

ROBERT CHARLIER

SYNERGIE UND KONVERGENZ

Tradition und Zukunft historischer Semantik am Beispiel des Goethe-  
Wörterbuchs.

In: Begriffsgeschichte im Umbruch?

Hg. von Ernst Müller

**Archiv für Begriffsgeschichte**; Sonderheft Jg. 2004 (2005), S. 167-183

ISSN 1617-4399 · ISBN 3-7873-1693-0

# Synergie und Konvergenz

## Tradition und Zukunft historischer Semantik am Beispiel des Goethe-Wörterbuchs

Die wachsende Zahl (retro)digitaler Wörterbücher und lexikographischer Online-Ressourcen zeigt: Wortschatzforschung, Begriffsgeschichte und allgemeine Enzyklopädik wachsen strukturell und technisch zusammen. Die immensen Möglichkeiten von intertextueller Verlinkung, Korpusintegration und Volltextsuche sorgen für systematisch nutzbare Synergien. Neue Verfahren der methodischen und materiellen Arbeitsteilung zwischen Wörterbüchern, Enzyklopädien und Lexika eröffnen bislang ungeahnte Dimensionen der komplementären Ergänzung und wechselseitigen Durchdringung der einzelnen Nachschlagewerke. Anhand des Akademienvorhabens Goethe-Wörterbuch (GWb) möchte ich im folgenden einige wesentliche Aspekte dieser Synthese vergegenwärtigen, die sich mit der Verbindung der wissenschaftlichen Fachlexikographie zu einer virtuellen Plattform für Historische Semantik abzeichnet.

### *I. Historische Enzyklopädik und Lexikographie*

Zunächst ist zwischen historischer Sach- und Sprachlexikographie zu unterscheiden. Der Anschaulichkeit halber spreche ich im ersten Fall von Enzyklopädik,<sup>1</sup> im zweiten von deutscher Wortschatz-Lexikographie (z. B. die Wörterbücher von Adelung, Campe oder der Gebrüder Grimm). Die ideen- und begriffsgeschichtliche Lexikalik nimmt in dieser Systematik eine Zwischenstellung ein, die sie zugleich für eine besondere Mittlerrolle prädestiniert.

---

<sup>1</sup> Im deutschsprachigen Raum handelt es sich vor allem um JOHANN HEINRICH ZEDLERS *Grosses vollständiges Universal-Lexicon aller Wissenschaften und Künste*, 1732-54 und die *Oekonomisch-technologische Encyclopädie* von JOHANN GEORG KRÜNITZ. Dieses Mammutwerk des Berliner 'Enzyklopädisten' verstand sich zunächst als Übersetzung entsprechender französischsprachiger Vorbilder. Es erschien seit 1773 über eine Zeitspanne von 85 Jahren in 242 Teilen, zunächst – bis einschließlich Band 32 – mit dem Titel: „Oeconomische Encyclopädie, oder allgemeines System der Land-, Haus und Staats-Wirthschaft“. Mit Band 33 nannte Krünitz sein Lexikon „Oekonomisch-technologische Encyclopädie“, im Untertitel mit dem modifizierten Zusatz: „allgemeines System, der Staats-, Stadt-, Haus- und Landwirtschaft, und der *Kunst-Geschichte*“. Dieser Fall ist ein schönes Beispiel für das unmittelbare Wechselspiel zwischen zeitgenössischer Sach- und Sprachlexikographie. Denn das Wort „Technologie“ erhielt in dieser Zeit eine Bedeutung im Sinne von „Kunstlehre, das Wort Kunst in weiterer Bedeutung [...] Genauer, die Kunst- und Handwerk- oder Gewerblehre“, vgl. JOACHIM HEINRICH CAMPE: Wörterbuch zur Erklärung und Verdeutschung der unserer Sprache aufgedrungenen

