernen Gesellschaften herauszuarbeiten. (0152)

Es muß immer Rätselhaftigkeit in der Poesie geben, und das Ziel der Literatur - andere Ziele gibt es nicht - ist es, die Gegenstände erstehen zu lassen, hervorzurufen, zu evozieren. (0108)

³³ So einfach wie es hier dargestellt wird ist die Problematik nicht. In beiden Quellen findet man Hinweise auf zahlreiche Faktoren, die sich auf das Auftreten von Korrelaten auswirken können, wie beispielsweise Komplexität der Satzglieder, Vorerwähtheit, Kontrast (vgl. Dudenredaktion, 2009) usw. Diese werden hier ignoriert.

Um einen engeren Komplexitätsbegriff zu ermöglichen, was schlussendlich das **Ziel** dieser Arbeit **darstellt**, muss man sich nun sozusagen eine Stufe tiefer begeben. (0017)

VERB + Ziel als Subjekt

bestehen in, liegen in

Die vorrangigen **Ziele** dieses ersten Teiles **bestehen** darin, werbespezifische Handlungstheorien in die allgemeine Sprechakttheorie zu integrieren und geeignete Werkzeuge zur Analyse von perlokutionären Kräften in Werbeanzeigen zu ermitteln. (0144)

Das Ziel der Einstellungsforschung **besteht** im Allgemeinen **darin**, eine Erklärung für menschliches Verhalten zu finden bzw. zukünftiges Verhalten genauer vorhersagen zu können (vgl. ebd.). (0025)

Ein weiteres **Ziel** **liegt** darin, die Ausformung zwischenmenschlicher Beziehungen und Beziehungsstörungen im Privat- sowie Berufsleben kognitiv und emotional zu erfahren und sich bewusst zu machen. (0129)

Wichtig zu unterscheiden ist hierbei jedoch, dass das **Ziel** dieses Satzes nicht sofort **in** der Verschmelzung mit Gott **liegt**, sondern im Hören des Wortes und der Lehre Gottes. (0105)

(Adjektiv)ATTRIBUT + Ziel

erklärt, vorrangig, grundlegend, weiter, primär, wichtig

Deshalb ist ein **erklärtes Ziel** der Analyse auch, traditionelle als auch individuelle Elemente der Künstlerautobiografie bzw. der Inszenierung des Künstlerlebens herauszufiltern, um zu einem genaueren Verständnis der Autobiografien dieser Zeit zu gelangen. (0091)

Auch wenn es nicht das **vorrangige Ziel** der Arbeit ist, die drei Sitzungen, aus denen das Korpus besteht, miteinander zu vergleichen, so kann man dennoch festhalten, dass es bei einer der Sitzungen zu einer auffälligen Abweichung kommt. (0114)

So formulieren es Perkuhn et al. und die Autoren bringen damit das **grundlegende Ziel** von Korpuslinguistik auf den Punkt: Die Sprache selbst ist Gegenstand der Untersuchung und mit der Sammlung von authentischen Sprachgebrauchsdaten soll realen sprachlichen Phänomenen nachgespürt werden. (0024)

Dies führt zu einem **weiteren Ziel** der Persönlichkeitsentwicklung im Adoleszenzroman, nämlich dem eigenständigen Aufbau eines sozialen Umfelds. (0028)

In jenen Vorstellungen werden 'transactional discourse types', welche fokussiert sind auf die optimal effiziente Informationsvermittlung, und 'interactional discourse types', deren **primäres Ziel** auf der Beziehungsarbeit beruht, unterschieden. (0115)

Neben der Herstellung von Öffentlichkeit kann das Verkaufen des eigenen Medienprodukts als **wichtigstes Ziel** publizistischer Medien betrachtet werden (vgl. Perrin 2015: 143). (0123)

Ziel + (Genitiv)ATTRIBUT

Arbeit, Untersuchung, Analyse, Autor

Ziel meiner Arbeit ist es, aufzuzeigen, wie Sprachplanung in Bezug auf das Belarussische und Slowenische durchgeführt worden ist bzw. wird und inwiefern sich diese im Laufe der Zeit auf den Status der beiden Sprachen ausgewirkt hat. (0026)

Diese Bewegung auf struktureller wie auch thematischer Ebene nachzuzeichnen, ist das **Ziel der** vorliegenden **Untersuchung**. (0088)

Ziel der Analyse ist, zu diskutieren, ob das bewegte Objekt oder die Bewegungsart größeren Einfluss auf die Verbwahl haben. (0116)

Abschließend ist festzuhalten, dass die **Ziele der Autoren** durchaus erreicht worden sind. (0021)

ANDERE KONSTRUKTIONEN

mit dem Ziel

Overt translations werden beispielsweise **mit dem Ziel** angefertigt, einen breiteren Rezipient/innenkreis zu erreichen, z.B. mittels stark gekürzten Übersetzungen von literarischen Werken. (0120)

Andere Überschriften wurden offenbar **mit dem Ziel** verfasst, potenzielle Leserinnen und Leser aufgrund rätselhafter oder ungewöhnlicher Formulierungen neugierig auf den restlichen Text zu machen (z.B. "Wien tanzt! (0123)

zum Ziel HABEN

Abschließend soll abermals darauf verwiesen werden, dass die Analyse Hofers Sprachverwendung weder die Denunzierung der politischen Person Norbert Hofer noch dessen Demaskierung **zum Ziel hat**. (0146)

Zu früh war es für ein solches Werk, das eine Aufarbeitung der Exilerfahrungen des Autors **zum Ziel hatte**. (0152)

5.9 Meinung

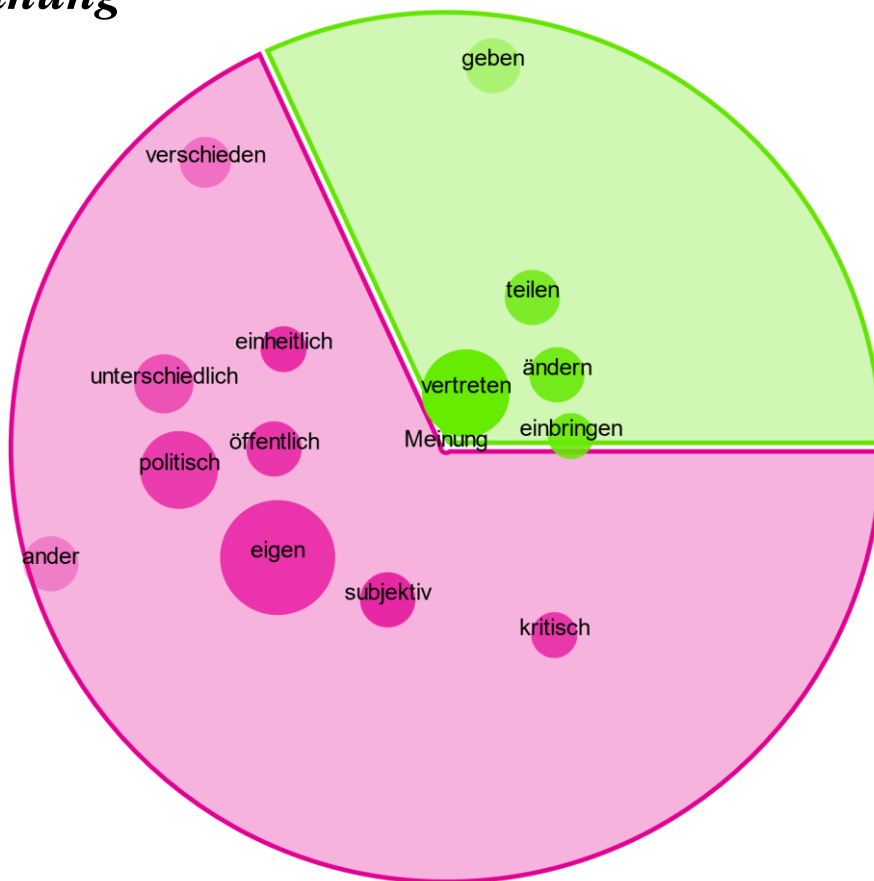


Abbildung 35: *Meinung* L1 **VERB+AKK** **ADJ-ATTRIBUT**

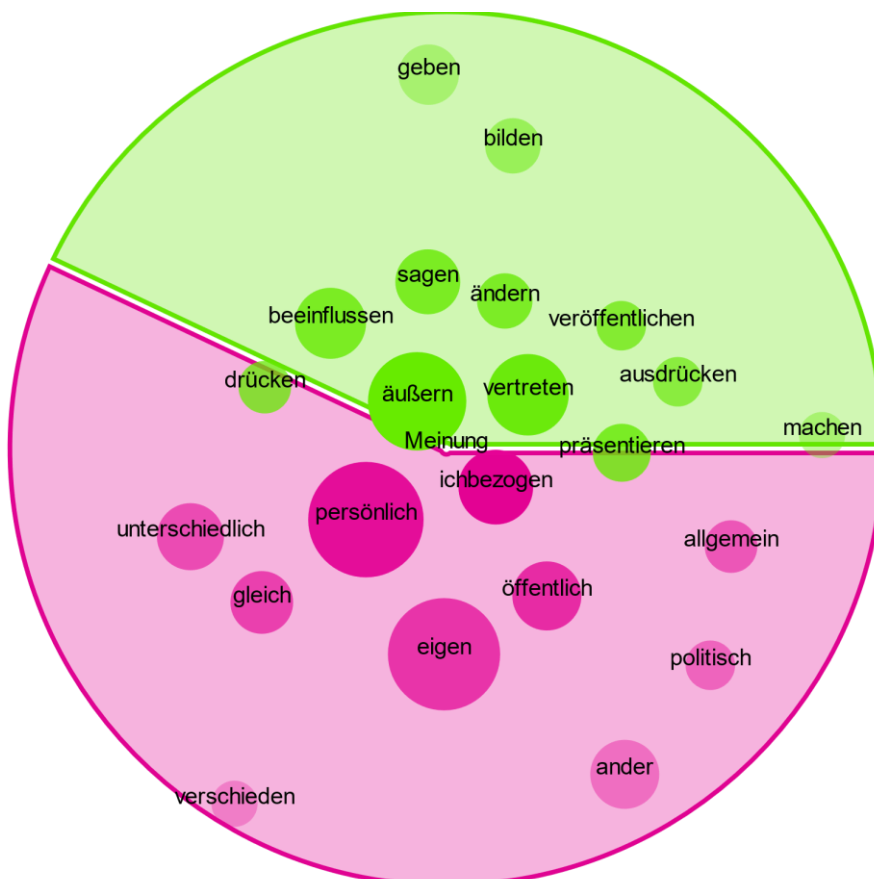


Abbildung 34: *Meinung* L2 **VERB+AKK** **ADJ-ATTRIBUT**

<i>Meinung</i>	Linguistik	Literatur	Σ	DP_{norm}
L1	289	275	564	0,42
L2	523	590	1113	0,60
Σ	812	865	1677	
X^2	(1, $N = 1677$) = 2,708, $p < 0,100$			

Tabelle 16: *Meinung* – statistische Übersicht

Kommentar

Bei diesem Substantiv kommen vor allem die Unterschiede zwischen Muttersprachlern und Nichtmuttersprachlern zum Ausdruck (RR = 0,675; CI [0,631; 0,722]). *Meinung*, genauer gesagt die syntagmatische Kette *meiner Meinung nach*, ist neben *ich glaube* und *ich denke* wohl einer der erst erlernten Sprachmittel, mit denen man seine Stimme einbringen kann und „die man eher aus kolloquialen Zusammenhängen kennt“ (Steinhoff, 2007, S. 20). So kann sich das Wort bzw. die feste Verbindung für die Nichtmuttersprachler als eine neutrale Bezeichnung etablieren und ihre stilistischen Besonderheiten in bestimmten Textsorten werden ignoriert.

Für *meiner Meinung nach* findet man im muttersprachlichen Subkorpus 71 Treffer (20,82 pM), im Nichtmuttersprachlichen 161 (46,84 pM). Die Werte korrespondieren fast perfekt mit der Verteilung des alleinständigen Substantivs. Steinhoff bietet sowohl eine Erklärung als auch eine wertvolle Kontextualisierung des Phänomens, die es angesichts der Limitationen im Fokus dieser Arbeit (nur studentische Wissenschaftssprache) wert sind, vollständig zitiert zu werden:

Die Verwendung solcher Muster kann jedoch nach Durchsicht der Korpora nicht einfach als „nicht wissenschaftlich“ bewertet werden. Mit dem ich kookkurrierende, für konzeptionell mündliche Meinungsäußerungen typische Verben kommen zwar hauptsächlich in den studentischen Texten vor, einige Belege finden sich jedoch auch bei den Experten. Ob derartige Ausdrücke jeweils angemessen sind, hängt sehr davon ab, ob die mit ihrer Hilfe vorgebrachte Position sachlich fundiert ist, hat aber auch damit zu tun, wie oft von diesen Ausdrücken Gebrauch gemacht wird. Die Experten verwenden solche Ausdrücke i. d. R. nur sehr selten, an exponierten Stellen ihrer Texte; einzelne studentische Schreiber hingegen gebrauchen sie recht häufig, was dazu führen kann, dass die Texte zu subjektiv und mithin nicht wissenschaftlich wirken. (Steinhoff, 2007, S. 20)

Festzuhalten ist, dass wenn Verbindungen wie *meiner Meinung nach* schon in L1 studentischen Texten streckenweise als nicht sachdienlich empfunden werden, kann dasselbe über ihren Gebrauch in L2 studentischen Texten geschlussfolgert werden, zumal sich ihre Zahlen auf fast das Doppelte belaufen.

VERB + *Meinung* als Akkusativobjekt

vertreten, teilen, es gibt

Eugenio Coseriu *vertritt* demgegenüber die **Meinung**, "daß [sic!] Gabelentz in Wirklichkeit einen sehr bedeutsamen Einfluß [sic!] ausgeübt hat, insbesondere auf de Saussure, und daß er daher als einer der Mitbegründer der modernen synchronischen Sprachwissenschaft angesehen werden" müsse. (0138)

Auch Ermen (1996) *teilt* diese **Meinung** und führt an, dass im Sprechakt 'aggressive Aufforderung' selten individuelle Formulierungen vorkommen. (0147)

*Auch sie erwähnt, dass es viele **Meinungen** zu diesem Thema **gibt** und dass diese nicht unterschiedlicher sein könnten. (0136)*

ATTRIBUT + Meinung

unterschiedlich, verschieden, einheitlich

*Diesbezüglich gibt es **unterschiedliche Meinungen** bzw. Ansätze, welche wir nicht im Detail behandeln brauchen, aber es ist natürlich auffällig, dass etwa das Thema in den meisten Fällen eine Anfangsstellung im Satz einnimmt. (0138)*

*Es wird akzeptiert, daß es **verschiedene Meinungen** über die Gestaltung des Zusammenlebens geben kann, und jeder hat die Möglichkeit, seinem Charakter und seinen Ansichten entsprechend zu leben, solange er damit dem Allgemeinwohl keinen ernsthaften Schaden zufügt. (0127)*

*Dadurch wird der Eindruck erweckt, sämtliche Österreicherinnen und Österreicher vertreten eine **einheitliche Meinung**. (0146)*

ANDERE KONSTRUKTIONEN

meiner Meinung nach (71)

*Grundsätzlich können in diesem Werk **meiner Meinung nach** zwei Personen als Täter in Betracht gezogen werden. (0110)*

*Diese Argumentation ist **meiner Meinung nach** jedoch nicht ganz schlüssig. (0032)*

SEIN der Meinung (31)

*Viele Probandinnen **sind der Meinung**, dass Kinder ohne die Beherrschung des "Dialekts" Nachteile im sozialen Bereich erfahren könnten. (0025)*

*Sie **war der Meinung**, jüngere Kinder könnten sehr wohl mit dem Thema Tod umgehen und das Werk auch richtig verstehen (vgl. Wiesli 2013, S. 21). (0085)*

5.10 Aussage

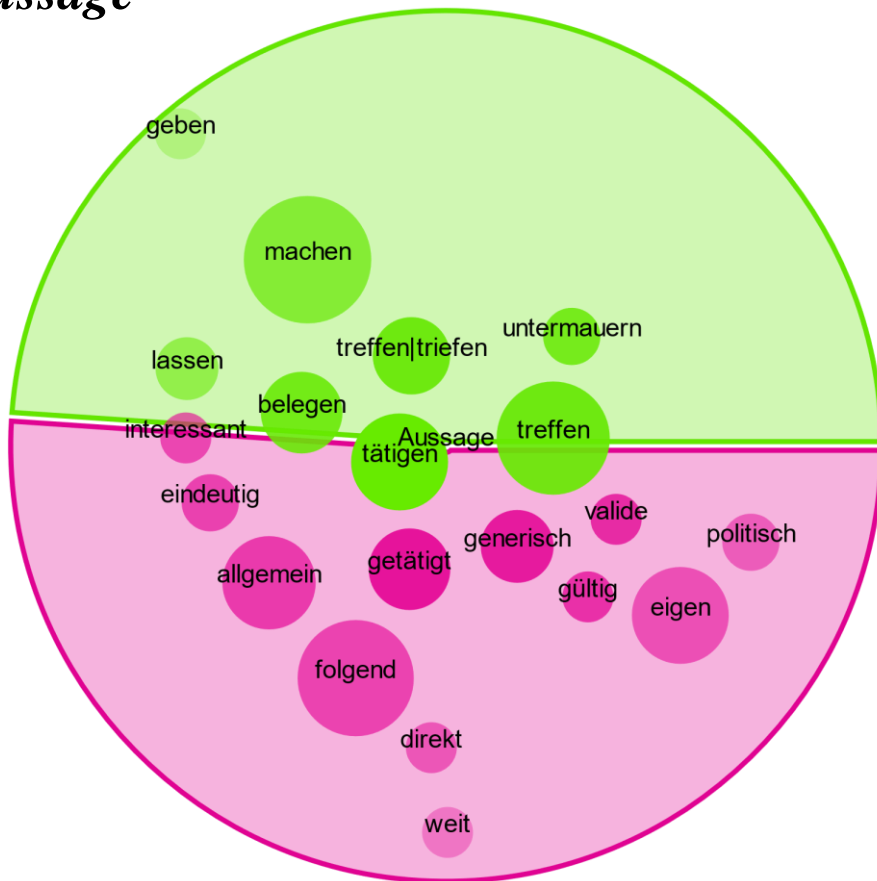


Abbildung 36: Aussage L1 VERB+AKK ADJ-ATTRIBUT

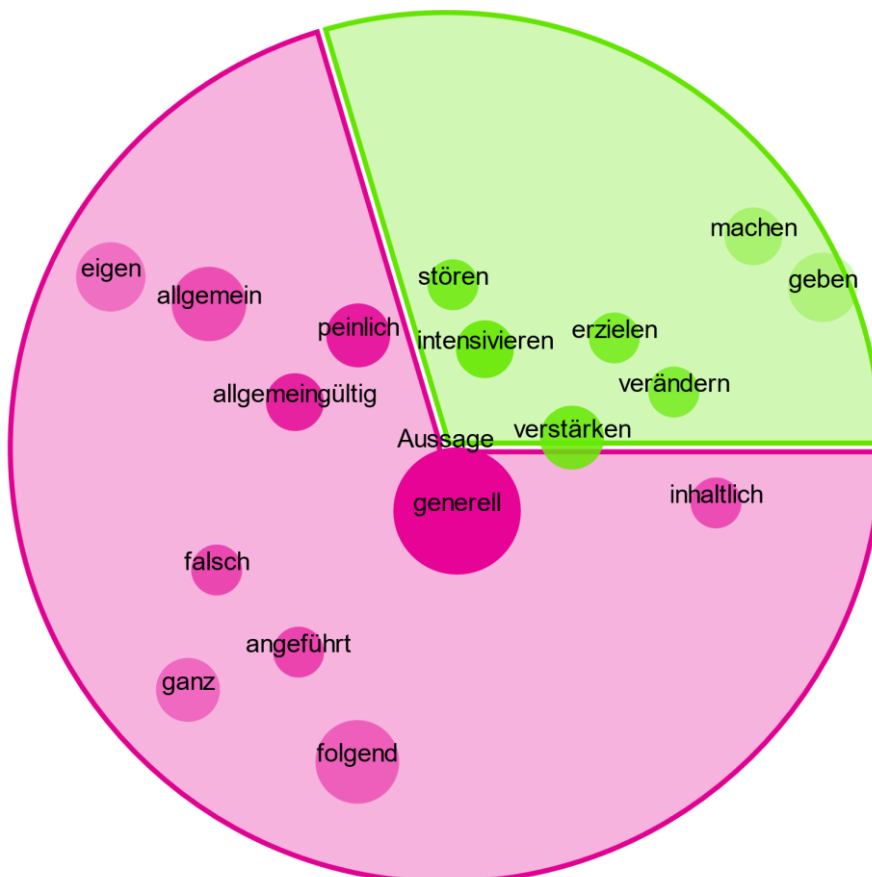


Abbildung 37: Aussage L2 VERB+AKK ADJ-ATTRIBUT

<i>Aussage</i>	Linguistik	Literatur	Σ	DP_{norm}
L1	475	319	794	0,38
L2	422	223	645	0,67
Σ	897	542	1439	
X^2	(1, $N = 1439$) = 4,758, $p = 0,029$			

Tabelle 17: *Aussage* – statistische Übersicht

Kommentar

Wie auch bei einigen anderen Substantiven, waren auch hier die Kollokationen aus anderen Gramrels unzuverlässig und würden der Analyse und dem Profil nur wenig bringen. Was an den Visualisierungen doch auffällt ist die Verteilung von Attributen und Verben mit Akkusativ. Auch wenn man die einfache, gemischte Kollokationsliste aufruft, stellt sich fest, dass in der L1-Gruppe weniger Verben vertreten sind als in der L2-Gruppe. Mit diesem Umstand, zusammen mit der überhaupt höheren Anzahl von *Aussage* mag die ungleiche Dispersion zusammenhängen. Zum Beispiel im L1-Subkorpus kommt das Kollokat *generell* 28-Mal vor, aber 19 Treffer davon stammen aus einer Arbeit.

