

Auf der Suche nach historischen* Phrasemen – oder: Wörterbücher als Korpora

Marcel Dräger (Basel)

Abstract

The treatment of phrasemes is problematic both in a lexicographic view and a corpus analytic view, even though the efforts of the collocation research provide enormous amounts of output. Finding and automatically extracting phrasemes, that are longer than two or three word collocations, is still not satisfactorily possible, because there are no markers in the text, that indicate the borders of the expressions. If in addition the perspective is an historical one, the problems arise as the amount of available historical texts decreases. For an historical phraseographic analysis realistic examples are extremely help- and needful, so that other sources for assembling examples have to be found. The following text proposes some approaches, how dictionaries can be used as corpora, and shows, which results (examples, illustrations, links etc.) can be expected. The point of origin of the idea is to use dictionaries not in the classical way – that means: searching the headword and reading the related article –, but to use the whole dictionary as a corpus. This inverse use of dictionaries allows us to find a lot of implicit lexicographic information, that in a normal way of using dictionaries could only be found by coincident.

1 Einleitung

Lexikographische Projekte im weitesten Sinne haben einen großen Anteil an empirischer Korpusarbeit – oder anders formuliert: Die meisten Wörterbücher entstehen unter Zuhilfenahme eines Korpus (Krishnamurthy 2006: 250). Diese Korpora entsprechen jedoch nicht immer einer engeren korpuslinguistischen Definition, wie sie etwa Hunston (2006: 234) formuliert:

A corpus is an electronically stored collection of samples of naturally occurring language. Most modern corpora are at least 1 million words in size and consist either of complete texts or of large extracts from long texts.

Weder elektronisch noch besonders ausführlich im Belegtextschnitt sind beispielsweise die Belegzettel des Deutschen Wörterbuchs (Grimm/Grimm 1854–1960/1999 (im Folgenden *DWB*)), doch in ihrem Umfang – geschätzte 6.4 Millionen an der Zahl¹ – entsprechen sie durchaus den Anforderungen an ein modernes Korpus. Die lexikographische Arbeit mit einem solchen Zettelkastenkorpus ist ungemein zeitintensiv (das Erstellen der Zettel nicht eingerechnet) und das Ergebnis ist direkt von der Qualität der exzerpierten Belegtexte abhängig. Die enormen Vorteile heutiger lexikographisch nutzbarer Korpora (z. B. des

* Vorab zu vermerken ist, dass sich im Folgenden *historisch* vorwiegend auf das ältere Neuhochdeutsche bezieht – wobei eine Übertragung der vorgestellten Herangehensweise auf das Mittelhochdeutsche durchaus möglich scheint, aber vom Autor bisher nicht ausgiebig getestet wurde.

¹ <http://150-grimm.bbaw.de/>. Die Digitalisierung der Belegzettel ist laut Angaben auf der Website geplant.

DWDS-Korpus²) und der (teil-)automatischen Analyseinstrumente gegenüber der manuellen Zettelkastenkorpora liegen auf der Hand – die Vorzüge der alten Methode (Texte wurden bei der Belegsuche gelesen, Belege wurden dadurch schon vorsortiert und Varianten ließen sich leichter finden) aber gerieten im euphorischen Strudel der schnellen Verfügbarkeit von Belegen aus digitalisierten Texten in den Hintergrund. So hat sich die Lexikographie entsprechend der technischen Möglichkeiten und der überzeugenden Geschwindigkeit bei der Datengewinnung von einer "corpus-based-lexicography" zu einer "corpus-driven-lexicography" (Ooi 1998: 47f.) entwickelt. Korpusdaten haben nun nicht mehr nur die Funktion, den lexikographischen Artikel zu belegen, Steyer (2004: 93) spricht hier vom nach wie vor dominierenden "traditionellen [...] Konsultationsparadigma", sondern werden konstitutiv in die Bedeutungsfindung einbezogen ("Analyseparadigma").

2 Exkurs: Terminologie des Belegs

Egal ob Analyseparadigma oder Konsultationsparadigma, der Beleg spielt in beiden Fällen eine zentrale Rolle. Das allgemeine Verständnis von *Beleg* birgt jedoch einige Undifferenziertheiten in sich, weshalb es als Basis für die folgenden Überlegungen ungenügend ist. So muss der Begriff *Beleg* vor allem hinsichtlich der Betrachterperspektive und des Umfangs genauer gefasst werden. Haß (1991: 272) unterscheidet "zwischen Belegen mit Datenbasisstatus und Belegen mit Beispiel- und Nachweisstatus", wobei erstere reine "Textsegmente" als Ergebnis einer Korpusanalyse seien und letztere "das Ergebnis von Verfahren der Wortartikelproduktion, denen eine Teilmenge der Textsegmente nach bestimmten adressatenbezogenen Kriterien unterzogen wurde". In Anlehnung an Reichmann spricht Haß im zweiten Fall vom "Beispielbeleg" verstanden als "Ausschnitt[e] aus originalen Texten der Sprache oder der Sprachvarietät, die lexikographisch zu beschreiben ist." (Reichmann 1988: 413). Für das folgende Vorhaben – und ich meine auch für die "Theorie des lexikographischen Beispiels" (Hermanns 1988, Harras 1989) – scheinen einige Ergänzungen angebracht. Der Weg eines Ausdrucks von einer Parole-Äußerung hin zu seiner lexikographischen Kodifizierung lässt sich in drei Stadien einteilen, welche auch begrifflich differenzierbar sein sollten.

Stadium 1: Ausdrücke kommen in der schriftlichen und mündlichen Sprachverwendung (parole) einfach vor, ohne dass wir retrospektiv von der Existenz jeder einzelnen Verwendung Kenntnis haben (können). Im Nachhinein lässt sich die tatsächliche Existenz dieser Parole-Äußerungen mit korpuslinguistischen Methoden nur anhand eines kleinen Ausschnitts überlieferter und gespeicherter Äußerungen prüfen, welche überwiegend in schriftlicher Form vorliegen (Briefe, literarische Texte, Fachtexte, Zeitungen, Tonbandaufnahmen, Filmaufnahmen etc.). Ein Ausdruck wird also wesentlich häufiger realisiert, als er korpuslinguistisch erfassbar ist. Ich schlage daher vor, dieses präkorpuslinguistische Stadium als *Realisierung* eines Ausdrucks zu bezeichnen.

