

CLIP 1



**Korpuslinguistik und interdisziplinäre
Perspektiven auf Sprache**

**Corpus Linguistics and
Interdisciplinary Perspectives on Language**

Bd./Vol. 1

Herausgeber / Editorial Board:

Holger Keibel, Marc Kupietz, Christian Mair

Gutachter / Advisory Board:

Heike Behrens, Mark Davies, Martin Hilpert,
Reinhard Köhler, Ramesh Krishnamurthy, Ralph Ludwig,
Michaela Mahlberg, Tony McEnery, Anton Näf,
Michael Stubbs, Elke Teich, Heike Zinsmeister

Marek Konopka/Jacqueline Kubczak
Christian Mair/František Štícha
Ulrich H. Waßner (Hgg.)

Grammatik und Korpora 2009

Dritte Internationale Konferenz

Grammar & Corpora 2009

Third International Conference

Mannheim, 22.-24.09.2009

narr |
VERLAG

Bibliografische Information der Deutschen Nationalbibliothek

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

© 2011 Narr Francke Attempto Verlag GmbH + Co. KG
Dischingerweg 5 · D-72070 Tübingen

Das Werk einschließlich aller seiner Teile ist urheberrechtlich geschützt.
Jede Verwertung außerhalb der engen Grenzen des Urheberrechtsgesetzes ist ohne
Zustimmung des Verlages unzulässig und strafbar. Das gilt insbesondere für
Vervielfältigungen, Übersetzungen, Mikroverfilmungen und die Einspeicherung und
Verarbeitung in elektronischen Systemen.
Gedruckt auf chlorfrei gebleichtem und säurefreiem Werkdruckpapier.

Internet: www.narr.de
E-Mail: info@narr.de

Layout: Tröster, Mannheim
Printed in Germany

ISSN 2191-9577
ISBN 978-3-8233-6648-5

Inhalt / Contents

Vorwort / Preface 9

I. Plenarvorträge / Keynote Speeches

Bruno Strecker: Korpusgrammatik zwischen reiner Statistik und „intelligenter“ Grammatikografie 23

Douglas Biber / Bethany Gray: Is conversation more grammatically complex than academic writing? 47

Mirjam Fried: Grammatical analysis and corpus evidence 63

Françoise Gadet: What can be learned about the grammar of French from corpora of French spoken outside France 87

II. Korpusgestützte Grammatikforschung / Corpus-based Grammar Research

Christa Dürscheid / Stephan Elspaß / Arne Ziegler: Grammatische Variabilität im Gebrauchsstandard: das Projekt „Variantengrammatik des Standarddeutschen“ 123

Reinhard Fiehler: Korpusbasierte Analyse von Univerbierungsprozessen 141

Hagen Hirschmann: Eine für Korpora relevante Subklassifikation adverbialer Wortarten 157

Franziska Münzberg: Korpusrecherche in der Dudenredaktion: Ein Werkstattbericht 181

Per Bærentzen: Einige neue Regularitäten im Gebrauch der Pronominalformen *deren* und *derer* 199

Geert Stuyckens: Zum Wesen der Subjektlücken in Verbzweitkoordination auf der Grundlage eines deutsch > niederländischen Übersetzungskorpus 213

Elma Kerz: The role of low-level schemas in English academic writing. A usage-based constructionist approach.....	229
Markéta Malá: Copular clauses in English and in Czech – a comparative corpus-based approach	253
Svetlana Gorokhova: The role of frequency effects in the selection of inflected word forms: A corpus study of Russian speech errors.....	267
Francesca Strik Lievers: Constructing Judgments. The Interaction between Adjectives and Clausal Complements in Italian	287
Lisa Brunetti / Stefan Bott / Joan Costa / Enric Vallduví: A multilingual annotated corpus for the study of Information Structure.....	305

III. Methodologie korpuslinguistischer Grammatikforschung / Methodologies of corpus-linguistic Grammar Research

Holger Keibel / Cyril Belica / Marc Kupietz / Rainer Perkuhn: Approaching grammar: Detecting, conceptualizing and generalizing paradigmatic variation	329
Oliver Mason: Reconciling Phraseology and Grammar	357
Milena Hebal-Jeziarska / Neil Bermel: Frequency and oppositions in corpus-based research into morphological variation	373
Stella Neumann: Contrasting frequency variation of grammatical features.....	389
Thomas Herbst / Susen Faulhaber: Optionen der Valenzbeschreibung. Ein Valenzmodell für das Englische.....	411
Amir Zeldes: On the Productivity and Variability of the Slots in German Comparative Correlative Constructions.....	429
Cyril Belica / Marc Kupietz / Andreas Witt / Harald Lungen: The Morphosyntactic Annotation of DeReKo: Interpretation, Opportunities, and Pitfalls	451

IV. Einblicke in die aktuelle Forschung / Insights into current studies

František Štícha: Der kommunikative und der systembezogene Status grammatischer Phänomene mit niedriger Häufigkeit	473
Said Sahel: Monoflexion als Erklärung für Variation in der Nominalphrasenflexion des Deutschen	485
Georg Albert: Innovative Sprachverwendungen: Verbreitung und Kontext.....	495
Eva Breindl / Maik Walter: Kausalverknüpfungen im Deutschen. Eine korpusbasierte Studie zum Zusammenspiel von Konnektorbedeutung, Kontexteigenschaften und Diskursrelationen.....	503
Manfred Stede / Uwe Küßner: Kausale Konnektoren in der Automatischen Textanalyse.....	513
Julia Richling: Diachrone Analyse eines Newsgroup / Webforum-Korpus	521
Tomas By: The Prolog version of the Tiger Dependency Bank.....	531
Silke Scheible / Richard Jason Whitt / Martin Durrell / Paul Bennett: Investigating diachronic grammatical variation in Early Modern German. Evidence from the <i>GerManC</i> corpus.....	539
Christopher Cox: Quantitative perspectives on syntactic variation: Investigating verbal complementation in a corpus of Mennonite Plautdietsch	549
Silvia Hansen-Schirra: Empirical profiling of LSP grammar	557
Olga O. Boriskina: Noun Cryptotype Analysis as an Approach to Corpus-driven Modelling of N+V Collocations	567
Siaw-Fong Chung / Yu-Wen Tseng: Learning Prepositions: A Corpus-based Study in Taiwan EFL Contexts.....	575
Svetlana Savchuk: The <i>Russian National Corpus</i> as a Tool for Research on Grammatical Variability.....	585
Ruska Ivanovska-Naskova: Italian-Macedonian parallel corpus	599