Neben *Frage* ist auch *Aussage* ein vielseitiges Hilfsmittel für die Kontextualisierung, wie auch für metasprachliche Kommentare. Sie gibt dem Verfasser die Möglichkeiten, auf Zitate anderer zu verweisen und konkrete Sprachbeispiele zu topikalieren.

VERB + *Aussage* als Akkusativobjekt

machen, treffen, tätigen, untermauern

*Noch dazu ist die Stichprobe von fünf Wörtern für das Wortfeld "Küche" eindeutig zu klein, um seriöse **Aussagen machen** zu können.* (0024)

*Um jedoch **Aussagen** über eine ganze Disziplin **treffen** zu können, müssen die Untersuchungen weitergeführt werden und es muss ein größeres Korpus untersucht werden.* (0148)

*Diese Einschränkung verschafft ihm die Möglichkeit, auch radikale **Aussagen** zu **tätigen**, ohne sie in der Realität beweisen zu müssen.* (0105)

*Stöckl **untermauert** diese **Aussagen** wiederum mit einem Verweis auf die semiotischen Eigenschaften der beiden Zeichensysteme: Während Bilder aufgrund ihres "Realität simulierenden Charakters" und einer "biophysikalischen Vernetzung mit dem Gefühl" Emotionen ansprechen und Dinge visuell präsentieren können, dient Sprache aufgrund ihrer eindeutigeren und stärker festgelegten Bedeutungsstruktur in erster Linie zur präzisen Analyse und Vermittlung von Information.* (0144)

ATTRIBUT + *Aussage*

folgend, eigen, allgemein, getätigt, eindeutig, valid, (allgemein) gültig

*Folgende **Aussage** ist aus diesem Grund von Interesse: Ein Fragesatz kann auch eine Aufforderung ausdrücken.* (0030)

*Seine Leidenschaften zügelt Voigt nach **eigener Aussage** schon immer, er verzichtet freiwillig auf Alkohol (21) und hat auch kein Interesse an Trost durch käufliche Liebe (29).* (0018)

Eine **allgemeine Aussage** über die Häufigkeit der Verwendung bzw. die Gebräuchlichkeit einzelner Varianten zu treffen, ist davon abgesehen schwierig, da entsprechende Untersuchungen fehlen. (0033)

Der Begriff "Aussagesatz" betrifft den Inhalt der **getätigten Aussage**, deren Wahrheitsgehalt idealerweise bejaht oder verneint werden kann. (0030)

Zusammenfassung Auch in Bezug auf die interpretative Verfasser*innenreferenz kann keine **eindeutige Aussage** in Bezug auf die Erweiterung des Ausdrucksspektrums getroffen werden. (0141)

Bei all diesen Einschätzung ist freilich das niedrige Frequenzniveau zu bedenken, welches **valide Aussagen** erschwert. (0036)

Versucht man nun anhand jener Zahlen eine Korrelation zwischen Länge und Struktur der NPs abzuleiten, so stellt man schnell fest, dass es zwar Tendenzen gibt, man aber nicht von **allgemein gültigen Aussagen** sprechen kann. (0017)

5.11 Tatsache

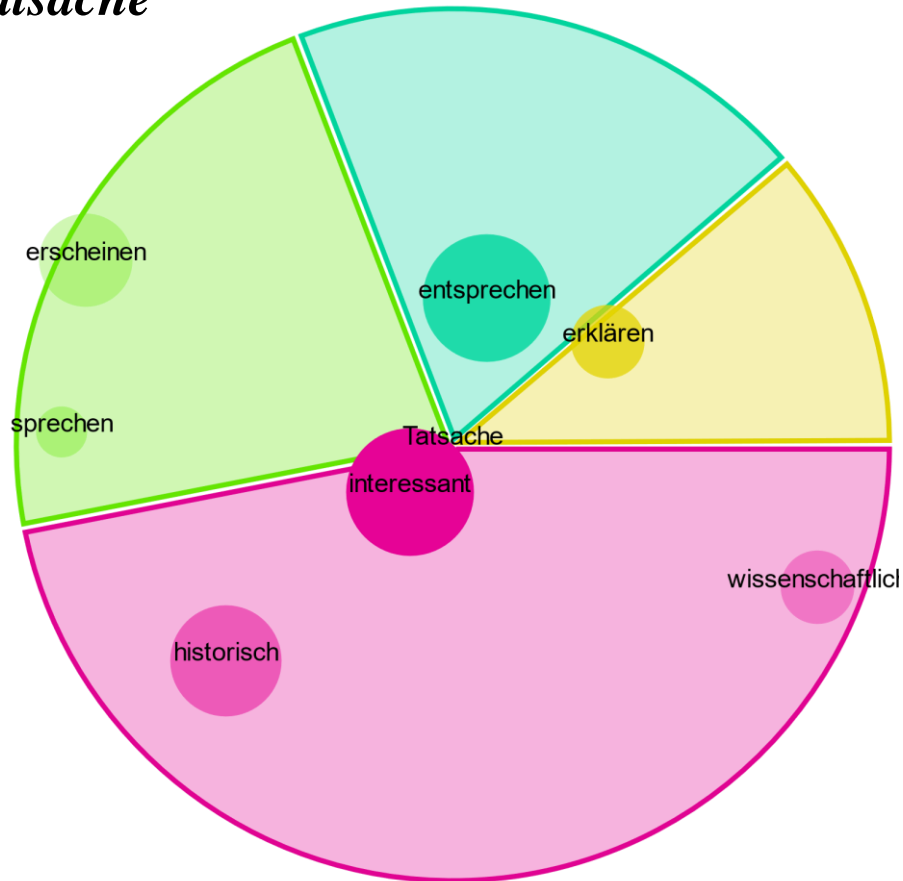


Abbildung 38: *Tatsache* L1 **VERB+AKK** **VERB+DAT** **VERB+SUBJEKT** **ATTRIBUT**

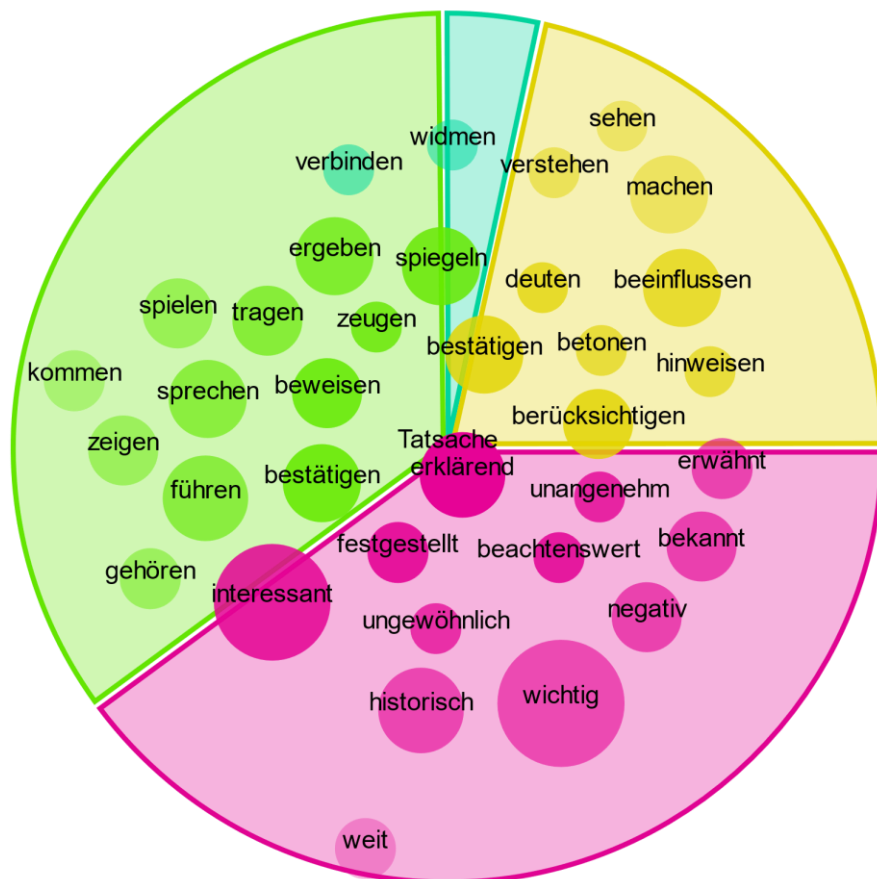


Abbildung 39: *Tatsache* L2 **VERB+AKK** **VERB+DAT** **VERB+SUBJEKT** **ATTRIBUT**

<i>Tatsache</i>	Linguistik	Literatur	Σ	DP_{norm}
L1	254	328	582	0,35
L2	334	445	779	0,64
Σ	588	773	1361	
χ^2	(1, $N = 1361$) = 0,080, $p = 0,777$			

Tabelle 18: *Tatsache* – statistische Übersicht

Kommentar

Was bei *Tatsache* besonders ins Auge sticht ist, dass Visualisierung aus dem L1-Subkorpus deutlich weniger Kollokate enthält als das L2-Subkorpus. Im Folgenden versuche ich kurz zu klären, warum dies der Fall ist. Das Diagramm unten zeigt die gleichen Gramrels, die Datenbasis entstammt jedoch einem Subkorpus des German Web 2013 Korpus, nämlich alle Dokumente, die zu der Zeit der Korpuserstellung unter der Domäne diplomarbeiten24.de eingespeichert waren (Größe – zirka 31 Mio. Wörter).

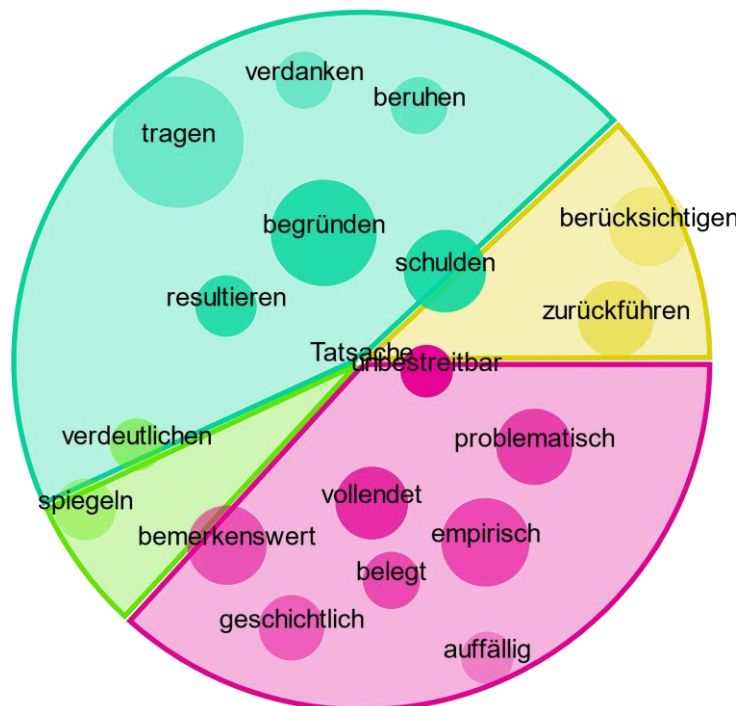


Abbildung 40: *Tatsache* (diplomarbeiten24.de) VERB+AKK VERB+DAT VERB+SUBJEKT ATTRIBUT

Es zeigt sich, dass obwohl das Subkorpus ungefähr 10-Mal größer als das L2-Teil von DESTWIKO (v0.2) ist, jede Gramrel beheimatet weniger als das Maximum (12 Kollokate). Außerdem sind die Häufigkeiten der entsprechenden Bigramme bestimmt nicht das Zehnfache von unserem L2-Subkorpus (siehe Tab. 19).

Attribut (modifier) von Tatsache				Verb mit Tatsache als Subjekt			
DESTWIKO (v0.2) [L2]		diplomarbeiten24.de		DESTWIKO (v0.2) [L2]		diplomarbeiten24.de	
3,4 Mio. W.		30,9 Mio. W.		3,4 Mio. W.		30,9 Mio. W.	
Lemma	n	Lemma	n	Lemma	n	Lemma	n
<i>wichtig</i>	14	<i>empirisch</i>	15	<i>führen</i>	7	<i>spiegeln</i>	7
<i>interessant</i>	12	<i>bemerkenswert</i>	12	<i>bestätigen</i>	6	<i>verdeutlichen</i>	5
<i>erklärend</i>	7	<i>problematisch</i>	11	<i>spiegeln</i>	6	<i>implizieren</i>	4
<i>historisch</i>	7	<i>vollendet</i>	10	<i>ergeben</i>	6	<i>resultieren</i>	4
<i>bekannt</i>	5	<i>geschichtlich</i>	8	<i>sprechen</i>	6	<i>deuten</i>	4
<i>negativ</i>	5	<i>belegt</i>	6	<i>beweisen</i>	5	<i>erschweren</i>	3
<i>festgestellt</i>	4	<i>unbestreitbar</i>	5	<i>tragen</i>	5		
<i>erwähnt</i>	4	<i>auffällig</i>	5	<i>spielen</i>	5		
<i>beachtenswert</i>	3	<i>geschuldet</i>	4	<i>zeigen</i>	5		
<i>unangenehm</i>	3	<i>belegbar</i>	3	<i>gehören</i>	4		
<i>ungewöhnlich</i>	3	<i>unumstößlich</i>	3	<i>kommen</i>	4		
<i>grundlegend</i>	3	<i>nachteilig</i>	3	<i>zeugen</i>	3		
<i>genannt</i>	3	<i>beobachtet</i>	3	<i>erweisen</i>	3		

Tabelle 19: Vergleich zwischen Gramrels in DESTWIKO (v0.2) [L2] und German Web 2013 [diplomarbeiten.de]

Nach diesem Vergleich scheint plausibel, dass *Tatsache* von den L2-Verfassern anders verwendet wird. Dass diese Verwendung fast frei von Kollokationen ist oder mehr Gebrauch von Konstruktionen aus anderen Gramrels macht, ist den Muttersprachlern nicht zu unterstellen. Sie könnte seine Ursache aber in der Phorik haben. Aus der ersten Version von DESTWIKO, wo alle SNV-Tags auch die Information über die Verweisrichtung (anaphorisch oder anaphorisch) beinhalten, wissen wir, dass von 146 Instanzen von *Tatsache* im L1-Subkorpus 22 anaphorisch und 124 kataphorisch gebraucht sind. Das L2-Subkorpus weist ein etwas unterschiedliches Verhältnis auf – 47 anaphorisch und 79 kataphorisch gebrauchte Exemplare. So subtil sie auch sein mag, ist hier die Verweisrichtung für die Bestimmung von Kollokationen auf zweierlei Weise relevant. Erstens wenn *Tatsache* einen nachfolgenden *dass*-Satz regiert, werden sehr oft die Verben „abgeschattet“, mit denen *Tatsache* in einem Valenzverhältnis steht, wie in den folgenden Beispielen:

- (31) *Da die **Tatsache**, dass die beiden Sprachen Genussprachen sind und Kongruenzerscheinungen zeigen, allerdings noch nicht sehr viel über die Beschaffenheit der einzelnen Genussysteme **aussagt**, bemüht sich dieser Abschnitt, die syntaktische Leistungsfähigkeit des Genus im Deutschen und Spanischen kontrastiv zu analysieren, um anschließend die Funktionalität der Kategorie zu beurteilen.* (0002D)
- (32) *Die **Tatsache**, dass die Ausgabe von 1795 keinerlei Angaben zum Autor bzw. der Autorin enthält, soll daher im Kontext der damaligen Verlegerstrategien nicht weiter **verwundern**.* (0004D)

Bedenkt man, wie sich die Ermittlung der Gramrels in Sketch Engine vollzieht (siehe 4.2.1), überrascht es kaum, dass verbale Kerne wie eben *aussagen* und *verwundern* hier nicht identifiziert werden können. Wenn dieser Befund auch auf die zweite Version extrapoliert wird, verliert man wegen der nachfolgenden *dass*-Sätze den Überblick über die eventuellen Kollokate. Zweitens da die Anzahl der anaphorisch gebrauchten (= leichter identifizierbaren)

Instanzen von *Tatsache* vermutlich höher ist die für die Nichtmuttersprachler und die absolute Anzahl von *Tatsache* in den Subkorpora gleichzeitig ebenfalls auf das Konto derselben Gruppe geht, gelangt man zu niedrigeren Zahlen von Kollokaten überhaupt. Andere Ursachen kann man zum Beispiel auf der Ebene der Konstruktionen sehen: [ADJ] + SEIN + *Tatsache* kommt in der L1-Basis relativ häufig vor und erhöht demnach die Kollokatenanzahl in der Attributgrammel:

- (33) **Wichtig ist aber die Tatsache**, dass es einige Denkmäler schon früher gab, und trotzdem werden sie nur von Becher erwähnt. (0042T)
- (34) **Interessant ist die Tatsache**, dass Sinti diese Arbeit immer im Sitzen gemacht haben. (0056T)
- (35) **Beachtenswert ist die Tatsache**, dass Weiskopf sich weniger das Leiden der Völker Europas (der Juden auch) und die innere Motivation der Deutschen, die sich von Nazismus missbrauchen ließ, interessierte. (0268T)

Aus den vorstehenden Ausführungen folgt, dass man ein deutlich größeres L1-Korpus bräuchte, um die Kollokationen von *Tatsache* repräsentativ erfassen zu können.³⁴ Zum Vergleich führe ich ein auf das gesamte German Web 2013 Korpus (16,5 Milliarden Wörter) basierendes Beispiel:

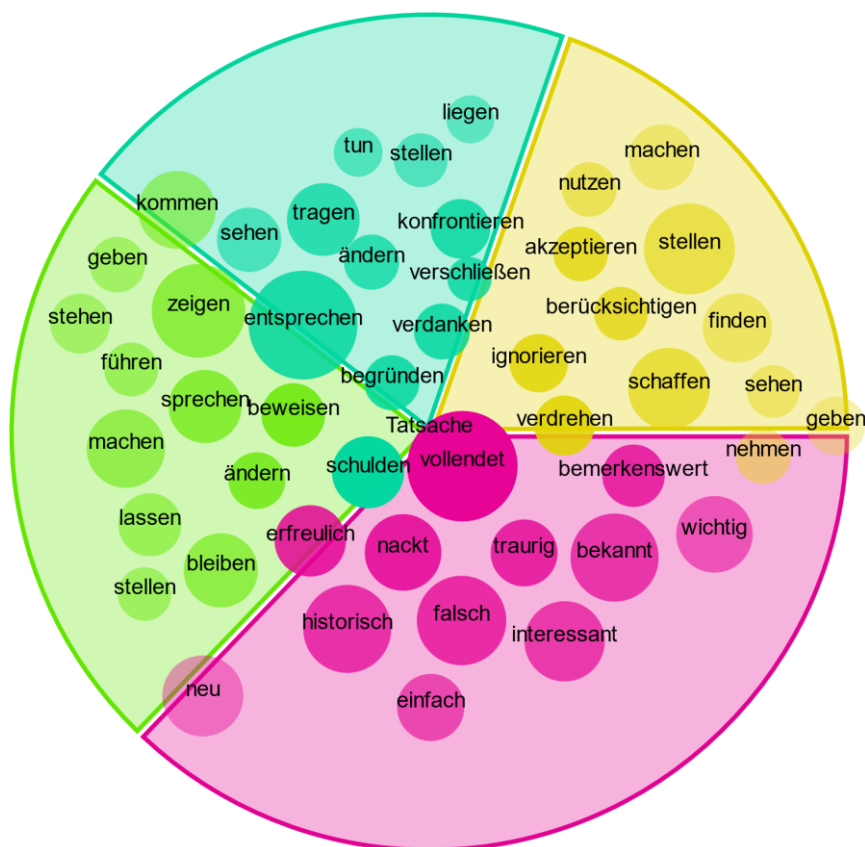


Abbildung 41: *Tatsache* (German Web 2013) **VERB+AKK** **VERB+DAT** **VERB+SUBJEKT** **ATTRIBUT**

³⁴ Auch in diesem Fall können aber die Kollokate nach dem *dass*-Satz nicht miteinbezogen werden. Unter diesem Gesichtspunkt könnte man sogar diese Darstellung als ebenfalls für nicht repräsentativ halten.

