

Studien zur deutschen Sprache

FORSCHUNGEN DES INSTITUTS FÜR DEUTSCHE SPRACHE



Wolfgang Teubert
(Hrsg.)

Neologie und Korpus



Gunter Narr Verlag Tübingen

Studien zur deutschen Sprache

FORSCHUNGEN DES INSTITUTS FÜR DEUTSCHE SPRACHE



Herausgegeben von

Bruno Streckler, Reinhard Fiehler und Hartmut Günther

Band 11 · 1998

Wolfgang Teubert
(Hrsg.)

Neologie und Korpus



Gunter Narr Verlag Tübingen

Die Deutsche Bibliothek – CIP-Einheitsaufnahme

Neologie und Korpus / Wolfgang Teubert (Hrsg.). –

Tübingen : Narr, 1998

(Studien zur deutschen Sprache ; Bd. 11)

ISBN 3-8233-5141-9

© 1998 · Gunter Narr Verlag Tübingen

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 säurefreiem und alterungsbeständigem Werkdruckpapier.

Gesamtherstellung: Hubert & Co., Göttingen

Printed in Germany

ISSN 0949-409X

ISBN 3-8233-5141-9

Inhalt

Vorwort	7
<i>Irmhild Barz</i>	
Neologie und Wortbildung. Zum Neuheitseffekt von Wortneubildungen	11
<i>Cyrl Belica</i>	
Statistische Analyse von Zeitstrukturen in Korpora	31
<i>Dieter Herberg</i>	
Neues im Wortgebrauch der Wendezeit. Zur Arbeit mit dem IDS-Wendekorpus	43
<i>Michael Kinne</i>	
Der lange Weg zum deutschen Neologismenwörterbuch. Neologismus und Neologismenlexikographie im Deutschen: Zur Forschungs- geschichte und zur Terminologie, über Vorbilder und Aufgaben	63
<i>John Sinclair</i>	
Korpus typologie. Ein Klassifikationsrahmen	111
<i>Wolfgang Teubert</i>	
Korpus und Neologie	129

Vorwort

Korpusbasierte Lexikographie steckt in Deutschland immer noch in den Kinderschuhen. Das hängt vielleicht mit der Wertschätzung des Strukturalismus Saussurescher Prägung zusammen, der dem ehernen Sprachsystem, der *langue*, einen olympischen Vorzugsplatz einräumt (was, wie wir heute wissen, nicht de Saussures Intentionen entsprach), während die *parole* als Emanation menschlicher Fehlerhaftigkeit gilt. Vergleichbar steht für die recht zahlreichen Kognitivisten der Chomsky-Schule das genetisch determinierte Sprachorgan und nicht die korrumpierte Performanz im Mittelpunkt ihres Interesses. Korpora können zur Erklärung der Funktionsweise metaphysischer Realitäten oder biologischer Phänomene wenig beitragen.

Doch muß es weitere Gründe geben, warum es für das Deutsche, verglichen mit anderen westeuropäischen Sprachen, so wenig korpusbasierte Wörterbücher gibt. Große Wörterbücher würden große Korpora erfordern, und die kosten Geld, mehr Geld, als deutsche Verlage zu investieren bereit sind. Auch in anderen Ländern sieht man vor allem aus Kostengründen die Erstellung nationaler Sprachkorpora als öffentliche Aufgabe an, mit der zentrale Sprachinstitute beauftragt werden. Allerdings gibt es in England, in Skandinavien, auf der iberischen Halbinsel oder in den Niederlanden eine Tradition echter Zusammenarbeit zwischen akademischer Forschung und kommerzieller Lexikographie, die in Deutschland leider selten ist. Bei uns fehlt ein Äquivalent zum *British National Corpus*, und ebensowenig gibt es den verlegerischen Ehrgeiz, lexikologische Erkenntnisse kompetenter in die lexikographische Praxis umzusetzen als die nationale und internationale Konkurrenz. Vielleicht mangelt es hierzulande am Wettbewerb.

Deswegen gibt es nur einige Spezialwörterbücher für das Deutsche, die für sich in Anspruch nehmen können, korpusbasiert zu sein; und sie sind nicht Ergebnisse unternehmerischer Initiative, sondern Produkte akademischen Schweißes. Leider wenden sie sich denn auch in erster Linie an den Fachkollegen und daneben allenfalls an andere professionell interessierte Benutzer, aber kaum an ein breiteres Publikum. Spezialwörterbücher erfordern Spezialkorpora, und die können, je nach Gegenstand, deutlich kleiner (und damit billiger) sein als Referenzkorpora, die den Sprachgebrauch einer Sprachgemeinschaft einigermaßen erschöpfend abbilden sollen. Geld für Spezialkorpora läßt sich im Rahmen üblicher Verfahren aufreiben. Für nationale Sprachkorpora reicht das nicht. Referenzkorpora, wie diese Textsammlungen heute gern genannt werden, siedelt man in der Größenordnung von einigen hundert Millionen Wörtern Textlänge an. Das dem Wörterbuch *Schlüsselwörter der Wendezeit* zugrundeliegende Wendekorpus hat demgegenüber nur einen Umfang von gut drei Millionen Wörtern.

Nun ist ein Neologismenwörterbuch zwar ein Spezialwörterbuch, aber es beschreibt keinen eingrenzenden Sprachausschnitt, sondern die ganze Standardsprache, also die ganze deutsche Sprache, sieht man einmal von den Fach- und Sondersprachen ab. Für ein korpusbasiertes Neologismenwörterbuch würde man also eigentlich ein Referenzkorpus benötigen, und mehr als das, nämlich ein Referenzkorpus für jeden zeitlichen Vergleichsabschnitt. Eine solche Kette gleich aufgebauter Korpora (man spricht hier mit John Sinclair von einem *Monitorkorpus*) gibt es bislang auch anderswo nur als Wunschvorstellung. An die Realisierung hat sich, vor allem aus Kostengründen, noch keiner gewagt. Bei der Planung des neuen Arbeitsvorhabens *Neologismenwörterbuch* am Institut für deutsche Sprache war deshalb die grundlegende Frage zu diskutieren, ob dieses Vorhaben wirklich korpusbasiert oder nur unter gelegentlicher Benutzung ohnehin existierender Korpora, und im übrigen auf der Grundlage manueller Exzerptionsarbeit, zu realisieren ist. Die Entscheidung darüber hängt von dem Anspruch an das zu erwartende Produkt ab. Soll es im wesentlichen verzeichnen, was dem wachen Sprachteilhaber, dem Lexikographen, als Neologismus auffällt, oder sind Gegenstand des Wörterbuchs auch solche Neologismen, die dem Leser zunächst nicht als solche auffallen, die aber automatische Verfahren als neue Ausdrücke und (was ungleich komplizierter ist) alte Ausdrücke in neuer Verwendungsweise aus zeitlich gestaffelten Korpora extrahieren und dem Lexikographen zur Prüfung vorlegen?

Das Wort *Besitzstände* etwa hat es als sprachlichen Ausdruck immer schon gegeben, und kaum einer würde dabei an einen Neologismus denken. Nur wenn man Korpora analysiert, erkennt man, daß dieser Ausdruck ziemlich genau seit Herbst 1994 auf gesetzlich und vertraglich geregelte Ansprüche angewendet wird, die Bürger, besonders Arbeitnehmer und Wohlfahrtsempfänger, gegenüber staatlichen Stellen, dem Sozialversicherungswesen oder ihren Arbeitgebern haben. Diese Bedeutung ist neu. Ist *Besitzstände* deshalb ein Neologismus? Ein anderes Beispiel wäre das Wort *Lohnabstandsgebot*. Es ist ebenfalls erst seit einigen Jahren belegt. Es ist, anders als *Besitzstände*, ein wirklich neues Wort, das es vorher nicht gegeben hat. Aber es ist ein Nominalkompositum, wie es prinzipiell unendlich viele gibt. Sie werden ständig ad hoc gebildet, und niemand würde daran denken, sie alle im Wörterbuch zu verzeichnen. Dorthin, so heißt es, gehören nur lexikalisierte Komposita. Was aber sind die Kriterien, nach denen über Lexikalisierung entschieden werden sollte?

In der Vorbereitung des Langzeitprojekts *Neologismenwörterbuch* galt es also, zwei Fragen zu klären. Zum einen ist das die Frage, ob strikte Korpusbasiertheit, also die systematische Auswertung eines Korpus, wirklich so unabdingbar ist oder angesichts der Kosten lieber fallengelassen werden sollte. Zum andern geht es darum, den Neologismenbegriff zu klären und für die lexikographische Praxis eine nachvollziehbare Definition zu erarbeiten. Schließlich ging es darum, die Neologismendefinition so mit der Korpusfrage zu ver-

knüpfen, daß sichtbar wird, ob und wie der Computer wirklich die Lexikographen von ermüdender Routine befreien kann, so daß sie sich möglichst ausschließlich der kreativen Gestaltung von Wörterbucheinträgen widmen können. Wie läßt sich das Verhältnis von Aufwand und Ertrag optimieren?

Die Vorbereitungsphase für das inzwischen angelaufene Neologismenprojekt liegt nun schon lange zurück. An ihrem Anfang stand ein Kolloquium zum Thema *Neologie und Korpus*, das im Oktober 1993 am Institut für deutsche Sprache stattfand. Die hier vorgelegten Beiträge sind aus diesem Kolloquium hervorgegangen. Eine ganze Reihe unglücklicher Umstände hat dazu geführt, daß der Band erst jetzt veröffentlicht wird, Umstände, an denen ich selbst zum nicht geringen Teil schuld bin. Den Autoren danke ich für die nachsichtige Langmut, die sie mit mir hatten. Alle Aufsätze sind, davon bin ich überzeugt, heute noch so aktuell wie damals. In der Tat interessiert sich die Korpuslinguistik heute mehr denn je für das Thema lexikalischen Wandels. Der vorliegende Band wird diese Diskussion anregen.

Wolfgang Teubert

Irmhild Barz

Neologie und Wortbildung

Zum Neuheitseffekt von Wortneubildungen

Der Beitrag will untersuchen, welche sprachlichen Faktoren die Intensität des „Neuheitseffektes“ (Olsen 1986, S. 50) neugebildeter substantivischer Komposita bei Rezipienten beeinflussen können und wie die Faktoren zusammenwirken.

Diese Fragen ordnen sich in aktuelle Probleme der Neologie ein, insbesondere in die Diskussion um den Neologismusbegriff sowie um Methoden zur Ermittlung neologischer lexikalischer Einheiten und ihrer lexikographischen Darstellung. In der Wortbildungsforschung spielen sie im Zusammenhang mit Entwicklungstendenzen, und zwar vor allem mit der Produktivität von Modellen und der Wortbildungsaktivität sprachlicher Einheiten eine Rolle, aber auch bei der Modellierung von Produktions- und Verstehensprozessen und bei stilistischen Analysen.

Komplementäre Produktionsaspekte sowie die Lexikalisierung von Neubildungen können hier nicht erörtert werden. Ein Zusammenhang zwischen bewußter und unbewußter Wortproduktion und Auffälligkeitsgrad der Produktionsergebnisse einerseits und deren Lexikalisierungsaffinität andererseits ist jedoch mit Sicherheit anzunehmen.¹

Nach einer knappen theoretischen Fundierung möglicher Beziehungen zwischen Neuheitseffekt und Worteigenschaften werden im folgenden empirisch erhobene Vertrauheitsurteile² über komplexe Wörter beschrieben und interpretiert. Die Ergebnisse werden als Antworten auf die eingangs gestellten Fragen zusammengefaßt.

¹ Vgl. dazu vor allem Wilss (1985, 1986 und 1992); zur Verbreitung sprachlicher Neuerungen grundsätzlich Große/Neubert (1982), des weiteren Grimm (1991); Helfrich (1993); zum Zusammenhang zwischen Auffälligkeit und Lexikalisierung Erben (1981) und Fleischer (1983).

² Die Bezeichnung ist von Plank übernommen. Er behandelt Vertrauheitsurteile im Zusammenhang mit Wohlgeformtheitsurteilen über Neubildungen (1981, S. 29) und zeigt, daß Urteile über die Wohlgeformtheit von Wörtern von deren Bekanntheit bestimmt sein können. Dieser Gedanke wird hier nicht weiter verfolgt, vielmehr wird Wohlgeformtheit als gegeben angenommen.

1. Neuheit und Neuheitseffekt

Neue Wörter einerseits und Wortbildungsprodukte andererseits sind verschiedene Teilmengen lexikalischer Einheiten, die nach jeweils unterschiedlichen Merkmalen aus dem Gesamtbestand an Wörtern ausgegliedert werden können: neue Wörter nach ihrer zeitlichen Markierung, Wortbildungsprodukte nach ihrer morphologischen und semantischen Durchschaubarkeit bzw. Motiviertheit. Wenn beide Merkmale zusammen bei ein und demselben Wort auftreten, liegt eine Neuprägung (WDG, Vorwort 014) oder „Wortneubildung“ (Plank 1981, S. 250) vor.

Wortneubildungen stellen eine zeitlich markierte Untermenge aller Wortbildungsprodukte des Wortbestandes dar, daneben existieren – unter zeitlichem Aspekt gesehen – veraltende, veraltete und unmarkierte. Gleichzeitig sind Wortneubildungen auch eine Untermenge aller neuen lexikalischen Einheiten. Diese können außer durch Wortbildung auch durch Entlehnung, Bedeutungsbildung (Munske 1990, S. 389) oder Phraseologisierung entstanden sein. Auf eine einfache Formel gebracht, heißt das: Nicht alle neuen Wörter sind Wortbildungsprodukte, und nicht alle Wortbildungsprodukte sind zu einem bestimmten Zeitpunkt (objektiv) neu.

Wortneubildungen sind demzufolge zweifach abzugrenzen: von Entlehnungen, Neosemantismen und Phraseologismen einerseits und von lexikalisierten Wortbildungsprodukten andererseits. Während ersteres aufgrund relativ sicherer Kriterien (fremdsprachliche Herkunft bei Entlehnungen, gleiches Formativ bei Bedeutungsveränderungen, Wortgruppenstruktur bei Phraseologismen) kaum Schwierigkeiten macht – vom Mischcharakter der Lehnwortbildung einmal abgesehen – ist die einem Wort nur zeitweise anhaftende Eigenschaft Neuheit weder formal noch semantisch zuverlässig am Wort zu erkennen.

Neue Wörter sind nicht im Langzeitgedächtnis der Angehörigen einer Sprachgemeinschaft gespeichert. Sie können oft nicht unreflektiert verstanden werden, sondern fallen bei der Rezeption als ungewohnt auf. Wie hier zu zeigen sein wird, existiert diese „Reflexionshürde“ allerdings nicht in gleicher Intensität bei allen Wortneubildungen, so daß wohl kaum verallgemeinert werden kann, Muttersprachler seien *stats* in der Lage zu unterscheiden, welche Wörter „zum festen Bestandteil des Wortschatzes ihrer Muttersprache gehören und welche [...] neu gebildet werden“.³

Obwohl jede lexikalische Einheit irgendwann neu aufkommt, wird ihre Neuheit nicht von allen Sprachbenutzern gleichermaßen erkannt, d.h. das Wort wird nicht von allen eine Zeitlang „als neu empfunden“ (Herberg 1988, S. 9).

³ Siebert-Ott (1991, S. 83); Hervorhebung I. B. Wie Siebert-Ott vermutet auch Olsen (1986, S. 51f.) diese ausgeprägte diachrone Kompetenz; zur Kompetenz grundsätzlich Coseriu (1988).

Vielmehr kann es nahezu unauffällig bleiben, auch wenn es sich nicht über die einmalige Verwendung hinaus verbreitet. Spontane Urteile über die Neuheit lexikalischer Einheiten, aber auch die Stichwortlisten einsprachiger Wörterbücher offenbaren erhebliche Normunsicherheiten bei der Entscheidung neu/nicht neu. Dieser Sachverhalt legt es nahe, in der Neologie zwischen der objektiven Neuheit der Wörter einerseits und deren Neuheitseffekt, ihrer diesbezüglichen Wirkung auf die Rezipienten andererseits, zu unterscheiden.⁴ Während die Neuheit über Erstbelege und Vorkommensanalysen relativ sicher, wenn auch mit erheblichem Aufwand, ermittelt werden könnte, sind Neuheitseffekte nur aus Rezipientenaussagen über neue Wörter zu erschließen. Solche Bewertungen erweisen sich in zweifacher Hinsicht als heterogen. Zum einen kann der Neuheitseffekt eines objektiv neuen Wortes von verschiedenen Probanden verschieden stark empfunden werden, zum anderen beurteilt auch ein und derselbe Proband die Neuheit von Wörtern eines Textes graduell verschieden.

Den genannten Differenzen soll hier nachgegangen werden. Dabei ist grundsätzlich zu bedenken, daß sie, wie metasprachliche Bewertungen immer, durch vielfältige, einander überlagernde sprachliche und außersprachliche Einflußgrößen zustande kommen können und ganz gewiß nicht auf einer linearen Ursache-Folge-Beziehung beruhen (Barz 1996). Von dieser Multikausalität wird im folgenden zugunsten einer vereinfachenden, methodisch motivierten Idealisierung abgesehen. Untersucht werden die Wortbildungseigenschaften von Neubildungen sowie – in Einzelfällen – Aspekte ihrer Semantisierung durch die Textumgebung.

2. Determinanten des Neuheitseffektes

Man kann davon ausgehen, daß die Unsicherheiten bei der Bestimmung der Neuheit von Wortbildungsprodukten auf dem Wesen der Sprachverarbeitungsprozesse beruhen. Nach kommunikationstheoretischen und psycholinguistischen Untersuchungen ist Sprachverarbeitung ein deutlich „erwartungsgeleiteter“ Vorgang.⁵ Wenn man Rezipientenerwartungen als „Einstellungen des Empfängers gegenüber Sprachhandlungen“ versteht (Lerchner 1984, S. 254), die sich aus Sachkenntnis, Spracherfahrung und Sprachwissen, also aus der im Langzeitgedächtnis verankerten Sach- und Sprachkompetenz entwickeln, dann liegt auf der Hand, daß sich Rezipientenurteile über die Neuheit

⁴ Diese Unterscheidung knüpft an Planks Unterscheidung von primären und sekundären Sprachdaten an. Unter den sekundären Daten versteht er Urteile kompetenter Sprecher über Sprache (1981, S. 6).

⁵ Vgl. Fix (1987); Hörmann (1987, S. 127f.); Rickheit/Strohner (1989, S. 233); Schwarz (1992, S. 159).

lexikalischer Einheiten aus individuellem Wissen ergeben und demnach ganz verschieden ausfallen können.

Mit der Erklärung des Neuheitseffektes eines neuen Wortes aus seinem Verhältnis zu Sach- und Sprachwissen wird eine proportionale Erklärung der Intensität dieses Effektes möglich: Je stärker eine Wortneubildung zu dem Wissen des Rezipienten in Widerspruch gerät, um so stärker ist ihr Neuheitseffekt. Umgekehrt ist er um so geringer, je geringer der Widerspruch aufgrund gespeicherter Wissens Elemente ist, je mehr „Bekanntes“ das neue Wort enthält bzw. aktiviert.

Als sprachliche Faktoren, die diesen Widerspruch ausgleichen können, kommen solche in Betracht, die die Wortneubildung in vorhandene Wissenszusammenhänge stellen: (1) die Regelmäßigkeit der Bildung einschließlich der Bekanntheit der motivierenden lexikalischen Einheiten (Bezug auf lexikalisches Regelwissen), (2) die kontextfreie Verständlichkeit, d.h. Selbstdeutigkeit, des neuen Wortes (Bezug auf Wortwissen und Wortschatzstrukturwissen sowie Textwissen), und (3) die allmählich einsetzende Lexikalisierung.

- (1) Regelmäßig gebildete neue Wörter wirken nicht so neu wie unregelmäßig gebildete. Darauf ist immer wieder hingewiesen worden.⁶ So wie der Sprachteilhaber regelmäßig gebildete Sätze nicht als neu empfindet, weil „jeder Satz in diesem Sinn ‘neu’ ist“ (Olsen 1986, S. 50), kann er auch bestimmte regelmäßige Wortneubildungen nicht ohne weiteres als neu einstufen. Sie signalisieren ihre Neuheit weniger stark als nicht regelmäßig gebildete.

Damit dieser Zusammenhang beschrieben werden kann, muß eindeutig festgelegt werden, was mit ‘regelmäßig gebildetes Wort’ gemeint ist. Darunter verstehe ich solche Wortneubildungen, die an eine Reihe gleichstrukturierter lexikalisierter Wörter, an ein Wortbildungsparadigma, angeschlossen werden können. Diese „Paradigmenstütze“ (Ortner/Ortner 1984, S. 167) baut den Neuheitseffekt eines neuen Wortes um so stärker ab, je weniger das betreffende Wort von seinem „Reihenbildungsmuster“ (Wilss 1992, S. 232) abweicht.⁷ Ich betrachte diese Reihenbildung als eine mögliche Eigenschaft von Wortneubildungen und nenne sie Serialität, neue Komposita mit dieser Eigenschaft serielle

⁶ Vgl. u.a. Ortner/Ortner (1984, S. 167). Wilss unterscheidet in diesem Sinn zwischen strukturmusterorientierten und innovativen Verfahren (1985, S. 279).

⁷ Die Serialität reduziert den Neuheitseffekt mitunter so stark, daß entsprechende Wortneubildungen bei der Untersuchung von Okkasionalismen kaum beachtet werden, vgl. z.B. Seppänen (1986, S. 87f.); ausdrücklich berücksichtigt dagegen jetzt unter stilistischem Aspekt bei Handler (1993).

Komposita.⁸ Ihnen stehen die nicht reihenhaft gebildeten, die singulären, gegenüber.

Um die Eigenschaft Serialität bei neuen Komposita diagnostizieren zu können, ist festzulegen, was man bei substantivischen Komposita unter einer Reihe verstehen möchte, um so mehr, als Reihenbildung im allgemeinen als Domäne der Derivation betrachtet und auch vorrangig an dieser Wortbildungsart behandelt wird. Dabei empfiehlt es sich, den Begriff auf einer sehr niedrigen Abstraktionsstufe anzusiedeln, weil, wie zu sehen sein wird, gerade relativ konkrete Muster offenbar jene Ebene darstellen, auf der sich die unreflektierte, automatische Rezeption von neuen Komposita vollzieht (Wilss 1992, S. 231). Die Annahme der mentalen Präsenz von Mustern eines geringen Abstraktionsgrades läßt auch eher Raum für die Erklärung von Wortneubildungen, die nach dem Vorbild von Einzelwörtern entstehen.

Eine Kompositionsreihe wird hier als ein Paradigma angesehen, das aus Komposita mit einer identischen und einer variablen Konstituente sowie identischer Wortbildungsbedeutung besteht. Als weiteres Merkmal kommt hinzu, daß die Stelle der variablen Konstituente nur von solchen Wörtern besetzt sein soll, die durch paradigmatische semantische Beziehungen miteinander verbunden sind, wie z. B. von Kohyponymen wie *Pflaume(n)*, *Apfel*, *Kirsch(e)* in Komposita mit *-baum*.⁹ Die letztgenannte Bedingung hat den Sinn, die Erklärungskraft dieses Paradigmas für Schwankungen des Neuheitseffektes von Wortneubildungen zu gewährleisten, wie in 4. deutlich werden wird. Für andere Untersuchungsziele, etwa die Ermittlung der generellen Kompositionsaktivität von Wörtern, ist sie möglicherweise überflüssig.

Für die mit *Baum* lexikalisierten Komposita ergibt sich beispielsweise aus dem hier festgelegten Reihenbegriff folgende Reihenbindung. *Baum* ist sowohl als Erst- als auch als Zweitglied kompositionsaktiv. In beiden Positionen prägt es Reihen aus, als Zweitglied z.B. in Verbindung mit Früchten und Samen wie in *Apfel-*, *Birn-*, *Feigenbaum* oder in Verbindung mit bestimmten Baumarten wie *Eich(en)-*, *Fichten-*, *Lindenbaum* (vgl. WDG, S. 436). Daneben existiert eine Vielzahl singularer, nicht seriell gebildeter Komposita wie *Mai-*, *Purzel-*, *Zwergbaum*.

- (2) Mit der Fixierung der Einflußgröße Verständlichkeit soll hier nicht eigens die unstrittige Tatsache behandelt werden, daß unverständliche Wörter mit Sicherheit einen Neuheitseffekt auslösen. Vielmehr wird ihre grundsätzliche Verständlichkeit als gegeben vorausgesetzt. Das ist inso-

⁸ Dieser Terminus ist nicht zu verwechseln mit „serial combinations“ bei Ortner/Ortner (1984, S. 185).

⁹ Zur Angemessenheit diese Reihenbegriffs vgl. ausführlicher Barz (1989, S. 317ff.).

fern legitim, als Wortneubildungen normalerweise immer in Texten und nur in Ausnahmefällen isoliert vorkommen. Nur in wenigen Textarten, z.B. in poetischen Texten, wird semantische Unterbestimmtheit von Wortneubildungen bewußt erzeugt und beibehalten; davon kann abgesehen werden. Hier geht es darum, die verbreitete Differenzierung von Komposita in „selbstverständliche“ (selbstdeutige, lokal interpretierbare, kontextfreie) und „pragmaverständliche“ (nicht selbstdeutige, nicht lokal interpretierbare, kontextgebundene) (Gersbach/Graf 1984, S. 32ff.) in bezug auf ihre Relevanz für den Neuheitseffekt der betreffenden Komposita zu prüfen.¹⁰

Für die Selbstdeutigkeit der neuen Komposita sorgen nach Fandrych/Thurmair (1994, S. 38f.) im wesentlichen lexikalisch-semantische Eigenschaften der Konstituenten, wie z.B. die Relationalität des Zweitgliedes, und Grundrelationen zwischen Erst- und Zweitglied, wie 'Situation', 'Konstitution' oder 'Zweck'. Bei seriellen Komposita kommt m.E. als weiterer verständlichkeitsfördernder Faktor der Paradigmenbezug hinzu, so daß serielle Komposita weitgehend selbstdeutig sein müßten. Aber auch bei der Interpretation singulärer Komposita ist der Rezipient keineswegs immer auf Textinformationen angewiesen. Auch sie können aus sich selbst heraus verstanden werden. Insbesondere Wortneubildungen nach Einzelwortmustern werden durch diese Wort-Assoziation verständlich, ohne daß sie unbedingt in einen Text eingebettet sein müssen, vgl. die Bildung von *Stammtischtäter* zu *Schreibtischtäter*.¹¹ Meistens jedoch braucht der Rezipient Textwissen für die Interpretation solcher singulären Bildungen (Heringer 1984; Schröder 1985; Wellmann 1993). Die Konstituenten des neuen Kompositums können im Text z.B. wörtlich wiederholt werden oder auch semantisch, z.B. durch Synonyme oder Kontrastwörter, vor- oder nachbereitet sein. Die Bedeutung kann sich dem Rezipienten aber auch aus dem Textsinn insgesamt erschließen (Wildgen 1982). Nach allem, was über Okkasionalismen bekannt ist (vgl. Fleischer 1983), sollte man von einem proportionalen Verhältnis zwischen der Textbindung der Bedeutung neuer Komposita und deren Neuheitseffekt ausgehen: je stärker die Textbindung, um so höher der Neuheitseffekt.

- (3) Die dritte in Zusammenhang mit dem Neuheitseffekt wichtige Wort-eigenschaft ist der Lexikalisierungsgrad. Dessen Verhältnis zum Neuheitseffekt erweist sich zunächst als unproblematisch, denn fehlende Lexikalisiertheit ist ein Definitionsmerkmal von Wortneubildungen. Dar-

¹⁰ Vgl. Boase-Beier/Toman (1986, S. 62); Fanselow (1981); Fandrych/Thurmair (1994).

¹¹ Zum Phänomen Akzeptabilität in der Wortbildung vgl. Wildgen (1982); Stepanowa/Fleischer (1985, S. 78f.); Matussek (1994); Fleischer/Barz (1995, S. 59f.); zum Zusammenwirken verschiedener Wissensbereiche bei der Verarbeitung von Wortbildungsprodukten vgl. Wilss (1986, S. 72ff.); Motsch (1995).

aus ergibt sich, daß die einsetzende Lexikalisierung eines Wortes zwangsläufig zum Abbau seines Status als Neubildung und damit auch zur Verringerung des Neuheitseffektes beiträgt (Fleischer 1983, S. 46). Es ist nur schwierig, den Nachweis über die Lexikalisiertheit zu führen (Helfrich 1993).

3. Zur Lexikalisiertheit von Wortneubildungen

Nach ihrem Lexikalisiertheitsgrad werden Wortneubildungen gewöhnlich in okkasionelle (Ad-hoc-)Bildungen und usuelle Neologismen gegliedert. Ein Teil der Okkasionalismen findet unter bestimmten Bedingungen Verbreitung und geht allmählich in den Wortschatz ein. Gemessen an ihrem Gesamtaufkommen in Texten – bei substantivischen Komposita in Kochbüchern wurden z.B. mehr als 20% Okkasionalismen gezählt (Müller-Bollhagen 1985, S. 234)¹² – nehmen jedoch nur relativ wenige diese Entwicklung.¹³ Nach Wellmann (1995, S. 400) enthalten Wörterbücher nur rund zwei Drittel der in einem durchschnittlichen Zeitungstext vorkommenden Wortbildungen, wobei das fehlende Drittel natürlich auch andere periphere Wörter, wie z.B. regional markierte, einschließen kann. Im WDG, das als einziges großes Bedeutungswörterbuch seine Lemmata durchgängig chronologisch markiert, beträgt der Anteil der Neologismen am Gesamtbestand der Stichwörter nur etwa 3% (vgl. Fleischer 1983, S. 49).

Die Prozentangaben erwecken den Anschein, als handle es sich bei Okkasionalismen, Neologismen und dauerhaft lexikalisierten Wörtern um deutlich diskrete Gruppen lexikalischer Einheiten. Das trifft bekanntermaßen nicht zu, vielmehr ist mit einer differenzierten Skala von Lexikalisiertheitsgraden zu rechnen, deren Pole Okkasionalität einerseits und Lexikalisiertheit andererseits darstellen.¹⁴ Da die eingangs aufgeworfenen Fragen grundsätzlich für alle neuen Wörter zutreffen und zunächst einmal unabhängig von einer beginnenden Usualisierung behandelt werden können, verzichte ich hier auf die Unterscheidung zwischen Okkasionalismen und Neologismen und fasse die durch Wortbildung entstandenen neuen Wörter mit dem Terminus Wortneubildung zusammen.

Auch wenn durch diese Verallgemeinerung eine vereinfachende Gliederung des gesamten Wortbildungsinventars in nur drei Gruppen, in neue, alte und

¹² Als ähnlich kreativ erweisen sich Autoren von Fachtexten, vgl. dazu detailliert Stein (1993).

¹³ Vgl. Erben (1981); Fleischer (1983); Ortner/Ortner (1984, S. 166ff.)

¹⁴ Ein differenziertes Schema zur „inneren Dynamik der Wortbildungsverhältnisse“ hat jetzt Handler vorgelegt. Es sieht fünf Stufen bis zur Lexikalisiertheit vor (1993, S. 291). Matussek verzichtet auf einen Übergangsbereich zwischen Neologismen und lexikalisierten (usuellen) Wortbildungsprodukten und rechnet mit vier Stufen (1994, S. 38).

zeitlich unmarkierte entsteht, bleibt die Zuordnung einzelner Beispiele zur Gruppe der neuen Wörter, um die es hier ausschließlich geht, sehr oft eine Ermessensentscheidung, und zwar deshalb, weil – wie oben angedeutet – Neuheit und Neuheitseffekt bei Wortneubildungen nicht zusammenfallen und die Identifizierung neuer Wörter damit weitgehend kompetenzabhängig ist.

Zuverlässige Hilfsmittel als Ersatz für das unzulängliche „diachronische Bewußtsein“ (Coseriu 1988, S. 135) gibt es bislang nicht und wird es wohl auch kaum geben können. Das Fehlen lexikalischer Einheiten in den Wörterbüchern ist kein sicheres Indiz für ihre Neuheit, wie man weiß.¹⁵ So verzeichnet z.B. der DUR (1991) *Studienabbrecher*, nicht aber *Studienanfänger* und *errätbar*, nicht aber *erwartbar*, ohne daß zwischen den Wörtern intuitiv Lexikalisiertheitsunterschiede zu entdecken wären. Nichtsdestoweniger bieten einsprachige Bedeutungswörterbücher Sprachdaten in einem Umfang und einer Tiefe, wie sie dem Einzelnen unter keinen Umständen mental zur Verfügung stehen können. In den einleitenden Bemerkungen zur zweiten Auflage des DGW wird in bezug auf die Stichwortauswahl angekündigt, „individualsprachliche Prägungen und Augenblicks- oder Situationsbildungen, d.h. Wörter, die jederzeit gebildet werden können, die aber nicht fester Bestandteil unseres Wortschatzes sind“, nicht aufzunehmen (DGW 1993, S. 7), aber den Wortschatz der deutschen Gegenwartssprache „mit allen Ableitungen und Zusammensetzungen so vollständig wie möglich“ (ebd.) zu erfassen. Angaben dazu, wie die Qualität eines Wortes, „fester Bestandteil unseres Wortschatzes“ zu sein, ermittelt wird, fehlen zwar, so daß auch hier mit Ermessensspielräumen zu rechnen ist. Dennoch kann gerade ein Wörterbuch mit einer so großen korpusbasierten Stichwortzahl wie das DGW bei der Beschreibung der Wortneubildungen mit Gewinn genutzt werden, und zwar vor allem für das Erkennen von lexikalisierten Wortbildungsreihen und bei der Suche nach lexikalisierten Musterwörtern für singuläre Wortneubildungen.

Für das Vorhaben, Reihenbildung zu erfassen, erweist sich neben dem DGW das WDG als besonders nützlich, weil es in den einzelnen Artikeln auch systematisch typische Komposita auflistet, in denen das betreffende Stichwort Zweitglied ist.

4. Exemplarische Wortanalysen

Vorstehende Überlegungen sollen nun an Vertrautheitsurteilen über Wortneubildungen eines kurzen, abgeschlossenen Textes überprüft werden. Der Text wurde 27 Germanistikstudentinnen und -studenten eines Hauptseminars

¹⁵ Vgl. Fleischer (1983), Ortner/Ortner (1984), Müller-Bollhagen (1985, S. 228), Matussek (1994, S. 33f.); zur Notwendigkeit, Neologismen lexikographisch zu erfassen vgl. Müller (1987); Heller u. a. (1988).

im Sommersemester 1994 an der Universität Leipzig mit der Bitte vorgelegt, „neue Wörter“ zu identifizieren. Auf Thema und Aufgabe waren die Probanden nicht direkt vorbereitet, Wörterbücher standen nicht zur Verfügung.

Der Text stammt aus dem „Kölner Stadtanzeiger“ und wird in Olsen (1986, S. 52f.) als Übungstext zum Thema Okkasionalität angeboten. Er wird hier vollständig wiedergegeben, wobei die behandelten Beispiele, substantivische Komposita, hervorgehoben sind (Hervorhebungen I. B.). Wortneubildungen anderer Wortart (*igelmüde*) oder anderer Wortbildungsart (*Stachelpelziger*) bleiben außer Betracht.

Besorgt über die *Igel-Flut*

Es gab eine Zeit, da glaubten die Tierschützer in Köln, man müßte etwas mehr für das Überleben einer *Tierart* tun, die viele Bürger nur als breitgewalzte leblose Masse auf den Straßen kennen: Igel. Es wurden *Igel-Merkblätter* gedruckt mit Tips über *Igelfutter*, *Igeldesinfektion* und *Igelhygiene*. *Igel-Auffangstellen* wurden zu jedem *Winterbeginn* eingerichtet. *Igelfreilassungen* und *Igelberichte* in den Zeitungen folgten, sobald im *Frühjahr* die Sonne wieder wärmte.

Es gab bald *Igelspezialistinnen* und *-spezialisten*, und die kellerten ihre *Fund-Igel* pfundweise ein, wie andere Bürger die Kartoffeln. Sie pflegten sie aufopfernd und hofften bald Nachahmer zu finden, auf daß die Arbeit nicht allein an ihnen hängen bleibe.

Als dann aber jahrein, jahraus im Herbst eine unberechenbare *Igel-Flut* über ihr Heim hereinbrach und das *Igel-Nottelefon* nicht mehr verstummen wollte, warfen einige der Pflegerinnen und Pfleger verständlicherweise das *Handtuch* und drangen auf Veröffentlichung, daß sie leider ihre *Igelherberge* dichtgemacht hätten. Auch der Zoo verlaublich, er sei igelmüde.

Doch für die Geister, die sie einst riefen, so scheint es, folgt nun mal auf die *Pflaumen-* und die *Pilzzeit* die Zeit der armen Igel. Und wehe, einer der Stachelpelzigen schlufft im Bergischen Land nur ein wenig blaß durchs *Herbstlaub*. Schwupp, hat ihn einer seiner ausgeschwärmte Igel und Erholung suchenden Kölner Freunde schon am Schlafittchen und ab mit ihm in die gute Stube. Und da muß er natürlich blitzschnell wieder weg, sobald er anfängt, streng zu riechen.

Kurz, die *Igelfinder* sind inzwischen Legion, Pfleger dagegen muß man, wie schon zu Beginn der *Igelbewegung* vor ungefähr einem *Jahrzehnt* mit der Lupe suchen. Und im *Tierheim* Dellbrück stöhnt man, weil dort die Igel jetzt schon weit vor Einbruch der Kälte eintrudeln. Immer häufiger werden dort auch „*Präventiv-Igel*“ abgeliefert, feiste Kameraden, die mit ihren tausend Gramm locker jeden Winter gesund überschlafen hätten, wenn ..., ja wenn sich nicht ein *Igel-Sammler* ihrer erbarmt hätte.

Mit dem Instrumentarium der Wortbildungsanalyse wird im folgenden untersucht, ob und wie die Probandenurteile über die Neuheit der Komposita des Textes mit deren Wortbildungsmerkmalen korrelieren. Berücksichtigt werden die Merkmalpaare Serialität/Singularität und Selbstdeutigkeit/Textgebundenheit.

Nach den Entscheidungen der Probanden zwischen „neu“ und „nicht neu“ gliedern sich die substantivischen Komposita des Textes in drei Gruppen: Es sind Komposita, die einheitlich als nicht neu empfunden werden, Komposita, die einheitlich als neu gelten, und solche, deren Beurteilung uneinheitlich ausfällt. Ich bezeichne sie nach ihrem Neuheitseffekt als unauffällige, auffällige und indifferente Wortneubildungen.¹⁶ Sie werden nun in dieser Reihenfolge beschrieben.

Die Gruppe der UNAUFFÄLLIGEN KOMPOSITA bilden – entsprechend der Abfolge im Text – *Tierart, Igelfutter, Winterbeginn, Frühjahr, Handtuch, Pflaumenzeit, Pilzzeit, Herbstlaub, Jahrzehnt, Tierheim*. Die meisten von ihnen sind zweifellos lexikalisiert, bezeichnen etablierte Begriffe (Hoffmann 1986, S. 34) und sind gebräuchlich, folglich für das hier erörterte Problem nicht von Interesse. Unerwartet ist jedoch die Einordnung von *Igelfutter, Winterbeginn, Pflaumenzeit* und *Pilzzeit* in diese Gruppe. Deren Lexikalisiertheit scheint mir nicht in gleicher Weise wie bei den erstgenannten Komposita gegeben zu sein. Die zu Rate gezogenen Wörterbücher DGW, DUW und WDG bestätigen diesen Eindruck, sie verzeichnen die vier Komposita nicht. Wie kommt es dennoch zur Vertrautheit der Rezipienten mit den Bildungen? Zur Beantwortung der Frage analysiere ich die Wörter einzeln nach den obengenannten Gesichtspunkten.

Igelfutter: Das WDG verzeichnet *Futter* als reihenbildend „in Verbindung mit Tierbezeichnungen [...], mit Bezeichnung von Futtermitteln“ und führt weitere fünfzehn Einzelkomposita mit *Futter* als Zweitglied an (WDG, S. 1438f.) Die Reihen werden im WDG als offene, erweiterungsfähige Paradigmen mit „unbegrenzter Auffüllbarkeit“ (Wilss 1985, S. 287) ausgewiesen, indem entsprechende Komposita ausdrücklich als Beispiele zitiert sind. Exemplarisch für die erstgenannte Reihe erscheinen *Fisch-, Hühner-, Schweine-, Vieh-, Vogelfutter*. Demnach ist *Igelfutter* eine regelmäßige Bildung, die offenbar nur wegen der geringeren lebenspraktischen Bedeutsamkeit des bezeichneten Gegenstandes nicht üblich ist. Das Kompositum ist selbstdeutig, es kann aus den Konstituentenbedeutungen verstanden werden, denn in der lexikalischen Bedeutung von *Futter* ist der Bezug auf Tiere enthalten, er ist ein Bestandteil der Bedeutung ‚Nahrung für Tiere‘.¹⁷

¹⁶ Die Merkmalopposition auffällig – unauffällig wird in der Wortbildung für unterschiedliche Markiertheitsphänomene verwendet, anders als hier z.B. bei Matussek (1994, S. 38f.). – Die Bindestrichschreibung bleibt bei der Materialuntersuchung unberücksichtigt, da sie im analysierten Text keinem einheitlichen Prinzip folgt.

¹⁷ Die Unauffälligkeit der Wortneubildung *Igelfutter* läßt den Schluß zu, daß auch weitere Komposita mit einer Tierbezeichnung als Erstglied gebildet werden könnten (*Hunde-, Kaninchen-, Katzen-, Pferdefutter*), wobei die Bedingung für Unauffälligkeit offenbar im pragmatischen Bereich liegt. Ausschlaggebend ist der Bezug der Tiere zum Menschen, so daß Bildungen wie *Frosch-, Libellenfutter* vermutlich von Probanden als neu eingestuft würden. Vgl. entsprechende Überlegungen zur Movierung bei Plank (1981, S. 96ff.).

Winterbeginn: Das Kompositum gehört zu einer Wortbildungsreihe mit *-beginn* als Zweitglied, die im WDG bei *Beginn* ebenfalls als offenes Paradigma verzeichnet ist. *Beginn* wird dort in drei verschiedenen Reihen, darunter „in Verbindung mit Jahreszeiten, z.B. *Frühlingsbeginn*“ (WDG, S. 466) und weiteren vierzehn Einzelbildungen als hochgradig kompositionsaktiv ausgewiesen. Wortbildungsprodukte dieser Reihe konkurrieren mit entsprechenden Komposita aus Bezeichnungen für die Jahreszeiten und *Anfang* als Zweitglied (*Herbstbeginn*, *-anfang*). Auch bei anderen Erstgliedern besteht diese Konkurrenz (*Ferien*-, *Semesteranfang*, *-beginn*). Das WDG und das DUW verzeichnen alle vier Komposita mit *-anfang* als selbständige Lemmata, die mit *-beginn* nicht. Das läßt zwar eine geringere Usualität der Komposita mit *-beginn* vermuten, auffällig in dem hier bestimmten Sinn werden sie dadurch jedoch nicht. Eine Blockierung der Kompositabildung durch lexikalisierte Synonyme, wie sie bei Derivaten beschrieben worden ist (vgl. z.B. Motsch 1977), tritt demnach bei der Komposition nicht unbedingt ein. Dieser Beobachtung soll hier nicht weiter nachgegangen werden. Sie zu verfolgen könnte für die funktionale Beschreibung von Komposition und Derivation nützlich sein, wobei zunächst ein für beide Wortbildungsarten gleichermaßen angemessener Modellbegriff zu erarbeiten wäre. *Winterbeginn* ist ebenfalls selbstdeutig. Das deverbale Zweitglied ist eine relationale Konstituente, die eine Interpretation des Erstgliedes als Argument des Basisverbs nahelegt.¹⁸ Der Text leistet für das Verständnis von *Igelfutter* und *Winterbeginn* nicht mehr als bei jedem anderen Substantiv, er fügt „situationsbedingte Bestimmungen“ hinzu (Seppänen 1986, S. 95) und sichert die Referenz.

Pflaumen-/Pilzzeit: Die Wortbildungseigenschaften von *Igelfutter* und *Winterbeginn* lassen sich auch bei diesen Komposita nachweisen. Sie schließen sich an eine entsprechende Wortbildungsreihe mit *Zeit* als Zweitglied in der Bedeutung 'Zeitraum' an. Die Erstglieder lexikalisierter Komposita dieser Reihe stehen ebenfalls in einem feldartigen Zusammenhang: *Beeren*-, *Obst*-, *Rosenzeit*. Weitere Gruppen von Wörtern, die sich nach den Angaben im WDG als Erstglieder mit *Zeit* verbinden, sind Bezeichnungen für Tätigkeiten (*Redezeit*), Jahreszeiten (*Frühlingszeit*), Feste (*Osterzeit*), Zeitabschnitte (*Jahreszeit*). Hinzu kommen über 80 weitere einzelne Komposita mit *Zeit* als Zweitglied in der Bedeutung 'Zeitraum', die *Zeit* in dieser Position als hochgradig wortbildungsaktiv belegen. Wie *Igelfutter* und *Winterbeginn* sind auch *Pflaumen*- und *Pilzzeit* selbstdeutig, und zwar mit Hilfe einer Relation, die zwischen Erst- und Zweitglied, ausgehend von der Bedeutung des Zweitgliedes, hergestellt wird. Sie kann als Stereotyprelation gelten (Fanselow 1981, S. 192ff): 'ein Zeitraum ist durch bestimmte Naturerscheinungen/-produkte charakterisiert bzw. wird durch diese Erscheinungen als Zeitabschnitt begrenzt'. Se-

¹⁸ Vgl. dazu die Argumentation zu *Reisebeginn* bei Gersbach/Graf (1984, S. 33). Zu den relationalen Nomina auch Boase-Beier/Toman (1986, S. 63ff).

mantisch gestützt wird die Interpretation der beiden Komposita durch das im Text vorherwähnte referenzidentische *Herbst*.

Die als unauffällig eingestufenen Komposita lassen den Schluß zu, daß neue Komposita dann keinen (oder einen geringen) Neuheitseffekt auslösen, wenn der Rezipient sie einer Wortbildungsreihe mit identischem Zweitglied zuordnen kann, ihr Zweitglied in weiteren Reihen als hochgradig kompositionsaktiv kennt und wenn er ihre Bedeutung durch diese Paradigmenstütze konstituieren kann. Vor allem lexikalisches Wissen bildet demnach die Grundlage für die Unauffälligkeit neuer Komposita, wobei dieses Wissen alternativ stärker grammatisch, semantisch oder pragmatisch bestimmt sein kann. Der Neuheitseffekt dieser Wortneubildungen bleibt offenbar deshalb so gering, weil der Rezipient bei ihrer Interpretation weitgehend ohne Textinformationen auskommt. Voraussetzung dafür ist selbstverständlich die Kenntnis der Bedeutung der jeweils nicht reihenbildenden Konstituente. Dieses Ergebnis bestätigt die Angemessenheit des Notierungsverfahrens für Komposita im WDG: Komposita offener Paradigmen brauchen im einsprachigen Bedeutungswörterbuch nicht in ihrer ganzen Vielfalt zu erscheinen, was vermutlich auch deshalb nicht optimal gelänge, weil die Rezeption entsprechender Bildungen (und gewiß auch die Produktion) zu den unbewußt ablaufenden Sprachverarbeitungsprozessen gehört (Schwarz 1992, S. 153); im Wörterbuch genügen typische Beispiele als Repräsentanten.

Einheitlich als neu bezeichnen die Probanden die Komposita *Igel-Flut* (sowohl in der Überschrift als auch bei der Wiederholung im Text), *Igel-Merkblatt*, *-desinfektion*, *-hygiene*, *-nottelefon*, *-finder*; *Fund-Igel*, *Präventiv-Igel*, die nach der oben vorgenommenen Gliederung AUFFÄLLIGE KOMPOSITA sind. Hinsichtlich der für die Erklärung der Auffälligkeit herausgearbeiteten Merkmale handelt es sich dabei allerdings um verschiedene Komposita.

In bezug auf die Serialität der Bildungen und die Kompositionsaktivität der Zweitglieder ergeben sich folgende Abstufungen: Nicht seriell gebildet sind *Igel-Merkblatt*, *-desinfektion*, *-Nottelefon*, *-finder*. Von diesen ist zunächst *Igel-Nottelefon* als Sonderfall abzuheben. Hier wird der Neuheitseffekt ganz offensichtlich dadurch ausgelöst, daß bereits die unmittelbare Konstituente *Nottelefon* eine Neubildung darstellt. Deren Interpretation ist durch die Beziehung zu *Notausgang*, *Notbeleuchtung*, *Notrufsäule* wohl gewährleistet, die des Kompositums *Igel-Nottelefon* jedoch erst durch den Gesamttext, der die gemeinte Notsituation erklärt. Als weiteres „Neuheitssignal“ kommt hinzu, daß andere Komposita mit diesen komplexen Wörtern als Zweitglied nicht lexikalisiert sind (**Theater-Notausgang*), so daß die Determination von *Nottelefon* durch *Igel* auch deshalb als neu empfunden wird. Als singuläres

textgebundenes Kompositum mit einer okkasionellen Konstituente hat *Igel-Nottelefon* einen besonders hohen Neuheitseffekt.¹⁹

Mit *Merkblatt*, *Desinfektion* und *Finder* als Zweitglied existieren ebenfalls keine Determinativkomposita, so daß die mit *Igel* gebildeten für den Rezipienten relativ isoliert, d. h. ohne Paradigmenstütze bleiben und nicht auf eine Reihe bezogen werden können. *Igel* ist zwar als Erstglied in metaphorischer Bedeutung kompositionsaktiv, vgl. *Igelfisch*, *-frisur*, *-ginster*, *-kaktus*, *-schnitt*, *-stellung* (lediglich in zwei Beispielen mit wörtlicher Bedeutung: *Igelborste*, *Igelstachel*), die Assoziation dieser Wörter wird aber im hier vorliegenden Text von vornherein durch das Thema (Tierschutz, Sichern des Überlebens der Tierart Igel, übertriebene Vorsorge), das diese Metaphorik nicht erwarten läßt, ausgeschlossen. Die neuen Komposita können deshalb ebenfalls als singular gelten.

Die Singularität der Bildungen führt zu ihrer Auffälligkeit. Der Rezipient muß einen für ihn völlig neuen Begriff bilden, d. h. genauer, dessen Bildung nachvollziehen, was er normalerweise bei der Rezeption usueller oder paradigmengestützter neuer Wörter in normgerechter Verwendung nicht in gleicher Intensität tut. Daß er dazu ohne weiteres in der Lage ist, erklärt die Psycholinguistik mit der Flexibilität natürlicher Begriffsbildungen. Danach kann sich der Mensch „mit Leichtigkeit immer wieder neue Klassenbildungen ausdenken“, und es bedarf auch „keines langwierigen Lernprozesses“ (Hoffmann 1986, S. 34), um mit solchen Bezeichnungen umgehen zu können. Dennoch bedürfen diese Wörter bei der Rezeption einer Verarbeitungsbewußtheit (Wilss 1992, S. 231). Die zusätzlich nötige Denkarbeit, ausgelöst durch den Widerspruch zwischen neuem Wort und gespeichertem lexikalischem Wissen, ist es also, die für Auffälligkeit sorgt.²⁰

Textgebunden ist lediglich *Igelmerkblatt*. *Desinfektion* und *Finder* bewirken als relationale deverbale Konstituenten Selbstdeutigkeit der Komposita mit *Igel*. Hier entsteht die Frage, ob man tatsächlich relationale Konstituenten immer als Signal für die Selbstdeutigkeit des komplexen Wortes werten kann. Wenn das zuträfe, wäre der einheitliche Neuheitseffekt von *Igeldesinfektion* und *Igelfinder* ein Nachweis dafür, daß Selbstdeutigkeit den Neuheitseffekt nicht zwingend reduziert. Das erscheint plausibel, bedarf jedoch genauerer Untersuchungen, vor allem zu Bestimmungsstücken für die Eigenschaft Selbstdeutigkeit.

¹⁹ Nach Manuskriptabschluß stieß ich auf einen weiteren Beleg für *Nottelefon*: Unter der Überschrift *Nottelefon ist jetzt einsatzbereit* berichtet die Leipziger Volkszeitung am 22.11.1994 über die Einrichtung eines telefonisch erreichbaren Wachdienstes für Notsituationen in der Nacht. Dieses Vorkommen läßt vermuten, daß für den Gebrauch des Wortes *Nottelefon* ein Bedarf entsteht, den das polyseme *Notruf* nicht klar (genug?) abdeckt.

²⁰ Groeben/Christmann (1989, S. 177); auch de Knop (1987, S. 37).

Im Unterschied zu den bisher genannten Beispielen ist *Hygiene* durchaus kompositionsaktiv. Das WDG nennt neun lexikalisierte Komposita mit *-hygiene* als Zweitglied, die allerdings keine Reihe in dem hier bestimmten Sinn darstellen, z.B. *Körper-, Schul-, Sozialhygiene*. Unter diesen befindet sich außerdem kein Kompositum mit einer Tierbezeichnung als Erstglied, so daß *Igelhygiene* trotz der Kompositionsaktivität des Zweitgliedes relativ isoliert bleibt und vom Rezipienten wahrscheinlich ebensoviel Rezeptionsaufwand wie die vorher genannten Komposita erfordert. Daraus läßt sich ableiten, daß die Kompositionsaktivität des Zweitgliedes allein, das heißt ohne Reihenausprägung, den Neuheitseffekt nur wenig reduzieren kann.

Für *Fund-Igel* und *Präventiv-Igel* gilt das Gesagte analog, allerdings ist *Igel* als Zweitglied kompositionsinaktiv. Es besteht offenbar, zumindest in nicht-fachlicher Kommunikation, kein sachlicher Bedarf nach einer Subklassifizierung der Tierart Igel. Die hier durch die Determinativkomposita vorgenommene Differenzierung des Begriffs *Igel* erwächst Textgestaltungsprinzipien, hat folglich nur für diesen einen Text Relevanz. Entsprechend ist die Interpretierbarkeit textgebunden.

Analogieschlüsse des Rezipienten bei der Verarbeitung der beiden Wörter sind nur bedingt möglich. *Fund-Igel* kann in Beziehung zu *Fundsache, -gegenstand* gebracht werden, wobei jedoch die Verbindung von *Fund-* mit Tierbezeichnungen nicht belegt und begrifflich kaum relevant ist. Für das Kompositum *Präventiv-Igel* läßt sich kein passendes Vorbild finden. Bildungen wie *Präventivmaßnahme, -medizin* haben andere Wortbildungsbedeutungen. In *Präventiv-Igel* wird eine ganze Aussage ironisierend kondensiert – darauf weisen auch die distanzierenden Anführungszeichen im Text hin – und im Anschluß an den Wortgebrauch syntaktisch ausgeführt: 'Die Igel werden vorsorglich abgegeben, obwohl sie gesund sind.'

Das einzige Exemplar der auffälligen Komposita, das man bedingt an eine Wortbildungsreihe anschließen könnte, ist *Igel-Flut*, vgl. *Bücher-, Farben-, Haar-, Licht-, Locken-, Menschenflut*; bedingt deshalb, weil sich die Erstglieder kaum zu einer semantischen Gruppe zusammenfassen lassen und auch deshalb, weil Tierbezeichnungen als Erstglied bei den lexikalisierten Belegen nicht vorkommen. Es ist in jedem Fall aber von einem wortbildungsaktiven Zweitglied auszugehen. Da der Referent der Wortneubildung, eine plötzlich vorhandene große Menge von Igel, allerdings höchst ungewöhnlich ist und nicht als alltäglich angesehen werden kann, wird das Wort als unbekannt eingestuft.²¹ Interessant ist, daß alle Probanden das Wort auch bei seinem zweiten Auftreten im Text einheitlich als neu bezeichnen. Wörtliche Wiederholung im Text reduziert den Neuheitseffekt demnach – zumindest bei diesem Wort – nicht.

²¹ Dazu ausführlich Müller-Bollhagen (1985).

Die auffälligen Wortneubildungen sind singular und daher die eigentlich innovativen Prägungen. Sie bedürfen der bewußten Rezeption, lassen sich nur unter Einbeziehung wortexterner Informationen aus dem allgemeinen semantischen Rahmen des Textes erschließen und sind – jedenfalls gilt das für die hier untersuchten Komposita – nicht „auf mental verfestigte Strukturmuster [zu] beziehen“ (Wilss 1986, S. 279). Ihre Textbindung und die nur lockeren paradigmatischen Bezüge führen zu einem deutlichen Neuheitseffekt.

Als INDIFFERENTE KOMPOSITA gelten die von den Probanden uneinheitlich beurteilten Wortneubildungen. Mehr als 50% der Probanden halten *Igelauffangstelle*, *-freilassung*, *-bericht*, *-herberge*, *-bewegung*, *-sammler* für neue Wörter, *Igelspezialist* und *-spezialistin* werden dagegen von weniger als 50% der Probanden als neu bestimmt. Da alle diese Komposita ausnahmslos für den Text gebildet und damit nachweislich neu sind, überrascht das Ergebnis insofern, als das Urteil 'bekannt' über neue Wörter für teilweise groteske Sachverhalte doch unerwartet ist. Es wäre eher mit dem einheitlichen Urteil 'neu' zu rechnen gewesen. Erklärungsmöglichkeiten bieten wiederum die Wortbildungseigenschaften der beurteilten Wörter.

Daß die Wörter *Igelspezialist* und *Igelspezialistin* von der Mehrzahl der Probanden für lexikalisiert gehalten werden, liegt zum einen an der gut ausgebauten Reihe mit *-spezialist/-in*, die, ähnlich wie die von *-zeit*, nahezu beliebig erweiterbar zu sein scheint, wobei offenbar kaum mit Restriktionen in bezug auf die Semantik der Erstglieder gerechnet werden muß.²² Zum anderen ist *Spezialist* ein relationales Substantiv, dessen Bedeutung eine Ergänzung fordert: *Spezialist für etwas*. Das Kompositum ist demzufolge selbstdeutig. Hinzu kommt, daß das bezeichnete Spezialistentum auch sachlich wahrscheinlich ist. Diese „Wahrscheinlichkeit“ wird hier durch den Gesamttext aufgebaut, die Bildung wäre aber auch in anderen Texten akzeptabel. Beide Wörter werden offenbar deshalb von der Mehrzahl der Probanden als bekannt eingestuft, weil ihre Serialität, die auch bei der hochproduktiven Movierung vorliegt, als „Bekanntheitsindikator“ über die Textbindung als „Neuheitsindikator“ dominiert.

Bei *-auffangstelle*, *-freilassung*, *-bericht*, *-bewegung*, *-sammler* stellt sich dieses Dominanzverhältnis gerade umgekehrt dar. Auch hier handelt es sich aufgrund der grammatisch-semantisch bedingten Ergänzungsbedürftigkeit der deverbalen Basen um relationale Zweitglieder. Daher sind die Komposita selbstdeutig und müßten eigentlich als bekannt eingestuft werden können. Dennoch entscheidet sich bei ihnen die Mehrheit der Probanden für 'neu'. Die Gründe liegen wohl vor allem darin, daß analoge Kompositionsreihen nicht entwickelt sind und dem Rezipienten deshalb keine aktivierbaren „Denk-

²² Dieses Beispiel weist wie auch schon *Zeit* darauf hin, daß innerhalb der Komposita Reihen verschiedener Abstraktionsstufen unterschieden werden sollten.

muster“ zur Verfügung stehen. Die Zweitglieder dreier Komposita sind zwar hochgradig kompositionsaktiv, vgl. *-bericht*: *Bank-, Börsen-, Reise-, Tatsachenbericht*; *-bewegung*: *Friedens-, Frauen-, Studentenbewegung*; *-sammler*: *Briefmarken-, Käfer-, Kunst-, Lumpen-, Pilzsammler*; in allen Gruppen, von *Schmetterlings-* und *Käfersammler* abgesehen, fehlen jedoch Komposita mit Tierbezeichnungen als Erstglied. Entsprechende Begriffe werden als dauerhafte Fixierungen nicht gebraucht. Insofern könnte man die Kombinationen mit *Igel-* als Abweichungen von den lexikalisierten Paradigmen ansehen, und das Abweichen für den trotz wortbildungsaktiver Zweitglieder erhöhten Neuheitseffekt verantwortlich machen. Wie schon bei *Desinfektion* und *Finder* zeigt sich auch an diesen Zweitgliedern, daß ihre Relationalität zwar die Selbstdeutigkeit weitgehend sichert, den Neuheitseffekt jedoch nicht aufzuheben vermag. Bei *Igelbewegung* kann zudem wegen der Polysemie des Wortes nicht ohne weiteres von Selbstdeutigkeit ausgegangen werden.

Nicht kompositionsaktiv sind *-auffangstelle* und *-freilassung*. Ihre Singularität sorgt dafür, daß sich bei der Mehrzahl der Probanden ein Neuheitseffekt einstellt, dennoch führen wiederum Selbstdeutigkeit aufgrund relationaler Konstituenten sowie Textinformationen bei einigen Probanden zu dem Urteil ‘bekannt’.

