

Homographen-Wörterbuch
der deutschen Sprache



Homographen-Wörterbuch der deutschen Sprache

von

Heinz J. Weber

Walter de Gruyter · Berlin · New York
1996

© Gedruckt auf säurefreiem Papier,
das die US-ANSI-Norm über Haltbarkeit erfüllt

Die Deutsche Bibliothek – CIP-Einheitsaufnahme

Weber, Heinz Josef:

Homographen-Wörterbuch der deutschen Sprache / von Heinz
J. Weber. – Berlin ; New York : de Gruyter, 1996

Zugl. Teildr. von Saarbrücken, Univ., Diss. 1973 u. d. T.:

Weber, Heinz Josef: Mehrdeutige Wortformen im heutigen
Deutsch

ISBN 3-11-014641-X

© Copyright 1996 by Walter de Gruyter & Co., D-10785 Berlin

Dieses 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.

Printed in Germany

Satz und Druck: Arthur Collignon GmbH, Berlin
Buchbinderische Verarbeitung: Lüderitz & Bauer-GmbH, Berlin

Vorwort

Die Konzeption dieses Homographen-Wörterbuchs und ein Teil des hier ausbreiteten Materials stammen aus meiner 1973 fertiggestellten Dissertation. Darin habe ich ein Verfahren zur maschinellen Erzeugung und Erfassung mehrdeutiger Wortformen entwickelt, das in ein maschinenzugängliches Lexikon-system integriert und bei der automatischen Analyse deutscher Texte eingesetzt werden kann.

Ich danke allen, die mir damals zur Seite gestanden haben, insbesondere Hans Eggers, der mich zu diesem Vorhaben ermutigte. Seinem Andenken ist dieses Wörterbuch gewidmet.

Trier, im Herbst 1994

Heinz J. Weber

Inhalt

Vorwort	V
-------------------	---

1. Teil Das Homographen-Wörterbuch (HWB)

1	Warum ein Homographen-Wörterbuch?	3
1.1	Homonymie – Homographie – Homophonie – Heterophonie . .	5
1.2	Homonymie – Polysemie	6
1.3	Homonymen-Wörterbücher	9
2	Das Homographen-Wörterbuch (HWB)	11
2.1	Worin unterscheidet sich das Homographen-Wörterbuch von anderen Homonymen-Verzeichnissen?	11
2.2	Die Schreibung im HWB: Lexikon-Schreibung und Text-Schreibung	11
2.3	Welche Homographien werden im HWB verzeichnet?	14
2.4	Homographie und Homographen-Reihen	18
2.5	Zum Vokabular des HWB	19
2.6	Zum Aufbau des Homographen-Wörterbuchs	20
3	Homographen-Klassifikation	22
3.1	Aufbau der Klassifikation	22
3.2	Wortformen-Zerlegung	23
3.3	Typisierung der Grundformen nach Flexionsbereichen	24
3.4	Gegenüberstellung der Konstituenten zu Vergleichspaaren	25
3.4.1	Gegenüberstellung der Konstituenten: Ermittlung von GD	26
4	Wie wird das Homographen-Wörterbuch benutzt?	29
4.1	Identifikation einer Wortform als Homograph	30
4.2	Integration einer Homographie ins HWB	32
5	Zur Häufigkeit von Homographien im HWB	34
5.1	Anzahl der Wortformen und Grundformen	34
5.2	Zur Anzahl der Lesarten	34
5.3	Zur Häufigkeit einzelner Homographie-Typen	35
5.3.1	Konstituenten-Homographie – Konstruktions-Homographie	35
5.3.2	Heterophone, Vornamen, Funktionswörter	36

5.4	Überlegungen zur Reduktion von Homographien durch Großschreibung und Wortklassen-Bestimmung	37
5.4.1	Schlußfolgerungen	39
5.5	Weitere Beobachtungen zur Häufigkeit von Homographien	39
5.5.1	Häufigkeiten und Wortlaut (Wort-Länge und Graphie).	39
5.5.2	Homographie und Wortklassen	40
5.5.3	Homographie und Wort-Struktur	40
6	Literaturhinweise	44
7	Konsultierte Wörterbücher	48

2. Teil

Die Einzelverzeichnisse des HWB

0.1	Abkürzungen, Sonderzeichen	51
0.2	Die Wortklassen im HWB	52
0.2.1	Erläuterungen zu einzelnen Wortklassen	53
1	Der Wortformen-Index	55
1.1	Erläuterungen zum Wortformen-Index	57
	Wortformen-Index	61
2	Der Grundformen-Index	231
2.1	Erläuterungen zum Grundformen-Index	233
	Grundformen-Index	235
3	Die Homographen-Klassen	355
3.1	Anordnung der Klassen und Subklassen	357
3.2	Der Aufbau der Einträge im Homographen-Klassen-Verzeichnis	357
3.3	Weitere Hinweise zur Kodierung	359
3.3.1	Imperativformen	359
3.3.2	Substantivierungen und Partizipien	359
3.3.3	Subgruppe NS	360
3.3.4	Subgruppe HS	361
3.3.5	Zur Homographie Verb / Verb	362
3.4	Übersicht über die an Homographien beteiligten Grundformvarianten	363
3.5	Übersicht über die Homographen-Klassen	368
	Verzeichnis der Homographen-Klassen	377
4	Hilfsverzeichnisse	525
4.1	An Homographien beteiligte Doppelformen und Schreibvarianten	527
4.2	Flexions-Muster	532
4.3	Heterophone im HWB	539
4.4	An Homographien beteiligte Funktionswörter	550
4.5	An Homographien beteiligte Vornamen	552
4.6	Homographien mit mehr als zwei Lesarten	553

1. Teil

Das Homographen-Wörterbuch (HWB)

1 Warum ein Homographen-Wörterbuch?

Als Homonyme (bzw. Homographe – zu dieser Unterscheidung später mehr) bezeichnen wir Schriftfolgen, die als Formen verschiedener Wörterbuch-Einträge interpretiert werden können: z. B. *Balls* als Form zu den Einträgen ‚Ball‘ („kugelförmiger Körper“) und ‚Ball‘ („Tanzfestlichkeit“), *vornehme* als Form zu den Einträgen ‚vornehm‘ (Adjektiv) und ‚vornehmen‘ (Verb), *Dosen* als Form zu ‚Dose‘ und ‚Dosis‘, *gefällt* als Form zu ‚gefallen‘ und ‚fällen‘ oder *Fächern* als Form zu ‚Fächer‘ und ‚Fach‘ etc.

Nun stehen bei Problemen, die mit unserem Vokabular zusammenhängen, bekanntlich entsprechende Wörterbücher bereit: zur Rechtschreibung, zur korrekten Aussprache, zur Wortgeschichte, zur Bedeutungserklärung, zur Synonymie, um nur die wichtigsten zu nennen. Auf Wörterbuch-Informationen kann man jedoch in der Regel nicht unmittelbar zugreifen. Zuerst muß der Benutzer zu der vorliegenden Schriftfolge (z. B. *gehst*) – unter Beachtung bestimmter Regularien – ein Stichwort konstruieren (z. B. ‚gehen‘), unter dem er dann die gesuchten Informationen aufsuchen kann (z. B. starkes Verb, intransitiv). Erst wenn die Stichwort-Ermittlung gelungen ist, stehen die zugeordneten Wörterbuch-Informationen zur Verfügung.

Bei Homographie gestaltet sich die Konsultation eines Wörterbuchs jedoch schwieriger, da es mit der Ermittlung **eines** Stichwortes nicht getan ist. Dazu muß man erst einmal wissen, **daß** es sich um einen Fall von Mehrfachinterpretation handelt, damit nicht bereits nach dem ersten Sucherfolg die Stichwort-Bestimmung abgebrochen wird.

Die gängigen Wörterbücher fürs Deutsche geben zur Ermittlung ihrer Stichwörter nur wenig Hilfestellung: Sie setzen meist voraus, daß der Benutzer die aufzusuchenden Stichwörter bereits kennt oder überlassen es seinem Spürsinn, die Suche fortzusetzen. Die Zuordnung **mehrerer** Stichwörter kann begünstigt sein, wenn die fraglichen Stichwörter sich mehr oder weniger gleichen und bei alphabetischer Anordnung nahe beieinanderstehen – das wäre für ‚Ball‘ und ‚Ball‘ zur Wortform *Balls* oder für ‚vornehm‘ und ‚vornehmen‘ zur Wortform *vornehme* der Fall. Schwierigkeiten treten vor allem dann auf, wenn sich die Stichwörter zu einer Wortform voneinander unterscheiden, wie z. B. ‚Dose‘ und ‚Dosis‘ für die Wortform *Dosen* oder ‚fällen‘ und ‚gefallen‘ für die Wortform *gefällt*.

Das betrifft vor allem – aber nicht allein – Deutschlernende, die auf Wörterbücher angewiesen sind, um sich über Bedeutungsumfang, Aussprache oder Rechtschreibung einer bestimmten Wortform Klarheit zu verschaffen und denen die ausgeprägte Flexionsmorphologie des Deutschen Schwierigkeiten

macht. Sie – die auf präzise Hinweise angewiesen sind – müssen sich diese aus Grammatiken und Flexionstabellen selbst erschließen. Damit werden sie bei der Identifikation mehrdeutiger Wortformen häufig überfordert und das betreffende Wörterbuch erfüllt hier seinen Zweck als Nachschlagewerk nur unzureichend.

Dem versucht das HWB abzuhelpen. Es verzeichnet die wichtigsten Wortformen-Mehrdeutigkeiten des Deutschen und beschreibt deren formales Potential. Über diese Kodierung können den homographen Wortformen die zulässigen Wörterbuch-Stichwörter zugeordnet werden. In der Regel wird es sich dabei um **zwei** Stichwörter handeln, in manchen Fällen um drei oder vier; aber auch fünf oder sechs sind möglich. Konsultiert eine Person zuerst das Homographen-Wörterbuch, wird sie auf mögliche Mehrfach-Interpretation einer Wortform hingewiesen, erhält Aufschluß über die entsprechenden Stichwörter und kann dementsprechend ihre Suche in einem anderen Wörterbuch gestalten, falls noch weitere Informationen gewünscht werden.

Neben der Erleichterung der Wörterbuch-Suche für Deutschlernende oder unsichere Durchschnittsbenutzer ergeben sich für das Homographen-Wörterbuch vielfältige Anwendungsmöglichkeiten in der Computerlinguistik, und zwar vorrangig in den Gebieten, die sich mit der Ausdrucksseite des Deutschen befassen und mit computerzugänglichen Wörterbüchern arbeiten. Dazu sind zu rechnen: die Entwicklung von Korrekturhilfen, Synonymendiensten, Lemmatisierungsverfahren; ebenso die Bereiche Spracherkennung und maschinelle Übersetzung. Was Deutschlernenden Schwierigkeiten macht, wird erst recht bei einem sprachverarbeitenden Computerprogramm zum Problem, dem ja in der Regel auch der sprachliche und situative Kontext fehlt, um mit Mehrdeutigkeiten umzugehen.

In neuerer Zeit stellt sich (erneut) das Problem des sogenannten ‚tagging‘, d. h. die maschinelle Zuweisung grammatischer Kategorien zu Text-Wortformen auf der Grundlage stochastischer, distributioneller oder morphographematischer Analysemethoden – diesmal für sehr umfangreiche Textcorpora (nicht selten mit mehreren Millionen laufender Wortformen). Gerade in Verbindung mit der Aufbereitung massenhafter Sprachdaten ohne Wörterbuch-Unterstützung müssen die Systeme mit einer Vielzahl **artificialer** Wortformen-Mehrdeutigkeiten fertigwerden. Hier dürfte eine vorweggenommene Markierung **echter** Homographien, wie sie im Homographen-Wörterbuch vorliegt, zu einer Entlastung der Analyse beitragen.

Hervorzuheben ist, daß das hier vorgelegte Homographen-Verzeichnis auch nach der **Orthographiereform** seine Aktualität in vollem Umfang behält. Die Auswirkungen der Neuregelung auf das Vokabular des HWB sind verhältnismäßig gering. So ist eine **Vermehrung** von Homographien nicht zu erwarten. **Homographie-vermeidend** wirkt allenfalls die Zulassung von Dreifach-Konsonanten bei Komposita (z. B. *Bettuch* vs. *Betttuch* oder *griffest* vs. *grifffest*).

Die Zulassung von Schreibvereinfachungen für bestimmte Fremdwörter führt bei einigen Homographen zu einem Nebeneinander von **Schreibvarianten**:

anthrazit – antrazit, isotherm – isoterm, homograph – homograf, homophon – homofon, differential – differenzial, potential – potenzial etc.

Das HWB weist auf derartige Fälle ausdrücklich hin (vgl. 2.5). Auch Änderungen in der Schreibung von ß und ss sind entsprechend markiert (siehe dazu noch 2.2).

1.1 Homonymie – Homographie – Homophonie – Heterophonie

Die Beschreibung der Mehrdeutigkeit sprachlicher Zeichen nimmt in den verschiedenen sprachwissenschaftlichen Disziplinen breiten Raum ein. Es existiert eine Vielzahl von Begriffen und Betrachtungsweisen, die hier nicht ausgebreitet werden sollen. Soweit es die Spielarten von Mehrdeutigkeit angeht, die im Rahmen des HWB behandelt werden, erscheint es jedoch geboten, einige Vereinbarungen zur Sprechweise zu treffen: Eine wichtige Unterscheidung in Zusammenhang mit sprachlicher Mehrdeutigkeit ist die Trennung in **lexikalische** und **strukturelle** Mehrdeutigkeit, wobei die erste Bezeichnung sich auf Mehrdeutigkeit auf der Ebene des Vokabulars bezieht, während letztere für die Mehrdeutigkeit von syntaktischen Konstruktionen gilt:

1. lexikalische Mehrdeutigkeit: *er vermisßt den Golfplatz*
vermisßt: zwei lexikalische Einheiten sind als Interpretation möglich:
 ‚vermessen‘ / ‚vermissen‘
2. strukturelle Mehrdeutigkeit: *kleine Kinder mögen alte Leute gern*
kleine Kinder kann Subjekt des Satzes sein oder Akkusativ-Objekt
 bzw. *alte Leute* ist als Objekt oder Subjekt möglich.