Stadium 2: Korpusanalytisch lassen sich nun die im Korpus erfassten Realisierungen eines Ausdrucks ermitteln, wobei diese schon einem gewissen Filtereffekt abhängig von der Korpuszusammensetzung und der verwendeten Analyseverfahren unterzogen sind. Mit Haß könnte man von *Textsegmenten* sprechen, wobei die genaue Ausdehnung eines solchen Segments undefiniert bliebe. Häufig wird das Textsegment durch die key-words-in-context (KWIC) festgelegt oder durch das nächste Satzschlusszeichen. Hingegen vernachlässigt bleibt die äußere Grenze des eigentlich relevanten Ausdrucks, welche beispielsweise bei Phrasemen nicht immer eindeutig ist³. Überlegungen zum Komponentenbestand eines Ausdrucks sind auch deshalb sinnvoll, da nur der Ausdruck selbst als Beleg im engeren Sinne des Wortes

² www.dwds.de/.

³ In einem Satz *Der Vater las dem Sohn heftig die Leviten* wäre zu entscheiden, ob *heftig* Bestandteil des Phrasems ist oder nicht.

betrachtet werden kann, wenn man – wie es im Folgenden geschehen soll – *Beleg* als 'Nachweis über die Existenz eines Ausdrucks' versteht. Nur auf der Basis dieses Verständnisses von *Beleg* ist eine Aussage wie *Das Lexem X ist im Korpus Y 100 mal belegt* zu rechtfertigen. Gemeint sind hier keine einhundert Textsegmente sondern nur die tatsächlichen im entsprechenden Korpus auffindbaren Realisierungen dieses Wortes. Entsprechend lässt sich das *Textsegment* in *Beleg* und *Belegkontext* unterteilen.

Stadium 3: Im dritten Stadium schreitet der Lexikograph ein, indem er die geeignetesten Belege für sein Wörterbuch auswählt und diese samt einem gewissen ihm sinnvoll erscheinenden Kontext exzerpiert. Es bietet sich an, diese einer subjektiven Auswahl unterzogenen und eventuell auch durch das Übertragen und Einpassen in die Artikelstruktur verfälschten Textpassagen mit Wiegand (1994: 242f.) *Belegtexte* zu nennen, was ihren eigenständigen, von der Quelle losgelösten Charakter unterstreicht. Diese Belegtexte werden durch den *Belegtextschnitt* in ihrer Länge begrenzt. Der zum Lemma in Bezug stehende Ausdruck innerhalb des Belegtextes soll – analog zu Stadium 2 – *Beleg* heißen, da eine weitere terminologische Unterscheidung auf dieser Ebene nicht zwingend notwendig erscheint.

3 Korpusanalyse und Phraseologie

Im Folgenden durchstreife ich die lexikographische Korpusrecherche in Bezug auf die in einigen Aspekten als Spezialfall zu behandelnde Phraseographie, denn gerade hinsichtlich der Auffindbarkeit von Phrasemen ist die Entwicklung der Korpusanalyseinstrumente ambivalent zu betrachten. Mit Rothkegel (2007: 1027) lassen sich die Bemühungen der phraseologischen Korpuslinguistik folgendermaßen umreißen: "Die Basisfrage ist, ob und aufgrund welcher Bedingungen eine Wortkette – sei sie kontinuierlich oder diskontinuierlich im laufenden Text angeordnet – eine lexikalisch-semantische Einheit darstellt und durch welche Art Verfahren dies festgestellt werden kann."⁴ Die derzeit funktionalste und erfolgversprechendste Methode zur Auffindung von Mehrworteinheiten ist die Kollokationsanalyse (oder Kookkurrenzanalyse), welche in kürzester Zeit riesige Mengen mittelbar und unmittelbar benachbarter Lexeme liefert und damit auch wichtige Hinweise auf stärker phraseologisierte Wortverbindungen. Heid (2007: 1042) sieht aber gerade bei den längeren und vor allem variantenreichen Mehrworteinheiten noch Entwicklungsbedarf und geht davon aus, dass sich auf automatisiertem Wege lediglich einige idiomatische Verbindungen finden ließen (Heid 2007: 1042). Auch sollte man sich angesichts des in kürzester Zeit zu gewinnenden Datenmaterials einmal vergegenwärtigen, was eine Formulierung, wie sie bei Rothkegel (2007: 1028) zu lesen ist, hinsichtlich des Zeitaufwandes bedeutet: "Die Feststellung, ob es sich bei einer Kombination um ein Phrasem handelt, bleibt in diesen Fällen der Interpretation durch einen Phrasem-Experten überlassen". Der Phrasem-Experte muss also über den semantischen Status entscheiden, um die Idiome von den Kollokationen zu separieren, er muss zudem die Grenzen bzw. den Komponentenbestand des Phrasems bestimmen und er muss die Varianten eines Phrasems zusammenführen⁵. All das lässt sich nicht einfach aus der Expertenintuition heraus beurteilen, sondern bringt einen nicht zu unterschätzenden zusätzlichen Rechercheaufwand mit sich⁶. Phraseme und in besonderem Maße Idiome besitzen aufgrund ihrer Motivierungsgeschichte und ihrer kulturellen Einbettung eine historische Dimension, welche

⁴ Die infrage kommenden und aktuell entwickelten Verfahren, die alle mehr oder weniger das Ziel haben, Phraseme halb-automatisch oder automatisch aus Texten zu extrahieren, werden in Rothkegel (2007) und Heid (2007) überblickend vorgestellt.

⁵ Auch Steyer (2004: 90) hält trotz ihres Plädoyers für eine unvoreingenommene Akzeptanz computerlinguistisch durchgeführter Analysen "die deutende und interpretierende Hand des Linguisten [...] für viele Zwecke letztlich immer unabdingbar."