Daß schließlich auch *Igelherberge* von manchen als bekannt eingestuft wird, läßt sich auf die gegenseitige Überlagerung von Singularität, möglicher Interpretation als Wortanalogie zu *Jugendherberge* und Selbstdeutigkeit zurückführen. Bei diesem Kompositum, das im Text als referenzidentische Bezeichnung zu *Igel-Auffangstelle* fungiert, sorgt zusätzlich eventuell dessen Vorerrwähntheit für den Abbau des Neuheitseffektes.

5. Fazit

Die Analyse von Rezipientenurteilen über den Neuheitseffekt von Wortneubildungen läßt in bezug auf das bewertete Material durchaus plausible Antworten auf die eingangs formulierten Fragen zu.²³

Der Neuheitseffekt von Wortneubildungen wird in starkem Maße von den Merkmalen Serialität bzw. Singularität bestimmt. Dafür sprechen die jeweils einheitlichen Urteile über serielle und singuläre Wortneubildungen. Einheitlich als bekannt gelten serielle Bildungen, einheitlich als neu singuläre aus völlig kompositionsinaktiven Konstituenten. In einem Zwischenbereich kann die

²³ Ergänzend zu der hier vorgelegten systemorientierten Untersuchung des Neuheitseffektes werden in Barz (1996) – primär rezipientenorientiert – Wissensstrukturen und Sprecherurteile aufeinander bezogen.

Selbstdeutigkeit singularer Bildungen den Neuheitseffekt partiell reduzieren, jedoch nicht völlig aufheben.

Der Neuheitseffekt ist eine graduelle Größe. Ein Kompositum löst den geringsten Neuheitseffekt aus, wenn es eine Reihe von lexikalisierten Komposita mit jeweils identischem Zweitglied, identischer Wortbildungsbedeutung und semantisch „verwandtem“ Erstglied ergänzt und wenn es selbstdeutig ist. Am stärksten ist der Neuheitseffekt bei singularen, nicht selbstdeutigen Wortneubildungen. Neuheitseffekt und wortbildungsparadigmatische Einbindung einer Wortneubildung verhalten sich demnach zueinander umgekehrt proportional.

Ob diese primär sprachsystematisch gewonnenen Verallgemeinerungen auch auf andere Arten komplexer Wörter übertragbar sind, kann hier nicht mehr ermittelt werden. Ebenso müssen Fragen zur gegenseitigen Überlagerung der den Neuheitseffekt abbauenden Faktoren, insbesondere zur Textbindung von Wortneubildungen, aber auch zum Einfluß sozialen und kulturellen Wissens auf deren Dekodierung, unbeantwortet bleiben. Sie sollten für komplexe Wörter verschiedener Wortbildungsarten und verschiedener Wortarten erneut gestellt werden.

6. Literatur

- Barz, Irmhild (1989): Die Relevanz des Merkmals Reihenbildung für die Morphemklassifikation. In: Heimann, Sabine/Lerchner, Gotthard/Müller, Ulrich/Reiffenstein, Ingo/Störmer, Uta (Hg.): Soziokulturelle Kontexte der Sprach- und Literaturentwicklung. Festschrift für Rudolf Große zum 65. Geburtstag. Stuttgart. S. 317-325.
- Barz, Irmhild (1996): Die Neuheit von Wörtern im Urteil der Sprecher. In: Hertel, Volker/Barz, Irmhild/Metzler, Regine/Uhlig, Brigitte (Hg.): Sprache und Kommunikation im Kulturkontext. Beiträge zum Ehrenkolloquium aus Anlaß des 60. Geburtstages von Gotthard Lechner. Frankfurt a.M./Berlin/Bern/New York/Paris/Wien. S.299-313
- Boase-Beier, Jean/Toman, Jindrich (1986): Komposita im Text. Überlegungen zur Unterscheidung zwischen grammatischem und textuellem Wissen. In: Zeitschrift für Literaturwissenschaft und Linguistik, 64, S. 61-72.
- Coseriu, Eugenio (1988): Sprachkompetenz. Grundzüge der Theorie des Sprechens. Bearb. u. hrsg. v. Heinrich Weber. Tübingen.
- DGW (1993) – Duden. Das große Wörterbuch der deutschen Sprache in acht Bänden. 2., völlig neu bearb. Aufl. Hrsg. u. bearb. vom Wiss. Rat u. d. Mitarbeitern d. Dudenredaktion unter d. Leitg. v. Günther Drosdowski. Mannheim/Leipzig/Wien/Zürich.
- DUR (1991) – Duden. Rechtschreibung der deutschen Sprache. 20., völlig neu bearb. u. erw. Auflage. Mannheim/Leipzig/Wien/Zürich.
- DUW (1989) – Duden. Deutsches Universalwörterbuch. 2., völlig neu bearb. u. stark erw. Aufl. (1. Aufl. 1983). Hrsg. u. bearb. vom Wiss. Rat u. d. Dudenred. unter Leitg. v. Günther Drosdowski. Mannheim/Leipzig/Wien/Zürich.

- Erben, Johannes (1981): Neologismen im Spannungsfeld von System und Norm. In: *Logos Semantikos* 5. Berlin/New York/Madrid. S. 35-43.
- Fandrych, Christian/Thurmair, Maria (1994): Ein Interpretationsmodell für Nominalkomposita: Linguistische und didaktische Überlegungen. In: *Deutsch als Fremdsprache*, 31, S. 34-45.
- Fanselow, Gisbert (1981): *Zur Syntax und Semantik der Nominalkomposition*. Tübingen.
- Fix, Ulla (1987): Erwartung in der Linguistik – Anmerkungen zum Verhältnis von Erwartung, Norm und Adäquatheit. In: *Beiträge zur Erforschung der deutschen Sprache* 7. S. 62-80.
- Fleischer, Wolfgang (1983): Dynamik in Wortbildung und Wortschatz der deutschen Gegenwartssprache: Okkasionalismen und Neologismen. In: *Germanistisches Jahrbuch DDR-UVR*. Hrsg. v. Deutschlektorat beim KIZ der DDR in Budapest II. S. 41-53.
- Fleischer, Wolfgang/Barz, Irmhild (1995): *Wortbildung der deutschen Gegenwartssprache*. Unter Mitarb. v. Marianne Schröder. Tübingen.
- Gersbach, Bernhard/Graf, Rainer (1984): *Wortbildung in gesprochener Sprache*. 1. Tübingen.
- Grimm, Ursula (1991): *Lexikalisierung im heutigen Englisch am Beispiel der -er Ableitungen*. Tübingen.
- Groeben, Norbert/Christmann, Ursula (1989): Textoptimierung unter Verständlichkeitsperspektive. In: Antos, Gerd/Krings, Hans P. (Hg.): *Textproduktion. Ein interdisziplinärer Forschungsüberblick*. Tübingen. S. 165-196.
- Große, Rudolf/Neubert, Albrecht (1982): Sprachwandel in soziolinguistischer Sicht. Thesen. In: *Sitzungsberichte der Akademie der Wissenschaften der DDR* 10 G. S. 5-14.
- Handler, Peter (1993): *Wortbildung und Literatur. Panorama einer Stilistik des komplexen Wortes*. Frankfurt a.M. u.a.
- Heller, Klaus u.a. (1988): *Theoretische und praktische Probleme der Neologismenlexikographie. Überlegungen und Materialien zu einem Wörterbuch der in der Alltagssprache der DDR gebräuchlichen Neologismen*. In: *Linguistische Studien A* 184. Berlin.
- Helfrich, Uta (1993): *Neologismen auf dem Prüfstand. Ein Modell zur Ermittlung der Akzeptanz französischer Neologismen*. Wilhelmsfeld.
- Herberg, Dieter (1988): Neologismen – lexikologisch und lexikographisch betrachtet. In: *Sprachpflege*, 37, S. 109-112.
- Heringer, Hans Jürgen (1984): *Wortbildung: Sinn aus dem Chaos*. In: *Deutsche Sprache*, S. 1-13.
- Hörmann, Hans (1987): *Einführung in die Psycholinguistik*. Darmstadt.
- Hoffmann, Joachim (1986): *Die Welt der Begriffe*. Berlin.
- Knop, Sabine de (1987): *Metaphorische Komposita in Zeitungüberschriften*. Tübingen.
- Lerchner, Gotthard (1984): *Sprachform von Dichtung. Linguistische Untersuchungen zu Funktion und Wirkung literarischer Texte*. Berlin/Weimar.
- Matussek, Magdalena (1994): *Wortneubildung im Text*. Hamburg.

- Motsch, Wolfgang (1977): Ein Plädoyer für die Beschreibungen von Wortbildungen auf der Grundlage des Lexikons. In: Brekle, Herbert Ernst/Kastovsky, Dieter (Hg.): *Perspektiven der Wortbildungsforschung*. Bonn. S. 180-202.
- Motsch, Wolfgang (1995): Semantische und pragmatische Aspekte der Wortbildung. In: Popp, Heidrun (Hg.): *Deutsch als Fremdsprache. An den Quellen eines Faches. Festschrift für Gerhard Helbig zum 65. Geburtstag*. München. S. 513-530.
- Müller, Wolfgang (1987): „Schlamm Schlacht“. Schon gehört? Ein Desiderat: Das deutsche Neologismenwörterbuch. In: *Sprache und Literatur in Wissenschaft und Unterricht*, 18, S. 82-90.
- Müller-Bollhagen, Elgin (1985): Überraschungsfrikadelle mit Chicorée Gemüse und Folienkartoffel. In: Koller, Erwin/Moser, Hans (Hg.): *Studien zur deutschen Grammatik. Johannes Erben zum 60. Geburtstag*. Innsbruck. S. 225-237.
- Munske, Horst Haider (1990): Über den Wandel des deutschen Wortschatzes. In: Besch, Werner (Hg.): *Deutsche Sprachgeschichte. Grundlagen, Methoden, Perspektiven. Festschrift für Johannes Erben zum 65. Geburtstag*. Frankfurt a.M. u.a. S. 387-402.
- Olsen, Susan (1986): *Wortbildung im Deutschen. Eine Einführung in die Theorie der Wortstruktur*. Stuttgart.
- Ortner, Hanspeter/Ortner, Lorelies (1984): *Zur Theorie und Praxis der Kompositafor-schung*. Tübingen.
- Plank, Frans (1981): *Morphologische (Ir-)Regularitäten. Aspekte der Wortstrukturtheorie*. Tübingen.
- Rickheit, Gert/Strohner, Hans (1989): *Textproduktion*. In: Antos, Gerd/Krings, Hans P. (Hg.): *Textproduktion. Ein interdisziplinärer Forschungsüberblick*. Tübingen. S. 220-256.
- Schröder, Marianne (1985): Überlegungen zur textorientierten Wortbildungsforschung. In: *Linguistische Studien A 123*. Berlin. S. 69-94.
- Schwarz, Monika (1992): *Einführung in die kognitive Linguistik*. Tübingen.
- Seppänen, Lauri (1986): Leere Bedeutungen, falsche Namen? Okkasionelle (und andere) Komposita aus referenzsemantischer Sicht. In: *Zeitschrift für germanistische Linguistik*, 14, S. 82-97.
- Siebert-Ott, Gesa (1991): *Wortbildung und sprachliche Kreativität. Wissenschaftliche Begleitung des Sprachservice-Telefons an der Universität Düsseldorf*. In: *Sprache und Literatur in Wissenschaft und Unterricht*, 67, S. 83-99.
- Stein, Achim (1993): *Nominalgruppen in Patentschriften: Komposita und prädikative Nominalisierungen im deutsch-französischen Vergleich*. Tübingen.
- Stepanowa, Maria D./Fleischer, Wolfgang (1985): *Grundzüge der deutschen Wortbildung*. Leipzig.
- WDG (1961-1977) – *Wörterbuch der deutschen Gegenwartssprache*. 1-6. Klappenbach, Ruth/Steinitz, Wolfgang (Hg.). Berlin.
- Wellmann, Hans (1993): Die Bedeutungen der Nominalkomposita. In: Ders. (Hg.): *Synchrone und diachrone Aspekte der Wortbildung im Deutschen*. Heidelberg. S. 147-168.
- Wellmann, Hans (1995): *Die Wortbildung*. In: Drosdowski, Günther u.a. (Hg.): *Duden 4. Grammatik der deutschen Gegenwartssprache*. Mannheim/Wien/Zürich. S. 399-539.

- Wildgen, Wolfgang (1982): Makroprozesse bei der Verwendung nominaler Ad-hoc-Komposita im Deutschen. In: *Deutsche Sprache*, 10, S. 237-257.
- Wilss, Wolfram (1985): Zur Produktion und Rezeption von Wortbildungserscheinungen. In: *Zeitschrift für germanistische Linguistik*, 13, S. 278-294.
- Wilss, Wolfram (1986): Wortbildungstendenzen in der deutschen Gegenwartssprache. Tübingen.
- Wilss, Wolfram (1992): Schematheorie und Wortbildung. In: *Deutsch als Fremdsprache*, 29, S. 230-234.

Cyril Belica

Statistische Analyse von Zeitstrukturen in Korpora¹

1. Einführung

Durch die stets wachsenden Textmengen, die den Linguisten für ihre empirische Sprachforschung in maschinenlesbarer Form zur Verfügung gestellt werden, wird auch die Entwicklung neuer Untersuchungsmethoden beschleunigt. Den gemeinsamen formalen Kern vieler dieser Methoden bilden mathematisch-statistische Aussagen über den Sachverhalt von Zeichenketten im Text. In einfachen Fällen werden diesen Äußerungen Größen wie z.B. Wortformen- oder Kollokationshäufigkeiten, im allgemeinen ein multidimensionaler Raum komplexer statistischer Charakteristiken zugrundegelegt. Durch die Interpretation dieser exakten mathematisch-statistischen Größen formuliert der Mensch Aussagen über die im Text belegten sprachlichen Phänomene, oder – wenn er einen Grund hat zu glauben, daß der Text auch einen anderen, allgemeineren Sprachausschnitt repräsentiert – extrapoliert er sie auch über die Grenzen der zugrundeliegenden Sprachstichprobe hinaus. Ist eine der Dimensionen des analysierten Raumes die Zeit, so kann das Verfahren als Analyse der zeitlichen Entwicklung von Sprachphänomenen bezeichnet werden – ein Begriff, der zweifelsohne dem Gebiet der Neologie zuzuordnen ist.

In der traditionellen computergestützten empirischen Forschung von Neologismen formuliert der Mensch, geführt von der eigenen Intuition, eine A-priori-Hypothese und versucht, diese für konkrete Sprachphänomene zu verifizieren bzw. zu widerlegen, indem er eine relevante, vom Rechner erstellte Synopse des Sachverhalts manuell analysiert und interpretiert. Die A-priori-Hypothese wurzelt fest in der Entscheidung, nach dem Konkreten zu suchen, und in der damit verbundenen Hoffnung, daß sich diese Suche – zielorientiert – produktiv auswirkt. Als Determinante dieser Vorgehensweise sind somit konkrete, aus der sprachlichen Intuition stammende Erwartungen zu erkennen. Als Ergebnis erhofft man sich die Vertiefung bzw. Abschwächung des Glaubens an die Ausgangshypothese. Dabei werden en passant die eigenen manuell-analytischen Methoden optimiert, und die sprachliche Intuition wird verfeinert.

Ein Beispiel für diese Arbeitsweise bietet ein Lexikograph, der – gewisse Veränderungen im alltäglichen Gebrauch des Wortes *Erneuerung* in der Wendezeit vermutend – sich vom Computer z.B. eine zeitlich sortierte KWIC-Liste

¹ Ist auch erschienen in: LDV-INFO 8, Institut für deutsche Sprache. Mannheim, 1996.

lt und durch Betrachtung der Kollokationen des Wortes versucht, den Sucher nachzugehen.

Diese Technologie hat zwei Mängel: Im ersten Arbeitsschritt – bei der Wahl der zu verifizierenden Hypothesen – ist der Mensch völlig auf eigenes Gespür angewiesen, er wird nicht vom Computer durch Hinweise auf eventuelle weitere interessante Sachverhalte im Text unterstützt. Ebenso einsam kämpft sich der Mensch im zweiten Arbeitsschritt durch die unter Umständen hoffnungslos wachsende Flut von Bildschirm- oder Druckerausgaben.

Alternativ wird versucht, eine sprachgebundene Hypothese in eine formale mathematisch-statistische Bedingung ohne Bezug auf konkrete Sprachelemente umzuformulieren. Es wird dem Computer überlassen, diese Bedingung für ausgewählte Sprachelemente zu verifizieren. Zum Beispiel werden eine statistische Analyse und eine automatische Klassifizierung von Kollokationen des Wortes *Erneuerung* veranlaßt. Auch diese Methode birgt den Nachteil in sich, daß sie im ersten Arbeitsschritt von der menschlichen Intuition abhängig ist. Das vorgeschlagene Verfahren ist eine Verallgemeinerung des zweiten Falles. Die in einen Algorithmus umgesetzte und programmtechnisch implementierte Hypothese wird in dem *gesamten* maschinell gespeicherten Sprachausschnitt verifiziert. In den akzeptierten Fällen ist dann auf das gesuchte linguistische Phänomen zu schließen. Hat die Hypothese ein geeignetes Skalenniveau, so kann vom Computer eine Sekundäranalyse akzeptierter Fälle vorgenommen werden, um weitere Zusammenhänge aufzuspüren. Der Ausgangspunkt dieser Methode ist also die mathematisch-statistische Approximation und algorithmische Implementierung einer sprachgebundenen Fragestellung. Als Resultat werden konkrete Sprachelemente identifiziert, die der Fragestellung entsprechen.

Ein ähnliches Modell ist in Programmen zur automatischen Filterung von Texten für Monitorkorpora implementiert. Mit zunehmender Verfügbarkeit maschinenlesbarer Texte wird die Gefahr groß, daß Monitorkorpora durch Ballasttexte (z.B. aus „billigen“ Satzbindern) übersättigt werden, d.h. durch Texte, die relativ zu ihrem Umfang oder im Vergleich zu den bereits vorhandenen Texten deutlich ohne zusätzliche verwertbare Information sind und nutzlos die Computerressourcen auslasten würden. Die Filterprogramme versuchen, die „interessanten“ Texte zu erkennen und möglichst alle Texte zu tilgen, deren Informationsgehalt in klarem Mißverhältnis zu ihrem Umfang steht. Auch Programme zur Überwachung externer Textquellen und zur automatischen Aktivierung von Aufzeichnungseinrichtungen von Nachrichtendiensten bedienen sich eines solchen Konzepts.

Die Zielsetzung unserer Arbeit war die Entwicklung und Implementierung einer Methode zur Berechnung, Analyse und Auswertung zeitrelevanter statistischer Parameter von Texten, die als Indikatoren des Begriffs „Neo-

logismus“ verwendet werden können. Die Bestimmung optimaler Indikatoren stand in dieser Etappe nicht im Mittelpunkt unseres Interesses – für die Verifizierung dieser Methode haben wir die zeitliche Verteilung von Wortformen- und Kollokationshäufigkeiten analysiert. Dadurch hofften wir, auf dem festgesetzten Signifikanzniveau Fälle zu isolieren, deren zeitliche Verteilung im Text der „natürlichen“ Verteilung widerspricht und auf Neologismen²-Kandidaten hinweist.

2. Das Problem

Die Erarbeitung der in der Einführung skizzierten Methode wurde durch ein spezielles Lexikographie-Projekt am IDS, „Wendewortschatz“, angeregt und an den Daten dieses Projekts, dem „Wendekorpus“, feingeschliffen und exemplifiziert.

Das Projekt soll die Neologismen und die Neosemantismen der Wendezeit zutage fördern. Das zugrundegelegte Textkorpus (ca. 4 Mio. laufende Wörter) umfaßt Texte, die in der ehemaligen DDR produziert wurden, sowie solche aus den alten Bundesländern. Es ist in insgesamt 7 (Ost-) bzw. 6 (West-) Zeitphasen gegliedert, die durch die Lexikographen vorgegeben und aufgrund politischer Bewertung der Zeitgeschichte gewonnen wurden.

Der Anlaß unserer Arbeit war die Erfahrung, daß die Fülle des Materials, das im Recherchesystem COSMAS online zur Verfügung stand, das systematische Durchdringen sehr aufwendig, wenn nicht unmöglich macht. Dagegen stand die Ansicht, daß dort, wo der Wald vor lauter Bäumen nicht gesehen wird, der Computer ersatzweise strukturieren müsse.

3. Die Methode

3.1 Im Allgemeinen

Wie in der Einführung bereits angedeutet, ist der Kern dieser Methode die mathematisch-statistische Analyse zeitlicher Verteilung von Häufigkeiten verschiedener Textmerkmale. Nun stellt sich die berechtigte Frage, was soll *linguistisch* mit diesen Berechnungen erreicht werden?

² Unter Neologismen verstehen wir hier hilfsweise auch Neosemantismen, Ad-hoc-Wortbildungen, Okkasionalismen u.ä.

In diesem Kapitel zeigen wir, daß es möglich ist

- „Auffälligkeiten“ in den Texten aufzuspüren,
- die Signifikanz solcher Auffälligkeiten zu beurteilen,
- Hinweise auf die Interpretation dieser Auffälligkeiten und auf ihre Einordnung in das Sprachsystem zu errechnen.

An dieser Stelle wollen wir noch offen lassen, was eigentlich unter diesen „Auffälligkeiten“ zu verstehen ist. Die Antwort auf diese Frage wird weiter unten ersichtlich.

Und ohne Argumentation möchten wir auch die Vermutung stehen lassen, daß diese oder sehr ähnliche Berechnungen noch ein weiteres linguistisches Anliegen betreffen, daß sie nämlich die Prädiktion von Sprachentwicklungstendenzen erlauben.

3.2 Im Einzelnen

3.2.1 Das Korpus und die Zeit

Die erste Voraussetzung für Anwendung der Methode ist das Vorhandensein eines geeigneten Textkorpus C . Die Eignung des Korpus besteht in erster Linie darin, daß sich jeder Korpustext aufgrund seiner Datierung einer von N disjunkten Phasen P_i , d.h. einem Sampling-Intervall, zuordnen läßt. Nun stellt sich die Frage, welche weiteren Eigenschaften diese Zeitgliederung haben soll, und wie sie festgelegt wird.

Den Ausgangspunkt bei der Festlegung der Zeitgliederung bilden externe, problemorientierte Kriterien. Es bleibt dem Linguisten überlassen, diejenige Granularität der Zeit festzusetzen, die am besten der Art des zu untersuchenden Phänomens entspricht. Es müssen jedoch einige – durch die Natur der Methode vorgegebene – interne Kriterien beachtet werden. So ist zu überprüfen, ob die festgelegten Sampling-Intervalle diesen Kriterien entsprechen. In den folgenden Absätzen werden diese Einschränkungen kurz beschrieben.

3.2.1.1 Textmenge

Zunächst ist zu berücksichtigen, daß die Sampling-Intervalle in Abhängigkeit von den *verfügbaren Textmengen* (Stichproben) zu sehen sind. Je schmaler die Textmenge, desto unwahrscheinlicher werden Unregelmäßigkeiten mit akzeptabler statistischer Signifikanz entdeckt. Geringe Textmengen innerhalb eines Sampling-Intervalls weisen wahrscheinlich einen geringen Sättigungsgrad auf.

So besteht die Gefahr, daß nicht die zeitrelevanten, sondern die textspezifischen Eigenschaften der Texte gemessen werden.

3.2.1.2 Dispersivität

Unter dem Aspekt der Analyse von Zeitstrukturen kann zwecks Datierung der Texte das jeweilige Erscheinungsdatum nur als Hilfsangabe angesehen werden. Als maßgeblicher Parameter gilt die Zeitspanne, für die die im Text erfaßte *Sprache* charakteristisch ist. Diese Zeitspanne kann durch zwei Angaben spezifiziert werden: durch ihre Position in der Zeit und durch ihre Dauer, hier Dispersivität genannt. So ist z.B. einem Artikel aus einer Tageszeitung das Erscheinungsdatum und die Dispersivität von ein paar Tagen, einem Roman dagegen die Entstehungszeit und Dispersivität von einigen Jahren zuzuweisen.

Es ist offensichtlich, daß die Zeitintervalle, die bei der jeweiligen Untersuchung über der Zeitachse gebildet werden, also das Sampling, mit Rücksicht auf diese Dispersivität der einzelnen Texte festgesetzt werden müssen. Je kleiner die Dispersivität, desto feiner können die Zeitintervalle gewählt werden. Es bedarf einer sorgfältigen Prüfung, ob die Dispersivität das beabsichtigte Sampling trägt. So macht es z.B. wenig Sinn, bei Texten, die u.a. aus Wochenzeitschriften bezogen werden, eine tageweise Analyse zugrundezulegen. Als Faustregel gilt, daß die charakteristische Dispersivität der zu analysierenden Texte im Vergleich mit dem Sampling deutlich kleiner sein muß.

3.2.1.3 Homogenität

Im Idealfall ziehe man zur Analyse von Zeitstrukturen in Korpora nur solche Texte heran, die sich in dem o.g. multidimensionalen Raum ausschließlich durch ihre Position auf der Zeitachse voneinander unterscheiden, aber in anderen Parametern, wie z.B. Textsorte, Sprachstil, Textgenre, Themenbereich, Autorenprofil usw. identisch sind. Ein Korpus mit dieser Eigenschaft nennen wir *homogen* in bezug auf das beabsichtigte Sampling. In der Praxis ist jedoch der aus der internen Inhomogenität stammende Einfluß auf die statistische Analyse von Zeitstrukturen zu erkennen und durch Veränderungen in der Korpuskomposition zu eliminieren. Die statistischen Eigenschaften, an denen die Homogenität gemessen wird, können in Abhängigkeit vom Untersuchungsziel definiert werden. Es ist z.B. möglich, lexikalische, morphosyntaktische sowie syntaktische Merkmale auszuwählen, von denen man annehmen kann, daß ihre Verteilung in Hinsicht auf das Untersuchungsziel entlang der zu prüfenden Dimensionen invariant ist.

Bislang haben wir die Verteilung von verschiedenen Kollokationen ausgewählter Funktionswörter auf ihre Eignung als Invarianten untersucht. Wir haben

jedoch vor zu überprüfen, ob verallgemeinerte Verifizierungsverfahren entwickelt werden können, die für verschiedene Dimensionen und möglichst unabhängig vom Untersuchungsziel den Homogenitätsgrad in Hinsicht auf das vorgenommene Sampling bestimmen.

3.2.2 Approximation der linguistischen Fragestellung

In den vorherigen Abschnitten haben wir Arbeitsschritte beschrieben, die die Vorbedingungen der Methode sichern. Mit einem solchen für unsere Zwecke geeigneten Korpus ausgestattet, müssen wir nun unsere linguistische Fragestellung: „Was gibt es in diesem Korpus Interessantes unter dem Aspekt der Neologie?“ in statistisch faßbare und meßbare Merkmale des Textes umsetzen, die als Indikatoren von Neologismen dienen können. Der Erfolg des vorgeschlagenen Verfahrens hängt davon ab, wie gut dies gelingt. Da jeder Indikator nur eine partielle Abbildung des Begriffs „Neologismus“ leistet, ist dieser Schritt in der Methode als ein austauschbarer LEGO-Stein konzipiert, was wiederum eine iterative Umsetzung unseres linguistischen Vorhabens in die Sprache der Statistik möglich macht. Mehrere Indikatoren können nach einem empirischen Gewichtungsmo­dell in einen Neologismus-Vektor zusammengefaßt und als eine Einheit betrachtet und ausgewertet werden.

Es hat sich gezeigt, daß bereits die Auswertung einfacher Statistiken wie „Häufigkeit von Wortformen“, „Häufigkeit von vorkommenden rechten und linken Nachbarn zu Wortformen“ aussagekräftige Ergebnisse liefern. Wir verweisen auf die Beispiele unter 5.

3.2.3 Berechnung der Indikatoren

Nachdem die Approximation der linguistischen Fragestellung festgelegt ist, werden die Indikatoren berechnet. Als Ergebnis erhält man eine Matrix M von $N \times P$ Angaben, wobei P die Anzahl der Sampling-Intervalle und N die Gesamtzahl des Vorkommens des Indikators im Korpus ist. Es wird zum Beispiel eine Matrix mit so vielen Zeilen erstellt, wie verschiedene Wortformen im Korpus vorkommen. Die Anzahl der Spalten entspricht der Anzahl von Sampling-Intervallen. Die Zellen der Tabelle geben dann die Intervallhäufigkeiten der Wortformen an.

3.2.4 Das Prüfverfahren

Als nächster Schritt wird mit Hilfe von gängigen statistischen Verfahren geprüft, ob man unter den in Matrix M erfaßten Beobachtungsdaten solche finden kann, deren Verteilung über die Sampling-Intervalle nur mit einer empirisch nicht akzeptablen Irrtumswahrscheinlichkeit als „normal“ bezeichnet werden kann.

3.2.4.1 Die Nullhypothese

Für diese Zwecke wird die Nullhypothese H_0 postuliert: „Die beobachtete Statistik weist entlang der Zeitachse keine Besonderheiten auf.“ Zum Beispiel entsprechen die Frequenzen aller Wortformen in allen Sampling-Intervallen den Intervallgrößen. Erweist sich die Nullhypothese im nächsten Schritt als falsch, so wird diese abgelehnt und die Alternativhypothese H_1 angenommen: „Diese Verteilung ist nicht mehr durch Zufall zu erklären.“

3.2.4.2 Auswahl des statistischen Tests und des Signifikanzniveaus

Nun ist der geeignete statistische Test für die Überprüfung der Nullhypothese auszuwählen. Prinzipiell wählt man den Test aus, dessen statistisches Modell am besten den Eigenschaften der ausgewählten Statistik entspricht. In den Analysen der Häufigkeitsverteilung von Wortformen und Kollokationen haben wir für die Verifizierung der Nullhypothese den χ^2 -Test benutzt:

$$\chi_y^2 = \frac{N \left(\left| AD - BC \right| - \frac{N}{2} \right)^2}{(A + B)(C + D)(A + C)(B + D)} \quad df = 1$$

für die Häufigkeiten der Wortform bzw. Kollokation W_j im Sampling-Intervall P_i aus der Kontingenztabelle:

	W_j	W_{res}	Total
P_i	A	B	A+B
P_{res}	C	D	C+D
Total	A+C	B+D	N

Als Akzeptanzkriterium für eine weitere Bearbeitung haben wir die Bedingung formuliert, daß die Beobachtung in mindestens k Phasen die Nullhypothese auf dem empirisch festgesetztem Signifikanzniveau α widerlegt. Der Wert von k ist ein einstellbarer Parameter.

Aus der Gesamtheit der Beobachtungen (z.B. Häufigkeitsverteilung von Wortformen, Kollokationen) eines Indikators werden also diejenigen akzeptiert, die aufgrund der ausgewählten Approximation unserer linguistischen Fragestellung auf Neologismen schließen lassen.

3.2.5 Sekundäranalyse und Klassifikation

Dieser Schritt kann zum gegenwärtigen Zeitpunkt nur problemorientiert formuliert werden. Es soll eine Sekundäranalyse der Verteilung der akzeptierten Fälle vorgenommen werden mit der Zielsetzung, aus den Eigenschaften der Verteilungsfunktion auf die Natur des zugrundeliegenden Phänomens zu schließen.

Stellt man die Häufigkeitsverteilung der akzeptierten Beobachtungen z.B. graphisch dar, so kann man sie unter Anwendung der üblichen Verfahren zur Bewertung von Kurven wie Diskussion von Minima und Maxima, Steigung, Stetigkeit usw. klassifizieren. Es ist zu erwarten, daß unterschiedlichen linguistischen Begriffen wie z.B. Neologismen, Okkasionalismen, aussterbenden Wörtern auch unterschiedliche *charakteristische* Kurven zuzuordnen sind.

Desweiteren können die akzeptierten Fälle unter sich auf Wechselwirkungen untersucht werden mit der Zielsetzung, eventuelle Zusammenhänge aufzudecken. Als Methoden zur Durchleuchtung solcher Interaktionen bieten sich u. a. an: Korrelationsanalyse, Faktorenanalyse, Pattern Recognition, Clusteranalyse usw.

Auf diesen Analysen kann nun der Lexikograph aufbauen und weiterführende linguistische Interpretationen vornehmen.

4. Implementierung

Die vorgeschlagene Methode ist vollständig in C unter UNIX implementiert. Die Errechnung der zwei aufgeführten Beispiele hat auf einem mit 128-MB-Hauptspeicher, 15-GB-Disk-Array und zwei Motorola-88110-Prozessoren ausgestatteten Data General AViiON-6225-Rechner im Institut für deutsche Sprache, Mannheim, jeweils 4 Stunden Realzeit in Anspruch genommen.

5. Beispiele

Die folgenden Beispiele wurden mit den oben beschriebenen Verfahren ohne jegliche Intervention des Menschen errechnet. Es zeigt sich, daß alle automatisch errechneten Neologismen-Kandidaten auch aus der Sicht der Lexikographen untersuchenswert sind. Da die vorgeführten Beispiele lediglich das vorgeschlagene Verfahren verdeutlichen sollen, verzichten wir hier völlig auf ihre linguistische Interpretation und Kommentierung. Für die graphische Darstellung wurden die unterschiedlichen Längen der Sampling-Intervalle linearisiert und die Phasenhäufigkeiten durch den entsprechenden Differenzenkoeffizienten³ normalisiert.

Beispiel

Korpus:	Wendekorpus-Ost
Indikator:	Kollokationen aller Wortformenpaare
Signifikanztest:	χ^2 -Test
Signifikanzniveau:	$\alpha=0.01$
Akzeptanzkriterium:	H_0 abgewiesen in mindestens 3 Phasen

Insgesamt wurden 1784 Wortpaare akzeptiert, davon ca. 70 Eigennamen, z.B.:

Andreas Voigt	Christa Wolf
Berliner Zeitung	Erich Honecker
Bernd Langnickel	Gregor Gysi
Brandenburger Tor	Günter Mittag
Bundesrepublik Deutschland	Günter Schabowski
Bündnis Grüne	Günter Krause

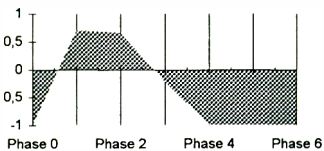
³ Der Differenzenkoeffizient c_s gibt auf der Skala $\langle -1,1 \rangle$, abstrahiert von absoluten Werten, die Abweichung zwischen der Anzahl f_t der beobachteten und der Anzahl f_e der erwarteten Fälle an:

$$c_s = \frac{f_e - f_t}{f_e + f_t}$$

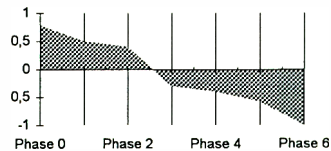
Einige weitere Wortpaare:

Einheit Deutschlands	Ex-DDR
Noch-DDR	Kampf gegen
Obhutspflicht für	europäischen Hauses
Osten Deutschlands	freie Wahlen
Runde Tisch	führende Rolle
Runden Tisches	fünf neuen
Wirtschaftsunion/ Währungsunion	gesellschaftlichen Dialog
alle Bürger	gesellschaftlichen Eigentums
dem Einigungsvertrag	kein Geld
der Treuhandanstalt	neuen Bundesländern
der ehemaligen	nicht unbedingt
die Arbeitslosigkeit	noch Esel
die Besetzer	weder Ochs
ehemaligen DDR	personenbezogenen Daten
wir fordern	unseres sozialistischen

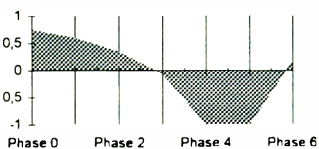
noch Esel



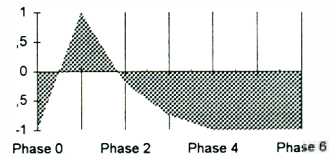
alle Bürger



europäisches Haus



Obhutspflicht für



Beispiel

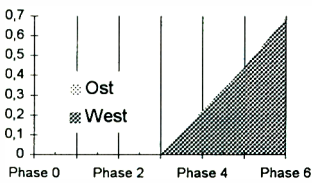
Korpus: Wendekorpus-Ost + Wendekorpus-West
 Indikator: Häufigkeit aller Wortformen (Ost/West-
 kontrastiv)
 Signifikanztest: χ^2 -Test
 Signifikanzniveau: $\alpha=0.01$
 Akzeptanzkriterium: H_0 abgewiesen in mindestens 4 Phasen

Als Ergebnis wurden insgesamt 282 Wortformen aus dem West-Korpus und 418 Wortformen aus dem Ost-Korpus akzeptiert, z.B.:

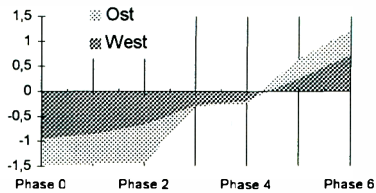
Ausreise
 Beitritt
 Demonstration
 Erneuerung
 Tisch
 Treuhandanstalt

Werkstätigen
 Wiedervereinigung
 Zusammenarbeit
 alten
 ehemaligen
 gesellschaftlichen

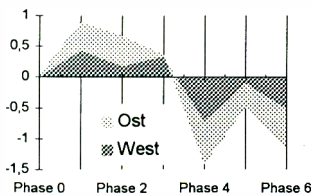
Treuhandanstalt



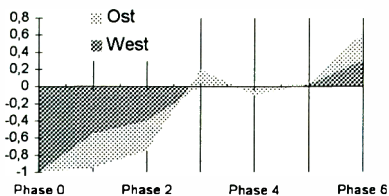
ehemaligen



Demonstranten



alten



6. Literatur

- Aarts, Jan/Meijs, Willem (1984): *Corpus Linguistics*. Amsterdam.
- Belica, Cyril/Neumann, Robert (1993): *Zeitstrukturen in Korpora*. Vortrag auf dem Internationalen Kolloquium Neologie. Mannheim.
- Biber, Douglas (1988): *Variation across Speech and Writing*. Cambridge.
- Clear, Jeremy H. (1987): *Trawling the Language: Monitor Corpora*. In: *ZurILEX 1986 Proceedings*. Tübingen.
- Garside, R./Leech, Geoffrey/Sampson, G. (1987): *The Computational Analysis of English*. London.
- Hofland, Knut/Johansson, Stig (1982): *Word Frequencies in British and American English*. Bergen.
- Kucera, Henry/Francis, Nelson W. (1967): *Computational Analysis of Present-Day American English*. Providence.
- McCarthy, M. J. (1988): *Naturalness in Language*. In: *English Language Research Journal*, Vol. 2. Birmingham.
- Phillips, M. (1989): *Lexical structure of text*. In: *Discourse Analysis Monographs*, 12: *English Language Research*. Birmingham.
- Reed, A. (1977): *CLOC: A Collocation Package*. *ALLC Bulletin* 5.
- Siegel, Sidney (1956): *Nonparametric Statistics for the Behavioral Sciences*. Tokyo.
- Sinclair, John M. (1985): *Lexicographic Evidence*. In: *Dictionaries, Lexicography and Language Learning*.
- Sinclair, John M. (1987): *Looking Up*. London.
- Yang, H. Z. (1986): *A new technique for identifying scientific/technical terms and describing science texts*. In: *Literary and Linguistic Computing*, 1.2, S. 93-103.

Dieter Herberg

Neues im Wortgebrauch der Wendezeit

Zur Arbeit mit dem IDS-Wendekorpus

Dieser Beitrag nimmt Bezug auf ein lexikologisches Arbeitsprojekt des Instituts für deutsche Sprache (Mannheim) und will einen Einblick in die Voraussetzungen und Ziele dieses Vorhabens sowie in die Arbeitsweise der Projektmitarbeiter geben. Dabei soll Aspekten der Korpus- und Computernutzung in den einzelnen Arbeitsetappen besondere Aufmerksamkeit gelten.

1. Ein Arbeitsprojekt und sein Ausgangspunkt

Die einschneidenden historischen Veränderungen, die 1989/90 – aufs engste verbunden mit parallelen internationalen Entwicklungen – in der DDR vor sich gingen, werden häufig als Wende bezeichnet. In einem engeren Verständnis ist damit die demokratische Revolution des Herbstes 1989 mit dem Ergebnis der Mauer- und Grenzöffnung im November 1989 gemeint; im weiteren Sinne kann aber auch der Zeitraum bis zur Herstellung der staatlichen Einheit Deutschlands am 3. Oktober 1990 mit dem damit verbundenen Untergang des deutschen Teilstaates DDR darunter verstanden werden. Einschließlich einer Vorphase im Sommer 1989 (Stichwort: Massenflucht aus der DDR) ergibt sich ein Zeitraum von ca. eineinhalb Jahren von Mitte 1989 bis Ende 1990, den wir als Wendezeit bezeichnen wollen.

Die friedliche Revolution des Herbstes 1989, die schrittweise erfolgende demokratische Umgestaltung vieler Lebensbereiche in der Noch-DDR des Jahres 1990 und schließlich der Beitritt der DDR zur Bundesrepublik Deutschland waren jeweils begleitet von gravierenden Änderungen im Sprachgebrauch, wie sie in einem so kurzen Zeitraum und so intensiv nur selten zu erleben und zu beobachten sind. Es ist daher nicht verwunderlich, daß die Wendezeit nicht nur bei Zeithistorikern, Politologen und Soziologen, sondern auch bei germanistischen Linguisten auf nach wie vor ungebrochenes Forschungsinteresse stößt. Obwohl schon zahlreiche sprachliche Wendezeit-Phänomene analysiert und dargestellt worden sind (vgl. den zusammenfassenden Forschungsbericht von v. Polenz 1993), kann doch von einer allseitigen und erschöpfenden wissenschaftlichen Aufarbeitung der sprachlichen Aspekte der Wendezeit – ganz zu schweigen von den Ost-West-Kommunikationskonflikten der Nach-Wendezeit (vgl. Fraas 1994) – noch nicht die Rede sein.

Am deutlichsten treten die sprachlichen Begleiterscheinungen der Wendezeit naturgemäß im Bereich der Lexik zutage. Die Veränderungen, die fast ausschließlich den Wortgebrauch in der DDR bzw. – ab 3. Oktober 1990 – in den sog. neuen Bundesländern betreffen, sind Ergebnisse vor allem von drei Arten von Prozessen, die den drei allgemeinen Kategorien entsprechen, die Munske (1990, S. 388) für Veränderungen im Wortschatz annimmt:

- 1) Vermehrung,
- 2) Abwandlung,
- 3) Schwund.

Diesen allgemeinen Kategorien lassen sich in der Wendezeit die folgenden Prozesse zuordnen:

- 1) *Aufkommen von Neologismen*, also von Neulexemen und Neubedeutungen, sowie *die Übernahme von BRD-Lexik* in den DDR-Sprachgebrauch. Gegenüber der Zahl der Übernahmen aus der BRD-Lexik ist die der wendezeitspezifischen „echten“ Neologismen relativ gering. Zum Teil waren sie nur in einer bestimmten Phase der Wendezeit aktuell und unterliegen ihrerseits bereits der Historisierung (z.B. *Ausreisepressdruck*, *Mauerspecht*, *Allianz für Deutschland*, *Währungs-, Wirtschafts- und Sozialunion*, *Zwei-Plus-Vier-Gespräche*).

Im Erwerb der Kompetenz für den aktiven Umgang mit einer Fülle von BRD-Wortgut liegt die einseitig von den Ostdeutschen zu erbringende Leistung der Anpassung an die neuen Kommunikationsbedingungen, die sich im Zusammenhang mit der Herstellung der staatlichen Einheit ergeben. Das gilt für alle Lebensbereiche; besonders aber im Erwerbsleben, im Sozialbereich und im Rechtswesen war und ist die möglichst schnelle Aneignung der lexikalischen Normen der BRD von A (wie *ABM*, *Abmahnung*, *Arbeitslosenhilfe*) bis Z (wie *Zeitarbeit*, *Zertifikat*, *Zuwendungsgeber*) für den einzelnen zumeist von existentieller Wichtigkeit.

- 2) *Verschiebungen im Gebrauch usueller Lexik*, besonders in Form von Frequenzverschiebungen zugunsten bestimmter Bedeutungen bzw. Verwendungsweisen. In manchen Fällen kann es schwer zu entscheiden sein, ob es sich „nur“ um eine neue Verwendungsweise eines usuellen Lexems handelt, oder ob bereits von einer neuen Bedeutung(svariante) gesprochen werden sollte (vgl. z.B. *Akte*), was eine Zuordnung zu 1) zur Folge hätte.
- 3) *Historisierung*, also das Außer-Gebrauch-Kommen, großer Teile DDR-spezifischer bzw. DDR-typischer Lexik, die in der aktuellen Kommunikation keine Rolle mehr spielt und nur bei der Auseinandersetzung mit der Vergangenheit noch – als Historismen – benötigt

wird. Das gilt vor allem für Lexik aus ideologie- und systemnahen Bereichen wie Politik (z.B. *antifaschistischer Schutzwall*, *Blockpolitik*, *Nationale Front*), Militärwesen (z.B. *Ehrendienst*, *NVA*, *Volksmarine*), Wirtschaft (z.B. *Gestattungsproduktion*, *Kombinat*, *Valutamark*), Justiz (z.B. *Arbeitsgesetzbuch*, *gesellschaftliche Gerichte*, *Konfliktkommission*), Bildungswesen (z.B. *Aspirantur*, *erweiterte Oberschule*, *polytechnische Erziehung*), betrifft aber auch Teile der Alltagslexik (z.B. *Menüladen*, *Westfernsehen*, *Wohngebiet*).

Im Institut für deutsche Sprache läuft ein mehrteiliges Arbeitsprojekt unter der Bezeichnung „Sprachwandel in der Wendezeit“ (Leitung: W. Teubert), das der Untersuchung und Dokumentation insbesondere der unter 1) und 2) genannten lexikalischen Erscheinungen gewidmet ist. Von den drei Teilprojekten – es gehören dazu auch die Erarbeitung eines korpuserschließenden Wörterbuches mit einer Vielzahl standardisierter Einzelartikel (vgl. Hellmann 1996) und die Untersuchung der Bedeutungsvarianz in Texten zur deutschen Einheit – wird hier auf dasjenige Bezug genommen, das sich die lexikologische Analyse und die Darstellung *wende(zeit)bedingter lexikalischer Veränderungen und Spezifika im öffentlichen Sprachgebrauch* zur Aufgabe gemacht hat.

Um das ausgedehnte Untersuchungsfeld sinnvoll zu begrenzen, konzentriert sich die Arbeit auf eine Anzahl sogenannter „Schlüsselwörter“ der Wendezeit. Ohne hier dieses Konzept im einzelnen erläutern zu können, sei soviel gesagt, daß die Rolle von Schlüsselwörtern der Wendezeit solchen lexikalischen Einheiten (Simplizia, Wortbildungen, Wortgruppen, Phraseologismen) zugeschrieben wird, die dominanten Inhalten des wendezeitbezogenen öffentlichen Sprachgebrauchs in einer, mehreren oder allen Phasen dieses Zeitraumes typischen sprachlichen Ausdruck geben, sie gleichsam kondensiert auf den Begriff bringen, und die daher in dem betreffenden Zeitabschnitt häufig gebraucht werden. Als Ergebnis ist eine Buchpublikation vorgesehen, die versucht, Sprachinformationen zur Bedeutung und zur Verwendung von Schlüsselwörtern der Wendezeit 1989/90 einschließlich ihrer Dokumentation mit der Vermittlung des nötigen zeithistorischen Hintergrundwissens zu verbinden (Herberg/Steffens/Tellenbach 1997).

Unter dem Aspekt der Thematik dieses Bandes sind nun diejenigen lexikalischen Einheiten von besonderem Interesse, deren Rolle als Schlüsselwörter mit dem Umstand verbunden ist, daß es sich entweder um Neulexeme handelt oder um usuelle Lexeme, an denen Bedeutungs- bzw. Gebrauchsveränderungen nachweisbar sind.

Bevor jede dieser Arten von Innovationen am Beispiel demonstriert wird, ist kurz auf die korpus- und LDV-bezüglichen Voraussetzungen der analytischen und dokumentierenden Arbeiten an diesem Projekt einzugehen.

2. Das Wendekorpus

Das Vorhandensein eines geeigneten Textkorpus ist für Projekte wie dieses eine unerläßliche Vorbedingung. Sein Zustandekommen und seine Existenz ist in diesem Falle der sog. „Gesamtdeutschen Korpusinitiative“ (GKI) zu verdanken, einem auf Anregung des IDS 1990/91 mit dem damals noch bestehenden Zentralinstitut für Sprachwissenschaft der Ostberliner Akademie der Wissenschaften durchgeführten Kooperationsprojekt (vgl. Herberg/Stickel 1992, Herberg 1993). Das Hauptziel des Vorhabens war, arbeitsteilig ein maschinell gespeichertes Korpus im Umfang von ca. 4 Millionen Wörtern aufzubauen, das ausgewählte Texte aus der DDR und der Bundesrepublik zu den Themenkreisen 'Politischer Umbruch in der DDR' und 'Annäherung und Vereinigung der beiden deutschen Staaten' enthält und die Beschreibung von Veränderungen im öffentlichen Sprachgebrauch der DDR im Vergleich zu dem der Bundesrepublik von Mitte 1989 bis Ende 1990 ermöglicht.

Diese Zielsetzung war von der Einsicht bestimmt, daß es das atemberaubende Tempo der politischen Ereignisse erforderte, auf diese Ereignisse bezogene Texte möglichst umgehend zu sammeln, für linguistische Analysen aufzubereiten und verfügbar zu machen. Der Bedarf an einer breiten Textbasis zur Untersuchung der sprachlichen Symptome und Folgen der Wende war unverkennbar und dringlich.

Das Resultat der GKI liegt im IDS als sog. Wendekorpus (WK), also als dokumentarisch aufbereitete, computergespeicherte Sammlung ausgewählter Textdokumente aus dem öffentlichen Sprachgebrauch in Deutschland von Mitte 1989 bis Ende 1990 zu den o.g. wendezeitbezogenen Themenkreisen vor und ergänzt die im IDS bereits seit Mitte der 60er Jahre angelegten Textkorpora zur deutschen Gegenwartssprache. Es besteht aus zwei analog strukturierten und nach identischen formalen Konventionen erfaßten Teilkorpora: aus dem Wendekorpus DDR (WKD) und dem Wendekorpus BRD (WKB).

Das WKD enthält 1632 Texte aus der DDR bzw. aus den neuen Bundesländern mit zusammen rund 1,5 Millionen Wortformen. Die Texte wurden unter drei leitenden Gesichtspunkten ausgewählt: chronologische Zuordnung, übergreifende Thematik und Textsortenvielfalt. Unter dem Gesichtspunkt der chronologischen Zuordnung wurden Texte mit erkennbarer Beziehung zu wendezeitrelevanten politischen Ereignissen ausgewählt (z.B. Fluchtbewegung aus der DDR, 40. Jahrestag der DDR, Montagsdemonstrationen, erste freie Volkskammerwahl, Währungs-, Wirtschafts- und Sozialunion, Beitritt der DDR zur Bundesrepublik). Unter dem Gesichtspunkt der übergreifenden Thematik waren Texte zu berücksichtigen, die den Diskurs zu wendezeitrelevanten Themen über einen längeren bzw. über den gesamten Zeitraum widerspiegeln (z.B. Stasi-Problematik, nationale Frage). Unter dem Gesichtspunkt der Textsortenvielfalt schließlich ging es darum, neben Zeitungstexten auch andere Text-

sorten und Äußerungsformen zu erfassen, die für die öffentliche Kommunikation in der DDR im Zusammenhang mit der Wende charakteristisch und wichtig waren (z.B. Aufrufe, Flugblätter, öffentliche Erklärungen, Wahlkampfmaterialien, verschriftlichte Volkskammerdebatten). Dennoch machen Texte aus Tages- und Wochenzeitungen (z.B. *Berliner Zeitung*, *Neues Deutschland*, *Junge Welt*, *Leipziger Andere Zeitung*, *Wochenpost*) rund 50% der Texte des WKD aus.

Das WKB enthält 1755 Texte aus den alten Bundesländern mit zusammen rund 1,8 Millionen Wortformen. Ausgewählt wurden Texte, die einen inhaltlichen Bezug zu den historischen Ereignissen 1989/90 in der DDR und zur staatlichen Vereinigung haben. Als Textquellen wurden in erster Linie Tages- und Wochenzeitungen und Zeitschriften genutzt. Hinzu kamen Protokolle von Sitzungen des Bundestages. Anders als für das WKD mußten keine weiteren Textsorten berücksichtigt werden, da etwa Flugblätter, Aufrufe, offene Briefe u.ä. für die öffentliche Diskussion der DDR-Ereignisse in der alten Bundesrepublik keine Rolle spielten. Allerdings finden sich in einem Teil der WKB-Texte Zitate oder indirekte Wiedergaben aus solchen Dokumenten. Texte aus Presseerzeugnissen wie *Mannheimer Morgen*, *Rheinischer Merkur*, *Frankfurter Rundschau*, *Der Spiegel*, *Die Zeit*, *Stern*, *die tageszeitung* machen rund 80% der Texte des WKB aus.

Jedes der beiden Teilkorpora gliedert sich – dokumentationstechnisch gesprochen – in Dokumente, die unterschiedliche textuelle Erscheinungsformen oder verschiedene Zeitungen und Zeitschriften voneinander abgrenzen. Die kleinste dokumentarische Einheit ist das Subdokument (z.B. der einzelne Zeitungsartikel, das einzelne Flugblatt), das jeweils durch eine Subdokumentennummer gekennzeichnet ist, die seine bibliographische Verifizierung bei den Computer-Recherchen gewährleistet. So verbergen sich z.B. hinter der Kennzeichnung WKB/RM3.20917 die Informationen, daß es sich um ein Subdokument aus dem WKB handelt, und zwar aus einer Nummer der Tageszeitung *Rheinischer Merkur* (RM), die dem 2.Halbjahr 1990 entstammt (3), präzise um den Artikel „Die Nacht der Deutschen“ aus der Nummer vom 05.10.1990, S. 3 (20917).

Mit Hilfe des in Zusammenarbeit mit der Arbeitsstelle LDV des IDS von einer Softwarefirma entwickelten Recherchesystems COSMAS (= Corpus Storage, Maintenance and Access System) sind die Texte nach Einzelwortvorkommen (z.B. *Wende*), nach Wortfolgen (z.B. *friedliche Revolution*) und nach Wortkombinationen (z.B. *sozial* nicht ausreichend *abgedeckt*) erschließbar (vgl. al-Wadi 1994). Ein im IDS entwickeltes Lemmatisierungsprogramm ermöglicht die Suche nach zu einer Grundform gehörenden Flexionsformen, Komposita und sonstigen Wortbildungsformen.

In den einzelnen Arbeitsphasen werden die unterschiedlichen Möglichkeiten der Ergebnispräsentation, die COSMAS bietet, genutzt: als Textübersicht mit absoluter oder relativer Häufigkeitsangabe, als Wortformenübersicht mit Häufigkeitsangabe, als einzeilige KWIC-Übersicht oder als KWIC-Darstellung in beliebig großem Kontext mit Quellennachweis. COSMAS kann die Ergebnismenge sortieren, gruppieren und randomisieren. Besonders wichtig für die textdokumentative Arbeit: Der Benutzer hat die Möglichkeit, Belege auszuwählen. Verschiedene Ergebnispräsentationen und beliebige Ergebnismengen können zur weiteren Bearbeitung auf einen lokalen Rechner geladen und anschließend auf Datenträger gesichert bzw. ausgedruckt werden.

In bezug auf das hier betrachtete Wendexik-Projekt liegen die Ergebnisse der COSMAS-Recherchen im Wendekorpus für die zu untersuchenden Stichwörter als Dokumente im Textverarbeitungssystem HIT vor. Speziell für die Belange dieses Projektes wurde von der Arbeitsstelle LDV des IDS die Segmentierung aller Quellen nach sechs Zeitphasen vorgenommen. Die Phasensegmentierung des Zeitraumes von Mitte 1989 bis Ende 1990 erfolgte nach dominanten historischen, den Gesamtzeitraum strukturierenden Ereignissen, wobei die Phasen so geschnitten wurden, daß auch die mediale Widerspiegelung des zeitlichen Umfeldes der betreffenden Ereignisse einbezogen ist:

Phase 1: 01.07.89 – 17.10.89 (Vorphase: Massenflucht, Montagsdemonstrationen, 40. Jahrestag der DDR)

Phase 2: 08.10.89 – 28.11.89 (Maueröffnung)

Phase 3: 29.11.89 – 18.03.90 (erste freie Volkskammerwahlen)

Phase 4: 19.03.90 – 22.08.90 (Währungs-, Wirtschafts- und Sozialunion)

Phase 5: 23.08.90 – 15.10.90 (Weg zur staatlichen Einheit)

Phase 6: 16.10.90 – 31.12.90 (das geeinte Deutschland)

Die Nummer der jeweiligen Phase (1 bis 6) wird in die Quellenangabe anstelle des K von WKD bzw. WKB aufgenommen, so daß das o.g. Beispiel als W5B/RM3.20917 erscheint, denn die Quelle, die vom 05.10.1990 stammt, fällt in die von uns mit der Zahl 5 bezeichnete Zeitphase vom 23.08.90 – 15.10.90.

Diese Phasensegmentierung der Textquellen ist ein wichtiges heuristisches Hilfsmittel zum Erkennen und Überprüfen von Wortgebrauchsveränderungen relativ zu vorherliegenden oder nachfolgenden Phasen bzw. zur Vor- oder Nach-Wendezeit.

3. Analysebeispiele

Im folgenden soll an je einem Beispiel für die Bildung eines Neulexems (NL), für die Herausbildung einer Neubedeutung (NB) und für die Entstehung neuen Wortgebrauchs (NG) während der Wendezeit demonstriert werden, auf welche Weise, in welchem Umfang und mit welchen Resultaten die oben umrissenen Möglichkeiten des Rechneinsatzes die Korpusauswertung und die linguistische Analyse unterstützen und fundieren. Dabei wird nur auf die für dieses Ziel wesentlichen Fakten und Gesichtspunkte eingegangen, während Fragen der Präsentation der Ergebnisse hier nur gestreift werden können (vgl. dazu Herberg 1996).

Die Festlegung, aus dem großen Bereich wendezeitbezogener lexikalischer Erscheinungen nur sog. Schlüsselwörter (s.o.) bzw. bei der Behandlung von Schlüsselthemen gebrauchte Lexik zu analysieren, reduzierte zwar den Kreis der Stichwortkandidaten, bot aber für die konkrete Auswahl der Analysewörter immer noch einen beträchtlichen Ermessensspielraum. Die endgültige Auswahl nach dem Kriterium besonderer Relevanz erfolgte auf der Grundlage der fachlichen und der kommunikativen Kompetenz der drei Bearbeiter (-innen), die während der gesamten Wendezeit in der DDR bzw. in den neuen Bundesländern gelebt haben.

Die Rechnerdaten zur Frequenz des subjektiv ausgewählten lexikalischen Materials im Wendekorpus bestätigten in den allermeisten Fällen auch unter quantitativem Gesichtspunkt die kommunikative Relevanz des Ausgewählten.

Von Anfang an bestand die Übereinkunft, daß das Ergebnis nicht ein Wörterbuch herkömmlicher Art mit einer Vielzahl knapper, stark standardisierter Einzelartikel sein sollte, sondern eher ein methodisch an textdokumentative Arbeiten (wie z.B. das Lexikon zum öffentlichen Sprachgebrauch *Brisante Wörter von Agitation bis Zeitgeist*, s. Strauß/Haß/Harras 1989) anknüpfendes Informationsmittel, in dem eine begrenzte Zahl wendezeitrelevanter sprachlicher Ausdrücke in ihrem wendezeitspezifischen Gebrauch relativ ausführlich analysiert, interpretiert und dokumentiert wird. Demgemäß nehmen die Textbausteine, die sich auf die Interpretation und auf die Dokumentation durch Belege beziehen, den zentralen Platz innerhalb der Kapitel ein. Im Baustein „Interpretation“, um dessen Inhalt es bei den folgenden Beispielen im wesentlichen geht, wird alles das mit Mitteln der narrativen, diskursiven Darstellung mitgeteilt, was sich an linguistisch relevanten Erkenntnissen aus der Analyse der Belege aus dem Wendekorpus in bezug auf das Analysewort gewinnen läßt. Die Interpretation kann sich je nach dem Gewicht bzw. der Einschlägigkeit der einzelnen Gesichtspunkte unterschiedlich akzentuiert und angeordnet auf folgende Gesichtspunkte beziehen:

- Auffälligkeiten in bezug auf die Verteilung in den Teilkorpora, in einzelnen Phasen
- Bedeutungs- und Sacherläuterungen
- Ursachen und Art der wendezeitbedingten Veränderungen (z.B. neues Lexem, neue Bedeutung, Frequenzänderung)
- wendezeitrelevante Synonyme/Antonyme zum Stichwort mit anderem Wortstamm
- Fragen der Wertung (z.B. in Abhängigkeit vom politischen Standort)
- Bindung an bestimmte Textsorten, Quellentypen, Sachgebiete
- auffällige neue Gebrauchsweisen (vor allem spezifische Kollokationen)
- grammatische Besonderheiten
- Wortbildungskonstruktionen
- enzyklopädische und sprachgeschichtliche Zusatzinformationen
- Sekundärliteratur.