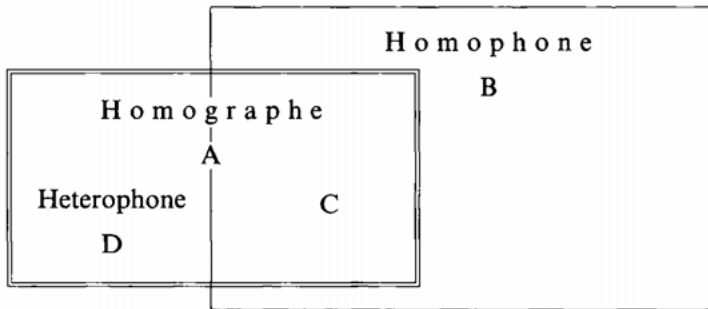
Das HWB behandelt nur Mehrdeutigkeit auf der Ebene des Vokabulars. Diese wird häufig als **Homonymie** bezeichnet. Von Homonymie spricht man in der Regel dann, wenn die betreffende Vokabel sowohl in gesprochener als auch in geschriebener Form mehrdeutig ist. Dies ist bei *vermisßt* ja auch der Fall. Häufig müssen die lautliche und die schriftliche Ausdrucksform jedoch getrennt betrachtet werden, denn es gibt Mehrdeutigkeiten im Bereich der Lautung, die in der Schreibung keine Entsprechung finden und umgekehrt. Man spricht daher bei Mehrdeutigkeiten im Bereich der Lautung von **Homophonie** („Gleichheit der Lautung“), bei Mehrdeutigkeiten im Bereich der Schreibung von **Homographie** („Gleichheit der Schreibung“). In einer engeren Auslegung werden mit Homophonie bzw. Homographie Diskrepanzen zwischen Lautung und Schreibung bezeichnet. Dabei gilt für Homophonie, daß eine Wortform in beiden Lesarten gleiche Lautung – aber verschiedene Schreibung aufweist (für [mo:r]: *Moor* – *Mohr*). Die umgekehrte Entsprechung von Homophonie wäre dann Homographie: z. B. *modern*. Die gesprochenen Korrelate unterscheiden sich jeweils hinsichtlich der Betonung ([mo'dɛrn] bzw. [ˈmodern]). Diese Unterscheidung wird jedoch nicht durchgängig praktiziert und wir schließen uns ihr nicht an.

Das HWB behandelt ausschließlich lexikalische Mehrdeutigkeit auf der Ebene der Schreibung, d. h. Homonymie in geschriebener Form, die hier

durchgängig als **Homographie** bezeichnet wird. Den Fall, daß gleichgeschriebene Wortformen in lautlicher Form verschieden sein können, bezeichnen wir als **Heterophonie** („Verschiedenheit der Lautung“). Lexikalische Mehrdeutigkeit, die ausschließlich in lautlicher Form auftritt und in schriftlicher Form keine Entsprechung findet, wird im HWB nicht behandelt. Wir verweisen dazu auf Mentrup 1971.

Das nachstehende Diagramm (Fig. 1) stellt die im HWB behandelten Typen lexikalischer Mehrdeutigkeit in doppelter Umrandung dar:

Fig. 1



Die Menge der Homonyme setzt sich zusammen aus den Teilmengen A (Homographie) und B (Homophone). Die Schnittmenge von A und B enthält die Elemente, die sowohl homograph als auch homophon sind (C). Die Teilmenge D enthält die Heterophone, i. e. die Elemente, die homograph, aber nicht homophon sind. Natürlich sagen die Proportionen der einzelnen Teilmengen im Diagramm nichts aus über die tatsächlichen Häufigkeitsverhältnisse (siehe dazu Abschnitt 5).

1.2 Homonymie – Polysemie

Das HWB verzeichnet lexikalische Mehrdeutigkeit explizit bereits auf der Ebene der **Wortformen**. Dagegen werden in den allgemeinen Wörterbüchern (wie dem Duden-Universalwörterbuch oder Wahrigs Deutschem Wörterbuch – um nur einige zu nennen) lexikalische Mehrdeutigkeiten explizit fast ausschließlich auf der Ebene der jeweiligen **Grundformen** behandelt, genauer gesagt: auf der Ebene identischer Grundformen (siehe dazu Weber 1974: 10–15).

Bei der Darstellung lexikalischer Mehrdeutigkeit identischer Grundformen haben sich in unserer Wörterbuch-Tradition im wesentlichen zwei Möglichkeiten herausgebildet, die man mit dem Begriffspaar **Homonymie-Polysemie** umschreiben kann. Von Homonymie-Notierung spricht man, wenn bei überein-

stimmendem Ausdruck und differierendem Inhalt von **verschiedenen** sprachlichen Zeichen (d. h. Wörterbuch-Einträgen) ausgegangen wird. Von „Polysemie“-Notierung, wenn differierende Inhalte als Varianten **desselben** Zeichens (Wörterbuch-Eintrags) verzeichnet werden. Zwischen diesen beiden Alternativen muß bei der Darstellung lexikalischer Mehrdeutigkeit eine Entscheidung getroffen werden.

Homonymie-Notierung:

Jede Bedeutung unter einem eigenen Stichwort.

Stichwort: Informationsteil:

Zug 1:	Substantiv maskulinum	„Verkehrsmittel“
Zug 2:	Substantiv maskulinum	„Bewegung einer Figur beim Brettspiel“
Zug 3:	Substantiv maskulinum	„stetige Luftbewegung“
Zug 4:	Substantiv maskulinum	„geschlossen ziehende Menschenmenge“

usw.

Polysemie-Notierung:

Alle Bedeutungen unter **einem** Stichwort.

Stichwort: Informationsteil:

Zug:	Substantiv maskulinum	1. „Verkehrsmittel“, 2. „Bewegung einer Figur beim Brettspiel“, 3. „stetige Luftbewegung“, 4. „geschlossen ziehende Menschenmenge“
------	-----------------------	---

usw.

Die Nachteile beider Notierungen liegen auf der Hand: Die Homonymie-Notierung – durchgängig angewendet – führt zu einer enormen Vermehrung der Stichwörter und einer Aufblähung des Wörterbuchs, da für **jede** Bedeutung ein **eigenes** Stichwort anzusetzen ist. Die Polysemie-Notierung (unter **einem** Stichwort sind **alle** Bedeutungen notiert) führt – durchgängig angewendet – zu einer Aufblähung des einzelnen Eintrags. Es verwundert daher nicht, daß sich keines der gängigen Standardwörterbücher einseitig für die Polysemie- oder die Homonymie-Notierung entschieden hat. Eine Auflösung dieses Dilemmas wird darin gesucht, zwischen beiden Extrem Lösungen einen gangbaren Mittelweg finden; d. h. die Entscheidung, ob lexikalische Mehrdeutigkeit als Polysemie oder Homonymie darzustellen ist, wird nicht generell für alle Einträge, sondern für jede einzelne Mehrdeutigkeit getroffen. In der praktischen Durchführung wirft dies jedoch einige Schwierigkeiten auf, was sich daraus ersehen läßt, daß die Entscheidung über Homonymie- oder Polysemie-Notierung in verschiedenen Wörterbüchern unterschiedlich ausfällt (siehe dazu Weber 1974: 17–20). Sprachliche Bedeutungen können oft nicht scharf konturiert werden, so daß eine simple Entweder-Oder-Entscheidung in Bezug auf die Homonymie-Polysemie-Notierung häufig weder möglich noch angemessen ist.

Auch eine graduelle Unterscheidung dürfte nicht ohne Widersprüche möglich sein, denn ein allgemein akzeptiertes Repertoire von Beschreibungselementen, mithilfe dessen diese Abstufung verbindlich durchgeführt werden könnte, liegt bislang für keine Sprache vor.

Es sind mehrere Versuche unternommen worden, die Notierung lexikalischer Mehrdeutigkeiten auf eine tragfähige theoretische und experimentell nachvollziehbare Grundlage zu stellen. Wir möchten sie hier nicht im einzelnen abhandeln, sondern verweisen auf die Diskussion dieser Problematik bei Wichter (1988) und Zöfgen (1989). Diese Versuche bestehen – kurz zusammengefaßt – darin, das Problem Bedeutungsunterscheidung nicht direkt anzugehen, sondern auf andere Beschreibungsebenen zu verlagern.

Eine Beschreibungsebene, auf die viele Standardwörterbücher zurückgreifen, ist die Herkunft von Wörtern (Etymologie). Beispiele dafür sind die verschiedenen Einträge zu ‚Kater‘ („Unwohlsein nach Alkoholgenuß“ und „männliche Katze“) oder ‚Kohl‘ („Gemüse“ oder „Unsinn“). Eine andere Möglichkeit besteht in der Verlagerung auf eine formalgrammatische Beschreibungsebene. Bevorzugt werden dabei relativ unstrittige flexionsbezogene Angaben, deren Kenntnis bei den meisten Wörterbuch-Benutzern vorausgesetzt werden kann: Wortklasse, Genus, Flexionstyp. Ob ein Substantiv als Femininum oder Neutrum beschrieben werden kann, ist leichter zu entscheiden als die Frage, ob ein „Konkretum“ oder ein „Abstraktum“ vorliegt. Grammatische Kategorien dieser Art sind nicht nur in weiterem Ausmaß akzeptiert als inhaltliche, ihr Umfang ist auch weitaus begrenzter als etwa ein Repertoire feinkörniger inhaltlicher Kategorien, wie es für eine Bedeutungsbeschreibung in Zusammenhang mit der Homonymie-Polysemie-Unterscheidung erforderlich wäre.

Aber auch dieser Weg führt zu Ergebnissen, die der sprachlichen Intuition zuwiderlaufen, so daß festzuhalten bleibt: Bei der Bedeutungsunterscheidung auf der Grundlage von Etymologie oder flexionsgrammatischen Kategorien handelt es sich um Behelfsmaßnahmen. Die Tatsache, daß zwei Bedeutungen gemeinsamer Herkunft sind oder mit übereinstimmenden grammatischen Angaben beschrieben werden können, spricht nicht zwingend dafür, von **einem Zeichen** (d. h. von Polysemie) auszugehen.

Stichwort: Informationsteil:

Absatz 1	Substantiv maskulinum	„Teil der Schuhsohle“,
Absatz 2	Substantiv maskulinum	„Gliederungselement eines Schriftstücks“,
Absatz 3	Substantiv maskulinum	„Menge verkaufter Waren“

können – trotz übereinstimmender grammatischer Informationen wie Genus und Flexion – in **unterschiedlichen** Wörterbuch-Einträgen notiert werden. Ebenso wenig müssen differierende grammatische Angaben dazu führen, von **verschiedenen Zeichen** (d. h. von Homonymie) auszugehen, wie folgendes Beispiel zeigt.

Stichwort: Informationsteil:

Dossier	1. Substantiv neutrum
	2. Substantiv maskulinum

dürfte angemessen in **einem** Wörterbuch-Eintrag zusammengefaßt werden – trotz des Genus-Unterschieds.

Daraus folgt, daß über die Konstitution eines Wörterbuch-Eintrags, d. h. seine Abgrenzung gegenüber anderen Einträgen und die Festlegung seiner internen Struktur, als letzte Instanzen das Sprachgefühl und die linguistische Kompetenz der bearbeitenden Person(en) entscheiden. Und daraus folgt wiederum, daß Notierungsunterschiede zwischen Wörterbüchern nicht zu vermeiden sind (dazu auch Lyons 1992: 139 und Zöfgen 1989: 785). Wir beziehen diesen Umstand auch auf das eigene, hier vorgelegte Projekt.

In der Uneinheitlichkeit des Aufbaus von Wörterbüchern oder in differierenden Eintragsnotierungen jedoch grundsätzlich einen Mangel zu sehen, wird der Sache nicht immer gerecht. Wörterbücher sind nicht allein unter zeichentheoretischen oder allgemeinlinguistischen Aspekten zu bewerten, sondern sollten auch an ihren selbstgesteckten Zielen gemessen werden. Etymologische Wörterbücher stellen andere Anforderungen als Wörterbücher zur Rechtschreibung oder zum Erlernen von Fremdsprachen (dazu Zöfgen 1989: 785), so daß eine an die Ziele und die Adressaten angepaßte Darstellung lexikalischer Mehrdeutigkeiten von Fall zu Fall unterschiedlich sein kann.

1.3 Homonymen-Wörterbücher

Versuche, lexikalische Mehrdeutigkeiten des Deutschen zu erfassen, sind schon sehr früh unternommen worden, waren jedoch meist auf Grundformen beschränkt. Eine wichtige Zielsetzung war die Vermeidung von Bedeutungsverwechslungen (Verwechslungsprophylaxe).

Als Beispiele wären zu nennen (in der Reihenfolge ihres Erscheinens):

Johann Fabritius 1532. Eyn nutzlich buchlein ettlicher gleich stymender wort-her Aber ungleichs verstandes. Erfurt (Neuaufgabe in ‚Ältere deutsche Grammatiken in Neudrucken‘. John Meier (ed.). Straßburg: Trübner, 44 Seiten)

Ch. L. Pape 1752. Verzeichnis einiger Wörter, welche dem Tone nach gleichlautend, aber dem Verstande nach sehr unterschieden sind. Berlin

Michael Kunitsch 1803. Grammatisch-orthographisches Wörterbuch der Homonyme der deutschen Sprache. Ein Handbuch für Kanzelleyen und Schullen nach Adelungs Grundsätzen. Erster Band. Grätz (208 Seiten)

Friedrich E. Petri 1807. Gleich- und ähnlich lautende Wörter der deutschen Sprache. Pirna (176 Seiten)

Nikolaus Bärmann 1810. Homonymicon der Deutschen. Oder vollständiges Verzeichnis aller gleichlautenden, dem Sinne nach aber verschiedenen Wörter von ganz verschiedener Bedeutung. Hamburg (128 Seiten)

G. C. Müller 1814. Gemeinnütziges, homonymisches Wörterbuch. Nürnberg (244 Seiten)

Victor J. Dewora 1817. Die meisten gleich- und ähnlich lautenden, aber der Bedeutung und Abstammung nach verschiedenen Wörter der deutschen Sprache, zum Gebrauche bei dem Dictir-Schreiben in Schulen. Coblenz (167 Seiten)

Karl F. Dietrich 1823. Versuch einer vollständigen, alphabetisch geordneten Sammlung der gleich- und ähnlich lautenden Wörter deutscher Sprache. Freiburg (528 Seiten)

Hermann Robolsky 1830. Deutsches Homonymicon, oder vollständige Sammlung der gleich und ähnlich lautenden Wörter unserer Sprache. Dessau

C.Friedrich Weiß 1847. Die Homonymen. Laut- und klangverwandte Wörter der deutschen Sprache. Nürnberg

K.W. Eulenhaupt 1888. 800 gleich- und ähnlichlautende Wörter mit kurzen Andeutungen über Abstammung oder Bedeutung nebst Anwendung derselben in 100 Sätzen. Nürnberg (17. Auflage)

F.W. Flachsmann 1889. 1700 gleich und ähnlich lautende Wörter. Zürich

Diese Wörterbücher sind zitiert nach Zaunmüller 1958 und Hausmann 1990. Daneben existieren Spezial-Listen, die in Grammatiken integriert sind:

J. G. Schottel. Wörter, die zwar der Ausrede nach ganz gleich aber in ihrer Deutung ganz unterschieden seyn. In: Ausführliche Arbeit von der Teutschen Haupt-Sprache (S. 198 ff. und 648–686).

Duden-Grammatik. Liste gleichlautender nichtverwandter Substantive mit verschiedenem Geschlecht und verschiedener Bedeutung. (Spalten 1395 ff.) Mannheim/Wien/Zürich: Dudenverlag

Jean M. Zemb 1984. Essentielle und akzidentielle Homographie. (Eine Auflistung von etwa 600 französischen und deutschen Homonymen) in: Vergleichende Grammatik Französisch-Deutsch. Teil 2 (490–507). Mannheim/Wien/Zürich: Dudenverlag.

2 Das Homographen-Wörterbuch (HWB)

2.1 Worin unterscheidet sich das Homographen-Wörterbuch von anderen Homonymen-Verzeichnissen?

Die genannten Homonymen-Wörterbücher entsprechen in Aufbau und Struktur der Einträge weitgehend den Standardwörterbüchern. Die Stichwörter (Grundformen) sind alphabetisch angeordnet, die Informationsteile enthalten grammatische Angaben mit Definitionen und/oder Beispielsätzen (Hausmann 1990a und 1990b). Die in Grammatiken integrierten Homonymenlisten führen die entsprechenden Formen lediglich alphabetisch sortiert und ohne Informationsteile auf.

