

Britt-Marie Schuster & Christopher Georgi

Beispiele geben, Belege anführen und Daten erheben: Diachrone Untersuchungen zur Herstellung empirischer Evidenz

Abstract: In der Linguistik ist die Bezugnahme auf sprachliches Material, das in Abhängigkeit vom jeweiligen Erkenntnisinteresse selektiert und präsentiert wird, charakteristisch. Im folgenden Beitrag stehen textuelle Mikropraktiken im Zentrum, die den Umgang mit Sprachmaterial prägen. Diese dienen dazu, Evidenz herzustellen. Sprache wird nicht allein sichtbar gemacht, sondern die (vermeintliche) Offenkundigkeit eines sprachlichen Phänomens leitet ebenso linguistische Kategorisierungen wie die wissenschaftliche Argumentation. Die Praktiken erschließen sich, wie der Beitrag zeigen möchte, durch die Verwendung der Ausdrücke *Beispiel*, *Beleg* und *Daten*. Ziel des Beitrags ist es, zu skizzieren, ob und inwieweit sich der Umgang mit sprachlichem Material in den letzten fünfzig Jahren gewandelt hat und sich unterschiedliche linguistische Subdisziplinen hinsichtlich der Verwendung von sprachlichem Material unterscheiden lassen. Der Beitrag nutzt das RGL-Korpus: Einerseits wurden diesem 30 Werke entnommen, die qualitativ auf ihre Verwendung von *Beispiel*, *Beleg* und *Daten* hin untersucht wurden, andererseits wurden sowohl gezielte, hypothesengeleitete Suchabfragen an das Korpus gestellt als auch korpusbasiert zentrale paradigmatische und syntagmatische Relationen ermittelt. Obgleich die RGL nicht für die Gesamtentwicklung der Linguistik stehen kann, zeigt sich an ihr zum einen exemplarisch ein zunehmend stärker werdendes Selbstverständnis der Linguistik als empirisch arbeitende Disziplin. Zum anderen wird deutlich, dass sich unterschiedliche Verwendungstraditionen für *Beispiel*, *Beleg* oder *Daten* herausgebildet haben, die gleichermaßen Indiz für unterschiedliche Zugangsweisen zum Sprachmaterial sind. So korrespondiert Beispielgebung mit Untersuchungen, die sich an wenigen Kommunikaten orientieren, während das Anführen von Belegen eher mit Quantifizierungen auf der Basis von Korpora einhergeht. Datenerhebungen wiederum haben sich in vielen Subdisziplinen durchgesetzt.¹

¹ Wir danken Mechthild Habermann und Heiko Hausendorf für hilfreiche Hinweise und Anregungen zur Überarbeitung einer vorausgehenden Fassung dieses Beitrags sowie Anna Abbenhaus für die Unterstützung im Rahmen der Formatierung und Finalisierung des Beitrags.

Britt-Marie Schuster, Christopher Georgi, Universität Paderborn

Keywords: *Beispiel, Beleg, Daten*, authentisches Beispiel, konstruiertes Beispiel, Repräsentativität, Kollokation, Word2Vec-Modell, empirische Evidenz

1 Einleitung

Für die Linguistik ist die Präsentation von Sprache, auch in dezidiert sprachtheoretischen Schriften, charakteristisch. Die Auseinandersetzung mit als *Beispiel, Beleg* oder *Daten*² bezeichnetem Sprachmaterial prägt Selbstverständnis, Forschungspraxis und Image der Disziplin. Ob es sich dabei um Wörter, Phrasen, Sätze, um Textausschnitte oder Auszüge aus Gesprächstranskripten handelt, hängt davon ab, welcher linguistischen Subdisziplin die jeweilige Veröffentlichung entstammt.

Durch die Präsentation von Sprache wird grundsätzlich Evidenz hergestellt. Sprache wird nicht allein sichtbar gemacht, sondern die präsentierte Sprache hilft bei der Schärfung linguistischer Kategorien, bei der Demonstration von Analyseverfahren oder bei der Stützung von Argumentationen. Die damit verbundenen Praktiken bilden unseres Erachtens ein zentrales Element der linguistischen Wissenschaftskommunikation, das ebenso wie die linguistische Wissenschaftskommunikation selbst kaum hinsichtlich zentraler Muster und Funktionen untersucht worden ist.³ Deshalb ist bisher unklar, auf welche Weise die Präsentation von Sprache je nach linguistischer Subdisziplin, gewähltem Thema und Erkenntnisinteresse variiert und unterschiedliche Schreibtraditionen begründet. Das heißt ausdrücklich nicht, dass nicht in unterschiedlichen Teildisziplinen über die Güte und den Wert von Beispielen, Belegen oder Daten nachgedacht würde: In der Argumentationstheorie spielt das Begründen durch Beispiele ebenso eine Rolle wie die Auseinandersetzung mit der statistischen Signifikanz von Belegen in der Korpuslinguistik. In der Konversationsanalyse ist der Stellenwert von Fallbeispielen und „Kollektionen“ häufiger diskutiert worden und auch in der Diskurslinguistik spielen Kriterien wie Reliabilität bei der Korpuszusammenstellung eine größere Rolle. Darüber hinaus zeigen viele Einzeluntersuchungen, die Teil des

2 Aus der explorativen Durchsicht von 30 RGL-Monografien, die in den letzten 50 Jahren erschienen sind, hat sich ergeben, dass die genannten Lexeme und entsprechende Kompositionen wie etwa *Belegmaterial, Fallbeispiel* u. ä. für die Kennzeichnung von Sprachmaterial zentral sind. Sicher sind dies nicht die einzigen Lexeme, auch muss Sprachmaterial nicht notwendigerweise gekennzeichnet werden. *Datum* wird sehr selten verwendet, weshalb wir in diesen Beitrag nahezu durchgängig von *Daten* sprechen.

3 Vgl. jedoch die Ausführungen von Kalwa, Merten und Bojarski, Huber & Bubenhofer in diesem Band sowie Roelcke (2013).

RGL-Korpus' sind, dass über die Rolle, etwa von authentischen und konstruierten Beispielen, reflektiert wird und versucht wird, Korpora und die spezifische Auswahl von Beispielen zumindest übergreifend hinsichtlich ihrer Repräsentativität sowie weiterer Standards zu begründen. Was indes zu fehlen scheint, sind empirische Untersuchungen zu Verwendungen von *Beispielen*, *Belegen* und *Daten* und eine übergreifende, theoretische Diskussion des Sprachmaterials in der Linguistik überhaupt, die erkennen ließe, was ein gutes Beispiel oder ein überzeugender Datensatz ist und an welchen Kriterien dies/dieser zu messen ist.

Grundsätzlich dürfte es allein schon einen Unterschied machen, ob etwas als *Beispiel* oder *Beleg* oder *Daten* bezeichnet wird. Ein Beispiel legt die Existenz ähnlich gearteter Kommunikate nahe. Seine Auswahl dürfte danach erfolgen, ob es für eine Gesamtmenge als prototypisch zu erachten ist. Mit der Bezeichnung *Beleg* lässt sich eine andere Perspektive verbinden: Ein sprachliches Phänomen wird dokumentiert (etymologisch verwandt mit *testimonium*, *documentum*). Da ein sprachliches Phänomen zumeist nicht ausreichend durch einen Beleg dokumentierbar ist, dürfte es sich als notwendig erweisen, einen Einblick in die Gesamtmenge gleichgearteter Belege zu geben. Diese Unterschiede, die sich aus der Semantik der Ausdrücke ergeben, spiegeln sich auch, wie zu sehen sein wird, in ihrer Verwendung: *Belege* werden häufig quantifiziert, was bei *Beispielen* kaum der Fall ist. Auf *Daten* wiederum wird rekuriert, wenn auf eine Gesamtmenge von Kommunikaten Bezug genommen wird, die unter einem bestimmten Gesichtspunkt entstanden bzw. erhoben und/oder die in einer bestimmten Weise behandelt bzw. aggregiert worden sind.

Im Folgenden soll nun gezeigt werden, wie sich der Umgang mit *Beispielen*, *Belegen* und *Daten* von 1975 bis in die unmittelbare Gegenwart entwickelt hat. Mit der Reihe Germanistische Linguistik (RGL) wird eine thematisch offene Reihe gewählt. Die Reihe zeichnet, wie auch der folgende Beitrag darstellen wird, aus, dass in vielen Publikationen mit authentischem Sprachmaterial gearbeitet wird, das oft, wengleich nicht immer, in seiner situativen, kulturellen oder medialen Bedingtheit kontextualisiert wird. Da nun in der RGL Autor*innen unterschiedlicher linguistischer Fächer zu Wort kommen, lassen sich unseres Erachtens unterschiedliche Arbeitsweisen herausarbeiten. Die folgenden Ausführungen beziehen sich zum einen auf die Nutzung des RGL-Korpus⁴ (Bojarski, Huber & Bubenhofer 2024), zum anderen auf 30 ausgesuchte Monografien und Sammelbände, die sich über die fünf

4 Das RGL-Korpus umfasst nahezu alle Bände seit 1975 und wurde mit freundlicher Genehmigung durch den de Gruyter-Verlag als Analysegrundlage für die Tagung „50 Jahre RGL: Germanistische Linguistik – in Vergangenheit, Gegenwart und Zukunft“ vom 15. bis 16. Juni 2023 erstellt. Da die PDF-Dateien bereits mittels OCR generierten Text enthielten, konnte dieser extrahiert werden. Anschließend wurde der Text semi-automatisiert bereinigt, indem etwa Literaturverzeich-

Jahrzehnte des Erscheinens der RGL erstrecken und sich auf zentrale Themenfelder der RGL wie Grammatik, Gesprächs- und Textanalyse, Variationslinguistik, Pragmatik und auch die Schreibforschung beziehen (zu den zentralen Themen siehe auch den Beitrag von Bojarski, Huber & Bubenhofer im vorliegenden Band). Anhand beider Zugangsweisen soll gezeigt werden, dass allein schon die Bezeichnung von Sprachmaterial als *Beispiel*, *Beleg* oder *Daten* Aufschluss über Entwicklungstendenzen in der Sprachwissenschaft gibt. Darüber hinaus lässt sich ermitteln, dass sich jeweils typische syntagmatische und paradigmatische Relationen mit diesen Bezeichnungen verbinden, die Muster offenlegen, die für bestimmte Forschungstraditionen, -gebiete und -methoden charakteristisch sind.

Der Beitrag ist folgendermaßen gegliedert: Zunächst soll die wissenschaftstheoretische Relevanz des Themas betont werden (s. u. Abschnitt 2). Dann sollen die Methoden (s. u. Abschnitt 3) sowie empirische Studien zum Anführen von *Beispielen*, *Belegen* und *Daten* präsentiert werden (s. u. Abschnitt 4). Schließlich sollen in Abschnitt 5 die unterschiedlichen wissenschaftlichen Traditionen in der RGL diskutiert werden.

2 Die Relevanz des Themas

Die Bezugnahme auf *Beispiele*, das Anführen von *Belegen* und der Rekurs auf *Datensätze* finden sich in vielen wissenschaftlichen Disziplinen. So ist Beispielgebung nicht nur in der Sprachwissenschaft verbreitet: Große Bedeutung hat etwa das Fallbeispiel in der medizinischen, v. a. psychiatrischen, und psychologischen Literatur. Auch in der Philosophie spielt das Beispiel, etwa in Wittgensteins späterer Philosophie, eine größere Rolle. In der Literaturwissenschaft ist die Arbeit mit Beispielen ebenfalls gängig und dient unterschiedlichen Zwecken. Erst seit gut zwei Jahrzehnten wird etwa die Beispielgebung in den Kulturwissenschaften intensiver beforscht, zuvor war sie „meist ein blinder Fleck der Wissensproduktion“ (Willer 2013: 271) und soll zu den „unmerklichen Formen der Plausibilisierung von Diskursen“ (Willer 2013: 271) gehört haben.⁵

Unseres Erachtens lassen sich Beispiele, Belege oder Daten aus zwei unterschiedlichen Perspektiven betrachten. Zum einen können ihre Rolle bei der Ausar-

nisse, Anhänge und Kopfzeilen entfernt bzw. zum Teil gesondert untersucht wurden sowie getrennte Wörter zusammengeführt wurden.