VERB + *Tatsache* als Akkusativobjekt

erklären

*Diese Ausführungen **erklären** die wichtige **Tatsache**, dass Einstellungen und Einstellungsäußerungen nicht einfach gleichgesetzt werden können. (0025)*

VERB + *Tatsache* als Subjekt

sprechen für/gegen, ansprechen

***Gegen** eine Lektüre als Märchen **spricht** die **Tatsache**, dass das Drama keinen versöhnlichen, sondern einen offenen Schluss hat. (0018)*

*Wie bereits weiter oben erwähnt wirken sich besonders jene Einstellungen, die in den ersten Jahren des Lebens erworben werden, besonders prägend aus. Diese **Tatsache spricht** ebenfalls **für** die Mütter als Probandinnen. (0025)*

*Todorov **spricht** die **Tatsache an**, dass Van Dines Regeln nur teilweise für jeden Kriminalroman gelten. (0110)*

VERB + *Tatsache* als Dativobjekt

entsprechen

*Entsprechend apokalyptisch sind auch die Bilder, die in den Romanen beschrieben werden, jedoch **entsprechen** diese Beschreibungen den historischen **Tatsachen**. (0106)*

ANDERE KONSTRUKTIONEN

aufgrund der Tatsache (26)

***Aufgrund der Tatsache**, dass sich ein und derselbe Begriff in vielen Sprachen auf beide Konzepte bezieht, kommt es sehr oft zu Verwirrungen. (0002)*

*Sie galt aufgrund mangelnder Recherche und **aufgrund der Tatsache**, dass sie in deutschen Verlagen publizierte, als Deutsche. (0084)*

5.12 Vorstellung

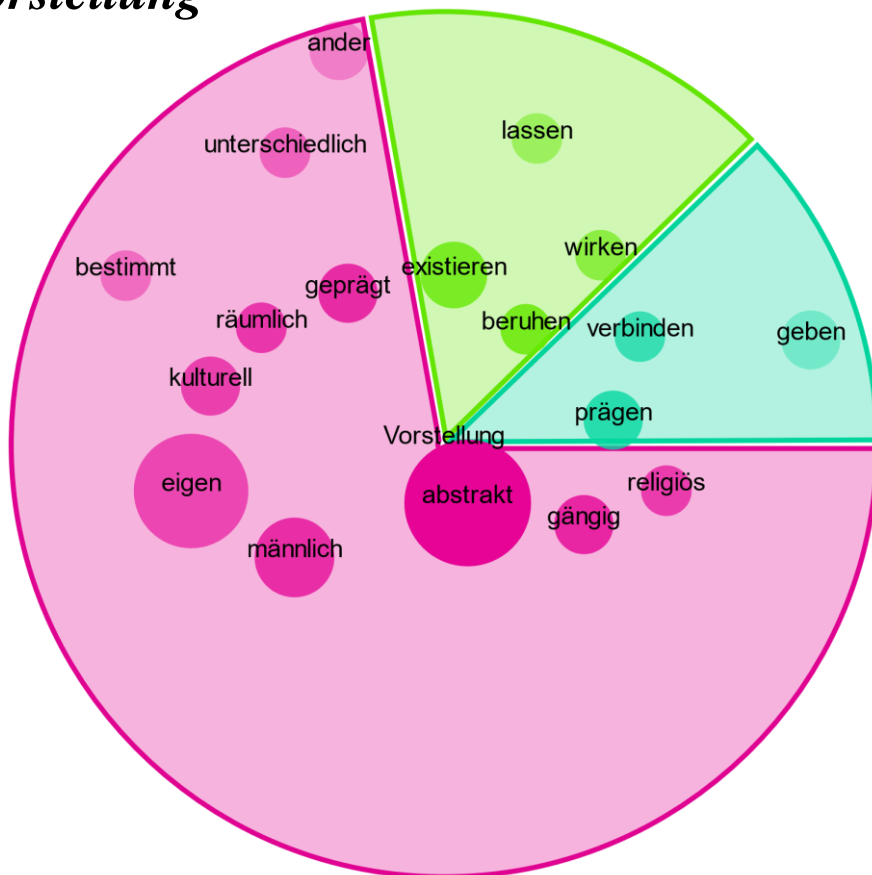


Abbildung 42: Vorstellung L1 VERB+AKK VERB+SUBJEKT ATTRIBUT

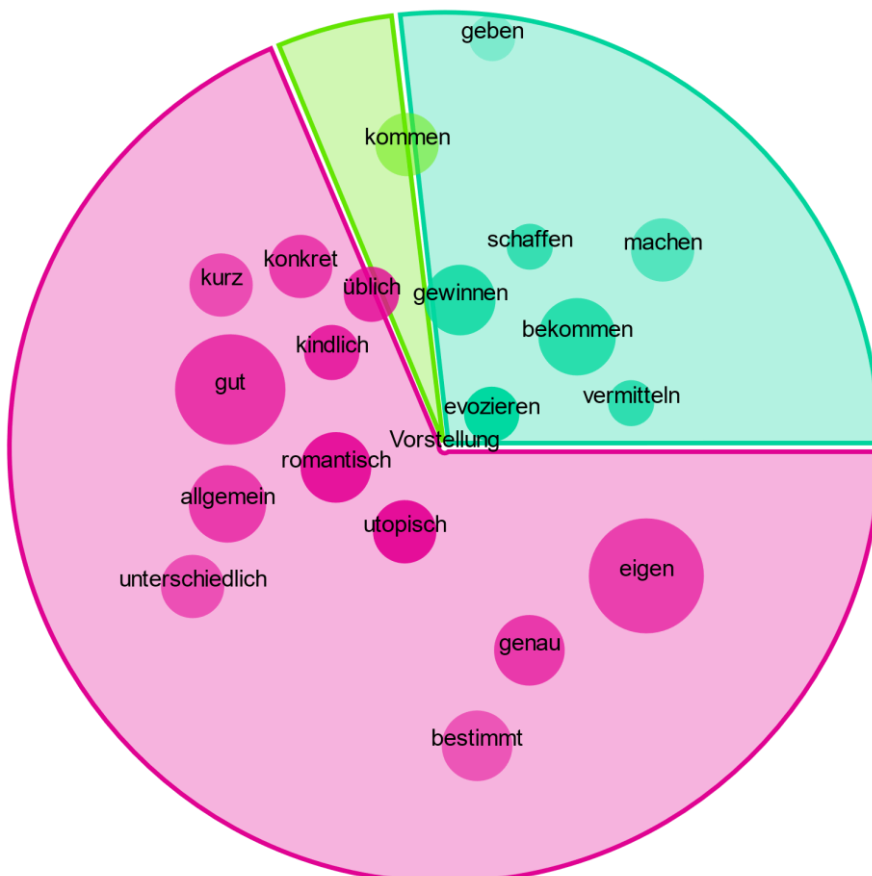


Abbildung 43: Vorstellung L2 VERB+AKK VERB+SUBJEKT ATTRIBUT

<i>Vorstellung</i>	Linguistik	Literatur	Σ	DP_{norm}
L1	156	381	537	0,37
L2	147	400	547	0,62
Σ	303	781	1084	
X^2	(1, $N = 1084$) = 0,637, $p = 0,425$			

Tabelle 20: *Vorstellung* – statistische Übersicht

Kommentar

Es liegt keine Interaktion von Muttersprache und Disziplin vor. Im Gegensatz dazu hat man mit einer klaren Gebrauchspräferenz zugunsten von Literatur zu tun (RR = 0,598; CI [0,544; 0,658]). Ähnlich dem Abstraktum *Gefühl* wird mithilfe von *Vorstellung* in den vielen Fällen auf Verfassers Deutungen innerlicher Zustände und Weltanschauungen der jeweiligen Gestalten Bezug genommen.

(36) *Ein derartiger Schritt zeigt auf, wie stark die autoritäre Wirkung Michael Kohlhaas' bis zur Auflösung des Haufens war. Seine Anhänger folgen seinen **Vorstellungen** von Recht und Gerechtigkeit.* (0004D)

Sporadisch benutzen es die Verfasser auch, um Ansichten aus sekundären Quellen in den Text zu integrieren.

(37) *O'Hara, der wie andere New Yorker Dichter der europäischen Kultur näher stehe, sollte nach Brinkmann eine neue Subjektivität illustrieren. Der/die KünstlerIn überwindet nach dieser **Vorstellung** die Grenze zwischen Kunst und Leben, wodurch literarische Konventionen hinsichtlich Stil, Form und Gestaltung gemieden werden.* (0019D)

In den linguistischen Arbeiten fungiert das Wort auch metatextuell im Sinne von Vorstellen.

(38) *Das zweite Kapitel diente zur **Vorstellung** von Schwerpunkten der deutschen Wortfolge.* (0054T)

Außerdem sieht man es in der Linguistik gemeinhin in entweder soziolinguistischen Arbeiten oder Arbeiten wo Probanden nach Meinungen befragt werden, um eine gewisse Distanz von den Äußerungen zu zeigen oder auf ihre Subjektivität hinzudeuten.

(39) *Es hängt mit der **Vorstellung** zusammen, dass was genau in dieser Kombination bekannt und gebräuchlich ist, müsse abgenutzt sein.* (0043T)

Es sei schließlich noch angemerkt, dass die Gesamthäufigkeiten schon relativ niedrig sind und die DP -Werte der Disziplin ungeachtet verhältnismäßig hoch. Die hier aufgezeichneten Tendenzen sind daher bestenfalls nicht mehr als disziplin- oder muttersprachbezogen, sondern eher als themaspezifisch zu verstehen.

VERB + *Vorstellung* als Akkusativobjekt

prägen, geben (von), verbinden

*Und jede Migration wirkt in anderer Weise auf das Leben der Migrant_innen ein, **prägt** ihre **Vorstellungen** und ihre Tätigkeiten.* (0159)

...gibt aber auch eine exhaustive Liste mit Ausnahmen dazu an, die ich im Folgenden besprechen möchte, um eine **Vorstellung** davon zu **geben**, wie viele (oder wie wenige) Wörter es gibt, die gegen Vennemanns Regeln verstoßen. (0149)

Mit Engeln **verbindet** man gemeinhin **Vorstellungen** von flügelbewehrten Wesen, die Gottes Glorie preisen und seine Botschaften übermitteln. (0108)

VERB + Vorstellung als Subjekt

existieren, beruhen auf, wirken

Besonders feste **Vorstellungen** bezüglich einer Sprachnorm **existieren** jedoch im Bereich der Kommunikation mit Kindern. (0025)

So **beruht** ihre **Vorstellung** von einer Beziehung zwar **auf** einem asymmetrischen Machtverhältnis zwischen den Geschlechtern, jedoch ist die Voraussetzung dafür eine freiwillige Unterwerfung der Frauen. (0152)

Diese **Vorstellung** von einer biologisch fundierten Volkszugehörigkeit **wirkt** als Gegenmodell zu einer pluralistischen, repräsentativen, demokratischen Gesellschaft und **äußerst[sic!]** sich in Ethnozentrismus, Ethnopluralismus sowie in der Ausgrenzung des "Fremden". (0146)

ATTRIBUT + Vorstellung

eigen, gängig, (irgendwie)geprägt, andere, unterschiedlich, bestimmt

Emma und Peter sprechen auch über ihre **eigenen Vorstellungen**, was nach dem Tod kommt. (0085)

Eine Begutachtung über **gängige Vorstellungen**, was maskulines Verhalten ausmacht, und wie dieses von den Femme fatale Frauenfiguren in der Literatur übernommen wird, ist in diesem Kontext allerdings annehmbar. (0098)

Aus den erwähnten typischen Elementen dieser Künstlerbiografien kann man auch implizit auf kulturell **geprägte Vorstellungen** einer Künstler- bzw. Autorenlaufbahn schließen, durch die die Autoren in der Erzählzeit geprägt waren. (0091)

Ihre völlig **anderen Vorstellungen** vom Leben und die Wucht, mit der Irmi dies deutlich macht, erschrecken Kajetan derart, dass er fluchtartig und weinend das Weite sucht. (0111)

Es enthält **unterschiedliche Vorstellungen** wie die vom Mann als Verführer und der Frau als Verführten wie umgekehrt die vom Mann als Verführtem und der Frau als Verführender. (0096)

In diesem Beispiel zeigt sich, wie **bestimmte Vorstellungen** über das Geschlecht auch durch Kleidung geprägt werden. (0152)

5.13 Aufgabe

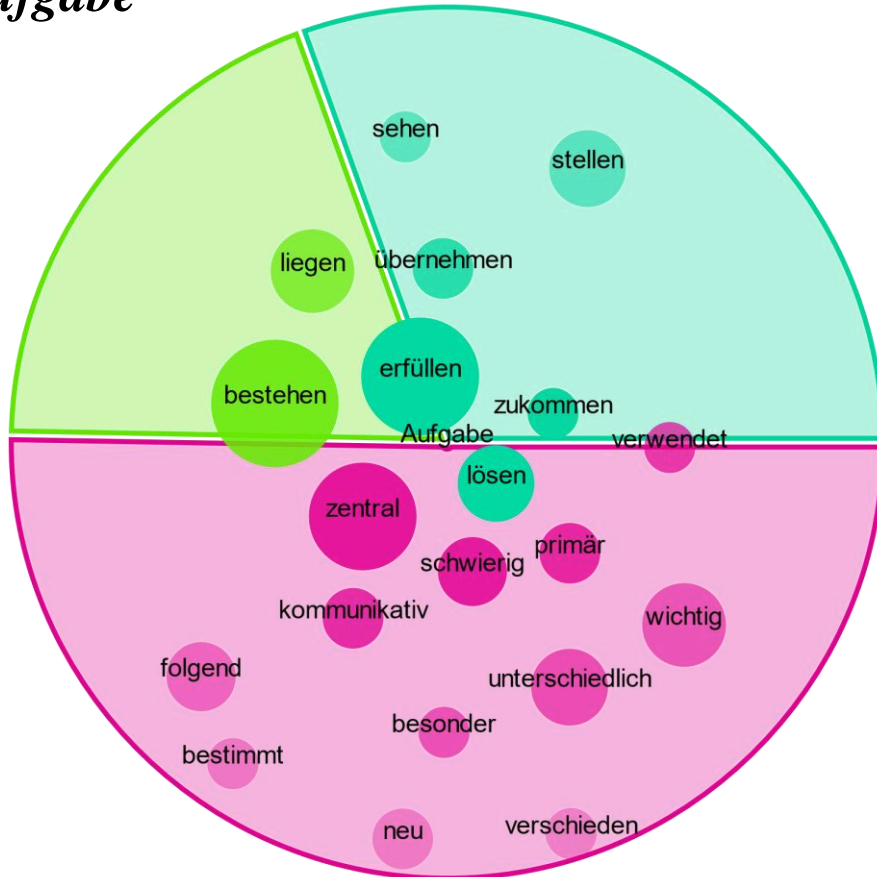


Abbildung 44: Aufgabe L1 VERB+AKK VERB+SUBJEKT ATTRIBUT

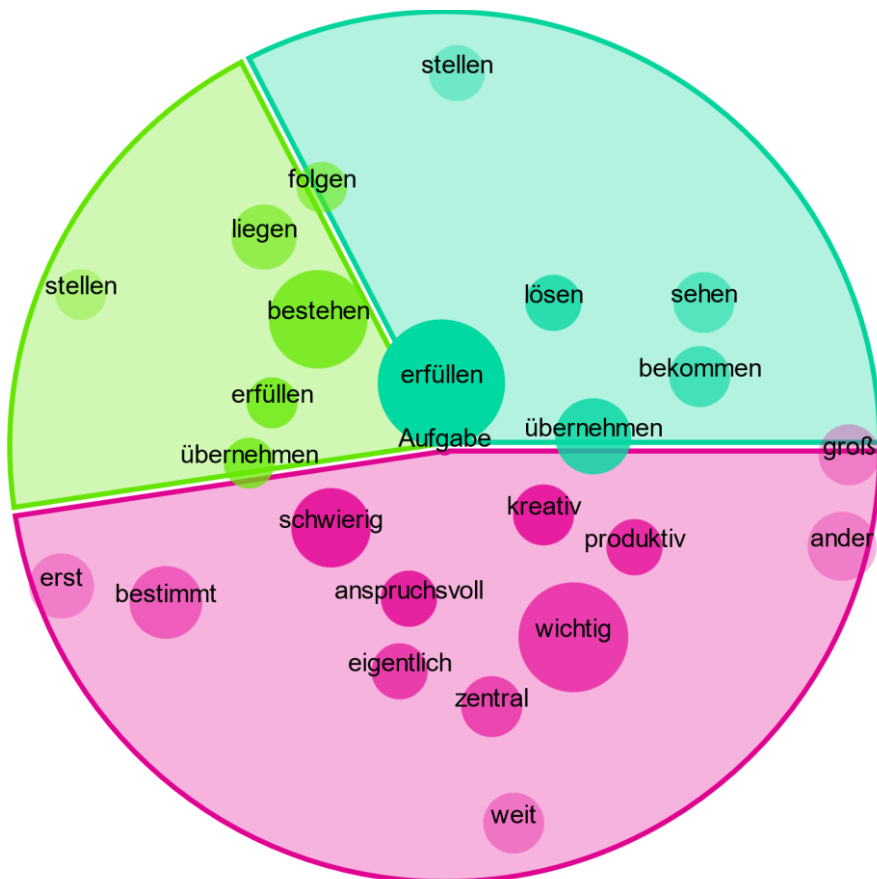


Abbildung 45: Aufgabe L2 VERB+AKK VERB+SUBJEKT ATTRIBUT

<i>Aufgabe</i>	Linguistik	Literatur	Σ	DP_{norm}
L1	214	229	443	0,38
L2	312	293	605	0,67
Σ	526	522	1048	
X^2	(1, $N = 1048$) = 1,089, $p = 0,297$			

Tabelle 21: *Aufgabe* – statistische Übersicht

Kommentar

Die Okkurrenz von *Aufgabe* ist ein wenig prominenter in den L2-Texten (RR = 0,849; CI [0,791; 0,911]). Dem soll aber keine große Bedeutung beigemessen werden. Für die Begründung ist eher auf die Stichprobenentnahme zurückzugreifen; einige der linguistischen Arbeiten befinden sich nämlich an der Schnittstelle zwischen Didaktik und Sprachwissenschaft und in diesen Arbeiten werden fast sämtliche Fälle von *Aufgabe* im Sinne von Lernaufgabe gebraucht. Ansonsten dient das Lexem dazu, Funktionen und Zwecke von Konzepten sowie Leuten (Gestalten) zu übermitteln.