Das HWB soll als Identifikationswörterbuch eine vermittelnde Rolle zwischen den in Texten vorkommenden Wortformen und den Stichwörtern herkömmlicher Wörterbücher spielen, mit dem Ziel, mehrdeutige Wortformen den möglichen Wörterbuch-Einträgen zuzuordnen. Bei erfolgreicher Zuordnung kann auf die dort verzeichneten Inhaltsbeschreibungen zugegriffen werden. Die lexikalischen Informationen des HWB brauchen deshalb nicht so umfassend und explizit zu sein wie die der Standardwörterbücher. In der Hauptsache gehören dazu Wortklassen- und Flexionsangaben, verschiedentlich Bedeutungs- oder Aussprachehilfen, die auch nur soweit entfaltet werden, wie zur Unterscheidung von Lesarten notwendig. Auch im Aufbau (der Makrostruktur) gibt es Unterschiede zu anderen Homonymen-Verzeichnissen. Das HWB setzt sich aus mehreren aufeinander bezogenen Einzelverzeichnissen zusammen. Dadurch kann die Bestimmung einer Homographie stufenweise vorgenommen werden (siehe dazu 2.6).

2.2 Die Schreibung im HWB: Lexikon-Schreibung und Text-Schreibung

Wenn wir sagen „Adjektive werden kleingeschrieben“, dann soll das heißen: „Adjektive werden **im Lexikon** kleingeschrieben“. **Im Text** dagegen gelten in mancher Hinsicht andere Schreibgepflogenheiten: Hier können Adjektive auch

großgeschrieben werden. So wird z. B. das Adjektiv ‚arm‘ im Duden-Recht-schreib-Wörterbuch als *arm* notiert. Im Text ist diese Vokabel aber auch noch als *Arm* oder *ARM* zulässig:

Arm am Beutel, krank am Herzen ...
ARM AM BEUTEL, KRANK AM HERZEN ...

Wir fassen diese Art der Schreibung zusammen als **Text-Schreibung** und stellen sie der **Lexikon-Schreibung** gegenüber. Die Lexikon-Schreibung ist normiert, verallgemeinert, die Text-Schreibung ist variabel, spezifisch. Da es sich beim HWB um ein Wörterbuch handelt, vereinbaren wir im Anschluß eine **Lexikon-Schreibung**, die als Generalisierung gegenüber der Text-Schreibung der Wortformen aufzufassen ist.

Wir gehen von insgesamt vier Ebenen des Schreibens aus. Unterhalb der beiden Ebenen Lexikon-Schreibung und Text-Schreibung unterscheiden wir zwei weitere Ebenen.

Lexikon-Schreibung.

1. Basisebene (Morphematische Schreibung):

Sie bezieht sich auf Morpheme als kleinste bedeutungstragende Einheiten und die entsprechenden Allomorphe. Die morphematische Schreibung wird von den anderen Ebenen vorausgesetzt:

seh, sieh, sah, steh, stand, ...
tisch, bein, leser, leserin, ...
fern, dumm, gescheit, ...

2. Lexikalische Ebene:

Sie betrifft die Schreibung der lexikalischen Zitierformen und der zugehörigen Flexionsformen, differiert von der morphematischen Schreibung im wesentlichen am Wortanfang (Groß-Klein-Schreibung):

sehen, stehen, ...
Hilfe, Tisch, Bein, Leser, Leserin, ...
fern, lieb, gescheit, ...

Text-Schreibung.

3. Syntagmatische Ebene:

Sie betrifft die Schreibung in Konstruktionen und weicht – vor allem am Wortanfang – von der morphematischen bzw. der Lexikonschreibung ab. So kann z. B. ‚Tisch‘ in Wortzusammensetzungen kleingeschrieben, ‚fern‘ kann großgeschrieben werden:

Wortbildung:
Tischbein, Fernseh Tisch, ...
Sehhilfe, ...
fernstehen, ...

Für die Schreibung im Satz gelten noch weitere Vorschriften, z. B. Großschreibung am Satzanfang oder bei Substantivierung:

Satz:

*Er kam nicht zum Stehen.
Getrennt von Tisch und Bett.
Liebe Leserinnen, liebe Leser
Fern der Heimat, ...
Es kam nichts Gescheites.*

Durch syntagmatische Schreibung von Konstruktionen können Homographien entstehen, die bei lexikalischer Schreibung der Einzelkonstituenten nicht vorliegen:

Bei Wortbildung:

Syntagmatische Schreibung:

*Weinkrampf
Platzangst*

Lexikalische Schreibung
der Konstituenten:

*Wein-Krampf / wein-Krampf
Platz-Angst / platz-Angst*

Im Satz:

Sie kaufte Würste zum Braten. ... zum Braten / ... zum braten.

4. Gestaltungsebene:

Sie bestimmt das „Schriftbild“ eines geschriebenen Textes mit; Grundlage der Variationen ist die morphematische Schreibung. Die Schreibung der anderen Ebenen ist optional:

*Liebe Leser(innen) ...
liebe leser(innen) ...*

*Liebe LeserInnen ...
LIEBE LESER(INNEN)*

Die bekannten Orthographie-Regeln beziehen sich auf die Ebenen 1, 2 und 3. Die 4. Ebene ist – mit Ausnahme der morphematischen Schreibung – orthographie-unabhängig. In Orthographiediskussionen werden diese Ebenen nicht selten unzulässig miteinander vermischt, z. B. in der leidigen Frage der Schreibung des Wortanfangs (groß oder klein?), die ja auf den Ebenen 2 und 3 angesiedelt werden muß; auf der ersten Ebene ist sie irrelevant und auf der vierten Ebene besitzt sie häufig ornativen Charakter.

Diskrepanzen zwischen Lexikon-Schreibung und Text-Schreibung bereiten Sprach- und Schriftkundigen in der Regel keine Schwierigkeiten. So wird der Unterschied zwischen Groß- und Kleinbuchstaben als nicht so gravierend aufgefaßt wie etwa der Unterschied zwischen ‚b‘ und ‚p‘ oder zwischen ‚B‘ und ‚P‘. Auch eine (lexikalisch und syntagmatisch) unzulässige Schreibung der Art

arM ..., ARm ... oder aRm für arm

dürfte eher akzeptiert werden als die morphematisch fehlerhafte Schreibung

arn ..., orm ... oder awm.

Die gleiche Toleranz besteht auch bei der Darstellung von ‚ß‘:

groß, gross, Groß oder GROSS

sind übliche Schreibvarianten im Text, die sich einigermaßen problemlos auf Lexikon-Schreibung (*groß*) beziehen lassen. Auch bei Wortformen mit Akzenten sind Varianten möglich:

fête, Fête, Fete, FETE für *Fête*.

Dem Sprach- und Schriftkundigen tritt geschriebene Sprache häufig in Text-Schreibung gegenüber (Ebene 3 oder 4). Bei Konsultation eines Wörterbuchs trifft er auf Lexikon-Schreibung (meist Ebene 2). Sind die entsprechenden Orthographie-Regeln nicht geläufig, kann sich der Transfer schwierig gestalten. Für das HWB legen wir daher eine Schreibregelung fest, bei der sich Text- und Lexikon-Schreibung besser aufeinander beziehen lassen, als dies üblicherweise der Fall ist. Sie ist der morphematischen Schreibung (Basisebene) sehr nahe und unterscheidet sich von der lexikalischen Schreibung (2. Ebene) in folgenden Punkten:

1. Kleinschreibung:

Wortformen werden im HWB durchgängig kleingeschrieben. Bei Fällen, in denen eine Homographie nur aufgrund syntagmatisch bedingter Großschreibung zustande kommt, wird ein Hinweis gegeben.

2. *ß* wird zu *ss*:

Das *ß* wird durchgängig als *ss* dargestellt, nicht nur in den Fällen, in denen die reformierte Schreibung das *ß* zugunsten von *ss* aufgegeben hat (wie in ‚Fluss‘). Dadurch werden einige schreibungsbedingte Homographien verursacht: *schlossen* für *schloßen* und *schlossen*.

3. Umlaute:

Umlaute werden – wie häufig praktiziert – als Vokal + *e* notiert: *raeubern* für *räubern*.

4. Akzente:

Akzente (z. B. bei Fremdwörtern aus dem Französischen) werden nicht dargestellt, so daß auch hier Homographien anfallen können: z. B. *pate* für ‚Pate‘ «Taufonkel» und ‚pâté‘ «Pastete».

2.3 Welche Homographien werden im HWB verzeichnet?

Die Ausführungen zur Unterscheidung von Polysemie und Homonymie dürften deutlich gemacht haben, daß es einen beträchtlichen Bereich gibt, in dem eine Entscheidung darüber, um welchen Typ von lexikalischer Mehrdeutigkeit es sich jeweils handelt, von Wörterbuch zu Wörterbuch verschieden ausfallen kann.

In diesem Bereich ist auch von einem Identifikationswörterbuch wie dem HWB, das vor der Benutzung anderer Wörterbücher konsultiert werden soll,

kein entscheidender Beitrag zur Trennung von Homonymien und Polysemien zu erwarten. Das HWB konzentriert sich deshalb auf die Fälle, die in den bekannten Allgemeinwörterbüchern des Deutschen unstrittig als Homonyme (Homographie) verzeichnet werden. Es handelt sich dabei um Mehrdeutigkeiten, bei denen inhaltliche Unterschiede mit Unterschieden auf anderen Beschreibungsebenen einhergehen. Dazu gehören Homographien

A) mit Unterschieden im **Wortlaut** der Grundformen:

lauter: laut (Adjektiv) / lauter (Adjektiv)

B) mit Unterschieden in der **Wortklasse**:

billige: Adjektiv / Verb

C) mit Unterschieden in der **Flexion**; flexivische Unterschiede können sich mehrfach zeigen, im Genus, in Flexionstyp, in Flexionsmerkmalen oder in der partiellen oder vollständigen Übereinstimmung der betreffenden Flexionsparadigmen:

Band: Substantiv maskulinum / neutrum
väterlich: Adjektiv steigerbar / nicht steigerbar
Bank: Substantiv Plural umlautend / nicht umlautend
senden: Verb stark flektierend / schwach flektierend

Gleichgeschriebene Formen innerhalb eines Flexionsparadigmas (sogenannte interne Homographie)

singt: 3. Singular / 2. Plural Präsens Indikativ von ‚singen‘

und syntaktisch bedingter Wortklassen-Übertritt (z. B. von Verb zu Adjektiv (Partizip))

bestimmte: Verb, Präteritum' / Adjektiv, Positiv von ‚bestimmen‘

werden im HWB durch Flexions-Muster bestimmt (siehe 2. Teil: Hilfsverzeichnisse, 4.2) und durch spezielle Markierungen hervorgehoben (2. Teil: Homographen-Klassen, 3.3.2).

D) mit Unterschieden in der **Wortstruktur**:

Gründung: grün-Dung / gründ-ung,
vertier: vert-ier / ver-Tier

E) mit Unterschieden in der **Lautstruktur** (Heterophone):

durchfahren: 'durchfahren / durch'fahren

In den aufgezählten Fällen strebt das HWB fürs Deutsche Vollständigkeit an – mit Ausnahme der Homographien zwischen trennbaren Verben und Simplex-Verben, die nicht verzeichnet werden (siehe dazu 2. Teil: Homographen-Klassen, 3.3.5). Dagegen werden die nun folgenden Fälle nur gelegentlich berücksichtigt (und durch eine entsprechende Markierung gekennzeichnet). Es handelt sich um Elemente offener Klassen, die auch in den gängigen Wörterbüchern nur unvollständig verzeichnet sind:

a) Homographie mit Beteiligung von Namen:

Konstanz: ,Konstanz‘ «Beständigkeit» / ,Konstanz‘ «Stadt»

b) Homographie mit Beteiligung von Abkürzungen:

RIAS: ,RIA‘ / ,RIAS‘ «Sender in West-Berlin»

c) Systematische Homographien bei Wortableitungen (z. B. Diminutiva; siehe dazu die Beschreibung in Weber 1974: 138–147):

Öhrchen: ,Ohr‘ / ,Öhr‘
Knöchelchen: ,Knöchel‘ / ,Knochen‘

d) Um das Verzeichnis in einem vertretbaren Rahmen zu halten, werden auch die folgenden Kuriositäten ausgeklammert. Auch sie sind zahlenmäßig nicht zu begrenzen, da es sich – bei mindestens einer der Lesarten – um nichtlexikalisierte Komposita bzw. um sprachlich unsinnige Konstruktionen handelt. Mit artifiziellen Mehrdeutigkeiten (absichtsvoll konstruiert, z. B. in Wortspielen, oder zufällig entstanden – z. B. durch maschinelle Wortzerlegung) ließe sich ein eigenes Verzeichnis füllen. Wir führen dazu nur einige der bekannteren Beispiele auf, da sich bei etwas Übung jede(r) Interessierte eigene Exemplare fabrizieren kann. Die unterschiedlichen Wortstrukturen werden durch lexikalische Schreibung der Wortbestandteile (siehe dazu 2.2) angedeutet. Weitere Beispiele in Weber 1974: 104–113, 140 ff.

