

Lemmatisierungspraxis und -problematik im Autorenwörterbuch am Beispiel des *Goethe-Wörterbuchs*

Christiane Schlaps

Thomas Schares

Akademie der Wissenschaften zu Göttingen

The macrostructure of an author's dictionary as determined by (the rules of) lemmatization so far has attracted little attention in practical and theoretical lexicography, with most publications on the topic covering small-sized dictionaries that represent only specific segments of a writer's works and vocabulary. The Goethe Dictionary (GWb), in contrast, endeavors to treat the complete vocabulary used by Johann Wolfgang Goethe in his prolific literary, scientific, philosophical, historiographic etc. writings as well as in his letters and, to some extent, his conversations. The word list contains over 90,000 headwords, thus making the project the largest dictionary on an author's idiolect worldwide. The form and placement of lemmas in the GWb is determined by a set of rules that over the decades of work on this dictionary have been collected in a style manual but which are far from comprehensive or even static. Examples from the published parts of the GWb will demonstrate a number of decisions that typically arise in author's dictionaries and the way they are solved in the GWb, including questions of an author's idiolectal aberration from the orthographic norm, the frequency of hapax legomena, and morphological idiosyncrasies that in the GWb led to specialized sublemmas. In our paper, we would like to emphasize the need for research, both practical and theoretical, into the special problems of lemma presentation in author-centred lexicography and its methodological foundation as we believe that to show the full range of an individual's lexicon will not only contribute to the field of metalexicography in particular but will also help to gain valuable insights into the lexicological, semantic, grammatical, and, incidentally, historical dimensions of language in general.

1. Einleitung

Das vorliegende Papier beschäftigt sich mit einigen Problemen und Prinzipien der Makrostrukturorganisation von Autorenwörterbüchern, indem es aus der Praxis heraus am Beispiel des Goethe-Wörterbuchs (GWb)¹ darstellt, wie dort einige grundsätzliche Lemmatisierungsfragen, die sich aus dem Wörterbuchtyp ergeben, behandelt werden. Untersuchungen zu Lemmatisierungsproblemen in umfassenderen Individualwörterbüchern, sei es methodischer Art, sei es als Werkstattberichte aus einzelnen Projekten, scheinen bisher nicht vorzuliegen, so daß unsere Überlegungen als ein erster Beitrag zum Thema gelten können.²

¹ Einführend zum GWb s. etwa Niedermeier/Objartel/Welter (2001). Das GWb publiziert seit 1966 in Einzellieferungen; inzwischen sind vier Bände erschienen und die ersten fünf Lieferungen des fünften Bandes bis zum Stichwort *Körper*. Die Online-Version des GWb ist im Internet unter www.goethe-woerterbuch.de unentgeltlich zugänglich, z.Zt. (Stand: März 2008) bis zum Lemmazeichen *inhaftieren*. An der Erstellung sind Lexikographen und Lexikographinnen aus Arbeitsstellen in Berlin, Hamburg, Leipzig und Tübingen beteiligt; die aktuelle Artikelarbeit ist inzwischen beim Buchstaben M angelangt.

² Auskunft von H. E. Wiegand im April 2007 aus seiner Arbeit an der im Druck befindlichen *Internationalen Bibliographie zur germanistischen Lexikographie und Wörterbuchforschung* (Wiegand 2006 ff.). Zur wörterbuchtypbezogenen Lemmaauswahl, nicht aber zur Lemmaansetzung (mit einem kurzen Überblick über die neuere germanistische Literatur zur Frage) vgl. zuletzt Brückner (2007). Der Beitrag von Arnold (2006) behandelt de facto nicht die Entstehung eines Autoren-, sondern vielmehr eines Werkwörterbuchs und bringt keine neuen Erkenntnisse zu Lemmatisierungsfragen (vgl. dort S. 8).

Allgemein läßt sich feststellen, daß Individualwörterbücher als Spezialfälle wissenschaftlicher Lexikographie insbesondere in der Germanistik kein häufig realisierter Wörterbuchtyp sind; besonders rar sind dabei Autorenwörterbücher, die an einem hinreichend großen Corpus nach Thesaurusprinzip verfahren. So gibt es zur Zeit kein einziges fertiggestelltes Wörterbuch zur Gesamtlexis eines deutsch(sprachig)en Autors, was bedeutet, daß zu Anlage und Ausführung eines solchen Werkes nicht viele Vorbilder und damit verbundene metalexikographische Überlegungen verfügbar sind.³ Gerade aus derartigen Projekten heraus, d.h. wenn nicht nur sog. „Grund- und Wesenswörter“ (nach Schadewaldt 1978: 4* und 7*) behandelt werden, sondern tatsächlich der gesamte, hoch- wie niedrigfrequente, semantisch auffällige wie auch (zunächst) weniger ergiebige Wortschatz eines Autors gleichermaßen erforscht wird, sind aber für die wissenschaftlich interessierten Benutzer neue Erkenntnisse sowohl allgemeinerer linguistischer Art, sei es lexikologischer oder morphologischer Ausrichtung, wie auch metalexikographischer Natur zu erwarten.