⁵ An dem in literatur-, kulturwissenschaftlichen und philosophischen Sammelbänden geführten Diskurs über das Beispiel hat sich die (germanistische) Linguistik kaum beteiligt (vgl. u. a. Ruchatz, Willer & Pethes 2007 sowie Lück, Niehaus & Schneider 2013).

beitung wissenschaftlicher Erkenntnisse und die Qualität des durch sie erzeugten Wissens beleuchtet werden. Zum anderen können sie unter dem Gesichtspunkt ihrer Verwendung betrachtet werden.

Für die erste Perspektive ist das Spannungsfeld zwischen Besonderem (dem Beispiel, der Belegreihe etc.) und Allgemeinem, etwa der Formulierung von Regeln und Regularitäten oder der Annahme von Modellen, zentral. Außerdem verbindet sich mit der Frage nach dem Verhältnis zwischen Besonderem und Allgemeinem auch die hermeneutische Frage, wie Beispiele, Belege oder Daten ausgewählt werden, um Verallgemeinerungen zu ermöglichen. Dabei ergibt sich das grundsätzliche Problem, dass die Materialgrundlage zumeist nicht vollständig sichtbar werden kann. Darüber hinaus kann man sich dem Kontext, aus dem Sprachmaterial entstammt, zumeist nur annähern, und es stellt sich die grundsätzliche Frage, wie es durch diese unvermeidliche Dekontextualisierung angemessen verstanden werden kann.

Für die zweite Perspektive, die wir nun in diesem Beitrag verfolgen, ist hingegen zentral, wie Beispiele, Belege und Daten verwendet werden. Verwendungen können, wie aus dem folgenden Zitat hervorgeht, auf unterschiedlichen Ebenen ermittelt werden:

Das Beispiel, verstanden als Mittel der Plausibilisierung, Initiierung und Suggestibilisierung von Aussagen, Geschichten und Diskursen, ist in seiner wandelbaren Geschichte auf verschiedenste Weise in Gebrauch gewesen. Man könnte sich also eine Systematik vorstellen, die Verwendungszusammenhänge voneinander abgrenzt und auf diese Weise entschlüsselt, wie Beispiele jeweils unterschiedlich in wissenschaftlichen Argumentationen, in poetischen Texten und in populären Ikonisierungen zum Tragen kommen. Denkbar wäre ebenfalls eine Sortierung nach textuellen und paratextuellen Verfahren bzw. Erscheinungsformen, die zur Herauspräparierung von Beispielen beitragen: Fabeln und Fallgeschichten, Listen und Zettelkästen, Parenthesen und Doppelpunkte. (Ruchatz, Willer & Pethes 2007: 8)

Der Verweis auf Texttypen, Handlungsmuster („in wissenschaftlichen Argumentationen“), „textuelle und paratextuelle Verfahren“, Einzelelemente und nichtsprachliche Artefakte („Zettelkästen“) macht deutlich, welche Ebenen bei der Skizze von Verwendungszusammenhängen zu berücksichtigen wären und wie umfangreich das nicht nur mit Beispielen, sondern auch mit Belegen und Daten verbundene Forschungsfeld sein könnte. In diesem Beitrag stehen v. a. Äußerungen im Vordergrund (s. zur Begründung Abschnitt 3), die auf Sprachmaterial verweisen, dieses kategorisieren und reflektieren. Aus ihrer Untersuchung lässt sich auf Mikropraktiken schließen, die für die Präsentation sprachlichen Materials und die Evidenzherstellung charakteristisch sind. Davon könnten Mesopraktiken (etwa das Argumentieren), jedoch auch Makropraktiken wie etwa die o. g. paratextuellen Verfahren unterschieden werden.

Außerhalb der Linguistik, in kulturwissenschaftlichen Veröffentlichungen beispielsweise, wird die Verwendung von *Beispielen* mit bestimmten Mikropraktiken identifiziert:

So gehört es zu den zwar umstrittenen, aber trotzdem unentwegt praktizierten Gewohnheiten in sprachwissenschaftlichen Darlegungen, für grammatische Muster – demonstrativ beliebige – Beispiele aufzubieten, um die Funktionsweise jener Muster vorzuführen. Damit sind linguistische Beispielsätze ihrerseits Musterbeispiele für die Funktionsweise nachträglicher, erläuternder, illustrierender Exemplifikation, insbesondere für die Frage, inwiefern Beispiele als Belege für gängige Regeln gelten können. (Ruchatz, Willer & Pethes 2007: 21)

Harald Weinrich wird zudem angeführt und dieser findet es „bemerkenswert, daß sich die Linguisten offenbar nur höchst ungern dazu herbeilassen, auf ihre kleinen und häufig läppischen Beispielsätze zu verzichten“ (Weinrich 1975: 439–440), plädiert für eine Text- und Pragmalinguistik und bemerkt weiterhin, dass „für jeden kleinen Beispielsatz die Reduktionsbedingungen [...], unter denen man ihn aus einem Text und einer komplexen Kommunikations-Situation herauspräpariert hat“ (Ruchatz, Willer & Pethes 2007: 22), mit zu erfassen sind. Die Sprachwissenschaft wird 2007 mit dem Verweis auf eine Kontroverse aus den späten 1970er Jahren als eine Disziplin wahrgenommen, in der Beispiele das Instrument sind, um grammatische Regeln zu formulieren, womit die Arbeitsfelder der Sprachwissenschaft selbstverständlich verengt wahrgenommen werden.

Die von Weinrich geforderte Pragmatisierung der Disziplin ist längst im Mainstream der Disziplin angekommen. Das heißt auch, dass die Linguistik wesentlich mehr als Beispielsätze in den Blick nimmt, Beispielsätze geradezu in Verfall gekommen sind und die Formulierung von Regeln nicht mehr das zentrale und durchgängige Interesse der Linguistik ist. Das durchnummerierte Anführen von Beispielsätzen, auf das in der entsprechenden Forschungsliteratur angespielt wird, ist eine Praktik in der Linguistik. Der Umgang mit Sprachmaterial ist jedoch wesentlich ausdifferenzierter.

Der folgende Beitrag möchte diesen Differenzierungsgrad beleuchten und konzentriert sich auf die Mikropraktiken der Evidenzherstellung, geht vom syntagmatischen Kontext aus und orientiert sich an den folgenden Fragen:

- Welche sprachlichen Einheiten (Lexem, Satz, Text etc.) werden überhaupt als *Beispiel*, *Beleg* oder *Daten* bezeichnet?
- Welche Qualitäten und Gütekriterien werden *Beispielen*, *Belegen* und *Daten* zugeschrieben? Werden diese etwa als *authentisch* oder *konstruiert* bezeichnet?
- Wird die Auswahl von *Beispielen*, *Belegen* oder *Daten* metakommunikativ reflektiert?
- Welcher Grad von Typizität in Bezug auf welche Bezugsgröße wird von den Forscher*innen angenommen?

- Welche Praktiken und Sprachgebrauchsmuster mit *Beispielen, Belegen* oder *Daten* zur Evidenzherstellung schälen sich heraus?
- Ist eine Bindung an bestimmte Arbeitsweisen und Methoden der Sprachwissenschaft erkennbar?
- Lässt sich die Entwicklung der Linguistik und linguistischer Teildisziplinen durch den Umgang mit dem entsprechenden Sprachmaterial erkennen?

3 Methodische Zugänge zur RGL

Um die genannten Fragen zu beantworten, wurde ein Zugang gewählt, der ein Close Reading von einzelnen Texten mit (halb)automatisierten Zugängen zum Gesamtkorpus verbindet. Es wurden zunächst 30 Monografien und Sammelbände durchgesehen, deren Zusammenstellung nach folgenden Kriterien erfolgte: Es sollten zum einen die unterschiedlichsten Zeiträume und zum anderen zentrale in der RGL vertretene linguistische Disziplinen wie Grammatik, Gesprächs- und Textanalyse, Variationslinguistik, Pragmatik und Schreibforschung vertreten sein. Durch dieses Vorgehen ergaben sich folgende Orientierungen für weitere korpuslinguistische Erhebungen:

- Durch die Bezeichnungen selbst und durch die mit ihnen gebildeten Komposita wird Sprachmaterial grundsätzlich identifizierbar. Dies kann als ein erstes Ergebnis verbucht werden.
- Auch Nummerierungen geben einen Hinweis auf das Vorliegen des entsprechenden sprachlichen Materials.
- Ferner lässt sich durch die Lektüre der einzelnen Monografien vermuten, dass sich für die jeweiligen Ausdrücke Formulierungsmuster mit einem besonderen Funktionspotential ergeben.
- Zudem stellte sich der Eindruck ein, dass sich allmählich etablierende Korpora auch eine zunehmend bedeutendere Rolle spielen könnten.

Diese Orientierungen bildeten den Ausgangspunkt für die Formulierung verschiedener Suchabfragen an das RGL-Korpus (vgl. Tab. 1). Ein selbst erstelltes Perl-Skript lieferte die Frequenz und die Einzelkontexte des jeweiligen Suchmusters im RGL-Korpus. Die Suchabfragen basierten auf verschiedenen regulären Ausdrücken, die es erlaubten, nach mehr oder weniger spezifischen Mustern zu suchen und etwa auch Komposita einzuschließen. Bei der Suche nach mehrteiligen Ausdrücken wurde lediglich die direkte Abfolge ohne mögliche Füllwerte berücksichtigt, da variable Abstände zwischen den Einzelbestandteilen des Suchmusters, selbst beschränkt auf den jeweiligen Einzelsatz, zu unspezifische Ergebnisse lie-

ferten. Die folgende Tabelle liefert einen Überblick über die verschiedenen Suchabfragen, geordnet nach Erkenntnisinteresse und formalen Aspekten:

Tab. 1: Linguistische Suchabfragen für das RGL-Korpus.