Die Unterscheidung zwischen Real- und Begriffslexikographie erweist sich keineswegs als selbstverständlich. So bemerkt Goethe in seinen sog. Maximen und Reflexionen: „Wenn einem Autor ein Lexikon nachkommen kann, so taugt er nichts“.<sup>2</sup> Diese Invektive diskreditiert keineswegs den zeitgenössischen Positivismus der Enzyklopädisten allein.<sup>3</sup> Vielmehr rührt Goethe hier an das Inkommensurable des autonomen Dichterwortes überhaupt: Dichtung ist letztlich nie rational oder eindimensional übersetzbar – weder in den Bereich analytischer Sacherklärung noch in den Wortschatz einer fremden Sprache. Versucht man, das Diktum spontan in eine Fremdsprache zu übertragen, z. B. das Englische, wäre wohl zweigliedrig zu übersetzen, etwa mit „lexicon“ und „dictionary“.

Die ‘französische’ Enzyklopädistik – von DENIS DIDEROT/JEAN LE ROND D’ALEMBERTS Jahrhundertwerk bis zur Schweizer *Encyclopédie* d’Yverdon des in Rom geborenen FORTUNÉ BARTHÉLEMY DE FÉLICE (1770-80) – begründete um die Mitte des 18. Jahrhunderts eine neue Ordnung und Verteilung des Wissens im Zeichen der europäischen Aufklärung. Lexikographische Produktion und Lexikon-Rezeption beförderten das Erwachen bürgerlicher Öffentlichkeit zum Jahrhundertwechsel. Dieser ‘Strukturwandel’ wird dann den sog. Konversationslexika<sup>4</sup> in der ersten Hälfte des 19. Jahrhunderts eine wahre Konjunktur bescheren und eine neue Popularisierung und Medialisierung des Wissens vorantreiben.

---

fremden Ausdrücke (Braunschweig <sup>2</sup>1813) 582.

<sup>2</sup> „Maximen und Reflexionen über Literatur und Ethik. Aus dem Nachlaß“, Hecker-Nr. 1059. Vgl. *Goethes Werke*. Hg. im Auftrage der Großherzogin Sophie von Sachsen. 4 Abteilungen in 143 Bdn. Weimar 1887-1919. I. Abteilung, Bd. 42.2, S. 252, Z. 1f. (= Erstdruck, 1907) [im folgenden WA]

<sup>3</sup> Goethes Skepsis gegenüber dem Lexikon- und Wörterbuchwesen ist ambivalent. Zum einen bedenkt er die Enzyklopädik mit viel Spott. Andererseits ist Goethe ein fleißiger Benutzer der Hand- und Wörterbücher seiner Zeit (z. B. der Werke von Adelung, Bayle, Campe, Sulzer, Mellin oder Riemer u. a.). Die ein- und zweisprachige Lexikographie seiner Zeit fundiert sein Schaffen dabei ebenso wie viele bedeutende begriffsgeschichtliche und enzyklopädische Wörterbücher. Allerdings wettert er vor allem gegen die „Verewigung“ von Irrtümern in naturwissenschaftlichen Lexika (vgl. z. B. Maximen und Reflexionen, Hecker-Nr. 426, 1261 und 1293). Dabei hat er aber in erster Linie Lehr- und Nachschlagewerke im Sinn, die z. B. Elemente der Newtonschen Physik als vermeintliche ‘Fehler’ perpetuieren. In seiner Bearbeitung der *Vögel* des Aristophanes die komische Hauptfigur lästern: „Hier sind die großen Lexica, die großen Krambuden der Literatur, wo jeder einzeln sein Bedürfnis pfennigweise nach dem Alphabet abholen kann!“ (WA I, Bd. 17, S. 95)

<sup>4</sup> Goethes widersprüchliche Haltung zur Lexikographie kulminiert in seinen Äußerungen zu einem der wichtigsten Bildungsmedien des bürgerlichen Zeitalters: „Conversations-Lexikon heißt’s mit Recht, / Weil, wenn die Conversation ist schlecht, / Jedermann / Zur Conversation es nutzen kann.“ (WA I, Bd. 3, S. 317, ‘Zahme Xenien’ V, 1225-28) Dagegen heißt es an anderer Stellen in den ‘Maximen und Reflexionen’: „Geschiedte Leute sind immer das beste Conversationslexikon“ (WA I, Bd. 42.2, S. 131, 8f.; Hecker Nr. 196)