⁶ Die Hauptschwierigkeiten liegen dabei in der fehlenden alphabetischen Ordnung älterer Sammlungen und dem nicht vorhandenen Wissen darüber, welche Variante aufgeführt sein könnte.

von keinem Experten ad hoc beurteilt und ausgewertet werden kann. Bezüglich historischer Sprachstufen besteht ein Kompetenzproblem (Burger/Linke 1998: 743, Filatkina 2007: 225), das mit allen zur Verfügung stehenden Hilfsmitteln (zeitgenössische Texte, Nachschlagewerke etc.), so gut es die Überlieferungslage erlaubt, kompensiert werden muss. Je weiter man in die Sprachgeschichte zurückgeht, desto spärlicher werden die zur Verfügung stehenden Quellen, doch auch schon für die Aufarbeitung der Phraseologie der neuhochdeutschen Sprachgeschichte fehlt heute das zeitgenössische Wissen (Dräger 2008). Eine automatische Extraktion phraseologischer Materials per Knopfdruck scheint also – zumindest für Phraseme mit mehr als drei Komponenten – kaum realistisch. Für den historischen Bereich erschwerend hinzukommt, dass aussagekräftige Ergebnisse mittels der Kollokationsanalyse nur bei ausreichender Datenlage möglich sind, eine Voraussetzung, die schon für das 19. Jahrhundert nicht mehr uneingeschränkt gegeben ist, da mündliche Zeugnisse kaum vorliegen. Mit der auf Häufigkeiten basierende Kollokationsanalyse lasse sich das Normale und Usuelle finden (Stubbs 2001: 211), das Idiosynkratische, das gerade den varianten Phrasemen eigen ist, bleibt statistisch irrelevant. Evert/Krenn (2001: 192) gehen davon aus, dass statistische Aussagen bei weniger als fünf Belegen dem Zufall gleichkommen – gerade im historisch-phraseologischen Bereich liegen aber oft nur ein oder zwei Belege einer Variante vor. Alles in allem verschlechtert sich das Kosten-Nutzen-Verhältnis also deutlich, je stärker man sich auf idiomatische Phraseme – und das sind gerade jene, welche in einem semantisch ausgerichteten Nachschlagewerk zentral sind – und auf eine diachrone Perspektive fokussiert.

Eine erfolgversprechende (historisch-)phraseologische Korpusanalyse benötigt einen von der Expertenintuition unabhängigen Input, mit welchem große Korpora zielgerichtet – durchaus mit Hilfe der Kollokationsanalysemethoden – durchsucht werden können. Das phraseographische Ziel dabei ist stets, aussagekräftige Belegtexte für ein Phrasem zu finden, um es a) zeitlich datieren und b) seine semantische Struktur anhand des Belegkontextes ermitteln, verifizieren und beschreiben zu können. Soweit ist das Verfahren mit jenem der (historischen) Einwortlexikographie identisch (Reichmann 1988, Harras 1989, Wiegand 1989), speziell in der Phraseographie jedoch kommt die Suche nach Varianten hinzu. Und hier entsteht das Problem: Ohne besagten Input wissen wir nicht, welche Phraseme und welche Varianten in der Vergangenheit überhaupt existierten – die Suche danach wäre also relativ wahllos.

Ein exemplarisches Beispiel veranschaulicht das Geschilderte: Die Redewendung *jemandem die Leviten lesen* kann aus heutiger Sicht als gebräuchlich gelten, lexikalische Varianten sind mutmaßlich keine gängigen bekannt. Die Kollokationssuche im DWDS-Kerncorpus liefert *Levit* als Kollokator zu *lesen* an 139. Stelle, weitere Kollokatoren, welche auf eine lexikalische Variante des besagten Phrasems hindeuten, waren beim Überfliegen der 2400 gelisteten Kollokationen nicht zu entdecken. Auf diesem Weg lassen sich also in einem überschaubaren zeitlichen Umfang kaum Varianten finden⁷. Im historischen *Deutschen Sprichwörter-Lexikon* (1867–1880) sind zu besagtem Phrasem als lexikalische Varianten für *Leviten* mindestens die folgenden aufgeführt: *Epistel*, *Text*, *Hund*, *Kapitel*, *Psalter*, *Vers(e)* und als Fehllesung dazu *Ferse*. Das Problem dieser mit Sicherheit nicht vollständigen Liste ist, dass sie keinerlei Aussage über die tatsächliche Gebräuchlichkeit dieser Varianten enthält – eine Rückprüfung anhand eines Textkorpus ist also unerlässlich. Angesichts der großen Ergiebigkeit bei der Konsultation eines Sprichwörterlexikons scheint dieser Weg der Korpusanalyse deutlich überlegen, wäre da nicht noch ein Haken. Rein durch das Nachschlagen in einem Lexikon, beispielsweise dem *Idiomatik-Duden* (2008) findet man nicht wesentlich mehr als das, was man schon vor der Suche wusste, denn ein weiterführender Verweis vom Stichwort *Leviten*

⁷ Einschränkung ist zu erwähnen, dass das DWDS-Kerncorpus nur Texte des 20. Jahrhunderts enthält und die meisten Varianten des Phrasems sehr wahrscheinlich älter sind.

auf *Text* fehlt, sodass man die dort durchaus aufgeführte Variante *jemandem den Text lesen* nicht finden würde⁸. Auch im *Deutschen Sprichwörter-Lexikon* wäre man auf diese Weise nicht weit gekommen, denn die hohe Trefferquote hat eine andere Ursache: *Jemandem die Leviten lesen* – als gängigste lexikalische Variante des Phrasems – steht bei allen anderen Varianten als Erläuterung aufgeführt. Daher lassen sich durch die Volltextsuche mit *Levit** in der digitalisierten Ausgabe des *Deutschen Sprichwörter-Lexikons* alle Einträge finden, in welchen das Wort entweder in der Nennform des Phrasems, in einem Belegtext oder in einem Erklärungstext vorkommt. Auf diese Weise stößt man nicht nur auf lexikalische Varianten, sondern auch auf phraseologische Synonyme, wie *Ich hab' ihm den Kümmel gerieben* (*Sprichwörter-Lexikon*: II, 1704) oder *Ich hab' em de Leber geschleimt* (*Sprichwörter-Lexikon*: II, 1867). Genau betrachtet ist dieser Prozess aber nicht mehr das Nachschlagen in einem Lexikon – also der Einstieg über ein bestimmtes Lemma – sondern eine Korpusrecherche. Durch den digitalen Zugriff auf den Volltext wird das ursprünglich gedruckte Nachschlagewerk zu einem (elektronischen) Korpus. Ausgehend von dieser Feststellung will ich im Folgenden beispielhaft einige Konsequenzen und Möglichkeiten aufzeigen, welche sich speziell für die Phraseologie ergeben, wenn man digitalisierte Wörterbuchtexte korpuslinguistisch verwendet.