Lexeme (mit Nummerierung)	Typische Beispiele (Pragmatik)	Nummerierung	Metakommunikative Marker	Korpora
- Beispiel NUM (PUNCT)*	- Milchmann	- PUNCT NUM (PUNCT)*	- (illustrieren exemplifizieren	- (DTA DWDS)
- .*[B b]eleg.*?	- DET Eis sein	- (NUM)*	zeigen)	- TIGER
- belegen	- dünn	- [a-z]?	- beispielhaft	- Spiegel
- .*[B b]eispiel.*?	- Zitrone	(PUNCT)*	- in folgender Weise	- Zeit
- .*[D d]aten.*?		- PUNCT [a-z]?	- sehen wir uns an	- .*?(K k)
- Text.*?		PUNCT	- mögen genügen	- COSMAS
- .*?text		- römische	- (vergleichbar	
- Illustration		Zahlen	vergleichen)	
- .*?(T t)reffer.*?				
- Datum				
- Signifikanz				

Neben den Ausgangslexemen, die im Falle von *Beispiel NUM (PUNCT)** auch Nummerierungen wie *Beispiel 1* einschließen, wurden mit Hilfe der regulären Ausdrücke auch Komposita für *Beispiel*, *Beleg* und *Daten* ermittelt.⁶ Außerdem wurde etwa anhand des Ausdrucks *Signifikanz* geprüft, ob sich im unmittelbaren Kontext u. a. Verweise auf *Treffer* und *Daten* finden lassen. Die Suchbegriffe *Signifikanz* und *Treffer* sind insgesamt nur in einer sehr geringen Frequenz im RGL-Korpus anzutreffen und werden im Folgenden ausgeklammert.⁷ Die Suche nach Nummerierungen erfolgte ausgiebig mit Hilfe verschiedener regulärer Ausdrücke, da wir uns auf diese Weise einen zuverlässigen Zugang zu konkreten Beispielen erhofften. Leider waren über die korrekten Treffer hinaus viele Fehlerkennungen in den Ergebnissen enthalten, die z. T. auf unwillkürliche Zeilennummerierungen im Zuge der

⁶ Bezüglich der Ermittlung der Komposita beispielsweise durch *.*?[D|d]aten.*?* ist anzumerken, dass es sich hierbei nicht um eine morphologische Analyse handelt. So treten Fehlerkennungen wie *Soldatenbriefe* auf, die sich jedoch nach einer manuellen Sichtung der Ergebnisse zahlenmäßig vernachlässigen lassen.

⁷ Auffällig ist hierbei, dass die Arbeiten, die diese Lexeme gehäuft enthalten, einen didaktischen Fokus haben und sich der Sprachentwicklung widmen (s. Klotz 1996; s. Müller 2016), oder eine soziolinguistische Fragestellung untersuchen (s. Ebert 1998; s. Domenech 2019).

OCR-Erkennung zur Textdigitalisierung bzw. Korpuserstellung zurückzuführen waren und eine Sichtung oder Weiterverarbeitung unmöglich machten.

Es wurde auch untersucht, ob es typische Beispiele oder Belege gibt, die in den Texten wiederholt auftauchen. Auf der Basis von Einführungswerken in die Pragmatik wurden versuchsshalber bekannte Äußerungen herausgesucht (bspw. „Nun, der Milchmann ist schon gekommen“ als Illustration für eine konversationelle Implikatur). Diese Probebohrung zu einigen pragmatischen Beispielen führte allerdings zu nur sehr wenigen Treffern. Ebenfalls auf Sprachmaterial aus der Pragmatik zielten Suchabfragen zu *Sprechakt*, *sprachliche Handlung*, *Praktik* und *pragmatische Wende*, die zwar zu keinem umfangreichen Material führten, aber in Hinsicht auf die Entwicklung der Germanistischen Linguistik aufschlussreich sind. Zudem trog auch der erste Eindruck, dass es sich bei den unter „metakommunikativen Markern“ angeführten Formulierungen um sehr frequente handeln könnte. Sie sind durchaus, jedoch nur mit einer vergleichsweise niedrigen Belegzahl, vorhanden und sind mit *Beispiel* und *Beleg* verzahnt, allerdings zeigen sich keine eindeutig lexikalisch spezifizierten Muster, sondern eher Schemata mit lexikalischen Formulierungsalternativen. Ertrageicher waren die Suchabfragen zu mehr oder weniger bekannten Korpora bzw. zur Benennung von Korpora überhaupt, wobei sich, was nicht erstaunen mag, eine immer größere Orientierung an größeren bzw. Referenzkorpora ergibt.

Darüber hinaus erfolgte die Berechnung von Kollokationen zu den Basislexemen *Beispiel*, *Beleg* und *Datum/Daten*.⁸ Dabei gingen wir von den folgenden Hypothesen aus:

- Kollokationspartner (Kookkurrenzen) geben einen Einblick, ob ein *Beispiel* oder *Beleg* als fingiert oder authentisch eingestuft wird.
- Durch Kollokationspartner können insgesamt die Güte, Relevanz und Gewichtung von *Beispielen*, *Belegen* und *Daten* ermittelt werden.
- Sie sind auch Indiz für Sprachgebrauchsmuster und diese wiederum geben Hinweise auf Praktiken der Evidenzherstellung

Aus den über die linguistischen Suchabfragen erzielten Datenkollektionen werden wiederum typische kommunikative Funktionen ersichtlich, die sich auf unterschiedliche Subdisziplinen und ihre typischen Bearbeitungsstrategien von Sprachphänomenen beziehen lassen.

⁸ Der Abstand zur Basis betrug jeweils 10 Token links und rechts. Als Kollokatoren (mit der Mindestfrequenz 5) wurden Nomen, Adjektive und Verben erfasst. Die Signifikanzmessung erfolgte mittels des Log-Likelihood-Ratio-Tests, der auch bei niedrigen Frequenzen bzw. kleinen Korpora zuverlässige Ergebnisse liefert. Der Kollokationsberechnung liegen außerdem die Lemmata zugrunde, sodass ggf. auch die Pluralformen erfasst werden.

Über die Ermittlung von Kollokationen hinaus wurde das RGL-Korpus in ein Word2Vec-Modell überführt, um anschließend die ähnlichsten Lexeme bzw. die Nearest-Neighbors (NNs) zu den Basislexemen *Beispiel*, *Beleg* und *Daten* ermitteln zu können.⁹ Bubenhofer, Calleri & Dreesen (2019: 226) sprechen bezüglich der NNs von „semantisch-funktionale[n] Äquivalenzen in einem spezifischen Geltungsbereich“.

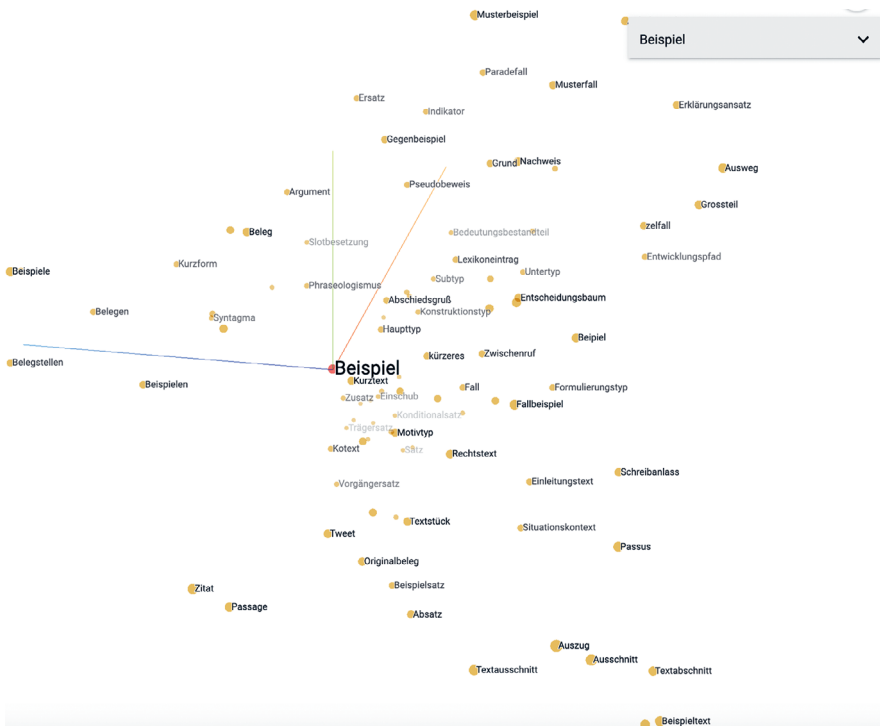


Abb. 1: Die NNs zum Vektor „Beispiel“ in der Tensorflow-Visualisierung¹⁰, zur besseren Übersicht auf 101 Vektoren reduziert.

Die obige Abb. 1 zeigt den Vektor des Lexems *Beispiel* im Zentrum sowie semantisch-funktional äquivalente Lexeme wie *Kurztext*, *Textstück*, *Originalbeleg* oder *Fallbeispiel* im Vektorraum. Schon diese äquivalenten Lexeme lassen die Einbin-

⁹ Hierfür wurde das Python-Modul Gensim genutzt: <https://radimrehurek.com/gensim/> (letzter Zugriff 23.05.2024).

¹⁰ <http://projector.tensorflow.org/> (letzter Zugriff 09.06.2024).

dung von Beispielen in solche linguistischen Disziplinen wie die (diachrone) Text- und Kommunikationsanalyse erkennen. Lexeme wie *Kurztext*, *Zwischenruf*, *Rechtstext* oder *Tweet* verweisen auf bestimmte Kommunikate, wohingegen etwa *Textausschnitt*, *Textabschnitt* oder *Passage* auf die Gliederung von Kommunikaten, *Kotext* oder *Situationskontext* auf deren kommunikative Einbettung und ein Lexem wie *Zitat* auf die intertextuelle Dimension verweisen. Aus der Darstellung wird auch deutlich, dass *Beispiel* und *Beleg* semantische Ähnlichkeiten aufweisen können, aber nicht bedeutungsgleich sind.

Um einen Einblick in diachrone Entwicklungen bzw. unterschiedliche wissenschaftliche Traditionen der RGL zu erhalten, wurden sowohl für die Kollokations- als auch für die Vektorenberechnung insgesamt sechs Subkorpora, bestehend aus den einzelnen Jahrzehnten von 1970 bis 2020, gebildet.¹¹ Die Berechnung der NNs für die einzelnen Jahrzehnte erfolgte mit der TWEC-Methode (Training Temporal Word Embeddings with a Compass), die Di Carlo, Bianchi & Palmonari (2019) beschreiben und über ein Git-Repositoryum¹² zur Verfügung stellen.¹³ Der Vorteil dieser Methode besteht darin, dass sie auch für vergleichsweise kleine Korpora robuste bzw. zuverlässige Ergebnisse liefert, indem zunächst ein Word2Vec-Modell als Kompass auf der Basis des Gesamtkorpus erstellt und anschließend mit Submodellen, die auf den spezifischen Zeitabschnitten trainiert wurden, abgeglichen wird. Auf diese Weise können die NNs zu einem Ausgangslexem, in unserem Fall u. a. *Beispiel*, für einzelne Zeitabschnitte ausgegeben und so miteinander verglichen werden (vgl. Di Carlo, Bianchi & Palmonari 2019: 6328–6329). Die Ergebnisse, die im Folgenden wiedergegeben werden, beziehen sich folglich zum einen auf den gesamten Zeitraum und darüber hinaus auf die einzelnen Jahrzehnte. Da im Rahmen des Beitrages nur auf einen Ausschnitt der Ergebnisse eingegangen werden kann, stehen die übrigen Ergebnisse in einem Repositoryum zur Verfügung.¹⁴ Ferner ist es möglich, die auf der Basis des RGL-Gesamtkorpus erzeugte Word2Vec-Visualisierung selbst zu explorieren.¹⁵

11 Dabei ist für die sich anschließende Auswertung zu beachten, dass die erste RGL-Ausgabe im Jahr 1975 erschienen ist. Ferner waren für die 2020er Jahre noch nicht alle Bände im Korpus enthalten, sodass die Jahrzehnte 1970 sowie 2020 in ihrem Umfang geringer waren als die übrigen Jahrzehnte.

12 <https://github.com/valedica/twec> (letzter Zugriff 22.05.2024).

13 An dieser Stelle sei Xenia Bojarski, Sonja Huber und Noah Bubenhofer gedankt, die diese Methode in ihrem Tagungsbeitrag vorgestellt und damit die Anregung für den vorliegenden Beitrag gegeben haben.

14 https://git.uni-paderborn.de/cgeorgi/rgl_beispiele_geben_diachrone_untersuchung (letzter Zugriff 25.08.2024).