Die Goethezeit schied also – noch ganz im Sinne der Spätaufklärung – ‘Diktionär’<sup>5</sup> und ‘Lexicon’ nur sehr unscharf. Dies nicht zuletzt deshalb, weil sie deren gemeinsamem Ursprung zeitlich näher war. Auf diese genetische Verwandtschaft von Lexikon und Wörterbuch verweist DIDEROT/D’ALEMBERTS Epochenwerk schon titelimmanent: *Encyclopédie ou Dictionnaire raisonné des sciences, des arts et des métiers* [...], Paris 1751-80 [72].<sup>6</sup>

Historisch entwickelten sich die einsprachigen Wörterbücher aus den zweisprachigen. Die (früh)mittelalterliche Glossographie bildete den Ursprung des Diktionärs aus dem Geiste der zwei- oder mehrsprachigen Wortsynopse. Noch das *Teutsch-Lateinische Wörterbuch* von JOHANN LEONHARD FRISCH (1741) markiert als gleichsam ‘erster’ Wortschatz des Neuhochdeutschen und Dokument der Fremdsprachenlexikographie zugleich diesen Verzweigungsprozeß.<sup>7</sup> Die fremdsprachenbezogene Herkunft aller Wörterbücher ist in der Goethe-Lexikographie noch deutlich zu erkennen. So besteht PAUL FISCHERS *Goethe-Wortschatz* von 1929 – erster, noch unvollständiger Versuch eines Goethe-Wörterbuchs – aus zwei Teilen: Auf eine Lemmatisierung der deutschen Wörter im Hauptteil folgt ein separater, verknappter und ebenfalls alphabetisierter „Fremdwörterteil“.<sup>8</sup> Auch das neuere, eigentliche *Goethe-Wörterbuch* basiert auf diesem Genotypus, indem es alle Vorlagenwörter der Übersetzungen und Bearbeitungen bucht, die Goethe ins Deutsche übertragen hat. Fremdsprachige Wörter, von Goethe appellativ oder terminologisch gebraucht, werden nach genau festgelegten Kriterien sogar eigens angesetzt.

Innerhalb der historischen Wortschatzforschung ist des weiteren zwischen älteren, zumeist abgeschlossenen Wörterbüchern und jüngeren zu trennen, die sich noch in der Bearbeitung befinden. Allein diese neueren Projekte können die Möglichkeiten der neuen Medien progressiv nutzen. Neuaufgaben oder -bearbeitungen mit (lokaler) digitaler Nutzungskomponente gehören inzwischen zum Nutzungsstandard einschlägiger ‘Seller’ unter den allgemeinsprachlichen Wörterbüchern wie DUDEN oder

---

<sup>5</sup> Goethe verwendet das eingedeutschte Neutrum im doppelten Sinn: „Enzyklopädie, (Real-)Lexikon“ und „Wörterbuch“, vgl. *Goethe Wörterbuch*. Hg. von der Berlin-Brandenburgischen Akademie der Wissenschaften, der Akademie der Wissenschaften in Göttingen und der Heidelberger Akademie der Wissenschaften (Stuttgart/Berlin/Köln [Mainz] 1966ff.), Bd. 2, Sp. 1205 s. v. ‘Diktionär’ [im folgenden GWb]

<sup>6</sup> Zur Wortgeschichte seit dem 17. Jahrhundert vgl. *Deutsches Wörterbuch* von JACOB GRIMM und WILHELM GRIMM. Leipzig 1854-1971 Bd. 14.2, Sp. 1559f. s. v. Wörterbuch [im folgenden DWb]

<sup>7</sup> Vgl. ULRIKE HAß-ZUMKEHR: *Deutsche Wörterbücher – Brennpunkt von Sprach- und Kulturgeschichte* (Berlin; New York 2001) 100-05.