2. Lemmatisierung im GWb

2.1. Stichwortaufnahme

Das GWb ist ein historisches, semasiologisches Autorenbelegwörterbuch, das lemmatische Vollständigkeit anstrebt, d.h. es verzeichnet vollständig alle im Archiv⁴ belegten lexikalischen Einheiten, also etwa auch Funktionswörter wie Partikeln, Präpositionen, Konjunktionen und Interjektionen, um „das ganze *Panorama der Goetheschen Sprache* zu entfalten“ (Schadewaldt 1978: 7*). Ausnahmen betreffen lediglich Numeralia ohne besonderes semantisches (z.B. phraseologisches) Interesse, rein identifizierend verwendete Eigennamen ohne zusätzliche semantische Funktion, fremdsprachliche lexikalische Einheiten, die nur in Exzerpten, Listen oder Zitaten aus Fremdtexten überliefert sind, sowie nicht-Goethische Werkstitel. Das Wörterbuch beinhaltet also alle Wortbildungen (Komposita und Derivate) Goethes, die vielfältigen (fachlichen) Teilwortschätze, diejenigen fremdsprachlichen Einheiten, die von Goethe in seinen eigenen Schriften genuin verwendet und nicht aus anderen Texten adaptiert wurden sowie Figurennamen, mythologische Namen und historische Namen in charakteristischer (metonymischer) Funktion. Spezialsprachliche Teilwortschätze aus verschiedenen Fach- und Wissenschaftssprachen und naturwissenschaftlichen Nomenklaturen, in der beruflichen Tätigkeit Goethes auch juristische und amtssprachliche Diktion sowie handwerklich-technische Ausdrücke z.B. aus der Bergmannssprache führen zu einer bemerkenswerten lexikalischen Ausweitung: Damit „ist die normale Sprachkompetenz eines Individuums weit überschritten“ (Unterberger 1985: 151). Gleichzeitig ergibt sich aus diesem Befund auf der Ebene der Semantik eine besondere Differenzierung der Sprechweise, was sich im GWb in einer stark hierarchisierenden Bedeutungsgliederung und —nuancierung widerspiegelt (worauf wir hier nicht weiter eingehen können). So wird das GWb nach aktueller Schätzung und laut unserer internen Wortliste bei seiner Vollendung etwa im Jahr 2025 einen Lemmabestand von rund 93.000 Einheiten enthalten. Die in der jahrzehntelangen konkreten

Für einen Überblick zum praktischen wie theoretischen Stand der Autorenlexikographie zum Zeitpunkt der konzeptionellen Anlage des GWb s. Dill (1959).

³ Zur Einordnung des GWb in die internationale Autorenlexikographie s. zuletzt Objartel (2005), wo für andere lexikographisch erfaßte Schriftstellerwortschätze Zahlen zwischen 12.000 (Milton, Cervantes) und noch unter 30.000 (Shakespeare, Ibsen) angegeben werden. Schillers Wortschatz wird auf etwa 30.000 Wörter geschätzt; genaueren Aufschluß soll einmal das noch nicht erschienene, seit einigen Jahren angekündigte Schiller-Wörterbuch geben. Zum Lemmabestand im GWb s. unten, II 1.

⁴ Archivgrundlage sind die Weimarer oder Sophien-Ausgabe (in 143 Bänden 1887 bis 1919 erschienen) als bislang einzige annähernd vollständige historisch-kritische Ausgabe von Goethes Werken sowie für bestimmte Werkbereiche (z.T. ergänzend) andere Ausgaben wie etwa im Teilbereich der naturwissenschaftlichen Schriften die Leopoldina-Ausgabe (1947 ff.), für die amtlichen Schriften die Ausgabe von Flach (1950-1978) und für die Tagebücher die neue historisch-kritische Ausgabe von Golz (1998 ff.); vgl. dazu auch Nutt-Kofoth (2005). Das Archiv hat mit seinen rund 3,3 Mio. Belegen einen —nicht nur für ein Autorenwörterbuch— enormen Umfang.

