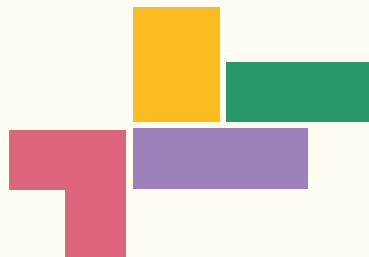


Herausgegeben von
Marek Konopka
Angelika Wöllstein
Ekkehard Felder

BAUSTEINE EINER KORPUSGRAMMATIK DES DEUTSCHEN

Band 1 | 2020



HEIDELBERG
UNIVERSITY PUBLISHING

Bausteine einer Korpusgrammatik des Deutschen
Band 1 | 2020

Wissenschaftlicher Beirat

Noah Bubenhofer (Zürich)

Peter Gallmann (Jena)

Mechthild Habermann (Erlangen-Nürnberg)

Mathilde Hennig (Gießen)

Martin Hilpert (Neuchâtel)

John Nerbonne (Freiburg)

Damaris Nübling (Mainz)

Oliver Schallert (München)

Barbara Schlücker (Leipzig)

Bausteine einer Korpusgrammatik des Deutschen


Band 1 | 2020

Herausgegeben von

Marek Konopka
Angelika Wöllstein
Ekkehard Felder

HEIDELBERG
UNIVERSITY PUBLISHING

Redaktion

Kristin Kopf  <https://orcid.org/0000-0002-0902-7918>

Ulrich H. Waßner

Koordination der Ausgabe

Thilo Weber

IDS

LEIBNIZ-INSTITUT FÜR
DEUTSCHE SPRACHE

EZS

EUROPÄISCHES
ZENTRUM FÜR
SPRACHWISSENSCHAFTEN



UNIVERSITÄT
HEIDELBERG
ZUKUNFT
SEIT 1386

Bibliografische Information der Deutschen Nationalbibliothek

Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über <http://dnb.dnb.de> abrufbar.



Dieses Werk ist unter der Creative Commons-Lizenz CC BY-SA 4.0 veröffentlicht. Die Umschlaggestaltung unterliegt der Creative-Commons-Lizenz CC BY-ND 4.0.

Publiziert bei Heidelberg University Publishing (heiUP)
Heidelberg 2020.

Die Online-Version dieser Publikation ist auf den Verlagswebseiten von Heidelberg University Publishing <https://heiup.uni-heidelberg.de> dauerhaft frei verfügbar (Open Access).

doi: <https://doi.org/10.17885/heiup.bkgd.2020.1>

Text © 2020. Das Copyright der Texte liegt beim jeweiligen Verfasser.

ISSN 2702-2080
eISSN 2702-2099

ISBN 978-3-96822-032-1 (Softcover)
ISBN 978-3-96822-033-8 (PDF)

Inhalt

<i>Marek Konopka</i> Einleitung	7
<i>Marek Konopka</i> Grundlegende Aspekte der Wortbildung	29
<i>Patrick Brandt</i> Bau von und Umbau zu Adverbien: Präpositionen, Vergleiche und Flexion	65
<i>Franziska Münzberg & Sandra Hansen</i> Starke vs. schwache Flexion aufeinanderfolgender attributiver Adjektive: <i>mit hohem technischen/technischem Aufwand</i>	99
<i>Franziska Münzberg & Felix Bildhauer</i> Reihenfolge attributiver Adjektive	131
Über die Autorinnen und Autoren	157

Marek Konopka 

Einleitung

1 In aller Kürze

Eine deskriptive, konsequent korpuslinguistisch fundierte und an das Fachpublikum gerichtete Grammatik des Deutschen gibt es bisher nicht. Es wird sie auch noch nicht so schnell als abgeschlossenes Werk geben, denn die Aufgabe ist enorm, wenn man sich zum Ziel setzt, mit modernen korpuslinguistischen Methoden die Vielfalt und Variabilität der deutschen Grammatik in großer Detailschärfe zu erfassen, und gleichzeitig für die Validierbarkeit der Ergebnisse sorgen will. Zeitnah möglich ist aber die Erarbeitung einzelner Bausteine einer solchen Grammatik. Mit den „Bausteinen einer Korpusgrammatik des Deutschen“ werden nach und nach korpusgestützte Untersuchungen im Open-Access-Format vorgelegt, die als einzelne Kapitel einer neuen wissenschaftlichen Grammatik des Deutschen gestaltet sind. Die Idee der „Bausteine“ ist im Projekt „Korpusgrammatik – grammatische Variation im standardsprachlichen und standardnahen Deutsch“¹ des Leibniz-Instituts für Deutsche Sprache in Mannheim (IDS) entwickelt worden. Die „Bausteine“ entstehen am IDS und werden in Kooperation mit dem Germanistischen Seminar der Universität Heidelberg unter Trägerschaft des Europäischen Zentrums für Sprachwissenschaften (EZS) herausgegeben. Die Reihe setzt sich dabei das ausdrückliche Ziel, einen Beitrag zur offenen Wissenschaft im Sinne der Open-Science-Bewegung zu leisten, indem Bestandteile des wissenschaftlichen Prozesses frei zugänglich und auch nachnutzbar gemacht werden: Neben Texten in Form von Grammatikkapiteln werden Ergebnisse korpuslinguistisch-statistischer Analysen und bei ausgewählten Themen die zugrunde liegenden Sprachdaten veröffentlicht. Geplant sind jährlich bis zu zwei Ausgaben jeweils in einer Online- und einer Print-on-Demand-Version. Bei der Qualitätssicherung werden die Herausgebenden der

1 Vgl. <https://www.ids-mannheim.de/gra/projekte/korpusgrammatik.html>, abgerufen am 27.12.2019.

„Bausteine“ durch den [Wissenschaftlichen Beirat](#) unterstützt, der u. a. für die Begutachtung der Beiträge sorgt.

2 Warum braucht man eine neue wissenschaftliche Grammatik des Deutschen?

In der neueren wissenschaftlichen Grammatikschreibung zum Gegenwartsdeutschen hat sich ein besonderes Paradigma herauskristallisiert. Demnach sollen Grammatikwerke deskriptiv und gleichzeitig theoretisch fundiert sein, ohne jedoch dabei nur einer Theorie folgen zu müssen (vgl. z.B. Eisenberg 2013a: 7). Ein gewisser theoretischer Eklektizismus scheint sowieso notwendig zu sein, wenn man eine konsequente Deskription anstrebt, da Theorien nicht lückenlos sind (dazu Hoffmann 1997: 4). Im Zweifelsfall soll man versuchen, dem grammatischen Phänomen und nicht einer bestimmten Theorie gerecht zu werden. Diesen Ideen fühlen sich auch die Autor/inn/en der „Bausteine“ verpflichtet. Allerdings gibt es bereits einige Grammatiken, in denen eine weitgehend theorieneutrale Deskription beherzigt wird (dazu z.B. Paranhos Zitterbart 2010: 67–70, Eisenberg 2013a: 1–2). Warum soll also eine weitere wissenschaftliche Grammatik des Deutschen entstehen?

In den folgenden Ausführungen werden einige ausgewählte Grammatiken besonders häufig als Beispiele herangezogen. Es sind dies die Dudengrammatik (2016), der „Grundriss der deutschen Grammatik“ von Eisenberg (2013a, 2013b), die „Grammatik der deutschen Sprache“ (GDS) von Zifonun et al. (1997) und die „Grammatik des Deutschen im europäischen Vergleich“ (GDE) von Gunkel et al. (2017). Zu betrachten ist zunächst, an wen sich diese Werke richten. Im populärsten, der Dudengrammatik, die als einzige das Prädikat *wissenschaftlich* nicht für sich beansprucht, werden als Einsatzorte Schulen und Universitäten und als Zielgruppe beruflich mit Sprache Beschäftigte, persönlich an sprachlichen Fragen Interessierte sowie Lehrende und Lernende des Deutschen als Fremdsprache genannt (Dudengrammatik 2016: 5). Eisenberg (2013a: IX–X) empfiehlt seine Grammatik als Lehrbuch und hebt unter ihrer Leserschaft Studierende hervor. Hoffmann (1997: 2) sieht die Zielgruppe der GDS in Sprachinteressierten mit Vorkenntnissen: Linguisten, Akademikern aus Nachbardisziplinen, Lehrenden und Studierenden. In der GDE (2017: 1) schließlich werden sprachwissenschaftliche Vorbildung, aber keine speziellen grammatiktheoretischen Kenntnisse vorausgesetzt. Man könnte hier eine von Werk zu Werk zunehmende Spezialisierung der Kenntnisse herauslesen, die von den Adressierten erwartet werden. Der nächste Schritt wäre es, direkt die Sprachwissenschaftler/innen zum anvisierten Kernpublikum zu erklären.

Ein Bekenntnis zu einem solchen Schritt ist aus unserer Sicht im Sinne des wissenschaftlichen Fortschritts notwendig. So soll die neueste Auflage der Dudengrammatik (2016) dem Vorwort zufolge den heutigen Stand des Wissens über Formen und Funktionen der deutschen Standardsprache präsentieren, und sie erreicht tatsächlich einen Grad an deskriptiver Präzision und theoretischer Durchdringung, der auch einer deklariert wissenschaftlichen Grammatik gut zu Gesicht stünde. Doch als traditionelles Nachschlagewerk „für alle“ muss die Dudengrammatik den Spagat zwischen deskriptiver Adäquatheit und Allgemeinverständlichkeit, zwischen Deskription und Normierung versuchen. Autor/inn/en rein wissenschaftlicher Grammatiken dagegen brauchen sich um die Allgemeinverständlichkeit nicht so große Sorgen zu machen und dürfen die Normierung sogar – so der Konsens der allermeisten Linguist/inn/en – gar nicht anstreben (dazu differenziert Hoffmann 1997: 6, Habermann 2010: 11–12, Eisenberg 2013a: 8–10). Unter den Grammatiken mit wissenschaftlichem Anspruch gibt es einige, in denen explizit didaktische Zwecke verfolgt werden (z. B. Eisenberg 2013a, 2013b, tendenziell Hentschel & Weydt 2013 und Engel 2009). Es gibt aber auch welche, in denen die Grammatikbeschreibung mit weniger Rücksicht auf die Didaktik betrieben wird (z. B. GDS, GDE). Gäbe es darunter Grammatiken, die sich explizit an Sprachwissenschaftler/innen als Kernpublikum richteten, so hätten ihre Autor/inn/en endgültig die Berechtigung, die Grenzen der grammatikografischen Deskription auszureizen. Sie könnten es auch zu einem ihrer zentralen Ziele erklären, mit der tiefer gehenden Beschreibung direkte Anknüpfungspunkte für Spezialforschung zu bieten, die den wissenschaftlichen Fortschritt noch weiter vorantreibt.

Eine Steigerung der Deskriptionstiefe, die in größerer Detailschärfe resultiert, passt also nicht nur zum Konzept einer deskriptiven Grammatik, sie ist auch dem wissenschaftlichen Fortschritt zuträglich. Sie lässt aber den Umfang einer Grammatik anwachsen und kann in der Praxis der Werkerstellung auf Kosten der Vollständigkeit gehen, die oft von Grammatiken erwartet wird. Allerdings ist das Vollständigkeitspostulat in seiner Absolutheit ohnehin utopisch, sodass sich wissenschaftliche Grammatiken mit mehr oder weniger fragmentarischen Lösungen begnügen müssen (dazu Hoffmann 1997: 4). Unter den weiter oben genannten Grammatiken sind die GDS und die GDE besonders umfangreich. Sie bestechen auch durch eine besondere Detailschärfe (zur GDS vgl. z. B. Paranhos Zitterbart 2010: 70) und eine beeindruckende theoretische Fundiertheit. Sie weisen aber zwangsläufig Lücken auf oder behandeln nur Teilbereiche.²

2 Hoffmann (1997: 11) weist in der Einführung der GDS u. a. auf das Fehlen „eine[r] wirklich ausgearbeitete[n] Morphosyntax“ und einer Behandlung der deutschen Wortbildung hin. Die GDE (2017: 1) ist zunächst nur als Grammatik des Nominals angelegt.

In beiden Grammatiken wird die Zielgruppe zwar noch nicht entsprechend eingeschränkt, jedoch kommen sie in der Praxis unserem Postulat einer Grammatik für Sprachwissenschaftler/innen schon sehr nahe. Dabei ist die GDE 2017 erschienen und damit ziemlich aktuell. Die Autor/inn/en nehmen darin allerdings eine spezielle Perspektive ein, die auf die Grammatik des Deutschen im Vergleich mit europäischen Kontrastsprachen ausgerichtet ist, und beschreiben dabei nur nominale Kategorien und Konstruktionen. Die GDS wiederum konzentriert sich zwar auf das Deutsche allein und bietet dabei eine überaus breit angelegte Beschreibung, sie ist aber bereits vor 23 Jahren erschienen. Was also offensichtlich fehlt, ist eine Grammatik, die einerseits dadurch, dass sie nur das Deutsche in den Fokus nimmt, ein entsprechend höheres Maß an Deskriptionsbreite und -tiefe erreichen kann und die andererseits einen Beitrag im aktuellen Forschungsdiskurs darstellt.

Warum ist die Aktualität so wichtig? Es liegt zunächst auf der Hand, dass die Forschung nicht stillsteht und immer wieder neue Erkenntnisse hervorbringt, die auch in wissenschaftliche Grammatiken einfließen sollen. Es ist außerdem zu prüfen, ob in neuester Zeit nicht sprachliche Entwicklungen stattgefunden haben, die zu Phänomenen geführt haben, die in wissenschaftlichen Grammatiken noch nicht beschrieben sind. Und schließlich sind in der grammatischen Beschreibung neue Forschungsparadigmen zu berücksichtigen, die bisher in Grammatiken noch nicht angewendete Methoden, noch nicht eingenommene Blickwinkel und noch nicht erörterte Fragestellungen relevant machen. So wurde in der grammatischen Forschung der letzten Jahrzehnte die empirische Fundierung immer wichtiger. Wesentlich dazu beigetragen hat die rasante Weiterentwicklung von Sprachkorpora, die computerbasiert aufgebaut, aufbereitet, bereitgestellt, durchsucht sowie analysiert werden und heute zum Teil Größen von Milliarden von Wörtern erreichen können. Was die Arbeit mit Korpora dabei anderen Empiriearten wie Befragung oder Experiment voraushat, ist neben der potenziell sehr großen Menge an Sprachdaten deren natürlichere Entstehungssituation (vgl. Konopka 2018: 152–153). Übrigens verstehen sich schon alle weiter oben aufgezählten Grammatiken als empirisch orientiert. Die Dudengrammatik (2016), die GDS und die GDE nutzen auch in größerem Maße Korpora.³ Sie ziehen sie vor allem als Quelle für Belege heran, die theoretische Feststellungen exemplifizieren. In der neueren grammatischen Forschung werden Korpora hingegen immer häufiger dazu benutzt, theoretische Feststellungen systematisch zu evaluieren (vgl. Meurers & Müller 2009). Sie werden außerdem zunehmend als Quelle für erklärungsbedürftige Phänomene und neue

3 Eisenberg (2013a: 9) dagegen äußert sich kritisch zum konsequenten „Datensammeln“, das sich schnell verselbständige und damit das Entstehen einer Grammatik verhindere.

Fragestellungen exploriert.⁴ Korpusorientierte Grammatikforschung heißt also heute Untersuchung großer Mengen authentischen, elektronisch durchsuchbaren und mit Metadaten angereicherten Textmaterials. Sie hat in den letzten Jahren über Einzelphilologien hinweg zu einer regelrechten Welle interessanter Studien geführt (vgl. z. B. die Beiträge in Konopka et al. 2011 und Fuß et al. 2018). In der Grammatikografie des Englischen hat sie auch bereits vor einiger Zeit ihren Niederschlag in korpusgestützten Grammatiken gefunden, etwa der „Cambridge Grammar of English“ von Carter & McCarthy (2006) und vor allem der „Longman Grammar of Spoken and Written English“ von Biber et al. (2006 [1999]).⁵ In der Germanistik jedoch wurde dieser Ansatz bisher nur im Rahmen von kleineren Studien und Monografien bzw. von Untersuchungen zu überschaubaren Grammatikausschnitten verfolgt. Eine breit angelegte Grammatik des Deutschen auf korpuslinguistischer Basis stellt immer noch ein Desiderat dar.

Korpuslinguistik ermöglicht genauere, detailliertere Deskription. Besiegelt ist damit nicht nur die endgültige Abwendung von der normativen Grammatik oder von einer Grammatik, die einen einheitlichen Sprachgebrauch vorgaukelt. Unausweichlich ist jetzt konsequenterweise auch die Hinwendung zu einem Grammatikkonzept, gemäß dem die Sprache in ihrem ganzen Facettenreichtum, ihrer Variabilität und in ihrem tatsächlichen Gebrauch beschrieben werden soll (vgl. Conrad 2010: 228). Variation ist mittlerweile zu einem der zentralen Themen der Grammatikschreibung geworden, während das Konzept der Standardsprache an Bedeutung eingebüßt hat bzw. jetzt anders gedacht werden muss (dazu Eichinger 2017: v. a. 17–19, Felder 2016: 23–26). Von einer neuen, korpuslinguistisch fundierten Grammatik des Deutschen ist in diesem Kontext zu erwarten, dass sie die Variation nicht nur festhält, sondern auch im Hinblick auf ihre Abhängigkeit von verschiedenen sprachinternen und/oder -externen Einflussfaktoren erklärt (vgl. Konopka 2018: 166).⁶ Dies stellt besondere Anforderungen an die Methodik der Untersuchungen, die der Grammatikbeschreibung zugrunde liegen. Die Bedeutung korpuslinguistisch-statistischer Verfahren, mit denen man die Datenfülle und -vielfalt zu fassen bekommen kann, wächst. Die Anwendung statistischer Verfahren verpflichtet dabei schon dazu, die Datenerhebung, -aufbereitung und -analyse für die Leserschaft zu dokumentieren, was sich sehr gut in das Konzept einer wissenschaftlichen Grammatik einfügt, die die Validierbarkeit der zugrunde liegenden Untersuchungen sichert. Damit wird die direkte Anschlussforschung ermöglicht, die unter Umständen sogar

4 Zur Evidenz aus dem Korpus, Validierung am Korpus und Induktion mithilfe des Korpus vgl. Konopka (2018: 151–164).

5 Zur kritischen Diskussion vgl. Mukherjee (2006).

6 Felder (2016: v. a. 7–19) zeigt instruktiv die Möglichkeiten, sprachliche Variation im Allgemeinen zu beschreiben und zu erfassen.

mit vergleichbaren Methoden und denselben sprachlichen Daten betrieben werden kann.

Die „Bausteine“ sollen es möglich machen, nach und nach eine wissenschaftliche Grammatik „aufzubauen“, die dem skizzierten Bedarf gut begegnen kann. Sie wird im Weiteren als *Korpusgrammatik* bezeichnet.

3 Eine *Korpusgrammatik* in spe

Die ersten konzeptionellen Überlegungen zu einer *Korpusgrammatik* wurden in Bubenhofer et al. (2014) vorgestellt. Von 2011 bis 2019 entstand am IDS außerdem eine Reihe von Pilotstudien, in denen anhand detaillierter Untersuchungen von Variationsphänomenen aus verschiedenen Grammatikbereichen exemplarische Fragestellungstypen und dazu passende korpuslinguistische Herangehensweisen herausgearbeitet wurden. Einen Überblick über diese Arbeiten geben Fuß et al. (2019). Nun werden die ersten Versionen von Kapiteln der künftigen *Korpusgrammatik* als „Bausteine“ vorgestellt. Es handelt sich dabei gewissermaßen um Probekapitel – die Anlage der Kapitel und die Darstellungsformen werden sich mit weiteren „Bausteinen“ weiterentwickeln: Es muss ja einiges ausprobiert werden, denn die *Korpusgrammatik* hat auch Neues vor. Zu beginnen ist aber mit den eher traditionellen Rahmeninformationen.

Gegenstand der *Korpusgrammatik* ist die Grammatik des geschriebenen Deutsch, wie es sich in Texttypen manifestiert, die Standardsprachliches erwarten lassen (vgl. Bubenhofer et al. 2014: 23–32). Der besondere Fokus liegt auf der Variation. Den grammatischen Gegenstandsbereich bilden Morphologie und Syntax, zuzüglich der Ausschnitte benachbarter Gebiete (Phonologie und Graphematik einerseits, „Textgrammatik“ andererseits), die morphologische und syntaktische Variation direkt begründen oder beeinflussen. Die *Korpusgrammatik* strebt eine empirisch fundierte Deskription an und stützt sich dabei hauptsächlich auf das sogenannte Untersuchungskorpus, das ein speziell aufbereitetes Teilkorpus des Deutschen Referenzkorpus (DeReKo, vgl. Kupietz et al. 2018)⁷ darstellt (Genaueres in Abschnitt 5 weiter unten). Das primär anvisierte Publikum der *Korpusgrammatik* sind Sprachwissenschaftler/innen.

Da es offenkundig utopisch ist, die gesamte Grammatik einer Einzelsprache in ihrer Breite und Tiefe zu erfassen, wurden in der obigen Kurzcharakteristik für die *Korpusgrammatik* von vorn herein gewisse Einschränkungen vorgenommen. In der Tradition neuerer Grammatiken beschränken wir uns auf das Standarddeutsche. Zwei weitere Einschränkungen sind durch die korpuslinguistische

⁷ Siehe auch <https://www.ids-mannheim.de/kl/projekte/korpora.html> (abgerufen am 29.11.2019).

Herangehensweise mitbedingt. Die eine ist medialer Natur: Die empirisch fundierte Deskription beschränkt sich auf das geschriebene Deutsch. Ein wichtiger Grund dafür ist, dass es sehr umfangreiche, inzwischen auch ausdifferenzierte und gut aufbereitete Korpora des geschriebenen Deutsch gibt, aber noch keine vergleichbaren Sammlungen des gesprochenen Deutsch.⁸ Die andere Einschränkung resultiert aus der Umgewichtung der grammatischen Themen und Perspektiven, wie sie in Abschnitt 2 beschrieben wurde: Besondere Aufmerksamkeit gilt grammatischer Variation. Selbstverständlich wird dabei davon ausgegangen, dass die Standardsprache Variation zulässt. So sollen sich die behandelten Phänomene auf das Standarddeutsche beziehen, da aber die Abgrenzung der Standardsprache ein Problem an sich darstellt (vgl. Bubenhofer et al. 2014: 21–32), werden zuweilen auch Übergänge zum Substandard ausgeleuchtet. Um solche Übergänge zu erfassen, werden ergänzend zum Untersuchungskorpus diejenigen Teilkorpora des Web-Korpus DECOW16 (Schäfer & Bildhauer 2012; Schäfer 2015) herangezogen, die „Grenzwertiges“ erwarten lassen. Im Weiteren soll sich die *Korpusgrammatik* weitgehend auf die Morphologie (Wortbildung und Flexion) und Syntax beschränken. Ein zentraler Grund für diese Einschränkung sind Machbarkeitsüberlegungen. Ausgewählt wurden mit Morphologie und Syntax aber Gebiete, die traditionell als Kernbereiche der Grammatik aufgefasst werden. Zuallerletzt führt die Tatsache, dass sich die *Korpusgrammatik* primär an Sprachwissenschaftler/innen richtet, zusammen mit der Priorisierung der Variation dazu, dass Phänomene, die schwer zu erklären bzw. in der Forschung umstritten sind, „einfachere“ und durch wissenschaftlichen Konsens geprägte Themen in den Hintergrund drängen bzw. im Falle von unstrittigen „Basics“ manchmal ganz verschwinden lassen.

Wie von einer empirisch fundierten Grammatik dieser Art zu erwarten, wird sich die *Korpusgrammatik* zunächst an Oberflächenphänomenen orientieren. Die zugrunde liegenden Untersuchungen von Variationsphänomenen werden unter Suspendierung des Konzepts „freie Variation“ durchgeführt, wobei versucht wird, inner- und außergrammatische Einflussgrößen der Variation zu ermitteln. Die Ermittlung ersterer soll grundlegend durch Korpusannotationen nach grammatischen Merkmalen ermöglicht werden, die Erfassung letzterer durch Dokumentmetadaten, die die Korpora etwa nach zentralen Textklassen (Registern), geografisch oder zeitlich strukturieren. Trotz der primären Orientierung an Oberflächenphänomenen werden durchaus theoretische Erklärungen anvisiert,

8 Zum geschriebenen Deutsch (insbesondere in den Druckmedien) sind außer DeReKo vor allem die Korpora des Digitalen Wörterbuchs der deutschen Sprache (DWDS) zu nennen (vgl. <https://www.dwds.de/d/korpora>, abgerufen am 20.6.2020). Ein Beispiel für ein im Aufbau begriffenes ausdifferenziertes Korpus der gesprochenen Sprache ist Folk (vgl. <http://agd.ids-mannheim.de/folk.shtml>, abgerufen am 26.5.2020).

und es wird theoretische Fundiertheit angestrebt, allerdings ohne die Theorieutralität aus den Augen zu verlieren. Letztere ist überhaupt die Voraussetzung dafür, dass die *Korpusgrammatik* von einem breiteren Fachpublikum umfassend benutzt werden kann. Einem solchen Publikum soll sie nämlich nicht nur Interpretationen sprachlicher Daten liefern. Sie soll auch quantitative Übersichten und differenzierte Datensammlungen bereitstellen, auf denen theoretische Forschung verschiedener Prägung aufsetzen kann. Außerdem hat sie so beschaffene Voruntersuchungen von Variationskontexten zu liefern, dass die empirische Spezialforschung direkt daran anschließen kann, auch wenn sie in spezielleren Frameworks oder mit anderen Methoden arbeitet.

Die *Korpusgrammatik* zeigt sich in noch einem Bereich anders als die bisherigen Grammatiken. So beeinflusst die Benutzung korpuslinguistischer Methoden die Inhalte und die Anlage der Grammatikkapitel. Diesen liegen oft statistische Untersuchungen zugrunde, die transparent dargestellt sein müssen, um methodisch nachvollziehbar zu sein, eine Bewertung der statistischen Befunde zuzulassen und idealerweise Replikation zu erlauben. Eine adäquate Darstellung von Ergebnissen solcher Untersuchungen führt dazu, dass viele Kapitel der *Korpusgrammatik* stärker durch methodische Argumentation geprägt sind und einen weniger deklarativen Stil aufweisen als für andere Grammatiken üblich. Dafür sind sie auch stärker problemlösungsorientiert und zwar in dem speziellen Sinne, dass sie nicht nur das Resultat zeigen, sondern auch den Weg zur Erkenntnis nachzeichnen, was dem forschenden Publikum nur entgegenkommen kann.

So hat die *Korpusgrammatik* zwar einige Unterschiede zu den in Abschnitt 2 genannten Grammatiken des Deutschen aufzuweisen, dennoch sind diese für die *Korpusgrammatik* einerseits Fundament und andererseits eine große Inspirationsquelle. Sie sind eine Fundgrube des grammatischen Wissens, das als weitgehend gesichert gilt. Sie liefern mit ihren Beschreibungen auch Ausgangspunkte für Vertiefungen und laden zur Evaluation und Weiterentwicklung von Generalisierungen ein. Sie bleiben in vielerlei Hinsicht – etwa bei den Darstellungsformen, der präzisen Sprache und dem grammatikografischen Duktus – ein Vorbild. Letztlich machen sie es möglich, dass die *Korpusgrammatik* als Handbuch für Sprachwissenschaftler/innen mehr in die Tiefe gehen und sich nicht nur der Präsentation von bereits Bekanntem, sondern auch verstärkt der Problemlösung widmen kann.

4 Warum „Bausteine“?

Vom Beginn der Arbeiten an war allen Beteiligten klar, dass sehr viel Zeit vergehen würde, bis eine Grammatik mit diesen Ansprüchen als umfassendes Buch vorgelegt werden könnte, und es sprachen mehrere gewichtige Argumente

dagegen, mit der Publikation der Teilergebnisse zu warten, bis die gesamte *Korpusgrammatik* auf einmal erscheinen kann.

Wie andere Grammatiken kann auch die *Korpusgrammatik* nie wirklich vollständig werden – besonders offensichtlich wird dies angesichts der Ziele, die Methodik nachvollziehbar zu gestalten und auf die Deskriptionstiefe zu setzen (die im Prinzip vor keiner natürlichen Grenze haltmachen muss). Es ist sinnvoll, sich nicht nur damit abzufinden, sondern sich an den Umstand auch anzupassen: Da eine Grammatik mehr oder weniger fragmentarisch bleiben muss und sowieso immer ergänzt werden kann, spricht wenig dagegen, die *Korpusgrammatik* von vornherein in nach und nach erscheinenden Einzelteilen, sprich Bausteinen, zu veröffentlichen, die sich mit der Zeit zu einem immer vollständigeren (aber nie gänzlich vollständigen) Bild der deutschen Grammatik zusammensetzen.

Die zeitnahe Veröffentlichung von Grammatikkapiteln sichert deren Aktualität. Die ihnen zugrunde liegenden Untersuchungen orientieren sich am aktuellen Forschungsstand, und ihre Ergebnisse sind eigentlich Beiträge im laufenden Forschungsdiskurs. In einigen Jahren allerdings werden die Untersuchungen ihre innovative Wirkung verloren und die Ergebnisse an Aktualität eingebüßt haben. Im ungünstigsten Fall werden die Ergebnisse schon überholt sein, wenn man mit ihrer Veröffentlichung darauf wartet, bis alle vorgesehenen Kapitel der Grammatik vorliegen. Das spricht für die „Zwischenveröffentlichung“ bereits erstellter Kapitel.

Wird im Team gearbeitet, übernehmen einzelne Autor/inn/en typischerweise verschiedene thematische Bereiche. Sie bearbeiten sie dann oft parallel. Die Grammatik entsteht also nicht unbedingt linear entlang einer angestrebten Gliederung. Die einzelnen Kapitel können außerdem – auch wenn man eine formale und inhaltliche Einheitlichkeit anstrebt – spürbar durch die Persönlichkeit der jeweiligen Autor/inn/en geprägt sein.⁹ Beides begründet oft eine gewisse Eigenständigkeit der Kapitel, die zu ihrer separaten Publikation in den „Bausteinen“ gut passt. Die Kapitel, die für eine Ausgabe der „Bausteine“ ausgewählt werden, müssen nicht thematisch zusammenhängen. Auch eignen sich nicht alle Kapitel dafür, in die „Bausteine“ aufgenommen zu werden. Die Kriterien, nach denen Kapitel für die „Bausteine“ ausgesucht werden, sind ein ausreichender linguistischer Informationsgehalt und eben gerade ein gewisses Maß an Eigenständigkeit.¹⁰

9 Vgl. die Erfahrungen der Autor/inn/en der GDE (2017: 8).

10 So dienen manche Überblickskapitel (vgl. Abschnitt 5) nicht so sehr der linguistischen Beschreibung wie einer Vorstrukturierung der Themen, die in den darauf folgenden Kapiteln beschrieben werden.

5 Sprachliche Daten, Untersuchungsmethoden, Präsentationsformen und Qualitätssicherung

Die den „Bausteinen“ zugrunde liegenden Untersuchungen basieren vor allem auf dem Untersuchungskorpus des IDS-Projekts „Korpusgrammatik – grammatische Variation im standardsprachlichen und standardnahen Deutsch.“ Das Untersuchungskorpus umfasst zurzeit 23 Mio. Dokumente, 5,8 Mrd. Wörter sowie 7 Mrd. Tokens und stellt einen Ausschnitt des Deutschen Referenzkorpus (DeReKo, vgl. Kupietz et al. 2018), Release 2017-II, dar. Es enthält vor allem klassische Druckerzeugnisse und wurde im Rahmen des Projekts mit Dokumentmetadaten wie ‚Land‘, ‚Region‘, ‚Register‘, ‚Medium‘, ‚(inhaltliche) Domäne‘ und ‚Entstehungszeit‘ angereichert und stratifiziert (Bubenhofer et al. 2014: 54–76). Seine aktuelle Zusammensetzung ist unter <https://grammis.ids-mannheim.de/korpusgrammatik/6616> (abgerufen am 23.7.2020, DOI: 10.14618/korpusgrammatik) beschrieben. Mitunter werden Teilkorpora des Untersuchungskorpus gebildet, die in den einzelnen Bausteinen an einschlägigen Stellen spezifiziert werden. Um Internettexte besser in den Blick zu bekommen, werden ergänzend Teilkorpora des DECOW16 (Schäfer & Bildhauer 2012; Schäfer 2015) benutzt.

Das Untersuchungskorpus verfügt über zahlreiche Annotationsebenen. In dieser Ausgabe der „Bausteine“ wurden insbesondere Annotationen nach Lemma und Wortart (v. a. TreeTagger, vgl. Schmid 1994, und SMOR, vgl. Schmid et al. 2004), morphologischen Merkmalen (MarMoT, vgl. Müller et al. 2014), Wortbildung (Canoo)¹¹ sowie Eigennamen (Stanford NER, Finkel et al. 2005) verwendet (vgl. <https://grammis.ids-mannheim.de/korpusgrammatik/6625>, abgerufen am 23.7.2020, DOI: 10.14618/korpusgrammatik). Für die einzelnen Untersuchungsthemen werden die sprachlichen Daten auch spezifisch nachannotiert, was an entsprechenden Stellen beschrieben wird.

In den statistischen Analysen werden in dieser Ausgabe der „Bausteine“ verschiedene Maße und Verfahren benutzt, darunter Chi-Quadrat-Tests und Pearson-Residuen mit entsprechenden Visualisierungen, Baayens (2009: 902) *potenzielle Produktivität* als morphologisches Produktivitätsmaß, Modellierungen von Worthäufigkeitsverteilungen mithilfe des Softwarepakets zipfR (Evert & Baroni 2007) sowie binäre logistische Regressionen. Bei der statistischen Modellierung und Visualisierung kommt insbesondere die freie Open-Source-Software R (R Core Team 2020) zum Einsatz. Neben den statistischen Analysen werden umfangreiche und vielschichtige qualitative Sichtungen der Daten vorgenommen, die öfter auch explorative Vorstufen für quantitative Analysen bilden. Die

11 Siehe <https://web.archive.org/web/20200326230620/http://canoonet.eu/> (abgerufen am 10.5.2020).

Kenndaten der Analysen sowie Hintergrunderläuterungen befinden sich an einschlägigen Stellen im Haupttext oder sind bei logistischen Regressionen in die Anhänge ausgelagert.

Die Arbeit mit der korpuslinguistischen Methodik und die Notwendigkeit, die statistischen Methoden genau darzustellen, werden die Kapiteltypen in der *Korpusgrammatik* und damit die Präsentationsformen der „Bausteine“ beeinflussen. Wie in den meisten anderen Grammatiken sollen auch in der *Korpusgrammatik* größere Themenbereiche soweit möglich in der Abfolge vom Allgemeinen zum Besonderen behandelt werden. Dies wird sich in verschiedenen Typen von Kapiteln spiegeln. An einem Ende der Skala steht dann ein Überblickskapitel, in dem ein ganzer Phänomenbereich unter Verwendung korpuslinguistischer Perspektiven gesichtet wird und in dem auf interessante Variationsbereiche hingewiesen wird, am anderen Ende ein Kapitel, in dem ein spezielles Variationsphänomen eingehend analysiert wird. Ein solches Kapitel kommt schon einer Fallstudie nahe: Darin werden oft die Einflussgrößen des konkreten Variationsfalls ermittelt, statistisch geprüft und in ihrer Effektstärke bestimmt. Zwischen diesen beiden Kapiteltypen stehen einerseits Überblickskapitel, die auf niedrigeren Abstraktionsstufen angesiedelt sind, und andererseits solche Kapitel, die die Variabilität der Grammatik auf abstrakteren Stufen untersuchen als der eines konkreten Variationsphänomens. Tendenziell werden in Überblickskapiteln mehr deskriptiv-statistische und explorative Methoden angewendet. Die inferenzstatistischen Methoden nehmen zu, je näher das Kapitel an die Fallstudie herankommt. Hier wächst auch der Bedarf an Dokumentation der Methoden. In dieser Ausgabe der „Bausteine“ erscheinen zwei Überblickskapitel unterschiedlichen Abstraktionsgrades und zwei Kapitel, die einer Fallstudie nahekommen (vgl. Abschnitt 7 weiter unten).

Mit den „Bausteinen“ wird eine Veröffentlichungsstrategie verfolgt, die den Mehrwert einer Online-Publikation ausnutzt, um die zugrunde liegenden Untersuchungen im intendierten Sinne nachvollziehbar und damit anschlussfähig zu machen. Anders als es bei bisherigen Grammatiken des Deutschen der Fall ist, werden bei ausgewählten Themen die ausgewerteten Sprachdaten online veröffentlicht und mit den „Bausteine“-Texten verlinkt. Ergebnisse komplexerer korpuslinguistisch-statistischer Analysen werden direkt in den Bausteinen dokumentiert oder – bei größerem Umfang – ebenfalls online zugänglich gemacht. Mit der Wahl eines Open-Access-Formats sowie mit der Veröffentlichung von Datensätzen (Open Data) und Dokumentationen statistischer Analysen (Open Methodology) trägt die Reihe zur offenen Wissenschaft bei.

Durch die geschilderten Maßnahmen wird die Entstehung der *Korpusgrammatik* weitgehend transparent gehalten, sodass den Autor/inn/en bei ihren grammatikografischen Bemühungen gewissermaßen zugesehen werden kann. So können auch die Reaktionen aus der Forschungsgemeinde auf das Vorhaben

Korpusgrammatik zurückwirken. Eine besondere Rolle im Entstehungsprozess der „Bausteine“ kommt dem [Wissenschaftlichen Beirat](#) zu. Er setzt sich aus namhaften Kolleg/inn/en zusammen, die über besondere Expertise in der Grammatikografie des Deutschen einerseits und/oder in der Korpus- und Computerlinguistik andererseits verfügen und sich bereit erklärt haben, uns auf dem Weg zu einer *Korpusgrammatik* zu begleiten.

6 Geplanter Aufbau der *Korpusgrammatik*

Die Bausteine müssen modular in eine Makrostruktur einpassbar sein, die der anvisierten Gliederung der künftigen *Korpusgrammatik* entspricht. Geplant ist eine in ihren Grundzügen traditionelle Gliederung mit zwei Hauptteilen. Im ersten wird die Grammatik von den herkömmlichen grammatisch bestimmbareren Ausdrucksklassen her betrachtet, die unterhalb der Satzebene liegen, wie Wort, Substantiv oder Nominalphrase. Im zweiten Hauptteil bildet der Satz als Realisierung der kommunikativen Minimaleinheit den Ausgangspunkt, und es werden dort syntaktische Funktionen und Relationen untersucht, die typischerweise auf der Satzebene zu beobachten sind. Wenn man umschließende Gliederungseinheiten wie Vorwort, eine allgemeine Einführung, ein einleitendes Kapitel zur korpuslinguistischen Basis und Methodik, ein Literaturverzeichnis und mögliche Register erst einmal beiseitelässt, sieht die Gliederung für den ersten Hauptteil der *Korpusgrammatik* folgende Großkapitel vor:

1. Wortbildung
2. Wörter und Wortgruppen
3. Substantiv, Nomen und Nominalphrase
4. Adjektiv und Adjektivphrase
5. Adverb und Adverbphrase
6. Andere nicht flektierbare Wortarten
7. Verb, Verbalkomplex, Verbgruppe (Verbalphrase)

Der erste Hauptteil ist somit als eine „Wort- und Phrasengrammatik“ konzipiert (zu einer in Grundzügen ähnlichen Anlage vgl. z.B. die korpusgestützte Grammatik des Englischen von Biber et al. 2006 [1999]: v.a. 44). Im Hinblick auf die grammatisch bestimmten Ausdruckseinheiten folgt der Aufbau im Kern einer aszendenten Betrachtungsweise, die die Reihenfolge Morphologie – Syntax vorgibt. Dabei wird zuerst in einem getrennten Großkapitel die Wortbildung behandelt. In jedem der darauf folgenden Großkapitel zu flektierbaren Wörtern und darauf basierenden Wortgruppen sollen dann die Bereiche Flexion und Syntax nach einer einführenden Übersicht über Flexionskategorien und ihre Realisierungsformen zu einer Morphosyntax zusammenfließen, sobald es um die

Betrachtung konkreterer Variationsfälle geht. Damit steht im ersten Hauptteil eine eher an formalen Aspekten orientierte Perspektive im Vordergrund.

Die Gliederung des zweiten Hauptteils soll durch Themen wie die folgenden bestimmt sein:

1. Kommunikative Minimaleinheit, Satz, Satzkomponenten
2. Komplementation
3. Supplementation
4. Kongruenz und Rektion
5. Wortstellung
6. Informationsstruktur
7. Komplexere Strukturen (Asyndese, Syndese, Koordination, Subordination, komplexe Sätze)

Zunächst wird mit dem Satz als einer auf dem finiten Verb basierenden, prototypischen Realisierung der kommunikativen Minimaleinheit (dazu Zifonun 1997a: 86; 1997b: 608) die zentrale Bezugsgröße der Beschreibung etabliert. Man kann somit beim zweiten Hauptteil tendenziell von einer „Satzgrammatik“ sprechen. Allerdings werden darin im Weiteren syntaktische Funktionen und Relationen thematisiert, die zwar typischerweise, aber nicht nur in Sätzen eine Rolle spielen. Der Aufbau ist bis in die Großkapitel Komplementation und Supplementation hinein deszendend. In diesen beiden Großkapiteln rückt dann langsam die funktionale Perspektive als Ordnungsprinzip in den Vordergrund, denn syntaktische Funktionen werden nicht nur für die Satzebene postuliert, sondern auch für Ausdruckseinheiten, die unterhalb der Satzebene liegen (vgl. z. B. Verbkomplemente einerseits und Komplemente in der Nominal- oder Adjektivphrase andererseits). So sind auch die Relationen der Kongruenz und der Rektion für Ausdruckseinheiten unterschiedlichen Ranges relevant.¹² Die folgenden, ebenfalls relational bzw. funktional interpretierbaren Themen werden (wie „Wortstellung“ und „Informationsstruktur“) üblicherweise im Kontext von Sätzen behandelt oder betreffen (wie „Komplexere Strukturen“) ein Spektrum von Ausdruckseinheiten, die die Satzebene sowohl deutlich unter- als auch überschreiten können. Der Schwerpunkt im zweiten Hauptteil liegt somit insgesamt auf dem Vergleich der verschiedenen Realisierungsformen für syntaktische Funktionen und Relationen, die unterschiedlichen grammatischen Ebenen angehören können.

12 Zur Kongruenz vgl. z. B. die Numeruskongruenz zwischen Subjekt und finitem Verb auf der Satzebene und die Numerus- und Kasuskongruenz zwischen Determinierer und Nomen in einer Nominalphrase, zur Rektion etwa die Rektion der Verbkomplemente durch das Verb auf der Satzebene oder die Rektion des Kasus einer Nominalphrase durch die Präposition in einer Präpositionalphrase.

Die vorgestellte Gliederung ist tentativ. Sie ist modifizierbar und vor allem in Bezug auf den zweiten Hauptteil ergänzbar. Die Formulierungen der Kapitelüberschriften sind vorläufig. Die Gliederung muss einigermaßen flexibel sein, damit die Autor/inn/en mit Änderungen und Erweiterungen auf sich eventuell durch die Forschungsdynamik ergebende Entwicklungen reagieren können, die in ihrer Arbeit an der *Korpusgrammatik* oder in der aktuellen Forschung im Allgemeinen entstehen.

Die Zweiteilung der *Korpusgrammatik* spiegelt die Doppelperspektivik, nach der sich die Darstellung richten soll: Zum einen wird bei der Grammatikbeschreibung von formbezogen bestimmten Einheiten ausgegangen (so tendenziell im ersten Hauptteil), zum anderen von funktional bestimmten Einheiten bzw. bestimmten Funktionen (so tendenziell im zweiten Hauptteil). Die eine Perspektive involviert, dass formbezogene Einheiten in Bezug auf ihre verschiedenen Funktionen untersucht werden, d.h. eine Art semasiologische Vorgehensweise verfolgt wird, die andere Perspektive, dass für bestimmte Funktionen verschiedene Realisierungsformen ermittelt werden, d.h. quasi onomasiologisch vorgegangen wird.¹³ Diese Doppelperspektivik bekommt unter dem Aspekt der Variation einen besonderen Sinn: Beide Male wird nach Varianten geforscht, nur sind diese in ersterem Fall funktional und in letzterem Fall formal ausdifferenziert (vgl. Fuß et al. 2017: 229–233). Interessanterweise hat sich in der modernen Linguistik unter dem Begriff Variation nur die zweite Perspektive richtig etablieren können, wozu entscheidend die soziolinguistischen Arbeiten Labovs (1960: 15–17) zur „linguistischen Variante“ beigetragen haben. Wir allerdings wollen auf das innovative Potenzial einer systematischen „reziproken“ Variationsbetrachtung setzen und den bewussten Perspektivenwechsel zum festen Bestandteil unserer grammatikografischen Herangehensweise machen. Die beiden Perspektiven lassen sich auf der Ebene der Hauptteile und der Großkapitel bis zu einem gewissen Grad auseinanderhalten. In kleineren Gliederungseinheiten soll demgegenüber ein zyklischer Perspektivenwechsel es möglich machen, in der grammatischen Erkenntnis immer weiter vorzudringen (dazu Fuß et al. 2017: 235).

7 Bausteine dieser Ausgabe

Geplant sind jährlich bis zu zwei Ausgaben der „Bausteine“ in einer Online- und einer Print-on-Demand-Version. Nur diese vorliegende Ausgabe enthält eine längere Einleitung – sie führt in die gesamte Reihe ein. Die Bausteine der

13 Zur Übertragung der semasiologischen und onomasiologischen Perspektive auf die Betrachtung der Variation vgl. Ammon (2005: 30) und schon von der Gabelentz (1995 [1901]: 84–86).

vorliegenden Ausgabe sind als Kapitel verschiedener Hierarchieebenen aus den Großkapiteln „Wortbildung“ und „Substantiv, Nomen und Nominalphrase“ angelegt. Wie sie sich in die bisher geplante Struktur der Großkapitel einfügen, zeigen die folgenden Übersichten. Auf der oberen, nummerierten Ebene erscheinen die Titel aller bisher vorgesehenen Kapitel dieser Ebene. Auf der unteren, nicht nummerierten Ebene hingegen sind nur diejenigen Kapitel aufgeführt, die als Bausteine publiziert sind. Die Titel aller vier Kapitel, die den Bausteinen dieser Ausgabe entsprechen, sind hervorgehoben:

Wortbildung

1. **Grundlegende Aspekte der Wortbildung**
2. Wortbildungsarten
3. Spezialbereiche der Wortbildung
 - **Bau von und Umbau zu Adverbien: Präpositionen, Vergleiche und Flexion**

Substantiv, Nomen und Nominalphrase

1. Flexion des Nomens
2. Aufbau der Nominalphrase
3. Determination und Quantifikation
4. Attribution
 - **Starke vs. schwache Flexion aufeinanderfolgender attributiver Adjektive: mit hohem technischen/technischem Aufwand**
 - **Reihenfolge attributiver Adjektive**
5. Syntaktische Funktionen von Nominalphrasen

Der Baustein **Grundlegende Aspekte der Wortbildung** eröffnet das Kapitel Wortbildung. Er besitzt dementsprechend einen expositorischen Charakter und ist als Überblickskapitel angelegt. Er zeigt grundlegende Mechanismen der Wortbildung im Deutschen, erkundet vor allem deskriptiv-statistisch Relationen (u. a. die Produktivitätsverhältnisse betreffend) zwischen Wortbildungen und primären Wörtern und präsentiert allgemeine Wortbildungsprofile der Hauptwortarten. Der Baustein **Bau von und Umbau zu Adverbien: Präpositionen, Vergleiche und Flexion** ist als Überblickskapitel einer etwas tieferen Ebene angelegt. Er behandelt ein für die empirische wie auch die theoretische Forschung besonders herausforderndes Thema. Die Wortbildung des Adverbs ist nicht allzu häufig Forschungsgegenstand und ist daher besonders geeignet, in die „Spezialbereiche der Wortbildung“ aufgenommen zu werden. Der Baustein gibt einen deskriptiv-statistisch fundierten Überblick über die Bildungsmuster des komplexen Adverbs. Besonders fokussiert werden Adverbien, die mit Formativen enden, die in der Flexion vorkommen, und solche mit präpositionalen

Bestandteilen (vgl. für beide Adverbklassen z.B. *zunächst*). Auf quantitativen und qualitativen Korpusauswertungen bauen theoretische Überlegungen auf, die grundlegende Bedingungen für die Entstehung komplexer Adverbien zum Gegenstand haben und Prozesse aufdecken, die komplexe Adverbien mit Präpositionalphrasen korrelieren. Mit **Starke vs. schwache Flexion aufeinanderfolgender attributiver Adjektive: mit hohem technischen/technischem Aufwand** wird ein Baustein aus dem Bereich Nominalphrase vorgelegt. Darin wird ein konkreter Variationsfall behandelt, der es auch zu einem notorischen Zweifelsfall des Sprachgebrauchs gebracht hat. Mithilfe logistischer Regressionsanalysen wird nach Einflussgrößen für die Wahl zwischen starker und schwacher Flexion des zweiten Adjektivs gesucht. Berücksichtigt werden sprachsystematische Bedingungen wie die Möglichkeit einer Koordination der beiden Adjektive sowie extralinguistische Einflüsse wie das Herkunftsland des Belegs. Im Anhang erscheint die Dokumentation der statistischen Modellierung, die die Analysen nachvollziehbar macht, aber aus dem Haupttext ausgelagert werden musste, um ihn besser lesbar zu machen. Ein Teil der Sprachdaten, die der Untersuchung zugrunde liegen, ist – angereichert um die untersuchungsrelevanten Annotationen – online veröffentlicht (vgl. Abschnitt 8 unten). Die Sprachdaten wurden im Rahmen dieser Untersuchung bereits so weit aufbereitet, dass sich auch der nächste Baustein **Reihenfolge attributiver Adjektive** darauf stützen konnte. In diesem Baustein wird geprüft, wie bei Abfolgen aus zwei Adjektiven wie in *mit begeistertem, minutenlangem Beifall* die Reihenfolge gesteuert wird. Es zeigt sich, dass dabei semantische Aspekte mit anderen Einflussgrößen interagieren wie etwa der Silbenanzahl oder der Häufigkeit der involvierten Adjektive. Die eingesetzte Regressionsanalyse wird auch bei diesem Baustein in einem Anhang genauer beschrieben.

Die Beiträge der Reihe „Bausteine einer Korpusgrammatik des Deutschen“ werden weitgehend wie Kapitel einer korpusgestützten, variationssensitiven Grammatik gestaltet, aber dennoch nicht ganz so, wie sie in einem bereits fertiggestellten Werk erscheinen würden. Vor allem bei dieser ersten Ausgabe muss vieles an Ort und Stelle erklärt werden, was ansonsten in ein anderes Kapitel hätte ausgelagert werden können. Dies betrifft in besonderer Weise Ausführungen zur Methodik, deren Explizitmachung uns so wichtig erscheint. Ein Teil dieser Ausführungen wird sich in einem Kapitel zur korpuslinguistischen Basis und Methodik der ganzen *Korpusgrammatik* wiederfinden, da einige methodische Aspekte konstant bleiben und einige Verfahren immer wieder angewendet werden. Jetzt schon einen Baustein zu schreiben, der einem solchen Kapitel entspricht, ist jedoch verfrüht, da sich das Spektrum der eingesetzten Methoden mit weiteren Themen der *Korpusgrammatik* und damit einhergehenden neuen Problemstellungen um einiges erweitern wird. Dadurch, dass die *Korpusgrammatik* mit der Publikation der „Bausteine“ portionsweise

entsteht, sind also gewisse Redundanzen vorprogrammiert. Die Leser/innen mögen durch die Aktualität der präsentierten Forschungsergebnisse, die relative Eigenständigkeit einzelner Bausteine und die Möglichkeit, das Entstehen der *Korpusgrammatik* mit frühzeitigen Reaktionen zu begleiten, entschädigt sein.

Die einzelnen Bausteine haben verschiedene Autor/inn/en, die sich auf ein gemeinsames inhaltliches Programm und die wichtigsten Aspekte der formalen Gestaltung geeinigt haben. Auf eine stilistische Vereinheitlichung der Texte wurde bewusst verzichtet. Wir hoffen, dass Stilunterschiede die Leser/innen von den Inhalten nicht ablenken werden.

8 Veröffentlichte Forschungsdaten zu dieser Ausgabe und Verlinkung mit Internetquellen

Die Bausteine **Starke vs. schwache Flexion aufeinanderfolgender attributiver Adjektive: mit hohem technischen/technischem Aufwand** und **Reihenfolge attributiver Adjektive** basieren in wichtigen Teilen auf dem aufbereiteten Datensatz `ATTRIBUTIVE_ADJEKTIVE_1`. Dieser ist die Grundlage für die „Datenbank attributive Adjektive“, die unter der folgenden Adresse öffentlich zugänglich ist:

<https://grammis.ids-mannheim.de/attributiveadjektivedb>
(abgerufen am 11.7.2020)
DOI: 10.14618/attributiveAdjektiveDB

Die Datenbank besteht aus 1.598 Belegen zu artikellosen Nominalphrasen mit je zwei attributiven Adjektiven im Dativ Singular Maskulinum oder Neutrum, angereichert durch zahlreiche Annotationen. Die komplette Datensammlung kann auch heruntergeladen und für eigene Anschlussuntersuchungen verwendet werden.

In den in einem Open-Access-Format erscheinenden „Bausteinen“ soll der Leserschaft nach Möglichkeit der direkte Zugang zu allen zitierten Internetquellen ermöglicht werden. Alle Internetadressen sind mit einem Abrufdatum versehen und in den online erscheinenden Texten als Links aktiv. Sofern den Quellen ein Permalink oder ein DOI zugeordnet sind, bemühen wir uns, diese anzugeben. Zusätzlich werden alle zitierten Internetseiten in [Internet Archive](#)¹⁴ gespeichert. Dort sollten alle Versionen mit dem angegebenen Abrufdatum einsehbar sein.

14 Vgl. <https://web.archive.org/>, abgerufen am 27.7.2020.

9 Danksagung

Die erste Ausgabe der „Bausteine“ hätte ohne die tatkräftige und inhaltlich sehr anregende Hilfe vieler Personen nicht und vor allem nicht in der vorliegenden Form zustande kommen können. Der Dank der Autor/inn/en und Herausgebenden gilt den Mitgliedern des wissenschaftlichen Beirats der Reihe für die gesamte Unterstützung der Unternehmung und insbesondere für die Begutachtung der Bausteine dieser Ausgabe, Kristin Kopf und Ulrich H. Waßner für die aufwendige Redaktion der Beiträge, Thilo Weber für die Organisation der reibungslosen Zusammenarbeit mit Heidelberg University Publishing, der Leitung und dem gesamten Team von Heidelberg University Publishing für die Ermöglichung und die Umsetzung der Open-Access-Publikation, einschließlich des Satzes, des Designs und Aufbaus des Web-Auftritts sowie der Realisierung der Print-Ausgabe, Stefan Falke für die Programmierung der „Datenbank attributive Adjektive“ und unseren Hilfskräften Nagehan Cetin, Frederik Dumele, Oskar Jabs, Anna Konovalova, Dang Hoang Nguyen und Viktoria Arendt für die Unterstützung in allen Phasen der Untersuchungen und der Beitragserstellung.

ORCID®

Marek Konopka  <https://orcid.org/0000-0001-9627-0918>

Literatur

- Ammon, Ulrich. 2005. Standard und Variation. Norm, Autorität, Legitimation. In Ludwig M. Eichinger & Werner Kallmeyer (Hgg.), *Standardvariation. Wie viel Variation verträgt die deutsche Sprache?* (Jahrbuch des Instituts für Deutsche Sprache 2004), 28–40. Berlin & New York: De Gruyter
- Baayen, Harald R. 2009. Corpus linguistics in morphology: Morphological productivity. In Anke Lüdeling & Merja Kytö (Hgg.), *Corpus Linguistics: An International Handbook* (Handbücher zur Sprach- und Kommunikationswissenschaft 29), Bd. 2, 899–919. Berlin: De Gruyter.
- Biber, Douglas, Stig Johansson, Geoffrey Leech, Susan Conrad & Edward Finegan. 2006 [1999]: *Longman grammar of spoken and written English*. Harlow: Longman.
- Bubenhofer, Noah, Marek Konopka & Roman Schneider. 2014. *Präliminarien einer Korpusgrammatik*, unter Mitwirkung von Caren Brinckmann, Katrin Hein & Bruno Strecker (Korpuslinguistik und interdisziplinäre Perspektiven auf Sprache, CLIP 4). Tübingen: Narr.