VERB + *Aufgabe* als Akkusativobjekt

erfüllen, lösen, stellen (vor), übernimmt

*Das Märchen eine psychologische **Aufgabe erfüllen**, ist bereits hinlänglich bekannt; zu diesem Thema gibt es bereits einige (Diplom)Arbeiten.* (0095)

*Gemein ist den beiden Formen, dass eine Figur eine **Aufgabe lösen** muss, wobei sie von magischen Requisiten umgeben ist.* (0029)

*Oft ist es auch ein Erwachsener, der vor eine **Aufgabe gestellt** ist, diese lösen muss und Unterstützung bei der Lösungsfindung benötigt.* (0095)

*Die Rede ist lediglich von seiner Mutter, welche noch immer an seiner Seite ist und die **Aufgaben** seiner verstorbenen Frau **übernimmt**.* (0029)

VERB + *Aufgabe* als Subjekt

bestehen in, liegen in

*Die **Aufgabe** des Mentors **besteht** allerdings **darin**, auf den Bewährungsraum vorzubereiten, aber nicht einzugreifen, da die Adoleszenz einen Prozess darstellt, der selbstständig bewältigt werden muss.* (0028)

*Diese pathetische Gedankenstimme des Kardinals über den Führer lässt sich fast als Gebet interpretieren, in dem der Führer als ein von Gott gesandtes Werkzeug bezeichnet wird, dessen **Aufgabe darin liegt**, Gottes Werk auf der Erde zu verrichten.* (0109)

ATTRIBUTE + *Aufgabe*

zentral, wichtig, unterschiedlich, schwierig, folgend, primär, neu

***Zentrale Aufgabe** wird sein, dass Sie in einem Satz beschreiben sollen, was die Person auf dem Video macht" (vgl. BOWERMAN [u. a.] 2004: 12).* (0116)

*Eine **wichtige Aufgabe** der Autobiografie ist nach Peter-André Alt die Erfassung der Schriftstellerexistenz als eigenes Kunstwerk.* (0091)

Deutlich zeichnen sich in den Zeitschriften die **unterschiedlichen Aufgaben** ab, die der Krieg von der Jugend gefordert hat. (0013)

Eine der **schwierigsten Aufgaben** innerhalb der Valenzgrammatik ist es, Aktanten von freien Angaben zu unterscheiden, die meisten Probleme treten im Bereich der Adverbialbestimmungen auf. (0136)

So gibt die GfdS auf ihrer Website **folgende Aufgaben** an: ... (0117)

Der Detektiv, dessen **primäre Aufgabe** es wäre, die bestehende Ordnung zu erhalten, setzt sich in diesem Fall nicht das einzige Mal über das Gesetz hinweg und folgt seinem subjektiven Verständnis von Recht. (0111)

Überforderung Sowohl der Vater als auch die Kinder sind mit der Situation und den damit verbundenen **neuen Aufgaben** oft überfordert. (0085)

ANDERE KONSTRUKTIONEN

sich zur Aufgabe MACHEN (12)

Die Politolinguistik versteht sich seines Erachtens als Teil der deskriptiven Linguistik, die es **sich zur Aufgabe gemacht** hat, sprachliche Phänomene ohne jedwede Wertung zu beschreiben und zu erklären. (0146)

Um eine Erleichterung für diese Gruppe zu schaffen, haben es **sich** viele Linguisten **zur Aufgabe gemacht**, herauszufinden, wie Muttersprachler einer Einzelsprache bei der Genuszuordnung vorgehen. (0002)

5.14 Tendenz

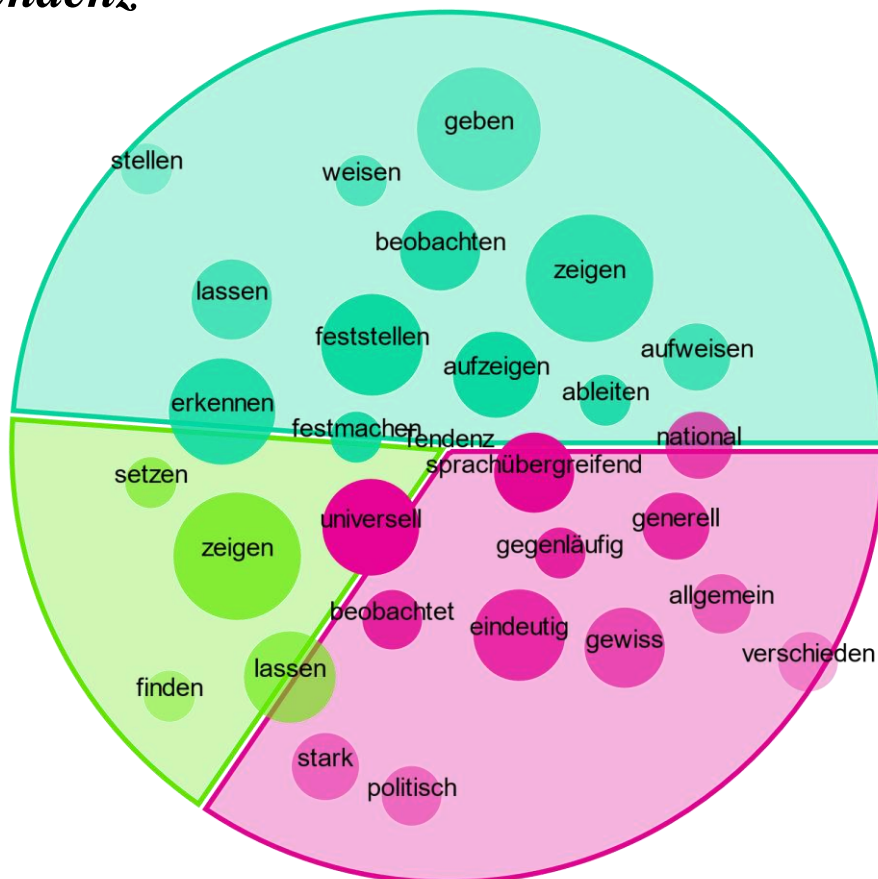


Abbildung 47: Tendenz L1 **VERB+AKK** **VERB+SUBJEKT** **ATTRIBUT**

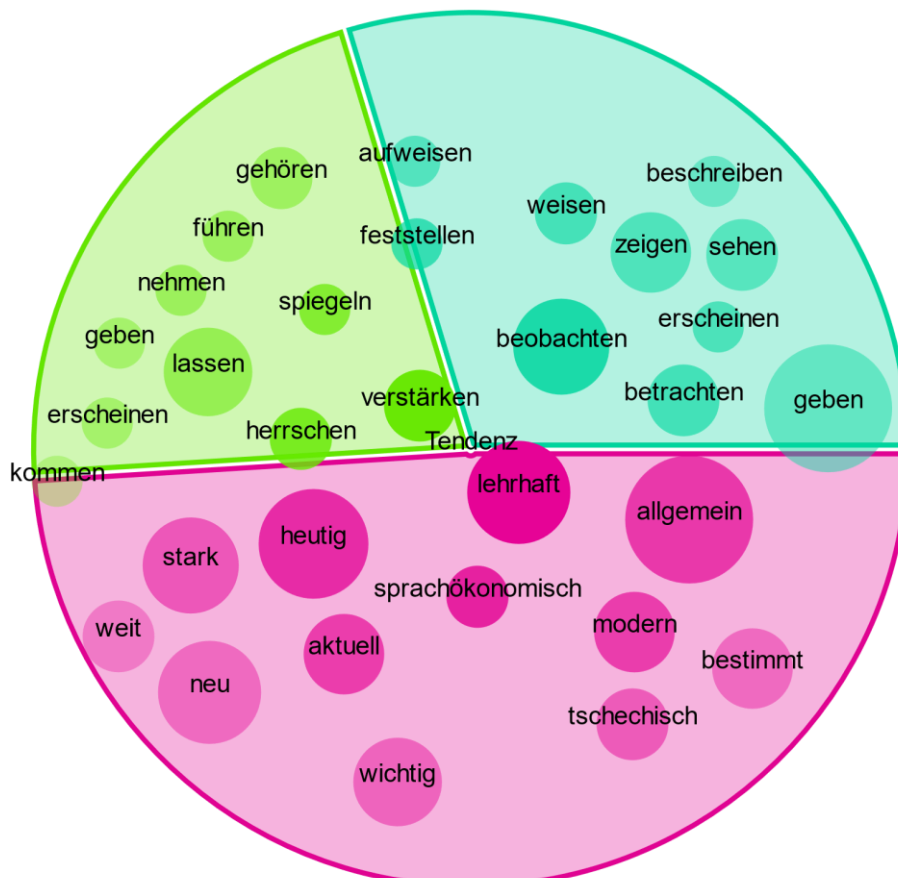


Abbildung 46: Tendenz L2 **VERB+AKK** **VERB+SUBJEKT** **ATTRIBUT**

<i>Tendenz</i>	Linguistik	Literatur	Σ	DP_{norm}
L1	262	147	409	0,45
L2	313	146	459	0,78
Σ	575	293	868	
X^2	(1, $N = 868$) = 1652, $p = 0,199$			

Tabelle 22: *Tendenz* – statistische Übersicht

Kommentar

Die unausgewogene Verteilung zwischen Linguistik und Literatur macht *Tendenz* zu einem eher disziplinspezifischen Substantiv (RR = 1,418; CI [1,352; 1,487]). Aufgrund des hohen DP -Werts könnte sein Auftreten vielleicht sogar von der Methodologie oder dem Thema abhängig sein. Damit man über Tendenzen sprechen kann, muss es ein Kontinuum geben, das im Idealfall auch auf die eine oder andere Art quantifizierbar ist. So konzeptualisiert stellt eine Tendenz grundsätzlich eine durch begrenzte Anzahl konträrer Belege abgeschwächte Tatsache dar. Aus allen Funktionen, die Hedges innerhalb des wissenschaftlichen Diskurses annehmen können, fallen die meisten Beispiele von *Tendenz* unter „sachlich notwendige Vorsicht“ (Graefen, 2000, S. 7). Diese wird als angebrachte Skepsis aufgefasst, die sich aus den Wissenschafts- bzw. Forschungsprinzipien herleitet. Anders ausgedrückt wird *Tendenz* wegen Geltungs- oder Objektivitätsanspruchs seitens des Verfassers benutzt und nicht beispielsweise aus Höflichkeit.

VERB + *Tendenz* als Akkusativobjekt

zeigen, es gibt, erkennen, feststellen, aufzeigen, beobachten, aufweisen, festmachen, ableiten

*Finale Sequenzen scheinen außerdem weniger sprachübergreifende **Tendenzen** zu zeigen als initiale.* (0130)

*In der Jugendsprache **gibt es** jedoch die **Tendenz**, sprachliche Superlative aller Art zu bilden.* (0125)

*Basierend auf GILBERTS (1972) Fragebögen, um eine bessere Vergleichbarkeit zu gewährleisten, lässt sich beim Vergleich mit GILBERTS (1972) Ergebnissen keine eindeutige **Tendenz erkennen.*** (0124)

*Im Hinblick auf die zwischenmenschlichen Beziehungen lässt sich in beiden Städten keine einheitliche **Tendenz feststellen.*** (0104)

*Es konnten hier lediglich **Tendenzen aufgezeigt** werden, die in gewisser Weise auch dem entsprechen, was in der vorhergehenden theoretischen Diskussion bereits aufgegriffen wurde.* (0017)

*Jedoch lässt sich hier, besonders anhand der N-Wort-Debatte, die **Tendenz beobachten**, entweder gewisse Ausdrücke von Anfang an zu umschreiben, zu verharmlosen oder den Text bzw. dessen Übersetzung nachträglich zu korrigieren.* (0090)

*Die Neuinterpretation der Führerfigur im Roman, die unverkennbare **Tendenzen** zur Hitler-Darstellung dieser Zeit **aufweist**, muss somit als politisches Bekenntnis des Autors begriffen werden.* (0109)

Über Gründe kann im Zuge dieser Arbeit nur spekuliert werden, da die Stichprobe für ein Wortfeld nur bei fünf Wörtern lag und damit klare **Tendenzen** in diesem Bereich nicht genau **festgemacht** werden können. (0024)

In diesem ersten Kapitel werden die absoluten Längen der NPs berechnet, Häufigkeiten festgestellt und mögliche **Tendenzen abgeleitet**. (0017)

VERB + Tendenz als Subjekt

sich zeigen, sich fortsetzen

Außerdem **zeigt sich** hier die typische **Tendenz** unzuverlässiger, geistig zerrütteter Erzähler, ihre Mitmenschen zu verdächtigen, böse Absichten zu hegen, ohne für diese Unterstellungen ein überzeugendes Motiv nennen zu können. (0087)

Diese **Tendenz setzt sich** im 20. Jh. **fort**, wenn das Präteritum auch in beiden Jahrhunderten auf Grund der Unsicherheit beim Ablautvokal gemieden wird. (0122)

ATTRIBUT + Tendenz

universelle, eindeutig, gewiss, generell, stark, beobachtet, allgemein, etwaig, negativ

Dies zeigt dass es zwar **universelle Tendenzen** gibt, wie natürliche Laute mit sprachlichen Lauten imitiert werden, das Ergebnis in den Einzelsprachen aber trotzdem sehr unterschiedlich aussehen kann. (0130)

Auch bei den weniger häufiger auftretenden Akkusativ-, Dativ- und auch Genitivergänzungen konnten keine **eindeutigen Tendenzen** beobachtet werden. (0017)

Wie schon bei den Kopfwörtern erwähnt, scheint es eine **gewisse Tendenz** zu geben, Initialwörter bei nicht eindeutigen Begriffen zu verwenden. (0132)

Eine **generelle Tendenz** des Übergangs von "nichtschwach" zu "schwach" ist prognostizierbar, ihr können sich letztlich wohl nicht einmal die "nichtschwachen" Verben der Suppletionsdomäne entziehen. (0122)

Für aktuelle Instanzen ist jedenfalls eine **starke Tendenz** in Richtung deutschsprachiger Namen gegeben. (0132)

Auch können **beobachtete Tendenzen** in einzelnen Regionen mit größeren Stichproben besser belegt oder widerlegt werden. (0024)

Im Sprachgebrauch ist aber häufig Vermischung der Formen intrans./trans. anzutreffen. Dabei ist die **allgemeine Tendenz** beobachtbar, dass die distinkte Symbolisierung durch die Übertragung der schwachen Flexion der trans. Verben auf ihre jeweiligen starken intrans. Pendanten abgebaut wird (vgl. bleichen). (0122)

In weiterer Folge finden **etwaige Tendenzen** und Auffälligkeiten, die sich aus der Stichprobe zwecks Extrapolation ergeben haben, Besprechung. (0036)

Die argumentative Ebene wird im Laufe der Schreibkarriere bei allen Teilnehmer*innen von der metakommunikativen "überholt" und zeigt immer eine **negative Tendenz**. (0041)

ANDERE KONSTRUKTIONEN

EINE/DIE Tendenz zu

*Es kann daraus also zumindest **eine Tendenz zu** einer wahrheitsgemäßen Beantwortung der Frage nach der subjektiven Dialektkompetenz geschlussfolgert werden. (0016)*

*Auch wenn die unterschiedliche Belegzahl berücksichtigt wird, bleibt **eine Tendenz zu** weniger sprachlichen Realisierungen pro Beleg. (0141)*

5.15 Versuch

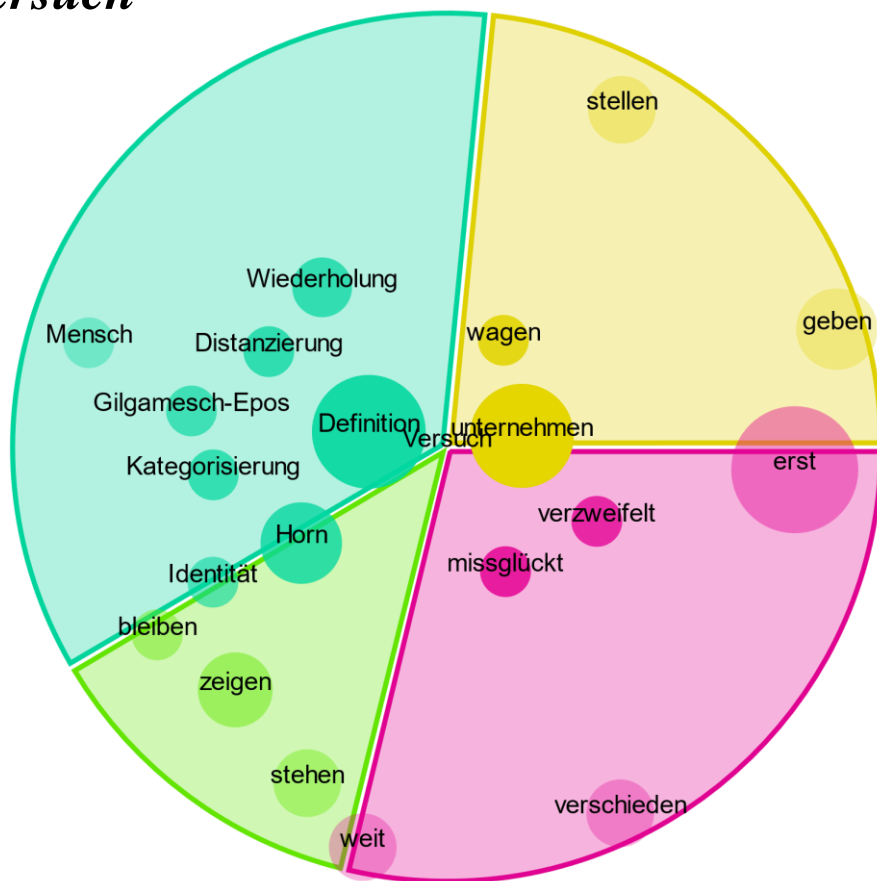


Abbildung 48: Versuch L1 **VERB+AKK** **VERB-SUBJEKT** **ADJ-ATTRIBUT** **GEN-ATTRIBUT**

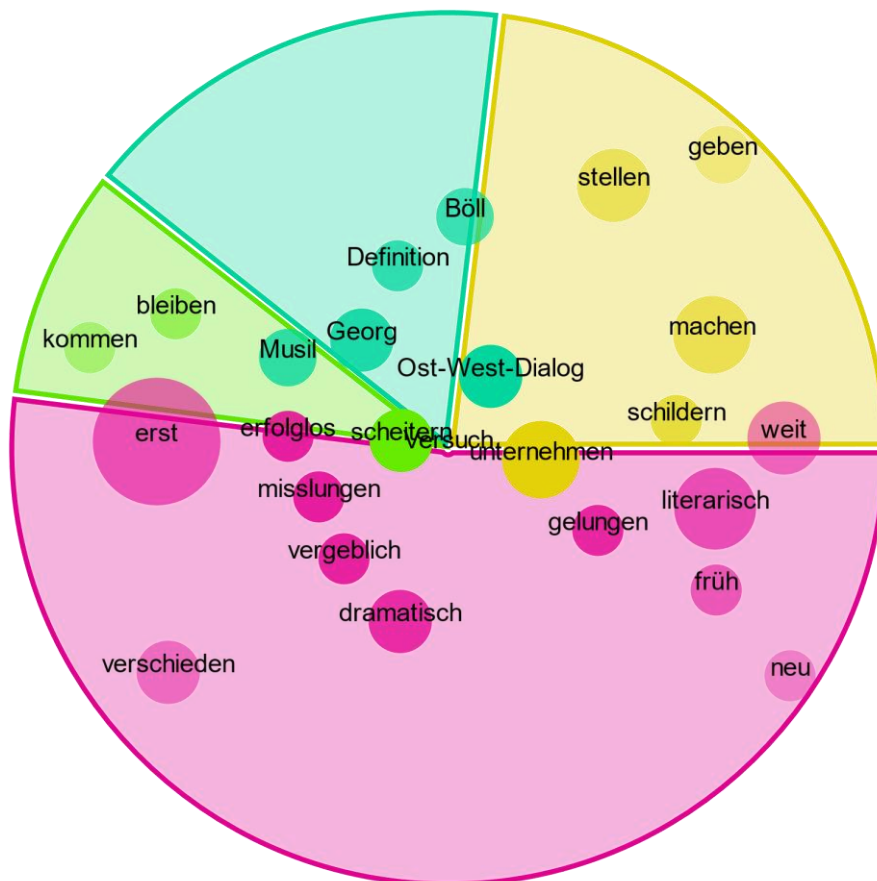


Abbildung 49: Versuch L2 **VERB+AKK** **VERB-SUBJEKT** **ADJ-ATTRIBUT** **GEN-ATTRIBUT**

<i>Versuch</i>	Linguistik	Literatur	Σ	DP_{norm}
L1	116	348	464	0,40
L2	92	259	351	0,58
Σ	208	607	815	
X^2	(1, $N = 815$) = 4,758, $p = 0,029$			

Tabelle 23: *Versuch* – statistische Übersicht

Kommentar

Die Ergebnisse der deskriptiven Statistik für *Versuch* waren freilich überraschend, denn dermaßen niedrige Anzahl in sprachwissenschaftlichen Texten wurde gar nicht erwartet (RR = 0,546; CI [0,486; 0,614]). Diese anfängliche Fehleinschätzung resultierte daraus, dass ich das recht vielfältige Funktionalitätspotenzial des Lexems in den zwei Disziplinen nicht in Kauf genommen hatte. In den literarischen Arbeiten sind es die Autoren oder Gestalten, die bestimmte Versuche unternehmen, was wiederum bedeutet, dass das Substantiv nur sehr selten metadiskursiv fungiert:

- (40) *Coninx wertete eher negativ Frischs **Versuche**, Schriftsteller zu werden oder als Journalist zu arbeiten.* (0003T)
- (41) *Seine Theaterstücke und Gedichte sind ein missglückter **Versuch**.* (0044T)

Diese „narrative“ Gebrauchsweise ist aber nicht nur den literarischen Arbeiten vorbehalten:

- (42) *Der Spiegel beschäftigt sich mit dem Thema Ursula von der Leyen, die einen **Versuch** unternommen hat, bestimmte Websites zu sperren.* (0014T)
- (43) *Unter den belarussischen Nationalisten wurden ebenso die ersten **Versuche** unternommen, ihre Muttersprache als Literatursprache zu verwenden.* (0026D)
- (44) *Ein **Versuch**, ein österreichisches phraseologisches Wörterbuch zu erstellen, stammt aus dem russischen Sprachraum.* (0033D)

Eine vermutlich in der Linguistik üblichere Funktion entspricht wiederum einem Hedge. Diesmal ist es allerdings im Hinblick auf die Objektivität nicht zwingend erforderlich. Die Beispielsätze unten hätten die Autoren genauso gut ohne *Versuch* formulieren können. Mit dem Entschluss, das Forschungsvorhaben als Versuch zu verkündigen, ändert sich gar nichts an dem Wahrheitsgehalt der Proposition, denn es handelt sich um einen Plan, dessen Erfüllung zu der Zeit des Lesens nicht objektiv beurteilt werden kann.