Arbeit gesammelten Erfahrungen mit Fragen der Lemmatisierung im GWb wurden ab 1999 in einem internen Regelwerk (gültige Fassung vom Juli 2005) dokumentiert, das zunächst alte Einzelregelungen (aus Protokollen und Briefwechseln auch über die Zeit der politischen Teilung der beiden deutschen Staaten hinweg) zusammenfaßte und das nun laufend durch die Diskussionsergebnisse der Mitarbeiterversammlungen aus den Arbeitsstellen ergänzt und überarbeitet wird.

Die konzeptionelle Entscheidung für den Thesaurusansatz des GWb, d.h. die Entscheidung, daß (so gut wie) allen in den Texten Goethes belegten linguistischen Einheiten im GWb Artikelstatus zukommen soll und also alle Lexeme durch entsprechende Lemmata repräsentiert werden müssen, ließ übliche selektionsbezogene Überlegungen bei der Konzeption allgemeiner Wörterbücher (zum Grundwortschatz einer Sprachepoche o.ä.) wie Erstellung einer Stichwortliste in Relation zu Bearbeitungsdauer, Werkumfang oder Ausgewogenheit im Verhältnis einzelner Buchstaben zum Gesamtumfang überflüssig werden. Als einzig gültiges Aufnahmekriterium gilt: Das Wort ist bei Goethe belegt. Goethe nicht (sicher) zuschreibbare Lexeme, etwa solche aus Gesprächsmitschriften, aus mit anderen Personen gemeinschaftlich verfaßten oder veröffentlichten Texten bzw. aus von anderen übernommenen Text(teil)en, werden dagegen mit einem Sternchen vor dem Lemma versehen, z.B. **Brezelverkäufer*, **Kegelbahnvergnügen* sowie **Lumpenkerl* bzw. **Kompositionsleiste*, **kopflös* oder **Lupercus*.

2.2. Lemmaansatz und Makrostruktur

Mit Lemmatisierung wird der Prozeß der Festlegung der Form des Lemmzeichens (orthographisch, morphologisch) und der Anordnung der Lemmzeichen innerhalb der Makrostruktur bezeichnet (Wiegand 1983: 446). Das GWb ist vom makrostrukturellen Typ her ein striktalphabetisches Wörterbuch mit Gruppierungen (im Sinne von Wiegand 1989): Simplizia und Komposita werden als Volllemmata aufgenommen; in speziellen Fällen kann jedoch eine Lemmaposition ein Hauptlemma und Nebenlemmata (Nischen- bzw. Nestartikel, vgl. unten II 3 e; s. auch Wolski 1989: 364 ff.) aufweisen.

Prinzipiell erfolgt der Lemmaansatz im GWb nach den üblichen Grundformen: Verben im Infinitiv, Substantive und Pronomina im genusspezifischen Nominativ Singular, Adjektive in der unflektierten starken Form des Nominativs Singular. Als orthographische Regelung gilt die Ansetzung jenseits aller schwankenden Rechtsschreibreformen nach der Duden-Norm *vor* 1998 (d.h. nach dem achtbändigen Duden von 1993-1995). Schwierigkeiten bei der Ansetzung einer kanonischen Form ergeben sich in der Praxis zum einen sprachepochenbezogen, zum anderen funktional (aus dem Wörterbuchtyp), da in der Autorenlexikographie als notwendig textbezogener Wörterbucharbeit immer aus der *parole* heraus (und keineswegs auf eine *langue* zu) lemmatisiert wird (hierzu ausführlich anhand eines mediävistischen Beispiels Reichmann 1996).

2.3. Lemmatisierungspraxis im GWb — Beispiele

2.3.1. Orthographische Normabweichungen und Hapax Legomena

Verwendungen, die bei Goethe durchgehend oder (bei höher frequenten Wörtern) fast ausschließlich abweichend von der oben genannten Duden-Norm belegt sind, werden in der autortypischen Nennform lemmatisiert, so etwa bei ¹*bevorworten* (statt *befürworten*) oder bei *Hülfe* und zugehörigen Wortbildungen (statt *Hilfe* etc.). Bezeichnend für die Ansetzungsentscheidungen und das damit einhergehende System alphabetisierungsbedingter Verweise in einem Autorenwörterbuch ist dabei der hohe Anteil von gering belegten Lemmata oder gar Hapax legomena. So zeigt die Arbeit am und der Blick ins GWb, daß etwa zwei Drittel von Goethes Lexik aus Gelegenheitswortschatz besteht und dabei über ein Drittel aus reinen Einmalwörtern, d.h. über einen Zeitraum von rund 65 Jahren Schreibtätigkeit extrem selten vorkommt. Je nach Fall (Belegmenge oder auch Herkunft der Belege aus einer bestimmten Textsorte) kann die Lemmatisierungsentscheidung gerade bei diesen Lemmata schwierig sein. Beispiele für entsprechende (Verweis-)Ansetzungen sind etwa *Jütland* s. *Juitland* (als Hapax legomenon), *Kader* s. *Cadre* (mit drei Briefbelegen), *Cacadou* (mit zwei Tagebuchbelegen) oder die Form *Löwenstain* als onomastisches Hapax legomenon.