- Carter, Ronald & Michael McCarthy. 2006. *Cambridge grammar of English*. Cambridge: Cambridge University Press.
- Conrad, Susan. 2010. What can a corpus tell us about grammar? In Anne O’Keeffe & Michael McCarthy (Hgg.), *The Routledge handbook of corpus linguistics* (Routledge Handbooks in Applied Linguistics), 227–240. London & New York: Routledge.
- Dudengrammatik 2016 = Wöllstein, Angelika & Dudenredaktion (Hgg.). 2016. *Dudenband 4 – Die Grammatik*, 9. Aufl. Berlin: Dudenverlag.
- Eichinger, Ludwig M. 2017. Standarddeutsch – die beste aller möglichen Sprachen. In Marek Konopka & Angelika Wöllstein (Hgg.), *Grammatische Variation. Empirische Zugänge und theoretische Modellierung* (Jahrbuch des Instituts für Deutsche Sprache 2016), 3–22. Berlin & Boston: De Gruyter.
- Eisenberg, Peter. 2013a. *Grundriss der deutschen Grammatik. Bd. 1: Das Wort*, 4. Aufl. Stuttgart & Weimar: Metzler.
- Eisenberg, Peter. 2013b. *Grundriss der deutschen Grammatik. Bd. 2: Der Satz*, 4. Aufl. Stuttgart & Weimar: Metzler.
- Engel, Ulrich. 2009. *Deutsche Grammatik*, Neubearbeitung, 2. Aufl. München: Iudicium.
- Evert, Stefan & Marco Baroni. 2007. zipfR: Wordfrequency distributions in R. In Sophia Ananiadou (Hg.), *Proceedings of the 45th Annual Meeting of the Association for Computational Linguistics Companion Volume Proceedings of the Demo and Poster Sessions*. 29–32. Prag: ACL. <https://www.aclweb.org/anthology/P07-2008.pdf> (abgerufen am 7.7.2020).
- Felder, Ekkehard. 2016. *Einführung in die Varietätenlinguistik* (Germanistik kompakt). Darmstadt: WBG.
- Finkel, Jenny Rose, Trond Grenager & Christopher Manning. 2005. Incorporating non-local information into information extraction systems by Gibbs sampling. In Kevin Knight, Hwee Tou Ng & Kemal Oflazer (Hgg.), *Proceedings of the 43rd Annual Meeting of the Association for Computational Linguistics (ACL’05)*, 363–370. Ann Arbor: Association for Computational Linguistics. <http://nlp.stanford.edu/~manning/papers/gibbscrf3.pdf> (abgerufen am 18.7.2020).
- Fuß, Eric, Marek Konopka & Angelika Wöllstein. 2017. Perspektiven auf syntaktische Variation. In Marek Konopka & Angelika Wöllstein (Hgg.), *Grammatische Variation. Empirische Zugänge und theoretische Modellierung* (Jahrbuch des Instituts für Deutsche Sprache 2016), 229–254. Berlin & Boston: De Gruyter.
- Fuß, Eric, Marek Konopka & Angelika Wöllstein (Hgg.). 2019. Grammatik im Korpus. Korpuslinguistisch-statistische Analysen morphosyntaktischer Variationsphänomene (Studien zur deutschen Sprache 80). Tübingen: Narr.

- Fuß, Eric, Marek Konopka, Beata Trawiński & Ulrich Hermann Waßner (Hgg.). 2018. *Grammar and corpora 2016*. Heidelberg: Heidelberg University Publishing. <https://heiu.uni-heidelberg.de/catalog/book/361> (abgerufen am 20.7.2020).
- GDE = Gunkel Lutz, Adriano Murelli, Susan Schlotthauer, Bernd Wiese & Gisela Zifonun (Hgg.). 2017. *Grammatik des Deutschen im europäischen Vergleich. Das Nominal* (Schriften des Instituts für Deutsche Sprache 14). 2 Bände. Berlin: De Gruyter.
- GDS = Zifonun, Gisela, Ludger Hoffmann & Bruno Strecker (Hgg.). 1997. *Grammatik der deutschen Sprache* (Schriften des Instituts für Deutsche Sprache 7). 3 Bände. Berlin & New York: De Gruyter.
- Habermann, Mechthild. 2010. Was ist eigentlich „Grammatik“? – Eine Einführung. In Mechthild Habermann (Hg.), *Grammatik wozu? Vom Nutzen des Grammatikwissens in Alltag und Schule* (Thema Deutsch 11), 9–14, Mannheim & Zürich: Dudenverlag.
- Hentschel, Elke & Harald Weydt. 2013. *Handbuch der deutschen Grammatik*, 4. Aufl. Berlin & Boston: De Gruyter.
- Hoffmann, Ludger. 1997. Einführung. In Gisela Zifonun, Ludger Hoffmann & Bruno Strecker (Hgg.), *Grammatik der deutschen Sprache* (Schriften des Instituts für Deutsche Sprache 7), 3 Bände. 1–16. Berlin & New York: De Gruyter.
- Institut für Deutsche Sprache (2017): *Deutsches Referenzkorpus / Archiv der Korpora geschriebener Gegenwartssprache 2017-II* (Release vom 1.10.2017). Mannheim: Institut für Deutsche Sprache. <https://www.ids-mannheim.de/DeReKo> (abgerufen am 20.7.2020).
- Konopka, Marek. 2018. Korpuslinguistik, Grammatiktheorie, Grammatikographie. In Angelika Wöllstein, Peter Gallmann, Mechthild Habermann & Manfred Krifka (Hgg.), *Grammatiktheorie und Empirie in der germanistischen Linguistik* (Germanistische Sprachwissenschaft um 2020 1), 151–184. Berlin & Boston: De Gruyter.
- Konopka, Marek, Jacqueline Kubczak, Christian Mair, František Štícha & Ulrich Hermann Waßner (Hgg.). 2011. *Grammatik und Korpora 2009. Dritte internationale Konferenz. Mannheim, 22.–24.09.2009* (Korpuslinguistik und interdisziplinäre Perspektiven auf Sprache, CLIP 1). Tübingen: Narr.
- Kupietz, Marc, Harald Lünen, Paweł Kamocki & Andreas Witt. 2018. The German Reference Corpus DeReKo: New Developments – New Opportunities. In Nicoletta Calzolari, Khalid Choukri, Christopher Cieri, Thierry Declerck, Sara Goggi, Koiti Hasida, Hitoshi Isahara, Bente Maegaard, Joseph Mariani, Hélène Mazo, Asuncion Moreno, Jan Odijk, Stelios Piperidis & Takenobu Tokunaga (Hgg.), *Proceedings of the Eleventh International Conference on Language Resources and Evaluation (LREC 2018)*, 4353–4360. Miyazaki:

- European Language Resources Association (ELRA). <https://www.aclweb.org/anthology/L18-1689.pdf> (abgerufen am 20.7.2020).
- Labov, William. 1966. The linguistic variable as a structural unit. *Washington Linguistic Review* 3, 4–22.
- Meurers, W. Detmar & Stefan Müller. 2009. Corpora and syntax. In Anke Lüdeling & Merja Kytö (Hgg.), *Corpus Linguistics: An International Handbook* (Handbücher zur Sprach- und Kommunikationswissenschaft 29), Bd. 2, 920–933, Berlin: De Gruyter.
- Mukherjee, Joybrato. 2006. Corpus linguistics and English reference grammars. In Antoinette Renouf & Andrew Kehoe (Hgg.), *The changing face of corpus linguistics. Papers from the 24th International Conference on English Language Research on Computerized Corpora (ICAME 24)*, 337–354. Amsterdam: Rodopi.
- Müller, Thomas, Helmut Schmid & Hinrich Schütze. 2013. Higher-order CRFs for morphological tagging. In David Yarowsky, Timothy Baldwin, Anna Korhonen, Karen Livescu & Steven Bethard (Hgg.), *Proceedings of the 2013 Conference on Empirical Methods in Natural Language Processing*, 322–332. Seattle: Association for Computational Linguistics.
- Paranhos Zitterbart, Jussara. 2010. Grammatiken des Deutschen im Vergleich. In Mechthild Habermann (Hg.), *Grammatik wozu? Vom Nutzen des Grammatikwissens in Alltag und Schule* (Thema Deutsch 11), 66–81, Mannheim & Zürich: Dudenverlag.
- R Core Team. 2020. *R: A language and environment for statistical computing*, Wien: R Foundation for Statistical Computing. <http://www.R-project.org/> (abgerufen am 26.6.2020).
- Schäfer, Roland. 2015. Processing and Querying Large Web Corpora with the COW₁₄ Architecture. In Piotr Bański, Hanno Biber, Evelyn Breiteneder, Marc Kupietz, Harald Lungen, & Andreas Witt (Hgg.), *Proceedings of the 3rd Workshop on Challenges in the Management of Large Corpora (CMLC-3)*, 28–34. Mannheim: Institut für Deutsche Sprache.
- Schäfer, Roland & Felix Bildhauer. 2012. Building Large Corpora from the Web Using a New Efficient Tool Chain. In Nicoletta Calzolari, Khalid Choukri, Thierry Declerck, Mehmet Uğur Doğan, Bente Maegaard, Joseph Mariani, Asuncion Moreno, Jan Odijk & Stelios Piperidis (Hgg.), *Proceedings of the Eighth International Conference on Language Resources and Evaluation (LREC'12)*, 486–493. Istanbul: European Language Resources Association (ELRA).
- Schmid Helmut. 1994. Probabilistic part-of-speech tagging using decision trees. *Proceedings of International Conference on New Methods in Language Processing*. Manchester: o. V. <https://www.cis.uni-muenchen.de/~schmid/tools/TreeTagger/data/tree-tagger1.pdf> (abgerufen am 11.7.2020).
- Schmid, Helmut, Arne Fitschen & Ulrich Heid. 2004. SMOR: A German Computational Morphology Covering Derivation, Composition, and Inflection. In

- Maria Teresa Lino, Maria Francisca Xavier, Fátima Ferreira, Rute Costa & Raquel Silva (Hgg.), *Proceedings of the Fourth International Conference on Language Resources and Evaluation (LREC'04)*, 1263–1266. Lissabon: European Language Resources Association (ELRA). <https://www.cis.uni-muenchen.de/~schmid/papers/smor.pdf> (abgerufen am 18.6.2020).
- von der Gabelentz, Georg. 1995 [1901], Die Sprachwissenschaft, ihre Aufgaben, Methoden und bisherigen Ergebnisse, 2. Aufl. In: Christopher Hutton (Hg.): *History of linguistics. 18th and 19th Century German linguistics*, Bd. 7. London: Routledge/Thoemmes Press.
- Zifonun, Gisela. 1997a. Kommunikative Minimaleinheit und Satz. In Gisela Zifonun, Ludger Hoffmann & Bruno Strecker (Hgg.), *Grammatik der deutschen Sprache* (Schriften des Instituts für Deutsche Sprache 7), 3 Bde., 85–97. Berlin & New York: De Gruyter.
- Zifonun, Gisela. 1997b. Der Modus kommunikativer Minimaleinheiten. In Gisela Zifonun, Ludger Hoffmann & Bruno Strecker (Hgg.), *Grammatik der deutschen Sprache* (Schriften des Instituts für Deutsche Sprache 7), 3 Bde., 605–675. Berlin & New York: De Gruyter.

Marek Konopka 

Grundlegende Aspekte der Wortbildung

Abstract This chapter begins with a sketch of the specifics of our approach, an overview of the contents of the chapters on word formation and some methodological notes. It then discusses the general characteristics of word formations and of their overall inventory, comparing word formations to primary words. Furthermore, the chapter explores the relative frequencies of word formations in different vocabulary areas and traces the word formation profiles of individual parts of speech. Finally, it compiles the characteristic word formation rules for different parts of speech.

Keywords word formation, simple word, complex word, productivity, part of speech, word formation types

1 Fragestellung und Methodisches

Die deskriptiv-grammatische Perspektive auf die Wortbildung¹ verlangt es, dass wir uns auf die Darstellung von Regularitäten konzentrieren und nicht auf die Behandlung von Idiosynkrasien. Dazu gehört es, dass ins Zentrum der Beschreibung Klassen von Wörtern gerückt werden, die auf relativ hohen Abstraktionsebenen definiert sind, und nicht etwa Bildungen auf der Basis eines bestimmten Lexems oder einiger weniger Lexeme eines bestimmten Wortfeldes. In diesem Sinne nehmen wir eine regelorientierte und nicht eine lexemorientierte Perspektive ein. Dennoch bleiben bei Kategorisierungen – etwa in primäre Wörter und Wortbildungen – lexembasierte Entscheidungen nicht aus. Die beiden Perspektiven sind damit nicht strikt getrennt. Allerdings wird die lexemorientierte Sicht nur so weit verfolgt, wie es für die regelorientierte Beschreibung notwendig ist.

Bei unserer Beschreibung von Wortbildungen können verschiedene Abstraktionsebenen unterschieden werden. Auf der obersten Ebene werden

1 Zum Wesen der Wortbildung als Gegenstand der Grammatik vgl. z.B. Fleischer & Barz (2012: 1–3), Eisenberg (2013: 201–208), Dudengrammatik (2016: 645–649).

Wortbildungen² nicht weiter analysiert, sondern als Gesamtgruppe den übrigen (nicht „wortgebildeten“) Wörtern gegenübergestellt. Die weiteren wichtigen Beschreibungsebenen sind – mit absteigendem Abstraktionsgrad:

- die Ebene der **Wortbildungsarten**, auf der wortartenübergreifende Bauweisen unterschieden werden können wie Komposition, Derivation oder Konversion (vgl. z. B. Fleischer & Barz 2012: 83–90),
- die Ebene der **Wortbildungsregeln**, auf der die Wortbildungsprodukte und ihre Bausteine kategorial spezifiziert werden, oft in Bezug auf Wortart oder Affixkategorie wie in den schematischen Darstellungen SBST + SBST → SBST³ (z. B. *Umsatz|steuer*) oder ADJ + SUFF → SBST (z. B. *Aufmerksam|keit*),
- die Ebene der **Wortbildungsmuster**, auf der die Wortbildungsprodukte teils kategorial, teils formal spezifiziert bzw. als individuelle Ausdrücke angegeben werden wie in der Formulierung „Adverbbildungen auf -weise“ (*behelfs|weise, haufen|weise* etc.) oder in der entsprechenden schematischen Darstellung SBST + *weise* → ADV.

Unsere empirische Exploration von Wortbildungsphänomenen orientiert sich vor allem an folgenden Fragen:

1. Welche grundlegenden Regularitäten sind in der Wortbildung des Gegenwartsdeutschen zu beobachten?
2. Wie verteilen sie sich auf verschiedene Klassen von Wörtern?
3. Welche Wortbildungsarten und Wortbildungsregeln sind in den behandelten Klassen von Wörtern besonders frequent?
4. Welche Wortbildungsmuster sind besonders produktiv und wie ist dies zu erklären?
5. Welche sprachlichen und außersprachlichen Variablen beeinflussen die Realisierung produktiver Wortbildungsmuster?
6. Welche Regeln stehen hinter den relevanten Wortbildungsmustern, welchen Wortbildungsarten entsprechen die Regeln?

Die Beschreibung betrifft dabei ausdrücklich nur das Gegenwartsdeutsche. Jedoch macht die Erklärung von Regularitäten im Aufbau von Ausdrücken, die als Wortbildungen kategorisiert werden, gelegentlich auch kurze Exkurse in die Sprachgeschichte notwendig.

2 *Wortbildung* wird hier also nicht nur für die verschiedenen Aspekte des Wortaufbaus und für die linguistische Disziplin gebraucht, sondern auch für ein Wort, das infolge von Wortbildungsprozessen entstanden ist. Letzteres wird zuweilen verdeutlichend als *Wortbildungsprodukt* bezeichnet.

3 SBST = Substantiv, ADJ = Adjektiv, ADV = Adverb, SUFF = Suffix.

Eine ausführliche Behandlung aller Bereiche der deutschen Wortbildung ist in dieser Grammatik nicht möglich.⁴ Wir beschränken uns auf einen empirisch fundierten Gesamtüberblick in den Kapiteln „Grundlegende Aspekte der Wortbildung“ und „Wortbildungsarten“ und widmen uns im Anschluss vertiefenden und auf besondere Detailschärfe abzielenden Beschreibungen ausgewählter Spezialbereiche.

Im vorliegenden Kapitel werden die allgemeinen Charakteristika der einschlägigen Wortbildungen und ihres Gesamtinventars vorgestellt, die relativen Häufigkeiten von Wortbildungen in verschiedenen Wortschatzbereichen erörtert, die Wortbildungsprofile einzelner Wortarten nachgezeichnet und die für einzelne Wortarten charakteristischen Wortbildungsregeln genannt (Fragen 1 bis 3 oben). Das darauffolgende Kapitel „Wortbildungsarten“⁵ liefert eine formal-kompositionale Beschreibung der Wortbildungsarten, ergänzt um die Ermittlung wortartenübergreifender Regularitäten und um die Erörterung von Prinzipien der semantisch-relationalen Interpretation komplexer Wörter (Fragen 3 bis 4). Die weiteren Kapitel behandeln einerseits besonders Variationsträchtiges, andererseits wenig Untersuchtes und schwer Klassifizierbares (Fragen 4 bis 6). So wird im Kapitel „Substantivkomposition und Fugenelemente“ ein Phänomenbereich aufbereitet, der besonders stark durch ein Nebeneinander verschiedener formaler Möglichkeiten gekennzeichnet ist. In den Kapiteln „Bau von und Umbau zu Adverbien“ und „Desubstantivische Zweitglieder in der Wortbildung des Adverbs“ werden wiederum Wortbildungsregeln und -muster behandelt, die sich mit den Wortbildungsarten wie Komposition, Derivation und Konversion in ihrem klassischen Verständnis nur schwer in Einklang bringen lassen. Im Kapitel „Partikel- und Präfixverben und ihre Komplemente“ schließlich werden Wortbildungsphänomene beschrieben, deren Auswirkungen weit in die Syntax hineinreichen.

In methodischer Hinsicht haben die Ausführungen ihren Ausgangspunkt in Betrachtungen quantitativer Verhältnisse zwischen Klassen unterschiedlich gebildeter Wörter. Besonders wenn sich diese Betrachtungen auf das gesamte Untersuchungskorpus (ca. 7 Mrd. Tokens, davon knapp 6 Mrd. Worttokens) beziehen, ist mit einer gewissen Unschärfe zu rechnen. Mit einer derart großen Datenmenge kann man nämlich nur dann effektiv arbeiten, wenn Ergebnisse automatischer Annotationen genutzt werden. Die Güte solcher Annotationen lässt sich in einem großen Korpus allerdings nur stichprobenartig überprüfen. Korrekturen können

4 Für umfassende Darstellungen des Gegenwartsdeutschen vgl. u. a. Deutsche Wortbildung (1973–1992, Bd. 1–5), Motsch (1999), Eichinger (2000), Donalies (2002), Fleischer & Barz (2012), nicht unerwähnt bleiben soll hier auch die online verfügbare, lexikographische Quelle *elexiko* (<https://www.owid.de/docs/elex/start.jsp>, abgerufen am 18.6.2020, vgl. Haß & Klosa 2015). Als wichtige ältere Werke sind vor allem Grimm (1826), Wilmanns (1899), Paul (1920) und Henzen (1965) zu nennen.

5 In dieser Ausgabe der „Bausteine einer Korpusgrammatik des Deutschen“ nicht mehr enthalten.

oft nur punktuell vorgenommen werden, sodass meist ein Recall- und evtl. auch ein Präzisionsverlust in Kauf genommen werden muss. Das Gesamtkorpus wird daher vor allem zur Exploration der Phänomenvielfalt und für globale Hypothesen zur Reichweite und Produktivität von Wortbildungsregeln genutzt. Bei spezielleren Fragestellungen werden öfter Teilkorpora herangezogen, und zwar

- das Korpus *40k*, in dem 40.000 zufällig ausgewählte Sätze aus dem Gesamtkorpus zusammengestellt wurden (750.643 Tokens) und
- die zufällig zusammengestellten Achtel des Gesamtkorpus (Korpora *8th-1* bis *8th-8*, je über 877 Mio. Tokens).

Manche komplexen Korpusrecherchen sind erst durch die Benutzung dieser Teilkorpora in einer vertretbaren Zeit durchführbar. Ihre Verwendung macht es möglich, die Qualität der Recherchen zu steigern, da die Ergebnisse automatischer Suchen auf Umfänge reduziert werden, bei denen sie genauer überprüft und im Bedarfsfall korrigiert bzw. durch Nachrecherchen ergänzt werden können. Im Folgenden machen wir im Allgemeinen Gebrauch von der TreeTagger⁶-basierten Lemmatisierung, einer Wortbildungsannotation, die auf den Wortbildungsanalysen von Canoo⁷ aufbaut, und den Wortarten-Annotationen von TreeTagger und Canoo. Da sich unsere Darstellung der substantivischen Wortbildungen auf die den Kern dieses Bereichs bildenden Appellative beschränkt, wird außerdem bei den Recherchen grundsätzlich mithilfe von Stanford NER (Finkel et al. 2005) eine Einschränkung auf Nicht-Eigennamen vorgenommen.

Im Teilkorpus *40k*, das im vorliegenden Kapitel eine besondere Rolle spielt, finden sich der Canoo-Klassifizierung zufolge 623.130 Wörter. Nach Ausschluss von Eigennamen, von Wörtern, deren Wortart nicht erkannt wurde oder die falsch bzw. gar nicht lemmatisiert wurden, sowie nach manuellen Korrekturen⁸ werden 558.374 Tokens berücksichtigt, die sich auf 48.901 TreeTagger-Lemmata verteilen. Bei diesen Daten wurde die Wortbildungsannotation an neuralgischen Stellen überprüft. Vor allem wurden alle Lemmata in Augenschein genommen, die Canoo nicht als Wortbildungen analysierte. Handelte es sich bei einem Lemma doch zweifelsfrei um eine synchron analysierbare Wortbildung, wurde es manuell als solche annotiert.⁹

6 Vgl. Schmid (1995) – unterstützend wurden ein Lexikon (<https://github.com/rsling/cow/blob/master/src/de/data/lexicon>, abgerufen am 7.6.2020) aus dem COW-Webkorpusprojekt (Schäfer & Bildhauer 2012, Schäfer 2015) und ein mithilfe von SMOR (Schmid et al. 2004) generiertes Ad-hoc-Lexikon verwendet.

7 Siehe <http://canoonet.eu/services/WordformationRules/ueberblick/index.html?lang=de> (abgerufen am 6.12.2019).

8 Manuell wurden weitere Eigennamen ausgeschlossen und Kardinalzahlwörter annotiert, die in der Wortartenklassifizierung von Canoo nicht erfasst wurden.

9 Solche Lemmata waren z.B. *Orkanbö*, *abrechenbar*, *gestreng*, *beprobten*, *derweil*, *sowas*.

2 Der Beschreibungsgegenstand in Theorie und Praxis

Bei der Bestimmung des Gegenstands der linguistischen Disziplin Wortbildung ist eine Abgrenzung der zu fokussierenden Bildungen von sonstigen Wörtern gefragt. Hilfreich ist hier eine Unterscheidung von primären und sekundären Wörtern (vgl. Dudengrammatik 2016: 645, Meibauer et al. 2007: 33–34). Erstere sind nicht auf andere sprachliche Einheiten zurückzuführen, letztere hingegen bauen auf anderen sprachlichen Einheiten auf und sind damit Wortbildungen. Diese Unterscheidung ist in der Praxis manchmal schwer anzuwenden, weil ältere Wortbildungen nicht immer als solche erkannt werden. Das Wort *heute* etwa wird von Fachleuten auf die althochdeutsche Phrase *hiu tagu* ‚an diesem Tag‘ zurückgeführt (vgl. z.B. DWB: Lemma „heute“). Für „normale“ Sprachbenutzer/innen ist die Wortbildungs­vergangenheit von *heute* aber nicht mehr zu erkennen. Verdunkelte Wortbildungen sind auch *Adler*, *Welt*, *erst*, *nein*¹⁰ und viele andere Ausdrücke, die auf den ersten Blick wie primäre Wörter aussehen. Überhaupt schrumpft die Gruppe der Kandidaten für den Status eines primären Wortes immer weiter, je genauer man den Ursprung der Wörter erforscht.

Solche alten, kaum oder gar nicht mehr als analysierbar erkennbaren Bildungen stehen jedoch nicht im Fokus von Beschreibungen der gegenwarts­sprachlichen Wortbildung. In diesen wird vielmehr eine „synchrone Erklärung des gegenwärtigen Wortbildungssystems“ (Fleischer & Barz 2012: 4) angestrebt. Das kann man so verstehen, dass in den zu behandelnden Ausdrücken Bausteine erkennbar sein sollen, die heute (das meint *synchron* in diesem Fall) interpretierbar sind und gleichzeitig nach heute nachvollziehbaren Regeln verbaut werden. Mit einer derartigen Vorgabe neigt man dazu, diejenigen Wortbildungen behandeln zu wollen, die von unseren Zeitgenoss/inn/en als solche empfunden werden. Das kann aber zu keiner klaren Konturierung des Beschreibungs­gegenstands führen. Das Wortbildungswissen und die wortbildungsanalytischen Fähigkeiten Einzelner sind nämlich sehr unterschiedlich, und gleichzeitig ist es unklar, welche Sprachbenutzer/innen die Referenzgruppe bilden sollen: Manche erkennen schon in *heute* eine Wortbildung, einige mehr, jedoch bei Weitem nicht alle, werden in *Fall*, *Wurf*, *statt* und *während*¹¹ eine Wortbildung sehen, ein breiter Konsens, dass es sich um Wortbildungen handelt, wird sich jedoch womöglich nur in Bezug auf eine zu sehr eingeschränkte Gruppe besonders transparenter Bildungen einstellen, die aus mehreren Konstituenten bestehen, welche entweder Wörtern gleichen (z.B. *Haus|tür*, *Schnell|zug*, *honig|gelb*, *klein|kariert*,

10 Vgl. Nübling et al. (2017: 185) und entsprechende Einträge in Pfeifer (1993).

11 Entstanden wohl durch Umdeutung syntaktischer Verhältnisse: *während der Zeit* < (in) *währender Zeit* (vgl. Lemma „während“ in Pfeifer 1993 und DWB).

nach|dem, so|mit) oder häufige Wortbildungsmorpheme darstellen (z. B. *Zeit|ung*, *ver|arbeiten*, *Ge|red|e*).

Indessen könnte postuliert werden, einschlägig seien bei der Zielsetzung „synchrone Erklärbarkeit“ diejenigen Wortbildungen, die sich sinnvoll und vollständig in solche sprachlichen Einheiten (Wörter, Morpheme, Fugenelemente u. ä.) analysieren lassen, die heute regulär verwendet werden. Eine korpusorientierte Ideallösung würde dann nicht auf das individuell variierende Sprachempfinden setzen, sondern versuchen, die Korpusdaten sprechen zu lassen. Ein Kriterium für eine sichere Klassifizierung einer Wortbildung als einschlägig wäre etwa der Befund, dass sich jeder ihrer Bausteine in der gleichen Bedeutung bzw. Funktion frei oder innerhalb von anderen Lexemen im Korpus wiederfindet. Einen entsprechenden intelligenten Abgleich von Korpuswörtern automatisch vorzubereiten und durchzuführen erscheint für unsere Zwecke aber als zu aufwendig, und „manuell“ ist er für größere Datenmengen nicht durchführbar.

Übrigens wären bei dieser Lösung Wortbildungen mit unikalen lexikalischen Morphemen ausgeschlossen worden. Betroffen hätte dies bekannte Fälle wie *Brom|beere*, *Lind|wurm* oder *Un|flat*, aber auch viele weniger bekannte Fälle wie *Mein|eid*, *gewieft*, *reüss|ieren*. Unikale Morpheme treten in genau einer Verbindung auf¹² und sind in diesem Sinne unproduktiv. Da ihre Bedeutung verblasst ist, können alle oben genannten Fälle als synchron nicht vollständig erklärbar betrachtet werden. Aber eigentlich sollte die Bestimmung der einschlägigen Wortbildungen im Sinne unserer regelorientierten Perspektive dazu dienen, dasjenige sprachliche Material zu sichern, in dem heute wirksame Regularitäten bzw. erkennbare Wortbildungsregeln zu finden sind, und die genannten Wörter mit unikalen Morphemen scheinen gerade nach solchen Regeln aufgebaut zu sein. Eine Lösung, die sie daher doch mit einbezieht, würde allerdings nicht mehr auf eine synchrone Erklärbarkeit der Bildungen abzielen, sondern nur auf ihre synchrone Analysierbarkeit. Bei dieser neuen Lösung wäre es für die Einschlägigkeit einer mehrgliedrigen Wortbildung unter Umständen ausreichend, wenn nur **eine** ihrer unmittelbaren Konstituenten heute regulär¹³ vorkäme. Allerdings lässt sich auch diese Lösung korpuslinguistisch nicht direkt umsetzen.

Vor dem Hintergrund solcher theoretischen Optionen muss nach einem pragmatischen Ausweg gesucht werden. Dabei wollen wir die zuletzt genannte Lösung nicht gänzlich aus den Augen verlieren: Wir möchten die Regularitäten

12 Evtl. auch in den auf dieser Verbindung aufbauenden Wortbildungen (z. B. *Brom-beer|saft*).

13 Also nicht nur im vorliegenden Wort und in den darauf aufbauenden Wortbildungen. So erscheint das Morphem {beere} nicht nur in Verbindung mit dem unikalen Morphem {brom-} in *Brombeere* oder *Brombeersaft*, sondern auch frei in *Beere* sowie in *Blaubeere*, *Stachelbeere* etc. Entsprechend ist *Brombeere* eine einschlägige Wortbildung.

in den heute synchron analysierbaren Wortbildungen beschreiben. In der Praxis greifen wir dann, um die großen Datenmengen zu bewältigen, auf die Canoo-basierte Annotation zurück. Diese verwendet vorzugsweise gelistete Wortbildungsanalysen und leitet ggf. nicht gelistete Bildungen mithilfe von Regeln ab. Die automatische Annotation wird, wann immer möglich und notwendig, durch manuelle Korrekturen präzisiert bzw. – sofern Recall-Prüfungen dies nahelegen – durch zusätzliche, etwa formorientierte Recherchen ergänzt. Angemerkt werden muss dabei, dass die manuelle Analyse bzw. Annotation in einigen Fällen Interpretationsspielraum zulässt (etwa bezüglich der Zugehörigkeit verschiedener Formen zu ein und demselben Morphem). Dennoch glauben wir uns insgesamt der auf synchrone Analysierbarkeit der Bildungen ausgerichteten Ideal-lösung anzunähern. Gemäß dieser sind von den in diesem Abschnitt genannten Beispielen die Wörter *Fall*, *Wurf*, *statt*, *während*, *Haustür*, *Schnellzug*, *honiggelb*, *kleinkariert*, *nachdem*, *somit*, *Zeitung*, *verarbeiten*, *Gerede*, *Brombeere*, *Lindwurm*, *Unflat*, *drollig*, *Meineid*, *gewieft*, *reüssieren* als einschlägige Wortbildungen einzu-stufen, *heute*, *Adler*, *Welt*, *erst*, *nein* dagegen nicht.

3 Primäre Wörter und Wortbildungen – Simplizia und komplexe Wörter

Die einschlägigen Wortbildungen im Sinne von Abschnitt 2 werden im Weiteren der Kürze wegen schlicht als „Wortbildungen“ bezeichnet. Ähnlich wird auf die Komplementärgruppe mit „primäre Wörter“ Bezug genommen, auch wenn eigentlich „tendenziell primäre Wörter“ passender gewesen wäre, da sich in dieser Gruppe neben echten primären Wörtern auch verdunkelte Wortbildungen finden. Einige dem Korpus entnommenen Vertreter beider Gruppen enthält Tab. 1. Sie ist die Hauptquelle für die Beispiele, die weiter unten unsere Ausführungen veranschaulichen.

Tabelle 1: Primäre Wörter vs. Wortbildungen: Beispiele.

Primäre Wörter		Wortbildungen	
<i>Tor</i>	<i>Vater</i>	<i>Torfrau</i>	<i>Schwiegermutter</i>
<i>Frau</i>	<i>Kabarett</i>	<i>Landesgrenze</i>	<i>Stiefvater</i>
<i>Land</i>	<i>heute</i>	<i>gelblich</i>	<i>(das) Heute</i>
<i>Grenze</i>	<i>laufen</i>	<i>Volkstum</i>	<i>(der) Lauf</i>
<i>gelb</i>	<i>werfen</i>	<i>vernehmen</i>	<i>(der) Wurf</i>
<i>Volk</i>	<i>leben</i>	<i>Gerede</i>	<i>(das) Leben</i>
<i>nehmen</i>	<i>grün</i>	<i>Dialekt</i>	<i>Grüne(r)</i>
<i>reden</i>	<i>das</i>	<i>Telearbeit</i>	<i>demnach</i>
<i>Arbeit</i>	<i>nach</i>	<i>kabarettophil</i>	<i>Unflat</i>
<i>Mutter</i>	<i>[...]</i>	<i>identisch</i>	<i>[...]</i>

Konstitutive Bausteine von Wortbildungen sind kleinste bedeutungs- bzw. funktionstragende Formen, sog. Morpheme, die **wortfähige Morpheme**, **gebundene lexikalische Morpheme** und gebundene grammatische Morpheme der Kategorie **Wortbildungsaffixe** repräsentieren.¹⁴

I. Wortfähige Morpheme

Wortfähige Morpheme im engeren Sinn können freistehend als (primäre) **Wörter** realisiert werden – sie stellen somit sog. freie Morpheme dar. Die meisten freien Morpheme weisen eine selbständige Bedeutung auf – als Beispiele können {tor}, {frau}, {land}, {grenze}, {gelb}, {volk}, {arbeit}, {mutter}, {vater}, {kabarett}, {heute} oder {grün} dienen. Sie werden als lexikalische Morpheme¹⁵ angesehen (vgl. z. B. Brandt et al. 2006: 136–137) und bilden sog. Inhaltswörter¹⁶ bzw. Autosemantika (vgl. Meibauer et al. 2007: 133), die mehrheitlich flektierbar sind. Wortfähig im engeren Sinn ist aber auch eine Reihe von Morphemen, die wie {das}¹⁷ und öfter auch {nach} als grammatisch bezeichnet werden, da ihre Bedeutung unselbständiger erscheint und sie ziemlich spezifische grammatische Funktionen ausüben. Freistehend werden solche Morpheme als sog. Funktionswörter¹⁸ bzw. Synsemantika eingestuft. Bereits am Übergang zu gebundenen lexikalischen Morphemen sind schließlich Basismorpheme von Verben zu verorten. Sie können nämlich standardsprachlich nur im Imperativ Singular schwacher Verben und starker Verben ohne *e/i*-Wechsel frei – d. h. hier auch – ohne eine explizite Flexionsmarkierung realisiert werden (Dudengrammatik 2016: 441–446), vgl. *leb!* bzw. *lauf!* Linear nicht segmentierbare Verbformen¹⁹ wie *nahm*, *warf* oder *nimm!* *wirf!* hingegen können bereits als Kombinationen aus Basismorphem und Flexionsmarkierung mittels Vokalwechsel aufgefasst

14 Zur Spezifik der Kurzwörter, die sich nicht direkt in Morpheme zerlegen lassen, siehe weiter unten.

15 Lexikalische Morpheme sind nicht mit Lexemen zu verwechseln, zu denen hier alle Wörter im Sinne von Lexikoneinheiten gezählt werden, also auch Funktionswörter, vgl. Dudengrammatik (2016: 136–140).

16 Zu den Inhaltswörtern werden Substantive, Adjektive, die meisten Verben und Adverbien sowie manchmal auch Präpositionen gezählt.

17 Als Determinierer und Pronomen.

18 Zu den Funktionswörtern werden vor allem Artikel, Partikeln, subordinierende und koordinierende Konnektoren (zuzüglich Adverbkonnektoren, vgl. HDK 1), meist auch Präpositionen, Hilfs- und Modalverben, bestimmte Modalwörter und Pronomina gezählt, vgl. Willems (2013).

19 Der Vokalwechsel begründet kein nur als Flexionsmorphem interpretierbares Segment des Wortes.

werden.²⁰ Basismorpheme von Verben können dennoch als wortfähig in einem erweiterten Sinn betrachtet werden: Sie bilden (primäre) Wörter ohne Zuhilfenahme eines anderen Morphems der Gruppen I bis III.

II. Gebundene lexikalische Morpheme

Lexikalische Morpheme, die nur gebunden vorkommen, sind fast immer als **Konfixe** einzustufen.²¹ Sie sind typischerweise aus anderen europäischen Sprachen entlehnt und wurzeln besonders häufig im Griechischen, wie {dia}, {lekt}, {tele}, {phil}, oder Lateinischen, wie {ident}. Den Konfixen können außerdem heimische Morpheme wie {schwieger} oder {stief} zugeschlagen werden. Zu gebundenen lexikalischen Morphemen sind am Rande auch die in Abschnitt 2 bereits angesprochenen unikalen Morpheme wie {flat} in *Unflat* zu zählen.²²

Die Morpheme der Gruppen I und II können miteinander sowie mit gebundenen grammatischen Morphemen der Kategorie Wortbildungsaffixe kombiniert werden.

III. Wortbildungsaffixe

Wir sprechen hier einfach von **Affixen** (vgl. z.B. Fleischer & Barz 2012: 52–63) bzw. spezifischer von Präfixen wie in *ver|nehmen* oder *Un|flat*, Suffixen wie in *gelb|lich* oder *ident|isch* und Zirkumfixen wie in *Ge|red|e*. Diese Bezeichnungen verwenden wir in diesem Kapitel, um Verwechslungen zu vermeiden, nicht für Flexionsmorpheme (wie in *nimm|st*), die übrig gebliebene Kategorie der gebundenen grammatischen Morpheme.

Nur die wortfähigen Morpheme (Gruppe I) können ganz alleine in einer Wortbildung eingesetzt werden, vgl. (*das*) *Heute* oder (*der*) *Lauf* (zur sog. morphologischen Konversion vgl. z.B. Eisenberg 2013: 280). Baut eine Wortbildung demgegenüber auf zwei oder mehr Morphemen der Gruppen I bis III auf, betrachten wir sie als **komplexes Wort**. Jedes komplexe Wort muss mindestens ein Morphem enthalten, das entweder ein wortfähiges Morphem oder ein gebundenes lexikalisches Morphem repräsentiert.

20 Donalies (2002: 20) fasst dennoch die „Verbstämme des Typs *ess-*, *geb-* oder *les-*“ (folglich wohl auch *nehm-* oder *werf-*) als frei vorkommend auf.

21 Vgl. Donalies (2002: 21–23), Fleischer & Barz (2012: 63–64). Verschiedene Konfixtheorien werden in Donalies (2009) diskutiert.

22 Zum Unterschied zwischen unikalen Morphemen und Konfixen vgl. Meibauer et al. (2007: 31).

In ihrer Minimalform basieren komplexe Wörter auf zwei Morphemen. Die Zusammensetzungs- und Reihenfolgeoptionen des Morphempaars sind folgender Auflistung zu entnehmen:²³

1. zwei wortfähige Morpheme, z. B. *Tor|frau*
2. ein gebundenes lexikalisches Morphem (A) und ein wortfähiges Morphem (B):

- a. A + B, z. B. *Tele|arbeit*
- b. B + A, z. B. *kabarett(o)|phil*²⁴

3. zwei gebundene lexikalische Morpheme, z. B. *Dia|lekt*
4. ein wortfähiges Morphem (A) und ein Affix (B):²⁵

- a. A + B, z. B. *gelb|lich* (Suffigierung)
- b. B + A, z. B. *ver|nehm-* (Präfigierung)

5. ein gebundenes lexikalisches Morphem (A) und ein Affix (B):

- a. A + B, z. B. *ident|isch* (Suffigierung)
- b. B + A, z. B. *Un|flat* (Präfigierung)

Es werden alle Reihenfolgemöglichkeiten genutzt, auch wenn die Optionen 2b und 5b weit seltener erscheinen als die anderen.²⁶ Die Verbindungstypen 1 bis 3 werden der Komposition zugerechnet, 4 und 5 der Derivation. Komposition und Derivation werden zusammen mit Konversion in weiteren Abschnitten differenzierter betrachtet und im Kapitel „Wortbildungsarten“ genauer analysiert.

Wortbildungsprozesse sind im Allgemeinen rekursiv. Das heißt hier, dass sich Verbindungen von Morphemen der Gruppen I bis III zum einem mit weiteren Morphemen der Gruppen I bis III, zum anderen mit weiteren Verbindungen von

23 Für eine Übersicht, die sich nicht nur auf komplexe Wörter in ihrer Minimalform beschränkt und auch Fugenelemente zur Unterscheidung von Kombinationsmöglichkeiten verwendet, vgl. Donalies (2002: 49).

24 Eingeklammert ist das Fugenelement, das keinen Morphemstatus besitzt – dazu weiter unten.

25 Weggelassen werden hier der Einfachheit halber Verbindungen mit Zirkumfixen, die als Kombinationen aus Präfigierung und Suffigierung betrachtet werden können.

26 Wird ein wortfähiges Morphem vor ein gebundenes lexikalisches Morphem gesetzt (2b), muss das vorangehende Wort häufig erst durch ein Fugenelement (vgl. -o- in *kabarett(o)|phil*) anschlussfähig für das darauf folgenden Konfix gemacht werden (wobei die Zusammensetzung oft immer noch ungewöhnlich klingt); die Verbindung eines Präfixes mit einem gebundenen lexikalischen Morphem (5b) scheint sich wiederum auf unikale Morpheme wie {flat} zu beschränken.

Morphemen der Gruppen I bis III zu noch komplexeren Bildungen vereinen können. Etwas pauschaler ausgedrückt: Komplexe Wörter können sich zusammen mit anderen Wörtern sowie mit Konfixen und Affixen zu noch komplexeren Wörtern verbinden.²⁷ Dabei können die Verbindungen im Deutschen bekanntlich eine sehr hohe Komplexität erreichen wie in den Korpusbeispielen *Gemeinnützigkeit(s) entbürokratisierung(s)gesetz* oder *Bund(es)verfassung(s)gericht(s)entscheidung(s)sammlung*. Wie solche Bildungen zu analysieren und hinsichtlich ihres Komplexitätsgrades zu bewerten sind, wird später im Kapitel „Wortbildungsarten“ diskutiert.

Neben den Morphemen der Gruppen I bis III können komplexe Wörter auch Fugenelemente und Flexionsmarkierungen enthalten. **Fugenelemente**²⁸ wurden in den gerade genannten Beispielen eingeklammert (vgl. dementsprechend auch *Land(es)|grenze*, *Volk(s)|tum*, *kabarett(o)|phil*), da ihnen der Morphemstatus meist abgesprochen wird (dazu Kopf 2018: 8–9). Sie sorgen nicht für eine Komplexitätserhöhung, sondern sind lediglich deren mögliche Begleiterscheinung.²⁹ Die heimischen Fugenelemente wie *-s-* oder *-es-* haben ihren Ursprung in Flexionsendungen, werden aber aus synchroner Sicht in der Regel nicht als deren Realisierungen eingestuft (vgl. z. B. Nübling & Szczepaniak 2009: 197).³⁰

Flektierbare Wörter gehen in die Wortbildungen meist als unflektierte Stämme ein. **Flexionsmarkierungen** sind für Wortbildungen auch nicht konstitutiv. Das heißt vor allem, dass Wortformen, die als Verbindungen aus einem lexikalischen Morphem und einem Flexionsmorphem erscheinen (z. B. *Tor|es*, *Länd|er*, *nimm|st*, *leb|te*) oder durch Vokalwechsel markiert sind (z. B. *nahm*, *warf* oder *nimm! wirf!*), nicht als komplexe Wörter betrachtet werden. Entsprechend werden alle Konversionen aus primären Wörtern nicht als komplexe Wörter gewertet. Dies gilt für die schon erwähnten morphologischen Konversionen (vgl. Eisenberg 2013: 280–284) wie (*der*) *Lauf*, bei denen eine unflektierte Wortform (ein Stamm) direkt zur unflektierten Grundform einer anderen Wortart „konvertiert“. Das gilt aber auch für sog. implizite Ableitungen wie *Wurf* (vgl. Eisenberg 2013: 280)³¹ oder offenkundig syntaktisch entstandene Infinitivkonversionen

27 Vgl. Donalies (2002: 15) zu den Möglichkeiten der Wortschatzerweiterung.

28 Sie werden genauer in einem getrennten Kapitel „Substantivkomposition und Fugenelemente“ behandelt.

29 Dies ist bei Minimalpaaren wie *Vermögen|steuer* vs. *Vermögen(s)|steuer* unmittelbar sichtbar. Zur Funktion von Fugenelementen s. die Übersicht in Nübling & Szczepaniak (2009: 201–208).

30 Speziell zu Fugenelementen als Nicht-Kasus-Suffixen s. Gallmann (1999). Für eine differenzierte Sicht auf die Semantik der Fugenelemente vgl. Donalies (2002: 46–48) und Kopf (2018: 54–72).

31 Donalies (2002: 135) wertet solche Fälle als Konvertate aus spezifischen, u. U. veralteten Verbstämmen (z. B. dem mhd. Präteritumstamm *wurf*, vgl. auch Fleischer & Barz 2012: 89). Wir werden sie im Kapitel „Wortbildungsarten“ diskutieren.

wie (*das*) *Leben*, die einen (ehemals) flexionsrelevanten Vokalwechsel respektive die Infinitivendung konservieren. Substantivierte Adjektive wie (*die/der*) *Grüne* (etwa im Sinne einer Person, die durch ihre politische Orientierung charakterisiert wird) stellen hier nur insofern einen Sonderfall dar, als sie nicht lediglich ein bestimmtes Flexionsmerkmal des Ausgangswortes behalten, sondern große Teile des adjektivischen Flexionsparadigmas in den nominalen Gebrauch integrieren (zur syntaktischen Konversion vgl. z. B. Eisenberg 2013: 280). Produkte aller genannten Konversionsarten gehören in unserer Auffassung also nicht zu den komplexen Wörtern, sondern sind zusammen mit primären Wörtern zu den einfachen Wörtern (den Simplicia) zu zählen.³² Sie bleiben nichtsdestoweniger Wortbildungen, denn sie gehen auf andere Lexeme bzw. auf Wörter einer anderen Wortart zurück.

In die weiter oben beschriebene, aus formaler Sicht rein additive „Aufbau-mechanik“ komplexer Wörter fügen sich auch einige etwas weniger typische Bildungen ein. In diesen können an die Stelle eines Wortes, welches das Erstglied eines Kompositums bildet, zum einen nicht wortfähige Gruppen von Morphemem (*Kurzzeit|gedächtnis*), Wortgruppen (*Saure-Gurken-Zeit*) und sogar satzartige Konstrukte (*Trimm-Dich-Pfad*), zum anderen etwa einzelne Buchstaben bzw. ihre Bezeichnungen (*S-Kurve*) treten.³³ Einige dieser Ausdrucksformen können auch als Basis für Konversionen dienen, vgl. (*das*) *Stelldichein*, das auf einen Imperativsatz zurückgeführt werden kann (Fleischer & Barz 2012: 65, Henzen 1965: 83–84).

Es gibt jedoch auch Bildungen, die den additiven Regularitäten nicht direkt folgen. Zu denken ist hier vor allem an Kurzwörter und Kontaminationen (vgl. z. B. Fleischer & Barz 2012: 91, 93–94). Bei den eher randständigen Kontaminationen sind die Ausgangskomponenten meist immer noch leicht rekonstruierbar (vgl. *jein*, *Denglisch* oder – aus dem Englischen – *Brexit*). Insofern liegt die Klassifizierung der Kontaminationen als komplex nahe. Die viel häufigeren Kurzwörter (vgl. Fleischer & Barz 2012: 91, Balnat 2011: 287–289) sind ein komplizierterer

32 Konversionen bzw. sog. implizite Ableitungen aus primären Wörtern als einfache Wörter zu betrachten, hat eine lange Tradition. Schon nach Grimm hebt die „[i]nnere Wortbildung [...] die Einfachheit des Wortes nicht auf“ (Grimm 1826: 1).

33 Die Reihe kann mit weiteren, eher randständigen Erstgliedtypen fortgesetzt werden, z. B. Kombinationen aus Zahlwörtern und Präpositionen (geschrieben als Ziffern mit Trennzeichen) wie in *1:8-Debakel*. Beachtenswert ist der häufige Gebrauch des Bindestriches zwischen Erst- und Zweitglied des Kompositums bei Wortgruppen, Sätzen, Buchstaben etc. als Erstglied. Der Bindestrich wird bei Wortgruppen und Sätzen auch innerhalb des Erstglieds eingesetzt. Zu den Gründen für die Bindestrichsetzung in ähnlichen Fällen vgl. Kopf (2018: 119–120).

Fall.³⁴ Sie stellen zwar Reduktionen von fast ausschließlich komplexen Wörtern oder von Wortgruppen dar, die Ausgangskomponenten sind aber selten aus dem Kurzwort erschließbar und den Sprachbenutzer/inne/n in der Kommunikationssituation öfter nicht präsent (vgl. Dudengrammatik 2016: 680). Außerdem kann ein Kurzwort die Langform weitgehend aus dem Sprachgebrauch verdrängen und sich damit noch stärker einem primären Wort annähern (tendenziell bei *Labor* gegenüber *Laboratorium*). Dies zeigt einerseits, dass Kurzwörter tatsächlich eigenständige Wörter sind (auch wenn sie die Bedeutung der Langform und ihre grammatischen Merkmale behalten).³⁵ Es macht andererseits die Entscheidung schwer, sie pauschal den Simplizia oder den komplexen Wortbildungen zuzuschlagen (vgl. Dudengrammatik 2016: 748).

Wir fassen Kurzwörter als mit ihren Langformen teilweise synonyme Lexeme auf und verwenden zur Bewertung ihrer Komplexität wie bei anderen Bildungen das Kriterium der formalen Repräsentation der Ausgangskomponenten. So zählen wir unisegmentale Kurzwörter,³⁶ die auf ein Morphem oder einen Morphemteil der Langform gekürzt sind – wie *Labor[atorium]*, *Ober[kellner]*, *Zoo[logischer Garten]* bzw. *Alu[minium]*, *Krimi[nalroman/-film]* –, zusammen mit primären Wörtern zu den einfachen Wörtern bzw. den Simplizia. Alle anderen Kurzwörter betrachten wir als komplex. Hierzu gehören alle multisegmentalen Kurzwörter – unabhängig davon, ob sie als Silben- oder Buchstabenwörter (z. B. *Schiri*, *Kita* bzw. *GmbH*, *Kfz*, *GAU*, *TÜV*) erscheinen und ob sie gebunden (*Schiri*, *Kita*, *GAU*, *TÜV*) oder mit den Buchstabennamen (*GmbH*, *Kfz*) ausgesprochen werden. Da alle Kurzwörter aus unserer Sicht Wortbildungen sind, geht die Unterscheidung von primären Wörtern und Wortbildungen auch an dieser Stelle nicht mit der Unterscheidung von Simplizia und komplexen Wörtern einher.

34 Fleischer & Barz (2012: 277) sehen in der Kurzwortbildung eine Wortbildungsart. Eichinger (2000: 34) schreibt differenzierter: „In gewisser Weise haben [...] die Kurzwörter mehr vom Phraseologismus als von der eigentlichen Wortbildung – das Kurzwort ist eine Art Metapher der Ausdrucksform“.

35 Bei Wortgruppen richten sich Kurzwörter grammatisch meist nach dem Kopf der Langform. Zwischen Kurzwort und Langform kann es aber Bedeutungsverschiebungen und konnotative Unterschiede geben (vgl. Balnat 2011: 110–117). Allerdings ist deren Entstehung dem eigentlichen Wortbildungsprozess nachgeordnet. So kann *BAföG* anders als *Bundesausbildungsförderungsgesetz* auch die aufgrund des *Bundesausbildungsförderungsgesetzes* gezahlte Unterstützung bedeuten (<https://www.duden.de/rechtschreibung/BAfoeG>, abgerufen am 8.6.2020). Wie nah sich ein Kurzwort und seine Langform semantisch stehen, könnte korpuslinguistisch automatisiert geprüft werden, indem in einem Word-Embedding-Modell die Distanz der beiden Formen zueinander gemessen würde (vgl. Bubenhofer et al. 2019: 222–226).

36 Vgl. Fleischer & Barz (2012: 277–279) und – etwas differenzierter – Donalies (2002: 144–146).



Abbildung 1: Primäre Wörter und Wortbildungen vs. Simplizia und komplexe Wörter.

Abb. 1 veranschaulicht zusammenfassend die Unterschiede zwischen den beiden Begriffspaaren anhand ausgewählter Beispiele.

4 Allgemeine quantitative Verhältnisse und Wortschatzerweiterung

Die automatisch als Wortbildungen analysierten Wörter machen im gesamten Untersuchungskorpus 28,5% der über 5,8 Milliarden Tokens aus, die in der Canoo-basierten Annotation als Wörter behandelt werden. Der tatsächliche Anteil der Wortbildungen im Untersuchungskorpus muss höher liegen, da automatisch – insbesondere bei seltenen Wörtern – einige Wortbildungen nicht als solche analysiert werden. Um die quantitativen Verhältnisse zwischen Wortbildungen und primären Wörtern genauer zu untersuchen, wird daher das Teilkorpus *4ok* herangezogen, das manuell bearbeitet werden konnte und über eine halbe Million relevante Worttokens umfasst, die sich auf knapp 50.000 TreeTagger-Lemmata³⁷ verteilen (vgl. Abschnitt 1 oben). Der Anteil der Wortbildungen beträgt in diesem Teilkorpus bei den Tokens 31,8% (177.202). Bei den Lemmata liegt er mit 91,5% (44.746) viel höher. Der sehr deutliche Unterschied zwischen Tokens und Lemmata erklärt sich daraus, dass Wortbildungslemmata

37 Wir verwenden hier den Terminus *Lemma* anstatt des evtl. zu erwartenden *Typ* (engl. *type*), um klarzustellen, dass es sich um eine im Tagging zugewiesene Kategorie handelt, die einem Lexem entsprechen soll. *Typ* lässt ungünstigerweise die Ambiguität zwischen einer Wortform (der die Tokens im Korpus entsprechen) und einem Lexem (das etwa bei flektierenden Wortarten mehrere Wortformen umfasst) zu.