- (45) *Mittels der Methoden einer ausführlichen textlinguistischen Analyse wird ein **Versuch** unternommen, den textuellen Aufbau der Glosse zu erfassen.* (0014T)
- (46) *Hauptbestandteil dieser Arbeit bildet der **Versuch**, die empirischen Daten aus den einzelnen Studien und die linguistischen Modelle miteinander in Verbindung zu setzen und zu überprüfen, inwieweit diese linguistischen Konzepte die Leistungsmuster der zum Teil sehr heterogenen Patientengruppen erfassen können.* (0032D)

Rolek (2018) in ihrer Analyse von Hedges und ihrer Funktion in studentischer Wissenschaftssprache deutet das Verb *versuchen* wie folgt: „[d]ie Verfasserin bringt dadurch sowohl die Ungewissheit zum Ausdruck, ob die Durchführung der beabsichtigten Handlung gelingt, als auch die Bemühungen, die beabsichtigte Handlung so erfolgreich wie möglich durchzuführen“ (S. 233). Gleiches trifft auf die Substantivierung zu, indem im Falle von *Versuch* hauptsächlich folgende Ziele realisiert werden: „Einschränkung der eigenen

Verantwortung für die Wahrheit des propositionalen Gehalts“ und „Vermeidung möglicher Kritik seitens der Rezipienten“. (Breitkopf, 2006, S. 115 zit. n. Rolek, 2018, S. 228)

VERB + Versuch als Akkusativobjekt

unternehmen, es gibt, darstellen, wagen

Die literarische Bewegung **unternahm** ab etwa 1830 **Versuche**, auf öffentlichem Wege auf ihre politischen Forderungen aufmerksam zu machen. (0156)

Doch **gab es** immer wieder **Versuche**, die Zensur durch Doppeldeutigkeiten und subtile Anspielungen soweit wie möglich zu umgehen, um sich wenn schon nicht auf politischem, zumindest auf literarischem Wege gewisse Freiheiten zu verschaffen. (0101)

Meine Arbeit **stellt** den Versuch einer systematischen **Analyse** des Gilgamesch-Epos als Prätext von "Fluß ohne Ufer" **dar**. (0158)

Mag. Bito hingegen wäre bereit, einen **Versuch** zu **wagen** und mit wissenschaftlichen Grammatiken zu arbeiten, mit einer Oberstufenklasse hat sie dies bereits erprobt. (0136)

VERB + Versuch als Subjekt

sich zeigen, stehen

Im Hinblick auf die Literatur ab dem Polenfeldzug **zeigt sich** auch deutlich der **Versuch** einer völkischen Erziehung, deren zentrales Element der Rassegedanke war. (0013)

Die deutschsprachigen adoleszenten Hauptfiguren bleiben entweder ihren Prinzipien treu oder entwickeln diese erst im Verlauf der Handlung, am Ende **steht** jedoch stets der **Versuch**, die Welt mittels politischem Engagement und Diplomatie zu verbessern. (0028)

Demgegenüber stand aber auch der **Versuch** eines inhaltlichen Neuanfangs, der darin bestand, die Leiden des Kriegs nicht zu verschweigen, sondern gerade auch in der Kinderliteratur zu thematisieren. (0084)

(Adjektiv)ATTRIBUT + Versuch

erst, verschieden, weiter, missglückt, verzweifelt

Der **erste Versuch** der Autorin, ihr Werk zu veröffentlichen, schlug allerdings fehl, als der Herausgeber der Zeitschrift Atlantic Monthly den Text als "zu unerfreulich" ablehnte. (0087)

Daher wurden **verschiedene Versuche** unternommen, die Gemeinsamkeiten der zu einer Textsorte gehörenden Texte zu spezifizieren. (0123)

Im Zuge seiner umfangreichen Darstellung kann dies als **weiterer Versuch** der Systematisierung angesehen werden. (0157)

In diesem Kapitel soll die Geschichte des **missglückten Versuchs** des Folgesymposiums dargestellt werden, ebenso wie Inhalt, Ablauf und Rezeption der tatsächlich stattgefundenen Tagung. (0020)

Das Gegenteil dieser Behauptung lässt sich anhand des Textes nachweisen: Die unzähligen **verzweifelten Versuche** der Männerfiguren, Rahels autonomer Weiblichkeit Herr zu werden,

deuten auf die Brüchigkeit des patriarchalen Herrschaftssystems und die Notwendigkeit eines neuen Frauenbildes voraus. (0092)

Versuch + (Genitiv)ATTRIBUT

Definition, Kategorisierung

*Im Folgenden sollen ein paar **Versuche einer Definition** von "Parenthesen", die aus rezenten Arbeiten zum Thema entnommen sind, diesen ersten Definitionstypen beispielhaft illustrieren. (0133)*

*Es gibt zahlreiche Untersuchungen und unterschiedliche **Versuche der Kategorisierung** der Merkmale von schauerlichen Werken. (0103)*

6. Didaktische Implikationen

Wie ich schon in der Einleitung bemerkt habe, nimmt die vorliegende Arbeit nicht nur auf das Produkt – die Analyse – Bedacht, sondern auch auf den Prozess. Idealerweise soll sie zugleich durchgängig ein Beispiel dafür setzen, wofür sie plädiert. Auf die Prozess-Produkt-Dichotomie stützt sich auch die Gliederung dieses Abschnitts. Im Kontext der Fremdsprachendidaktik wende ich mich nämlich den folgenden drei Fragen zu. Was bringen die Wortprofile? Was bringt ihr Erhebungsprozess? Und welche Vorannahmen liegen deren Didaktisierung zugrunde?

6.1 Produkt

Ich resümiere zuerst die didaktischen Implikationen der in der Analyse gewonnen Erkenntnisse und danach widme ich mich dem Potenzial des praktischen didaktischen Einsatzes der Profile.

Erstens zu den Unterschieden zwischen den Kollokaten von L1- und L2-Studierenden. In ihrer Gesamtheit betrachtet könnte man sich im Glauben wähnen, es bestehen keine, wenn die Dispersion nicht eingerechnet worden wäre. Diese hat gezeigt, dass die Verteilung der untersuchten Abstrakta, und daher auch ihrer Kollokate, in der L2-Gruppe alles andere als uniform ist. Kurzum sind die Wörter zwar nicht unbekannt, sondern sie werden in ganz unterschiedlichem Ausmaß verwendet. Dies mag wiederum von den einleitend erwähnten Niveauvariabilität zeugen. Zweitens wurde klar, dass einige Abstrakta eindeutige Präferenzen für eine der zwei Disziplinen aufweisen. *Fall*, *Unterschied*, *Ergebnis*, *Aussage*, und *Tendenz* kamen alle tendenziell in den sprachwissenschaftlichen Texten vor, während *Vorstellung* und *Versuch* begegnete man eher in den literaturwissenschaftlichen. Daneben stellte man fest, es können auch disziplinbedingte Gebrauchsspezifika hinsichtlich der Funktion vorkommen. Unter diesem Gesichtspunkt ist vor allem die Notwendigkeit von Berücksichtigung des Kontexts hervorzuheben. Verfassen Studierende beispielsweise eine sprachwissenschaftlich ausgerichtete Arbeit, bedeutet dies nicht nur die Inklusion der entsprechenden Ausdrucksmittel, sondern auch einen vorsichtigen Umgang mit diesen Mitteln, der die disziplinären bzw. thematischen Eigenarten reflektiert – Brommer (2019) spricht von einem (produktiven) „Kontextualisierungspotenzial“: „Indem der Produzent auf sprachlich Verfestigtes zurückgreift, signalisiert er die Zugehörigkeit des Texts zu einem bestimmten Kontext oder einem spezifischen Sprachausschnitt“ (S. 11). Drittens unterstreichen die Ergebnisse die Bedeutung differenzierter Herangehensweise an sprachliche Phänomene. Um eine möglichst hilfreiche und umfassende Darstellung zu produzieren, sind nicht nur N-Gramme, sondern auch Kollokationen unentbehrlich. Daher sind bei einigen Abstrakta auch N-Gramme angeführt worden. Letztens nur eine Bemerkung hinsichtlich der Wichtigkeit von Abstrakta in der Wissenschaftssprache. In Anbetracht deren Häufigkeit sowie Funktion innerhalb des wissenschaftlichen Diskurses sind sie unbestritten ein Bestandteil der AWS und als solchen soll ihnen in Lehrveranstaltungen mit Schwerpunkt Wissenschaftssprache/ Schreiben Aufmerksamkeit geschenkt werden. Ob dies ihrer Form nach wie hier, oder eher anhand ihrer Funktion geschieht, ist dabei m. E. nicht so wichtig, solange genug Kontext und Kontext gegeben wird.

Nun zum praktischen Gebrauch der Wortprofile. Sie sind, ähnlich wie die Einträge in einem Kollokationswörterbuch, eine Abkürzung. Wenn man sie vor sich hat, ist die langwierige Prozedur der Korpusabfrage nicht mehr durchzulaufen. Im Unterschied zu einem

Kollokationswörterbuch bieten die Profile je einen Beispielsatz pro Kollokat. Der Kontext dieser Sätze ist dem Kontext des erwünschten Ergebnisses so nah, wie es möglich ist; es handelt sich um die gleiche Textsorte, sogar um die gleichen Disziplinen. Andererseits setzt sich das Korpus ausschließlich aus studentischen Texten zusammen und zudem ist die ganze Datenbank relativ klein. Daraus geht hervor, dass nur die häufigsten Kollokate durch das Sieb gedungen sind. Da diese bis auf einige Ausnahmen wohl bekannte³⁵ Lexeme sind, besteht die Herausforderung darin, alte Wörter auf neue Weise zu benutzen zu lernen.

Um dieses Ziel zu erreichen, können Studierende die Sätze und die Abstrakta auf ihre Funktion analysieren. Es bietet sich ebenfalls die Möglichkeit, die Sätze nur zu lesen, sie später im eigenen schriftlichen Gebrauch versuchsweise nachzuahmen und somit die Wortprofile als Konzentrat sprachlicher Erscheinungen anzusehen, womit es wert ist, konfrontiert zu werden. Des Weiteren steigt bei solcher gezielter Konfrontation die Wahrscheinlichkeit, dass Studierende sich nicht nur die Kollokationen merken, sondern dass sich bei wiederholter Begegnung auch das kolligationelle Umfeld einprägt, d. h. die grammatischen Konstruktionen in unmittelbarer Nähe. Es sind diese mehrfachen Begegnungen, insbesondere wenn sie als Komponent der „wiederholten Verteilung“ (spaced repetition) angewandt werden (Joe, 2010), was im Wortschatzerwerb eine Schlüsselrolle zu spielen scheint. In dieser Hinsicht warnen Laufer & Rozovski-Roitblat (2015) jedoch davor, dass obwohl die Anzahl von Begegnungen tatsächlich ein Faktor ist, dürfte er wahrscheinlich dadurch beeinflusst werden, auf welche Art sich die Begegnungen abspielen. Abschließend behaupten die Autoren, „[i]f one or two good exercises with new words lead to better learning than 18–21 encounters with the words in texts, then such exercises should be an integral part of course and material design“ (S. 708). Demgemäß soll die Arbeit von Profilen immer um eine analytische oder kreative Aufgabe erweitert werden. Im Kontext von der Wissenschaftssprache bietet sich als Beispiel eine Schreibaufgabe, wo die Lernenden Kollokationen oder N-Gramme aus anderen Arbeiten in ihre eigenen einzubauen versuchen können.³⁶

6.2 Prozess

Zunächst wird eine knappe, exkursartige Kontextualisierung des DDL-Konzepts vorgenommen, woran sich ein konkreter, mit der obigen Analyse zusammenhängender Didaktisierungsentwurf anschließt.

6.2.1 Exkurs: Korpora im Fremdsprachenunterricht

Unter „Prozess“ wird hier gezieltes Arbeiten mit korpusanalytischen Instrumenten im Kontext des Fremdsprachenerwerbs verstanden. Die Idee hat sich in der englischen Fremdsprachendidaktik bereits in den 90er Jahren als einer der vielen Wege des „prinzipiellen Eklektizismus“ (principled eclecticism) in der Post-Methoden-Ära³⁷ unter dem Namen Data-Driven-Learning (DDL) etabliert.³⁸ Im Einklang mit dem oft-zitierten Postulat Kumaravadivelu (2001), „[t]he postmethod learner is an autonomous learner“ (S. 545) plädiert der Gründer des Ansatzes, Tim Johns, fürs Umdenken der Rolle des Fremdsprachenlernenden:

³⁵ Im Sinne, dass sie zum passiven Wortschatz gehören.

³⁶ Siehe Pecorari (2003) für die Diskussion zum Thema Plagiat und *patchwriting* als eine Überbrückungsstrategie.

³⁷ Dazu u. a. Kumaravadivelu (2001).

³⁸ Schon viel früher gab es allerdings Bemühungen, Technologie und Fremdsprachenlernen zu vereinbaren (siehe etwa Higgins, 1983). Seitdem werden alle auf dieses Ziel bedachten Ansätze und Methoden unter dem Namen CALL (Computer-assisted language learning) bekannt, innerhalb dessen auch DDL zu verorten ist.

„the language-learner is also, essentially, a research worker whose learning needs to be driven by access to linguistic data“³⁹ (Johns, 1991, S. 2). Da es völlig unrealistisch wäre, sich hier mit dem Ansatz, der inzwischen sehr facettenreich geworden ist, auseinanderzusetzen⁴⁰, werden nur drei zusammenfassende Prinzipien der DDL erwähnt:

- the data is primary, and the teacher does not know in advance exactly what rules or patterns the learners will discover: indeed, they will often notice things that are unknown not only to the teacher, but also to the standard works of reference on the language.
- a teacher, who has to learn to become a director and coordinator of student-initiated research
- placing the learner's own discovery of grammar at the centre of language-learning, and by making it possible for that discovery to be based on evidence from authentic language use (S. 3)

Obwohl der DDL-Ansatz ursprünglich als universell und mit praktisch allen Zielgruppen von Sprachlernenden im Blick konzipiert wurde, erfreute er sich besonderer Beliebtheit im Bereich Wissenschaftssprache; vermutlich auch deswegen, weil die Lernenden schon über genügend Sprachkenntnisse und Fachwissen verfügten, die einige mit DDL verbundene Hürden zu überwinden helfen. Welche Vorteile hat der datengeleitete Ansatz also für die Wissenschaftssprache oder Uni-Studierenden überhaupt?

Bei Lernenden (sowie teilweise auch Lehrkräften), die nach dem GER Kenntnisse mindestens⁴¹ auf der guten Mittelstufe (B2) demonstrieren, wird sich die Rolle der Intuition bei der Beurteilung verschiedener sprachlicher Erscheinungen (je nach dem Niveau) wohl zunehmend verstärkt haben. Eine mögliche Komplementierung von Intuition mit Korpora wird von Káňa (2014) folgendermaßen skizziert:

Das „Bauchgefühl“ funktioniert bei muttersprachlichen und/ oder erfahrenen Lehrern/-innen zwar einwandfrei. Besser ist es jedoch, wenn dieses „Gefühl“, bzw. längst in Vergessenheit geratene Schulformeln, gleich auch durch mehrere Belege objektiviert werden. (Káňa, 2014, S. 115)

Ogleich die Behauptung, Intuition funktioniere bei Muttersprachlern einwandfrei, als absolut haltlos zurückzuweisen ist (siehe Hanks, 2013, S. 20; Wray, 2005, zit. n. Flowerdew L., 2012, S. 42), ist dem zweiten Teil des Zitats völlig zuzustimmen. Es ist nämlich oft die Intuition, die angezweifelt wird, wodurch der Grundstein für die entsprechende Fragestellung gelegt wird, die man nachher einer Korpusüberprüfung unterziehen kann. So gelangt man an Daten, die interpretiert und strukturiert werden – man erhält Informationen. Im nächsten Schritt werden diese in unser Denken integriert, um das Wissen zu bilden (Thomas, 2017, S. 2). So schön sich das auch anhören mag, bleibt festzuhalten, dass konkrete Vorschläge fürs Deutsche immer auf

³⁹ Angemerkt sei, dass es sich bei der praktischen Umsetzung streng genommen eher um eine korpusbasierte Analyse handelt. Da diese als Anreiz zu weiteren Fragestellungen fungieren kann und oft dies auch tut, verfließen bei DDL die Grenzen zwischen dem Korpusgesteuerten im wahren Sinne des Wortes und dem Korpusbasiertem.

⁴⁰ Für die Entwicklung von DDL siehe z. B. Boulton (2011).

⁴¹ Drumbl (2011) macht auf die Beziehung zwischen dem individuellen Sprachniveau und dem Gefühl sprachlicher Unzulänglichkeit aufmerksam: „[j]e fortgeschrittener die Sprachkompetenz und die Sprachbewusstheit des Sprechenden, desto klarer die Erkenntnis, dass seine Sprachkompetenz auf eine ganz eigene, auf eine persönliche Weise „lückenhaft“ geblieben ist“ (S. 64). Es sei allerdings darauf hingewiesen, dass dies ein erfolgreicher DDL-basierter Unterricht nicht ausschließt (vgl. Geist & Hahn, 2012).

der Ebene von separaten linguistischen Fallstudien bleiben, die – um für didaktische Zwecke angewandt werden zu können – viel zu wünschen übrig lassen. Erstens geht es oft um ausgeprägte linguistische Forschungen, wo das didaktische Potenzial mit bloßem Hinweis auf dessen Möglichkeit kenntlich gemacht wird, zweitens erfordern solche Arbeiten in der Regel spezialisierte Vorkenntnisse, mitunter sogar von den Lehrkräften/Dozenten. Etwas zugänglicher und praktischer sind die Aufsätze von Dubova & Proveja (2016) und Ylönen (2012), die immerhin mehr als (zugegebenermaßen nützliche) Berichte konzipiert sind und daher eine mögliche Replikation von anderen Lehrkräften kaum erleichtern. Materialien, die es eben tun, werden im nächsten Absatz aufgeführt.