<i>Astrologe:</i>	Astrologe / Astro-Loge
<i>Automaten:</i>	Automat-en / Au-Tomate-n
<i>Ballast:</i>	Ballast / Ball-Ast
<i>Barkasse:</i>	Barkasse / bar-Kasse
<i>Barbier:</i>	Barbier / Bar-Bier
<i>Generalisten:</i>	Generalist-en / Genera-Liste-n
<i>Gendarm:</i>	Gendarm / Gen-Darm
<i>Instrumentarien:</i>	Instrumentar-ien / Instrument-Arie-n
<i>Kurtage:</i>	Kurtage / Kur-Tag-e
<i>Neurose:</i>	Neurose / neu-Rose
<i>Urlaub:</i>	Urlaub / Ur-Laub
<i>Messbecher:</i>	Mess(e)-Becher / mess-Becher

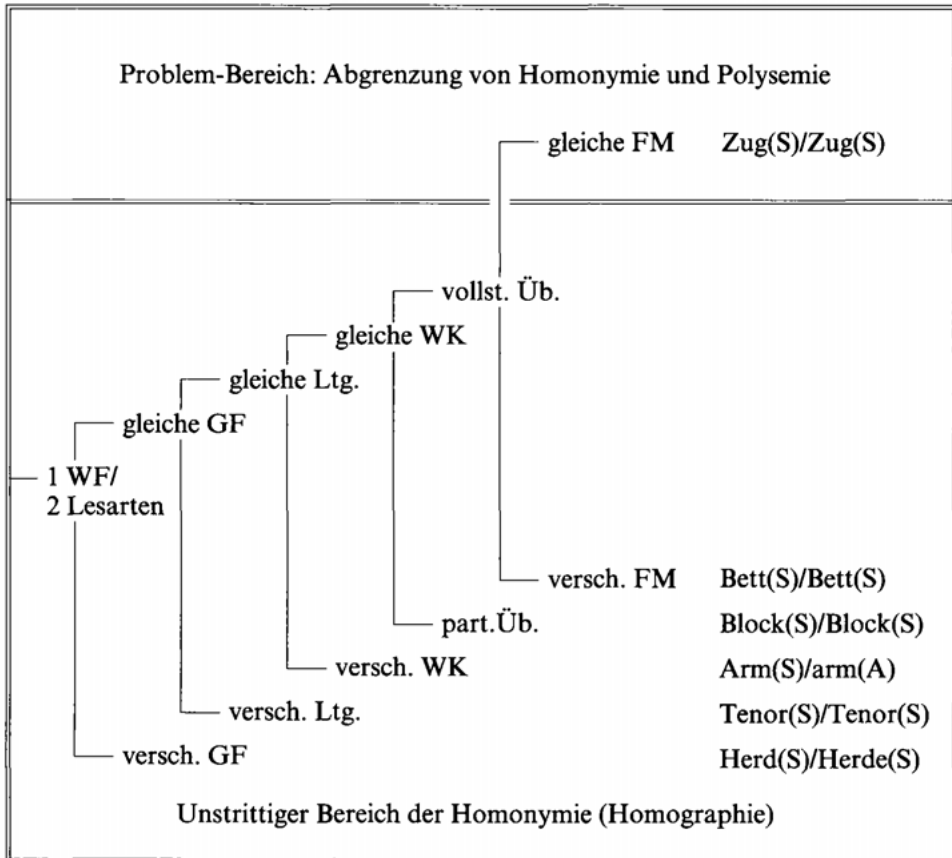
<i>Stillehre:</i>	Stil-Lehre / still-Ehre (in alter Schreibung noch: stil(l)-Lehre)
<i>Platzangst:</i>	Platz-Angst / platz-Angst
<i>Steuersystem:</i>	Steuer-System / steuer-System
<i>Weinkrampf:</i>	Wein-Krampf / wein-Krampf
<i>Fächerkombination:</i>	Fa(e)ch-er-Kombination / Faecher-Kombination
<i>Kastengeist:</i>	Kaste-n-Geist / Kasten-Geist
<i>Schuppentier:</i>	Schuppe-n-Tier / Schuppen-Tier
<i>Rechengerät:</i>	rech(e)n-Gerät / rech-en-Gerät
<i>Zeichenlehrer:</i>	zeich(e)n-Lehrer / Zeichen-Lehrer
<i>Bilderfolge:</i>	Bild-Erfolg-e / Bilder-Folge
<i>Druckerzeugnis:</i>	Druck-Erzeugnis / Drucker-Zeugnis
<i>Gliedersatz:</i>	Glied-Ersatz / Glieder-Satz
<i>Musikerleben:</i>	Musik-Erleben / Musiker-Leben
<i>Texterfassung:</i>	Text-Erfassung / Texter-Fassung
<i>Verbumbstellung:</i>	Verb-Umstellung / Verbumb-Stellung
<i>Urinsekt:</i>	Ur-Insekt / Urin-Sekt
<i>abtrennen:</i>	ab-trenn-en / Abt-Rennen
<i>angelernte:</i>	an-gelernt-e / Angel-Ernte
<i>herzeigen:</i>	her-zeig-en / Herz-eigen
<i>Jungfernfahrt:</i>	Jungfer-n-Fahrt / jung-fern-Fahrt
<i>Kulturinfiltration:</i>	Kultur-Infiltration / Kult-Urin-Filtration
<i>Liebestempel:</i>	Liebes-Tempel / Liebe-Stempel
<i>mißlungen:</i>	miß-lung-en / Miß-Lunge-n
<i>Staubecken:</i>	Stau-Becken / Staub-Ecke-n
<i>Wachstube:</i>	wach-Stube / Wachs-Tube
<i>Wachtraum:</i>	wach-Traum / Wacht-Raum
<i>zugreifen:</i>	zu-greif-en / Zug-Reifen
<i>Altarschelle:</i>	Altar-Schelle / alt-Arsch-Elle
<i>Beinhaltung:</i>	Be-Inhalt-ung / Bein-Haltung
<i>Epigeneselehre:</i>	Epigenese-Lehre / epigen-Esel-Ehre
<i>europaeischen:</i>	europäisch-en / Europa-Eis-chen
<i>Landeskinderklinik:</i>	Landes-Kinderklinik / Landeskinder-Klinik
<i>Weltschmerzkongreß:</i>	Welt-Schmerzkongreß / Weltschmerz-Kongreß
<i>Lebensmittelpunkt:</i>	Leben-s-Mittelpunkt / Lebensmittel-Punkt
<i>Rechnungshofbericht:</i>	Rechnungshof-Bericht / Rechnung-s-Hofbericht

Unberücksichtigt bleiben auch alle Mehrdeutigkeiten aus dem Problembereich zwischen Homonymie und Polysemie, die ausschließlich inhaltlich fundiert sind. Nachstehendes Diagramm faßt den Gegenstandsbereich des HWB nochmals zusammen. Die Doppellinie grenzt den unstrittigen Bereich der Homonymie (bzw. Homographie) vom Problembereich Homonymie/Polysemie ab.

Die Abkürzungen bedeuten:

WF	Wortform
(gleiche, verschiedene) GF	Grundform (Flexionsstamm)
(gleiche, verschiedene) Ltg.	Lautung
(gleiche, verschiedene) WK	Wortklasse
(partielle, vollständige) Üb.	graphische Übereinstimmung zweier Flexionsparadigmata, Flexionstyp
(gleiche, verschiedene) FM	Flexionsmerkmale: Genus, Kasus, Numerus, Person, Modus

Fig. 2 Zum Gegenstandsbereich des HWB



2.4 Homographe und Homographen-Reihen

Das HWB verzeichnet die Homographe in sogenannten **Homographen-Reihen**. Das sind Mengen von Wortformen, die in gleicher Weise homograph sind.

Als **Homographe** haben wir Schriftfolgen bezeichnet, die als Wortformen verschiedener Wörterbuch-Einträge interpretiert werden können:

Herde kann als Form des Wörterbuch-Eintrags ‚Herd‘ oder des Eintrags ‚Herde‘ interpretiert werden.

Da sich Wörterbuch-Einträgen oft mehrere Wortformen zuordnen lassen, die in gleicher Weise homograph sein können, sprechen wir von **Homographen-Reihen**:

Die Einträge ‚Herd‘ und ‚Herde‘ stimmen nicht nur in der Form *Herde*, sondern auch in der Form *Herden* überein.

Eine Homographen-Reihe läßt sich als Durchschnittsmenge zweier (oder mehrerer) Paradigmata darstellen:

‚Herd‘ / ‚Herde‘ = {*Herde*, *Herden*}

Sie kann auch aus nur **einer** Wortform bestehen:

‚Rind‘ / ‚Rinde‘ = {*Rinde*}.

Je nachdem, wie weit die Paradigmata homographier Einträge übereinstimmen, wird auch von **absoluter** oder **partieller Homographie** gesprochen. Den gleichen Sachverhalt bezeichnen die Ausdrücke **vollständige** bzw. **unvollständige Homographie** (Homonymie). Ausführlicher dazu Weber 1974: 60–62.

2.5 Zum Vokabular des HWB

Der Großteil der hier vorgelegten Homographien wurde mithilfe des in Weber 1974: 127–138 vorgestellten Homographen-Erzeugungs-Algorithmus ermittelt. Das Rohmaterial stammt aus verschiedenen Quellen. Eine davon ist das von Hans Eggers konzipierte RDE/FAZ-Corpus, eine Zusammenstellung von über 11.000 Einzelsätzen aus populärwissenschaftlicher Prosa („rowohlts deutsche enzyklopädie“) und Zeitungstexten („Frankfurter Allgemeine Zeitung“), die mehr als 200.000 laufende Wortformen enthält. Aus der algorithmischen Ermittlung rühren einige artifizielle Homographien her, die im Homographen-Klassen-Verzeichnis als (kon) – konstruiert – markiert sind (siehe dazu auch 2.3 d). Die Materialbasis wurde im Laufe der Zeit durch Fundstücke aus Fachzeitschriften und Nachrichtenmagazinen erweitert. Das gesamte Vokabular wurde mithilfe bekannter deutschsprachiger Wörterbücher (Duden-Rechtschreibung, Wahrigs Deutsches Wörterbuch, Duden-Universalwörterbuch, Klappenbach/Steinitz' Wörterbuch der deutschen Gegenwartssprache) überprüft.

Das **Klassifikationsprinzip** für Homographie hat sich gegenüber der 1974 veröffentlichten Fassung – mit Ausnahme der Notierung der Klassen in Dezimalen – nicht verändert. Allerdings sind durch Hinzunahme neuen Materials

auch neue Klassen aufgestellt worden, was sich wiederum auf die Anordnung der Homographen-Klassen ausgewirkt hat.

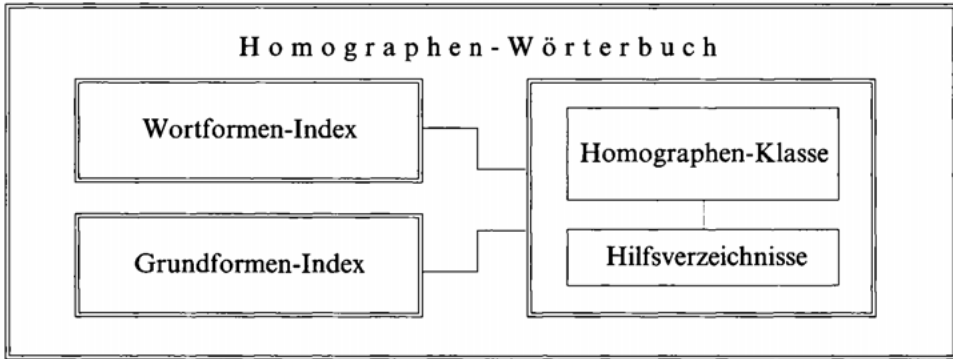
Veraltetes (obsoletes) Vokabular wurde teilweise berücksichtigt, obwohl es sich um Wortschatz-Elemente handelt, die nicht mehr auf der Höhe der Zeit sind. Sie werden gelegentlich noch benutzt, um bestimmte Wirkungen zu erzielen; z. B. um etwas zu ironisieren oder um damit zu imponieren. Wie auch immer – es ist damit zu rechnen, daß auch in Gegenwartsdeutsch abgefaßte Texte derartige Formen aufweisen. Obsolete Vokabeln sind dann aufgenommen worden, wenn sie zu anderen Formen homograph sind (wie ‚dermalen‘ oder ‚just‘). Der Vermerk (obs) weist auf solche Fälle hin. **Idiomatische Ausdrücke** sind dann berücksichtigt worden, wenn die darin vorkommenden Wortformen abweichend flektieren und mit Flexionsformen anderer Einträge homograph sind; z. B. *muttern* als Dativ Singular von ‚Mutter‘ in ‚wie bei Muttern‘, das homograph ist zu *muttern* als Pluralform von ‚Mutter‘ «Schraubenteil». Auch **Personen-Vornamen** wurden aufgenommen, sofern sie zu anderen Vokabeln homograph sind. Wie schon erwähnt, bleiben die übrigen Eigennamen unberücksichtigt (vgl. 2.3 a). **Fremdwörter** sind verzeichnet, wenn sie in einem der konsultierten Wörterbücher aufzufinden sind. **Lexikalische Varianten** werden als DF (Doppelformen) gekennzeichnet. Schriftliche Übereinstimmungen zwischen Varianten **desselben** Wörterbuch-Eintrags werden im Verzeichnis der Homographen-Klassen nicht berücksichtigt (siehe dazu auch Hilfsverzeichnisse, 4.1). **Reformierte Schreibung**, wie sie für die neue amtliche Schreibregelung vorgeschlagen ist (z. B. *homograf*, *antrazit*, *differenzial* neben *homograph*, *anthrazit*, *differential*) wird einbezogen und als RS markiert (ergänzend dazu die Markierung GS für die gängige Schreibung).

2.6 Zum Aufbau des Homographen-Wörterbuchs

Das HWB setzt sich zusammen aus mehreren aufeinander bezogenen Verzeichnissen: einem Wortformen-Index, einem Grundformen-Index, einem Verzeichnis der Homographen-Klassen und mehreren Hilfsverzeichnissen. Dadurch kann die Bestimmung einer Homographie stufenweise vorgenommen werden. Wortformen- und Grundformen-Index sind alphabetisch sortiert und dienen zur ersten Orientierung; über Adressen wird auf das Verzeichnis der Homographen-Klassen verwiesen. Hier wird eine Spezifikation der betreffenden Homographie vorgenommen, die durch die angeschlossenen Hilfsverzeichnisse präzisiert werden kann.

Die Klassifikationskriterien sind gestaffelt; sie beruhen auf der graphemischen Grundformen-Differenz GD, die sich aus der paarweisen Gegenüberstellung von Grundformen (bzw. Flexiven) ergibt, weiterhin gehören dazu die Wortklassen der beteiligten Grundformen und die Flexionsmerkmale der Wortformen für die einzelnen Lesarten. Aufgrund der binären Einteilung ist die Klassifikation offen, d. h. beliebig erweiterbar.

Fig. 3 Verbund der HWB-Verzeichnisse



3 Homographen-Klassifikation

3.1 Aufbau der Klassifikation

Die Klassifikation homographischer Wortformen baut sich in mehreren Schritten auf:

- 1) **Feststellung der Homographie** von Wortformen aufgrund ihrer Zugehörigkeit zu mehreren Wörterbuch-Einträgen.
- 2) **Zusammenfassung der Homographie**, die den gleichen Wörterbuch-Einträgen zuzuordnen sind, zu **Homographen-Reihen**.
- 3) **Zerlegung der Wortformen** einer Homographen-Reihe in Abhängigkeit von den möglichen Lesarten in Grundform und Flexiv, **Typisierung** der Wortformen-Bestandteile (Konstituenten).
- 4) **Paarweise Gegenüberstellung** der einzelnen Zerlegungsversionen und graphematischer Vergleich der entsprechenden Konstituenten: Grundform 1 mit Grundform 2 bzw. Flexiv 1 mit Flexiv 2; Ermittlung der **Grundform-Differenz GD**.
- 5) **Bildung der Homographen-Klassen** in Abhängigkeit von GD.

Die ersten beiden Schritte sind in 1.1 und 2.4 bereits behandelt worden. Dieser Abschnitt befaßt sich mit den drei nächsten, der Wortformen-Zerlegung, der Gegenüberstellung der Flexionsformen-Bestandteile und der Bildung der Homographen-Klassen. Wir stützen uns auch hier weitgehend auf die Darstellung in Weber 1974: 73–87 und 98–147.

Es wird davon ausgegangen, daß in einem deutschen Text jede Wortform als eine **Flexionsform** beschrieben und einem Wörterbuch-Eintrag zugeordnet werden kann. Die Gesamtmenge der Wortformen (sprich: Flexionsformen) zu einem Wörterbuch-Eintrag nennen wir **(Flexions-)Paradigma**. Wortformen nichtflektierender Wortklassen (z. B. Präposition, Konjunktion) gelten als Sonderfälle, bei denen das Paradigma aus jeweils nur **einer** Wortform (Flexionsform) besteht.

Wir gehen weiter davon aus, daß jede Wortform als Flexionsform in zwei Bestandteile (Konstituenten) zerlegt werden kann: in eine **Grundform** und ein **Flexiv**. Auch bei nichtflektierenden Klassen dürfte diese Annahme nachvollziehbar sein. So ergibt z. B. eine Zerlegung der Konjunktion ‚weil‘ die Grundform weil und ein leeres Flexiv (dargestellt als -#).