15 Dabei können die TSV-Dateien der Vektoren und der Metadaten, die auf folgender Unterseite des Repositoryums zu finden sind: https://git.uni-paderborn.de/cgeorgi/rgl_beispiele_geben_dia

Grundsätzlich ist anzumerken, dass die Trefferanzahl und deren Signifikanz nicht die nötige Durchsicht der Datenkollektionen und die Re-Lektüre der in die RGL aufgenommenen Monografien bzw. Sammelbände ersetzt, die auf der Basis der gewonnenen Erkenntnisse zu den Suchmustern erfolgte. Close und Distant Reading sind in jeder Phase der Ausarbeitung unseres Beitrags aufeinander bezogen worden. Die Berechnungen wurden dahingehend angepasst, dass sie stets die genauen Kontexte und Belegstellen lieferten.

4 Ergebnisse

4.1 Allgemeine Tendenzen

Um zunächst einen Überblick über die Relevanz und mögliche Trends im diachronen Verlauf bezogen auf *Beispiel*, *Beleg* und *Daten* zu gewinnen, wurde die Auftretenshäufigkeit dieser Lexeme für die einzelnen Jahre ermittelt und in einem Diagramm (s. Abb. 2) gegenübergestellt. Außerdem wurde für jedes Lexem eine lineare Trendlinie berechnet, die den Grad der Zunahme (oder Abnahme) der Werte über die Zeit hinweg verdeutlicht.

Der Überblick durch die Jahre zeigt, dass *Beispiel* relativ konstante Werte aufweist und sich erst um 2007 eine (vage) Trendwende abzeichnet, wo in Jahrgängen wie 2015 erstmals *Belege* mit *Beispielen* gleichziehen. Gerade in den 1970er Jahren ist die Rede von *Belegen* noch randständig und das Sprachmaterial wird fast ausschließlich als *Beispiel* gekennzeichnet, *Datum* und *Daten* spielen zunächst, wie auch unter 4.3 noch gezeigt wird, nahezu keine Rolle. Der Ausreißer im Jahr von 2008 ist auf die erste dezidiert korpuslinguistische Arbeit von Kamber (2008) zu Funktionsverbgefügen zurückzuführen und zeigt – zumindest in Umrissen – eine neuere Wissenschaftskultur, in der *Belege* einen anderen Stellenwert erhalten.¹⁶ Kamber thematisiert die Rolle der *Belege* explizit, indem er der deutschen Sprachwissenschaft grosso modo ein Empiriedefizit attestiert und fol-

chrone_untersuchung/-/tree/main/Word2Vec_Visualisierung (letzter Zugriff 25.08.2024), im Tensorflow-Projector (<https://projector.tensorflow.org/>, letzter Zugriff 25.08.2024) über den „Load“-Button auf der linken Seite hochgeladen werden. Anschließend lässt sich in der rechten Spalte neben der Visualisierung nach Wortformen suchen. Zudem können die NNs zur jeweiligen Wortform angezeigt werden.

¹⁶ Während die relative Frequenz pro 100.000 Token für das Lexem *Beleg* von 1975 bis 2007 im Mittelwert rund 21,1 betrug, sticht die Frequenz für das Jahr 2008 mit rund 487,1 deutlich heraus. Auch das Jahr 2014 stellt mit der Arbeit von Brünjes (2014) und einer Frequenz von rund 128,3 einen Sonderfall dar.

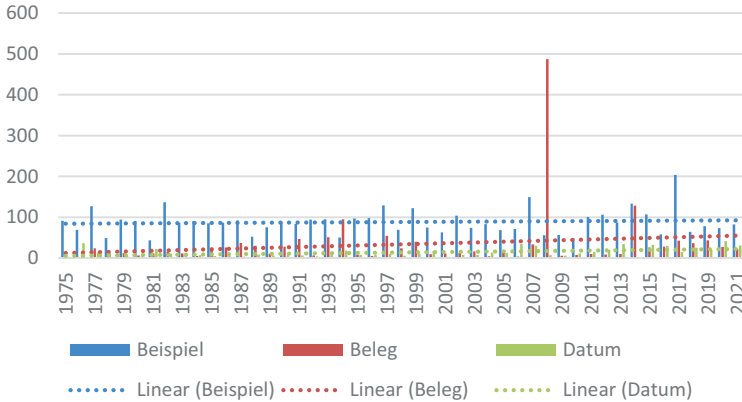


Abb. 2: Auftretenshäufigkeit der Lexeme *Beispiel*, *Beleg* und *Datum* nach Jahren mit relativer Frequenz pro 100.000 Token und linearer Trendlinie.

gende Gründe nennt. Wie unten noch dargestellt wird, stellt eine Positionierung zu konstruierten Belegen selbst eine eigene Praktik dar (s. Abschnitt 4.2):

Der erste Grund ist, dass die deutsche Sprachwissenschaft [...] über drei Jahrzehnte lang sehr stark mit spekulativen Sprachmodellen, insbesondere mit der generativen Grammatik beschäftigt war. Es wurde nicht in wohldefinierten Korpora nach Belegen (und Gegenbelegen!) für die postulierten Regularitäten gesucht, sondern die „Belege“ wurden von den Autoren gleich selbst konstruiert. [...] Der zweite Grund ist etwas anderer Natur. Es ist natürlich sehr viel mühsamer und zeitraubender, die aufgestellten Behauptungen empirisch zu belegen als diese gleich selbst durch ein selbstkonstruiertes Beispiel zu „legitimieren“ (ein Verfahren, das den Standards der wissenschaftlichen Forschung kaum gerecht wird). (Kamber 2008: 37–38)

Verweise auf *Daten* sind im Vergleich zur Thematisierung von *Beispielen* und *Belegen* weniger frequent verzeichnet, weisen aber ebenfalls analog zu *Beleg* eine, wengleich weniger stark steigende, Trendlinie im zeitlichen Verlauf auf. Prüft man die Ausschläge in der Frequenz auf ihre Kontexte hin, sind wiederum Arbeiten mit empirischem Fokus anzutreffen wie die Arbeit von Zimmer (2018).

Die Übersicht der NNs zu *Beispiel* in Abb. 1 verdeutlicht bereits, dass über die Einzelllexeme hinaus mittels verschiedener Komposita auf Beispiele, Belege oder Daten rekuriert wird, wobei Komposita wie *Schlüsselbeispiele*, *Gegenbeispiel*, *Korpusdaten*, *Datensatz*, *Majuskelbelege* oder *Zufallsbelege* die Bedeutung zusätzlich spezifizieren. Um den Grad der Variation und somit den Anteil der Komposita gegenüber den Einzelllexemen zu bestimmen, wurde die Type-Token-Ratio berechnet. Die Anzahl unterschiedlicher Types, die jeweils neben dem Einzelllexem die verschiedenen Komposita umfassen, wird dabei durch die Anzahl der Token,

also das Gesamtvorkommen des Einzellexems und der Komposita, geteilt. Je größer der Wert, umso mehr Variation gibt es.

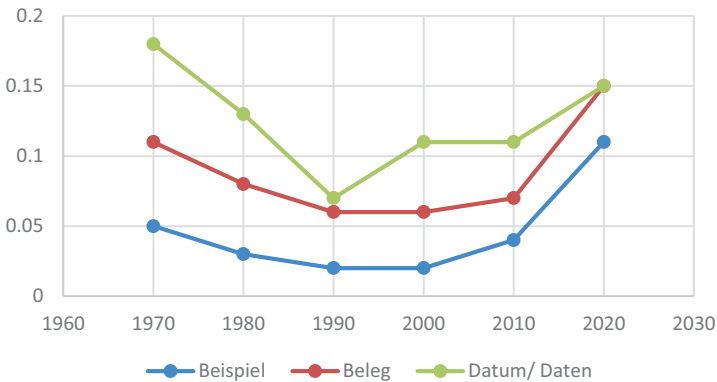


Abb. 3: Type-Token-Ratio der Einzellexeme und der Komposita mit *Beispiel*, *Beleg* oder *Daten*.

Vergleicht man die Type-Token-Ratio der verschiedenen Lexeme, wie in Abb. 3 erfolgt, wird deutlich, dass bei den Komposita zu *Beleg* und *Datum/Daten* mehr Variation als bei *Beispiel* zu verzeichnen ist. Bei allen drei Lexemen ist hinsichtlich der Komposita über die verschiedenen Jahrzehnte hinweg zunächst eine abnehmende und schließlich eine zunehmende Tendenz zu beobachten. Für die 2020er Jahrgänge und somit für die gegenwärtige Entwicklung in der RGL lässt sich konstatieren, dass eine Ausdifferenzierung hinsichtlich der Komposita insbesondere für *Beleg* und *Datum/Daten* aber auch für *Beispiel* stattgefunden hat.

Konstanz und Ausdifferenzierung lassen sich besonders gut an *Beispiel* nachzeichnen. Dabei zeigt sich grundsätzlich, dass die Bandbreite der Komposita mehrere Jahrzehnte nicht besonders groß ist und sich gruppieren lässt. Hier ist zunächst auffällig, dass äußerst selten von *Beispielwort* oder *Wortbeispiel* die Rede ist, was damit zusammenhängen könnte, dass *Beleg* in der Lexikologie und Lexikografie einen terminologischen Status besitzt. Umso häufiger ist von *Beispielsatz*, *-äußerung* oder *Beispieltext (Textbeispiel)*, etwas weniger frequent von *Fallbeispiel* und *Gesprächs*-, *Dialog*- oder *Interaktionsbeispiel*, die Rede, womit auch noch weitere Spezifizierungen wie *Nachrichtentextbeispiel* einhergehen können; gerade die zunehmende Spezifizierung begründet die stärkere Variation in den 2010er Jahren. Neben dem Ausweisen des linguistischen Status des Beispiels wird der Zweck des Beispiels bis heute als *Analysebeispiel* oder *Beispielanalyse* gekennzeichnet. Dass die angeführten Beispiele auf eine größere Menge gleichgearteter Kommunikate verweisen, wird an Bildungen wie *Muster*- oder *Paradebeispiel* ersichtlich. Dies zeigt auch, wie wichtig es den Autor*innen ist, den repräsentativen

Charakter ihrer Beispiele zu behaupten, wie dies auch an den Sprachgebrauchsmustern des Vergleichens ersichtlich ist (vgl. 4.2). Schließlich ist an den Kompositionen auffällig, dass von *Beispielsammlung*, *-liste*, *-reihe* oder *-reihung* die Rede ist. Wie die Frequenz eines Lexems wie *Textbeispiel* deutlich macht, verweist das Beispiel insgesamt eher auf bestimmte linguistische Disziplinen wie Text- und Gesprächslinguistik und auf die linguistische Pragmatik und Stilistik. Dieser Befund bestätigt sich auch im Kontrast zu *Beleg* (vgl. 4.3).