4 Suche nach bislang nicht phraseographisch erfassten Phrasemen bzw. Varianten

Eine ergiebige Möglichkeit der phraseologischen Belegsuche fernab jeglicher Korpora wäre, historische Texte zu lesen und jene Passagen, welche syntaktisch oder semantisch auffällig erscheinen, auf ihren phraseologischen Gehalt hin zu überprüfen.⁹ Auf diese Weise vermag ein geschultes Auge tatsächlich zahlreiche Phraseme zu entdecken, welche heute unbekannt oder nur noch in einer varianten Form gebräuchlich sind. Doch will man diese quasi manuelle Korpusrecherche in einem repräsentativen Umfang durchführen, dürfte der Zeitaufwand für die Schulung der Augen und die eigentliche Textrezeption dem oben beschriebenen Expertenverfahren in nichts nachstehen¹⁰. Zu Nutzen machen könnte man sich aber, dass das Lesen und Exzerpieren schon für zahlreiche historische Werke erledigt wurde, nämlich von den Zuarbeitern, welche den großen Wörterbuchprojekten die Belege und Textschnipsel lieferten, die in den Zettelkästen archiviert wurden und teilweise in die Wörterbücher als Belegtexte einfließen. Hier wurde ein bislang nicht systematisch ausgewertetes aber für den historisch orientierten Sprachwissenschaftler äußerst wertvolles Wissen konserviert, welches greifbar zu machen durchaus lohnend erscheint. Die Belegtexte in historischen Wörterbüchern sind Dokumente einer vergangenen Sprachsituation, welche vom Bearbeiter als exemplarisch eingeschätzt wurden und – im Gegensatz zum bloßen Text aus der gleichen Zeit – mit zeitgenössischen Kommentaren und Erklärungen versehen sind.

Doch eine korpuslinguistische Auswertung historischer Wörterbücher wird durch die darin fehlenden Konzepte von Phraseologie stark erschwert (Stantcheva 2003). Eine Suchanfrage "Gib alle Phraseme aus" würde daran scheitern, dass Phraseme in Wörterbüchern bis in die jüngste Zeit nicht (einheitlich) markiert werden (cf. Burger 1989, Korhonen 2004) und daher nicht automatisch adressierbar sind. Der Grund hierfür ist, dass die am gedruckten Buch orientierten Wörterbuchkonzepte derartige Fragestellungen des Benutzers bislang nicht vorsahen. Phraseme werden im Großen und Ganzen nur in sehr seltenen Fällen als Lemmata

⁸ Nicht so in umgekehrter Reihenfolge, aber das brächte für die Fragestellung keinen Gewinn.

⁹ Dieses Verfahren wurde erfolgreich zur Ermittlung des erklärungsbedürftigen Differenzwortschatzes für das "Klassikerwörterbuch" angewendet (cf. Brückner/Knoop 2003: 72).

¹⁰ Čermák (2006: 33) sieht im Lesen ebenfalls eine Möglichkeit, Phraseme zu erkennen, schließt aufgrund der heutigen Korpusgrößen aber jegliche manuelle Annäherung aus.

angesetzt¹¹ und haben daher keine feste mikrostrukturelle Position. Zudem ist ihre Ausdehnung nicht eindeutig abgegrenzt, denn während bei Einwortlexemen klar ist, wo das Wort aufhört – nämlich vor dem Leerzeichen –, ist das Ende eines Phrasems nicht schematisch zu ermitteln. Etwas besser sieht es mit dem Anfang der Phraseme aus, der häufig durch diakritische Zeichen (z. B. den Asterisken*), durch Abkürzungen (z. B. *R.A.* bei Adelung 1793–1801/1970, im Folgenden nur Adelung) oder auch durch sonstige metasprachliche Hinweise (z. B. *sprichwörtlich*) markiert ist. Allerdings bezieht sich eine einzige metasprachliche Markierung häufig auf mehrere hintereinander aufgelistete Phraseme, was eine (voll-)automatische Analyse wiederum verhindert. Eine stichprobenartig durchgeführte Suche im *Grammatisch-kritischen Wörterbuch der Hochdeutschen Mundart* (1793–1801) zeigte jedoch, dass sich der manuelle Aufwand der Separation der einzelnen Phraseme in Grenzen hält, der Gewinn einer solchen Auswertung hingegen vielversprechend ist. Die gängigsten, für eine erste Auswertung ausreichenden Marker für Phraseme lassen sich in den Wörterbüchern recht leicht finden. Weitere, bei der Suche aufgefundene Marker können dann in einem induktiven Verfahren in die Suche integriert werden. Im *Adelung* ist beispielsweise die Suche nach *R.A.* (910 Treffer) und *Redensart** (345 Treffer) erfolgreich. Ein Marker kennzeichnet in der Regel mehrere Phraseme, sodass die 1255 genannten Treffer auf ein Vielfaches an Sprichwörtern und Redewendungen verweisen. Der große Gewinn dieses Verfahrens liegt darin, dass sich so zahlreiche historische und heute unbekannte Phraseme finden lassen, die in einem historisch-phraseologischen Nachschlagewerk aufgrund ihrer Vorkommenshäufigkeit in historischen Texten unbedingt erklärt werden müssen. Beispielsweise findet sich im *Adelung* unter *Achse* markiert mit *R.A.* das Phrasem "etwas auf der Achse herzu führen, eine Waare auf der Achse kommen lassen, fortbringen" (Adelung: I, 147) mit der Bedeutung 'etwas auf dem Landweg transportieren'¹². In den heute gängigen phraseologischen Nachschlagewerken ist dieses Phrasem nicht aufgeführt, in Röhrichs *Lexikon der sprichwörtlichen Redensarten* ist lediglich *auf Achse* erklärt als "(dienstlich, geschäftlich) unterwegs sein; auch: immer beschäftigt, immer in Aktion, Bewegung sein. >Auf Achse< ist eine saloppe Umschreibung für die Fortbewegung mittels eines Eisenbahnzugs oder Autos; sie ist seit Ende des 19. Jahrhunderts belegt" (Röhrich 2002: I, 64).