FRANZISKA MÜNZBERG

Korpusrecherche in der Dudenredaktion: Ein Werkstattbericht

Abstract

Thema des Beitrags ist der Einsatz des Dudenkorpus in der Zusammenarbeit von Grammatikautoren und Dudenredaktion. Das annotierte Korpus und die Recherchemöglichkeiten, die es bietet, werden anhand aktueller Beispiele aus der Werkstatt einer Dudenredakteurin beschrieben. Einen Schwerpunkt bildet neben einfachen Vergleichen zwischen zwei oder drei morphologischen Varianten die komplexere Frage, ob temporales *wo* (*der Zeitpunkt, wo; jetzt, wo*) in der Dudengrammatik weiterhin als standardsprachlich bezeichnet werden soll. Zugleich wird versucht, die Attraktivität alternativer Konstruktionen (*der Zeitpunkt, zu dem; jetzt, da*) für Schreibende und Lesende zu messen. Diese 'Alternativen' verhalten sich jedoch keineswegs wie die eingangs erwähnten morphologischen Varianten zueinander – zu unterschiedlich sind semantische und syntaktische Leistungen, zu unterschiedlich die Restriktionen, die für ihre Verwendung im Satz gelten, zu unterschiedlich sind schließlich die untersuchten Texte, aus denen die mittels Hochrechnung ausgewerteten über 30 000 Sätze stammen. Zur Diskussion steht, welche Konsequenzen in einer Grammatik für ein breites Publikum zu ziehen sind. Diese Frage wird für die 'Wortgrammatik' anders beantwortet als für die 'Regelgrammatik'.

1. Anlass für die Erhebung empirischer Daten

Undenkbar wäre die Arbeit der Dudenredaktion und ihrer Autor(inn)en ohne Korpora. Aussagen über Morphologie, Syntax und Semantik in Wörterbüchern, Grammatiken und Ratgebern wurden früher anhand von Exzerptsammlungen formuliert, während sie heute mit den Ergebnissen aus der Recherche in elektronischen Korpora konform sein sollen. Das setzt voraus, dass immer wieder anhand neuer Datensätze, mit neuen Recherchemethoden und offen für unterschiedliche Formen von Normbewusstsein alte Aussagen überprüft und Aussagen über neue Themen gemacht werden. Das Ideal ist hier noch nicht die sich selbst schreibende korpusbasierte Grammatik oder die 'emergente' Grammatik, die eine reale Sprachentwicklung in einer Gesellschaft, bei einem Kind, in einer bestimmten Sprechergruppe imitieren kann, indem sie Frequenzen in Regeln umrechnet. Eine solche Grammatik würde

beispielsweise anhand der vorliegenden Korpora geschriebener Sprache voraussichtlich bei einer großen Anzahl von Verben keine 2. Person und keinen Imperativ erzeugen oder das Paradigma zwar aufstellen, aber eine Warnung ausgeben, dass diese Phänomene selten und daher vermutlich nicht standardsprachlich seien. Von der Dudenredaktion erwartet die Öffentlichkeit vielmehr Antworten auf häufig gestellte Fragen, die viele Benutzer(innen), Laien ebenso wie Sprachprofis, brennend interessieren. Besonders was in der telefonischen Sprachberatung erfragt wird, müssen die Dudenbände 4, „Die Grammatik“, und 9, „Richtiges und gutes Deutsch“, oder die Wörterbücher beantworten. Damit geraten Themen in den Vordergrund, die als Zweifelsfälle bewusst diskutiert werden.

2. Arbeit mit dem Dudenkorpus (Grammatik)

Das Dudenkorpus gehört mit 1,79 Milliarden Wortformen und Satzzeichen (Stand September 2009) zu den größten der deutschen Sprache. Neben Tages- und Wochenzeitungen aus Deutschland, Österreich und der Schweiz enthält es Zeitschriften, Fach- und Sachbücher sowie belletristische Werke unterschiedlicher Genres. Die Teilkorpora sind lemmatisiert, flach annotiert und einzeln oder in beliebiger Zusammenstellung mithilfe der Abfragesprache CQP (Corpus Query Processor) durchsuchbar. Zu den hinterlegten Metadaten gehören neben Quelle und Datum auch etwa die Angabe, ob es sich um eine Übersetzung handelt oder ob in alter Rechtschreibung geschrieben wurde, und eine Zuordnung zu Themen und Sachgebieten. Da das Korpus noch nicht die angestrebten ausgewogenen Anteile einzelner Korpora enthält, werden in regelmäßigen Abständen kleinere ausgewogene Teilkorpora erstellt. Die Redaktion verwendet sie, um aus dem Gesamtkorpus gewonnene Ergebnisse zu kontrollieren. Das Korpus ist nicht öffentlich verfügbar und wird ausschließlich von der Dudenredaktion und ihren Autor(inn)en genutzt.

In der Lexikografie wird das Korpus genutzt, um Einzelbedeutungen zu unterscheiden, Belege zu finden und grammatische Eigenschaften wie Genus und Flexionsklasse zu bestimmen. Zu den typischen Rechercheaufgaben bei der Redaktion einer Grammatik gehört der Vergleich zweier Varianten. Die einfachste Form einer solchen Recherche besteht z.B. in der Anfrage, wie oft das Adjektiv *gesund* im Komparativ und Superlativ umgelautet wird. Dazu muss lediglich festgestellt werden, wie häufig die Formenreihe *gesündere.**, *gesündest.** (2 125 Treffer im Gesamtkorpus) im Vergleich zu *gesundere.**, *ge-*

*sundest.** vorkommt (72 Treffer im Gesamtkorpus; theoretisch könnte sich hierunter ein das Ergebnis leicht verwässerndes *du gesundest* befinden). Ergänzt werden kann diese Anfrage durch die Suche nach der undeklinierten Form in Wortfolgen wie *ist gesünder / gesunder* und *gesünder / gesunder ist* (jeweils ohne Unterbrechung durch Satzzeichen); ein falsch positives Ergebnis erhält man dabei durch die Wortfolge *und der Appetit ein gesunder ist*. Raffinierter und unmöglich ohne Annotation ist die Suche nach [finites Verb, gefolgt von] *gesunder* und *gesunder* [, gefolgt von finitem Verb]; hier ist jedoch mit mehr falsch positiven Ergebnissen zu rechnen. Schon diese Routineanfrage zeigt, um wieviel komfortabler eine Korpusrecherche im Vergleich zu einer Internetrecherche ist. Sie veranschaulicht auch, auf welche Klippen googelnde Laien stoßen könnten. So würde eine Suche nach der Form *gesunder* im Vergleich zur Form *gesünder* die Ergebnisse verfälschen: *Birnen sind ein gesunder Teil unserer Ernährung*, weiß www.misterinfo.de.

Besonders hilfreiche Features des Dudenkorpus und seiner Benutzeroberfläche *Koala* sind, dass die Groß- und Kleinschreibung bei der Suche beachtet wird und dass nach Satzzeichen gesucht werden kann. Der Bereich von Trunkierungen (Platzhaltern, optionalen Buchstaben und Wortformen) und anderen Bedingungen ist steuerbar. Den Nutzer(inne)n bleibt es überlassen, ob sie nach Lemmata oder nach konkreten Wortformen oder Endungen suchen möchten, nach Wortarten bzw. 'Parts of Speech' oder aber nach Kategorisierungen wie etwa Numerus und Kasus. Schließlich kann die Ergebnisanzeige erweitert werden, wenn sich herausstellt, dass der gewählte Kontext von ein oder zwei Sätzen nicht ausreicht, um den Beleg zu interpretieren.