<sup>8</sup> PAUL FISCHER: *Goethe-Wortschatz* (Leipzig 1929) 787ff.

KLUGE.<sup>9</sup> Insbesondere haben sich (retro)digitale Wörterbücher zu global verfügbaren Erkenntnishelfern und Informationsinstanzen entwickelt. Als feste Forschungsgrundlage dienen sie Netzphilologie, Kulturwissenschaften und interessiertem Fachpublikum. Zugleich garantieren sie eine stetig wachsende Wirksamkeit und Nützlichkeit der meist akademiegestützten Forschung in den Augen einer breiteren Öffentlichkeit.<sup>10</sup>

Nichts fordert das ganz Neue so wie das sehr Alte – nach diesem Grundsatz beeindruckt die historische deutsche Sprachlexikographie mit einer einmaligen Synthese von geisteswissenschaftlicher Tradition und technischer Innovation. Den konsequent beschrittenen Weg vom lexikographischen Monument zum gemeinnützigen Webwörterbuch dokumentiert hier nur eine Auswahl der wichtigsten Beispiele (Stand bei Drucklegung):

- JOHANN CHRISTOPH ADELUNG, Grammatisch-kritisches Wörterbuch der Hochdeutschen Mundart, 1793-1801; elektronische Volltextedition auf CD-ROM (= Digitale Bibliothek; 40), Berlin 2000; digitale Rekonstruktion der Ausgabe Wien 1808: [www.ub.uni-bielefeld.de/diglib/adelung/grammati/](http://www.ub.uni-bielefeld.de/diglib/adelung/grammati/); parallele Image- und Volltextversion derselben Ausgabe in der Auflage von 1811 (im Rahmen der „Verteilten digitalen Forschungsbibliothek“ der Staatsbibliothek München): <http://mdz.bib-bvb.de/digbib/lexika/adelung>
- Deutsches Wörterbuch von JACOB GRIMM und WILHELM GRIMM, 1854-1971; anspruchsvolle Retrodigitalisierung u. a. mit Gliederungsansicht: [www.dwb.uni-trier.de/index.html](http://www.dwb.uni-trier.de/index.html); elektronische Volltextedition auf 2 CD-ROMs, u. a. mit komplexen Suchmöglichkeiten, rückläufigem Stichwortindex und sog. Random-Reading-Funktion: DER DIGITALE GRIMM, hg. vom Kompetenzzentrum Trier in Verbindung mit der Berlin-Brandenburgischen Akademie der Wissenschaften, Frankfurt/M. (Zweitausendeins) 2004: [www.Zweitausendeins.de/Grimm-Forum](http://www.Zweitausendeins.de/Grimm-Forum)
- MATTHIAS LEXER, Mittelhochdeutsches Handwörterbuch, 3 Bde. Leipzig 1872-78; vernetzte Volltextdigitalisierung im Wörterbuchverbund (Benecke/Müller/Zarncke u.

---

<sup>9</sup> Vgl. z. B. Etymologisches Wörterbuch der deutschen Sprache, Neuaufl. Berlin; New York 2002 mit CD-ROM

<sup>10</sup> Einen Überblick über die virtuelle Wörterbuchlandschaft bietet: „Lexilinks – kommentierte Linkliste zur Lexikographie“ unter <http://grimm.adw-goettingen.gwdg.de/lexilinks> (verantwortet von M. Schlaefer, M. Runte, W. Arndt und C. Hohmann). Zum Überblick über aktuelle netzgestützte Wörterbuchprojekte vgl. Wissenschaftliche Lexikographie im deutschsprachigen Raum, hg. von THOMAS STÄDTLER (Heidelberg 2003) 439-86 sowie U. HAß-ZUMKEHR: Deutsche Wörterbücher, a. a. O. [Anm. 7] 362-70.