Nur in einem relativ kleinen Teil der Fälle wird einem einzelnen Lexem ein spezielles Kapitel gewidmet (z.B. *Akte*, *Runder Tisch*, *Treuhand*). Die weitaus meisten Lexeme werden in zusammenfassenden Kapiteln behandelt, sofern sie nämlich in einem paradigmatischen Zusammenhang stehen (z.B. *Beitritt* / *Anschluß* / *Angliederung* / *Annexion* / *Einverleibung* / *Übernahme* / *Vereinnahmung*) oder sofern sie sich auf ein Schlüsselthema beziehen (z.B. „Vorbildwirkung der Sowjetunion“ mit Lexemen wie *Glasnost*, *Perestroika*, *Umgestaltung*, *neues Denken*, *Tapetenwechsel*).

3.1 Neulexem: *Runder Tisch*

Dieses Beispiel steht für die Entstehung eines neuen Lexems – in diesem Falle eines Wortgruppenlexems – während der Wendezeit. Unter vergleichender Hinzuziehung von standardsprachlichen Wörterbüchern der deutschen Gegenwartssprache (den sog. „Vergleichswörterbüchern“ HDG und DUW) und eines von uns so genannten „Vergleichskorpus“ (das sich aus mehreren im IDS vorhandenen gegenwartssprachlichen Korpora von Texten der Vor-Wendezeit mit insgesamt ca. 20 Millionen Wortformen zusammensetzt: Mannheimer Korpora 1 und 2, Bonner Zeitungskorpus, vier Handbuchkorpora, Limas-Korpus, Thomas-Mann-Korpus und Grammatik-Korpus 1), läßt sich anhand des Wendekorpus feststellen, verfolgen und dokumentieren, daß es sich um ein NL handelt und wie sich die Herausbildung dieses Lexems in der Wendezeit vollzogen und sein Gebrauch entwickelt haben.

In der deutschen Standardsprache kann usuell mit der metaphorischen präpositionalen Wendung *am runden Tisch* mit Bezug auf Gespräche, Verhandlungen u.ä. ausgedrückt werden, daß es sich um Gespräche, Verhandlungen u.ä. „unter gleichberechtigten Teilnehmern“ (HDG), die an einem Tisch sitzen, handelt. Die konkrete Form des Verhandlungstisches – rund, oval, rechteckig oder anders – ist dabei gleichgültig. Nur vereinzelt anzutreffen ist die Wortgruppe *runder Tisch* in metaphorischer Bedeutung. (Das Bild vom „runden Tisch“ hat eine kultur- und sprachhistorische Vorgeschichte, die hier ausgespart werden kann.)

Das Adjektiv *rund* in der Wortgruppe *runder Tisch* wird auch in deren metaphorischer Verwendung (‘Beratungsrunde gleichberechtigter Teilnehmer’) – so wie in der präpositionalen Wortgruppe *am runden Tisch* auch – der orthographischen Regelung nach klein geschrieben.

Die metaphorische Verwendung bildet den Ausgangspunkt für eine im folgenden darzustellende wendezeitspezifische Entwicklung. Im Herbst 1989 wurde darüber diskutiert, ob eine außerparlamentarische Institution von Vertretern der verschiedenen politischen Richtungen nach polnischem und ungarischem Vorbild nicht auch in der DDR etabliert werden sollte. In den entsprechenden Texten wird diese möglicherweise zu schaffende Institution *runder Tisch* genannt. Die Kleinschreibung des Adjektivs, der mitunter vorkommende Einschluß in Anführungszeichen („*runder Tisch*“) sowie die Verwendung des unbestimmten Artikels (*ein runder Tisch*) sind Kennzeichen des zu dieser Zeit noch nicht auf eine tatsächlich existierende DDR-Institution bezogenen Gebrauchs. Mit der Etablierung dieser Institution geht auf sprachlicher Ebene die Entstehung des NL *Runder Tisch* einher. Der Darstellung der sprachlichen Besonderheiten, die mit dem Gebrauch des NL *Runder Tisch* zusammenhängen, sollen einige historische Daten als Verständnishilfe vorausgeschickt werden.

Ende 1989 konstituierte sich in Ost-Berlin diese Institution: der (zentrale) Runde Tisch. Teilnehmer waren Vertreter der umstrukturierten ehemaligen Blockparteien, von neugegründeten Parteien und von zahlreichen Bürgerrechtsbewegungen. Sie bildeten ein Forum der öffentlichen Kontrolle der Arbeit der Regierung Modrow auf zentraler politischer Ebene in der krisenhaften Phase der DDR zwischen der Wende und der Etablierung demokratisch legitimierter Machtverhältnisse. Moderiert von drei Vertretern der Kirchen der DDR, hat der Runde Tisch in der reichlich drei Monate währenden Zeit seiner Tätigkeit (7.12.89 – 12.3.90) insgesamt 16 Beratungen abgehalten. Zu seinen wichtigsten Aktivitäten gehörten die Bestimmung des Termins für die ersten freien Wahlen in der DDR, die Bemühungen um die Liquidierung des Apparates des Ministeriums für Staatssicherheit und die Vorlage eines Verfassungstextes für eine neugestaltete DDR. Kurz vor der ersten freien Wahl zur

Volkskammer der DDR (18.3.1990) löste sich der Runde Tisch, der nun seinen selbstgewählten Auftrag als erledigt ansehen konnte, auf.

Es gab jedoch nicht nur den einen, zur eindeutigen Kennzeichnung öfter so genannten zentralen Runden Tisch in Ostberlin, sondern parallel dazu eine Fülle Runder Tische auf den verschiedensten Ebenen, die sich jeweils für bestimmte Bereiche um konsensfähige Lösungen anstehender Probleme bemühten, weil neue, demokratisch legitimierte Institutionen und Entscheidungsgremien zumeist noch nicht existierten. In den Bezeichnungen für diese Runden Tische wird fast immer der jeweilige Zuständigkeitsbereich in einem voran- oder nachgestellten Attribut genannt. Die attributiven Zusätze nennen lokale Zuordnungen, z.B.

*der Runde Tisch im Bezirk Leipzig
am Parchimer Runden Tisch
die örtlichen Runden Tische*

oder/und sachliche Zuordnungen, z.B.

*ein Runder Tisch für Bildungsfragen in Magdeburg
ein Runder Tisch der Jugend
ein Grüner Runder Tisch
am betriebseigenen Runden Tisch.*

Die Wortgruppe *Runder Tisch* hat naturgemäß keinen Plural, wenn damit der zentrale Runde Tisch gemeint ist; sie hat die Pluralform nur dann, wenn über die nicht zentralen Runden Tische, ggf. unter Einschluß des zentralen Runden Tisches, gesprochen wird.

Die folgende sprachliche Analyse des Korpusmaterials konzentriert sich auf die Verwendung des Wortgruppenlexems *Runder Tisch* als Bezeichnung für den in Phase 3 etablierten zentralen Runden Tisch, der aufgrund seiner politischen Bedeutsamkeit im Zentrum der öffentlichen Aufmerksamkeit stand. Gegenüber dem Gebrauch der metaphorischen Wortgruppe *runder Tisch* zeigen sich im Gebrauch des Wortgruppenlexems *Runder Tisch* deutliche Veränderungen.

Die äußerliche, sofort ins Auge fallende Begleiterscheinung der Verwendung als Bezeichnung für die etablierte Institution ist die nun nahezu durchgängige Großschreibung des adjektivischen Bestandteils (*Runder Tisch*). Damit entfällt die Notwendigkeit, zur Hervorhebung der Funktion dieses zweigliedrigen Ausdrucks als Benennungseinheit Anführungszeichen zu verwenden.

In Phase 3, der Zeit der Existenz und der politischen Wirksamkeit des Runden Tisches, stehen natürlicherweise die Zusammensetzung, die Aufgaben, die

Arbeitsweise, die Befugnisse und die Erfolge dieser Institution im Zentrum der öffentlichen Aufmerksamkeit, und die Verwendungen des Lexems Runder Tisch beziehen sich häufig auf diese Aspekte, z.B.

*Teilnehmer, Mitglieder, Vertreter, Beauftragte des Runden Tisches
die Moderatoren des Runden Tisches
die Gruppierungen und Parteien des Runden Tisches
Arbeitsgruppen des Runden Tisches, beim Runden Tisch
die Geschäftsordnung, das Vetorecht des Runden Tisches
der Regierungsbeauftragte für den Runden Tisch
ohne den Runden Tisch läuft nichts
den Runden Tisch akzeptieren, aufwerten.*

Lexeme mit der Bedeutung 'Institution' haben häufig eine mit dieser Bedeutung metonymisch verbundene weitere Bedeutung 'Gremium von Menschen, das diese Institution repräsentiert'. So beziehen sich auch manche der folgenden Verwendungen von Runder Tisch mehr auf das konkrete Gremium als auf die abstrakte Institution, z.B.

*Beratungen, eine Sitzung des Runden Tisches
Forderungen, Empfehlungen, Vorschläge, Beschlüsse des Runden Tisches
der Verfassungsentwurf des Runden Tisches
die Meinung, der Wille des Runden Tisches
den Runden Tisch leiten, auflösen
dem Runden Tisch etw. vorlegen, ein Angebot machen.*

Die Bedeutung 'Gremium von Menschen' liegt immer dann vor, wenn Runder Tisch Agens ist, z.B.

*der Runde Tisch macht Vorschläge, bildet Arbeitsgruppen
der Runde Tisch beschließt, fordert, verweigert etw.
der Runde Tisch unterbricht seine Beratungen
jmd. wird durch den Runden Tisch nominiert
etw. wird vom Runden Tisch eingebracht, beschlossen.*

In solchen und ähnlichen Kollokationen ist die semantische und syntaktische Verselbständigung des neuen Lexems gegenüber der metaphorischen Wortgruppe der Vor-Wendezeit am weitesten vorangeschritten.

Demgegenüber gibt es Verwendungen in Präpositionalgruppen, in denen der Bezug zum usuellen Gebrauch der Vor-Wendezeit (bewußt) hergestellt ist; es handelt sich vor allem um bestimmte Konstruktionen mit den Präpositionen *an* und *auf*, z.B.:

*Diskussionen am Runden Tisch
am Runden Tisch sitzen, diskutieren, erscheinen
an den Runden Tisch kommen, gehören
etw. auf den Runden Tisch bringen, legen.*

In den Phasen 4 bis 6 spielten der zentrale Runde Tisch und die anderen Runden Tische in der öffentlichen Diskussion eine stetig geringer werdende Rolle, da sie ihre Aufgaben weitgehend erfüllt hatten. Entsprechend seltener sind Belege in WK zu finden; sie bestehen fast ausschließlich in Rückblicken auf die Zeit der Wirksamkeit der Runden Tische, z.B.:

*der aufgelöste, frühere, ehemalige Runde Tisch
zur Zeit, zu Zeiten des Runden Tisches/der Runden Tische
der Runde Tisch der DDR im Frühjahr, der Anfangszeit.*

Als Kontextsynonym zum Wortgruppenlexem *Runder Tisch* wird aus sprachökonomischen Gründen mitunter nur *Tisch* verwendet. *Tisch* kann wiederum unmittelbar – ohne daß das Adjektiv *rund* erscheint – mit anderen Adjektiven verbunden werden, die den speziellen Aufgabenbereich des so bezeichneten Runden Tisches charakterisieren, z.B. *Grüner Tisch* ('Runder Tisch für Umweltfragen').

Als Resümee der Analyse und Interpretation des hier nur ausschnitthaft vorgeführten reichen Korpusmaterials kann festgehalten werden, daß in der Wendezeit auf der Grundlage der usuellen metaphorischen Wortgruppe *am runden Tisch* das Neulexem *Runder Tisch* als Bezeichnung für eine außerparlamentarische Institution in der DDR zur Beratung gesellschaftlich relevanter Probleme und zur informellen Beteiligung an der Machtausübung in der krisenhaften Phase zwischen der Wende und der Etablierung demokratisch legitimer Machtverhältnisse sowie für das entsprechende Gremium entsteht. Mit der Bezeichnung *Runder Tisch* kann sowohl auf den zentralen Runden Tisch in Ost-Berlin als auch auf Runden Tische auf nichtzentralen Ebenen, deren Zuständigkeitsbereich jeweils in einem voran- und nachgestellten Attribut genannt wird, Bezug genommen werden.

3.2 Neubedeutung: *Akte*

Für die Herausbildung einer wendezeitspezifischen NB – oder genauer gesagt einer neuen Bedeutungsvariante zu einem usuellen Lexem – mag *Akte* als Beispiel stehen.

Die usuelle Bedeutung von *Akte* wird in den Vergleichswörterbüchern wie folgt angegeben:

- HDG: 1.1. *Schriftstück(e) über einen speziellen Fall, zu einem bestimmten Vorgang*
 1.2. */nur im Pl./geordnete) Sammlung von Schriftstücken, Unterlagen*
- DUW: (bes. Verwaltung, Gericht): *[Sammlung von] Unterlagen zu einem geschäftlichen oder gerichtlichen Vorgang*

Die in WK gespeicherten Texte zeigen nur ganz vereinzelt *Akte* in der von den Wörterbüchern kodifizierten allgemeinen Bedeutung, und nur in wenigen Fällen beziehen sie sich auf den umfassenden Bereich 'Verwaltung' oder auf den besonderen Bereich 'Gericht'. Die Mehrzahl der Belege von *Akte* läßt vielmehr wendezeitbedingte Veränderungen im Gebrauch des Wortes deutlich werden.

In der öffentlichen Auseinandersetzung um die Art und den Grad der individuellen Verstrickung in das politische System der DDR, insbesondere von im öffentlichen Leben bzw. im öffentlichen Dienst stehenden Personen, gewann von Ende 1989 an die Beschäftigung mit den vorhandenen Akten zur Person zunehmend an Bedeutung. Das Interesse richtete sich in erster Linie auf die unter Verschuß befindlichen brisanten Stasi-Akten, aber auch auf die bedingt zugänglichen Kaderakten (Personalakten). Die kommunikative Relevanz dieses Themas führte zu einer deutlichen Frequenzsteigerung von *Akte* und der mit *Akte* gebildeten Zusammensetzungen in Verbindung mit Verschiebungen hinsichtlich der Semantik des Wortes *Akte*.

In der überwiegenden Zahl der Belege wird *Akte* in bezug auf Akten speziell des Staatssicherheitsdienstes, der Stasi, gebraucht. Dieser Gebrauch ist – verglichen mit dem in den Wörterbüchern verzeichneten usuellen Gebrauch – neu, da das Thema der von dieser Institution geführten Akten bis zur Wende tabuisiert war. Er setzt in Phase 3 ein und findet sich gleichbleibend häufig in beiden Teilkorpora bis zum Ende des Untersuchungszeitraumes. Die Texte thematisieren die Erstürmung der Stasi-Zentrale, die Forderung der Demonstranten nach Sicherstellung der Akten und ihren Protest gegen deren Vernichtung, die Volkskammertagungen zum Thema MfS/AfNS, die parlamentarischen Diskussionen über die Schaffung eines Gesetzes zur Aufbewahrung und Nutzung der Akten und die Überprüfung z.B. von Volkskammerabgeordneten anhand ihrer Akten. Aus den Kontexten geht hervor, daß sich *Akte* in der Regel speziell auf die Akten des Staatssicherheitsdienstes, die über Personen geführt wurden, d.h. auf die personenbezogenen Akten bezieht.

Es ergeben sich Gruppen jeweils sehr häufig vorkommender wendezeitspezifischer Kollokationen mit *Akte*, die oft, ohne daß der weitere Belegkontext zur Interpretation nötig wäre, deutlich werden lassen, daß es sich dabei um Stasi-Akten handelt.

Sehr oft, z.B. bei der Thematisierung solcher Aspekte wie dem des Zugangs zu oder des Umgangs mit den Akten, wird *Akte* als personenbezogene Akte ohne Bezug zu einer bestimmten Person, zu bestimmten Personen verstanden; das Wort steht dann meist im Plural, z.B.:

die Akten sichten, offenlegen
die Akten aufarbeiten, untersuchen, durchforsten, überprüfen
keinen Zugriff auf die Akten haben
Akten entfernen, verschwinden lassen, vernichten
die Sicherung, Lagerung, Nutzung der Akten.

So dominant in den Texten das Thema des generellen Umgangs mit den personenbezogenen Akten ist, so wendezeitspezifisch sind doch auch die Texte, in denen von *Akten* als von personenbezogenen Akten einer bestimmten Person oder bestimmter Personen die Rede ist. Wird über das Dossier einer bestimmten Person gesprochen, so wird überwiegend der Singular verwendet, oft in Verbindung mit einem Possessivpronomen (*meine, deine, seine, ihre Akte*), z.B.:

es gibt eine, keine Akte über jmdn.
von jmdm. existiert eine Akte
jmd. hat eine Akte
seine Akte sehen wollen, anfordern
in die Akten einsehen dürfen
etw. steht in jmds. Akte.

Es fällt auf, daß die Zahl der Belege mit einem die Institution Staatssicherheitsdienst benennenden Genitivattribut (*Akte der Stasi*) und die Zahl der Belege, in denen *Akte* ein entsprechendes Bestimmungswort erhält (*Stasi-Akte*) relativ gering ist im Verhältnis zu der großen Menge von Belegen, in denen *Akte* allein, ohne Nennung von Bezeichnungen für die Institution Staatssicherheitsdienst, eindeutig die *Akte* des Staatssicherheitsdienstes meint. Das spricht dafür, daß *Akte* eine Bedeutungsverengung erfahren hat. *Akte* hat in der Wendezeit die neue Bedeutungsvariante 'personenbezogene Stasi-Akte' erhalten. Diese Bedeutungsverengung wird besonders deutlich, wenn *Akte* im Singular gebraucht wird. Tritt in Belegen von WK das Wort *Akte* im Singular auf, wird es in der Regel sofort als Stasi-Akte verstanden. Indirekt wird das durch den Befund bestätigt, daß es keinen Beleg gibt, in dem ein die Institution Staatssicherheitsdienst bezeichnendes Genitivattribut verdeutlichend zu einem Singular von *Akte* tritt (**die Akte der Stasi*) und daß sich nur ganz vereinzelt Belege finden, in denen das semantisch eindeutige Wort *Stasi-Akte* im Singular steht. Das heißt, *Akte* im Singular bedarf nicht der Präzisierung durch *Stasi(-)*, um als Stasi-Akte interpretiert zu werden.

Ergänzend sei bemerkt: Im Unterschied zur kommunikativ dominierenden Stasi-Akte ist die Verselbständigung einer neuen Bedeutungsvariante von *Akte* im Sinne von 'Kaderakte' oder 'Personalakte' nicht eingetreten; *Akte* wird lediglich als Kontextsynonym zu *Kaderakte* bzw. *Personalakte* verwendet. Es ist zu resümieren, daß das Lexem *Akte* aufgrund von Bedeutungsverengung eine wendezeitspezifische neue Bedeutungsvariante hinzugewonnen hat: 'vom Staatsicherheitsdienst angelegte Akte(n) mit zusammengetragenen Informationen über bestimmte Personen, besonders über solche, die als politisch unzuverlässig galten'.

3.3 Neugebrauch: *Mauer*

Zu den oben aufgeführten Hauptarten von wendezeitbedingten Bewegungen in der Lexik gehören auch Verschiebungen im Gebrauch usueller Lexik. Hier soll als Beispiel für Frequenzsteigerung in bezug auf eine Bedeutung, in der das betreffende Lexem bis zur Wendezeit im offiziellen und damit auch weiterhin im öffentlichen Sprachgebrauch der DDR tabuisiert war, das Lexem *Mauer* dienen.

Bei den aufgrund subjektiver Entscheidung ausgewählten Kandidaten für Lexeme mit wendezeitbedingten Gebrauchsänderungen ist die Konfrontation des Gebrauchs in den WK-Texten mit dem in Texten des Vergleichskorpus von besonderer Bedeutung, weil damit sowohl auffällige Frequenzänderungen als auch Verschiebungen in bezug auf bevorzugte Kollokationspartner festgestellt werden können. Das Recherchesystem COSMAS bietet verschiedene Möglichkeiten für die alphabetische Sortierung in bezug auf Vorgänger und Nachfolger eines Suchobjektes (al-Wadi 1994, S. 181f.), höchstens aber für drei Wörter in beliebiger Kombination (Suchwort und/oder Vorgänger/Nachfolger und/oder Vorgänger/Nachfolger). Damit können auf effizientem Wege sowohl qualitative als auch quantitative Daten gewonnen werden, die die linguistische Analyse sowohl fundieren als auch erleichtern.

Doch zum Analysewort: Am 13. August 1961 errichtete die DDR-Regierung eine mitten durch Berlin und um West-Berlin verlaufende militärisch gesicherte Sperranlage, um die Fluchtbewegung in den Westen zu unterbinden. Schon bald nachdem begonnen worden war, den Stacheldraht, der West-Berlin zunächst abriegelte, durch Betonplatten zu ersetzen, wurde diese Sperranlage *Berliner Mauer*, häufiger aber einfach *Mauer* genannt. Das usu-elle Lexem *Mauer* gewinnt mithin eine Neubedeutung 'mitten durch Berlin und um West-Berlin verlaufende, aus Betonplatten bestehende militärisch gesicherte Sperranlage mit Grenzfunktion'. Es gibt aber auch noch eine zweite Lesart von *Mauer*, nämlich in Kontexten, in denen die Funktion der Mauer, Grenze zu sein, dominiert; in ihnen wird sowohl auf die Grenze um West-

Berlin als auch auf die – sonst üblicherweise mit *Grenze* oder umgangssprachlich mit *Zaun* bezeichnete – zwischen der DDR und der Bundesrepublik Bezug genommen, die nur beide zusammen die Abschottung nach Westen bewirken konnten. In diesen Fällen ergibt sich für *Mauer* die Bedeutungsvariante ‘mitten durch Berlin, um West-Berlin und im Westen der DDR verlaufende Grenze in Form einer militärisch gesicherten Sperranlage’. Beide Lesarten von *Mauer* wurden in der Bundesrepublik seit 1961 durchgängig verwendet. In der DDR dagegen hatte im offiziellen Sprachgebrauch für die Mauer in Berlin die Bezeichnung *antifaschistischer Schutzwall* zu gelten, mit der suggeriert werden sollte, daß die Sperranlage errichtet worden war, um die DDR und ihre Bevölkerung vor einer Gefahr aus dem Westen zu schützen. Später traten die neutralen Bezeichnungen *Staatsgrenze (West)*, *Staatsgrenze der DDR* bzw. *DDR-Staatsgrenze* auf, die sich sowohl auf die Berliner als auch auf die innerdeutsche Grenze bezogen. Sie wurden dann auch, abgesehen von Anlässen, bei denen man aus propagandistisch-demonstrativen Gründen auf die Bezeichnung *antifaschistischer Schutzwall* zurückgriff, bevorzugt verwendet.

Auch in der DDR wurde aber *Mauer* gebraucht, jedoch nur in der Alltagssprache, und zwar fast ausschließlich von denen, die sich in bewußte Opposition zum offiziellen Sprachgebrauch begeben wollten, weil sie die mit *Mauer* bezeichnete Sache ablehnten. Offiziell war die Bezeichnung *Mauer* tabuisiert, weil sie die Sache anschaulich und ungeschönt beim Namen nannte und mit negativer Wertung verbunden war. Nur ganz vereinzelt wurde bis zum Ende der achtziger Jahre in offiziellen Texten *Mauer* – und dann immer unter ausdrücklichem Bezug auf den westlichen Sprachgebrauch (häufig in Anführungszeichen) – verwendet.

Die krisenhafte Entwicklung in der DDR brachte es mit sich, daß zunehmend offener über das Thema „Mauer“ gesprochen werden konnte. Dementsprechend nimmt in der Wendezeit der Gebrauch des Wortes *Mauer* in der betreffenden Bedeutung immer stärker zu, und zu den wenigen bereits vorhandenen Wortbildungen mit *Mauer* (z.B. *Mauerbau*) tritt eine große Zahl neuer (auf die an dieser Stelle jedoch nicht eingegangen werden kann).

Beide Vergleichswörterbücher verzeichnen die hier genannten Bedeutungen von *Mauer* nicht. Daß sie in HDG fehlen, hat politische Gründe: Der Eintrag dieser Bedeutungen von *Mauer* war in einem DDR-Wörterbuch nicht opportun. Ihr Fehlen in DUW scheint dagegen ein lexikographisches Manko zu sein, denn schließlich war in der Bundesrepublik die Verwendung von *Mauer* in diesen Bedeutungen seit langem allgemein üblich.

Die in den WK-Texten vorkommenden Kollokationen mit *Mauer* lassen sich zu mehreren thematischen Gruppen zusammenfassen. Ob *Mauer* in der Bedeutung ‘Sperranlage’ oder in der Bedeutung ‘Grenze’ verwendet wird, geht aus

den weiteren Kontextpartnern hervor. Die Kollokationsgruppen können hier nicht im Detail entfaltet, sondern nur genannt und mit jeweils wenigen Beispielen illustriert werden. *Mauer* tritt in Kollokationen auf,

- die sich auf die Errichtung der Mauer beziehen
(*der Bau der Mauer; die Mauer wurde errichtet*)
- die sich auf das Öffnen, das Beseitigen der Mauer beziehen
(*die Öffnung, der Fall der Mauer; die Mauer wird abgerissen*)
- die sich auf solche von Personen in unmittelbar räumlichem Bezug zur Mauer ausgeführten Tätigkeiten beziehen, die deren ursprünglicher Funktion entgegenstehen
(*auf der Mauer stehen, herumturnen, tanzen; der erste Gang durch die Mauer; die Mauer mit Hammer und Meißel bearbeiten*)
- die sich auf den räumlich trennenden Charakter der Mauer beziehen
(*diesseits und jenseits der Mauer; auf der anderen Seite der Mauer*).

Neben diesen Gruppierungen gibt es zwei umfangreiche Gruppen von Kollokationen, für die andere als die bisher genannten inhaltlichen Aspekte gelten, nämlich ein bewertender und ein temporaler Aspekt. *Mauer* tritt in Kollokationen auf, die sich auf die negative Bewertung der Mauer beziehen (*durch die Mauer gelitten haben, getrennt worden sein; die Opfer der Mauer; schändliche Mauer*). Außerdem lassen sich Kollokationen aus mehreren der o.g. Gruppen unter temporalem Aspekt zusammenfassen (*bevor es die Mauer gab; als die Mauer gebaut wurde, fiel; vor, seit, nach (der) Öffnung der Mauer*); daß der Bau und dann wieder der Fall der Mauer so häufig zum Bezugspunkt der Situationsbeschreibung gemacht werden, zeigt, um wie einschneidende Ereignisse im Leben der Menschen es sich gehandelt hat.

Die bisher erwähnten Kollokationen zeigten *Mauer* in den Bedeutungen 'Sperranlage' und 'Grenze'. Auffällig sind darüber hinaus solche Kollokationen, die sich auf ein geistiges Begrenztsein beziehen: *Mauer im Kopf; innere, geistige Mauer*. Durch die Kollokationspartner wird deutlich, daß *Mauer* bildlich zu verstehen ist. Die genannten Kollokationen bezeichnen solche psychischen Zustände, die durch eine eingeschränkte, verkrampfte und von Vorurteilen geprägte Weltsicht bestimmt sind.

Die Kollokation *Mauer im Kopf* (mit den Varianten *Mauern im Kopf, Mauer in den Köpfen* und *Mauern in den Köpfen*) ist nicht neu. Im Vergleichskorpus finden sich Belege aus den achtziger Jahren. Aus der Tatsache, daß sie die Kollokationen in politisch irrelevanten Zusammenhängen zeigen, kann geschlossen werden, daß nicht die hier behandelte Bedeutung, sondern die usuelle Bedeutung von *Mauer* die Ausgangsbasis für ihre Entstehung war. Erst später, und zwar erstmalig in Phase 3 der Wendezeit, der Phase nach der Öffnung der Mauer, tritt diese Kollokation auch in politisch relevanten Zusam-

menhängen auf. In diesen Fällen liegt es nahe, sie als durch *Mauer* in der wendezeitrelevanten Bedeutung motiviert zu sehen. Die Kollokation tritt dann verstärkt ab Phase 4 auf, in der die durch jahrelange Abschottung verursachten Probleme des Zusammenwachsens, die inzwischen deutlich geworden waren, auch sprachlich reflektiert werden.

Die Kollokationen *innere Mauer* und *geistige Mauer* fehlen im Vergleichskorpus und kommen in politischen Texten der Wendezeit erstmalig vor. Sie zeigen, daß *Mauer* auf den mentalen Bereich übertragen wird. In ihnen wird *Mauer* als Bild für eine Barriere gebraucht, die zum einen durch die Trennung zwischen den Systemen und zum anderen durch Argwohn und Angst in einem Staat voller Spitzel bedingt ist und die die zwischenmenschliche Kommunikation erschwert. Der politische Bezug macht es sehr wahrscheinlich, daß die wendezeitrelevanten Bedeutungen von *Mauer* Ausgangsbasis für ihre Entstehung waren.

Die Wortverbindung *eine Mauer vor dem Kopf haben* – um ein letztes Verwendungsbeispiel anzuführen – demonstriert den wortspielerischen Umgang mit Phraseologismen. Ihr liegt der Phraseologismus *ein Brett vor dem Kopf haben* zugrunde. *Brett* wird durch *Mauer* ersetzt, d.h. es wird der aktuelle politische Bezug auf *Mauer* in der Bedeutung ‘Sperranlage’ hergestellt. Auf dieser Grundlage wird dann die phraseologische Bedeutung von *ein Brett vor dem Kopf haben* assoziiert und konkretisiert: ‘aufgrund der Auswirkungen der Mauer in seiner Weltsicht beschränkt sein’.

Die drei Beispiele *Runder Tisch*, *Akte* und *Mauer* sollten andeutungsweise demonstrieren, wie Textkorpus und Computer entsprechend den gegenwärtig gegebenen Möglichkeiten für die linguistische Analyse des Gebrauchs von zuvor ausgewählten Neulexemen, Lexemen mit einer Neubedeutung und Lexemen mit neuen Gebrauchsweisen bei der Bearbeitung eines konkreten Forschungsprojektes – der Untersuchung von Schlüsselwörtern der Wendezeit 1989/90 – eingesetzt werden.

4. Literatur

- DUW (1989): – Duden. Deutsches Universalwörterbuch. 2., völlig neu bearb. u. stark erw. Aufl. (1. Aufl. 1983.). Hrsg. u. bearb. v. Wiss. Rat u. d. Dudenred. unter d. Leitg. v. Günther Drosdowski. Mannheim/Leipzig/Wien/Zürich.
- Fraas, Claudia (1994): Kommunikationskonflikte vor dem Hintergrund unterschiedlicher Erfahrungswelten. In: *Zeitschrift für germanistische Linguistik*, 22, S. 87-90.
- Hellmann, Manfred W. (1996): Lexikographische Erschließung des Wendekorpus [Werkstattbericht]. In: Weber, Nico (Hg.): *Semantik, Lexikographie und Computeranwendungen*. (= Sprache und Information 23). Tübingen. S. 195-216.
- Herberg, Dieter (1993): Die Sprache der Wendezeit als Forschungsgegenstand. Untersuchungen zur Sprachentwicklung 1989/90 am IDS. In: *Muttersprache*, 103, S. 264-266.
- Herberg, Dieter (1996): Schlüsselwörter der Wendezeit. Ein lexikologisch-lexikographisches Projekt zur Auswertung des IDS-Wendekorpus. In: Zettersten, Arne/Pedersen, Viggo Hjørnager (Hg.): *Symposium on Lexicography VII*. (= Reihe Lexicographica. Series Maior 76). Tübingen. S. 119-126.
- Herberg, Dieter/Stickel, Gerhard (1992): Gesamtdeutsche Korpusinitiative. Ein Dokumentationsprojekt zur Sprachentwicklung 1989/90. In: *Deutsche Sprache*, 20, S. 185-192.
- Herberg, Dieter/Steffens, Doris/Tellenbach, Elke (1997): *Schlüsselwörter der Wendezeit. Wörter-Buch zum öffentlichen Sprachgebrauch 1989/90*. Berlin/New York.
- HDG (1984): – Handwörterbuch der deutschen Gegenwartssprache. In 2 Bden. Von einem Autorenkollektiv unter d. Ltg. v. Günter Kempcke. Berlin.
- Munske, Horst Haider (1990): Über den Wandel des deutschen Wortschatzes. In: Besch, Werner (Hg.): *Deutsche Sprachgeschichte: Grundlagen, Methoden, Perspektiven*. Festschrift für Johannes Erben zum 65. Geburtstag. Frankfurt a.M./Bern/New York/Paris. S. 387-401.
- Polenz, Peter v. (1993): Die Sprachrevolte in der DDR im Herbst 1989. Ein Forschungsbericht nach drei Jahren vereinter germanistischer Linguistik. In: *Zeitschrift für germanistische Linguistik*, 21, S. 127-149.
- Strauß, Gerhard/Haß, Ulrike/Harras, Gisela (1989): *Brisante Wörter von Agitation bis Zeitgeist*. Ein Lexikon zum öffentlichen Sprachgebrauch. Berlin/New York.
- al-Wadi, Doris (1994): *COSMAS*. Ein Computersystem für den Zugriff auf Textkorpora. Version R. 1.3-1. Benutzerhandbuch. Mannheim.

Michael Kinne

Der lange Weg zum deutschen Neologismenwörterbuch

Neologismus und Neologismenlexikographie im Deutschen: Zur Forschungsgeschichte und zur Terminologie, über Vorbilder und Aufgaben*

0. Vorbemerkungen

Das Thema „Neologismus im Deutschen“ hat innerhalb der germanistischen Linguistik ebenso wie das Neologismenwörterbuch im Rahmen der deutschen Lexikographie bis in die jüngere Vergangenheit hinein bekanntlich keine nennenswerte und schon gar keine gewichtige Rolle gespielt. Das ist erstaunlich, wenn man bedenkt, daß in unserem Jahrhundert gerade für den lexikalischen Bereich des Deutschen immer wieder markante Wandlungen charakteristisch waren – Wandlungen, die sich an neuen Wörtern und Bedeutungen wie auch an Wörtern, die weitgehend außer Gebrauch kamen, ablesen lassen und die häufig an bestimmte historische Phasen und Ereignisse gebunden sind: Man denke an die beiden Weltkriege, an die Weimarer Republik, die Nazi-Ära, an die Zeit der deutschen Zweistaatlichkeit und an die neue Einheit.

Im Umkreis dieser Zeitläufe haben immerhin zwei Perioden, jeweils geprägt durch staatstragende totalitäre Ideologien, besonderes lexikologisch-lexikographisches Interesse auf sich ziehen können. Zum spezifischen Wortschatz und Wortgebrauch sowohl der Nationalsozialisten als auch der DDR-Sozialisten wurden viele Untersuchungen, mehrfach auch Wörterbücher vorgelegt,¹ lexikographisch mehr oder auch weniger solide, meist jedoch von informativem Wert, zumindest für die einschlägig interessierte Öffentlichkeit. Bei diesen Wörterbüchern, in der linguistischen Diskussion nicht selten eher als fragwürdige und ihrem Thema nicht umfassend gerecht werdende denn als nützliche Produkte eingestuft, handelte es sich im Grunde genommen jeweils zumindest um eine Art Vorform eines deutschen Neologismenwörterbuches, wenn auch der spezifische Aspekt der Neologie für die Intentionen der Verfasser in der Regel eher eine untergeordnete Rolle gespielt haben mag. Aber

* Eine überarbeitete und in Teilen des 2. Kapitels gestraffte Fassung dieses Aufsatzes ist in *Deutsche Sprache* (H. 4, 1996) erschienen.

¹ Zur NS-Zeit z.B.: Berning (1964); Brackmann/Birkenhauer (1988); Kammer/Bartsch (1982); zur DDR z.B.: Ahrends (1986/1989); Constantin (1982); Kinne/Strube-Edelmann (1980).

immerhin: Diese Spezialwörterbücher, wenn auch nicht als Neologismenwörterbücher konzipiert, verzeichnen in erster Linie Neologismen.

Hingewiesen wird hier auf diese Spezialwörterbücher auch aus Gründen einer Art Ehrenrettung der deutschen Lexikographie, zumal der wissenschaftlich-seriösen, der es bisher nicht gelingen wollte, der interessierten Öffentlichkeit der Gegenwart und den Fachkollegen der Zukunft für bestimmte Zeiträume konzipierte Neologismenwörterbücher vorzulegen, wie das weltweit in vielen Ländern inzwischen längst gang und gäbe geworden ist.²

Nicht nur, aber doch auch der Kuriosität wegen sei erwähnt, daß es sowohl in Dänemark (Paaske 1981/83) wie in China (Dou Xuefu 1988) Neologismenwörterbücher mit Deutsch als Ausgangssprache gibt,³ klein und lexikographisch wenig anspruchsvoll zwar, aber immerhin: Das Erscheinen solcher Wörterbücher markiert deutlich den Bedarf einer bestimmten Interessentengruppe am Neologismenwörterbuch, nämlich den der ausländischen Deutschler. Hier finden sie offenbar die Lexik, die sie in den üblichen ihnen zur Verfügung stehenden ein- und zweisprachigen Wörterbüchern (noch) nicht finden.

Das ist ein wesentlicher Punkt, der für Zweck und Nützlichkeit von Neologismenwörterbüchern spricht. Ihm soll – ohne hier auf das konstitutive Thema der Begründung und Zweckbestimmung der Neologismenlexikographie systematisch eingehen zu wollen – noch ein zweiter hinzugefügt werden, und zwar mit Bezug auf die bereits erwähnten eher unfreiwillig neologischen Spezialwörterbücher zum NS- bzw. zum DDR-Sprachgebrauch. Alle diese Wörterbücher haben Wortschätze im Blickfeld, die den Ablauf der typischen Prozesse des Wortschatzwandels – vom Neologismus zum Archaismus, vom Neuaufkommen und von der Vermehrung zur Etablierung bis hin zum Außergebrauchkommen und Schwund⁴ – in verhältnismäßig kurzer Zeit absolviert haben, so daß sie, wenn überhaupt, dann kaum länger als nur kurzfristig Eingang in die großen Wörterbücher gefunden haben.⁵ In den aktuellen gegenwarts-

² Vgl. hierzu etwa in Heller u.a. (1988), Kap. 3 „Zur internationalen Neologismenlexikographie“, S. 16-69.

³ Paaske (1981): *Neologismer i moderne journalistisk tysk* enthält neue Wörter und Wortbildungen, die unter bestimmten formalen und inhaltlichen Gesichtspunkten zusammengestellt sind (nicht alphabetisch sortiert). Eine 2. Aufl. erschien 1983.

Dou Xuefu (1988): *Das kleine deutsch-chinesische Neologismenwörterbuch* ist ein zweisprachiges alphabetisch sortiertes Wörterverzeichnis (knappe grammatische Angaben) mit ca. 5000 Einträgen; Auflage: 31 000! Der Autor arbeitet seit vielen Jahren an einem umfassenden deutsch-chinesischen Neologismenwörterbuch.

⁴ Vgl. hierzu Munske (1990).

⁵ Der Vollständigkeit wegen muß in diesem Zusammenhang noch auf die jüngsten lexikographischen Bemühungen um den sogenannten Wendewortschatz (aus dem Umfeld der Ereignisse der Jahre 1989/1990) hingewiesen werden, dessen Neulexemen und Neubedeutungen größtenteils vermutlich extreme Kurzlebigkeit im Sprachgebrauch beschieden

sprachlichen Wörterbüchern hatte bzw. hat der größere Teil dieser Lexik über kurz oder lang sinnvollerweise keinen Platz mehr. Aus wortgeschichtlicher Sicht und somit im Hinblick auf sprachhistorische Fragestellungen gewinnen also Spezialwörterbücher dieser Art und damit auch die reinen Neologismenwörterbücher über den Rahmen ihres ursprünglichen Aktualitätswertes hinaus im nachhinein einen weiteren Wert, eben den der Möglichkeit der fundierten wortgeschichtlichen Information.

Kritik am Fehlen von Produkten einer deutschen Neologismenlexikographie ist in jüngerer Zeit wiederholt vorgebracht und inzwischen beinahe zu einer Art Allgemeinplatz geworden. Der kritische Hinweis auf das Desiderat berührt jedoch nur die eine, die negative Seite der Angelegenheit, die sehr wohl auch eine positive Seite aufzuweisen hat, der sich die Kritiker des Negativums wohl durchaus auch bewußt sind. Denn wenn die deutsche Sprachwissenschaft auch keine lange Tradition in der Neologieforschung und der Neologismendokumentation aufweisen kann – etwa im Unterschied zu Frankreich, England und den USA, worauf hier noch zurückzukommen sein wird – so sind in der zweiten Hälfte unseres Jahrhunderts schließlich doch wichtige Weichenstellungen vorgenommen, entscheidende Fundamente gelegt worden.

Der lange und, wie sich immer deutlicher zeigt, beschwerliche Weg einer wissenschaftlich fundierten deutschen Neologismenlexikographie begann in den fünfziger Jahren mit den Konzeptionsarbeiten für das dann in der Folgezeit an der Ostberliner Akademie der Wissenschaften erarbeitete *Wörterbuch der deutschen Gegenwartssprache* (WDG). In ihm wurde – in Anlehnung an ein frühes Vorbild deutscher Lexikographie –, an Joachim Heinrich Campe's *Wörterbuch der Deutschen Sprache* (S. 1807ff.), vermutlich aber auch an jüngere Vorbilder der osteuropäischen, insbesondere der sowjetrussischen Lexikographie⁶ – bei der Darstellung des Wortschatzes die Neologie als wichtiger Gesichtspunkt berücksichtigt und fand in spezifischen Markierungen der betreffenden Einträge (*Neuwort, Neuprägung, Neubedeutung; Modewort*) ihren Niederschlag.⁷ Vergleichbares hatte in der deutschen Lexikographie bis dahin einzig Campe geleistet, der in seinem Wörterbuch den neuen Wortschatz erstmals durch Markierungssymbole hervorhob. Lexikographische Nachfolger hatte er darin nicht gefunden. Und so sehr das WDG in vielen seiner lexikographischen Details in der Folgezeit als richtungweisend anerkannt

sein wird. Die Wendewortschatz-Lexikographie im Rahmen der Arbeiten des Instituts für deutsche Sprache sieht sich wissenschaftlich allerdings bewußt im Vor- bzw. Umfeld neologismenlexikographischer Bemühungen (vgl. dazu den Beitrag von Herberg in diesem Band).

⁶ Klappenbach/Steinitz (Hg.) (1961-1977). Zur Markierung der Neologismen vgl. das Vorwort zum 1. Bd. (1964), S. 14 (Abschnitt „Die Kennzeichnung der zeitlichen Zuordnung“).

⁷ Vgl. hierzu: Klappenbach/Malige-Klappenbach (1978) und Malige-Klappenbach (1986, bes. S. 1-55).

wurde, so hat die hier im Rahmen eines großen allgemeinsprachlichen Wörterbuchs erstmals wieder praktizierte Neologismenmarkierung in den nachfolgenden größeren deutschen Wörterbuchunternehmungen ebenfalls leider keine Nachfolge gefunden.

Eine Kontinuität in den neologismenlexikographischen Bemühungen in der DDR (Vergleichbares hat es in der Bundesrepublik zu keiner Zeit gegeben)⁸ wurde dann deutlich mit der Ankündigung des *Wörterbuches der in der Allgemeinsprache der DDR gebräuchlichen Neologismen*, das bis zu den Wendeereignissen des Jahres 1989 gut vorangekommen⁹ und als innovatives lexikographisches Projekt auch im Westen wiederholt vorgestellt worden war.¹⁰ Es gab sicherlich gute Gründe, dieses Unternehmen aufgrund seiner makrostrukturellen Vorgaben mit dem Ende der DDR abzubrechen. Die Bemühungen um das erste deutsche Neologismenwörterbuch blieben damit zunächst allerdings auf der Strecke.

Im Laufe der vergangenen zehn Jahre wurden einige kleinere Beiträge zu den Themenbereichen Neologismus, Neologismenforschung und -lexikographie vorgelegt, in denen der dringende Bedarf am deutschen Neologismenwörterbuch wiederholt überzeugend formuliert wurde.¹¹ Der bisher eher schmale und ergebnisarme Weg der deutschen Neologismenarbeit soll nun in einem neuen Angang im Institut für deutsche Sprache (IDS) in kontinuierliche Bahnen gelenkt werden, unter anderem mit der längerfristigen Arbeit an einem Wörterbuch der Neologismen der neunziger Jahre. Der vorliegende Beitrag ist im Vorfeld der für dieses Projekt notwendigen Erkundungen entstanden.

Gedacht war hierbei ursprünglich an einen Überblick über wichtige Beiträge und deren Ergebnisse zum Thema „Neologismus im Deutschen“ sowie zur deutschen Neologismenlexikographie. Es stellte sich schnell heraus – auch bei der Umschau in den linguistischen Teildisziplinen, die den Gegenstand „Neologie/Neologismus“ näher berühren oder auch einschließen (Wortgeschichte, Wortbildung, Sprachwandel) –, daß das Thema in der germanistischen Linguistik einer ausführlicheren Darstellung weitgehend entbehrt. Eine Ausnahme bildet der Forschungsbericht *Theoretische und praktische Probleme der Neologismenlexikographie*, vorgelegt von der Lexikographengruppe, die in den achtziger Jahren an der Ostberliner Akademie der Wissenschaften am oben genannten Neologismenwörterbuch arbeitete (Heller u. a. 1988). Die

⁸ Hingewiesen sei in diesem Zusammenhang jedoch auf die Neologismenwörterbücher von Heberth (1977/1992) (Österreich) und von Hellwig (1972/1983), die wissenschaftslexikographischem Anspruch allerdings nur bedingt gerecht werden. Vgl. dazu Herberg (1988a) und Wiegand (1990).

⁹ Vgl. Heller u. a. (1988), Herberg (1991).

¹⁰ Vgl. Herberg (1988a, 1988c).

¹¹ Vgl. dazu Müller (1987), Dou Xuefu (1989), Kinne (1989), R. u. C. Barnhart (1990), Wiegand (1990) und Herberg (1991).

Untersuchungen dieses Forschungsberichtes haben in ihren Aussagen und Ergebnissen inzwischen weder an Aktualität und kaum an Richtigkeit verloren, so daß sie hier im einzelnen nicht wiederholt zu werden brauchen. Sie bilden jedoch die Basis für jede thematische Weiterarbeit und somit auch für die folgenden Ausführungen.

Neben dem Mangel an grundlegenden thematisch einschlägigen Untersuchungen ist andererseits die Fülle kleinerer, aber auch umfangreicherer Untersuchungen beachtenswert, in denen zwar nicht über Neologie und den Neologismus im allgemeinen reflektiert wird, in denen vielmehr Neologismen im besonderen, und zwar meist in beachtlichem Umfang, Gegenstand der Darstellung sind. In der Regel gehören diese Arbeiten zum Themenkreis „Sprachentwicklung im gegenwärtigen Deutsch“.¹² Untersuchungen dieses Typs haben vor allem den Vorteil, daß sie für die Praxis der Neologismenlexikographie eine Fülle von Beispiel-, oft auch Belegmaterial bieten. Als solche entbehren sie hier jedoch der Notwendigkeit weiterer Erörterung.

Aufgrund dieser vorgefundenen Sachlage habe ich mich entschlossen, die Akzente meiner Darstellung dort zu setzen, wo die Fachliteratur und die neologismenlexikographische Praxis sinnvolle und ausreichende Anknüpfungspunkte bieten. Das ist zum einen der Umgang mit dem Terminus *Neologismus*, wie er vor allem in linguistischen Fachwörterbüchern und Sachlexika greifbar ist, und zum anderen die Einschätzung der Rolle und der Bedeutung des Neologismus im Rahmen von Sprachwandel und Sprachentwicklung. Auf dieser Grundlage wird eine Klassifizierung der Neologismen vorgenommen bzw. vorgeschlagen, werden entsprechende terminologische Festlegungen bzw. Definitionen zur Diskussion gestellt, deren Verbindlichkeit für zukünftige neologismenlexikographische Projekte für notwendig erachtet wird.

Im zweiten Teil wird die bemerkenswert ergebnisreiche angloamerikanische Neologismenlexikographie der letzten Jahrzehnte in einem knappen Überblick kommentierend vorgestellt. Sie kam, wie ich meine, bisher noch nicht ihrem hohen Stellenwert entsprechend ins Blickfeld der deutschen Lexikographie. Zu ihr werden abschließend die Aufgaben einer zukünftigen deutschen Neologismenlexikographie in Beziehung gesetzt.

¹² Vgl. dazu neben vielen Beiträgen der Zeitschrift *Der Sprachdienst* unter anderen die Veröffentlichungen von Braun (1987), Carstensen (1986), Fleischer (1987), Glück/Sauer (1990) und Steger (1989). Hingewiesen sei in diesem Zusammenhang auch auf Schippan (1992), die in ihrer *Lexikologie ...* in einem eigenen Kapitel „Neologismen“ (S. 243-247) vergleichsweise ausführlich auf die einzelnen Typen der Neologie eingeht, sich kritisch mit Klassifizierungen (Heller u.a. 1988) auseinandersetzt und auf linguistisch wesentliche Merkmale bei der Bildung von Neologismen in der Gegenwartssprache hinweist.

1. Neologismus

1.1 Zur Wortgeschichte im Französischen und Deutschen

Das Wort *Neologismus* wurde in der 2. Hälfte des 18. Jhs. aus dem Französischen ins Deutsche entlehnt. In Frankreich war *néologisme* eine Wortneuschöpfung der mittleren 1. Hälfte des 18. Jahrhunderts.¹³ Es handelt sich dabei um eine neoklassische Lehnwortbildung (also ohne direktes griechisches Pendant), zusammengesetzt aus den Komponenten [néos] 'neu' und [logos] 'Wort, Lehre'. Jeweils vergleichbaren Zeiträumen des 18. Jhs. entstammen als Teile der umfangreichen neuen Wortfamilie unter anderen *néologique/neologisch*, *néologie/Neologie*, *néologue/Neologe* und *néologist/Neologist*. Angesichts dieser umfanglichen Wortfamilie kann darauf geschlossen werden, daß seinerzeit der Benennungsbedarf im Französischen groß gewesen sein muß. Bevor hier Wortbedeutung und Wortgebrauch im Deutschen sowie deren Wandlungen näher betrachtet werden, erscheint ein kurzer Blick auf Wort und Sache im französischen Ursprungsbereich sinnvoll,¹⁴ und zwar sowohl gewisser Zusammenhänge als auch deutlicher Unterschiede wegen.

Neologismus im Französischen

Das Aufkommen der Wortfamilie um *néologique* und *néologisme* im 18. Jh. ist in Frankreich von vornherein und wohl auch ausschließlich verankert im sprachwissenschaftlich-lexikographischen Umfeld, und zwar im Spannungsfeld der bis ins 16. Jh. zurückreichenden und sich bis weit ins 19. Jh. fortsetzenden heftigen wissenschaftlichen und öffentlichen Diskussionen um das Pro und Contra in der Einstellung sprachlichen Innovationen gegenüber. Diese genuin sprachwissenschaftliche Fragestellung weitete sich schon vor dem Zeitalter der Aufklärung durchaus auch in politische Bereiche aus. Daß die Wortfamilie diesem Zeitalter entstammt, markiert den Stellenwert und die Aktualität der sprachlichen Neuerung für die Aufklärung.

Einer bereits im 17. Jh. stark ausgeprägten traditionalistisch-konservativen, sprachpuristischen Haltung entsprach die unerbittliche Ablehnung jeglicher sprachlichen Innovation, die nicht nur als ein Zeichen sprachlichen, sondern auch eines generellen Verfalls nationaler und kultureller Werte galt. Soweit Verfechter dieser Richtung auf lexikographischem Gebiet wirkten, verweigerten sie der neuen Lexik strikt die Aufnahme ins Wörterbuch. Befürworter der sprachlichen Innovation andererseits, die in ihr vorrangig Anzeichen für sprachliche Bereicherung und natürliche Sprachentwicklung erkannten und sie

¹³ Der erste Beleg im *Trésor de la Langue Française* (Bd.12, Paris 1986) entstammt dem Jahr 1734.

¹⁴ Ich stütze mich hier auf die detailreichen Darstellungen von Klare (1977) und Ricken (1977).

darüber hinaus durchaus als gesellschaftlich progressive Erscheinung eingeschätzt, fanden sich unter den Aufklärern ebenso häufig wie dann vor allem im unmittelbaren Gefolge der Französischen Revolution. Je nachdem, ob die sprachliche Innovation als negativer oder positiver Vorgang eingeschätzt wurde, verbanden sich mit dem Neologismus, mit Wort und Sache, durchgängig sowohl negative als auch positive Wertungen. Interessant ist in diesem Zusammenhang die im 18. Jh. gelegentlich anzutreffende feine Differenzierung zwischen *néologie* als Bezeichnung für die berechtigte, sinnvolle sprachliche Innovation und *néologisme* für die mißbräuchliche, abzulehnende sprachliche Neuerung,¹⁵ und dementsprechend *néologue* für den kreativen sprachlichen Neuschöpfer und *néologiste* (man beachte die formale Anlehnung an *puriste*) für den Produzenten überflüssig-unsinniger Neuwörter.¹⁶

Das heftige Pro und Contra um die Neologismen in Frankreich erinnert in manchem an die Auseinandersetzungen, die in Deutschland über Generationen hinweg um die sogenannten Fremdwörter geführt wurden. Immerhin haben diese Kontroversen um die – aus sprachpuristisch-doktrinärer Sicht – lexikalischen Schmutzkinder Neologismus und Fremdwort in beiden Ländern wohl mit dazu beigetragen, die Traditionen jeweils spezifischer lexikographischer Produkte, nämlich des Neologismenwörterbuches hier und des Fremdwörterbuches dort, zu begründen.

Es entbehrt nicht einer gewissen Kuriosität, daß das erste französische Neologismenwörterbuch (weltweit vermutlich eines der ersten überhaupt) alles andere als ein Verfechter der sprachlichen Neuerung oder gar ein bewußter Wegbereiter einer neuen lexikographischen Spezies war. Pierre-François Desfontaines' *Dictionnaire néologique à l'usage des beaux du siècle*, zuerst 1726 in Paris erschienen und dann vielfach neu aufgelegt (9. Aufl. 1798), führt mit dem neuen Wort *néologique* nicht nur das erste Glied der ganzen neuen Wortfamilie ein. Vor allem werden die aus Desfontaines' Sicht verruchten neomodischen Wörter und Wortverwendungen aufgelistet, und zwar allein zu dem Zweck, um sich lustig über sie zu machen und sie höhnischer Kritik zu unterziehen.

Zur Ehre der französischen Lexikographie muß nachdrücklich darauf verwiesen werden, daß in der Nachfolge von Desfontaines' (Anti-)Neologismenwörterbuch schon bald, und zwar zunächst mit den seriösen Neologismenwörterbüchern von Alletz (1770) und Mercier (1801), die Tradition der französischen Neologismenlexikographie eröffnet wurde, die in ihrem Ursprung, in

¹⁵ Diese Differenzierung sowie die damit auch verbundene positive Haltung gegenüber der lexikalischen Neuerung entstammt dem *Dictionnaire de l'Académie* (1762). Vgl. Ricken (1977, S. 82).

¹⁶ Diese Unterscheidung findet sich in Roubauds *Nouveaux synonymes françois* (1785). Vgl. Ricken (1977, S. 82).

ihrer Begründung und in ihrer Fülle wohl einzigartig ist.¹⁷ Die Tatsache, daß in großen französischen Wörterbuchunternehmungen aufgrund sprachpuristischer Begründung den Neologismen immer wieder bewußt die Aufnahme verwehrt worden war, hatte eine isolierte lexikographische Füllung dieser Lücke zur Notwendigkeit gemacht.

Neologismus im Deutschen

Im Jahre 1754 veröffentlichte Christoph Otto Freiherr von Schönaich ein Buch mit dem Titel *Die ganze Aesthetik in einer Nuss oder Neologisches Wörterbuch*. Es ist eine Publikation, die wohl in Anlehnung an Desfontaines' *Dictionnaire néologique* entstand. Wie dieser, so zieht auch Schönaich spöttisch-kritisch vorwiegend über diverse belletristische Wortneuschöpfungen her, von denen die meisten Okkasionalismen waren. Auch sein Buch, ebenfalls aus einer sprachpuristischen Anti-Haltung der sprachlichen Neuerung gegenüber entstanden, ist kein Wörterbuch im traditionellen Sinne. Dennoch bildet es den zeitlich ausgesprochen isoliert positionierten Ausgangspunkt neologismenlexikographischer Bemühungen in Deutschland, der langfristig ohne nennenswerte Nachfolge bleiben sollte. Eine Tradition deutscher Neologismenlexikographie begründete der Freiherr von Schönaich bekanntlich nicht.

Schönaichs *Neologisches Wörterbuch* ist für das Deutsche zugleich die früheste (mir bekannte) Belegquelle für die dem Französischen entlehnten Neologismen *neologisch* (im Buchtitel oder etwa „neologische Wörterley“, S. 230), *Neologie* (in einer Kapitelüberschrift: „Neumodisches Wörterbuch für angehende Dichter oder die Neologie“, S. 15) und *Neologist* („Denn die schönsten Redensarten unserer heiligen Neologisten gleichen einem Bündel Disteln“, S. 345). Die beiden mir vorliegenden frühesten Belege für *Neologismus* sind zeitlich nur geringfügig jüngeren Datums:

„Man schreibt also nicht wie die Litteratoren in Berlin, denn bey diesen ist der Neologismus noch mäßig“ (1767);

„Dem Neologismus (in sprachlichen Dingen) mit geziemender Mäßigung bin ich nicht so ganz abgeneigt“ (1768).¹⁸

Alle hier aufgeführten Belege befinden sich zwar eindeutig in sprach- bzw. literaturbezogenen Kontexten. Dennoch kann im Beleg aus dem Jahre 1768 der Klammerzusatz „Neologismus (in sprachlichen Dingen)“ gewisse Auf-

¹⁷ Alletz, Pons Augustin: *Dictionnaire des richesses de la langue française et du néologisme qui s'y est introduit*. Paris (Neuauf. Genf 1968); Mercier, Louis-Sébastien: *Néologie ou vocabulaire de mots nouveaux, à renouveler ou pris dans des acceptions nouvelles*. 2 Bde. Paris.

¹⁸ Beleg 1767: Riedel, Friedrich Just: *Theorie der schönen Künste*. Jena. S. 383; Beleg 1768: Sonnenfels, Josef v. (1884): *Briefe über die wienerische Schaubühne*. Wien (Neudruck). S. 703.

merksamkeit für sich beanspruchen und mit einer leicht irritierten Fragestellung im Sinne von „ja wo denn sonst?“ verbunden werden. Diese Fragestellung korrespondiert nun mit aufschlußbringenden Beobachtungen, die sich beim Blick auf die Buchung der Wortfamilie um *Neologismus* in deutschen Wörterbüchern machen lassen. Diese erfolgte 1804 zuerst und im 19. Jh. dann durchgängig zunächst ausschließlich in Fremdwörterbüchern.

Die ersten Buchungen von *Neolog/Neologe*, *Neologie*, *neologisch* (hinzu kommt das ebenfalls dem Französischen entlehnte Verb *neologisieren*) finden sich 1804 in Heyses *Fremdwörterbuch* sowie 1813 bei Campe. *Neologismus* erscheint erstmals 1816 bei Oertel und 1817 bei Petri.¹⁹ Bei diesen wie auch bei den nachfolgenden Buchungen fällt auf, daß bei den Bedeutungsangaben im Unterschied zum französischen Vorbild der Bezug auf Sprache entweder ganz fehlt,

Neolog(e): „Neuerer (in irgend einer alten Lehre), Neulehrer“, Heyse (1804), entsprechend Campe (1813), dazu bei ihm u. a. der Zusatz „Ist von Neuerungen in Glaubenssachen die Rede, so kann man auch *Neugläubiger* dafür sagen“;

Neologie: „a. Neulehre, Lehrneuerung, b. Neuerungssucht in Glaubenssachen“, Oertel (1816);

neologisch: „neuerungssüchtig, neulehrig, neugläubig“, Campe (1813);

Neologismus: „Neuerungssucht“, Oertel (1816),

oder daß der Sprach-Bezug – sofern überhaupt existent – eher einer allgemeineren Bedeutung nach- oder untergeordnet erscheint

Neologie: „Neuerungssucht, Neusüchtigkeit (bes. in der Erfindung und dem Gebrauch neuer ungewöhnlicher Wörter und Redensarten), Neu- oder Irrgläubigkeit“, Heyse (1804); „Neuerungssucht. *Wortfindung*, welches Catel dafür hat, ist viel zu enge“, Campe (1813);

neologisch: „neuerungssüchtig, neulehrig (bes. in der Sprache und in Glaubenssachen)“, Heyse (1804); „a. neulehrig, neuerungssüchtig, b. neue Wörter erklärend“, Oertel (1816).