3.2 Wortformen-Zerlegung

Die Grundformen entsprechen weitgehend den üblichen lexikalischen Zitierformen, die in den Standardwörterbüchern als Stichwörter dienen. Dies gilt vor allem bei Substantiven und Adjektiven. Bei Verben ist die Grundform meist identisch mit dem Infinitiv abzüglich der Infinitiv-Endung (-en bzw. -n).

Bei der Ermittlung der Wortbestandteile können die Formen eines Paradigmas in der Regel **einer** Grundform zugeordnet werden. Diese ist dann als die allen Wortformen eines Paradigmas gemeinsame Schriftfolge anzusehen, die variablen Schriftfolgen entsprechend als Flexive. Als Beispiele die Flexionsformen zum Paradigma von ‚lieben‘, zerlegt in Grundform und Flexive:

Tab. 1

Grundform:	Flexive:				
	Präsens Indikativ:	Präteritum	Präsens Partizip:	Präteritum	
lieb	-e	-te	-end	ge- t	
	-st	-test	-ende	ge- te	
	-t	-te	-endem	ge- tem	
	-en	-ten	-enden	ge- ten	
	-t	-tet	-ender	ge- ter	
	-en	-ten	-endes	ge- tes	
			-endere	ge- tere	
			-enderem	ge- terem	
			-enderen	ge- teren	
			-enderer	ge- terer	
			-enderes	ge- teres	
			-endste	ge- teste	
			-endstem	ge- testem	
			-endsten	ge- testen	
			-endster	ge- tester	
			-endstes	ge- testes	
		Konjunktiv:			
		-e	-te		
		-est	-test		
		-e	-te		
	-en	-ten			
	-et	-tet			
	-en	-ten			
	Imperativ:		Substantivierung:		
	Singular	Plural			
	-#	-(e)t	-en		
			-ens		
			-en		
			-en		

Die Regelung „eine einzige Grundform pro Paradigma“ ist im Deutschen jedoch nicht für alle Flexionsformen in der geschilderten Weise durchführbar. In der Vielzahl der Fälle unproblematisch, bringt sie gelegentlich bizarre Folgen

mit sich. So wäre z. B. für das Verb ‚sein‘ die Grundform eine leere Buchstabenkette, die Flexive entsprächen den Schriftfolgen *bin, bist, war* etc. Für ‚tun‘ bestünde die Grundform nur aus *t*, für ‚werden‘ müßte innerhalb eine Leerstelle verbleiben (*w—r*), um den flexionsbedingten Wechsel bei *werde, wird, wurde, würde, geworden* berücksichtigen zu können. Es dürfte schwerfallen, diese Reste als bedeutungstragende Einheiten anzusehen. Vergleichbare Verhältnisse ergäben sich bei einigen Adjektiven und Substantiven: Die Grundform des Paradigmas von ‚gut‘ wäre leer. Die Grundform zu ‚Seemann‘ lautete *See*; Flexive wären entsprechend die Schriftfolgen *mann, leute*.

Im Deutschen sind es vor allem die starken Verben, bei denen Flexion nicht als einfache Aneinanderreihung einer konstanten Grundform mit variablen Flexiven darstellbar ist. Auch die häufig vorkommenden substantivischen Fremdwörter, wie ‚Dosis‘ (Plural: *Dosen*), ‚Index‘ (Plural: *Indices*) oder ‚Museum‘ (Plural: *Museen*), weichen davon ab. Hier verblieben als Grundformen nur noch *Dos, Ind* oder *Mus*.

Eine Möglichkeit, auch bei unregelmäßiger Flexion die Grundformen konstant zu halten, ohne sie bis zur Unkenntlichkeit zu verstümmeln, besteht in der Zulassung mehrerer Grundformen für ein Paradigma. In Weber (1974: 73–87) ist eine Möglichkeit beschrieben, Varianten einer Grundform zu ermitteln und auf festgelegte Teile eines Paradigmas (Flexionsbereiche) zu beziehen. Die dort vorgelegte Typisierung für Verben, Substantive und Adjektive wird hier übernommen. Dabei sind zwei Punkte wichtig: Innerhalb eines Paradigmas darf keine Grundform-Variante Teilmenge einer anderen Variante sein und der Satz der Flexive, der sich aus der Wahl einer Grundform-Variante ergibt, ist möglichst allgemein zu halten, damit er auf weitere Paradigmata zutrifft.

3.3 Typisierung der Grundformen nach Flexionsbereichen

Für die einzelnen Grundformen-Typen werden folgende Flexionsbereiche festgesetzt:

Tab. 2

Substantiv (a):	regelmäßige Flexion in Singular und Plural: <i>kind (Sn)</i> unregelmäßige Flexion im Singular: <i>lexikon (Sn)</i>
Substantiv (b):	unregelmäßige Flexion im Plural: <i>lexika (Sn)</i> <‚Lexikon‘>
Adjektiv (a):	regelmäßige Flexion in prädikativ Positiv, Komparativ und Superlativ: <i>lieb (A)</i> unregelmäßige Flexion in prädikativ Positiv: <i>hoch (A), gut (A), nah (A), links (A)</i>
Adjektiv (b):	unregelmäßige Flexion in attributiv Positiv: <i>hoh (A)</i> <‚hoch‘>, <i>link (A)</i> <‚links‘>

Tab. 2 (Fortsetzung)

Adjektiv (c):	unregelmäßige Flexion in prädikativ Komparativ: besser (A) <‚gut‘>
Adjektiv (d):	unregelmäßige Flexion in prädikativ Superlativ: best (A) <‚gut‘>, nach (A) <‚nah‘>
Verb (a):	regelmäßige Flexion in Infinitiv, Indikativ Präsens Singular, in Indikativ Präteritum, Konjunktiv Präteritum, Partizip Präteritum: lieb (V) unregelmäßige Flexion im Infinitiv: sei <‚sein‘>.
Verb (b):	unregelmäßige Flexion im Indikativ Präsens Singular: bin (V) <‚sein‘>, bist (V) <‚sein‘>, wirst (V) <‚werden‘>
Verb (c):	unregelmäßige Flexion im Indikativ Präteritum: nahm (V) <‚nehmen‘>, wog (V) <‚wiegen‘>, ging (V) <‚gehen‘>
Verb (d):	unregelmäßige Flexion im Konjunktiv Präteritum: wurf (V) <‚werfen‘>
Verb (e):	unregelmäßige Flexion im Partizip Präteritum: nomm (V) <‚nehmen‘>, troff (V) <‚treffen‘>

Eine Übersicht über die an Homographien beteiligten Grundform-Varianten wird im 2. Teil, Abschnitt 3.4 gegeben.

3.4 Gegenüberstellung der Konstituenten zu Vergleichspaaren

Der nächste Klassifikationsschritt besteht in der Gegenüberstellung gleicher Konstituenten: Grundform zu Grundform, Flexiv zu Flexiv. Dabei werden alle Lesarten einer homographen Wortform paarweise angeordnet. Sind zwei Grundformen voneinander unterschieden, steht die kürzere von beiden links:

sucht wird beschrieben durch das Grundformen-Paar *such* / *sucht*,
gefahren durch das Grundformen-Paar *fahr* / *gefahr*.

Bei paarweiser Gegenüberstellung ergibt bzw. ergeben sich für eine Wortform

- bei zwei Lesarten ein Vergleichspaar,
- bei drei Lesarten drei (1 + 2) Vergleichspaare,
- bei vier Lesarten sechs (1 + 2 + 3) Vergleichspaare etc.

Die Klassifikation der Homographen durch paarweise angeordnete Grundformen hat mehrere Vorteile: Einmal läßt sich eine bereits klassifizierte Homographie bei Entdeckung einer neuen Lesart durch Hinzufügen weiterer Vergleichspaare vervollständigen, ohne daß die Klassifikation für die schon bestehenden

Vergleichspaare revidiert werden müßte. Zum andern ermöglicht erst die paarweise Gegenüberstellung von Grundformen eine genaue Feststellung von Heterophonie.

So ist z. B. die Wortform *wachst* im Hinblick auf die beiden Grundformen
 wachs (V) «größer werden» / wachs (V) «mit Wachs versehen»
 homograph aber nicht heterophon;
 für die Vergleichspaare
 wachs (V) «größer werden» / wach (V) und
 wachs (V) «mit Wachs versehen» / wach (V)
 kommt noch Heterophonie hinzu: [wakst] / [wachst].

3.4.1 Gegenüberstellung der Konstituenten: Ermittlung von GD

Beim graphematischen Vergleich der Konstituenten lassen sich unmittelbar zwei Fälle unterscheiden. Wir nennen den ersten Fall „Konstituenten-Homographie“, weil hier bereits auf der Ebene der Konstituenten Homographie vorliegt, und den zweiten „Konstruktions-Homographie“, weil Homographie erst bei der Zusammenfassung beider Konstituenten in einer Konstruktion Homographie entsteht.

1. Die Konstituenten beider Lesarten stimmen überein; dies gilt für die Wortform *billige*, die in ihren beiden Lesarten zerlegt werden kann in

Lesart:	Grundform:	Flexiv:
1. „nicht teuer“	billig (A)	-e
2. „zustimmen“	billig (V)	-e

2. Die Konstituenten beider Lesarten überschneiden sich; dies trifft zu bei der Wortform *mildere*, die in ihren beiden Lesarten zerlegt werden kann in

Lesart:	Grundform:	Flexiv:
1. „sanft“	mild (A)	-ere
2. „mild machen“	milder (V)	-e

Bei der Anlage der Homographen-Klassen dient diese Unterscheidung als oberstes Einteilungskriterium.

Nächstes Einteilungskriterium ist die Spezifizierung der schriftlichen Überschneidung der Konstituenten in einem Vergleichspaar, die sogenannte **Grundformen-Differenz** (abgekürzt: GD).

Bei den Konstituenten-Homographien ist GD leer (notiert als #). Bei Konstruktions-Homographien ermöglicht dieses Kriterium eine weitere Untertei-

lung in Klassen. Auch auf anderen Ebenen der Wortstruktur (Wortableitung, Wortkomposition) kann die graphematische Differenz zwischen gleichartigen Konstituenten als Klassifikationskriterium eingesetzt werden. Wir werden dies hier nicht ausführen, sondern verweisen auf Weber 1974: 98–115.

Gleichzeitig kann GD ausgenutzt werden, um über die Ausprägung der Homographien in beiden Gruppierungen Voraussagen zu machen. Dies beginnt bei relativ allgemeinen Feststellungen, z. B. mögliche Grundformen-Konstellationen, geht über die Beteiligung von Wortklassen an den einzelnen Homographie-Typen, bis zu den Flexiven.

Beginnen wir mit der Betrachtung von GD1 in Zusammenhang mit den **Schriftfolgen** der Grundformen: Da für GD gilt, daß sie (ganz oder teilweise) mit Flexiven übereinstimmen muß, kann die Menge der zulässigen Grundformen-Vergleichspaare weiter eingeschränkt werden:

So ist z. B. das Vergleichspaar *trab* (Sm) und *trabant* (Sm) unplausibel, weil *-ant* als GD1 aufgrund der genannten Einschränkung unzulässig ist.

Gleiches gilt für das Vergleichspaar *turm* (Sm) und *sturm* (Sm): GD1 = *s-* ist unzulässig, denn es gibt kein Flexionspräfix *s-*.

Weitere unzulässige Vergleichspaare in diesem Zusammenhang sind: *fang* (V)/*fango* (Sm); *rad* (Sn)/*radau* (Sm); *futter* (Sn)/*futteral* (Sn) etc. In Betracht käme dagegen *tablett* (Sn)/*tablette* (Sf), da GD1 = *-e* mit mindestens einem Flexiv (ganz oder teilweise) übereinstimmt.

In Zusammenhang mit den **Wortklassen** der gegenübergestellten Grundformen lassen sich über GD folgende Beschränkungen ableiten.

Konstituenten-Homographie (GD1 = # / GD2 = #): Da GD1 und GD2 leer sind, also keinerlei Beschränkungen von ihnen ausgehen, ist diese Gruppe auch am variationsreichsten. Sie umfaßt alle Wortklassen, flektierende und nicht-flektierende. Auch hinsichtlich der Anzahl der Homographie ist dies die stärkste Gruppierung (siehe dazu 5.3.1 und 5.5).

Konstruktions-Homographie (GD1 < > # / GD2 = #): Da gilt: GD1 < > #, muß die Wortklasse an erster Stelle (d. h. die Wortklasse der kürzeren Grundform) flektierend sein. Zugelassen sind Wortklassen, deren Flexive mit GD1 (ganz oder teilweise) übereinstimmen. Z. B. ist für GD1 = *-em* nur eine Grundform der Wortklasse Adjektiv möglich (vgl. Klassen 8 und 9). Nur ein Teil der Möglichkeiten für GD1 ist mit Homographen-Klassen vertreten; als GD1 kommen u. a. **nicht** vor: *-u-en*, *-e-est*, *-e-test*, *-e-etest*, *-e-et*, *-e-te*, *-e-ten*, *ge-ten*.

Konstruktions-Homographie (GD1 < > # / GD2 < > #): In dieser Gruppierung sind nur noch flektierende Wortklassen zulässig, da sowohl GD1 als auch GD2 < > # sind.

Schließlich lassen sich bei der Abschätzung von möglichen Homographien auch die **Flexive** über GD zueinander in Beziehung setzen:

Das Vergleichspaar kanon (Sm)/kanone (Sf) mit GD1 = -e ist von der Schriftfolge und der Wortklasse der Grundformen her zulässig. Trotzdem sind die Flexionsformen beider Grundformen nicht homograph, da GD1 = -e nicht mit den für kanon (Sm) zulässigen Flexiven (-s, -s) übereinstimmt, wohingegen das Vergleichspaar patron (Sm)/patrone (Sf) mit GD1 = -e an einer Homographie beteiligt ist: GD1 = -e stimmt mit einem für die Grundform patron (Sm) zulässigen Flexiv überein.

In Weber 1974: 125–138 sind mehrere aus GD abgeleitete Bedingungen formuliert und in einen Algorithmus zur maschinellen Erzeugung von Homographen integriert.

4 Wie wird das Homographen-Wörterbuch benutzt?

Hier einige Probleme, bei deren Lösung dieses Wörterbuch helfen soll. Für jeden Lösungsweg sind die einzelnen Stationen des Suchpfades angegeben:

- 1) Ist eine bestimmte Wortform homograph und in welcher Weise?

Beispiel: *Gehört*

Suchpfad: Wortformen-Index → Homographen-Klassen-Verzeichnis
→ Flexions-Muster (→ Allgemeinwörterbuch)

Siehe dazu 4.1 und Fig. 4.

- 2) An welchen Homographien ist eine bestimmte Grundform beteiligt?

Suchpfad: Grundformen-Index → Homographen-Klassen-Verzeichnis
→ Flexions-Muster (→ Allgemeinwörterbuch)

- 3) Welche homographen Wortformen werden je nach Lesart unterschiedlich ausgesprochen?

Suchpfad: Hilfsverzeichnis Heterophone → Homographen-Klassen-Verzeichnis (→ Aussprachewörterbuch)

- 4) Welche Funktionswörter sind an Homographien beteiligt?

