

KS und BS = Bezugssubstantiv als Kern der Substantivgruppe und Basissubstantiv, deren Spezifik nach Möglichkeit semantisch gekennzeichnet wird.

- 3) Innerhalb des possessiven Funktionsparadigmas ist je nach Semantik des Basis- und Bezugswortes mit semantischen Nuancen der elementaren Prädikationen mit dem Klassenmerkmal *haben* zu rechnen, z.B.: *erdige Rüben* → 'Rüben, die voll Erde sind' (*voll sein* im Sinne von 'behaftet sein mit' einem äußeren ornativen Merkmal), *kalkiger Boden* → 'Boden, der (viel) Kalk enthält' (solche quantitativen Abstufungen finden sich vor allem bei den Suffixoiden: *kalkreich* - *kalkhaltig* - *kalkarm*), und schließlich der elliptische Typ *übellauniges Gesicht* → 'Gesicht, das üble Laune (des X) verrät/das verrät, daß X üble Laune hat' (der Träger des im Basiswort genannten Zustands ist ausgespart).

9. Manfred Hellmann, FÖS (Bonn): Sprachliche Differenzen zwischen Bundesrepublik und DDR

9.0. Seit Jahren wird unter Sprachwissenschaftlern, Journalisten, Philologen und interessierten Zeitungslesern über die sprachlichen Unterschiede zwischen Bundesrepublik und DDR diskutiert - mit sehr unterschiedlichen Ergebnissen. Es war von "Sprachspaltung die Rede und von unvermeidlichen Wortschatzdifferenzen, von "klassenspezifischen Bedeutungen" und "Parteiargon"; neuerdings sprechen DDR-Linguisten von vier "nationalsprachlichen Varianten", nämlich dem Deutsch der DDR, der Bundesrepublik Deutschland, der Schweiz und Österreichs.

9.1. Die Bonner Forschungsstelle für öffentlichen Sprachgebrauch (F.ö.S.) hat seit ihrer Gründung (im Jahre 1964 auf Veranlassung und unter Leitung von Prof. H. Moser) die Aufgabe übernommen, dieses umstrittene, unter dem Aspekt der "Einheit der Nation" und der Verständigung zwischen Bürgern hüben und drüben gewiß auch politisch relevante Forschungsgebiete mit einem möglichst hohen Grad an Objektivität zu bearbeiten und für weitere Forschungen zu erschließen.

9.1.1. Dazu mußten zunächst einige Grundlagen geschaffen werden, aus denen sich die laufenden Grundaufgaben der Bonner Forschungs-

stelle entwickelten:

1. Beobachtung und bibliographische Bereitstellung der Literatur zum öffentlichen Sprachgebrauch in der BRD und der DDR (Zeitschriften-Aufsatz-Archiv).

Auf der Grundlage dieses bibliographischen Archivs, das durch eine entsprechend spezialisierte Handbibliothek ergänzt wird, wurde umfangreicher bibliographischer Service für eine Reihe externer Forschungsvorhaben geleistet; ferner wurde daraus die kommentierte "Bibliographie zum öffentlichen Sprachgebrauch in der BRD und der DDR" (1976) entwickelt.

2. Dokumentation und maschinelle Bereitstellung von west- und ostdeutschen Zeitungstexten.

Im Rahmen dieser Teilaufgabe wurde das sogenannte "Bonner Zeitungskorpus" erstellt, eine in Fünfjahresschritten zeitlich gestufte Sammlung von repräsentativen Auswahltexten aus zunächst nur je einer ost- und westdeutschen Zeitung (WELT und NEUES DEUTSCHLAND). Zeitungstexte bieten den Vorteil einer hohen Aktualität und einer sehr großen thematischen, inhaltlichen und formalen Streuung des Materials. Gerade diese Heterogenität macht aber auch eine sorgfältige Klassifikation der Texte erforderlich, weil keineswegs jede der in den Zeitungen einer Seite vorkommenden Textsorten mit den Textsorten der anderen Seite voll vergleichbar ist. So z.B. kommen Werbung und Stellenanzeigen im ND fast gar nicht vor, umgekehrt sind die meisten Sparten im ND weit stärker politikorientiert als in der WELT. Aus diesem Grunde wird jeder Artikel unseres Zeitungsmaterials sorgfältig außer nach Herkunft, Datum, Autor und Umfang auch nach den Kategorien Sparte, Sachgebiet und Mitteilungsform klassifiziert.

Aus diesem Material wurde mit Hilfe geeigneter Programme für eigene und viele externe Untersuchungen maschineller Service geleistet.