'Interpretieren' ist ein bewusst gewähltes Schlagwort. Tatsächlich wird bei dieser Form der Korpusrecherche nicht nur gezählt und gerechnet, sondern auch gelesen und ausgelegt. Nur so können geeignete Filter gefunden werden, um in einer neuen Suche oder in einer Hochrechnung unerwünschte positive Ergebnisse (im Folgenden „Fehlbelege“ genannt) auszuschließen und um schließlich eine geeignete Darstellung der Ergebnisse in einem Wörterbuch oder einer Grammatik herauszuarbeiten. Um beispielsweise herauszufinden, ob und wo es regelmäßig gebildete Steigerungsformen zum Adjektiv *nah(e)* gibt, muss die Anfrageroutine im Vergleich zu *gesundest.** geändert werden. Sonst werden als Fehlbelege die Wortformen, die mit *nahestehend* beginnen, eingerechnet. Eine unter mehreren Möglichkeiten ist die Suche nach den einzelnen Superlativformen *naheste*, *nahesten*, *nahestem*, *nahester* – mit immerhin 20 erstaunlichen Treffern und ohne jegliche Entsprechung im Komparativ. Genauso gut

hätte man aber auch die optionalen Buchstaben am Ende der Wortform *nahes-te* auf höchstens einen einschränken können. Wegen der verschiedenen Suchmöglichkeiten, die die Ergebnisse unter Umständen beeinflussen könnten, wird meist zusammen mit den Ergebnissen auch die Anfrage gespeichert, bevor die Suchergebnisse an die Autor(inn)en weitergeleitet werden. Anschließend werden die Ergebnisse diskutiert und bewertet; nicht alles irgendwie Auffällige wird gleich nach seiner Entdeckung berücksichtigt. Die *nahesten*-Belege beispielsweise sind noch in keinem Dudenband veröffentlicht. Viel komplexer als die Suche nach morphologischen Eigenschaften wird die korpusbasierte Arbeit an einer Grammatik, wenn syntaktische und semantische Leistungen von Varianten beschrieben werden sollen.

3. Beispiel für eine Recherche: *wo* temporal

Im Abschnitt über Nebensätze (Randnummer 1659) in der 8. Auflage der Dudengrammatik (Duden 2009) beschreibt Peter Gallmann, dass Pro-Adverbien (*wie*, *wo*) anstelle einer Verbindung aus Präposition und Relativpronomen Nebensätze anschließen können:

Die einfachen Pro-Adverbien entsprechen relativen Präpositionalphrasen:

[...] Über die Art, wie (= in der) man schwermütige, tolle und rasende Menschen behandeln müsse, sollte billig ein philosophischer Arzt ein eigenes Werk schreiben. (A. Knigge) [...] – (Auch temporal:) Wenn in unserem Leben etwas Tragisches passiert, kommt der Zeitpunkt, wo (= zu / in / bei dem) wir das Dunkel ins Auge fassen müssen. (Internetbeleg) [...]

Neben Nebensätzen mit temporal zu verstehendem *wo* finden sich auch Nebensätze mit *da* sowie mit *als* und *wenn*, wobei *als* und *wenn* wohl nicht als Adverbien, sondern als Subjunktionen zu bestimmen sind. Alle vier Konstruktionen sind standardsprachlich korrekt:

Aber auch der Zeitpunkt, wo das Kind zum ersten Mal „Nein“ sagen wird, rückt näher. – Dies war der Zeitpunkt, da der Vorstand der Paul-Martini-Stiftung anregte, das Konzept der Stiftungsarbeit zu aktualisieren. – Just zum Zeitpunkt, als das neu errichtete SOS-Kinderdorf hätte besiedelt werden sollen, brach in Liberia Bürgerkrieg aus. – Am besten stellt man Fragen zu dem Zeitpunkt, wenn sie aktuell auftreten. (Internetbelege)

Die Darstellung entspricht seit der 7. Auflage in wesentlichen Zügen der in den großen wissenschaftlichen Grammatiken, vgl. Zifonun et al. (1997, Bd. I: 42), und Eisenberg (2006, Bd. II: 277), mit berechtigter Kritik an der 6. Auflage der

Dudengrammatik (Duden 1998, Randnummer 1332) und einleuchtender Erklärung (das Lokaladverb *wo* kann den Charakter eines „universellen Relativadverbs“ annehmen). Im Folgenden (Duden (2009), Randnummer 1660) zeigt Gallmann, dass man in regionalen Varietäten noch wesentlich kreativer mit *wo* umgeht. Die Vielfalt der regionalen und gesprochensprachlichen Verwendungen von *wo* beschreiben über die genannten Standardwerke hinaus eindrücklich z.B. Günthner (2002), Pittner (2004), Fleischer (2004) und Elspaß/Möller (http://www.philhist.uni-augsburg.de/lehrstuehle/germanistik/sprachwissenschaft/ada/runde_3/f12a-b/ Stand: 08/2010). Trotz seines guten Standings in den Grammatiken fällt das Relativadverb *wo* jedoch regelmäßig Aufsatz- und Manuskriptkorrekturen zum Opfer, sobald es andere als lokale Bezüge herstellen soll. Besonders Lehrer(innen) sind nach Davies (2007: 56f.) dafür verantwortlich, dass die Liebe der Grammatiker(innen) insbesondere zum „temporalen *wo*“, wie es auch hier im Folgenden verkürzend genannt wird, bis jetzt eine eher unglückliche Liebe geblieben ist. Als dieses Auseinanderklaffen zwischen Referenzgrammatik und Grammatikalitätsurteilen nun auch noch auf der IDS-Jahrestagung 2008 offen angesprochen wurde, beschlossen Autor und Redaktion, ihre Meinung nochmals zu überdenken. Dass temporales *wo* in gesprochener Sprache, auch im elaborierten akademischen Vortrag jenseits der Nähesprache, wenigstens in bestimmten Kontexten völlig unmarkiert und sprachlicher Brillanz keineswegs abträglich zu sein scheint, zeigten einige der Tagungsbeiträge selbst. Aber von der Dudengrammatik wird ja unter anderem gefordert, dass sie den geschriebenen Standard beschreibt. Dieser geschriebene Standard wird höchstwahrscheinlich durch Aufsatzkorrekturen in der Schule beeinflusst.

Zunächst muss eingegrenzt werden, welches temporale *wo* überhaupt zum Gegenstand einer sinnvollen Diskussion über die Standardnähe werden kann. Unbestritten nächstsprachlich sind freie Adverbialsätze mit *wo* als einer Art Konjunktion; gleich drei Belege finden sich im folgenden Gesprächsausschnitt:

S2: Ja, ja, wo wir haben fort müssen, anno vierzig. [hier vielleicht noch Relativadverb: Bezug zwischen *wo* und *anno vierzig*]

[...]