2.3.2. Eigennamen

Auch in der Frage der Aufnahme von Eigennamen kann das Autorenwörterbuch als ein Spezialfall gelten, denn in ihm wird im Gegensatz zu den verschiedenen Formen des *Langue*-Wörterbuchs der onomastische Wortschatz als konstitutiver Teil des behandelten Idiolektivs und der auch enzyklopädischen Aufgabe von Individuallexikographie allgemein begriffen. Grundsätzlich werden Eigennamen im GWb aufgenommen, wenn sie nicht lediglich rein (individuell) identifizierend gebraucht werden; dies betrifft Anthroponyme, d.h. Personen(gruppen)namen, ebenso wie Toponyme, wobei die Behandlung der Örtlichkeitsnamen im GWb z.Zt. nicht einheitlich und abschließend geklärt ist.

Eine Besonderheit sind die Figurennamen aus dem literarischen Werk Goethes; sie können gelegentlich gleichzeitig als Werksnamen fungieren (*Faustus*, *Götz* u.a.). Bezüglich der Aufnahme nicht genuin Goethescher Figurennamen hat sich die Aufnahmepraxis im Verlauf der Erstellung des GWb gewandelt. Laut neuerer Regelung (im Zuge von Straffungsmaßnahmen 2005/2006) finden diese dann keine Aufnahme in das Wörterbuch mehr, wenn sie nicht von Goethe unter tatsächlicher produktiver Aneignung rezipiert wurden, d.h. in seinen fiktionalen Werken und nicht etwa nur als rein notizhafte Nennung etwa in Paralipomena vorkommen. Eine solche dichterische Anverwandlung ist beispielsweise bei der Vielfalt der Personen- und Ortsnamen aus dem *Westöstlichen Divan* gegeben. Verzeichnet werden ebenso die in Goethes *Ceuvre* entsprechend aufgenommenen mythologischen Namen sowie reale Eigennamen, wenn Goethe sie appellativisch gebraucht, etwa *Ali* im Sinne eines treuen Anhängers (nach dem loyalen Schwiegersohn Mohammeds von Goethe für Eckermann gebraucht) oder *Fauntleroy* als Name eines berüchtigten englischen Wechselfälschers, den Goethe in einem Gedicht-Titel zugleich als Prototyp eines Betrügers verwendet.

2.3.3. Doppellemmata und Klammerungen

Unterschiedliche orthographische Verwendungsweisen eines Wortes bei Goethe führen gelegentlich zu mehr oder weniger gleichberechtigt nebeneinander stehenden Variantenformen, deren normgerechte Lemmatisierung im Laufe der Bearbeitungsdauer des GWb wechselt; dies betrifft insbesondere die Ansetzung von Doppellemmata oder Klammerungen innerhalb eines Lemmas. Als Doppellemma gebucht sind z.B. *Fels*, *Felsen* mit einer Vorbemerkung, daß von rund 800 Belegen 620mal „Felsen“ belegt sei mit einer deutlichen Präferenz der Kurzform in der ersten angegebenen Bedeutung, sonst aber ohne semantische Unterscheidung; die erste Bedeutung ist nur singularisch belegt und wird als „kompaktes hartes Gestein, bes[onders] als fester Unter-, Baugrund“ definiert (GWb III: 649). Solche sprachgeschichtlich interessanteren Doppelformen und Klammerungen wie auch etwa bei *ahn(d)en* stehen im GWb neben eher zu vernachlässigenden wie bei *Kassegeld*, *Kassageld* oder *Kellertür(e)* bzw. *Gesumm(e)*, oder (in der Kombination von Doppellemma und Klammer): *Kehricht*, *Kehrig(t)*, wozu die Vorbemerkung des Artikels erläutert, daß alle drei Formen noch in der „Italienischen Reise“ nebeneinander vorkommen. Die zu diesem Beispiel gehörigen (zwei) Komposita werden allerdings nur mit *Kehricht*-lemmatisiert, obwohl bei *Kehrichthaufen* zwei von drei Belegen in der Form *Kehrigt*- vorliegen. Diese uneinheitliche Praxis wurde inzwischen vereinheitlicht: Während ein Protokoll von 1968 die synoptische Lemmaansetzung für Ausnahmefälle ausdrücklich vorsieht, allerdings auf Ausnahmefälle beschränkt wissen will, ist im Protokoll der letzten Mitarbeiterversammlung vom April 2007 geregelt, daß sowohl Klammerungen als auch Doppelformen künftig vermieden werden sollen und für die abweichenden Formen makrostrukturell entsprechende Verweisartikel und mikrostrukturell erläuternde orthographische Vorbemerkungen im Artikel vorgesehen sind.