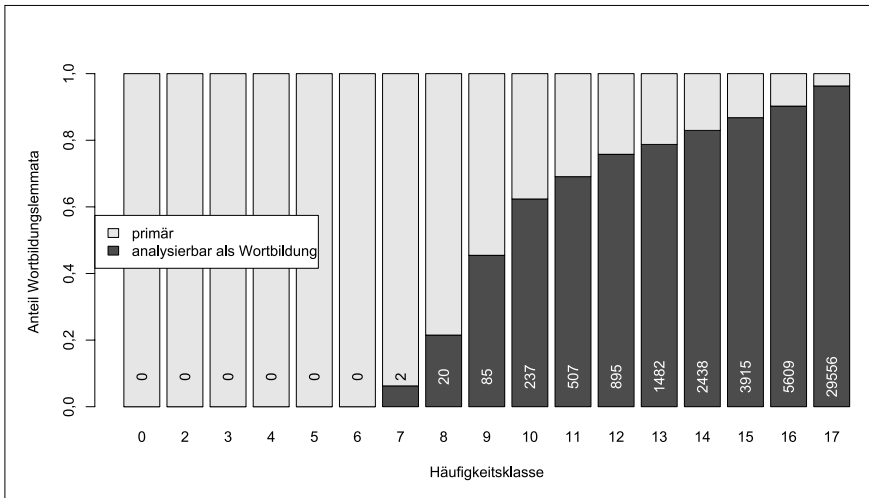


Abbildung 2: Anteil der Wortbildungslemmata nach Häufigkeitsklassen im Teilkorpus 40k: Je höher die Häufigkeitsklasse, desto tokenärmer die Lemmata, wobei: $hk \approx \lfloor \log_2 (f(\text{der, die, das}) / f(\text{Lemma})) + 0,5 \rfloor$. Die Zahlen in den Säulen geben die Anzahl der Wortbildungslemmata in der Häufigkeitsklasse an.

tendenziell tokenarm sind (66,0% davon sind Hapaxlegomena³⁸) und es unter den besonders tokenreichen Lemmata gar keine Wortbildungen gibt. Teilt man alle Lemmata in Häufigkeitsklassen ein, so lässt sich in den Klassen, die Wortbildungen enthalten, Folgendes beobachten: Je weniger Tokens die Lemmata einer Häufigkeitsklasse aufweisen, desto größer der Anteil der Wortbildungen (vgl. Abb. 2). Es ist davon auszugehen, dass sich das auch in größeren Korpora so wiederfindet.³⁹

In der Rangliste der 100 häufigsten Lemmata des Teilkorpus (vgl. Tab. 2 unten) findet sich das erste als Wortbildung analysierte Wort erst auf Rang 77. Das Fremdwort *Prozent* ist Canoo zufolge eine Zusammensetzung aus zwei „neoklassischen Formativen“.⁴⁰ Der erste Bestandteil der Bildung kann (allerdings anders als ein Konfix) auch frei auftreten wie in *pro Jahr* oder *pro Liter Benzin* (neben weiteren gebundenen Verwendungen wie in *prowestlich* oder *pro(-) amerikanisch*), der zweite Bestandteil lässt sich in *Zentimeter*, *Zentiliter* u. Ä.

38 Bei den primären Wörtern beträgt der Anteil der Hapaxlegomena 27,4%. Auch in Baayens (vgl. 2001: 135–138) Untersuchungen zu den Häufigkeiten niederländischer Wörter ist der Anteil der Hapaxlegomena bei den Ableitungen auf *-heid* viel höher als bei substantivischen Simplizia.

39 Vgl. auch dazu die prinzipiell in die gleiche Richtung zeigenden Ergebnisse von Baayen (2001: 135–138).

40 Vgl. <http://www.canoonet.eu/services/WordformationRules/Neoklassisch/Intro.html?MenuId=WordFormation112> (abgerufen am 3.11.2019).

Tabelle 2: Die Wortbildungen (hervorgehoben) unter den 100 häufigsten Lemmata im Teilkorpus 40k (HK=Häufigkeitsklasse).

Rang	Lemma	Frequenz	HK	Rang	Lemma	Frequenz	HK
1	<i>die</i>	68.695	0	34	<i>noch</i>	1.966	6
2	<i>in</i>	17.759	2	35	<i>vor</i>	1.916	6
3	<i>und</i>	14.508	3	36	<i>so</i>	1.863	6
4	<i>eine</i>	13.814	3	37	<i>alle</i>	1.765	6
5	<i>sein</i>	11.714	3	38	<i>über</i>	1.652	6
6	<i>zu</i>	8.628	3	39	<i>bis</i>	1.592	6
7	<i>werden</i>	7.406	4	40	<i>nur</i>	1.564	6
8	<i>von</i>	6.776	4	41	<i>müssen</i>	1.555	6
9	<i>haben</i>	6.658	4	42	<i>sollen</i>	1.516	6
10	<i>mit</i>	5.924	4	43	<i>sagen</i>	1.484	6
11	<i>an</i>	5.894	4	44	<i>geben</i>	1.476	6
12	<i>für</i>	4.825	4	45	<i>Uhr</i>	1.330	6
13	<i>auf</i>	4.779	4	46	<i>Sie</i>	1.318	6
14	<i>sich</i>	4.722	4	47	<i>mehr</i>	1.245	6
15	<i>nicht</i>	4.274	5	48	<i>man</i>	1.236	6
16	<i>es</i>	4.199	5	49	<i>neu</i>	1.228	6
17	<i>auch</i>	3.828	5	50	<i>wollen</i>	1.216	6
18	<i>er</i>	3.453	5	51	<i>oder</i>	1.203	6
19	<i>bei</i>	3.361	5	52	<i>gut</i>	1.200	6
20	<i>als</i>	3.109	5	53	<i>keine</i>	1.194	6
21	<i>dass</i>	2.790	5	54	<i>aber</i>	1.137	6
22	<i>aus</i>	2.705	5	55	<i>kommen</i>	1.093	6
23	<i>sein</i>	2.693	5	56	<i>unter</i>	1.041	7
24	<i>können</i>	2.665	5	57	<i>groß</i>	1.021	7
25	<i>dies</i>	2.660	5	58	<i>wenn</i>	1.016	7
26	<i>nach</i>	2.578	5	59	<i>gegen</i>	998	7
27	<i>ihr</i>	2.522	5	60	<i>gehen</i>	993	7
28	<i>wir</i>	2.477	5	61	<i>durch</i>	975	7
29	<i>ich</i>	2.344	5	62	<i>machen</i>	968	7
30	<i>sie</i>	2.328	5	63	<i>schon</i>	950	7
31	<i>um</i>	2.159	5	64	<i>andere</i>	938	7
32	<i>wie</i>	2.130	6	65	<i>viele</i>	889	7
33	<i>Jahr</i>	2.114	6	66	<i>dann</i>	860	7

Tabelle 2. Fortsetzung

Rang	Lemma	Frequenz	HK	Rang	Lemma	Frequenz	HK
67	<i>erst</i>	836	7	84	<i>jede</i>	561	7
68	<i>zwei</i>	821	7	85	<i>nun</i>	549	7
69	<i>wieder</i>	815	7	86	bereits	545	7
70	<i>was</i>	787	7	87	<i>heute</i>	538	7
71	<i>stehen</i>	756	7	88	<i>unser</i>	529	8
72	<i>aber</i>	730	7	89	damit	525	8
73	<i>lassen</i>	700	7	90	dabei	518	8
74	<i>immer</i>	699	7	91	<i>beide</i>	515	8
75	<i>ab</i>	685	7	92	<i>Frau</i>	515	8
76	<i>seit</i>	670	7	93	<i>Kind</i>	512	8
77	Prozent	650	7	94	<i>sehr</i>	509	8
78	<i>sehen</i>	641	7	95	<i>hier</i>	507	8
79	<i>weit</i>	632	7	96	<i>stellen</i>	498	8
80	<i>finden</i>	622	7	97	<i>Tag</i>	489	8
81	<i>jetzt</i>	620	7	98	<i>Mensch</i>	487	8
82	<i>drei</i>	611	7	99	<i>Zeit</i>	483	8
83	<i>ganz</i>	583	7	100	<i>weil</i>	477	8

erkennen. Die restlichen automatisch als Wortbildungen analysierten Lemmata der Rangliste aus Tab. 2 stellen Adverbien dar, deren synchrone Analysierbarkeit offensichtlich ist (*bereit|s, da|mit, da|bei*). Nicht als Wortbildungen analysiert wurden übrigens Fälle wie *dass* (Rang 21) oder *weil* (Rang 100), die man in einer historischen Betrachtung mit dem Pronomen *das* respektive dem Substantiv *Weile* hätte in Verbindung bringen müssen. Wir fassen sie aber als nicht einschlägig auf, da sie sich von den Basiswörtern bedeutungsmäßig bzw. funktional und teilweise auch formal zu sehr isoliert haben, als dass sie ohne Weiteres als synchron analysierbar betrachtet werden können.

Die obersten Positionen der Rangliste werden dominiert durch Funktionswörter, vor allem Determinierer, Präpositionen und Pronomina.⁴¹ Auch die ersten auftauchenden Verblemmata (*sein, werden, haben, können, müssen, sollen*) werden zum großen Teil als Kopula-, Hilfs- oder Modalverben realisiert, die ebenfalls zu den Funktionswörtern gezählt werden können. Die häufigsten

41 Das Lemma *eine* (Rang 4) steht für Determinierer, Pronomen und Zahlwort, *sein* (Rang 5) für Verb, *sein* (Rang 23) für Determinierer und Pronomen und *dies* (Rang 25) ebenfalls für Determinierer und Pronomen.

Lemmata verweisen auf einsilbige Tokens – Ausnahmen finden sich unter den ersten 30 Lemmata bei einzelnen Wortformen der oben genannten Verben (z.B. *sein, bin, bist* etc. – einsilbig vs. *waren, gewesen* etc. – mehrsilbig) und des Determinierers/Pronomens/Zahlworts *ein(e)*. Die ersten sonstigen Lemmata, die für Inhaltswörter stehen, erscheinen auf der Liste erst später und gehören wie die Substantive *Jahr* (Rang 33) und *Uhr*⁴² (Rang 45) oder die Verben *sagen* (Rang 43) und *geben* (Rang 44) den primären Wörtern an.

Festzuhalten ist, dass besonders häufige Wörter funktionale Ausdrücke darstellen, die kurz und primär sind. Auch die häufigsten Inhaltswörter sind kurz und primär. Bei besonders seltenen Wörtern hingegen überwiegen bei Weitem die Wortbildungen. Dabei sind unter den besonders seltenen Wörtern, insbesondere unter den Hapaxlegomena, die meisten Neuprägungen⁴³ zu erwarten (vgl. Baayen 2009: 905–906), und so ist die hohe Anzahl niederfrequenter Wortbildungen ein Indiz dafür, dass der Gesamtwortschatz durch Wortbildung immer weiter ausgebaut werden kann. Demgegenüber erscheint der Bestand an primären Wörtern nur beschränkt erweiterbar. Die Schöpfung neuer Wörter ohne Anschluss an vorhandenes Sprachmaterial (*Urschöpfung* oder auch einfach *Wort-schöpfung* genannt) spielt in der Gegenwartssprache kaum eine Rolle.⁴⁴ Sie ist weitgehend beschränkt auf Onomatopoetika⁴⁵ und Produktnamen (vgl. Donalies 2002: 17–18, Fleischer & Barz 2012: 18–20). An der Wortschatzerweiterung sind neben der Wortbildung vor allem die Entlehnung (*Computer, cool, Interface, Slot, Smartphone*) und die Bedeutungsveränderung bzw. -erweiterung (*geil* für ‚großartig‘, *Netz* für ‚Internet‘, *Rechner* für ‚Computer‘) beteiligt. Wohlgermerkt,

42 Das Lemma *Uhr* ist nur deshalb so häufig, weil es in Uhrzeitangaben wie *um 14.30 Uhr* auftritt. Gerade dort verhält es sich aber nicht wie ein typisches Substantiv (es ist nicht flektierbar, nicht erweiterbar etc.).

43 Neu- bzw. Spontanbildungen sind etwa in folgenden Hapaxlegomena aus dem Korpus zu vermuten: *ananasbeglückt, Bibliothekszwilling, kohlhäsisch, Kohl-Lafontaine-Weizsäcker-Vollmer-Mischung, valutadeutsch, wohlfahrtsvernünftig, wurmisch, (das) Zerlesenswerden*.

44 Anders als eine neue Wortbildung, die Assoziationen zu bereits bestehenden Wörtern wecken kann, kann ein gerade eingeführtes primäres Wort nur in einem absolut eindeutigen Kontext verstanden werden – etwa gestisch begleitet oder versehen mit einer Bedeutungserklärung wie „**Sitt** [...] ist ein Kunstwort, das als Adjektiv das Gegenteil von *durstig* bedeuten [...] soll“ (<https://de.wikipedia.org/wiki/Sitt>, abgerufen am 8.6.2020).

45 Allerdings sieht Donalies (vgl. 2002: 18, 91) in Verben aus Onomatopoetika (z.B. *miauen* oder *platschen*) bereits Wortbildungen. Hingegen betrachtet sie die Substantive *Kuckuck* und *Wauwau* als onomatopoetische Urschöpfungen (vgl. auch Fleischer & Barz 2012: 19). Henzen (vgl. 1965: 6–7) geht bei onomatopoetischen Verben dann von Wortbildungen aus, wenn sie verbale Suffixe *-r-* bzw. *-l-* wie in *plappern* oder *bimmeln* aufweisen.

die entlehnten Wörter können auf primäre Wörter (*cool, slot*) oder Wortbildungen (*computer, interface, smart phone*) der Gebersprache zurückgehen. Letztere werden nicht selten auch im Deutschen als Wortbildungen empfunden (*Interface, Smartphone*), da sie mit anderen geläufigen Wortbildungen (*Intercity, international* bzw. *Smartboard, Smartcard, Smart-TV*) leicht assoziierbar sind. Sie sind somit synchron analysierbar und auch erklärbar. Der Wortbildungsursprung eines Fremdworts kann im Deutschen manchmal aber auch kaum zu erkennen sein wie bei *Aids*⁴⁶ oder *Laser*, die im Englischen als multisegmentale Kurzwörter entstanden sind. Solche Entlehnungen können wie primäre Wörter wahrgenommen werden (vgl. dazu Dudengrammatik 2016: 744) und sind als synchron nicht analysierbar zu betrachten.

In einem umfangreichen Korpus können nicht alle Fremdwörter aufgespürt und daraufhin überprüft werden, ob sie synchron analysierbar sind, und selbst wenn das möglich gewesen wäre, hätte es bei einzelnen Entscheidungen öfter Interpretationsspielraum gegeben.⁴⁷ Wir stützen uns daher beim Gesamtkorpus auch bei Fremdwörtern auf die Ergebnisse der automatischen Wortbildungsanalysen von Canoo, die punktuell überprüft und korrigiert werden. In diesen werden von den oben genannten Fremdwörtern nur die ältere Bildung *international* sowie die neueren *Smartcard* und *Smartboard* als Wortbildungen dargestellt. Vereinbar mit unseren Überlegungen werden *Aids, Radar* und *Computer* nicht als Wortbildungen analysiert, ebenso wenig aber – diesmal nicht mit unseren Überlegungen konform – *Intercity, Interface, Smartphone* und *Smart-TV*. Die Annotation von Belegen wie den letzteren wird, sofern die Fehler gefunden werden, korrigiert. Was sich hier zeigt, gilt für die Nutzung der Wortbildungsannotation zur Suche nach Wortbildungen im Fremdwortbereich insgesamt. Die Annotation ist relativ präzise, zeigt jedoch Recall-Schwächen.

5 Quantitative Verhältnisse in verschiedenen Wortklassen

Die relative Häufigkeit von Wortbildungen variiert von Wortart zu Wortart und auch von einer Teilkategorie innerhalb einer Wortart zur anderen. Näher betrachtet werden hier in diesem Zusammenhang die Wortarten Substantiv, Adjektiv, Verb und Adverb, deren Lexembestand in größerem Maße erweiterbar ist. Sie

46 Durch die ebenfalls gebräuchliche Großschreibung *AIDS* wird das Wort als Kurzwort (Buchstabenwort) gekennzeichnet, was es formal wieder zur Wortbildung macht.

47 So könnte *Computer* von manchen als synchron nicht analysierbar und nicht erklärbar aufgefasst werden, von anderen aber als analysierbar in ein Wortbildungsmorphem {-er} und ein Konfix {comput-}, das in unserem Korpus noch in dem Ausdruck (*ubiquitäres*) *Computing* vorkommt.

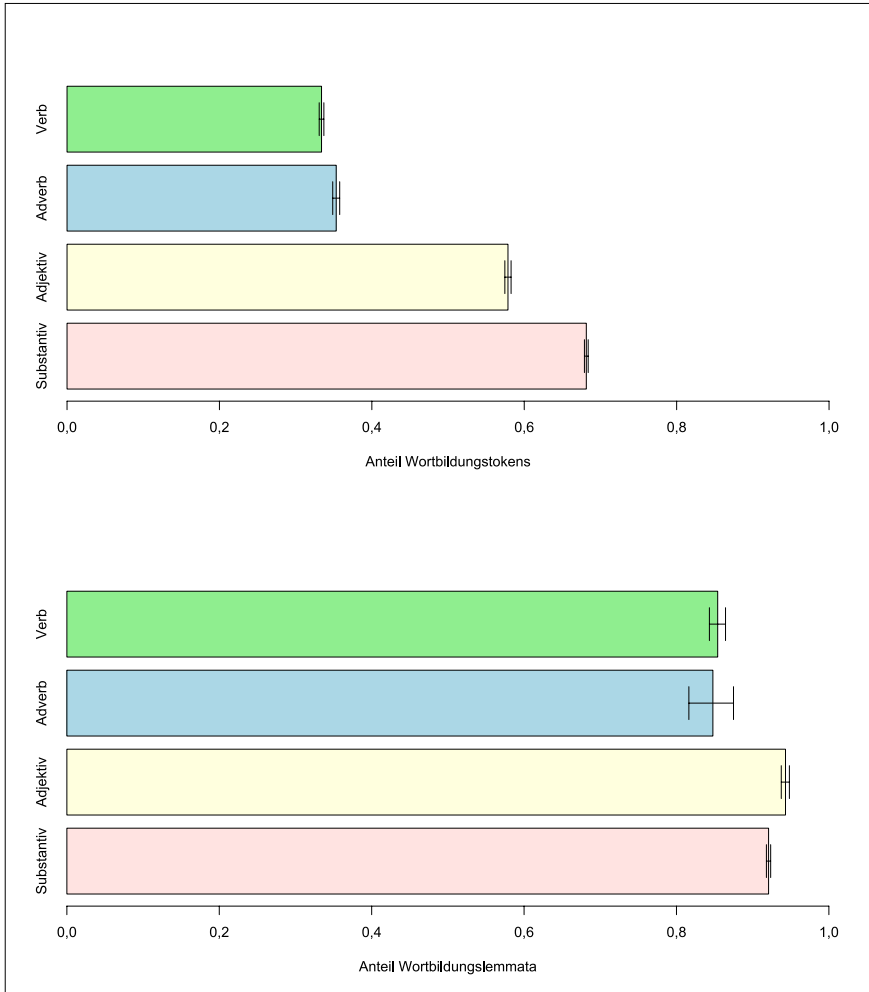


Abbildung 3: Anteil der Wortbildungen bei den erweiterbaren Wortarten im Untersuchungskorpus anhand des Teilkorpus *40k* (jeweils mit einem 95 %-Konfidenzintervall).

setzen sich in erster Linie aus Inhaltswörtern zusammen und werden allgemein als offene Klassen angesehen.

Substantive und Adjektive sind deutlich wortbildungsfreundlicher als Adverbien und Verben. Abb. 3 zeigt den für das Teilkorpus *40k* ermittelten Anteil der Wortbildungen an allen Tokens bzw. Lemmata der vier Wortarten.⁴⁸

48 Zu der Zählweise ist zu bemerken, dass Partikelverben nur dann als *ein* Wort eingestuft werden, wenn sie im Text zusammengeschrieben erscheinen. Bei der Distanzstellung der Partikel (bzw. Getrennschreibung des Partikelverbs) wird die abgetrennte

Der Unterschied zwischen dem verhältnismäßig niedrigen Anteil der Wortbildungen an den Tokens und ihrem durchgehend deutlich höheren Anteil an den Lemmata ist besonders ausgeprägt bei Verben und Adverbien. Dies kann schon daran liegen, dass es einige hochfrequente Verb- und Adverb-lemmata gibt, die primäre Wörter sind (z. B. *sein*, *werden*, *haben* bzw. *auch*, *noch*, *so*, vgl. Tab. 2 weiter oben). Bei allen vier Wortarten ist der Anteil der Wortbildungen an allen Lemmata allerdings abhängig von der Größe des Teilkorpus. Wird diese erhöht, steigen die Anteile der Wortbildungen im Bereich der Lemmata bei allen vier Wortarten an, während ihre Anteile im Bereich der Tokens relativ unveränderlich bleiben. Hier wirkt sich die Produktivität der Wortbildungen aus: Mit wachsender Anzahl berücksichtigter Tokens kommen bei Wortbildungen schnell neue Lemmata dazu, wogegen bei primären Wörtern die Anzahl der Lemmata viel verhaltener ansteigt. Abb. 4 zeigt die Wortschatzwachstumskurven⁴⁹, die mithilfe des Softwarepakets zipfR unter Anwendung des finiten Zipf-Mandelbrot-LNRE-Modells (vgl. Evert & Baroni 2007) allen Wortbildungen und allen tendenziell primären Wörtern im Teilkorpus 40K zugeordnet werden können.

Während die Wachstumskurve für die Wortbildungslemmata auch am rechten Rand der Abb. 4 immer noch relativ steil ansteigt, flacht die Wachstumskurve für die Lemmata primärer Wörter sehr schnell ab, sodass ihr Anstieg bald kaum mehr bemerkbar ist. Ähnliches zeigt sich beim Vergleich der Kurven für die Hapaxlegomena (unter diesen sind die meisten Neuprägungen zu erwarten): Die Kurve für die Wortbildungen steigt am rechten Rand der Abbildung immer noch deutlich an, während die Kurve für die primären Wörter da sogar bereits abfällt. Aus diesen Vergleichen kann man erneut auf die grundsätzliche Produktivität der Wortbildungsprozesse und die vergleichsweise geringe Bedeutung der Wortschatzerweiterung durch primäre Wörter schließen.⁵⁰ Zu prinzipiell ähnlichen Ergebnissen kommen Roth (vgl. 2014: 177–182) beim Vergleich verschiedener

Partikel separat gezählt (z. B. in *Verspätungen setzen sich wie in einer Kettenreaktion weiter fort*) und das übrig gebliebene Verb wird, sofern es nicht weiter analysierbar ist, als primär klassifiziert. Die Überprüfung einer Stichprobe ergab, dass sich unter den als primär eingestuften Verbtokens $5,8\% \pm 2,3\%$ ($n=400$, 95%-Konfidenzintervall) Verben befinden, zu denen eine Verbpartikel in Distanzstellung gehört. Wenn man die entsprechenden Partikelverben zu den Wortbildungen zählen würde, würde der Anteil der Wortbildungen an allen Verbtokens also leicht (um ca. 4 Prozentpunkte) ansteigen.

49 Auch *Vocabulary growth curve* (so bei Evert & Lüdeling 2001: 168), *Vokabularwachstumskurve* (bei Roth 2014: 169).

50 Die potenzielle Produktivität P (Quotient aus der Anzahl der Hapaxlegomena einer Kategorie und der Anzahl aller Tokens dieser Kategorie, vgl. Baayen 2009) beträgt bei 382.108 Tokens (beobachtete Tokenanzahl der primären Wörter) 0,0049 für primäre Wörter und 0,1213 für Wortbildungen.

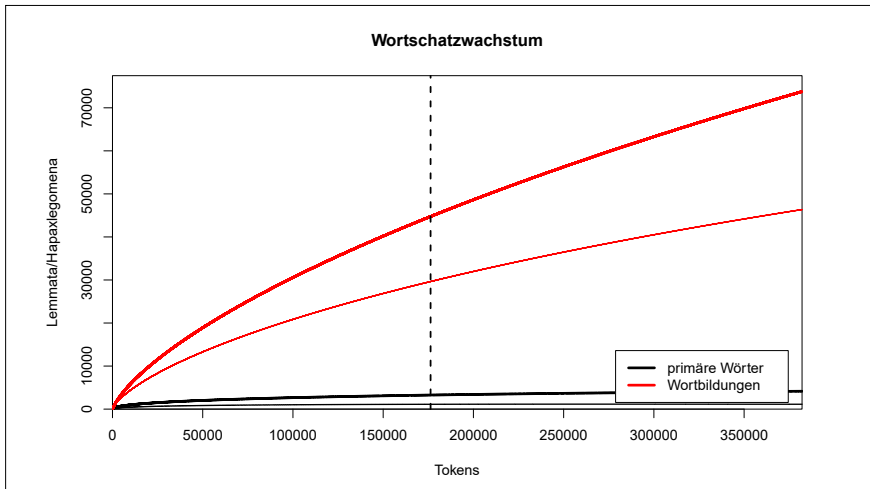


Abbildung 4: Anzahl der Lemmata (dickere Kurvenlinien) und der Hapaxlegomena (dünnere Kurvenlinien) bei steigenden Tokenzahlen berücksichtigter Wortbildungen bzw. primärer Wörter im Teilkorpus *40k* (gemäß dem finiten Zipf-Mandelbrot-Modell in *zipfR*). Bei den Wortbildungen wurden die Wachstumskurven über die beobachtete Anzahl der Tokens (176.266, gestrichelte Linie) hinaus so weit extrapoliert, bis die beobachtete Anzahl der Tokens der primären Wörter (382.108, Ende der Skala) erreicht wurde.

Kompositionsmuster mit Simplizia im Deutschen sowie bereits Baayen (vgl. 2001: 135–138) beim Vergleich von Derivationen auf *-heid* mit simplizischen Substantiven im Niederländischen.⁵¹

Wie zu erwarten, sind Wortbildungen verschiedener Wortarten unterschiedlich produktiv (zur Kompositionsproduktivität vgl. Roth 2014: 177–178). Die Wortschatzwachstumskurven in Abb. 5 unten lassen auf deutliche Unterschiede schließen. Im Teilkorpus *40k* sind die Wortbildungen besonders produktiv bei den Substantiven. Die anderen Wortarten folgen in der Reihenfolge Adjektive⁵², Verben und Adverbien. Die Wachstumskurve der Lemmata für Adverbien zeigt in Abb. 5 kaum mehr Steigung und die Kurven der Hapaxlegomena für Adverbien und Verben sinken am rechten Rand der Abbildung sogar schon.⁵³

51 In Baayens Modellen steigen die Wachstumskurven der Worttypes und der Hapaxlegomena bei Derivationen auf *-heid* immer weiter an, während bei Simplizia die Anzahl der Worttypes nicht mehr wächst und die Anzahl der Hapaxlegomena abnimmt, sobald höhere Tokenzahlen erreicht werden.

52 Die Hapaxlegomena-Kurve für Adjektive ist in Abb. 5 durch die Lemmata-Kurve für Verben verdeckt.

53 Die potenzielle Produktivität P für das Ende der Skala (59.561 Tokens) erreicht für Substantive 0,2980, für Adjektive 0,0758, für Verben 0,0212 und für Adverbien 0,0009. Die Extrapolation der Kurven für Adverbien auf das Vierfache der beobachteten

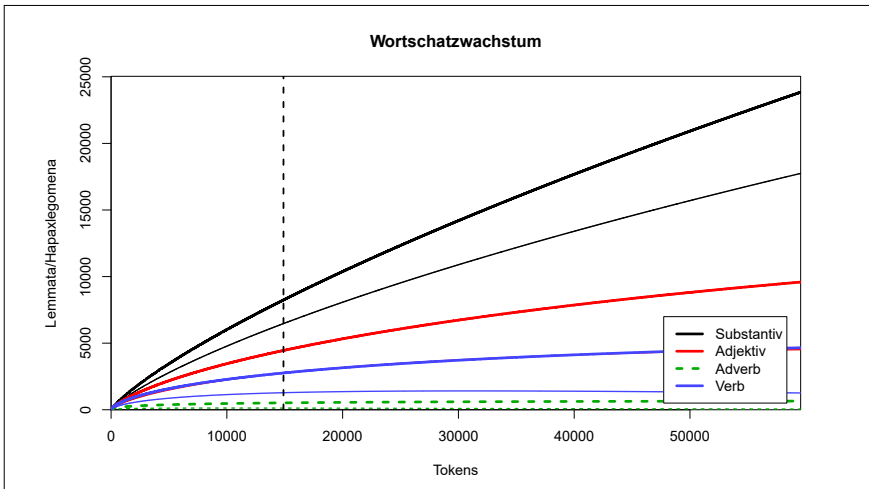


Abbildung 5: Anzahl der Lemmata (dickere Kurven) und der Hapaxlegomena (dünnere Kurvenlinien) für verschiedene Wortarten bei steigenden Tokenzahlen anhand des Teilkorpus *40k* (gemäß dem finiten Zipf-Mandelbrot-Modell in *zipfR*). Die Wachstumskurven wurden auf das Vierfache der beobachteten Anzahl der Tokens der seltensten Wortart Adverb extrapoliert. Die gestrichelte Linie zeigt die beobachtete Anzahl der Tokens für Adverbien (14.890).

Unsere Betrachtungen bedürfen einiger teilweise einschränkender Kommentare. Die traditionellen Wortarten sind auf hohen Abstraktionsniveaus definiert und stellen in mancher Hinsicht sehr grobe Oberkategorien dar. Die Unterkategorien können Teilklassen konstituieren, die bezüglich der Wortbildung differenziert zu betrachten sind. So ist die Wortart Substantiv wie bereits erwähnt durch die Einteilung in Appellative und Eigennamen geprägt. Beide Untergruppen unterscheiden sich deutlich in ihrer Semantik, Pragmatik und Schreibung. Eigennamen weisen auch Besonderheiten hinsichtlich der gegenwartssprachlich-synchronen Analysierbarkeit und Erklärbarkeit der Wortbildungsmuster auf (vgl. Fleischer & Barz 2012: 179). Hier ist noch einmal zu betonen, dass wir in den quantitativen Übersichten nur die „regulärere“ Wortbildung der Appellative fokussiert und die Eigennamen weitestmöglich ausgeschlossen haben. Ferner können in den Wortbildungen im Allgemeinen Bausteine erscheinen, die stark reihenbildend sind. Dabei handelt es sich nicht nur um Affixe. In appellativen Substantiven etwa sind es auch manche kompositionalen Zweitglieder wie das polyseme *Werk* in *Kraft|werk*, *Kunst|werk* oder *Studierenden|werk*

Tokenanzahl entspricht der Empfehlung von Evert & Baroni (2005) für noch plausible Ergebnisse.

(man vergleiche dazu auch das meist schon als Suffix betrachtete *-werk*⁵⁴ in *Mauer|werk*, *Netz|werk*, *Regel|werk*), *Stelle* in *Bau|stelle*, *Geschäfts|stelle* oder *Halte|stelle* und *Mann* in *Ehe|mann*, *Kauf|mann* oder *Vorder|mann*. Als Adjektive wiederum sind viele Kardinal- und Ordinalzahlverwendungen einzustufen, z. B. in *fünfzehn Jahre* oder *das fünfzehnte Jahrhundert*. Sie bilden potenziell unendliche Reihen aus Formen, die nach einigen wenigen, teilweise idiosynkratischen Wortbildungsregeln gebildet werden. Die Zahlwörter können im Übrigen außer mit Buchstaben auch mit Ziffern, u. U. unter Zuhilfenahme der Punktsetzung wie in *der 5. Mai* ausgedrückt werden, allerdings wurden solche mit Ziffern verschriftlichten Zahlwörter von der Untersuchung ausgenommen. Ausgeschlossen wurden dabei auch die von einem Punkt gefolgt Ziffern, die sich unter Umständen als aus Ordinalzahlen abgeleitete Adverbien interpretieren lassen (z. B. *1.* im Sinne von *erstens*) – solche Adverbkandidaten waren durch den Tagger nämlich ziemlich unzuverlässig klassifiziert worden.⁵⁵ Die Adverbien insgesamt zerfallen in verschiedene Teilklassen, die hinsichtlich der Wortbildung ganz unterschiedlich zu charakterisieren sind. So beträgt etwa der Anteil der Wortbildungen an Präpositionaladverbien – sie setzen sich aus einem Pro-Element und einer Präposition zusammen – per definitionem 100 % (Tokens und Lemmata). Allerdings bilden Präpositionaladverbien unter allen Adverbien eher eine kleine Gruppe (knapp 7 % der Lemmata), sodass sie die Zählungen zu Adverbien insgesamt nur geringfügig beeinflussen. Diese erscheinen also dennoch verhältnismäßig wenig wortbildungsfreudig (vgl. Abb. 3 weiter oben).

Anders als in den oben behandelten Wortarten ist der Lexembestand in denjenigen Wortarten, die sich aus Funktionswörtern zusammensetzen, sehr beschränkt ausbaufähig. Determinierer, Pronomina und Konjunktionen werden oft als geschlossene Klassen betrachtet (dazu kritisch Waßner 2002). Charakteristisch für diese Wortarten ist eine im Vergleich zu den berücksichtigten vier Wortarten überschaubare Anzahl an Lexemen, die teilweise mit sehr hohen Tokenzahlen (vgl. Tab. 2 weiter oben) realisiert werden. So entsprechen etwa den nur zwölf Formen des bestimmten und des unbestimmten Artikels bzw. der gleichlautenden Pronomen (*die*, *der*, *das*, *des*, *dem*, *den* einerseits und *eine*, *einer*, *ein*, *eines*, *einem*, *einen* andererseits) im Gesamtkorpus (knapp 6 Mrd. Wort-tokens) fast 700 Mio. Tokens.

54 So z. B. in Fleischer & Barz (2012: 56).

55 Sie werden im Kapitel „Bau von und Umbau zu Adverbien“ (Brandt in dieser Ausgabe) dennoch stichprobenartig untersucht.

6 Wortbildungsprofile verschiedener Wortarten im Vergleich

Die einzelnen Wortarten sind u. a. dadurch charakterisiert, dass sie verschiedene Wortbildungsarten unterschiedlich stark gewichten⁵⁶ und diese mit jeweils spezifischen Wortbildungsregeln realisieren. Die Verteilung der anhand der Wortbildungsannotation ermittelten Wortbildungsarten auf die Wortarten Substantiv, Adjektiv, Verb und Adverb im Teilkorpus *40k* wird für Tokens und Lemmata in Abb. 6 illustriert. Im Anschluss diskutieren wir die wichtigsten Gemeinsamkeiten und Unterschiede, die sich beim Vergleich der Wortarten zeigen. Dabei werden auch die jeweils auffälligsten Wortbildungsregeln genannt und durch Beispiele veranschaulicht. Zum Schluss wird in den Tab. 3 bis 6 eine Übersicht über die typischen Wortbildungsregeln der einzelnen Wortarten gegeben.

In einer ersten formal-kompositionalen Annäherung unterscheiden wir an dieser Stelle als Wortbildungsarten lexikalische Verbindung, Suffigierung, Präfigierung und Konversion. Die übrigen Bildungsweisen fassen wir unter „Rest“ zusammen. Es sind einige Spezifika dieser Unterteilung zu erläutern.

Unter *lexikalischer Verbindung* verstehen wir – provisorisch zusammengefasst – alle Verbindungen aus wortfähigen Einheiten⁵⁷ und/oder Konfixen als unmittelbaren Konstituenten (Fugenelemente werden hier vernachlässigt). Dazu gehören vor allem „klassische“ Komposita (vgl. z. B. Schlücker 2012: 8), in denen das Zweitglied die gleiche Wortart hat wie die gesamte Wortbildung (*Haus|tier*, *zahl|reich*, *darauf|hin*). Hier werden aber auch einige abweichende Bildungen berücksichtigt, etwa solche, in denen die Verbindung nicht die Wortart des Zweitglieds hat (*da|mit*, *dem|nach*) und die auch als Univerbierung⁵⁸ betrachtet werden können (*trotz|dem*, *in|zwischen*, vgl. Donalies 2018: 225–243). Unter Präfigierung findet sich außer den üblicherweise dazugerechneten Fällen (*Un|glück*, *Ge|flügel*, *er|klären*, *Ge|bet*) – bedingt durch die Spezifik der automatischen Annotation – auch die Zirkumfigierung (*Ge|länd|e*, *Ge|web|e*). Die Suffigierung wiederum umfasst neben den üblichen Fällen (*Musik|er*, *beruf|lich*) die meisten Zusammenbildungen mit Suffixen (*Vier|bein|er*, *groß|züg|ig*) und manche

56 Eichinger (2000: 74–114) beobachtet aus der umgekehrten Perspektive Präferenzen bei der Verteilung der Wortbildungsarten auf die Wortarten Substantiv, Adjektiv, Verb und erörtert eingehend die Gründe für solche Präferenzen.

57 Hierzu gehören wohlgermerkt auch Komponenten, denen lexikalische Bedeutung abgesprochen wird (zumindest *dem* in *trotz|dem*). Wir verwenden dennoch die Bezeichnung *lexikalische Verbindung*, weil wir hier Komponenten mit lexikalischer Bedeutung für prototypisch halten.

58 Mit *Univerbierung* kann der Prozess der Zusammenrückung von Wörtern einer Wortgruppe zu einem Einzelwort bezeichnet werden, aber auch ein komplexes Wort, das formal einer Wortgruppe gleicht – mit dem Unterschied, dass die Wortgruppe auseinander geschrieben wird.

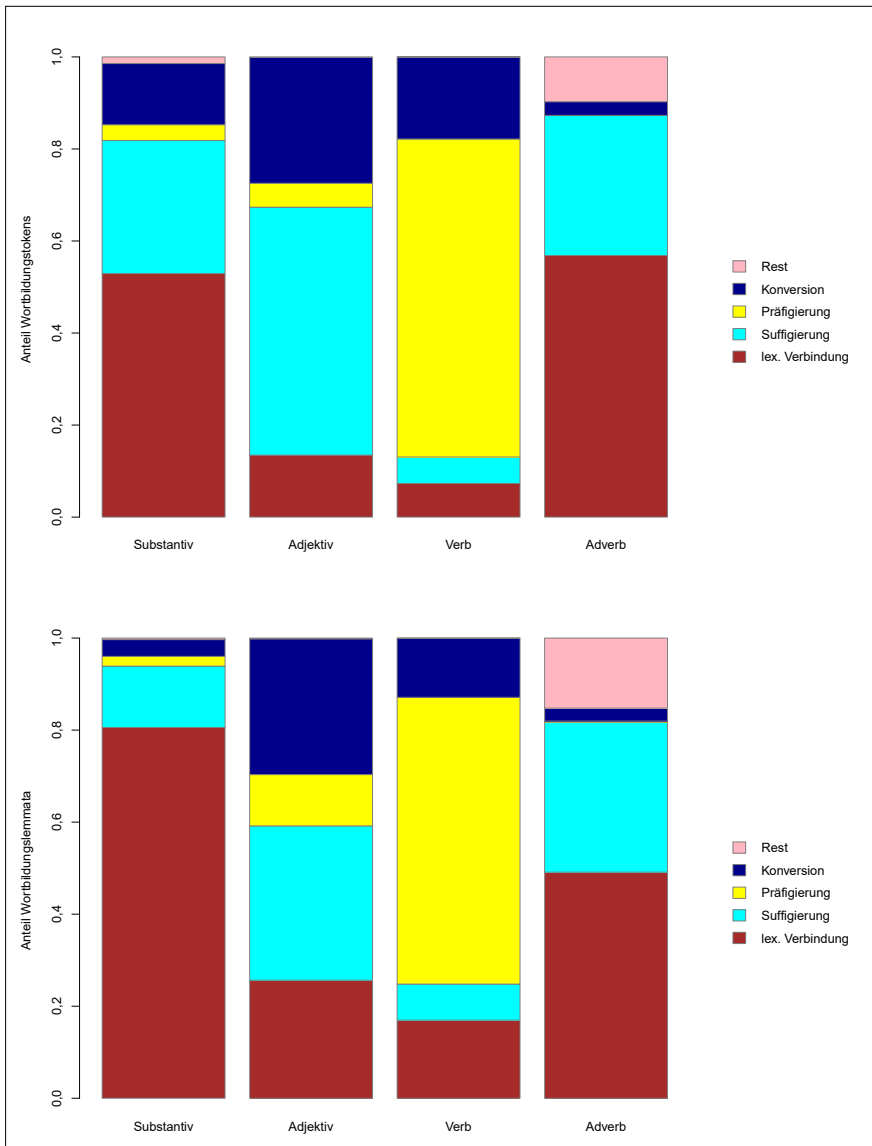


Abbildung 6: Wortbildungsprofile verschiedener Wortarten: Anteile der Wortbildungsarten.

Rektionskomposita (*Feuer|spuck|er*, *Graffiti-|Sprüh|er*).⁵⁹ Unter Zusammenbildungen verstehen wir dabei – ähnlich wie u. a. Neef (2015: 583–585, zurückgehend auf Leser 1990: 4) – Dreierverbindungen, bei denen weder die ersteren beiden noch die letzteren beiden Komponenten mit der in die gesamte Bildung integrierbaren Bedeutung wortfähig sind (vgl. **Vierbein*, **Beiner* bzw. **Großzug*, **zügig*)⁶⁰ – in solchen Fällen geht die Basis der Suffigierung scheinbar auf eine Wortgruppe oder Fragmente einer größeren Wortgruppe zurück. Die Konversion umfasst an dieser Stelle neben den typischen Fällen wie *fertigen* oder (*das*) *Leben* auch departizipiale Adjektive und Substantive (*spannend*, (*der/ein*) *Betroffene(r)*) sowie auch implizite Ableitungen ((*das*) *Verbot* < *verbieten*, (*der*) *Auftritt* < *auftreten*). Zum „Rest“ schließlich gehören einige analysierbare, aber von Canoo aus verschiedenen Gründen nicht mit einer Regel verbundene Bildungen ((*das*) *Alter*, *ausfindig*, *schließlich*, *übrigens*). Eine genauere Differenzierung der Wortbildungsarten wird später im gleichnamigen Kapitel ausgearbeitet.

Der Vergleich des Anteils der lexikalischen Verbindung bei verschiedenen Wortarten legt auf den ersten Blick ihre paarweise Gruppierung nahe. **Substantive** und **Adverbien** ähneln sich darin, dass sie besonders durch lexikalische Verbindungen geprägt sind. Bei den Substantiven fallen insbesondere determinativ analysierbare Komposita aus zwei substantivischen Konstituenten wie *Wochen|ende*, *Stadt|rat*, *Verzichts|planung* ins Gewicht,⁶¹ in größerem Abstand gefolgt von Komposita mit einer adjektivischen Erstkonstituente wie *Alt|papier*, *Neu|verschuldung*, *Halb|zeit* und solchen mit einer verbalen Erstkonstituente wie *Warte|zeit*, *Fahr|zeug*, *Halte|stelle*. Etwas seltener, aber immer noch sehr zahlreich sind Komposita mit einer präpositionalen Erstkonstituente wie *Mit|glied*, *Auf|takt*, *Gegen|rede* und Konfixverbindungen wie *Tele|fon*, *Dia|lekt*, *Gastro|nom*.

Bei den Adverbien sind Verbindungen aus Adverb und Präposition (Präpositionaladverbien u. ä.) wie *da|durch*, *her|aus*, *wo|nach* am auffälligsten – sie entsprechen nicht den typischen Komposita (zu diesen vgl. etwa Schlücker 2012: 8 oder Ortner & Ortner 1984: 36) – nicht zuletzt, weil das rechts stehende Element nicht die Wortart der gesamten Wortbildung aufweist. Häufig sind außerdem Komposita aus zwei Adverbien wie *wo|her*, *darauf|hin*, *immer|hin*. Zu weiteren relevanten lexikalischen Verbindungen, die übrigens ebenfalls

59 Die Zusammenbildungen aus (Zahl-)Adjektiv, Substantiv und Suffix werden oft auch zu den Komposita gezählt (so bei Neef 2015: 584–585, anders bei Donalies 2002: 95–96). Vor allem aber hätten Rektionskomposita auch der lexikalischen Verbindung zugeordnet werden können.

60 Somit zeigen sie eine Bildungsweise, in der mehrere Wortbildungsschritte zu einem zusammengefasst werden. Sie hat eine größere Reichweite als bisher angenommen und wird im Kapitel „Wortbildungsarten“ ausführlicher behandelt.

61 Hier sind auch Verbindungen mit Bindestrich(en) sehr zahlreich.

nicht den typischen Komposita entsprechen, gehören vor allem solche aus einer Präposition und einem Pronomen wie *trotz|dem, unter|dessen, ohne|dies* (unter ihnen sind auch Verbindungen mit einem komplexen Pronomen als Zweitglied zu finden wie *in|einander, gegen|einander, mit|einander*) und solche aus zwei Präpositionen wie *vor|bei, in|zwischen, mit|unter*. Die Ähnlichkeit der Substantive und der Adverbien hinsichtlich ihrer Prägung durch lexikalische Verbindungen entpuppt sich somit schnell als eine nur scheinbare. Die Positionierung der adverbialen Bildungen gegenüber den kanonischen, substantivischen Komposita wird ein wichtiger Punkt im Kapitel „Wortbildungsarten“ sein. Dem Aufbau von Adverbien aus Lexemen unterschiedlicher Provenienz, bei dem neben der lexikalischen Verbindung auch die Suffigierung eine sehr wichtige Rolle spielt, widmet sich das Kapitel „Bau von und Umbau zu Adverbien“ (Brandt in dieser Ausgabe).

Substantive und Adverbien haben darüber hinaus, was die Tokens angeht, einen ähnlichen Anteil an Suffigierungen, aber auch hier sind die Bildungen eigentlich grundsätzlich verschieden. Bei Substantiven dominieren die „klassischen“ Suffigierungen wie *Erklär|ung, Musik|er, Gewiss|heit*, bei denen das komplexe Substantiv mithilfe eines spezialisierten Suffixes vor allem von einem Verb, sehr häufig aber auch von einem anderen Substantiv oder einem Adjektiv abgeleitet wird. Bei Adverbien kommen die Suffixe *-s* und *-ens* hingegen in Verbindung mit sehr vielfältigen Basen zum Einsatz (*angesicht|s, bereit|s, zusehend|s, vergeben|s, erst|ens, höchst|ens*). Die Verwendung von *-s* überlagert dann auch einen Teil der Bildungen mit suffixartigen desubstantivischen Zweitgliedern wie *andern|fall|s, dies|seit|s, heim|wärt|s*. Zu solchen Bildungen gehören ansonsten auch etwa *beispiels|weise, glücklicher|weise* oder *gleicher|maßen*. Die sehr produktiven und theoretisch schwer einzuordnenden Adverbien mit desubstantivischen Zweitgliedern sind ebenfalls Gegenstand eines eigenen Kapitels.

Im restlichen Teil ihrer Wortbildungsprofile sind sich Substantive und Adverbien schon auf den ersten Blick unähnlich. Auffällig ist bei Substantiven ein merklicher Anteil an Konversionen im Bereich der Tokens (*(das) Leben, (der) Anfang, (der) Auftritt*), worin sie mit den Adjektiven und Verben vergleichbar erscheinen,⁶² bei den Adverbien hingegen ein beträchtlicher Anteil an Bildungen, die nicht den etablierten Wortbildungsarten entsprechen wie *schließ|lich, übrig|ens, zu|nächst* (in Abb. 7 dem „Rest“ zugeordnet) – einige von ihnen können als Univerbierungen betrachtet werden (*seiner|zeit, zur|zeit, so|zu|sagen*, vgl. Donalies 2018: 176–262). Präfigierungen sind bei Adverbien absolut

62 Es ist aber nur eine Vergleichbarkeit hinsichtlich der Anteile und nicht in Bezug auf die Bildungsweisen. Letztere werden im Kapitel „Wortbildungsarten“ genauer dargestellt.

randständig,⁶³ bei Substantiven machen sie sich vor allem im Hinblick auf den Anteil der Tokens bemerkbar (*Miss|erfolg, Un|menge, Ur|aufführung*).

Bei **Adjektiven** und **Verben** sind die lexikalischen Verbindungen anteilmäßig schwächer ausgeprägt. Bei Adjektiven fallen vor allem die Verbindungen mit einem Substantiv als Erstglied ins Gewicht wie *zahl|reich, sinn|voll, welt|weit*, in einem deutlichen Abstand gefolgt von den Verbindungen mit einem Adjektiv als Erstglied wie *weit|gehend, offen|sichtlich, gleich|gültig*. Bildungen beider Typen sind oft (wie die meisten lexikalischen Verbindungen bei den Substantiven) als typische Komposita zu betrachten. Unter den Verbindungen mit einem Adjektiv als Erstglied finden sich aber auch viele, die auch als Suffigierungen bzw. Zusammenbildungen interpretiert werden könnten wie *recht|zeit|ig, voll|ständ|ig, frei|will|ig*. Bei Verben wiederum kommen spezielle morphosyntaktische Regularitäten zum Tragen. So machen sich hier die in Verbletzstellung befindlichen und daher zusammengeschriebenen Partikelverben besonders bemerkbar. Dabei fallen vor allem Verben mit depräpositionalen und deadverbialen Erstgliedern wie *mit|teilen, fort|setzen, hin|weisen* ins Gewicht, dann in einem größeren quantitativen Abstand (mit absteigender Häufigkeit) solche mit einem deadjektivischen (*bekannt|geben, sicher|stellen, offen|legen*), desubstantivischen (*statt|finden, gewähr|leisten, teil|haben*) und deverbalen (*kennen|lernen, stehen|lassen, liegen|bleiben*) Erstglied.

Was die Verben im Gesamtvergleich auszeichnet, ist der besonders hohe Anteil an Präfigierungen. Es dominieren die Ableitungen eines Verbs aus einem anderen Verb wie *er|klären, be|stehen, ent|scheiden*. Quantitativ sehr deutlich nachgeordnet erscheinen Ableitungen mit einer desubstantivischen (*be|fristen, ver|ursachen, ent|puppen*) oder deadjektivischen Basis (*be|treuen, er|ledigen, ver|öffentlichen*). Bei den Adjektiven ist wiederum der Anteil der Suffigierung im Vergleich mit den drei anderen Wortarten am höchsten. Dabei erfolgt die Ableitung vor allem aus Substantiven wie in *beruf|lich, beispiel|haft, witz|ig*, schon deutlich seltener aus Verben wie in *deut|lich, denk|bar, abhäng|ig* und noch seltener aus Adjektiven wie in *kürz|lich, stet|ig, gemein|sam*. Auf die Ableitung aus Adjektiven folgen, was die Häufigkeit angeht, Zusammenbildungen wie *gleich|zeit|ig, wahr|schein|lich, lang|frist|ig*. Sie können als Ableitungen aus Adjektiv-Substantiv-Verbindungen interpretiert werden, welche nicht als Wörter (Komposita) etabliert sind (**Gleichzeit, *Wahrschein, *Langfrist*).

Adjektive weisen im Weiteren auch den im Vergleich der Wortarten höchsten Anteil an Konversionen auf, was vor allem darauf zurückgeht, dass hierzu die große Anzahl von adjektivisch gebrauchten Partizipien bzw. partizipbasierten Bildungen wie *vergangen, entsprechend, bewusst* gehört. Bei Verben ist der Anteil

63 Zu Bildungen wie *zu|tiefst* als Präfigierungskandidaten vgl. Fleischer & Barz (2012: 370) und das Kapitel „Bau von und Umbau zu Adverbien“.

Tabelle 3: Wortbildungsregeln beim Substantiv.

Substantiv		
Wortbildungsart	Typische Regeln	Beispiel
Lexikalische Verbindung	1. SBST+SBST 2. ADJ+SBST, V+SBST 3. PRÄP+SBST 4. KONF+KONF, KONF+SBST 5. ADV+SBST, [ADJ+SBST]+SBST	<i>Wehrmachts ausstellung</i> <i>National rat, Turn halle</i> <i>Ab glanz</i> <i>Disko thek, Atmo sphäre</i> <i>Auswärts spiel, Lang zeit projekt</i>
Suffigierung	1. V+SUFF 2. SBST+SUFF 3. ADJ+SUFF	<i>Besuch er</i> <i>Herr schaft</i> <i>Aufmerksam keit</i>
Konversion	1. V → SBST 2. ADJ → SBST	<i>Beginn</i> <i>Intellektuelle(r)</i>
Präfigierung	1. PRÄF+SBST 2. PRÄF+V(+SUFF)	<i>Un glück</i> <i>Ge sang, Ge bild e</i>

der Konversionen wie *landen*, *nutzen*, *sichern* ebenfalls beachtlich, wobei sich vor allem bei Stämmen, die sowohl substantivisch als auch verbal gebraucht werden, zumindest synchron gesehen öfter die Frage der Konversionsrichtung stellt, z. B. *Spiel* < *spielen* oder *spielen* < *Spiel* (vgl. Eisenberg 2013: 284). Die Präfigierung ist bei Adjektiven (*miss|mutig*, *un|schön*, *ur|alt*) anteilmäßig etwa so schwach ausgeprägt wie die Suffigierung bei Verben. Letztere wird durch das Suffix *-(is)ier* im Anschluss an fremde Basen dominiert. Gänzlich heimische Bildungen wie *fest|igen*, *rein|igen*, *genehm|igen* sind deutlich seltener.

In den Tab. 3 bis 6 werden die für die vier Wortarten typischen Wortbildungsregeln nach Wortbildungsarten geordnet gelistet und jeweils mit einem Beispielbeleg veranschaulicht. Die häufigeren Wortbildungsarten und Regeln werden zuerst aufgeführt.⁶⁴ Sind bestimmte Regeln innerhalb einer Wortbildungsart etwa gleich häufig, erscheinen sie in derselben Zeile.

64 Bei den Häufigkeiten für Tokens und für Lemmata ergeben sich in der Regel gleiche Reihenfolgen wie schon weiter oben in der Abb. 6 für Wortbildungsarten.

Tabelle 4: Wortbildungsregeln beim Adjektiv.

Adjektiv		
Wortbildungsart	Typische Regeln	Beispiel
Suffigierung	1. SBST+SUFF 2. V+SUFF 3. ADJ+SUFF 4. ADV+SUFF	<i>farb ig</i> <i>bedenk lich</i> <i>eigent lich</i> <i>heut ig</i>
Konversion	1. V → ADJ 2. SBST → ADJ	<i>bestehend</i> <i>ernst</i>
Lexikalische Verbindung	1. SBST+ADJ 2. ADJ+ADJ 3. PRON+ADJ, ADV+ADJ	<i>jahre lang</i> <i>halb voll</i> <i>all gegenwärtig, wohl bekannt</i>
Präfigierung	1. PRÄF+ADJ 2. PRÄF+[V+SUFF]	<i>in direkt</i> <i>un ermüd lich</i>


Tabelle 5: Wortbildungsregeln beim Verb.

Verb		
Wortbildungsart	Typische Regeln	Beispiel
Präfigierung	1. PRÄF+V 2. PRÄF+SBST 3. PRÄF+ADJ	<i>ver treten</i> <i>be flügeln</i> <i>ent schuldigen</i>
Konversion	1. SBST → V 2. ADJ → V	<i>planen</i> <i>sichern</i>
Lexikalische Verbindung	1. PRÄP/ADV+V 2. ADJ+V 3. SBST+V	<i>mit teilen, fort setzen</i> <i>fest stellen</i> <i>teil nehmen</i>
Suffigierung	1. SBST+SUFF 2. ADJ+SUFF	<i>bilanz ieren</i> <i>mobil isieren</i>

Tabelle 6: Wortbildungsregeln beim Adverb.

Adverb		
Wortbildungsart	Typische Regeln	Beispiel
Lexikalische Verbindung	1. ADV+PRÄP 2. ADV+ADV 3. PRÄP+PRON 4. PRÄP+PRÄP	<i>da bei</i> <i>wo anders</i> <i>zu dem</i> <i>zu vor</i>
Suffigierung	1. ADJ+SUFF 2. [ADJ+SBST]+SUFF 3. SBST+SUFF 4. ADV+SUFF, ADJ _{super} +SUFF	<i>bereit s</i> <i>neuer ding s</i> <i>teil s</i> <i>damal s, mindest ens</i>
Konversion	1. ADJ _{super} → ADV 2. SBST → ADV	<i>möglichst, längst</i> <i>morgen</i>
Präfigierung	1. PRÄF+ADJ _{super}	<i>zu vorderst</i>

ORCID®

Marek Konopka  <https://orcid.org/0000-0001-9627-0918>

Literatur

- Baayen, Harald R. 2001. *Word Frequency Distribution* (Text, Speech and Language Technology 18). Dordrecht: Springer.
- Baayen, Harald R. 2009. Corpus linguistics in morphology: Morphological productivity. In Anke Lüdeling & Merja Kytö (Hgg.), *Corpus Linguistics: An International Handbook* (Handbücher zur Sprach- und Kommunikationswissenschaft 29), Bd. 2, 899–919. Berlin: De Gruyter.
- Balnat, Vincent. 2011. *Kurzwortbildung im Gegenwartsdeutschen* (Germanistische Linguistik: Monographien 26). Hildesheim, New York & Zürich: Olms.
- Brandt, Patrick, Rolf-Albert Dietrich & Georg Schön. 2006. *Sprachwissenschaft: Ein roter Faden für das Studium der deutschen Sprache*, 2. Aufl. Köln: Böhlau.
- Bubenhofer, Noah, Selena Calleri & Philipp Dreesen. 2019. Politisierung in rechtspopulistischen Medien: Wortschatzanalyse und Word Embeddings. *OBST. Osnabrücker Beiträge zur Sprachtheorie* 95, 211–242.
- Deutsche Wortbildung 1 = Kühnhold, Ingeburg & Hans Wellmann. 1973. *Deutsche Wortbildung: Erster Hauptteil: Das Verb* (Sprache der Gegenwart 29). Düsseldorf: Schwann.
- Deutsche Wortbildung 2 = Wellmann, Hans. 1975. *Deutsche Wortbildung: Zweiter Hauptteil: Das Substantiv* (Sprache der Gegenwart 32). Düsseldorf: Schwann.
- Deutsche Wortbildung 3 = Kühnhold, Ingeburg, Oskar Putzer & Hans Wellmann. 1978. *Deutsche Wortbildung: Dritter Hauptteil: Das Adjektiv* (Sprache der Gegenwart 43). Düsseldorf: Schwann.
- Deutsche Wortbildung 4 = Ortner, Lorelies, Elgin Müller-Bollhagen, Hanspeter Ortner, Hans Wellmann, Maria Pümpel-Mader & Hildegard Gärtner. 1991. *Deutsche Wortbildung: Vierter Hauptteil: Substantivkomposita* (Sprache der Gegenwart 79). Berlin & New York: De Gruyter.
- Deutsche Wortbildung 5 = Pümpel-Mader, Maria, Elsbeth Gassner-Koch & Hans Wellmann. 1992. *Deutsche Wortbildung: Fünfter Hauptteil: Adjektivkomposita und Partizipialbildungen* (Sprache der Gegenwart 80). Berlin & New York: De Gruyter.
- Donalies, Elke. 2002. *Die Wortbildung des Deutschen* (Studien zur deutschen Sprache 27). Tübingen: Narr.
- Donalies, Elke. 2009. Stiefliches Geofaszintainment: Über Konfixtheorien. In Peter O. Müller (Hg.), *Studien zur Fremdwortbildung* (Germanistische Linguistik 197–198), 41–64. Hildesheim, New York, Zürich: Olms.