Durch diese zum klassischen phraseographischen Arbeitsprozess¹³ inverse (analyseparadigmatische) Vorgehensweise lassen sich alleine in Adelungs Wörterbuch hunderte bislang nicht phraseographisch erfasste, aber historisch relevante Phraseme finden, und bestehende Phraseme können um eine Vielzahl an Varianten ergänzt werden. Grund für diesen großen Gewinn an heute unbekanntem oder unbeachteten Lemmata ist das Aufbrechen der phraseographischen Abschreibetradition¹⁴ denn die Stichwortlisten (und nicht selten auch die Erklärungen) neuer Lexika und Sammlungen speisen sich zu einem Großteil aus bestehenden Sammlungen und Lexika – allerdings eben nur aus den Teilen, welche direkt (also per Stichwortsuche) zugänglich sind. Die ergiebigen, aber nur indirekt auffindbaren Auflistungen und auch Erläuterungen von Redewendungen und Sprichwörtern, die nicht eigens lemmatisiert wurden, blieben daher weitestgehend unbeachtet. Vieles spricht deshalb dafür, für die Suche nach Phrasemen allgemeinsprachliche Wörterbücher und gerade nicht die vielleicht auf den ersten Blick geeigneter erscheinenden Sprichwörter-sammlungen bzw. -lexika zu verwenden, da dort mehr individuelles Wissen zu finden ist. Als weiterer Vorteil einer

¹¹ Das ist nur dann der Fall, wenn sie Unikalia oder im Wörterbuch nicht geführte Komponenten enthalten (z. B.: "fix und fertig").

¹² Das Phrasem ist auch in mehreren historischen Texten des 18. und 19. Jahrhunderts belegt und sollte daher in einem historisch-phraseologischen Nachschlagewerk des Neuhochdeutschen erklärt werden.

¹³ Nämlich jenem, der dem Konsultationsparadigma folgt, also Belege zur Untermauerung und Beschreibung der lexikographischen Erklärung verwendet.

¹⁴ Nachgewiesen wurde diese im Raum stehende Vermutung bislang nicht, aber beispielsweise Wiegands (1994: 264f.) eher amüsierende als ernsthafte Suche nach Karl-Kraus-Zitaten kommt genau zu diesem Ergebnis.

Auswertung von Wörterbüchern lässt sich folgender nennen: In den rein phraseologischen Nachschlagewerken wurde häufig eine Überfülle an nicht am Sprachgebrauch rückgeprüften Varianten aufgenommen (und dies zudem ohne Bezug auf mögliche Quellen), oder Nennformen aus älteren Sammlungen wurden unreflektiert übernommen, sodass die tatsächliche Existenz eines Phrasems oder einer Variante nicht aus ihrer phraseographischen Kodifizierung geschlossen werden kann. In den Wörterbüchern hingegen – und speziell in klassischen Belegwörterbüchern wie dem *Deutschen Wörterbuch* nach Jacob und Wilhelm Grimm – gewährleisteten Belegtexte und sprachnah formulierte Beispiele einen engeren Bezug zur Sprachrealität¹⁵.

5 Suche nach sprachnahen Nennformen – die Belege

In der (historischen) Phraseologie spielen Realbelege eine wichtige Rolle, da nur durch sie erschließbar ist, in welchen Varianten und morphosyntaktischen Variationen ein Phrasem überhaupt existierte. Es ist daher notwendig, das tatsächliche Vorkommen der gefundenen Phraseme in großen historischen Korpora zu überprüfen, wobei ein Nichtvorhandensein noch nicht gegen die prinzipielle Existenz des Phrasems spricht. Dieser Schritt soll allerdings bis auf einen Spezialfall ausgeblendet bleiben: Belegwörterbücher listen im Allgemeinen zu den einzelnen Lemmata und oft auch zu den Phrasemen reelle Sprachbelege auf, die – und dafür sind sie ja auch per definitionem gedacht – das tatsächliche Vorkommen des Lemmas unter Beweis stellen. Während das *Grammatisch-kritische Wörterbuch der Hochdeutschen Mundart* aufgrund des lexikographischen Konzepts (cf. Haß-Zumkehr 2001) arm an Belegen ist – Adelung hat in der Regel auf selbst formulierte Beispielsätze zurückgegriffen –, liefert das *DWB* eine unermessliche Sammlung an Belegtexten. Dort würden sich korpusanalytische Verfahren schon für das Durchsuchen einzelner Artikel lohnen angesichts der Menge an Bedeutungen und vor allem an den Bedeutungen zugeordneten Belegtexten. Andererseits ist die Markierung von phraseologischen Einträgen nicht annähernd so einheitlich wie bei Adelung¹⁶. Die realen Belegtexte des *DWB* stammen überwiegend aus literarischen Werken, und sind ein Indiz für die prinzipielle Existenz eines Ausdrucks. Sein phraseologischer Status (Festigkeit bzw. Gebräuchlichkeit) lässt sich erst mit mehreren Belegen beweisen, die dann im Idealfall auch gleich eine Auflistung real existierender Nennformen sind. Hier kommt dem Phraseographen Sammelleidenschaft und die Weitschweifigkeit der Autoren des *DWB* im Umgang mit Belegtexten zu Gute. Ein verhältnismäßig knappes Beispiel zu *einem das Herz abfressen*, was mit dem heute noch geläufigen *jemandem das Herz brechen* verwandt sein dürfte, illustriert das Gesagte:

einem das herz (*vom leibe*) abfressen *von herznagendem kummer, wobei man wol ursprünglich an zauber dachte*: es will mir aber das herz abfressen. *GÖTTE* 57, 253; es hätte mir das herz abgefressen. 8, 56. *roher ist die redensart*, sich das herz abfressen: der henker danke dir, dasz ich mir deinetwegen das herze und das leben abfressen musz. *WEISE erz.* 10; ich fresse mir das herz ohne noth ab (*gräme mich ohne noth*). *WEISE kl. leute* 292; meinstu das herz müsse mit sorge und kummer abgefressen sein. *das*. 304. den leuten das ir fressen ab. *H. SACHS* 1, 453^c (*DWB*: I, 41).

Ausgehend von den belegten Phrasemen ließe sich nun nach weiteren phraseologischen Varianten suchen. Der teilweise recht knappe textuelle Kontext wäre durch einen Abgleich mit den oft nur in gedruckter Form vorliegenden Quellen zu erweitern. Aufgrund des Zeitaufwandes ist das allerdings nur bei vielversprechenden Belegen sinnvoll, also solchen, bei welchen der Kontext weitere Erkenntnisse über die Semantik erwarten lässt. Über die digitalisierte Version des *DWB* lassen sich auf diese Art nicht nur literarische Quellen

¹⁵ Auch hier sind natürlich Erfindungen des Autors nicht auszuschließen, wobei davon vor allem die Erläuterung der Phraseme betroffen ist, wie es das Beispiel "kurz angebunden" zeigt (Dräger 2008).