2.3.4. Mehrwortlemmata

Doppellemmata und Klammerfügungen durchbrechen nicht formal den für das GWb geltenden Grundsatz des Einwort-Ansatzes. In unserer lexikographischen Praxis finden sich jedoch auch Mehrwort-Ansetzungen, etwa im Bereich der Präfixverben mit iterativer Komponente: so im vierten Band auf 2½ Spalten mit insgesamt 47 Lemmata von *Hin-und-her-Dissertieren* bis *Hin-und-wi(e)der-Zerren*. Der den Arbeitsstellen bereits vorliegende, bis zum letzten Wort (*Zyrus*)

reichende und detaillierte sog. Index verborum des GWb, d.h. die Gesamtliste der zu bearbeitenden Lexeme, verzeichnet eine solche Ballung jedoch als Ausnahme. Weitere Beispiele für Mehrwortlemmata sind nicht sinnvoll auflösbare (semantische) Einheiten wie z.B. *Anton Reiser* oder *Adels- und Hofnatur* und *Kreuz-und-Quer-Bewegung*, *Kreuz-und-Quer-Fahrt*, *Kreuz-und-Quer-Gang*, *Kreuz-und-Quer-Zug* u.ä.

Eine andersartige und ungleich häufigere Ausprägung Goethe-spezifischer Mehrwort-Ansetzungen, das Kopulativ-Kompositum, wird dagegen makrostrukturell auffällig als Nebenlemmata realisiert (s. folgender Abschnitt).

2.3.5. Nebenlemmata

Die Lemmaanordnung im GWb erscheint auf den ersten Blick glattalphabetisch, zumal die vielen Komposita(reihen) als Vollemmata gebucht werden, nicht als Nischenlemmata. Trotz des hohen Platzaufwandes läßt sich methodisch rechtfertigen, daß diese oft gering belegten bzw. okkasionell verwendeten Lexeme in einem Autorenwörterbuch vollgültig behandelt werden sollen, um die idiolektale Besonderheit eines Einzelwortschatzes (auch typographisch) deutlich werden zu lassen. Ferner ist diese Lösung der makrostrukturellen Anordnung im GWb u.E. auch im Hinblick auf den Benutzungskontext des Wörterbuchs angemessen: Der Rezipientenkreis gerade des GWb ist ausgesprochen heterogen; neben den linguistischen und literaturwissenschaftlichen Fachleuten soll es möglichst, wie im Verlagsprospekt von Kohlhammer formuliert, eine „Fundgrube für alle [sein], die [...] professionell mit Sprache zu tun haben oder einfach Liebhaber der Literatur und der Wörter sind: Schriftsteller, Übersetzer, Journalisten, Lektoren, Werbetexter, Politiker, Philosophen, Lehrer, Leser“.

Eine partielle Durchbrechung dieses benutzerfreundlichen, aber platzaufwendigen Ordnungsprinzips findet sich im GWb jedoch bei verschiedenen Arten von Nebenlemmata, von denen im folgenden einige typische Formen genannt werden sollen. Eine explizite Regelung für die Ansetzung eines gültigen Nebenlemmas gibt es für das GWb dabei nicht; in der Praxis wird der Status eines Nebenlemmas im Regelwerk verschiedenen semantischen oder morphologischen Kriterien zugeordnet.

Relativ häufig auftretende multilemmatische Artikel im GWb entstehen aus einer für den Autor Goethe typischen Sprachverwendung in der Reihung von Adjektiven, d.h. „Kopulativkomposita, die bei Goethe zum Teil anreihend ohne Bindestrich erscheinen und damit einen eigentümlichen Schwebezustand aufweisen“ (so in einem Rundschreiben der Berliner Arbeitsstelle 1986). Dabei können auch längere Reihen von Nebenlemmata entstehen (die hier in sich wiederum striktalphabetisch geordnet sind), beispielsweise bei den Hauptlemmata *historisch*, *lebendig* oder *literarisch*. Bei *historisch* zählt man 28 Nebenlemmata mit zweigliedrigen bis hin zu viergliedrigen (*historisch-polemisch-kritisch-metaphysisch*) Adjektivkomposita, außerdem Substantivierungen wie *Historisch-Positive*.