Die Ergebnisse im Rahmen der Type-Token-Ratio deuten zumindest vage darauf hin, dass sich in linguistischen Publikationen Praktiken der Präsentation von Sprachmaterial dahingehend verändern, dass die Qualität und fachliche Güte von Sprachmaterial betont wird und dass ein möglicher Wandel des Selbstverständnisses einer nunmehr vornehmlich empirisch arbeitenden Disziplin stattgefunden hat. Dies soll nicht heißen, dass die Germanistische Sprachwissenschaft nicht auch schon vor 2000 mit größeren Korpora oder mit authentischen sprachlichen Materialien gearbeitet hat, was sich auch direkt in den Veröffentlichungen der RGL spiegelt, jedoch sind die Zugänglichkeit von Korpora und die zunehmenden Möglichkeiten, eigenständig Material zu digitalisieren, gestiegen und bilden eine Grundlage aktueller linguistischer Untersuchungen. Allgemein lässt sich über die Jahrzehnte feststellen, dass sich entlang der Veröffentlichungen der RGL jeweils die Entwicklung einschlägiger Korpora (bspw. das SPIEGEL-Korpus) und die Bedeutung von Referenzkorpora (bspw. COSMAS) nachvollziehen lässt. Ferner versuchen viele Autor*innen, ihr Untersuchungsmaterial, bezogen auf die jeweiligen Möglichkeiten ihrer Zeit, zugänglich zu machen, teils finden sich etwa umfangreiche Anhänge mit transkribiertem Material. Aus der RGL lässt sich, wie auch die folgenden Abschnitte zeigen werden, die Wertschätzung authentischen Sprachmaterials und eine Ablehnung von fingierten Beispielen herauslesen.

4.2 Beispiele und ihre Verwendung

Betrachtet man die Kontexte der Suchabfragen zu *Beispiel* genauer, wird deutlich, dass es immer wieder Auseinandersetzungen mit der Qualität von Beispielen gibt. Es wird ersichtlich, dass schon in den 1970er Jahren fingierte Beispiele kritisiert werden:

- (1) Es ist ein Punkt erreicht, wo man die Grenzen von immer neuen, selbst konstruierten Linguisten-Beispielen und die Grenze einer immer neuen Diskussion alter Beispiele sehen und überschreiten muß (Dobnig-Jülch 1977: 99),

wobei diese Kritik in den nachfolgenden Jahrzehnten fortgesetzt wird:

- (2) Um Zusammenhänge in der Komplexität erfassen zu können, die sie in der natürlichen Sprache haben, und nicht zu sehr der Versuchung zu der Vereinfachung zu verfallen, dienen soweit als möglich nicht konstruierte Beispiele zu der Illustration, sondern solche, die aus konkreten Texten stammen. (Peyer 1997: 26)
- (3) Heine kann sich für die Annahme der zahlreichen ambigen Übergangsformen auf einigermassen zuverlässig erhobene Beobachtungsdaten nicht stützen (er arbeitet nur mit konstruierten Beispielen), und die zuletzt zitierte Aussage mutet ziemlich spekulativ an. (Leirbukt 1997: 212)
- (4) Der Mangel an empirischer Abstützung auf dem Gebiet der FVG treibt manchmal – um bei Sinclairs Metapher zu bleiben – seltsame Blüten. Die Folgen sollen hier kurz an drei Beispielen belegt werden, die teilweise an der Sprachrealität vorbeiziehen, aber leider recht repräsentativ für die Arbeitsweise einer gewissen Linguistik der zweiten Hälfte des 20. Jahrhunderts sind. (Kamber 2008: 38)

Es handelt sich hierbei um eine spezifische Praktik, nämlich um *die Positionierung zum Evidenzwert sprachlichen Materials*. Hiermit wird nicht nur das konstruierte Beispiel als solches, sondern die damit verbundene Qualität des wissenschaftlichen Arbeitens in Zweifel gezogen. Dass die Orientierung am authentischen Sprachmaterial, an der „Sprachrealität“ ein wichtiges Ziel ist, wird auch daran deutlich, dass die Entwicklung von Korpora und Datenbanken wiederholt im Fokus steht und deren Durchsetzung schon früh als Fortschritt bezeichnet wird – auch hier lässt sich wieder eine bestimmte Praktik annehmen, und zwar *die Bezugnahme auf die Art der Aufbereitung von Sprachmaterial*:

- (5) Die nachfolgend aufgelisteten Kongruenzbeispiele sind der Kartei des bibliographischen Institutes in Mannheim entnommen. [...] Das Ergebnis ist zwar noch keine linguistische Datenbank, doch hat der Bereich der syntaktischen Kongruenz mit der folgenden Zusammenstellung immerhin das Stadium der »Dampfkartei« überwunden. (Jaeger 1992: 219)¹⁷

17 In diesem Zusammenhang sei nur erwähnt, dass gerade das konstruierte Beispielmaterial offensichtlich zu einem lexikalischen „Doing Gender“ einlädt, was noch einmal ein ganz eigenes Untersuchungsthema wäre. Einschlägige Äußerungen sind: (a) Sie weint wie ein kleines Mädchen. - Er weint wie ein kleines Mädchen. (Jaeger 1992: 6; konstruiert); (b) Das charmante Mädchen, das du gestern bei mir getroffen hast, ist noch am Abend verunglückt. Sie liegt jetzt im Krankenhaus. (Jaeger 1992: 39; konstruiert); (c) Das nette Mädchen von nebenan, das so lustig pfeifen kann, verbringt hier seinen Auslandsaufenthalt. (Jaeger 1992: 54; konstruiert); (d) Wenn also einer ein Mädchen zu dem Tanz bittet, ändert es doch nichts an der Bitte, wenn er selbst zu dem Beispiel den Tanz gar nicht bis zu dem Ende mitzumachen vorhat, weil er sie vorher verführen will. (Beck 1980: 106; konstruiert)

Die Kategorisierung von Sprachmaterial als *Beispiel* erfolgt durch die Kennzeichnung einer Äußerung als *Beispiel*, jedoch auch mit höheren Werten durch die Muster *zum Beispiel*, *dazu ein Beispiel* oder *ein Beispiel*, wobei ein Beispiel bzw. Beispielreihen entweder durch Absätze oder durch Nummerierungen gegliedert werden; die Häufigkeit von Beispielen wird äußerst selten angegeben. Ferner ist auffällig, dass es charakteristische Praktiken gibt, *die Evidenz eines Beispiels abzusichern*, so durch die Qualifizierung eines Beispiels als *typisch* als häufige Kollokation, durch das Kennzeichnen einer Ähnlichkeit zu genannten Beispielen durch *vgl.* oder durch den Verweis auf *viele (andere)* oder *weitere* Beispiele. Dabei ist das, was jeweils als typisch gilt, vom Erkenntnisinteresse der Einzelnen abhängig. Es kann sich um ein Wort handeln:

- (6) Typisches Beispiel dafür ist *Pulli*, die im Deutschen gebildete und gebrauchte Kurzform für Pullover. (Yang 1990: 12),

oder um einen Text, der als gleich geartet mit einem anderen Text innerhalb einer Textserie zu betrachten ist:

- (7) Abschließend werde ich eine Wegbeschreibung (WB) (verfaßt von VP 12 WB) aus dem Korpus dokumentieren, die als ein typisches Beispiel der in dieser Versuchsreihe entstandenen Texte anzusehen ist. (Rau 1994: 64)

Schließlich können auch andere Kategorien, so eine abstrakte Kategorie wie „Deutungsmuster“, und ihr sprachlicher Ausdruck als typisch deklariert werden:

- (8) Besonders wichtig im Terrorismus-Zusammenhang ist das Deutungsmuster der Eskalation. Es findet seinen typischen Ausdruck in Nominalphrasen wie »die Eskalation der Gewalt« oder in Sätzen wie »Die Anarchisten eskalieren den Terror« (Welt 23.05.1972, S. 3). (Steinseifer 2011: 265)

Der Verweis auf Typizität ist vor dem Hintergrund zu verstehen, dass ein*e Textproduzent*in ein Analyseergebnis präsentiert, dessen Zustandekommen nicht notwendig in seinen einzelnen Untersuchungsschritten offengelegt wird. Wie schon beschrieben, sind die Verweise auf eine Vergleichbarkeit des einen *Beispiels* mit einem *weiteren Beispiel* häufig, wobei die anderen Beispiele nicht notwendig aufgeführt werden müssen:

- (9) *vgl., man vergleiche auch, man vergleiche dazu, dieses Beispiel wird auch bei X behandelt, vergleichbar mit den Beispielen X, Y, ein weiteres Beispiel geben, die Beispiele ließen sich noch um viele vermehren, diese wenigen Beispiele mögen genügen, in meinem Korpus findet sich kein wirklich vergleichbares Beispiel.*

Jedes dieser Muster ist aufschlussreich. Mit dem Muster „diese wenigen Beispiele mögen genügen“ wird die ausreichende Evidenz der angeführten Beispiele betont. Dadurch dass wenig Sprachmaterial als ausreichend festgestellt wird, wird die problemlose Nachvollziehbarkeit für Rezipient*innen betont, die einzig durch den Gebrauch des Konjunktivs relativiert wird. Schließlich legen Verfasser*innen mit diesem Muster eine souveräne Textorganisation nahe.

Interessant ist nun der Gesamtbereich der ermittelten Kollokationen, die mit Beispielen verbunden sind. Substantivische Kollokationen wurden nicht in die Tab. 2 aufgenommen, weil sie auf die unterschiedlichsten Themen, die mit der Analyse von Kommunikaten zusammenhängen, verweisen. Hier zeigt sich ein bedeutender Unterschied zu *Beleg* (s. Abschnitt 4.3), weil dort die substantivischen Kollokationen eine Affinität zu grammatischen und lexikographischen Traditionen zeigen.

Tab. 2: Liste zu verbalen und adjektivischen Kollokationen des Lexems *Beispiel*.

Kollokationen (Verb)	semantisch-funktionale Kategorie	Kollokationen (Adjektiv)	semantisch-funktionale Kategorie
<i>zeigen</i> <i>demonstrieren</i> <i>erläutern</i> <i>belegen</i> <i>diskutieren</i>	Sichtbar-Machen	<i>folgend</i> <i>angeführt</i> <i>ander</i> <i>obig</i> <i>zitiert</i> <i>neu</i> <i>aufgeführt</i> <i>diskutiert</i>	Textdeixis, diskursive Funktion
<i>illustrieren</i> <i>verdeutlichen</i> <i>veranschaulichen</i>	Stützung	<i>konstruiert</i> <i>typisch</i> <i>sprachlich</i>	Qualität
<i>nennen</i> <i>anführen</i> <i>betrachten</i> <i>vorführen</i>	Beispiel-Geben	<i>sprachreflexiv</i> <i>deutsch</i> <i>sozial</i> <i>weit</i>	Relationalität
<i>stammen</i> <i>zitieren</i> <i>finden</i>	Ressource		

Die Verbkollokatoren zeigen Folgendes: Eine Gruppe bilden Kollokationen, die darauf abheben, dass ein Beispiel etwas erkennen bzw. sichtbar werden lässt, also v. a. etwa *zeigt*, *demonstriert* oder *belegt*, was zuvor festgestellt oder behauptet wor-

den ist. Dies kann bestätigt werden, Erkenntnisse absichern oder erst den Auftakt zu weiteren Überlegungen bilden. Es sind Formulierungen dafür, wofür ein Beispiel verwendet wird oder wofür es steht. Darüber hinaus kann ein Beispiel etwas *illustrieren*, *verdeutlichen* oder *veranschaulichen*, was zuvor als abstraktes Konzept bzw. als eine linguistische Kategorie eingeführt worden ist. Ferner können Beispiele dahingehend betrachtet werden, wovon sie ein Beispiel sind, sie *stammen* aus einer Quelle, werden *zitiert* oder lassen sich irgendwo *finden*. Damit wird die Herkunft eines Beispiels offengelegt. Schließlich kann die Aktivität des Beispielnennens unterschiedlich bezeichnet werden, ohne auf die Funktionalität des Beispielmaterials schon hinzudeuten (*nennen*, *anführen*, *betrachten*, *vorführen*).