Suchpfad: Hilfsverzeichnis Funktionswörter → Wortformen-Index
→ Homographen-Klassen-Verzeichnis (→ Allgemeinwörterbuch)

- 5) Welche Vornamen sind homograph?

Suchpfad: Hilfsverzeichnis Vornamen → Wortformen-Index
→ Homographen-Klassen-Verzeichnis (→ Allgemeinwörterbuch)

- 6) Welche Homographie weisen mehr als zwei Lesarten auf?

Suchpfad: Hilfsverzeichnis Lesarten → Wortformen-Index
→ Homographen-Klassen-Verzeichnis (→ Allgemeinwörterbuch)

Dazu zwei Beispiele, die Identifikation einer Wortform als Homograph und die Integration einer noch nicht verzeichneten Homographie ins HWB:

4.1 Identifikation einer Wortform als Homograph

Die Identifikation einer homographen Wortform kann in mehreren Schritten vorgenommen werden. Angenommen, die Wortform sei *Gehört* in folgendem Zusammenhang

... *Gehört hat ihn niemand.* ...,

dann ergibt sich folgender Ablauf:

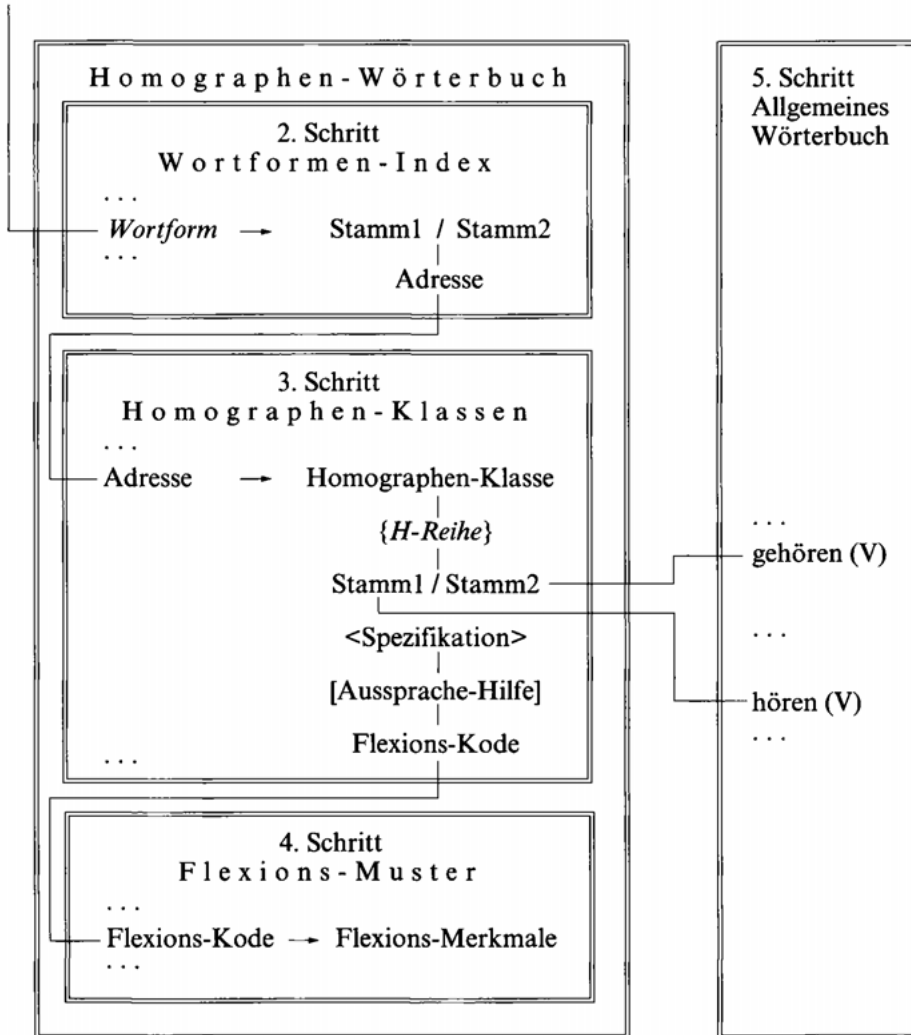
1. **Normieren** auf HWB-Schreibung (siehe Abschnitt 2.4): Kleinschreibung, Umlaut → Vokal + e.
Für das Beispiel *Gehört* ergibt sich daraus: *gehoert*
2. Nachschlagen im **Wortformen-Index** (alphabetisch angeordnet); ist die Wortform verzeichnet, werden folgende Angaben gemacht:
 - a) Wenn weitere Wortformen in gleicher Weise homograph sind, liegt eine mehrwortige Homographen-Reihe vor. Für *gehoert* umfaßt die Homographen-Reihe die Wortformen {*gehoert, gehoerte, gehoerten*}.
 - b) Die Anzahl der Lesarten ist verzeichnet: <3>, <4> etc.; im unmarkierten Fall handelt es sich um zwei Lesarten.
 - c) Heterophone sind markiert: <H>, ebenso Personen-Vornamen: <N>.
 - d) Verzeichnet ist der Wortlaut der möglichen Grundformen (Stichwörter) und ihre Wortklassen. Für obiges Beispiel kommen als Grundformen in Frage: *hoer* (V) und *gehoer* (V).
 - e) Angegeben ist die Adresse im Verzeichnis der Homographen-Klassen. Für unser Beispiel lautet die Adresse: II, 2.2.1 d1)
3. Nachschlagen im Verzeichnis **Homographen-Klassen** gemäß Index-Adresse. Im Klassen-Verzeichnis werden angegeben:
 - a) die Homographen-Reihe, zu der die betreffende Wortform gehört;
 - b) zwei Grundformen (Stichwörter) und ihre Wortklassen
 - c) bei morphologischen Alternanzformen ggf. die lexikalische Zitierform: z. B. *wuchs* (V), <,'wachsen'>
 - d) ggf. eine Bedeutungsumschreibung: <<größer werden>>

Fig. 4 Identifikation einer Wortform als Homograph

... *Gehört hat ihn niemand...*

1. Schritt: Normierung der Schriftfolge:

gehört



- e) bei Heterophonie eine Aussprachehilfe für beide Lesarten:
z. B. *hinzufahren*: [*ˈhɪnzufahren*] / [*hɪnˈzʊfahɪrən*]
- f) weitere Angaben: Doppelform, obsolet, idiomatisch ...
- g) die Flexionskodierung der Wortformen dieser betreffenden Homographen-Reihe. Für unser Beispiel ergeben sich folgende Kodierungen:

<i>gehört</i>	V: 50 / V: 16
	A: 18
<i>gehörte</i>	A: 1 / V: 32
<i>gehörten</i>	A: 9 / V: 39

4. Nachschlagen in den **Hilfsverzeichnissen**:

- a) **Flexions-Muster** gemäß der Flexionskodierung aus dem Verzeichnis der Homographen-Klassen
- b) Verzeichnis der an Homographien beteiligten **Doppelformen**

5. Nachschlagen der ermittelten Grundformen (Stichwörter) in einem **Allgemein-Wörterbuch**, um weitere Informationen einzuholen; dabei ggf. Konversion der HWB-Schreibung in allgemeine Lexikon-Schreibung. Für das Beispiel *gehört* ergäbe dies als mögliche Stichwörter ‚gehör(en)‘ und ‚hör(en)‘.

Dazu können im Rechtschreibungs-Duden (Ausgabe 1991) folgende Eintragungen ermittelt werden:

- ‚gehören‘: das Haus gehört mir;
 die mir gehörenden Häuser;
 ich gehöre zur Familie;
 südd., österr., schweiz. auch:
 ihm gehört (gebührt) eine Strafe
- ‚hören‘: er hat von dem Unglück heute gehört;
 sie hat die Glocken läuten hören od. gehört;
 von sich hören lassen

4.2 Integration einer Homographie ins HWB

Im folgenden wird die Aufnahme einer noch nicht verzeichneten homographen Wortform ins HWB – zusammen mit weiteren Wortformen derselben Homographen-Reihe – anhand eines konstruierten Beispiels durchexerziert. Gehen wir von folgender (erfundener) Textstelle aus:

Beim Entflöhen Ihres Hundes ist zu beachten, daß ...

Nun ist die Wortform *Entflöhen*, die in vorliegendem Zusammenhang als Flexionsform (genauer: als substantivierter Infinitiv) von ‚entflöhen‘ (umschrieben als «Flöhe entfernen») – analog zu ‚entlausen‘) zu verstehen ist, auch als Flexionsform *entflöhen* zu ‚entfliehen‘ möglich – es liegt demnach ein Fall von Homographie vor. Die Aufnahme ins HWB geschieht in folgender Weise:

1. **Normieren** auf HWB-Schreibung (wie oben). Für Beispiel *Entflöhen* ergibt sich daraus: *entfloehen*

2. Ermittlung der möglichen **Grundformen** (Flexionsstämme): Wortlaut, Wortklasse und Grundformen-Typ:
 1. Lesart ‚entfliehen‘: entfloh (V); Grundformen-Typ c
 2. Lesart ‚entflöhen‘: entfloeh (V); Grundformen-Typ a
3. Paarweise Gegenüberstellung der Grundformen und Ermittlung von **GD** und Einordnung in die entsprechende Homographen-Klasse: entfloh / entfloeh: GD = **-e-**; es handelt sich also um Konstruktions-Homographie. Die betreffende Homographen-Klasse: **II, 1.5.2**
4. Ermittlung möglicher weiterer Homographien zur Bildung einer **Homographen-Reihe**:
{*entfloeh*, *entfloehst*, *entfloehen*, *entfloehest*, *entfloehet*, *entfloeht*}
5. Ermittlung der entsprechenden **Flexive** für beide Grundformen:
-e-e, -e-st, -e-en, -e-est, -e-et -e-t / -e, -st, -en, -est, -et, -t;
Einordnen in entsprechende Subklasse: **II, 1.5.2 a)**
6. Ermittlung der entsprechenden **Flexionsmerkmale**:
43, 44, 46, 44, 47, 47 / 4, 7, V: 18 – S: 27, 26, 29, V: 16 – A: 18.
Einordnen in entsprechende Subklasse: **II, 1.5.2 a2)**
7. **Eintragen ins Verzeichnis der Homographen-Klassen**; ggf. weitere **Spezifikationen**
8. Übertragen der gesamten Homographen-Reihe in **Wortformen-Index**:
als Grundformen-Paarung mit Wortklassen, Spezifikation, Adresse:
{*entfloeh*, *entfloehst*, *entfloehen*, *entfloehest*, *entfloehet*, *entfloeht*}:
entfloh (V) / entfloeh (V) **II, 1.5.2 a2)**
9. Übertragen jeder Grundform in **Grundformen-Index**:
mit Grundform, Wortklasse, lexikalische Zitierform, Adresse;
entfloeh (V) **II, 1.5.2 a2)**
entfloh (V) <‚entfliehen‘> **II, 1.5.2 a2)**
10. Ggf. **Aktualisieren** der übrigen Verzeichnisse: z. B. Heterophone, Häufigkeitsangaben im 1. Teil, Abschnitt 5.

5 Zur Häufigkeit von Homographien im HWB

Bei den nun folgenden Häufigkeitsangaben beziehen wir uns auf die im HWB **explizit** verzeichneten Homographien. Wie bereits ausgeführt, handelt es sich bei Homographen nicht um eine abgeschlossene Menge von Wortformen. Viele Fälle lassen sich in ihrem Umfang nur abschätzen; so z. B. die hier nicht verzeichneten systematischen Homographien bei bestimmten Wortbildungen oder bei trennbaren Verben und Simplex-Verben (siehe dazu 1. Teil, 2.3 und 2. Teil: Homographen-Klassen, 3.3.5).

5.1 Anzahl der Wortformen und Grundformen

Das HWB verzeichnet insgesamt ca. 14.000 Wortformen-Einträge (token). Dabei handelt es sich um ca. 11.300 verschiedene Wortformen (sogenannte types), da H-Reihen sich in ihren Wortformen teilweise überschneiden können. Die Zahl der Grundformen (token) beläuft sich auf ca. 11.400 (mehr als 8.000 types). Geht man davon aus, daß ein Standardwörterbuch fürs Deutsche um die 120.000 Stichwörter enthält, dann sind etwa 10% der Einträge an Homographien beteiligt.

Die ca. 14.000 Wortformen lassen sich ca. 5.700 Grundformen-Paarungen bzw. Homographen-Reihen zuordnen, so daß pro H-Reihe im Durchschnitt etwa 2,5 Wortformen anzusetzen sind.

Tab. 3

Insgesamt: WF (token)	WF (types)	H-Reihen (GF-Paare)	Grund- formen	WF/H-Reihe
14.000	11.300	5.700	11.400	2,5

5.2 Zur Anzahl der Lesarten

Der weitaus überwiegende Teil der etwa 14.000 Homographie weist zwei Lesarten auf (ca. 13.000); die Zahl der Wortformen mit mehr als zwei Lesarten

beläuft sich auf etwa 980; das sind deutlich unter 10% der Homographe. Davon entfallen ca. 830 auf Homographe mit drei Lesarten; ca. 120 auf Wortformen mit vier Lesarten, 26 Homographe weisen fünf Lesarten auf und nur 4 Homographe haben sechs Lesarten. Wir werden auf diesen Sachverhalt noch eingehen. Homographe mit mehr als sechs Lesarten sind im HWB nicht explizit verzeichnet, dürften aber bei bestimmten Verb-Verb-Homographien häufiger vorkommen; siehe dazu 2. Teil: Homographen-Klassen, 3.3.5.

Tab. 4

Insgesamt: WF (token)	2 LA	3 LA	4 LA	5 LA	6 LA
14.000	13.000	830	120	26	4

5.3 Zur Häufigkeit einzelner Homographie-Typen

5.3.1 Konstituenten-Homographie – Konstruktions-Homographie

Die meisten Homographen-Reihen entfallen auf grundform-homographie Typen (Konstituenten-Homographie). Von ca. 5.700 H-Reihen sind dies etwa 3.200. Den deutlich kleineren Teil machen die sogenannten Konstruktions-Homographien aus: ungefähr 2.500. Entsprechendes gilt für die Anzahl der Grundformen: An den ca. 11.400 Grundformen (token) sind die Konstituenten-Homographien mit ca. 6.400 Grundformen beteiligt; für die Konstruktions-Homographien sind dies etwa 5.000.

Bei der Zahl der Wortformen verschiebt sich die Relation noch stärker zu Ungunsten der Konstruktions-Homographien: So entfallen von den 14.000

Tab. 5

Insgesamt: H-Reihen (GF-Paare)	Grundformen	Wortformen	WF/H-Reihe
5.700	11.400	14.000	2,5
Homographie-Typ I:			
3.200	6.400	10.000	3,1
Homographie-Typ II, III:			
2.500	5.000	4.000	1,6

Wortformen auf die Grundform-Homographien ca. 10.000 Wortformen, auf die Konstruktions-Homographien nur etwa 4.000 Wortformen. Daraus ergibt sich für die Anzahl der Wortformen pro H-Reihe bei den Konstituenten-Homographien ein Verhältnis von 3,1 Wortformen pro H-Reihe gegenüber 1,6 Wortformen bei den Konstruktions-Homographien (bei einem Gesamtdurchschnitt von 2,5).