Nur in Ausnahmefällen ist der Bezug auf sprachliche Erscheinungen eindeutig und dominant, so zum Beispiel im Eintrag *Neologismen* bei Petri (1817): „Neuwörter, neue (bes. fehlerhaft neugebildete) Wörter und Redensarten“.

¹⁹ Heyse, Johann Christoph August (1804): Allgemeines Wörterbuch zur Verdeutschung und Erklärung der in unserer Sprache gebräuchlichen fremden Wörter und Redensarten. Oldenburg; Campe, Joachim Heinrich (1813): Wörterbuch zur Erklärung und Verdeutschung der unserer Sprache aufgedrungenen fremden Ausdrücke. Braunschweig (Neudruck Hildesheim/New York 1970); Oertel, Eucharius Ferdinand Christian (1816): Gemeinnütziges Wörterbuch zur Erklärung und Verteutschung der im gemeinen Leben vorkommenden fremden Ausdrücke ... 3. Aufl. Ansbach; Petri, Friedrich Erdmann (1817): Gedrängtes Deutschungs-Wörterbuch der unsere Schrift- und Umgangssprache, selten oder öfter, entstellenden fremden Ausdrücke ... 3. Aufl. Dresden.

Zusammenfassend kann festgestellt werden, daß in diesen frühen wie in den späteren regelmäßig erfolgten Buchungen in deutschen (Fremd-)Wörterbüchern des 19. Jhs. die Wortfamilie um *Neologismus* anders als im Französischen ein deutlich über den Rahmen 'Sprache' hinausgehendes Beziehungsfeld aufweist (Lehre, Religion) und – mit Ausnahme der Personenbezeichnung *Neologe*, die zunächst ganz aus dem Sprach-Umfeld herausgelöst erscheint („Neuerer“, „Neulehrer“, „Lehrneuerer“, „Neugläubiger“) – vorwiegend negativ konnotiert ist („neuerungssüchtig“, „Neuerungssucht“, „Irrgläubigkeit“), dies vor allem auch durch die Andeutung der Normabweichung bei der Bezugnahme auf sprachliche Erscheinungen („Neuerungssüchtigkeit bes. in der Erfindung und dem Gebrauch neuer ungewöhnlicher Wörter“, „neue, bes. fehlerhaft neugebildete Wörter“).

In den frühen Konversationslexika erfolgt im Gegensatz zu den Fremdwörterbüchern die Darstellung von Neologie und Neologismus – in oft recht umfangreichen Artikeln – insgesamt wesentlich korrekter und differenzierter. Zur Veranschaulichung folgen Ausschnitte aus dem Lexikonartikel *Neologie* im *Conversations-Lexikon* des Verlages Brockhaus in der 9. Aufl. von 1846 (Bd. 10)²⁰:

Neologie heißt Sprachneuerung, besonders im tadelnden Sinne, wenn man ohne dringende Veranlassung neue Wörter, Redensarten und Wendungen statt derer einführt, welche die classische Periode der Sprache in ausreichender Weise bereits darbietet. ... Allerdings müssen von der bloßen Sucht, neue Wörter zu schaffen, die Fälle ausgenommen werden, wenn z.B. neue Erfindungen, Entdeckungen und Einrichtungen auch neue Benennungen wünschenswerth machen, oder wenn Dichter zur Steigerung der Kraft des Ausdrucks und des Affects u.s.w. dergleichen bilden. ... Man wird daher neue Ausdrücke gewiß treffend und bezeichnend finden [genannt werden als Beispiele Klopstocks *wildedel* und Matthissons *heimatsiech*]. Verwerflich aber sind solche Wörter und Formen, die ... in ihrer Bildung der Etymologie und Analogie geradezu widerstreiten und in ihrer Einzelheit *Neologismen* genannt werden. In keiner der neueren Sprachen finden wir ein so verschiedenes Streben gegen Neologien, als im Französischen. ... In einer abgeleiteten Bedeutung bezeichnet man mit *Neologie* jede andere Neuerung, jedoch gewöhnlich mit einer gehässigen Nebenbedeutung des Gefährlichen, Verderblichen und Werthlosen. – In der Mitte des 18. Jahrh. bezeichneten die orthodoxen Lehrer der christlichen Kirche die Meinungen der Heterodoxen mit dem Worte *Neologie* und nannten jene deshalb *Neologen*.

Differenziert wird hier also deutlich zwischen 'Neologismen' als der überflüssigen und regelwidrig gebildeten Neologieerscheinung und den positiv einzuschätzenden, nicht exakter benannten 'neuen Ausdrücken', die gemäß dieser Darstellung also den Kategorien *Neologie/Neologismus* nicht zugeordnet werden könnten.

²⁰ (1846): Allgemeine deutsche Real-Encyclopädie für die gebildeten Stände. Conversations-Lexikon. 9. Originalaufl. Bd. 10. Leipzig. S. 210.

Der *Neologe* – noch immer in der Form *Neolog* – wird dann erstmals von Sanders (1863, kein Fremdwörterbuch!) dem sprachlichen Bereich zugeordnet, und zwar wertneutral: „Neuerer, nam. auf sprachl. Gebiet (wie *Neograph* im Gebiet der Orthographie)“. Kehrein (1876) weist unter *Neologismus* als erster korrekt darauf hin, daß es sich um eine dem Französischen entlehnte neoklassische Bildung handelt (bis dahin galt die gesamte Wortfamilie in der Regel als „griechisch“) und leitet – ebenso wie Sanders im Falle von *Neolog* – mit der sachlich-neutralen Bedeutungsangabe „neugemachtes Wort“ zur mehr und mehr wertungsfrei werdenden Behandlung des Wortfeldes in den Buchungen der Folgezeit, vor allem im 20. Jh. über. Immerhin vermerkt aber noch Mackensen (1986) zu *Neologismus* nichts anderes als ‘verkrampte Wortneubildung’²¹!

Den Weg von der ursprünglich negativ konnotierten zur wertneutralen Bezeichnung und zum Fachterminus teilt *Neologismus* (inklusive seiner ganzen Wortfamilie) unter anderem mit den kunsthistorischen Stilbegriffen *Gotik* und *Barock*.

1.2 Neologismus als Gegenstand linguistischer Reflexion

1.2.1 Darstellung in linguistischen Nachschlagewerken

Der Hinweis darauf, daß *Neologismus* ein der Fachsprache der Linguistik zugehöriger Terminus ist, begegnet im allgemeinsprachlichen Wörterbuch vergleichsweise spät, nämlich erst 1978 im 4. Band von *Duden. Das große Wörterbuch der deutschen Sprache* (GWDS). Die hier dem fachsprachlichen Hinweis folgende Bedeutungsangabe „in den allgemeinen Gebrauch übergegangene sprachliche Neuprägung (Neuwort od. Neubedeutung)“ übernimmt die Begriffe *Neuprägung*, *Neuwort* und *Neubedeutung* zwar offensichtlich aus der Terminologie des WDG²², verunklart die sachlichen Zusammenhänge aber dadurch, daß hier Neuwort und Neubedeutung als Formen der Neuprägung aufgefaßt werden können, was natürlich falsch ist, denn weder das Neuwort noch die Neubedeutung sind Subkategorien der Neuprägung. Diese wenig korrekte Bedeutungsangabe findet sich unverändert leider auch in der jüngsten Neuauflage des GWDS (1994, Bd. 5). Exakter hingegen ist die Bedeutungsangabe im ebenfalls vom Bibliographischen Institut vorgelegten *Duden. Das große Fremdwörterbuch* (auch 1994): „[in den allgemeinen Sprachgebrauch überge-

²¹ Sanders, Daniel (1863): Wörterbuch der Deutschen Sprache. 2 Bde. Leipzig; Kehrein, Joseph (1876): Fremdwörterbuch. Stuttgart; Mackensen, Lutz (1986): Deutsches Wörterbuch. 12. Aufl. München.

²² Im WDG, dem im Hinblick auf die Neologismen-Markierung so außerordentlich innovativen und verdienstvollen lexikographischen Unternehmen, fehlt übrigens der Eintrag *Neologismus*. Die Autoren des *Handwörterbuches der deutschen Gegenwartssprache* (Berlin 1984) haben diese Lücke geschlossen.

gangene] sprachliche Neubildung (Neuwort bzw. Neuprägung)“, wobei hier jedoch die ebenfalls dem Neologismus zuzuordnende Neubedeutung unerwähnt bleibt.²³

Der vergleichsweise spät in den allgemeinsprachlichen Wörterbüchern auftauchende Hinweis darauf, daß *Neologismus* der linguistischen Fachsprache zugehört, mag auch damit zusammenhängen, daß die Darstellung des Terminus in den einschlägigen linguistischen Sachwörterbüchern und Nachschlagewerken bisher insgesamt recht knapp und in den Definitionen uneinheitlich ausgefallen ist.

Das *Lexikon der Germanistischen Linguistik* (1980, 2. Aufl. Tübingen) – um mit dem Unergiebigsten zu beginnen – enthält dem Neologismus seine Aufmerksamkeit weitestgehend vor: Der Terminus kommt so gut wie überhaupt nicht vor, lediglich einmal als Klammerzusatz in „eigene, neue Wortbildungen (Neologismen)“ (S. 177) sowie zweimal in anderen entlegeneren Zusammenhängen (Sprachstörung, S. 444; Nietzsche, S. 742). Wiederum nur in einem nicht weiter kommentierten Klammerzusatz wird für die Termini *Neuwort*, *Neubedeutung*, *Neuprägung* auf das WDG verwiesen (S. 630).

Insgesamt positivere Eindrücke vermittelt der Blick in fünf zwischen 1976 und 1993 erschienene linguistische Sachwörterbücher, die hier exemplarisch auch für andere stehen, und zwar die Sachwörterbücher von Lewandowski (1976), Conrad (1985), Sommerfeldt/Spiewok (1989), Bußmann (1990) und Glück (1993).²⁴ Generell ist festzuhalten, daß es keinerlei negative Wertung der Er-

²³ Bei den jüngsten Buchungen der Wortfamilie um *Neologismus* in den beiden großen Wörterbuchunternehmungen des Bibliographischen Institutes (jeweils 1994) fällt auf, daß der (angebliche) Aspekt des Negativen/Krankhaften in der Neologie bei einigen Bedeutungsangaben – von altersher tradiert – gewissermaßen immer noch mitgeschleppt wird, was dem aktuellen Sprachgebrauch meines Erachtens nicht mehr entspricht: *neologisch*: „... 2. neuerungssüchtig“ (GWDS, *Duden-Fremdwörterbuch*); *Neologismus*: „1. Neuerungssucht, bes. auf religiösem od. sprachlichem Gebiet ...“ (*Duden-Fremdwörterbuch*). Das GWDS rückt diese meines Erachtens nicht nur veraltende, sondern tatsächlich bereits veraltete Bedeutung von *Neologismus* immerhin an die 2. Stelle hinter die der Sprachwissenschaft zugeordnete Bedeutungsangabe. Der bei diesen jüngsten Buchungen in den Einträgen *Neologie* und *neologisch* ebenfalls von altersher beigehaltene Bezug auf eine Neuerungsbewegung der evangelischen Theologie des 18. Jhs. (vgl. dazu oben das Brockhaus-Zitat von 1846) ist zwar korrekt, findet außerhalb der historischen Bezugnahme im fachsprachlichen Gebrauch der Theologie heute aber kaum noch anderweitige sprachliche Realisierung. Beide Gegebenheiten des heutigen Sprachgebrauchs lassen sich anhand der großen computergespeicherten Textkorpora der letzten Jahrzehnte verifizieren.

²⁴ Lewandowski, Theodor (1976): *Linguistisches Wörterbuch*. Bd. 2. Heidelberg. S. 473f.; Conrad, Rudi (Hg.) (1985): *Lexikon sprachwissenschaftlicher Termini*. Leipzig. S. 161; Sommerfeldt, Karl-Ernst/Spiewok, Wolfgang (1989): *Sachwörterbuch für die deutsche Sprache*. Leipzig. S. 154f.; Bußmann, Hadumod (1990): *Lexikon der Sprachwissen-*

scheinung des Neologismus gibt. Im Hinblick auf eine Differenzierung des Neologismus, etwa nach Ursprüngen, Typen, Geltungsbereichen und zeitlichen Faktoren, sind die Darstellungen recht uneinheitlich. Als allzu knapp und kaum informativ fällt die Darstellung bei Conrad etwas aus dem Rahmen. Kurz, aber durchweg ausreichend informativ sind die Darstellungen von Lewandowski, Bußmann und Glück, am differenziertesten sind die Erscheinungen bei Sommerfeldt/Spiewok beschrieben.

Es fällt auf, daß die Kategorien bzw. Klassifizierungen des Neologismus besonders uneinheitlich sind. Bei Lewandowski werden die verwendeten Bezeichnungen *Neubildung*, *Wortneuschöpfung* und *neues Wort* bzw. *neuer Ausdruck* nicht definitorisch klassifiziert bzw. systematisch miteinander in Beziehung gesetzt. Conrad gibt sich lediglich mit der einen Kategorie *Neuwort* zufrieden. Bußmann unterscheidet zwar korrekt vier Formen des Neologismus, vermeidet jedoch deren Benennung, vermutlich aus dem gutem Grund fehlender Einheitlichkeit in der bisherigen Forschung. Glück folgt der zuletzt vorgelegten Klassifizierung in Heller u. a. (1988) (*Neulexem/Neuwort*, *Neusemem/Neubedeutung*, *Neubezeichnung*).²⁵ Sommerfeldt/Spiewok wiederum lehnen sich zwar deutlich an die Klassifizierung des WDG an (*Neuwort*, *Neuprägung*), klammern dabei die *Neubedeutung* als Kategorie des Neologismus jedoch aus. Allerdings räumen sie ihr einen separaten Eintrag ohne Bezugnahme auf den Neologismus ein, wollen die Neubedeutung also vermutlich aus dem Bereich Neologismus ausgeklammert wissen.²⁶ Bei Lewandowski und Conrad bleibt die Kategorie Neubedeutung gänzlich unerwähnt. Bußmann und Glück stellen sie korrekt als Kategorie des Neologismus dar.

Fazit dieser Darstellungen: Eine übergreifende Vereinheitlichung der Klassifizierung und der Terminologie des Neologismus auf der Basis des bisher Vor-

schaft. 2. Aufl. Stuttgart. S. 520; Glück, Helmut (1993): Metzler-Lexikon Sprache. Stuttgart/Weimar. S. 415.

²⁵ Mit der Neologismen-Klassifizierung von Heller u. a. (1988) übernimmt Glück auch den Neologismtyp der *Neubezeichnung*, vermeidet allerdings die bei Heller u. a. dafür eingeführte Bezeichnung *Neuformativ* (S. 8), worunter neue Bezeichnungen für bereits benannte Erscheinungen oder Sachverhalte der objektiven Realität (z.B. *Raumpflegerin* für *Reinemachfrau*) verstanden werden. Als selbständige Hauptkategorie neben *Neulexem* und *Neusemem* ist das *Neuformativ* jedoch nicht aufrechterzuhalten, da es sich formal bei ihm um eine (inhaltlich allerdings zusammengehörende) Untergruppe der *Neulexeme* handelt. Mit dem *Neuformativ* bei Heller u. a. (1988) hat sich bereits Schippan (1992, S. 245f.) kritisch auseinandergesetzt. Das *Neuformativ* als selbständiger Neologismtyp wird als solcher inzwischen übrigens auch von den Autoren nicht mehr aufrechterhalten (mündliche Mitteilung).

²⁶ Auch für Schippan (1992) stellt sich die Zuordnung der Neubedeutung (Neosemem/Neosemantismus) zum Neologismus als problematisch dar: „Die Annahme von Neosemantismen berücksichtigt zu wenig die Beweglichkeit der Wortbedeutung und birgt die Gefahr einer statischen Zeichenauffassung in sich. Andererseits entstehen neue Benennungseinheiten durch die bewußte Zuordnung einer neuen Bedeutung zu einem Formativ. ... Hier erscheint 'Neosemantismus' als angebracht“ (S. 245). Vgl. auch Anm. 31.

liegenden (und das heißt vor allem im Anschluß an das WDG und an Heller u. a. 1988) wäre – zumal als Instrumentarium für eine zukünftige Neologismenlexikographie – ebenso sinnvoll wie wünschenswert.

Bevor hier (in Kap. 1.3) ein diesbezüglicher Vorschlag gemacht wird, sollen zunächst noch einige Anmerkungen zur Rolle und zur Behandlung des Neologismus in dem ihm übergeordneten Forschungsbereich des Sprachwandels angeschlossen werden.

1.2.2 Neologismus und Sprachwandel-Forschung

Nur Lewandowskis Sachwörterbuch-Artikel *Neologismus* enthält einen Querverweis auf das Stichwort *Sprachwandel* und markiert damit den übergeordneten Sachbereich, in den der Neologismus als lexikalische Erscheinung einzuordnen ist. Generell kann festgehalten werden, daß in früheren wie neueren Publikationen dieses Forschungsbereiches, wie bedeutend ihr Stellenwert für die diachronische Linguistik im einzelnen auch immer sein mag, den Erscheinungsformen der lexikalischen Innovation im allgemeinen und der Neologie im besonderen lediglich eine marginale Rolle eingeräumt wird, daß es somit innerhalb dieses Bereichs so gut wie keine grundlegend-systematischen Darstellungen gibt.²⁷

Bei den hier (auch in diesem Falle exemplarisch) ausgewählten Untersuchungen der Sprachwandel-Forschung fällt als erstes auf, daß in den meisten von ihnen Sprachwandel vorwiegend am Beispiel des Lautwandels dargestellt und diskutiert wird.²⁸ Cherubim (1975) etwa weist in der Einleitung seines Sammelbandes darauf hin, daß in der in diesem Band dokumentierten Sprachwandel-Forschung die einzelnen sprachlichen Ebenen unterschiedlich bearbeitet wurden, um dann festzustellen: „Während Untersuchungen zum Lautwandel das Übergewicht haben, treten Arbeiten zum Bedeutungswandel, erst recht aber zum syntaktischen Wandel dahinter zurück“ (S. 7). Bezeichnenderweise

²⁷ Einen wichtigen Neuansatz bietet Munske (1990). Seine Untersuchung, auf die hier noch zurückzukommen sein wird, will zeigen, „welche Phänomene am Wandel des Wortschatzes beteiligt sind, wie sie miteinander in Beziehung stehen und welche Rolle die Sprachteilhaber dabei spielen“, und versteht sich als „ersten Schritt zu einer Neukonzeption deutscher Wortgeschichte“ (S. 399). – v. Polenz (1991) behandelt in der neuen, stark erweiterten Fassung seiner *Deutschen Sprachgeschichte* im Rahmen der Grundbegriffe der Sprachentwicklung Erscheinung und Erscheinungsformen des Neologismus im Abschnitt „Sprachliche Innovation“ (Bd. 1, S. 37-58), ohne den Terminus als solchen zu benutzen und als spezifisches Phänomen des Wortschatzwandels zu diskutieren. Im Überblick behandelt werden von ihm die dem Neologismus zugrundeliegenden Prozesse Wortbildung, Entlehnung, Integration, Lehn-Wortbildung und Bedeutungswandel.

²⁸ Die folgenden Ausführungen greifen zurück auf Cherubim (Hg.) (1975), Cherubim (1984), Coseriu (1958/1974), Itkonen (1984), Mattheier (1984), Munske (1985) und (1990), Schank (1984) und Wurzel (1991).

kommt der lexikalische Wandel hier überhaupt nicht vor. In modifizierter Form können jedoch viele Beobachtungen und Aussagen der Forschung zum Lautwandel im Hinblick auf den Sprachwandel generalisiert werden, so daß sie partiell durchaus auch auf den Wortschatzwandel bezogen werden können und für diesen Bereich Gültigkeit besitzen.

Alle bisherigen Versuche, die Bedingungen von Sprachwandel, Sprachentwicklung und damit auch von sprachlichen Neuerungen zu motivieren, bewegen sich bekanntlich zwischen zwei Möglichkeiten: einerseits den Deutungen, die vom inneren Bereich der Sprache, von den systemimmanenten Gegebenheiten der Sprachstruktur ausgehen, zum anderen den Erklärungen, die im systemexternen, außersprachlichen Bereich angesiedelt sind, im instrumentalen Charakter der Sprache, in ihren Funktionszusammenhängen. Bedingungen und Motivation des lexikalischen Wandels und der lexikalischen Neuerung sind grundsätzlich im letzteren Bereich positioniert. Das Interesse der lexikalischen Sprachwandel-Forschung ist dabei durchgängig zweifach konzentriert: zum einen auf die Motivation und Spezifik des Aufkommens einer sprachlichen Neuerung, zum anderen mehr aber noch auf das Prozedere des Fest- und Durchsetzens sowie letztendlich der Akzeptanz der Neuerung. Coseriu (1958) begründet die klare Unterscheidung zwischen dem „individuellen schöpferischen Akt“ der **Neuerung** oder, wie er auch sagt, der **Anfangsneuerung** (S. 127) und dem der **Übernahme**, wobei er als das eigentliche Kernproblem des Sprachwandels nicht das punktuelle Problem und Faktum der Anfangsneuerung, sondern das Problem der Übernahme, also der Traditionskonstituierung, der Ausbreitung und Festigung verstanden wissen will.²⁹ Für die Neologismenlexikographie sind beide Gesichtspunkte, Neuerung wie Übernahme, von unmittelbarer Relevanz.

Jede sprachliche und somit auch jede lexikalische Neuerung entspringt zunächst immer einem individuell-subjektiven (Schöpfungs-/Prägungs-)Akt. „Alle Sprachneuerungen sind notwendigerweise individuell, doch entsprechen die Neuerungen, die angenommen und verbreitet werden, sicher interindividuellen Ausdrucksbedürfnissen“ (Coseriu 1958/1974, S. 127). Den allermeisten Produkten des individuellen schöpferischen Uraktes ist eine weitere Existenz nicht vergönnt. Der, wenn man so will, Spielball 'lexikalische Neuerung' wird von seinem Eigner hochgeworfen, von vielen auch wahrgenommen, meist aber nur von ihm selbst wieder aufgefangen; kaum ein anderer sonst will diesen Ball haben, ihn übernehmen. Dies ist das „Schicksal“ der Okkasionismen, die nicht das Zeug zum richtigen Neologismus haben, die allerdings mit ihrer unüberschaubaren Anzahl immer wieder von der Lebendigkeit und

²⁹ Coseriu (1958) strebte vor allem eine grundsätzliche Revision der theoretischen Erörterungen über den Sprachwandel an, indem er den Nachweis zu führen versucht, daß die Kausalität eine inadäquate Kategorie für die Erklärung von Sprachwandel darstellt. Coseriu geht es vorrangig um die mit dem Übernahmeprozess verknüpften Bedingungen sprachlichen Wandels.

Leistungsfähigkeit der Sprache Zeugnis geben. Dabei bleibt der Okkasionalismus in der Regel ein solcher nicht etwa wegen seiner Fragwürdigkeit oder Unverständlichkeit, er bleibt es, weil er im Kommunikationsprozeß offenbar nicht gebraucht wird.

Wird nun dieser Spielball 'lexikalische Neuerung' mehr und mehr auch von anderen und über kurz oder lang von vielen und dann sehr vielen aufgefangen, weitergeworfen und wieder aufgefangen, wird er als nützlich und sinnvoll, vielleicht auch als „schick“ oder „schön“ und letztendlich sogar als unentbehrlich empfunden, dann hat er alles, was zu einem richtigen Neologismus gehört. Als von wesentlich größerem linguistischen Interesse als der zufällig-punktueller Schöpfungsakt des Neologismus (der ursprünglich zunächst immer „nur“ ein Okkasionalismus ist) erweist sich der Prozeß der Aneignung, der Durchsetzung und Akzeptierung der lexikalischen Neuerung durch die Sprachgemeinschaft, der immer der Weg vom Okkasionalismus zum Neologismus ist, ein Weg, der zeitlich sehr kurz oder aber sehr lang sein und der allerdings plötzlich auch wieder ins Nichts führen kann. Der interessanten Beobachtung, daß selbst viele nahezu akzeptierte, teilweise bereits lexikalisierte Neologismen nach vieljähriger Gebrauchsphase dennoch mehr oder weniger unerwartet und plötzlich wieder außer Gebrauch kommen, geht anhand der englischen Lexik in der zweiten Hälfte unseres Jahrhunderts ein Beitrag von Algeo (1993) nach.

Als motivierend und letztendlich ausschlaggebend für den Sprachwandel generell wie für das Aufkommen lexikalischer Neuerungen und für den Durchsetzungsprozeß des Neologismus wird in der Sprachwandel-Forschung immer wieder auf die konstituierende Notwendigkeit bestimmter außersprachlicher Faktoren verwiesen, seien diese nun (historisch-)zeitbedingter, sozialer, ökonomischer, technischer, kultureller oder vor allem auch kommunikativer Natur. Es geht hier also um das Verhältnis zwischen Sprache und außersprachlicher Realität. Vachek (1962/1975) betont zu Recht, daß „die zunehmende Komplexität der außersprachlichen Realität, die sich in der entsprechenden Erweiterung, Anreicherung und Differenzierung des Wortschatzes einer Sprache zeigt, überwiegend durch externe Faktoren begründet ist“ (in Cherubim 1975, S. 190). Itkonen (1984), der Sprache in erster Linie als „social phenomenon“ auffaßt, betrachtet den Sprachwandel nur als Sonderfall des Sozialwandels. Als Form des Kollektivverhaltens ist jeder Sprachwandel für ihn teleologisch und nicht-nomisch. Häufig geäußert wird in der Forschung schließlich die Auffassung, daß Grundvoraussetzung für jeden Sprachwandel und damit auch für die Durchsetzung des Neologismus eine kommunikative Notwendigkeit (auf sozialer, ökonomischer ... Basis) ist; der lexikalische Wandel spiegele vorrangig jeweils bestimmte Benennungsbedürfnisse der Gesellschaft wider und diene im Rahmen der Sicherung der Verständigung zwischen den Sprachteilhabern deren Mitteilungsabsichten. Daß sich Hinweise dieser Art in der bisherigen Forschung häufig mehr im Bereich hypothetischer Allgemein-

plätze bewegen, könnte unter anderem damit zusammenhängen, daß die für detaillierte Forschungen in diesem Umfeld unerläßliche Fundierung langfristiger, differenzierter neologismenlexikographischer Kärnerarbeit für die jeweils neue deutsche Lexik bisher noch für keine Zeitphase geleistet worden ist.

Mattheier und Schank (beide 1984) knüpfen in ihren innovativen Beiträgen über Probleme des Sprachwandels nicht direkt an außersprachliche oder auch systeminterne Gesichtspunkte an, sondern bewegen sich jeweils in einem zumindest mehr sprachnahen, aber durchaus von beiden Komponenten des Außen und Innen geprägten Umfeld. Mattheier lenkt dabei das Interesse auf das für Sprachveränderungsprozesse wichtige Reservoir der sprachlichen Variabilität. Er sieht im komplexen Varietätengefüge der Sprache mit ihren verschiedenen stilistischen, regionalen und sozialen Ausdrucksmöglichkeiten die Hauptquelle für sprachliche Variabilität und somit auch für sprachliche Veränderungen. Die Variabilität, so wird ausgeführt, gewinnt für die Sprachveränderungsprozesse besondere Bedeutung, „weil für den Einzelsprecher in der Pluralität der Varietäten ein unerschöpfliches Varianten-Reservoir bereitsteht, aus dem er je nach den soziosituativen Regeln, der gesellschaftlichen Bewertung der Varietät oder der Sprechintention die angemessene Form bewußt oder routiniert, d.h. einer momentan bewußten Praxis folgend, auswählen kann“ (S. 773). Mattheier betont, daß seine Überlegungen zur sprachlichen Variabilität und den darauf basierenden Sprachwandelprozessen bisher in der Sprachgeschichtsschreibung des Deutschen (und somit auch in der Wortgeschichte) so gut wie noch keinen Niederschlag gefunden haben (S. 776).

Schank (1984) geht es um neue theoretische Ansätze einer pragmatisch orientierten Sprachwandelforschung, die auf dem Konzept der Textsorte basiert. Er hebt hervor, daß Sprachwandel durch die Heterogenität von Sprachstadien ausgelöst wird. Der Vorgang der Akzeptierung einer sprachlichen Neuerung, so führt Schank aus, „kann nur zufriedenstellend beschrieben werden, wenn ein Beschreibungsmodell für Sprachstadien erarbeitet ist, das die Heterogenität derselben greifbar macht. Weiter führt hier der Begriff der Textsorte, da er Sprachwandelprozesse in pragmatische Verwendungskontexte situiert“ (S. 763). Schank will Sprache vorrangig „als ein Gesamt des menschlichen, kommunikativen Handelns thematisiert“ wissen. Dementsprechend sei unter Sprachwandel der Wandel menschlichen Handelns zu verstehen. Aufgrund der Tatsache, daß kommunikative Handlungskontexte mit dem Begriff der Textsorte thematisiert werden können, werde im Wandel der Textsorten Sprachwandel als Wandel greifbar, und zwar sowohl als Wandel zeichen-linguistischer Elemente wie auch pragmatischer Muster (S. 762).

Die einzige jüngere Untersuchung aus dem Bereich der Sprachwandel-Forschung, die sich ausschließlich und dezidiert Fragen des Wortschatzwandels zuwendet, ist der Beitrag von Munske (1990), dessen Ergebnisse auch für den nachfolgenden terminologischen Teil relevant sind und dort entsprechend Be-

rücksichtigung finden werden. Munske weist einleitend auf das bekannte Dilemma hin, nämlich darauf „daß es für das Deutsche keine zusammenfassende Darstellung darüber gibt, mit welchen Mitteln die Sprachteilhaber den Wortschatzwandel bewirken, in welchen Beziehungen diese zueinander stehen, welche sprachexternen und -internen Ursachen den Wortschatzwandel initiieren und wie schnell eigentlich die Regeneration des Wortschatzes vor sich geht“ (S. 387). Munske wendet sich vor allem der Frage zu, „welche Mittel dem Sprachteilhaber zur Vermehrung oder Abwandlung lexikalischer Einheiten und ihrer Bedeutungen dienen, was sie leisten, in welchen Beziehungen sie zueinander stehen“ (S. 388).

Die Formen des Wortschatzwandels ordnet Munske drei allgemeinen Kategorien zu, der **Vermehrung**, der **Abwandlung** und dem **Schwund**. Als Prototypen des lexikalischen Wandels erstrecken sie sich auf die Inhalts- oder die Ausdrucksseite der Lexeme oder auf beide. Relevanz für die Neologie besitzen dabei natürlich vorrangig Prozesse im Bereich Vermehrung sowie teilweise im Bereich Abwandlung. Für die Vermehrung von Lexemen benennt und erläutert Munske drei Prozesse: die Wortbildung, die Phraseologiebildung und die Entlehnung. Die Vermehrung von Sememen lexikalischer Einheiten, die Munske (in Parallelität zum Prozeß der Wort- und Phraseologiebildung) als „Bedeutungsbildung“ bezeichnet und der er als „Kern allen Bedeutungswandels“ (S. 391) eine besonders wichtige Rolle zumißt, differenziert er als metaphorische, als metonymische und als syntagmatische Bedeutungsbildung. Als weitere Arten des Bedeutungswandels werden von Munske unter anderem die Bedeutungsverengung und Bedeutungserweiterung sowie die Pejoration diskutiert.

Von den Wortschatzwandel-Prozessen, die Munske unter der Kategorie „Abwandlung“ erörtert, erscheint mir der Prozeß „Abwandlung der Ausdrucksseite lexikalischer Einheiten“ für die Neologie von besonderem Interesse. Dazu zählen neben der nach wie vor typischen Erscheinung der morphologischen Kürzung (*Fundamentalist* > *Fundi*) Abkürzungen und Initialwörter mit den für diese charakteristischen Erscheinungen, die Munske zu Recht als „die moderne Form der Wortschöpfung“ bezeichnet (S. 398).

Nach diesem Überblick über neuere Untersuchungsansätze im Rahmen der Sprach- und Wortschatzwandel-Forschung, die in ihren Ergebnissen empirisch zu erproben und zu untermauern wären, komme ich zurück auf die für den Sprachwandelprozeß generell charakteristische (in der Forschung allgemein akzeptierte) Differenzierung in Anfangsneuerung (Bildung vieler Neuerungen) einerseits und Ausbreitung/Durchsetzung einzelner Neuerungen andererseits, und dabei insbesondere auf den Weg der Neuerung zur gesellschaftlich akzeptierten Spracherscheinung, der für die meisten Forscher den eigentlichen Prozeß des Sprachwandels markiert. Als Produkte der lexikalischen Anfangsneuerung können Okkasionalismen für Untersuchungen im Bereich Wortbildung

durchaus interessante Objekte sein, für die Aufnahme ins Wörterbuch taugen sie aufgrund ihres Mangels an Allgemeingebrauchlichkeit natürlich nicht. Von hohem lexikographischen Interesse hingegen sind die temporal und lokal exakt fixierten Ursprünge eines Okkasionalismus dann, wenn er sich zum Neologismus weiterentwickelt bzw. entwickelt hat. Da die Anfänge des Aufkommens der dann später gebrauchsetablierten Neologismen in der Regel ins Dunkle zurückgefallen sind, muß sich der Lexikograph anstelle des in den Zeitläufen verschütteten Urbelegs meist mit den frühesten ihm erreichbaren Belegen zufriedengeben. Die Kenntnis des Urbelegs ist ein extremer Ausnahmefall.

Der Prozeß der Durchsetzung einer lexikalischen Neuerung, also der Weg vom Okkasionalismus zum Neologismus, hat, sofern er anhand von Belegen verfolgt und beschrieben werden kann, für den mikrostrukturellen Bereich der Neologismenlexikographie einen spezifischen Stellenwert. Hier können den Etablierungsprozeß erhellende Details beschrieben werden, die bei der späteren Aufnahme des Neologismus ins allgemeinsprachliche Wörterbuch verständlicherweise keinen Platz mehr haben werden. Typische Merkmale der Durchsetzungsphase eines Neologismus sind häufig Gebrauchsmarkierungen wie Anführungsstriche oder Kursivsetzung – beide können sowohl die Neuheit wie auch die Statusunsicherheit signalisieren; weiterhin typisch für den Zeitraum der Gebrauchsetablierung sind dem Neologismus in Parenthese hinzugefügte Bedeutungserläuterungen. Belege für diese letztere Art sind für die lexikographische Erfassung von besonderem Wert. Mit der Aufnahme ins Wörterbuch, mit seiner Lexikalisierung ist der Prozeß der Entwicklung eines Neologismus (als Neologismus) abgeschlossen. Der ins allgemeinsprachliche Wörterbuch integrierte Neologismus ist im Grunde genommen bereits keiner mehr, er wird allenfalls noch kurzfristig als solcher empfunden.

1.3 Zur Terminologie

Im folgenden knüpfe ich an Bemühungen und Darstellungen in den wenigen wichtigen vorliegenden Untersuchungen zu dieser Thematik³⁰ an. Wiegand (1989), der den Neologismus (vielleicht zu Unrecht) als „notorisch schwer zu definierenden Terminus“ charakterisiert (S. XI), verweist auf die Notwendigkeit solcher definitorischer Bemühungen, zumal im Vorfeld neologismenlexikographischer Arbeiten. Das ist exakt der Ausgangspunkt der hier folgenden terminologischen Überlegungen. Wichtig sind diese nicht nur intern als gemeinsame einheitliche Orientierung für alle an einem neologismenlexikographischen Projekt Beteiligten, sondern auch nach außen zum kritischen Dialog und zur Abstimmung mit allen an der Neologismenarbeit Interessierten und

³⁰ Wichtigste Grundlage ist Heller u.a. (1988), wo wiederum an Vorgaben des WdG angeknüpft worden war. Berücksichtigt wurden daneben Müller (1987), Dou Xuefu (1989), Munske (1990) und Schippan (1992).

Beteiligten. Die terminologischen Überlegungen orientieren sich dabei zum einen vorrangig an der Praktikabilität im Rahmen neologismenlexikographischer Arbeit, sie sind aber zum anderen auch bemüht um eine angemessene Balance zwischen fachspezifischer Exaktheit und einer Verständlichkeit nach außen. Dies vor allem deshalb, weil die Neologismenlexikographie, das kann aufgrund der Erfahrungen in anderen Ländern festgestellt werden, mit einem Interessentenkreis zu rechnen hat, dessen weitaus größter Teil nicht zur linguistisch-lexikographischen Fachwelt zu rechnen ist.

Ausgangs- und Bezugsebene ist der *allgemein-sprachliche Wortschatz der gegenwärtigen deutschen Standardsprache*. In Anlehnung an Munske (1990) (er bezeichnet ihn als „Kernwortschatz“, S. 387) ist dieser definiert als die Schnittmenge lexikalischer Zeichen, die gegenwärtig in den deutschsprachigen Ländern Mitteleuropas in nicht ausschließlich fach- oder gruppensprachlicher Kommunikation benutzt werden. Ausgeklammert bleibt neben Regionalismen somit vor allem die rein fach- und gruppensprachliche Lexik, wobei, zumal im Rahmen der Neologismenforschung, die intensive Wechselwirkung zwischen der Standardsprache einerseits und den Fach- und Gruppensprachen andererseits jedoch keinesfalls aus dem Blickfeld geraten darf.

In einem ersten Schritt erfolgt die definitatorische Annäherung an den Neologismus aufgrund seiner beiden Haupterscheinungsformen, dem Neulexem und der Neubedeutung. In einem zweiten Schritt wird dann anhand eines erweiterten definitatorischen Angangs, der vor allem auch auf den prozessualen Charakter der Erscheinung der Neologie Bezug nimmt, die Einbindung in den Gesamtbereich sprachlicher Innovation vorgenommen.

Ausgangspunkt ist die lexikalische Einheit, die als bilaterales Zeichen aus Ausdrucks- und Inhaltsseite, aus *Form* und *Bedeutung* konstituiert ist. Von den etablierten, seit langem gebräuchlichen Wortschatzeinheiten unterscheidet sich der *Neologismus* dadurch, daß

1. entweder seine *Form* **und** seine *Bedeutung* oder
2. **nur** seine *Bedeutung*

von der Mehrheit der Angehörigen der Kommunikationsgemeinschaft von einem (mehr oder weniger exakt bestimmbar) Zeitpunkt an und über eine längere oder auch nur kurze Zeitphase hinweg als **neu** empfunden wird. Die Neuheitsspezifik, die eine Wortschatzeinheit als Neologismus ausweist, kann sich also 1. auf das Zeichen in seiner formalen und inhaltlichen Gesamtheit oder 2. nur auf eine seiner beiden Seiten, nämlich auf seine inhaltliche Seite (Bedeutung) beziehen.

Zu unterscheiden sind demgemäß grundsätzlich **zwei Neologismustypen**: das Neulexem und die Neubedeutung.

(1) NEULEXEM:

= neue lexikalische Einheit (Einwortlexem, Wortgruppenlexem, Phraseologismus, Abkürzung/Kurzwort sowohl indigener als auch entlehnter Natur), bei der Form und Bedeutung neu sind,

Auschwitz-Leugnung, Infopost, Infotainment, Osterweiterung, kollektiver Freizeitpark, ICE.

Grundsätzlich neu können die Erscheinungen, Sachverhalte etc. sein, die dementsprechend erstmals zu benennen sind. Oder die neue Wortschatzeinheit ist eine Neubezeichnung, also eine neue weitere Bezeichnung für bereits vorhandene, bisher jedoch anders benannte Erscheinungen, Sachverhalte etc. In letzterem Fall ist bei dem Neulexem (vgl. z.B. neugebildet *Auszubildender/Raumpflegerin* gegenüber *Lehrling/Putzfrau*) jeweils nach einer eventuellen euphemistischen, pejorativen, meliorativen oder anderen Komponente zu fragen, die von konstitutiver Bedeutung für die Entstehung des Neulexems gewesen sein kann.

Je nach Bildungsart bzw. -modell der Neulexeme können als Untergruppen unterschieden werden:

Neuschöpfung:

ohne analoge Wortbildungsmuster; sehr selten;

Neuprägung oder Neubildung:

nach indigenen oder entlehnten Wortbildungsmustern; aus indigenen und/oder entlehnten Morphemen/Wörtern, in der Regel Komposita und Derivate; sofern es sich um Entlehnungen aus anderen Sprachen handelt, kann auch von **Neuentlehnung** gesprochen werden.

Die Bezeichnung *Neuwort* wird oft ohne weitere Differenzierung entweder als Synonym für *Neologismus* oder – noch häufiger – im Sinne der hier vorgeschlagenen Bezeichnung *Neulexem* verwendet. Zur Gebrauchsvielfalt kommt hinzu, daß das WDG in seiner Terminologie mit *Neuwort* die äußerst seltene Form der weitgehend vorbildlosen Neuschöpfung bezeichnete. So erweist sich eine Gebrauchsdifferenzierung zwischen *Neuwort* und *Neulexem* als sinnvoll, ja als notwendig: Da *Neuwort* deutlich auf das (neologische) Einzelwort abhebt, das neue Wortschatzelement aber durchaus auch mehrteilige Strukturen aufweisen kann (Wortgruppenlexem, Phraseologismus), erscheint die hier gewählte Bezeichnung *Neulexem*, in der solche unterschiedlichen Strukturen integriert sind, der Sache angemessener. *Neuwort* kann allerdings durchaus als nicht fachspezifische allgemeine Bezeichnung sowohl für einen Okkasionalismus wie für das *Neulexem* (und zwar ausschließlich in der Form des Ein-

zellextens) verwendet werden. Es ist jedoch darauf zu achten, daß beide Bezeichnungen, die allgemeinere *Neuwort* und die spezifische *Neulexem* nicht miteinander verwechselt werden.

(2) NEUBEDEUTUNG oder *Neusemem* (auch *Neosemantismus*³¹):

= neue Bedeutung, die einer bereits vorhandenen mono- oder polysemen lexikalischen Einheit hinzugefügt wird (Bedeutungserweiterung)

abfackeln, Optik, Schlüssel, alternativ.

Die Neubedeutung ist das Produkt eines semantischen Wandels, der meist die Form der semantischen Erweiterung bei einer etablierten mono- oder polysemen Wortschatzeinheit hat (*Bedeutungserweiterung*). Diese semantische Erweiterung kann auch mit grammatischem Wandel, d.h. mit neuen grammatischen Gegebenheiten verbunden sein.

Zu beachten sind bei der Neubedeutung jeweils mögliche Differenzierungen wie: gleichwertig, dominant, untergeordnet etc. im Hinblick auf die bisher vorhandene Bedeutung bzw. auf mehrere schon etablierte Bedeutungen.

Neubedeutungen können auch auf fremdsprachliche Einflüsse zurückgehen (Untergruppe der Neubedeutung: die *Bedeutungsentlehnung*). Vor allem bei Verben kann die Neubedeutung auch mit grammatischem Wandel (Valenzänderung) verbunden sein. Ausschließlich grammatischer Wandel wird jedoch in der Regel nicht als Neologismus eingestuft.

Der Prozeß der Herausbildung von Neubedeutungen ist in der Regel wesentlich langwieriger, diffiziler und somit oft auch schwerer erkennbar im Vergleich zum Entstehungs- und Durchsetzungsprozeß von Neulexemen.

Lediglich kontextuell bedingte Varianten einer etablierten Bedeutung sind keine Neubedeutung. Weiterhin gilt: Sofern bei einer polysemen Wortschatzeinheit lediglich eine bereits vorhandene (möglicherweise sogar bisher nur untergeordnete bzw. im Sprachgebrauch selten realisierte) Bedeutung gegenüber anderen Bedeutungen im Rahmen einer Neuentwicklung in auffälliger Weise dominant wird und somit andere Bedeutungen möglicherweise stark (oder auch völlig) zurückdrängt, liegt gemäß unserer Neologismus-Definition

³¹ Der der romanistischen Linguistik entstammende Terminus *néosemantisme/Neosemantismus*, sehr wahrscheinlich in Analogie zu *néologisme/Neologismus* gebildet, betont wohl mehr die eigenständige Rolle der Neubedeutung und somit eher das Trennende als das Verbindende der Neologierscheinungen. Wird – wie in unserem Definitionsvorschlag – die Neubedeutung als lexikalische Neuerungsform dem Neologismus zu- bzw. untergeordnet, erscheinen die Bezeichnungen *Neusemem* oder *Neubedeutung* eindeutig und angemessen. Vgl. auch Anm. 26.

kein Neologismus des Typs Neubedeutung vor, sondern eine im Rahmen lexikalischer Innovation untergeordnete Neuerung.

Neologismus ist aufgrund der (durch die Komponente *neo-* bedingte) kategorialen Einordnung der bezeichneten lexikalischen Einheiten nach ihrer Entstehungszeit ein relativer und historisch gebundener Begriff. Die Bezugnahme auf den (mehr oder weniger exakt zu eruiierenden) Zeitpunkt der Schöpfung sowie auf die sich anschließende Entfaltungs- und Durchsetzungsphase (bis zur Lexikalisierung) hat für jeden Neologismus einen hohen definitionsbestimmenden Stellenwert. Folglich ist es für die lexikographische Bearbeitung von Neologismen wichtig, sich relativ genau über die Zeitphase (bzw. ihre Eingrenzung) im klaren zu sein, in deren Rahmen Neologismen erfaßt und lexikographisch dargestellt werden sollen.

Wie bereits ausgeführt, wird der Neologismus-Begriff hier auf den allgemeinsprachlichen Wortschatz der gegenwärtigen deutschen Standardsprache als den lexikographisch vorrangig interessierenden Bereich bezogen. Er kann in modifizierter Form aber weitgehend auch für Fach-, Gruppen- und Sondersprachen mit ihren jeweiligen Geltungsbereichen angewendet werden.

Im folgenden Definitionsvorschlag werden alle bisher erörterten Gesichtspunkte zusammengeführt:

Ein **Neologismus** ist

eine **ganz neue** lexikalische Einheit in ihrer Gesamtheit aus Form und Bedeutung (NEULEXEM),

oder

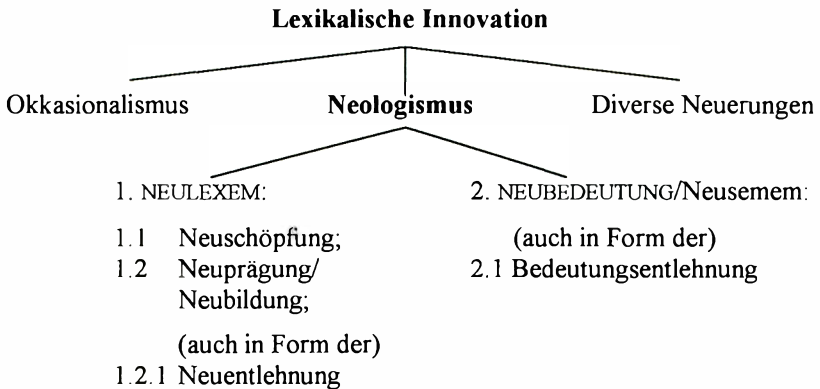
eine **ganz neue** (zum Vorhandenen hinzukommende) Bedeutung einer etablierten lexikalischen Einheit (NEUBEDEUTUNG),

- die zunächst noch in keinem Wörterbuch steht;
- die in einem bestimmten Abschnitt der Sprachentwicklung in einer Kommunikationsgemeinschaft vor allem (aber nicht ausschließlich) aufgrund kommunikativer Bedürfnisse aufkommt und sich ausbreitet,
- die in den allgemeinsprachlichen Wortschatz der Standardsprache übernommen (**Usualisierung**),
- als sprachliche Norm allgemein akzeptiert (**Akzeptierung**),
- sodann lexikographisch gespeichert (**Lexikalisierung**)
- und die – innerhalb dieses gesamten Entwicklungsprozesses – von der Mehrheit der Sprachbenutzer über eine gewisse Zeit hin als **neu** empfunden wird.

Jeder Neologismus durchläuft also folgende Phasen: Entstehung → Usualisierung → Akzeptierung → Lexikalisierung/Integration (Speicherung als Bestandteil des allgemeinen Wortschatzes). Somit ist der Neologismenbegriff nicht statischer, sondern durchaus prozessual-dynamischer Natur.

Jeder Neologismus ist in seiner Ur- bzw. Entstehungsphase zunächst ein Okkasionalismus (eine Individualbildung, eine Ad-hoc-Bildung). Einsetzend mit Usualisierung und Akzeptierung und schließlich und endlich mit der Lexikalisierung kommt der Integrationsprozeß des Neologismus zum Abschluß. Der bisherige Neologismus kann oder sollte von diesem Zeitpunkt an nicht mehr als solcher bezeichnet werden (es sei denn im historisch relativierenden Zusammenhang: 'ein Neologismus der achtziger Jahre'). Der bisherige Neologismus ist nun Bestandteil des allgemeinsprachlichen Wortschatzes der Standardsprache geworden; er wird **nicht mehr** als **neu** empfunden. Der Prozeß vom Aufkommen eines Neologismus bis zu seiner Akzeptierung durch die Sprachgemeinschaft kann von unterschiedlicher Dauer sein.

Es erscheint sinnvoll, den Begriff des Neologismus im Rahmen der lexikalischen Innovation, das heißt: der Gesamtheit lexikalischer Neuerungsmöglichkeiten exakt zu positionieren. Im Zusammenhang der im Prozeß von Wortschatzwandel und -entwicklung auftretenden verschiedenartigen lexikalischen Neuerungen können unterschieden werden:



Usualisierung³², Akzeptierung, Lexikalisierung und somit Integration sind wesentliche Abgrenzungskriterien des Neologismus gegenüber dem **Okkasio-
nialismus** (andere Bezeichnungen, bei teilweise unterschiedlicher Akzentuierung: *ad hoc-*, *Einmal-*, *Individual-Bildung*), der – was aber auch bei Neologismen der Fall sein kann – zudem seine Existenz meist keinem allgemeinen kommunikativen Bedürfnis verdankt.

³² Vgl. hierzu auch Jordanova (1994b, S. 166). Anstelle von *Usualisierung* wird hier der Begriff *Aktivierung* verwendet.

Zu den **diversen Neuerungen** können u.a. gezählt werden Wandel- bzw. Entwicklungserscheinungen in der:

- Grammatik (z.B. Valenzwandel, *jmdn. motivieren*);
- Frequenz (Gebrauchszunahme/-abnahme von Lexemen),
- Bedeutungsdominanz (Schwankungen bei polysemen Lexemen),
- Konnotation, Bewertung (Pejorisation, Meliorisation).

Eine bei einem etablierten Wortschatzelement neu aufgekommene starke Frequenzzunahme (z.B. durch gesteigerte kommunikative Relevanz oder wegen des Gebrauchs als Modewort) ist also ebenso wie ein neu zu beobachtender deutlicher Konnotations- oder Wertungswandel **kein** ausreichendes Kriterium für den Status eines Neologismus. Diese Erscheinungen zählen im Rahmen der lexikalischen Innovation zu den untergeordneten Neuerungen.

Bisher generell oder – sofern neu aufgekommen – zunächst streng fach- oder anderssprachliche Wortschatzelemente, die dabei sind, auch im allgemeinsprachlichen Wortschatz der Standardsprache integriert zu werden, zählen zu den allgemeinsprachlichen Neologismen, sofern sie alle weiteren der genannten Kriterien des Neologismus erfüllen.

Über Orte und Zeiten, die (scheinbar) für das Aufkommen lexikalischer Innovationen, vor allem natürlich von Neologismen besonders disponiert sind, ist vielfach nachgedacht worden. Permanente Motive für das Entstehen von Neologismen wirken in nahezu allen der ständig von Wandlungen und Entwicklungen geprägten Sach-, Fach- oder Lebensbereiche. Für die Gegenwart wären als besonders markant vielleicht Wirtschaft, Technik, Medizin, Sport, Mode, Unterhaltung und Freizeit zu nennen – oder eben andere vergleichbare Bereiche. Letztlich können Neologismen immer und überall entstehen, und es darf bezweifelt werden, ob Zeit und Ort besonders intensiver Produktivität aufgrund intensiven Nachdenkens tatsächlich exakt oder eben nur spekulativ benennbar sind.

Richtig ist wohl, daß vor allem in Zeiten einschneidender gesellschaftlicher, politischer oder wirtschaftlicher Umbrüche offenbar eine dringliche Notwendigkeit zur Benennung und Bezeichnung neuer Sachverhalte, Einrichtungen, Vorgänge etc. oder auch (etwa aus ideologischen Gründen) zur Um- und somit Neubenennung von (längst sprachlich benanntem) Vorhandenem gegeben ist. Aufkommen und Entwicklung von Neologismen sind aber nicht nur und vermutlich auch gar nicht vorrangig an solche historischen Umbruchsituationen gebunden. Viel eher handelt es sich um einen kontinuierlichen Prozeß, der – einmal deutlicher, dann wieder weniger intensiv – an beliebigem Ort und zu jeder Zeit stattfinden kann. Ganz entscheidend innerhalb dieses Prozesses ist allerdings immer, daß die sprachliche Innovation über den Urproduzenten

bzw. Erstverwender hinaus innerhalb einer mehr oder weniger kurzen Zeitphase von ganz unterschiedlichen und immer größeren Sprachteilnehmer-Kreisen aufgenommen und weitertransponiert wird.

Dabei können die Gründe für diese Transporttätigkeit ganz unterschiedlicher Art sein: Es kann sich um eine sachlich oder auch politisch-gesellschaftlich erforderliche bzw. unumgängliche Bezeichnungsnotwendigkeit handeln (neues Denotat), andererseits aber ebenso um eine gar nicht unbedingt notwendige sprachliche Neuerung, die als solche aber Sprechern und Schreibern ganz einfach gefällt, und die von ihnen weitertransponiert wird, weil sie diese für besonders passend, für originell, schick und modern, eben für neu und somit für auffällig und unabgenutzt, vielleicht auch für praktisch oder provozierend halten. Träger der Transportfunktion im heutigen mediengeprägten Zeitalter sind in allererster Linie die massenmedialen Textproduzenten, die Vorbildfunktion für alle möglichen anderen Bereiche gesprochener und geschriebener Sprache haben. Ohne sie liefe im heutigen „Neologismengeschäft“ vermutlich sehr viel weniger.

Der Prozeß der Etablierung eines Neologismus läßt sich unter anderem an seiner Gebrauchshäufigkeit und deren Kontinuität beobachten. Hinsichtlich der Aussagekraft der Belegfrequenzen von Neologismen muß aber berücksichtigt werden, daß der Status von Neologismen und der Grad ihrer Etablierung nicht automatisch und unbedingt mit einer hohen Belegdichte einhergehen müssen. Diese wird jeweils unmittelbar vom Denotat des Neologismus abhängen, wobei sich erhebliche Unterschiede in der Gebrauchsnotwendigkeit des Neologismus zwangsläufig aus der Natur der bezeichneten Sache(n) und dem jeweiligen Zugehörigkeitsbereich ergeben, was sich sodann in erklärbar unterschiedlich hohen oder eben auch niedrigen Frequenzen auswirken wird. Über den Etablierungsgrad von Neologismen können im übrigen auch zum Neologismus gebildete Ableitungen und Komposita Aussagewert besitzen.

Die Frage nach den besonders neologismusträchtigen Quellen und Sprachbereichen hat ihren Sinn natürlich vor allem auch im Hinblick auf die Zusammenstellung von spezifischen, die Neologismenlexikographie unterstützenden Computerkorpora. Über eher generalisierende Antworten (etwa: Texte hochfrequenter, möglichst benutzer-unspezifischer Massenmedien) wird man dabei vermutlich aber kaum hinauskommen können. Es sei denn, man begibt sich in den Bereich des Spekulativen. Für die eigentliche Arbeit des Neologismenlexikographen ist die Frage nach den üppigsten Quellen für Neologismen ohnehin keine von Priorität.

2. Neologismenlexikographie

Die lexikographische Bearbeitung und Darstellung der Neologismen einer bestimmten Zeitphase ist Aufgabe des Neologismenwörterbuchs. Dieses kann definiert werden als ein einsprachiges Bedeutungswörterbuch in der Form des Spezialwörterbuchs zu einem bestimmten Wortschatz-Ausschnitt, nämlich dem der Neologismen, das heißt der in einem bestimmten (von den Lexikographen festgelegten) Zeitabschnitt aufgetretenen und sich im allgemeinen sprachlichen Wortschatz der Standardsprache weithin „ablesbar“ festigenden Neulexeme und Neubedeutungen.

Neologismenwörterbücher, die es für das Deutsche, von den hier einleitend genannten Vorformen abgesehen, bisher nicht gegeben hat, haben in vielen Ländern eine mehr oder weniger lange Tradition und nehmen dort einen wichtigen Platz unter den lexikographischen Produkten ein. Die längste neologismenlexikographische Tradition, auch darauf wurde hier bereits hingewiesen, kann Frankreich vorweisen. Auch für die frühesten englischen Wörterbücher wird geltend gemacht, so von Algeo (1993), daß sie im Grunde Neologismenwörterbücher gewesen seien:

The history of English lexicography begins with the study of neology. The first monolingual English dictionaries, the „hard word“ books, recorded words entering the language as a result of the knowledge explosion of the Renaissance. Robert Cawdrey's *Table Alphabeticall* of 1604, usually cited as the first English dictionary, was intended to teach „the true writing, and vnderstanding of hard vsuall English wordes, borrowed from the Hebrew, Greeke, Latine, or French, & c.“ It and its congeners were, in effect, new-word books. (S. 281)

Der entlehnte Neologismus, seinerzeit die dominante Erscheinung und lexikalischer Ausgangspunkt für die frühe englische Lexikographie, spielt, wie sich zeigen wird, in der jüngeren englischen Neologismenlexikographie nur noch eine untergeordnete Rolle.

Als Auslöser der modernen englischen Neologismenlexikographie betrachtet Algeo den 2. Weltkrieg. In seinem zeitlichen Umfeld erschienen mehrere Neologismenwörterbücher (u.a. Berg 1953, Reifer 1955),³³ die für das Englische eine bis heute ununterbrochene Folge von Neologismenwörterbüchern unterschiedlichen Typs und Anspruchs eröffneten. Für die 2. Hälfte unseres Jahrhunderts ist aber auch generell im internationalen Rahmen eine Zunahme neologismenlexikographischer Bemühung und Produktion (mit bestimmten lokalen Schwerpunkten) zu registrieren. Sie läßt Rückschlüsse zu zum einen im Hinblick auf offenbar starke lexikalische Wandlungsprozesse im Rahmen neuer, weltweiter politischer, wirtschaftlicher, technischer und sozialer Entwicklungen und auf deren kommunikative Vermittlung auf der Grundlage

³³ Bereits während des Krieges erschienen: Taylor, A. Majorie (1944): *The Language of World War II*. 2. Aufl. 1948. New York.

neuartiger publizistisch-medialer Instrumente, zum anderen aber auch auf ein gesteigertes Öffentlichkeitsinteresse am neologischen Prozeß und an seinen Ergebnissen, ein Interesse, das über den fachinteren Bereich weit hinausreichen dürfte.

Heller u. a. (1988, S. 16-69) geben einen umfassenden, detaillierten Überblick über die gesamte jüngere internationale Neologismenlexikographie, und zwar anhand von nicht weniger als 60 fremdsprachigen Neologismenwörterbüchern. Nun hat sich gerade die englische Neologismenlexikographie seitdem weiterhin durch eine Fülle neuer Publikationen profiliert, so daß sie im internationalen Maßstab heute eine führende Position einnimmt: „Neology has become a growth industry“, schreibt Algeo (1993, S. 281). Deshalb soll hier im Anschluß an die Ausführungen bei Heller u. a. (1988) die jüngere Entwicklung der englischen Neologismenlexikographie anhand ihrer Ergebnisse im Überblick skizziert werden.

2.1 Englische Neologismenlexikographie als Vorbild

Die Publikationen der englischen Neologismenlexikographie können in vier Typen unterschieden werden, für die im einzelnen sich teilweise überschneidende Kriterien unterschiedlicher Art konstitutiv sind. Es handelt sich zum einen um die formalen Kriterien der Eigenständigkeit (2, 3, 4), der Unselbstständigkeit (1) oder der Periodizität (4) der Publikationen, zum anderen um inhaltliche Kriterien der mehr (1, 2, 4) oder weniger (3) wissenschaftlich exakten Präsentation der Neologismen.

1. (Unselbständig:) Ergänzungsbände zu großen umfassenden Wörterbuchunternehmen:

- Zum *Oxford English Dictionary* (OED):
Burchfield (1972-1986); Simpson u. a. (1993);
- Zu *Webster's Third New International Dictionary* (Webster):
Mish u. a. (1976, 1983, 1986);
- Zu *The Macquarie Dictionary*, Sydney (1981):
Butler (1990).

Die Ergänzungsbände zu den großen historischen und gegenwartssprachlichen Wörterbüchern entsprechen deren Standards und damit ihrem hohen lexikographischen Niveau. Neologismenwörterbücher sind sie bedingt in dem Sinne, daß diese Ergänzungsbände **auch** und durchaus in nicht geringem Umfang Neulexeme und Neubedeutungen präsentieren, die sich seit dem Erscheinen des Stammwörterbuchs im allgemeinen Sprachgebrauch etabliert haben. Daneben finden sich in diesen Ergänzungsbänden in größerem oder auch geringe-

rem Umfang aber auch Ergänzungen bzw. Korrekturen im Hinblick auf Lücken, Mängel oder auch Fehler in den jeweiligen Stammwörterbüchern. Aufgrund dieser makrostrukturellen Heterogenität kann bei diesen Ergänzungsbänden in der Regel nicht vom reinen Typ des Neologismenwörterbuchs gesprochen werden, obwohl sie insgesamt sehr reiches neologisches Material enthalten.