S1: (gleichzeitig) Und wer hat hingesehen wo wir heimgekommen sind? Wo ich mal heimgekommen bin, hat ganz NAME hat gebrannt. [keinerlei relativer Anschluss]

(Datenbank gesprochenes Deutsch, Zwirner-Korpus: Alemannisch, 1955; Interaktion ZW038, <http://dsav-wiss.ids-mannheim.de/DSAv/KORPORA/ZW/ZW0/ZW038/ZW038TRA.HTM>; zur regionalen Zuordnung vgl. auch Interaktion OS023 und den bei Pittner (2004: 365) zitierten Hamburger Beleg aus dem Pfeffer-Korpus, PF022).

Interessanter für die Frage nach der Standardkonformität von temporalem *wo* ist die Verwendung nach Adverbphrasen mit *jetzt* oder *heute* als Kern. Pittner (2004: 63) ist der Ansicht, dieser Gebrauch von *wo* müsse schon deshalb als standardsprachlich gelten, weil es in der Standardsprache „keinen gleichwertigen Ersatz“ gebe. So könnten manche Sprecher(innen) *da* anstelle von *wo* in den folgenden Belegen vielleicht als „gehoben“ oder „veraltend“ empfinden; semantisch würde *da* vermutlich dasselbe leisten:

Dieser Aspekt wurde in diesem Kapitel bereits erwähnt (siehe Abbildung 16.3), aber es ist wichtig, ihn jetzt, wo Sie mehr über Transkription und Translation wissen, nochmals aufzugreifen. (Übersetzung eines Biologie-Lehrbuchs von Campbell/Reece) [rein temporal oder auch kausal lesbar?]

Wieso soll der Gaspreis jetzt, wo Öl wieder billiger wird, weiter steigen? (taz, 01.08.2005) [rein temporal oder auch konzessiv bzw. äußerungsbegründend lesbar?]

Zum Vergleich ein Originalbeleg mit *da*:

Zwischen 1995 und 2003 hat sich der Umsatz des schwäbischen Familienunternehmens verdreifacht, und das in einem Zeitraum, da der deutsche Hemdenmarkt von 90 Millionen verkauften Stück auf knapp 62 Millionen geschrumpft ist. (SZ, 31.12.2004)

Überprüft werden sollte, ob die Konstituentenstruktur (Adverbphrase) des Bezugsausdrucks eine Rolle für die Frequenz von *wo*-Relativsätzen spielt. Schließlich sollte untersucht werden, unter welchen Bedingungen *wo* nach Bezugsausdrücken mit *Jahrhundert*, *Tag*, *Augenblick*, *Moment*, *Zeitpunkt* steht, auch um festzustellen, ob die für die Dudengrammatik ausgewählten Beispielsätze, die Nominalphrasen mit *Zeitpunkt* enthalten, typisch und damit geeignet sind.

4. Erste Sichtungen und Arbeitshypothesen

Die unterschiedliche Beurteilung des temporalen *wo* durch Grammatiker(innen) und Lehrkräfte könnte nicht nur auf Unterschieden im Normbewusstsein beruhen, sondern es könnte auch zahlenmäßig erfassbare syntaktische oder semantische Gründe dafür geben, dass *wo* und speziell temporales *wo* als relativer Anschluss einmal markiert, einmal unmarkiert erscheint. Zu Beginn der Untersuchung sah es freilich so aus, als müsste diese Hypothese verworfen werden: Ausgetestet wurden jeweils unterschiedliche Bezugsausdrücke – in verschiedenen Kasus, mit und ohne Präposition (*der Augenblick*, *wo* vs. *in dem Augenblick*, *wo*), mit Wechsel im Numerus (*der Augenblick*, *wo* vs. *die Augenblicke*, *wo*), mit Variation in der Struktur bzw. Wortstellung (*die langen Au-*

*genblicke des Wartens, wo; die Augenblicke, wo ... [Verb] und [Verb]), determiniert (der, dieser Augenblick) oder unbestimmt (ein Augenblick) usw. Meistens dienen die Bezugsausdrücke als Adverbialien bzw. als Teile von Adverbialien – aber das gilt nicht speziell für *wo*-Belege. Keine der Datenreihen lieferte eine sensationell einfache Erklärung.*

Ferner könnte temporales *wo* ja tatsächlich eher in Interviews und umgangssprachlich gefärbten Glossen als in Berichten und Analysen vorkommen. Wenn es wirklich klar als nächstsprachlich markiert ist, dann lässt sich das nicht nur anhand von Metadaten, sondern eventuell auch anhand häufiger Personalpronomina und anhand typisch nächstsprachlicher Wortwahl und Grammatik im selben Satz nachweisen. Auch hier ließen erste Stichproben allerdings noch keine Beweiskraft erwarten – der eine *wo*-Beleg gibt einen Mannheimer Fährmann möglichst im Original wieder, ein anderer einen Berliner Taxifahrer, die nächsten Belege aber stammen aus Fachbüchern über mittelalterliche Geschichte oder über Biotechnologie, aus Theaterkritiken oder Wirtschaftsprognosen. Es konnte keine Textsorte ausgemacht werden, in der temporales *wo* niemals vorkommt, und es wird zu oft auch im gehobenen und im akademischen Stil verwendet, um schlicht als umgangssprachlich abgetan zu werden.

Immerhin aber kristallisierten sich aus diesen ersten Versuchsreihen einige „*wo*-Konkurrenten“ heraus, die den Schreibenden je nach Kontext als Alternativen zur Verfügung stehen. So wurden für die Vergleiche zwischen der Nebensatzeinleitung durch *wo* und alternativen Ausdrucksmöglichkeiten schließlich folgende Kombinationen ausgewählt (die Datenreihen mit Bezugsausdruck *jetzt* stehen als voraussichtlicher Sonderfall gleichsam außer Konkurrenz daneben):

Augenblick	} direkt gefolgt von Komma und	} <i>wo</i>	
Moment			<i>in (der / dem)</i>
Zeitpunkt			<i>an (der / dem)</i>
Jahr (mit Jahreszahl)			<i>zu (der / dem)</i>
Jahrhundert			<i>wenn</i>
Nacht			<i>als</i>
Tag			<i>da</i>
damals			<i>während (deren / derer / dessen)</i>
[Sonderfall: jetzt]			
tagsüber			
tags			

Gefunden wurden über 30 000 Sätze, Belege mit Bezugsausdruck *jetzt* nicht eingerechnet. 1 780 davon enthielten die gesuchte Nebensatzeinleitung mit *wo*. Aus jeder Datenreihe wurden wenn nicht alle, so doch mindestens je 50 Sätze aus unterschiedlichen Quellen gelesen und ausgewertet. Der Anteil an Fehlbelegen wurde hochgerechnet, d.h., bei 10 Fehlbelegen in 100 durchgesehenen Sätzen wurden vom Gesamtergebnis 10 % Fehlbelege abgezogen. Als Fehlbelege wurden beispielsweise gewertet:

Aber jetzt, als Rentner, zieht es ihn auch ab und zu weit fort. (Mannheimer Morgen 19.04.2000) Auch das Schnüffeln des Hundes klingt seltsam jetzt, als wäre es nicht wirklich da, nur dumpfe Erinnerungen an lautere Tage. (SZ, 05.09.1995)

Sie schwieg einen Moment, während sie beschleunigte und in einer uneinsehbaren Kurve zu einem besonders waghalsigen Überholmanöver ansetzte. (P. Mayle [Übersetzung])

Erhellend erscheint die Frage, der derzeit Ulrike Stölzel, IDS Mannheim, in einer eigenen Erhebung nachgeht: ob der Bezugsausdruck von *wo* eher einen Zeitpunkt oder einen Zeitraum bezeichnet. Im Zusammenhang damit steht allgemein die Frage nach Tempus und Modus im übergeordneten Satz und im *wo*-Nebensatz, die beim genaueren Lesen der Belege aus dem Dudenkorpus aufkam. So könnte generalisierendes Präsens („immer wenn“ mit Betonung nicht des Iterativen, sondern des Regelhaften) die Verwendung von *wo* begünstigen. Ein typischer Beleg wäre der folgende; referiert wird auf einen Zeitpunkt:

Anmeldung ist neun Monate vor der Einschulung, das ist der Moment, wo in meinen Augen Schule losgehen kann. (Zeit, 13.02.2002)

Entsprechend im Präteritum (auch *augenblick*, *wo* bezieht sich auf ein für sich genommen punktuell Ereignis, das einer „immer wenn“-Regel nach immer dieselbe Folge hat):

der zuschlag erfolgte z.b. in dem augenblick, wo ein angezündetes licht erlosch oder das ins licht gesteckte geldstück beim herunterbrennen zu boden fiel (Deutsches Wörterbuch, „Versteigerung“, gefunden über ein Zitat in der Computerzeitschrift c't, 08/2001)

Als mündlich, umgangssprachlich oder stärker regional markiert könnten demgegenüber Belege empfunden werden, in denen *Augenblick*, *wo* sich eher auf ein einzelnes erzähltes (auch hier wieder punktuell) Ereignis bezieht. Charakteristisch ist dann möglicherweise das Auftreten der im Korpus ansonsten selteneren Pronomina *ich*, *wir*, *Sie*:

Und drei Monate später kam der Augenblick, wo ich die Zehen ganz schwach bewegen konnte. (K. Meyer 1996)

Auf der anderen Seite im Varietätenspektrum würde man wohl einige der zahlreichen Belege ansiedeln, in denen aus der „immer wenn“-Beziehung eine regelrecht konditionale geworden ist – hier wird nun nicht erzählt, sondern vermutet, behauptet und analysiert:

Doch in dem Augenblick, wo eine entsprechende technische Vorrichtung handlungsbestimmend wird, verändern sich die Bedingungen der Bewertung. (K. Weber 2005)

In dem Moment, wo man den Raum des Rechts betritt, muss man nicht nur den anderen anerkennen, sondern sich auch selbst rechtfertigen. (Zeit, 08.12.2004)

Sucht man hingegen mit Bezugsausdrücken wie *Jahrhundert* gezielt nach der Referenz auf einen längeren Zeitraum, so stößt man auf beinahe lokale Bezüge, in denen das „Jahrhundert“ wie die „Gewölbe“ und „Katakomben“ eher als Bühne oder Setting, mehr als virtueller Raum denn als Ereigniszeit erscheint. Anders als in anderen Datenreihen sind hier auch nichtrestriktive relative Anschlüsse typisch:

Nach einem Ausflug in das 18. Jahrhundert, wo er in den „Gewölben des Dr. Hahnemann“ die Geschichte des Urvaters der Homöopathie unterhaltsam und spannend erforschte, kehrt er jetzt zurück in das Reformationszeitalter: „Der Bader von St. Denis“ heißt sein neuester Roman, in dem er auf den Spuren von Ambroise Par [sic], dem Wegbereiter der modernen Chirurgie wandelt. (Mannheimer Morgen, 02.09.2004)

Oder er stürzt in die dunklen Katakomben der Weltgeschichte und landet im 15. Jahrhundert, wo er mit den blutrünstigen Hussiten für die tschechische Unabhängigkeit kämpfen soll. (SZ, 14.01.1995)

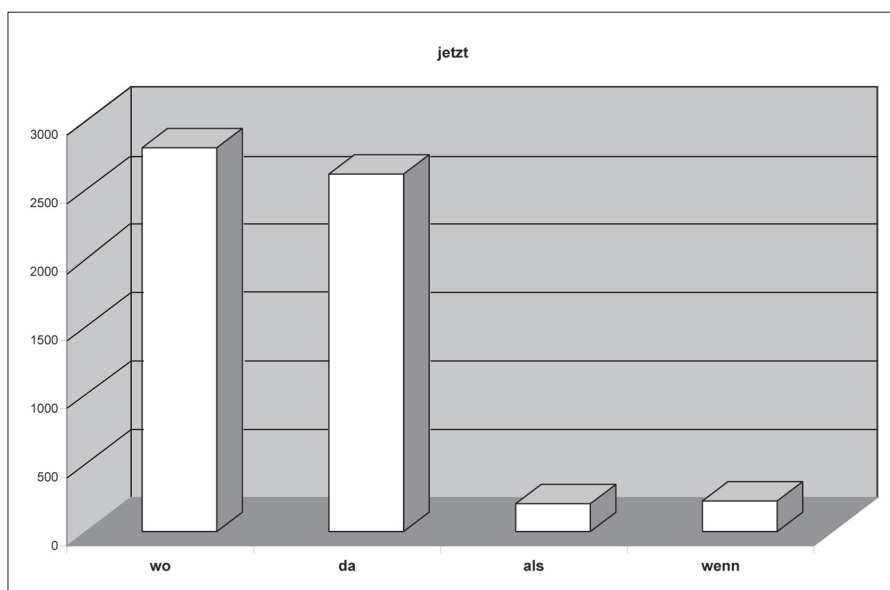
Folgende Schwierigkeiten ergeben sich für die Auswertung:

- 1) Die wichtigsten ‚Konkurrenten‘ von *wo* unterliegen anderen syntaktischen Restriktionen als *wo* und leisten – abgesehen wohl von *da* – semantisch nicht dasselbe wie *wo*. So sind *als* und *wenn* im Deutschen kaum füreinander einsetzbar, und der präpositionale Anschluss ist nicht möglich nach einem Adverb (**tagsüber, an dem*).
- 2) Textausschnitte können aus anderen Quellen zitiert sein, vgl. den oben zitierten Artikel aus dem Grimm'schen Wörterbuch.