5.3.2 Heterophone, Vornamen, Funktionswörter

Für Heterophone liegen insgesamt etwa 610 H-Reihen vor mit ca. 4.100 Wortformen-Einträgen/token (bzw. etwa 3.900 types). Die Relation von ungefähr 6,7 Wortformen pro H-Reihe liegt weit über dem Durchschnitt von 2,5 und dürfte auf die hohe Beteiligung von Reihen mit ‚durch-‘, ‚um-‘, ‚über-‘ und ‚unter-‘ präfigierten Verben zurückzuführen sein, die in der Regel mehr als 10 Wortformen umfassen.

Verzeichnet sind 57 H-Reihen, an denen Vornamen beteiligt sind. Wortformen-Einträge belaufen sich auf 72 token bzw. 58 types. Die Relation von knapp 1,3 Wortformen pro H-Reihe liegt deutlich unter dem Durchschnitt (2,5). Ein Grund dafür liegt in den Paradigmata für Vornamen: es gibt keine Flexionsformen im Plural.

Homographie mit Beteiligung von Funktionswortklassen ist in 16 Klassen und 107 Subklassen verzeichnet. Es liegen insgesamt 652 H-Reihen vor. Die Zahl der Wortformen-Einträge (token) ist annähernd so hoch (671), weil nur wenige H-Reihen mehr als eine Wortform umfassen. Da der Anteil der Wort-

Tab. 6

Insgesamt:				
WF (token)	WF (types)	H-Reihen (GF-Paare)	Grundformen	WF/H-Reihe
14.000	11.300	5.700	11.400	2,5
Heterophone:				
4.100	3.900	610	1.220	6,7
Vornamen:				
72	58	57	114	1,2
FWK:				
670	350	650	1.300	1,03

formen, die mehr als zwei Lesarten aufweisen, bei Funktionswortklassen überdurchschnittlich hoch ist, fällt die Anzahl der verschiedenen Wortformen (types) relativ niedrig aus: 345 (siehe auch 5.5.2).

5.4 Überlegungen zur Reduktion von Homographien durch Großschreibung und Wortklassen-Bestimmung

Auf der Grundlage der vorliegenden Homographen-Reihen und Homographen-Klassen lassen sich einige Häufigkeitsangaben zur disambiguierenden bzw. homographie-reduzierenden Rolle der Großschreibung machen. Wir haben das strittige Thema in Abschnitt 2.2) bereits angeschnitten. Deshalb hier einige Zahlen:

Das Verzeichnis der Homographen-Klassen umfaßt 55 Klassen mit 305 Subklassen. Von diesen 305 Subklassen betreffen 176 permanent homographie Wortformen (wie z. B. *vornehme* als Adjektiv/Verb- oder *rinde* als Substantiv/Substantiv-Homograph) und 129 Subklassen großschreibungs-abhängige Homographie, d. h. Homographie liegt nur vor, wenn eine Lesart im Text substantiviert auftritt oder in Satzanfangstellung steht (z. B. *extreme* als Adjektiv-Substantiv oder *streifen* als Verb/Substantiv-Homograph).

Tab. 7

Homographen-Klassen:		
Insgesamt:	permanent:	großschreibungsabhängig:
305 Subklassen	176 Subklassen	129 Subklassen

Demnach beziehen sich deutlich mehr als die Hälfte (ca. 58%) der Subklassen auf permanent homographie Wortformen, die von der herrschenden Großschreibungsregelung unberührt bleiben. Ca. 42% der Subklassen sind dagegen von der Text-Schreibung abhängig; hier wirkt die Großschreibungsregelung homographie-reduzierend.

Stärker als bei den Homographen-Klassen zeigt sich die Auswirkung der Großschreibungsregelung bei den homographen Wortformen. Von den ca. 14.000 Homographen, die das HWB verzeichnet, sind etwa 7.000 permanent homograph, die anderen 7.000 hängen von der Text-Schreibung ab,

Am deutlichsten zeigt sich die Wirksamkeit der Großschreibungsregelung bei den ca. 5.700 Homographen-Reihen bzw. den Grundformen-Paarungen: Etwa 1.800 Reihen beinhalten permanent homographie Wortformen; die Zahl der Homographen-Reihen mit großschreibungsabhängiger Homographie beläuft sich auf ca. 3.900.

Tab. 8

Insgesamt:		
WF (token)	H-Reihen (GF-Paare)	WF/H-Reihe
14.000	5.700	2,5
Permanent:		
7.000	1.800	3,9
GS-abhängig:		
7.000	3.900	1.8

Über das Ausmaß der Auswirkungen der Großschreibungsregelung im laufenden Text ist damit jedoch nichts ausgesagt. Dazu wird auf Weber/Zimmermann 1973 verwiesen.

Nun zu den Möglichkeiten der Homographie-Reduktion durch Ermittlung der Wortklasse: Von den 305 Subklassen des HWB sind der überwiegende Teil Paarungen mit verschiedenen Wortklassen (z. B. S/V oder A/VZS), und zwar 277; lediglich 28 Subklassen weisen Paarungen mit gleicher Wortklasse auf. Von den ca. 5.700 Homographen-Reihen gehen etwa 1.100 auf Grundformen-Paarungen mit gleicher Wortklasse zurück. Bei den übrigen 4.600 Homogra-

Tab. 9

Homographen-Klassen:		
Insgesamt:	gleiche WK:	verschiedene WK:
305 Subklassen	28 Subklassen	277 Subklassen

Tab. 10

Insgesamt:	H-Reihen (GF-Paare):	WF/H-Reihe:
14.000 WF (token)	5.700	2,5
gleiche WK:		
5.600 WF (token)	1.100	5,5
verschiedene WK:		
8.400 WF (token)	4.600	1,8

phen-Reihen liegen Paarungen mit verschiedenen Wortklassen zugrunde. Die Verhältnisse bei den 14.000 Wortformen (token): etwa 5.600 homographe Wortformen lassen sich Paarungen mit gleicher Wortklasse, ca. 8.400 Homographe lassen sich Paarungen mit verschiedenen Wortklassen zuordnen.

5.4.1 Schlußfolgerungen

Die Beachtung der **Groß-Klein-Schreibung** des Wortanfangs reduziert die Fälle von Homographie in folgendem Ausmaß:

Tab. 11

Von 305 Subklassen	auf 176 Subklassen,
von 14.000 Wortformen (token)	auf 7.000 Wortformen (token),
von 5.700 H-Reihen	auf 1.800 H-Reihen.

Von diesem Rest an Problemfällen könnten durch Ermittlung der **Wortklasse** (über Distributionsanalysen oder stochastische Verfahren) weitere Homographien reduziert werden

Tab. 12

auf 28 Subklassen,
auf 5.600 Wortformen (token),
auf 1.110 H-Reihen.

Dieser Rest dürfte dann den harten Kern der Homographien ausmachen, dem weder mit Groß-Klein-Schreibung noch mit Wortklassen-Bestimmung beizukommen ist.

5.5 Weitere Beobachtungen zur Häufigkeit von Homographien

Es folgen einige Beobachtungen über Zusammenhänge zwischen weiteren Wortformen-Eigenschaften und der Häufigkeit von Homographien. Wir können uns dabei auf eine Reihe von Forschungsergebnissen der Quantitativen Linguistik stützen, bewegen uns aber zu einem großen Teil im Rahmen von Vermutungen, die detaillierte Untersuchungen nicht ersetzen können.

5.5.1 Häufigkeiten und Wortlaut (Wort-Länge und Graphie)

Der Anteil von Wortformen mit geringer Wortlänge dürfte bei Homographen höher sein als im Gesamtvokabular. Dieser Annahme liegt die Beobachtung

zugrunde, daß Homographe mit längeren Flexiven relativ selten sind (Beispiele dafür: ern, ten, test, -etest; siehe 5.5.3). Die überwiegende Zahl der homographen Flexionsformen weist Flexive mit einem oder zwei Buchstaben Länge auf (z. B. e, t bzw. en, te, st). Der Zusammenhang von Wort-Länge und Zahl der Bedeutungen gehört zum gesicherten Wissen in der Quantitativen Linguistik. Eine Formulierung dieser Gesetzmäßigkeit und ihre Erklärung im Rahmen eines Modells der linguistischen Synergetik liefert Köhler 1986: 57–62.

Im Falle von Homographie gilt: Kurze Wortformen weisen mehr Lesarten auf als längere. Ein Blick auf die Zusammenstellung der Homographien mit mehr als zwei Lesarten (Hilfsverzeichnisse, Abschnitt 4.6) zeigt, daß bei höherer Zahl der Lesarten der Anteil einsilbiger Wortformen ebenfalls höher liegt (die Silbenzahl ist dabei als Behelfsmaß für die Wortformen-Längen verwendet):

Tab. 13

3 Lesarten: 18%	4 Lesarten: 28%
5 Lesarten: 39%	6 Lesarten: 75%

Zur Graphie der Wortformen allgemein ist festzustellen: Die relativen Häufigkeiten der Anfangsbuchstaben beim Vokabular des Gesamt-Wörterbuchs dürften sich auch bei den Homographen in ähnlicher Weise wiederfinden; so sind Homographien mit Anfangsbuchstaben A oder B häufiger vertreten als solche mit O oder P. Homographe, die mit X oder Y beginnen, fehlen ganz. Allerdings sind bei Homographen die Anfangsbuchstaben D und U relativ stärker vertreten als im Gesamtwortschatz, was auf die große Zahl von Verb/Verb-Homographien mit ‚durch‘-, ‚über‘-, ‚um‘- und ‚unter‘-Präfigierung zurückzuführen ist.

5.5.2 Homographie und Wortklassen

Die absoluten Häufigkeiten bei Homographien sind proportional zum Umfang der beteiligten Wortklassen; d. h. am häufigsten ist die Wortklasse Substantiv vertreten, gefolgt von Verb und Adjektiv. Weist eine Wortklasse nur eine geringe Zahl von Wortformen auf, ist die relative Häufigkeit von Homographen und Lesarten höher. So sind Elemente von Funktionswortklassen im Durchschnitt häufiger homograph und weisen mehr Lesarten auf als die Elemente der Haupt-Wortklassen.

5.5.3 Homographie und Wort-Struktur

Weitere Tendenzen lassen sich im Hinblick auf die morphologische Struktur homographischer Wortformen beobachten (siehe dazu Weber 1974: 114): Homographien, die eine höhere Konstituenten-Anzahl aufweisen, sind seltener als solche mit niedriger Anzahl von Konstituenten:

Wortformen wie *drucke* – dargestellt als *druck-e* (Sm) / *druck-e* (V) – sind häufiger an Homographien beteiligt als Wortformen wie *ausdrucke* – dargestellt als *aus-druck-e* (Sm) / *aus-druck-e* (V).

Außerdem: Homographien, bei denen sich die Morphemgrenzen in beiden Lesarten decken, sind häufiger als solche mit sich überschneidenden Morphemgrenzen.

Wortformen wie *ausdrucke* – dargestellt als *aus-druck-e* (Sm) / *aus-druck-e* (V) – sind häufiger an Homographien beteiligt als Wortformen wie *montage* – dargestellt als *mon-tag-e* (Sm) / *mont-age* (Sf).

Das dürfte auch ein Grund sein für das häufigere Vorkommen von Homographien, die gleichzeitig Homophonien sind, gegenüber heterophonen Homographien.

In einigen Beobachtungen zur Häufigkeit von Homographien spielt die Schriftfolge GD eine wichtige Rolle. GD drückt die Differenz zweier Grundformen in einer Homographen-Reihe aus (siehe 3.4.1). Stimmt GD mit einem Wortbildungs-Morph (Derivativ) überein und gleichzeitig mit einem zulässigen Flexiv, dann besteht eine höhere Wahrscheinlichkeit zur Homographie, als wenn GD nicht mit einem Derivativ übereinstimmt. Tabelle 14 stellt einige Häufigkeitsangaben bei Klassen mit schriftlicher Übereinstimmung von GD mit Derivativ und Flexiv den Fällen gegenüber, in denen GD keinem Derivativ entspricht (Zahlen meist aufgerundet):

Tab. 14

GD	Beispiele:	H-Reihen:	Wortformen:
Übereinstimmung von Derivativ und Flexiv:			
GD = -e:	red (V) / rede (Sf)	1200	2100
GD = -en:	brat (V) / braten (Sm)	300	500
GD = -t:	fahr (V) / fahrt (Sf)	130	250
Keine Übereinstimmung von Derivativ und Flexiv:			
GD = -em:	tot (A) / totem (Sn)	2	2
GD = -se:	bus (Sm) / busse (Sf)	1	2
GD = -ten:	gar (V) / garten (Sm)	4	4

Homographen-Reihen und Homographie sind häufiger, wenn GD kürzer ist, d. h. wenn sich die aufeinander bezogenen Grundformen weniger voneinander unterscheiden.

Nehmen wir den Extremfall GD = -#, in dem ein Null-Derivativ (Ableitungsmorph) bei Grundform 1 mit einem Null-Flexiv zu Grundform

2 übereinstimmt: druck (Sm)/druck (V). Dieser Homographie-Typ ist der bei weitem häufigste.

Auch für den Fall GD < > # läßt sich die Verteilung der Homographien auf die verschiedenen Klassen in diesem Zusammenhang sehen. Tabelle 15 zeigt, daß Klassen, in denen GD nur **einen** Buchstaben lang ist, im Durchschnitt mehr Belege aufweisen als die anderen.

Tab. 15

GD	Beispiele:	H-Reihen:	Wortformen:
GD = -e-:	pflug (Sm) / pflug (V)	130	380
GD = -e:	red (V) / rede (Sf)	1200	2100
GD = -t:	fahr (V) / fahrt (Sf)	130	250
GD = ge-:	biss (V) / gebiss (Sn)	40	110
GD = -en:	brat (V) / braten (Sm)	300	500
GD = -er:	mild (A) / milder (V)	80	150
GD = ge-e:	trieb (V) / getriebe (Sn)	3	3
GD = -e-er:	blatt (Sn) / blaetter (V)	25	58
GD = -e-ern:	holz (Sn) / hoelzern (A)	3	3

Es gibt allerdings Ausnahmen. Zwar gilt der erwähnte Zusammenhang von Wort-Länge und Zahl der Bedeutungen nicht nur für lexikalische Einheiten, sondern auch für grammatische Konstituenten, wie Derivative oder Flexive: Vgl. etwa das funktionale Potential von -e (Derivativ und Flexiv in allen Hauptwortklassen) gegenüber dem von -eten oder -etest (Flexive nur für Verben) oder -esten (Flexiv für Adjektive). Doch fällt -en in dieser Hinsicht aus dem Rahmen. Die relativ hohe Zahl von Belegen hängt vermutlich mit seinem funktionalen Potential zusammen, das sich auf alle Hauptwortklassen erstreckt. So könnten im Gegensatz dazu die Belegzahlen für GD = -t, die unter denen für GD = -en liegen, auf das geringere Potential von -t zurückzuführen sein, das sich hauptsächlich auf die Wortklasse Verb beschränkt.