Das gilt zunächst in vollem Umfang für die Supplement-Bände zur 1. OED-Auflage (Burchfield 1972–1986). Sie sind unentbehrlich aufgrund ihrer detaillierten Angaben und der Fülle der Belege zu den hier gebuchten (seit 1882 aufgetauchten) „Neologismen“. Sie bearbeiten lexikographisch jedoch eine Periode von ca. 100 Jahren, insgesamt also eine viel zu lange Zeitdauer, um im Hinblick auf Neologismen von entschieden aktueller Qualität sein zu können und somit dem lexikographischen Typ Neologismenwörterbuch exakt zu entsprechen. Algeo (1990) verweist für Burchfield zudem auf Lücken: „The *Supplement* necessarily misses a good many new words, because they are too recent for its publication dates or do not fit the profile that Burchfield uses in choosing words for inclusion“ (S. 211).

Im Anschluß an die 1989 erschienene 2. Auflage des OED soll nun etwa im 2-Jahres-Rhythmus jeweils ein neuer Band der OED *Additions Series* ediert werden, wie es im Vorwort des vorliegenden 1. Bandes (Simpson u. a. 1993) angekündigt wird. Auch bei diesen Bänden der OED-Ergänzungsserie handelt es sich erklärtermaßen keineswegs um Neologismenwörterbücher im engeren Sinne:

„The *Additions* volumes should be regarded as the result of work-in-process on new entries for the *OED*. They are not intended individually to present a balanced picture of recent additions to the language; a fuller picture will establish itself as the series unfolds. It will in fact be noted that the majority of entries in each volume are not neologisms, but represent a heterogeneous collection of accessions to the language over the past few centuries. Some are surprising omissions from earlier editions of the Dictionary, others cover subjects which have recently gained in prominence (...).“ (S. X)

Eindeutiger auf Neologismen orientiert sind die Ergänzungsbände zum *Webster* (Mish u. a. 1976, 1983, 1986). Repräsentiert das OED den Typ des großen historischen Wörterbuchs, so ist der *Webster* das deutlich gegenwartssprachbezogene lexikographische Großunternehmen. Je größer der Abstand vom Erscheinungsjahr 1961 wurde, desto spürbarer machten sich für den *Webster* die lexikographischen Lücken im Hinblick auf Wortschatzneuerungen bemerkbar. Für Abhilfe sorgen die drei zwischen 1976 und 1986 erschienenen Supplement-Bände, in denen in kumulativer Vorgehensweise zunächst 6 000, dann 9 000 und zuletzt 12 000 (neue) Wörter erfaßt und beschrieben sind – die Tausenderzahlen werden jeweils wirkungsvoll im Titel genannt. Jeder Folgebände enthält also den größten Teil des Wortmaterials seines Vorgängerbandes.

des und fügt diesem den inzwischen neu in Gebrauch gekommenen Wortschatz hinzu, so daß die Folgebände im Grunde genommen jeweils (allerdings wesentlich) erweiterte Neuauflagen des Vorgängerbandes darstellen. Da die *Webster-Supplements* im Unterschied zu denen des OED anderweitige Änderungen, Ergänzungen und Verbesserungen zum Stammwörterbuch nur in recht geringem Umfang enthalten, geben sie insgesamt einen konzentrierteren Überblick über Neuerungen im Wortschatz seit dem Erscheinungsjahr des Stammwörterbuchs (1961), also in einem im Hinblick auf Neologismen angemessenen, überschaubaren Zeitabschnitt. Somit stehen sie dem Typ Neologismenwörterbuch deutlich näher.

Hingewiesen sei im Zusammenhang der unselbständigen Neologismenwörterbücher schließlich noch auf den Ergänzungsband zum 1981 in Sydney erschienenen *Macquarie Dictionary* (Butler 1990), der in verdienstvoller Weise die Neologismenrecherche und -präsentation auf die australische Varietät des Englischen ausweitet. Ayto vermerkt in seiner Rezension zu Recht (1992, S. 238f.), daß Butler (1990) zwar formal zu den an große Stammwörterbücher gebundenen Ergänzungsbanden zu zählen ist, in der makro- und mikrostrukturellen Darbietung seines Materials dem Stammwörterbuch gegenüber aber eher modifizierte und eigenständige Wege geht, so daß es durchaus auch zu den eigenständigen anspruchsvollen Neologismenwörterbüchern gerechnet werden kann. Der *Macquarie*-Band enthält im Unterschied zu den OED-Supplements keinerlei anderweitige Ergänzungen oder Verbesserungen zum Stammwörterbuch und ist somit ein weitgehend eigenständiges reines Neologismenwörterbuch.

2. Eigenständige Neologismenwörterbücher auf lexikographisch hohem Niveau:

Ayto (1989/1990), R. Barnhart u.a. (1990 [1973, 1980]), Mort (1986).

Algeo (1990, S. 211) spricht in bezug auf die englische Lexikographie vom „broad genre of new-word-glossaries“. Er unterscheidet dabei „two main subcategories: the evidential and the popular“. Als entscheidendes Kriterium des beweiskräftigen Wörterbuchs („evidential new-word dictionary“) benennt Algeo die in die Wortartikel integrierten Belege, und zwar „quotations with adequate bibliographical identification – not merely to illustrate but especially to document the existence of the word and what the lexicographer has said about it. Evidential dictionaries are intended for those with a professional, scholarly, or at least serious, empirical interest in the lexis“ (S. 211). Algeos Typ des „evidential new-word dictionary“ ist weitgehend deckungsgleich mit dem hier angesetzten Typ des eigenständigen Neologismenwörterbuchs auf lexikographisch hohem Niveau. Er erstreckt sich bei ihm (natürlich zu Recht) allerdings auch auf die von mir separat unter 1 zusammengefaßten (un-

selbständigen) Ergänzungswörterbücher, von denen Algeo besonders die OED-Supplements (Burchfield 1972-1986) würdigt.

Unter den selbständigen englischen Neologismenwörterbüchern ragen als mustergültig die Publikationen des „Barnhart enterprise“ (Algeo 1990, S. 211) heraus. Die drei in sich selbständigen Bände C.L. Barnhart u.a. (1973,³⁴ 1980) und R. Barnhart (1990)³⁵ erfassen in makro- wie mikrostruktureller Hinsicht auf vorbildliche Weise die allgemeinsprachlichen Neologismen der (vom Erscheinungsdatum aus gerechnet) jeweils letzten zehn Jahre (in etwa). Dabei verstehen sich die Nachfolgebände (1980, 1990) jeweils als Fortsetzung der vorausgegangenen Ausgaben. In gewisser Weise stellen sie stark erweiterte und revidierte Neuauflagen dieser vorausgegangenen Bände dar, indem sie viele Einträge dieser früheren Bände – soweit erforderlich – nach dem jeweils neuesten Entwicklungsstand im behandelten zehnjährigen Erfassungszeitraum fortschreiben.

Die Barnhart-Neologismenwörterbücher geben anhand von jeweils mehr als 10 000 Neulexemen und Neubedeutungen, die in der Regel vorher noch nicht in Wörterbüchern verzeichnet waren,³⁶ einen umfassenden und zuverlässigen Überblick über die Neologismen des jeweils abgelaufenen Jahrzehnts. Sie dürften nicht zuletzt aufgrund der ausführlichen, exakten Gebrauchsbeschreibungen und der vorzüglichen Qualität ihrer Belegteile auch für andere lexikologische und lexikographische Unternehmungen von hohem Interesse und Wert sein.³⁷

Zu den eigenständigen Neologismenwörterbüchern auf lexikographisch hohem Niveau zählen weiterhin die *Longman Register of New Words* (Ayto 1989³⁸ und Ayto 1990), die im *Longman Guardian New Words* (Mort 1986) eine Art Vorläufer hatten.³⁹ Ayto zählt seine Wörterbuch-Unternehmungen zu den

³⁴ Zu C.L. Barnhart (1973) vgl. ausführlich bei Heller u.a. (1988, S. 26-30).

³⁵ Zu R. Barnhart u. a. (1990) vgl. auch die Rezension von Linda L. Rapp (*American Speech* 1992, H. 2, S. 205-207).

³⁶ Zu ihren Auswahlkriterien im Hinblick auf die Aufnahme eines Neologismus ins Neologismenwörterbuch äußern sich R. und C. Barnhart an anderer Stelle (1990, S. 1160): „In the process of selection, ten of thousands of terms collected over a set period are carefully examined and evaluated for their appropriateness as dictionary entries, their frequency of appearance, the range of sources in which they appear, their cruciality and importance, the length of time over which they reoccur, their topicality, and similar properties. Items that do not meet the standards set by these tests of importance and interest are excluded.“

³⁷ Zum Periodikum *The Barnhart Dictionary Companion: A Quarterly to Update General Dictionaries* (C.L. Barnhart 1988ff.) vgl. unter 4.

³⁸ Zu Ayto (1989) vgl. die Rezensionen von Algeo (1990) sowie von Uwe Carls (*Zeitschrift für Anglistik und Amerikanistik* 1992, H. 1, S. 65-66).

³⁹ Mort (1986) enthält Neologismen des Zeitraums 1985/1986 größtenteils aus britischen Quellen, die in Ayto (1989 und 1990) nicht nochmals erfaßt sind; in ihrem Wortbestand sind die Longman-Publikationen also jeweils selbständig.

„‘freestanding’ neologism dictionaries“ (1992, S. 238) und charakterisiert sie wie folgt:

The aim of the *Longman Register of New Words* is to ... build up a rounded picture of the ways in which English has grown and developed over the years 1986 to 1988. ... Here are gathered together 1200 new pieces in the never-finished jigsaw of the English language. ... The *Longman Register of New Words* is a representative sample plotting the peaks in the graph of lexical change. Virtually every medium of human communication has made its contribution – conversation, books, radio and television broadcasts, films – but newspapers and magazines naturally assume a leading part, capturing innovation as they do and enshrining it in print. (1989, „Introduction“, o.S.)

Den von den Barnhart-Wörterbüchern gesetzten hohen Standards entsprechen die Longman-Publikationen allerdings nicht im vollen Umfang. Algeo (1990, S. 213) moniert beispielsweise für Ayto (1989) bei den Belegen die in den Quellenangaben durchgängig fehlenden Seitenzahlen, die eine Überprüfung der Belege in den meisten Fällen erheblich erschweren, wenn nicht unmöglich machen. Algeo verweist weiterhin (S. 215) auf die unscharfe Trennung zwischen echten Neubedeutungen und den nur kontextuellen Varianten einer etablierten Bedeutung, die nicht zu den Neologismen gezählt werden können. Die beiden Bände von Ayto erfassen in einheitlicher Form und Struktur Neologismen jeweils unterschiedlicher Zeiträume; Ayto (1990) ist also keine erweiterte Neuauflage von Ayto (1989). Eine gewisse Nähe der *Longman Register of New Words* zu den „popular new-word dictionaries“ (zu unserem Typ 3) konstatiert wiederum Algeo (1990, S. 214), und zwar vor allem im Hinblick auf die diskursiven Bemerkungen in vielen Wortartikeln.

3. Eigenständige Neologismenwörterbücher populärwissenschaftlichen Charakters:

LeMay u. a. (1985/1988); Mager u. a. (1982); Tulloch (1991);
bereits früher: Berg (1953), Reifer (1955).

Besonderheiten und Aufgaben der populärwissenschaftlichen Neologismenwörterbücher werden von Algeo (1990, S. 213) treffend so beschrieben: [Popular new-word dictionaries are] „glossaries that list and define new words and often other kinds of information about them, especially interesting etymological stories, but are without documentary citations. The popular glossaries are also sometimes chatty in tone and in other ways try to appeal to the general educated reader with an interest in words“.

Die ersten – zeitlich im Umfeld des 2. Weltkrieges angesiedelten – englischen Neologismenwörterbücher erfüllen lexikographisch nicht die Standards der späteren hier als Typ 1 und 2 vorgestellten Wörterbücher. Als populärwissenschaftliche Unternehmungen versuchten sie erstmals, auf das in weiteren (als

nur den fachinternen) Kreisen aufkommende oder auch bereits vorhandene Interesse an neuen Wortschatzentwicklungen einzugehen. Das Kriterium der unterschiedlichen Benutzer- bzw. Interessentenkreise, hier ein breiteres wohl an der Sprache, aber weniger an lexikographischer Feinheit und Gelehrsamkeit interessiertes Publikum, dort die Lexikologen und Lexikographen, war also von vornherein wesentlich für die Differenzierung in populärwissenschaftliche (popular) und lexikographisch anspruchsvolle (evidential) Neologismenwörterbücher. Solidität und Wert der (meisten) populärwissenschaftlichen Neologismenwörterbücher sollen durch solcherlei Differenzierung allerdings keinesfalls in Frage gestellt werden. Zumal ihnen ja die zeitlich führende Rolle vor den anspruchsvolleren Unternehmungen zukommt, waren und sind sie auch für die Fachleute von durchaus hohem Wert. So wird etwa Reifers frühes Neologismenwörterbuch (Reifer 1955) von Algeo gewürdigt als „substantial work of 234 double-columned pages and, discriminatively used, still a useful source for mid-century new words“ (1990, S. 213).

Spätestens seit den achtziger Jahren waren das fachliche und das öffentliche Interesse an neueren Wortschatzentwicklungen des Englischen so stark etabliert, daß es – in einer Art Wachstumsbranche⁴⁰ – kontinuierlich durch eine Fülle unterschiedlich strukturierter Neologismenwörterbücher befriedigt werden mußte. Den „potential market for new-word books“ (Algeo 1990, S. 213) erkundete als frühes populärwissenschaftliches Unternehmen dieser Periode zunächst Mager u.a. (1982). Es folgten die kleineren Wörterbücher von LeMay u.a. (1985) und (verbessert und erweitert 1988).

Das jüngste Produkt unter den populärwissenschaftlichen Neologismenwörterbüchern ist die Arbeit von Tulloch (1991)⁴¹, die in selektiver Auswahl mehr als 1000 Neulexeme und Neubedeutungen des vorangegangenen Jahrzehnts aus dem gesamten englischen Sprachbereich behandelt. Das lexikographisch sorgfältig erarbeitete Buch, das zu jedem Eintrag, wenn auch in nur geringem Umfang, Belege mit Quellenangaben enthält, kann als populärwissenschaftliches Wörterbuch auf sehr hohem Niveau gelten, das den hier als Typ 2 behandelten Wörterbüchern qualitativ nahesteht. Es erregt Interesse unter anderem auch durch Neuerungen im mikrostrukturellen Bereich. So werden beispielsweise jedem Eintrag piktogrammartige Zeichen („subject icons“) für den jeweiligen Herkunftsbereich des behandelten Neologismus zugeordnet. Wie für Tullochs Wörterbuch insgesamt, so trifft für die meisten populärwissenschaftlichen englischen Neologismenwörterbücher im allgemeinen zu, was Wolski in seiner Rezension so formuliert: „Es ist dies ein Wörterbuch nicht zum raschen Nachschlagen, sondern eines, in dem man gern liest, dessen Artikeltexte aus der Perspektive ganz unterschiedlicher Erfahrungsbereiche ein Schlaglicht auf das gegenwärtige bzw. kurz zurückliegende Zeitgeschehen werfen“ (S. 296).

⁴⁰ Algeo (1990, S. 213): „It was not until the 1980s that neology became a popular growth industry.“

⁴¹ Vgl. die Rezension von Wolski (*Lexicographica*, 1991, S. 295-296).

4. Periodisch erscheinende Neologismen-Publikationen auf lexikographisch hohem Niveau:

Algeo (1991) (= *Among the New Words* 1941ff.); C.L. Barnhart u.a. (1982ff.); Simpson u.a. (1993ff.)

Die englische Neologismenlexikographie kann sich glücklich schätzen, zu ihrer Thematik über inzwischen nicht weniger als drei anspruchsvolle Periodika gänzlich unterschiedlicher Art verfügen zu können.

Auf die wichtige Rolle des 2. Weltkrieges für das Interesse am englischen Neologismus wurde hier bereits mehrfach hingewiesen. Auch die erste der periodischen Publikationen gründet und beginnt in dieser Zeit, nämlich 1941, und sie ist somit die am längsten (nämlich bis heute) währende Dokumentation englischer Neologismen. Die Rede ist von der Kolumne *Among the New Words*, die (annähernd) regelmäßig in der Zeitschrift der American Dialect Society *American Speech* erscheint, und in der jeweils eine größere Zahl von Neologismen behandelt und anhand von Belegen dokumentiert wird. In Algeo (1991) sind alle bis dahin, also in nicht weniger als 50 Jahren erschienenen Folgen der Rubrik in chronologischer Reihenfolge als Reprint in Buchform zusammengestellt, alle behandelten Neologismen in einem alphabetischen Register erfaßt und in einem aufschlußreichen Vorwort in zusammenfassender Übersicht wissenschaftlich kommentiert und aufbereitet worden. Algeo wertet das Unternehmen *Among the New Words* als „the most devoted and successful effort ever to chronicle continuously the rise of new words in the language“ (1990, S. 212) und beschreibt die Kolumne als „a record of such new words as have not yet made their way into dictionaries at the time the column is being prepared“ (1991, S. 3). Als solcher liegt der Wert dieser Kolumne (und der des Reprints) mit den hier behandelten Okkasionalismen und Neologismen vor allem darin, rückblickend differenzieren zu können, welche neuen Wörter okkasional geblieben sind, welche neuen Wörter ihren Weg in den allgemeinen Sprachgebrauch und in die Wörterbücher gemacht haben, und welche trotz zeitweiliger Etablierung letztendlich dennoch wieder außer Gebrauch gekommen sind.⁴²

Das zweite Periodikum ist eine Vierteljahresschrift aus dem Umkreis der Barnhart-Neologismenwörterbücher (C.L. Barnhart u.a. 1973, 1980 und R. Barnhart 1990). Es handelt sich dabei, so Algeo (1990), um „the only periodical devoted exclusively to neology“ (S. 212). Der *Barnhart Dictionary Companion* (C.L. Barnhart u.a. 1982ff.) enthält zum einen u.a. kürzere Essays zu unterschiedlichen lexikologischen und lexikographischen Themen, zum anderen aber, und das macht den Hauptinhalt der Zeitschrift aus, Wortartikel zu Neologismen (jeweils mit ausführlichem Belegteil). Jedes Heft enthält ein al-

⁴² Zu letzterem vgl. vor allem den Aufsatz von Algeo (1993).

phabetisches Register der in ihm behandelten Neologismen. Die Register werden jahrweise und im Zeitschritt von jeweils vier Jahren kumuliert.

Das dritte Periodikum schließlich ist die bereits oben unter Typ 1 vorgestellte *Additions Series* zum OED, deren Erscheinen auf einen Zeitrhythmus von ca. zwei Jahren konzipiert ist (Simpson u. a. 1993ff.). Hier werden Neologismen und Fragen der Neologie allerdings nur als ein Themenbereich unter vielen anderen behandelt.

In allen drei Periodika wird Neologismenlexikographie auf hohem Niveau praktiziert. Was die aufwendige Eruiierung der Neologismen betrifft, so stützen sich alle drei Unternehmungen auf sehr umfangreiche und offenbar ebenso differenziert ausgewählte wie gut instruierte und organisierte (ehrenamtliche) Zuträgerkreise, auf die sie zuverlässig zurückgreifen können.⁴³

Die sich inzwischen über einige Jahrzehnte erstreckende intensive englische Neologismenlexikographie ist aufgrund ihrer Ergebnisse befähigt, zusammenfassende und gesicherte Aussagen über Art, Aufkommen und Herkunft der von ihr behandelten Neologismen zu machen. Dies geschieht vorrangig in den häufig umfangreichen und aufschlußreichen Einleitungen der Neologismenwörterbücher, zum geringen Teil auch in Zeitschriftenbeiträgen. Insgesamt kann festgestellt werden, daß das Bedürfnis nach methodisch-theoretischer Beschäftigung mit dem Thema Neologismus und Neologismenlexikographie eher gering ausgeprägt zu sein scheint; vorherrschend ist der deutlich spürbare lexikographisch-praktische Zugriff.

In den Wörterbucheinleitungen dominieren (häufig sehr differenzierte) Aussagen zum einen über Entstehungs-, Bildungs-, Herkunftsart und zum anderen über Bereichszuordnungen der englischen Neologismen, teilweise auch im Vergleich der im jeweiligen Wörterbuch behandelten Neologismen mit denen früherer Perioden.

Von den sechs produktiven etymologischen Typen der Bildung von englischen Neologismen (meist mit mehreren Subtypen) ist die echte Neuschöpfung (CREATING) durchgängig der offensichtlich nicht nur im Englischen am wenigsten produktive Typ.⁴⁴ Der allergrößte Teil der Neologismen greift in der

⁴³ Vgl. hierzu z.B. Algeo (1991, S. 3) sowie Simpson u. a. (1993, S. ix-xi).

⁴⁴ Die folgenden Ausführungen greifen vor allem zurück auf Algeo (1991: „The making of new words“, S. 3-14), Ayto (1989: „Introduction“, o.S.) und auf C.L. Barnhart u. a. (1980), und hier speziell auf die 22 alphabetisch als Einträge in den Neologismen-Lexikonteil eingeordneten, zum Teil recht umfangreichen *Language Notes*, mit Titeln wie *Abbreviations, Coinages, Idioms, Technical Terms*. – Die Vermischung des alphabetischen Neologismenteils mit diesen linguistischen Beiträgen in Barnhart (1980) scheint mir ein kaum gelungener Versuch der Autoren, die Benutzer zusammenhängend und übersichtlich mit den wesentlichen linguistischen Daten des Themas Neologismus

einen oder anderen Art auf vorhandene, geläufige (indigene oder entlehnte) Wörter oder Wortsegmente zurück. Algeo (1991) hierzu: „Although the English vocabulary continues to grow at a prodigious rate, most new words are constructed from its existing resources“ (S. 14). Und an anderer Stelle: „The purest kind of creating would be to make a word completely from scratch, creation ex nihilo. Theoretically, it should be possible to make words in that fashion; but in fact there are no such words of which we can be sure. To make something out of nothing does not seem to be a human talent.“ (S. 4). Als (äußerst seltene) Neuschöpfungen gelten in der Regel laut- und klangimitierende (onomatopoetische) Wörter. Ansonsten kommt das dem Typ der Verkürzung zugehörnde Akronym der echten Neuschöpfung wohl noch am nächsten (vgl. das aus den Anfangsbuchstaben mehrerer Wörter gebildete Kurzwort *yuppie*, für ‘young urban professional’ + das Suffix *-ie*). Vom Gesamt der englischen Neologismen machen die Produkte der Neuschöpfung kaum 1% aus.

Die Entlehnung aus anderen Sprachen (BORROWING) spielt – wohl im deutlichen Unterschied zum Gegenwartsdeutschen und zu vielen anderen Sprachen – im Rahmen der englischen Neologieproduktion zwar eine kontinuierliche, heute insgesamt aber eher untergeordnete Rolle. Lehnwörter und Lehnübersetzungen machen insgesamt nur knapp 5% der englischen Neologismen aus. Nach Algeo (1990, S. 215) ist aufgrund eines Frequenzvergleichs der gebuchten Entlehnungen bei C.L. Barnhart u. a. (1973) und Ayto (1989) die Tendenz bei der Neuentlehnung rückläufig.

Den weitaus größten Anteil an den englischen Neologismen haben, hierin dem Deutschen vergleichbar, als Produkte der gängigen Wortbildungsprozesse (COMBINING) mit insgesamt ca. 62% die Komposita (ca. 35%) und die Derivate (ca. 27%), wobei nach Algeo (1990, S. 215) aufgrund von Frequenzvergleichen zwischen C.L. Barnhart u. a. (1973) und Ayto (1989) die Tendenz bei den Komposita zunehmend, bei den Derivaten eher rückläufig ist. Auch im Englischen bilden Substantive den weitaus größten Teil unter den Komposita und Derivaten.

Den, zwar mit deutlichem Abstand, aber immerhin zweitgrößten Anteil haben mit ca. 17% die Produkte von Wandelprozessen (SHIFTING) in den Bereichen Grammatik und – mit deutlichem Übergewicht (ca. 11%) – Bedeutung (Neubedeutungen). Zu einem nicht geringen Teil handelt es sich dabei um Metaphern und Metonymien.

vertraut zu machen. Ein übersichtlich gegliedertes Vorwort oder die Zusammenstellung der im übrigen aufschluß- und materialreichen *Language Notes* in einem eigenständigen alphabetisch sortierten linguistischen Lexikonteil wäre im Interesse der Benutzer wohl sinnvoller gewesen. R. Barnhart u. a. (1990) ist erfreulicherweise von dieser Praxis der Vermischung von Neologismeneinträgen und *Language Notes* wieder abgerückt.

Knapp 10% der englischen Neologismen sind Produkte von diversen Verkürzungsprozessen (SHORTENING), zu denen Abkürzungen, Akronyme, Initialwörter, Rückbildungen u.a. zählen.

Bei den verbleibenden ca. 5% handelt es sich um Neologismenprodukte aus Mischprozeduren (BLENDING), an denen die Prozesse COMBINING und SHORTENING simultan beteiligt sind. Neologismen dieser Art sind verschiedenartige Kompositionsformen der Verkürzung oder Kontraktion von (geläufigen) Wörtern und Wortteilen (*documentary + drama > docudrama*) oder ausschließlich von Segmenten zweier oder mehrerer (geläufiger) Wörter (*smoke + fog > smog, motor + hotel > motel*).

Die beiden zuletzt genannten Bildungstypen von Neologismen (SHORTENING/Verkürzung und BLENDING/Mischung) markieren mit ihren vielfach witzig-sprachspielerischen, dabei nicht selten lediglich okkasionellen Produkten einen jüngeren, vermutlich sprachmodischen, mit Sicherheit aber sprachökonomischen Trend vor allem im dominanten medien sprachlichen Bereich des Englischen. Dieser Trend kann für die derzeitige Neologismenproduktion im Englischen als ausgesprochen typisch und markant gelten, und zwar weniger im Hinblick auf die Frequenz (wo die konventionellen Formen der Zusammensetzung und Ableitung weiterhin sehr deutlich dominieren), als vielmehr in der formalen Spezifik. Bezugnehmend auf die Gesamtheit der etymologischen Typen innerhalb der englischen Neologie kommt Algeo (1990) beim Vergleich von C.L. Barnhart u.a. (1973) und Ayto (1989) resümierend zu dem Ergebnis: „The relative frequency of etymological types is fairly constant“ (S. 215).⁴⁵

Untersuchungen über die Dominanz bestimmter Herkunftsbereiche mit überdurchschnittlicher Neologismenproduktivität bleiben – was auch für andere Sprachen generell symptomatisch sein könnte – in ihren Ergebnissen für das Englische eher verschwommen und somit wenig aussagekräftig. Das wiederum dürfte vor allem für die lexikographische Neologismenrecherche kaum hilfreich sein.

⁴⁵ Für die Neologie im Deutschen macht Dou Xuefu (1989, S. 52ff.) ebenfalls fünf grundlegende etymologische Typen geltend. Dabei korrespondieren bei ihm mit dem Englischen die Neuentlehnung (BORROWING), die – bei Dou getrennt als zwei Typen behandelt – Ableitung und die Neuprägung (COMBINING), die Abkürzung (SHORTENING) und die Neubedeutung (SHIFTING). Der vor allem für das heutige Englisch charakteristische Typ BLENDING spielt für das Deutsche nur in der Form von Entlehnungen eine Rolle, also keine eigenständige, und fehlt dementsprechend bei Dou Xuefu. Zu den deutschen Neologismen zugrundeliegenden Prozessen Wortbildung, Entlehnung, Integration, Lehnwortbildung und Bedeutungswandel vgl. auch den Überblick bei v. Polenz (1991, S. 37–58).

Zwar scheint gesichert, daß der Bereich Wirtschaft mit Spitzenstellung vor allen anderen Bereichen sowie in der weiteren Rangabfolge dominante Bereiche wie Medizin, Technik, Politik, Unterhaltung und Soziales als besonders neologismenreich eingestuft werden können. Algeo (1990, S. 216; 1991, S. 15f.) benennt im weiteren eine derart große und entsprechend wenig differenzierte Fülle neologismenproduktiver Sachbereiche (insgesamt mehr als 30), daß man im Endeffekt eigentlich nur zu dem Allgemeinplatz gelangen kann: Neologismen können in mehr oder weniger großer Anzahl überall und immer auftreten. Andererseits werden (so etwa bei Ayto 1989) einzelne zeitgeschichtliche Ereignisse (Watergate, Politik der Thatcher-Ära, Reformen in Osteuropa, das gesamte Umfeld der Krankheiten Krebs und Aids) oder mit *lifestyle* und *behaviour* bezeichnete Lebensbereiche mit ihren spezifischen zeittypischen Verhaltens- und Umgangsformen als stark neologismenproduktiv herausgestrichen. Im Grunde genommen ist aber auch das eher trivial: denn natürlich hat jedes herausragende zeitgeschichtliche Geschehen, jedes neue Verhaltensmuster, hat jedes historische Ereignis sprachliche Neuerung notwendigerweise im Gefolge.

In ihrer Arbeit und deren Ergebnissen kann sich die englische Neologismenlexikographie durch ein reges öffentliches (und natürlich auch durch das fachinterne) Interesse an ihrem Thema bestätigt fühlen. Optimistisch resümierend stellt einer ihrer Repräsentanten fest: „As we head toward the end of the century, we can expect popular interest in language – focused upon lexis – to continue. Doubtless the attention to neology will intensify“ (Algeo 1990, S. 217).

2.2 Aufgaben der deutschen Neologismenlexikographie

Das Wörterbuch der Neologismen ist, wie R. u. C. Barnhart (1990) feststellen, „primarily a twentieth-century development in lexicography“ (S. 1162), genauer gesagt ist es ein Kind der mittleren zweiten Hälfte unseres Jahrhunderts. Über Anstöße, Vorformen und auch Angänge innerhalb der deutschen Lexikographie wurde einleitend berichtet. Knapp vor dem Ende dieses 20. Jahrhunderts könnte es ihr vielleicht doch noch gelingen, auf dem langen Weg zum deutschen Neologismenwörterbuch ein gutes Stück weiter voranzukommen und somit Anschluß an die (vielerorts aufkeimende oder bereits etablierte) internationale Neologismenlexikographie zu finden.⁴⁶

⁴⁶ Vor kurzem ist ein Neologismenwörterbuch des Tschechischen erschienen: Sochová, Zdeňka/Poštolková, Běla (1994): Co v slovnících nenajdete. Novinky v současné slovní zásobě [Was nicht in den Wörterbüchern zu finden ist. Neuerungen im gegenwärtigen Wortschatz]. Prag. Für das Rumänische berichtet Jordanova (1994a) über laufende Arbeiten an einem Neologismenwörterbuch. Die neologismenlexikographischen Bemühungen in diesen Ländern imponieren um so mehr, als sie in der derzeitigen wirtschaftlich desolaten Phase stattfinden.

Es ist wenig sinnvoll, über versäumte Notwendigkeiten im Bereich der deutschen Lexikographie weiter zu lamentieren.⁴⁷ Vielleicht lohnt es sich nicht einmal, über die Gründe zu spekulieren, warum es sich ausgerechnet die deutsche Lexikographie (zumal die der Bundesrepublik Deutschland) bisher leisten konnte, sich vom internationalen Forschungsstandard abzukoppeln. Wenn es nicht an mangelnden materiellen Voraussetzungen lag (und davon kann ausgegangen werden) und schon gar nicht am mangelnden fachlichen und öffentlichen Interesse, so muß von einer mehr oder weniger motivierten Absicht ausgegangen werden.

Was die großen kommerziellen Wörterbuchunternehmen betrifft, so vermutet Wiegand (1990, S. 2185), daß diese aufgrund deutscher Marktanalysen davon ausgehen, daß für Neologismenwörterbücher keine Marktlücke vorhanden sei. Dem könnte zweierlei widersprechen. Zum einen ist auch für Deutschland ein öffentliches Interesse am Wortschatzwandel, am neuen Wort mühelos nachweisbar, und zwar anhand von vielen regelmäßig in Massenmedien anzutreffenden thematisch einschlägigen Beiträgen (z.B. tägliche Rubriken wie das „Aktuelle Lexikon“ der *Süddeutschen Zeitung*, Sprachglossen u. ä.); vor allem aber ist es nachweisbar anhand der Vielzahl der oft in Taschenbuchform vorgelegten, meist satirisch-kritischen kleinen Wörterbücher zum aktuellen deutschen Sprachgebrauch mit Titeln wie *Amideutsch*, *Bösdeutsch*, *Dummdeutsch*, *In-Deutsch*, *Jubeldeutsch*, *Neudeutsch*, *Schicki-Micki-Sprache* oder *Szene-Sprache*.⁴⁸ Diese lexikographisch meist wenig anspruchsvollen und stets nur

⁴⁷ Auf die Unterentwicklung der deutschen Neologismenlexikographie verweisen mehr oder weniger nachdrücklich Müller (1987), Wiegand (1990) sowie Hausmann (1985, S. 390). Der Terminus *Neographie*, den Hausmann wohl als Kurzform für 'Neologismenlexikographie' verwendet und der verschiedentlich übernommen wurde, scheint mir unkorrekt und für die Sache nicht geeignet, zumal er seit dem frühen 19. Jh. semantisch anderweitig besetzt ist: vgl. 1813 bei Campe: „*Neographie*, die Neuschreibung, d.i. die Abweichung von dem Alten und Gewöhnlichen im Schreiben“ (zu Campe vgl. Anm. 19).

⁴⁸ Wir haben es hier mit einem Typ Wörterbuch zu tun, der allenfalls im Vorfeld seriöser, anspruchsvoller Lexikographie einzuordnen ist. Immerhin erhebt aber mancher Autor dieser in ihrer Machart meist anspruchslosen und unsystematischen Büchlein zumindest den Anspruch ernstzunehmender Sprachkritik (wobei es sich in der Regel eher um sehr subjektive Zeit- und Sachkritik als tatsächlich um Sprachkritik handelt, so z.B. bei Henscheid und Röhl, s.u.) oder den Anspruch der angemessenen Dokumentation eines Ausschnittes aus dem aktuellen deutschen Sprachgeschehen (Horx, Schönfeld, s.u.). Für die neologismenlexikographische Arbeit bieten alle diese Wörterbücher viel interessantes Material. Für die achtziger Jahre läßt sich etwa ein Dutzend solcher Wörterbücher nachweisen. Jüngere Titel der neunziger Jahre sind: Bittermann, Klaus/Henschel, Gerhard (Hg.) (1994): *Das Wörterbuch des Gutmenschen. Zur Kritik der moralisch korrekten Schaumsprache*. Berlin; Bruno, Rainer (1991): *In-Deutsch. Ein Schnellkurs für Hochstapler*. Reinbek; Henscheid, Eckhard (1993): *Dummdeutsch. Ein Wörterbuch*. Unter Mitwirkung von Carl Lierow und Elsemarie Maletzke. Stuttgart; Horx, Matthias (1994): *Trendwörter von Acid bis Zippies. Lexikon*. 2. Aufl. 1995. Düsseldorf u.a.; Lengenfelder, Werner G. (1992): *Das Szene-Schimpfwörterbuch*. Nidderau. (= *Das merkwürdige Wörterbuch Band 2 [sic!]*); Röhl, Klaus Rainer (1995): *Deutsches Phrasenlexikon*.

sehr kurzfristig aufgelegten Taschenwörterbücher kommentieren zwar nicht ausschließlich, aber doch stets in größerem Umfang auch Neologismen und neue Wortgebräuche.

Zum anderen widerspricht Wiegands Vermutung das merkwürdig-auffällige Kokettieren der Werbung der großen Wörterbuchverlage mit den Abertausenden von neuen Wörtern, die angeblich in ihren Neuauflagen oder Neuerscheinungen jeweils enthalten sein sollen. Nur sind sie dort, weil als solche nicht gekennzeichnet, eben kaum auffind- oder nachweisbar. Mit vollem Recht also spricht Sauer von dem „Schnickschnack mit den ‘neuen Wörtern’“, und er „fragt sich verwundert, welche Wörter in welcher Form auf welche Weise ‘dudenfähig’ werden“⁴⁹. Der Werbeträchtigkeit der neuen Wörter (und damit des öffentlichen Interesses an ihnen) ist man sich also sehr wohl bewußt und nutzt diese demgemäß auch kräftig, der Mühe ihrer differenzierten Behandlung will man sich merkwürdigerweise aber nicht unterziehen. Die Logik einer solchen Verfahrensweise ist kaum nachvollziehbar, auch nicht mit dem Hinweis auf die mögliche Ersatzfunktion der im Deutschen (und eben nur im Deutschen) etablierten und beliebten mittelgroßen Fremdwörterbücher (so bei Wiegand 1990, S. 2186).

Fremdwörterbücher und Neologismenwörterbücher haben im Inventar ihrer jeweiligen Einträge nur eine vergleichsweise kleine Schnittmenge, nämlich die Neuentlehnungen. Im mittelgroßen Fremdwörterbuch sind diese nicht als solche markiert und somit auch nicht erkennbar. An dieser Teilmenge der Neologismen läßt sich die unterschiedliche Funktion der beiden Wörterbuchtypen gut veranschaulichen. Während es im Fremdwörterbuch aufgrund der spezifischen Benutzererwartung bei der Neuentlehnung in erster Linie um die Worterklärung, also die Bedeutungsangabe und dazu von Fall zu Fall auch um orthographische Fragen geht, ist der Aufgabenbereich des Neologismenwörterbuchs entschieden weiter gefaßt. Ohne schon hier auf mikrostrukturelle Details der Einträge im Neologismenwörterbuch im einzelnen eingehen zu wollen – dies bleibt einer folgenden Publikation vorbehalten –, so kann doch festgehalten werden, daß im Wortartikel des Neologismenwörterbuchs für den jeweiligen Neologismus und somit auch für die Neuentlehnungen detaillierte Ausführungen zur Herkunft, Angaben zum jeweiligen etymologischen Typ, Fragen des Aufkommens und der Etablierung, Hinweise zum Gebrauch und gute Belegdokumentation verbindlich sein müssen, und damit insgesamt lexikographische Qualitäten, die über den Rahmen der Darstellung im mittelgroßen Fremdwörterbuch deutlich hinausgehen. Das Fremdwörterbuch ist in er-

Lehrbuch der Politischen Korrektheit für Anfänger und Fortgeschrittene. Berlin/Frankfurt a.M.; Schönfeld, Eike (1995): Alles easy. Ein Wörterbuch des Neudeutschen. (= Beck'sche Reihe 1126) München.

⁴⁹ Sauer, Wolfgang Werner (1988): Der „Duden“. Geschichte und Aktualität eines „Volkswörterbuchs“. Stuttgart. S. 174; S. 173.

ster Linie wörterklärendes Hilfsmittel; das Neologismenwörterbuch will vor allem umfassende Dokumentation leisten.

Natürlich wäre eine systematische deutsche Neologismenlexikographie am leichtesten bei den großen Wörterbuchverlagen mit ihren etablierten, gut ausgestatteten und funktionierenden Arbeitsinstrumentarien anzusiedeln. Das legt überzeugend auch das angloamerikanische Vorbild nahe, und hier insbesondere der oben beschriebene Wörterbuchtyp 1 (unselbständige Ergänzungsbände zu großen umfassenden Wörterbuchunternehmen). Da im Rahmen der deutschen Lexikographie derartige Supplement-Bände über kurz oder lang nicht zu erwarten sind, müssen die entsprechenden Bemühungen um die Neologismen anderweitig angesiedelt werden.

Seit seinem Bestehen hat das *Institut für deutsche Sprache* im lexikalischen Bereich immer wieder und zum Teil auch langfristig an Projekten zu ausgewählten Bereichen der aktuellen Wortschatzentwicklung gearbeitet. Zunächst waren das die weithin wortschatzbezogenen Untersuchungen zur Sprachentwicklung im geteilten Deutschland, zuletzt die Arbeiten zur spezifischen Lexik der Wendephase um 1989/1990.⁵⁰ Beide Unternehmungen hatten in Teilbereichen auch enge Berührungspunkte mit Fragen der Neologie im Deutschen. So scheint es sinnvoll, wenn nun im Anschluß an diese Arbeiten in einem neu zu etablierenden Arbeitsbereich „Neologismenlexikographie“ die Erfahrungen aus früheren lexikographischen Arbeiten⁵¹ mit den Bemühungen um eine Dokumentation deutscher Neologismen gekoppelt werden.

Wenn hier vergleichsweise ausführlich die englische Neologismenlexikographie in ihrer Spezifik und in ihren Ergebnissen bzw. Produkten vorgeführt wurde, so nicht nur, um ihre vorbildliche und führende Rolle herauszustreichen, sondern vor allem, um anhand dieser Ergebnisse Erkenntnisse sowie Erfahrungs- und Richtwerte für die eigenen neologismenlexikographischen Bemühungen gewinnen zu können. Es sollen aber auch die beiden mehr oder weniger detaillierten Neologismenwörterbuch-Pläne, die bisher für das Deutsche vorgelegt, aber nicht realisiert wurden (Müller 1987, Heller u. a. 1988), im Auge behalten werden.

Anzuknüpfen ist auch an Wiegands bilanzierende Feststellung, daß die deutschen Lexikographen, „was die Neologismen und Neosemantismen angeht, ihren Dokumentationsaufgaben bisher nicht gerecht geworden“ sind (1990, S. 2185), wobei der Akzent auf Dokumentationsaufgabe zu setzen ist. Es kann in Zukunft nicht mehr damit getan sein, daß sich das „Ergebnis“ dieser Dokumentationsaufgabe lediglich in den vollmundig-verheißungsvollen Tau-

⁵⁰ Vgl. hierzu den Beitrag von Herberg in diesem Band.

⁵¹ Miteinbezogen sind dabei die Erfahrungen am früheren Neologismenprojekt des Zentralinstituts für Sprachwissenschaft bei der Akademie der Wissenschaften der DDR in Berlin.

senderzahlen der bunten Werbebroschüren für neue Wörterbücher erschöpft und daß somit gewissermaßen ein neologismenlexikographisches Spiel mit stets verdeckten Karten und magerem Ausgang gespielt wird. Für das fachliche wie für das öffentliche Interesse müßten von der deutschen Lexikographie zum einen kurzfristig die Fülle der gerade neu aufgekommenen Wörter und Wendungen – durchaus auch inklusive Okkasionalismen – dokumentiert werden sowie zum anderen langfristig, für exakt zu definierende Zeiträume, die sich etablierenden bzw. die nahezu etablierten Neologismen. Dabei können beide Aufgaben durchaus auf verschiedene (institutionelle) Hände verteilt sein.

Im Blick auf die angloamerikanischen Vorbilder sind dabei von den unter 2.1 beschriebenen Typen vor allem der Typ 2 (eigenständige Neologismenwörterbücher auf lexikographisch hohem Niveau) und der Typ 4 (periodisch erscheinende Neologismen-Publikationen auf lexikographisch hohem Niveau) als mustergültig heranzuziehen. Während der Typ des kürzerfristigen Neologismenperiodikums auf eine noch wenig differenzierte Fülle von Neuerungen abhebt und die Ergebnisse neologismenlexikographischer Werkstattarbeit belegen soll, dokumentiert das längerfristig orientierte Neologismenwörterbuch die in lexikographischer Feinarbeit gezielt vorgenommene Auswahl der nahezu etablierten lexikalischen Neuerungen. Daß es sich dabei um den Typ des „evidential new-word dictionary“ handeln muß, versteht sich von selbst.

Beide Typen neologismenlexikographischer Arbeit sind sinnvoll nur im operativen Miteinander realisierbar, wobei die Materialien der kürzerfristigen periodischen Neologismenpräsentation als Vorstufe und Zuarbeit zum längerfristig konzipierten Wörterbuch zu fungieren haben. Mögliche Formen der kurzfristigen Neologismendokumentation können (nach dem Vorbild von *Among the New Words* der Zeitschrift *American Speech*) feste Sparten in Sprachfachzeitschriften sein, gegebenenfalls auch selbständig in kürzeren Abständen (etwa des Ein- oder Zwei-Jahresrhythmus) erscheinende Reihenbroschüren. Wesentliche Aufgabe dieser kurzfristig orientierten neologismenlexikographischen Arbeit ist natürlich die Neologismenrecherche als Grundlage jeglicher Neologismenlexikographie. Neben der neu aufkommenden DV-gestützten Recherche,⁵² deren objektiven Ergebnissen der bisher vorwiegend Zetteldatei-orientiert und somit subjektiv-zufallsunterworfen arbeitende Lexikograph mit ziemlicher Spannung entgegenblicken wird, verbleibt vorerst der individuellen Recherche ein noch immer hoher Stellenwert. Feste neologismenlexikographische Rubriken innerhalb von Sprachzeitschriften, zumal solchen, die neben Fachwissenschaftlern auch fachexterne Sprachinteressenten zu ihren Lesern zählen, bieten zudem die Möglichkeit, die Leser, in mehr oder weniger lockerer Form, in die Rechercharbeit einbinden zu können. Im angloamerikanischen Bereich wurden bei der (allerdings straff strukturierten und organisier-

⁵² Vgl. hierzu die Beiträge von Belica und Teubert in diesem Band

ten) Nutzung solcher Möglichkeiten für die Neologismenrecherche gute Erfahrungen gemacht.

Erfahrungsgemäß wird im Umfeld gezielter kurzfristiger neologismenlexikographischer Arbeit sehr schnell eine Fülle von zunächst diffusem Neologismen- und Okkasionalismenmaterial zusammengetragen. Letztendlich gefragt ist dann die Kompetenz der Lexikographen, und zwar sowohl im Hinblick auf eine überzeugende Auswahl der Stichwortkandidaten für das längerfristige Neologismenunternehmen als auch im Hinblick auf die begründete Entscheidung betreffs der zeitlichen Eingrenzung des Erfassungszeitraums.

Lange Zeit galt bei den neologismenlexikographischen Überlegungen für das Deutsche der historische Einschnitt der Jahre 1945 (Kriegsende) und 1949 (Gründung der beiden Nachkriegsrepubliken) als bevorzugter Ausgangspunkt für die Erfassung der im Wörterbuch zu dokumentierenden Neologismen.⁵³ Jüngere historische Entwicklungen einerseits und der für eine aktuelle Neologismendokumentation mit inzwischen nahezu fünfzig Jahren viel zu lange Zeitraum andererseits lassen jedoch eine Neuorientierung im Hinblick auf mögliche Abgrenzungen von Erfassungszeiträumen sinnvoll und notwendig erscheinen.

So könnte jetzt das einen wichtigen historisch-politischen Einschnitt markierende Datum des 3. Oktober 1990 und somit das Jahr 1990 für neuere Festlegungen von zeitlichen Ausgangspunkten in der deutschen Neologismenlexikographie Priorität besitzen. Für die abgeschlossene Periode der deutschen Zweistaatlichkeit wurde Neologismenlexikographie allenfalls in Ansätzen und in Teilbereichen geleistet. Aus heutiger Sicht ist diese Periode lexikalisch nicht mehr von aktuellem, sondern nur noch von sprachhistorischem Interesse. Für die Neologismenlexikographie stellt sie einen inzwischen viel zu weit zurückliegenden Zeitraum dar, um noch das leisten zu können, was diese ja vor allem auszeichnen sollte, nämlich aktuell zu dokumentieren und zu informieren.

Für die Nützlichkeit ihrer Dokumentationsarbeit kann jede, und so auch die deutsche Neologismenlexikographie letztendlich viele gute Gründe vorweisen. Von ihrem Nutzen für die Sprachgeschichtsschreibung im allgemeinen wie für die Wortforschung im besonderen wurde bereits gesprochen, auf den Nutzen vor allem für ausländische Sprachlerner wie für die zweisprachige Lexikographie wurde auch in der Literatur immer wieder hingewiesen.⁵⁴ Jedoch nicht nur für den fachinternen Bereich, auch im weiteren zeitgeschichtlichen Kon-

⁵³ So noch bei Müller (1987): „Als ungefähre Neologismugrenze bietet sich das Jahr 1945 an, denn bei dieser zeitlichen Festlegung können auch die oft revolutionären Innovationen der Nachkriegszeit auf technischem, wirtschaftlichem, politischem, kulturellem und wissenschaftlichem Gebiet lexikographisch dokumentiert werden“ (S. 89).

⁵⁴ Vgl. dazu unter anderem bei R. u. C. Barnhart (1990), Dou Xuefu (1989) und Wiegand (1990).

text und für den an ihm interessierten Zeitgenossen hat die Neologismenlexikographie ihren Sinn, Wert und Zweck. In resümierenden Passagen über die Aussagekraft englischer Neologismenwörterbücher kommt das, wie ich meine, sehr schön und überzeugend zum Ausdruck:

The introduction of new words – and of new meanings for old ones – reflects developments and innovations in the world at large and in society. So the vocabulary items in this collection present in microcosm the concerns that have impinged on speakers of English worldwide in the late 1980s. (Ayto 1989, Introduction, o.S.)

The new words – recorded in this volume – tell us something about the world and our reactions to it during the ... period during which they were gathered. They reflect changes in material and intellectual culture. And they show us something of the way human beings cope with problems and laugh at absurdities of life. (Algeo 1991, S. 15)

3. Literatur

Einen bibliographischen Gesamtüberblick zum Thema „Neologismus und Neologismenlexikographie im Deutschen“ wird ein Band der IDS-Reihe Studienbibliographien – Sprachwissenschaft geben, der gegenwärtig vorbereitet wird. (Herausgeber: D. Herberg und M. Kinne; ersch. 1998).

Ahrends, Martin (1986/1989): Trabbi, Telespargel und Tränenpavillon. Das Wörterbuch der DDR-Sprache. München 1986. Überarb. Neuauf.: Allseitig gefestigt. Stichwörter zum Sprachgebrauch der DDR. München 1989.

Algeo, John (1990): [Rezension] John Ayto: *The Longman Register of New Words*. Harlow/Essex 1989. In: *International Journal of Lexicography*, 3, S. 211-218.

Algeo, John (Hg.) (1991): *Fifty Years Among the New Words*. A Dictionary of Neologisms, 1941-1991. New York u.a. Introduction S. 1-16.

Algeo, John (1993): Desuetude among New English Words. In: *International Journal of Lexicography*, 6-4, S. 281-293. [Mit guter Bibliographie zur englischsprachigen Neologismenlexikographie.]

Algeo, John (1995): A Short History of New-Word Study. In: *Dictionaries. Journal of The Dictionary Society of North America*, 16, S. 3-15. (Vgl. Neology Forum).

Ayto, John (1989): *The Longman Register of New Words*. 1. Harlow/Essex.

Ayto, John (1990): *The Longman Register of New Words*. 2. Harlow/Essex.

Ayto, John (1992): [Rezension] *The Macquarie Dictionary of New Words*. Hrsg. v. Susan Butler. In: *International Journal of Lexicography*, 1, S. 238-241.

Barnhart, Clarence L./Barnhart, David K. (Hg.) (1982ff.): *The Barnhart Dictionary Companion: A Quarterly to Update General Dictionaries*. 1ff. Cold Spring/New York.

Barnhart, Clarence L./Steinmetz, Sol/Barnhart, Robert K. (1973): *The Barnhart Dictionary of New English since 1963*. Bronxville/New York.

- Barnhart, Clarence L./Steinmetz, Sol/Barnhart, Robert K. (1980): The *Second* Barnhart Dictionary of New English. Bronxville/New York.
- Barnhart, David K. (1995): Clarence L. Barnhart and Quotations for New Words in American Dictionaries. In: Dictionaries. Journal of The Dictionary Society of North America, 16, S. 65-78. (Vgl. Neology Forum).
- Barnhart, Robert K./Steinmetz, Sol/Barnhart, Clarence L. (1990): *Third* Barnhart Dictionary of New English. New York.
- Barnhart, Robert/Barnhart, Clarence (1990): The Dictionary of Neologisms. In: Hausmann, Franz Josef u.a. (Hg.): Wörterbücher – Dictionaries – Dictionnaires. Ein internationales Handbuch zur Lexikographie. (= Handbücher zur Sprach- und Kommunikationswissenschaft. 5.2, 2. Teilbd.). Berlin/New York. S. 1159-1166.
- Berg, Paul Charles (1953): A Dictionary of New Words in English. New York.
- Berning, Cornelia (1964): Vom *Abstammungsnachweis* zum *Zuchtwart*. Vokabular des Nationalsozialismus (= Die kleinen de Gruyter Bände 6). Berlin. (Urspr. Diss. Bonn 1958. Auch in Forts. veröffentl. in: Zeitschrift für deutsche Wortforschung von 1960. 1/2 bis 1963, 1/2.)
- Brackmann, Karl-Heinz/Birkenhauer, Renate (1988): NS-Deutsch. „Selbstverständliche“ Begriffe und Schlagwörter aus der Zeit des Nationalsozialismus (= Europäisches Übersetzer-Kollegium Straelen. Glossar Nr. 4). Straelen.
- Braun, Peter (1987): Tendenzen in der deutschen Gegenwartssprache. 2. veränd. u. erw. Aufl. (Erstauf. 1979). Stuttgart u.a.
- Burchfield, Robert (Hg.) (1972-1986): A Supplement to the Oxford English Dictionary. 4 Bde. Oxford.
- Busse, Ulrich (1996): Neologismen. Der Versuch einer Begriffsbestimmung. In: Gellerstam, Martin u.a. (Hg.): EURALEX '96. Proceedings I-II. Part II. Göteborg. S. 645-658. [Nach Abschluß meines Manuskripts erschienen.]
- Butler, Susan (Hg.) (1990): The Macquarie Dictionary of New Words. Dee Why/Sydney.
- Carstensen, Broder (1986): Beim Wort genommen. Bemerkenswertes in der deutschen Gegenwartssprache. Tübingen.
- Cherubim, Dieter (Hg.) (1975): Sprachwandel. Reader zur diachronischen Sprachwissenschaft. Berlin/New York.
- Cherubim, Dieter (1984): Sprachgeschichte im Zeichen der linguistischen Pragmatik. In: Besch, Werner u.a. (Hg.): Sprachgeschichte. Ein Handbuch zur Geschichte der deutschen Sprache und ihrer Erforschung (= Handbücher zur Sprach- und Kommunikationswissenschaft 2.1.). Berlin/New York. S. 802-815.
- Constantin, Theodor (1982): Plaste und Elaste. Ein deutsch-deutsches Wörterbuch. Berlin.
- Coseriu, Eugenio (1958/1974): Synchronie, Diachronie und Geschichte. Das Problem des Sprachwandels. (=Internationale Bibliothek für Allgemeine Linguistik 3). München. (Originalausg.: Montevideo 1958).
- Dou Xuefu (1988): Das kleine deutsch-chinesische Neologismenwörterbuch. Peking.
- Dou Xuefu (1989): Neologismus und Neologismenwörterbuch. In: Wiegand, Herbert Ernst (Hg.): Wörterbücher in der Diskussion. Vorträge aus dem Heidelberger Lexikographischen Kolloquium. (= Lexicographica. Series Maior 27). Tübingen. S. 39-75.

- Dou Xuefu (1994): Neologismus als Merkmal der deutschen Gegenwartssprache. In: Dokumentation der Tagungsbeiträge des chinesisch-japanischen Germanistentreffens Beijing 1990. Beijing. S. 241-252.
- Fleischer, Wolfgang (u. Autorenkollektiv) (1987): Wortschatz der deutschen Sprache in der DDR. Fragen seines Aufbaus und seiner Verwendungsweise. Leipzig.
- Glück, Helmut/Sauer, Wolfgang Werner (1990): *Gegenwartsdeutsch*. (= Sammlung Metzler 252). Stuttgart.
- Hausmann, Franz Josef (1985): Neologismenwörterbücher. In: Schwarze, Christoph/Wunderlich, Dieter (Hg.): *Handbuch der Lexikologie*. Königstein/Ts. S. 390.
- Heberth, Alfred (1977): *Neue Wörter. Neologismen in der deutschen Sprache seit 1945*. Wien.
- Heberth, Alfred (1982): *Neue Wörter 2. Neologismen in der deutschen Sprache seit 1945*. Wien.
- Heller, Klaus u.a. (1988): *Theoretische und praktische Probleme der Neologismenlexikographie. Überlegungen und Materialien zu einem Wörterbuch der in der Alltagsprache der DDR gebräuchlichen Neologismen*. (= Linguistische Studien, Reihe A, Arbeitsberichte 184). Berlin.
- Hellwig, Gerhard (1972/1983): *Kennen Sie die neuesten Wörter?* (Erw. TB-Ausg. 1983). München.
- Herberg, Dieter (1988a): *Stand und Aufgaben der Neologismenlexikographie des Deutschen*. In: Harras, Gisela (Hg.): *Das Wörterbuch – Artikel und Verweisstrukturen*. IdS-Jahrbuch 1987. (= Sprache der Gegenwart 74). Düsseldorf. S. 265-283.
- Herberg, Dieter (1988b): *Ein Wörterbuch der DDR-Neologismen: Prinzipien seiner inhaltlichen und formalen Gestaltung*. In: Hyldgaard-Jensen, Karl/Zettersten, Arne (Hg.): *Symposium on Lexicography IV. Proceedings of the Fourth International Symposium on Lexicography April 20-22, 1988 at the University of Copenhagen*. (= Lexicographica. Series Major 26). Tübingen. S. 143-162.
- Herberg, Dieter (1991): *Neologismen als Forschungsgegenstand. Aktuelle Aufgaben und Ziele der Neologismenlexikographie*. In: Sommerfeldt, Karl-Ernst (Hg.): *Sprachwissenschaft und Sprachkultur. Tagungsband der Konferenz in Neubrandenburg am 10. u. 11. Mai 1990*. (= Sprache – System und Tätigkeit 1). Frankfurt a.M. u.a. S. 111-119.
- Itkonen, Esa (1984): *On the 'Rationalist' Conception of Linguistic Change*. In: *Diachronica. International Journal for Historical Linguistics*, 2, S. 203-216.
- Jordanova, Ljubima (1994a): *Historischer und kultureller Hintergrund der Lexik in der neologischen Lexikographie*. In: Hyldgaard-Jensen, Karl/Pedersen, Viggo Hjørnager (Hg.): *Symposium on Lexicography VI. Proceedings of the Sixth International Symposium on Lexicography May 7-9, 1992 at the University of Copenhagen*. (= Lexicographica. Series Major 57). Tübingen. S. 175-181.
- Jordanova, Ljubima (1994b): *Bulgarisches Neologismenwörterbuch*. In: Hyldgaard-Jensen, Karl/Pedersen, Viggo Hjørnager (Hg.): *Symposium on Lexicography VI. Proceedings of the Sixth International Symposium on Lexicography May 7-9, 1992 at the University of Copenhagen*. (= Lexicographica. Series Major 57). Tübingen. S. 165-174.
- Kammer, Hilde/Bartsch, Elisabet (1982/1992): *Jugendlexikon Nationalsozialismus. Begriffe aus der Zeit der Gewaltherrschaft 1933-1945*. Hamburg 1982. (Neuauf. unter dem Titel: *Nationalsozialismus. Begriffe aus der Zeit der Gewaltherrschaft*. Hamburg 1992.)