- Donalies, Elke. 2018. *Wetterbeobachter, Zeitlang, wahrsagen, zartfühlend, kurzerhand, dergestalt: Handbuch zur Univerbierung*. Heidelberg: Winter.
- Dudengrammatik. 2016 = Wöllstein, Angelika & Dudenredaktion (Hgg.). 2016. *Dudenband 4 – Die Grammatik*, 9. Aufl. Berlin: Dudenverlag.
- DWB = Grimm, Jacob & Wilhelm Grimm. 1854–1961. *Deutsches Wörterbuch von Jacob und Wilhelm Grimm*. 16 Bände in 32 Teilbänden. Quellenverzeichnis 1971. Leipzig: Hirzel. http://woerterbuchnetz.de/cgi-bin/WBNetz/wbgui_py?sigle=DWB (abgerufen am 6.6.2020).
- Eichinger, Ludwig M. 2000. *Deutsche Wortbildung: Eine Einführung*. Tübingen: Narr.
- Eisenberg, Peter. 2013. *Grundriss der deutschen Grammatik. Bd. 1: Das Wort*, 4. Aufl. Stuttgart & Weimar: Metzler.
- Evert, Stefan & Marco Baroni. 2005. Testing the extrapolation quality of word frequency models. In Pernilla Danielsson & Martijn Wagenmakers (Hgg.), *Proceedings from the Corpus Linguistics Conference Series*, Bd. 1. Birmingham: University of Birmingham. <https://www.birmingham.ac.uk/Documents/college-artslaw/corpus/conference-archives/2005-journal/Lexiconodf/EvertBaroniz2005.pdf> (abgerufen am 7.7.2020).
- Evert, Stefan & Marco Baroni. 2007. zipfR: Wordfrequency distributions in R. In Sophia Ananiadou (Hg.), *Proceedings of the 45th Annual Meeting of the Association for Computational Linguistics Companion Volume Proceedings of the Demo and Poster Sessions*. 29–32. Prag: ACL. <https://www.aclweb.org/anthology/P07-2008.pdf> (abgerufen am 7.7.2020).
- Evert, Stefan & Anke Lüdeling. 2001. Measuring morphological productivity: Is automatic preprocessing sufficient? In Paul Rayson, Andrew Wilson, Tony McEnery, Andrew Hardie & Shereen Khoja (Hgg.), *Proceedings of the Corpus Linguistics 2001 conference*, 167–175. Lancaster: UCREL.
- Finkel, Jenny Rose, Trond Grenager & Christopher Manning. 2005. Incorporating non-local information into information extraction systems by Gibbs sampling. In Kevin Knight, Hwee Tou Ng & Kemal Oflazer (Hgg.), *Proceedings of the 43rd Annual Meeting of the Association for Computational Linguistics (ACL'05)*, 363–370. Ann Arbor: Association for Computational Linguistics. <http://nlp.stanford.edu/~manning/papers/gibbscrf3.pdf> (abgerufen am 18.7.2020).
- Fleischer, Wolfgang & Irmhild Barz. 2012. *Wortbildung der deutschen Gegenwartssprache*, 4. Aufl. Berlin & Boston: De Gruyter.
- Gallmann, Peter. 1999. Fugenmorpheme als Nicht-Kasus-Suffixe. In Matthias Butt & Nanna Fuhrhop (Hgg.), *Variation und Stabilität in der Wortstruktur (Germanistische Linguistik 141–142)*, 177–190. Hildesheim, New York & Zürich: Olms.
- Grimm, Jakob. 1826. *Deutsche Grammatik*. Bd. 2. Göttingen: Dietrich.

- Haß, Ulrike & Klosa, Annette (2015): *elexiko – Geschichte und Zukunft. Sprachreport* 3. 10–20.
- HDK 1 = Pasch, Renate, Ursula Brauße, Eva Breindl & Ulrich H. Waßner. 2003. *Handbuch der deutschen Konnektoren: Linguistische Grundlagen der Beschreibung und syntaktische Merkmale der deutschen Satzverknüpfers (Konjunktion, Satzadverbien und Partikeln)* (Schriften des Instituts für deutsche Sprache 9). Berlin & New York: De Gruyter.
- Henzen, Walter. 1965. *Deutsche Wortbildung*, 3. Aufl. Tübingen: Niemeyer.
- Kopf, Kristin. 2018. *Fugenelemente diachron: Eine Korpusuntersuchung zu Entstehung und Ausbreitung der verfügbaren N+N-Komposita* (Studia linguistica Germanica 133). Berlin & Boston: De Gruyter.
- Leser, Martin. 1990. *Das Problem der ‚Zusammenbildungen‘: eine lexikalistische Studie* (Fokus 3). Trier: WVT.
- Meibauer, Jörg, Ulrike Demske, Jochen Geilfuß-Wolfgang, Jürgen Pafel, Karl Heinz Ramers, Monika Rothweiler & Markus Steinbach. 2007. *Einführung in die germanistische Linguistik*, 2. Aufl. Stuttgart & Weimar: Metzler.
- Motsch, Wolfgang. 1999. *Deutsche Wortbildung in Grundzügen* (Schriften des Instituts für deutsche Sprache 8). Berlin & New York: De Gruyter.
- Neef, Martin. 2015. Synthetic compounds in German. In Peter O. Müller, Ingeborg Ohnheiser, Susan Olsen & Franz Rainer (Hgg.), *Word-Formation: An International Handbook of the Languages of Europe* (Handbücher zur Sprach- und Kommunikationswissenschaft 40), Bd. 1, 581–592. Berlin: De Gruyter.
- Nübling, Damaris, Antje Dammel, Janet Duke & Renata Szczepaniak. 2017. *Historische Sprachwissenschaft des Deutschen: Eine Einführung in die Prinzipien des Sprachwandels*, 5. Aufl. Tübingen: Narr.
- Nübling, Damaris & Renata Szczepaniak. 2009. *Religion+s+freiheit, Stabilität+s+pakt und Subjekt(+s)+pronomen: Fugenelemente als Marker phonologischer Wortgrenzen*. In Peter O. Müller (Hg.), *Studien zur Fremdwortbildung* (Germanistische Linguistik 197–198), 197–222. Hildesheim, New York & Zürich: Olms.
- Ortner, Hanspeter & Lorelies Ortner. 1984. *Zur Theorie und Praxis der Komposita-forschung* (Forschungsberichte des Instituts für deutsche Sprache 55). Tübingen: Narr.
- Paul, Hermann. 1920. *Deutsche Grammatik. Bd. 5: Wortbildungslehre*. Halle a. S.: Niemeyer.
- Pfeifer, Wolfgang (Hg.). 1993–heute. *Etymologisches Wörterbuch des Deutschen.: Digitalisierte und v. Wolfgang Pfeifer überarbeitete Version im Digitalen Wörterbuch der deutschen Sprache*. Bearbeitet von Wilhelm Braun, Gunhild Ginschel, Gustav Hagen, Anna Huber, Klaus Müller, Heinrich Petermann, Gerlinde Pfeifer, Dorothee Schröter & Ulrich Schröter. <https://www.dwds.de/d/wb-etymwb> (abgerufen am 8.5.2020).

- Roth, Tobias. 2014. *Wortverbindungen und Verbindungen von Wörtern: Lexikografische und distributionelle Aspekte kombinatorischer Begriffsbildung zwischen Syntax und Morphologie* (Basler Studien zur deutschen Sprache und Literatur). Tübingen: Francke.
- Schäfer, Roland. 2015. Processing and Querying Large Web Corpora with the COW₁₄ Architecture. In Piotr Bański, Hanno Biber, Evelyn Breiteneder, Marc Kupietz, Harald Lungen, & Andreas Witt (Hgg.), *Proceedings of Challenges in the Management of Large Corpora (CMLC-3)*, 28–34. Mannheim: Institut für Deutsche Sprache.
- Schäfer, Roland & Felix Bildhauer. 2012. Building Large Corpora from the Web Using a New Efficient Tool Chain. In Nicoletta Calzolari, Khalid Choukri, Thierry Declerck, Mehmet Uğur Doğan, Bente Maegaard, Joseph Mariani, Asuncion Moreno, Jan Odijk & Stelios Piperidis (Hgg.), *Proceedings of the Eighth International Conference on Language Resources and Evaluation (LREC'12)*, 486–493. Istanbul: European Language Resources Association (ELRA).
- Schlücker, Barbara. 2012. Die deutsche Kompositionsfreudigkeit: Übersicht und Einführung. In Livio Gaeta & Barbara Schlücker (Hgg.), *Das Deutsche als kompositionsfreudige Sprache: Strukturelle Eigenschaften und systembezogene Aspekte* (Linguistik – Impulse & Tendenzen 46), 1–25. Berlin & Boston: De Gruyter.
- Schmid, Helmut. 1995. Improvements in Part-of-Speech Tagging with an Application to German. *Proceedings of the ACL SIGDAT-Workshop*. Dublin (Irland). <http://www.cis.uni-muenchen.de/~schmid/tools/TreeTagger/data/treetagger2.pdf> (abgerufen am 9.2.2020).
- Schmid, Helmut, Arne Fitschen & Ulrich Heid. 2004. SMOR: A German Computational Morphology Covering Derivation, Composition, and Inflection. In Maria Teresa Lino, Maria Francisca Xavier, Fátima Ferreira, Rute Costa & Raquel Silva (Hgg.), *Proceedings of the Fourth International Conference on Language Resources and Evaluation (LREC'04)*, 1263–1266. Lissabon: European Language Resources Association (ELRA). <https://www.cis.uni-muenchen.de/~schmid/papers/smor.pdf> (abgerufen am 18.6.2020).
- Waßner, Ulrich H. 2002. Geschlossene Klassen? In Reinhard Rapp (Hg.), *Sprachwissenschaft auf dem Weg in das dritte Jahrtausend: Akten des 34. Linguistischen Kolloquiums in Germersheim 1999. Teil II: Sprache, Computer, Gesellschaft*, 635–643. Frankfurt a. M., Berlin, Bern u. a.: Lang.
- Willems, Klaas. 2013. Funktionswort. In Stefan J. Schierholz & Pál Uzonyi (Hgg.), *Grammatik: Formlehre: Wörterbücher zur Sprach- und Kommunikationswissenschaft (WSK) Online*. Berlin & Boston: De Gruyter. https://db.degruyter.com/view/WSK/wsk_id_wsk_artikel_artikel_21552 (abgerufen am 18.6.2020).
- Wilmanns, Wilhelm. 1899. *Deutsche Grammatik: Gotisch, Alt-, Mittel- und Neuhochdeutsch*. Straßburg: Trübner.

Patrick Brandt

Bau von und Umbau zu Adverbien: Präpositionen, Vergleiche und Flexion

Abstract This chapter focuses on the formation of adverbs from a corpuslinguistic perspective, providing an overview of adverb formation patterns in German that includes frequencies and hints to productivity as well as combining quantitative methods and theoretically founded hypotheses to address questions that concern possible grammaticalization paths in domains that are formally marked by prepositional elements or inflectional morphology (in particular, superlative or superlative-derived forms). Within our collection of adverb types from the project corpus, special attention is paid to adverbs built from primary prepositions. The data suggest that generally, such adverb formation involves the saturation of the internal argument slot of the relation-denoting preposition. In morphologically regular formations with the preposition in final position, pronominal forms like *da* ‘there’, *hier* ‘here’, *wo* ‘where’ as well as *hin* ‘hither’ and *her* ‘thither’ serve to derive adverbs. On the other hand, morphologically irregular formations with the preposition – in particular: *zu* ‘to’ or *vor* ‘before, in front of’ – in initial position show traits of syntactic origin such as (remnants of) inflectional morphology. The pertaining adverb type dominantly saturates the internal argument slot by means of universal quantification that is part and parcel as well of the derivation of superlatives and demonstrably fuels the productivity of the pertaining formation pattern.

Keywords adverb formation, argument saturation, inflection, primary preposition, superlative, zero derivation

1 Ausgangspunkt und Ziel

Ziel des vorliegenden Kapitels ist es, ausgehend von tatsächlich vorkommenden Sprachdaten mittels Korpusanalyse Verfahren in der Wortbildung zu erhellen, die den adverbialen Gebrauch von Wörtern und damit so etwas wie die Wortart Adverb bestimmen. Vor dem Hintergrund gängiger Annahmen über argumentstrukturelle Repräsentationen zeigt der korpuslinguistische Zugang zahlreiche

Zusammenhänge von Form und Funktion (Bedeutung) auf, die in der bisherigen Forschung nur am Rande oder gar nicht behandelt worden sind.

Gegenstand der Betrachtung ist zunächst eine Liste von Wörtern aus einem Teil des Deutschen Referenzkorpus (DeReKo), die der TreeTagger als Adverb erkennt. Die formbasierte Ordnung und quantitative Auswertung der so erhaltenen Lemmata zielt einerseits auf eine empirisch fundierte Typologie der verschiedenen Bildungsweisen von Adverbien. Andererseits geben die vorgefundenen quantitativen und qualitativen Bezüge und Eigenschaften Anlass zu einer genaueren Sichtung, Erhebung und Analyse von sekundären Adverbien, die bestimmte Formelemente enthalten; dabei handelt es sich um präpositionale Formen, die den 20 primären (oder „alten“) Präpositionen des Deutschen entsprechen sowie um Formen aus dem Flexionsinventar des Deutschen, die als solche oder miteinander kombiniert interessante synchrone oder diachrone Bezüge zu den präpositionalen Formen oder sie als Kern enthaltenden syntaktischen Konstruktionen aufweisen. Die Natur dieser Bezüge veranlasst uns, zwischen wortbildungsregelgerechtem „Bau von“ und auf syntaktische Quellen rekurrierenden „Umbau zu“ Adverbien zu unterscheiden und liefert Hinweise auf mögliche Grammatikalisierungspfade im betrachteten Bereich.

1.1 Formale und funktionale Bestimmung der Wortart

Wie keine andere Wortart ist das Adverb funktional bestimmt: Gängigen Auffassungen zufolge sind Adverbien Wörter, die adverbial gebraucht werden (z. B. Helbig & Buscha 2001:306, Meibauer et al. 2007: 132). Anhand der ausführlicheren, auf Gebrauchsmöglichkeiten basierenden Charakterisierung von Schmöe (2002: 159) und Schäfer (2013: 19), die in der folgenden Liste (1. bis 4.) zusammengefasst wird, soll hier die mit der Kategorisierung als Adverb verbundene Problematik skizziert und ein Ausblick auf das Kommende gegeben werden:¹

1. Adverbien werden nicht flektiert.
2. Adverbien können als Satzglied fungieren.
3. Adverbien sind einzelne Wörter.
4. Adverbien können nicht als Subjekt fungieren.

1 Für oder gegen die Notwendigkeit der Unterscheidung von Lexemen (vs. Morphemen), möglichen (vs. tatsächlichen) Gebräuchen, lexikalischen (vs. syntaktischen) Wörtern oder bzgl. ihrer Wortartenzugehörigkeit unbestimmten Wurzeln und damit verbunden einer mehr oder weniger direkten Verbindung zwischen Laut- und Bedeutungsstrukturen wird hier nicht vorrangig argumentiert. Vgl. dazu Aronoff (1994: xiii), 9, Dudengrammatik (2009: 132–134), Halle & Marantz (1993) sowie unten Abschnitt 5.4.

Bezüglich 1 werden wir feststellen, dass Adverbien häufig kompariert werden bzw. komparierten Formen von Adjektiven entsprechen; insbesondere Superlativierung trägt offenbar zur Brauchbarkeit als Adverb bei (vgl. Abschnitt 5). Darüber hinaus enden zahlreiche Adverbien in Formen, die auch in der Flexion des Deutschen vorkommen (insbesondere: *-e*, *-en*, *-s*, *-st* und Kombinationen derselben wie *-stens*, s. Abschnitt 3.1 und 5.3), was eine wenigstens historisch gegebene Flektierbarkeit vorauszusetzen scheint. Das funktionale Kriterium 2 markiert die Grenze zwischen Adverb und Partikel und anderen in der Regel nicht selbstständig auftretenden Elementen wie Präpositionen; tatsächlich ist diese Grenze für viele Elemente in diesem Bereich nicht klar zu ziehen (vgl. Abschnitt 4). Kriterium 3 ist bezüglich sogenannter Zusammenrückungen (auch: Univerbierungen) problematisch; zum einen handelt es sich dabei um lexikalisierte Relikte mehr oder weniger direkt aus der Syntax übernommener, zu Wörtern umgebauter phrasaler Einheiten wie z.B. *vonnöten*, *imstande*, die auch flexivische Elemente beinhalten können (vgl. wiederum Kriterium 1); zum anderen sind Zusammenrückungen im Bereich superlativischer Bildungen, wie wir zeigen werden, auch synchron produktiv (Abschnitt 5.2). Das wiederum funktionale Kriterium 4 ist im Fall lokaler und temporaler Adverbien problematisch, die wenigstens in kopolativen Konstruktionen als Subjekte erscheinen können:²

(1) **Vorn** ist da, wo die Chinesen sind. (Berliner Zeitung, 22.3.2006, S. 16)

1.2 Adverbial gebrauchte Adjektive oder Adjektivadverbien

Die Abgrenzung von Adjektiven und Adverbien wirft grundlegende Probleme für die Wortartenklassifizierung auf (vgl. Telschow 2014). Entscheidend ist die Frage, ob Adjektive unter flexionslosem, weder attributivem noch prädiaktivem und mithin adverbialem Gebrauch zu Adverbien werden und also die Rede von „Adjektivadverbien“ rechtfertigen wie bei Helbig & Buscha (2001: 280, 306). Weiter verbreitet erscheint die Auffassung, dass Adjektive unter

- 2 Direktionale Ausdrücke spreizen die Grenze zwischen Argument und Adverbial.
- (i) illustriert Variation zwischen einer direktionalen Präpositionalphrase mit Akkusativ und einem Dativobjekt.
- (i) a. Sie könnten die Schüler direkt **an die Mittelschule** zuweisen. (Die Südostschweiz, 12.8.2010, o. S.)
- b. Sie habe den Jungen **der Grundschule** zugewiesen; dies sei für die Stadt Koblenz bindend. (Rhein-Zeitung, 27.8.2003, o. S.)

Im Englischen sind direktionale Präpositionalphrasen neben nominativischen Nominalphrasen die einzigen Konstituenten, die unter Lokativinversion die Position vor dem finiten Verb besetzen können, vgl. z.B. Bresnan (1994).

adverbialer Verwendung ihre Kategorie bewahren; es ist dann die Rede von „adverbial gebrauchten Adjektiven“ (z.B. Heidolph et al. 1981: 621–622, Eisenberg 2013: 227–228). Zur Annahme eines Wortartenwechsels passt die Annahme von Konversion, d.h., der Null-Ableitung von Adverbien aus Adjektiven. Interessant ist in diesem Zusammenhang, dass offenbar nur graduierbare und steigerbare Adjektive wie z.B. *schnell* oder *schön* als Adverbiale fungieren, nicht aber sogenannte absolute Adjektive wie z.B. *rund* oder *lebendig*, die nicht ohne Weiteres steigerbar sind.³ Die Bedingungen für Steigerbarkeit einerseits und Adverbierbarkeit andererseits haben also etwas miteinander zu tun. Graduierbare Eigenschaften taugen bei ansonsten geeigneter Semantik ohne Weiteres für adverbialen Gebrauch, d.h. solchen, der das bezeichnete Geschehen (s.u. Abschnitt 1.3) modifiziert (*Gianna läuft schnell*).⁴ Der nicht-attributive Gebrauch absoluter Adjektive ist hingegen in aller Regel prädikativ, d.h., es wird nicht die Art und Weise eines Geschehens modifiziert, sondern es wird einem der Argumente eine Eigenschaft zugesprochen:

- (2) Diese Tiere gehen lebendig in die entsprechenden Brühbehälter.
 ‚Diese Tiere sind lebendig, wenn sie in die Brühbehälter gehen.‘
 (Protokoll der Sitzung des Parlaments Deutscher Bundestag, 29.6.2012,
 o. S.)

Sofern absolute Adjektive doch adverbial verwendet werden, handelt es sich um metaphorischen Gebrauch:⁵

- (3) Ein Schulmuseum im Eifeldörfchen Immerath (be-)schreibt **lebendig** Geschichte.
 Nicht: ‚Ein Schulmuseum im Eifeldörfchen Immerath ist lebendig, wenn es Geschichte (be-)schreibt.‘ (Rhein-Zeitung, 12.7.1997, o. S.)

3 Mit Eisenberg (2013: 228) bezeichnet ein „Prototyp“ von Adjektiv „einfache“ Eigenschaften von Dingen; dieser ist „in seiner Grundbedeutung attributiv und prädikativ, nicht aber adverbial verwendbar [...]“

4 Über zugrundeliegende Relationalität hinaus spielt z.B. die Unterscheidung zwischen „stage-level“ und „individual-level“-Prädikaten eine Rolle (Milsark 1974, Carlson 1977), indem nur raumzeitlich kontingente (also „stage-level“) Prädikate sinnvoll auf Geschehen applizieren.

5 Vgl. wiederum Eisenbergs (2013: 228) Beispiele für die metaphorisch verschobene Interpretation einfacher Adjektive unter adverbialen Gebrauch:

- (i) a. Der Motor läuft rund.
 b. Karl denkt grün.

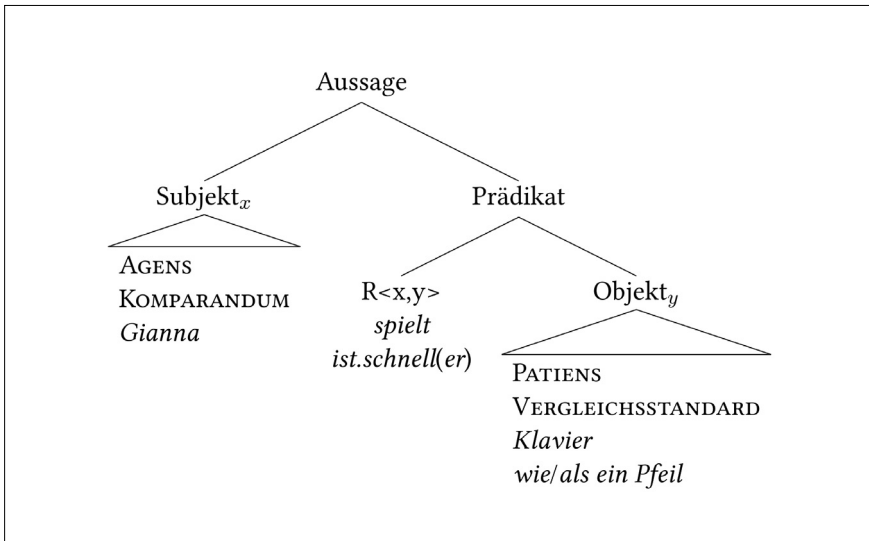


Abbildung 1: Externe und interne Argumente einer Relation.

Absolute Adjektive bezeichnen Eigenschaften von Individuen, während graduierbare Adjektive Individuen auf der durch sie induzierten Skala in Beziehung zu einem VERGLEICHSTANDARD setzen. Graduierbare Adjektive sind damit fundamental relational; sie teilen ihre logische Form mit zweistelligen Verben, die einem universellen semantischen Prinzip unterliegen: Sofern in der Beziehung ein AGENS- und ein PATIENS-Argument aufeinander bezogen werden, wird ersteres als grammatisches Subjekt und letzteres als grammatisches Objekt realisiert.⁶ Mit Blick auf die zugrundeliegende Konstituentenstruktur folgen wir hier der Sprechweise von relativ zum Prädikat „externen“ (AGENS, Subjekt) und „internen“ (PATIENS, Objekt) Argumenten. VERGLEICHSTANDARDS verhalten sich analog zu (internen) Patiens-Argumenten, die in ihr Prädikat inkorporieren (*Gianna ist pfeilschnell* bzw. *Gianna darf klavierspielen*) oder unausgedrückt bleiben können, im Gegensatz zu KOMPARANDA oder AGENS-Argumenten, die aufgrund ihrer Externalität syntaktisch durch Prädikation lizenziert werden müssen. Die Baumstruktur in Abb. 1 veranschaulicht die externe (als Subjekt) vs. interne (als Objekt) Realisierung der genannten thematischen Bausteine relativ zum Prädikat.

Wir nehmen zusammenfassend an, dass geschehensbezogenen Adverbien relationale Einheiten zugrundeliegen, die im Fall von Verben bzw. Adjektiven ein

6 Wir verstehen unter AGENS den belebten Verursacher einer Handlung und unter PATIENS eine von der Handlung betroffene und in der Regel durch sie veränderte Entität (vgl. Primus 1999).

AGENS- bzw. KOMPARANDUM-Argument und ein PATIENS bzw. VERGLEICHSTANDARD-Argument in Beziehung zueinander setzen. Allgemeineren grammatischen Prinzipien folgend werden diese strukturell extern (als Subjekt) bzw. intern (als Objekt) realisiert. Einen roten Faden für die vorliegende Untersuchung bildet der für die Herstellung von Adverbien wesentliche Umbau von Relationen zu Eigenschaften: Die Sättigung der internen Argumentstelle des relationalen Ausdrucks (einer Funktion von Paaren von Individuen in Wahrheitswerte) bringt ein einstelliges Prädikat, d. h. eine Eigenschaft (eine Funktion von Individuen in Wahrheitswerte) hervor, die dann von einem Geschehen (s. u. Abschnitt 1.3) präzisiert werden kann. Wie wir sehen werden, kommen als relationale Einheiten über Verben und Adjektive hinaus vor allem Präpositionen in Frage.⁷

1.3 Externe und interne Rollen: Partizipien

Die Relevanz der Unterscheidung zwischen extern und intern realisierten semantischen Rollen für den Bau von Adverbien wird bei der vergleichenden Betrachtung von Partizipien I versus Partizipien II deutlich. Beide Typen können wie Adjektive gebraucht werden, wobei sie bei zugrundeliegenden transitiven Verben eindeutig auf semantische Rollen bezogen sind: Partizipien I beziehen sich auf das AGENS (externe Rolle, 4a) und Partizipien II auf das PATIENS (interne Rolle, 4b).

- (4) a. Ebernbург spielte **enttäuschend**. (Rhein-Zeitung, 21.2.2000, o. S.)
 b. Berlins Mineiro (Nr. 7) dreht **enttäuscht** ab. (Berliner Zeitung, 29.1.2008, o. S.)

In (4a) enttäuscht Ebernburg (AGENS bzw. Subjekt, externe Rolle), es bleibt unausgedrückt, wer enttäuscht wird (PATIENS bzw. Objekt, interne Rolle). In (4b) ist bzw. wurde Mineiro enttäuscht (PATIENS bzw. Objekt, interne Rolle), wer oder was ihn enttäuscht, bleibt unausgedrückt (AGENS bzw. Subjekt, externe Rolle). Partizipien I können ohne Weiteres adverbial gebraucht werden, Partizipien II nicht.

Im Anschluss an Davidson (1967) bestimmen wir Adverbien – unter Ausklammerung satzkommentierender Adverbien wie *leider* oder *erfreulicherweise* – als Ausdrücke für Eigenschaften von Geschehen. Propositionen können wahr oder falsch sein, Geschehen dagegen „haben statt“, „passieren“,

7 Die hier im Fokus stehenden sogenannten primären Präpositionen realisieren in ihrer lokativen Funktion das Locandum-Argument extern und das Locatum-Argument intern (vgl. die Abschnitte 3.2 und 4.1).

„verlaufen“ oder „ereignen sich“. Davidson spricht von „action sentences“ und hat damit offenbar vor allem spezifisch durch AGENS-Argumente verursachte Geschehen im Sinn. Die Paraphrase von (4a) in (5) bringt die Nähe zwischen AGENS-Argumenten und den durch sie in Gang gebrachten Geschehen heraus; eine solche Paraphrase steht für (4b) nicht zur Verfügung. Vielmehr ist das enttäuschende Geschehen, dessen PATIENS Mineiro ist, gar nicht Teil des propositionalen Gehalts von (4b), sondern z. B. in einem vorhergehenden Satz kodiert oder irgendwie inferierbar.

- (5) Es gibt ein Spielgeschehen e, das AGENS von e ist Ebernburg und Ebernburgs Agieren ist enttäuschend.

Partizipien liefern Hinweise darauf, dass AGENS-orientierte Prädikate sich unter Unterschlagung des PATIENS-Arguments ohne Weiteres in Eigenschaften von (agentivischen) Geschehen umbauen und mithin als Adverbien gebrauchen lassen; Adverbien sind also verwandt mit AGENS-orientierten Prädikaten. Umgekehrt gibt es keinen unmarkierten Weg, PATIENS-orientierte Prädikate unter eventueller Unterschlagung des AGENS-Arguments in Eigenschaften von Geschehen umzubauen und mithin als Adverbien zu gebrauchen.

2 Korpus, Suchanfragen und Methoden

Aus einem Achtel des Untersuchungskorpus (ca. 900 Mio. Tokens) wurden sämtliche als Adverb getaggten Wortformen extrahiert (TreeTagger, STTS). Der so gewonnene Datensatz ADVERBIEN_1 enthält 3.753 unterschiedliche Typen (Lemmata) mit über 40 Mio. Tokens.⁸ Die Daten wurden nachträglich nach Suffix (29 Klassen) und Bauform (primäre, flexivische, präpositionale und suffixale

8 Eine formbasierte oder negativ taggingbasierte Extraktion mit dem Ziel einer höheren Quote der berücksichtigten Adverbien aus der Gesamtheit aller tatsächlichen Adverbien im Korpus ist aufgrund der Abwesenheit des Ausdrucks umgebungsabhängiger morphosyntaktischer Merkmale an Adverbien (z. B., Kongruenz) bei der gleichzeitigen Vielfalt einzelner Adverbmarker und gegebenenfalls deren Multifunktionalität (z. B. des -s oder -en) hier nicht darstellbar. Im Bereich der genauer betrachteten präpositionalen und flexivischen Formen wurden ergänzende formbasierte Suchen durchgeführt und die Typen händisch auch hinsichtlich ihrer Konstituentenstruktur überprüft; durch den Herausfall von Formen, deren Letztglied bereits ein Adverb ist (z. B. *geradeheraus*) reduzierte sich die Zahl der Typen dabei nur leicht (um jeweils 3 bei den Typen mit Präpositionen als Letztglied bzw. Erstglied). Schreibvarianten (z. B. *-ermaßen* vs. *-ermassen*) sowie offensichtliche Fehlschreibungen wurden nach Möglichkeit zusammengefasst bzw. aussortiert, allerdings nicht in der großen Gruppe der *mal*-Typen, in der Schreibungen und Frequenzen stark variieren.

Typen) annotiert, womit sich 94 % des Bestandes erfassen ließen (zum Überblick s. Tab. 1). Nicht kategorisiert wurden z.B. Adverbien, deren Letztglied ein freies Morphem ist (z.B. *kurzerhand*, *nunmehr*, *gleichwohl*), konvertierte Formen wie *gemach* oder *getrost*, Lehnwörter (z.B. *brutto*, *unisono*, *offshore*) sowie Abkürzungen (z.B., *ca.*, *evtl.*).

Für eine Detailstudie (Abschnitt 3.3) wurden außerdem im selben Korpusachtel wortformbasiert sämtliche Bildungen auf *-erweise* und *-ermaßen* extrahiert, sie bilden den Datensatz ADVERBIEN_2. Grund für den Verzicht auf das Tagging war der bessere Recall, d. h., eine höhere Quote der aus allen tatsächlich validen Belegen korrekt als solchen erkannten Belegen: Die Zahl der Typen steigt verglichen mit den taggingbasierten Zahlen bei Bildungen mit *-erweise* von 756 auf 873 an, bei Bildungen mit *-ermaßen* von 67 auf 77; die Zahl der Hapaxlegomena steigt auf 405 (von 308) bei *-erweise* und auf 30 (von 22) bei *-ermaßen*.

3 Beschränkt heterogene Bildungsweisen

Tabelle 1 zeigt die Verteilung der Formen nach ihrer Bildungsweise.⁹ Wir unterscheiden neben tendenziell primären Wörtern, die mit 64 Typen etwa 2 % des Bestands ausmachen, drei übergeordnete, zum Teil aufeinander bezogene sekundäre Bauformen:

- „Flexivische“ Typen, die auf in der Flexion des Deutschen vorkommende Formative enden (287 oder 8 % der Typen insgesamt).¹⁰

9 Um bei Schreibungen wie <1.>, <2.>, ... <174.> den wahrscheinlichen Anteil an aus Ordinalzahlen abgeleiteten Adverbien (*erstens*, *zweitens*, *hundertvierundsiebzigstens*) zu ermitteln, wurde eine Stichprobe von 250 Belegen genommen und hochgerechnet. Insgesamt gibt es im Korpus 1024 solcher Schreibungen (davon 100 Hapaxlegomena), die mit einer durchschnittlichen Tokenzahl von 208,9 insgesamt 213.905 Tokens liefern. Die manuelle Auszählung ergibt, dass 6,4 % davon Ordinalzahlen, also *stens*-Fälle sind. Extrapoliert ergeben sich 66 zusätzliche *stens*-Typen, 13.787 zusätzliche Tokens und sechs zusätzliche Hapaxlegomena (jeweils gekennzeichnet mit einem Plus-Zeichen).

10 Die in Frage stehenden Formen haben wohl eine ehemalige flexionsbezogene Funktion verloren (z.B.: Anzeige von Abhängigkeit durch Kasus durch *-e* oder *-en*). Das heißt m. E. jedoch nicht, dass die in Frage stehenden Elemente im eigentlichen Sinne funktionslos sind, denn sie markieren ja qua Argument (wiewohl möglicherweise unter anderem) Adverbien. Die eindeutig Adverbien markierende und synchron produktive *-stens* Form (z.B.: *hundertsiebenundachtzigstens*, *hundertachtundachtzigstens*, *allerallerschnellstens* ...) hat regulär elativische, also eindeutig semantische Funktion (vgl. jedoch Abschnitt 5.2). Die Diskussion in Abschnitt 5 deutet darauf hin, dass die Komparation und insbesondere Superlativierung markierenden Formen als solche Adverbierung begünstigen und also auch in diesem (neuen) Sinne funktional sind.

Tabelle 1: Bauformen, Typen, Tokens, Hapaxlegomena und TTRs (Type-Token-Verhältnisse).

		Typen	Tokens	Hapaxlegomena	TTR×10⁶
primär		64	26.652.241	0	2
flexivisch	-stens	60	287.606	6	209
		+ 66	+ 13.787	+ 6	418
	-s	74	1.670.678	10	44
	-en	49	396.643	2	124
	-st	33	1.086.391	4	30
	-ens	5	58.883	0	85
	-e	31	486786	2	64
präpositional	-präp	244	1.052.093	23	232
	präp-	181	2.423.535	7	75
suffixal	-mal	1.201	726.604	609	1.653
	-erweise	756	111.320	308	6.791
	-wärts	251	32.093	109	7.821
	-seits	158	88.066	78	1.794
	-ermaßen	67	33.102	22	2.024
	-hin	49	330.930	7	148
	-tags	34	109.977	0	318
	-lich	38	1.203.809	0	32
	-her	33	411.661	4	80
	-weise	28	180.371	0	155
	-falls	27	312.041	6	87
	-wegs/en	22	106.388	5	207
	-orts/en	18	13.021	2	1.382
	-mals	14	411.840	0	34
	-teil(s)	14	36.158	2	387
	-willen	10	123	3	81.301
	-zeit(en)	9	285.842	0	31
	-lings	6	907	0	9.923
	-dings	6	360.075	3	17
	-halben	6	1.957	2	3.066
-halber	5	3.819	0	1.309	
Summe		3.528	38.401.961	1.218	92
ADV gesamt		3.753	41.223.616	1.228	91

- „Präpositionale“ Formen, die auf unabhängig (d.h. in nicht adverbierender Funktion) vorkommenden präpositionalen Elementen aufbauen (425 oder 11 % der Typen insgesamt).¹¹
- „Suffixale“ Bildungen, die aus der Anfügung meist aus Substantiven abgeleiteter Suffixe wie *-mal*, *-(er)weise*, *-wärts* oder *-teil(s)* hervorgehen. Mit 73 % machen sie den Großteil (2.752) des Gesamtbestands an Typen aus.

Es ist ersichtlich, dass die Bildungsweisen von Adverbien – verglichen etwa mit den Produkten von Derivation und erst recht Komposition im Bereich von Substantiven – nicht sonderlich komplex sind. Mehrfachableitung, d.h. die Einbettung bereits komplexer Formen, kommt praktisch nur insoweit vor, wie bestimmte Bestandteile von Adverbien innerhalb ihrer (Hauptwortart-) Kategorie bereits komplex sind (z.B. *aufeinander-zu*, *braunkohletagebauabbruchkanten-wärts*). Adverbien sind in diesem Sinne im Kontrast zu Substantiven oder Adjektiven nicht-rekursiv, d.h., ein einmal als Adverb kategorisiertes Element taugt in aller Regel nicht zu weiteren Ableitungen (vgl. aber Abschnitt 5.2).

3.1 Flexion oder Suffigierung

Flexion gehört per Definition nicht zur Wortbildung, indem sie – mit Ausnahme intuitiv klar bedeutungstragender Plural- oder Komparativmorphologie – erst durch die Einsetzung lexikalischer Einheiten in einen bestimmten syntaktischen Kontext bestimmt wird. Die Tatsache, dass ein nicht geringer Teil von Adverbtypen auf Formen endet, die auch in der Flexion des Deutschen vorkommen, deutet darauf hin, dass der Bestand an Adverbien sich in nicht geringem Maße aus der Syntax speist. Wir unterscheiden zwischen flexivischen Formen, die bei Adverbien häufig sind, aber nicht unbedingt eindeutig Adverbien ableiten und Formen, die eindeutig zur Kategorisierung als Adverb führen.

3.1.1 Bauweisen mit *-e*, *-en* und *-st*

Von den 80 Typen, die auf *-e* oder *-en* enden, enthalten 37 (46 %) eine Präposition als Erstglied, die in der entsprechenden syntaktischen Konstruktion für das Kasusmorphem verantwortlich ist. Das kasusmarkierte Element trägt zur Sättigung einer der Argumentstellen der Präposition bei. Wir haben im Einzelnen:

11 Wie im einleitenden Kapitel bemerkt, kommen präpositionale Elemente ansonsten praktisch nur im verbalen Bereich als sog. Verbpartikeln vor.

- (6) a. *insbesondere, beinahe, zustande, zugute, zuhause, zugrunde, zunichte, beiseite, beileibe, imstande, zumute, zunutze, außerstande, zuwege, zuleide, zuhilfe*
 b. *zusammen, ansonsten, bisweilen, zuweilen, abhanden, mitnichten, vonstatten, vonnöten, beisammen, zuhanden, zuschanden, außen, mitsammen, vondannen, ohnesorgen, insgleichen, vonnutzen*

Die Häufigkeit der Endung *-st* unter den als Adverbien getaggtten Typen ist auffällig. Von 49 einschlägigen Typen sind 33 eindeutig Adverbien, d. h., sie kommen im Korpus nicht weitergehend flektiert vor (z. B. *vorerst, demnächst, dereinst* oder *tunlichst*). Von diesen 33 Typen beginnen 20 mit der präpositionalen Form *zu* (z. B. *zuletzt, zunächst, zuoberst*). Wenn *zu* Erstglied ist, liegt eindeutig ein Adverb vor, vgl. (7) und Abschnitt 5.2.

- (7) a. Das (*zu)oberste Brett
 b. *(Zu)oberst liegt ein Brett

3.1.2 Bauweisen mit *-s*

Besonders frequent unter den flexivischen Elementen im Bereich der Adverbien ist suffigiertes *-s*, dem traditionell zumeist genitivischer Ursprung zugeschrieben wird (vgl. z. B. Paul 1920: 129 und Beispiele für den sogenannten adverbialen Genitiv wie *eines Tages, des Wegs, meines Erachtens*). Daneben speist sich das für Adverbierung typische *-s* aus verschiedenen Quellen. Erstens endet ein guter Teil der adverbaleitenden Suffixe auf *-s*:

- (8) *himmelwärts, mancherorts, keinesfalls, seinerseits*

Zweitens tritt *-s* zu Formen, die schon Adverbien waren oder ganz überwiegend adverbial gebraucht wurden:¹²

- (9) *bereits, öfters, insbesondere, jeweils*

12 Potentiell einschlägig sind auch die sich aus verschiedenen Quellen speisenden Formen *mitschiffs, vollends, nirgends, zusehends* sowie *eilends*. Nach dem DWB (1854–1961) ist bei den letzten drei Typen *-d-* sekundär eingefügt; Paul (1897) ordnet *eilends* als Partizip I mit Genitiv-*s* ein. Die Nähe von Partizipien I zu Adverbien ist Gegenstand von Abschnitt 3.3.

Drittens kann -s gegen Präpositionen in adverbial gebrauchten Präpositionalphrasen getauscht werden:¹³

(10) -s ADV	PP ADV
<i>anfangs, flugs</i>	<i>am Anfang, im Fluge</i>
<i>übrigens, rechtens</i>	<i>im Übrigen, im/zu Recht</i>
<i>mindestens, schärfstens</i>	<i>im Mindesten, am schärfsten</i>
<i>erstens ... siebzigstens</i>	<i>zum Ersten ... Siebzigsten</i>

Namentlich die synchron produktiven Bildungen aus Superlativformen (*schärfstens*, *mindestens*) haben Entsprechungen mit den präpositionalen Elementen *im*, *am* oder *zum*. Unter Abzug der nicht-flexivischen Suffixbildungen auf -s (-wärts, -seits, -orts etc., s. 9) ist die Verteilung der flexivisch endenden Formen wie folgt (wiederholt aus Tab. 1, s. dort für Erläuterungen).

Tabelle 2: Flexivische Typen, Tokens, Hapaxlegomena und TTRs.

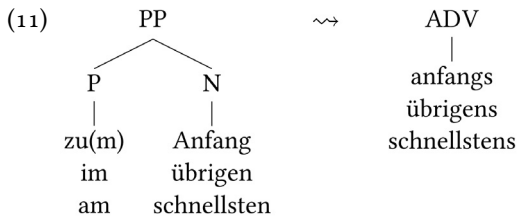
	Typen	Tokens	Hapaxlegomena	TTR×10 ⁻⁶
-stens	60	287.606	6	209
	+ 66	+ 13.787	+ 6	418
-s	74	1.670.678	10	44
-en	49	396.643	2	124
-st	33	1.086.391	4	30
-ens	5	58.883	0	85
-e	31	486.786	2	64

-stens liefert also von allen flexivisch endenden Formen die meisten Typen und hat auch die höchste Type-Token-Ratio, die als ein erster Anhaltspunkt für höhere Produktivität gesehen werden kann. Wilmanns (1896: 610) hat zu Recht die Besonderheit des Superlativs, in prädikativer Position nicht „nackt“, sondern in Begleitung einer mit dem definiten Artikel verschmolzenen präpositionalen Form zu erscheinen (*am schnellsten*), für den Sonderweg verantwortlich gemacht, den der Superlativ im Bereich der Morphosyntax und Wortbildung verglichen mit dem Positiv und Komparativ genommen hat.¹⁴ Das Vorhandensein des

13 In ähnlicher Weise sind adverbiale Nominalphrasen im Genitiv (= adverbiale Genitive) oft gegen Präpositionalphrasen austauschbar, vgl. z.B. *eines Abends* – *an einem Abend*; *meines Wissens* – *nach meinem Wissen*.

14 In Richtung der Vermutung, dass der Wortartenwechsel zwischen Adjektiven und Adverbien durch einen Ableitungsprozess (Suffixierung bzw. Konversion) geregelt ist bemerkt Wilmanns (1896: 610):

präpositionalen Elements eröffnet für den Superlativ die Möglichkeit der Teilhabe an breit applizierenden Bildungsmustern, die Konstruktionen mit bestimmten Präpositionen mit flexivischen Wortbildungsprodukten verbinden; in diesem Zusammenhang einschlägige Prozesse reichen zum Teil weit zurück (im Fall des Umbaus von Fügungen wie *im günstigen/besten Fall* zu *günstigen-/bestenfalls*), zum Teil sind sie rezent (wie im Fall von *anbetrachts* aus *in Anbetracht*, analog zu *angesichts* aus *im Angesicht* oder das angesichts der Herkunft der Belege offenbar in der Schweiz geläufige *auftrags*, aus *im Auftrag*). (11) veranschaulicht den Umbau von adverbialen Präpositionalphrasen zu auf -s endenden Adverbien.



Die Superlativformen exemplifizieren ein typisches Grammatikalisierungsmuster, indem Abbau der Markierung (Präposition gegen -s) mit semantischer Schwächung einhergeht: Die Bedeutung wird von der Quantifikation über Individuen in Vergleichen zu einer nurmehr elativischen Modifikation von Geschehen verschoben. Die *s*-Suffigierung liefert ein eindeutiges Adverb, das im Standarddeutschen nicht prädikativ erscheinen kann:

(12) Ben war am schnellsten/*schnellstens. SCHNELLER.ALS(Ben,jeder)

In Richtung der Davidson'schen Intuition (vgl. Abschnitt 1.3) denotieren Subjekte von marginal prädikativ erscheinenden *stens*-Adverbien Geschehen, vgl. den folgenden grammatisch fragwürdigen Beleg aus einem Internetforum:

(13) Die Lieferung war schnellstens [...]. SEHR.SCHNELL(die Lieferung)
(<https://www.ebay.de/usr/ctec48flo>, abgerufen am 4.12.2019)

Im Positiv und Comparativ führte bei den meisten Adjectiven die Lautentwicklung zum Zusammenfall des Adverbiums und des prädicativen Adjectivs, im Positiv durch die Apokope der Adverbial-Endung, im Comparativ durch Apokope der Adjectiv-Endung; im Superlativ dagegen, wo sich das Adjectivum gern mit dem Artikel verbindet, blieben eben dadurch die beiden Formen unterschieden und das Adverbium ging seine eigene Bahn. Schliesslich freilich wirkte der Mangel an Unterscheidung im Positiv und Comparativ auch auf den Superlativ, und zwar in der Weise, dass das mit *am* gebildete Adverbium das Adjectivum zurückdrängte.

3.2 Präpositionen als Bausteine

Wenn man bedenkt, dass Präpositionen in der Syntax die prototypischen adverbialen Relatoren sind und die Syntax eine wichtige Quelle für sekundäre Wörter darstellt, überrascht es nicht, dass mittels Präpositionen gebildete Adverbien eine Gruppe von beträchtlicher Größe ausmachen.¹⁵ Tatsächlich werden vornehmlich als Präpositionen kategorisierte Einheiten mitunter als Adverbien analysiert, da sie insbesondere in Bildungen mit dem Pro-Adverb *da* (sog. Präpositional- oder Pronominaladverbien) mit eindeutig als Adverb zu kategorisierenden Einheiten konkurrieren und hier auch „abgespalten“ für sich stehen können (Krause 2007: 476):¹⁶

- (14) a. d(a)r-in – dr-innen, d(a)r-aus – dr-außen
 b. Davon hält sie nichts – da hält sie nichts von.

Darüber hinaus können einzelne Präpositionen, ihre Nähe zu Adverbien unterstreichend, auch ohne Ergänzung wie Adverbien gebraucht werden (vgl. Abschnitt 4.1).¹⁷ Präpositionen erscheinen in Adverbbildungen als Erst- oder Letztglieder; ihrer Nähe zu Adverbien und allgemeinen Prinzipien der Wortbildung gemäß – kategorien- und bedeutungsbestimmende Elemente stehen rechts – überwiegen jedoch Bildungen mit nachgestellten Präpositionen wie *darum*, *weitaus* oder *zweifelsohne*. Die ganz überwiegende Zahl der Bildungen basiert auf den 20 sogenannten primären Präpositionen:

- (15) *ab, an, auf, aus, bei, durch, für, gegen, hinter, in/ein, mit, nach, neben, über, um, unter, von, vor, zu, zwischen*

15 Die 37 Typen, die Präpositionen sowohl als Erst- wie als Letztglieder enthalten (z. B. *durchaus, mitunter, voran*), wurden zu den Bildungen mit Letztgliedern geschlagen. Bildungen mit Präpositionen bzw. mit ihnen gleichlautenden Einheiten in Mittelstellung wie *allzuviel, sozusagen* wurden nicht gezählt. Die Form *zu* fungiert in *allzuviel* als Gradpartikel, in *sozusagen* als Infinitivpartikel. Diachron hängen diese Formen mit der Präposition *zu* zusammen, die funktionale Abspaltung liegt allerdings weit vor der Bildung von *allzuviel* und *sozusagen*. Vgl. zu möglichen semantischen Gemeinsamkeiten der verschiedenen *zu*-Formen Brandt (2019).

16 Geuder (2019:222) beschreibt Präpositionaladverbien als durch Sättigung der Argumentstelle der Präposition abgeleitete Adverbien, sieht sie jedoch als isolierte Fälle. Die Betrachtung von Adverbien als intransitive Präpositionen ist in der generativen Grammatik nicht unüblich, vgl. Lee (1999).

17 Parallel dazu können semantisch transparente Verbsätze in sog. Partikelverben auch prädikativ gebraucht werden: *die Tür zumachen* > *die Tür ist zu*; vgl. dagegen *mit etwas aufhören* > **es ist auf*.

Einzig die primären Präpositionen sind an der Bildung von Präpositionaladverbien beteiligt (*daran, darauf, ... dazu*). Präpositionaladverbien machen den für Adverbierung wesentlichen Umbau von Relationen zu Eigenschaften anschaulich (vgl. 4.1): Eine der Proformen *da, hier* oder *wo* realisiert das interne Argument der Präposition formal, stellt aber anaphorisch oder deiktisch Bezug zu etwas anderem her; textfunktional sind also nach wie vor Relationen gegeben.¹⁸ Das Verbauen von Proformen ist in der Wortbildung ansonsten unüblich, eröffnet allerdings attraktive Möglichkeiten: So können Präpositionaladverbien insbesondere Bezüge zwischen propositionalen Gehalten herstellen, deren Trägerausdrücke als direkte Ergänzungen von Präpositionen nicht in Frage kommen (vgl. **Daher rechne er mit, dass ...*). Unter Austausch von *da* (bzw. dem selteneren *hier*) durch *wo* kehren sich dabei die Einbettungsverhältnisse zwischen einbettender und eingebetteter Proposition um:¹⁹

- (16) a. Daher rechne er damit, dass Ende Oktober tatsächlich produziert werden kann. (die tageszeitung, 1.10.2007, o. S.)
 b. Das Bruttoinlandsprodukt hat sich zum Vorquartal um 0,7 Prozent erhöht, womit Experten gerechnet hatten. (dpa, 20.10.2006, o. S.)

Die Bildungen mit den Proformen *da, hier* und *wo* stellen den regelmäßigsten Typ der Bildungen mit Präpositionen überhaupt dar, indem 19 der 20 primären Präpositionen (Ausnahme: *ab*) daran in gleicher Weise, wenn auch unterschiedlich frequent beteiligt sind. Die niedrigen Type-Token-Ratios in Tab. 1 deuten schon darauf hin, dass der Hauptanteil der Bildungen mit Präpositionen auf vergangener Produktivität beruht. Dabei bilden Formen mit präpositionalen Letztgliedern eine gegenüber Bildungen mit präpositionalen Erstgliedern insgesamt homogenere Gruppe (vgl. Abschnitt 4.2). Abschnitt 4 bietet eine eingehendere, binnendifferenzierte Darstellung der Bildungen mit primären Präpositionen.²⁰

18 Die Formen *da, hier* und *wo* sind jeweils semantisch bzw. pragmatisch unmarkiert in dem Sinne, dass ihr außersprachlicher Bezug mehr oder weniger garantiert ist: *da* ist die in der Dimension der Nähe bzw. Ferne zur sprechenden Person gegenüber *dort* bzw. *hier* neutrale Form, *hier* kann immer mit dem Äußerungsort identifiziert werden. *Wo* fungiert als W-Form mit lokaler Grundbedeutung auch als „universelles Relativadverb“ (Eisenberg 2013: 277), was plausiblerweise mit der Allgemeinheit seiner Bedeutung und dem Metaphorisierungspotential lokaler Ausdrücke zusammenhängt.

19 Dieselben 20 primären Präpositionen kombinieren als Erstglieder mit *einander*; eingeschränkt ist daran auch die Präposition *ohne* beteiligt.

20 Neben den primären Präpositionen finden vor allem *ohne* und *seit* Eingang in komplexe Wortbildungstypen, außerdem selbst schon komplexe präpositionale Formen wie *zufolge*.

3.3 Externe und interne Rollen: Suffixbildungen mit *-erweise* und *-ermaßen*

Der Bau von Adverbien mittels Suffixen wird an anderer Stelle behandelt und steht hier nicht im Mittelpunkt (s. auch Fleischer & Barz 2012: 366–370, Heinle 2004, Ronca 1975). Die korpuslinguistische Betrachtung bestimmter Ausschnitte erhellt jedoch hier zentral stehende Zusammenhänge und liefert Hinweise zur Produktivität einzelner Bausteine; dazu werden im Folgenden die Suffixe *-erweise* und *-ermaßen* genauer miteinander verglichen. Grundlage dieses Teils stellt der Datensatz ADVERBIEN_2 dar (s. Abschnitt 2).

Im Hinblick auf einen für die Adverbierung erforderlichen relationalen Input ist bemerkenswert, dass ein großer Teil der nominalen Basen der adverbableitenden Suffixe relational ist: *x-wärts* bedeutet, dass die Richtung von etwas *x* ist; *x-weise*, dass die Weise von etwas *x* ist, *x-orten*, dass der Ort von etwas *x* ist. Die Basis *x* sättigt in diesen Bildungen die interne Argumentstelle der Beziehung und stellt damit eine Eigenschaft her, die als Prädikat eines Geschehens fungieren kann.

Weiteres Licht auf die relevanten Strukturen wirft nun der Vergleich von Suffixbildungen mit *-erweise* und *-ermaßen*. Wie oben in Abschnitt 1.3 diskutiert, sind Partizipien I durch ihren AGENS-Bezug für adverbialen Gebrauch prädestiniert. Bildungen aus Partizipien I und *-erweise* haben oft pleonastischen Charakter (vgl. Elsner 2016: 215, Ronca 1975), d. h., die Bildung eines Partizips I macht die *erweise*-Suffixierung für den adverbialen Gebrauch eigentlich überflüssig. Der Abzug des *erweise*-Letztglieds im repräsentativen Beispiel (18) etwa würde kaum ins Gewicht fallen.