Hierbei handelt es sich nur um die Rechtserweiterungen; entsprechende Linkserweiterungen werden dagegen in einem Wortbildungsblock am Artikelende angegeben. Bei *historisch* finden sich weitere zehn Linkserweiterungen sowie eine Rechts- und Linkserweiterung *poetisch-historisch-kritisch-bildlich*. Zum Stichwort *lebendig* kommen zehn jeweils zweigliedrige Nebenlemmata hinzu, wobei diese Reihe zeigt, wie im GWb die unterschiedlichen Befunde der Textausgaben normiert werden: So erscheinen *lebendig- Augenblicklich*, *lebendig-beweglich* oder *lebendig-fertig* sowie *lebendig-nährend* und *lebendig-unförmlich* auch in der Weimarer Ausgabe mit Bindestrich, *lebendig-aufmerksam*, *lebendig-frisch*, *lebendig-heiter* dagegen dort ohne Bindestriche, während bei *lebendig-farbig* sowie *lebendig-praktisch* in der Edition Änderungen Goethes von Getrenntschreibung (einmal mit Kopula) zu Bindestrichschreibung dokumentiert sind. Hinzu kommen noch 14 Linkserweiterungen (alle doppelt). Bei *literarisch* wiederum folgen insgesamt 13 zwei- bis dreigliedrige Nebenlemmata und weitere 11 Linkserweiterungen, davon eine als Dreifachreihung. —Diskutabel scheint u.E., inwiefern durch diese Lemmatisierungsvorschrift möglicherweise eine Art von Übersystematisierung und damit sogar ein unangemessener Eindruck des dargestellten Autoren-Idiolekts erzielt wird.

Der semantische Bezug in solchen für Goethe typischen Adjektivreihungen kann im übrigen verschieden realisiert werden: bei *literarisch-ökonomisch* etwa differenzierend; dort geht es um eine Postsendung an Goethe, die aus einem Buch zur Göttinger Universität und aus Erbsensaatgut bestand. Bei *literarisch-poetisch* dagegen fungiert die Reihung, die hier Wielands dichterische Tätigkeit charakterisiert, verdoppelnd-verstärkend.

Eine weitere typische Form der Realisierung von Nebenlemmata im GWb betrifft fremdsprachige Entsprechungen zum Hauptlemma, z.B. *Justitiar* mit Sublemma *Iustitarius*, *Justiz* mit *iustitia*, *Kanon* mit *canon*, *Kalk* mit *chalk*. Dabei können die Nebenlemmata durchaus zu umfangreichen Artikeln werden, etwa beim Stichwort *Historia*, dessen sechs Hauptbedeutungen meist denjenigen beim Hauptlemma *Historie* zuzuordnen sind, aber zusätzlich auch eine partielle Bedeutungsverschiebung und eine Personifikation (Muse bzw. Allegorie der Geschichtsschreibung) buchen.

Darüberhinaus werden in gewissem Umfang etwa Substantivierungen von Adjektiven als Nebenlemmata realisiert, sofern sie auf Personen bezogen oder semantisch eigenständig sind, etwa *Gute* (als Maskulinum wie als Femininum; mit immerhin noch acht Hauptpunkten in diesem Nebenlemma sowie einer kompositionalen Rechtserweiterung *Gute-Schöne*) zu *gut* oder *Innere* zu *inner* (mit drei Hauptpunkten zu den entsprechenden Punkten des Hauptlemmas, die jeweils umfassendes Belegmaterial aufweisen; das Nebenlemma umfaßt mehr als 2½ Spalten).

Problematisch erscheinen die Regelungen, Partizipien, „die zum adjektivischen Charakter tendieren“, bzw. substantivierte Partizipien, die „zusätzliche Semantikmerkmale besitzen“, (so jeweils laut GWb-Regelwerk) als Nebenlemmata zu buchen. Dies führt in der Praxis etwa dazu, daß unter dem Hauptlemma *komplizieren* (worunter zwei Bedeutungen mit insgesamt drei Belegen auf knapp 5 Zeilen zu finden sind) ein Nebenlemma *kompliziert* anschließt, für das über 150 Belege in vier Hierarchiestufen mit insgesamt 14 (Teil-)Bedeutungen auf über einer Spalte aufgefächert wurden, bevor das Hapax legomenon *Kompliziertheit* dann wieder als Hauptlemma folgt, oder anderswo dazu, daß der *Gekerkerte* als Nebenlemma unter *kerkern* etwas untergeht, während der *Kennende* als eigenständiges Lemma (direkt nach *kennen*) behandelt wird.