Zu den ersten englischsprachigen, für die Verwendung im Unterricht geeigneten Materialien gehören m. W. Tim Johns' Kibbitzers.⁴² In den „Kibbitzers“ wurde anhand eines problematischen Beispiels aus einer studentischen Arbeit eine Frage aufgeworfen, die dann schrittweise mithilfe von Korpora beantwortet wird. Aus den neulich erschienenen Materialien nennen wir beispielsweise Thomas (2017) mit über 300 Fragen (siehe Beispiele unten), die mit Korpora zu beantworten sind, oder Viana (2020) mit bereitgestellten Unterrichtsplänen, die auf Korpusabfragen beruhen.

Is *not only* always completed with *but also*?

How is *rather* used apart from the phrase *would rather*?

Is *data* used as a singular noun, plural, or both?

Im Deutschen lassen sich Tendenzen in diese Richtung vor allem in Arbeiten von Wallner (z. B. 2013) ausmachen. Nicht nur in der tschechischen Germanistik beispielhaft sind m. E. aber die Fallstudien in Káňa (2014), indem sich der Autor auch mit der didaktischen Seite die Hände schmutzig macht; ähnlich wie in Thomas werden konkrete und – was vielleicht noch wichtiger ist – reale Fragen zum Sprachgebrauch gestellt, beispielsweise:

Welche Grundwörter verbinden sich mit den adjektivischen Suffixen *-ig*, *-lich* und *-isch*?

Welche Form überwiegt: *Peking* oder *Beijing*? Gibt es auch regionale Unterschiede? Seit wann ist die Form *Beijing* belegt?

Im Anschluss daran folgen lernfreundliche Anweisungen zur Handhabung des gewählten Korpus(instruments) mit dem Ziel, zur Beantwortung der Frage anzuleiten.

In den obigen Ausführungen sind das Potenzial und mögliche Realisierungsformen der prozessuellen Auseinandersetzung mit sprachlichen Daten umrissen worden. Als besonders fruchtbar sehe ich in diesem Kontext den Einsatz von Fragen, und zwar weil er die oft kritisierten Unzulänglichkeiten des traditionellen Unterrichts, Lehrerzentriertheit und uneigentliche Lehrerfragen (Huneke & Steinig, 2013, S. 118), vollends beseitigt. Gleichwohl bleibt festzuhalten, dass bislang nur sehr wenige nutzbare Beispiele Eingang in die Unterrichtspraxis gefunden haben. Im nächsten Abschnitt wird – so möchte ich wenigstens glauben – ein solches Beispiel vorgestellt.

⁴² Siehe Johns (o. J.): <https://lexically.net/TimJohns/>.

6.2.2 Praktische Umsetzung

Es wird davon ausgegangen, dass die Gelegenheiten zum Fremdsprachenerwerb u. a. durch Korpusabfragen geschaffen werden können. Dafür sind als Anstoß einleitende Fragen vonnöten. Um die Prozedur noch lernerzentrierter zu gestalten, könnte es genauso gut den Lernenden überlassen werden, die Fragen zu entwerfen. Neben den offensichtlichen externen Beschränkungen wie dem Lehrplan läuft man bei diesem Modus Operandi Gefahr, vor der Geist & Hahn (2012) warnen: „the students also complained about not being able to find a language issue which would be worth looking up in a corpus“ (S. 130). Ich fürchte, dass die Sorge – zumindest in den ersten Jahren – ebenfalls bei Philologiestudierenden berechtigt ist. Vor diesem Hintergrund stellen die von den Lehrkräften aufgeworfenen Fragen einen tragfähigen Kompromiss dar. Zugleich sollen jene Fragen nichtsdestoweniger indirekt die Kompetenzen fördern, eigene Fragen zu stellen.

Die Zielgruppe in dieser Arbeit sind Germanistik-Bachelorstudierende des zweiten/dritten Studienjahres. Sie verfügen über Grundkenntnisse über Sprachwissenschaft und Literatur, werden höchstwahrscheinlich eine begrenzte Anzahl von Seminararbeiten auf Deutsch verfasst haben, ihre Deutschkenntnisse sind mehrheitlich wohl irgendwo zwischen B2 und C1 (GER), und in einem Jahr werden die meisten ihre Bachelorarbeit zu schreiben haben. Im Unterschied zu „normalen“ DaF-Lernenden kann man in dieser Gruppe ein sekundäres Ziel voraussetzen, die Studierenden für bestimmte methodische Konzepte und sprachwissenschaftliche Begrifflichkeiten zu sensibilisieren, denn – wie Thomas (2017) zutreffend bemerkt – „[w]e learn so much more from corpus study than mere language!“ (S. 180) So lernt man zum Beispiel die Begriffe *Diathese* und *Mehrwortverbindung* kennen, analysiert die Interaktion von zwei Variablen, verwendet absolute und relative Häufigkeiten, macht sich Gedanken über die Struktur einer Abschlussarbeit und vor allem erfährt eine große Menge über die 15 Abstrakta.

Auf der anderen Seite ist in Wirklichkeit derzeit nicht damit zu rechnen, die Studierenden haben vorher mit einem Korpus gearbeitet. Korpuslinguistik als Lehrveranstaltung wird an den meisten germanistischen Instituten in Tschechien – wenn überhaupt – erst später (im Magisterstudium) angeboten.⁴³ Obwohl sich die einzelnen Fragen m. E. gut dazu eignen, um einige Funktionen von Korpusinstrumenten zu veranschaulichen und einzuüben, ist der gesamte Fragenpool als eigenständiges Einführungsmaterial nicht zu empfehlen. Erstens sind die Fragen zu spezifisch – sie fokussieren sich lediglich auf die Abstrakta, zweitens wurde während der Aufbereitung auf eine gewisse Einstufung nach Aufwand/Schwierigkeit zwar Rücksicht genommen, sie ist aber wegen des Fokus sehr ungenau. Trotzdem werden für die Zwecke dieser Arbeit all jenen Studierenden hypothetisch einige Kompetenzen in Bezug auf die Bedienung des Korpusmanagers zugetraut. Genauer gesagt werden Grundkenntnisse⁴⁴ des Sketch-Engine-Interface vorausgesetzt. Immerhin ist Arbeit mit einem anderen Korpusmanager prinzipiell auch denkbar. Was die Logistik angeht, wäre es von erheblichem Vorteil, wenn die Fragen im Rahmen einer Lehrveranstaltung bearbeitet werden könnten. Mit den Fragen selbständig

⁴³ Diese Situation ist u. a. auf die Sonderstellung von Korpuslinguistik innerhalb anderer Disziplinen der Sprachwissenschaft zurückzuführen. Es fragt sich daher, ob man zuerst die als Teil von Korpuslinguistik erworbenen Fertigkeiten im Studium der anderen Disziplinen anwenden soll, oder umgekehrt in den Disziplinen nach einer Grundlage und Fragen suchen, die sich erst später durch Korpuslinguistik beantworten lassen. Dieses Henne-Ei-Problem wird hier zwar nicht weiter ausgeführt, es sollte bislang aber klargestellt sein, wie schlüssig es ist für die Überlegungen zur praktischen Umsetzung von auf Korpora basierenden Lernaufgaben.

⁴⁴ Diese werden hier grob als praktische Kenntnisse der Funktionen Konkordanz (mit der Ausnahme von einigen komplizierten CQL-Abfragen), Word Sketch und N-Gramme aufgefasst.

Arbeiten ist nicht zwar ausgeschlossen, man beraubt sich aber der einzigartigen Möglichkeit, mit anderen zu diskutieren und sich von den Dozenten beraten lassen zu können.

Nun zu den Fragen selbst. Sie reichen von kurzen geschlossenen Fragen, deren Beantwortung nur einer einfachen Suche bedarf, bis hin zu komplexen Aufgaben, die aufwendiger sind und tiefere Kenntnisse erfordern. Es gibt sogar einige Fragen, die „nur“ als Denk- bzw. Diskussionsanstoß dienen und dadurch die theoretische Vorbereitung auf die Recherchen gewährleisten. Außerdem regen sie die Studierenden zur Aufstellung eigener Hypothesen, was die nachfolgende Suche und demzufolge auch Überprüfung zweckmäßiger macht. Die meisten Lösungen könnte man in DESTWIKO (v0.2) finden. Wird man dort allerdings nicht fündig, wird oft auf German Web 2013 (deTenTen2013) oder sein Teilkorpus – die Domain www.diplomarbeiten24.de – zurückgegriffen. Die Studierenden werden hiermit aufgefordert, nach Lust und Laune auch Korpora außerhalb von Sketch Engine (z. B. von IDS Mannheim) zu benutzen, falls Interesse oder Bedarf besteht. Außerdem kann man die angefertigten Wortprofile nach Antworten durchsuchen. Dazu noch eine Anmerkung: Manche Fragen gehen über den Rahmen der Wortprofile hinaus und lassen sich somit nur mithilfe der Profile bei Weitem nicht zufriedenstellend beantworten. Zu guter Letzt sei darauf hingewiesen, dass viele der Fragen eher leichtherzig und informell formuliert worden sind. Auf diese Weise wird der Perspektive der Studierenden Rechnung getragen, was eine motivierende Wirkung haben sollte. Darüber hinaus werden die Sätze etwas verständlicher, wenn man auf die formelle Ausdrucksweise verzichtet.

6.2.3 Fragen zur Korpusarbeit

Fall

- Frage 1.* Wird *Fall* über die zwei Disziplinen gleichmäßig verteilt oder gibt es eine Disziplin, wo es häufiger vorkommt?
- Frage 2.* Was sind die Funktionen von *Fall* in den zwei Disziplinen? Unterscheiden sie sich im Hinblick auf die Disziplin?
- Frage 3.* Warum kollokieren so viele Nummern mit *Fall*?
- Frage 4.* Ich habe die Form *Falle* gesehen. Wie häufig ist sie im Vergleich zu *Fall*? Gibt es einige Restriktionen hinsichtlich ihres Gebrauchs?
- Frage 5.* Die Phrase *auf jeden Fall* hat eine ziemlich hohe Frequenz. Wie sieht sie im Vergleich zum Adverb *jedenfalls*? Gibt es in dieser Hinsicht Unterschiede zwischen Muttersprachlern und Nichtmuttersprachlern?
- Frage 6.* Welche Konstruktion folgt der Verbindung *im Falle*? Genitiv, *von* + Dativ, oder beides?
- Frage 7.* *In diesem Fall* scheint minimale Varianz aufzuweisen. Gibt es doch Beispiele von dieser Konstruktion, wo dem Substantiv *Fall* ein Adjektiv vorausgeht?
- Frage 8.* Das Kopulaverb *sein* tritt oft mit *Fall* zusammen. Wie kann man sie zusammen benutzen? Welche anderen Satzglieder hängen von dem Kopulaverb ab?
- Frage 9.* Wird *Fall* im Singular und Plural mit gleichen Adjektiven gebraucht?
- Frage 10.* Da *in diesem Fall* recht häufig vorkommt, hat es etwa eine bestimmte Funktion im Text?/Worauf bezieht sich das Substantiv?

Frage

- Frage 11.* Wird *Frage* über Disziplin und Muttersprache gleichmäßig verteilt?
- Frage 12.* Warum spricht man überhaupt über Fragen? Sie sind denn nicht durch das Fragezeichen gekennzeichnet?
- Frage 13.* Was machen wir mit Fragen im wissenschaftlichen Diskurs (Germanistik)?
- Frage 14.* Sehen sie sich eine größere Menge vom Kotext an. In welchen Teilen von Abschlussarbeiten kommt *Frage* besonders häufig vor?

- Frage 15.* Man interessiert sich *für* etwas, aber hat Interesse *an* etwas. Man kann auch *nach* etwas fragen. Welche Präposition benutzt man nach dem Substantiv *Frage*?
- Frage 16.* Das Verb *gehen* und *Frage* kollokieren, aber Fragen gehen doch nicht und niemand kann eine Frage gehen. Wie werden die zwei Wörter denn zusammen gebraucht?
- Frage 17.* Welche verbalen Kollokationen von *Frage* werden mit dem Substantiv auch als Partizipien gebraucht?
- Frage 18.* In meinem Aufsatz habe ich *wichtige Frage* schon mehrmals verwendet. Welche Alternativen gibt es?
- Frage 19.* Wie erscheinen Fragen in wissenschaftlichen Arbeiten? Schauen sie sich die Verben mit *Frage* als Subjekt an. Was ist vielen dieser Verben gemeinsam?/Ist *Frage* wirklich das Subjekt?
- Frage 20.* Ich bin auf das Substantiv *Fragestellung* gestoßen. Verhält es sich anders als *Frage*? Sind die Bedeutungen gleich?

Grund

- Frage 21.* Studierende, die sich mit einem sprachwissenschaftlichen Thema auseinandersetzen, sprechen gewiss mehr über Ursachen als diejenigen, die sich mit Literatur beschäftigen. Bestätigt die Verteilung von *Grund* in den Subkorpora diese These?
- Frage 22.* *Lassen* ist ein häufiges Kollokat von *Grund*, aber Konstruktionen wie *einen Grund lassen* oder *ein Grund lässt etwas* ergeben doch keinen Sinn. Wie benutzt man die zwei zusammen?
- Frage 23.* In welchen festen Mehrwortverbindungen kommt *Grund* vor?
- Frage 24.* Wie üblich ist es, nach *diesem* in *aus diesem Grund* ein Adjektiv zu haben?
- Frage 25.* Ich habe die Phrase *aus Verständlichkeitsgründen* gesehen. Wie häufig ist das Muster *aus* + [Nomen]*gründen* in Abschlussarbeiten vertreten? Welche Formen hat eventuell der zweite Teil? Vergleichen Sie die Ergebnisse ihrer Recherche mit einem größeren Korpus bzw. einem Referenzkorpus.
- Frage 26.* Während meiner Recherche bin ich der Form *auf Grund* begegnet. Ich habe immer gedacht es soll *aufgrund* heißen. Ermitteln sie die Häufigkeiten beider Formen. Spielt die Muttersprache etwa eine Rolle?

- Frage 27.* Ich kenne niemanden, der die Präposition *aufgrund* benutzt. In wie vielen Arbeiten von Nichtmuttersprachlern ist sie zu finden? Wird sie dort mehr oder weniger gebraucht als in den von Muttersprachlern verfassten Arbeiten?
- Frage 28.* *Der Grund ist ...* sagt es klipp und klar. Gibt es aber Alternativen zu dem Kopulaverb, wo aber gleichzeitig *Grund* an der Stelle des Subjekts bleibt? (Ich will ja, dass der Satz mit *Grund* auf den vorhergehenden Satz anknüpft.)
- Frage 29.* Wenn man Gründe nennt, wie werden diese orthographisch dargestellt? Werden sie durch einen Doppelpunkt, ein Komma oder als selbstständiger Satz eingeführt? Als ein langer Satz oder mehrere Sätze? Wenn die Gründe länger sind, wie sieht die Textorganisation aus (Nummerierung usw.)?
- Frage 30.* Welche adjektivischen Kollokate von *Grund* könnten wir als textorganisierend bezeichnen? Wieso?
- Frage 31.* Gibt es außer *aus* eine andere Präposition, die mit *Grund* häufig auftritt?
- Frage 32.* Das Substantiv *Grund* kann sich in einem Text sowohl auf die Ursache als auch den Effekt beziehen. Versuchen Sie Beispiele von beidem zu finden.
- Frage 33.* Welche Adjektive kollokieren mit dem Plural von *Grund*? Gibt es Unterschiede zum Singular?

Unterschied

- Frage 34.* Über Unterschiede sprechen bedeutet vergleichen. Wer vergleicht denn mehr? Muttersprachler oder Nichtmuttersprachler? Linguisten oder Literaten? Wie würden Sie die eventuellen Unterschiede erklären?
- Frage 35.* Schlägt sich die Disziplin in der Wahl der Kollokate von *Unterschied* nieder?
- Frage 36.* Was machen wir mit *Unterschiede*? Denken Sie daran, in welcher Forschungsphase wir diese Sachen machen.
- Frage 37.* Haben alle Verben mit *Unterschied* als Akkusativobjekt ein belebtes Subjekt?
- Frage 38.* Besteht in der Gruppe von Verben mit *Unterschied* als Akkusativobjekt ein Unterschied in der Diathese (aktiv/passiv) wenn der Verfasser der Arbeit das Subjekt ist und wenn die Subjektposition von einem zitierten Autor besetzt wird?
- Frage 39.* Wie wird das Verb *erkennen* mit *Unterschied* benutzt? Welche Form nimmt das Verb häufig an?
- Frage 40.* Wenn *Unterschied* mit dem Verb *bestehen* verwendet wird, welche Struktur schließt sich daran?

Frage 41. In einer Bachelorarbeit finden sie Folgendes:

Ich habe große Unterschiede gefunden.

Der Satz ist zwar richtig, etwas damit stimmt aber nicht. Was ist es und was kann man dagegen tun?/Der Satz ist zwar richtig, für die Textsorte jedoch unangemessen. Sehen Sie sich die möglichen Kollokate an und verbessern sie den Satz.

Frage 42. Unter den adjektivischen Kollokaten von *Unterschied* befindet sich auch *etwaig*, was mich an *etwas* erinnert. Daher bedeuten *etwaige Unterschiede* so etwas *wie einige* oder *wenige Unterschiede*... Wie stichhaltig ist diese Schlussfolgerung?

Frage 43. Welche Präposition(en) wird/werden mit *Unterschied* benutzt?

Frage 44. Ich habe die Konstruktion *im Unterschied zu* gesehen. Überprüfen Sie ihre Festigkeit? Ist es üblich, andere Wörter hinzuzufügen? Gibt es die Präposition nur in der verschmolzenen Form?

Frage 45. Wie wird *im Unterschied zu* verwendet? Welche anderen Sprachmittel kennen Sie, die über eine ähnliche Funktion verfügen?

Frage 46. Vergleichen wir häufiger mit dem Verb *unterscheiden* oder mit dem Substantiv *Unterschied*? Sind die Ergebnisse für sowohl Muttersprachler als auch Nichtmuttersprachler gültig? Wie sieht die Situation aus in einem Vergleichskorpus?

Frage 47. Wie bringt man zur Sprache, dass zwei Sachen nicht gleich sind. Unterscheiden wir oder vergleichen? Schreiben wir mehr über das Ergebnis (*Unterschied, unterscheiden*) oder dem den Prozess (*Vergleich, vergleichen*)?

Möglichkeit

Frage 48. Welche Verben kollokieren mit *Möglichkeit*. Was ist ihre semantische Valenz/Kompatibilität? Unterscheidet sich diese in Bezug auf die Disziplin?

Frage 49. Welche Verben mit *Möglichkeit* als Akkusativobjekt können mit dem Substantiv auch als Subjekt auftreten (mit *sich*)? (Ein größeres Korpus mag hierfür erforderlich sein.)

Frage 50. Welche der Adjektivkollokate von *Möglichkeit* können wir auch prädikativ verwenden?

Frage 51. Mehr als eine Möglichkeit haben ist immer gut. Mit welchen Adjektiven wird diese Tatsache in Abschlussarbeiten versprochen?

Frage 52. Möglichkeiten sollen sich aus etwas ergeben. Was ist dieses Etwas? Welchen Charakter hat es – Substantiv, Satz...? Ein Vergleichskorpus mag hilfreich sein.

Frage 53. Was *bietet* Möglichkeiten? Wie häufig werden sie passivisch geboten?

Problem

Frage 54. Wer problematisiert mehr? L1- oder L2-Studierende? Sprachwissenschaftler oder Literaten? Manchmal lohnt es sich, sich beides gleichzeitig anzuschauen, da es eine Interaktion geben kann. Ist dies hier der Fall?

Frage 55. Rufen Sie die Liste aller adjektivischen Kollokate von *Problem* ab. Verlassen Sie sich auf Ihre Intuition und bestimmen Sie in welcher Disziplin das jeweilige Adjektiv häufiger ist. Waren Ihre Annahmen richtig?

Frage 56. Eines der häufigen Kollokate von *Problem* ist das Verb *bereiten*. Wie sieht sein Valenz aus? Falls Sie das Verb nicht kannten, oder es einfach nicht benutzen würden, stellen Sie fest ob das nur Ihnen so geht – vergleichen Sie die L1- und L2-Datenbasen.

Frage 57. Sind Probleme wirklich Probleme? Mit *Problem* wird normalerweise etwas Negatives assoziiert. Inwiefern stimmt das in wissenschaftlichen Arbeiten?

Frage 58. Ist *Problematik* nur die große Schwester von *Problem*? Vergleichen Sie die Häufigkeiten. Spielen Disziplin und Muttersprache eine Rolle? Werfen Sie einen Blick auch auf die Wortformen. Worauf deuten die Ergebnisse hin? Kann man daraus auch einige Schlüsse über die Bedeutung ziehen?