Insgesamt deuten die Verbkollokationen darauf hin, dass Sprachbeispiele nicht nur eine These im Sinne einer Stützung belegen, sondern mit einem größeren Funktionsspektrum verbunden sind. Neben den Adjektiven sind auch die Präpositionen *anhand* und *mittels* auffällig, die zumeist das Beispiel als Element des Wissens- und Erkenntnisgewinns kennzeichnen. Bei einem Großteil der frequenten Adjektive handelt es sich um Textdeiktika (*folgend*, *angeführt*, *aufgeführt*, *obig*), die zeigen, dass Beispiele ein eigenes Verweissystem ausbilden können. Neben der kohäsiven Funktion der Adjektiv-Kollokationen wird auch die diskursive Funktion des Beispiels deutlich (*diskutiert*). Wie schon betont, ist für das wissenschaftliche Handeln wichtig, sich des Charakters, der Nähe des Beispiels zur „Sprachrealität“, zu vergewissern. Die letzte Gruppe bilden Kollokationen, die die Zugehörigkeit zu einer Einzelsprache, einer sozialen Varietät betonen oder auf reflexive Sprachthematizierungen aufmerksam machen.

Die Kollokationen geben interessante Hinweise auf relevante Kotexte, in denen sich Beispiele bewegen. Als wesentliche Praktiken können neben den schon genannten Praktiken – Positionierung zum Evidenzwert sprachlichen Materials, Bezugnahme auf die Art der Materialbearbeitung, Absicherung der Evidenz des angeführten Beispiel – unterschieden werden:

ein Phänomen registrieren und seine Bandbreite erfassen

- (10) Beispiele: „jn veräppeln“, „jn auf den Arm nehmen“, „jn verkackeiern“, „jn lächerlich machen“, „jn hereinlegen“, „jn zum besten halten“, „jn necken“, „jn zu dem Narren halten“, „jn für dumm verkaufen“, „bloßstellen“. (Koller 1977: 113)

ein Verständnis präzisieren bspw. hinsichtlich dessen, was unter einen Begriff fällt

- (11) Da ich einerseits den ohnehin schon schwierigen Begriff der Metakommunikation nicht überdehnen möchte, andererseits aber auch Beispiele wie die oben jeweils unter (b) angeführten einbeziehen möchte, werde ich in dem folgenden nicht von «Metakommunikation» oder «metakommunikativen Anmerkungen» sprechen. (Peyer 1997: 251)

ein Verfahren bzw. eine Untersuchungsmethode demonstrieren

- (12) Das oben beschriebene Verfahren für die Ermittlung von Kenngrößen soll an einem Beispiel illustriert werden. (Lemnitzer 1997: 159)

eine Auffassung bestätigen und widerlegen

- (13) Wie kommt es, fragen wir noch einmal, daß, trotz so offensichtlicher Gegeninstanzen, die zu der Widerlegung die Mühe einer umständlichen Analyse eigentlich sollten überflüssig gemacht haben (und durch die Searle auch z. T. oberflächliche Kritik gegen sich heraufbeschworen hat, welche oberflächlich auch abgetan werden konnte), Searle sein Beispiel und seine daran angeknüpfte Behauptung dennoch hat aufrechterhalten können? (Beck 1980: 83)

Die von der kulturwissenschaftlichen Forschung postulierte Ermittlung von Regeln findet sich nur gelegentlich in dezidiert grammatischer Forschungsliteratur:

- (14) Hierzu zunächst die folgenden Beispielsätze:
 (2 a) Er hat doch das Auto geklaut.
 (2 b) Er hat das Auto doch geklaut.
 (2 c) Er hat doch ein Auto geklaut.
 (2 d) *Er hat ein Auto doch geklaut. (Hentschel 1983: 47)

Etwas später formuliert die Verfasserin die Regel: „Die Regel, die sich aus der Gesamtheit der Beobachtungen ableiten läßt, lautet somit einfach: Abtönungspartikeln müssen vor dem Rhema stehen.“ (Hentschel 1983: 50)

Es wurden auch die Kollokationen ermittelt, die *nicht* bei *Beleg* oder *Daten* auftreten. Dabei handelt es sich nicht nur um den erwartbaren Kollokationspartner *geben*, sondern auch um die Verben *illustrieren*, *verdeutlichen* und *zeigen*. Während also Beispiele etwas illustrieren, gilt dies offensichtlich nicht für Belege oder Daten. Es ist ebenso bemerkenswert, dass das Adjektiv *diskutiert* nur bei *Beispiel* verbucht werden kann. Über die Tab. 2 hinaus ist auch auffällig, dass bei *illustrieren*, *verdeutlichen* u. Ä. die Wortfamilie genutzt wird.

4.3 Belege und ihre Verwendung

Die Art und Verteilung der mit *Beleg* realisierten Kollokationen zeigt grundsätzlich andere Verwendungstraditionen als bei *Beispiel*. In Tab. 3 werden anders als bei *Beispiel* auch die frequentesten substantivischen Kollokationen aufgeführt. Aus der Auflistung wird erkennbar, dass *Korpus* an erster Stelle steht. Das heißt zunächst, dass die Belege aus unterschiedlichen Korpora entstammen, woraus sich allerdings noch kein notwendiger Bezug zur Korpuslinguistik ergibt. Auffälliger ist unseres Erachtens, dass es sich bei den nachfolgenden Kollokationen fast ausschließlich um

solche handelt, die auf die Grammatik verweisen. Die nächsten Kollokationen, nämlich *Kollektivbegriffe*, *Präpositionalobjekt*, *Dativobjekt*, *Datenserie*, *Erweiterbarkeit*, *Abstrakta*, *Quelle*, *Beschreibungsschema* und *Kommunikation* bestätigen, dass die Kennzeichnung von Sprachmaterial als *Beleg* v. a. in der Grammatik und auch in der Lexikologie zentral ist. Auch die Kollokationen, die nicht parallel bei *Beispiel* und *Daten* vorhanden sind, lassen dies erkennen. Es handelt sich größtenteils um grammatische Termini wie *Teilsatz*, *Konditionalsatz* oder *Konstruktionstyp*, oder um lexikographische Termini wie *Wörterbuchartikel* oder *Wörterbucheintrag*. Ferner wird ersichtlich, dass die Angabe einer Quelle als *Korpus*, die Kennzeichnung von Sprachmaterial als *Beleg* und die Charakterisierung eines Forschungsergebnisses als *Befund* zusammenhängen. *Befund* gehört auch zu den Lexemen, die bei *Beispiel* nicht nachzuweisen sind. Es ist sicherlich zu spekulativ, an der Verwendung von *Beispiel* und *Befund* wissenschaftliche Kulturen und ‚Denkstile‘ festmachen zu wollen, doch erscheint uns dieses Zusammenspiel interessant, da die Evidenz und das problemlose Auffinden grammatischer Kategorien zumindest nahegelegt werden.

Die adjektivischen Kollokatoren, die nicht in Tab. 3 erscheinen, gleichen denen, die bei *Beispiel* aufgeführt worden sind, mit dem Unterschied, dass es teils wieder grammatische Termini wie *intraphrastisch* sind, teils auch Spezifizierungen wie *häufig* oder *frequent* sind, die angeführt werden. Unter den verbalen Kollokationen befinden sich nicht *veranschaulichen*, *illustrieren* oder *verdeutlichen*.

Tab. 3: Verbale und substantivische Kollokationen mit *Beleg*, zugeordnet zu semantisch-funktionalen Kategorien.

Kollokationen (Verben)	semantisch-funktionale Kategorie	Kollokationen (Substantiv)	semantisch-funktionale Kategorie
<i>finden</i> <i>entnehmen</i> <i>(sich) vorfinden</i>	Erhebung	<i>Korpus</i> <i>Sprache</i> <i>Syntagma</i>	Kategorisierung von Sprachmaterial als grammatisch
<i>stammen</i> <i>liefern</i> <i>enthalten</i> <i>kommen</i>	Ressource	<i>Definitartikel</i> <i>Anzahl</i> <i>Präposition</i> <i>Vorderglied</i> <i>Subjektstelle</i>	
<i>entfallen</i> <i>verteilen</i>	Verteilung	<i>Verbformen</i> <i>Modalpartikel</i>	
<i>belegen</i> <i>weisen</i> <i>anführen</i> <i>aufführen</i>	Funktion		

Das oben angedeutete Wissenschaftsverständnis zeigt sich auch daran, dass das sprachliche Material etwas *liefert* oder *enthält*, was sich der Sprachwissenschaftlerin unmittelbar zu erschließen scheint und in seiner Verteilung bestimmt werden kann (*entfallen, verteilen*). Zu den weiteren Kennzeichen der Verwendungsweise von *Beleg* gehört, dass häufig die Anzahl von *Belegen* genannt wird, was bei *Beispielen* nicht der Fall ist. Interessant ist, dass – anders gegenüber *Beispiel* – keine Muster erkennbar sind, die auch auf die Textorganisation des Schreibenden abzielen (also nicht: „ein Beleg muss genügen“).

Zu den kommunikativen Praktiken, die mit dem Beleg verbunden sind, gehören die folgenden, wobei *Beleg* anders als *Beispiel* nicht mit den unter 4.2. genannten Praktiken, insbesondere mit der Positionierung zur Güte des sprachlichen Materials verbunden ist. Auch die jetzt angeführten Praktiken sind nicht mit denen unter 4.2 ansonsten angeführten Praktiken zu vergleichen. Es finden sich:

die Häufigkeit und Verteilung eines Phänomens registrieren, wobei Beleggruppen auch miteinander verglichen oder kontrastiert werden können

- (15) Das fnhd. Korpus (Veit Dietrich) umfasst insgesamt 593 Belege, die sich wie folgt verteilen: wollen 193 (32,5 %), sollen 189 (31,95 %), können 110 (18,5 %), müssen 57 (9,65 %), mögen 31 (5,2 %), dürfen 13 (2,2 %). (Diewald 1999: 10)

einen Prototyp ermitteln

- (16) Dagegen wird in mehr als 500 Belegen aus einem Jahrgang des „Mannheimer Korpus“ des Instituts für deutsche Sprache das Wort *Gewalt* nahezu ausschließlich im Sinne von *roher*, häufig sogar *brutaler körperlicher Gewalt* verwendet. (Busse 1991: 259)

eine Kategorie entwickeln oder schärfen

- (17) Dabei zeigen die bereits besprochenen Belege, dass eine nicht nur semantisch begründete, sondern auch formal bei genauerem Hinsehen naheliegende Subdifferenzierung von Nöten ist. (Merten 2018: 433)

eine Auffassung oder Behauptung bestätigen oder widerlegen

- (18) Diese Aussage stützt sich auf eine Analyse der (ca. 900) Belege des Wortes in den extern recherchierbaren Korpora für geschriebene Sprache des Mannheimer Instituts für deutsche Sprache. (Steinseifer 2011: 52)

Die Wortbildungen stützen die skizzierten, hauptsächlichen Verwendungsweisen. Anders als bei *Beispiel*, wo v. a. die Derivationen *beispielsweise* und *beispielhaft* gebraucht worden sind, werden mit *belegbar, belegen, Belegung* die Möglichkeiten der Wortfamilie voll ausgeschöpft. Es wird allenfalls vereinzelt *Satz-* oder *Textbeleg* verwendet. Die Spezifizierung durch ein Determinans ist generell selten und

wird dann häufiger durch grammatische Termini besetzt, etwa *Plural-* oder *Genitivbelege*. In der überwiegenden Mehrzahl der Fälle ist <beleg> Determinans. Mit diesen Bildungen wird überwiegend gekennzeichnet, woher sprachliche Phänomene stammen (*Belegstellen*, *Belegmaterial*, *Belegsammlung*, *Belegreihe*, *Belegblock* oder *Belegdatensätzen*) oder in welcher Häufigkeit sie erscheinen (*Belegzahlen*, *Belegdichte* oder *Belegfrequenz*). Auch dies deutet auf unterschiedliche sprachliche Bearbeitungsstrategien hin. Die paradigmatische Bedeutungsähnlichkeit, die durch Word2Vec ermittelbar ist, zeigt zwar, dass auch Lexeme wie *Textausschnitt* oder *Beispieltext* verwendet werden, zeigt in unmittelbarer Nähe zum Lexem jedoch auch *Satztypus*, *Befund*, *Satz* oder *Teilsatz*.