Bei allen genannten Fällen der im GWb zulässigen Ansetzung von Nebenlemmata, insbesondere aber bei den Partizipien, ergibt sich die Frage, wie effizient der Zugriff auf diese Artikel in der Praxis bei den verschiedenen Nutzergruppen ist. In der Printversion läßt sich etwa gelegentlich durch die Zufälle des Zeilenfalls nur schwer im Satzspiegel erkennen, ob es sich um ein Nebenlemma oder doch um ein Vollemma handelt. (Ein sicheres Anzeichen für den Benutzer ist jedoch, daß etwa alle eventuellen Synonyme, in der Zuordnung aufgefächert nach Haupt- und Nebenlemma bzw. -lemmata, jeweils ganz am Ende des Gesamtartikels erscheinen.) Aus Gründen der Homogenität des Wörterbuchs ist es bei einem so weit im Alphabet fortgeschrittenen Projekt wie dem GWb zwar nicht sinnvoll, die makrostrukturelle Anordnung für den verbleibenden Teil des noch zu bearbeitenden Materials zu ändern, doch ist u.E. aus der am GWb gewonnenen Erfahrung die metalexikographische Forderung abzuleiten, wenigstens bei (der Neukonzeption von zukünftigen) Autorenwörterbüchern auf die Ansetzung von Nebenlemmata zu verzichten, zumal sich deren unterschiedliche Formen dem durchschnittlichen Benutzer kaum intuitiv erschließen und diese implizite Strukturierung (nach auch in den Paratexten des Wörterbuchs nirgendwo genannten Kriterien) des behandelten individuellen Wortschatzes etwa dem fachlich versiertem Lexikologen insbesondere bei elektronischen Versionen von Wörterbüchern, die ja doch ohnehin zur Norm zu werden versprechen, auf andere Art ebenso oder besser, weil auf spezifische Fragestellungen ausgerichteter, zugänglich sind.

3. Zusammenfassung

Die Lemmatisierungspraxis des GWb beruht auf einer größeren Menge von Lemmatisierungsvorschriften, die hier nicht alle behandelt oder auch nur genannt werden konnten. Die an Beispielen aus verschiedenen wörterbuchtypgenuinen Bereichen dargelegten Prinzipien orientieren sich insgesamt eher an semantischen als an grammatischen Kriterien, aus dem

Zweck des GWb als semasiologischem, auch textexegetischem Wörterbuch und Werkzeug heraus. Abweichungen vom semantischen Prinzip kann es bei der Behandlung von Funktionswörtern geben oder z.B. im Fall von Auxiliärverben, vgl. etwa sub voce *haben* unter Punkt D. Je weiter das Goethe-Wörterbuch als eines der letzten wissenschaftlichen Wörterbuch-Großprojekte im deutschsprachigen Raum im Alphabet voranschreitet, ein desto zuverlässigerer, lexikographisch aufbereiteter Überblick wird sich uns schließlich über einen einzigartig umfassenden und schöpferischen Idiolekt in einer sprachgeschichtlich formativen Sprachstufe des Deutschen darbieten. Auffälligkeiten in der makrostrukturellen Anordnung und Füllung im Vergleich mit historischen Sprachwörterbüchern, die nicht einzelpersonenbezogenen Corpora gewidmet sind, ergeben sich aus verschiedenen idiolektalen Eigenheiten Goethes, etwa bei den geringfrequenten lexikalischen Einheiten oder den Bindestrichnebenlemmata. Es konnte gezeigt werden, daß zunächst Fragen der Aufnahme bzw. Nichtaufnahme von Lemmata in bestimmten Wortschatzbereichen, etwa dem onomastischen Wortschatz, bedenkenswert sind; daneben konnten Fragen zur Form der Lemmataufnahme gestellt werden, die unter anderem eine Neubewertung der Nebenlemmaklassifizierungskriterien im Lichte neuerer Entwicklungen der (elektronischen) Lexikographie nahelegen. Kriterien der Lemmatisierung, auch für die Ansetzung von Haupt- und Nebenlemmata, bedürfen sowohl im Autorenwörterbuch als auch im Sprach(stadien)wörterbuch insgesamt jedoch der genaueren Überprüfung und Diskussion besonders im Hinblick auf eine mögliche Steigerung der Nutzerfreundlichkeit der lexikographischen Produkte (s. zuletzt etwa Gouws 2007: 59 f.).