Bei den genannten Zusammenhängen sind bewußt „inhaltsneutrale“ Faktoren, wie Wort-Länge oder Konstituenten-Zahl in Betracht gezogen worden. Ergänzend dazu sei aber auch darauf hingewiesen, daß in vielen Fällen von Homographie, in denen Derivative einerseits und Flexive andererseits übereinstimmen, zwischen den beteiligten Grundformen inhaltliche Verwandtschaft angenommen werden kann; in Tabelle 16, erste Spalte sind einige dieser Konstellationen aufgeführt. Zufälle und Kuriositäten, wie sie in der zweiten Spalte gegenüberstehen, sind relativ selten (schätzungsweise weniger als 10%). Es ist daher zu vermuten, daß viele Homographien dieser Art tieferliegende systematische Züge aufweisen.

So wären wir zum Schluß wieder bei den eingangs behandelten Problemen der Beschreibung von Mehrdeutigkeiten, insbesondere bei der Unterscheidung Homonymie-Polysemie, angekommen.

Tab. 16

Inhaltlich verwandte GF	nicht verwandte GF
red (V) / rede (Sf)	sehn (V) / sehne (Sf)
fahr (V) / fahrt (Sf)	fluch (V) / flucht (Sf)
mild (A) / milder (V)	scheu (A) / scheuer (V)
knot (V) / knoten (Sm)	back (V) / backen (Sm)

Fassen wir die oben aufgeführten Beobachtungen zusammen: Die Wahrscheinlichkeit, daß eine Wortform an einer Homographie beteiligt ist, hängt ab

1. von der **Länge** der Wortform insgesamt,
2. von **Anzahl und Graphie** der darin vorkommenden **Morpheme**
3. vom **semantischen Potential** der Kern-Morpheme.

Im einzelnen heißt dies: Die Wahrscheinlichkeit einer Beteiligung an Homographien nimmt zu, wenn eine Wortform

- eine (unter)durchschnittliche Länge aufweist;
- morphologisch einfach ist (z. B. Grundform oder Grundform + Flexiv) – also keine komplexe Konstruktion vorliegt, wie Derivation oder Komposition;
- wenn sie homographische Konstituenten aufweist, die ihrerseits wieder eine geringe Länge aufweisen; dies gilt einmal für die Übereinstimmung zwischen Grundformen, aber auch wenn Flexive homographisch sind mit Derivativen, wie -#-, -e-, -en-, -er-, -t (dagegen spielt der Fall, daß Flexive oder Derivative homographisch sind mit Kern-Morphemen, wie -test-, -end-, bzw. -bar-, -ei-, -ion-, -sam-, nur eine unbedeutende Rolle);
- wenn ihre semantische Basis (das sind bei Flexionsformen die Grundformen, bei Derivationen und Kompositionen die Kern-Morpheme) in mehreren Konstruktionen vorkommt (einer sogenannten ‚Wortfamilie‘ angehört, wie z. B. ‚kühl‘ in ‚kühl‘, ‚kühlen‘, ‚Kühle‘, ‚Kühler‘ etc.).

Umgekehrt nimmt die Wahrscheinlichkeit ab, wenn eine Wortform

- überdurchschnittlich lang ist
- morphologisch komplexer ist (z. B. Derivation und/oder Komposition in Verbindung mit Flexion)
- keine homographischen Konstituenten aufweist oder die Konstituenten überdurchschnittlich lang sind;
- wenn ihre semantische Basis keiner oder nur einer kleinen Wortfamilie angehört.

6 Literaturhinweise

- Althaus, Hans Peter (1980).** Graphemik. In: Althaus, Hans Peter/Helmut Henne/Herbert Ernst Wiegand (eds.) (1973). *Lexikon der germanistischen Linguistik* (2. Auflage), 142–151. Tübingen: Niemeyer
- Augst, Gerhard (1975).** *Lexikon zur Wortbildung*. 3 Bde. Tübingen: Narr
- Bergmann, Rolf (1977).** Homonymie und Polysemie in Semantik und Lexikographie. *Sprachwissenschaft* 2/1977, 27–60
- Bierwisch, Manfred (1973).** Lautstruktur und Schriftstruktur. *Studia Grammatica* XI
- Coulmas, Florian (1981).** *Über Schrift*. Frankfurt a. M.: Suhrkamp
- Deutsche Rechtschreibung. Vorschläge zu ihrer Neuregelung (1992).** Herausgegeben vom Internationalen Arbeitskreis für Orthographie. Tübingen: Narr
- Ducháček, Otto (1962).** L'homonymie et la polysémie. *Vox Romanica* 21/1962, 49–56
- Eggers, Hans (1966).** Untersuchungen zum Homographen ‚sein‘. In: *Syntax und Datenverarbeitung* 2 (Forschungsbericht 8 der Deutschen Forschungsgemeinschaft), 140–155. Wiesbaden: Franz Steiner
- Eggers, Hans (1969).** Das Vorhaben ‚Elektronische Syntaxanalyse‘. In: Hans Eggers und Mitarbeiter (1969). *Elektronische Syntaxanalyse der deutschen Gegenwartssprache*, 1–19. Tübingen: Niemeyer
- Eisenberg, Peter (1983).** Orthografie und Schriftsystem. In: Günther, Klaus B./Günther, Hartmut (eds.) (1983). *Schrift, Schreiben, Schriftlichkeit. Arbeiten zur Struktur, Funktion und Entwicklung schriftlicher Sprache*, 41–68. Tübingen: Niemeyer
- Fries, Norbert (1980).** *Ambiguität und Vagheit. Einführung und kommentierte Bibliographie*. Tübingen: Niemeyer
- Gallmann, Peter (1986).** *The Graphic Elements of German Written Language*. In: Augst, Gerhard (ed.) (1986). *New Trends in Graphemics and Orthography*, 43–79. Berlin, New York: de Gruyter
- Harras, Gisela (1977).** Überlegungen zur Anlage von Artikeln in einsprachigen Wörterbüchern. *Germanistische Linguistik* 5–6/1977, 153–188

- Harweg, Roland (1973).** Phonematik und Graphematik. In: Koch, Walter A. (ed.) (1973). *Perspektiven der Linguistik*, Band 1, 37–64. Stuttgart: Kröner
- Hausmann, Franz Josef (1990a).** Les homonymes et les paronymes dans le dictionnaire monolingue. In: Hausmann, Franz Josef/Oskar Reichmann/Herbert Ernst Wiegand/Ladislav Zgusta (eds.) (1990). *Wörterbücher. Dictionnaires. Dictionnaires. Ein internationales Handbuch zur Lexikographie*. Zweiter Teilband, 640–641. Berlin, New York: de Gruyter
- Hausmann, Franz Josef (1990b).** Das Wörterbuch der Homonyme, Homophone und Paronyme. In: Hausmann, Franz Josef/Oskar Reichmann/Herbert Ernst Wiegand/Ladislav Zgusta (eds.) (1990). *Wörterbücher. Dictionnaires. Dictionnaires. Ein internationales Handbuch zur Lexikographie*. Zweiter Teilband, 1120–1125. Berlin, New York: de Gruyter
- Heger, Klaus (1963).** Homographie, Homonymie und Polysemie. *Zeitschrift für Romanische Philologie* 79/1963, 471–491
- Heller, Klaus (1994).** *Rechtschreibreform*. Sprachreport (Extra-Ausgabe) Dezember 1994. Mannheim: Institut für deutsche Sprache
- Henne, Helmut/Helmut Rehbock (1980).** Sprachzeichenkonstitution. In: Althaus, Hans Peter/Helmut Henne/Herbert Ernst Wiegand (eds.) (1980). *Lexikon der germanistischen Linguistik* (2. Auflage), 151–159. Tübingen: Niemeyer
- Heringer, Hans Jürgen (1981).** Die Unentscheidbarkeit der Ambiguität. In: Wolf Dietrich/Horst Geckeler (eds.) (1981). *Logos semantikos: studia linguistica in honorem Eugenio Coseriu 1921–1981*. Vol. III. *Semantik*, 93–126. Berlin, New York: de Gruyter und Madrid: Gredos
- Klappenbach, Ruth (1971).** Homonymie oder polysemes Wort? *Deutsch als Fremdsprache* 8/1971, 99–104
- Koch, Walter A. (1963).** Zur Homonymie und Synonymie. Eine kritische Zusammenfassung. *Acta Linguistica Acad. Scient. Hung.* XIII, 65–91
- Köhler, Reinhard (1986).** *Zur linguistischen Synergetik: Struktur und Dynamik der Lexik*. Bochum: Brockmeyer
- Kooij, Jan G. (1971).** *Ambiguity in Natural Language*. Amsterdam, London: North Holland
- Lyons, John (1992).** *Die Sprache*. (4. Auflage). München: C.H. Beck
- Messelaar, Petrus Adrianus (1985).** Polysémie et homonymie chez les lexicographes. Plaidoyer pour plus de systématisation. *Cahiers de Lexicologie* 46/1985, 45–56
- Nerius, Dieter (1980).** Zu orthographischen Problemen der Eigennamen im Deutschen. *Zeitschrift für Phonetik, Sprachwissenschaft und Kommunikationsforschung* 33 (1980), 93–102
- Öhmann, Emil (1959).** Über die Wirkung der Homonymie im Deutschen. *Wirrendes Wort* 9/1959, 2–12

- Pacak, Milos (1963).** Homographs: Classification and Identification, Part I. *Vistas in Information Handling*, Vol 1. Chapter 6. Washington: Spartan Books
- Pacak, Milos (1967).** Homographs: Classification and Identification, Part II. *Etudes de Linguistique Appliquée* 5/1967, 89–105
- Piirainen, Ilpo Tapani (1986).** Die Autonomie der Graphematik in historischer Sicht. In: Augst, Gerhard (ed.) (1986). *New Trends in Graphemics and Orthography*, 97–104. Berlin, New York: de Gruyter
- Pinkal, Manfred (1985).** Logik und Lexikon – Die Semantik des Unbestimmten. Berlin, New York: de Gruyter
- Schaeder, Burkhard (1987).** Germanistische Lexikographie. Tübingen: Niemeyer
- Scharnhorst, Jürgen (1969).** Zur semantischen Struktur des Wortschatzes der deutschen Gegenwartssprache. *Zeitschrift für Phonetik, Sprachwissenschaft und Kommunikationsforschung* 22/1969, 502–517
- Schelp, Hanspeter (1960).** Homographie im heutigen Deutsch. Die neueren Sprachen (NF) 8/1960, 388–396
- Schildt, Joachim (1969).** Gedanken zum Problem Homonymie – Polysemie in synchronischer Sicht. *Zeitschrift für Phonetik, Sprachwissenschaft und Kommunikationsforschung* 22/1969, 352–359
- Ströbl, Alex (1971).** Polysemie-Homonymie (Kurzfassung). In: von Stechow, Arnim (ed.) (1971). *Beiträge zur generativen Grammatik. Referate des 5. Linguistischen Kolloquiums Regensburg 1970*, 258–263. Braunschweig: Vieweg
- Trnka, Bohumil (1931).** Bemerkungen zur Homonymie. *Travaux du Cercle Linguistique de Prague* 4/1931, 152–156
- Ullmann, Stephen (1959).** Grundzüge der Semantik. Berlin: de Gruyter
- Vachek, Josef (1939).** Zum Problem der geschriebenen Sprache. *Travaux du Cercle Linguistique de Prague* 8/1939, 94–104
- Vachek, Josef (1948).** Written language and printed language. *Recueil Linguistique de Bratislava* 1/1948, 67–75
- Weber, Heinz J./Volker Wolf (unter Mitwirkung von Rainer Rath) (1966).** Zur Homographie in der deutschen Gegenwartssprache. Eine morphologische Untersuchung. *Linguistische Arbeiten Saarbrücken* Nr. 3
- Weber, Heinz J. (1969).** Die Bestimmung der Wortklassen. In: Hans Eggers und Mitarbeiter (1969). *Elektronische Syntaxanalyse der deutschen Gegenwartssprache*, 62–89. Tübingen: Niemeyer
- Weber, Heinz J./Harald Zimmermann (1973).** Zur Verwertbarkeit der Großschreibung bei der automatischen Reduktion syntaktischer Wortformen-

- Mehrdeutigkeiten im Deutschen. In: Festgabe für Paul Grebe, Bd. 2, 238–261. Düsseldorf: Schwann
- Weber, Heinz J. (1974).** Mehrdeutige Wortformen im heutigen Deutsch. Studien zu ihrer grammatischen Beschreibung und lexikographischen Erfassung. Tübingen: Niemeyer
- Weber, Heinz J. (1976).** Verfahrensweisen bei der automatischen Auffindung von Wortformen-Mehrdeutigkeiten im Text und im Lexikon. Zeitschrift für Phonetik, Sprachwissenschaft und Kommunikationsforschung 3 (1976), 264–274
- Weydt, Harald (1975).** Mehrfachverständnis sprachlicher und nichtsprachlicher Zeichen. Systematische Analyse der Ambiguität. In: Drachman, Gaberell (ed.) (1975). Akten der 1. Salzburger Frühlingstagung für Linguistik, 239–254. Tübingen: Narr
- Wichter, Sigurd (1988).** Signifikantgleiche Zeichen. Tübingen: Narr
- Wiegand, Herbert Ernst (1985).** Fragen zur Grammatik in Wörterbuchbenutzungsprotokollen. Ein Beitrag zur empirischen Erforschung der Benutzung einsprachiger Wörterbücher. In: Bergenholtz, Henning /Mugdan, Joachim/Wiegand, Herbert Ernst (eds.) (1985). Lexikographie und Grammatik. Akten des Essener Kolloquiums zur Grammatik im Wörterbuch, 28.-30.6.1984, 20–98. Tübingen: Niemeyer
- Zaunmüller, Wolfram (1958).** Bibliographisches Handbuch der Sprachwörterbücher. Stuttgart: Hiersemann
- Zemb, Jean M. (1984).** Vergleichende Grammatik Französisch-Deutsch. Teil 2. L'économie de la langue et le jeu de la parole. Mannheim, Wien, Zürich: Dudenverlag
- Zöfgen, Ekkehard (1976).** Polysemie oder Homonymie? Zur Relevanz und Problematik ihrer Unterscheidung in Lexikographie und Textlinguistik. In: Barrera-Vidal, Alberto/Ernstpeter Ruhe/Peter Schnuck (eds.) (1976). Lebendige Romania. Festschrift für Hans-Wilhelm Klein, 425–464. Göppinger Akademische Beiträge
- Zöfgen, Ekkehard (1989).** Homonymie und Polysemie im allgemeinen einsprachigen Wörterbuch. In: Hausmann, Franz Josef/Oskar Reichmann/Herbert Ernst Wiegand/Ladislav Zgusta (eds.) (1989). Wörterbücher. Dictionaries. Dictionnaires. Ein internationales Handbuch zur Lexikographie. Erster Teilband, 779–787. Berlin, New York: de Gruyter

7 Konsultierte Wörterbücher

Bahlow, Hans (1980). Deutsches Namenlexikon. München: Keyserische Verlagsbuchhandlung

Duden: Deutsches Universalwörterbuch (1989). Mannheim, Wien, Zürich: Dudenverlag

Duden Band 1: Rechtschreibung der deutschen Sprache (1991). Mannheim, Wien, Zürich: Dudenverlag

Klappenbach, Ruth/Wolfgang Steinitz (eds.) (1961–1977). Wörterbuch der deutschen Gegenwartssprache. Berlin: Akademie-Verlag

Wahrig, Gerhard (1991). Deutsches Wörterbuch (Jubiläumsausgabe). Gütersloh, München: Bertelsmann Lexikon-Verlag

2. Teil

Die Einzelverzeichnisse des HWB