- Kinne, Michael (1989): Endlich: Ein deutsches Neologismenwörterbuch. In: Der Sprachdienst, 4, S. 115-117.
- Kinne, Michael/Strube-Edelmann, Birgit (1980): Kleines Wörterbuch des DDR-Wortschatzes. 2. Aufl. 1981. Düsseldorf.
- Klappenbach, Ruth/Malige-Klappenbach, Helene (1978): Das Wörterbuch der deutschen Gegenwartssprache. Entstehung, Werdegang, Vollendung. In: Kopenhagener Beiträge zur Germanistischen Linguistik, 14, S. 5-47.
- Klappenbach, Ruth/Steinitz, Wolfgang (Hg.) (1961-1977): Wörterbuch der deutschen Gegenwartssprache. 6 Bde. Berlin 1964-1977 (1. Lieferung 1961).
- Klare, Johannes (1977): Neologismus und Neosemantismus als lexikographisches Problem unter besonderer Berücksichtigung des französischen gesellschaftspolitischen Wortschatzes. In: Linguistische Arbeitsberichte, 17, S. 41-50.
- LeMay, Harold/Lerner, Sid/Taylor, Marian (1985): New Words Dictionary. New York.
- LeMay, Harold/Lerner, Sid/Taylor, Marian (1988): The New New Words Dictionary. New York.
- Mager, N. H./Mager, Sylvia K. (1982): The Morrow Book of New Words. New York.
- Malige-Klappenbach, Helene (1986): Das „Wörterbuch der deutschen Gegenwartssprache“ Bericht, Dokumentation und Diskussion. Tübingen.
- Mattheier, Klaus J. (1984): Sprachwandel und Sprachvariation. In: Besch, Werner u.a. (Hg.): Sprachgeschichte. Ein Handbuch zur Geschichte der deutschen Sprache und ihrer Erforschung. (= Handbücher zur Sprach- und Kommunikationswissenschaft 2.1.). Berlin/New York. S. 768-779.
- Matussek, Magdalena (1994): Wortneubildung im Text. (= Beiträge zur germanistischen Sprachwissenschaft 7). Hamburg. [Nach Abschluß meines Manuskriptes erschienen.]
- Mish, Frederick C. u.a. (Hg.) (1976): 6 000 Words: A Supplement to Webster's Third New International Dictionary. Springfield/Mass.
- Mish, Frederick C. u.a. (Hg.) (1983): 9 000 Words: A Supplement to Webster's Third New International Dictionary. Springfield/Mass.
- Mish, Frederick C. u.a. (Hg.) (1986): 12 000 Words: A Supplement to Webster's Third New International Dictionary. Springfield/Mass.
- Mort, Simon (Hg.) (1986): Longman Guardian New Words. Harlow/Essex.
- Müller, Wolfgang (1987): „Schlammschlacht“. Schon gehört? Ein Desiderat: Das deutsche Neologismenwörterbuch. In: Sprache und Literatur, S. 82-90.
- Munske, Horst Haider (1985): Lexikologie und Wortgeschichte. In: Stötzel, Georg (Hg.): Germanistik – Forschungsstand und Perspektiven. Vorträge des Deutschen Germanistentages 1984. Teil I. Berlin/New York. S. 27-43.
- Munske, Horst Haider (1990): Über den Wandel des deutschen Wortschatzes. In: Besch, Werner (Hg.): Deutsche Sprachgeschichte: Grundlagen, Methoden, Perspektiven. Festschrift für Johannes Erben zum 65. Geburtstag. Frankfurt a.M. u.a. S. 387-401.
- Neology Forum (1995): Themenheft mit Beiträgen von Agnes, Michael, vgl. Algeo, John, vgl. Barnhart, David K.; Barnhart, Robert K.; Cannon, Garland, Creamer, Thomas B.I.; Jost, David/Rowan, Beth/Schwartz, Susan; Neufeldt, Victoria; vgl. Paton, Berna-

- dette; Sheidlower, Jesse T. (= Dictionaries. Journal of The Dictionary Society of North America 16). [Nach Abschluß meines Manuskriptes erschienen.]
- Paaske, Gunhild (1981): Neologismen i moderne journalistisk tysk. (2. Aufl. 1983). (Det erhvervsproglige Fakultet Handelshøjskolen). Kopenhagen.
- Paton, Bernadette (1995): New-Word Lexicography and the OED. In: Dictionaries. Journal of The Dictionary Society of North America, 16, S. 79-89. (Vgl. Neology Forum)
- Polenz, Peter v. (1991): Deutsche Sprachgeschichte vom Spätmittelalter bis zur Gegenwart I: Einführung. Grundbegriffe. Deutsch in der frühbürgerlichen Zeit. Berlin/New York. S. 37-58.
- Reifer, Mary (1955): Dictionary of New Words. With an Introduction by E. Patridge. New York.
- Ricken, Ulrich (1977): Zur Neologie-Diskussion des 18. Jahrhunderts und ihre Fortsetzung nach der Revolution. In: Linguistische Arbeitsberichte, 18, S. 81-95. (Auch in: Wissenschaftliche Zeitschrift der Universität Halle, S. 109-118.)
- Schank, Gerd (1984): Ansätze zu einer Theorie des Sprachwandels auf der Grundlage von Textsorten. In: Besch, Werner u.a. (Hg.): Sprachgeschichte. Ein Handbuch zur Geschichte der deutschen Sprache und ihrer Erforschung. (= Handbücher zur Sprach- und Kommunikationswissenschaft 2.1). Berlin/New York. S. 761-768.
- Schippa, Thea (1992): Lexikologie der deutschen Gegenwartssprache. Tübingen. (Zuerst: Leipzig 1984). S. 240-247.
- Schönaich, Christoph Otto Freiherr von (1754/1900): Die ganze Aesthetik in einer Nuss oder Neologisches Wörterbuch. Neudruck Berlin 1900 (mit Einleitung u. Anm. hrsg. v. Albert Köster).
- Simpson, John/Weiner, Edmund (Hg.) (1993/1994): Oxford English Dictionary Additions Series 1 (1993) u. 2 (1994). Oxford.
- Steger, Hugo (1989): Sprache im Wandel. In: Sprache und Literatur in Wissenschaft und Unterricht, 63, S. 3-31. (Auch in: Benz, Wolfgang (Hg.): Die Geschichte der Bundesrepublik Deutschland 3. Frankfurt a.M. S. 15-46.)
- Tulloch, Sara (1991): The Oxford Dictionary of New Words. A Popular Guide to Words in the News. Oxford/New York.
- Vachek, Josef (1962/1975): Zum Zusammenspiel von internen und externen Faktoren bei der Sprachentwicklung. In: Cherubim (Hg.) (1975), S. 190-207. Zuerst veröffentl. 1962.
- Wiegand, Herbert Ernst (Hg.) (1989): Wörterbücher in der Diskussion. Vorträge aus dem Heidelberger Lexikographischen Kolloquium. (= Lexicographica. Series Maior 27). Tübingen.
- Wiegand, Herbert Ernst (1990): Neologismenwörterbücher. [= Kapitel 2.4.1 aus: Die deutsche Lexikographie der Gegenwart]. In: Hausmann, Franz Josef u.a. (Hg.): Wörterbücher – Dictionaries – Dictionnaires. Ein internationales Handbuch zur Lexikographie. (= Handbücher zur Sprach- und Kommunikationswissenschaft 5.2., 2. Teilbd.). Berlin/New York. S. 2185-2187.
- Wurzel, Wolfgang Ullrich (1991): Faktoren des Sprachwandels. In: Papiere zur Linguistik, 1/2, S. 159-173.

John Sinclair

Korpustypologie. Ein Klassifikationsrahmen*

1. Einführung

Elektronische Korpora sind eine Neuheit und bisher gibt es wenig Konsens darüber, was als Korpus zählt und wie Korpora klassifiziert werden sollen. Die wichtigsten Bereiche sind:

- a) die Mindestanforderungen, die eine Sammlung von Sprachdaten erfüllen soll, um als Korpus eingestuft zu werden;
- b) die Unterscheidung von Korpora, die die Allgemeinsprache abbilden, von solchen, die spezialisiertes Sprachverhalten dokumentieren.

Beides sind umstrittene Bereiche. Wenn die Fachwelt die Kriterien, die hier vorgeschlagen werden, als angemessen für den derzeitigen Bedarf erachtet, wird das ein beachtlicher Fortschritt sein, denn gegenwärtig gibt es noch viele Sprachdaten-Sammlungen, die als Korpora bezeichnet werden, obwohl sie diese Voraussetzungen nicht erfüllen; andererseits sind einige Korpora erhältlich, die spezielles und künstliches Sprachverhalten erfassen, ohne dies für den Laien ausdrücklich zu kennzeichnen.

Darüber hinaus entwickelt sich die Disziplin Korpuslinguistik sehr schnell, und entsprechend werden Normen und Vorgaben ständig überarbeitet. Kategorien müssen also besonders flexibel sein, um solchen instabilen Bedingungen Rechnung zu tragen.

Daher sind die Klassifikationen in diesem Beitrag nur insoweit vorgenommen worden, wie es derzeit als klug erachtet werden kann. Sie bieten eine solide und relativ gut nachvollziehbare Art, Korpora zu klassifizieren, mit deutlich abgegrenzten Kategorien – wo immer möglich – und fundierten Vorschlägen in Zweifelsfällen. Dieser Beitrag wurde von vielen Experten auf dem Gebiet kritisch durchgesehen, die weitgehend darin übereinstimmen, daß eine engere Klassifikation intellektuell nicht gerechtfertigt ist und bei den meisten Wissenschaftlern keine Beachtung finden würde. Der gegenwärtige Beitrag kann auf Akzeptanz hoffen, da er relevante Themen anspricht und verwendbare Klassifikationen anbietet.

* Der vorliegende Beitrag ist eine Übersetzung des internen EAGLES-Berichts "Corpus Typology. Preliminary Recommendations", EAG-CSG/IR-T1.1, October 1994 (Work in Progress). Die Übersetzung besorgten Ann Lawson, Susanne Lenz und Norbert Volz (alle IDS Mannheim).

Für die Dauer von fast zwanzig Jahren wurden die ursprünglichen Ziele des *Brown Corpus* (Kucera und Francis 1967) als Standard erachtet:

- a) 1 Million Wörter
- b) ungefähr gleiche Aufteilung in Genres
- c) 500 Texte
- d) 2000 Wörter pro Text (*sample*)
- e) geschriebene, veröffentlichte Quellen

Das *Brown Corpus* gilt häufig immer noch als Bezugspunkt, obwohl die Umstände, die seine Urheber zur Aufstellung dieser Parameter veranlaßt haben, sich wesentlich von den gegenwärtigen unterscheiden.

Es erscheint hilfreicher, die Prinzipien des ursprünglichen Entwurfs, die den jeweiligen Entscheidungen zugrundelagen, herauszufiltern:

- a) Das Korpus sollte so groß wie nur möglich sein angesichts der Technologie jener Zeit. Das traf auf das Brownsche Korpus mit einer Million Wörtern zu, und sein Erscheinen glich einem Wunder: zum erstenmal hatte man so viele Wörter zur Verfügung. Aber bis Mitte der siebziger Jahre stieg die Größenordnung des angestrebten Ziels an, und der Bestand von *The Birmingham Collection of English Text* erreichte 1985 schließlich an die 20 Millionen Wörter. Jetzt, Mitte der 90er Jahre, steuert er wiederum auf eine andere Größenordnung zu: *The Bank of English* ist mittlerweile bei 320 Millionen Wörtern Textlänge angelangt.
- b) Es sollte 'Textproben' (*samples*) von großer Vielfalt beinhalten, um in irgendeiner Form Repräsentativität zu gewährleisten.
- c) Es sollte eine dazwischenliegende Klassifikation in Genres zwischen dem Korpus als Ganzem und den einzelnen 'Textproben' (*samples*) bereitstellen.
- d) Die 'Textproben' (*samples*) sollten von gleicher Größe sein.
- e) Das Korpus als Ganzes sollte nach Quellen dokumentiert sein.

Der obengenannte Punkt d), nämlich daß 'Textproben' (*samples*) von gleicher Größe sein sollten, ist kontrovers und wird in diesen Vorschlägen nicht übernommen (weitere Ausführungen dazu siehe unten). Die Begrenzung des *Brown Corpus* auf schriftliches Datenmaterial (immer noch regelmäßig übernommen in späteren Arbeiten) wird für ein Modell, obgleich verständlich in seinem historischen Kontext, als irreführend angesehen. Tatsächlich waren die

ersten alternativen Modelle zum *Brown Corpus* europäische Sammlungen transkribierter Sprache, so z.B. das *Edinburgh-Birmingham Corpus* in den frühen 60er Jahren (Jones & Sinclair 1974) (zur Vordringlichkeit gesprochener Korpora und ihrem speziellen Beitrag zur Korpusarbeit siehe unten). Es ist zu bemerken, daß immer noch eine erhebliche Zurückhaltung unter den Korpusdesignern besteht, gesprochenes Datenmaterial einzubeziehen; im Planungsstadium des europäischen Infrastrukturprojekts *Network of European Textual Reference Corpora* (NERC) 1990 wurde es fast ausgespart, und es gibt in der Europäischen Union Anzeichen dafür, daß gesprochene und geschriebene Korpora getrennt entwickelt werden sollen und daß die Konfusion um 'gesprochene Sprache' (*spoken language*) und 'Korpora zu Sprachprozessen' (*speech*) wieder aufgetreten ist.

Die frühen Korpusdesigner arbeiteten mit leistungsschwachen Computern, die hauptsächlich für die Arbeit mit Zahlen ausgerichtet waren und deren Software sich für die Verarbeitung von alphanumerischen Zeichen als problematisch erwies. Texte bestanden aus großen Lochkartenstapeln, und Suchprogramme wurden nachts im Batchbetrieb ausgeführt. Das gesamte Datenmaterial mußte Stück für Stück per Tastatur eingegeben werden.

Im letzten Jahrzehnt gab es eine wahre Revolution, was die Verfügbarkeit von maschinenlesbaren Texten angeht, nämlich das Aufkommen einer neuen schriftlichen Form, der E-mail, die nur in dieser Form existiert, sowie die Erfindung des Scanners als Hilfsinstrument bei der Texterfassung. Die Arbeitsgeschwindigkeit von Maschinen hat zugenommen, während die Kosten für Speicherkapazität gleichzeitig stark zurückgingen, so daß auch für die Benutzer handelsüblicher PCs der Zugriff auf große Korpora möglich ist, während man mit größeren Rechnern Millionen von Wörtern auf einmal bearbeiten kann. So hat sich mittlerweile das Problem von Engpässen beim Zusammentragen und Speichern von Korpusmaterial hin zu einem der Bewältigung ganzer Daten 'fluten' aus einer Vielfalt unkoordinierter Quellen verlagert. In Anbetracht dieser Veränderung werden Monitorkorpora entwickelt, um ein Korpus eher in Form eines Datenflusses, denn als unveränderliches Archiv neu zu konzeptualisieren (siehe unten 6.).

2. Definitionen

2.1 Korpus und Computerkorpus

Ein *Korpus* ist eine Sammlung von Sprachstücken, die nach expliziten linguistischen Kriterien ausgewählt und geordnet sind, um als 'Probe' (*sample*) der Sprache verwendet zu werden.

Es ist jedoch anzumerken, daß hier der unverfängliche Begriff 'Stücke' (*pieces*) anstelle von 'Texte' verwendet wurde. Das hängt mit den verwendeten Sample-Techniken zusammen. Wenn die 'Proben' (*samples*) alle von gleicher Größe sein sollen, können sie nicht alle Texte darstellen. Die meisten werden Textfragmente sein, die willkürlich aus ihrem Kontext genommen wurden.

Ein *Computerkorpus* ist ein Korpus, das in einer standardisierten und homogenen Weise für eine offene Zahl von Abfrage-Aufgaben kodiert ist. Die 'Sprachstücke' (*language pieces*), aus denen es besteht, sind nach ihren Ursprüngen und nach ihrer Herkunft dokumentiert.

2.1.1 Unangemessene Begriffe

Wörter wie *Sammlung* oder *Archiv* beziehen sich auf Mengen von Texten, die weder ausgewählt noch geordnet zu sein brauchen oder bei denen Auswahl bzw. Ordnung nicht nach linguistischen Kriterien erfolgt sein muß. Sie sind deshalb Korpora ganz unähnlich.

'Belege' (*citations*) sind individuelle Vorkommen von Wörtern in konkreter Verwendung; Belegsammlungen haben auch keinen Anspruch, Korpora zu sein. Die genauen Kriterien für eine geeignete Probengröße eines Korpus stellen noch einen Diskussionspunkt dar (siehe unten), aber niemand, der etwas von Korpora versteht, würde erwägen, eine Belegsammlung anzulegen und sie als Korpus zu bezeichnen. Besitzer von früher angelegten Belegsammlungen haben jedoch versucht, sie als Brücke zwischen traditioneller Praxis (insbesondere in der Lexikographie) und korpusbasierter Arbeit zu verwenden.

Es hilft nicht, Kategorien in dieser Weise durcheinanderzuwerfen, und es ist wichtig, die minimalen Kriterien für den Gebrauch des Begriffs *Korpus* klarzustellen.

2.2 *Text*

Die Definition von *Text* wird das Thema zukünftiger Arbeiten bilden. Hier soll deshalb lediglich darauf verwiesen werden, daß der Begriff sowohl für gesprochene als auch für geschriebene Kommunikation verwendet wird.

2.3 Subkorpus, Komponente und Sondersprache

Korpusaufteilungen: ein Korpus kann in *Subkorpora* aufgeteilt werden. Ein Subkorpus hat alle Eigenschaften eines Korpus, ist aber daneben Teil eines größeren Korpus. Korpora und Subkorpora können in *Komponenten* aufgeteilt werden. Eine Komponente muß nicht unbedingt eine angemessene 'Probe' (*sample*) einer Sprache sein, und darin unterscheidet sie sich von einem Korpus und einem Subkorpus. Es ist eine Sammlung von Sprachstücken, die nach genau umrissenen linguistischen Kriterien ausgewählt und geordnet sind, die die sprachliche Homogenität dieser Stücke charakterisieren helfen. Während ein Korpus (und bis zu einem gewissen Grad auch ein Subkorpus) auch Heterogenität illustrieren kann, illustriert die Komponente einen besonderen Typ von Sprache. Was man 'Sondersprachen' (*sublanguages*) nennt, sind nach dieser Definition Komponenten, aber es gibt für sie noch andere Einschränkungen, die später behandelt werden.

2.4 Linguistische Klassifikationskriterien

Klassifikationskriterien können auf Korpora, Subkorpora und Komponenten angewendet werden. Linguistische Kriterien können:

- *extern* sein, wenn sie Teilnehmer, Gelegenheit, soziale Situation, kommunikative Funktion der Sprachstücke betreffen. Dies sind die bekannten Kriterien von Francis und Kucera, Hofland und Johansson, Sinclair, Atkins et al. und des NERC Reports (1994);
- *intern* sein, wenn sie Okkurrenz und Rekurrenz von Sprachmustern innerhalb von Sprachstücken betreffen. Diese Kriterien sind neuer und beanspruchen wachsendes Interesse – s. Biber (1988) und Nakamura (1993).

Auf Korpora werden zumeist die externen Kriterien der Texttypologie abgebildet, worauf hier aber nicht genauer eingegangen wird, da dies – ebenso wie interne Kriterien Gegenstand einer zukünftigen Arbeit sein wird.

2.5 Charakteristika

Ein Korpus sollte bestimmte Charakteristika als 'Standardbedingungen' (*defaults*) erfüllen. Sie müssen gelten, wenn sie nicht für ein Korpus explizit ausgeschlossen werden. Ein Korpus, das ein oder mehrere Werte enthält, die nicht den Standardbedingungen dieser Charakteristika entsprechen (*non-default values*), ist ein *Spezialkorpus*: sein Name sollte die Abweichungen von

den Vorgaben – die auf diesen Standardbedingungen gründen – deutlich machen (siehe auch 2.5.2 für weitere Informationen zu Spezialkorpora).

2.5.1 Quantität

Der Wert für den Parameter *Quantität* ist 'groß'. Ein Korpus muß eine große Menge Wörter enthalten. Der Sinn bei der Zusammenstellung eines Korpus besteht darin, quantitativ große Datenmengen zu sammeln. Die Größe von Korpora wächst weiter rapide, und es wäre unsinnig, bestimmte Zahlen anzusetzen. Mit dem Aufkommen von *Monitorkorpora* (s.u.) ändert sich auch die Berechnungsgrundlage von Gesamtsumme zu Durchflußrate. Die Größe eines Korpus stellt lediglich die Summe der Größen seiner Komponenten dar. Fragen der Größenordnung werden am besten in bezug auf eine Komponente behandelt.

In der Praxis gibt die Größe einer Komponente dazu Auskunft, wie leicht oder schwer es ist, Datenmaterial zu bekommen. Umgekehrt kann dieser Faktor in lockerem Zusammenhang mit der öffentlichen Zugänglichkeit des Datenmaterials gebracht werden; dies impliziert einen relativ großen Einfluß dieses Sprachtyps gegenüber solchem Datenmaterial, das schwieriger zugänglich ist, z.B. wegen seiner vergleichsweise geringen Verbreitung. Für gesprochene Sprache ist das Verhältnis umgekehrt; hier wird das einflußreichste und überzeugendste Datenmaterial durch informelle, spontane Konversation gestellt, die normalerweise nicht aufgezeichnet wird.

Eine sinnvollere Beziehung läßt sich zwischen der Größe einer Komponente und der Anzahl von Personen, die sie betrifft, aufstellen. Da Millionen von Menschen Zeitungen lesen und Radio hören, sollte es verhältnismäßig leicht fallen, eine große Menge an Daten zu diesem Typ zu sammeln, wovon – hier – auch ausgegangen wird. Das Lokalradio erreicht schon weniger Menschen, aber immer noch Hunderttausende, wie auch Illustrierte und Bestseller in Taschenbuchform. Ansprachen bei großen Kundgebungen, sowie Flugblätter und Handzettel erreichen Tausende von Menschen. Dabei handelt es sich jedoch nur um Richtwerte, denn natürlich gibt es auch einige Flugblätter, die von Regierungen und Inserenten in Millionenaufgabe gedruckt werden, und auf der anderen Seite einige Illustrierte, die nur in sehr kleiner Auflage erscheinen und einige Radioprogramme, die nur einen sehr kleinen Zuhörerkreis haben. Aber es gilt zu bemerken, daß, wenn eine Ansprache bei einer Kundgebung im Fernsehen wiederholt wird, oder ein Artikel aus einer Illustrierten nochmals in den Tageszeitungen abgedruckt wird, der angesprochene Adressatenkreis immer noch der ursprünglich intendierte ist, was danach mit einem Text geschieht, beeinflußt die ursprüngliche linguistische Klassifikation nicht.

Man kann sicherlich argumentieren, daß diese linguistische Klassifikation ausschlaggebend für die Übernahme eines bestimmten Textes in ein anderes Medium ist, oder auch, daß dem/der Sprecher(-in)/Schreiber(-in) daran lag, einen weiteren Adressatenkreis zu erreichen, weshalb er/sie seine/ihre Sprache entsprechend angelegt hat und der Text folglich besser von seiner Bestimmung her klassifiziert werden sollte. In der Praxis kommt es ständig vor, daß ein Text von einem Medium ins andere übertragen wird, und ein derartiger Text kann auch zweimal in einem Korpus auftauchen.

Datenmaterial mit geringerer Verbreitung, das lediglich einige hundert Menschen betrifft, läßt sich z.B. am Arbeitsplatz und in institutionellen Dokumentationen finden und im gesprochenen Medium z.B. bei Vorlesungen, Diskussionsrunden und manchen Predigten. Wenn es sich bei dem Publikum nur noch um einige Dutzend handelt, ist schließlich die Ebene von Rundschreiben mit Laufzetteln und von Seminaren erreicht. Die einzelnen Rezipienten können identifiziert werden. Auf diesen Ebenen ist das Datenmaterial nicht einfach zu klassifizieren, aber natürlich gibt es daneben auch private Korrespondenz mit einer Leserschaft von lediglich einem oder zwei Lesern. Hier gibt es eine große Vielfalt, was das gesprochene Medium betrifft, mit allen möglichen Arten von Diskussionen, Interviews, Treffen und Konversationen mit einer relativ geringen Anzahl von Teilnehmern. Korrespondenz mittels E-mail ist oft in eher kleinen Gruppen anzutreffen, wobei auch das hier zirkulierende Datenmaterial zuzunehmen scheint.

Zwei Aspekte kristallisieren sich aus dieser Darstellung heraus. Der eine besagt, daß gesprochene Sprache aufgezeichnet, an ein Massenpublikum weitergegeben und in elektronischer Form zugänglich gemacht werden kann. Das widerspricht der in der Korpusforschung oft vertretenen Ansicht, daß gesprochene Sprache schlecht nutzbar und teuer in der Umsetzung sei. Dieser Punkt wird später eingehender behandelt.

Der andere besagt, daß die Klassifikation von Texten anhand der ungefähren Publikumsgröße allein nicht ausreicht, um die Standardbedingung 'Größe' zu spezifizieren. Insbesondere wirft man diesem Aspekt oft vor, daß er aus der Not eine Tugend macht. Andere geeigneterere, aber genauso realistische Kriterien bieten sich an, so daß wir auf die allererste Frage zurückgeworfen werden, die sich jeder Neuling in der Korpuslinguistik stellt: Wie groß muß das Korpus sein, das ich brauche?

2.5.2 Qualität

Der Wert für den Parameter *Qualität* ist 'authentisch'. Alles gesammelte Datenmaterial muß aus echten Kommunikationssituationen stammen, in denen

Menschen ihren normalen Geschäften nachgehen. Geht der Eingriff des Sprachwissenschaftlers über das reine Datensammeln hinaus, können die erhobenen Daten allenfalls als Spezialkorpus bezeichnet werden. Nur so kann dem Anspruch derer entsprochen werden, die Aussagen über die Art machen möchten, wie Sprache in normaler Kommunikation verwendet wird, und die sich ansonsten irrtümlicherweise dazu veranlaßt fühlen könnten, Daten, die unter experimentellen oder anderweitig künstlich beeinflussten Bedingungen ermittelt wurden miteinzubeziehen.

Es ist schwierig, die Grenze zu ziehen. Beispielsweise legen es einige Fernsehshows bewußt darauf an, ihre Kandidaten in ungewohnte und sogar bizarre Situationen zu versetzen und so zu sehr seltsamen Antworten zu verleiten. Ebenso wirkt zwanglose Konversation meist improvisiert, sie kann aber durchaus von einer oder mehreren Gesprächspartnern einstudiert worden sein.

Wenn auch die Grenzen oft schwierig zu ziehen sind, so ist es doch wichtig, ein ernsthaftes Eingreifen des Sprachwissenschaftlers oder die Schaffung besonderer Szenarien bei der Benennung des Korpus eigens zu kennzeichnen. Diese gesamte Kategorie könnte beispielsweise *Experimentalkorpus* genannt werden.

Ein weithin bekannter Typ des *Experimentalkorpus* ist z.B. das 'Korpus zu Sprachprozessen' (*speech corpus*), das zusammengestellt wird, um besondere Eigenschaften gesprochener Sprache zu ermitteln. So ein Korpus kann sehr klein sein, wenn etwa Probanden willkürliche Äußerungen in schallschluckenden Räumen zu lesen haben. Die Klassifikation von Korpora zum Bereich Sprachvermögen ist jedoch kein Gegenstand dieses Berichts. (Mehr Informationen zu sog. *speech corpora* sind in den Berichten der EAGLES *Spoken Language Working Group* zu finden.)

Eine besondere Kategorie ist ferner das *literarische Korpus*, von dem viele Arten existieren. Biblische und literarische Gelehrtenkreise hatten schon seit langem die Disziplin der Korpuslinguistik begründet, und in literarischen Zirkeln existiert eine Menge Fachwissen dazu, wie man z.B. einen Kanon der Werke eines bestimmten Autors erstellt. Klassifikationskriterien beinhalten außer dem *Autor* auch *Gattung* (Oden, Kurzgeschichten usw.), *Epoche*, *Gruppe* (augustinische Dichter, Bildungsroman), *Thema* (revolutionäre Literatur). *Drama* wird normalerweise getrennt von Prosa und Lyrik behandelt.

Korpora zur Kindersprache, zur Sprache alter Menschen oder zur Sprache von Nicht-Muttersprachlern und von Dialektsprechern, sowie Korpora, die sehr spezialisierte Kommunikationsgebiete (wie die Sondersprachen der Heraldik, der Strickmuster oder der Versteigerung) umfassen, sollten aufgrund der mangelnden Repräsentativität der verwendeten Sprache auch als *Spezialkorpora* bezeichnet werden.

Es gilt hinzuzufügen, daß das Spezialkorpus sich dem Prinzip nach anders verhält als ein Korpus, in dem die eine oder andere Varietät einer normalen, authentischen Sprache vorkommt. Ein Konversationskorpus ist kein Spezialkorpus, noch ist dies ein Korpus von Zeitungstexten, nicht einmal, wenn es auf eine bestimmte Zeitung beschränkt ist. Es wird hier unterschieden zwischen einer Varietät, die innerhalb der Grenzen des für einen Muttersprachler sinnvoll zu Erwartenden liegt und Varietäten, die aus dem einen oder anderen Grund sehr stark vom allgemeinen Konsens abweichen. Spezialkorpora sind solche, die nicht zur Beschreibung der Alltagssprache beitragen und das entweder, weil sie viele unübliche Merkmale beinhalten, oder weil sie ihrer Herkunft nach nicht verläßlich das Normalverhalten von Menschen widerspiegeln.

In einem solchen Fall illustriert jede Komponente eine bestimmte Art von Sprache, und deshalb sollte jede Komponente so gekennzeichnet sein, daß die Homogenität des darin enthaltenen Datenmaterials angegeben wird. Die Besonderheit der Sprache kann auf Korpus- oder Subkorpusebene beibehalten werden, ohne daß das Korpus der Kategorie *Spezialkorpus* zugeordnet werden muß.

Es wird somit vorgeschlagen, daß *Spezialkorpora* wie folgt gekennzeichnet werden:

Spezialkorpus: Aphasische Lyrik

Ein Korpus, das eine bestimmte Varietät der Normalsprache illustriert, wird wie folgt gekennzeichnet:

Subkorpus der Tageszeitung *The Times*, 1991

2.5.3 Einfachheit

Der Wert für den Parameter *Einfachheit* ist 'einfacher Text'. Das bedeutet, daß der Benutzer eine ununterbrochene Kette von ASCII-Zeichen erwarten kann, in denen jede Auszeichnung klar identifizierbar und abtrennbar vom Text ist. Viele Texte sind heute im SGML-Format verfügbar und in Zukunft eventuell in TEI. Diese sorgfältig entworfenen 'Auszeichnungen' (*mark-ups*) fügen dem Text keine zusätzliche linguistische Information hinzu. Ihre Rolle bei der Textrepräsentation besteht im großen und ganzen allein darin, bei der Umsetzung in eine lineare Kodierung einige Merkmale zu erhalten, die andernfalls verlorengehen. Sie sind durchaus hilfreich, aber ihr Vorhandensein muß verzeichnet werden und der ursprüngliche Text muß leicht wiederherstellbar sein.

Die Auszeichnungskonventionen können auf vielfältige Annotationen ausgedehnt werden, die zusätzliche Informationen wissenschaftlich-linguistischer Art enthalten. Der Grundsatz der Einfachheit richtet sich nicht gegen Annotation und nicht gegen die Verwendung dieser Auszeichnungsvereinbarungen. Es wäre hilfreich, eine klare Unterscheidung zu treffen zwischen 'Auszeichnungen' (*mark-ups*), was soviel bedeutet wie zusätzliche Kodierungen, die ausschließlich Oberflächenmerkmale von Texten beinhalten und *Annotationen*, d.h. zusätzlichen Kodierungen, die Analysen und Interpretationen beinhalten. Diese Unterscheidung muß besonders sorgfältig für Transkriptionen gesprochener Sprache vorgenommen werden, denn man kann erwarten, daß der Transkribent über eine analytische Notation irgendeiner Art verfügt, aber eine, die so konventionell und verbreitet ist, daß man sie wie eine gut durchdachte 'Auszeichnung' (*mark-up*) behandeln kann und die sich eindeutig von Intonationsmarkierungen oder grammatischen *Tags* unterscheidet (siehe 3.).

Schwieriger ist die Frage *annotierter Korpora*. Wir schlagen vor, daß dieser Ausdruck für jedes Korpus verwendet wird, das zusätzliche Informationen in kodierter Form enthält – Herkunft, analytische Zeichen usw. Auch hier sollten die Annotationen auf eine einfache und klar definierte Weise vom ursprünglichen Text abtrennbar sein. Eine Reihe von Konventionen, was das Wegnehmen, Wiederherstellen und Verändern von Annotationen angeht, sind nötig, insbesondere weil zu erwarten ist, daß in den nächsten Jahren immer mehr annotierte Korpora zur Verfügung stehen werden. Es wäre naiv zu glauben, daß große Korpora voller verschiedener Annotationen einfach zu bearbeiten seien; die Abfragezeiten sind jetzt schon kritisch.

2.5.4 Dokumentiertheit

Der Wert für den Parameter *Dokumentiertheit* ist 'dokumentiert'. Das bedeutet, daß, wie im NERC-Bericht vorgeschlagen, alle Einzelheiten über die Konstituenten einer Komponente getrennt von der Komponente selbst abgespeichert werden. Das Modell dafür ist der DTD oder 'Kopf' (*header*) in SGML, und, im Anschluß daran, in TEI. Im Gegensatz zu den bekannten Richtlinien dieser Institutionen scheinen es die Korpusbenutzer vorzuziehen, die Dokumentation von Texten getrennt von den Texten aufzubewahren und nur einen Minimalheader als Verweis auf die Dokumentation hineinzunehmen. Was die Verwaltung von Korpora angeht, so steht diese Praxis in Einklang mit der Trennung von Ursprungstext und Annotation, und bedarf nur eines sehr geringen Programmieraufwandes; da DTDs extrem umfangreich sind, vermindern sie die Leistungsfähigkeit von Suchläufen im Echtzeitmodus, wenn sie nicht abkoppelbar sind.

3. 'Korpus gesprochener Sprache' (*Spoken Corpus*)

Dieser Begriff sorgt oft für beträchtliche Verwirrung, und eine Klärung der Terminologie ist deshalb angebracht. Zunächst einmal ist ein 'Korpus gesprochener Sprache' von einem 'Korpus zu Sprachprozessen' (*speech corpus*) zu unterscheiden, das bereits oben beschrieben wurde (vgl. 2.5.2). Außerdem existieren zwei verschiedene Definitionen: Im Jargon einiger Autoritäten (z.B. Chafe 1995) bezeichnet dieser Begriff ein Korpus von informeller, spontaner Konversation unter Ausschluß sonstiger Medien. Andererseits wird dieser Begriff oft umfassender verwendet, nämlich, um Sprache zu bezeichnen, deren ursprüngliche Präsentation in mündlicher Form erfolgte, d.h., daß die jeweils beteiligten Sprecher sich mündlich äußern. Wenn ein solcher Text später in schriftlicher Form ohne jegliche Veränderung, außer der Transkription, dargestellt wird, so sollte er immer noch als *spoken* eingestuft werden (z.B. eine BBC Reith Vorlesung). Wenn ein 'Korpus gesprochener Sprache' (*spoken corpus*) in absehbarer Zeit sowohl in Form von Schallwellen als auch als Transkript gespeichert wird, so wird ein solcher Text demnach in zwei Versionen existieren, was letztendlich zur Einführung eines speziellen Parallelkorpus führen wird.

Umgekehrt kann jeder Text, der schriftlich niedergelegt wurde, vorgelesen werden, seine Ausdrucksform muß lediglich entsprechend dem neuen Medium verändert werden. Deshalb bleibt er dennoch in erster Linie ein geschriebener Text.

Wir geben hier der weiter gefaßten Definition des 'Korpus gesprochener Sprache' (*spoken corpus*) den Vorzug. Dabei gibt es sicherlich Grenzbereiche, abhängig von der Bedeutung, die man wählt: Ab wann spricht man von spontan, ab wann von informell und so weiter? Wie weiß man, ob ein Textproduzent einen Text überhaupt schriftlich oder mündlich oder in beiden Formen konzipiert hat? Den Begriff hingegen nur auf eine kleine Klasse von Texten, bestehend aus gesprochener Sprache zu beschränken, scheint zur Verzerrung der Bedeutung der betreffenden Wörter zu führen.

Das Problem besteht darin, daß informelles spontanes Sprechen von vielen Gelehrten als die wichtigste Varietät von allen betrachtet wird, die dem Zentrum der Sprache am nächsten steht und somit die charakteristischen Muster einer Sprache in einer Art offenbart, wie es sonst keine Varietät vermag. Darüber hinaus ist sie eine der schwierigsten und teuersten in der Analyse, sie ist schwer klassifizierbar und schlecht zu bearbeiten. Die Widrigkeiten der Transkription machen ein 'Korpus gesprochener Sprache' (*spoken corpus*) zu einem unbefriedigenden Bereich, und es gibt bisher noch keinen Konsens über Konventionen zur Transkription. Am weitesten ist man bisher im NERC Report mit den Empfehlungen fortgeschritten.

4. 'Proben' (*samples*)

Besonders in der Nachfolge des *Brown Corpus* (wie in der Einführung zusammengefaßt) ist es üblich gewesen, eine Auswahltechnik bei der Zusammenstellung von Korpora anzuwenden. Die ausgewählten Textstücke (Proben) sind klein, bezogen auf Texte wie Zeitungen, Bücher, Radioprogramme, und von konstanter Größe und somit, wie bereits zuvor erwähnt, nicht als 'Texte' i. e. S. zu betrachten. Aus diesem Grund wird nicht *Text*, sondern *Sprachstücke* in der Definition von Korpus (s. o.) verwendet. Die demnächst erscheinende Definition von Text wird eine klare Unterscheidung zwischen *Textkorpus*, 'Volltextkorpus' (*whole text corpus*) und *Probenkorpus* treffen. Wir würden *Volltext* gerne als Standardbedingung sehen, und somit Probenkorpora als eine Kategorie der Spezialkorpora klassifizieren, aber es gibt immer noch viele Korpora, die aus kleinen Proben zusammengesetzt sind. Es sollte jedoch klar sein, daß dieses Merkmal nur ein Überbleibsel der beschränkten Möglichkeiten der Vergangenheit darstellt und nicht zur Qualität des Korpus beiträgt. Die Berücksichtigung von Proben konstanter Größe ist daher pseudowissenschaftlich.

5. 'Sondersprachen' (*sublanguages*)

'Sondersprache' (*sublanguage*) ist ein wichtiger Begriff in der Verarbeitung natürlicher Sprache. Man nimmt an, daß durch die Einengung auf eine Untervarietät, üblicherweise in einem fachsprachlichen Kontext, die vorgefundene Sprachstruktur einfacher wird und sich dadurch leichter verarbeiten läßt. Auch das Vokabular ist begrenzt und oft spezifisch. Entsprechende Beschränkungen finden sich auch auf semantischer, konzeptueller und pragmatischer Ebene (Harris 1988).

Sondersprachen werden von internen und externen Kriterien zugleich bestimmt, wobei die internen wichtiger sind. Es bleibt abzuwarten, ob interne und externe Kriterien sich in der Praxis wirklich zuordnen lassen. Die Untersuchung von Genre und von Fachsprachen zeigt, daß sich Schreiber in einem technischen oder professionellen Kontext bei der Textproduktion oft an ausgefeilte Vorschriften halten; somit ist es nicht überraschend, so viele Ähnlichkeiten vorzufinden.

Sondersprachen machen das andere Extrem des sprachlichen Spektrums gegenüber dem Referenzkorpus aus. Ihre strukturelle Homogenität macht sie nutzbar für *natural language processing* (NLP; Verarbeitung natürlicher Sprache), erlaubt es, die Datenmenge niedrig zu halten und zeigt damit typischerweise gute Abgeschlossenheitseigenschaften. Sondersprache muß in ihrer Konzeption von *künstlichen* oder *reduzierten* Sprachen unterschieden werden.

Letztere sind künstlich konstruierte Sprachen, während Sondersprachen sich natürlich entwickeln (obwohl sie auf der Ebene der Terminologie bewußt festgelegt werden können).

Zunehmend stellen Wissenschaftler des Bereichs NLP fest, daß sie Zugang zu Korpora brauchen, die Sondersprachen enthalten, um Spezialtexte behandeln zu können (McNaught 1993). Entsprechend unserer vorangehenden Definition nennen wir Korpora, die aus Sondersprachen bestehen, *Spezialkorpora*.

6. Referenzkorpora

Ein *Referenzkorpus* ist ein Korpus, das konzipiert worden ist, um umfassende Informationen über eine Sprache zu liefern. Es hat den Anspruch, groß genug zu sein, um alle wichtigen Varietäten einer Sprache einschließlich des jeweiligen Wortschatzes zu repräsentieren, so daß es als Grundlage für verlässliche Grammatiken, Wörterbücher, Thesauri und andere Nachschlagewerke über Sprache dienen kann. Das Auswahlmodell definiert gewöhnlich eine Reihe von Parametern für den Einschluß von möglichst vielen soziolinguistischen Variablen und schreibt die Größenverhältnisse für jeden auszuwählenden Texttyp vor. Ein großes Referenzkorpus kann eine hierarchisch geordnete Struktur von Komponenten und Subkorpora aufweisen.

Fragen der *Ausgewogenheit* und *Repräsentativität* kommen in der Diskussion über Referenzkorpora ständig vor. Sie sind sehr schwer zu definieren; dennoch läßt sich gut mit ihnen arbeiten. Während üblicherweise nicht behauptet wird, es gäbe eine „Kernvarietät“ der Sprache, scheint es eine große Zahl von sich stark überschneidenden Varietäten zu geben, denen im Überschneidungsbereich die Masse des Wortschatzes und fast alle Syntaxregeln gemeinsam sind. Sie unterscheiden sich durch marginale Wortschatzeinheiten und eine Individualität in der Phraseologie. Einige allgemeine Eigenschaften, die mit Punkten wie Förmlichkeit, Stil, Elaboriertheit und dem allgemeinen Themenbereich zu tun haben, und eine grobe Repräsentativität werden erreicht, indem man sicherstellt, daß eine große Menge Text jeden dieser Parameter exemplifiziert.

Spezialkorpora bestehen aus Texten, die sich nicht so weitgehend mit dem großen Zentralrepertoire überschneiden. Um erkennbar „in einer Sprache“ zu sein, müssen sie eine ganze Reihe von grammatischen und lexikalischen Eigenschaften dieser Sprache darstellen. Die Markiertheit der speziellen Muster hilft, sie deutlich von den allgemeinen Sprachvarietäten abzusetzen. In absehbarer Zeit und mit dem wachsenden Einfluß interner Kriterien werden Referenzkorpora verwendet werden können, um die Abweichung von Spezialkorpora zu bemessen.

Referenzkorpora stehen im Zentrum künftiger Entwicklung korpusbasierter Arbeit in Europa und anderswo. Referenzkorpora in unterschiedlichen Sprachen, die nach ähnlichen Prinzipien aufgebaut sind, bilden die Gruppen sogenannter *vergleichbarer Korpora* (siehe 9.).

Beispiel: *The Bank of English*

The Bank of English ist ein Referenzkorpus. Aus dem gesamten Datenmaterial wird in Birmingham von Zeit zu Zeit ein Korpus erstellt und mit Hilfe des Internet zusammen mit passender Software weltweit zugänglich gemacht. Gegenwärtig beinhaltet das Korpus um die 320 Millionen Wörter online, wird aber bald die 400 Millionengrenze erreicht haben. Die 320 Millionen setzen sich aus mehreren Subkorpora zusammen:

Tageszeitungen:	133 Millionen Wörter		
Bücher:	75	"	"
Illustrierte:	48	"	"
Radio:	41	"	"
Kurzlebiges:	6	"	"
Informell Gesprochenes:	20	"	"

Die ersten vier Subkorpora sind Proben aus ziemlich reichhaltigem Datenmaterial (obwohl Illustrierte erst unlängst zugänglich gemacht wurden). Es ist eine ganze Menge mehr in petto, um es einmal so auszudrücken, aber nur ein Bruchteil davon ist online zugänglich.

Kurzlebiges, basierend auf einer großen Vielfalt von Pamphleten und Broschüren, muß immer wieder neu eingegeben werden, und das ist mühsam und teuer. Das informelle *spoken corpus*, obwohl winzig im Vergleich zum Rest, ist wahrscheinlich das bisher größte seiner Art.

Diese Subkorpora lassen sich wiederum in weitere Subkorpora zerlegen, so ist z.B. im Bereich der Tageszeitungen *Großbritannien* ein weiteres Subkorpus; dieses gliedert sich wiederum in die Bestandteile *The Times* (10 Millionen) und *The Guardian* (12 Millionen). Das Radiosubkorpus ist unterteilt in *BBC* (18 Millionen), das amerikanische *NPR* (21 Millionen), jedes dieser Subkorpora ist nach chronologischen Kriterien aufgegliedert. Das Subkorpus „Illustrierte“ enthält *The Economist* (7 Millionen) und den *New Scientist* (3 Millionen) und einen Rest, der als allgemeines Subkorpus zusammengefaßt wird.

Die Abfragesoftware erlaubt dem Benutzer, das gesamte Korpus oder jede mögliche Gruppierung seiner Bestandteile abzurufen, und das entweder zeitweilig oder als permanente Grundeinstellung für diesen Benutzer. Es gibt keine Begrenzung, was die Kombination von Zeichen und Kodierungen angeht,

die gewählt werden können, ob mit Lücken, Jokerzeichen oder einer Kombination von beiden. Kombinationen sind besonders einfach zu untersuchen, entweder in einer mutmaßlichen Sondersprache oder im Rahmen des gesamten Korpus, um Übersetzungshilfen oder Lexika zusammenzustellen.

7. Monitorkorpora

Die Einsicht in die Notwendigkeit großer Korpora ist mit der Erfahrung im Umgang mit ihnen gewachsen, und der zunehmende Einzug elektronischer Anwendungen im Verlagswesen hat Sprachdaten in sehr großen Mengen verfügbar gemacht. Vor einigen Jahren schon wurde deutlich, daß die Annahme einer fixierten Grenze für ein Korpus, das einen längeren Zeitraum umfaßt, eine unnötige Einschränkung darstellt. Für bestimmte Zwecke ist eine feste Korpusgröße und Zusammensetzung unabdingbar; aber das läßt sich leicht innerhalb einer großen und ständig in Bewegung befindlichen Sammlung bewerkstelligen. Die Frage, um die es ging, war, wie man die großen Datenmengen würde handhaben können, die für ein sog. *Monitorkorpus* vorgesehen waren.

Das erste Modell war ein Korpus einer konstanten Größe, damit die derzeit verfügbare Software damit umgehen könnte. Dieses Korpus würde ständig durch neues Datenmaterial aufgefrischt, während gleichgroße Mengen alten Datenmaterials in ein externes Archiv ausgelagert würden. Die jeweilige Zusammensetzung des Korpus wäre also vergleichbar mit der Zusammensetzung zu einem früheren oder späteren Zeitpunkt.

Dieses Modell bildete die Grundlage zu dem Gedanken einer *Durchflußrate* als dem besten Weg, das Korpus zu handhaben. Anstelle von angenommen 10 Millionen Wörtern als der richtigen Größenordnung für ein bestimmtes Genre könnte man genausogut 10 Millionen Wörter pro Jahr oder pro Monat oder pro Woche festlegen. Die Sprache würde durch den Computer fließen, so daß zu jedem Zeitpunkt eine geeignete 'Probe' (*sample*) vorrätig wäre, die mit früheren und kommenden Zuständen vergleichbar wäre.

Solch ein Modell würde denen, die an der Verarbeitung natürlicher Sprache interessiert sind, neue Aussichten eröffnen; und es würde den gegenwärtigen Korpora eine zusätzliche Dimension geben, nämlich die diachrone. Neue Wörter könnten aufgefunden werden, ebenso Entwicklungen im Sprachgebrauch, die zu Bedeutungsveränderungen führen. Aus einem solchen Korpus könnten Langzeitnormen der Häufigkeitsverteilung sowie eine Vielzahl anderer Informationen abgeleitet werden.

Einige Wissenschaftler waren mit der Vorstellung, alte Texte auszulagern, wenn neue dazukämen, nicht einverstanden. Das Problem läßt sich jedoch durch die ständig steigende Rechnerleistung und Speicherkapazität neutralisieren. Es gibt keine Notwendigkeit mehr, Texte auszulagern. Will man indessen ein Monitorkorpus möglichst ökonomisch verwalten, empfiehlt es sich auch weiterhin, es in Abschnitte annähernd gleicher Größe und Zusammensetzung zu zerlegen.

Die Größenverhältnisse der Komponenten eines Monitorkorpus werden sich langfristig ändern. Neue Datenquellen werden verfügbar sein, und neue Verfahren werden dazu führen, daß bislang spärliches Datenmaterial zukünftig in großer Menge zur Verfügung steht. Die Durchflußrate muß daher von Zeit zu Zeit angepaßt werden.

8. Parallele Korpora

Ein paralleles Korpus ist eine Sammlung von Texten, von denen neben der originalsprachlichen Fassung eine oder mehrere Übersetzung(en) vorliegen. Im einfachsten Fall sind nur zwei Sprachen beteiligt: ein Korpus ist eine exakte Übersetzung des anderen. Einige Parallelkorpora existieren jedoch in mehreren Sprachen. Außerdem muß die Übersetzungsrichtung nicht notwendigerweise konstant sein, so daß einige Texte eines Parallelkorpus von Sprache A in Sprache B übersetzt werden, andere jedoch umgekehrt. Die Übersetzungsrichtung kann sogar völlig unbekannt sein.

Die Nachfrage nach parallelen Korpora ist sehr groß, da sie die Möglichkeit bieten, Original und Übersetzung in direkte Beziehung zueinander zu setzen und dadurch neue Einsichten in das Wesen der Übersetzung zu gewinnen. Auf der Grundlage solcher Untersuchungen hofft man, Werkzeuge für computergestützte Übersetzungsverfahren entwickeln zu können. Mit Hilfe solcher Korpora können maschinelle Übersetzungssysteme auf der Basis probabilistischer Verfahren erprobt werden. Parallele Korpora kommen im Kommunikationsalltag mehrsprachiger Einrichtungen, wie der Vereinten Nationen, NATO, der EU und mehrsprachiger Länder wie z.B. Kanada zustande.

9. Vergleichbare Korpora

Ein *vergleichbares Korpus* ist eines, das ähnliche Texte in mehr als einer Sprache oder Varietät enthält. Bisher jedoch gibt es keine Übereinstimmung bezüglich der Kriterien für Vergleichbarkeit, denn es gibt nur wenige Beispiele für vergleichbare Korpora. Ein typischer Vertreter ist das ICE, das *International Corpus of English* (Greenbaum, 1991). Korpora von ungefähr einer Mil-

lion Wörtern in jeder der vielen Varietäten des Englischen rund um den Erdkreis werden nach demselben Muster, das die Genres und eine Zielgröße vorschreiben, aufgebaut. Ursprünglich sollten die Korpora alle im selben Jahr zusammengestellt werden.

Die Möglichkeiten eines vergleichbaren Korpus liegen darin, verschiedene Sprachen oder Varietäten in ähnlichen Kommunikationsbedingungen zu vergleichen und dennoch die unvermeidliche Verzerrung durch Übersetzung, wie man sie in parallelen Korpora findet, zu vermeiden.

Anmerkung: **Multilinguale Korpora**

Gegenwärtig existieren keine *multilingualen Korpora*, abgesehen von parallelen und vergleichenden Korpora. Tatsächlich gibt es viele Zentren, die Textmaterial in mehreren Sprachen gesammelt haben, von denen einige eigenständige Korpora darstellen. Aber solange diesen Sammlungen zumindest auf der Ebene vergleichbarer Korpora keine gemeinsamen Auswahlkriterien zugrundeliegen, stellen sie lediglich Textressourcen in verschiedenen Sprachen dar. Es scheint daher nicht dienlich zu sein, darauf den Begriff *multilinguales Korpus* anzuwenden.

10. Literaturverzeichnis

- Atkins, Sue/Clear, Jeremy/Ostler, Nick (1992): Corpus design criteria. In: *Literary and Linguistic Computing*, 7, S. 1-16.
- Biber, Douglas (1988): *Variation across speech and writing*. Cambridge.
- Chafe, Wallace (1995): Conference Contribution. In: Leech, Geoffrey/Myers, Greg/Thomas, J. (Hg.): *Spoken English on Computer*. London.
- Greenbaum, Sydney (1991): The Development of the International Corpus in English. In: Aijmer, Karin/Altenberg, Bengt (Hg.): *English Corpus Linguistics*. London.
- Harris, Zellig S. (1988): *Language and Information*. New York.
- Hofland, Knuu/Johansson, Stig (1982): *Word frequencies in British and American English*. London.
- Jones, Susan/Sinclair, John (1974): *English Lexical Collocations*. In: *Cahiers de Lexicologie*, Institut des Professeurs de Français à l'étranger.
- Kucera, Henry/Francis, W. Nelson (1967): *Computational analysis of present day American English*. Providence.
- M^cNaught, Jock (1993): User needs for textual corpora in natural language processing. In: *Literary and Linguistic Computing*, 8/4, S. 227-234.
- Nakamura, Junsaku (1993): Text Types. In: Baker, Mona/Francis, Gill/Tognini-Bonelli, Elena (Hg.): *Text and Technology*. Amsterdam.

NERC (1997): Feasibility of a network of European reference corpora. Report submitted to the CEC. Pisa. (Ersch. demn. als Buch.)

Sinclair, John (Hg.) (1988): Looking up. London.

Wolfgang Teubert

Korpus und Neologie

1. Neologismen und Korpus

In Deutschland gibt es, anders als in England, Rußland oder Frankreich, keine ausgeprägte Tradition der Neologismen-Lexikographie. Dafür gibt es sicher mehrere Gründe. In erster Linie dürfte es indessen mit der spezifisch deutschen Einstellung zum eigenen Wortschatz zu tun haben. Immer noch herrscht die Vorstellung vor, daß dieser Wortschatz in zwei distinkte Teile zerfällt. Da sind einmal die autochthonen deutschen Wörter, und da sind andererseits die Fremdwörter, die nicht wirklich zu unserer Sprache gehören und nur in Ausnahmefällen naturalisiert, zu Lehnwörtern erklärt werden können. Die deutschen Wörter beruhen auf einer begrenzten Zahl alter Wurzeln, sind Ableitungen oder Zusammensetzungen von ihnen. Neue autochthone Wörter können durch Anwendung produktiver Wortbildungsmuster gebildet werden. Sie lassen sich deshalb, so glaubt man, unmittelbar verstehen. Danach vergeht in aller Regel eine geraume Zeit, bis sie schließlich lexikalisiert, also wörterbuchfähig werden. Ein besonderes Wörterbuch für solche „deutschen“ Neologismen braucht man also nicht.

Neologismen im eigentlichen Sinn, also wirklich neue lexikalische Ausdrücke zur Bezeichnung neuer Denotate, finden sich nach herkömmlicher Anschauung eher als Fremdwörter, also entweder als direkte Übernahmen aus anderen Sprachen oder als Lehnbildungen mit überwiegend lateinischem und griechischem Wortmaterial. Einen positiven Beitrag zur Sprachkultur liefern Fremdwörter, wie viele meinen, nicht. Daß es besondere Wörterbücher für sie gibt, ist eher ein Indiz dafür, daß sie uns immer fremd und unverständlich bleiben. Denn wir Deutsche benutzen unsere Gebrauchswörterbücher kaum, um Bedeutungen nachzuschlagen. Wir wenden uns an sie mit Rechtschreibproblemen und Fragen der Flexion. Anders bei Fremdwörterbüchern: Diese benutzt man, um Bedeutungen zu ermitteln.

Neologismenwörterbücher, Wörterbücher also, die Sprachwandel (und damit ja auch gesellschaftlichen Wandel) beschreiben, sind folglich in dieser Sicht schlechthin überflüssig. Entweder verstehen sich neue Wörter gleichsam von selbst, wenn sie nämlich von deutschen Wörtern und Wurzeln abgeleitet sind, oder sie sind unserer Sprache und unserem Wesen eh fremd – und dann sollte man ihnen keine spezielle Aufmerksamkeit widmen, sondern sie bei Gelegenheit in Fremdwörterbüchern verzeichnen.

Diese Vorstellung von Neologismen ist natürlich falsch. Im nächsten Abschnitt werden wir deshalb zu klären haben, was wir sinnvollerweise unter einem Neologismus zu verstehen haben. Daraus wird indessen keine erleuchtete Klarheit resultieren können, wohl aber ein gesteigertes Bewußtsein von Problemen, die sich, je nach Erkenntnisinteresse, begründet entscheiden lassen.

Setzen wir zunächst einmal voraus, wir wüßten, welche lexikalischen Einheiten wir als neue Wörter zu betrachten haben. Es stellt sich dann die Frage, wie man sie in Texten identifiziert. Zwei Extreme sind denkbar. Entweder der Lexikograph verläßt sich auf seine professionelle Erfahrung und auf seine Kompetenz als Sprachteilhaber und notiert sich in der alltäglichen sprachlichen Interaktion mit Mitgliedern seiner Sprachgemeinschaft, beim Lesen oder sonstigem Medienkonsum jedes Wort, das er für einen Neologismus hält. Die Auswahl, die er so trifft, ist subjektiv. Er wird kaum vermeiden können, Neologismen zu übersehen. Manche Wörter, die er für Neologismen hält, hat es vielleicht schon lange gegeben. Dafür mögen andere Wörter in der Tat neu sein, die ihm längst etabliert vorkommen. Vor allem aber wird er nur zu solchen lexikalischen Einheiten Zugang haben, die ihm in seiner sprachlichen Interaktion begegnen, in seiner privaten und beruflichen Lektüre, in Gesprächen mit Freunden und Kollegen, in individuell gewählten Fernsehprogrammen. Die Texte, mit denen der Lexikograph solchermaßen konfrontiert ist, sind eine winzige Auswahl aus dem gesamten Textuniversum einer Sprache, eine Auswahl, die zudem von subjektiven Interessen gesteuert und damit alles andere als ausgewogen ist. Mit dem gesamten Gesamtwortschatz einer Sprache wird es kein Mensch je zu tun haben. Manche Linguisten behaupten, ausgebaute Sprachen wie das Deutsche oder Französische verfügten über einen Wortschatz von etwa einer halben Million. Gleichzeitig wird der passive Allgemeinwortschatz sprachlich aktiver und vielseitiger Menschen auf nicht mehr als 50.000 bis 100.000 Wörtern geschätzt, und allein dadurch wird das Mißverhältnis evident. Eine Neologismenliste, die sich wesentlich auf die zufälligen Erkenntnisse eines einzelnen Menschen stützt, kann nicht befriedigen, wenn sie den lexikalischen Wandel der deutschen Sprache belegen soll.

Eine breitere Basis der Textauswahl ist also unbedingt erforderlich. So könnten sich beispielsweise mehrere Lexikographen zusammentun und im übrigen verfahren, wie oben beschrieben. In der Tat ist das wohl das methodische Verfahren, das den meisten (englischen und französischen) Neologismenwörterbüchern immer noch zugrundeliegt.

Das entgegengesetzte Extrem im methodischen Sinne läge in einem rein empiriegesteuerten Verfahren, das im Prinzip wenigstens alle subjektiven Einflüsse ausschließen würde. Dazu gehört zweierlei: Zum einen müßten Neologismen in beliebigen Texten automatisch erkannt werden, etwa auf der Grundlage formaler, algorithmisierbarer Kriterien. Zum anderen müßte das gesamte Text-

universum einer Sprache zugrundegelegt werden, ein gewiß utopisches Unterfangen.

Alles spricht dafür, beide Verfahren in ihren jeweiligen Stärken zu kombinieren. Einerseits ist und bleibt die Kompetenz und Professionalität des Lexikographen unverzichtbar, wenn es darum geht, zu entscheiden, ob mögliche Kandidaten für Neologismen wirklich buchenswert sind. Andererseits sollte die Auswahl der Kandidaten auf einer möglichst breiten und objektiven Basis stattfinden. Wenn es nicht das ganze Textuniversum sein kann, muß es ein Korpus sein, eine Textsammlung also, deren Texte nach nachvollziehbaren Kriterien ausgewählt sind.

Seit den sechziger Jahren gibt es nun schon maschinenlesbare Korpora als empirische Datensammlung für die sprachwissenschaftliche Arbeit. Die ersten beiden Jahrzehnte der Korpuslinguistik waren von der Suche nach dem sogenannten repräsentativen Korpus bestimmt, einem Korpus also, das repräsentativ für die Gesamtsprache ist, sie quasi im Kern abbildet. Inzwischen hat sich die Einsicht, daß es Repräsentativität in diesem Sinn nicht geben kann, weithin durchgesetzt. Doch immer noch gibt es viele Linguisten, die nur das Adjektiv *repräsentativ* durch *ausgewogen* ersetzt haben, für die aber das Konzept dasselbe geblieben ist. Wir werden uns deshalb im zweiten Abschnitt mit dem heutigen Verständnis der Korpusproblematik beschäftigen. Besondere Aufmerksamkeit schenken wir dabei dem Prinzip der Sättigung, das helfen kann, die optimale Größe von Korpora zu bestimmen.

Im vierten Abschnitt geht es zunächst um die Ermittlung von Parametern, die einem möglichen speziellen Neologismenkorpus zugrunde gelegt werden könnten. Wenn Repräsentativität oder Ausgewogenheit keine Parameter sein können, wo lassen sie sich dann finden? Wie läßt sich eine bewußt nicht auf Ausgewogenheit abzielende Auswahl dennoch objektivieren? Wir plädieren am Ende unseres Beitrags für ein komplementäres Miteinander von opportunistischem Monitorkorpus und intellektuell zusammengestelltem Diskurskorpus, um mit diesen beiden Datenspeichern das Phänomen der Neologie sowohl in der Breite zu beschreiben als auch in der Tiefe zu erklären.