- 3) Häufig zitierte Titel von Filmen und Theaterstücken, auch Zitate etwa aus Liedtexten können Ergebnisse verfälschen.
- 4) Besonders in Interviews und in humoristischen Beiträgen wird manchmal der Versuch unternommen, 'unterschichtige' oder anders stigmatisierte Redeweisen nachzuahmen.
- 5) Aussagen über regionale Variation können aus überregional erscheinenden Blättern nur unter Vorbehalt gefolgert werden, weil die Schreibenden oder die Interviewten selbst keineswegs am Erscheinungsort des Blattes geboren oder aufgewachsen sein müssen. Auch könnte die Nachbearbeitung etwa von Interviews in einzelnen Redaktionen und sogar bei einzelnen freien Journalisten, die einer Zeitung besonders viele Beiträge liefern, rigoroser sein als bei anderen.
- 6) Einzelne Faktoren, denen ein Einfluss auf die Verwendung von *wo* zugeschrieben wird, könnten (auch unbemerkt) miteinander korrelieren. Wenn etwa zwei Zeitungen aus unterschiedlichen Regionen als Beispiele für regionale Schreibsprache gegeneinandergehalten werden, so unterscheiden sich diese beiden Zeitungen notwendigerweise noch durch andere Merkmale als den Erscheinungsort. Dieselben Zeitungen könnten z.B. unterschiedlich viele Tokens aus nächstsprachlich geprägten Interviews mit biografischem Schwerpunkt und insgesamt unterschiedlich viele Tokens mit Merkmal 1. oder 2. Person, Präsens usw. enthalten.
- 7) Im Zusammenspiel von Bezugsausdruck und Nebensatzeinleitung können neben der temporalen Bedeutung weitere Bedeutungen mitschwingen. Es ist nicht sicher, ob alle Sprecher(innen) den Wert „temporal“ für alle gefundenen Belege vergeben würden. Eine Auswertung nach dem genauen semantischen Wert, nach Aktionsarten u.Ä. („punktuell“, „immer wenn“; „eher lokal“, „eher konditional“) trägt auch bei möglichst exakter Arbeitsanweisung subjektive Züge. Es ist auch nicht sicher, ob sie alle dieselben Sätze als 'Konkurrenten' zu Relativsätzen mitzählen würden. Als besonders schwierig erwies sich diese Entscheidung in den Datenreihen mit der Nebensatzeinleitung *wenn*; gezählt wurde hier nach dem Grundsatz „im Zweifel für den Beleg“.
- 8) Es stehen keine Hilfskräfte zur Verfügung, um alle Daten auf diese Weise semantisch auszuwerten, alle Korrelationen auszutesten und dabei auch satzübergreifend den Kontext einzubeziehen.

5. Korpusdaten

Alle nachfolgend präsentierten Zahlen sind mit entsprechender Vorsicht zu lesen. Zunächst werden die Datenreihen mit Bezugsausdruck *jetzt* gesichtet – hier lautete die Arbeitshypothese, dass *wo* der einzig mögliche standardsprachliche Anschluss eines Relativsatzes sei –, danach die übrigen Datenreihen, in denen *wo* wohl nur unter bestimmten Umständen unmarkiert wirkt.



Die Ergebnisse zu *jetzt*, *wo* und seinen Konkurrenten (insgesamt 5 834 Belege) bestätigen die Annahme, dass *wo* hier standardsprachlich ist. Allerdings sprechen die Zahlenverhältnisse im Duden-Gesamtkorpus (Stand Frühjahr 2008, 1, 3 Mrd. Tokens) dafür, *da* ebenfalls als standardsprachlich zu bewerten.

Aus den *jetzt*-Belegen selbst zeichnen sich wiederum Muster ab. Ein großer Teil der *wo*- und *da*-Sätze kann nicht nur temporal, sondern zugleich auch kausal oder konzessiv verstanden werden. Referiert wird regelmäßig auf einen Sachverhalt, den die Leser(innen) voraussichtlich nicht bestreiten werden. Die unanfechtbare Aussage im *wo*- / *da*-Nebensatz soll der im Matrixsatz vorgetragenen (unter Umständen durchaus kontroversen) Meinung mehr Nachdruck verleihen. Begünstigt werden diese Lesarten durch Partikeln wie *gerade*, *ausgerechnet*, *selbst*. Wirklich ersetzbar sind *da* und *wo* aber nicht durch die Subjunktionen *weil* und *obwohl*, denn *da* und *wo* sind besser geeignet, einen relativen Bezug herzustellen:

Gerade jetzt, da [weil] in jeder Klinik das Zauberwort „Kostensenkung“ die Runde mache, seien Ärzte mit kaufmännischer Ausbildung bitter nötig. (SZ, 11.04.1996)

Gerade jetzt, wo [obwohl(?)] Reklameaussagen als Eigenschaftszusagen verbindlich werden, sind die Werbetexte seltsam aussagelos. (Computerzeitschrift c't 10/2001)

Die wenigen Nicht-Fehlbelege (nur etwa die Hälfte von 442 Treffern) für *jetzt*, *als* entstammen eher Texten, die im Präteritum oder im historischen Präsens aus der Perspektive einer Figur erzählen. Auch hier kann der auf *jetzt* folgende Nebensatz zugleich auf Bekanntes, Unbestreitbares referieren und eine begründende oder kontrastierende Funktion haben:

Erst jetzt, als nichts mehr getan werden konnte, merkte ich, wie meine Hände zitterten. (B. Jaumann 2002)

Mit dem ebenfalls seltenen *wenn* wird – wiederum in einer Datenreihe mit nur gut 200 validen von insgesamt über 300 gefundenen Belegen – eine Vorzeitigkeit, Bedingung oder Begründung ausgedrückt.

Ein wahres Schnäppchen sei das Unternehmen doch jetzt, wenn es in Insolvenz gehe. (SZ, 29.09.2006)

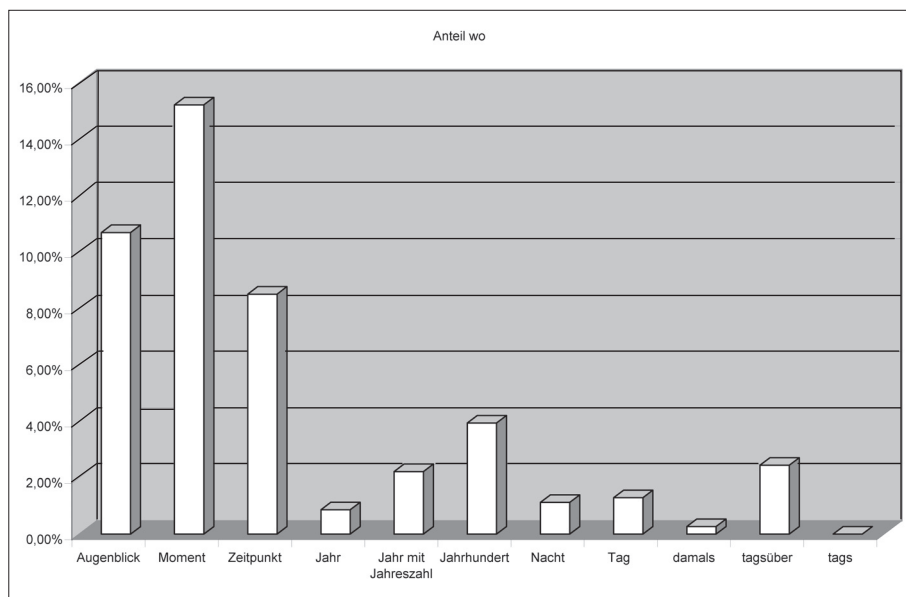
Was für *jetzt* gilt, trifft aber nicht auf andere Adverbien wie *tagsüber*, *tags*, *nachts* (präferierte Nebensatzeinleitung: *wenn*) oder *damals* (präferierte Nebensatzeinleitung: *als*) zu.