Das Spektrum der Verwendungsweisen von *Datum/Daten* ist insgesamt kleiner als das von *Beispiel* und *Beleg*. Eine Sichtung ergab, dass dieser Verweis auf Sprachmaterial v. a. im Rahmen methodischer Reflexionen erscheint, was im folgenden Abschnitt detaillierter beleuchtet wird.

5 Diachrone Entwicklungen in der RGL

Um die diachrone Entwicklung in der RGL nachzuzeichnen, sollen kontrastierend die Ergebnisse der Kollokationen in den 1970er/1980er Jahre und der 2020er Jahre gegenübergestellt werden. Daraus erschließt sich auch, wie und in welcher Weise der Ausdruck „Daten“ in der Linguistik verwendet wird. Über den gesamten Zeitraum, die letzten fünfzig Jahre Germanistischer Linguistik, ergibt sich das in Abb. 4 dargestellte Bild, wiederum als Word2Vec-Darstellung realisiert.

Die Abbildung zeigt ein bestimmtes Verständnis wissenschaftlichen Arbeitens, bei dem Daten im Blickpunkt stehen. Es wird erkennbar, dass *Daten*, *Sprachdaten* und *empirisch* eng aneinander gebunden sind. Ferner erschließt sich, von welcher Art Daten sein können und welche die *Datengrundlage* ist: Es sind entweder *Korpusdaten*, oder sie gehen aus *Erhebung*, *Befragung* oder *Beobachtung* hervor. Dabei verweisen einige Kernlexeme wie *Statistik* auf die Art der Auswertung und die Art und Weise, wie Aussagen über sprachliche Phänomene getroffen werden. Allerdings geht aus der Darstellung auch hervor, dass der Gebrauch des Ausdrucks *Daten* auch in der Gesprächs- und interaktionalen Linguistik gebräuchlich ist. Bemerkenswert ist, dass man sich auch andere wissenschaftliche Disziplinen wie die Soziologie oder die Bildungswissenschaften vorstellen könnte, die ein ähnliches Verständnis von *Daten* besitzen.

Das Verständnis einer derart datenbasierten Disziplin hat sich, wie die zwei Abbildungen 5 und 6 zeigen, in den letzten fünfzig Jahren erst herausgebildet.



Abb. 5: Kollokationen in den 1970er Jahren.

Die Verwendung von *Beispiel* und *Beleg* fügt sich in die schon genannten Tendenzen ein. So wird an den Kollokationen wiederum deutlich, dass *Beleg* eher in der Grammatikschreibung verankert ist als *Beispiel*, wovon etwa *Passiv*, *Definitartikel* oder *Morphemgrenze* zeugen. Das *Beispiel* ist mit hochfrequenten Kollokationen vertreten, wie *folgend*, *haben*, *zeigen*, *geben*, *illustrieren* oder *veranschaulichen*, wobei der Gebrauch von *Daten* wiederum an methodische Fragestellungen anknüpft.

In der nachfolgenden Tab. 4 werden die NNs, die *ausschließlich* bei *Daten*, nicht aber bei *Beispiel* und *Beleg* vorkommen, noch einmal nach Jahrzehnten getrennt, aufgelistet. Außerdem werden nur die für das jeweilige Jahrzehnt neuen NNs dargestellt, um den Fokus auf mögliche Veränderungen zu lenken. Sichtet man die Aufstellung der vorliegenden Tabelle, zeigt sich, dass ab den 1980er Jahren im Zusammenhang mit *Hypothesen* und *Tests* sowie verstärkt ab den 1990er Jahren *Daten* im Sinne einer korpusbasierten und -geleiteten Vorgehensweise, also im Zusammenhang mit *Kollokationsanalysen*, *Korpora*, *Erhebungen* sowie

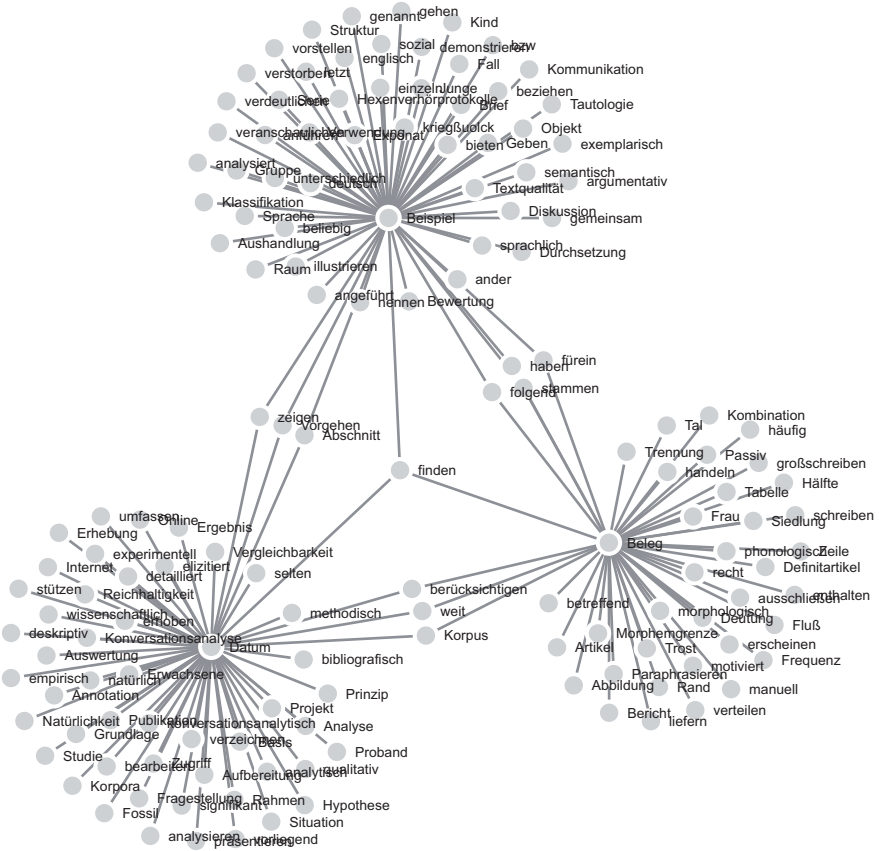


Abb. 6: Kollokationen in den 2020er Jahren.

Vorgehensweisen gebraucht werden. Ab den 2000er Jahren wird zudem beispielsweise über *Quellen*, *Fundstellen*, *Stichproben* sowie konkret über *Videoaufnahmen* verstärkt die Herkunft bzw. Beschaffenheit der Daten thematisiert und es erfolgt eine Einordnung der verwendeten Daten vor dem Hintergrund der *Datenlage*.

Tab. 4: NNs der Wortform *Daten* nach Jahrzehnten gemäß der TWEC-Methode; bereinigt um die NNs, die auch bei *Beispiel* und *Beleg* oder in den vorangegangenen Jahrzehnten auftreten.

Jahrzehnt	NNs zu Daten
1970	<ul style="list-style-type: none"> handlungsvorkommen Zusammenhänge Deutungen Informationen Bedingungen Möglichkeiten Zielsetzungen Interaktionen Kriterien Gegebenheiten Analysen Kommunikationssituationen Phänomene Beschreibungen Faktoren
1980	<ul style="list-style-type: none"> Interaktionssituationen Schwächen Beurteilungskriterien Auswertungsergebnisse Techniken Hypothesen Tests Interpretationen Sprachdaten Fakten Auswertungen Texteigenschaften
1990	<ul style="list-style-type: none"> Kollokationsanalysen Transkripte Ergebnisse Befragungen Korpora Auswertung Resultate Vorgehensweisen Erhebungen Datenbasis Beobachtungen

Tab. 4 (fortgesetzt)

Jahrzehnt	NNs zu Daten
2000	Datenmaterial Ausschnitte Textstellen Quellen Fundstellen Textbeispiele Schlagzeilen Statistiken Befunde Entwicklungsniveaus
2010	Auswertungsverfahren Stichproben Datenlage Datenanalyse Wissensressourcen Therapien
2020	Fragestellungen Videoaufnahmen Erkenntnisse Datengrundlage Lehrkräften Studien Untersuchung

6 Fazit

Die RGL ist eine Reihe mit großer thematischer Offenheit. Diese Offenheit bedeutet grundsätzlich, dass Untersuchungen mit unterschiedlichen Erkenntnisinteressen und unterschiedlichen Methoden in der Reihe präsent sind. In der Reihe finden sich bis heute Arbeiten, die sich wenigen Texten hermeneutisch („entziffernde Lektüre“) nähern:

- (19) Die vorliegende Studie leistet eine entziffernde Lektüre sprachreflexiver und poetologischer Gedichte des österreichischen Schriftstellers Hugo von Hofmannsthal, dessen Texte als chiffrierte Epochensignaturen interpretiert werden. Hofmannsthals lyrisches Werk wird im Rahmen dieser Arbeit in einer Doppelperspektive in seiner sprachkünstlerischen Gestaltung erschlossen und als Dokument der Sprachgeschichte gedeutet. (Heinz 2009: V)

Mehrere übergreifende Trends lesen wir aus dem Korpus und den Auswertungen heraus:

- Die Präsentation von Sprache und damit die Kategorisierung des Sprachmaterials als *Beispiel*, *Beleg* oder *Daten* sind für die Sprachwissenschaft unentbehrlich, wobei sich ansatzweise eine Entwicklung zeigt, etwas weniger von *Beispielen*, etwas mehr jedoch von *Belegen* und *Daten* zu sprechen. Anhand von Komposita wird erkennbar, dass die Verfasser*innen dazu übergehen zu scheinen, das Sprachmaterial immer spezifischer zu kennzeichnen (etwa *Nachrichtentextbeispiel* statt *Textbeispiel*).
- Es zeigt sich, dass das Vorliegen authentischen Sprachmaterials gegenüber selbst erdachtem, konstruiertem Material als Vorzug der eigenen Untersuchungen betont wird. Selbst Beispiele zu erfinden, wird in der Reihe nicht mehr der guten wissenschaftlichen Praxis zugerechnet.
- Die zunehmende Wertschätzung authentischer Sprache korrespondiert ebenso mit der allgemeinen Zugänglichkeit von Korpora, aus denen Daten gewonnen werden, wie mit Datenerhebungen und Verfahren der Datenauswertung. Der Nachweis, woher authentische Sprache stammt und wie diese erhoben worden ist, ist mittlerweile selbstverständlich.
- Das RGL-Korpus zeigt unseres Erachtens, dass sich die Germanistische Linguistik immer stärker als empirisch arbeitende Disziplin profiliert. Ob diese Entwicklung sprachtheoretische Beiträge und Überlegungen verdrängt (hat), müsste gesondert ermittelt werden. In der Zitationspraxis der RGL – und diese wäre zumindest ein Indiz für die geringere Bedeutung der Sprachtheorie – ist jedenfalls auffällig, dass fachfremde pragmatische Klassiker weniger zitiert werden (vgl. den Beitrag von Bojarski, Huber & Bubenhofer in diesem Band). Es könnte jedoch sein, dass die Streubreite der Zitation fachfremder Schriften zugenommen hat.¹⁸

Neben diesen allgemeinen Trends konnte gezeigt werden, dass sich *Beispiele*, *Belege* und *Daten* hinsichtlich ihrer Kollokationen unterscheiden. Die mit ihnen gebildeten Kompositionen sind ebenso wenig identisch wie die über Word2Vec ermittelten paradigmatischen Relationen. Die anhand der Kollokationen und Suchabfragen ermittelten Muster überschneiden sich nur teilweise. Ein Muster wie „wenige/einige

18 Durch die Durchsicht der Datenkollektionen, die durch die Suchabfragen erzeugt wurden, ergeben sich noch weitere Tendenzen, die auf Wandlungen der Wissenschaftssprache hindeuten, die nicht detailliert untersucht worden sind, hier jedoch nicht unerwähnt bleiben sollen: Dazu gehört die Bezeichnung von einzelnen Verfasser*innen als *wir*, was im Laufe der Zeit deutlich abnimmt. Konfrontative und polemische Textpassagen, die sich gegen die Annahmen anderer Wissenschaftler*innen richten, scheinen ebenso zurückzugehen.