3. Auswertung von Wörterbüchern und Sekundärliteratur

Insbesondere das Ostberliner "Wörterbuch der deutschen Gegenwartssprache" sowie eine Reihe wichtiger Aufsätze wurden im Hinblick auf die dort gebuchten Hinweise auf lexikalische Ost- bzw. West-Spezifika ausgewertet.

4. Vergleichende Untersuchungen zu Besonderheiten im öffentlichen Sprachgebrauch der BRD und der DDR.

Bevorzugt wurden hier Arbeiten zu bestimmten Begriffsfeldern, aber auch wortmonographische Untersuchungen sowie konzeptionelle Arbeiten für die Entwicklung vergleichender Wörterverzeichnisse.

9.1.2. DFG-Projekt "Ost-West-Wortschatzvergleiche"

Nachdem die Deutsche Forschungsgemeinschaft im Juni 1976 das

Projekt "Ost-West-Wortschatzvergleiche" bewilligt hatte (Projektleiter: Dr. Manfred Hellmann), konnten die genannten Arbeitsbereiche in weiterem Umfang und mit konkreter Zielsetzung in Angriff genommen werden.

Das Projekt "Ost-West-Wortschatzvergleiche" hat folgende Ziele:

- a) Erweiterung der vorhandenen Textbasis aus ost- und westdeutschen Zeitungstexten durch Aufnahme zusätzlicher Regionalzeitungen und Überregionalzeitungen (DDR: Der Morgen, Norddeutsche Neueste Nachrichten; BRD: Bonner General-Anzeiger, Frankfurter Rundschau) sowie durch die Aufnahme sehr früher und sehr aktueller Jahrgangsauswahlen (1949 und 1974), so daß nach Fertigstellung ein Zeitraum von 25 Jahren Nachkriegs-Sprachgeschichte dokumentiert sein wird. (Vgl. schematische Darstellung am Schluß des Berichts).
- b) Entwicklung von Verfahrensweisen zur (teil-)maschinellen Verarbeitung von Texten zu Wörterbüchern. Hier soll, ausgehend von den Textdateien, versucht werden, mit Hilfe des Rechners dem lexikographischen Bearbeiter möglichst viel und möglichst präzise sprachliche Daten an die Hand zu geben, die zur Beschreibung eines Stichworts in einem Wörterbucheintrag erforderlich sind ("maschinelles Korpus-Wörterbuch" - MKWb).
- c) Erstellung vergleichender Wörterverzeichnisse zur ost- und westdeutschen Zeitungssprache im Hinblick auf Art, Ausmaß, Schwerpunkte und Entwicklungstendenzen der jeweiligen lexikalischen Besonderheiten. Dabei geht es weniger um die oft diskutierten Bedeutungsunterschiede bei ideologischen Kernbegriffen wie z.B. "Freiheit", "Demokratie", "Sozialismus", "fortschrittlich", für deren umfassende Beschreibung unser Zeitungsmaterial vielleicht keine ausreichende Grundlage bietet, sondern vor allem um die Vielzahl neuer, unterschiedlicher Sachbezeichnungen, die nach dem Krieg in jeweils einem der beiden deutschen Staaten neu entstanden und geläufig geworden sind. Daneben wird der zahlenmäßige Vergleich der Texte auch Gebrauchsbesonderheiten aufzeigen, die nicht unbedingt auf Bedeutungsunterschiede, wohl aber auf thematische oder textfunktionale Präferenzen zurückzuführen sind. Der maschinelle Vergleich zeitlich gestuften Materials erlaubt erstmals auch die Feststellung quantitativer Wortschatzverschiebungen über mehrere Jahrzehnte hin.

Die Projektkonzeption sieht eine Laufzeit von vier Jahren vor, davon die ersten zwei Jahre (bis April 1978) im wesentlichen für die Textdokumentation sowie die Festlegung und Erprobung der Verfahren, die zweiten für die Erstellung der vergleichenden Wörterverzeichnisse.

Der Projektplanung folgend werden zur Zeit die Arbeitsgebiete (1) Textdokumentation, (2) Sekundäres Material, (3) Programmierung intensiv bearbeitet; ferner werden im Rahmen einer (4) Pilotstudie bestimmte Teilfragen der Textverarbeitung, der Lexembearbeitung und der Konzeption des "maschinellen Korpus-Wörterbuchs" behandelt.

Zu (1) Textdokumentation (verantwortlich: B. Schaefer):

Die Arbeiten zur Erfassung und Bereitstellung der Zeitungstexte sind weit fortgeschritten und werden termingerecht, d.h. bis Ende April 1978, den in der Graphik dargestellten Stand erreicht haben. (Graphik siehe nächste Seite.) Auf die wünschenswerte Verstärkung des Anteils von Regionalzeitungen (z.B. Jahrgänge 1954) mußte aus Kapazitäts- und Zeitgründen verzichtet werden; aus dem gleichen Grunde mußte auch die Textklassifikation vereinfacht werden.