- (17) So wurde man **sitzenderweise** Zeuge einer Materialschlacht, aus O-Tönen und Kamerafahrten, die um sich selber kreiste. (die tageszeitung, 23.9.1991, S. 18)

Partizipien II müssen dagegen für den adverbialen Gebrauch markiert werden.²¹ In aller Regel geschieht dies durch *-erweise* (43 Typen) oder insbesondere durch das auf Partizipien II spezialisierte Suffix *-ermaßen*: 60 von insgesamt 77 Typen oder 78 % der Bildungen mit *-ermaßen* haben Partizipien II als Basen, wobei sich die Ausnahmen einerseits bei „originellen“ Typen mit wenigen Tokens finden

21 Die im Korpus vorkommenden adverbial nutzbaren „nackten“ Partizipien II *bestimmt*, *ungelogen* sowie das komplexere *saisonbedingt* beziehen sich als Adverbien auf Propositionen und nicht auf Geschehen, d. h., sie modifizieren die Art des Gegebenseins propositionaler Gehalte und nicht die Art (und Weise) von Geschehen. Wie die meisten Partizip-II-*ermaßen*-Wortbildungen und im Kontrast zu Partizip-II-*erweise*-Wortbildungen können sie z. B. als Antworten auf Satzfragen fungieren.

(z.B. *pikantermaßen*, *stolzermaßen* mit je einem Token) und andererseits bei offenbar alten, hochfrequenten Typen mit sehr allgemeinbedeutenden, funktionswortartigen Basen (z.B. *gleichermaßen*: 11.220 Tokens; *einigermaßen*: 8.935 Tokens; *dermaßen*: 2.663 Tokens). Bekanntlich beziehen sich Partizipien II auf interne (PATIENS-)Argumente; dass dies ebenso bei Bildungen mit *-ermaßen* der Fall ist zeigt sich auch daran, dass es hier keine inkorporierten Objekte gibt, wie sie bei Bildungen aus Partizipien I und *-erweise* durchaus vorkommen (z. B. *ehrenrettenderweise*, *erbsenzählenderweise*). Außer in *folgendermaßen* (1.622 Tokens) und *zwingendermaßen* (5 Tokens) kombiniert *-ermaßen* nicht mit Partizipien I; pleonasmusverdächtige Bildungen aus Partizipien I und *-erweise* sind mit 142 Typen dagegen recht häufig. Abb. 2 vergleicht die Produktivität der beiden Suffixe anhand von Wortschatzwachstumskurven, die den Suffixbildungen mithilfe des Softwarepakets *zipfR* unter Anwendung des finiten Zipf-Mandelbrot-LNRE-Modells (vgl. Evert & Baroni 2007) zugeordnet werden können.

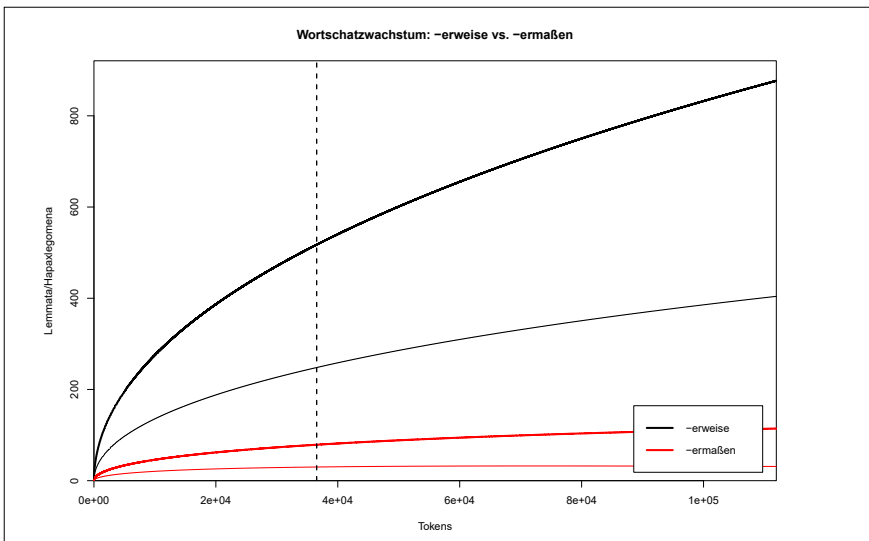


Abbildung 2: Anzahl der Lemmata (dickere Kurvenlinien) und der Hapaxlegomena (dünnere Kurvenlinien) bei steigenden Tokenzahlen von Bildungen auf *-erweise* bzw. *-ermaßen* (gemäß dem finiten Zipf-Mandelbrot-Modell in *zipfR*). Bei Bildungen auf *-ermaßen* werden die Wachstumskurven über die beobachtete Anzahl der Tokens (36.538, gestrichelte Linie) hinaus so weit extrapoliert, bis die beobachtete Anzahl der Tokens von Bildungen auf *-erweise* (111.937, Ende der Skala) erreicht wird.

Während die Wachstumskurve für die Lemmata auf *-erweise* am rechten Rand der Abb. 2 immer noch steil ansteigt, flacht sie für die Lemmata auf *-ermaßen* deutlich schneller ab. Ähnliches zeigt sich beim Vergleich der Kurven für die Hapaxlegomena: Die Kurve für *-erweise* steigt am rechten Rand der Abbildung

immer noch deutlich an, während die Kurve für *-ermaßen* dort kaum mehr Steigung zeigt. Das offenbar kaum beschränkte Suffix *-erweise* ist also entschieden produktiver als *-ermaßen*. Dabei ist bemerkenswert, dass die spezifische Produktivität „redundanter“ *erweise*-Ableitungen aus Partizipien I noch größer zu sein scheint: Fast zwei Drittel der entsprechenden Typen (90 von 142) sind Hapaxlegomena gegenüber weniger als der Hälfte der Typen (408 von 873) insgesamt (die p-Werte am Ende der Skala sind $3,55e-3$ für *-erweise* und $3,27e-4$ für *-ermaßen*).

4 Bildungen mit primären Präpositionen

Als prototypische adverbiale Relatoren in der Syntax vermitteln Präpositionen zwischen vor allem durch Verben kodierten Geschehen und Denotata von Nomen, die als Komplemente von Präpositionen fungieren und deren interne Rolle sättigen. Die externe Rolle entspricht dem Geschehen, das der Satzausdruck insgesamt beschreibt, der die adverbiale Präpositionalphrase enthält.²²

4.1 Relation und Eigenschaft

Mit Lehmann (2002: 78) besteht die Beziehung zwischen z.B. der Präposition *über* und dem Adverb *oben* darin, dass Quantifikation der internen Argumentstelle an deren explizite Sättigung durch eine Nominalphrase tritt:

- (18) a. Er ist **über** dem Dach. ÜBER(er,das.dach)
 b. Er ist **oben**. $\exists x$ ÜBER(er,x)

Die interne Bindung von Argumentstellen im Zuge der Adverbierung lässt sich auch beobachten, wenn z.B. in syntaktischen Phrasen präsenste pronominale Elemente zugunsten einer generischen Interpretation gelöscht werden:

- (19) zeit meines/deines/seines Lebens → zeitlebens

Der Schritt von vollen Nominalphrasen zu Pronomen bildet womöglich eine Vorstufe zur internen Bindung von Argumentstellen, vgl. insbesondere die nicht standardgemäßen, aber im Korpus recht häufigen Formen aus (20c) und (20d):

²² Man vergleiche attributiv verwendete Präpositionalphrasen wie *der Mann auf der Brücke*, die plausiblerweise auf relativ angeschlossene Konstruktionen zurückgehen: *der Mann, (der) auf der Brücke (ist)*.

- (20) a. wegen der Tochter → ihretwegen
 b. wegen des Wetters → deswegen
 c. dem Papst zufolge → ihmzufolge
 d. dem Papst nach → ihmnach

Präpositionen sind praktisch die einzigen Elemente, die innerhalb der Wortbildung mit Proformen verknüpft werden.²³ Proformen sind formal referentielle und also argumentstellensättigende Ausdrücke, die aber kaum zur Bedeutung beitragen, vgl. die Diskussion in Abschnitt 3.2.

Zwischen den 20 primären Präpositionen bestehen wesentliche Unterschiede bezüglich der Sättigung der internen Argumentstelle: Manche Präpositionen können auch ohne overte Sättigung adverbial fungieren, andere nicht. Der exemplarische Vergleich von Typen, in denen die Präpositionen *durch* bzw. *um* als Letztglieder fungieren, liefert Hinweise auf zugrundeliegende Mechanismen. Es fällt zunächst auf, dass *um* weit mehr Typen liefert als *durch* (47 vs. 21). Die nähere Betrachtung zeigt, dass der Vorsprung an Typen mit *um* v. a. der Existenz von Varianten mit bzw. ohne overt realisiertes Richtungsadverb (oder dessen Kurzform) geschuldet ist;²⁴ bei *durch* fehlen solche Varianten weitgehend (varierende Formen sind durch Verknüpfung mit „-“ markiert):

- (21) a. *rundum – rundherum; ringsum – ringsherum; andersum – andersrum – andersherum; rum – herum; weitem – weitherum; hintenrum – hintenherum; linksrum – linksrum – linksherum; rechtsrum – rechtsrum – rechtsherum; untenrum – untenherum; außenrum – außenherum; vorderum – vorderherum; hierum – hierherum; wiederum – wiederherum; darum – daherum, rumherum – darumherum*
 b. *kurzum, reihum, hinwiederum, ebendarum, verkehrtherum, obenherum, falschherum, richtigherum, vorderum, irgendwarum, weiterherum, umeinanderherum, worum*
- (22) a. *zwischen-durch – zwischenhindurch; mittendurch – mittenhindurch; dadurch – dahindurch; wodurch – wohindurch*
 b. *hindurch, untendurch, querdurch, ebendadurch, hintendurch, unterdurch, obendurch, vordurch, drunterdurch, halbdurch, zickzackkursdurch, weiterdurch, hierdurch*

23 Mögliche Ausnahmen bilden vereinzelte Rektionskomposita wie *Allesfresser/-köpfer*, sofern das Element *alles* hier als Pronomen analysiert wird.

24 Als Quelle für die Kurzform kommen auch Bildungen mit *da* in Frage; s. u. zur pronominalen Qualität der Richtungsadverbien *hin* und *her*.

Die Varianten in (21) sind größtenteils synonym verwendbar, bezüglich der Varianten in (22) besteht völlige Synonymie in der lokalen Lesart. Im Fall von *durch* besteht offenbar die Möglichkeit der nur oberflächlichen Löschung einer formal leerstellensättigenden Proform. Das auch pronominal fungierende Richtungsadverb *hin* scheint in Vorkommen von *durch* funktional oder bedeutungsseitig schon enthalten zu sein:

1. *durch* und *hindurch* sind bei trennbaren Partikelverben in lokaler Lesart salva veritate austauschbar (Šimečková 2002: 102).
2. *durch* ist nur mit vorangestelltem *hin* kompatibel (weshalb z. B. für *um* vorhandene Varianten mit *her* fehlen, s. o.).
3. In den Adverbtypen finden wir viel *durch* ohne Richtungsadverb, aber nicht *um* ohne Richtungsadverb.²⁵

Die interne Stelle von *um* muss, bevor es als Adverb fungieren oder weiter komponiert werden kann, transparent gesättigt werden. Dies geschieht zumeist durch das Richtungsadverb *her*, aber auch substantivische oder adjektivische Stämme, die mit Richtung bzw. Lokalisation zu tun haben, können diese Funktion ausüben (*reihum*, *ringsum*, *rechtsum*). Die tendenzielle Funktionsfähigkeit von *durch* gegenüber *um* zeigt sich dabei auch syntaktisch in der Vorfeldfähigkeit – alleiniges Auftreten im Vorfeld ist für *um* im Teilkorpus nie, für *durch* dagegen 53 mal der Fall, vgl. z. B. (23).

(23) Durch kam er nie. (dpa, 10.07.2009, o.S.)

Dafür, dass *her* in Kombination mit *um* argumentstrukturell wirksam ist, spricht im Übrigen, dass *um* kombiniert mit dem origo-orientierten *her* als Verbpartikel zu einer Verallgemeinerung der Bedeutung führt. Das geht systematisch mit der internen Bindung der Stelle des direkten Objekts einher, vgl. z. B. (24), wo die

25 Ähnlich argumentiert auch die GDS (1997: 2087) für bestimmte Verwendungen von *durch*; vgl. ebenfalls Behaghel (1924: 23):

Die Präposition ist in der Regel ein Adverb von relativer Bedeutung, das seine Bestimmung durch eine unmittelbar davon abhängige Größe erfährt, und zwar durch eine solche, die für sich allein in dem gegebenen Satzzusammenhang nicht erscheinen könnte: *mit Recht*, *der Eier wegen*, *bis jetzt*. In einer Wendung wie *er schläft die Nacht durch* ist *durch* noch nicht Präposition, da es auch heißen könnte: *er schläft die Nacht*; [...]. Doch ist eine völlig strenge Scheidung zwischen Präposition und Adverb nicht durchführbar.

Die pronominale Funktion von Richtungsadverbien scheint in Substitutionen wie den folgenden durch: *es kommt dazu* – *es kommt hinzu*; *er kommt heran* – *er kommt zu mir*; *x ist drüber* – *x ist hinüber*.

Unmöglichkeit, die Präposition *an* wegzulassen, aufzeigt, dass *Schokolade* nicht als direktes Objekt von *herumbasteln* lizenziert ist.

- (24) Und es gibt viele hier, die mehr oder weniger erfolgreich *(an) Schokolade herumbasteln. (Süddeutsche Zeitung, 28.9.2012, S. 11)

Die äußerst produktive Kombination von *herum* mit nahezu beliebigen Verben verweist wiederum auf einen Zusammenhang zwischen semantischer Harmlosigkeit (Redundanz bzw. Schwächung) und Produktivität (vgl. Abschnitt 3.3).

4.2 Erst- und Letztglieder

Sofern präpositionale Elemente maßgeblich für Adverbierung sind, erwarten wir dem Prinzip der Rechtsköpfigkeit gemäß, dass sie als Letztglieder erscheinen (vgl. oben Abschnitt 3.2). Das Gros der Typen in ADVERBIEN_1, die präpositionale Elemente enthalten, bestätigt diese Erwartung; *zu* und *vor* sind jedoch als Erstglieder signifikant überrepräsentiert, vgl. Abb. 3.²⁶

zu und *vor* sind in je verschiedenen Dimensionen diversifiziert: *zu* ist bezüglich seiner Form stabil, variiert aber funktional, indem es in seinen Hauptverwendungen als Präposition, als Gradpartikel und als Infinitivmarker fungiert. In ähnlicher Weise, wenn auch weniger augenfällig, haben die Präpositionen *in* und *an* Funktionen im Bereich der Graduierung bzw. im Bereich Tempus und Aspekt: Sie begleiten insbesondere Superlative (*am besten*, *im mindesten*) und haben in Kombination mit Infinitiven aspektuelle Funktion, vgl. die Beispiele für die sogenannte rheinische Verlaufsform:

- (25) Es muss ein bisschen was passieren, wir sind am verhandeln. (Braunschweiger Zeitung, 3.3.2011, o. S.)
- (26) Der Linksextremismus ist im abklingen, das kann man sagen. (Zeit Online, 28.2. 2001, o. S.)

Auf eine rekonstruierte Wurzel **per* zurück gehen *vor* und die nicht-lokative Form *für* sowie das verbale Präfix *ver-*,²⁷ das systematisch aspektuell wirksam ist.

26 Vgl. zu *zu* als Erstglied auch Abschnitt 3.1.1. Beispiele für *vor* als Erstglied sind *vorab*, *voran*, *voraus*, *vorbei*, *vorerst*, *vor(vor)gestern*, *vorher*, *vorhin*, *vormals*, *vorüber*.

27 Vgl. Seebold (2002). Als verbale Präfixe sind *zu-* und *ver-* sehr produktiv und in der aspektuellen Dimension auf die Kodierung von Zustandswechseln spezialisiert. *zu* und *vor* lizenzieren als einzige verbale Präfixe bzw. Partikeln systematisch Dativar-

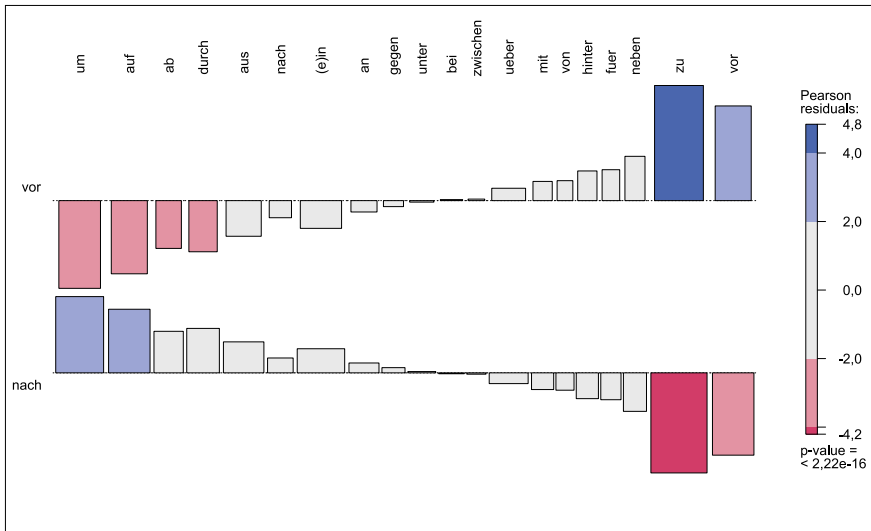


Abbildung 3: Assoziationsplot für primäre Präpositionen als Erst- und Letztglieder in ADVERBIEN_1 ($\chi^2 = 147$, $df = 19$, $p < 2,2e-16$, Cramérs $V = 0,52$). Die Höhe der Balken visualisiert den Grad der standardisierten Abweichungen (Residuen) von den erwarteten Werten, wobei Balken oberhalb der gepunkteten Linie höhere, Balken unterhalb niedrigere Werte als erwartet bedeuten. Die Breite der Balken ist proportional zu den erwarteten Häufigkeiten. Blaue Einfärbung steht für signifikante Abweichung (+/- 1,97) nach oben, rote Einfärbung für signifikante Abweichung nach unten.

Die Vielfalt der Typen mit *zu* als Präfix ist vor allem zwei Bauweisen zu verdanken, von denen die erste direkt als Umbau aus der Syntax erkennbar ist, indem das nominale oder adjektivische Letztglied in der Regel eine von *zu* verantwortete Dativendung trägt (sortiert nach absteigender Häufigkeit):

- (27) *zurück, zudem, zurzeit, zugute, zustande, zurecht, zugrunde, zuweilen, zuhause, zuteil, zueinander, zunichte, zuliebe, zuhauf, zunutze, zumute, zutage, zuhänden, zuwege, zuleide, zuschanden, zugut, zuhülfe, zupasse, zuweil, zualledem*

Die zweite Bauform basiert auf *zu* und Superlativen (sortiert nach absteigender Häufigkeit):²⁸

*gumente (jemandem zuarbeiten ggüb. *jemandem arbeiten, jemandem vorlesen vs. *jemandem lesen)*, was dafür spricht, dass die Elemente eine Variable ins Spiel bringen, die durch das Dativargument gebunden werden kann (vgl. Brandt 2003, Hole 2014).

²⁸ Die übrigen Typen mit vorangestelltem *zu* sind: *zunächst einmal, zuvor, zugleich, zumal, zuviel, zuwenig, zusehends, zuwider*. Als einzige präpositionale Form kann

- (28) *zunächst, zuletzt, zumindest, zuerst, zumeist, zutiefst, zuallererst, zuoberst, zuvorderst, zumindestens, zuvörderst, zuhinterst, zuallerletzt, zuunterst, zuinnerst, zuallermeist, zuhöchst, zualleroberst, zuäußerst, zuallernächst, zuallermindest, zuallervorderst, zuallervörderst*

Die Möglichkeit der rekursiven – und semantisch redundanten – Einbettung des quantifizierenden Elements *aller* macht die in (29) exemplifizierte Bauweise zur Basis prinzipiell unendlich vieler neuer Typen (vgl. unten Abschnitt 5.2).

Wenn wir die Typen mit präpositionalen Letzt- und Erstgliedern vergleichen, wird klar, dass erstere entlang typischer Wortbildungsregeln erzeugt werden: Die vorgefundenen Typen sind relativ gleichmäßig über die verschiedenen Präpositionen – mit Ausnahme von *zu* und *vor* – verteilt, sie enthalten kaum flexionsverdächtige Elemente,²⁹ und der präpositionale Kopf steht rechts. Insbesondere die Typen mit Proformen als Erstglied – *da, hier, wo, dem, her, hin, so* – applizieren breit über alle Präpositionen. Die Typen mit Substantiven, Adjektiven oder Adverbien als Erstglied enthalten regelmäßig ein optionales Bindeelement in Mittelstellung, das offenbar den Vollformen der Richtungsadverbien *hin* bzw. *her* bzw. auch der Proform *da(r)* zu verdanken ist, vgl. z.B. *berg(he)runter, oben(dr)auf*. Wieder muss die interne Stelle der Präposition gesättigt werden, bevor das Produkt ein Kompositum mit einem anderen freien Morphem bilden kann. Bei den Typen mit Präpositionen als Erstglied dagegen fehlen solche Bindeglieder. Stattdessen weisen insbesondere die Typen mit Substantiven, Adjektiven oder Adverbien die Endungen *-e* oder *-en* auf, die offenbar Reflexe eines durch die Präposition in der Syntax zugewiesenen Dativkasus sind (s. auch oben 6 und Abschnitt 3.1.1). Es handelt sich also nicht um eigentliche Wortbildungen bzw. nach Wortbildungsregeln erzeugte Typen, sondern um mehr oder weniger transparente Umbauten aus der Syntax.

5 Vergleiche und Adverbierung

Die Werte in Tab. 1 geben einen Hinweis darauf, dass superlativische Bildungen im Bereich des Adverbs besonders stark vertreten und produktiv sind; dies gilt insbesondere für eindeutig adverbmarkierende, auf *-stens* endende Formen,

zu mit gradierender Bedeutung in Mittelstellung vorkommen, nämlich in *allzugern, allzuoft, allzuviel* oder *vielzusehr, vielzuoft*. Die quantifizierenden Elemente *all* bzw. *viel* scheinen hier die Zusammenschreibung zu ermöglichen, vgl. unten Abschnitt 5.2. Weitere Formen sind *dazumal, dannzumal, annodazumal, sozusagen*.

29 In Fällen wie *hint-en-an* oder *tag-s-über* sind *-en-* bzw. *-s-* als Teil des als Erstglied fungierenden Adverbs bzw. als Fugenelement anzusehen.

die als semantisch schwächere (nämlich elativische) Formen auf Superlative bezogen sind, die durch Präpositionalphrasen ausgedrückt werden (vgl. oben Abschnitt 3.1.2). Wie fügt sich die Popularität superlativischer Formen in das hier entstehende Bild der Adverbierung?

5.1 Schwache Komparative

Der Bereich der Komparation ist variationsträchtig, da eine Vielzahl an Ausdrucksmöglichkeiten einer offenbar stark beschränkten Auswahl an Bedeutungen gegenübersteht. Die Sätze in (29) etwa bedeuten je dasselbe, und die Bedeutungsunterschiede der Sätze in (30) können je nach Kontext minimal ausfallen.³⁰

- (29) a. Ben ist schwerer als / nicht so leicht wie Bo.
 b. Bo ist leichter als / nicht so schwer wie Ben.

(30) Ben ist zu/etwas/aber schwer / nicht leicht genug / schwerer als zulässig.

Im Bereich des adverbialen Gebrauchs bewirken superlativische Formen von Adjektiven gegenüber flexionslosen vor allem Emphase oder Intensivierung (z.B. 31a), während wahrheitsfunktionale Unterschiede eine eher untergeordnete Rolle spielen (z.B. 31b).

- (31) a. Er grüßte herzlich. ≈ Er grüßte herzlichst.
 b. Er kam schnell. ≈ Er kam schnellstens.

Auch für den adverbialen Gebrauch von Adjektiven gilt, dass absolut vorkommende komparierte Formen ohne kontextuelle Verankerung zu schwächeren Bedeutungen führen als nicht komparierte Formen:³¹

- (32) a. Das habe ich aber schon länger gewusst. (Saale-Zeitung, 11.7.2012, S. 11)
 b. Bert Schreiber hat's schon lange gewusst. (Mannheimer Morgen, 5.5.2006, o. S.)

Der Umstand, dass komparierte Formen bei einem Mehr an Ausdruck tatsächlich ein Weniger an Bedeutung liefern, mag zu ihrer Beliebtheit für den adverbialen

³⁰ *Bedeutung* verstehen wir wahrheitsfunktional, d.h., wir abstrahieren von eventuellen Unterschieden bezüglich des Gebrauchs (Pragmatik) und eventueller Kollokationen etc.

³¹ Siehe dazu auch Becker (2005).

Gebrauch beigetragen haben, indem er die verwendete Morphologie dafür freimacht, eben anderes als stärkere Bedeutungen zu markieren.³² Bemerkenswert ist, dass die im Falle des Superlativs in adverbialer Funktion notwendig erscheinende präpositionale Form *am ... -sten* und die mit ihr verbundene, eindeutig Adverbien markierende *-stens*-Form (vgl. Abschnitt 3.1.2) dort erscheinen, wo der in adverbialer Funktion unflektierte Positiv des entsprechenden Adjektivs erscheint. In Anbetracht der adverbierenden Funktion des produktiven *-stens*-Suffixes liegt es nahe, auch im Falle des produktiv adverbialen Gebrauch anzeigenden unflektierten Positivs ein adverbierendes, allerdings unsichtbares Morphem anzusetzen, d. h., von Konversion und entsprechend einer Kategorie „Adjektivadverb“ auszugehen (s. o. Abschnitt 1.2 und unten Abschnitt 5.3).³³

Wir gehen davon aus, dass Ausdrücke, die graduierbare Eigenschaften bezeichnen, Relationen zwischen zu vergleichenden Entitäten kodieren (s. Abschnitt 1.2). Die erste Stelle ist je durch das **KOMPARANDUM** bzw. den durch es instantiierten Grad der relevanten Eigenschaft besetzt; die zweite Stelle durch einen **VERGLEICHSTANDARD**, der je nach Steigerungsstufe unterschiedlich realisiert werden kann (vgl. Tab. 3).

Tabelle 3: **KOMPARANDA** und **VERGLEICHSTANDARDS** der Steigerungsstufen.

	KOMPARANDUM	VERGLEICHSTANDARD
Positiv	Ben ist/läuft (so) schnell	{ Ø, für ein Kind, wie Ben }
Komparativ	Ben ist/läuft schneller	als Bo
Superlativ	Ben ist/läuft am schnellsten	Ø

An verschiedenen Stellen haben wir Evidenz dafür gesehen, dass Adverbierung damit zu tun hat, aus relationalen Ausdrücken einstellige Ausdrücke zu machen (vgl. die Abschnitte 1.2, 1.3, 4.1). Im Fall von Superlativen spricht viel dafür, dass der **VERGLEICHSTANDARD** implizit durch Allquantifikation gebunden wird. Erstens sind Superlative oberflächlich betrachtet einstellige Prädikationen, deren **VERGLEICHSTANDARD** sich nicht unabhängig realisieren lässt:

32 Vgl. unten Abschnitt 5.3 und Paul (1920: 110):

Zu dem Adjektivum *mittel* hat man, weil darin etwas Vergleichendes liegt, einen an sich unnötigen Komparativ und Superlativ *mittlere*, *mittelste* gebildet; diese Formen haben allmählich den Positiv ganz verdrängt [...].

33 *Mittelste* ist heute allerdings unüblich; *mittlere* verhält sich synchron wie die lokalen Adjektive mit stammhaftem Ausgang *-er*, der auch bei Superlativierung erhalten bleibt (*oberste*, *vorderste* etc.); es handelt sich also synchron nicht mehr um Komparative.

(33) Ben ist am schnellsten (*als/wie Bo) / (*für ein Kind).

Zweitens gibt es auch sprachübergreifend einen engen Zusammenhang zwischen Superlativierung und Allquantifikation, indem Elemente wie *all* in einigen Sprachen die Basis für Superlativmorphologie abgeben bzw. in Kombination mit Komparativmorphologie Superlative bilden.³⁴ Im Deutschen lässt sich das Element *aller* ohne Änderung der wahrheitsfunktionalen Bedeutung rekursiv mit superlativischen Formen verknüpfen, was darauf hindeutet, dass Allquantifikation in Superlativen schon anderweitig kodiert ist (*Ben ist am (aller(aller...)) schnellsten*). Drittens machen Superlative ihr Argument analog zu allquantifizierenden Ausdrücken zu monoton fallenden Umgebungen, die negative polare Elemente wie deutsch *je* lizensieren:

(34) Der große Plan, lehrte Noam Chomsky, er existiere wirklich, und alle vorhandenen Sprachen, alles, was je gesagt und geschrieben wurde, sei nur die Aktualisierung eines immer schon vorhandenen Wissens [...]. (Süddeutsche Zeitung, 6.12.2008, o. S.)

(35) Auschwitz war das Niedrigste, was je der Mensch dem Menschen angetan hat. (Nürnberger Zeitung, 28.1.2005, o. S.)

Superlativische Formen sind womöglich deshalb zur Adverbierung prädestiniert, weil die entscheidende Sättigung der internen Argumentstelle ihre Gestalt schon qua Allquantifikation definiert. Es kommt hinzu, dass nicht-attributiv verwendete Superlative durch Präpositionen lizenziert werden müssen, die auch als Ausgangsmaterial für das im Bereich des Adverbs allgegenwärtige *-s* fungieren (Abschnitt 3.1.2). Nicht zuletzt ist Superlativierung bzw. Elativierung bei *-stens* Formen im semantischen Sinne weit(est)gehend harmlos, was, wie in anderen Bereichen beobachtet, zu ihrer Produktivität beitragen mag (vgl. Abschnitte 3.3, 4.1).

5.2 Harmonisches Zusammenrücken: *zu*, *all*, *-st(ens)*

Die Präposition *zu* wirkt in Richtung der Adverbierung mit Formen zusammen, die universale Quantifikation ausdrücken – dazu gehören das Element *all* bzw. *aller* und das Superlativsuffix *-st* (s. oben Abschnitt 4.2). Im Bereich der

34 So z.B. im Lettischen, Estonischen oder Amharischen, vgl. Heine & Kuteva (2001: 36–37). Vgl. zur Analyse von Superlativen als Komparative mit allquantifiziertem VERGLEICHSTANDARD auch Ultan (1972), Heine (1997: 124) sowie Bobaljik (2012: 61).

Adverbbildungen mit Präpositionen konstatieren wir neben dem Bestehen eines ‚morphologischen‘ Typs mit Letztgliedern und Proformen (*hinten(he)rum*, *mitten(hin)durch*, vgl. oben Abschnitt 4.1) einen alternativen ‚syntaktischen‘ Weg der Sättigung bzw. Bindung der internen Argumentstelle durch (universale) Quantifikation (*zu(aller)erst*, s. u.).

zu und *-* in geringerem Maße *- vor* kommen verglichen mit den anderen Präpositionen signifikant häufiger als Erstglieder vor (Abschnitt 4.2) und Superlative geben eventuell deshalb einen besonders brauchbaren Rohstoff für Adverbien ab, weil ihr internes Argument implizit quantifiziert ist. Besonders die Möglichkeiten, superlativische Formen alternativ durch Voranstellung von *zu*, durch Ersetzung einer Präposition durch suffigiertes *-s* oder auch stillschweigend *zu* adverbieren sowie rekursiv und damit beliebig oft das quantifizierende Element *aller* einzubetten, tragen zu Typenreichtum und Produktivität superlativischer Adverbien bei, vgl. die Übersicht der Bildungsmöglichkeiten anhand einschlägiger Formen aus ADVERBIEN_1 in Tab. 4.³⁵

Tabelle 4: Bildungen mit Superlativformen, *zu-*, *-aller-* und *-ens*.

Grundform	∅	zu-	aller-	-ens	zu--ens
<i>meist</i>	✓	✓	✓	✓	✓
<i>erst</i>	✓	✓	✓	✓	-
<i>mindest</i>	-	✓	✓	✓	✓
<i>höchst, letzt, tiefst</i>	-	✓	✓	✓	-
<i>äußerst, hinterst, nächst, oberst, unterst, vorderst</i>	-	✓	✓	-	-
<i>best, feinst, genauest, schärfst, schlimmst, schnellst, schönst, spätest, strengst, striktest, wärmst, wenigst</i>	-	-	✓	✓	-

Einbettung von *aller* ist ausnahmslos und sowohl in Kombination mit *zu-* wie auch mit *-stens* möglich. Vorangestelltes *zu-* und suffigiertes *-s* sind weitestgehend in komplementärer Distribution, wie zu erwarten ist, wenn *-s* vorangestellte Präpositionen ersetzt (vgl. Abschnitt 3.1.2). Ausnahmen bilden *zumindestens* mit 2.351 Belegen (gegenüber je ca. 700.000-mal *zumindest* und *mindestens*) sowie *zumeistens* mit zwei Belegen im Gesamtkorpus (gegenüber je ca. 100.000-mal *zumeist* und *meistens*). Möglicherweise werden die Grundformen hier nicht

35 Die „nackten“ Formen *höchst* oder *äußerst* werden als Gradpartikeln gebraucht, die aufgrund der fehlenden Funktionsfähigkeit (reflektiert in alleinigem Auftreten im Vorfeld) aus der hier zugrundegelegten Definition von Adverb herausfallen.

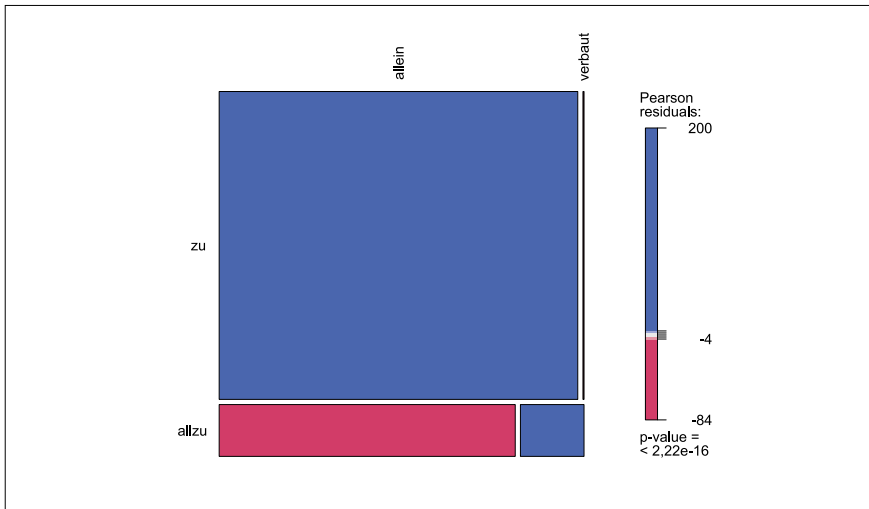


Abbildung 4: Mosaikplot für (un)verbautes (*all*)*zu* in ADVERBIEN_1 ($\chi^2 = 48425$, $df = 1$, $p < 2.2e-16$, $\Phi = 0,36$). Die Häufigkeitsverteilung wird durch Rechtecke visualisiert, deren Größe proportional zur Fallzahl ist. Rote Einfärbung bedeutet Unterrepräsentation, blaue Einfärbung Überrepräsentation des jeweiligen Falls.

mehr als Superlative empfunden, wodurch die Bildungen auch aus dem Bereich der Regel rücken, die die mit phrasalen Superlativen verknüpften Präpositionen gegen *-s* tauscht. Stillschweigende Adverbierung der Superlativform ist nur in den Fällen möglich, die auch eindeutige Adverbierung durch vorangestelltes *zu-* oder durch suffigiertes *-ens* erlauben. Nicht enthalten sind – mit Ausnahme von *erst* – auf Ordinalzahlen basierende Typen, die wie Superlative das semantische Merkmal „Einzigkeit“ tragen, dessen partielle Wiederholbarkeit durch *aller* für die beobachtete Variation wesentlich ist.

Auch andere Wortbildungen – bzw. Umbauten zu Adverbien – weisen auf eine für die Adverbierung entscheidende Sonderbeziehung zwischen *zu-* und *all-* hin. Insbesondere wird die Form *allzu*, aber nicht *zu* allein, mit den durch sie modifizierten Adjektiven bzw. Adverbien nicht selten (in etwa 20 % der Fälle) zusammengeschrieben und mithin als eigenständige Bedeutungseinheit (d. h. Wort) empfunden. Abb. 4 visualisiert die Überrepräsentation von „unverbautem“ *zu* (großes Feld links oben) gegenüber der Überrepräsentation von „verbautem“ *allzu* (Feld rechts unten) für solche Fälle, in denen die Formen als Adverbierer konkurrieren.³⁶

36 Dahinter stehen aus dem Gesamtkorpus extrahierte Kombinationen mit den Formen *sehr*, *oft*, *weit*, *lange*, *gern(e)*, *gut*, *schnell* sowie *bald*, die möglichst ausschließlich als Adverb fungieren; *zu* kommt „verbaut“ frequent in Kombination mit den

5.3 Wortbildung vs. Flexion

Die Frage „Wortbildung oder Flexion“ präsentiert sich unter der hier gemachten Unterscheidung zwischen dem Bau von und dem Umbau zu Adverbien in anderem Licht: die formalen Parallelen zwischen *bona-fide*-Zusammenrückungen und superlativischen Adverbformen (vgl. oben 6 sowie 27 und 28) ebenso wie die Beobachtungen zu Univerbierung und nicht-attributiver (adverbialer) Verwendung im letzten Abschnitt sprechen dafür, dass es sich bei superlativischen Adverbien um Umbauten aus der Syntax handelt; die synchron produktive und eindeutig Adverbien markierende *-stens* Form (vgl. Abschnitte 3.1.2, 5.2) präsentiert sich als ein Derivationsuffix mit syntaktischer Geschichte. Den Übergang von flexivischer Syntax zur Wortbildung bzw. die Durchmischung beider Ebenen markieren in Sonderheit die aus phrasal-präpositionalen Superlativen und *stens*-Adverbien bestehenden Reihen (z. B. *am schnellsten* – *schnellstens*): das *-en-* in den *stens*-Formen ist ein Überbleibsel aus der Syntax, das mit dem gegen eine es ursprünglich verantwortende Präposition getauschten *-s* zu *-ens* wird.

Entlang traditioneller Einteilungen sprechen die hier erzielten Ergebnisse dafür, dass Superlativmorphologie eher in den Bereich der Wortbildung gehört, indem sie eine vitale Funktion bei der Adverbierung innehat. Klar zutage tritt dies bei durch Umbau aus adverbialen *am ... -sten*-Phrasen entstandene *-stens* Adverbien (vgl. die Abschnitte 3.1.2, 5.1). Für die Zugehörigkeit der Komparation zur Wortbildung spricht auch ihre Konkurrenz zur Derivationsmorphologie, die deutlich im Englischen zutage tritt (Zwicky 1989), wo Komparativsuffixe und das systematisch Adverbien markierende Suffix *-ly* nicht in ein und derselben Wortbildung erscheinen (Zwicky 1989:139).³⁷ In derselben Richtung lässt sich eine Tendenz beobachten, Superlativmorphologie fälschlich nach außen zu „schieben“ bzw. diese am rechten Wortrand zu wiederholen, sofern die entsprechende

quantifizierenden Formen *viel* bzw. *wenig* als Partikel oder Pronomen vor. Die absoluten Werte sind:

	allein	verb.
zu	313.735	931
allzu	43.583	9.361

37 In der Geschichte des Deutschen haben wie oben erwähnt (Abschnitt 3.1.2) verschiedene Lautwandel zum Zusammenfall der im adverbialen Gebrauch endungslos erscheinenden Adjektivendungen *-is* und *-os* mit den Komparativformen *-ir* und *-or* geführt, die im Mittelhochdeutschen zu *-er* werden, das regelmäßig adverbiale Funktion markiert (Wilmanns 1896: 607, vgl. Abschnitt 3.1.2). Zeugen dieser Entwicklung sind z. B. rein adverbiales *eher* (von *air* (*airis*), s. o.), *leider* und evtl. *fürder* (zu *fort*, vgl. engl. *further*) oder adverbial gebrauchte Komparativformen wie *länger* oder *weiter*.

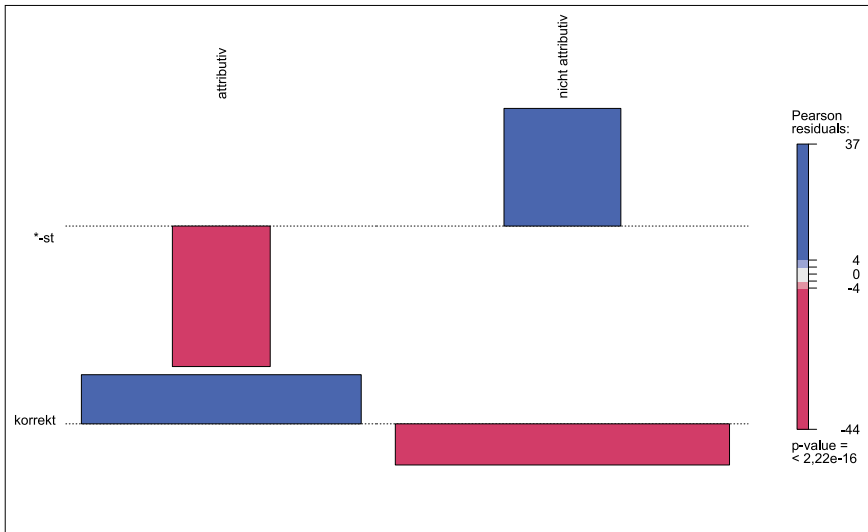


Abbildung 5: Assoziationsplot für (nicht-)attributiven Gebrauch (un)korrekter Superlativformen in ADVERBIEN_1 ($\chi^2 = 3751,4$; $df = 1$; $p < 2,2e-16$, $\Phi = 0,24$). Siehe für Erläuterungen Abb. 3.

Form adverbial verwendet wird. Während attributiver und nicht-attributiver Gebrauch der korrekten Formen etwa ausgeglichen ist, ist der nicht-attributive, zumeist klar adverbiale Gebrauch „falscher“ Formen wie *weitmöglichst* oder *bestmöglichst* mehr als zehnmal (13,45x) so hoch wie ihr attributiver Gebrauch. Der Assoziationsplot in Abb. 5 zeigt, dass mit inkorrekt *st*-Suffigierung eine Verschiebung hin zum nicht-attributiven Gebrauch einhergeht.³⁸

38 Der Vergleich basiert auf einer Liste der 20 häufigsten Adverbien auf *-lichst* aus KoGraUK gesamt. Ausgewählt wurden diejenigen adjektivischen Basen, für die es eine ‚falsche‘ und eine ‚korrekte‘ Superlativform gibt: *baldmöglichst* (vs. *baldestmöglichst*), *schnellstmöglichst* (vs. *schnellstmöglich*), *bestmöglichst* (vs. *bestmöglich*), *raschmöglichst* (vs. *raschestmöglich*), *weitmöglichst* (vs. *weitestmöglich*). Die absoluten Werte sind:

	ADJA	¬ADJA
*-st	477	6.414
korrekt	25.592	30.774

5.4 Zusammenfassung

Das Gros der Adverbien im Deutschen – 94% der Typen in ADVERBIEN_1 – fällt formal in 29 Klassen, die sich neben Simplizia zu drei Oberklassen bündeln lassen: Adverbien mit flexivischen Endungen, Adverbien mit präpositionalen Bestandteilen und Adverbien mit aus Substantiven gebildeten Suffixen. Die ersten beiden Oberklassen und Beziehungen zwischen ihnen standen hier im Fokus; ein zentrales Ergebnis ist, dass es im Deutschen einen Grammatikalisierungsprozess gibt, in dem die Präpositionen *in/im*, *an/am* und *zu/zum* eliminiert und durch suffigiertes *-s* ersetzt werden.

Relationalität ist eine grundlegende Bedingung für das Ausgangsmaterial, aus dem Adverbien gebaut werden. Im Zuge der Adverbierung wird das interne (PATIENS-, VERGLEICHSTANDARD-, Objekt-)Argument gesättigt bzw. gebunden und das externe (AGENS-, KOMPARANDUM-, Subjekt-)Argument zu Geschehen verschoben. Partizipien I und Superlative sind qua Bedeutung und der sie kodierenden Struktur besonders gut als Adverbien brauchbar: Partizipien I beziehen sich auf externe AGENS-Argumente, die eng an Geschehen geknüpft sind (Abschnitte 1.3, 3.3). Superlative beziehen sich auf KOMPARANDA, interne KOMPARATA werden hier per Allquantifikation gebunden (Abschnitt 5.1).³⁹

Innerhalb der Wortbildungen mit primären Präpositionen, die in der Syntax prototypische adverbiale Relatoren sind, gibt es wesentliche Unterschiede in syntaktischer und semantischer Dimension (vgl. Abschnitt 4):

- Die Pro-Adverbien *da*, *hier*, *wo* sowie *hin* und *her* sättigen die interne Argumentstelle bei Wortbildungen mit Präpositionen in Letztstellung und leiten so Adverbien ab.
- Präpositionen wie *durch* oder *mit* sind „näher“ an Eigenschaften (und mithin Adverbien), Präpositionen wie *um* „näher“ an Relationen; einiges spricht dafür, dass z. B. adverbiales *durch* eine stille Proform enthält.
- Wortbildungen mit den Erstgliedern *zu* und *vor* haben syntaktische Züge und eröffnen die Möglichkeit der Rekursion. Allquantifikation wie an Superlativformen geknüpft bindet die interne Argumentstelle in Bildungen mit *zu-*.

39 Bezüglich der in Abschnitt 1.1 angerissenen Grundsatzproblematik scheint das Bestehen von Mechanismen der Adverbierung bzw. von Mustern des Adverb-seins, die Kategorien und Bedeutungsbereiche übergreifen (viz. Reduktion der internen Stelle bestimmter Relationen), für einen lexembasierten Ansatz zu sprechen. Die Beobachtungen um Adverbbildung mittels *-stens* und – in möglicher Verlängerung – die Adverbableitung aus positiven Adjektiven mittels Nullmorphem, also Konversion, sprechen dagegen eher für morphem- bzw. wurzelbasierte Ansätze („distribuierte Morphologie“).

An verschiedenen Stellen finden sich Hinweise auf einen Zusammenhang zwischen semantischer Schwäche und Produktivität: Semantisch harmonisierende Elemente rücken produktiv zu Wörtern zusammen (Abschnitt 5.2) und redundante bzw. semantisch harmlose Operationen sind besonders produktiv (Abschnitte 3.3, 4.1, 5.1).

Literatur

- Aronoff, Mark. 1994. *Morphology by Itself*. Cambridge, MA: MIT Press
- Becker, Thomas. 2005. Warum eine alte Dame älter ist als eine ältere Dame: Zum absoluten Komparativ im Deutschen. *Deutsche Sprache* 33. 97–116.
- Behaghel, Otto. 1924. *Deutsche Syntax. Eine geschichtliche Darstellung. Band 2: Die Wortklassen und Wortformen*. Heidelberg: Winter.
- Bobaljik, Jonathan. 2012. *Universals in comparative morphology. Suppletion, superlatives, and the structure of words* (Current studies in Linguistics 50). Cambridge, MA: MIT Press.
- Brandt, Patrick. 2003. *Cipient Predication: Unifying Double Object, Dative Experiencer and Existential/Presentational Constructions* (LOT dissertation series 74). Utrecht, UiL-OTS: Dissertation.
- Brandt, Patrick. 2019. *Discomposition Redressed: Hidden Change, Modality, and Comparison in German* (Schriften des Instituts für Deutsche Sprache [SDS] 81). Tübingen: Narr.
- Bresnan, Joan. 1994. Locative Inversion and the Architecture of Universal Grammar. *Language* 70. 72–131.
- Carlson, Gregory N. 1977. A unified analysis of English bare plural. *Linguistics and Philosophy* 1. 413–456.
- Davidson, Donald. 1967. The Logical Form of Action Sentences. In Nicolas Rescher (Hg.), *The Logic of Decision and Action*, 105–122. Pittsburgh: University of Pittsburgh Press.
- Dudengrammatik 2009 = Dudenredaktion (Hg.). 2009. *Dudenband 4 – Die Grammatik*, 8. Aufl. Mannheim/Leipzig: Dudenverlag.
- DWB = Grimm, Jacob & Wilhelm Grimm. 1854–1961. *Deutsches Wörterbuch von Jacob und Wilhelm Grimm*. 16 Bände in 32 Teilbänden. Quellenverzeichnis 1971. Leipzig: Hirzel.
- Eisenberg, Peter. 2013. *Grundriss der deutschen Grammatik. Bd. 2: Der Satz*, 4. Aufl. Stuttgart & Weimar: Metzler.
- Elsner, Daniela. 2016. Empirisch basierte Überlegungen zu Ableitungen mit -weise/-erweise. In Eric Fuß, Marek Konopka, Beata Trawiński, & Ulrich H. Waßner (Hgg.), *Grammar and Corpora 2016*, 211–220. Heidelberg: Heidelberg University Publishing.

- Evert, Stefan & Marco Baroni. 2007. zipfR: Wordfrequency distributions in R. In Sophia Ananiadou (Hg.), *Proceedings of the 45th Annual Meeting of the Association for Computational Linguistics Companion Volume Proceedings of the Demo and Poster Sessions*. 29–32. Prag: ACL. <https://www.aclweb.org/anthology/P07-2008.pdf> (abgerufen am 7.7.2020).
- Fleischer, Wolfgang & Irmhild Barz. 2012. *Wortbildung der deutschen Gegenwartssprache*, 4. Aufl. Berlin & Boston: De Gruyter.
- GDS = Zifonun, Gisela, Ludger Hoffmann & Bruno Strecker. 1997. *Grammatik der deutschen Sprache*. (Schriften des Instituts für Deutsche Sprache 7). 3 Bände. Berlin & New York: De Gruyter.
- Geuder, Wilhelm. 2019. Eine Art Wortart: Das Adverb im Deutschen. *Zeitschrift für Sprachwissenschaft* 38(2). 191–241.
- Halle, Morris & Alec Marantz. 1993. Distributed Morphology and the Pieces of Inflection. In Kenneth Hale & Samuel J. Keyser (Hgg.) *The View from Building 20*, 111–176. Cambridge, MA: MIT Press.
- Heidolph, Karl Erich, Walter Flämig & Wolfgang Motsch. 1981. *Grundzüge einer deutschen Grammatik*. Berlin: Akademie Verlag.
- Heine, Bernd. 1997. *Possession. Cognitive sources, forces and grammaticalization*. Cambridge: Cambridge University Press.
- Heine, Bernd & Tania Kuteva. 2002. *World Lexicon of Grammaticalization*. Cambridge: Cambridge University Press.
- Heinle, Eva-Maria. 2004. *Diachronische Wortbildung unter syntaktischem Aspekt*. Heidelberg: Winter.
- Helbig, Gerhard & Joachim Buscha. 2001. *Deutsche Grammatik: Ein Handbuch für den Ausländerunterricht*. Stuttgart: Klett.
- Hole, Daniel. 2014. *Dativ, Bindung und Diathese* (Studia grammatica 78). Berlin: De Gruyter.
- Krause, Maxi. 2007. Überlegungen zu DA-+X. *Sprachwissenschaft* 32. 453–485.
- Lee, David. 1999. Intransitive Prepositions: Are they viable? In Peter Collins & David Lee (Hgg.), *The Clause in English: In Honour of Rodney Huddleston*, 133–148. Amsterdam: Benjamins.
- Lehmann, Christian. 2002. New reflections on grammaticalization and lexicalization. In ders. (Hg.), *New reflections on grammaticalization*, 1–18. Amsterdam: Benjamins.
- Meibauer, Jörg, Ulrike Demske, Jochen Geilfuß-Wolfgang, Jürgen Pafel, Karl Heinz Ramers, Monika Rothweiler & Markus Steinbach. 2007. *Einführung in die germanistische Linguistik*, 2. Aufl. Stuttgart & Weimar: Metzler.
- Milsark, Gary L. 1974. *Existential sentences in English*. Cambridge, Massachusetts Institute of Technology, Dissertation.

- Paul, Hermann. 1897. *Deutsches Wörterbuch*. 9. Auflage. Bearbeitet von Helmut Henne & Georg Objartel unter Mitarbeit von Heidrun Kämper-Jensen. Halle: Niemeyer.
- Paul, Hermann. 1920. *Deutsche Grammatik. Bd. 5: Wortbildungslehre*. Halle a. S.: Niemeyer.
- Primus, Beatrice. 1999. *Cases and Thematic Roles: Ergative, Accusative and Active* (Linguistische Arbeiten 393). Tübingen: Niemeyer.
- Ronca, Dorina. 1975. *Morphologie und Semantik deutscher Adverbialbildungen. Eine Untersuchung zur Wortbildung der Gegenwartssprache*. Universität Bonn: Dissertation.
- Schäfer, Martin. 2013. *Positions and Interpretations. German adverbial adjectives at the syntax-semantics interface* (Trends in Linguistics. Studies and Monographs [TiLSM] 245). Berlin: De Gruyter.
- Schmoe, Friederike. 2002. „Folglich trat Hubert barfuß und dennoch ungemein heftig gegen die zue Tür.“ Über einige Eigenschaften der deutschen Adverbien. In Peter Wiesinger (Hg.), *Akten des X. Internationalen Germanistenkongresses Wien 2000. Zeitenwende – Die Germanistik auf dem Weg vom 20. ins 21. Jahrhundert*, 157–164. Bern: Lang.
- Seebold, Elmar. 2002. *Kluge. Etymologisches Wörterbuch der deutschen Sprache*. 24. Auflage. Berlin & New York: De Gruyter.
- Šimečková, Alena. 2002. *Funktionalisierung von Trennbarkeit und Untrennbarkeit beim komplexen Verb*. Prag: Karls-Universität.
- Telschow, Claudia. 2014. *Die Adjektiv-Adverb-Abgrenzung im Deutschen: Zu grundlegenden Problemen der Wortartenforschung* (Reihe Germanistische Linguistik 299). Berlin & Boston: De Gruyter.
- Ulan, Russell. 1972. Some features of basic comparative constructions. *Working Papers on Language Universals* 12, 117–162. Stanford: Stanford University.
- Wilmanns, Wilhelm. 1896. *Deutsche Grammatik. Gotisch, Alt-, Mittel- und Neuhochdeutsch*. Straßburg: Trübner.
- Zwicky, Arnold M. 1989. Quicker, more quickly, *quicklier. In Geert Booij & Jaap van Marle (Hgg.), *Yearbook of Morphology* 2, 139–173. Dordrecht: Foris.

Franziska Münzberg  & Sandra Hansen 

Starke vs. schwache Flexion aufeinanderfolgender attributiver Adjektive: *mit hohem technischen/ technischem Aufwand*

Abstract Usually, weak inflection of an attributive or nominalized adjective occurs if the adjective is preceded by an inflected determiner: *mit diesem technischen Aufwand* ('at great technical expense'). Otherwise, the inflection of the adjective is strong: *mit technischem Aufwand*. Following this rule of thumb, we would expect strong inflection of an adjective following another adjective whenever the determiner is missing: *mit hohem technischem Aufwand*. But many German speakers opt for a weak dative singular ending *-en* following the strong ending *-em* on the first adjective: *mit hohem technischen Aufwand*. This chapter shows which explanatory variables play a role in this variation within standard German.

Keywords adjective inflection, strong inflection, weak inflection, Switzerland, Swiss, logistic regression

Open Data Der Datensatz `ATTRIBUTIVE_ADJEKTIVE_1` ist öffentlich zugänglich und durchsuchbar in der [Datenbank attributive Adjektive](#) (DOI: [10.14618/attributiveAdjektiveDB](https://doi.org/10.14618/attributiveAdjektiveDB)).

1 Grundlagen

Folgt ein attributives oder nominalisiertes Adjektiv (Partizip) auf einen Determinierer mit Flexionsendung, wird es schwach dekliniert, sonst stark: *mit diesem technischen Aufwand*, aber *mit technischem Aufwand*. Abweichungen von diesem Prinzip finden sich bei

- maskulinen und neutralen Adjektiven im Dativ Singular: Die starke Flexionsendung *-(e)m* kann nicht nur dann schwache Flexion an folgenden Adjekti-

ven auslösen, wenn sie am Determinierer zu finden ist (*mit diesem technischen Aufwand*), sondern auch als Flexionsendung eines typischen Adjektivs (*mit hohem technischen Aufwand*). Weil die beiden Adjektive unterschiedlich flektieren, sprechen wir von Wechselflexion. Im Gegensatz dazu steht die Parallelflexion, bei der beide Adjektive gleich flektieren: *mit hohem technischen Aufwand*.