¹⁶ Der häufigste Marker dürfte "Redensart(en)" sein, der immerhin auf 609 Einträge kommt.

erschließen, sondern auch große Teile historischer Sprichwörtersammlungen werden zugänglich¹⁷. In diesen älteren Sammlungen wird man normalerweise nur mit sehr hohem Nachschlageaufwand fündig¹⁸, ein mit Quellenangaben versehener Verweis im *DWB* erleichtert die phraseographische Arbeit daher ungemein. Mit wenig Aufwand wären die im *DWB* aus historischen Sprichwörtersammlungen entnommenen Phraseme aufzulisten, indem man das Wörterbuch hinsichtlich der Namen der Verfasser solcher Sammlungen (z. B. *Simrock*, *Agricola* oder *Eiselein*) auswertet, da diese als Belegstellenangabe verwendet wurden. Das Quellenverzeichnis gibt in etwa Auskunft darüber, welche Werke in das Wörterbuch eingeflossen sind. Gleichzeitig fungieren diese Autorennamen bei einer normalen Suche als Sprungmarken zu den phraseologischen Informationen innerhalb der einzelnen manchmal mehrere bedruckte Seiten umfassenden Artikel. Der größte Vorteil einer Belegsuche per *DWB* ist also der korpuslinguistisch nutzbare Zugang zu (noch) nicht digitalisierten Werken (literarischen Texten wie auch Nachschlagewerken), in welchen sich im Zweifelsfall aufgrund der genauen Stellenangabe zielgerichtet ein Beleg in seinem größeren textuellen Kontext manuell nachschlagen lässt.

6 Die Verwendung des Wörterbuchs als Sprachkorpus

Die beiden bislang vorgestellten korpuslinguistischen Zugänge zu digitalisierten Wörterbüchern ähneln noch recht stark der klassischen Herangehensweise an ein Wörterbuch, wobei die Fragestellung nicht einer speziellen stichwortbasierten Suche, sondern einer möglichst allgemeinen Analyse zur Gewinnung größerer Datenmengen galt. Metalexikographische Marker werden dabei ausgenutzt und teilweise umfunktioniert um direkt an jene Stellen im Wörterbuchtext zu gelangen, wo sich phraseologische Informationen befinden. In einer dritten exemplarischen Herangehensweise soll nun das Wörterbuch im Ganzen als Textkorpus verstanden werden, welches per Volltextsuche zugänglich ist. Die Kollokationssuche scheidet dabei allerdings aus, da das Ergebnis durch den segmentartigen Aufbau eines Wörterbuchartikels (fehlende Satzgrenzen, Stichwortreihungen, häufige Verwendung von Abkürzungen etc.) starken Verfälschungen unterworfen sein dürfte. Bei der Belegsuche im gesamten Wörterbuchtext ergibt sich ein Unterschied zu dem im vorigen Kapitel vorgestellten Verfahren, da nicht nur die als Beleg markierten Passagen, sondern der gesamte Text (inklusive Erläuterungen, Kommentaren etc.) analysiert werden. Dabei haben die ökonomischen Gesetzen unterliegenden Wörterbücher gegenüber den klassischen Textkorpora den Nachteil, dass darin vorkommende Textpassagen häufig sehr kurz, komprimiert und stellenweise auch aus dem Zusammenhang gerissen sind, weshalb sie sich nicht immer eindeutig erschließen lassen. Doch sie haben auch zahlreiche Vorteile:

Wie schon erwähnt bietet ein digitalisiertes Wörterbuch einen Zugang zu Belegtexten, die manuell aus zeitgenössischen oder historischen Texten exzerpiert wurden. Diese Belege lassen sich – sofern die Texte selbst nicht bereits digitalisiert wurden – anderweitig nur durch Lesen der entsprechenden Werke finden. Gerade bei selten vorkommenden Phrasemen ist dieser Gewinn an Belegmaterial nicht zu unterschätzen. Natürlich lassen sich auch aus gedruckten Wörterbüchern Belege exzerpieren, dadurch erhöht sich der Zeitaufwand (bspw. für das Nachschlagen unter mehreren Komponenten) allerdings erheblich (cf. Dräger 2004). Bei allen Vorzügen der digitalen Belegrecherche in lexikographischen Werken sei an dieser Stelle auch ein Nachteil erwähnt. Die Belegtexte gerade aus älteren Werken haben nur eine bedingt wissenschaftlich akzeptable Verlässlichkeit hinsichtlich der Übereinstimmung mit der tatsächlichen Quelle. Flüchtige Fehler, ungenaues Exzerpieren und vor allem das Zitieren aus dem Gedächtnis wirkten hier verfälschend, weshalb bei der weiteren nicht rückgeprüften

¹⁷ Es dürften, wenn man nach den Trefferzahlen bei der Suche nach den Autorennamen geht, über dreiviertel der Lemmata der jeweiligen Sammlung im Grimm'schen Wörterbuch enthalten sein.

¹⁸ Siehe dazu Fußnote 7.

Verwendung dieser Belegtexte – was aus Zeitgründen wohl das einzig Praktikable ist – stets die vermittelnde Quelle mit angegeben werden sollte.

In gedruckten Wörterbüchern lassen sich Belege nur über die Stichwortsuche nach den einzelnen Komponenten des betroffenen Phrasems finden. Die Belegtexte aber bestehen in der Regel nicht nur aus dem zu belegenden Ausdruck, sondern haben auch mehr oder weniger zusätzlichen Kontext. All jene Realisierungen eines Ausdrucks, welche im Belegtext eines anderen Lemmas vorkommen, lassen sich somit nicht nachschlagen. Im klassischen lexikographischen Sinn sind das auch keine Belege, da sie nicht mit dem Zweck aufgenommen wurden, etwas zu belegen. Dennoch bilden sie eine interessante Basis für die Belegsuche, was am folgenden Beispiel aus dem *DWB* deutlich wird: Die Volltextsuche *lange Bank* liefert 17 Treffer, worunter auch dieser ist:

war der unglaublichste verzögerer und trödler, der je unter der sonne gelebt, und ihm allein habe ich es schuld gegeben, dasz meine sache so auf die lange bank geschoben wurde LESSING w. 18, 88 L.-M.:¹⁹