Alle neu aufgenommenen Texte wurden und werden durch eine beauftragte Service-Firma erfaßt und vorkorrigiert.

Zu den fertigen Jahrgängen liegen vor:

- alphabetische Wortformen-Indices mit Stellenangaben,
- alphabetische und nach Rängen sortierte Register (ohne Stellenangaben)
- Vergleichsregister (Register mit parallelen Häufigkeitsangaben zu zwei oder mehr Texten) sowie
- diverse Konkordanzen (Textauszüge zu Suchbegriffen mit Kontext).

Außer dem Bonner Zeitungs-Korpus stehen zu Vergleichszwecken noch folgende weitere Korpora zur Verfügung:

- das Mannheimer Korpus I einschließlich umfangreicher Texte aus der BILD-Zeitung 1967
- zusätzliche Texte aus dem Jahrgang 1974 der WELT
- das LIMAS-Korpus, eine breitgestreute Sammlung unterschiedlicher Texte aus dem Jahre 1970
- beantragt ist die Korrektur und Bereitstellung des "Lunder Korpus", eine Sammlung von Zeitungstexten der WELT und der Süddeutschen Zeitung aus dem Jahre 1967, die von Frau Prof. Rosengren an der Universität Lund erstellt worden ist.

AUFBAU DES BONNER ZEITUNGSKORPUS

		WE 49 (200.000)	ND 49 (170.000)		
		WE 54 (240.000)	ND 54 (240.000)		
		WE 59 (260.000)	ND 59 (230.000)		
	RZW 64 (340.000)	WE 64 (400.000)	ND 64 (160.000)	RZO 64 (100.000)	
		WE 69 (370.000)	ND 69 (160.000)		
FR 74 (310.000)	RZW 74 (340.000)	WE 74 (310.000)	ND 74 (180.000)	RZO 74 (100.000)	MO 74 (140.000)
310.000	680.000	1.780.000	1.114.000	200.000	140.000

Erläuterungen:

- WE: Die Welt
- ND: Neues Deutschland
- RZW: Bonner General-Anzeiger
- RZO: Norddeutsche Neueste Nachrichten (Rostock)
- FR: Frankfurter Rundschau
- MO: Der Morgen (überregionale Tageszeitung der DDR)

Die in Klammern gesetzten Zahlen geben jeweils die Menge der laufenden Wörter (running words) an .

Umfang des Bonner Zeitungskorpus: 4.280.000 laufende Wörter (running words)

Zu (2) Sekundäres Material (verantwortlich: Dr. Kinne, Dr. Schmidt)

Das aus dem Wörterbuch der deutschen Gegenwartssprache (WDG) und der einschlägigen Sekundärliteratur exzerpierte Belegmaterial repräsentiert für uns zunächst den derzeitigen Stand der Forschung. Insofern ist es unerlässlich zur Strukturierung des Problemfeldes und zur Bildung von Hypothesen darüber, welche Typen von Wortschatzbesonderheiten und Bedeutungsunterschieden es gibt. Gleichwohl gehen wir nicht davon aus, daß dieses Sekundäre Material die tatsächlichen Wortschatzunterschiede vollständig oder richtig wiedergäbe. Entscheidend für die Feststellung und Beschreibung dieser Wortschatzunterschiede ist letztlich der Textvergleich, jedoch kann das Sekundäre Material den Prozess des Auffindens und auch der Beschreibung der Lexeme verkürzen.

Auch diese Arbeiten werden zum Ende der ersten Projektphase beendet sein.

Zu (3) Programmierung (W. Krause, H. Siebenbach)

Die projektbezogenen Rechenarbeiten werden zum größeren Teil bei der Gesellschaft für Mathematik und Datenverarbeitung (GMD) Bonn-Bad Godesberg, zum kleineren Teil auf der IdS-eigenen Rechananlage in Mannheim erledigt. Vorrangig erstellt wurden neue Index- und Registerprogramme, neue Suchprogramme sowohl für Listen wie auch für Texte, sowie Dienstprogramme zum DATA-Management und Arbeitsprogramme. Da, anders als bisher im Wissenschaftsbetrieb üblich, für alle Rechenarbeiten volle Kosten zu erstatten sind, mußten alle Programme, aber auch die Technik der Programmierung und der Auswertung selbst, auf möglichste Kostenoptimierung hin orientiert werden. Beim Aufbau der diversen Dateien und der Zugriffsverfahren konnte daher auch nicht ohne weiteres eines der vorhandenen, aufwendigen Datenbanksysteme eingesetzt werden. Wir haben uns für ein speziell zugeschnittenes, kostengünstiges eigenes System entschieden. Zur Frage der maschinellen Lemmatisierung sowie der maschinellen Umfeldanalyse als Hilfsmittel zur semantischen Analyse wurden mehrere fremde Systeme geprüft, teil-

weise wurden Kooperationsvereinbarungen getroffen.