- Adjektiven, die auf Pronominaladjektive¹ (z. B. *beid-*, *einig-*, *solch-*) folgen: Pronominaladjektive haben Eigenschaften sowohl von Determinierern als auch von typischen Adjektiven. Entsprechend findet sich die Variation *mit solcher vom Gemeinderat gewollten* vs. *gewollter Verkehrskanalisierung*. Im Dativ Singular Maskulinum/Neutrum findet sich besonders viel schwache Flexion. Diese Variation betrifft aber auch alle anderen Merkmalskombinationen, bei denen starke von schwachen Endungen unterschieden werden können, wie etwa hier den Dativ Femininum Singular.
- maskulinen und neutralen Adjektiven im Genitiv Singular: Die starke Endung *-es* gibt es noch bei Determinierern (*unseres Schrebergartens*), aber bei typischen Adjektiven ist sie untergegangen (*aufgrund schlechten Wetters*). Bei den Determinierern gibt es Schwankungsfälle (*dieses/diesen Jahres*).
- Adjektiven nach *deren* und *dessen*: Traditionell lösen diese Wörter keine schwache Flexion am folgenden Adjektiv aus: *mit deren famoser Idee, für dessen famose Ideen*. Vereinzelt kommt aber auch schwache Flexion vor, wohl weil die Formen *deren* und *dessen* flektierten Determinierern ähneln: *mit deren famosen Idee, für dessen famosen Ideen*. Hochinteressant, aber nicht Thema dieses Kapitels ist die Frage, ob auch hier der Dativ Singular Maskulinum/Neutrum besonders stark betroffen ist, vgl. *zu deren eigenen Schutz* (Freie Presse, Beispielsatz aus Zweifelsfäleduden 2016: 225).

Welche Eigenheiten der starken dativischen Endung *-(e)m* dazu führen, dass sie in bestimmten Kontexten durch die schwache Endung *-(e)n* ersetzt werden kann, ist nicht abschließend geklärt (Überblick in Gallmann 2016). Für andere starke Endungen gilt das nicht analog, der Sprachwandel scheint zumindest bei Adjektiven nach Pronominaladjektiven in entgegengesetzter Richtung zu verlaufen (Bildhauer et. al. 2019: 308): älter *sämtlicher verwendeten Materialien* (schwach), neuer *sämtlicher verwendeter Materialien* (stark). Insofern lohnt der Sprachberatungsklassiker „-m oder -n“ (vgl. z. B. Kubczak 2015, Zweifelsfäleduden 2016: 29, Grammatikfragen 2018) weitere wissenschaftliche Beobachtungen.

1 Zum Terminus Pronominaladjektiv vgl. Wiese (2009: 167).

2 Fragestellung

Im Folgenden sollen die Bedingungen, unter denen Wechselflexion im Dativ Singular Maskulinum/Neutrum wahrscheinlich wird, näher beschrieben werden. Bei der Variation vom Typ *hohem technischem/technischen Aufwand* sind in der bisherigen Forschung unter anderem folgende Einflussgrößen untersucht worden:

- interne Struktur der NP (z. B. Moulin 2000), *und*-Test (z. B. Heynatz 1770: 96 = Dritter Theil Nr. 40)
- Interpunktion (z. B. Moulin 2000)
- semantische Klassen der Adjektive (z. B. Nübling 2011)
- Genus des Substantivs (z. B. Nübling 2011)
- Rektionstyp (der Kasus der NP wird von einer Präposition regiert oder nicht; Sahel 2011)
- phonologische Merkmale wie /m/ in der Mitte/am Ende des Substantivs (Niebuhr 2017)
- Frequenz der beiden Adjektive (Schäfer 2017)
- Land (Peter 2013)
- Jahr (Peter 2013)

Mehrere Einflussgrößen in Kombination berücksichtigt Schäfer (2017), allerdings anders als im vorliegenden Kapitel auf der Basis von Internetkorpora. Auch uns geht es um das Zusammenspiel von Einflussgrößen. Schließlich soll die Frage, welche Einflussgrößen besonders stark wirken, beantwortet werden.

3 Korpora, Suchanfragen und Methoden

Die hier dargestellten Ergebnisse beziehen sich auf zwei unterschiedlich große Datensätze. Den kleineren Datensatz, `ATTRIBUTIVE_ADJEKTIVE_1`, haben wir aus einem Ausschnitt des Untersuchungskorpus von 122 Millionen Tokens erzeugt. Er enthält nach dem Aussortieren von Fehlbelegen 1.598 Belege und war damit klein genug, um komplett von Hand annotiert zu werden. In der „Datenbank attributive Adjektive“ (Münzberg et al. 2018) kann dieser Datensatz durchsucht werden.

Die folgende Suchanfrage wurde gestellt:

Präposition	∅	erstes Adjektiv mit Endung <i>-em</i>	∅	∅	zweites Adjektiv mit Endung <i>-em</i> oder <i>-en</i>	Nomen
	Adverb		,	Adverb		

Abbildung 1: Suchmuster für Adjektivflexion im Datensatz `ATTRIBUTIVE_ADJEKTIVE_1`.

Präpositionalphrasen, die diesem Muster entsprechen würden, sind etwa:

- (1) a. mit ziemlich langem, beinahe schwarzem/schwarzen Haar
- b. mit ziemlich langem schwarzem/schwarzen Haar
- c. mit langem, schwarzem/schwarzen Haar
- d. mit langem schwarzem/schwarzen Haar

Der größere der beiden Datensätze, ATTRIBUTIVE_ADJEKTIVE_2, umfasst nach dem Aussortieren von Fehlbelegen 119.376 Belege. Gesucht wurde im Untersuchungskorpus nach dem Muster:

Präposition	erstes Adjektiv mit Endung <i>-em</i>	∅	zweites Adjektiv mit Endung <i>-em</i> oder <i>-en</i>	Nomen
		,		
		<i>und</i>		
		<i>bzw.</i>		
		<i>beziehungsweise</i>		
		<i>, aber</i>		
		<i>, doch</i>		
		<i>, also</i>		
		<i>d. h.</i>		
		<i>d.h.</i>		
		<i>das heißt</i>		
<i>, und zwar</i>				

Abbildung 2: Suchmuster für Adjektivflexion im Datensatz ATTRIBUTIVE_ADJEKTIVE_2.

Präpositionalphrasen, die diesem Muster entsprechen würden, sind etwa:

- (2) a. mit langem und schwarzem/schwarzen Haar
- b. mit langem, schwarzem/schwarzen Haar
- c. mit langem schwarzem/schwarzen Haar

Dieser größere Datensatz ermöglicht die Untersuchung syndetischer Koordination und eine Darstellung der regionalen Verteilung sowie der zeitlichen Entwicklung.

Die nachträgliche Belegannotation umfasste in beiden Datensätzen die folgenden Einflussgrößen, die in Abschnitt 4.1 bis 4.3 näher erläutert werden:

- Frequenz des ersten und des zweiten Adjektivs
- Silbenzahl des ersten und des zweiten Adjektivs
- /m/ am Ende des Substantivs
- /m/ in der Mitte des Substantivs

Nur der Datensatz ATTRIBUTIVE_ADJEKTIVE_1 enthält Belegannotationen von Hand nach den Prädiktoren

- Genus des Substantivs
- Wortbildung der Adjektive (Partizip – lexikalisiertes Partizip – kein Partizip)
- semantische Klasse der Adjektive (11-Klassen-System)
- semantische Klasse der Adjektive (2-Klassen-System: intersektiv und nicht intersektiv)
- *und*-Test-Ergebnis

sowie die im Suchergebnis (vgl. Abb. 1) vorhandenen Werte für die Variablen

- Komma zwischen den beiden Adjektiven JA/NEIN
- Adverb vor erstem Adjektiv JA/NEIN
- Adverb vor zweitem Adjektiv JA/NEIN

und die nachträglich automatisch annotierte Information

- Wahrscheinlichkeit dafür, dass ein Adjektiv attributiv zu einem gegebenen Substantiv auftritt (= $P(\text{Adj|N})$, im Weiteren „substantivspezifische Adjektivfrequenz“; vgl. fürs Englische Wulff 2003: 267: „noun-specific frequency“)

Bei diesen Einflussgrößen stand die Frage im Vordergrund, ob semantisch-syntaktische Prädiktoren wie Komma JA/NEIN und semantische Klasse der Adjektive einen stärkeren oder einen schwächeren Effekt auf die abhängige Variable Flexion STARK/SCHWACH haben als phonologische Prädiktoren wie Silbenzahl und /m/ am Ende des Substantivs – oder ob nicht beides durch Frequenzeffekte ausgehebelt wird. Für den kleinen Datensatz haben wir mehrere logistische Regressionsmodelle spezifiziert. Am Ende eines Selektionsprozesses stand ein Regressionsmodell mit den folgenden Prädiktoren (die Modellspezifikation ist im Anhang dargestellt):

- Frequenz des ersten Adjektivs
- Silbenzahl des zweiten Adjektivs
- /m/ am Ende des Substantivs
- /m/ in der Mitte des Substantivs
- semantische Klasse (11-Klassen-System) des zweiten Adjektivs
- Komma JA/NEIN
- *und*-Test-Ergebnis
- substantivspezifische Adjektivfrequenz des zweiten Adjektivs
- Land

Um die Effektstärken besser miteinander vergleichen und visualisieren zu können, wurden auf dem kleinen Datensatz ATTRIBUTIVE_ADJEKTIVE_1 mit binär umkodierten Faktoren Koeffizienten geschätzt (s. Abschnitt 5).

Die Regressionsanalyse für den großen Datensatz ATTRIBUTIVE_ADJEKTIVE_2 umfasst folgende Prädiktoren (die Modellspezifikation ist im Anhang dargestellt):

- Frequenz des ersten Adjektivs
- Silbenzahl des zweiten Adjektivs
- /m/ am Ende des Substantivs
- /m/ in der Mitte des Substantivs
- Koordinationsmarker wie Komma oder Konjunktion oder kein Koordinationsmarker
- Land
- Jahr
- inhaltliche Domäne

Ergänzt wurden die logistischen Regressionen durch visuelle Überprüfungen von Teilen des großen Datensatzes mithilfe von Assoziationsplots, die auf Chi-Quadrat-Analysen basieren (vgl. Zeileis et al. 2007).

4 Ergebnisse der Korpusuntersuchung

Im Folgenden werden die Wirkungen der einzelnen Einflussgrößen beschrieben. Die genauen Kennzahlen zu den Regressionsanalysen sind im Anhang dargestellt. Einen Überblick über die Variation im Datensatz `ATTRIBUTIVE_ADJEKTIVE_1` gibt Abb. 7 (Abschnitt 5).

Viele, aber nicht alle Hypothesen aus der Literatur sind in die Untersuchung eingeflossen. Nach Sahel (2005: 364–365) ist zu erwarten, dass das zweite Adjektiv bei der Rektion der Nominalgruppe durch eine Präposition eher stark flektiert wird als bei anderen Rektionstypen. In der Tat kann in einer Regressionsanalyse bei Adjektiven nach Pronominaladjektiven ein schwacher Einfluss in diese Richtung nachgewiesen werden. Dies zeigte sich im Rahmen der Arbeiten an Bildhauer et al. (2019). Da artikellose Nominalphrasen im Dativ Singular Neutrum oder Maskulinum ohnehin meist nach Präposition vorkommen, wenn nicht gerade das erste Adjektiv ein Pronominaladjektiv ist, haben wir diesen Faktor in der vorliegenden Untersuchung konstant gehalten: Alle ersten Adjektive folgen auf eine Präposition.

4.1 Syntaktische und semantische Einflussgrößen

4.1.1 Möglichkeiten, auf die interne Struktur der Nominalphrase zu schließen

Wenn vor einem artikellosen Nomen mehr als ein attributives Adjektiv (Partizip) steht, sind vor allem drei unterschiedliche Strukturen denkbar (dazu und zu weiteren Analysemöglichkeiten vgl. Eisenberg 2013: Bd. 2, 391). Für die erste dieser drei Strukturen, die gereichte Struktur, steht Beispiel (3):

- (3) zu offenem und offensivem Dialog (Süddeutsche Zeitung, 30.10.2006, S. 11)

Jedes der beiden Adjektive modifiziert dasselbe Nomen. In diesem ersten Fall kann man sich vorstellen, dass es zwei Prädikationen über das vom Nomen Bezeichnete gibt: Ein „offener und offensiver Dialog“ ist ein Dialog, der sowohl offen als auch offensiv ist. Bei dieser Variante kann die Konjunktion *und* durch Komma ersetzt werden: *ein offener, offensiver Dialog*. Eine seltenere Variante dieser Struktur ist die Geminatio, eine rhetorische Figur, die der Intensivierung dient: Das Adjektiv wird wiederholt. In diesem Sonderfall gibt es nur eine Prädikation; Komma und Konjunktion sind nicht austauschbar:

- (4) mit großem, großem Engagement (Rhein-Zeitung, 31.5.2011, S. 6)

Ebenso wenig austauschbar sind Konjunktion und Komma in elliptischen² Strukturen, die wir ebenfalls zu den gereichten Strukturen rechnen (5).

2 Bei unserer Definition der Koordinationsellipse halten wir uns an Hennig (2015: 58): „Es müssen zwei bis auf eine Konstituente gleichen Typs identische Strukturen gleichen Typs koordiniert sein; die identischen Teile werden bei einem der beiden Vorkommen weggelassen.“ Der gleiche Typ ist hier: Nominalphrase mit attributivem Adjektiv (*privatem Leben, beruflichem Leben*). Der identische Teil ist *Leben*. Es liegt also nach Hennigs Klassifikation eine „Ellipse des Bezugsnomens“ (Hennig 2015: 59) vor; vgl. Dudengrammatik (2016: 909): „(c) [Die rote _] und [die schwarze Katze] schlafen.“ Es ist nicht immer leicht, solche Strukturen abzugrenzen von Strukturen mit koordinierten attributiven Adjektiven wie in *Die satte und zufriedene Katze schläft*. In die Interpretation gehen sowohl die Semantik der Adjektive (Antonymie spricht für Ellipse) als auch die Semantik des Bezugsnomens ein (ein Lexem wie *Balance* oder *Pol* spricht für Ellipse). Entscheidend ist u. E., ob nach der semantischen Interpretation eine Referenz auf eine einzige Entität (ein Wesen, ein Ding, z. B. eine Katze) oder eine Referenz auf zwei verschiedene Entitäten (z. B. zwei Katzen, eine rote und eine schwarze) vermutet wird; vgl. Dudengrammatik (2016: 913) sowie 4.1.4. Bei Abstrakta wie *Leben* ist diese Entscheidung schwieriger als bei Konkreta; als Paraphrase mag hier dienen: *zwischen Privatleben und Berufsleben* (jeder hat nur ein

- (5) die Balance zwischen privatem [hinzuzudenken: Leben] und beruflichem Leben (Angela Merkel, Ansprache zur Verabschiedung von Edmund Stoiber, 2007)

Die zweite denkbare Struktur, die gestufte, zeigt Beispiel (6):

- (6) mit riesigem technischen Aufwand (Der Prignitzer, 17.7.2013, S. 23)

Das zweite Adjektiv modifiziert das Nomen, das erste Adjektiv (*riesigem*) modifiziert nicht direkt das Nomen, sondern die gesamte folgende Gruppe aus zweitem Adjektiv und Nomen (*technischen Aufwand*). Die entsprechende Prädikation wäre also ‚Der technische Aufwand ist riesig‘. Von dieser zweiten Struktur wird seit dem 18. Jh. immer wieder vermutet, dass sie eher als die erste Struktur schwache Flexion am zweiten Adjektiv auslöst: „Wenn zwei Beywörter ohne Geschlechtswort, zwischen welchen kein *und* ausgelassen ist, vor einem Hauptworte stehen, so gehen sie so: [...] schönem rothen Weine. [...] schöner rothen Weine. [...] schönem grünen Grase“ (Heynatz 1803³/2006 159 = Wortforschung Nr. 50) (vgl. auch Behaghel 1923: 212–214 sowie den hier untersuchten Faktor *und*-Test-Ergebnis). Wie man an dem Beispiel für den Genitiv Plural sieht, beschränkt sich die schwache Flexion zu dieser Zeit noch nicht auf den Dativ Singular Maskulinum/Neutrum. Nur die Unterscheidung zwischen gereihter und gestufter Struktur ist bereits wirksam – jedenfalls nach Heynatz (1770: 96, 1803/2006 159). Behaghel (1923: 212–214) zeigt allerdings auch ältere Belege für Wechselflexion im Dativ Singular Maskulinum/Neutrum unabhängig von der Struktur. Auch zieht Adelung (1782: 634) die Wechselflexion in allen obliquen Kasus vor, ebenfalls unabhängig von der syntaktischen Struktur. Beides legt nahe, dass der Zusammenhang zwischen Struktur der Nominalphrase und Flexion tatsächlich erst spät entdeckt oder konstruiert (vgl. Moulin 2000: 88) worden ist. Dazu und zum zeitweiligen Überwiegen der schwachen Flexion im 17./18. Jh. vgl. den Überblick in Peter (2013: 188).

Für Beispiel (7) nehmen wir eine dritte Struktur an:

- (7) in angemessenem, das heißt raschem Tempo (Kleine Zeitung, 10.1.1998, o. S.)

Leben und trotzdem haben viele Menschen erstens ein Privatleben und zweitens ein Berufsleben).

- 3 Der *und*-Test wird bereits in der ersten Auflage (Heynatz 1770: 96 = Dritter Theil Nr. 40) unter der Flexionstabelle erwähnt.

Das zweite Adjektiv modifiziert eingrenzend oder paraphrasierend das erste Adjektiv. Das kann explizit gemacht werden etwa durch *und zwar, nämlich, genauer gesagt, das heißt*. Das zweite Adjektiv kann aber auch einen Oberbegriff zum ersten Adjektiv bezeichnen. Explizit machen kann man das z.B. durch *also* oder *das heißt*. In beiden Fällen kann statt der expliziten Formulierung ein Komma stehen:

- (8) Zucker aus deutschem, europäischem Boden (Süddeutsche Zeitung, 10.3.2010, S. 7)

Fälle wie (7) und (8) werden bei unserem *und*-Test zusammengefasst (Wert 2, vgl. Abschnitt 4.1.3).

In Nominalphrasen ohne explizite Koordination durch eine Konjunktion oder eine funktional vergleichbare Wortgruppe ist zu erwarten, dass die erste Struktur wie in (3) am häufigsten ist (vgl. Gallmann 2016: 9), am zweithäufigsten die zweite Struktur wie in (6) und am seltensten die dritte Struktur wie in (7).

Korpuslinguistisch lassen sich die drei Strukturen schlecht voneinander unterscheiden. Deswegen wurde diese Unterscheidung auch nicht als Einflussgröße in die statistischen Modelle aufgenommen. Es gibt aber eine exakt messbare Variable, die Interpunktion (s. Abschnitt 4.1.2), und den Heynatz'schen *und*-Test (s. Abschnitt 4.1.3), dessen Ergebnis sich als Variable operationalisieren lässt. Wir gehen davon aus, dass beide Variablen indirekte Rückschlüsse auf die interne Struktur der Nominalphrase erlauben.

4.1.2 Interpunktion

Nach § 71 (2) der amtlichen Rechtschreibregelung (Rat für deutsche Rechtschreibung 2018) zeigt ein Komma an, dass eine gereimte Struktur (3) gemeint ist, die beiden Adjektive also gleichrangig sind. Die Ergänzungsregel E1 legt noch einmal fest, dass ein Komma in der gestuften Struktur (6) falsch ist – verdeutlicht aber auch, dass es „gelegentlich“ (gemeint: wenn es die Semantik erlaubt) von der Äußerungsabsicht abhängt, ob ein Komma die zweite Struktur kennzeichnen soll. In *mit langem[,] schwarzem Haar* etwa wird beim Schreiben bestimmt, ob ein Komma angemessen ist: Beim Lesen kann man nicht beurteilen, ob ein Kommafehler gemacht wurde (vgl. Moulin 2000: 87). Bei *mit neuen[,] umweltfreundlichen Verfahren* (so das Beispiel aus E1) schließlich würde ein Komma anzeigen, dass ältere Verfahren nicht umweltfreundlich gewesen sind – hier kommt es also wieder auf die Äußerungsabsicht an, die korpuslinguistisch nicht ohne Weiteres erschlossen werden kann. Für die dritte Struktur (7) wird in § 71 der amtlichen Rechtschreibregelung nur auf § 77 (4) verwiesen, der die Kommasetzung für Adjektivgruppen mit expliziter Konjunktion (*also, das heißt, nämlich ...*) regelt:

Unabhängig davon, in welcher Reihenfolge Ober- und Unterbegriff stehen, ist nur ein Komma zwischen den Adjektiven zu setzen und keines zwischen dem zweiten Adjektiv und dem Nomen. Zur Parallel- und Wechselflexion macht die amtliche Regelung keine Aussage.

Im Ergebnis gleich ist die Dudenregelung, die vor der Rechtschreibreform gegolten hat (z.B. Rechtschreibduden 1991: 38). Die frühere Dudenregelung ist allerdings semantikbasiert und behandelt Semantik als etwas Objektivierbares („hängt es vom Sinn des Satzes ab, ob ein Gesamtbegriff vorliegt oder nicht“). Wir gehen davon aus, dass beide zitierten Normen den Gebrauch professioneller Schreiberinnen und Schreiber möglichst genau abbilden sollen.

Weil die Kommasetzung von der Äußerungsabsicht abhängig ist und weil natürlich trotz allen Bemühens um eine gebrauchsbasierte Regelung nicht gesagt ist, dass alle der offiziellen Regelung folgen wollen oder können, werten wir die Interpunktion nur als Indiz für eine mit gewisser Wahrscheinlichkeit intendierte Struktur der Nominalphrase.

Adjektivreihungen mit Komma, die u.E. meist der ersten Struktur aus Abschnitt 4.1.1 entsprechen, weisen eher Parallelflexion auf:

- (9) Denn der Rückverweis an die streitenden Parteien ist der einzig mögliche Weg, um jeden Ruch des Staatsinterventionismus zu vermeiden – und mit **sanftem, indirektem** Druck diese leidige, politisch kontraproduktive Affäre beizulegen. (die tageszeitung, 22.4.2004, S. 11)

Typisch für Strukturen ohne Komma (wir vermuten: mehrheitlich gestufte Strukturen, also die zweite Strukturvariante aus Abschnitt 4.1.1) ist die Wechselflexion:

- (10) Das Gericht hat trotz **fehlendem direkten** Beweis keine Zweifel an der Schuld des Angeklagten. (Die Südostschweiz, 9.6.2012, S. 7)
- (11) Nach **langem juristischen** Tauziehen kam es zu einem Vergleich. (Die Rheinpfalz, 25.3.2013, o. S.)

Parallelflexion kommt in dieser zweiten, nach amtlicher Regelung kommalosen Struktur ebenfalls vor, wenngleich seltener als Wechselflexion (zur Normdiskussion Nübling 2011: 178):

- (12) Die Ausweisung erfolgte nach **langem juristischem** Tauziehen (Berliner Zeitung, 21.6.2005, S. 1)

Ambige Beispiele (13) – zwar mit Komma, aber mit schwacher Flexion – sind noch seltener (vgl. Tab. 1).

Tabelle 1: Ergebnisse aus dem großen Datensatz ATTRIBUTIVE_ADJEKTIVE_2 zum Suchmuster Adjektiv auf *-em* – optionale Konjunktion (ggf. mit Komma) bzw. Komma – Adjektiv auf *-em/-en* – Nomen. (ADJ1: erstes Adjektiv, ADJ2: zweites Adjektiv.)

Muster	stark	in %	schwach	in %	gesamt
ADJ1, <i>also</i> ADJ2	64	97 %	2	3 %	66 = 100 %
ADJ1 <i>bzw.</i> ADJ2	205	95 %	11	5 %	216 = 100 %
ADJ1 <i>und</i> ADJ2	26.713	95 %	1.546	5 %	28.259 = 100 %
ADJ1 <i>beziehungsweise</i> ADJ2	110	93 %	8	7 %	118 = 100 %
ADJ1, <i>aber</i> ADJ2	984	89 %	121	11 %	1.105 = 100 %
ADJ1, ADJ2 (Komma dazwischen)	15.423	79 %	4.038	21 %	19.461 = 100 %
ADJ1 ADJ2 (nichts dazwischen)	24.557	35 %	45.675	65 %	70.232 = 100 %

- (13) Es ist ein kühler Blick auf die Kulissen einer Unterhaltungs- und Freizeitindustrie, auf verlassene Orte in **fahlem, bläulichen** Licht. (Süddeutsche Zeitung, 26.8.2000, S. 21)

Ein Komma erlaubt keine so deutliche Vorhersage der Parallelflexion wie die syndetische Koordination mit einer Konjunktion. In Tab. 1 stellen wir Nominalphrasen mit verschiedenen Konjunktionen (vgl. Meisner 2017), Nominalphrasen mit bloßem Komma und Nominalphrasen mit direkt aufeinanderfolgenden Adjektiven aus dem Datensatz ATTRIBUTIVE_ADJEKTIVE_2 einander gegenüber.

4.1.3 Zusammenspiel zwischen *und*-Test-Ergebnis (möglicher Koordination) und Interpunktion

Um die Merkmale der syntaktisch-semantischen Struktur quantifizierbar zu machen, haben wir den Datensatz ATTRIBUTIVE_ADJEKTIVE_1 von Hand nachannotieren lassen. Dabei wurde versucht, anhand des *und*-Tests eine mögliche Koordination auszuschließen (s. Abschnitt 4.1.1; Heynatz 1803/2006: 159; Eisenberg 2013: Bd. 2, 391). Es wurde jeweils getestet, ob sich zwischen die beiden Adjektive sinnvoll die Konjunktion *und* (*und*-Test-Wert 1, 396 Vorkommen) oder einer der Konnektoren *also*, *d. h.*, *nämlich* (*und*-Test-Wert 2, 48 Vorkommen) oder nichts (*und*-Test-Wert 0, 1.153 Vorkommen) einfügen ließe. Tatsächlich ist bei möglichem *und* eine Tendenz in Richtung der starken Flexion nachweisbar. Beim selteneren Fall, dass *also*, *d. h.* oder Ähnliches eingefügt werden könnte, ist keine Tendenz feststellbar. Eine Geminatio trat nur einmal auf (*mit großem, großem Engagement*; Rhein-Zeitung, 31.5.2011, S. 6) und ging nicht in die Analysen ein.

Diese Vorhersage ist aber nicht so gut wie die Vorhersage durch das Komma, was sich leicht erklären lässt: Gelingt der *und*-Test (Wert 1), ist Koordination zwar möglich – aber nicht unbedingt intendiert. Der einzige greifbare Ausdruck der Schreiberintention ist das Komma. Wenn der *und*-Test gelingt, das Komma aber fehlt, ist das ein Indiz dafür, dass eine gestufte Struktur gemeint war, auch wenn eine gereichte Struktur ebenfalls möglich gewesen wäre:

- (14) *und*-Test 1, kein Komma, schwache Flexion:
- a. in edlem zeitgemäßen Ambiente (Braunschweiger Zeitung, 5.7.2007, o. S.)
 - b. mit kurzem grauen Bart (die tageszeitung, 22.2.2012, S. 5)
- (15) *und*-Test 1, kein Komma, starke Flexion:
- a. in ärmellosem rotem Cocktailkleid (die tageszeitung, 25.6.2008, S. 23)
 - b. in weitem hügelumstelltem Tiefland (Victor Klemperer: Tagebücher 1952, 1999, S. 348)

Die Effektplots in Abb. 3 basieren auf einer Regressionsanalyse zum Datensatz ATTRIBUTIVE_ADJEKTIVE_1. Die Ergebnisse der Regressionsanalyse mit genauen Kennzahlen sind im Anhang dargestellt. Die beiden Plots veranschaulichen noch einmal, dass der Effekt der Interpunktion stärker ist als der des *und*-Test-Ergebnisses. Die Interpunktion erlaubt somit eine zuverlässigere Einschätzung als der *und*-Test.

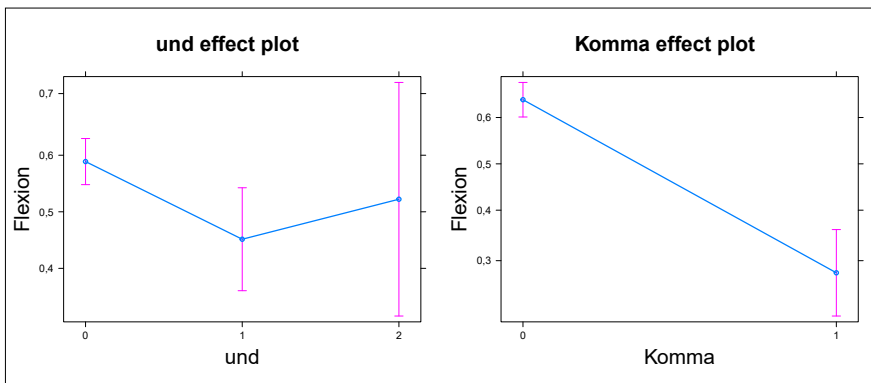


Abbildung 3: Geschätzte Werte der Regressionsanalyse für schwache Flexion des zweiten Adjektivs je nachdem, ob sich zwischen die beiden Adjektive ein Koordinationsmarker einfügen ließe (linker Plot, „und“ ist das *und*-Test-Ergebnis: 0 = kein Marker, 1 = *und*, 2 = anderer Marker) und je nachdem, ob ein Komma zwischen den beiden Adjektiven steht oder nicht (rechter Plot; 0: kein Komma, 1: Komma). (Grundlage: Datensatz ATTRIBUTIVE_ADJEKTIVE_1.) Nach oben ist die Wahrscheinlichkeit für schwache Flexion des zweiten Adjektivs abgetragen (hoher Wert: hohe Wahrscheinlichkeit).

Darüber hinaus haben wir Korrelationen zwischen den verschiedenen Einflussfaktoren überprüft. Erwartungsgemäß gibt es eine starke Korrelation zwischen *und*-Test-Ergebnis und Interpunktion (vgl. Anhang, Abschnitt 6.2).

4.1.4 Semantische Klassen der Adjektive

Wie Nübling (2011: 186) zeigt, hängen die Möglichkeiten der Reihung oder aber der Stufung von den semantischen Klassen der Adjektive ab: Steht an zweiter Stelle ein Zugehörigkeitsadjektiv (Classificativum nach Eichinger 1991: 325) wie *finanziell* oder *ärztlich*, ist die gereichte Struktur unwahrscheinlich. Damit wären in unseren Daten erstens eine Korrelation zwischen *und*-Test-Ergebnis und semantischer Klasse und zweitens eine Korrelation zwischen Kommasetzung und semantischer Klasse zu erwarten gewesen. Diese beiden Korrelationen konnten jedoch im Datensatz `ATTRIBUTIVE_ADJEKTIVE_1` nicht nachgewiesen werden.

Bei der Zuordnung von Lexemen zu semantischen Klassen gibt es immer Entscheidungen, die diskutiert werden und mit guten Gründen individuell unterschiedlich ausfallen können. Unsere Einteilung ist eine Synthese aus Eichinger (1987), Eichinger (1991), Trost (2006), Nübling (2011) und Dudengrammatik (2016: 342–347). Die folgende Übersicht lehnt sich in der Terminologie und bei den Beispielen an die Dudengrammatik an:

- Wörter, die nicht zwischen Artikel und Nomen passen, wie *einige*.⁴
- „relationale Adjektive, die“ Demonstrativa „nahestehen [...] *letztgenannte, besagte, obige, folgende*“. Nicht prädikativ verwendbar.
- „quantifizierende Adjektive, die“ Indefinita „nahestehen: [...] *andere, sonstige, weitere, derartige*“. Nicht oder schlecht prädikativ verwendbar.
- „Ordinalzahladjektive: *zweite, dritte, hundertste*“; außerdem *letzte*. Nur als Nominalisierung prädikativ verwendbar.
- „relationale Adjektive mit Bezug auf [die] räumliche Lage [...]: [...] *dortige, linke*“ (aber nicht ‚politisch links‘), „*vordere, äußere*“. Nicht oder schlecht prädikativ verwendbar.
- relationale Adjektive mit Bezug auf die zeitliche Situierung: „*damalige, heutige, gestrige, morgendliche, diesjährige, einstige*“; *monatlich, täglich*. Nicht oder schlecht prädikativ verwendbar.
- „relationale Adjektive, die die stoffliche Beschaffenheit ausdrücken (und meist von Stoffsubstantiven abgeleitet sind): *silbern, ledern, hölzern*“ (strenger eingegrenzt als in Nübling 2011: 186).
- „relationale Adjektive, die den Bereich [oder] die Zugehörigkeit [...] angeben [...]: *steuerlich, schulisch, technisch, wissenschaftlich, katholisch*. Diese

4 Das kommt im Datensatz `ATTRIBUTIVE_ADJEKTIVE_1` nicht vor, die morphologische Annotation war zuverlässig.

Adjektive sind in der aktuellen⁵ Lesart nicht oder schlecht prädikativ verwendbar und nicht oder schlecht steigerbar. Im Folgenden werden sie „Zugehörigkeitsadjektive (ohne solche, die die geografische Herkunft bezeichnen)“ genannt.

- Adjektive, die die geografische/politische Zugehörigkeit oder die Herkunft angeben und die typischerweise von Eigennamen abgeleitet sind: *französisch*, *bayrisch*.
- qualifizierende Adjektive, die nach Trost (2006: 127) in die Klasse der referenziellen Adjektive rücken: *stark* in *starke Raucherin* oder *neu* in *mein neues Auto*, wenn es sich um einen Gebrauchtwagen handelt. Schlecht prädikativ verwendbar.
- Farbadjektive; zusätzlich solche wie *blond*, *hell*, *dunkel*.
- sonstige Adjektive, die in der aktuellen Lesart auch prädikativ verwendbar und in der Regel steigerbar sind: *groß*, *hoch*, *schwer* etc. Hier wurden insbesondere typische qualifizierende Adjektive eingeordnet.

Die beiden Tests, die am häufigsten zur Abgrenzung der qualifizierenden Adjektive von den Mitgliedern anderer semantischer Klassen eingesetzt wurden, waren Steigerung und Einsatz als Prädikativ.

Für die unterschiedlichen semantischen Klassen aus der vorangehenden Liste gelten verschiedene Wortstellungspräferenzen (s. Münzberg & Bildhauer in dieser Ausgabe). Fasst man diese Klassen wie Münzberg & Bildhauer (in dieser Ausgabe) zu sechs Wortstellungsklassen zusammen, so ergibt sich für jede Wortstellungsklasse ein Wert zwischen 1 (sehr weit rechts, also geringe Distanz zum Nomen) und 6 (am linken Rand, also große Distanz vom Nomen). Adjektive mit Wert 1 (nahe am Nomen) sind in der Regel Substantivderivate, also „substantivverwandt“, Adjektive mit Wert 6 sind „den Artikelwörtern ähnlich“ (Eisenberg 2013: Bd. 2, 393–394). Wenn man nun den Distanzwert des ersten Adjektivs vom Distanzwert des zweiten Adjektivs abzieht, erhält man einen Wert für den Unterschied in der Nomennähe. Für sich betrachtet,⁶ ist dieser Unterschied signifikant für die Flexion des zweiten Adjektivs: Gehören beide Adjektive derselben Distanzklasse an (Distanzwert 0), wird das zweite Adjektiv eher stark flektiert. Schwache Flexion ist wahrscheinlicher bei der erwartbaren Reihenfolge von Adjektiven unterschiedlicher Distanzklassen (positiver Distanzwert). Wir haben den Einflussfaktor in

5 *technisch* in *technischer Aufwand* ist ein Zugehörigkeitsadjektiv. *technisch* in *Das Fußballspiel dieser Mannschaft ist schneller, technischer und moderner geworden* ist ein qualifizierendes Adjektiv.

6 Überprüft anhand eines Chi-Quadrat-Tests und eines Assoziationsplots mit Darstellung der Pearson-Residuen; $\chi^2 = 147.66$, $df = 7$, $p < 0,001$, Cramers $V = 0,31$ (nur Flexion und Distanzwert, ohne weitere Einflussfaktoren).

Tabelle 2: Semantische Klasse und Flexion des zweiten Adjektivs im Datensatz ATTRIBUTIVE_ADJEKTIVE_1 (* = ohne Adjektive, die eine geografische Herkunft bezeichnen; ** = ohne Farbadjektive).

Flexion des zweiten Adjektivs	Zugehörigkeit*		qualifizierend**	
schwach (-n)	483 =	62 %	146 =	31 %
stark (-m)	292 =	38 %	318 =	69 %
	775 =	100 %	464 =	100 %

Tabelle 3: Interpunktion und Flexion bei Zugehörigkeitsadjektiven (ohne Adjektive, die eine geografische Herkunft bezeichnen).

Flexion des zweiten Adjektivs	mit Komma		ohne Komma	
schwach (-n)	5 =	10 %	477 =	66 %
stark (-m)	45 =	90 %	248 =	34 %
	50 =	100 %	725 =	100 %

Tabelle 4: Interpunktion und Flexion bei qualifizierenden Adjektiven (ohne Farbadjektive).

Flexion des zweiten Adjektivs	mit Komma		ohne Komma	
schwach (-n)	47 =	17 %	99 =	52 %
stark (-m)	228 =	83 %	90 =	48 %
	275 =	100 %	189 =	100 %

einer logistischen Regressionsanalyse mit den anderen Prädiktoren überprüft und konnten keine signifikanten Einflüsse auf die Flexionsart feststellen. Die Modellgüte wurde durch die Aufnahme des Faktors nicht verbessert. Auch die semantische Klasse des ersten Adjektivs als Prädiktor zeigt in einem logistischen Regressionsmodell keinen signifikanten Einfluss. Erwartet hatten wir mehr schwache Flexion nach Adjektiven, die Demonstrativa oder Indefinita nahestehen, so wie *folgende* oder *andere*; vgl. Bildhauer et. al. (2019). Für solche Adjektive gibt es aber im Datensatz ATTRIBUTIVE_ADJEKTIVE_1 nur wenige Belege.

Was die semantische Klasse des zweiten Adjektivs angeht, so entsprechen die Zahlenverhältnisse in unserem kleinen Datensatz der von Nübling (2011: 187) beschriebenen Tendenz, dass Zugehörigkeitsadjektive als zweite Adjektive eher schwach flektieren als qualifizierende Adjektive (Tab. 2).

Sobald man aber die Interpunktion einbezieht, scheinen sich Zugehörigkeitsadjektive (Tab. 3) den qualifizierenden Adjektiven (Tab. 4) ähnlicher zu verhalten.⁷

7 Komma JA bei Zugehörigkeitsadjektiven tritt vor allem in den folgenden Fallgruppen auf: Zweifelsfälle bei der Annotation, Kommafehler (*auf sicherem, rechtlichen Boden*), und-Test mit Wert 2 (das zweite Adjektiv modifiziert das erste).

Verrechnet man semantische Klasse und Komma JA/NEIN in einer Regressionsanalyse mit weiteren grammatischen und außergrammatischen Einflussgrößen,⁸ zeigt sich, dass im Datensatz ATTRIBUTIVE_ADJEKTIVE_1 nicht Zugehörigkeitsadjektive, sondern Ordinalzahladjektive mit schwacher Flexion besonders hervorstechen. Auch für Farbadjektive ist schwache Flexion typisch (Tab. 7 im Anhang; ähnlich bei Nübling 2011: 188). Bei der Bewertung dieses Ergebnisses sollte allerdings nicht außer Acht gelassen werden, dass sie sich auf ein bestimmtes Korpus beziehen. Andere Korpora und andere Untersuchungsmethoden (etwa psycholinguistische) können zu anderen Ergebnissen führen, vgl. Schäfer (2017).

Weitere semantische Feinheiten könnten eine Rolle spielen. So stellt Schäfer (2017) einen Einfluss der Präpositionen *von* und *zwischen* fest, verzichtet aber darauf, diesen Einfluss inhaltlich zu interpretieren. (Wir selbst konnten keinen Einfluss der lexikalischen Wahl der Präposition nachweisen). Der Versuch, den von Schäfer (2017) nachgewiesenen Einfluss vielleicht doch grammatisch zu erklären, hat uns auf eine neue Hypothese zum Muster Adjektiv + *und* + Adjektiv + Nomen gebracht: Vor allem bei der Präposition *zwischen* ist wahrscheinlich, dass die beiden Adjektive sich auf unterschiedliche Entitäten beziehen (*zwischen roten und schwarzen Katzen unterscheiden*). Das lässt auf eine Koordination von Nominalphrasen schließen, also auf eine elliptische Struktur, zu der ein Nomen hinzugedacht werden kann. Im Gegensatz dazu scheinen sich in Formulierungen wie *bei sonnigem und trockenen Hochsommerwetter* (dpa, 11.6.2006, o. S.) beide Adjektive gleichermaßen auf eine einzige Entität zu beziehen. Koordiniert werden zwei Adjektivphrasen. Unsere Hypothese ist: Wenn beide Adjektive sich auf eine einzige Entität beziehen (also keine Ellipse vorliegt), wird die schwache Flexion etwas wahrscheinlicher.

Damit soll nicht gesagt werden, dass Ellipsen in den untersuchten Daten nur bei den Präpositionen *von* und *zwischen* auftreten: Bei der Entscheidung, ob eine Ellipse vorliegt oder nicht, müssen vor allem die Semantik der Adjektive (Antonymie?), die des Kopfnomens und eventuell auch die des regierenden Verbs (vgl. *unterscheiden*) berücksichtigt werden. Bei *an fließendem und stehendem Gewässer* (Rhein-Zeitung, 1.4.2006, o. S.) z. B. ist nach *fließendem* das Nomen *Gewässer* eingespart. Bei der Interpretation als Ellipse spielt die Präposition *an* keine Rolle: Die Vereinsmeisterschaft, über die im zitierten Zeitungsartikel berichtet wird, beinhaltet zwei Disziplinen, das Angeln in fließendem Gewässer und das Angeln in stehendem Gewässer. Die Antonymie der Adjektive und der weitere Kontext legen eine Ellipse nahe.

8 Frequenz des ersten Adjektivs, Silbenzahl des zweiten Adjektivs, *und*-Test, /m/ in der Mitte des Substantivs, /m/ am Ende des Substantivs.

Zwar ist bei Adjektiven nach der Konjunktion *und* die schwache Flexion insgesamt sehr selten (vgl. Tab. 1 in Abschnitt 4.1.2). Dennoch lässt sich ein kleiner Unterschied zwischen Nominalphrasen mit Ellipse und Nominalphrasen ohne Ellipse erkennen. Die Ergebnisse zu einer Stichprobe aus dem Datensatz ATTRIBUTIVE_ADJEKTIVE_2 sind in Tab. 5 aufgeführt.

Tabelle 5: Ellipse und Flexion bei Koordination mit *und*.

Flexion des zweiten Adjektivs (Muster X <i>und</i> Y + Nomen)	Ellipse		keine Ellipse	
schwach (-n)	67 =	5 %	53 =	9 %
stark (-m)	1.213 =	95 %	559 =	91 %
	1.280 =	100 %	612 =	100 %

In Konkurrenz zu den elliptischen Konstruktionen in unserem Datensatz ATTRIBUTIVE_ADJEKTIVE_2, die alle nur jeweils eine einzige Präposition enthalten (*an fließendem und stehendem Gewässer angeln*), stehen die oft leserfreundlicheren, aber weniger ökonomischen und auch nicht immer zulässigen Konstruktionen mit zwei Präpositionen (*an fließendem und an stehendem Gewässer angeln*, aber **Balance zwischen privatem und zwischen beruflichem Leben*). Die Bedingungen für das Auftreten oder den Wegfall der zweiten Präposition beschreibt Lotzow (2015). Die Frage, ob eine Ellipse der Präposition vorliegt (*an fließendem und [an] stehendem Gewässer angeln*) oder nicht (*Balance zwischen privatem und beruflichem Leben*; vgl. Lotzow 2015: 91), möchten wir trennen von der Frage, ob eine Ellipse des von der Präposition regierten Nomens (*Gewässer, Leben*) vorliegt. Wir vermuten, dass eine zweite Präposition die schwache Flexion des zweiten Adjektivs noch wesentlich unwahrscheinlicher macht als die von uns untersuchte Ellipse des Bezugsnomens: *an fließendem und an ?stehenden Gewässer angeln*). Eine Untersuchung steht noch aus.

Dass es für die Flexion einen Unterschied macht, ob die beiden Adjektive jeweils intersektiv sind oder nicht, konnte nicht nachgewiesen werden. Für die Wortstellung spielt dieser Faktor jedoch durchaus eine Rolle, s. Münzberg & Bildhauer (in dieser Ausgabe).

4.1.5 Partizip vs. Nichtpartizip und Genus des Substantivs

Ebenso wenig signifikant werden in den Regressionsanalysen Unterschiede zwischen (lexikalisierten oder nicht lexikalisierten) Partizipien und Nichtpartizipien bei den beiden Adjektiven und der Faktor Genus des Substantivs.

4.2 Phonologische Eigenschaften der Adjektive und des Substantivs

Nach Niebuhr (2017) erzeugen /m/ in der Mitte und vor allem /m/ am Ende des Substantivs⁹ einen Trend zur starken Flexion des zweiten Adjektivs: *in elegantem blauem Kostüm* (Rhein-Zeitung, 30.10.2008, o. S.). Die Vorhersage trifft auf beide Datensätze zu, der Unterschied ist aber nur im großen Datensatz ATTRIBUTIVE_ADJEKTIVE_2 signifikant.

Die Regressionsanalyse zeigt, dass die Einflussgröße „Silbenzahl des zweiten Adjektivs“¹⁰ im größeren Datensatz signifikant ist, im kleineren Datensatz ATTRIBUTIVE_ADJEKTIVE_1 ist sie hingegen bei der Verrechnung mit anderen Prädiktoren (inkl. der semantischen Klasse des zweiten Adjektivs und des *und*-Test-Ergebnisses) nicht signifikant. Das Sprachmaterial, das zwischen der letzten starken Endung und der fraglichen Flexionsendung steht, scheint jedenfalls mehr Einfluss in Richtung der starken Flexion zu haben als die Silbenzahl des ersten Adjektivs. Eine ähnliche Tendenz gibt es bei der Erweiterung der Adjektive durch Adverbien und Partikeln im Datensatz ATTRIBUTIVE_ADJEKTIVE_1: Die Erweiterung des ersten Adjektivs hat keinen Einfluss auf die Flexion des zweiten Adjektivs, während die Erweiterung des zweiten Adjektivs – also wiederum Material zwischen der letzten starken Endung und der fraglichen Endung – in Richtung der starken Flexion wirken kann:

- (16) a. von individuellem, **oft** skurrile**m** Zubehör (Süddeutsche Zeitung, 15.11.2004, S. 2)
 b. Nach monatelangem **sehr** harte**m** Training (Rhein-Zeitung, 15.1.1996, o. S.)

Im Zusammenspiel mit den anderen Prädiktoren ist dieser Trend aber nicht mehr signifikant.

9 Ermittelt mithilfe eines einfachen Algorithmus, der <mm> als ein /m/ rechnet, entsprechend <mm> am Ende des Substantivs als /m/ am Ende. In der Mitte des Substantivs können mehrere /m/ an verschiedenen Stellen stehen.

10 Ermittelt mithilfe eines Algorithmus, der grundsätzlich pro Vokalbuchstaben eine Silbe ausgibt, wobei aber bestimmte Vokalbuchstabenkombinationen als Zeichen für nur eine Silbe gerechnet werden.

4.3 Frequenzeffekte

Für die Frequenz¹¹ des ersten Adjektivs lautete die Hypothese: Je häufiger das erste Adjektiv ist, desto häufiger erscheint schwache Flexion am zweiten Adjektiv. Dahinter stand der Gedanke, dass die Häufigkeit des ersten Adjektivs mit weiteren Eigenschaften determiniererähnlicher Adjektive korrelieren könnte. Tatsächlich zeigt sich in beiden Datensätzen ein (schwacher) Effekt in die erwartete Richtung.¹² Die Frequenzeffekte überlagern aber nicht die semantisch-syntaktischen Faktoren.

Im Datensatz `ATTRIBUTIVE_ADJEKTIVE_1` ist auch die substantivspezifische Adjektivfrequenz berücksichtigt (s. Münzberg & Bildhauer in dieser Ausgabe). Bei den Lemmata *technisch* und *Gerät* etwa kommt die Kombination aus attributivem Adjektiv (ohne ein vorangehendes anderes attributives Adjektiv) und Nomen 5.114-mal vor, während *Gerät* allein 289.091-mal gefunden wird. (Gesucht wurden Lemmata bzw. Stämme, nicht Wortformen.) Damit ergibt sich als substantivspezifische Adjektivfrequenz ein Wert von $5.114 : 289.091 = 0,01769$. Für *technisch* und *Selbstbewusstsein* liegt der Wert nur bei 0,00001. Je höher der Wert ist, desto eher tritt Wechselflexion ein.

4.4 Geografische Verteilung: Land und Region

Sprachgeografische Varianz ist für die Adjektivflexion von Bedeutung. Nicht nur im großen, auch im kleineren Datensatz sind für die Schweiz im Vergleich zu den anderen Ländern signifikante Abweichungen zu verzeichnen. Die Grafik in Abb. 4 wurde basierend auf dem größeren Datensatz `ATTRIBUTIVE_ADJEKTIVE_2` aus der Regressionsanalyse (siehe Anhang) erzeugt. Das bedeutet u. a.: Wenn in den Texten aus der Schweiz die beiden Adjektive häufiger koordiniert sind, ist das bereits berücksichtigt und kann das Ergebnis nicht verfälschen.

11 Ermittelt anhand der DeReWo-Frequenzklassen, s. Belica et al. (2012). „Dabei hat eine Grundform die Häufigkeitsklasse N, wenn die häufigste Form etwa 2^N -mal häufiger vorkommt als diese Form. Für die Grundformenliste ist der Eintrag mit der höchsten Frequenz ‚der, die, das‘“ (Belica et al. 2013: Kap. 2.5). Das bedeutet: Je höher die Frequenzklasse, desto seltener das Wort.

12 Zum Zeitpunkt der Berechnung lagen nicht für alle Adjektive Frequenzklassen vor. Im Datensatz `ATTRIBUTIVE_ADJEKTIVE_1` wurden für die Untersuchung der Adjektivreihenfolge (Münzberg & Bildhauer in diesem Band) alle fehlenden Frequenzklassen nachgetragen. Eine Neuberechnung würde möglicherweise einen stärkeren Effekt zeigen. Da allerdings auch der Frequenzeffekt im Datensatz `ATTRIBUTIVE_ADJEKTIVE_2` schwach war, haben wir auf eine Neuberechnung verzichtet.

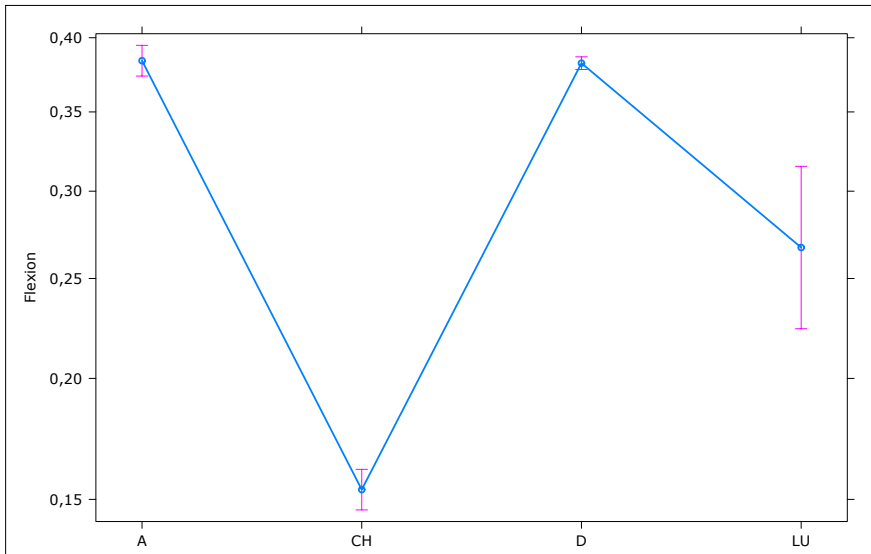


Abbildung 4: Modellschätzwerte aus der logistischen Regressionsanalyse für schwache Flexion des zweiten Adjektivs im Datensatz `ATTRIBUTIVE_ADJEKTIVE_2` nach Ländern. Nach oben ist die Wahrscheinlichkeit für schwache Flexion abgetragen (hoher Wert: hohe Wahrscheinlichkeit).

Betroffen sind direkt aufeinanderfolgende Adjektive ohne Komma oder Konjunktion.¹³ Hier ist in der Schweiz die starke Flexion überrepräsentiert, während sie in den anderen Ländern unterrepräsentiert ist. Dazu passend finden sich etwa im Datensatz `ATTRIBUTIVE_ADJEKTIVE_1` fünf Belege aus Deutschland für *mit anschließendem gemütlichen Beisammensein* (nur schwache Flexion des zweiten Adjektivs), während die Schweizer Daten zwei Belege für *mit anschliessendem gemütlichem Beisammensein* (nur starke Flexion) enthalten. Bei mit Komma verbundenen Adjektiven hingegen ist in allen Ländern gleichermaßen, nicht nur in der Schweiz, die starke Flexion überrepräsentiert.

Um zu überprüfen, ob dieser Unterschied nur Herkunftsländer betrifft oder auch Regionen, haben wir die Zahlen aus der angrenzenden Region Südwestdeutschland mit alemannischen Basisdialekten herangezogen. In den Assoziationsplot aus dem Datensatz `ATTRIBUTIVE_ADJEKTIVE_2` (Abb. 5) sind allerdings nur die Einflussgrößen Komma und Region abgetragen; die Chi-Quadrat-basierte Berechnung berücksichtigt also keine anderen Einflussfaktoren. Wie beim Vergleich der Länder untereinander zeigen sich Unterschiede bei ohne Komma

13 Rechnerisch und visuell überprüft mithilfe eines Assoziationsplots, der die Pearson-Residuen darstellt, zum Datensatz `ATTRIBUTIVE_ADJEKTIVE_1` mit den Einflussfaktoren Komma JA/NEIN und Land ($p < 0,001$).

direkt aufeinanderfolgenden Adjektiven, während sich bei mit Komma verbundenen Adjektiven keine Unterschiede zwischen den Regionen ausmachen lassen. Die beiden auffälligen Regionen sind Südwest (SW) und Mittelost (MO). Hier ist in Folgen ohne Komma die starke Flexion überrepräsentiert.

Die Vermutung, dass bei der Schweizer und südwestdeutschen Präferenz für Parallelflexion auch Lautsystem und Morphologie der Basisdialekte durchscheinen (Gallmann 2016: 2), liegt nahe. Abweichungen in der Region Mittelost (Thüringen, Sachsen) sind darüber aber kaum zu erklären. Peter (2013: 202) stellt die Hypothese auf, das „vergleichsweise häufige Auftreten der Parallelflexion in Adjektivreihungen in der Schweiz“ werde „vermutlich also zusätzlich durch die stärkere (präskriptive) Normorientierung der Schweizer Schreiberinnen und Schreiber begünstigt“. Unsere Ergebnisse bedeuten nicht, dass länderspezifisches Normbewusstsein (vgl. für die Schweiz Heuer et al. 2017: 139, die nur Parallelflexion akzeptieren) beim Schreiben und im Lektorat keine Rolle für die Adjektivflexion spielen würde – beides, der Basisdialekt und das Normbewusstsein, kann zusammenwirken.