So sieht es also im Gehirn aus, nachts, wenn wir schlafen, oder tags, wenn wir träumen. (SZ, 28.09.1998)

Das war damals, als noch der Guckkasten angesagt war. (taz, 17.06.2004)

Die Merkmale, die den ersten Vermutungen nach zur Verwendung von *wo* führen könnten, treten keineswegs ausschließlich in *wo*-Sätzen auf. Erst der Vergleich mit lexikalisch unterschiedlichen Bezugsausdrücken liefert die erhofften deutlichen und relativ interpretationsunabhängigen Zahlenunterschiede, die kein Vergleich grammatischer Merkmale erbringen konnte: Nach *Augenblick* und *Moment* ist *wo* häufiger – hier gewinnt es seinen Konkurrenten etwa 15% 'Marktanteil' ab –, wenn auch lange nicht so gebräuchlich wie nach *jetzt*. Das in der Dudengrammatik gewählte Beispiel *Zeitpunkt* kann wahrscheinlich ebenfalls als akzeptabel für ein nachfolgendes *wo* bestätigt werden, auch wenn der Anteil an *wo*-Sätzen gering im Vergleich zur Datenrei-

he *Moment, wo* erscheint und der Nebensatz in 92% der Fälle anders eingeleitet wurde. In allen anderen Datenreihen ist *wo* im Vergleich zu den Alternativen *als, wenn, in der / dem* usw. ausgesprochen selten:



Greift man nun *Augenblick* und *Moment* als typische Kerne des Bezugsausdrucks zu einem *wo*-Relativsatz heraus, so müssten sich regionale Unterschiede in der Verwendung von *wo* wohl recht gut innerhalb dieser vergleichsweise *wo*-affinen, ansonsten unmarkierten Umgebung nachweisen lassen. Aber abgesehen davon, dass die Prozentsätze zwischen *Augenblick* und *Moment* überhaupt variieren (hier gibt es möglicherweise regionale oder diastratische Unterschiede), liefert dieser Vergleich wiederum keine bahnbrechenden Ergebnisse. Die erwartbare süddeutsch-österreichisch-schweizerische Vorliebe für temporales *wo* lässt sich aus den Daten (Gesamtkorpus, Stand Herbst 2009, knapp 1,8 Mrd. Tokens, und Probe mit ausgewogenerem Teilkorpus) nicht ablesen. Eine künftige Umfrage mit leicht modifizierten Beispielen im *Atlas zur deutschen Alltagssprache* kann hier möglicherweise aussagekräftigere Ergebnisse liefern.

Schließlich kann die Verwendung von *wo* (oder einer Subjunktion) der stilistischen Variation geschuldet sein. Einzelne Schreiber, Lektorate oder Redaktionen könnten die Formulierung *in dem Moment, in dem* wegen der Wiederholung von *in dem* als ungenk empfinden (nur 734 Vorkommen im Dudenkorpus gegenüber immerhin 541 Vorkommen von *in dem Moment, wo* – bei insgesamt

2 571 Belegen für *Moment, in* und nur 974 für *Moment, wo*). Ein weiterer Grund für stilistische Variation ist eine ähnliche Formulierung im Kontext, der für solche Aspekte der Untersuchung eigentlich über den einzelnen Satz hinaus ausgewertet werden müsste:

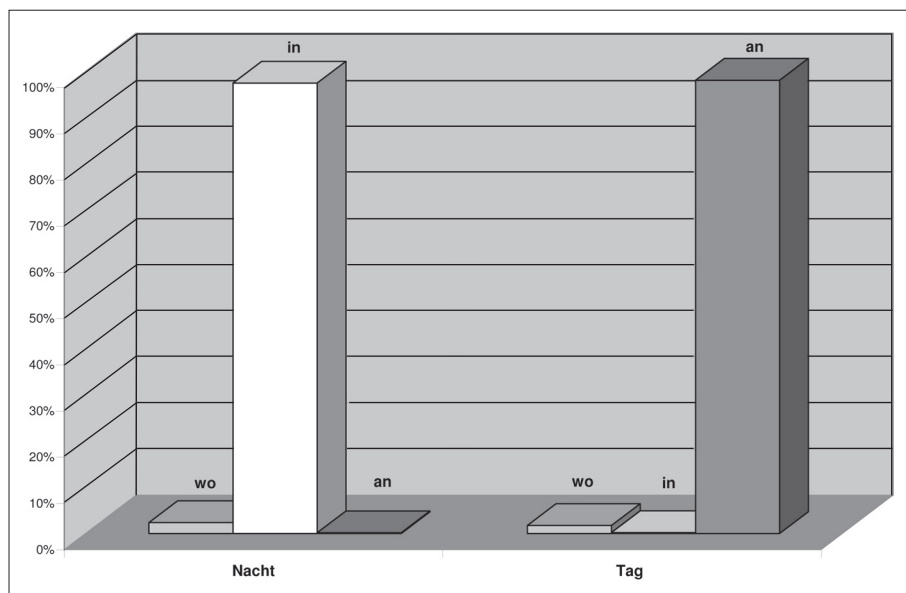
Über das Gefühl, wenn der Boden unter den Füßen wegbriecht, über den Moment, wo man sich entscheiden muss. (Zeit, 31.07.2002)

6. Diskussion

Diese Ergebnisse beweisen keine der Eingangshypothesen zur Semantik. Sie stehen aber auch nicht im Widerspruch dazu: Schließlich können Eigenschaften wie „auch konditional zu verstehen“ oder „nicht auf einen längeren Zeitraum in seinem Verlauf bezogen“ den Wortfolgen *jetzt / Moment / Augenblick, wo* meist leichter zugeschrieben werden als den Wortfolgen *Jahr / Jahrhundert, als / in dem*.