2. Was sind Neologismen?

Ein Abgleich eines Jahrgangs des *Wall Street Journal* mit einem führenden amerikanischen Wörterbuch ergab, daß etwa ein Drittel der vorgefundenen Wörter nicht im Wörterbuch zu finden war. Davon waren allerdings nur ein recht kleiner Teil wirkliche Neologismen. Die große Mehrheit bestand aus Eigennamen, also geographischen Namen und Personennamen, und ebenfalls befand sich unter diesen nicht belegten Wörtern eine große Zahl von Druck-

fehlern. Für das Deutsche ergibt sich ein ähnliches Bild. Am Centrum für Informations- und Sprachverarbeitung (CIS, Universität München) wird seit Jahren täglich für jede neue Nummer der *Süddeutschen Zeitung* automatisch eine Liste der Wörter (Wortformen) erstellt, die in den vorausgegangenen vier Jahrgängen nicht belegt waren. Nach Ausfilterung von Eigennamen und Druckfehlern enthält diese Liste im Durchschnitt etwa 500-550 Einträge.¹ Die Mehrzahl davon sind ad hoc gebildete Zusammensetzungen und Ableitungen, Okkasionalismen also, die nicht in ein Wörterbuch gehören. Dazu kommen eine Reihe von Wörtern, die bereits in Wörterbüchern verzeichnet sind, aber zufällig zuvor nicht belegt waren. Übrig bleiben pro Tag vielleicht ein oder zwei potentielle Neologismen. Was aber sind Neologismen?

2.1 Neologismen als neue lexikalische Ausdrücke

Neologismen sind zum einen neue Zeichenketten, die weder Eigennamen noch flektierte Formen bereits belegter Wörter noch falsch geschriebene Wörter sind. Anders gesprochen sind Neologismen bisher in Texten oder Wörterbüchern nicht belegte lexikalische Ausdrücke unabhängig von ihrer Bedeutung. Das Problem ist dabei die Belegung. Wörterbücher sind prinzipiell nicht vollständig. Verlässlich sind auch in unseren Großwörterbüchern nicht mehr als dreißig Prozent des Gesamtwortschatzes beschrieben; damit werden bei weitem nicht einmal die Vokabulare abgedeckt, die in der fachexternen Kommunikation, beispielsweise in Sachbüchern, Verwendung finden. Die Tatsache, daß ein Text Wörter enthält, die in keinem Wörterbuch stehen und die vielleicht auch bisher nicht in lexikalisch analysierten Texten zu finden waren, bedeutet folglich nicht, daß es das Wort nicht schon lange gegeben hat. Sind beispielsweise fast obsoletere Wörter wie *Schutzwaffen*, die über lange Zeit extrem selten belegt und in Wörterbüchern kaum zu finden waren, Neologismen, wenn sie plötzlich wieder in aller Munde sind?²

Eine andere Frage in diesem Zusammenhang betrifft die Dichotomie System und Gebrauch: Es wird gesagt, daß Neologismen erst dann vorliegen, wenn die betreffende lexikalische Einheit Teil des Sprachsystems geworden ist. Dazu gehört allgemein die Vorstellung, daß sich der Gebrauch in einer bestimmten Weise verfestigt hat, wofür in einem Korpus etwa Parameter wie Frequenz und Distribution maßgeblich wären. Eine lexikalische Einheit müßte erst in einer bestimmten Häufigkeit und über eine bestimmte Breite von Textsorten und Genres nachgewiesen sein, bevor ihr der Status eines Neologismus zukommt. Erreicht sie dieses Ziel nicht, hat sie als Okkasionalismus zu gelten

¹ Siehe CORDON-Proposal, 1996, S. 4. Diese Zahl wurde auf der Grundlage der Auswertung eines Tages im Januar 1993 ermittelt.

² Vgl. hierzu Teubert (1988).

und gehört als solche nicht zum Sprachsystem. Hierher gehören beispielsweise viele idiolektale Idiosynkrasien.

System und Gebrauch betreffen auch Parameter wie Regiolekte, Soziolekte, Gruppensprachen usw. In Randlagen oder in der Isolierung finden sich verstärkt fremdsprachliche Bestandteile in Texten, wie die Analyse deutschsprachiger Periodika im Elsaß, in Luxemburg, in Südtirol, in Australien oder Kanada zeigen. Sie gehören dort zu den gängigen und oft auch in Flexion und Schreibweise eingedeutschten Wörtern. Verzeichnet werden sie jedoch allenfalls in Spezialwörterbüchern. Dasselbe gilt auch für Soziolekte und Gruppensprachen. In bestimmter Weise gehören sie zum Sprachsystem, denn sie können – anders als echte fremdsprachliche Bestandteile – relativ problemlos in standardsprachliche literarische Texte integriert werden.

Was also wäre Sprachwandel oder lexikalischer Wandel auf der ‘Systemebene’? Gäbe es daneben auch einen Wandel auf der ‘Gebrauchsebene’? Zunächst einmal muß festgestellt werden, daß ‘System’ in diesem Sinn wenig mit System im Sinne Coserius (1970) zu tun hat. Was eher gemeint ist, ist, daß Wörter aus eher peripheren Bereichen wie funktional determinierten Sondersprachen oder aus sozial determinierten Gruppensprachen in den Bereich eindringen, der gemeinhin Standardsprache genannt wird, also den Kern der Allgemeinsprache. Was in den peripheren Bereichen geschieht, gehört nach Coseriu selbstverständlich auch zum Sprachsystem. Für ihn reduziert sich der Systembegriff indessen darauf, was sprachintern stattfindet, und das ist in erster Linie die Grammatik. Dagegen hat lexikalischer Wandel auch immer mit einem Wandel unserer Umwelt oder wenigstens unserer Wahrnehmung davon zu tun. Unter Systemgesichtspunkten wäre Coseriu zufolge allenfalls die Integration von fremdsprachigen Wörtern relevant, die nicht der deutschen Phonetik entsprechen, z.B. *mouse* als Zubehör für einen PC (korrekt ausgesprochen mit einem stimmhaften *s* im Auslaut, den es im deutschen Lautsystem nicht gibt). Wenn manche Lexikographen dagegen zwischen ‘Systemebene’ und ‘Gebrauchsebene’ unterscheiden, meinen sie damit etwas anderes, nämlich die Frage nach der Kanonisierung. Danach sind Wörter kanonisiert, wenn sie in Wörterbüchern aufgeführt werden, und damit gehören sie zum ‘System’. Nicht zum ‘System’ gehört demnach alles, was sich noch nicht etabliert hat. Aber damit erfahren wir nichts darüber, was geschehen sein muß, damit ein neues Wort als etabliert gelten kann, wenn es also in Wörterbücher aufgenommen werden kann. Als Hilfsmittel zur Identifikation von Neologismen taugt die Unterscheidung von ‘System-’ und ‘Gebrauchsebene’ also nicht. Wir werden gleichwohl auf diesen Punkt weiter unten zurückzukommen haben.

Nicht zu den Neologismen wollen wir die Okkasionalismen zählen. Das sind neue Zeichenketten, die entweder in den zu analysierenden Texten sehr selten, vielleicht nur als *hapax legomenon* vorkommen, zudem beschränkt sind auf nur wenige Texte, die nur zu einer oder zwei Textsorten bzw. Genres gehö-

ren. Damit ist implizit wenigstens ein Parameter angesprochen, der für das Neologismenproblem fundamental ist, nämlich die Dimension der Zeit. Wörter sind neu, wenn es sie bisher nicht gegeben hat (wenigstens nicht in dem Bereich, für den unsere Untersuchung gilt, also für den Kernwortschatz der Sprache unter Einbeziehung der Vokabulare der fachexternen Kommunikation, also solcher Fachtexte, die nicht nur für die eigentlichen Experten bestimmt sind). Es reicht jedoch nicht, daß ein Wort nur einmal belegt ist, um sich als Neologismus zu qualifizieren. Es muß über einen längeren Zeitraum verteilt auf mehrere Texte, Textsorten bzw. Genres vorkommen. Wenn eine dieser Bedingungen nicht erfüllt ist, zählt die lexikalische Einheit nur als Okkasionalismus. Das führt zu der paradoxen Einsicht, daß man von einem Neologismus erst dann sprechen kann, wenn das betreffende Wort, bezogen auf das Korpus, über einen längeren Zeitraum und in einer bestimmten Häufigkeit belegt ist – also nicht mehr neu ist.

Bisher nicht berücksichtigt haben wir die Wörter, die nach produktiven Wortbildungsmustern als zuvor nicht belegte Ableitungen oder Zusammensetzungen gebildet sind. Die Frage, ob sie als Neologismen gelten, stellt sich hier anders als bei den Okkasionalismen. Denn es entspricht üblichen lexikologischen Prinzipien und gängiger lexikographischer Praxis, solche Bildungen unabhängig von ihrer Vorkommenshäufigkeit erst dann ins Wörterbuch aufzunehmen, wenn sie 'lexikalisiert' sind. Sie gelten dann als 'lexikalisiert', wenn sie eine Bedeutung haben, die weiter oder enger oder einfach anders ist als die, die nach den Bildungsregeln zu erwarten wären. Metaphorisierung ist dafür ein Hauptgrund. Die Größenordnung von Fällen, wo die Metapher von Anfang an die Bedeutung konstituiert und die sogenannte wörtliche Bedeutung nie zum Tragen kommt, ist beträchtlich. So ist beispielsweise das während des deutschen Vereinigungsprozesses häufiger belegte Wort *Mauer-specht* immer nur als Metapher verstanden worden.³ Häufiger dürfte jedoch der Fall von neuen Fachwörtern sein, die nach produktiven Mustern gebildet sind. Ein Beispiel dafür ist das seit 1993 belegte Wort *Lohnabstandsgebot*. Im Prinzip läßt diese Bildung offen, ob der Lohn jemandem oder etwas gebietet, Abstand zu halten, oder ob jemand gebietet, daß der Lohn zu etwas Abstand halten soll, oder daß etwas zum Lohn Abstand halten soll. Gemeint (und damit die lexikalisierte Bedeutung) ist jedoch (in der Fachsprache der Verwaltung), daß die Sozialhilfe einen bestimmten Abstand zu den unteren Lohngruppen einhalten soll (Teubert 1996).

Damit haben wir eine Reihe von Kriterien behandelt, die für die Entscheidung wichtig sind, ob neue lexikalische Ausdrücke Neologismen sind oder nicht:

³ Zur Frage der Vorgängigkeit der Metapher vgl. auch Cooper, David E. (1986): *Metaphor*. Oxford (Blackwell). Giovanni Battista Vico ist meines Wissens der erste, der Neologismenbildung aus Metonymie und Metaphorik erklärt.

- Gegenstand sind Wörter, nicht Wortformen.
- Falsch geschriebene Wörter bleiben unberücksichtigt.
- Neue Namen natürlicher und juristischer Personen sowie geographische Namen können in der Regel nicht Neologismen sein.
- Daß ein bestimmtes Wort bisher nicht in Wörterbüchern verzeichnet war, ist kein hinreichendes Kriterium.
- Die Belegdauer ist wichtig. Neologismen müssen mindestens ein Jahr (oder zwei Jahre?) alt sein.
- Beleghäufigkeit ist wichtig.
- Distribution der Belege über Texte, Textsorten (oder Textsparten), Genres usw. ist wichtig.
- Lexikalischer Wandel betrifft zwar den Gesamtwortschatz einer Sprache. Neologismenlexikographie sollte sich jedoch auf die Alltagsprache unter Einschluß der Vokabulare der fachexternen Kommunikation beschränken. Wir wollen diesen Sprachausschnitt im weiteren als Standardsprache bezeichnen. Damit geht es um den Wortschatz, wie er in standardsprachigen Wörterbüchern verzeichnet ist. Diese Einschränkung ist normativ. Sie ergibt sich nicht schlechthin aus dem Untersuchungsgegenstand.
- Was Standardsprache ist, wird über Texte definiert, nicht über Wörter.
- Ableitungen und Zusammensetzungen gelten nur dann als Neologismen, wenn ihre Bedeutung nicht regelhaft abgeleitet werden kann.

2.2 Andere Neologismen

Auch bereits in Wörterbüchern und Texten belegte lexikalische Ausdrücke können Neologismen sein, wenn neben vorhandene Bedeutungen eine neue Bedeutung tritt oder ein bisher monosemes Wort neuerdings polysem zu einem Homonym wird. Solche Neologismen sind vermutlich weit häufiger als auf neuen lexikalischen Ausdrücken beruhende Neologismen.

Allerdings sind wir damit mit einem Problem konfrontiert, für das es keine einfache Lösung gibt. Was konstituiert eine neue Bedeutung? So scheint es beispielsweise, als würde sich in den letzten Jahre neben die bereits bestehenden Bedeutungen für Kind eine neue herausbilden, in der 'geborene und ungeborene Kinder' zusammengefaßt werden. Natürlich konnte man immer schon von *ungeborenen Kindern* sprechen, doch wurde dieser Ausdruck dann als ad hoc verwendete Metapher verstanden, um auf bestimmte mit dem Wort *Embryo* verbindbare Attribute aufmerksam zu machen. Heute hingegen gibt es in der Alltagssprache immer mehr Texte, in denen systematisch statt von *Embryos* von *Kindern* gesprochen wird.

Auch hier kann wieder nur gelten, daß solche neue Bedeutungen dann einen lexikalischen Ausdruck zum Neologismus machen, wenn sie über einen hinreichend langen Zeitraum in hinreichend vielen Texten verteilt auf mehrere Textsorten oder Genres belegt sind. Dasselbe gilt auch für Ableitungen und Zusammensetzungen. Wie wir oben gesehen haben, sind auch häufigere Bildungen nicht 'lexikalisiert', solange ihre Bedeutung regelhaft auf der Grundlage des einschlägigen Bildungsmusters abgeleitet werden kann. Erst eine neue Bedeutung, die nicht erschließbar ist, oder die Auswahl einer Bedeutung aus einer Vielzahl möglicher Bedeutungsvarianten rechtfertigt die Einordnung als Neologismus.

In diesem Zusammenhang müssen wir wieder auf das Stichwort Gebrauch oder Verwendung zurückkommen. Schon vor der Wiedervereinigung konnte gelegentlich von *sozialistischer Zwangswirtschaft* gesprochen werden. Als Bezeichnung für Wirtschaftssysteme, die im Gegensatz zu den Prinzipien der freien Marktwirtschaft organisiert sind, erschließt sich die Bedeutung von Zwangswirtschaft aus dem zugrundeliegenden Wortbildungsmuster. Seit der Wiedervereinigung aber ist zum einen die Verwendungshäufigkeit stark angestiegen und zum anderen hat sich in den meisten Belegen die Bedeutung verlagert auf 'zum Scheitern verurteiltes Wirtschaftssystem, wie es im Ostblock üblich war'. Natürlich läßt sich argumentieren, daß die Bedeutungseinengung kontextuell verursacht ist, das Wort also weiterhin auch etwa für die Kriegswirtschaften kapitalistischer Länder während des Zweiten Weltkriegs verwendet werden könnte. Um eine operationalisierbare Lösung zu finden, sollte man zwischen Bedeutung und Denotat unterscheiden. Sowohl das Wort Zwangswirtschaft wie das Wort Droge läßt sich in seiner Bedeutung beschreiben, ohne daß auf konkrete Denotate (z.B. 'Wirtschaftssystem der DDR' bzw. 'Heroin') verwiesen wird. Wenn aber Wörter immer konsequenter nur für ganz bestimmte Denotate verwendet werden, kann es soweit kommen, daß es einer Rechtfertigung bedarf, wenn sie zwar in ihrer habituellen Bedeutung, aber für ein unübliches Denotat verwendet werden. So muß man etwa sagen: *Auch Zigaretten sind Drogen, denn sie machen süchtig und sind der Gesundheit abträglich*. Daß *Zigaretten* aber nicht genuin als mögliches Denotat des Wortes Droge verstanden wird, zeigt sich beispielsweise an Zusammensetzungen mit *Droge*, beispielsweise *Drogenpolitik*, wo es ausdrücklich nicht um Zigaretten geht.

Bedeutungswandel und Denotatswandel hängen eng zusammen. Wird ein Wort habituell nur noch für ein bestimmtes Denotat verwendet, engt sich auch die Bedeutung entsprechend ein. So ist ein *Wagen* zwar immer noch eine Vorrichtung, die auf Rädern fährt und zum Transport von Menschen oder Sachen dient, aber daneben hat sich durch ständige Verwendung für das Denotat *Auto* eine zweite Bedeutung 'Auto' entwickelt, die sinnvollerweise neben der ursprünglichen Bedeutung zu verzeichnen ist. Wann dies zu geschehen hat, bemißt sich nach den oben aufgeführten Kriterien.

2.3 Bedeutung und Begriff

Wir können Wörter der Alltagssprache danach unterscheiden, ob sie Fachwörter oder *termini technici* sind oder nicht. Zwar gibt es keine klare Trennungslinie, und nicht wenige Wörter lassen sich sowohl als 'normale' Wörter wie auch als Fachwörter verwenden. Ein bekanntes Beispiel dafür sind die Ausdrücke *Besitz* und *Eigentum*, die natürlich allgemeinsprachlich gebraucht werden können (*Sie sind auf ihren Besitz ans Meer gefahren.*), aber auch in der Allgemeinsprache fachsprachlich (*Das Anwesen ist seit langem ihr Eigentum.*). Die häufigsten Namen für Tiere und Pflanzen sind beides gleichzeitig; nur selten wird unterschieden: Das ist doch keine *Pflaume*, das ist eine *Zwetschge*. Ein reines Fachwort, das auch in der Allgemeinsprache vorkommt, ist *Plutonium*, ein anderes wäre *CD-ROM*.

Was bedeutet eine solche Unterscheidung für Neologismen? Feststeht, daß Neologismen 'normale' Wörter oder eben Fachwörter sein können. Aus sprachwissenschaftlicher Sicht von Interesse sind aber in erster Linie die 'normalen' Wörter, Fachwörter dagegen nur insoweit, als sie auch semantische Eigenschaften haben, die 'normalen' Wörtern zukommen. Fachwörter wie *Bruttoinlandsprodukt*, *Fraktal*, *BASIC*, *Retrovirus*, die seit ein paar Jahrzehnten in der Standardsprache zu finden sind, sind für die Sprachlexikographie ohne Belang. Mit ihnen muß sich die Sachlexikographie beschäftigen. Zur Sprachlexikographie gehören dagegen Neologismen wie *organisierte Kriminalität*, *Begrüßungsgeld*, *Mobbing*, *affengeil* usw. Diese Wörter sagen anders als die Fachwörter etwas aus über die Welt, in der wir leben. Sie stellen eine intentionale Beziehung zwischen dem Bewußtsein und unserer Umgebung her, sie vermitteln die sensorische Wahrnehmung an die mentale Ebene. Fachwörter dagegen sind lediglich lexikalische Ausdrücke für Begriffe, die bereits in Schubladen abgelegt sind. Solche Begriffe sind vor jedem Wandel gefeit.

Der wesentliche Unterschied zwischen 'normalen' Wörtern und Fachwörtern liegt darin, wer darüber entscheidet, ob ein Wort richtig verwendet wird. Bei Fachwörtern ist dies der Experte. Er entscheidet darüber, ob ein Satz wie *Retroviren haben keine Chromosomen* richtig ist oder nicht. Dagegen ist ein Satz wie *Die Stadtverwaltung ist Teil der organisierten Kriminalität* eine Meinungsäußerung. Ob sie richtig ist oder nicht, kann kein Experte entscheiden. Es gibt keine verbindliche Definition dafür, was *organisierte Kriminalität* ist. Jeder kann bei der Frage mitreden, ob in einem gegebenen Fall organisierte Kriminalität vorliegt. Dagegen weiß nur der Experte, was mit dem Begriff *Retrovirus* gemeint ist. Wenn wir sagen, daß der Ausdruck *Retrovirus* den Begriff 'Retrovirus' bezeichnet, so grenzen wir damit Begriff (bei Fachwörtern) von Bedeutung (bei 'normalen' Wörtern) ab, wie das auch von Cherubim gefordert wird (1979, S. 320). Burkhardt führt diese Unterscheidung wie folgt aus:

Die Aufgabe der Bedeutung ist die intersubjektive Verständigung, der Begriff zielt auf Wahrheit bzw. Angemessenheit der durch ihn komprimierten Beschreibung. Insofern besteht der Unterschied zwischen Bedeutung und Begriff vor allem darin, daß Bedeutung auf der Übereinstimmung zwischen Menschen der alltäglichen Praxis innerhalb der Lebensform beruht, der Begriff aber abhängt von Annahmen über seine Übereinstimmung mit seinem Gegenstand oder vielmehr seinen intensionalen Merkmalen mit den Eigenschaften der Gegenstände der durch ihn begriffenen Klasse konkreter oder abstrakter Entitäten. Das Ideal des Begriffs ist eine nachweisliche Wahrheit und Angemessenheit, das der Bedeutung ihre allgemeine Anerkennung durch den Gebrauch. (1991, S. 12)

Diese Auffassung vom Unterschied zwischen allgemeinsprachlichen und fachsprachlichen Wörtern findet sich auch bei Putnam (1990). Für ihn liegt die wahre Bedeutung von Termini in der Übereinstimmung mit der Realität, die unabhängig von unserer Wahrnehmung existiert. Putnam bezeichnet sich konsequenterweise in diesem Zusammenhang als 'metaphysischer Realist'. Dabei vergißt er offensichtlich, daß es neben der Wirklichkeit der natürlichen Dinge auch eine soziale Wirklichkeit gibt, die sich der arbiträren Setzung durch eine Gesellschaft (oder ihre Repräsentanten) verdankt, für die daher Prädikate wie 'Wahrheit' oder 'Angemessenheit' unangemessen sind.⁴

Burkhardts Kennzeichnung von Bedeutung nimmt Bezug auf „die alltägliche Praxis innerhalb der Lebensform“ mit der die Frage von Moral untrennbar verbunden ist. Doch auch Hermanns überaus einleuchtende Darstellung der 'Deontik' der Wörter eignet sich nicht für die Unterscheidung von Fachsprache und Allgemeinsprache.⁵ Das *Lohnabstandsgebot*, ein Fachwort, muß wie jedes *Gebot* (ein Wort der Allgemeinsprache) eingehalten werden; es ist keineswegs neutral in Hinblick auf die alltägliche Praxis. Das gilt ebenso für *Retroviren* und für *Plutonium*, die man wegen ihrer Gefährlichkeit in der Alltagspraxis meiden sollte.

Sinclair macht in einem wichtigen Beitrag einen neuen Ansatz, Wörter der Allgemeinsprache von Wörtern der Fachsprache zu unterscheiden (1996b). Auch er unterscheidet Bedeutung und Begriff. Kriterium für ihn ist jedoch nicht der Bezug zur Alltagspraxis, sondern die Opazität der allgemeinsprachlichen Wörter, die der Sprecher im jeweiligen Kontext konkretisiert, im Unterschied zur konzeptuellen Setzung bei Fachwörtern, über die nur der Terminologe, nicht aber der normale Sprachteilnehmer Definitionshoheit besitzt.

⁴ Vgl. dazu beispielsweise Searle (1995). Allerdings ist auffällig, daß Searle hier zu guter Letzt doch wieder für einen „textual realism“ plädiert, offensichtlich, um im Gleichklang mit der vorherrschenden Tendenz der amerikanischen *philosophy of mind* zu bleiben. Dieser „externe Realismus“ ist denn auch weitgehend sinnentleert und entkräftet in keiner Weise die ausschließlich gesellschaftliche Begründung der „institutional facts“. Übrigens ist beschämend, daß John Searle in keiner Weise auf das (ihm sicher zu relativistische) Standardwerk von Peter L. Berger und Thomas Luckmann *The Social Construction of Reality* von 1966 Bezug nimmt, dessen Titel er in seinem variiert.

⁵ Vgl. dazu vor allem Hermanns (1989).

Demzufolge würden neue allgemeinsprachliche Wörter (bzw. die neuen Bedeutungen bereits vorhandener allgemeinsprachlicher Wörter) sowie die nicht-genormten Bedeutungsbestandteile von Termini, soweit sie sich gewandelt haben, Gegenstand der Neologismenlexikographie sein. Angesichts des Auseinanderklaffens zwischen Theorie und Praxis der Terminologiearbeit (vgl. Felber/Budin 1989) läßt sich letztendlich auch eine solche Unterscheidung kaum operationalisieren.

Eine klare Trennungslinie zwischen 'normalen' Wörtern und Fachwörtern kann es folglich nicht geben. Auch unter Fachwörtern gibt es eben viele, die neben ihrer fachspezifischen Definition auch noch eine weitergehende, auf Konvention beruhende, die Wirklichkeit interpretierende Bedeutung haben. So meint das neue Verb *faxen* nicht nur die Übermittlung von Signalen auf Papierseiten von einem sendenden zu einem empfangenden Gerät auf elektronischem Wege, sondern es drückt sich darin auch der Lebensstil einer Generation aus, deren Lebensziel weithin in einem sofortigen permanenten Informationsaustausch auf möglichst hohem technischen Niveau zu liegen scheint.

3. Korpus und Lexikographie

3.1 Voraussetzungen

Seit etwa dreißig Jahren gibt es in der Linguistik maschinenlesbare Korpora, Sammlungen von Texten oder Textsegmenten, die größere Sprachausschnitte bis hin zur Gesamtsprache abbilden und damit repräsentieren sollten. Das erste Korpus dieser Art war das *Brown Corpus*, dem folgende Parameter zugrundeliegen (Sinclair 1994):

- (a) Größe: 1 Million Wörter
- (b) Genres: ungefähr gleiche Aufteilung
- (c) Zahl der Texte: 500
- (d) Textgröße: je 2000 Wörter
- (e) Grundlage: geschriebene gedruckte Quellen

In Deutschland wurde nach ähnlichen Prinzipien das *LIMAS-Korpus* aufgebaut. Eine Million Wörter wurde seinerzeit für ausreichend gehalten, um Grundwortschatz und erweiterten Grundwortschatz (bis zu 5.000 Lemmata) darzustellen, um die Grammatik, vor allem die Syntax, exhaustiv zu beschreiben, und um, darauf aufbauend, sogar ein System für automatische Übersetzung zu

entwickeln. Es muß allerdings klargestellt werden, daß man bei allen Erwartungen, die man sich damals gemacht hat, schon seinerzeit lieber ein weit größeres Korpus zusammengestellt hätte. Man war sich durchaus bewußt, daß die Korpusgröße unbefriedigend war. Andererseits waren damals die Erfassungskosten weit höher als heute. Als Eingabemedium fungierten Lochkarten und Lochstreifen; Scanner gab es ebensowenig wie Setzbänder für die elektronische Druckaufbereitung. Speicherplatz war kostspielig; der Zugriff auf Magnetbänder, die damals allein als externe Speicher zur Verfügung standen, war langsam, zumal die Bänder nur sequentiell abgesucht werden konnten. Die Zentraleinheit war im Vergleich zu heutigen Möglichkeiten extrem klein und langsam. Mit einem Korpus wie dem *Brown Corpus* (oder dem *LIMAS-Korpus*) stieß man an die Grenzen des technisch Machbaren.

Gewandelt hat sich indessen in den vergangenen dreißig Jahren auch die Vorstellung, man könne eine Sprache auf einen Kern, den man Grunddeutsch, Basic English oder Français fondamental nannte, reduzieren, ohne die kommunikativen Möglichkeiten drastisch einzuschränken. In den siebziger Jahren vor allem wurden von Instituten und Verlagen sogenannte Grundwortschätze erarbeitet, die zwar alle mehr oder weniger vorgaben, nach nachvollziehbaren Kriterien zusammengestellt zu sein, die aber alle nicht von sich behaupten konnten, ihnen habe ein wirklich repräsentatives Korpus zugrundegelegen. Denn auch in einem Korpus von einer Million Wörtern sind es nicht mehr als gut tausend Wörter, die mindestens zehnmals, und weit weniger als 10.000 Wörter, die wenigstens fünfmal belegt sind. Alles aber, was darunter liegt, liegt im Bereich des Zufälligen, den man nur begrenzen kann, wenn man neben dem Parameter Häufigkeit auch noch einen Parameter Distribution oder Verteilung der Belege über eine Mindestzahl von Texten ansetzt. Dazu kommt noch, daß der Computer damals allenfalls Wortformen, nicht aber Lemmata nach Häufigkeit sortieren konnte, und daß es damals (wie heute) keine befriedigende Homographendisambiguierung gab. Zum einen also reichten die seinerzeit verfügbaren Korpora nicht aus, um einen Grundwortschatz der 2.000 häufigsten Lemmata zu ermitteln; zum andern hat sich auch durch die Beschäftigung mit dem Problem Grundwortschatz die Erwartung erschüttert, es ließe sich wirklich irgendwo sinnvoll eine Grenze zwischen einem Grundwortschatz und dem Vokabular der Alltagssprache ziehen (Kühn 1979). Für das Deutsche war bis in die frühen achtziger Jahre die verlässlichste Häufigkeitsliste die von Wilhelm Käding, die um die Jahrhundertwende zur Entwicklung eines Kurzschriftsystems auf der Grundlage einer Textsammlung von ca. 11 Millionen Wörtern erhoben wurde.⁶ Der Gebrauchswert dieser Liste war und ist unbefriedigend. Unter den zweitausend häufigsten Wörtern fehlt etwa da der *Samstag*, weil er damals noch mit dem *Sonnabend* konkurrierte, und als

⁶ Vgl. dazu Meier, Helmut (1978): Deutsche Sprachstatistik. 1. Bd. 2. Aufl. Hildesheim (Olms), besonders S. 7-16.

Lebensmittel findet sich dort nur das Wort *Kartoffel*, denn es wurden überwiegend norddeutsche Texte ausgewertet.

Um den ermittelten Häufigkeiten den Ruch des Zufälligen zu nehmen, galt Repräsentativität lange als die Zauberformel. Die Anwendung statistischer Verfahren sollte sicherstellen, daß das Korpus tatsächlich für den beanspruchten Sprachausschnitt repräsentativ ist. Viele Fragen mußten gleichwohl offenbleiben. Mochte man sich auch darauf verständigen, daß der Ausschnitt beispielsweise die gedruckte, veröffentlichte, geschriebene Sprache eines Jahres sein sollte, blieb beispielsweise ungeklärt, ob für die Bemessung der auf die einzelnen Genres entfallenden Teilkorpora die Maße des gedruckten Wortes oder das Rezeptionsverhalten der Leser wichtiger ist. Keinen Konsens gab es natürlich (und gibt es noch heute nicht), was ein Genre ist, wie man Textsorten sinnvollerweise definiert und manch anderes mehr.⁷

In den späteren achtziger Jahren wurde in Deutschland das Ziel, ein repräsentatives Korpus aufzubauen, nicht weiter verfolgt. Die verfügbaren Korpora schienen auszureichen, um die wenigen grammatischen und syntaktischen Phänomene, die angesichts kaum funktionierender Parser überhaupt maschinell recherchierbar waren, zu untersuchen. Das Ziel eines großen enzyklopädischen Wörterbuchs des Deutschen, das neben der Standardsprache auch Fach- und Sondersprachen berücksichtigen sollte und als völlig korpusbasiert konzipiert war, wurde mangels hinreichenden Interesses bei Geldgebern und in der Öffentlichkeit allmählich aufgegeben. übrigens ging man in der Planung davon aus, daß ein solches allumfassendes Wörterbuch mit einem Korpus von 50 Millionen Wörtern auskommen würde.⁸

Statt dessen benutzte man, wenn überhaupt, eigens zusammengestellte Korpora vorwiegend für ausgewählte lexikographische Spezialunternehmen. Das einzige größere deutsche Wörterbuchunternehmen, das dem Anspruch erheben darf, wirklich korpusbasiert zu sein, war das Projekt Maschinelles Korpuswörterbuch (MKWB), das unter der Leitung von Manfred W. Hellmann zunächst in Bonn, dann in Mannheim durchgeführt wurde. Dabei vergingen von der Planungsphase bis zur Fertigstellung etwa 20 Jahre.⁹ Es ging darum, Differenzen im Sprachgebrauch zwischen Zeitungstexten in der Bundesrepublik und in der DDR zu ermitteln und zu beschreiben. Zugrundegelegt wurde ein Spezialkorpus aus statistisch verteilten Textsegmenten aus der *Welt* und dem *Neuen Deutschland* der Jahrgänge 1949, 54, 59, 64, 69 und 74. Dieses Korpus hat einen Umfang von etwa 3,5 Millionen Wörtern. Ausgewählt wurden fast ausschließlich nach statistischen Verfahren etwa 600 Lemmata, deren

⁷ Vgl. dazu auch Teubert (1982).

⁸ So heißt es in These 18 der „20 Bad Homburger Thesen“: „Das Textcorpus für das Wörterbuch soll nicht weniger als 50 Millionen Textwörter enthalten.“ (Henne/Mentrup/Möhn/Weinrich 1988, S. 283).

⁹ Das fertige Produkt liegt inzwischen als Hellmann (1992) vor.

Beleghäufigkeit in West- und Osttexten am unterschiedlichsten war. Objektivität in der Lemmaauswahl hatte höchste Priorität, wobei selbstverständlich nur lexikalische Ausdrücke ohne jede Homographendisambiguierung berücksichtigt werden konnten. So findet sich im fertigen Produkt eine Reihe von Lemmata, die man vermutlich vergeblich suchen würde, wenn die Auswahl intellektuell (und dann allerdings auch subjektiv) durch professionelle Lexikographen vorgenommen worden wäre¹⁰.

Das MKWB-Projekt hat zum einen deutlich gemacht, daß für die meisten linguistischen, vor allem lexikographischen Unternehmen 'repräsentative' Korpora von der Art des *Brown Corpus* keine ausreichende Basis bilden. In der Regel benötigt jedes dieser Projekte ein Spezialkorpus, das genau auf das Untersuchungsziel zugeschnitten ist. Bei der Zusammenstellung eines solchen Korpus kann man vielleicht auf Texte oder Textteile bereits vorhandener Korpora zugreifen. Spezialkorpora selber können auch wieder Teile anderer Korpora werden. Zweitens hat das MKWB-Projekt gezeigt, daß auch Spezialkorpora viel größer sein müssen, wenn es um lexikographische Untersuchungen geht, als man dies noch in den siebziger Jahren glaubte. Mindestens fünf Belege sind erforderlich, um eine Gebrauchsvariante als usuell zu belegen. Soweit es um Vergleiche geht, müssen diese fünf Belege für jede Vergleichsgröße vorhanden sein. Daß etwa ein Unterschied einerseits zwischen *Neuem Deutschland 54* und *Welt 54*, andererseits zwischen *Neuem Deutschland 54* und *Neuem Deutschland 59* vorliegt, kann ich nur nachweisen, wenn das in Frage stehende Phänomen wenigstens für eine dieser Vergleichsgrößen fünf-fach belegt ist. Teilt man jedoch das MKWB-Korpus entsprechend in die zwölf Teilkorpora auf, hat jedes dieser Teilkorpora nur einen Umfang von etwa 300.000 Wörtern. Welche Lexeme in einem derart kleinen Korpus zu finden sind, ist höchst zufällig und kann von der Normalverteilung signifikant abweichen. Verteilt man dann noch die belegten Lexeme auf jeweils mehrere Gebrauchsvarianten oder Bedeutungen, bleibt an Substanz nichts Nennenswertes übrig. Drittens zeigt das MKWB-Projekt, daß eine maschinelle Korpuserschließung den Lexikographen zwar entlasten, ihn aber keineswegs ersetzen kann. In seiner Arbeit über Grundwortschatz hat Kühn (1979) gezeigt, welche Vollständigkeitsprinzipien der Lexikograph bei der korpusbasierten Lemmaauswahl zu beachten hat, damit eine gewisse Abrundung entsteht.

In anderen Ländern, die eine gefestigtere Tradition der Zusammenarbeit zwischen Lexikographen in der akademischen Welt und Wörterbuchverlagen haben, wurde die Utopie eines 'repräsentativen' Korpus als Grundlage für Wörterbücher der Standardsprache weitergeführt. Anstelle von *repräsentativ* sprach man nun allerdings von *ausgewogen* (*balanced*), was inhaltlich indessen nichts Neues brachte. Eine ausführliche, wenngleich nicht mehr ganz ak-

¹⁰ So sucht man vergeblich die Stichwörter *Ideologie*, *Klassenkampf* oder *Marktwirtschaft*, während Wörter wie *sofortig*, *ihr* oder *Pfandbrief* diesen Status haben.

tuelle, Übersicht über in Europa verwendete Korpora, die in irgendeiner Hinsicht einen Anspruch auf Ausgewogenheit machen, findet sich in *Towards a Network of European Reference Corpora* (1996) und in weiteren Forschungsberichten, die im Zusammenhang mit dem NERC-Projekt erstellt wurden. Bekanntestes und wichtigstes Ergebnis korpusgestützter Lexikographie der Standardsprache sind die COBUILD-Wörterbücher (z.B. COBUILD 1987). Diese Wörterbücher verzeichnen nur Wörter, die im seinerzeit zugrundeliegenden Korpus von ca. 50 Millionen belegt waren, und in der Regel nur in den Bedeutungen, in denen sie belegt waren. Das große COBUILD *English Language Dictionary* hat immerhin den Umfang des DUDEN *Universalwörterbuchs*. Das Korpus von 50 Millionen Wörtern wurde rasch als viel zu klein erkannt. Die Allgegenwart maschinenlesbarer Texte (etwa auf Setzbändern) und von leistungsfähigen Scannern, sowie die Entwicklung neuer Zugriffsverfahren und der Preisverfall für Massenspeicher führten zu einem Ausbau des COBUILD-Korpus auf derzeit über 300 Millionen Wörter (einschließlich gesprochener Sprache). Bei solchen Größenordnungen verlagert sich die Frage nach der Ausgewogenheit allmählich vom Korpusersteller zum Korpusbenutzer. Er muß sich aus dem Gesamtangebot von Korpustexten die für seine Untersuchung geeigneten Texte zu einem Spezialkorpus zusammenstellen.

Abgeschlossen ist inzwischen auch der Aufbau des *British National Corpus* im Umfang von 100 Millionen Wörtern. Dieses Korpus ist ein gemeinsames Projekt verschiedener britischer Universitäten und Wörterbuchverlage. Es ist noch mit dem traditionellen Konzept der Ausgewogenheit zusammengestellt. Inzwischen (April 1996) ist ein von der Kommission der Europäischen Gemeinschaft finanziertes Verbundprojekt LE-PAROLE mit Partnerinstituten in praktisch allen EU-Staaten angelaufen, das die Erstellung gleichartig zusammengesetzter und einheitlich kodierter sogenannter Referenz-Korpora für die Amtssprachen der EU (zuzüglich Katalanisch) zum Ziel hat. Indessen haben diese Korpora, wenigstens in der ersten Stufe, nur einen Umfang von 20 Millionen Wörtern; aus heutiger Sicht liegen sie daher an der untersten Grenze der Brauchbarkeit. Wichtiger als Ausgewogenheit dürfte auf jeden Fall die Frage der Vergleichbarkeit dieser Korpora untereinander sein. Wie im nächsten Abschnitt dargelegt wird, tritt in der Zwischenzeit an die Stelle des 'ausgewogenen' Korpus immer mehr das Konzept des 'Referenzkorpus'.

Aus dem Gesagten wären für das große Neologismenprojekt am Institut für deutsche Sprache (IDS), das 1996 angelaufen ist, konkrete Konsequenzen zu ziehen. Um den Kernwortschatz der Standardsprache im Bereich von 20.000 bis 25.000 Lemmata in den wesentlichen Gebrauchsweisen bzw. Bedeutungen zu dokumentieren, ist ein Korpus von ca. 50 bis 100 Millionen Wörtern erforderlich, und zwar pro Zeiteinheit, die miteinander verglichen werden soll. Wenn also der Sprachwandel in jährlichen Schritten untersucht werden soll, muß für jedes Jahr ein solches Korpus angelegt werden. Diese Korpora müssen so gleichartig wie möglich aufgebaut werden. Andererseits müssen sie

jedoch auch dem Wandel im Kommunikationsverhalten (z.B. Fernsehen heute statt früher Radio) und dem Wandel in inhaltlichen Vorlieben (heute mehr Tennis als Handball früher) Rechnung tragen. Im nächsten Abschnitt ist zu diskutieren, ob es möglich und sinnvoll wäre, die Korpusgröße ohne Qualitätsverlust zu reduzieren, wobei die Vergleichbarkeit demnach gewährleistet bliebe. Zunächst jedoch geht es darum, den gegenwärtigen Stand der Reflexion über Korpora darzustellen.

3.2 Zum Stand der Korpusdiskussion in Europa

In der oben erwähnten, von der Europäischen Gemeinschaft finanzierten Projektstudie *Towards a Network of European Reference Corpora* (1996), an der Institute in sechs Ländern teilgenommen haben, wurde auf der Grundlage des Entwicklungsstands in den einzelnen Ländern ein konkreter Vorschlag für den Aufbau nationaler Referenzkorpora erarbeitet, der in der ersten Stufe Korpora von 60 Millionen Wörtern pro Sprache vorsieht. Der Gesichtspunkt der 'Ausgewogenheit' wurde dabei insoweit berücksichtigt, als die Praxis der vergangenen Jahre bezogen auf die einzelnen Sprachen Gemeinsamkeiten in dieser Hinsicht aufwies. Im Grunde beruht dieser Plan noch auf dem gängigen Korpusverständnis. Was hier Referenzkorpus genannt wird, bildet eine Einheit von Sprachdaten, also von empirischem Material, die ausreichen soll, um eine Vielzahl von Aufgaben zu erfüllen: Empirische Begründung von Grammatiken und Wörterbüchern der Standardsprache, statistische Aussagen, Entwicklung von Lexikonkonzepten für NLP-Anwendungen und monolingualer Input für multilinguale lexikalische Datenbanken. Es ist also im Grunde die alte Konzeption des *Brown Corpus*, nur in einer den neuen technischen Möglichkeiten angepaßten Größenordnung. Dieser Plan wird jetzt in einem ersten Schritt in dem oben genannten Verbundvorhaben LE-PAROLE verwirklicht. Zur Zeit wird daran gearbeitet, in dieses Projekt auch Partner aus mittel- und osteuropäischen Ländern (einschließlich der neuen GUS-Staaten) einzubeziehen. Erfreulicherweise gibt es in den letzten Jahren auch im Nachdenken über Korpora einige Fortschritte. Das von der EG finanzierte Vorhaben EAGLES (Expert Advisory Group on Language Engineering Standards), eher eine konzertierte Aktion von Instituten und Experten aus vielen europäischen Ländern als ein Projekt im eigentlichen Sinn, hat auch ein *Subcommittee on Corpora* eingerichtet, das neue Konzepte und Empfehlungen im Hinblick auf allgemeine Wünschbarkeit, Realisierungsmöglichkeit und vor allem Standardisierungsnotwendigkeit entwickelt. Diesen Standardisierungsaktivitäten ist im neuen EG-Forschungsprogramm besondere Aufmerksamkeit gewidmet. Es ist damit zu rechnen, daß es künftig öffentliche Drittmittelförderung von Korpusprojekten, ganz gleich, ob im nationalen oder im europäischen Rahmen, nur geben

wird, wenn auf dem erreichten europäischen Konsens, wie er etwa in EAGLES-Papieren formuliert wird, hingewiesen wird.¹¹

John Sinclair hat im Rahmen von EAGLES ein Diskussionspapier mit dem Titel *Corpus Typology* vorgelegt (Sinclair 1996a). Eine Übersetzung findet sich in diesem Band. Das Neologismenkorpus, für das wir im nächsten Abschnitt eine Konzeption entwickeln, wird in manchen Punkten von Sinclairs Vorschlägen abweichen.

3.3 Vorschläge für eine neue Korpuspolitik in Deutschland

Anders als in England, Frankreich, Dänemark, den Niederlanden und weiteren Ländern wurde in Deutschland bisher nicht der Versuch gemacht, ein synchronisches oder historisches Wörterbuch der Standardsprache auf der Grundlage eines maschinenlesbaren Korpus zu erstellen. Auch für das *Grimm-sche Wörterbuch* (in seiner zweiten Auflage) hat es lange keine Anstrengungen gegeben, eine maschinenlesbare Sammlung der Quellen aufzubauen.

In den Ländern, in denen korpusgestützte Standardlexikographie betrieben wird, handelt es sich in der Regel bei den jeweiligen Projekten um *joint ventures*, d.h. um Kooperation zwischen akademischer Lexikologie/Lexikographie und Wörterbuchverlagen. In Westdeutschland ist es nie zu einer vergleichbaren Tradition der Zusammenarbeit gekommen. Das ist sicher der wichtigste Grund dafür, daß es auch im heutigen Deutschland keine korpusbasierte Standardlexikographie gibt. Für die meisten Projekte in anderen Ländern gab es dieselbe Arbeitsteilung: Die Erstellung des Korpus, die Entwicklung der Zugriffsprogramme und die inhaltliche Leitung der Vorhaben liegen in akademischer Hand; sie werden überwiegend oder gänzlich mit öffentlichen Mitteln gefördert. Die Erstellung der Wortartikel sowie die wirtschaftliche Planung und Leitung liegen in Verlagshand. Die Korpora, die in ihrer Zusammensetzung nach einem Ideal von Repräsentativität oder Ausgewogenheit streben, stehen (mit gewissen Einschränkungen) auch anderen Interessenten aus Forschung und Industrie zur Verfügung.

Bisher zeigten sich die deutschen Verlage an solchen *joint ventures* nicht besonders interessiert, und sie blieben selbst gegenüber finanziellen Verlockungen der Öffentlichen Hände standhaft. Daß die Sparte Informationstechnologie in Europa einen stark expandierenden Markt weniger für traditionelle Wörterbücher, aber mit sprunghaft steigender Tendenz für monolinguale und besonders multilinguale lexikalische und terminologische Datenbanken erwarten läßt, wurde von den Verlagen in vielen europäischen Ländern frühzeitig

¹¹ Zum gegenwärtigen Stand von EAGLES-Empfehlungen kann man sich informieren bei: <http://www.ilc.pi.cnr.it/EAGLES/home.html>

begriffen, im lexikographischen Oligopol Deutschlands aber weitgehend verdrängt und wenn überhaupt, so vor allem von einschlägig interessierten Großunternehmen wie Siemens, Mercedes-Benz oder IBM wahrgenommen. Auch gibt es hier im Vergleich zu anderen Ländern nur eine sehr junge und vielfach als unbefriedigend empfundene Tradition der Zusammenarbeit zwischen öffentlich-rechtlicher und industrieller Forschung und Entwicklung.

Dies ist sicher ein Grund dafür, daß Deutschland, was Sprachtechnologie angeht, in europäischen Projekten eine recht unbefriedigende Rolle spielt. Die meisten akademischen Institutionen verfügen nicht über die erforderlichen Ressourcen (Korpora und maschinenlesbare Wörterbücher), haben keine festen Partner auf der Seite der Verlage, haben entweder starkes Mißtrauen gegenüber jeder Zusammenarbeit mit der Industrie und fürchten um den Verlust ihrer wissenschaftlichen Integrität, während einige andere Institutionen um so intensiver nach Kontakten zur Industrie suchen, um damit einen möglichen Mangel an Substanz in ihrer akademischen Verortung zu überspielen. Beides macht sie zu Außenseitern in der europäischen Szene.

Das Institut für deutsche Sprache (IDS) war bis vor wenigen Jahren die zentrale Institution im deutschsprachigen Raum, die über nennenswerte Korpus-Ressourcen einschließlich der Software für den Zugriff verfügt. Zum Aufbau, zur Pflege und zur Erweiterung seiner Korpus-Ressourcen unterhält es seit den siebziger Jahren eine Arbeitsstelle für linguistische Datenverarbeitung. Lange hat es diese Ressourcen in erster Linie als Dienstleistung für die hausinternen Forschungsvorhaben betrachtet, sie aber auch im Service für externe akademische und industrielle Interessenten angeboten.

Wenn das IDS ein wichtiges Zentrum für deutsche Sprachressourcen bleiben will, muß es zur Drehscheibe eines Netzwerkes von Anbietern und Nutzern von Sprachressourcen werden, Dienstleistungen auch gegenüber der Sprachindustrie (Wörterbuchverlage, Softwarehäuser, Übersetzungsbüros usw.) erbringen und darf sich bei der Weiterentwicklung von Ressourcen (Korpora, lexikalische Datenbanken und Auswertungssoftware) nicht nur von der eigenen Forschungsplanung leiten lassen, sondern muß auch externe Bedürfnisse berücksichtigen. Das kostet Kapazität, die zunächst durch zusätzliche Finanzmittel nur unzureichend kompensiert wird, so daß der Aufwand zu Lasten der IDS-eigenen Forschung ginge.

Auch die damit zusammenhängende Rolle des IDS, zentraler deutscher Ansprechpartner für die europäische Zusammenarbeit im Bereich Sprachressourcen zu werden, kostet zunächst nur Kapazität und bringt erst langfristig einen Ertrag. Denn noch ist das IDS-eigene Forschungsprogramm weitgehend auf monolinguale Projekte fixiert. An vielen vergleichbaren Partnerinstituten in anderen europäischen Ländern hat seit einiger Zeit der Anteil multilingualer Vorhaben deutlich zugenommen. Ein zunehmend integriertes Europa, das sich

– anders als andere Weltregionen – seine sprachliche Vielfalt bewahren will, ist indessen auf die Entwicklung von multilingualen Hilfsmitteln für das Schreiben von Texten in Fremdsprachen, für die Übersetzung und Informationserschließung angewiesen. Für deren Qualität sind wiederum das aus abgestimmten, vergleichbaren Ressourcen derivierte Sprachwissen maßgebend.

4. Zu einigen Korpusarten

4.1 Anmerkungen zum Diskurskorpus

In einer Übersicht über Korpusarten nennt Sinclair zwar das Spezialkorpus (in diesem Band, S. 115 ff.), aber er erwähnt dort nicht einen interessanten Sonderfall des Spezialkorpus, nämlich das Diskurskorpus.

Wie bereits oben ausgeführt, gibt es historische Gründe, warum sich im deutschsprachigen Raum nicht die auf Ausgewogenheit zielende Korpusideologie der anderen europäischen Länder entwickelt hat. Da kein großes historisches oder synchronisches korpusbasiertes Wörterbuch der Standardsprache auf der Tagesordnung stand, gab es seither keine konkreten Pläne mehr für die Erstellung eines ausgewogenen Korpus.

Spezialkorpora waren statt dessen am IDS das Gebot der Stunde. Für eine Reihe von Projekten sind Spezialkorpora bis zur Entwicklung echter Referenzkorpora weithin unverzichtbar. Beispiele sind Untersuchungen über die Unterschiede in ostdeutscher gegenüber westdeutscher Zeitungssprache (mit zusätzlichem diachronischen Anspruch), über den Sprachwandel während der Wendezeit, 'brisante' Wörter der Bereiche Politik, Umwelt und Feuilleton, Beratungssprache, Aussiedlersprache und manches mehr. Spezialkorpora sind und bleiben auch künftig für viele Anwendungen längerfristig unverzichtbar.

In der korpuslinguistischen Literatur ist ein Sonderfall des Spezialkorpus bisher nicht beachtet worden: das Diskurskorpus. Meines Wissens gibt es erst einen Versuch, ein Diskurskorpus für linguistische Untersuchungen aufzubauen: das „Wendekorpus“ des IDS, das Texte aus der BRD und der DDR aus der Zeit von Mitte 1989 bis Ende 1990 enthält, die sich auf die Wende in der DDR und die im Anschluß daran vollzogene Wiedervereinigung beziehen. Dieses Korpus hat einen Gesamtumfang von ca. 3.3 Millionen Wörtern, davon entfallen ungefähr zwei Drittel auf die BRD und ein Drittel auf die DDR (Herberg/Steffens/Tellenbach 1996).

Dieses Wendekorpus ist bereits während des Beobachtungszeitraums konzipiert worden. Zielsetzung war, Sprachwandel erstmals kontrolliert zu dokumentieren, um so Aufschluß darüber zu gewinnen, welche Faktoren relevant

sind und welche Erklärungsmuster Plausibilität beanspruchen können. Das Ende dieses Sprachwandelprozesses, bezogen auf die DDR, stand schon Ende 1989 fest: die fast totale Angleichung an westlichen Sprachgebrauch. Dagegen werden Änderungen im Sprachgebrauch der BRD oder jetzt Gesamtdeutschlands erst mit geraumer Verspätung greifbar. Die Probleme, die der Sprachwandel in der DDR während der Wende und besonders das Wendekorpus in Hinblick auf mögliche Sprachwandeltheorien bietet, habe ich in einem früheren Aufsatz dargestellt (Teubert 1993). Der ursprüngliche Plan, das Wendekorpus als Diskurskorpus zu konzipieren, konnte aus mancherlei Gründen nur für das DDR-Teilkorpus und auch da nur in engen Grenzen realisiert werden.

Das entscheidende Kriterium für ein Diskurskorpus ist, daß durch die Festlegung bestimmter Parameter durch den Linguisten zunächst ein Diskurs konstituiert wird. Zu den Parametern, die hier wichtig sind, gehören: Thema oder inhaltliche Charakteristika, Zeitraum, Areal, Textsorte, Veröffentlichungsform, sowie von grundlegender Bedeutung: intertextuale Bezüge. Es ist der letzte Punkt, der Diskurskorpora von allen anderen Korpora unterscheidet. Die Texte eines Diskurskorpus müssen zueinander in einer intertextuellen Beziehung stehen. Diese Bezüge sind primär semantisch; sie können im Extremfall die Form eines expliziten Zitats haben. Häufiger sind explizite Verweise, und noch häufiger ist der implizite Verweis, ob er nun vom Hörer mitvollzogen werden soll oder nicht. Typisch für implizite Verweise, die vom Leser mitvollzogen werden sollten, sind Parodien. Dagegen geht es dem politischen Redner, wenn er sich, ohne zu zitieren, auf Programme und Beschlüsse seiner Partei bezieht, nicht darum, daß die Verweise vom Adressaten erkannt und verstanden werden – ihm geht es in erster Linie um Absicherung gegenüber seiner Partei.

Ein solcher Diskurs ist zunächst einmal ein virtuelles Korpus. In aller Regel fehlen die Mittel, fehlt aber auch die Notwendigkeit, alle verfügbaren einschlägigen Texte in ein konkretes Korpus aufzunehmen. Das Auswahlkriterium ist simpel genug: Je mehr auf einen vorausgegangenen Text von späteren Texten Bezug genommen wird, desto wichtiger ist er.

Diskurskorpora sind die konkreten Korpora, die die ausgewählten Texte eines Diskurses enthalten, wie er durch die oben genannten Parameter definiert worden ist. Diskurskorpora sind die empirische Grundlage für das junge linguistische Spezialgebiet *Historische Semantik*. Unter zwei Aspekten, nämlich unter einem theoretischen Aspekt und unter einem thematischen Aspekt, ist Historische Semantik Teil der linguistischen Beschäftigung mit Sprachwandel.

In der Betrachtung von Sprachwandel finden sich beide in der Menschheitsgeschichte verwurzelten Weisen der Deutung von Prozessen. Die eine Sichtweise ist es, einen Prozeß als naturgesetzlich ablaufenden Vorgang zu deuten, der

im Prinzip erklärt werden kann und dann erklärt ist, wenn man die zugrundeliegenden Naturgesetze aufgefunden hat. In der anderen Sichtweise wird ein Prozeß als eine dem intentionalen Bewußtsein des Menschen entspringende zielgerichtete Handlung gedeutet. In der neueren Linguistik gab es stets beide Strömungen. Einerseits wurde nach quasi ehernen Gesetzen geforscht, nach den allen Sprachen zugrundeliegenden Universalien, die Sprachwandel, wo immer er vorkommt, steuern, andererseits wurde Sprache als von einer Sprachgemeinschaft vereinbarte Konvention verstanden, die durch Sprachhandeln geändert werden kann. Beide Strömungen finden sich auch in der Sprachwandelforschung. Dabei interessieren sich die eher naturwissenschaftlichem Denken verpflichteten Linguisten vor allem für Phonetik, Morphologie und Syntax, während die geisteswissenschaftlich orientierten Linguisten sich in erster Linie für lexikalischen und damit semantischen Wandel interessieren. Dies ist Gegenstand der Historischen Semantik (Busse/Teubert 1994). Es ist auch die zentrale Frage für die Beschäftigung mit Neologie.

Diskurskorpora dienen der Untersuchung von lexikalischem und semantischem Wandel. Diskurskorpora enthalten die Texte, in denen die semantischen Kämpfe ausgefochten werden. Wer wissen will, wie sich das Wortfeld 'Schwangerschaftsabbruch' seit dem Zweiten Weltkrieg verändert hat, benötigt ein Diskurskorpus, in dem um die 'richtige' Bedeutung und Verwendung von Wörtern wie *Abbruch*, *Unterbrechung*, *Fötus*, *Embryo*, *Leibesfrucht*, *Indikation*, *Tötung*, *Mord*, *ungeborenes Leben*, *ungeborenes Kind* usw. gerungen wird.

Sind Diskurskorpora überhaupt legitimer Gegenstand linguistischer Analyse? Eine Frage, die im Zusammenhang mit Projekten zum Wendekorpus immer wieder gestellt wird, ist die, ob es hier überhaupt einen Sprachwandel auf Systemebene oder eben nicht nur eine Veränderung im Sprachgebrauch gegeben hat. Dazu habe ich oben schon im Abschnitt „Was sind Neologismen?“ Stellung bezogen. Nachzutragen bleibt, daß ein solchermaßen reduzierter Systembegriff das, was Bedeutung ist, auf die kurze Angabe reduziert, die unter einem Lemma in einem mittleren standardsprachigen Wörterbuch zu finden ist. Für diese Bedeutungsbeschreibungen mag sich in der Tat durch die Wiedervereinigung oder durch die Diskussion um den Schwangerschaftsabbruch nichts geändert haben oder ändern (bezogen auf die westdeutschen Wörterbücher). Aber in den Wörterbüchern finden sich allenfalls Abkürzungen der Bedeutung, nicht aber die Beschreibung des in einer Sprachgemeinschaft verfügbaren und präsenten Wissens um konkrete Verwendungen, d.h. Belege eines Wortes mit ihren jeweiligen semantischen, deontischen und emotiven Implikationen. Die Bedeutung eines Wortes ist die Interpretation der Summe aller im diskurssemantischen Sinn relevanten Belege für dieses Wort, die von der Sprachgemeinschaft kollektiv erinnert werden bzw. archiviert sind. Dazu gehört das gewichtete Ensemble der Kontexte, in denen das Wort auftritt. Darum geht es bei der semantischen Analyse des Wendekorpus. Die Bedeutungs-

beschreibung des Adjektivs *marktwirtschaftlich* in standardsprachigen Wörterbüchern mag auf lange Zeit unverändert bleiben. Aber seit der Wende (und seit dem Kollaps des RGW-Systems) bedeutet das Wort auch, daß es zur Marktwirtschaft keine erfolgreiche Alternative gibt, daß Marktwirtschaft zudem bessere Lebensbedingungen garantiert und daß sie gegenüber Zwangs- und Kommandowirtschaft für sich das Ideal der Freiheit beansprucht.

Diskurskorpora sind also die Grundlage für Untersuchungen über lexikalisch-semanticen Wandel und damit auch für die Beschäftigung mit Neologie. Sie gestatten es, das Sprachhandeln einzelner auf der Ebene eines funktional zu interpretierenden kollektiven Prozesses zu beschreiben, der seine Spuren im Sprachsystem hinterläßt. Bestimmendes Merkmal von Diskurskorpora ist das Vorhandensein intentionaler intertextueller Bezüge zwischen den Texten eines Korpus anstelle von einer nach dem Zufallsprinzip gesteuerten Auswahl für alle anderen Korpustypen.

4.2 Anmerkungen zum Referenzkorpus

Hohe Priorität bei der Korpusarbeit hat, so die übereinstimmende Auffassung der Autoren der bereits genannten NERC-Studie, Schaffung eines Referenzkorpus. Was aber die Natur eines solchen Referenzkorpus angeht, gibt es, besonders in der Theorie, kontroverse Vorstellungen. Bei aller Kritik an bestehenden „repräsentativen“ oder „ausgewogenen“ Korpora kann sich auch John Sinclair in seiner in diesen Band enthaltenen Studie über Korpustypologie von dem Konzept der Ausgewogenheit letztlich nicht lösen. Er glaubt, einen Überschneidungsbereich aller Varietäten in einer Zentralzone, einem Sprachkern erkennen zu können, und darin seien für alle Varietäten „die Masse des Wortschatzes und fast alle Syntaxregeln“ gemeinsam. Doch Vorsicht ist geboten. Natürlich gibt es diesen Kernbereich, was die Syntax der geschriebenen Sprache angeht, besonders, wenn man sich auf diejenigen Syntaxregeln beschränkt, die satzintern sind. Aber daneben gibt es auch eine satzübergreifende Textsyntax, etwa pronominale Verweisungen, aber auch bisher weit weniger untersuchte Phänomene, die die innere Einheit eines Textes als lineare Kette einzelner Sätze bestimmen. Dies ist bei narrativen Texten häufig die Tempusstruktur in Verbindung mit Temporalangaben und Temporalpartikeln, wie sie etwa zur Markierung von Rückblenden und Vorverweisen dienen. Bei argumentativen Texten gibt es andere Phänomene, die Texteinheit herstellen, und für viele auf Textsorten gegründete Varietäten wäre der Überschneidungsbereich vermutlich auf die pronominale Verweisung beschränkt. Dazu kämen noch die besonderen Phänomene gesprochener Sprache, mit von der Situation bestimmten vielfältigen satzinternen Syntaxregeln bis hin zu satzübergreifenden Regeln des *turn-taking*. Auch hier dürfte die Beschränkung eines 'ausgewogenen' Korpus nur auf eine „Zentralzone“ nicht ausreichen.