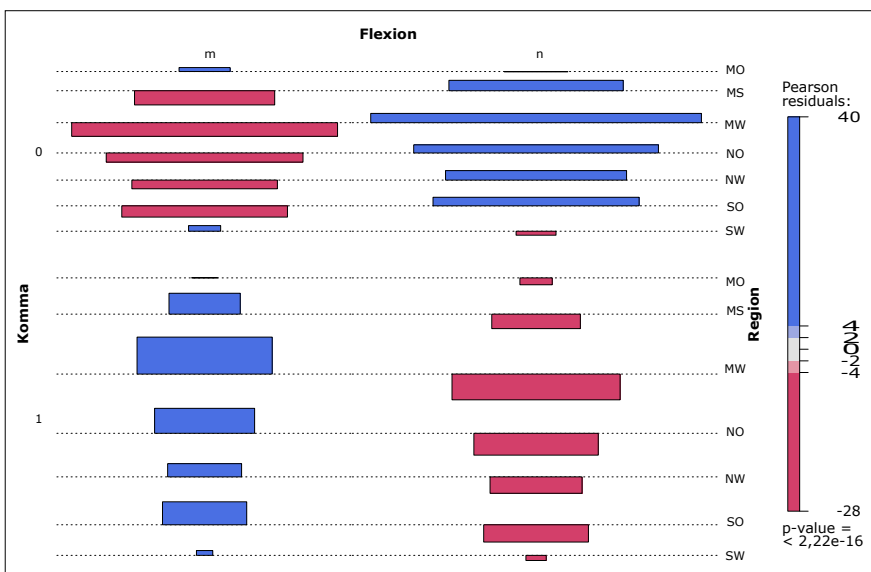


Abbildung 5: Assoziationsplot, der die Pearson-Residuen zur Verteilung starker und schwacher Flexion nach Regionen darstellt; $n = 48.727$ (ohne Region „überregional“, ohne Region „nicht zuzuordnen“ und ohne Belege aus der Schweiz, aus Österreich und aus Luxemburg. SW = Südwest, MO = Mittelost, MS = Mittelsüd, MW = Mittelwest, NO = Nordost, NW = Nordwest, SO = Südost). Der Unterschied zwischen diesen beiden Regionen einerseits und den übrigen Regionen andererseits zeigt sich im Fall Komma = 0 (kein Komma zwischen den beiden Adjektiven) und den Regionen MO und SW. Blaue Balken stehen für ‚überrepräsentiert‘, rote für ‚unterrepräsentiert‘. (Grundlage: Datensatz ATTRIBUTIVE_ADJEKTIVE_2.)

4.5 Sprachwandel (Einflussgröße Jahr) und Verwendungskontext (Einflussgrößen Medium und Domäne)

Schließlich zeigt die Entwicklung über die letzten hundert Jahre hinweg für den Dativ Singular Maskulinum/Neutrum (und nur für diese Merkmalskombination) im Datensatz `ATTRIBUTIVE_ADJEKTIVE_2` wieder einen Anstieg der im 17./18. Jahrhundert so geläufigen (vgl. Peter 2013: 188) Wechselflexion. Das wird aus Abb. 6 ersichtlich.

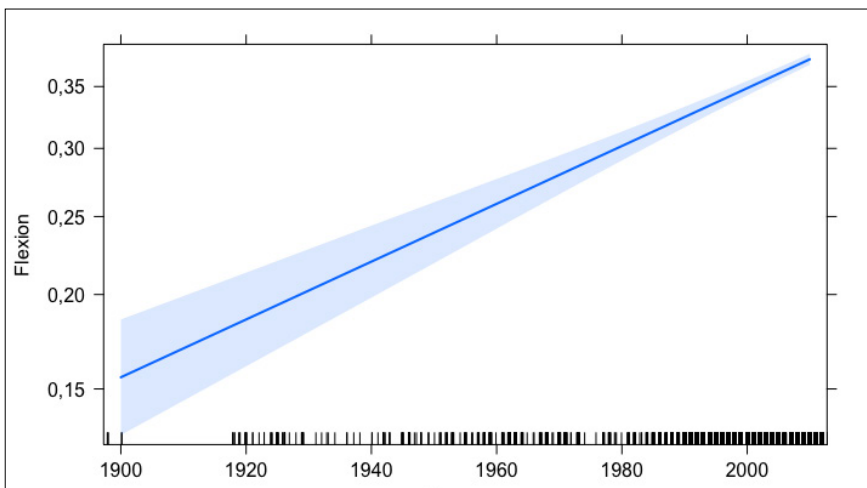


Abbildung 6: Modellschätzwerte aus der logistischen Regressionsanalyse zur Entwicklung hin zur Wechselflexion. Über die Jahre hinweg wird die schwache Flexion des zweiten Adjektivs immer wahrscheinlicher. (Grundlage: Datensatz `ATTRIBUTIVE_ADJEKTIVE_2`.)

Dieser Anstieg ist besonders steil in der Schweiz, es findet also eine Angleichung statt.

Im kleinen Datensatz `ATTRIBUTIVE_ADJEKTIVE_1` und in einer Stichprobe aus dem Datensatz `ATTRIBUTIVE_ADJEKTIVE_2`, die genauso klein ist wie der Datensatz `ATTRIBUTIVE_ADJEKTIVE_1`, deutet sich bei der Regressionsanalyse dieselbe Richtung des Einflusses an – also immer mehr schwache Flexion –, wobei der Prädiktor Jahr in der Stichprobe nicht signifikant ist.

Vielleicht der Normdiskussion und dem professionellen Lektorat zuzurechnen ist die im Datensatz `ATTRIBUTIVE_ADJEKTIVE_2` auffällige Tendenz zur Parallelflexion beim Medium Buch und in der Domäne Fiktion; damit ist die starke Endung *-m* auch im Register literarische Texte überrepräsentiert.

5 Zusammenfassung

Auf die Flexion des zweiten Adjektivs wirken sprachsystematische (Syntax, Semantik, Phonologie, Frequenz) und extralinguistische Einflussgrößen (geografische Verteilung, Alter des Belegs, Verwendungskontext). Regressionsanalysen erlauben es, mehrere Einflussgrößen zugleich zu berücksichtigen.

Abbildung 7 fasst die Ergebnisse aus dem Datensatz `ATTRIBUTIVE_ADJEKTIVE_1` zusammen. Das Schaubild zeigt geschätzte Koeffizienten aus einem logistischen Regressionsmodell, in dem die Werte der Prädiktoren zur besseren Vergleichbarkeit binär kodiert sind. Die genauen Koeffizienten und Ergebnisse des Regressionsmodells sind im Anhang dargestellt. Der Plot ist wie folgt zu lesen: Je weiter ein Punkt von der senkrechten Mittellinie (0) entfernt ist, desto stärker ist der Effekt. Nach links sind negative Koeffizienten abgetragen, die die schwache Flexion des zweiten Adjektivs weniger wahrscheinlich machen, nach

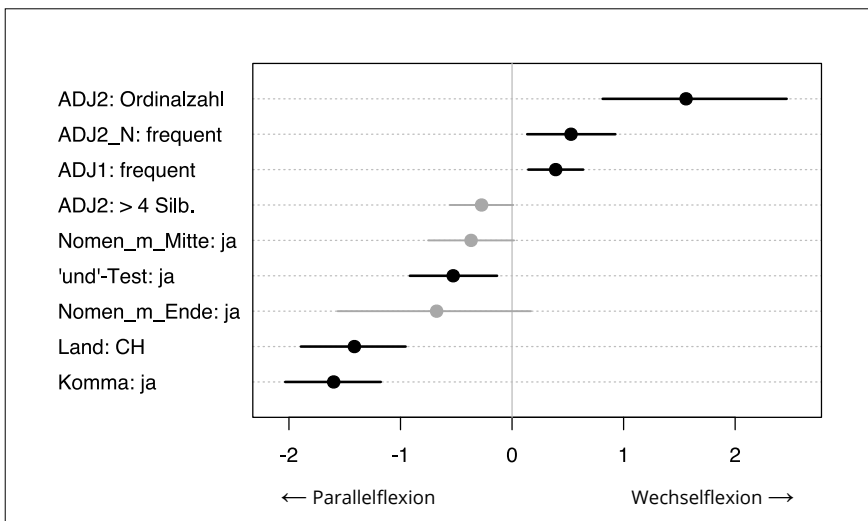


Abbildung 7: Geschätzte Koeffizienten aus einem logistischen Regressionsmodell zum Datensatz `ATTRIBUTIVE_ADJEKTIVE_1`. Legende: **ADJ2**: zweites Adjektiv. **ADJ1**: erstes Adjektiv. **ADJ2: Ordinalzahl**: Das zweite Adjektiv ist eine Ordinalzahl. **ADJ2_N: frequent**: Das zweite Adjektiv kommt außerhalb des vorliegenden Satzes besonders häufig als attributives Adjektiv zum betreffenden Substantiv vor (Häufigkeit der Adjektiv-Substantiv-Kombination geteilt durch Häufigkeit des Substantivs $\geq 0,02$). **ADJ1: frequent**: DeReWo-Frequenzklasse des ersten Adjektivs ≤ 8 , d. h., das erste Adjektiv ist besonders häufig. **ADJ2 > 4 Silb.**: Das zweite Adjektiv hat über 4 Silben (Flexionssuffix mitgerechnet); hier nicht signifikant, wohl aber im großen Datensatz (siehe Anhang). **Subst_m_Mitte, Subst_m_Ende**: /m/ in der Mitte des Substantivs (nicht signifikant, aber im großen Datensatz [siehe Anhang]) bzw. /m/ am Ende des Substantivs (nicht signifikant, aber im großen Datensatz [siehe Anhang]). **und-Test**: Zwischen die beiden Adjektive lässt sich *und* oder etwas wie *d. h., und zwar* einfügen.

rechts positive Koeffizienten, die die Wahrscheinlichkeit für schwache Flexion erhöhen. Die waagerechten Linien stehen für 95 %-Konfidenzintervalle.

Die stärksten Effekte haben also das Komma als Symptom einer möglichen Koordination, die Ausprägung „Schweiz“ des Faktors Land und die Ausprägung „Ordinalzahl“ des Faktors semantische Klasse des zweiten Adjektivs: In NPs mit Komma (*mit kurzem, rotbraunem Haar*; Süddeutsche Zeitung, 12.3.2002, S. 14) sowie in Schweizer Belegen (*mit dichtem schwarzem Haar*; St. Galler Tagblatt, 26.6.1999, o. S.) überwiegt Parallelflexion, ansonsten ist Wechselflexion häufiger (*mit dichtem schwarzen Haar*; Süddeutsche Zeitung, 23.4.2001, S. 15). Besonders Ordinalzahladjektive werden eher schwach flektiert (*nach verlorenem ersten Satz steigerte er sich jedoch und gewann verdient*; St. Galler Tagblatt, 22.5.2000, o. S.).¹⁴

14 Fast alle Belege zu Ordinalzahladjektiven an zweiter Stelle entstammen solchen Sportnachrichten.

6 Anhang

6.1 Modellspezifikation

Für die vermuteten Einflussfaktoren wurden logistische Regressionsanalysen durchgeführt. Die statistischen Analysen beziehen sich entweder auf den Datensatz `ATTRIBUTIVE_ADJEKTIVE_1` oder auf den Datensatz `ATTRIBUTIVE_ADJEKTIVE_2`. Der Bezug zum Datensatz wird an den jeweiligen Stellen im Text gekennzeichnet. Die Daten der beiden Datensätze wurden jeweils in mehreren logistischen Regressionsanalysen mit festen Effekten modelliert (generalisierte lineare Modelle; vgl. Jaeger 2008). Dabei wurde in Abhängigkeit der verschiedenen Prädiktoren die Wahrscheinlichkeit modelliert, dass in einem gegebenen Fall starke oder schwache Flexion auftritt. Der Einfluss einzelner Prädiktoren (unabhängig von anderen Prädiktoren) auf diese Wahrscheinlichkeit kann dabei jedoch nicht quantifiziert werden. Stattdessen wird der Einfluss einzelner Prädiktoren auf die logarithmierte Chance, das Logit, geschätzt.

Die Einflussfaktoren für beide Datensätze werden in Abschnitt 3 beschrieben. Ausgeschlossene Faktoren trugen entweder nicht zur Verbesserung des Modells bei oder wiesen eine zu starke Korrelation mit einem anderen Faktor auf. Damit ergeben sich folgende Modellspezifikationen:

Datensatz `ATTRIBUTIVE_ADJEKTIVE_2`

$$P(\text{Flexion} = \text{schwach}) = \text{logit}^1 (\beta_0 + \text{Conj, also} \times \beta_1 + \text{Conj0} \times \beta_2 + \text{Conj1} \times \beta_3 + \text{Conjbeziehungsweise} \times \beta_4 + \text{Conjzbzw.} \times \beta_5 + \text{Conjund} \times \beta_6 + \text{CO_DOMAIN1} \times \beta_7 + \text{CO_DOMAIN2} \times \beta_8 + \text{CO_DOMAIN3} \times \beta_9 + \text{CO_DOMAIN4} \times \beta_{10} + \text{CO_DOMAIN5} \times \beta_{11} + \text{CO_COUNTRYCH} \times \beta_{12} + \text{CO_COUNTRYD} \times \beta_{13} + \text{CO_COUNTRYLU} \times \beta_{14} + \text{scale(CO_YEAR)} \times \beta_{15} + \text{Silb2} \times \beta_{16} + \text{ADJ1.Lemma.FC} \times \beta_{17} + \text{NN_m_Mitte} \times \beta_{18} + \text{NN_m_Ende1} \times \beta_{19})$$

Datensatz `ATTRIBUTIVE_ADJEKTIVE_1`

$$P(\text{Flexion} = \text{schwach}) = \text{logit}^1 (\beta_0 + \text{LandCH} \times \beta_1 + \text{LandD} \times \beta_2 + \text{LandLU} \times \beta_3 + \text{Sem2ref-temp} \times \beta_4 + \text{Sem2rel-ber} \times \beta_5 + \text{Sem2rel-geo} \times \beta_6 + \text{Sem2quali-farb} \times \beta_7 + \text{Sem2quali-rel} \times \beta_8 + \text{Sem2ref-lok} \times \beta_9 + \text{Sem2ordi} \times \beta_{10} + \text{und1} \times \beta_{11} + \text{und2} \times \beta_{12} + \text{Komma1} \times \beta_{13} + \text{Frequ1} \times \beta_{14} + \text{NN_m_Mitte} \times \beta_{15} + \text{NN_m_Ende1} \times \beta_{16} + \text{Silb2} \times \beta_{17} + \text{f_ADJ2_N} \times \beta_{18})$$

Auflösung der Abkürzungen in der Reihenfolge ihres Vorkommens:

- Flexion: Flexionsendung des zweiten Adjektivs (stark oder schwach)
- Conj: Konjunktion (*aber, also* mit vorangehendem Komma, *beziehungsweise, bzw., und*), Komma (Wert 1) oder weder Konjunktion noch Komma (Wert 0) zwischen den beiden Adjektiven
- CO_DOMAIN: Domäne Fiktion (1), Kultur/Unterhaltung (2), Mensch/Natur (3), Politik/Wirtschaft/Gesellschaft (4), Technik/Wissenschaft (5) oder unklassifizierbar (0)
- CO_COUNTRY: Österreich (A), Schweiz (CH), Deutschland (D), Luxemburg (LU)
- CO_YEAR: Erscheinungsjahr
- Silb2: Anzahl der Silben des zweiten Adjektivs
- ADJ1.Lemma.FC: Frequenzklasse des ersten Adjektivs
- NN_m_Mitte: Anzahl der m-Phoneme zwischen erstem und letztem Phonem des Nomens
- NN_m_Ende: Endung des Nomens auf Phonem /m/ oder nicht
- Land: Österreich (A), Schweiz (CH), Deutschland (D), Luxemburg (LU)
- Sem2: semantische Klasse des zweiten Adjektivs nach einem 11-Klassen-System. Werte: art-dem (Adjektiv, das einem Demonstrativum nahekommmt: *einige*; kommt im Datensatz nicht vor); art-indef (Adjektiv, das einem Indefinitum nahekommmt: *besagt*); ordi (Ordinalzahladjektiv: *zweite*); quali-farb (Farbadjektiv: *rot*); quali-rel (ursprünglich qualitatives Adjektiv in relationaler Lesart, vgl. *stark* in *starke Raucherin*); quali-sonst (sonstiges qualifizierendes Adjektiv: *groß*); ref-lok (referenzielles Adjektiv mit lokalem Bezug: *dortig*); ref-temp (referenzielles Adjektiv mit temporalem Bezug: *damalig*); rel-ber (Zugehörigkeitsadjektiv, das weder eine geografische Herkunft noch eine materielle Beschaffenheit bezeichnet: *finanziell*); rel-geo (Zugehörigkeitsadjektiv, das eine geografische Herkunft oder Zugehörigkeit bezeichnet: *sächsisch*); rel-material (Adjektiv, das eine materielle Beschaffenheit bezeichnet: *hölzern*)
- Komma: Komma (Wert 1) oder kein Komma (Wert 0) zwischen den beiden Adjektiven
- Frequ1: Frequenzklasse des ersten Adjektivs
- f_ADJ2_N: substantivspezifische Adjektivfrequenz (Häufigkeit der Adjektiv-Substantiv-Kombination geteilt durch Häufigkeit des Substantivs im gesamten Untersuchungskorpus)

6.2 Ergebnisse

6.2.1 Datensatz: ATTRIBUTIVE_ADJEKTIVE_2

Tabelle 6: Regressionsanalyse zu ATTRIBUTIVE_ADJEKTIVE_2 (Nagelkerkes $R^2 = 0,41$)

Faktor/Ausprägung	geschätzter Koeffizient	Std.-Fehler	z	p
(Intercept)	-0,995	0,111	-8,928	< 2e-16 ***
Conj, also	-1,861	1,000	-1,860	0,063 .
Conj0	2,763	0,099	27,974	< 2e-16 ***
Conj1	0,787	0,100	7,868	3,60e-15 ***
Conjbeziehungsweise	-0,446	0,382	-1,168	0,243
Conjbnw.	-0,810	0,340	-2,383	0,017 *
Conjund	-0,719	0,102	-7,053	1,75e-12 ***
CO_DOMAIN1	-0,926	0,131	-7,054	1,73e-12 ***
CO_DOMAIN2	0,235	0,037	6,390	1,66e-10 ***
CO_DOMAIN3	0,129	0,060	2,167	0,030 *
CO_DOMAIN4	0,159	0,037	4,317	1,58e-05 ***
CO_DOMAIN5	0,124	0,044	2,803	0,005 **
CO_COUNTRYCH	-1,235	0,037	-33,665	< 2e-16 ***
CO_COUNTRYD	-0,007	0,024	-0,300	0,765
CO_COUNTRYLU	-0,538	0,121	-4,430	9,42e-06 ***
scale(CO_YEAR)	0,078	0,008	10,312	< 2e-16 ***
Silb2	-0,183	0,006	-29,220	< 2e-16 ***
ADJ1.Lemma.FC	-0,050	0,002	-23,505	< 2e-16 ***
NN_m_Mitte	-0,119	0,023	-5,140	2,75e-07 ***
NN_m_Ende1	-0,352	0,044	-7,942	1,99e-15 ***

6.2.2 Datensatz: ATTRIBUTIVE_ADJEKTIVE_1

Tabelle 7: Regressionsanalyse zu Datensatz ATTRIBUTIVE_ADJEKTIVE_1 (Nagelkerkes $R^2 = 0,29$)

Faktor/Ausprägung	geschätzter Koeffizient	Std.-Fehler	z	p	
(Intercept)	1,476e+00	3,817e-01	3,867	0,000110	***
LandCH	-1,309e+00	2,926e-01	-4,472	7,74e-06	***
LandD	1,330e-01	1,902e-01	0,700	0,484	
LandLU	-1,268e+00	1,048e+00	-1,209	0,227	
Sem2ref-temp	1,880e-01	6,533e-01	0,288	0,773	
Sem2rel-ber	2,204e-01	1,869e-01	1,179	0,238	
Sem2rel-geo	2,258e-01	2,644e-01	0,854	0,393	
Sem2quali-farb	5,447e-01	2,777e-01	1,961	0,050	*
Sem2quali-rel	1,088e+00	6,420e-01	1,695	0,090	.
Sem2ref-lok	7,097e-01	5,133e-01	1,383	0,167	
Sem2ordi	8,374e-01	5,296e-01	1,581	0,114	
und1	-5,567e-01	2,312e-01	-2,408	0,016	*
und2	-2,699e-01	4,477e-01	-0,603	0,547	
Komma1	-1,515e+00	2,267e-01	-6,681	2,38e-11	***
Frequ1	-7,128e-02	2,006e-02	-3,553	0,0004	***
NN_m_Mitte	-3,843e-01	1,897e-01	-2,025	0,043	*
NN_m_Ende1	-6,296e-01	4,385e-01	-1,436	0,151	
Silb2	-9,513e-02	5,974e-02	-1,592	0,111	
f_ADJ2_N	3,426e-05	1,364e-05	2,512	0,012	*

Außerdem wurde überprüft, ob es Korrelationen zwischen den Faktoren gab, die in die Regressionsanalysen eingingen. Lediglich zwischen den Faktoren „Komma“ und „und“ konnte eine starke Korrelation nachgewiesen werden (Cramers $V = 0,76$).

Um in Abschnitt 5 die Effekte durch die Einflussfaktoren vergleichbarer zu machen und die Erkenntnisse zusammenzufassen, wurden die Werte der einzelnen Einflussfaktoren in binäre Kategorien zusammengefasst. In einem weiteren logistischen Regressionsmodell wurden die Daten mit den binär kodierten Einflussfaktoren als feste Effekte modelliert. In der zusammenfassenden Darstellung in Abb. 7 sind die geschätzten Koeffizienten aus diesem Regressionsmodell dargestellt. Die Ergebnisse sind in Tab. 8 festgehalten.

Tabelle 8: Regressionsanalyse zu Datensatz ATTRIBUTIVE_ADJEKTIVE_1 mit binär kodierten Werten der Einflussfaktoren (Nagelkerkes $R^2 = 0,27$).


Faktor/Ausprägung	geschätzter Koeffizient	Std.-Fehler	z	p
(Intercept)	0,219	0,209	1,043	0,300
LANDLand_CH	-1,413	0,239	-5,920	3,23e-09 ***
Komma1	-1,599	0,218	-7,329	2,32e-13 ***
UNDUND_JA	-0,528	0,199	-2,650	0,008 **
SEM2Sem2_ORDI	1,560	0,415	3,759	0,0002 ***
FREQU1FREQU1_8	0,391	0,125	3,119	0,002 **
NN_M_MITTE_NN_m_Mitte_JA	-0,367	0,196	-1,872	0,061 .
NN_M_ENDE1	-0,676	0,438	-1,545	0,122
SILB2Silb2_groesser_4	-0,273	0,145	-1,884	0,060 .
FADJ2Nf_Adj2_N_hoch	0,527	0,201	2,627	0,009 **

Auflösung der Abkürzungen, die in den anderen beiden Modellen nicht vorkommen:

- UNDUND_JA: *und*-Test: Zwischen die beiden Adjektive lässt sich *und* oder etwas wie *d. h., und zwar* einfügen
- LANDLandCH: Beleg stammt aus der Schweiz
- FREQU1FREQU1_8: DeReWo-Frequenzklasse des ersten Adjektivs kleiner gleich 8, d. h., das erste Adjektiv ist besonders häufig.
- SILB2Silb2_groesser_4: Das zweite Adjektiv hat über 4 Silben (Flexionssuffix mitgerechnet)
- FADJ2Nf_Adj2_N_hoch: Das zweite Adjektiv kommt außerhalb des vorliegenden Datensatzes besonders häufig als attributives Adjektiv zum betreffenden Substantiv vor (Häufigkeit der Adjektiv-Substantiv-Kombination geteilt durch Häufigkeit des Substantivs $\geq 0,02$)

ORCID®

Franziska Münzberg  <https://orcid.org/0000-0002-4780-8440>

Sandra Hansen  <https://orcid.org/0000-0003-4243-0220>

Literatur

- Adelung, Johann C. 1782. *Umständliches Lehrgebäude der Deutschen Sprache, zur Erläuterung der Deutschen Sprachlehre für Schulen*. Leipzig: Breitkopf.
- Behaghel, Otto. 1923. *Deutsche Syntax I* (Germanische Bibliothek: Abteilung 1, Sammlung germanischer Elementar- und Handbücher: Reihe 1, Grammatiken 10,1). Heidelberg: Winter.
- Belica, Cyril, Marc Kupietz, Harald Lungen & Rainer Perkuhn. 2012. *DeReWo – Korpusbasierte Grund-/Wortformenlisten: derewo-v-ww-bll-320000g-2012-12-31-1.0*. <http://www1.ids-mannheim.de/kl/projekte/methoden/derewo.html> (abgerufen am 13.11.2019).
- Belica, Cyril, Marc Kupietz, Harald Lungen & Rainer Perkuhn. 2013. *Benutzerdokumentation (Technical Report IDS-KL-2013-02) zum Produkt Korpusbasierte Wortgrundformenliste DEREWO v-ww-bll-320000g-2012-12-31-1.0*. <http://www1.ids-mannheim.de/fileadmin/kl/derewo/derewo-v-ww-bll-320000g-2012-12-31-1.0.zip> (abgerufen am 28.7.2020).
- Bildhauer, Felix, Eric Fuß & Franziska Münzberg. 2019. Starke und schwache Adjektivflexion in neuem korpuslinguistischen Licht. In Ludwig M. Eichinger & Albrecht Plewnia (Hgg.), *Neues vom heutigen Deutsch. Empirisch – methodisch – theoretisch*. (Jahrbuch 2018 des Instituts für Deutsche Sprache), 293–312. Berlin: De Gruyter.
- Dudengrammatik 2016 = Wöllstein, Angelika & Dudenredaktion (Hgg.). 2016. *Dudenband 4 – Die Grammatik*, 9. Aufl. Berlin: Dudenverlag.
- Eichinger, Ludwig M. 1987. Die Adjektive auf -isch und die Serialisierungsregeln in deutschen Nominalgruppen. In Herbert E. Brekle, Brigitte Asbach-Schnitker & Johannes Roggenhofer (Hgg.), *Neuere Forschungen zur Wortbildung und Historiographie der Linguistik: Festgabe für Herbert E. Brekle zum 50. Geburtstag* (Tübinger Beiträge zur Linguistik Bd. 284), 155–176. Tübingen: Narr.
- Eichinger, Ludwig M. 1991. Ganz natürlich – aber im Rahmen bleiben. Zur Reihenfolge gestufter Adjektivattribute. *Deutsche Sprache* (4), 312–329.
- Eisenberg, Peter. 2013. *Grundriss der deutschen Grammatik. Bd. 2: Der Satz*, 4. Aufl. Stuttgart & Weimar: Metzler.
- Gallmann, Peter. 2016. *Die m-Verschwörung: Zum Abbau der starken Deklinationsendung -em*. http://www.personal.uni-jena.de/~x1gape/Wort/Wort_NP_m_Schwaeche.pdf (abgerufen am 10.12.2019).

- Grammatikfragen. 2018. https://grammatikfragen.de/showthread.php?1505-quot-m-quot-oder-quot-n-quot-in-der-endung&prefixid=adjektiv_flexion (abgerufen am 7.7.2020).
- Hennig, Mathilde. 2015. Explizite und elliptische Junktion in der Attribution: Eine Bestandsaufnahme. In: Mathilde Hennig & Robert Niemann (Hgg.), *Junktion in der Attribution. Ein Komplexitätsphänomen aus grammatischer, psycholinguistischer und praxistheoretischer Perspektive*. Berlin: De Gruyter.
- Heuer, Walter, Max Flückiger & Peter Gallmann. 2017. *Richtiges Deutsch: Vollständige Grammatik und Rechtschreiblehre*, 32. Aufl. Zürich: Verlag Neue Zürcher Zeitung.
- Heynatz, Johann F. 1770. *Deutsche Sprachlehre zum Gebrauch der Schulen*. Berlin: Mylius. <http://mdz-nbn-resolving.de/urn:nbn:de:bvb:12-bsb10583655-1> (abgerufen am 7.7.2020).
- Heynatz, Johann F. 1803/2006. *Deutsche Sprachlehre zum Gebrauch der Schulen*. Faksimilenachdruck der 5. Aufl. (Documenta Orthographica 7). Hildesheim: Olms.
- Jaeger, Florian T. 2008. Categorical Data analysis: Away From ANOVAs (Transformation or not) and towards Logit Mixed Models. *Journal of Memory and Language* 59(4). 434–446.
- Kubczak, Jacqueline. 2015. Mit echtem bayerischem Senf oder mit echtem bayerischen Senf? – Flexion bei artikellosen Folgen von Adjektiven (aus „Grammatik in Fragen und Antworten“). *Sprachreport* 31(3). 22–27. Vgl. <https://grammis.ids-mannheim.de/fragen/44> (DOI: 10.14618/grammatikfragen, abgerufen am 28.7.2020)
- Lotzow, Stephanie. 2015. Bedingungen der (Nicht-)Realisierung von Präpositionen in verknüpften Präpositionalattributen. In Mathilde Hennig & Robert Niemann (Hgg.), *Junktion in der Attribution. Ein Komplexitätsphänomen aus grammatischer, psycholinguistischer und praxistheoretischer Perspektive*. 85–122. Berlin: De Gruyter.
- Meisner, Philipp. 2017. *Zur Deklination attributiver Adjektive im Gegenwartsdeutschen*. Universität Gießen: Bachelorarbeit.
- Moulin, Claudine. 2000. Varianz innerhalb der Nominalgruppenflexion: Ausnahmen zur sogenannten Parallelflexion der Adjektive im Neuhochdeutschen. *Germanistische Mitteilungen* 52. 73–97.
- Münzberg, Franziska, Stefan Falke, Sandra Hansen-Morath, & Ulrich Herrmann Waßner. 2018. *Datenbank attributive Adjektive*. <https://grammis.ids-mannheim.de/attributiveadjektivedb> (DOI: 10.14618/attributiveAdjektiveDB, abgerufen am 10.6.2020).
- Niebuhr, Astrid. 2017. *Mit einem stetigem Anstieg – Variation in der Adjektivflexion* (Vortrag auf der DGFS-Jahrestagung 2017, 9.3.2017). Saarbrücken.

- Nübling, Damaris. 2011. Unter großem persönlichem oder persönlichen Einsatz? Der sprachliche Zweifelsfall adjektivischer Parallel- vs. Wechselflexion als Beispiel für aktuellen grammatischen Wandel. In Klaus-Michael Köpcke & Arne Ziegler (Hgg.), *Grammatik – Lehren, Lernen, Verstehen: Zugänge zur Grammatik des Gegenwartsdeutschen* (Reihe germanistische Linguistik 293), 175–195. Berlin: De Gruyter.
- Peter, Klaus. 2013. Steuerungsfaktoren für Parallel- und Wechselflexion in Adjektivreihungen. *Jahrbuch für Germanistische Sprachgeschichte* 4(1). 186–204.
- Rat für deutsche Rechtschreibung. 2018. *Deutsche Rechtschreibung: Regeln und Wörterverzeichnis*. http://www.rechtschreibrat.com/DOX/rfdr_Regeln_2016_redigiert_2018.pdf (abgerufen am 9.12.2019).
- Rechtschreibduden 1991 = Dudenredaktion (Hg.). 1991. *Dudenband 1 – Die deutsche Rechtschreibung*, 20. Aufl. Mannheim, Leipzig, Wien & Zürich: Dudenverlag.
- Sahel, Said. 2005. Die Variation der Adjektivflexion nach Pronominaladjektiven und einige ihrer Determinanten: Eine empirische Untersuchung. *Deutsche Sprache* 33(4). 355–381.
- Sahel, Said. 2011. Monoflexion als Erklärung für Variation in der Nominalphrasenflexion des Deutschen. In Marek Konopka, Jacqueline Kubczak, Christian Mair, František Štícha & Ulrich H. Waßner (Hgg.), *Grammatik und Korpora 2009: Dritte internationale Konferenz Mannheim 22.–24.9.2009* (Korpuslinguistik und interdisziplinäre Perspektiven auf Sprache 1), 485–494. Tübingen: Narr.
- Schäfer, Roland. 2017. *Parallel- und Wechselflexion in obliquen Adjektivreihen – Eine empirische Vervollständigung* (31.5.2017). Vortrag am IDS Mannheim, Abstract.
- Trost, Igor. 2006. *Das deutsche Adjektiv: Untersuchungen zur Semantik, Komparation, Wortbildung und Syntax* (Beiträge zur germanistischen Sprachwissenschaft 19). Hamburg: Buske.
- Wiese, Bernd. 2009. Variation in der Flexionsmorphologie: Starke und schwache Adjektivflexion nach Pronominaladjektiven. In Marek Konopka & Bruno Strecker (Hgg.), *Deutsche Grammatik – Regeln, Normen, Sprachgebrauch* (Jahrbuch Institut für Deutsche Sprache 2008), 166–194. Berlin: De Gruyter.
- Wulff, Stefanie. 2003. A multifactorial corpus analysis of adjective order in English. *International Journal of Corpus Linguistics* (8:2). 245–282.
- Zeileis, Achim, David Meyer & Kurt Hornik. 2007. Residual-based Shadings for Visualizing (Conditional) Independence. *Journal of Computational and Graphical Statistics*, 16(3). 507–525.
- Zweifelsfäleduden = Hennig, Mathilde (Hg.). 2016. *Dudenband 9 – Das Wörterbuch der sprachlichen Zweifelsfälle*, 8. Aufl. Berlin: Dudenverlag.

Franziska Münzberg  & Felix Bildhauer 

Reihenfolge attributiver Adjektive

Abstract The present chapter investigates the relative order of attributive adjectives in German. Based on corpus data, our results corroborate previous findings that semantics is the most important factor in accounting for adjective order. Going beyond previous studies, we also consider coordinated structures (such as *mit* [[**großem, verwildertem**] **Garten**] ‘with (a) large, overgrown garden’), where both adjectives are of equal rank. While adjective order in embedded structures (*mit* [**schwierigem** [**familiärem Hintergrund**]]) ‘with (a) difficult domestic background’) can be predicted rather accurately on semantic grounds, we show that predictions can also be made for coordinated structures, albeit with lower accuracy. Using regression analysis, we examine how semantic factors interact with a number of other explanatory variables.

Keywords word order, noun phrase, adjective, attribute, attributive participle, logistic regression, random effects model

Open Data Der Datensatz `ATTRIBUTIVE_ADJEKTIVE_1` ist öffentlich zugänglich und durchsuchbar in der [Datenbank attributive Adjektive](#) (DOI: 10.14618/attributiveAdjektiveDB).

1 Fragestellung

Für die Reihenfolge attributiver Adjektive links vom Kopfnomen sind vor allem semantische bzw. kognitive Kriterien und damit die folgende Grundreihenfolge beschrieben worden:

- Artikelklassifikatoren – Qualificativa – Nominalklassifikatoren (vgl. Eichinger 1991: 327)
- bzw. referenziell – qualitativ – klassifikatorisch (vgl. Schlotthauer 2017: 1646)

Nominalgruppen, in denen die Adjektivreihenfolge dieser Grundreihenfolge entspricht, wären etwa:

- (1) a. der **besagte** (referenziell) **gerechte** (qualitativ) **finanzielle** (klassifikatorisch) Ausgleich
 b. **andere** (referenziell) **beeindruckende** (qualitativ) **landschaftliche** (klassifikatorisch) Schönheiten

Unter den Adjektiven mit qualitativer Lesart stehen die wertenden Adjektive vor den übrigen (vgl. Adam & Schecker 2011: 164).

Aus formaler Perspektive werden die Möglichkeit der prädikativen Verwendung, die Komparierbarkeit und die Wortbildung der Adjektive als Kriterien für ihre Abfolge angeführt (vgl. Eisenberg 2013: 393–394): Prädikativ Verwendbares, Komparierbares und Deverbales steht mittig (und hat in der Regel eine qualitative Lesart). Substantivderivate haben meist eine klassifikatorische Lesart und stehen entsprechend rechts, also näher am Kopfnomen.

- (2) **andere** (nicht¹ prädikativ verwendbar, nicht komparierbar) **beeindruckende** (prädikativ verwendbar, komparierbar, deverbale) **landschaftliche** (nur eingeschränkt² prädikativ verwendbar, nicht komparierbar, desubstantivisch) Schönheiten

In einer Synthese von solchen semantischen und morphosyntaktischen Eigenschaften attributiver Adjektive gibt die Dudengrammatik (2016: 346) ähnlich wie Engel (vgl. 1988: 634–635) eine feiner gegliederte Abfolge an. Zu der nur leicht modifizierten Form, in der wir diese Eigenschaften als „11-Klassen-System“ in Abschnitt 3.1 übernommen haben, vgl. Münzberg & Hansen in dieser Ausgabe.

Diese gut gesicherten Erkenntnisse werden hier um Spezifikationen ergänzt, die sich nicht aus den beschriebenen semantischen und syntaktischen Eigenschaften ableiten lassen. So soll etwa überprüft werden, ob das Behaghel'sche „Gesetz der wachsenden Glieder“ (Behaghel 1909: 139; vgl. auch Behaghel 1932: 6 = § 1416) für aufeinanderfolgende attributive Adjektive gilt, ob also kürzere

1 Soweit man davon ausgeht, dass *anders* ein eigenes Lexem ist.

2 Anstelle einer auffälligen Formulierung mit prädikativem Gebrauch wie *Die Schönheiten sind landschaftlich* könnte man ausweichen auf Alternativen mit attributivem Gebrauch wie *Die Schönheiten sind landschaftlicher Natur / von landschaftlicher Art* o. Ä. Den prädikativen Gebrauch solcher Adjektive kategorisch auszuschließen, wäre falsch, vgl. fürs Englische Demonte (2019: 406–407). Das liegt nicht zuletzt an der Polysemie von Adjektiven. Das Adjektiv *landschaftlich* kann zum Beispiel auch ‚regional, areal, diatopisch markiert‘ bedeuten und ist dann uneingeschränkt prädikativ verwendbar: *Das Substantiv Sonnabend ist mittlerweile eindeutig landschaftlich*. Man könnte allerdings argumentieren, dass ursprünglich klassifikatorische Adjektive bei prädikativem Gebrauch ihre semantische Klasse wechseln und zu qualitativen Adjektiven werden.

Adjektive vor längeren stehen und ob durch ein Adverb modifizierte Adjektive nach nicht modifizierten Adjektiven stehen. Was die Einteilung in semantische Klassen angeht, wird zusätzlich zum 11-Klassen-System ein Faktor mit nur zwei Ausprägungen („2-Klassen-System“) etabliert: Die beiden Ausprägungen sind „intersektiv“ und „nicht intersektiv“ (s. Abschnitt 3.2). Untersucht wird auch, ob Partizipien im vorliegenden Datensatz eher vor Nichtpartizipien stehen (vgl. Sommerfeldt 1971: 15). Darüber hinaus soll die Wahrscheinlichkeit dafür, dass ein Adjektiv attributiv zu einem gegebenen Substantiv auftritt, einbezogen werden (= P(Adj|N), im Weiteren „substantivspezifische Adjektivfrequenz“; vgl. fürs Englische Wulff (2003: 267), dort als „noun-specific frequency“). Die Erwartung ist, dass ein Adjektiv dann eher nahe beim Substantiv steht, wenn es auch sonst häufig attributiv zu diesem Substantiv steht. Auch Effekte der Frequenzklassen der beiden Adjektive sollen geschätzt werden: Das linke Adjektiv könnte determiniererähnlicher und damit tendenziell häufiger sein als das rechte (*andere landschaftliche Schönheiten*).

Darüber hinaus gibt es noch mehr mögliche Einflussgrößen, die vor allem in Bezug auf das Englische untersucht worden sind. Viele sind nur schwer objektivierbar. Einen Überblick gibt Dyer (2017: 5–67).

2 Korpus, Suchanfrage und Methoden

Ausgewertet wurde der Datensatz `ATTRIBUTIVE_ADJEKTIVE_1` aus der Untersuchung der Adjektivflexion im Dat. Sg. Mask./Neutr. (vgl. Münzberg & Hansen in dieser Ausgabe). Er wurde aus einem Ausschnitt von 122 Millionen Tokens aus dem Untersuchungskorpus erzeugt und enthält nach dem Aussortieren von Fehlbelegen 1.598 Belegsätze zur Suchanfrage

Präposition	∅	erstes Adjektiv mit Endung <i>-em</i>	∅	∅	zweites Adjektiv mit Endung <i>-em</i> oder <i>-en</i>	Nomen
	Adverb		,	Adverb		

Wortgruppen, die diesem Muster entsprechen würden, sind etwa:

- (3) a. mit langem schwarzem/schwarzen Haar
- b. mit langem, schwarzem/schwarzen Haar
- c. mit ziemlich langem schwarzem/schwarzen Haar
- d. mit ziemlich langem, beinah schwarzem/schwarzen Haar

Die Besonderheit der Daten ist also, dass keine der einschlägigen Nominalphrasen einen Determinierer enthält. Das kann für die Interpretation der Ergebnisse eine Rolle spielen, vgl. Abschnitt 3.1 sowie die Einleitung zu dieser Ausgabe.

Die Beschränkung auf Nominalphrasen mit exakt zwei attributiven Adjektiven erklärt sich dadurch, dass schon zwei attributive Adjektive in unserem hauptsächlich Preetexte enthaltenden Untersuchungskorpus außergewöhnlich sind, mehr als zwei aber äußerst selten. Die Auswertung eines anderen Datensatzes aus einem Achtel des gesamten Untersuchungskorpus illustriert das: Von 2.388 Nominalphrasen, deren Kopf ein Appellativ ist (kein Pronomen, kein Eigennamen) und die eine der Funktionen Subjekt, Prädikativ, Akkusativobjekt, Dativobjekt haben, enthalten 26,00 % ein attributives Adjektiv, 1,59 % zwei attributive Adjektive und nur 0,08 % drei attributive Adjektive.

Bei der nachträglichen Belegannotation wurden die folgenden Einflussgrößen für jedes der beiden Adjektive berücksichtigt (in Abschnitt 3.1–3.7 werden sie näher erläutert):

- semantische Klasse (11-Klassen-System)
- semantische Klasse (2-Klassen-System)
- Partizip JA/NEIN
- Silbenzahl (inklusive der Silbe des Flexionssuffixes)
- Erweiterung durch ein Adverb
- Frequenz
- substantivspezifische Adjektivfrequenz

Zusätzlich wurde die Möglichkeit bewertet, die beiden Adjektive zu tauschen, ohne dass sich die Semantik der Nominalgruppe wesentlich ändert oder die Reihenfolge sehr ungewöhnlich wirkt.³ Wenn dieser Tausch „gelingt“, wurde der Wert 1 vergeben, sonst der Wert 0. Starke/schwache Flexion wurde bei diesem Tauschtest ignoriert: Die beiden Adjektive in der Präpositionalgruppe *in edlem zeitgemäßen Ambiente* sind also als tauschbar markiert worden. Das Ergebnis des Tauschs wäre dann nicht *in zeitgemäßen edlem Ambiente*, sondern *in zeitgemäßem edlem Ambiente* oder *in zeitgemäßem edlen Ambiente*. Diese Annotation sollte nicht als Einflussgröße in statistische Modelle übernommen werden. Vielmehr dient sie bei der Untersuchung der nicht semantischen Einflussgrößen der Differenzierung von kategorischen Kontexten und Wahlkontexten.

Auf der Grundlage dieser Daten haben wir ein logistisches Regressionsmodell spezifiziert. Dieses modelliert die Wahrscheinlichkeit dafür, dass ein bestimmtes der beiden Adjektive (der „Kandidat“) die rechte, dem Kopfnomen nähere Position einnimmt, in Abhängigkeit von einigen der oben benannten Einflussgrößen. Für diese Einflussgrößen werden dabei Koeffizienten geschätzt, welche die Größe des jeweiligen Effekts darstellen. Um die Daten mit einer logistischen Regression

3 Die Annotation wurde von zwei Personen vorgenommen. Das Interrater-Agreement war „moderate“ im Sinne von Landis & Koch (1977); Cohens Kappa = 0,5. Das Untersuchungsziel war den annotierenden Hilfskräften des IDS nicht bekannt.

zu analysieren, wurde in zufällig ausgewählten 799 Belegen (entspricht 50 % der Fälle) das zweite Adjektiv als „Kandidat“ ausgezeichnet. Die abhängige Variable („steht rechts“) ist in diesen Fällen „1“. In der anderen Hälfte der Fälle wurde dagegen das linke Adjektiv als „Kandidat“ ausgezeichnet, die abhängige Variable („steht rechts“) ist in diesen Fällen „0“. Das jeweils andere Adjektiv wird im Folgenden als „Konkurrent“ bezeichnet.

In einem gemischten logistischen Regressionsmodell haben wir folgende Faktoren berücksichtigt:

- semantische Klasse (2-Klassen-System) Kandidat: INTERSEKTIV, NICHT-INTERSEKTIV
- semantische Klasse (2-Klassen-System) Konkurrent: INTERSEKTIV, NICHT-INTERSEKTIV
- Kandidat ist Partizip: JA, NEIN
- Konkurrent ist Partizip: JA, NEIN
- Differenz Silbenzahl, dichotomisiert als POSITIV (Kandidat hat mehr Silben als Konkurrent), NICHTPOSITIV
- Differenz substantivspezifische Adjektivfrequenz (SAF), dichotomisiert als POSITIV (SAF des Kandidaten ist größer als die des Konkurrenten), NICHT-POSITIV
- Differenz Frequenzklasse, dichotomisiert als POSITIV (Kandidat hat höhere Frequenzklasse als Konkurrent), NICHTPOSITIV
- Kandidat ist durch ein Adverb erweitert: JA, NEIN
- Konkurrent ist durch ein Adverb erweitert: JA, NEIN

Außerdem wurde das Lemma des Kandidatenadjektivs als Gruppierungsfaktor verwendet.

Die Abfolge semantischer Klassen nach dem 11-Klassen-System wurde im Regressionsmodell nicht berücksichtigt, da durch die zahlreichen Ausprägungen dieser Variablen z. T. Gruppen mit sehr wenigen Fällen entstehen (und Koeffizientenschätzungen entsprechend ungenau werden). Stattdessen bietet Abschnitt 3.1 einen deskriptiven Überblick über den Zusammenhang zwischen semantischer Klasse und Reihenfolge. Die elfte Klasse kam nicht vor; sie war vorgesehen für den Fall, dass beim Part-of-Speech-Tagging Determinierer als attributive Adjektive ausgezeichnet würden. Bei der Auszählung haben wir z. B. festgestellt, dass im vorliegenden Datensatz 33-mal ein relationales Adjektiv mit Bezug auf die zeitliche Situierung (*derzeitig*, *früher*) vor einem Zugehörigkeitsadjektiv (*juristisch*, *politisch*) steht und die umgekehrte Reihenfolge nicht vorkommt. Genauso wurden auch alle anderen vorkommenden Kombinationen ausgezählt. Theoretisch hätte man für 10 semantische Klassen auch 10 verschiedene „Distanzklassen“ erhalten können, je nachdem, wie weit entfernt Adjektive einer bestimmten semantischen Klasse von ihrem Kopfnomen stehen. Dazu ist der Datensatz aber

zu klein – Adjektive, die eine materielle Beschaffenheit bezeichnen (*steinern, wollen*), kommen beispielsweise nur 5-mal vor, und für einige Kombinationen gibt es auch zu wenige Belege. Eine Beschränkung auf nur 6 verschiedene Distanzklassen erschien von daher sinnvoll.

3 Ergebnisse der Korpusuntersuchung

Im Folgenden werden die Wirkungen der einzelnen Einflussgrößen beschrieben. Dem Überblick über die in der Dudengrammatik (2016: 346) angesetzten semantischen Klassen und unsere 6 Distanzklassen folgen ab Abschnitt 3.2 die Ergebnisse der Regressionsanalyse. Abb. 1 in Abschnitt 4 fasst dann die wichtigsten Punkte zusammen.

3.1 Semantische Klasse der Adjektive (11-Klassen-System)

Da in der Suchanfrage nur artikellose Nominalphrasen berücksichtigt sind, nehmen Ordinalzahladjektive im vorliegenden Datensatz einen anderen Platz ein als gewöhnlich (weiter rechts als sonst erwartbar), wie z. B. in (4).

- (4) [...] nach verlorenem ersten Satz steigerte er sich jedoch und gewann verdient (St. Galler Tagblatt, 22.05.2000, o. S.)

Fast alle Belege zu Ordinalzahladjektiven an zweiter Stelle entstammen solchen Sportnachrichten. Diesen 72 Belegen steht ein einziger Beleg mit einem Ordinalzahladjektiv an erster Stelle gegenüber.

Suchen mit dem Beispiel *erst-* in einem Achtel des 7-Mrd.-Untersuchungskorpus ergaben die Zahlen in Tab. 1.

Tabelle 1: Artikellose Nominalphrasen mit Ordinalzahladjektiv *erst-* im Untersuchungskorpus.

Suchanfrage			Ergebnisse pro Mio. Tokens
Präposition	<i>erstem</i>	attributives Adjektiv	0,03
Präposition	attributives Adjektiv	<i>erstem</i>	0,04
Präposition	attributives Adjektiv	<i>ersten</i>	0,55

Das heißt, dass das Ordinalzahladjektiv *erst-* in artikellosen Nominalphrasen in nur 5 % der Fälle an erster Stelle vor einem anderen attributiven Adjektiv steht

($0,03 / (0,03 + 0,04 + 0,55) = 0,05$). Dagegen tritt in Nominalphrasen mit definitem oder indefinitem Artikel das Ordinalzahladjektiv in 73 % der Fälle vor das andere Adjektiv, wie Tab. 2 zeigt.

Tabelle 2: Nominalphrasen mit Artikel und Ordinalzahladjektiv *erst-* im Untersuchungskorpus.

Suchanfrage				Ergebnisse pro Mio. Tokens
Präposition	Artikel	<i>ersten</i>	attributives Adjektiv	8,62
Präposition	Artikel	attributives Adjektiv	<i>ersten</i>	3,22

Die Variation bei der Stellung von Ordinalzahladjektiven ist semantisch nicht frei: *Ein erster verlorener Satz* kann der zweite Satz nach einem gewonnenen ersten Satz sein, *ein verlorener erster Satz* ist immer der erste Satz. Offensichtlich ist in Nominalphrasen mit definitem oder indefinitem Artikel die zweite Referenzsituation häufiger (73 %) und in artikellosen Nominalphrasen die erste (95 %).

Ansonsten bestätigen unsere Distanzklassen die in der Dudengrammatik (2016: 346) beschriebene Reihenfolge. Von links nach rechts, d. h. nach abnehmender Entfernung vom Kopfnomen geordnet:

- Distanzklasse 6: Adjektive, die Indefinita oder Demonstrativa nahekommen: *besagt, weiter, sogenannt**: etwa im Beispiel *weiteres* (Distanzklasse 6) *wochenlanges* (Distanzklasse 5) *Ringens*
- Distanzklasse 5: relationale Adjektive mit Bezug auf die zeitliche Situierung: *früher, jahrelang, anschließend: das frühere* (Distanzklasse 5) *obere* (Distanzklasse 4) *Stockwerk*
- Distanzklasse 4: relationale Adjektive mit Bezug auf die räumliche Lage: *recht, inner, ober, dortig: mit dortigem* (Distanzklasse 4) *festem* (Distanzklasse 3) *Wohnsitz*
- Distanzklasse 3: sonstige qualifizierende Adjektive (im vorliegenden Datensatz ohne eindeutigen Unterschied zwischen im Kontext tatsächlich qualifizierenden Adjektiven und solchen, die als relationale Adjektive verwendet werden): *gemütlich, gut, angemessen, lang: mit langem* (Distanzklasse 3) *schwarzem* (Distanzklasse 2) *Haar*
- Distanzklasse 2: Ordinalzahl- und Farbadjektive (wobei diese Stellung der Ordinalzahladjektive so nah am Nomen im vorliegenden Datensatz eine Beson-

4 Der Auslassungsbindestrich zeigt an, dass die Form nicht oder nur eingeschränkt prädikativ verwendbar ist, sondern üblicherweise dekliniert und attributiv vorkommt. Einige Adjektive wie etwa *inner-* haben überhaupt keine frei vorkommende Grundform.

- derheit ist, s. o.): *erst, schwarz, bunt, dunkelblau: mit bunten* (Distanzklasse 2)
tibetischen (Distanzklasse 1) *Fähnchen*
- Distanzklasse 1: Zugehörigkeitsadjektive (einschließlich solcher, die die geografische Herkunft bezeichnen): *demokratisch, sozial, finanziell, tibetisch, sächsisch*

3.2 Semantische Klasse der Adjektive (2-Klassen-System)

Die hier angesetzte Einteilung in intersektive und nicht intersektive Adjektive orientiert sich an GDS (1997: 2000–2004), Russell (2012: 333) und Romero (2008). Der Grundgedanke ist, dass manche Adjektive schnittmengenbildend sind, manche nicht. So sind etwa „mehrstimmige Lieder“ als Schnittmenge aus mehrstimmigen Musikstücken („Musikstück“ als Genus proximum zu „Lied“) und Liedern beschreibbar. Das liegt daran, dass die Bedeutung von *mehrstimmig* unabhängig von der Bedeutung von *Lied* interpretiert werden kann: Das Adjektiv *mehrstimmig* ist intersektiv. Ist aber von *langen Liedern* die Rede, so funktioniert der Schnittmengentest nicht: Ein langes Lied ist nicht per se ein langes Musikstück, denn lange Lieder können kürzer sein als kurze Sinfonien. Um das Adjektiv *lang* richtig zu interpretieren, muss man die Bedeutung des Ausdrucks kennen, den es modifiziert, hier *Lied*. Damit ist *lang* wie viele im Kontext steigerbare Adjektive (*schwierig, hart, fein, breit, schwer*) nicht intersektiv.

Auch bei den privativen Adjektiven (die hier ebenfalls unter die nicht intersektiven Adjektive gerechnet werden) wie etwa *scheinbar* kann keine Schnittmenge gebildet werden: Ein „scheinbares Lied“ ist eben kein Lied und kann darum auch nicht in einer Schnittmenge aus scheinbaren Musikstücken und Liedern enthalten sein.

Unter „intersektiv“ fassen wir im Folgenden auch die klassifizierenden Adjektive, die im Sinne von Zifonun (vgl. 2017b: 25) teilmengenbildend (subsektiv) sind, wie etwa *familiär, körperlich, politisch* oder auch *privat*; vgl. *carnivorous* als Beispiel für Intersektivität bei Romero (2008: 16). Ein „familiärer Hintergrund“ ist in diesem Sinne die Schnittmenge aus familiären Angelegenheiten und möglichen Hintergründen, „körperliches Unwohlsein“ ist die Schnittmenge aus körperlichen Empfindungen/Zuständen und Unwohlsein, „gedrechseltes Holz“ ist die Schnittmenge aus gedrechselten Werkstücken und Holz.

Ein weiterer Test, mit dem sich die beiden Gruppen von Adjektiven abgrenzen lassen, ist: *Dieses Stück ist lang für ein Lied, Dieses Holz ist hart für Holz* (*lang, hart* sind nicht intersektiv), aber **Dieses Stück ist mehrstimmig für ein Lied, *Dieses Holz ist gedrechselt für Holz* (*mehrstimmig, gedrechselt* sind intersektiv; vgl. Demonte 2019: 389).

Aus der Beschreibung der Tests, die bei der manuellen Belegannotation durchgeführt wurden, wird ersichtlich: Wie viele semantische Beurteilungen hat auch diese eine starke subjektive Komponente. Vor allem muss zur Entscheidung in der Regel der Kontext herangezogen werden. Mit *augenblicklich wache Babys* kann die Schnittmenge aus Menschen, die augenblicklich nicht schlafen, und Babys gemeint sein; das Adjektiv *wach* unterscheidet sich in dieser Lesart nicht von *wach* in *augenblicklich wache Greise* und erscheint darum als intersektives Adjektiv. Aber in *waches Gesicht* bezeichnet das Adjektiv *wach* eine Eigenschaft, die in höherem oder geringerem Grade einem Gesicht zugesprochen werden kann und die nicht einfach das Gegenteil von *schlafend* ist. Ohne das Substantiv *Gesicht* kann nicht verstanden werden, welche Eigenschaft gemeint ist. Hier ist *wach* eher nicht intersektiv.

Nicht intersektive Adjektive stehen im vorliegenden Datensatz⁵ eher vor intersektiven (79% der Belege, in denen die Adjektive unterschiedlichen semantischen Klassen angehören; Tab. 3 in Abschnitt 3.3):

- (5) a. mit schwierigem familiärem Hintergrund (Süddeutsche Zeitung, 13.11.2004, S. V1/21)
- b. aus hartem, gedrechseltem Holz (Hannoversche Allgemeine, 1.3.2008, S. 5)
- c. mit geringfügigem körperlichen Unwohlsein (St. Galler Tagblatt, 20.5.2000, o. S.)