Insgesamt scheinen *wo* und *da* in den beschriebenen Kontexten die einzigen Nebensatzeinleitungen zu sein, die niemals ganz agrammatisch sind. Damit würde *wo* sich in beinahe überall einsetzbare Relativa in anderen Sprachen einreihen, vgl. neugriech. *pou* („wo“, anders als seine Konkurrenten *to opoio / i opoia* etc. ohne Bestimmung nach Kasus, Genus und Numerus und auch nicht, wie in *apo / stin / me / gia tin opoia*, direkt mit einer Präposition kombinierbar), oder es könnte mit temporalen/konditionalen Nebensatzeinleitungen wie frz. *au moment où* verglichen werden, die ebenfalls als Ausweitungen eines ursprünglich lokalen Gebrauchs gelesen werden können. Verfasser einer reinen Wort- oder Kollokationengrammatik würden unter den Stichwörtern *Augenblick* und *Moment* die verschiedenen Varianten des relativen und/oder temporalen (konditionalen) Anschlusses beschreiben. Die Nebensatzeinleitung mit *wo* dürfte man als standardsprachlich eher selten, aber nicht als falsch bezeichnen, und man müsste die ‚Arbeitsteilung‘ zwischen *wenn* und *als* erklären. Schließlich wäre auf die Wahl der Präposition (meist *in*, aber auch *an* ist möglich) und auf die Vermeidung von Wortwiederholungen einzugehen.

Unter den Stichwörtern *Tag* und *Nacht* wäre über die *als-oder-wenn-Regel* hinaus vor allem für Zweitsprachler der Hinweis von Nutzen, dass auch der präpositionale Anschluss nicht frei gewählt wird: Es heißt *der Tag, an dem*, aber *die Nacht, in der*. Die Nebensatzeinleitung mit *wo* hingegen ist so selten (< 2%), dass man sie sogar verschweigen könnte.



Ähnlich wären häufig nachgefragte Stichwörter wie *Jahr* und *Jahrhundert* zu behandeln. Bei *jetzt* und *heute* könnte man umgekehrt zwischen *wo* und *da* wählen lassen und die anderen Möglichkeiten als Ausnahmen darstellen oder gar vereinfachend unterschlagen. Für eine Grammatik, die über die Wortgrammatik und über beliebte Kollokationen hinaus allgemeine Regeln aufstellen soll, ist nun eine Formulierung zwischen zwei Extremen gesucht, die sich beide ausgehend von objektiven Zahlen begründen lassen:

Extrem Nr. 1: Temporales *wo* ist nie ganz falsch, es ist in allen Datenreihen zu finden. Darüber hinaus hat es sogar den Charme, einzelsprachübergreifend mit der Bedeutungsverlagerung von lokal über temporal zu konditional, kausal und konzessiv 'im Trend zu liegen'. Deshalb soll es in der Dudengrammatik uneingeschränkt empfohlen und seinen Konkurrenten gegenüber positiv herausgehoben werden.

Extrem Nr. 2: In den meisten Datensätzen ist temporales *wo* weitaus seltener als seine Konkurrenten und wirkt daher auch mehr oder weniger stark markiert. Deshalb ist die Dudenredaktion nicht bereit, es als standardsprachlich anzuerkennen.

Zwischen den beiden Extremen liegt der Status quo. Der Autor des Syntaxkapitels, Peter Gallmann, und die Dudenredaktion (Kathrin Kunkel-Razum, Franziska Münzberg) haben sich anhand der hier vorgestellten Ergebnisse dagegen entschieden, Grundlegendes an der Darstellung in der Dudengramma-

tik zu ändern. Würde man vor temporalem *wo* in bestimmten Kontexten warnen – denn auf eine Warnung laufen deskriptiv gemeinte Attribuierungen wie „selten“, „besonders in der gesprochenen Sprache“, „umgangssprachlich“ in der Praxis der Benutzer(innen) meist hinaus –, dann käme es womöglich vermehrt zu weitaus auffälligeren Formulierungen und Hyperkorrekturen wie *das Jahrhundert, an dem* oder *in dem Moment, zu dem*. Und solche auffälligen Formulierungen wären dem individuellen Ziel der Schreibenden, in einer Grammatik nachzuschlagen und sich daraufhin sprachlich möglichst konform und standardnah verhalten zu können, entgegengesetzt. Was im Dudenband 4 noch fehlt, ist eine Bemerkung zu den außergewöhnlichen Zahlenverhältnissen bei *jetzt, da* und *jetzt, wo*. In den Kontext von Randnummer 1659 passt *jetzt* als Adverb nicht. Nun bleibt zu diskutieren, ob das Thema besser nur im Dudenband 9 (*Richtiges und gutes Deutsch*) abgehandelt wird oder ob (und *wo*, zulasten welches anderen Themas) es in der nächsten Auflage der Dudengrammatik seinen Platz finden soll. Der je nach Kontext und vielleicht wiederum je nach Bezugsausdruck schwankenden Beurteilung von *da* als Relativsatzeinleitung trägt Damaris Nübling in ihrem Kapitel über die Nichtflektierbaren Rechnung (in Randnummer 943 mit einem als „gehoben veraltend“ markierten und einem unmarkierten Beispiel). In Randnummer 943 könnte man auch *wo* einordnen – was die Lehrkräfte aber kaum besänftigen dürfte.

Literatur

- Davies, Winifred V. (2007): Die Geschichte vom „schlechten“ Deutsch. In: Der Deutschunterricht 3: 52-62.
- Duden (1998): Duden – Die Grammatik. (= Duden 4). 6. Aufl. Mannheim u.a.: Dudenverlag.
- Duden (2009): Duden – Die Grammatik. (= Duden 4). 8. Aufl. Mannheim u.a.: Dudenverlag.
- Duden (2007): Duden – Richtiges und gutes Deutsch. (= Duden 9). 6. Aufl. Mannheim u.a.: Dudenverlag.
- Eisenberg, Peter (2006): Grundriss der deutschen Grammatik. 2 Bde. 3. Aufl. Stuttgart: Metzler.
- Elspaß, Stephan / Möller, Robert: Atlas zur deutschen Alltagssprache (AdA). Internet: <http://www.philhist.uni-augsburg.de/lehrstuehle/germanistik/sprachwissenschaft/ada/> (Stand: 09 / 2009).

- Fiehler, Reinhard/Wagener, Peter: Datenbank gesprochenes Deutsch. Internet: <http://dsav-wiss.ids-mannheim.de/DSAv/DSAVINFO.HTM> (Stand: 01 / 2010).
- Fleischer, Jürg (2004): A typology of relative clauses in German dialects. In: Bernd Kortmann (Hg.) *Dialectology meets typology. Dialect grammar from a cross-linguistic perspective*. Berlin / New York: de Gruyter, 211-243.
- Günthner, Susanne (2002): Zum kausalen und konzessiven Gebrauch des Konnektors *wo* im gesprochenen Umgangsdeutsch. In: *Zeitschrift für Germanistische Linguistik* 30, 3: 310-341.
- Pittner, Karin (2004): *Wo* in Relativsätzen – eine korpusbasierte Untersuchung. In: *Zeitschrift für germanistische Linguistik* 32, 3: 357-375.
- Zifonun, Gisela/Hoffmann, Ludger/Strecker, Bruno et al. (1997): *Grammatik der deutschen Sprache*. 3 Bde. (= Schriften des Instituts für Deutsche Sprache 7). Berlin / New York: de Gruyter.