Der Belegtextschnitt würde aufgrund seiner Subjektlosigkeit den Kriterien für einen guten Belegschnitt (Wiegand 1994: 244f.) sicher nicht standhalten, der Belegtext selbst ist aber sehr aussagekräftig: Das Lemma *Trödler* wird durch den semantisch ähnlichen Ausdruck *Verzögerer* und das Phrasem *etw. auf die lange Bank schieben* gut erläutert. Umgekehrt wäre der Belegtext natürlich auch prädestiniert für das Phrasem, doch dort wird er nicht aufgeführt, was an der subjektiven Auswahl des Lexikographen liegen kann oder einfach daran, dass der Phrasembeleg nicht oder – das ist in diesem Fall das Wahrscheinlichste²⁰ – zur Entstehungszeit des Artikels noch nicht verzettelt wurde. Die Volltextsuche bietet also einen einfachen Zugriff auf viele zusätzliche Realisierungen von Phrasemen in Belegtexten anderer Lemmata mit diversen positiven Begleiterscheinungen. Zum Ersten haben die für ein Wörterbuch ausgewählten Belegtexte in der Regel eine gewisse semantische Qualität. Es ist also davon auszugehen, dass man insgesamt geeignetere Textsegmente findet als in einem großen Textkorpus, wo sich auch viele kaum verwertbare Ergebnisse finden. Zum Zweiten sind die Belege schon mit entsprechenden Quellenangaben versehen, auch wenn diese gerade in historischen Wörterbüchern teilweise einem (unlösbaren) Rätsel gleichen. Doch nicht nur Quellenangaben werden en passant mitgeliefert, auch Varianten, grammatische Angaben und zeitgenössische Textstellenkommentare sind gelegentlich vorzufinden. Drittens sind die Belege anders kontextualisiert als in einem Textkorpus, denn während dort der größere textimmanente Zusammenhang den Kontext liefert, konstituiert sich dieser im Wörterbuch zusätzlich über das Lemma, die Bedeutungserläuterung und mögliche benachbarte Belegtexte. Die Hinweise auf Bedeutungsnuancen, wie sie der Verfasser des Artikels vielleicht mit dem ausgewählten Belegtext illustrieren wollte, sind im Wörterbuch daher deutlicher.

Der oder die Autor(en) sind auch diejenigen, welche den dritten Teil an korpusanalytisch auswertbarem Textmaterial in einem Wörterbuch liefern. Gemeint sind die Erläuterungen und Kommentare, welche – im Gegensatz zu den Belegtexten – ein sprachliches Zeitzeugnis der Entstehungszeit des Wörterbuchs sind. In Dräger (2009) habe ich das anhand eines Beispiels aus dem *Röhrich* (2002: II, 513) beschrieben: Das Phrasem *einen (scharf) ins Gebet nehmen* wird dort erläutert als "ihn zur Rechenschaft ziehen, ihm ins Gewissen reden." – Handlungen, die eher auf einen gesitteten Umgang schließen lassen. *Einem tüchtig einheizen* erklärt

¹⁹ Ein Subjekt ist in diesem Belegtext nicht enthalten. Die Sigle "Lessing w. 18, 88 L.-M." ist zu lesen als 'Sämtliche Schriften von Gotthold Ephraim Lessing, hrsg. von Karl Lachmann, 3., auf's Neue durchgesehene und vermehrte Auflage, besorgt durch Franz Muncker, Göschen, Stuttgart 1886–1924; hier Band 18, Seite 88'.

²⁰ Wahrscheinlich ist das deshalb, weil die Werkausgabe von Lachmann/Muncker erst ab 1886 erschienen ist, der Buchstabe *B* aber von den beiden Brüdern selbst bearbeitet wurde.

Röhrich (2002: II, 367) mit "einem scharf zusetzen, ihm durch Drohungen Angst einjagen, ihn hart ins Gebet nehmen". Er verwendet *jemanden ins Gebet nehmen* in seiner Erklärungssprache, also seinem aktualisierten Sprachgebrauch, anders als er das Phrasem in seinem Lexikon selbst erläutert. Zudem führt er noch eine zuvor nicht genannte Variante mit dem fakultativen Adverbial *hart* auf. Diese zweite Ebene, die den Sprachgebrauch der Autoren abbildet, steht in einem historisch ausgerichteten Nachschlagewerk in (deutlicher) zeitlicher Distanz zur erklärten Sprache. Damit werden – setzt man wie im geschilderten Beispiel erklärte Sprache und erklärende Sprache in Beziehung – innerhalb eines Wörterbuchs diachrone Wandelprozesse offensichtlich, die für die diachrone Beschreibung von Phrasemen von großer Bedeutung sind.

Alles in Allem spricht Einiges dafür, das (historische) Wörterbuch aus seiner Existenz als Nachschlagewerk zu befreien und es korpuslinguistisch zu benutzen und auszuwerten. Weiterführende Überlegungen müssten sich nun damit befassen, wie sich der noch sehr hohe manuellen Aufwand weiter reduzieren ließe. Auch in der – nach Möglichkeit automatisierten – Verknüpfung unterschiedlicher Recherchestrategien beispielsweise kombiniert mit induktiven Verfahren steckt bislang ungenutztes Potential. Die vorgestellte Methodik zur Unterstützung phraseographischer Arbeit hat sich bereits in der Praxis bewährt, und das vor allem bei historischen Fragestellungen, wo andere Recherchemöglichkeiten jenseits des Zufalls rar sind.