Beispiele mögen genügen“ korrespondiert mit der vorwiegend beim *Beispiel* nachweisbaren Mikropraktik. Das Registrieren der Häufigkeit eines sprachlichen Phänomens, das im Zusammenhang mit *Beleg* auftritt, ist wiederum bei *Beispielen* nicht nachweisbar, diese werden kaum quantifiziert. Im Zusammenspiel der hier gewählten Zugänge lässt sich zumindest eine Bindung des Beispielgebens und des Anführens von Belegen an unterschiedliche linguistische Subdisziplinen annehmen. Die Grammatikschreibung operiert v. a. mit Sprachmaterial, das als *Beleg* kategorisiert wird. Demgegenüber lässt sich erkennen, dass Beispielgebung ein Werkzeug ist, um in einem linguistischen Close Reading holistisch den Sinnpotentialen einzelner Kommunikate nachzugehen.

In diesem Beitrag standen lediglich Äußerungen mit *Beispiel*, *Beleg* oder *Daten* und Mikropraktiken der Evidenzherstellung im Vordergrund. Um den Stellenwert von Beispielen, Belegen oder Daten weiter zu ermitteln, müssten die unter Abschnitt 2 skizzierten Perspektiven weiter verfolgt werden und es müsste etwa herausgearbeitet werden, wie Sprachmaterial kontextualisiert oder dekontextualisiert wird oder ob und inwieweit das Verhältnis von Besonderem und Allgemeinem modelliert wird. Erst weitere Untersuchungen könnten zeigen, ob sich neben der Affinität linguistischer Subdisziplinen zu den aufgeführten Praktiken auch ein Umgang mit sprachlichem Material erschließen lässt, der auf unterschiedliche Wissenschaftskulturen unter dem Dach der Linguistik hindeutet oder ob dies nicht der Fall ist.

7 Literaturverzeichnis

7.1 Quellen

- Beck, Götz (1980): *Sprechakte und Sprachfunktionen. Untersuchungen zur Handlungsstruktur der Sprache und ihren Grenzen* (Reihe Germanistische Linguistik 27). Tübingen: Niemeyer.
- Bojarski, Xenia, Sonja Huber & Noah Bubenhofer (2024): *Korpus Reihe Germanistische Linguistik (RGL)*, Version 3. Unter Mitarbeit von Christopher Georgi. Zürich. <https://gitlab.uzh.ch/noah.bubenhofer/corpus-documentations/-/blob/master/corpora/rglv3.md> (letzter Zugriff 25.08.2024).
- Brünjes, Lena (2014): *Das Paradigma deutscher Modalpartikeln: Dialoggrammatische Funktion und paradigmenerinterne Oppositionen* (Reihe Germanistische Linguistik 301). Berlin, Boston: De Gruyter.
- Busse, Dietrich (1991): Der Bedeutungswandel des Begriffs “Gewalt” im Strafrecht. Über institutionell-pragmatische Faktoren semantischen Wandels. In Dietrich Busse (Hrsg.), *Diachrone Semantik und Pragmatik. Untersuchungen zur Erklärung und Beschreibung des Sprachwandels* (Reihe Germanistische Linguistik 113), 259–275. Tübingen: Niemeyer.
- Diewald, Gabriele (1999): *Die Modalverben im Deutschen. Grammatikalisierung und Polyfunktionalität* (Reihe Germanistische Linguistik 208). Tübingen: Niemeyer.

- Dobnig-Jülch, Edeltraud (1977): *Pragmatik und Eigennamen. Untersuchungen zur Theorie und Praxis der Kommunikation mit Eigennamen, besonders von Zuchttieren* (Reihe Germanistische Linguistik 9). Tübingen: Niemeyer.
- Domenech, Madeleine (2019): *Schriftsprachliche Profile von Fünftklässlern: Argumentative Briefe im Zusammenspiel unterschiedlicher textueller, familiärer und individueller Ressourcen* (Reihe Germanistische Linguistik 317). Berlin, Boston: De Gruyter.
- Ebert, Robert Peter (1998): *Verstellungswandel bei Jugendlichen, Frauen und Männern im 16. Jahrhundert* (Reihe Germanistische Linguistik 190). Tübingen: Niemeyer.
- Heinz, Tobias (2009): Hofmannsthals Sprachgeschichte. Linguistisch-literarische Studien zur lyrischen Stimme (Reihe Germanistische Linguistik 284). Tübingen: Niemeyer.
- Hentschel, Elke (1983): Partikeln und Wortstellung. In Harald Weydt (Hrsg.), *Partikeln und Interaktion* (Reihe Germanistische Linguistik 44), 46–53. Tübingen: Niemeyer.
- Jaeger, Christoph (1992): *Probleme der syntaktischen Kongruenz. Theorie und Normvergleich im Deutschen* (Reihe Germanistische Linguistik 132). Tübingen: Niemeyer.
- Kamber, Alain (2008): Funktionsverbgefüge – empirisch: Eine korpusbasierte Untersuchung zu den nominalen Prädikaten des Deutschen (Reihe Germanistische Linguistik 281). Tübingen: Niemeyer.
- Klotz, Peter (1996): *Grammatische Wege zur Textgestaltungscompetenz: Theorie und Empirie* (Reihe Germanistische Linguistik 171). Tübingen: Niemeyer.
- Koller, Werner (1977): *REDENSARTEN. Linguistische Aspekte, Vorkommensanalysen, Sprachspiel* (Reihe Germanistische Linguistik 5). Tübingen: Niemeyer.
- Leirbukt, Oddleif (1997): *Untersuchungen zum bekommen-Passiv im heutigen Deutsch* (Reihe Germanistische Linguistik 177). Tübingen: Niemeyer.
- Lemnitzer, Lothar (1997): *Akquisition komplexer Lexeme aus Textkorpora* (Reihe Germanistische Linguistik 180). Tübingen: Niemeyer.
- Merten, Marie-Luis (2018): *Literater Sprachausbau kognitiv-funktional. Funktionswort-Konstruktionen in der historischen Rechtsschriftlichkeit* (Reihe Germanistische Linguistik 311). Berlin, Boston: De Gruyter.
- Müller, Hans-Georg (2016): *Der Majuskelgebrauch im Deutschen: Gross- und Kleinschreibung theoretisch, empirisch, ontogenetisch* (Reihe Germanistische Linguistik 305). Berlin, Boston: De Gruyter.
- Peyer, Ann (1997): *Satzverknüpfung – syntaktische und textpragmatische Aspekte* (Reihe Germanistische Linguistik 178). Tübingen: Niemeyer.
- Rau, Cornelia (1994): *Revisionen beim Schreiben. Zur Bedeutung von Veränderungen in Textproduktionsprozessen* (Reihe Germanistische Linguistik 148). Tübingen: Niemeyer.
- Steinseifer, Martin (2011): ›Terrorismus‹ zwischen Ereignis und Diskurs. Zur Pragmatik von Text-Bild-Zusammenstellungen in Printmedien der 1970er Jahre (Reihe Germanistische Linguistik 290). Berlin, Boston: De Gruyter.
- Yang, Wenliang (1990): *Anglizismen im Deutschen. Am Beispiel des Nachrichtenmagazins DER SPIEGEL* (Reihe Germanistische Linguistik 106). Tübingen: Niemeyer.
- Zimmer, Christian (2018): *Die Markierung des Genitiv(s) im Deutschen: Empirie und theoretische Implikationen von morphologischer Variation* (Reihe Germanistische Linguistik 315). Berlin, Boston: De Gruyter.

7.2 Sekundärliteratur

- Bubenhofer, Noah, Selena Calleri & Philipp Dreesen (2019): Politisierung in rechtspopulistischen Medien: Wortschatzanalyse und Word Embeddings. In David Römer & Constanze Spieß (Hrsg.), *Populismus und Sagbarkeiten in öffentlich-politischen Diskursen* (Osnabrücker Beiträge zur Sprachtheorie (OBST) 95), 211–241. Duisburg: Universitätsverlag Rhein-Ruhr.
- Di Carlo, Valerio, Federico Bianchi & Matteo Palmonari (2019): Training Temporal Word Embeddings with a Compass. *Proceedings of the AAAI Conference on Artificial Intelligence* 33 (01), 6326–6334.
- Grewendorf, Günther (1980): Argumentieren in der Sprachwissenschaft. *Zeitschrift für Literaturwissenschaft und Linguistik* 10 (38), 129–150.
- Güsken, Jessica (2019): Beispiele geben. Zur Problematik einer unumgänglichen Praxis im Diskurs der Ästhetik (1750–1850). In Olaf Kramer, Carmen Lipphardt & Michael Pelzer (Hrsg.), *Rhetorik und Ästhetik der Evidenz*, 129–154. Berlin, Boston: De Gruyter.
- Hirschmann, Hagen (2019): *Korpuslinguistik. Eine Einführung*. Stuttgart: Metzler.
- Kindt, Walther (1978): Argumentation und Theoriebildung in der historischen Linguistik. Eine Untersuchung am Beispiel des Vernerschen Gesetzes. *Indogermanische Forschungen* 83, 1–39.
- Lück, Christian, Michael Niehaus & Manfred Schneider (Hrsg.) (2013): *Archiv des Beispiels*. Zürich: Diaphanes.
- Roelcke, Thorsten (2013): *Definitionen und Termini: Quantitative Studien zur Konstituierung von Fachwortschatz* (Linguistik – Impulse & Tendenzen 48). Berlin, Boston: De Gruyter.
- Ruchatz, Jens, Stefan Willer & Nicolas Pethes (2007): Zur Systematik des Beispiels. In: Jens Ruchatz, Stefan Willer & Nicolas Pethes (Hrsg.), *Das Beispiel. Epistemologie des Exemplarischen*, 7–59. Berlin: Kadmos.
- Weinrich, Harald (1975): Präsuppositionen in Sätzen und Beispielsätzen. In Harald Weinrich (Hrsg.), *Positionen der Negativität* (Poetik und Hermeneutik VI), 439–440. München: Fink.
- Willer, Stefan (2004): Was ist ein Beispiel? Versuch über das Exemplarische. In Gisela Fehrmann, Erika Linz, Eckhard Schumacher & Brigitte Weingart (Hrsg.), *Originalkopie. Praktiken des Sekundären*, 51–65. Köln: DuMont.
- Willer, Stefan (2013): Beispiel. In Roland Borgards, Harald Neumeyer, Nicolas Pethes & Yvonne Wübben (Hrsg.), *Literatur und Wissen. Ein interdisziplinäres Handbuch*, 271–276. Stuttgart: Metzler.