Zu (4) Pilotstudie:

Folgende Teilfragen, die zur Klärung des methodischen Vorgehens und zur Erstellung der späteren Wörterverzeichnisse im weiteren Projektverlauf wichtig sind, werden zur Zeit an Testmengen behandelt:

- Untersuchung von Frequenzdifferenz als Indiz für Gebrauchsunterschiede
- Prüfung verschiedener statistischer Verfahren zur Ermittlung signifikanter Vorkommens- und Verteilungsunterschiede
- Vergleich einer vorläufigen aus Texten gewonnenen Liste wahrscheinlich spezifischer Lexeme mit der aus dem Sekundären Material gewonnenen Liste angeblich spezifischer Lexeme
- Überprüfung der Sachgebietsgebundenheit des Vokabulars als Indiz für Ost-West-Spezifik
- Entwicklung von Modellen für die Gestaltung der Wörterbucheinträge, insbesondere für die unterschiedlichen Arten von Bedeutungsangaben.

Ziel der Pilotstudie ist es, anhand ausgewählten Textmaterials die Verfahren zu beschreiben, mit deren Hilfe die jeweiligen lexikalischen Besonderheiten unserer Ost- bzw. West-Texte möglichst vollständig erfaßt, geprüft und schließlich in den vergleichenden Wörterverzeichnissen beschrieben werden können, sowie die Festlegung der dafür erforderlichen Arbeitsschritte.

Sollten die vorgeschlagenen Verfahren die Zustimmung der Gutachter finden, erhoffen sich die Projektmitarbeiter eine wenigstens zweijährige statt der bisher sehr kurzatmigen einjährigen Bewilligungsdauer.

9.1.3. Projekt "Kleines Wörterbuch des DDR-Wortschatzes"

Das Bundesministerium für innerdeutsche Beziehungen hat dem Institut für deutsche Sprache (Projektleiter: Dr. Manfred Hellmann, mitverantwortlich: Dr. Wolfgang Mentrup) den Auftrag erteilt, innerhalb eines Zeitraums von 18 Monaten ein kleines, handliches,

allgemeinverständliches Wörterbuch der wichtigsten DDR-spezifischen Wörter herzustellen. Geplanter Umfang: ca. 450 Hauptstichwörter. Wissenschaftlicher Bearbeiter: Dr. Michael Kinne, Bonn, sowie zwei Hilfskräfte. Das Projekt ist personell und seiner Zielsetzung nach selbständig, jedoch organisatorisch und teilweise auch methodisch mit dem DFG-Projekt verzahnt und auf dieses angewiesen. Es trägt auf der anderen Seite auch zur konkreten lexikographischen Arbeit im DFG-Projekt bei.

9.1.4. Projektantrag "Lunder Korpus"

Ebenfalls bei der Deutschen Forschungsgemeinschaft beantragt worden ist ein Projekt, mit dessen Hilfe das von Frau Prof. Inger Rosengren, Universität Lund (Schweden) erstellten Korpus westdeutscher Zeitungstexte korrigiert sowie in standardisierter, benutzerfreundlicher Form als Klartextausdruck, alphabetische Indices sowie KWIC-Konkordanzen (auf Microfiche) dem Benutzer zur Verfügung gestellt werden soll. Interessant für das Projekt "Ost-West-Wortschatzvergleiche" sind diese Texte vor allem insofern, als sie schon in lemmatisierter Form vorliegen. Zur Realisierung des Vorhabens sind Kooperationsvereinbarungen getroffen worden zwischen Frau Prof. Rosengren als Urheberin, dem Projektleiter (Dr. Manfred W. Hellmann, F.ö.S., Bonn), Prof. Lenders vom Institut für Kommunikationsforschung und Phonetik der Universität Bonn und dem Rechenzentrum des IdS Mannheim. Die Arbeiten müssen spätestens Anfang 1978 beginnen und im Herbst des gleichen Jahres abgeschlossen sein.

10. Wolfgang Mentrup: Großes interdisziplinäres Wörterbuch der deutschen Sprache der Gegenwart

10.0. In dem auf der Jahrestagung des Instituts für deutsche Sprache (IdS), Mannheim, 1975 gehaltenen Vortrag "Die Wahrheit der Wörterbücher"¹⁾ geht Weinrich davon aus, daß das 20. Jahrhundert unter der Herrschaft der Wissenschaften und ihrer Fachsprachen steht und daß eine Verständigung zwischen den Wissenschaften und zwischen der Gemeinsprache und den Fachsprachen nicht