Noch fraglicher ist Sinclairs Bestimmung der „Masse des Wortschatzes“. Es fragt sich, ob der Überschneidungsbereich aller Varietäten wirklich so definiert werden kann, daß eine sinnvolle und nachvollziehbare Lexikauswahl zustandekommt. Denn bestimmt man die Varietäten auch in Hinblick auf den Parameter 'inhaltliche Domänen', so bleibt für den Kernbereich allein der inhaltübergreifende Wortschatz übrig, nämlich solche Wörter, wie man sie sowohl in Kochbüchern als auch in den politischen Nachrichten, in Trivialromanen oder in Kunstführern findet.

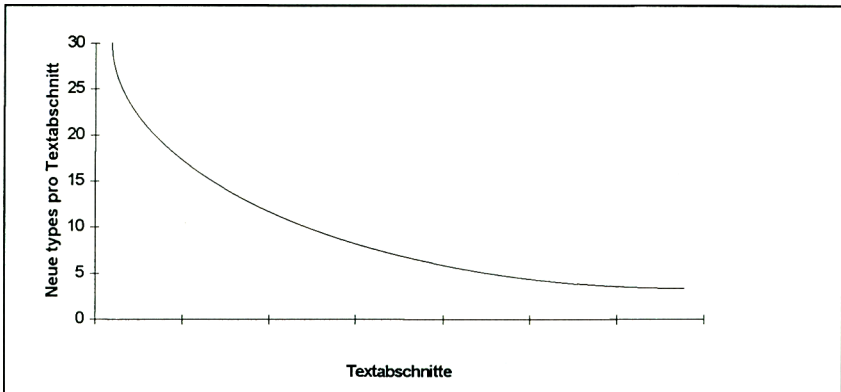
Referenzkorpora sind wichtig, und zwar aus drei Gründen. Einmal braucht man sie, um den Abstand von Spezialkorpora zur Standardsprache zu messen. So ist etwa zu erwarten, daß das Wort *Grenze* im oben genannten Wendekorpus prozentual erheblich häufiger vorkommt als in der Standardsprache, denn eines der bestimmenden Themen der Zeit war gerade die Überwindung der Grenzen. Zum anderen, und das ist vielleicht der wichtigste Gesichtspunkt, sind Referenzkorpora die unabdingbare Grundlage für Wörterbücher der Standardsprache, oder sie sollten es wenigstens sein. Schließlich sind Referenzkorpora als Vergleichskorpora oder vergleichbare Korpora, unabdingbare Voraussetzungen für die Entwicklung multilingualer sprachtechnologischer Anwendungen. Das bedeutet, daß der Aufbau von Referenzkorpora streng normiert sein muß.

Ausgewogenheit im Sinne von intellektueller Textauswahl entsprechend von vorgegebenen Parametern mit ihren Werten kann, wie wir oben gezeigt haben, nicht Prinzip eines Referenzkorpus sein. Was aber dann? Wenn man mit Referenzkorpora den Abstand von Spezialkorpora zur Standardsprache definieren will, welches Verständnis von Standardsprache ist damit vorausgesetzt?

Wir definieren Standardsprache im korpuslinguistischen Sinn als ein virtuelles Gebilde, das nur in bezug auf eine endliche Zahl von Parametern aktualisiert werden kann. Ein Korpus ist also nur standardsprachlich, insoweit es einen bestimmten Wert eines Parameters 'erfüllt'. In bezug auf alle sonstigen Werte und Parameter, insoweit sie nicht 'erfüllt' sind, bleibt das Korpus nicht-standardsprachlich. Ein Referenzkorpora liegt dann vor, wenn ein Korpus möglichst viele (im Prinzip alle) als relevant erachtete Werte von möglichst vielen als relevant erachteten Parametern 'erfüllt' oder realisiert. In bezug auf diese Werte und Parameter kann das Korpus dann als Referenzkorpus gelten. Referenzkorpora sind vergleichbar, wenn sie dieselben Werte derselben Parameter 'erfüllen'. Doch was heißt 'erfüllen'?

Hinter 'erfüllen' steht das von Herdan in der Frühzeit der Computertechnik entwickelte Konzept der *type-token*-Relation. Herdan war der erste, der die grundlegende Bedeutung der *type-token*-Relation für die Sprachstatistik erkannte. Er stellte fest, daß das numerische Verhältnis von *type* zu *token* von

der Textlänge abhängt.¹² Es ist leicht nachzuvollziehen, daß etwa, was die Wortformen angeht, im ersten Satz eines Textes fast so viele *types* wie *tokens* vorkommen. Danach nimmt, bezogen auf den Gesamttext, die Zahl der *types* ständig ab. Die 50 *types* der Funktionswörter machen etwa die Hälfte aller *tokens* aus. Am Anfang finden sich indessen noch recht viele neue *types* im Text. In einem Kochbuch etwa dürften beliebige 10 Prozent des Gesamttextes schon 95 Prozent aller *types* enthalten. Die restlichen 5 Prozent neuen *types*, die auf den restlichen 90 Prozent des Gesamttextes zu finden sind, sind Wörter für neue Zutaten oder neue Gerichte. Die meisten Kochverfahren und die wichtigsten Zutaten bleiben für das ganze Kochbuch dieselben. Die *type-token*-Relation kann als Kurve dargestellt werden, wobei die horizontale Achse die Zahl der Textabschnitte markiert, die vertikale Achse die Anzahl neuer *types*, die in jedem neuen Textabschnitt zu finden sind. Nach einer bestimmten Zahl von Textabschnitten tendiert die Kurve hin zu einer Geraden. Das bedeutet, daß dann mit jedem neuen Textabschnitt ungefähr die gleiche Zahl neuer *types* auftritt.



Dabei spielt im Prinzip keine Rolle, welche Phänomene als *types* und *tokens* betrachtet werden. Die Kurven bleiben ähnlich, ob man nun Wortformen, Wörter, Morpheme oder syntaktische Phänomene untersucht. Sind die Phänomene endlich, tendiert die Kurve irgendwann nach null; sind die Phänomene prinzipiell offen, wie das etwa für Substantive, Adjektive und Verben der Fall ist, wird sich die Kurve jedenfalls oberhalb von null auf einem bestimmten Niveau einpendeln, d.h. mehr oder weniger zur Geraden werden. Wie Herdan gezeigt hat, unterscheiden sich die Kurven je nach Textsorte und sind sogar spezifisch für unterschiedliche Autoren. Aber das interessiert hier nicht. Wo-

¹² Herdan (1966), besonders S. 322ff.

Nach der internationalen Norm (ISO/DIS-2, Terminology Work – Vocabulary – Part 2, 1996) ist *type* definiert als „linguistic unit in a text representing a defined class“; *token* ist definiert als „occurrence of a type“.

rum es geht, ist die Bedeutung der *type-token*-Relation für das Konzept des Referenzkorpus.

Denn nun kann man exakt definieren, wann ein Korpus, was die Erfüllung bestimmter Parameter angeht, als gesättigt anzusehen ist. Der Gedanke, in der *type-token*-Relation ein exaktes Instrument zur Bestimmung der Korpusgröße zu besitzen, wurde von Cyril Belica und Robert Neumann auf dem Kolloquium *Neologie und Korpus* im Herbst 1994 am Institut für deutsche Sprache vorgestellt; eine schriftliche Konkretisierung steht leider noch aus. Für die Korpuslinguistik ist das Prinzip der 'Sättigung' von großer Wichtigkeit, besonders wenn es darum geht, den Wortschatz eines 'Kernbereichs' der Sprache zu quantifizieren.

Ein Referenzkorpus ist also ein bezüglich der wesentlichen Parameter gesättigtes Korpus. Die Sättigung bezieht sich dabei zentral auf den Wortschatz. Es ist anzunehmen, daß ein nach Wortschatz gesättigtes Korpus auch im Hinblick auf Morphologie und Einzelsatzsyntax (wenngleich nicht in Hinblick auf textlinguistische Phänomene) gesättigt ist. Was aber sind die wesentlichen Parameter? Es sind natürlich wieder Textsorten, Texttypen, Domänen; im Grunde die ganze Palette der Kategorien, die in der jahrzehntelangen Diskussion um 'repräsentative' oder 'ausgewogene' Korpora eine Rolle gespielt haben.

Eine Aufgabe von Referenzkorpora ist es, Vergleiche mit Referenzkorpora anderer Sprachen zu ermöglichen. Ist das Ziel etwa der Aufbau einer multilingualen lexikalischen Datenbank, muß angestrebt werden, daß die Auswahl der Texte nach einheitlicher Parametrisierung erfolgt. Welche Parameter und welche jeweiligen Werte wesentlich für ein Referenzkorpus sind, ist letztlich eine Frage der Korpusanwendungen. Multilinguale lexikalische Datenbanken werden für Übersetzungen benötigt, und es läßt sich leicht ein Bild darüber herstellen, welche Texte, Texttypen, Textsorten, inhaltliche Domänen usw. bei der Übersetzungsarbeit in Europa besonders wichtig sind. Wie fruchtlos auch immer die korpuslinguistische Diskussion um Ausgewogenheit war und weiter sein wird, das NERC-Projekt hat gezeigt, daß es möglich ist, sich auf europäischer Ebene nach den genannten objektiven Kriterien auf Umfang und Zusammensetzung von Referenzkorpora zu verständigen. Gleichwohl zeigt sich im LE-PAROLE-Projekt,¹³ daß die Wirklichkeit der Korpuserstellung leider hinter dem bereits erreichten theoretischen Konsens über die Komposition von Referenzkorpora zurückbleibt. Maßgeblich dafür sind in erster Linie finanzielle Gründe. Je mehr Parameter gewählt werden und je mehr Werte pro Parameter unterschieden werden, desto größer muß das Referenzkorpus werden. Die Grenzen des Machbaren sind da schnell erreicht.

¹³ Weitere Angaben zu diesem Projekt findet man unter:

<http://www.ids-mannheim.de/le-parole/parole-d.html>

Die andere Aufgabe von Referenzkorpora ist es, durch den Vergleich mit Spezialkorpora die Abweichung, vor allem natürlich bezogen auf die Lexik, dieser Korpora von der Standardsprache zu messen. Das Referenzkorpus muß also auch ein Bild der Standardsprache sein. Was Standardsprache ist, läßt sich, wie oben gezeigt, nicht objektiv festlegen. Es handelt sich um einen bewußten Akt der Normung, die allerdings die Zustimmung derer finden muß, die mit dieser Normung arbeiten müssen. Betroffen werden von einer solchen Festlegung der Standardsprache alle potentiellen Benutzer des Korpus. Ist die Zusammensetzung erst einmal beschlossen, ergibt sich die Größenordnung automatisch nach dem Prinzip der Sättigung.

Ein Referenzkorpus, das nach den hier genannten Grundsätzen aufgebaut ist, enthält selbstverständlich nicht den Gesamtwortschatz einer Sprache, nicht einmal den der Varietät 'Standardsprache'. Die Tatsache, daß ein Korpus lexikalisch 'gesättigt' ist, bedeutet ja nicht, daß beim Hinzufügen neuer Texte keine neuen *types* mehr auftreten würden, sondern nur, daß von nun an die Zuwachsrate konstant bleibt. Alle dazukommenden Wörter würden aber natürlich ebenfalls zur Standardsprache gehören. Deshalb reicht ein Referenzkorpus nicht, um ein rein korpusbasiertes Wörterbuch der Standardsprache zu belegen.

Für ein solches Wörterbuch gibt es ein Modell, nämlich das *COBUILD English Language Dictionary* (s.o. S. 143). Es enthält nur Wörter, die im zugrundeliegenden Korpus von ca. 50 Millionen enthalten waren. So wurde auch in den Rezensionen das Fehlen „wichtiger“ Wörter moniert, und das nicht ohne Recht. Von einem Wörterbuch der Standardsprache wird mehr erwartet, als daß es ein Korpus abbildet. Nach Tradition und Gewohnheit stellt man an ein solches Wörterbuch den Anspruch einer bestimmten Vollständigkeit. So sollte es etwa alle Blumen verzeichnen, mit denen man im Sprachgebiet üblicherweise zu tun hat: neben *Rosen* auch *Petunien* und *Levkojen*, und sollten aus irgendwelchen dummen Zufällen im Korpus keine *Primeln* enthalten sein, sind sie selbstverständlich nachzutragen. Wie bereits erwähnt, hat Kühn die Vollständigkeitsansprüche an Wörterbücher gründlich analysiert. Sie sind berechtigt und setzen dem Dogma der Korpusbasiertheit, d.h. der wirklichen Belegtheit von Wörtern in Texten, Grenzen. Doch steht andererseits außer Frage, daß Korpusbasiertheit ein hoher Wert ist. Besonders deutschen standardsprachlichen Wörterbüchern ist anzusehen, daß sie viel zu sehr auf tradierten Wortbeschreibungen, die sich über ganze Wörterbuchgenerationen fortschreiben, als auf eigener Beleganalyse beruhen. Viele englische und französische Wörterbücher sind in dieser Beziehung den deutschen überlegen. Bis zu welcher Grenze also sollte ein Wörterbuch sinnvollerweise korpusbasiert sein? Was muß also umgekehrt auf jeden Fall in einem Referenzkorpus ausreichend belegt sein?

Im Abschnitt „Was sind Neologismen“ haben wir eine in diesem Zusammenhang sehr wichtige Unterscheidung getroffen, nämlich die zwischen normalen Wörtern und Fachwörtern. Wenigstens in bezug auf Substantive gilt, daß sich die Vollständigkeitsgebote (soweit es sich nicht um prinzipiell endliche Mengen wie beispielsweise die Wochentage handelt) ganz überwiegend auf Namen für natürliche Arten (wie etwa Blumen) oder auf Namen für Artefakte, die im Alltagsleben vorkommen und von jedermann gelegentlich benannt werden müssen, beziehen. Vollständigkeit hat hier mit übergeordneten Kategorien oder zusammenfassenden Konzepten zu tun. Beispiele sind ‘Werkzeug’ oder ‘Badezimmer’. Vollständig müssen die im Alltagsleben vorkommenden Bestandteile eines Badezimmers verzeichnet sein: *Badewanne, Dusche, Waschbecken, Wasserhahn, Mischbatterie, Kachel, Duschvorhang, Stöpsel, Abfluß, Klo, Klodeckel, (Wasser-)Spülung* usw., d.h. die Gegenstände, die in einem Bildwörterbuch auf dem entsprechenden Bild benannt werden. Diese Wörter gleichen Fachwörtern, insofern die Begriffsbildung auf reale konkrete Gegenstände zugreifen kann, es also exemplarische Denotate gibt. Sie sind insofern keine Fachwörter, als sie nicht normiert sind. Die im Alltagsleben vorkommenden Bestandteile erhalten ihre Bedeutung weniger durch ‘essentielle’ Eigenschaften als vielmehr durch Funktionszuschreibung. Ein Eimer Wasser kann als Dusche fungieren, eine Colaflasche als Blumenvase. Daneben kennt aber jeder Installateur sehr viel mehr Bestandteile eines Badezimmers (beispielsweise bestimmte Einheiten, aus denen Wasserleitungen zusammengesetzt werden). Für diese Bestandteile gibt es genormte Bezeichnungen, und nur der Experte kann entscheiden, ob diese Namen richtig auf die Gegenstände angewendet werden. Diese Wörter gehören nicht zur Standardsprache.

Soweit solche Artefakte oder natürliche Arten (Blumen) im Alltagsleben von jedermann benannt werden müssen, gehören sie selbstverständlich in ein Wörterbuch der Standardsprache. Aber man wird sie nicht alle in einem Referenzkorpus finden. In ein Wörterbuch gehören beispielsweise die Namen der Blumenarten, die man typischer Weise in Gärten findet, die man für den Garten oder als Schnittblumen kaufen kann oder die in der Natur vorkommen und auffällig genug sind, daß sie auch ein Laie kennt. Fehlen sie im Korpus, müssen sie vom Lexikographen aus anderen Quellen ergänzt werden. Wörter dieser Art gehören dann ins Wörterbuch, wenn sie Denotate bezeichnen, die jeder im Alltagsleben benennen muß und in bezug auf die auch von jedem Benennungskompetenz erwartet wird: also jedenfalls *Tiger* und *Löwen*, aber vielleicht keine *Makaken*.

Auf jeden Fall belegt sein müssen (und zwar in der notwendigen Belegdichte) im Referenzkorpus die ‘normalen’ Wörter. Das sind, wie wir oben gesagt haben, die Wörter, über deren richtige Anwendung die Sprachgemeinschaft, vertreten durch den einzelnen Sprachteilhaber, entscheidet. Hier gibt es keine Experten, die mehr Kompetenz für sich beanspruchen können. Doch lassen sich diese Wörter weitergehend charakterisieren? Es handelt sich hier um die

Wörter, mit denen wir unsere Wirklichkeit konstruieren, als Sprach- und Kulturgemeinschaft, und eben bis hin zu dem Ausmaß, als diese gesellschaftlich konstituierte Wirklichkeit in der Tat noch allgemein und nicht auf einen Kreis von Fachleuten reduziert ist. Wo die Anschaulichkeit am größten ist, ist der konstruktivistische Gestaltungsraum am geringsten. Das gilt für Denotate von realen und leicht zu unterscheidenden Gegenständen, etwa *Tiger* und *Löwen*. Bei Artefakten ist der Gestaltungsraum um so größer, je weniger festgelegt das Erscheinungsbild und je wichtiger demgegenüber und je allgemeiner die Funktion ist. Eine Glühlampenfassung eröffnet weniger konstruktivistische Spielräume als ein Sessel. Gesellschaftlich-kulturell determiniert sind besonders Substantive für mentale Konstrukte, wie *Arbeit*, *Strafe*, *Frieden*, *Vertrauen*, desgleichen Verben wie *sich entwickeln*, *angreifen*, *festlegen*, *singen*, Adjektive wie *groß*, *gut*, *selbstverständlich*, *tüchtig*, *geheimnisvoll*. Wörter wie diese, die dazu dienen, die Wirklichkeit zu konstituieren, können in ihrer Bedeutung nur durch Sprache, nicht durch Verweis auf Denotate, beschrieben werden. Ihre Bedeutung ist nichts anderes als die Summe ihrer bisherigen Verwendung in Texten, die gewichtende Interpretation der Belege. Die meisten Bezeichnungen von Artefakten und natürlichen Arten lassen sich, beispielsweise über ein Bildwörterbuch, auch anders als durch den Sprachgebrauch beschreiben. Für 'normale' Wörter ist der Sprachgebrauch konstitutiv.

Um wieder auf das Referenzkorpus zurückzukommen, so muß es ausreichen, den Wortschatz der 'normalen' Wörter so umfassend zu belegen, daß sich auf dieser Grundlage verlässliche Bedeutungsbeschreibungen erstellen lassen. Ein Korpus, das in bezug auf die im Alltag verwendete Standardsprache 'gesättigt' ist, wird dazu in der Lage sein. Den übrigen Wortschatz, Namen für anschauliche Dinge und Fachwörter, auch wenn sie zur Standardsprache gehören, braucht ein Referenzkorpus nur in zufälliger Auswahl enthalten. Er läßt sich mühelos aus anderen Quellen vervollständigen und kann durch den Verweis auf die Denotate problemlos beschrieben werden.

Der Aufbau eines Referenzkorpus muß das Ziel jedes nationalen Zentrums für Korpus-Ressourcen sein. In der heutigen Zeit sollten Referenzkorpora die Grundlage von Wörterbüchern der Standardsprache sein. Die Erstellung von Spezialkorpora läßt sich dagegen nicht monopolisieren, um so mehr, als zunehmend Korpustechnologie jedermann zugänglich ist. Ein Referenzkorpus setzt in bezug auf seine Parametrisierung Absprachen voraus, sowohl innerhalb des nationalsprachigen Raums, was die mögliche Nutzung des Korpus betrifft, als auch innerhalb Europas, insofern es um multilinguale Anwendungen geht. Ein Referenzkorpus setzt zugleich auch Maßstäbe. Wörterbücher, Grammatiken und andere Ergebnisse linguistischen Arbeitens müssen sich, soweit sie empirische Aussagen enthalten, den Vergleich mit den Daten des Referenzkorpus gefallen lassen.

4.3 Anmerkungen zum opportunistischen Korpus

Bis es ein Referenzkorpus gibt, das diese Bezeichnung wirklich verdient, muß eine Zwischenlösung verfolgt werden. Dies ist der Aufbau eines opportunistischen Korpus. Viele europäische Zentren für Korpus-Ressourcen konzentrieren ihre Aktivitäten in diesen Jahren auf die Erstellung opportunistischer Korpora. Sie tun dies zum einen, weil die andauernde Diskussion um Ausgewogenheit zu keiner wirklich überzeugenden Lösung und zu keiner neuen Zauberformel geführt hat. Wichtiger ist indessen, daß sich seit den siebziger Jahren die Lage, was die Verfügbarkeit maschinenlesbarer Texte angeht, von Grund auf geändert hat. Anstelle von früher üblicher Eingabe über Tastatur ist es heute oft billiger, Texte mit Scannern einzulesen und die Korrektur des oft fehlerhaften Einleseprozesses mit Rechtschreibprogrammen zu unterstützen, wenn nicht die Texte von vornherein in maschinenlesbarer Form zugänglich sind. Quellen dafür sind im wesentlichen Satzbänder für elektronische Setzverfahren, vor allem für Zeitschriften und Zeitungen, weniger für Bücher, weil für die unzähligen in Einsatz befindlichen Setzverfahren jeweils spezifische Reinigungsprogramme zur Elimination von druckspezifischen Sonderzeichen entwickelt werden müssen, die den Aufwand für ein einzelnes Buch nicht lohnen. Außer Satzbändern gibt es aber zunehmend auch elektronisch gespeicherte Textbanken, etwa von Nachrichtenagenturmeldungen, im juristischen Bereich von Urteilsbegründungen, von Einrichtungen wie Internet und World Wide Web, die ständig erweitert werden und jährlich sicher einen Umfang von weit über einer Milliarde Wörtern haben. Dazu kommen seit einigen Jahren Texte, die auf CD-ROM angeboten werden, beispielsweise der Inhalt ganzer Zeitungsjahrgänge, aber auch große Wörterbücher und Texte anderer Provenienzen. Interessant für opportunistische Korpora sind in erster Linie periodisch zur Verfügung stehende Texte auf Datenträgern.

Ein opportunistisches Korpus ist ein Korpus, das qualitative Ansprüche auf den Umweg über Quantität erfüllt, indem es auf solche maschinenlesbar vorhandenen Texte zugreift, die mit einem Minimum an Aufwand bezogen auf die Textlänge zu integrieren sind. Es speist sich also aus solchen Texten, die billig oder kostenlos angeboten werden und bei denen die Erfassung und Anpassung an die Kodierungskonventionen für die maschineninterne Repräsentation weitestgehend automatisiert werden kann. Ähnlich wie das Gesamtkorpus stellt das opportunistische Korpus in erster Linie ein Angebot an sprachwissenschaftliche Forschungsvorhaben dar, sich aus dem Gesamtrepertoire der verfügbaren Texte entsprechend den Projektvorgaben geeignete Texte zu einem Spezialkorpus zusammenzustellen. Sinnvoll und überhaupt erst möglich geworden ist das Konzept eines opportunistischen Korpus durch die Entwicklung der Speichertechnik, aufgrund derer Speicherplatz heute so billig ist, daß er finanziell kaum noch relevant ist.

Ein opportunistisches Korpus, das komplette Jahrgänge mehrerer Tageszeitungen und dazu noch mehrere Jahrgänge von Periodika wie *Die Zeit*, dem *Spiegel* oder *Stern* enthält, ist im Vergleich mit jetzt vorhandenen Korpora sicher von gigantischer Größe (mehrere hundert Millionen Wörter), und es würde natürlich auch kontinuierlich erweitert werden können, monatlich sicher in der Größenordnung jetziger Korpusbestände, also etwa von 20 bis 50 Millionen Wörtern. Es wäre alles andere als 'ausgewogen', denn es bestünde vorwiegend aus Zeitungssprache. Andererseits ist ein Vorteil von Zeitungssprache ihre Vielseitigkeit: Berichte und Kommentare in zahlreichen inhaltlichen Domänen, Leserbriefe, Essays, Literatur. Bestimmte Textsorten würden gleichwohl fehlen oder stark unterrepräsentiert sein: administrative Texte, Literatur, Scripts von TV-Serien, Reiseführer, Gebrauchsanleitungen, Ratgeberliteratur.

Wer es sich nicht leisten kann, für seine Forschungsvorhaben erst einmal die passenden Spezialkorpora anzulegen, ist gewohnt, mit solchen Defiziten zu leben. Auch das Problem, mit Datenmengen konfrontiert zu werden, die sich keinesfalls mehr intellektuell verarbeiten lassen, kann man bewältigen, indem man mit statistischen Verfahren eine Zufallsauswahl erstellt oder, was mit verbesserter Korpussoftware zu erreichen ist, indem man das interessierende Phänomen intelligent eingrenzt. Für viele seltenere lexikalische Ausdrücke läßt sich überhaupt nur mit einem derart gigantischen Korpus eine befriedigende Belegmenge erhalten. Es bleibt also letztendlich nur ein Manko, das sich in der traditionell üblichen Weise nicht lösen läßt, nämlich das der Dokumentation. Zwar kann man argumentieren, daß es reicht, Zeitung, Datum und Seite zu identifizieren, und daß man nicht jeden einzelnen Artikel bibliographisch erfassen muß (und in der Tat sind einige Zeitungskorpora nur auf diesem Minimalniveau dokumentiert). Darüber hinaus lassen sich meist automatisch Überschrift, Sparte, in manchen Fällen auch Verfasser (oder sonstige Quellenangaben) automatisch identifizieren. Solange diese automatisch kumulierten Angaben nicht manuell kontrolliert sind, sind sie nur mit Vorsicht zu genießen. Doch jede manuelle Intervention würde den Kosten Vorteil opportunistischer Korpora wieder zunichte machen. Das bedeutet dann auch, daß die Kodierung der einzelnen Texte in Hinblick auf die Kompositionsparameter wie Textsorte, Texttype, Areal, usw. beschränkt wäre auf die formalisierten Angaben, die sich automatisch verarbeiten lassen.

Dokumentation ist dann unabdingbar, wenn für ein Forschungsvorhaben aus dem Repositorium des opportunistischen Korpus ein Spezialkorpus zusammengestellt werden soll, und dies dürfte einer der häufigeren Benutzerwünsche sein. Die Investition für die Entwicklung eines vollautomatischen Dokumentationssystems wäre sicher nicht gering, aber es wäre eine einmalige Anschaffung.

4.4 Anmerkungen zum Monitorkorpus

Die Tatsache, daß sich ein opportunistisches Korpus primär aus auf Dauer periodisch zur Verfügung stehendem maschinenlesbarem Textmaterial speist, bietet sich für einen gezielten Schritt in Hinblick auf die historische Tiefe eines Korpus an. Darauf gründet sich die Idee des Monitor-Korpus, die Sinclair schon Ende der achtziger Jahre erstmals dargelegt hat.

Ein Monitorkorpus ist ein Korpus, das in regelmäßigen Abständen unter möglichst genauer Beibehaltung der Kompositionsparameter ergänzt wird. Monitorkorpora sind somit die ideale Grundlage jeglicher Neologismenlexikographie. Solange Referenzkorpora, die ausreichend parametrisiert sind und in ihrem Umfang dem Kriterium der Sättigung entsprechen, aus finanziellen Gründen eine Utopie bleiben, bietet sich in der Zwischenzeit der Typ des opportunistischen Korpus als zeitliche Grundeinheit des Monitorkorpus an. Auch hier ist es wichtig, die Zusammensetzung langfristig möglichst unverändert zu lassen.

Beim heutigen Stand der Technik bieten sich für ein Monitorkorpus vor allem Tages- und Wochenzeitungen mit möglichst weitgestreutem Inhalt an. Das Neologieprojekt des IDS sieht den Aufbau eines solchen Monitorkorpus vor. Mit vielleicht sechs Tageszeitungen und zwei bis vier Wochenzeitungen dürfte es pro Jahr eine Größe von 100 Millionen Wörtern erreichen. Das sollte ausreichen, um den Kernwortschatz der Standardsprache im Bereich von ca. 100.000 zu belegen und bezogen darauf den lexikalischen Wandel zu dokumentieren.

Wie oben bereits ausgeführt, bietet die Extraktion von echten Neologismen, also neuen Zeichenketten zwischen Leerzeichen, die weder Eigennamen noch Fehlschreibungen sind, am wenigsten Probleme. Aber auch Bedeutungsveränderungen schon länger existierender Wörter können in Monitorkorpora mit ziemlicher Treffsicherheit erkannt werden. Dazu wird das Korpus in Phasen, beispielsweise in Jahresabschnitte zerlegt, und die absolute und relative Häufigkeit jedes *types*, in diesem Fall jeder lemmatisierten Wortform, ermittelt. Ändert sich nun von einer Phase zur nächsten die Häufigkeit in statistisch relevanter Weise, ist ein Kandidat für eine Bedeutungsänderung gefunden. Diese Lemmakandidaten können dann in einem zweiten Schritt in Hinblick auf ihr Kontextverhalten überprüft werden. Eine Analyse der *types*, die in den zehn Wörtern links oder rechts unseres Neologismenkandidaten zu finden sind, kann beispielsweise ergeben, daß in Phase B statistisch relevant *types* zu finden sind, die in Phase A nicht belegt waren, oder daß *types*, die in Phase A häufig zu finden waren, in Phase B auffällig selten geworden sind. Die Wahrscheinlichkeit, daß ein Bedeutungswandel vorliegt, ist groß genug, um in diesen Fällen das Belegmaterial dem Lexikographen vorzulegen. Ein Monitor-

Korpus kann ihm die ermüdende Routine der intellektuellen Neologismensuche weitgehend ersparen.

Spezialkorpus (einschließlich Diskurskorpus), Referenzkorpus, opportunistisches Korpus und schließlich Monitor-Korpus sind für eine längerfristige Korpuspolitik keine Alternativen, sondern komplementäre Ziele. Unterschiedliche Anwendungen erfordern unterschiedliche Korpora.

5. Die Alternative: Ein Spezialkorpus für Neologismen

Ein auf dem Prinzip des opportunistischen Korpus beruhendes Monitorkorpus ist nur eine Möglichkeit, lexikalischen Wandel zu dokumentieren. Alternativ dazu bietet sich als Ausgangspunkt ein Spezialkorpus an, das, selbstverständlich unter Beibehaltung der zugrundegelegten Kompositionsparameter, in regelmäßigen Zeitabständen ergänzt wird. Auch dieses Neologismenkorpus ist also ein Monitorkorpus. Es ist jedoch, anders als das opportunistische Korpus, nicht zufällig, sondern durch bewußte Auswahl zusammengesetzt.

Gegenüber dem opportunistischen Monitorkorpus hat ein spezielles Neologismenkorpus bestimmte Vorteile:

- Es kann erheblich kleiner sein bei gleicher Leistungsfähigkeit.
- Es kann den eigentlichen Bedeutungswandel (als ‘Kampf um Begriffe’) transparenter machen, vor allem, wenn es als *Diskurskorpus* (s. o.) konzipiert ist.
- Es kann, als Diskurskorpus, wichtige intertextuelle Bezüge für den Sprachwandel illustrieren.
- Es kann, wiederum als Diskurskorpus, einen Beitrag zur ‘Historischen Semantik’ leisten.

Das oben bereits erwähnte Wendekorpus des IDS zur Dokumentation des Sprachwandels von Mitte 1989 bis Ende 1990 ist in nuce ein Beispiel für ein solches Korpus. Wenigstens, was seine ostdeutschen Bestandteile angeht, wurde Wert darauf gelegt, solche Texte auszuwählen, die die wichtigen Diskussionszusammenhänge in der Zeit der Wende und der Vorbereitung der Vereinigung belegen. Was die für diesen Zeitraum kennzeichnenden Phänomene für den Wandel im Sprachgebrauch angeht, ist dieses intellektuell zusammengesetzte Korpus ebenso mächtig wie ein opportunistisches Monitorkorpus, das zehnmal größer wäre. Diese Leistungssteigerung hängt eng mit der Frage der Kompositionsparameter zusammen. Welche Kompositionsparameter für die Neologismenlexikographie besonders wichtig wären, soll im folgenden diskutiert werden.

Die Zusammensetzung eines jeden Korpus läßt sich auf die zugrunde gelegten Kompositionsparameter abbilden. Manche dieser Parameter, z.B. Textsorte, Areal, Zeitsegment, Gattung, sind innerhalb der Korpuslinguistik weithin akzeptiert (wobei dennoch eine Vielzahl von Definitionen etwa für *Textsorte* nebeneinander stehen können). Welche Kompositionsparameter relevant sind, ist für Referenzkorpora und Spezialkorpora unterschiedlich zu beantworten. Aber auch für den einzelnen Korpusstyp gehen die Ansichten weit auseinander. Manche, wie etwa die Verantwortlichen für das *British National Corpus*, meinen heute etwa, bei einem Referenzkorpus (das ja, wie wir uns erinnern, ausgewogen sein soll), müsse nicht nur das Alter, sondern auch das Geschlecht der Sprecher-Schreiber berücksichtigt werden, während andere die Forderung nach einem solchen Geschlechtsparameter für unvernünftig oder unrealistisch halten werden. Bei einem Spezialkorpus mit Diskurscharakter wie etwa dem 'Historikerstreit' oder auch dem 'Wendekorpus' werden manche die politische Heimat der Sprecher/Schreiber für ein wichtiges Kriterium halten, an dem Ausgewogenheit bemessen werden kann; andere werden das speziell für das Wendekorpus für unmöglich halten.

Es liegt auf der Hand, daß für ein bestimmtes Korpus die Zahl der Parameter in engen Grenzen bleiben muß. Zehn Parameter mit jeweils zehn Werten hätten schon zehn Milliarden Kategorien zur Folge und wären also gänzlich unpraktikabel. Auch fünf Parameter mit je fünf Werten würden schon zu 15.625 Kategorien führen, wären also ebenfalls nicht handhabbar, wenngleich (wenigstens für ein Referenzkorpus) durchaus sinnvoll. Beispielsweise könnten diese Parameter sein: fünf Zeitabschnitte, fünf Areale, fünf Themenkreise, fünf Adressatenkreise und fünf Veröffentlichungsmodi. Was ist die Konsequenz daraus?

Die Kosten für ein in dieser Weise ausgewogenes Referenzkorpus sind auf lange Zeit prohibitiv, besonders wenn jede mögliche Parameterkonfiguration bis zur Sättigungsgrenze ausgebaut sein soll. Allenfalls könnte es Korpora geben, die in bezug auf zwei oder maximal drei Parameter ausgewogen sein könnten. Dabei liegt auf der Hand, daß die Frage der Relevanz dieser Parameter nicht generell zu klären ist, sondern vom Untersuchungszweck abhängt.

Welche Parameter für die Ausgewogenheit eines Spezialkorpus, beispielsweise des Neologismenkorpus, zu berücksichtigen sind, ergibt sich aus dem Untersuchungszweck. In Bezug auf Neologismenlexikographie scheint mir, daß die folgenden Parameter wichtig sind (deren Relevanz ich weiter unten diskutiere):

- **Domänen.** Mit *Domäne* bezeichne ich einen Themenkreis. Jeder Themenkreis kann mehrere Themen enthalten. Dieser Parameter ist kaum objektivierbar und notwendigerweise ein Kompromiß unter-

schiedlicher und widersprüchlicher Betrachtungsweisen. Die Liste der Domänen und Einzelthemen ist grundsätzlich offen.

Als *Domänen* sehe ich beispielsweise an:

- Lifestyle
- Technik/Wissenschaft
- Sport als Unterhaltung
- Lebenshilfe
- Natürliches Volksempfinden
- Frau
- Kultur und Kunst
- Bildung und Geisteswissenschaft
- Wirtschaft
- Politik
- Reise und Exotik
- Geschichte/Biographien
- Umwelt
- Mobilität
- Gesundheit

Die Domäne 'Natürliches Volksempfinden' könnte etwa folgende Themen enthalten:

- Verbrechen/Verbrechensbekämpfung
 - Sexualität
 - das Eigene und das Fremde
 - Sonstiges
- **Textmodus.** Mit *Textmodus* bezeichnen wir die Ebene, auf der ein Text anzusiedeln ist. Wir unterscheiden vornehmlich OBJEKTTEXTE und METATEXTE. Zu den OBJEKTTEXTEN in Zeitungen gehören Nachrichten, Anzeigen und gegebenenfalls literarische Texte, etwa ein Fortsetzungsroman. METATEXTE sind Kommentare, allgemeine Betrachtungen, Besprechungen und Werbung. Zwischen OBJEKTTEXTEN und METATEXTEN setzen wir LESERBRIEFE an. LESERBRIEFE beschäftigen sich typischerweise mit OBJEKTTEXTEN.
 - **Anpassung.** Der Parameter *Anpassung* bezeichnet das Maß, in dem ein Text sich in den Rahmen der medial vermittelten „öffentlichen Meinung“ einfügt bzw. dazu eine Alternative darstellen will. Das Maß der Anpassung beruht auf eigener wie auf fremder Festlegung. Wir unterscheiden die Werte: ANGEPASST PLURALISTISCH, ANGEPASST IDEOLOGISCH (wobei je nach Weltanschauung zu unterscheiden

ist), LIBERAL PLURALISTISCH/LIBERAL IDEOLOGISCH, ALTERNATIV PLURALISTISCH, ALTERNATIV IDEOLOGISCH. Texte in der *Süddeutschen Zeitung* sind in der Regel ANGEPASST PLURALISTISCH, in der *Frankfurter Allgemeinen Zeitung* ANGEPASST IDEOLOGISCH (*konservativ*), Texte oft ALTERNATIV PLURALISTISCH, im *Neuen Deutschland* ALTERNATIV IDEOLOGISCH (*links*). Wichtig ist dieser Parameter deshalb, weil Neologismen mutmaßlich in alternativen Texten häufiger sind als im angepaßten.

- **Medium.** Mit *Medium* bezeichnen wir die Art und Weise, wie der Text vom Sprecher/Schreiber zum Adressaten kommt. Übliche Werte für diesen Parameter sind (beschränkt auf Texte geschriebener Sprache): PRIVATBRIEF, GESCHÄFTSBRIEF, NEWSLETTER, BUCH, BROSCHÜRE, ZEITSCHRIFT, ZEITUNG, ÜBER DEN RUNDFUNK/DAS FERNSEHEN VERMITTELTE TEXTE usw.
- **Adressierung.** Mit *Adressierung* bezeichnen wir den Parameter, dessen Werte unterschiedlichen Adressatenkreisen entsprechen. Dazu gehören beispielsweise: DER GEBILDETE BÜRGER, DER NORMALBÜRGER, DIE JUGEND, FRAUEN, DER HOBBYIST, DER ARBEITNEHMER usw.
- **Verbreitung.** Der Parameter *Verbreitung* gibt an, wieviel mutmaßliche Rezipienten ein Text, bezogen auf einen zu definierenden Zeitraum hat.
- **Intertextuelle Relevanz.** Dieser Parameter gibt an, inwieweit auf diesen Text in nachfolgenden Texten explizit oder implizit Bezug genommen wird. Dieser Parameter könnte auch *Kanonisierung* genannt werden.
- **Textsorte.** Dieser Parameter ist immer kontrovers diskutiert worden. In Absetzung von den anderen Parametern wäre es sinnvoll, unter *Textsorte* ein Bündel textinterner linguistischer Merkmale zu verstehen, dem ein gegebener Text entspricht.

Wie groß müßte ein nach solchen Parametern zusammengestelltes Neologismenkorpus sein? Ein Korpus von 5 Millionen Wörtern jährlich wäre jedenfalls zu klein, wenn auf seiner Grundlage Neologismenlexikographie durchgeführt werden soll. Dafür gibt es viele Gründe; wichtig sind sicher die folgenden:

- In einem Korpus von 5 Millionen Wörtern sind nur 8.000 Wörter mindestens zehnmals belegt. Dies entspricht allenfalls einem erweiterten Grundwortschatz. Zehn Belege sind wohl das Minimum für eine (möglichst ausschließlich) korpusbasierte Bedeutungs- bzw. Verwendungsanalyse. Wenn M.A.K. Hallidays Vermutung stimmt, daß

binäre Ambiguitäten entweder im Verhältnis 0,5:0,5 oder im Verhältnis 0,9:0,1 vorkommen, würden schon 10 Belege nicht ausreichen, um Zweitbedeutungen zuverlässig erkennen und beschreiben zu können (Halliday 1993, S. 9).

- Neologismen sind (wie oben gezeigt) entweder neue Ausdrücke zwischen Leerzeichen oder es sind alte Ausdrücke mit neuer Bedeutung oder in neuer Verwendung. Aber nicht alle neuen Ausdrücke sind Neologismen. Daneben gibt es nämlich auch beispielsweise Personen- und geographische Namen, Markenzeichen, *ad hoc* gebildete Nominalkomposita, überhaupt zunächst alle neuen Resultate produktiver Wortbildung. Als lexikalisiert gelten solche Wortbildungen erst, wenn ihre Bedeutung vom Prognostizierbaren abweicht, z.B. dadurch, daß mögliche Ambiguität verlorengeht, daß sich die Bedeutung einengt (oder seltener erweitert), daß es Verwendungsrestriktionen gibt usw. Auch bei neuen Wortbildungen ist zur Feststellung etwaiger Lexikalisierung ein Minimum von 10 Belegungen anzusehen.
- Neue Belegungen alter Ausdrücke können bei 10 Belegen nur im eher unwahrscheinlichen Fall erkannt werden, daß die Bedeutungen etwa hälftig verteilt sind. Für den Fall, daß die neue Bedeutung nur im Verhältnis 1:9 vorkommt (was wahrscheinlich eher typisch ist), müßte der Ausdruck also mindestens 50 mal belegt sein, um den Neologismus nachweisen zu können.
- Von Neologismen sind Okkasionalismen zu unterscheiden. Okkasionalismen sind solche Alphaketten, die entweder extrem selten oder aber gehäuft während eines sehr kurzen Zeitraums in einer besonderen Domäne, einem besonderen Texttyp, einer besonderen Textsorte oder gar nur in einem besonderen Text vorkommen, die also von der Normaldistribution über diese Parameter extrem abweichen. Ein typisches Beispiel dafür wäre das Lemma *Blechtrommel*, das im gleichnamigen Roman von Günter Grass extrem häufig belegt ist, aber trotzdem nicht zum Grundwortschatz der deutschen Sprache gehört.

Diese vier Punkte sprechen eindeutig gegen ein opportunistisches Korpus von nur fünf Millionen Wortformen jährlich. Ein Korpus dieser Größe wäre allenfalls als Grundlage für die Beschreibung des erweiterten Grundwortschatzes ausreichend. Wenn Neologismen erst einmal zu diesem Wortschatzsegment gehören, haben sie in aller Regel aufgehört, Neologismen zu sein.

Welche Auswege bieten sich an? Die eine Möglichkeit wäre, wie oben gezeigt, mit einem wirklich großen opportunistischen Korpus zu arbeiten, das regelmäßig als Monitorkorpus ergänzt wird. Dies ist wohl die kostengünstigste

Möglichkeit für eine korpusbasierte Neologismenlexikographie im klassischen Sinn.

5.1 Relevanzanreicherung

Es läßt sich aber auch die Aussagekraft eines Neologismenkorporus verbessern, wenn man es als Spezialkorpus konzipiert, wobei die Auswahl der einzelnen Korpus-texte an genaue Vorgaben gebunden ist, die zu einer Anreicherung des Korpus an Neologismen bis hin zu einer Sättigung führen sollen. Ein solches Korpus bildet die Gemeinsprache nicht mehr vollständig und nur noch verzerrt ab. Daß das ein Nachteil ist, hat sich an der Arbeit mit dem 'Wendekorpus' gezeigt. Es ist dort nötig, einzelne (beileibe nicht alle) Ergebnisse durch den Vergleich mit neutraleren Korpora zu validieren. Andererseits kommt dieses Korpus, besonders das ostdeutsche Teilkorpus, seinem Ziel recht nahe, ausreichende Beleggrundlage für bis zu 1000 „wenderelevante“ Wörter zu sein.

Lexikalischer Wandel ist immer auch Ausdruck von gesellschaftlichem Wandel. Gesellschaftlicher Wandel indessen findet nicht überall statt. Es gibt Bereiche, in denen kaum Veränderungen zu finden sind (etwa im Bereich des Philosophierens), und es gibt Bereiche, wo sich ein rasanter Wandel abspielt (etwa in der Informationstechnologie). Ein Roman wird nur selten unmittelbar gesellschaftlichen Wandel bewirken (Ausnahme: *Werther*); dagegen gelingt es manchmal Sachbüchern, das Denken zu beeinflussen (Beispiel: *Global 2000*). In systematischer Hinsicht geht es darum, die Kompositionsparameter in Relation zur erwartbaren Häufigkeit von Neologismen zu konzipieren. Beispielsweise wäre für das frühere Serbokroatisch heute Arealität von vorrangiger Bedeutung, und zwar gemäß den Regionen serbisch-kontrolliert, kroatisch-kontrolliert und bosnisch-moslemisch-kontrolliert. Für den deutschsprachigen Raum indessen besteht keine besondere Gefahr der Auseinanderentwicklung. Was die Frage der Relevanzanreicherung für das geplante Neologismenkorporus angeht, ist es also erforderlich, entsprechende Gewichtungen der Kompositionsparameter zu definieren und innerhalb der einzelnen Parameter die besonders neologismenrelevanten Werte zu identifizieren.

5.2 Neologismenrelevante Kompositionsparameter

Parameter, bei denen eine Relevanz in Hinblick auf die Anreicherung eines Korpus mit Neologismen zu erwarten ist, sind *Domäne*, *Textmodus*, *Anpassung* und *intertextuale Relevanz*. Das bedeutet, daß Texte auszuwählen sind, die bestimmten Werten dieser Parameter entsprechen. Ein solches Korpus ist dann selbstverständlich in Hinblick gerade auf diese Parameter nicht mehr ausgewogen oder repräsentativ. Wenn man etwa unter dem Parameter *Domä-*

ne nur drei Themenkreise mit jeweils zwei Einzelthemen berücksichtigt, werden alle anderen Domänen und alle anderen Einzelthemen notwendigerweise ausgeklammert. Inwieweit auch andere Parameter neologismenrelevant sind, bedürfte einer gründlichen Analyse. Vieles spricht etwa dafür, daß Texte, die sich an einen relativ kleinen, aber recht homogenen Adressatenkreis wenden, anfälliger sind für sondersprachliche Einflüsse oder Gruppenjargon und daß hier eine wesentliche Quelle für gemeinsprachliche Neologismen sprudelt. Zu den drei genannten Parametern hier einige Anmerkungen:

- **Domäne.** Grundsätzlich sind für jeden der oben genannten Themenkreise Neologismen zu erwarten. Die Häufigkeit hängt weniger vom Themenkreis als vom Einzelthema ab. Wie Einzelthemen und Themenkreise in verschiedenen Zeitungen und Wochenblättern verteilt sind, ist den Anlagen (2), (3) und (4) zu entnehmen. Neologismen sind, wie gesagt, eher da zu erwarten, wo gesellschaftlicher Wandel stattfindet. Bei *Lifestyle* also eher bei den Einzelthemen *Mode* und *Beziehungsfragen* als bei *Prominentenklatsch*, bei *Technik/Wissenschaft* eher beim Einzelthema *Gen-Technologie* als beim Thema *Atomindustrie*; bei der Domäne *Frau* eher beim Einzelthema *Karriere* als beim Thema *Frau und Familie* usw. Unter dem Gesichtspunkt der Priorität der Neologismenanreicherung muß die Ausklammerung zahlreicher Einzelthemen toleriert werden. Zum einen sind indessen auch nachweislich viele Einzelthemen in existierenden Korpora nicht vertreten, die nach Repräsentativität und Ausgewogenheit streben. Andererseits läßt sich die Abweichung eines Spezialkorpus (wie des Neologismenkorpus) vom zugrundeliegenden virtuellen Textuniversum durch den Abgleich mit dem Referenzkorpus (wie opportunistisch es immer sein mag, wenn es nur groß genug ist) ziemlich genau ermitteln.
- **Textmodus.** Dieser Parameter ist aus verschiedenen Gründen sehr wichtig. Texte wie Kommentare, essayistische Betrachtungen und vor allem Berichte über Texte, also Rezensionen, beschäftigen sich weniger mit Fakten und Ereignissen als mit Veränderungen, mit dem Neuen, das sich in Nachrichten, in Sachbüchern und literarischen Werken manifestiert. Eine Rezension etwa zur *Blechtrommel* wird sicher alle relevanten Neologismen, die in dem Roman vorkommen, enthalten und wird ansonsten die Ausgewogenheit eines Korpus bezogen auf die Häufigkeit der Lemmata weniger stören als der Roman selbst. Kommentare und Betrachtungen nehmen einerseits Einfluß auf die „öffentliche Meinung“ und tragen so zum gesellschaftlichen Wandel bei; andererseits aber greifen sie auch gesellschaftlichen Wandel auf und verleihen ihm eine sprachliche Form. Aus diesen Gründen muß der Wert *Metatext* eine sehr hohe Priorität haben.

- **Anpassung.** Gesellschaftlicher Wandel und somit auch Sprachwandel vollzieht sich primär immer in den Segmenten der Gesellschaft, die am wenigsten angepaßt sind, die also der Gesellschaft mit alternativen Konzepten gegenüberstehen. Lange Jahre wurden neue Sichtweisen etwa in der *taz* vorgedacht, bevor sie zunächst in *liberalen* und danach auch in *angepaßten* Texten auftraten. Heute finden sich alternative Sichtweisen eher im *Neuen Deutschland*. Wichtiger sind *pluralistische* als *ideologisch* festgelegte Texte. Zumal wenn Texte *orthodox-ideologisch* sind, ist ihre Neologismenrelevanz äußerst gering, wie jede Analyse des *Neuen Deutschland* vor der Wende un schwer zeigen kann.
- **Intertextuelle Relevanz.** Für die Neologismenlexikographie sind die Texte am interessantesten, auf die umfassend von nachfolgenden Texten Bezug genommen wird. Hier (beispielsweise in Ansprachen von Bundespräsidenten) finden sich eher 'erfolgreiche' Neologismen (im Gegensatz zu Okkasionalismen) als in Privatbriefen oder auch in Kommentaren kleinregionaler Tageszeitungen, die nur in Ausnahmefällen von wichtigen Kommentatoren zur Kenntnis genommen werden. Der Parameter *Intertextuale Relevanz* ist besonders wichtig, wenn die Genese des eigentlichen Bedeutungswandels dargestellt werden soll. Texte mit geringer intertextueller Relevanz haben dagegen für die Neologismenlexikographie nur geringen Wert.

5.3 Ein Plädoyer für Komplementarität

Ein spezielles Monitorkorpus, das in seiner Zusammensetzung stark Texte bevorzugt, die solchen Parametern und vor allem solchen Werten Rechnung tragen, die neologismenrelevant sind, dürfte bei einem Umfang von nur fünf Millionen Wörtern pro Jahr in Hinblick auf Neologismenlexikographie gleichwertig mit einem opportunistischen Monitorkorpus von 50 bis 100 Millionen Wörtern sein. Darüber hinaus wäre es auch für vertiefende onomasiologische Unternehmungen zu semantischem Wandel ideal geeignet. Die intellektuelle Zusammenstellung und großenteils manuelle Eingabe der Texte dürften jedoch um einiges aufwendiger sein als ein überwiegend automatisch kompiliertes opportunistisches Monitorkorpus. Dazu kommt, daß ein solches Spezial-Neologismenkorpus nur bedingt für andere lexikographische oder linguistische Auswertungen geeignet ist, da es die Standardsprache nur in verzerrierter Form abbildet.

Andererseits haben die am IDS durchgeführten Wendeprojekte, die mit dem Wendekorpus gearbeitet haben, gezeigt, welcher Wert in intellektuell zusammengestellten Spezialkorpora liegt. Das Wendekorpus selbst ist den Weg zum

Typ des Diskurskorpus nur in ersten Ansätzen gegangen. Schon das hat ge-
reicht zu zeigen, daß für Untersuchungen auf dem Gebiet Begriffsgeschichte
oder Historische Semantik kein opportunistisches Korpus mit einem Dis-
kurskorpus konkurrieren kann.

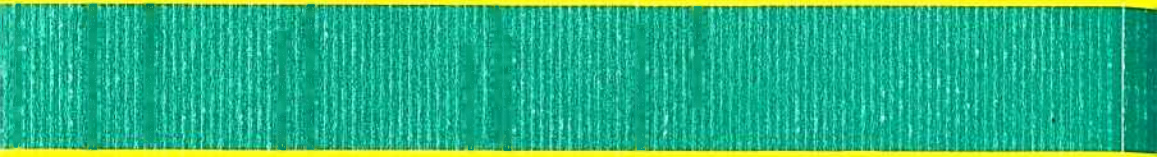
Neologismenlexikographie läßt sich einmal verstehen als Teil der Lexikogra-
phie der Standardsprache. In diesem Sinn ist ein opportunistisches Monitor-
koprus das geeignete Arbeitsinstrument. Neologismenlexikographie läßt sich
andererseits auch verstehen als systematische und umfassende Beschreibung
des semantischen Wandels, wie er sich im Wortschatz der Standardsprache
einer jeden Sprachgemeinschaft darstellt. In diesem Sinne ist die Beschäfti-
gung mit Neologismen ein Beitrag zur Beschreibung des kulturellen Wandels
einer Gesellschaft. Das opportunistische Monitorkorpus kann auf einem relativ
flachen Niveau die dazu notwendigen Daten liefern. Will man semantischen
Wandel jedoch nicht nur beschreiben, sondern auch erklären, ist komplementär
zu einem opportunistischen Monitorkorpus ein Spezialkorpus, von Hand
ausgewählt und zusammengestellt, unabdingbar.

6. Literatur

- Belica, Cyril: Statistische Analyse von Zeitungsstrukturen in Korpora. In: LDV-Info, 8, Institut für deutsche Sprache. Mannheim. S. 86-95.
- Burkhardt, Armin (Hg.) (1991): Vom Nutzen und Nachteil der Pragmatik für die diachrone Semantik. In: Busse, Dietrich (Hg.): Diachrone Semantik und Pragmatik. Untersuchungen zur Erklärung und Beschreibung des Sprachwandels. Tübingen. S. 7-36.
- Busse, Dietrich/Teubert, Wolfgang (1994): Ist Diskurs ein sprachwissenschaftliches Objekt? Überlegungen zu einer linguistischen Diskurssemantik. In: Busse, Dietrich/Herrmanns, Fritz/Teubert, Wolfgang (Hg.): Zeichengeschichte, Begriffsgeschichte, Diskursgeschichte. Opladen. S. 10-28
- Calzolari, Nicoletta/Baker, Mona/Kruijt, Johanna G. (Hg.) (1996): Towards a Network of European Reference Corpora. (= *Linguistica Computazionale XI*). Pisa.
- CED (1995) – Collins COBUILD English Dictionary. John Sinclair (Leitender Hg.). University of Birmingham.
- Cherubim, Dieter (1979): Zum Problem der Ursachen des Sprachwandels. In: Zeitschrift für Dialektologie und Linguistik, 46, S. 320-337.
- CORDON (1996): Proposal for TELEMATICS Application Programme: Corpus-Oriented Detection of Neologisms (CORDON). Proposal for Language Engineering Sector, January 1996.
- Coseriu, Eugenio (1970): System, Norm und 'Rede'. In: Ders. (Hg.): Sprache. Strukturen und Funktionen. XII Aufsätze. (= Tübinger Beiträge zur Linguistik 2). Tübingen. S. 193-212.

- DUW (1989) – DUDEN – Deutsches Universalwörterbuch A-Z. Hrsg. u. bearb. vom Wiss. Rat u. d. Dudenred. unter Leitung v. Günther Drodowski. 2., völlig neu bearb. u. stark erw. Aufl. (1. Aufl. 1983). Mannheim/Wien/Zürich.
- DWB (1854-1960) – Deutsches Wörterbuch. Bd. 1-16. Grimm, Jacob/Grimm, Wilhelm. Leipzig.
- Felber, Helmut/Budin, Gerhard (1989): Terminologie in Theorie und Praxis. (Forum für Fachsprachen-Forschung 9.) Tübingen.
- Halliday, M.A.K. (1993): Quantitative Studies and Probabilities in Grammar. In: Hoey, Michael (Hg.): Data, Description, Discourse. Papers on the English language in Honour of John McH Sinclair on his Sixtieth Birthday. London. S. 1–25.
- Hellmann, Manfred W. (1992): Wörter und Wortgebrauch in Ost und West. Ein rechnergestütztes Korpus-Wörterbuch zu Zeitungstexten aus den beiden deutschen Staaten. DIE WELT und NEUES DEUTSCHLAND 1949–1974. 3 Bde. (= Forschungsberichte des Instituts für deutsche Sprache Mannheim 69). Tübingen.
- Henne, Helmut/Mentrup, Wolfgang/Möhn, Dieter/Weinrich, Harald (Hg.) (1978): Interdisziplinäres Deutsches Wörterbuch in der Diskussion. Düsseldorf.
- Herberg, Dieter/Steffens, Doris/Tellenbach, Elke (1997): Schlüsselwörter der Wendezeit. Wörter-Buch zum öffentlichen Sprachgebrauch 1989/1990. Berlin.
- Herdan, Gustav (1966): The Advanced Theory of Language as Choice and Chance. (= Kommunikation und Kybernetik in Einzeldarstellungen 4). Berlin/Heidelberg/New York.
- Hermanns, Fritz (1989): Deontische Tautologien. Ein linguistischer Beitrag zur Interpretation des Godesberger Programms (1959) der Sozialdemokratischen Partei Deutschlands. In: Klein, Joseph (Hg.): Politische Semantik. Opladen. S. 69–149.
- Hermanns, Fritz (1994): Schlüssel-, Schlag- und Fahnenwörter: Zur Begrifflichkeit und Theorie der lexikalischen „politischen Semantik“. (= Arbeiten aus dem Sonderforschungsbereich 245, „Sprache und Situation“, Heidelberg/Mannheim 81). Heidelberg.
- Institut National de la Langue Française, Nancy (INaLF) (Hg.) (1971-1994): Trésor de la langue française. Dictionnaire de la langue du XIX^e et du XX^e siècle (1789-1960). 16 Bde. Centre National de la Recherche Scientifique. Nancy, Paris. [Imbs, Paul (leitender Hg.) (1971-): Trésor de la langue française. Dictionnaire de la langue du XIX^e et du XX^e siècle (1789-1960). Centre National de la Recherche Scientifique. Centre de Recherche pour un Trésor de la langue Française Nancy, Paris.]
- Keller, Rudi (1994): Sprachwandel. Von der unsichtbaren Hand in der Sprache. 2. überarb. Aufl. (1. Aufl. 1990). Tübingen.
- Kühn, Peter (1979): Der Grundwortschatz. Bestimmung und Systematisierung. Tübingen.
- Putnam, Hilary (1975a): The Meaning of „Meaning“. In: Gunderson, Keith (Hg.): Language Mind and Knowledge. Minneapolis. S. 131–193.
- Putnam, Hilary (1975b): Mathematics, Matter and Method. Philosophical Papers 1. Cambridge.
- Putnam, Hilary (1990): Die Bedeutung von „Bedeutung“. Frankfurt a.M.
- Scarle, John R. (1995): The Construction of Social Reality. New York.
- Sinclair, John M. (1996a): Corpus Typology. A Framework for Classification. EAGLES. URL: <http://www.ilc.pi.cnr.it/EAGLES96/corpusTyp/corpusTyp.html>

- Sinclair, John M. (1996b): The Empty Lexicon. In: *International Journal of Corpus Linguistics*, 1/1, S. 99–119.
- Teubert, Wolfgang (1982): Corpus and Lexicography. In: *Proceedings of the Second Scientific Meeting Computer Processing of Language Data*. Bled (Jugoslawien). S. 275–302.
- Teubert, Wolfgang (1988): Politische Vexierwörter. In: Klein, Joseph (Hg.): *Politische Semantik. Beiträge zur politischen Sprachverwendung*. Opladen. S. 51–68.
- Teubert, Wolfgang (1993): Sprachwandel und das Ende der DDR. In: Reiher, Ruth/Läzer, Rüdiger (Hg.): *Wer spricht das wahre Deutsch?* Berlin. S. 28–52.
- Teubert, Wolfgang (1996): Zum politisch-gesellschaftlichen Diskurs im Postsozialismus. In: Reiher, Ruth/Läzer, Rüdiger (Hg.): *Von Buschzulage und Ossinachweis. Ost-West-Deutsch in der Diskussion*. Berlin. S. 286–318.



Das in der Germanistik lange vernachlässigte Thema der Neologie und des lexikalischen Wandels wird in theoretischen, methodologischen und praktischen Aspekten beleuchtet. Es wird gezeigt, welchen Beitrag die Korpuslinguistik bei der Objektivierung des Bedeutungswechsels bereits vorhandener lexikalischer Ausdrücke leisten kann und welche Relevanzkriterien für die lexikographische Bearbeitung erfüllt sein müssen.