Frage 59. Gibt es Bedeutungsunterschiede zwischen *Problem* und *Problematik*? Führen Sie eine schnelle Analyse von Kollokationsrestriktionen durch. Man kann zum Beispiel Probleme lösen, kann man auch Problematik lösen usw.? Ein größeres Korpus wäre hier hilfreich – erstellen Sie eins: German Web 2013 (deTenTen13) enthält unter anderem alle Dokumente von der Domain diplomarbeiten24.de.

Frage 60. Gehen wir Probleme um oder umgehen wir Probleme?

Frage 61. Wenn Sie über Probleme anderer Bescheid wissen, können Sie vielleicht einigen davon vorbeugen. Wählen Sie passende verbale Kollokate von *Problem* und ermitteln Sie, was Probleme in Abschlussarbeiten bereitet? Sie sollen dabei immerhin die Frage 57 im Hinterkopf behalten. (Sie können Ihre Abfrage genauer machen, indem Sie konkrete adjektivische Kollokate benutzen, z. B. *methodisch*. Dafür mögen Sie aber ein umfangreiches Korpus benötigen.)

Ergebnis

- Frage 62.* Wessen Studien sind ergiebiger? Vergleichen Sie die Häufigkeiten hinsichtlich der Disziplin und Muttersprache. Nehmen sie die Ergebnisse Ihrer Recherche zum Anlass, um über folgende Punkte nachzudenken:
- Was stellen Sie sich im Kontext von Wissenschaft unter Ergebnissen?
 - Welche methodischen Vorgehensweisen kann man benutzen um Ergebnisse, wie Sie sie eben definiert haben, zu gewinnen?
 - Wie sind Ihre Antworten auf die zwei Fragen mit den Häufigkeitszahlen zu vereinbaren?
 - Beantworten Sie die Frage b, in dem Sie sich die Genitivattribute von *Ergebnis* ansehen.
- Frage 63.* Was ist wenn sich nur ein paar Leute über Ergebnisse seitenlang auslassen? Das heißt doch nicht, dass die ganze Gruppe über Ergebnisse spricht oder sie überhaupt hat. In wie vielen Arbeiten ist das Wort *Ergebnis* zu finden? Welche Rolle spielen die Disziplin und Muttersprache? Ändert sich etwas, wenn Sie einmalige Erwähnungen ausschließen?
- Frage 64.* Welche Rollen werden Ergebnissen zugeteilt? Was machen Ergebnisse und was machen wir mit Ergebnissen? Versuchen Sie die Frage zuerst ohne Korpushilfe zu beantworten. Stellen Sie daraufhin eine Liste von den verbalen Kollokaten zusammen. Gibt es einige, die nur im Zusammenhang mit eigenen Ergebnissen/ fremden Ergebnissen benutzt würden?
- Frage 65.* Wie benutze ich das Verb *ausfallen* zusammen mit *Ergebnis*?
- Frage 66.* In der Wissenschaftssprache tendiert man meistens dazu, definitive Aussagen zu vermeiden. Ordnen Sie die Verben mit *Ergebnis* als Subjekt nach wie viel Sicherheit sie ausdrücken. Unterscheidet sich in dieser Hinsicht der Umgang mit eigenen Ergebnissen von dem mit Befunden aus sekundärer Literatur?
- Frage 67.* Leute sehen *schön/gut/schlecht* aus. Wie sehen denn Ergebnisse in der Wissenschaft aus?
- Frage 68.* Finden Sie in einem größeren Korpus heraus, ob Ergebnisse häufiger *gegen* etwas oder *für* etwas sprechen? Sind die etwaigen Tendenzen mit dem Gebrauch des mit Menschen benutztes Verbs *sich aussprechen für/gegen* vergleichbar?
- Frage 69.* Denken Sie an das Thema ihrer Abschluss- oder Seminararbeit. Unter welchen Umständen würden Sie die Ergebnisse als *eindeutig/aussagekräftig/wichtig/interessant/überraschend* bezeichnen?

Ziel

- Frage 70.* Wer spricht mehr über Ziele? Muttersprachler, Nichtmuttersprachler? Sprachwissenschaftler oder Literaturwissenschaftler? Worauf sind die eventuellen Unterschiede zurückzuführen?
- Frage 71.* Welche Funktion(en) hat *Ziel* im Text? Wird es in der Literatur oder Linguistik anders gebraucht?
- Frage 72.* Beschränken sich Ziele auf Sätze, Absätze, Abschnitte, Kapitel, den empirischen Teil oder sogar die ganze Arbeit? Worauf ist das Substantiv bezogen? Schauen Sie sich die Genitivattribute an. Gibt es Unterschiede zwischen den Disziplinen?
- Frage 73.* Ich habe *Ziel* und *hinausschießen* nebeneinander gesehen? Hat die Kollokation auch andere Komponenten? Was bedeutet sie? Was kann im Internet zu ihrer Stilistik erfahren? Äußert sich dies in ihrem Gebrauch? Vergleichen sie die relativen Häufigkeiten der Kollokation in DESTWIKO (v0.2)/diplomarbeiten.de und in German Web 2013 (deTenTen2013) - wie typisch ist die Kollokation für die Textsorte Abschlussarbeit?
- Frage 74.* *Ziel ist, die Sprachen zu vergleichen* oder *Ziel ist es, die Sprachen zu vergleichen*. Benutzen Sie größere Korpora und andere Nachschlagewerke und recherchieren Sie die Rolle des Korrelats.
- Frage 75.* Zu welchen Zeitpunkten im Prozess des Verfassens wären die Verben mit *Ziel* als Akkusativ zu gebrauchen?
- Frage 76.* *Mein Ziel Nummer eins ist...* Ab und zu, vor allem wenn man mehr Ziele hat, will man die Priorität jener Ziele verdeutlichen. Mit welchen Adjektiven lässt sich so etwas machen?
- Frage 77.* Was kann eigentlich ein Ziel sein? Beschreiben, vorstellen? Mithilfe von welchen Verben werden Ziele formuliert?
- Frage 78.* Welche festen Verbindungen mit *Ziel* gibt es?

Meinung

- Frage 79.* Wie sieht die Verteilung des Substantivs *Meinung* über die Muttersprache und Disziplin aus? Erklären Sie die Unterschiede (falls vorhanden).
- Frage 80.* Wie häufig sind Sätze *ich glaube* und *ich denke* im Korpus vertreten. Gibt es Unterschiede zwischen den Muttersprachlern und den Nichtmuttersprachlern?
- Frage 81.* *Meiner Meinung nach* in der Wissenschaft ihrer Meinung nach – eine großartige Wahl oder ein absolutes No-Go? Vergleichen Sie anhand der

Muttersprache. (Versuchen Sie eventuell ein eigenes Korpus zu dieser Frage zu erstellen.) Heißt das, *Meinung* wird aus der Wissenschaftssprache verbannt? Wie trägt *meiner* dazu bei?

Frage 82. Welche Synonyme zu *Meinung* gibt es?

Frage 83. *Ich/jemand habe/hat die Meinung, dass...* Finden sie passende Kollokationen von *Meinung* als Ersatz. Behalten Sie aber die Funktion des Satzes - Meinungsäußerung.

Frage 84. Gibt es außer *meiner Meinung nach* andere feste Verbindungen mit *Meinung*? Überprüfen Sie ihre Akzeptabilität hinsichtlich der Textsorte.

Aussage

Frage 85. Worauf bezieht sich das Substantiv *Aussage*? Auf primäre oder sekundäre Quellen? Auf den Autor oder die literarischen Gestalten? Auf eigene oder fremde Meinungen?

Frage 86. In Arbeiten mit erheblich mehreren Treffer von einem Wort kann sich um einen Fachausdruck handeln? Ist dies der Fall für *Aussage* in DESTWIKO (v0.2)?

Frage 87. Wo befindet sich die eigentliche Aussage - nach oder vor dem Substantiv *Aussage*?

Frage 88. Jemand kann eine Aussage machen. Welche Synonyme zu *machen* können Sie finden? Was können sie von ihrer Stilistik behaupten?

Frage 89. Was bedeutet eine *Aussage untermauern*? Wie kommt diese "Untermauerung" zustande?

Frage 90. Wie gültig sind Aussagen? Manche Adjektive können durch weitere attributiv modifiziert werden. *Gültig* ist eins davon. Versuchen Sie dessen Attribute zu finden.

Tatsache

Frage 91. Sehen Sie sich die Unterschiede im L1- und L2-Gebrauch. Vergleichen Sie diese mit dem Teilkorpus diplomarbeiten24.de.

Frage 92. Welche grammatischen Strukturen finden sich um das Substantiv *Tatsache* herum?

Frage 93. Wo steht der Inhalt, auf den sich *Tatsache* bezieht. Vor oder nach dem Substantiv. Wie ist er mit dem Nomen sprachlich verbunden?

Frage 94. Wo befinden sich in der Regel die Verben, die *Tatsache* als Subjekt/Objekt regieren? Versuchen die Gramrels für Verben mit *Tatsache* als

Subjekt/Akkusativobjekt anhand von Beispielen aus einem Korpus zu erstellen.

Frage 95. Vergleichen Sie die Häufigkeit des folgenden Syntagmas anhand von der Muttersprache des Verfassers: [ADJ] + SEIN+ *die* + *Tatsache*.

Frage 96. Die L1-Verfasser benutzen recht viele Kollokationen mit *Tatsache*. Überprüfen Sie in German Web 2013 (deTenTen2013) die Existenz derjenigen davon, die Sie selber nicht benutzen würden.

Frage 97. Mit welchen Adjektiven kollokiert *Tatsache*. (im Korpus diplomarbeiten24.de)

Frage 98. Welche feste Mehrwortverbindungen gibt es mit *Tatsache*?

Vorstellung

Frage 99. In welcher Disziplin ist *Vorstellung* mehr vertreten? Raten Sie zunächst ohne ein Korpus zu benutzen.

Frage 100. In der Linguistik stellt sich doch keiner ständig etwas vor? Welche Funktionen dürfte das Substantiv in den zwei Disziplinen besitzen?

Frage 101. Finden Sie heraus, in wie vielen Arbeiten das Substantiv *Vorstellung* zu finden ist. Überlegen Sie anhand der Konkordanzzeilen, an welche Themen *Vorstellung* gebunden sein mag.

Frage 102. *Prägen* kollokiert mit *Vorstellung*. Wessen Vorstellungen werden *geprägt*? Wodurch sind und werden (unsere) Vorstellungen *geprägt*?

Frage 103. Heißt *beruhen auf* das Gleiche wie *berufen auf*?

Frage 104. Wie werden adjektivische Kollokate von *Vorstellung* wie *andere*, *unterschiedliche* modifiziert?

Aufgabe

Frage 105. Welche Aufgabe erfüllt *Aufgabe* in wissenschaftlichen Texten? Vergleichen Sie Ihrer Intuition nach die Rolle des Substantivs in Arbeiten mit Schwerpunkt Literatur, Sprachwissenschaft und Lehramt (DaF).

Frage 106. Was folgt den Verben *bestehen in* und *liegen in*? Eine nominale Gruppe oder Korrelat (*darin*) + Infinitivnebensatz?

Frage 107. Ich habe ein Ziel, ergo habe ich Aufgaben zu bewältigen. Wird *Aufgabe* ähnlich wie *Ziel* zur Ankündigung der Forschungsziele benutzt?

Frage 108. Es gibt *primäre* Aufgaben, wie wäre es mit *sekundären*?

Frage 109. Vergleichen sie die Richtung, in die auf Aufgaben verwiesen wird (*diese* vs. *folgende*)?

Frage 110. *Machen* ist eines der Kollokate von *Aufgabe*. In dieser Kombination tritt das Pronomen *sich* recht häufig auf. Wie lautet die ganze Verbindung?

Tendenz

Frage 111. Denken Sie über die Semantik von *Tendenz* ein bisschen nach. Hätten Sie etwa sogar eine Definition parat?

Frage 112. Es heißt doch Wissenschaftssprache, hier sollen nur Fakten zählen. Lässt sich der Gebrauch von *Tendenz* im wissenschaftlichen Diskurs Ihrer Meinung nach überhaupt sinnvoll begründen? Schauen Sie sich danach einige Beispiele an. Haben Sie die Funktionen, die Sie erwähnt haben?

Frage 113. Wie sieht die Verteilung des Substantivs *Meinung* über die Muttersprache und Disziplin aus? Erklären Sie die Unterschiede (falls vorhanden).

Frage 114. Mit welchen Verben mit *Tendenz* als Akkusativobjekt lässt sich die Passivumschreibung mit *lassen* gebrauchen? Mit welchen ist dies sogar häufig der Fall? (Diese Suche werden Sie wohl in einem größeren Korpus durchführen müssen.)

Frage 115. Sie wollen das Passiv nicht übernutzen, gleichzeitig wollen Sie in dem Satz unpersönlich bleiben. Welche Verben können Sie benutzen, damit *Tendenz* zum Subjekt wird.

Frage 116. Ordnen Sie die Attribute von *Tendenz* nach ihrer Vagheit. (Bei manchen, die eher sachlich sind, ist dies nicht möglich, z. B. *negativ*.)

Frage 117. Welche Präposition findet man um *Tendenz* herum?

Versuch

Frage 118. Ist *Versuch* in der Linguistik oder in Literatur gebräuchlicher?

Frage 119. Welche Funktionen hat *Versuch*? Unterscheiden sich diese im Hinblick auf die Disziplinen? Sammeln Sie einige Beispiele.

Frage 120. Warum wird in der Zielsetzung häufig von Versuchen gesprochen? Eine wissenschaftliche Arbeit soll möglichst sachlich sein. Zudem soll ich am Ende ziemlich gut einschätzen können, ob die Arbeit bloß ein Versuch darstellt oder das einzig Wahre ist, oder?

Frage 121. Was wäre die Alternative in Form des Funktionsverbgefüges zu dem Verb *versuchen*?

Frage 122. Worin bestehen Versuche im wissenschaftlichen Diskurs? Werfen Sie einen Blick auf die Genitivattribute.

Frage 123. Was ist häufiger – das Verb *versuchen* oder das Substantiv *Versuch*? Entspricht das beobachtete Verhältnis Daten aus anderen Korpora?

7. Zusammenfassung und Ausblick

Da auf konkrete Schlussfolgerungen, Limitationen und Ergebnisse bereits im Rahmen einzelner Kapitel eingegangen worden ist, resümiere ich an dieser Stelle vor allem im Hinblick auf die Zielsetzung der Arbeit. Diese bestand in der Analyse von Kollokationen ausgewählter Abstrakta und ihrer eventuellen Didaktisierung. Es wäre nicht realistisch gewesen das Ziel ohne folgende Schritte zu erreichen.

Erstens habe ich mir vorgenommen, ein für die vorliegenden Fragestellungen geeignetes Korpus zu erstellen, wobei ich auf Vorarbeiten im Rahmen der früheren Abschlussarbeit zurückgreifen konnte. Die jetzige Version des Korpus im Sketch Engine – DESTWIKO (v0.2) – wurde im Vergleich zur ersten Version um weitere 226 Masterarbeiten (zum Gesamtumfang von 280 Arbeiten) ergänzt, was detailliertere Recherchen der Lexik ermöglichte. Das manuelle Tagging der satzbezogenen nominalen Verweise (SNVs) aus der ersten Version wurde im Falle der neu hinzugefügten Dokumente nicht durchgeführt. Besonderes Augenmerk galt bei der Erstellung der Ausgewogenheit, Repräsentativität und den Metadaten. Die zwei entstandenen Subkorpora (Muttersprachler, Nichtmuttersprachler) sind fast gleich groß, beinhalten genug Material, um als mindestens teilweise repräsentativ für die Masterarbeiten der tschechischen und österreichischen Germanisten/Sprach-/Literaturwissenschaftler bezeichnet werden zu können. Zudem wird genug Informationen über die Texte angegeben, damit im Falle einer weiteren Vergrößerung noch tiefgründiger recherchiert werden kann. Schon im Laufe dieser Studie stellte sich dennoch wiederholt heraus, dass für das Studium der lexikalischen Ebene die aktuelle Größe nicht völlig ausreicht. In der Zukunft könnte die Datenbank somit durch Texte aus anderen Universitäten und, vor allem, Bachelor- und Doktorarbeiten erweitert werden.

Einen ebenfalls hohen Stellenwert nahm in dieser Arbeit das methodische Vorgehen. Die Arbeit kombiniert qualitative und quantitative Analysen der untersuchten Phänomene. Dank dieser Triangulation sind die Ergebnisse m. E. deutlich aussagekräftiger und objektiver. Als das wichtigste Beispiel sei die Berechnung von Dispersion genannt, die es ermöglichte, die Gesamtzahlen aus den Subkorpora in die richtige Perspektive zu rücken. Konkret bestand eine relativ breite Kluft zwischen den Muttersprachlern und den Nichtmuttersprachlern im Hinblick auf die Gleichmäßigkeit der Verteilung von den untersuchten Lexemen über die einzelnen Dokumente – die Muttersprachler waren deutlich konsistenter.

In den theoretischen Kapiteln wurden neben der theoretischen Abhandlung der Problematik von Wissenschaftssprache, Kollokationen und Verweisung auch die Visualisierungsmöglichkeiten in der neuen Version vom Korpusmanager Sketch Engine behandelt. Konkreter war ihr Potenzial für die Lexikografie von Interesse, das ohne menschlichen Eingriff immer noch relativ gering ausfällt; erwähnt sei diesbezüglich die mitunter ungenaue Identifikation der Gramrels, die eventuellen Gramrel-Überlappungen und eingeschränkte Möglichkeit weiterer Einstellungen (Größe, Maßstab, Bearbeitung). Wenn die Visualisierungen jedoch nur ein Teil eines größeren Ganzen, in diesem Fall der Wortprofile, sind, stellen sie eine nützliche Ergänzung, die zum schnellen Überblick verhilft. Am geeignetsten erwiesen sich die Kreisdiagramme, die folglich jedem Profil beiliegen. Neben den Visualisierungen enthält jedes Profil einen Kommentar, der dem Leser die statistische Auswertung sowie weitere Informationen über den Gebrauch näherbringt. Den Rest des Wortprofils bilden ausgewählte Beispielsätze mit Kollokaten des jeweiligen Lexems. Auf diese Weise wird der muttersprachliche Gebrauch anschaulich illustriert, ohne dass man unbedingt direkt auf ein Korpus zurückgreifen müsste.

Da die Wortprofilierung als didaktisch motiviert aufgefasst wurde, habe ich den letzten Teil der Arbeit ihrer Didaktisierung gewidmet. Es wurde ein Exkurs in den Gebrauch von Korpora im Fremdsprachunterricht unternommen und im Lichte dessen wird eine Sammlung von 123 Fragen zusammengestellt, die (nicht nur) den Gebrauch der 15 Abstrakta in der studentischen Wissenschaftssprache zum Ausgangspunkt haben.

Nach der eventuellen Erweiterung des Korpus durch Bachelor- und Doktorarbeiten wäre es in der Zukunft von Interesse, eine Längsschnittstudie durchzuführen, in der die Entwicklung der Schreibkompetenzen von Studierenden über die drei Bildungsstufen analysiert würde. Dies betrifft natürlich nicht nur Abstrakta, sondern auch andere Komponenten der allgemeinen Wissenschaftssprache.