Literatur

- Arnold, B. (2006). „Night Watches on the Computer: Creating an Author's Dictionary with Computational Means“. *Literary and Linguistic Computing* 21 (Suppl. Issue). 5-14.
- Brückner, D. (2007). „Zur Lemmaauswahl im Klassikerwörterbuch“. *Lexicographica* 22. 173-186.
- Dill, C. (1959). „Lexika zu einzelnen Schriftstellern“. *Forschungen und Fortschritte. Nachrichtenblatt der deutschen Wissenschaft und Technik* 33 (11). 340-346, 369-375.
- Goethe-Wörterbuch*. Hg. von der Akademie der Wissenschaften der DDR (ab Band III: der Berlin-Brandenburgischen Akademie der Wissenschaften), der Akademie der Wissenschaften in Göttingen und der Heidelberger Akademie der Wissenschaften. Stuttgart u.a.: Kohlhammer, 1978 ff.
- Gouws, R. H. (2007). „Sublemmata or main lemmata. A critical look at the presentation of some macrostructural elements“. In Gottlieb, H.; Mogensen, J. E. (Hgg.). *Dictionary Visions, Research and Practice. Selected papers from the 12th International Symposium on Lexicography, Copenhagen 2004*. Amsterdam: John Benjamins. 55-69.
- Niedermeier, M.; Objartel, G.; Welter, R. (2001). „O-Ton Goethe. Das „Goethe-Wörterbuch“ (GWb)“. *Zeitschrift für Germanistik. Neue Folge* 3. 596-600.
- Nutt-Kofoth, R. (2005). „Goethe-Editionen“. In Nutt-Kofoth, R.; Plachta, B. (Hgg.). *Editionen zu deutschsprachigen Autoren als Spiegel der Editions-geschichte*. Tübingen: Niemeyer (= Bausteine zur Geschichte der Edition; 2). 95-116.
- Objartel, G. (2005). „Johann Wolfgang von Goethe“. In Cruse, D. A. et al. (Hgg.). *Lexikologie. Lexicology. Ein internationales Handbuch zur Natur und Struktur von Wörtern und Wortschätzen*. [...] 2. Halbband. Berlin: de Gruyter. 1489-1493.
- Reichmann, O. (1996). „Neueste Autorenlexikographie: Problemerkörterung am Beispiel des Wörterbuchs zur Göttinger *Frauenlob*-Ausgabe“. In Wiegand, H. E. (Hg.). *Wörterbücher in der Diskussion II. Vorträge aus dem Heidelberger Kolloquium*. Tübingen: Niemeyer. 204-238.
- Schadewaldt, W. (1978). „Einführung“. In *Goethe-Wörterbuch*, Bd. I: *A-azurn*. Hg. von der Akademie der Wissenschaften der DDR, der Akademie der Wissenschaften in Göttingen und der Heidelberger Akademie der Wissenschaften. Stuttgart u.a.: Kohlhammer. 1*-14*.
- Unterberger, R. (1985). „Die Totalität des Individuellen: Über das Goethe-Wörterbuch“. *Jahrbuch für internationale Germanistik* 17. 147-168.
- Wiegand, H. E. (1983). „Was ist eigentlich ein Lemma? Ein Beitrag zur Theorie der lexikographischen Sprachbeschreibung“. In ders. (Hg.). *Studien zur neuhochdeutschen Lexikographie III*. Hildesheim/Zürich/New York: Olms. 391-474.
- Wiegand, H. E. (1989). „Aspekte der Makrostruktur im allgemeinen einsprachigen Wörterbuch: alphabetische Anordnungsformen und ihre Probleme“. In Hausmann, F. J. et al. (Hgg.). *Wörterbücher. Ein internationales Handbuch zur Lexikographie* [...]. Bd. 1. Berlin: de Gruyter. 371-409.
- Wiegand, H. E. (2006 ff). *Internationale Bibliographie zur germanistischen Lexikographie und Wörterbuchforschung*. Berlin: de Gruyter.
- Wolski, W. (1989). „Das Lemma und die verschiedenen Lemmatypen“. In Hausmann, F. J. et al. (Hgg.). *Wörterbücher. Ein internationales Handbuch zur Lexikographie* [...]. Bd. 1. Berlin: de Gruyter. 360-371.