Literatur

- Adelung, Johann Christoph (ed.) (1793–1801/1970): *Grammatisch-kritisches Wörterbuch der hochdeutschen Mundart. Mit beständiger Vergleichung der übrigen Mundarten, besonders aber der Oberdeutschen*. Nachdruck hrsg. v. Helmut Henne. Hildesheim/New York.
- Brückner, Dominik/Knoop, Ulrich (2003): "Das Klassikerwörterbuch. Begründung und Erläuterung eines digitalen Wörterbuchprojekts zum differenten Wortschatz in der klassischen Literatur". *Zeitschrift für germanistische Linguistik* 31: 62–86.
- Burger, Harald (1989): "Phraseologismen im allgemeinen einsprachigen Wörterbuch". In: Hausmann, Franz Josef et al. (ed.): *Wörterbücher. Dictionaries. Dictionnaires. Ein internationales Handbuch zur Lexikographie*. Bd. 1. Berlin/New York, de Gruyter: 593–599.
- Čermák, František (2006): "Statistical Methods for Searching Idioms in Text Corpora". In: Burger, Harald/Häcki Buhofer, Annelies (ed.): *Phraseology in Motion. Methoden und Kritik*. Bd. 1. Baltmannsweiler, Schneider Verlag Hohengehren: 33–42.
- Dräger, Marcel (2008): "Kurz angebunden. Historisch-lexikographische Betrachtungen einer Redewendung". Erscheint in: Földes, Csaba (ed.) (im Druck): *Phraseologie disziplinär und interdisziplinär*. Tübingen, Gunter Narr.
- Dräger, Marcel (2009): "Phraseologische Nachschlagewerke im Fokus". Erscheint in: Korhonen, Jarmo (ed.): *Phraseologie. Global – areal – regional*. Baltmannsweiler, Schneider Verlag Hohengehren.
- Dudenredaktion (ed.) (2008): *Redewendungen. Wörterbuch der deutschen Idiomatik*. 3., neu bearb. und aktual. Aufl. Mannheim etc.: Dudenverlag. (= *Der Duden in zwölf Bänden* 11).
- Evert, Stefan/Krenn, Brigitte (2001): "Methods for the Qualitative Evaluation of Lexical Association Measures". *Proceedings of the 39th Annual Meeting of the Association for Computational Linguistics*. Toulouse: 188–195.
- Filatkina, Natalia (2007): "Formelhafte Sprache und Traditionen des Formulierens (HiFoS). Vorstellung eines Projekts zur historischen formelhaften Sprache". *Sprachwissenschaft* 32/2: 217–242.
- Grimm, Jacob/Grimm, Wilhelm (eds.) (1854–1960/1999): *Deutsches Wörterbuch*. 16 Bde. Leipzig 1854–1960: Hirzel. Fotomechanischer Nachdruck. München 1999: dtv.

- Harras, Gisela (1989): "Zu einer Theorie des lexikographischen Beispiels". In: Hausmann, Franz Josef et al. (ed.): *Wörterbücher. Dictionaries. Dictionnaires. Ein internationales Handbuch zur Lexikographie*. Bd. 1. Berlin/New York, de Gruyter: 607–614.
- Haß-Zumkehr, Ulrike (2001): *Deutsche Wörterbücher. Brennpunkt von Sprach- und Kulturgeschichte*. Berlin/New York: de Gruyter.
- Haß, Ulrike (1991): "Textkorpus und Belege. Methodologie und Methoden". In: Harras, Gisela et al. (ed.): *Wortbedeutungen und ihre Darstellung im Wörterbuch*. Berlin/New York, de Gruyter: 212–292.
- Heid, Ulrich (2007): "Computational linguistic aspects of phraseology II". In: Burger, Harald et al. (ed.): *Phraseologie. Ein internationales Handbuch der zeitgenössischen Forschung*. Bd. 2. Berlin/New York, de Gruyter: 1036–1044.
- Hermanns, Fritz (1988): "Das lexikographische Beispiel. Ein Beitrag zu seiner Theorie". In: Harras, Gisela (ed.): *Das Wörterbuch. Artikel und Verweisstrukturen*. Düsseldorf, Schwann: 161–195.
- Hunston, Susan (2006): "Corpus Linguistics". In: Brown, Keith (ed.): *Encyclopedia of language and linguistics*. Bd. 3. 2. Aufl., Amsterdam, Elsevier: 234–248.
- Korhonen, Jarmo (2004): "Duden 11. Nutzungserfahrungen aus der DaF-Perspektive". In: Steyer, Kathrin (ed.): *Wortverbindungen. Mehr oder weniger fest. Vorträge der 39. Jahrestagung des Instituts für deutsche Sprache, 2003 in Mannheim*. Berlin, de Gruyter: 360–393.
- Krishnamurthy, Ramesh (2006): "Corpus Lexicography". In: Brown, Keith (ed.): *Encyclopedia of language and linguistics*. Bd. 3. 2. Aufl., Amsterdam, Elsevier: 250–254.
- Ooi, Vincent B. Y. (1998): *Computer corpus lexicography*. Edinburgh: Edinburgh University Press.
- Reichmann, Oskar (1988): "Zur Funktion, zu einigen Typen und zur Auswahl von Beispielbelegen im historischen Bedeutungswörterbuch". In: Hyldgaard-Jensen, Karl/Zettersten, Arne (ed.): *Symposium on Lexicography IV. Proceedings of the Fourth International Symposium on Lexicography, April 20–22, 1988 at the University of Copenhagen*. Tübingen, Niemeyer: 413–444.
- Röhrich, Lutz (ed.) (2002): *Das große Lexikon der sprichwörtlichen Redensarten*. 3 Bde. Darmstadt/Basel: WBG.
- Rothkegel, Annely (2007): "Computerlinguistische Aspekte der Phraseme I". In: Burger, Harald et al. (ed.): *Phraseologie. Ein internationales Handbuch der zeitgenössischen Forschung*. Bd. 2. Berlin/New York, de Gruyter: 1027–1035.
- Stantcheva, Diana (2003): *Phraseologismen in deutschen Wörterbüchern. Ein Beitrag zur Geschichte der lexikographischen Behandlung von Phraseologismen im allgemeinen einsprachigen Wörterbuch von Adelung bis zur Gegenwart*. Hamburg: Dr. Kovac. (= *Philologia* 53).
- Steyer, Katrin (2004): "Kookkurrenz. Korpusmethodik, linguistisches Modell, lexikografische Perspektiven". In: Steyer, Kathrin (ed.): *Wortverbindungen. Mehr oder weniger fest. Vorträge der 39. Jahrestagung des Instituts für deutsche Sprache, 2003 in Mannheim*. Berlin, Walter de Gruyter: 87–116.
- Stubbs, Michael (2001): *Words and phrases. Corpus studies of lexical semantics*. Oxford: Blackwell Publishers.
- Wiegand, Herbert Ernst (1989): "Die lexikographische Definition im allgemeinen einsprachigen Wörterbuch". In: Hausmann, Franz Josef et al. (ed.): *Wörterbücher. Dictionaries. Dictionnaires. Ein internationales Handbuch zur Lexikographie*. Bd. 1. Berlin/New York, de Gruyter: 530–588.
- Wiegand, Herbert Ernst (1994): "Kritische Lanze für Fackel-Redensartenwörterbuch. Bericht und Diskussion zu einem Workshop in der Österreichischen Akademie der Wissenschaften am 14.2.1994". *Lexicographica. International Annual for Lexicography* 9: 230–271.