Unter diese Regel fällt wohl auch die Tendenz der Farbadjektive, rechts von den übrigen qualitativen Adjektiven zu stehen (eine eingehende Diskussion, ob Farbadjektive tatsächlich intersektiv sind, findet sich in GDS (vgl. 1997: 2003–2004):

5 Dyer (2017: 24) behandelt diese Regularität sogar als eine mögliche Universalie. Er spricht ihr Erklärungskraft ab, räumt aber ein, dass ihre Einbeziehung seine eigenen Entropieberechnungen eventuell hätte exakter machen können (Dyer 2017: 150). Für das Englische postuliert Truswell (2009: 527): „Subjective adjectives precede interjective adjectives“ und „Interjective adjectives do not precede subjective adjectives“. Seine Ergebnisse basieren auf einer Google-Recherche. Gezählt wurden anders als in der vorliegenden Studie nur Belege, die als akzeptabel bewertet wurden, und die Informationsstruktur wurde kontrolliert (Truswell 2009: 526). Diesen Ergebnissen widersprechen Trotzke & Wittenberg (2019: 278–279): In ihrer Studie, die auf Google Ngram (also auf dem Datenbestand von Google Books) basiert, gibt es mehr Belege für die Reihenfolge intersektiv – subsektiv als für die Reihenfolge subsektiv – intersektiv. Woher diese sehr unterschiedlichen Ergebnisse kommen, diskutieren sie nicht. Möglicherweise ordnen sie andere Adjektive als intersektiv ein.

- (6) a. mit feinem weißen Sand (Rhein-Zeitung, 28.5.1998, o. S.)
 b. mit breitem grünen Streifen (Potsdamer Neuste Nachrichten, 10.4.2012, o. S.)
 c. aus schwerem, weinrotem Damast (Rhein-Zeitung, 23.1.2008, o. S.)

Aber auch die umgekehrte Reihenfolge intersektiv – nicht intersektiv kommt vor, ohne auffällig zu wirken:

- (7) mit gesenktem, wachem Gesicht (Lenz: Die Auflehnung, 1999 [1994], S. 343)

3.3 Interaktion semantische Klasse – Komma JA/NEIN

Allgemein kann man davon ausgehen, dass die Wortstellung freier ist, wenn beide Adjektive gleichermaßen das Nomen modifizieren (vgl. fürs Englische Davidse & Breban 2019: 347–348). Ein Anzeichen für eine solche Koordination von Adjektiven (eine gereimte Struktur) ist das Komma zwischen beiden Adjektiven (vgl. Münzberg & Hansen in dieser Ausgabe): *mit gesenktem, wachem Gesicht*. Demgegenüber ist bei der gestuften Struktur (das erste Adjektiv modifiziert die Gesamtheit aus zweitem Adjektiv und Nomen) die Reihenfolge der Adjektive semantisch festgelegt. Daher beschränken Eichinger (1991) sowie Kotowski & Härtl (2019) ihre Untersuchungen ausdrücklich auf gestufte Strukturen bzw. auf Strukturen ohne Komma.

Tabelle 3: Adjektivreihenfolge nach Semantik der Adjektive in Kontexten mit und ohne Komma.

Abfolge	Komma		
	nein	ja	Σ
nicht intersektiv < intersektiv (Normalabfolge)	711	93	804
intersektiv < nicht intersektiv	144	73	217
Σ	855	166	1.021

Bei der statistischen Modellierung wurde die Beziehung zwischen den Variablen „Komma“ und „Semantik“ (2-Klassen-System) als Interaktion berücksichtigt. Das Ergebnis fällt deutlich aus: Während in Fällen ohne Komma ein intersektives Adjektiv eine sehr starke Tendenz hat, nach einem nicht intersektiven Adjektiv aufzutreten, ist diese Tendenz für Fälle mit Komma deutlich weniger ausgeprägt (vgl. den geschätzten Effekt für die entsprechende Interaktion im Anhang, Tab. 5).

Ein Beispiel für die Normalabfolge nicht intersektiv – intersektiv in einer Nominalgruppe mit Komma zwischen den beiden Adjektiven ist (8).

- (8) mit dichtem, wollenem Unterfell (Rhein-Zeitung, 9.1.2006, o. S.)

Die seltenere Reihenfolge intersektiv – nicht intersektiv ist in Strukturen mit Komma (Beispiel 9) wahrscheinlicher als in Strukturen ohne Komma.

- (9) inmitten lautlosem, dichtem Schneefall (Grass: Die Blechtrommel, 1964 [1962], S. 103)

3.4 Partizip

Der Beleg (7) *mit gesenktem, wachem Gesicht* ist ein Beispiel dafür, dass Partizipien eher vor Nichtpartizipien stehen (vgl. Sommerfeldt 1971: 15). Bei Partizipien I in Kombination mit Nichtpartizipien macht in unseren Daten die Reihenfolge Partizip – Nichtpartizip 77% der Fälle aus, bei Partizipien II 69% der Fälle. Ein Beispiel für die Reihenfolge Partizip I – Nichtpartizip ist (10).

- (10) mit anschließendem gemütlichen Beisammensein (Mannheimer Morgen, 17.8.2013, S. 16)

Der Grund für diesen Wortstellungseffekt könnte im semantischen Potenzial von Partizipien gesucht werden. Partizipien können assertorisch (einen Sachverhalt behauptend) modifizieren und möglicherweise deswegen links stehen (zur Definition von „assertorisch“ vgl. Murelli 2016: 1808; im Fokus stehen dort ausgebaute Partizipialphrasen mit Komplementen und ggf. Supplementen, die in ihrer Funktion Relativsätzen nahekommen, vgl. Murelli 2016: 1811–1812). Etwa der Ausdruck *das von keinem der Zuschauer erwartete Tor* kann nur auf ein Tor zutreffen, für das gilt: Der Sachverhalt „Das Tor wurde von keinem der Zuschauer erwartet“ wird behauptet (vgl. Zifonun 2017a: 1737). Aber auch „unerweiterte Partizipialattribute“ (Murelli 2016: 1837) können assertorisch sein (*der verschossene Elfmeter*).

Im vorliegenden Datensatz gibt es nur 20 Beispiele für adverbial erweiterte Partizipialattribute. Eines davon, das als assertorisch gelesen werden kann, findet sich in (11).

- (11) nach nun vorgenommenem erfolgreichem Eingriff (Vorarlberger Nachrichten, 18.1.1997, S. D12)

Partizipien können allerdings auch wie andere attributive Adjektive qualifizierend oder klassifikatorisch gebraucht werden. Die klassifikatorische Lesart ist im Deutschen seltener (vgl. Murelli 2017: 1811; ein Beispiel dort: *stehende Gewässer*). Nach Engel (1988: 635) sind unter den qualitativen Adjektiven „auch die meisten attributiv gebrauchten Partizipien“ einzuordnen. Tatsächlich wird auch im Großteil der hier untersuchten Belege das Partizip eher qualifizierend gebraucht. Vor allem als Adjektive lexikalisierte Partizipien (*ansprechend, begeistert*) können qualifizierend sein (12).

- (12) a. mit ansprechendem kulinarischen Angebot (Berliner Morgenpost, 19.6.1999, S. 13)
 b. mit begeistertem, minutenlangem Beifall (Mannheimer Morgen, 22.5.1996, o. S.)

Wenn nun lexikalisierte Partizipien (steigerbare, mit *un-* präfigierbare und/oder als Adjektiv in den Duden online aufgenommene Partizipien) oft qualifizierend sind, wäre vielleicht zu erwarten, dass lexikalisierte Partizipien tendenziell weiter links stehen als nicht lexikalisierte. Andererseits wiederum – und das könnte diese Erwartung neutralisieren – sind lexikalisierte Partizipien als Adjektive den „gewöhnlichen“ Adjektiven, also den Nichtpartizipien, ähnlicher als Partizipien, die noch nicht als Adjektive lexikalisiert sind. Unsere Ausgangshypothese lautete daher: Die nicht lexikalisierten Partizipien stehen am häufigsten links, gefolgt von den lexikalisierten Partizipien und schließlich den Nichtpartizipien.

Wie Tabelle 4 zeigt, trifft dies auf ca. 72% der 370 Belege zu, in denen die beiden Adjektive zu unterschiedlichen Kategorien (lexikalisiertes Partizip, nicht lexikalisiertes Partizip, Nichtpartizip) gehören.

Tabelle 4: Verteilung der Abfolgen von lexikalisierten Partizipien, nicht lexikalisierten Partizipien und Nichtpartizipien („<“ zu lesen als „geht voran“). Einteilung in erwartete und unerwartete Abfolgen gemäß der Hypothese „nicht lexikalisierte Partizipien stehen am häufigsten links, gefolgt von lexikalisierten Partizipien, gefolgt von Nichtpartizipien“.

	Abfolge	Belege
erwartet	lexikalisiertes Partizip < Nichtpartizip	161
	nicht lexikalisiertes Partizip < Nichtpartizip	102
	nicht lexikalisiertes Partizip < lexikalisiertes Partizip	2
unerwartet	Nichtpartizip < lexikalisiertes Partizip	69
	lexikalisiertes Partizip < nicht lexikalisiertes Partizip	3
	Nichtpartizip < nicht lexikalisiertes Partizip	33
Σ		370

Der Datensatz enthält nur 5 Belege für die Kombination eines lexikalisierten Partizips mit einem nicht lexikalisierten. Bei der Kombination eines Partizips mit einem Nichtpartizip unterscheiden sich lexikalisierte und nicht lexikalisierte Partizipien kaum, der Anteil an erwarteten Abfolgen liegt bei 70 % respektive ca. 76 %. Die Differenz ist bei einer Stichprobengröße von 365 zu klein, um von einem tatsächlichen Unterschied auszugehen.⁶ Aus diesem Grund verzichten wir im Folgenden auf die Unterscheidung zwischen lexikalisierten und nicht lexikalisierten Partizipien und berücksichtigen bei der statistischen Modellierung nur die Unterscheidung zwischen Partizipien einerseits und Nichtpartizipien andererseits.

Diejenigen Partizipien, die qualifizierend gebraucht werden, stehen in der Grundreihenfolge referenziell – qualitativ – klassifikatorisch links vor den (klassifizierenden) Zugehörigkeitsadjektiven. Auch könnte man annehmen, dass Assertorisches in der Regel allgemein weiter entfernt vom Kopfnomen steht als Klassifikatorisches und wohl auch weiter entfernt als Attribute mit qualitativer Lesart, vgl. (jetzt vom Kopfnomen aus nach **rechts**): *die Kita in der Mühlthalstraße* (klassifikatorisch), *die unser Sohn besucht* (assertorisch); *eine Bergkulisse von unglaublichem Reiz* (qualitativ), *die schon oft fotografiert worden ist* (assertorisch). Dass assertorisch modifizierende Partizipialattribute in gestuften Strukturen **links** vom Kopfnomen meist vor anderen Adjektivattributen stehen, stimmt mit dieser Annahme überein. Vielleicht wirkt schon die grundsätzliche Möglichkeit, Partizipien assertorisch zu verwenden, auf die Wortstellung ein.

Eine zusätzliche Stichprobe im ersten Achtel des Untersuchungskorpus ergibt für die Struktur „erweitertes Partizipialattribut – kein Komma und keine Konjunktion – attributives Adjektiv“ (Beispiel: *ein an der Kreuzung Waterloostraße / Wabestraße abgestellter roter Alfa Romeo*) ca. 27 Belege / 1 Mio. Tokens, für die Reihenfolge „attributives Adjektiv – kein Komma und keine Konjunktion – erweitertes Partizipialattribut“ nur rund 2 Belege / 1 Mio. Tokens. Für die Struktur „erweitertes Partizipialattribut – Komma oder Konjunktion – Adjektiv“ sind es auch nur ca. 2 Belege / 1 Mio. Tokens, während die Reihenfolge „Adjektiv – Komma oder Konjunktion – erweitertes Partizipialattribut“ (Beispiel: *ein langjähriger, für das rot-rot-grüne Projekt 2017 fest eingeplanter Koalitionspartner*) rund 7-mal / Mio. Tokens vorkommt. Nun nehmen wir an, dass Strukturen ohne das Oberflächenmerkmal Komma in der Regel gestuft sind und Strukturen mit

6 $\chi^2(1, 365) = 1,04, p = 0,31$. Bei gleichzeitiger Betrachtung mit anderen möglichen Einflussgrößen in einem logistischen Regressionsmodell wird ebenfalls kein Unterschied zwischen lexikalisierten und nicht lexikalisierten Partizipien deutlich. Dies ändert sich auch dann nicht, wenn potenziell korrelierende Variablen (Intersektivität und Silbenzahl) aus dem Modell entfernt werden.

Komma in der Regel gereiht (Münzberg & Hansen in dieser Ausgabe). Das würde bedeuten: In gestuften Strukturen steht das erweiterte Partizip eher vor einem anderen Adjektiv, in gereihten eher danach.

3.5 Silbenzahl

Adjektive mit mehr Silben⁷ stehen im vorliegenden Datensatz in der Regel nach Adjektiven mit weniger Silben (13). Es scheint also das „Gesetz der wachsenden Glieder“ (Behaghel 1909:139; vgl. auch Behaghel 1932: 6 = § 1416) zu wirken.

- (13) a. mit großem, verwildertem Garten (Frankfurter Rundschau, 21.8.1997, S. 12)
 b. mit wildem, rücksichtslosem Gerangel (Die Zeit [Online-Ausgabe], 22.2.2007, o. S.)

Hier sind jedoch kategorische Kontexte und Wahlkontexte zu unterscheiden. So kann das erste Adjektiv die Gesamtheit aus zweitem Adjektiv und Nomen modifizieren (gestufte Struktur, vgl. Münzberg & Hansen in dieser Ausgabe). Wenn die Schreibenden eine solche Struktur intendieren, spielt die Silbenzahl keine Rolle. Es handelt sich um einen kategorischen Kontext, der von der Semantik bzw. von der internen Struktur der Nominalphrase bestimmt ist. Daher stehen qualifizierende Adjektive auch dann eher vor Zugehörigkeitsadjektiven (s. Abschnitt 3.1), wenn sie länger sind als diese:

- (14) mit augenzwinkerndem britischen Humor (Die Presse, 12.4.1994, o. S.)

Auch im Beispiel (15) wirkt wohl nicht das Gesetz der wachsenden Glieder.

- (15) mit enormem finanziellen Aufwand (Süddeutsche Zeitung, 21.12.1992, S. 0)

Zwar ist das zweite Adjektiv wie zu erwarten länger als das erste. Das ist aber wahrscheinlich nicht der Grund für die Wortstellung. Vielmehr ist *enorm* ein qualifizierendes Adjektiv und *finanziell* ein Zugehörigkeitsadjektiv. Auch hier gibt ein kategorischer Kontext die Reihenfolge vor. Nicht behauptet werden soll, dass alle Kontexte mit qualitativen Adjektiven und i. d. R. als Zugehörigkeitsadjektive

7 Ermittelt mithilfe eines Algorithmus, der grundsätzlich pro Vokal eine Silbe ausgibt, wobei aber bestimmte Vokalbuchstabenkombinationen als eine Silbe gerechnet werden. Beispiele für solche Kombinationen: /ai/ außer in *etwaig*, *archaisch*, /ee/ außer in *ideenreich*.

einzuordnenden Adjektiven per se kategorische Kontexte mit gestuften Strukturen sind; entscheidend ist immer die Äußerungsabsicht. Nun sind die für gestufte Strukturen typischen Zugehörigkeitsadjektive in der Regel länger (meist auch komplexer) als qualitative Adjektive; vgl. Eichinger (1991). Damit entsprechen solche Belege zwar dem „Gesetz der wachsenden Glieder“, aber in diesen kategorischen Kontexten wirken vor allem das erste Behaghel'sche Gesetz, nach dem „das geistig eng Zusammengehörige auch eng zusammengestellt wird“ (Behaghel 1932: 4 = § 1426), und das dritte, nach dem „das unterscheidende Glied dem unterschiedenen vorausgeht“ (Behaghel 1932: 5 = § 1426).

Irrelevant ist das „Gesetz der wachsenden Glieder“ darum nicht. Betrachten wir die insgesamt 307 Fälle, in denen die Adjektive als tauschbar markiert worden sind (z. B. bei aufeinanderfolgenden qualifizierenden Adjektiven). Das sind Wahlkontexte, in denen das „Gesetz der wachsenden Glieder“ nicht von den beiden semantischen Gesetzen eins und drei überlagert werden kann. In dieser Teilmenge des Datensatzes ist in 144 Belegen das zweite Adjektiv länger als das erste und nur in 77 Belegen das erste Adjektiv länger als das zweite (im Rest der Belege ist die Silbenzahl gleich).⁸

3.6 Erweiterung durch Adverbien und Partikeln

Nach dem „Gesetz der wachsenden Glieder“ wäre auch zu erwarten, dass komplexe Adjektivgruppen (mit Ausnahme der erweiterten Partizipialattribute in gestuften Strukturen, vgl. Abschnitt 3.4) einfachen eher nachfolgen. Ein Beispiel für eine gereichte Struktur, in der die beiden Adjektivgruppen als prinzipiell tauschbar (vgl. Abschnitt 2) markiert worden sind und die dieser Erwartung entspricht, ist (16).

- (16) mit neuem, teilweise gewagtem Glanz (Mannheimer Morgen, 7.11.2008, S. 21)

Wie im Falle der Silbenzahl (Abschnitt 3.5) muss auch diese Hypothese auf Fälle eingeschränkt werden, in denen die Semantik keine zwingende Reihenfolge vorgibt. Im folgenden Beleg könnte das intersektive Zugehörigkeitsadjektiv *therapeutischen* (bzw. die starke Form *therapeutischem*) kaum vor die Adjektivgruppe mit dem nicht intersektiven qualifizierenden Adjektiv *geringem* gestellt werden:

- (17) mit nur geringem therapeutischen Nutzen (Frankfurter Rundschau, 29.10.1999, S. 13)

⁸ $\chi^2(1, 221) = 20,3, p < 0,001$.

Daher kann aus den absoluten Zahlen – 56 erweiterten ersten Adjektiven stehen 27 erweiterte zweite Adjektive gegenüber – nicht direkt geschlossen werden, dass das „Gesetz der wachsenden Glieder“ hier unwirksam, die Hypothese also falsch ist. Denn bei den Adjektiven, die als tauschbar markiert worden sind, überwiegen wieder 19 erweiterte zweite Adjektive 12 erweiterte erste Adjektive. Kontrolliert man jedoch wie in dem Regressionsmodell, das im Anhang beschrieben ist, andere relevante Variablen, ist der Effekt anhand unseres Datensatzes nicht mehr nachweisbar.

3.7 Frequenzphänomene

Ein Unterschied in den Frequenzklassen⁹ der beiden Adjektive begünstigt das Auftreten des selteneren Adjektivs näher am Nomen. Frequente Adjektive stehen also tendenziell vor weniger frequenten (vgl. auch fürs Englische Trotzke & Wittenberg 2019: 278).

Das zweite verwendete Frequenzmaß ist die substantivspezifische Adjektivfrequenz, d. h. die bedingte Wahrscheinlichkeit für das Auftreten des Adjektivs mit einem gegebenen Substantiv. Für das Englische konnte hier ein deutlicher Einfluss auf die Wortstellung nachgewiesen werden: Wulff (vgl. 2003: 267) errechnet für englische Adjektiv-Substantiv-Folgen die Frequenz des Vorkommens von attributivem Adjektiv und Substantiv zusammen geteilt durch die Frequenz des Substantivs. Je höher der Wert für ein Adjektiv ist, desto größer die Wahrscheinlichkeit, dass dieses Adjektiv als zweites von zwei aufeinanderfolgenden Adjektiven (also nah beim Nomen) steht. Ein Fall aus dem vorliegenden Datensatz, auf den diese Vorhersage zutrifft, wäre (18).

- (18) Der klassische Hans-Poelzig-Bau aus dem Jahr 1929 wird endlich in altem, neuem Glanz erstrahlen. (Berliner Zeitung, 27.4.2000, S. 3)

Das Adjektiv *alt* kommt im 7-Mrd.-Untersuchungskorpus 2.835-mal attributiv zu *Glanz* vor. Das Adjektiv *neu* kommt viel häufiger zusammen mit *Glanz* vor, nämlich 16.973-mal. Das Substantiv *Glanz* ist 78.573-mal im 7-Mrd.-Korpus vertreten. 2.835 geteilt durch 78.573 ergibt gerundet 0,0361. Das ist die Wahrscheinlichkeit

9 Für die meisten Adjektivlexeme liegen DeReWo-Frequenzklassen vor: „Dabei hat eine Grundform die Häufigkeitsklasse N, wenn die häufigste Form etwa 2^N -mal häufiger vorkommt als diese Form. [...] Für die Grundformenliste ist der Eintrag mit der höchsten Frequenz ‚der, die, das‘“ (Belica et al. 2012, 2.5 „Häufigkeitsklassen“). Das bedeutet: Je höher die Frequenzklasse, desto seltener das Wort. Fehlende Frequenzklassen wurden nachgetragen.

für das Auftreten von *alt*, wenn *Glanz* gegeben ist. Für *neu* ist der Wert höher, er beträgt 0,2160. Das kann einer der Gründe dafür sein, dass das Adjektiv *neu* in Beleg (18) näher am Nomen steht – aber rechnerisch lässt sich das anhand des Datensatzes `ATTRIBUTIVE_ADJEKTIVE_1` nicht nachweisen. Für die vorliegende Studie wurde die Differenz der entsprechenden Werte für die beiden beteiligten Adjektive verwendet, jedoch konnten wir anhand unseres Datensatzes keinen Effekt feststellen.

Wichtiger ist die substantivspezifische Adjektivfrequenz im vorliegenden Datensatz für die starke/schwache Flexion des zweiten Adjektivs (vgl. Münzberg & Hansen in dieser Ausgabe).

Die beiden untersuchten Frequenzmaße (d.h. Frequenzklasse und substantivspezifische Adjektivfrequenz) können zu gegensätzlichen Vorhersagen hinsichtlich der Adjektivstellung führen. Der Grund dafür ist, dass die substantivspezifische Adjektivfrequenz nicht die absolute Frequenz des Adjektivs berücksichtigt. So ist für ein allgemein sehr häufig vorkommendes Adjektiv auch erwartbar, dass es zusammen mit einem gegebenen Substantiv häufiger vorkommt als ein allgemein selteneres Adjektiv. Die Frequenzklasse des häufigeren Adjektivs lässt dann vermuten, dass es an erster Position steht, die substantivspezifische Adjektivfrequenz würde dagegen vorhersagen, dass es die zweite Position einnimmt.

3.8 Lexikalisches

Auch wenn man alle anderen oben beschriebenen Einflussfaktoren berücksichtigt, zeigen verschiedene Kandidatenlexeme ein unterschiedliches Stellungsverhalten (hier modelliert als zufälliger Effekt; im Anhang unter 5.4 findet sich eine detailliertere Darstellung). Die Varianz zwischen individuellen Lexemen ist verhältnismäßig hoch, und welche Position ein Adjektiv einnimmt, wird nicht unwesentlich durch lexikalische Idiosynkrasien bestimmt. Möglicherweise ließe sich ein Teil dieser Variabilität weiter aufklären, indem zusätzliche Eigenschaften von Adjektivlexemen berücksichtigt werden. Beispiele könnten sein: der Wortbildungstyp auch bei Nichtpartizipien, das Alter des Adjektivs, der Polysemiegrad und andere mehr. Weitere Einflussgrößen könnten die relative Häufigkeit komparierter Formen im Gesamtkorpus und die relative Häufigkeit der prädikativen Verwendung im Gesamtkorpus sein (zu beidem vgl. Abschnitt 1) sowie das Vorkommen in bestimmten Textsorten. Bei der Berücksichtigung so vieler Einflussgrößen wären dann typische Merkmalbündel wie gut komparierbar, gut prädikativ verwendbar, Simplex, *alt*, in allen Textsorten vorkommend, hochgradig polysem zu bedenken, denn dadurch entstehen Korrelationen.

4 Zusammenfassung

Dafür, dass links vom Kopfnomen ein attributives Adjektiv B nach einem attributiven Adjektiv A steht, also näher am Nomen, spricht neben dem konkreten Adjektivlexem B (vgl. Abschnitt 3.8) (nach abnehmender Stärke der Effekte geordnet),

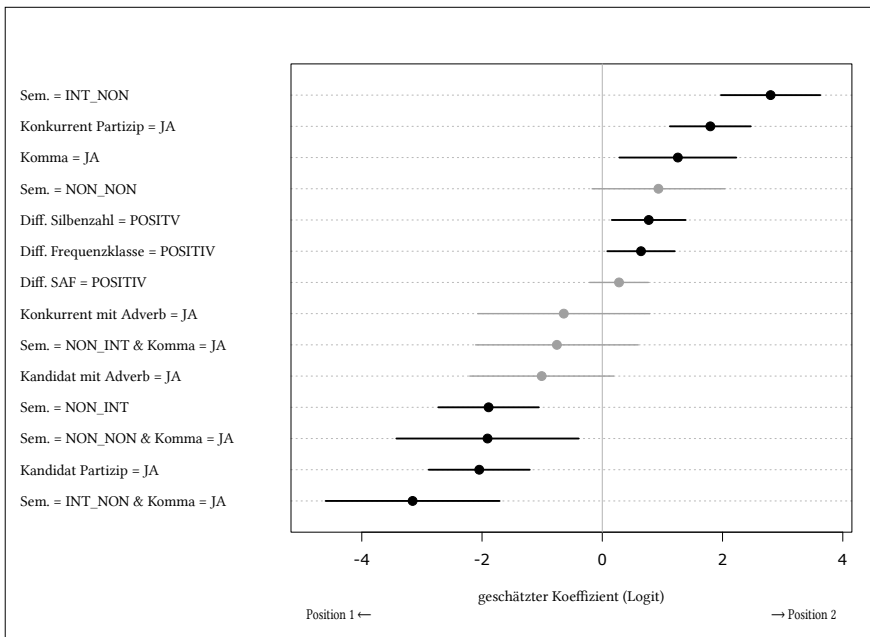


Abb. 1: Koeffizientenschätzungen und 95 %-Konfidenzintervalle für das gemischte logistische Regressionsmodell. Positive Koeffizienten erhöhen die Wahrscheinlichkeit, dass das Kandidatenadjektiv an zweiter Position steht. Erläuterungen zu den einzelnen Prädiktoren finden sich im Anhang. Die Faktoren **Sem.** und **Komma** beziehen sich auf die Referenzausprägung des jeweils anderen Faktors (**Komma = NEIN** bzw. **Sem. = INT_INT**); bei anderen Ausprägungen sind zusätzlich die entsprechenden Interaktionsterme zu beachten. Legende: **Sem. = INT_NON**: Der Kandidat (das betrachtete Adjektiv) ist intersektiv, der Konkurrent (das andere Adjektiv) nicht. **Konkurrent Partizip = JA**: Der Konkurrent ist ein Partizip. **Komma = JA**: Zwischen den beiden Adjektiven steht ein Komma. **Sem. = NON_NON**: Keines der beiden Adjektive ist intersektiv. **Diff. Silbenzahl = POSITIV**: Der Kandidat hat mehr Silben als der Konkurrent. **Diff. Frequenzklasse = POSITIV**: Der Kandidat ist seltener als der Konkurrent. **Diff. SAF = POSITIV**: Der Kandidat kommt häufiger als der Konkurrent attributiv zum Nomen vor. **Konkurrent mit Adverb = JA**: Erweiterung des Konkurrenten durch ein Adverb wie in *ziemlich großsem*. **Sem. = NON_INT & Komma = JA**: Interaktion semantische Klasse (Kandidat nicht intersektiv, Konkurrent intersektiv) mit Komma (JA). **Kandidat mit Adverb = JA**: Erweiterung des Kandidaten durch ein Adverb. **Sem. = NON_INT**: Der Kandidat ist nicht intersektiv, der Konkurrent ist intersektiv. **Sem. = NON_NON & Komma = JA**: Interaktion semantische Klasse (keines der beiden Adjektive ist intersektiv) mit Komma (JA). **Kandidat Partizip = JA**: Der Kandidat ist ein Partizip. **Sem. = INT_NON & Komma = JA**: Interaktion semantische Klasse (der Kandidat ist intersektiv, der Konkurrent nicht) mit Komma (JA).

- dass Adjektiv B intersektiv verwendet wird und Adjektiv A nicht; das Vorhandensein eines Kommas wirkt diesem Effekt jedoch stark entgegen,
- dass Adjektiv A ein Partizip ist,
- dass Adjektiv B mehr Silben hat als Adjektiv A,
- dass Adjektiv B seltener ist (höhere Frequenzklasse) als Adjektiv A.

Dafür, dass Adjektiv B an erster Stelle steht, also weiter entfernt vom Nomen, spricht (nach abnehmender Stärke der Effekte geordnet),

- dass Adjektiv B selbst ein Partizip ist,
- dass Adjektiv A intersektiv verwendet wird und Adjektiv B nicht.

5 Anhang

Für die statistische Modellierung mussten 5 Belege aus dem Datensatz ausgeschlossen werden, da die Frequenzklasse der Adjektive nicht ermittelbar war. Die semantische Klasse von Kandidat und Konkurrent (2-Klassen-System) wurde in einer Variable mit den vier Ausprägungen INT_INT, INT_NON, NON_INT und NON_NON zusammengefasst (jeweils Kandidat_Konkurrent, mit INT = intersektiv, NON = nicht intersektiv). Die Variablen „Differenz Silben“, „Differenz Frequenzklasse“ und „Differenz substantivspezifische Adjektivfrequenz“ wurden dichotomisiert und als Faktoren kodiert, jeweils mit den Ausprägungen POSITIV (Kandidat > Konkurrent) und NICHT-POSITIV (Kandidat ≤ Konkurrent).

5.1 Modellspezifikation

Es wurde ein gemischtes logistisches Regressionsmodell wie folgt spezifiziert:

$$P(y_i = 1) = \text{logit}^{-1}(\alpha_{j[i]} + \beta_0 + \text{DiffSilben}_i \times \beta_1 + \text{DiffSAF}_i \times \beta_2 + \text{DiffFreq}_i \times \beta_3 \\ + \text{KandAdv}_i \times \beta_4 + \text{KonkAdv}_i \times \beta_5 + \text{KonkPartizip}_i \times \beta_6 \\ + \text{Sem}_i \times \beta_7 + \text{Komma}_i \times \beta_8 + \text{Komma}_i \times \text{Sem}_i \times \beta_9)$$

mit

$$\alpha_j \sim N(\gamma_0 + \text{KandPartizip}_j \times \gamma_1, \sigma_\alpha)$$

Abbildung 2: Modellspezifikation, logistische Regression mit festen und zufälligen Effekten (hierarchisches Modell mit variablen Intercepts für die Gruppe der Kandidatenlemmata). Die Notation ist hier abgekürzt, für einen Faktor mit n Ausprägungen werden n-1 Koeffizienten geschätzt. Statt β_7 und β_9 werden also tatsächlich jeweils drei Koeffizienten geschätzt. Die Zusammenfassung in Tab. 5 macht dies deutlich.

Das Logit der Wahrscheinlichkeit, dass im Fall i das Kandidatenadjektiv an zweiter Position steht (also näher am Nomen), wird dabei modelliert als lineare Kombination der Prädiktoren „Differenz Silben“ (*DiffSilben*), „Differenz substantivspezifische Adjektivfrequenz“ (*DiffSAF*), „Differenz Frequenzklasse“ (*DiffFreq*), „Kandidat mit Adverb“ (*KandAdv*), „Konkurrent mit Adverb“ (*KonkAdv*), „Konkurrent ist Partizip“ (*KonkPartizip*), „semantische Klasse von Kandidat und Konkurrent“ (*Sem*), „Komma“ sowie dem Produkt von „semantischer Klasse“ und „Komma“ (Interaktion). Außerdem enthält das Modell eine je nach Kandidatenlexem j variierende Konstante α_j . Diese lexemspezifischen Konstanten werden ihrerseits mittels einer einfachen linearen Regression vorhergesagt, in der die Variable „Kandidat ist Partizip“ (*KandPartizip*) als Prädiktor dient.

5.2 Interpretation

Die Parameter in diesem Modell wurden in R (R Core Team 2019) mit dem *lme4*-Paket (Bates et al. 2015) auf einer Datengrundlage von 1.593 Belegen geschätzt. Tab. 5 zeigt die Koeffizientenschätzungen. Auf der hier dargestellten Ebene der Logits lassen sich die Beiträge eines Prädiktors unabhängig von den Ausprägungen der anderen Prädiktoren ablesen. Ausnahmen sind die Variablen *Komma* und *Sem*, deren Koeffizientenschätzungen wegen des Interaktionsterms nicht unabhängig voneinander sind und nur für die Referenzausprägung des jeweils anderen Faktors (*Komma* = NEIN bzw. *Sem* = INT_INT) gelten. Im Gegensatz zu den Logits ist die Abhängigkeit zwischen Wahrscheinlichkeit und Prädiktoren nicht linear. Über den Einfluss eines Prädiktors auf die Wahrscheinlichkeit lassen sich deshalb genaue Aussagen nur machen, wenn die Ausprägungen aller anderen Prädiktoren berücksichtigt werden, d.h. nur fallweise. Auf der Ebene der Logits können in diesem Modell das Vorzeichen und der Betrag interpretiert werden: Bei Koeffizienten > 0 steigt die vorhergesagte Wahrscheinlichkeit dafür, dass das Kandidatenadjektiv an zweiter Position (näher am Nomen) steht, bei negativen Koeffizienten nimmt sie ab.

5.3 Modellselektion

Für die Koeffizienten der Variablen „Differenz substantivspezifische Adjektivfrequenz“, „Kandidat mit Adverb“ und „Konkurrent mit Adverb“ liegt nicht genug Evidenz dafür vor, dass sie tatsächlich verschieden von null sind (die entsprechenden Signifikanztests haben p -Werte $> 0,05$). Bei einer Modellselektion auf der Grundlage von Likelihood-Ratio-Tests könnten diese drei Faktoren ohne nennenswerte Auswirkungen auf die Modellgüte entfernt werden. Sie wurden

Tabelle 5: Parameterschätzungen für das verwendete gemischte Modell.

Faktor	Ausprägung	geschätzter Koeffizient	Std.-Fehler	z	p
Diff. Silben (Kand-Konk)	POSITIV	0,773	0,313	2,47	0,014
Diff. SAF (Kand-Konk)	POSITIV	0,278	0,249	1,12	0,264
Diff. Frequenzklasse (Kand-Konk)	POSITIV	0,644	0,286	2,25	0,024
Kandidat ist Partizip	JA	-2,046	0,429	-4,77	< 0,001
Konkurrent ist Partizip	JA	1,797	0,343	5,24	< 0,001
Kandidat mit Adverb	JA	-1,008	0,610	-1,65	0,099
Konkurrent mit Adverb	JA	-0,640	0,730	-0,88	0,380
Komma	JA	1,257	0,496	2,53	0,011
Semantik (Kand_Konk)	INT_NON	2,799	0,421	6,65	< 0,001
	NON_INT	-1,891	0,426	-4,44	< 0,001
	NON_NON	0,934	0,563	1,66	0,097
Komma : Semantik (Kand_Konk)	JA : INT_NON	-3,155	0,738	-4,27	< 0,001
	JA : NON_INT	-0,755	0,690	-1,10	0,274
	JA : NON_NON	-1,990	0,774	-2,47	0,014
1.593 Beobachtungen, 726 Ausprägungen in der Gruppe Kandidatenlemma					
Standardabweichung für die lemmaspezifischen Konstanten $\sigma_a = 2,65$					

hier jedoch im Modell belassen, weil zu jedem dieser Faktoren eine sprachwissenschaftlich begründete Hypothese bestand (die in dieser Studie nicht erhärtet werden konnte) und die Auswahl eines optimalen statistischen Modells nicht im Vordergrund steht.

Da im Zusammenhang mit der Adjektivflexion Auffälligkeiten bei Daten aus der Schweiz feststellbar sind (vgl. Münzberg & Hansen in dieser Ausgabe), wurde tentativ der Faktor „Land“ (dichotomisiert als SCHWEIZ, NICHT-SCHWEIZ) mit ins Modell aufgenommen, um eventuelle Interaktionen mit anderen Einflussgrößen aufzudecken. Für keine der möglichen Interaktionen zwischen „Land“ und anderen Faktoren gibt es jedoch hinreichende Evidenz (gemäß Likelihood-Ratio-Tests), sodass der Faktor „Land“ nicht weiter im Modell berücksichtigt wurde. Für den Einfluss anderer nicht-sprachlicher Größen (wie „Medium“ oder

„Register“) auf die Reihenfolge attributiver Adjektive sind uns keine Hypothesen bekannt, sie wurden deshalb nicht berücksichtigt.

5.4 Zufälliger Effekt

Ein Lexem lässt sich einerseits als Individuum betrachten, andererseits lassen sich ihm aber auch inhärente Eigenschaften zuschreiben. Als eine inhärente Eigenschaft von Lexemen wurde in dieser Studie berücksichtigt, ob eines der Adjektive (der „Kandidat“) ein Partizip ist bzw. aus einem Partizip gebildet worden ist. Der Einfluss individueller Lexeme lässt sich in einer Regressionsanalyse als Zufallseffekt modellieren, sodass (in unserem Fall) für jedes Lexem ein konstanter Wert vorhergesagt und bei der Modellierung der Adjektivstellung berücksichtigt wird. Inhärente Eigenschaften eines Lexems können dazu beitragen, diese lexemspezifische Konstante genauer vorherzusagen: Solche Eigenschaften stellen die systematische Komponente bei der Vorhersage der lexemspezifischen Konstanten dar. Hinzu kommt eine zufällige Komponente u. a. aufgrund von wirklichen lexemspezifischen Idiosynkrasien und unberücksichtigten (aber im Prinzip systematisierbaren) inhärenten Eigenschaften der Lexeme (vgl. Gelman & Hill 2007: Kap. 12).

Die unerklärte Varianz zwischen den Kandidatenlexemen – also Varianz, die nicht durch die lexemspezifische Eigenschaft „Kandidat ist Partizip“ erklärt wird – ist verhältnismäßig hoch ($\sigma_\alpha = 2,65$). Ein möglicher Grund dafür ist, dass wichtige Einflussfaktoren auf der Gruppenebene (lexemspezifische Eigenschaften des Kandidaten) nicht berücksichtigt wurden.

Wenn, wie oben beschrieben, für jedes Kandidatenlexem eine eigene Konstante vorhergesagt und bei der Modellierung der Wahrscheinlichkeit verwendet wird, ist eine bessere Anpassung des Modells an die Daten zu erwarten. Deutlich wird das beispielsweise, wenn man die Vorhersagegenauigkeit des Modells für die beiden diskreten Klassen „1“ (Kandidat steht an zweiter Position) und „0“ (Kandidat steht nicht an zweiter Position) betrachtet. Dabei werden die vom Modell vorhergesagten Wahrscheinlichkeiten üblicherweise so in zwei Klassen eingeteilt, dass Wahrscheinlichkeiten $> 0,5$ auf „1“ abgebildet werden und Wahrscheinlichkeiten $\leq 0,5$ auf „0“. Durch den anschließenden Vergleich mit den beobachteten Daten ergibt sich der Anteil korrekt vorhergesagter Fälle. Für das beschriebene gemischte Modell liegt diese Vorhersagegenauigkeit bei 95%. Macht man jedoch Vorhersagen für ungesehene Daten, bei denen einige Kandidatenlemmata neu sind und das Modell entsprechend keinen Zufallseffekt vorhersagt, fällt die Vorhersagegenauigkeit auf 80% im Mittel (bei 10-facher Kreuzvalidierung). Das Modell generalisiert demnach weniger gut auf neue Daten und ein wesentlicher Teil der Variation zwischen den Kandidatenlexemen wird nicht

durch die lexemspezifische Eigenschaft „Kandidat ist Partizip“ erklärt. Berücksichtigt man den Zufallseffekt bei der Vorhersage nicht (oder verwendet man ein einfaches logistisches Regressionsmodell ohne Zufallseffekt), liegt die mittlere Vorhersagegenauigkeit bei 74 %.

Möglicherweise ließe sich die unerklärte Varianz zwischen den Kandidatenlexemen verringern, indem zusätzliche Eigenschaften des Kandidatenlexems berücksichtigt werden.

Andere Eigenschaften von Adjektivlexemen in Nominalphrasen mit zwei attributiven Adjektiven wurden in dieser Studie mit Bezug auf das jeweils andere Adjektivlexem im Beleg erfasst (z. B. die Differenz in der Silbenzahl und die Differenz in der Frequenzklasse). Die resultierenden Werte sind damit keine Eigenschaften eines einzelnen Lexems, sondern Eigenschaften eines spezifischen Belegs und tragen damit nicht zur Aufklärung der Varianz auf der Ebene der Lexeme bei.

5.5 Modellgüte

Ein Indikator für die Modellgüte ist der Anteil der vom Modell erklärten Varianz in den Daten. Für generalisierte lineare Modelle greift man üblicherweise auf eine Reihe von Pseudo- R^2 -Maßen zurück. Wir berichten im Folgenden das im MuMIn-Paket (Bartoń 2019) implementierte Pseudo- R^2 von Nakagawa et al. (2017). Dieses ist auch für gemischte Modelle definiert und unterscheidet zwischen Varianz, die durch die festen Effekte erklärt wird (marginales R^2), und Varianz, die durch das gesamte Modell einschließlich der zufälligen Effekte erklärt wird (konditionales R^2). Für das hier beschriebene gemischte Modell ist das marginale $R^2 = 0,32$ und das konditionale $R^2 = 0,78$. Ein Großteil der Varianz wird also durch den zufälligen Effekt erklärt, die festen Effekte alleine (d. h. die Prädiktoren im Modell) tragen nur mäßig dazu bei, die Varianz zu erklären.

Als weiterer Anhaltspunkt zur Modellevaluierung (zumindest, was den Beitrag des zufälligen Effekts betrifft) bietet sich die Vorhersagegenauigkeit an, wenn das Modell als Klassifikator verwendet wird. Wie in Abschnitt 5.4 angesprochen, werden dabei üblicherweise vom Modell vorhergesagte Wahrscheinlichkeiten $> 0,5$ auf „1“ abgebildet und alle anderen auf „0“. Durch den anschließenden Vergleich mit den beobachteten Daten ergibt sich der Anteil korrekt vorhergesagter Fälle. Dieser lässt sich noch ins Verhältnis setzen zum Anteil korrekt vorhergesagter Fälle, den man mit einem „Nullmodell“ erhalten würde, das immer die häufigere Kategorie (Kandidat steht an näher am Nomen oder Kandidat steht nicht näher am Nomen) vorhersagt. Im hier verwendeten Datensatz liegt dieser Anteil bei 50 %, weil die Daten nach dem Zufallsprinzip einer dieser beiden Kategorien zugeordnet wurden. Der PRE-Wert (proportional reduction in error) gibt an, um

Tabelle 6: Mittlere Vorhersagegenauigkeit und mittlere Reduzierung des Vorhersagefehlers (proportional reduction in error, PRE) bei 10-facher Kreuzvalidierung. GLMM entspricht dem in Abb. 2 und Tab. 5 dargestellten gemischten Modell; GLMM (ohne RE) entspricht demselben Modell, die Vorhersagen wurden jedoch ohne Berücksichtigung des zufälligen Effekts gemacht; GLM entspricht einem einfachen logistischen Regressionsmodell mit denselben unabhängigen Variablen wie in GLMM.

	Trainingsdaten		Testdaten	
	Genauigkeit %	PRE	Genauigkeit %	PRE
GLMM	94,9	0,89	80,2	0,60
GLMM (ohne RE)	74,6	0,49	74,1	0,48
GLM	74,0	0,48	73,6	0,47

welchen Anteil der Vorhersagefehler von ursprünglich 50 % falsch klassifizierter Belege reduziert wird, wenn Prädiktoren ins Modell aufgenommen werden. Um eine realistischere Einschätzung darüber zu bekommen, inwieweit das Modell Generalisierungen erfasst hat (anstatt übermäßig Idiosynkrasien des Datensatzes aufzunehmen), wurde hier eine 10-fache Kreuzvalidierung durchgeführt. Dabei wird (iterativ) ein Teil der Daten bei der Parameterschätzung ausgelassen, und Vorhersagen werden anschließend für diesen nicht verwendeten Teil der Daten erzeugt und abgeglichen. Auch hierbei zeigt sich, dass das Modell für ungesehene Datenpunkte, insbesondere solche mit bisher ungesehenen Kandidatenlemmata, weniger genaue Vorhersagen macht. Die Vorhersagegenauigkeit liegt im Mittel bei 80,2 % für die Testdaten (94,9 % für die Trainingsdaten), der mittlere PRE-Wert beträgt 0,6 für die Testdaten (0,89 für die Trainingsdaten), d. h., der Vorhersagefehler wird aufgrund der im Modell enthaltenen Prädiktoren um 60 % gegenüber dem Nullmodell reduziert. Werden die Vorhersagen ohne Berücksichtigung des zufälligen Effekts gemacht, liegt die Genauigkeit bei 74,1 % (Testdaten) bzw. 74,6 % (Trainingsdaten), der Unterschied fällt hier erwartungsgemäß deutlich geringer aus. Der PRE-Wert liegt in diesem Fall bei 0,48. Die verbesserte Vorhersagegenauigkeit bei Berücksichtigung des zufälligen Effekts ist vor allem darauf zurückzuführen, dass in den Testdaten in der Regel mehr als die Hälfte der Datenpunkte ein Kandidatenlemma haben, das auch in den Trainingsdaten vorhanden ist. Darüber hinaus hat die Modellierung mit einem zufälligen Effekt hier möglicherweise auch eine verbesserte Schätzung der festen Effekte zur Folge.

Tab. 6 fasst diese Daten zusammen und zeigt den Vergleich zu einem einfachen GLM ohne zufälligen Effekt.

ORCID®

Felix Bildhauer  <https://orcid.org/0000-0002-6567-5987>

Franziska Münzberg  <https://orcid.org/0000-0002-4780-8440>

Literatur

- Adam, Séverine & Michael Schecker. 2011. Position und Funktion: Kognitive Aspekte der Abfolge attributiver Adjektive. In Günter Schmale (Hg.), *Das Adjektiv im heutigen Deutsch. Syntax Semantik Pragmatik* (Eurogermanistik 29), 157–172. Tübingen: Stauffenburg.
- Bartoń, Kamil. 2019. *MuMIn: Multi-Model Inference*. R package version 1.43.6.
- Bates, Douglas, Martin Mächler, Ben Bolker & Steve Walker. 2015. Fitting linear mixed-effects models using lme4. *Journal of Statistical Software* 67(1). 1–48.
- Behaghel, Otto. 1909. Beziehungen zwischen Umfang und Reihenfolge von Satzgliedern. *Indogermanische Forschungen* (25). 110–142.
- Behaghel, Otto. 1932. *Deutsche Syntax IV* (Germanische Bibliothek: Abteilung 1, Sammlung germanischer Elementar- und Handbücher. Reihe 1, Grammatiken 10,4). Heidelberg: Winter.
- Belica, Cyril, Marc Kupietz, Harald Lungen, & Rainer Perkuhn. 2012. *DeReWo – Korpusbasierte Grund-/Wortformenlisten*. *derewo-v-ww-bll-320000g-2012-12-31-1.0*. <http://www1.ids-mannheim.de/kl/projekte/methoden/derewo.html> (abgerufen am 13.11.2019).
- Davidse, Kristin & Tine Breban. 2019. A cognitive-functional approach to the order of adjectives in the English noun phrase. In: *Linguistics* 57(2). 327–371.
- Demonte, Violeta. 2019. Adjectives. In Claudia Maienborn, Klaus von Heusinger & Paul Portner (Hgg.), *Semantics. Lexical structures and adjectives* (Mouton reader), 381–414. Berlin & Boston: De Gruyter.
- Dudengrammatik 2016 = Wöllstein, Angelika & Dudenredaktion (Hg.). 2016. *Dudenband 4 – Die Grammatik*, 9. Aufl. Berlin: Dudenverlag.
- Duden online: www.duden.de (abgerufen am 7.7.2020).
- Dyer, William Edward. 2017. *Minimizing Integration Cost: A General Theory of Constituent Order*. University of California, Davis, CA. <https://ucdavis.app.box.com/s/h4yk2plex1f4ls7d99zfomyr8ttxlkjv> (abgerufen am 7.7.2020).
- Eichinger, Ludwig M. 1991. Ganz natürlich – aber im Rahmen bleiben. Zur Reihenfolge gestufter Adjektivattribute. *Deutsche Sprache* 4. 312–329.
- Eisenberg, Peter. 2013. *Grundriss der deutschen Grammatik. Bd. 2: Der Satz*, 4. Aufl. Stuttgart & Weimar: Metzler.
- Engel, Ulrich. 1988. *Deutsche Grammatik*. Heidelberg: Groos.
- GDS = Zifonun, Gisela, Ludger Hoffmann & Bruno Strecker (Hg.). 1997. *Grammatik der deutschen Sprache* (Schriften des Instituts für Deutsche Sprache 7). 3 Bände. Berlin & New York: De Gruyter.
- Gelman, Andrew & Jennifer Hill. 2007. *Data Analysis Using Regression and Multi-level/Hierarchical Models*. Cambridge: Cambridge University Press.

- Kotowski, Sven & Holden Härtl. 2019. How real are adjective order constraints? Multiple prenominal adjectives at the grammatical interfaces. *Linguistics* 57(2). 395–427.
- Landis, J. Richard & Gary G. Koch. 1977. The Measurement of Observer Agreement for Categorical Data. *Biometrics* 33. 159–174.
- Murelli, Adriano. 2017. Partizipialattribute. In Lutz Gunkel, Adriano Murelli, Susan Schlotthauer, Bernd Wiese & Gisela Zifonun (Hgg.), *Grammatik des Deutschen im europäischen Vergleich. Das Nominal* (Schriften des Instituts für Deutsche Sprache 14), 1807–1838. Berlin: De Gruyter.
- Nakagawa, Shinichi, Paul C.D. Johnson & Holger Schielzeth. 2017. The coefficient of determination R^2 and intra-class correlation coefficient from generalized linear mixed-effects models revisited and expanded. *Journal of the Royal Society Interface* 134(14). 1–11.
- R Development Core Team. 2019. R: A Language and Environment for Statistical Computing. Wien: R Foundation for Statistical Computing.
- Romero, Maribel. 2008. *Ling215: Semantik II: Non-verbal predicates. Modifiers. Heim-Kratzer Chapter 4*. <http://ling.uni-konstanz.de/pages/home/romero/classes/wiseo809/215/chapter4-modifiers.pdf> (abgerufen am 7.7.2020).
- Russell, Gillian (Hg.). 2012. *The Routledge companion to philosophy of language* (Routledge philosophy companions). New York: Routledge.
- Schlotthauer, Susan. 2017. Nicht-possessive PP- und NP-Attribute. In Lutz Gunkel, Adriano Murelli, Susan Schlotthauer, Bernd Wiese & Gisela Zifonun (Hgg.), *Grammatik des Deutschen im europäischen Vergleich. Das Nominal* (Schriften des Instituts für Deutsche Sprache 14), 1638–1697. Berlin: De Gruyter.
- Sommerfeldt, Karl-Ernst. 1971. Zur Wortstellung in der Gruppe des Substantivs. *Deutsch als Fremdsprache* 8(1). 13–19.
- Trotzke, Andreas & Eva Wittenberg. 2019. Long-standing issues in adjective order and corpus evidence for a multifactorial approach. *Linguistics* 57(2). 273–282.
- Truswell, Robert. 2009. Attributive Adjectives and Nominal Templates. *Linguistic Inquiry* 40(3). 525–533.
- Wulff, Stefanie. 2003. A multifactorial corpus analysis of adjective order in English. *International Journal of Corpus Linguistics* (8:2). 245–282.
- Zifonun, Gisela. 2017a. Relativsyntagmen. In Lutz Gunkel, Adriano Murelli, Susan Schlotthauer, Bernd Wiese & Gisela Zifonun (Hgg.), *Grammatik des Deutschen im europäischen Vergleich. Das Nominal* (Schriften des Instituts für Deutsche Sprache 14), 1736–1805. Berlin: De Gruyter.
- Zifonun, Gisela. 2017b. Überblick. In Lutz Gunkel, Adriano Murelli, Susan Schlotthauer, Bernd Wiese & Gisela Zifonun (Hgg.), *Grammatik des Deutschen im europäischen Vergleich. Das Nominal* (Schriften des Instituts für Deutsche Sprache 14), 16–31. Berlin: De Gruyter.

Über die Autorinnen und Autoren

Felix Bildhauer  <https://orcid.org/0000-0002-6567-5987>

ist wissenschaftlicher Mitarbeiter am Leibniz-Institut für Deutsche Sprache in Mannheim (Abteilung Grammatik) sowie an der Humboldt-Universität zu Berlin (SFB 1412). Seine Forschungsschwerpunkte liegen in den Bereichen grammatische Variation, quantitative Korpuslinguistik, statistische Modellierung sowie Korpuserstellung.


Patrick Brandt ist wissenschaftlicher Mitarbeiter im Projekt „Korpusgrammatik“ in der Abteilung Grammatik des Leibniz-Instituts für Deutsche Sprache in Mannheim und Privatdozent in der Germanistischen Linguistik an der Universität Mannheim. Seine Forschungsinteressen umfassen Syntax, Morphologie und Semantik und ihre Schnittstellen sowie Korpus- und Psycholinguistik

Sandra Hansen  <https://orcid.org/0000-0003-4243-0220>

ist Mitarbeiterin im Projekt „Korpusgrammatik“ in der Abteilung Grammatik des Leibniz-Instituts für Deutsche Sprache in Mannheim. Ihre Forschungsinteressen umfassen Korpuslinguistik, Variation und Wandel, Grammatik, Quantitative Linguistik, empirische Methoden und Psycholinguistik.

Marek Konopka  <https://orcid.org/0000-0001-9627-0918>

ist Leiter des Projekts „Korpusgrammatik“ in der Abteilung Grammatik des Leibniz-Instituts für Deutsche Sprache in Mannheim. Seine Forschungsinteressen umfassen Grammatik, Korpuslinguistik, Sprachgeschichte, Computerlinguistik und Psycholinguistik.

Franziska Münzberg  <https://orcid.org/0000-0002-4780-8440>

ist Mitarbeiterin im Projekt „Korpusgrammatik“ in der Abteilung Grammatik des Leibniz-Instituts für Deutsche Sprache in Mannheim. Ihre Forschungsinteressen umfassen Grammatik, Korpuslinguistik, Computerlinguistik und Lexikografie.

„Bausteine einer Korpusgrammatik des Deutschen“ ist eine neue Schriftenreihe, die am Leibniz-Institut für Deutsche Sprache in Mannheim (IDS) entsteht. Sie setzt sich zum Ziel, mit korpuslinguistischen Methoden die Vielfalt und Variabilität der deutschen Grammatik in großer Detailschärfe zu erfassen und gleichzeitig für die Validierbarkeit der Ergebnisse zu sorgen. Die erste Ausgabe enthält eine Einführung in die Reihe sowie vier als Kapitel einer neuen Grammatik gestaltete Texte: 1. Grundlegende Aspekte der Wortbildung, 2. Bau von und Umbau zu Adverbien, 3. Starke vs. schwache Flexion aufeinanderfolgender attributiver Adjektive und 4. Reihenfolge attributiver Adjektive. Die Ausgabe ist mit einer interaktiven Datenbank zu attributiven Adjektiven verknüpft.

Marek Konopka ist Leiter des Projekts „Korpusgrammatik“ in der Abteilung Grammatik des Leibniz-Instituts für Deutsche Sprache in Mannheim.

Angelika Wöllstein ist Leiterin der Abteilung Grammatik des Leibniz-Instituts für Deutsche Sprache in Mannheim und Universitätsprofessorin für germanistische Linguistik an der Universität Mannheim.

Ekkehard Felder hat eine Professur für Germanistische Linguistik an der Universität Heidelberg und ist einer der Leiter des Europäischen Zentrums für Sprachwissenschaften (EZS).