Literaturverzeichnis

- Andresen, M., & Knorr, D. (2017). *KoLaS: Kommentiertes Lernendenkorpus akademisches Schreiben*. Abgerufen am 14. Februar 2018 von Universitätskolleg: Universität Hamburg: <https://www.universitaetskolleg.uni-hamburg.de/projekte/tp05/tp05-media/andresen-knorr-2015-kolas.pdf>
- Biber, D. (1993). Representativeness in Corpus Design. *Literary and Linguistic Computing*, 8(4), S. 243–257.
- Biber, D., Reppen, R., Schnur, E., & Ghanem, R. (2016). On the (non)utility of Juilland's D to measure lexical dispersion in large corpora. *International Journal of Corpus Linguistics*, 21(4), S. 439–464. doi:10.1075/ijcl.21.4.01bib
- Boulton, A. (2011). Data-driven learning: the perpetual enigma. In S. Roszkowski, & B. Lewandowska-Tomaszczyk (Hrsg.), *Explorations across Languages and Corpora* (S. 563–580). Frankfurt: Peter Lang.
- Brommer, S. (2018). *Sprachliche Muster: Eine induktive korpuslinguistische Analyse wissenschaftlicher Texte*. Berlin: De Gruyter.
- Brommer, S. (2019). Sprachlich Verfestigtes analysieren. Theoretische und methodische Überlegungen. *Linguistik Online*, 96(3), S. 9–23. doi:10.13092/lo.96.5531
- Bubenhofer, N. (2018). Visualisierungen in der Korpuslinguistik: Diagrammatische Operationen zur Gegenstandskonstitution, -analyse und Ergebnispräsentation. In M. Kupietz, & T. Schmidt (Hrsg.), *Korpuslinguistik* (S. 27–60). Berlin: De Gruyter.
- Buhofer, A. H., Dräger, M., Meier, S., & Roth, T. (2014). *Feste Wortverbindungen des Deutschen: Kollokationswörterbuch für den Alltag*. Tübingen: Francke Verlag.
- Bußmann, H. (2008). *Lexikon der Sprachwissenschaft*. Stuttgart: Alfred Kröner Verlag.
- Conte, M. (1996). Anaphoric Encapsulation. *Belgian Journal of Linguistics*, S. 1–10.
- Cvrček, V. (2020). *Calc: Korpusová kalkulačka*, 1.01. Abgerufen am 27. März 2020 von Český národní korpus: Portál: <https://www.korpus.cz/calc/>
- Davies, M. (2015). Corpora: an introduction. In D. Biber, & R. Reppen (Hrsg.), *The Cambridge Handbook of English Corpus Linguistics* (S. 11–31). Cambridge: Cambridge University Press.
- Drumbl, H. (2011). Sprachdidaktik und Korpora. In A. Abel, & R. Zanin (Hrsg.), *Korpora in Lehre und Forschung* (S. 55–100). Bolzano: Bozen-Bolzano University Press.
- Dubova, A., & Proveja, E. (2016). Datengeleitetes Lernen im studienbegleitenden Deutschunterricht am Beispiel des KoGloss-Ansatzes. *Zeitschrift für Interkulturellen Fremdsprachenunterricht*, 21(1), S. 85–109.
- Dudenredaktion (2009). *Duden: Die Grammatik* (8. Ausg.). Berlin: Bibliographisches Institut.

- Řurčo, M., & Řurčo, P. (2012). *German Word Sketches and prepositional Phrases in German*. Abgerufen am 5. March 2020 von academia.edu: https://www.academia.edu/22514580/German_Word_Sketches_and_prepositional_Phrases_in_German
- Řurčo, P. (2016). Zum Konzept der Kollokationsdidaktik und des Kollokationslernens bei Germanistikstudenten. In P. Řurčo (Hrsg.), *Kollokationsforschung und Kollokationsdidaktik* (S. 153–178). Münster: LIT Verlag.
- Ehlich, K. (1993). Deutsch als fremde Wissenschaftssprache. In *Jahrbuch Deutsch als Fremdsprache* (Bd. 19, S. 13–42). München: Iudicum.
- Ehlich, K. (1999). Alltägliche Wissenschaftssprache. *Info DaF*, 26(1), S. 3–24.
- Ehlich, K. (2000). Deutsch als Wissenschaftssprache für das 21. Jahrhundert. *German as a foreign language (GFL)*, 1(1), S. 47–63.
- Eisenberg, P. (2013). *Grundriss der Deutschen Grammatik Band 2: Der Satz* (4. Ausg., Bd. 2). Stuttgart: Metzler.
- FF CUNI. (2020). *Přijímací zkoušky - Německý jazyk a literatura - Bakalářské studium*. Abgerufen am 24. Januar 2020 von Ústav germánských studií filozofické fakulty Univerzity Karlovy: <https://german.ff.cuni.cz/cs/uchazec/prijimaci-zkousky/nemecky-jazyk-a-literatura-bakalarske-studium/>
- FF MU. (2020). *Německý jazyk a literatura – bakalářské studium*. Abgerufen am 24. Januar 2020 von Nabídka studia: <https://www.muni.cz/bakalarske-a-magisterske-obory/24300-nemecky-jazyk-a-literatura#cile>
- FF UJEP. (2020). *Podmínky přijímacího řízení na KGER*. Abgerufen am 24. Januar 2020 von Katedra germanistiky filozofické fakulty Univerzity Jana Evangelisty Purkyně v Ústí nad Labem: <http://ff.ujep.cz/kger/studium/prijimaci-rizeni2>
- FF UPOL. (2020). *Přijímací řízení*. Abgerufen am 24. Januar 2020 von Katedra germanistiky Filozofické fakulty Univerzity Palackého v Olomouci: <https://germanistika.upol.cz/uchazecum/prijimaci-rizeni/>
- Firth, J. R. (1957). Modes of Meaning. In J. R. Firth, *Papers in Linguistics, 1934-1951* (S. 190–213). London: Oxford University Press.
- Flowerdew, J., & Forest, R. W. (2015). *Signalling Nouns in English: A Corpus-Based Discourse Approach*. New York : Cambridge University Press.
- Flowerdew, L. (2012). *Corpora and Language Education*. Basingstoke: Palgrave Macmillan.
- Geist, M., & Hahn, A. (2012). Using a corpus for written production: A classroom study. In J. Thomas, & A. Boulton (Hrsg.), *Input, Process and Product: Developments in Teaching and Language Corpora* (S. 123–133). Brno: Masaryk University Press.
- Goldhahn, A. (2010). Wissenschaftstexte kontrastiv: Korpusdesign. In I. Kratochvílová, & N. R. Wolf (Hrsg.), *Kompendium Korpuslinguistik: eine Bestandsaufnahme aus deutsch-tschechischer Perspektive* (S. 229–234). Heidelberg: Winter.

- Goldhahn, A. (2016). *Tschechische und deutsche Wissenschaftssprache im Vergleich. Eine kontrastive Analyse tschechischer und deutscher Wissenschaftlicher Artikel der Linguistik*. Brno: Masarykova Univerzita, (Diss.).
- Graefen, G. (1997). *Der Wissenschaftliche Artikel – Textart und Textorganisation*. Frankfurt am Main: Peter Lang.
- Graefen, G. (2000). “Hedging” als neue Kategorie? Ein Beitrag zur Diskussion. S. 1–13. Abgerufen am 20. März 2020 von <https://www.daf.uni-muenchen.de/media/downloads/hedge.pdf>
- Graefen, G. (2009). Die Didaktik des wissenschaftlichen Schreibens: Möglichkeiten der Umsetzung. *German as a foreign language (GFL)*, 10(2-3), S. 106–127.
- Graefen, G., & Moll, M. (2011). *Wissenschaftssprache Deutsch: lesen – verstehen – schreiben: ein Lehr- und Arbeitsbuch*. Frankfurt am Main: Peter Lang.
- Gräfe, K., Lange, D., Sieradz, M., Meißner, C., Slavcheva, A., & Stoppel, D. (2013). *GeWiss: Handbuch zum Korpus*. Abgerufen am 20. Februar 2020 von uni-leipzig.de: <https://gewiss.uni-leipzig.de/fileadmin/documents/Handbuch.pdf>
- Gries, S. T. (2008). Dispersions and adjusted frequencies in corpora. *International Journal of Corpus Linguistics*, 13(4), S. 403–437. doi:10.1075/ijcl.13.4.02gri
- Gries, S. T. (im Druck). Analyzing dispersion. In M. Paquot, & S. T. Gries (Hrsg.), *The Practical Handbook of Corpus Linguistics* (S. 1–16). Springer. Abgerufen am 15. März 2020 von http://www.stgries.info/research/ToApp_STG_Dispersion_PHCL.pdf
- Halliday, M., & Hasan, R. (1976). Lexical Cohesion. In M. Halliday, & R. Hasan, *Cohesion in English* (S. 276–292). London: Longmans.
- Hanks, P. (2013). Chapter 1: Words and Meanings: The Need for a New Approach. In P. Hanks, *Lexical Analysis: Norms and Exploitations* (S. 1–23). Cambridge, MA: MIT Press.
- Hanzlíčková, A. (2016). Korpusbasierte Kollokationsforschung und ihre Realisierung in den Wörterbüchern. *Slowakische Zeitschrift für Germanistik*, 8(1), S. 41–53.
- Higgins, J. (1983). Computer assisted language learning. *Language Teaching*, 16(2), S. 102–114. doi:10.1017/S0261444800009988
- Hollós, Z. (2010). Auf dem schmalen Grad zwischen Kollokationen und festen Wortverbindungen. In P. Ďurčo (Hrsg.), *Feste Wortverbindungen und Lexikographie: Kolloquium zur Lexikographie und Wörterbuchforschung* (S. 85–96). Berlin: De Gruyter.
- Hollós, Z. (2014a). *KOLLEX. Deutsch-ungarisches KOLLOkationsLEXikon. Korpusbasiertes Wörterbuch der Kollokationen. Deutsch als Fremdsprache*. Szeged: Grimm Kiadó.
- Hollós, Z. (2014b). Syntagmatik im KOLLEX: Die lexikographische Darstellung grammatischer Syntagmatik in einem zweisprachigen Kollokationslexikon für Deutschlerner. In M. J. Vázquez, F. Mollica, & M. N. Curcio, *Zweisprachige Lexikographie zwischen Translation und Didaktik* (S. 113–129). Berlin/Boston: De Gruyter.

- Hotařová, M. (2019). *Typische Wortverbindungen in der Wissenschaftssprache. Eine korpuslinguistische kontrastive n-Gramm-Analyse*. (unv. Diss.). Abgerufen am 15. Februar 2020 von is.muni.cz: <https://is.muni.cz/auth/th/v2rwb/>
- Huneke, H.-W., & Steinig, W. (2013). Fremdsprachliche Lernwelten. In H.-W. Huneke, & W. Steinig, *Deutsch als Fremdsprache: Eine Einführung* (6. Ausg., S. 117–140). Berlin: Erich Schmidt Verlag.
- Hunston, S. (2002). *Corpora in Applied Linguistics*. Cambridge: Cambridge University Press.
- Ivanič, R. (1991). Nouns in search of a context: A study of nouns with both open- and closed-system characteristics. *International Review of Applied Linguistics in Language Teaching*, 2, S. 93–114.
- Joe, A. (2010). The quality and frequency of encounters with vocabulary in an English for Academic Purposes programme. *Reading in a Foreign Language*, 22(1), S. 117–138.
- Johns, T. (1991). Should you be persuaded: Two examples of data-driven learning. In T. Johns, & P. King (Hrsg.), *Classroom Concordancing. English Language Research Journal*, 4, S. 1–14.
- Johns, T. (o. J.). *Tim Johns' Kibbitzers*. Abgerufen am 1. April 2020 von WordSmith Tool: <https://lexically.net/TimJohns/>
- Káňa, T. (2014). *Sprachkorpora in Unterricht und Forschung DaF/DaZ*. Brno: Masarykova Univerzita.
- Kilgariff, A., Rychlý, P., Smrž, P., & Tugwell, D. (2004). The Sketch Engine. *Proceedings of the Eleventh EURALEX International Congress*, S. 105–116.
- Kilgarriff, A., Baisa, V., Bušta, J., Jakubíček, M., Kovář, V., Michelfeit, J., Rychlý, P., Suchomel, V. (2014). The Sketch Engine: ten years on. *Lexicography*, 1, S. 7–36.
- Kilgarriff, A., Baisa, V., Rychlý, P., & Jakubíček, M. (2015). Longest–commonest Match. *Proceedings of Corpus Linguistics (CL2015)*, (S. 397–404). Brighton.
- Kilgarriff, A., Husák, M., McAdam, K., Rundell, M., & Rychlý, P. (2008). GDEX: automatically finding good dictionary examples in a corpus. In E. Bernal, & J. DeCesaris (Hrsg.), *Proceedings of the XIII EURALEX International Congress* (S. 425–432). Barcelona: Documenta Universitaria.
- Kilgarriff, A., Kovář, V., Krek, S., Srdanovič, I., & Tiberius, C. (2010). A quantitative evaluation of word sketches. *Proceedings of the 14th EURALEX International Congress*, S. 372–379.
- Kocincová, L. (2015). *Interactive Visualisation Methods for Sketch Engine*. Brno: MU Brno, (unv. Magisterarbeit.). Abgerufen am 13. Oktober 2019 von <https://is.muni.cz/auth/th/odmai>
- Kocincová, L., Baisa, V., Jakubíček, M., & Kovář, V. (2015). Interactive Visualizations of Corpus Data in Sketch Engine. *Innovative Corpus Query and Visualization Tools* (S. 17–22). NODALIDA 2015.

- Koepfel, R. (1993). *Satzbezogene Verweisformen: Eine datenbankgestützte Untersuchung zu ihrer Distribution und Funktion in mündlichen Texten, schriftlichen Texten und schriftlichen Fachtexten des Deutschen*. Gunter Narr Verlag: Tübingen.
- Kumaravadivelu, B. (2001). Toward a Postmethod Pedagogy. *TESOL Quarterly*, 35(4), S. 537–560.
- Laufer, B., & Rozovski-Roitblat, B. (2015). Retention of new words: Quantity of encounters, quality of task, and degree of knowledge. *Language Teaching Research*, 19(6), S. 687–711. doi:10.1177/1362168814559797
- Lemnitzer, L., & Zinsmeister, H. (2010). *Korpuslinguistik: Eine Einführung* (2. Ausg.). Tübingen: Narr Verlag.
- Lijffijt, J., & Gries, S. T. (2012). Correction to Stefan Th. Gries' "Dispersions and adjusted frequencies in corpora". *International Journal of Corpus Linguistics*, 17(1), S. 147–149.
- Lüdeling, A., & Walter, M. (2010). Korpuslinguistik. In H.-J. Krumm, C. Fandrych, B. Hufeisen, & C. Riemer (Hrsg.), *Deutsch als Fremd- und Zweitsprache: Ein internationales Handbuch* (Bd. 1, S. 315–322). Berlin: De Gruyter.
- Lutzeier, P. R. (2007). *Lexikologie: Ein Arbeitsbuch*. Tübingen: Stauffenburg Verlag.
- Mach, T. (im Druck). Zur Leistung der Abstrakta in der studentischen Wissenschaftssprache: ein quantitativer L1-L2-Vergleich. *Linguistik Online*, S. 1–29 [nicht formatiert].
- McEnery, T., & Hardie, A. (2012). What is corpus linguistics? In T. McEnery, & A. Hardie, *Corpus Linguistics: Method, Theory and Practice* (S. 1–24). Cambridge: Cambridge University Press.
- McEnery, T., & Wilson, A. (2001). Quantitative data. In T. McEnery, & A. Wilson, *Corpus Linguistics: An Introduction* (2. Ausg., S. 75–101). Edinburgh: Edinburgh University Press.
- Mollica, F. (2010). Das Korrelat *es*₀ in Verbindung mit einer Adjektival- bzw. Nominalergänzung. In F. Mollica, *Korrelate im Deutschen und im Italienischen* (S. 85–98). Frankfurt am Main: Peter Lang.
- Nation, I. S. (2013). *Learning Vocabulary in Another Language* (2. Ausg.). Cambridge: Cambridge University Press.
- Pecorari, D. (2003). Good and original: Plagiarism and patchwriting in academic second-language writing. *Journal of Second Language Writing*, 13, S. 317–345. doi:10.1016/j.jslw.2003.08.004
- Pittner, K., & Berman, J. (2015). *Deutsche Syntax: Ein Arbeitsbuch* (6. Ausg.). Tübingen: Narr.
- Porzig, W. (1934). Wesenhafte Bedeutungsbeziehungen. *Beiträge zur Geschichte der deutschen Sprache*, 58, S. 70–97.
- Quasthoff, U. (2009). Korpusbasierte Wörterbucharbeit mit den Daten des Projekts Deutscher Wortschatz. *Linguistik Online*, 39(3). doi:10.13092/lo.39.484
- Quasthoff, U. (2011). *Wörterbuch der Kollokationen im Deutschen*. Berlin: De Gruyter.

- Reder, A. (2006). *Kollokationen in der Wortschatzarbeit*. Wien: Praesens Verlag.
- Reder, A. (2011). Kommen Kollokationen in Mode? Kollokationskonzepte und ihre mögliche Umsetzung in der Didaktik. *Linguistik Online*, 47(3), S. 131–140.
- Reznicek, M., Lüdeling, A., Krummes, C., Schwantuschke, F., Walter, M., Schmidt, K., Hirschmann, H., Andreas, T. (26. September 2012). Das Falko-Handbuch: Korpusaufbau und Annotationen Version 2.01. Berlin, Deutschland. Abgerufen am 23. Februar 2020 von <http://www.linguistik.hu-berlin.de/institut/professuren/korpuslinguistik/forschung/falko>
- Rolek, B. (2018). Hedging in wissenschaftlichen Artikeln polnischer und deutscher Studierender. In L. Cirko, & K. Pittner (Hrsg.), *Wissenschaftliches Schreiben interkulturell: Kontrastive Perspektiven* (S. 225–247). Berlin: Peter Lang.
- Römer, U. (2012). Corpora and teaching academic writing: Exploring the pedagogical potential of MICUSP. In J. Thomas, & A. Boulton (Hrsg.), *Input, Process and Product: Developments in Teaching and Language Corpora* (S. 70–82). Brno: Masaryk University Press.
- Rychlý, P. (2008). A Lexicographer-Friendly Association Score. In P. Sojka, & H. (. A., *Proceedings of Recent Advances in Slavonic Natural Language Processing* (pp. 6–9). Brno: Masaryk University.
- Schlaefler, M. (2009). *Lexikologie und Lexikographie: Eine Einführung am Beispiel deutscher Wörterbücher* (2. Ausg.). Berlin: Erich Schmidt Verlag.
- Schmid, H.-J. (1998). Constant and Ephemeral Hypostatization: THING, PROBLEM and Other "Shell Nouns". In B. Caron (Hrsg.), *Proceedings of the 16th International Congress of Linguists*. Paris: Elsevier.
- Schmid, H.-J. (2000). *English Abstract Nouns as Conceptual Shells: From Corpus to Cognition*. Berlin: Mouton de Gruyter.
- Sinclair, J. M. (1991). *Corpus, concordance and collocation*. Oxford: Oxford University Press.
- Steinhoff, T. (2007). Zum „ich“-Gebrauch in wissenschaftlichen Texten. *Zeitschrift für germanistische Linguistik*, 35, S. 1–26.
- Steyer, K. (2000). Usuelle Wortverbindungen des Deutschen. *Deutsche Sprache*, 28(2), S. 101–125.
- Steyer, K. (2008). Kollokationen in deutschen Wörterbüchern und in der deutschen Wörterbuchforschung. *Lexicographica*, 24, S. 185–207.
- Steyer, K. (2009). Zwischen theoretischer Modellierung und praxisnaher Anwendung. Zur korpusgesteuerten Beschreibung usueller Wortverbindungen. In C. Mellado Blanco (Hrsg.), *Theorie und Praxis der idiomatischen Wörterbücher* (S. 119–145). Berlin/New York: De Gruyter.
- Thomas, J. (2017). *Discovering English with Sketch Engine: a corpus-based approach to language exploration* (2. Ausg.). Brno: Versatile.

- Tomášková, S. (2016). Kollokationen in der Wissenschaftssprache am Beispiel von Astrophysik. In P. Ďurčo (Hrsg.), *Kollokationsforschung und Kollokationsdidaktik* (S. 127–142). Münster: LIT Verlag.
- Viana, V. (2020). *New ways in Teaching with Corpora*. Alexandria, VA: TESOL Press.
- Völz, I. (2015). *Lexikalische Textgliederung beim wissenschaftlichen Schreiben in der Fremdsprache Deutsch: Eine empirische Untersuchung zum Erwerb und Gebrauch textorganisierender Ausdrücke durch internationale DaF-Studierende*. Kassel: kassel university press.
- Wallner, F. (2013). Korpora im DaF-Unterricht – Potentiale und Perspektiven am Beispiel des DWDS. *Revista Nebrija de Lingüística Aplicada*, 13, o. S.
- Xiao, R. (2015). Collocation. In D. Biber, & R. Reppen (Hrsg.), *The Cambridge Handbook of English Corpus Linguistic* (S. 106–124). Cambridge: Cambridge University Press.
- Ylönen, S. (2012). Qualitative und quantitative Methoden datengeleiteten Lernens. *German as a foreign language*, S. 75–113.