a.) unter <http://gaer27.uni-trier.de/MWV-online/MWV-online.html>  
- Wörterbuch der deutschen Gegenwartssprache [von RUTH KLAPPENBACH und WOLFGANG STEINITZ], hg. von der Deutschen Akademie der Wissenschaften zu Berlin, 1961-77; restrukturiertes Volltextdigitalisat auf der Internetseite des Digitalen Wörterbuchs der deutschen Sprache des 20. Jahrhunderts unter [www.dwds.de](http://www.dwds.de), Hyperlink 'Wörterbuch'

Typologisch gruppieren sich innerhalb der einsprachigen historischen Lexikographie allgemein-, fach- und individualsprachliche Wörterbücher.<sup>11</sup> Genaugenommen gliedern sich die Wörterbücher zu den großen Repräsentanten einer Nationalsprache (z. B. zu SHAKESPEARE, PUSCHKIN, MICKIEWICZ) wiederum in umfassende Autoren- bzw. selektive Werkwörterbücher. Die Forschungsgeschichte des GWb zeigt die seltene Genese eines (vollständigen) Autorenwörterbuchs aus Teilwörterbüchern zu einzelnen herausragenden Werken wie 'Götz', 'Werther', 'Divan' oder 'Faust'. Zu nennen sind hier vor allem Unternehmungen aus dem Umfeld der (Ost-)Berliner Akademie.<sup>12</sup>

## *II. Wortschatz-Sammlung als Ausdruck von Wertschätzung*

Seit der Erschließung von Wortkonkordanzen zur Bibel im Paris des 13. Jahrhunderts war es, wie Katharina Mommsen erinnert, „die Überzeugung von der Kostbarkeit eines jeden Worts der Heiligen Schrift“<sup>13</sup>, die Schreiber und Gelehrte zur Verfertigung vollständiger Vers- und Verbalkonkordanzen motivierte. Der Begriff Wortschatz offenbart bereits eine etymologisch sinnfällige, programmatische Wahrheit: Thesaurus, griech. θησαυρός: „Schatzkammer“ oder „Vorrat, Schatz“, kontextuell übertragen auch für alles Teure, Wertvolle. Vor diesem bedeutungsgeschichtlichen Hintergrund erweist sich Wortschatzsammlung als ein Ausdruck von 'Wort-Schätzung' im Sinne höchster kollektiver Wertschätzung eines sprachlichen (Kunst-)Werkes.

Die Wiederbegründung der Goetheforschung unmittelbar nach dem Zweiten Weltkrieg an der damaligen Deutschen Akademie der Wissenschaften zu Berlin erhellt erst vor dem Hintergrund einer besonderen Klassikerwertschätzung zur deutschen „Stunde Null“. Mit Blick auf humanistische Traditionen annoncierte der klassische Philologe WOLFGANG SCHADEWALDT das Klassiker-Wörterbuch als „Magna Charta für

---

<sup>11</sup> Vgl. dazu ANGELIKA STORRER/K. FREESE: Wörterbücher im Internet. In: Deutsche Sprache (Mannheim 1996) 97-153, insbes. 107-16

<sup>12</sup> Vgl. die Werkwörterbücher von N.-FÜRSTENAU 1958; MERKER 1966 und DILL 1987

<sup>13</sup> KATHARINA MOMMSEN: Verskonkordanz zu Goethes 'Faust, Erster Teil'. Bearbeitet von Steven P. Sondrup und David Chisholm (Tübingen 1986) VII.

das neuere Deutsch”.<sup>14</sup> Verstand sich VICTOR KLEMPERERS *Lingua Tertii Imperii* (LTI) von 1946 noch als Versuch kritischer Sprachinventur ex negativo, so bildete das GWb den ersten nennenswerten Versuch der Nachkriegszeit, sich positiv(istisch) auf das deutsche Sprach- und Kulturerbe zu besinnen.

Eine personelle Verbindung aus der Gründerzeit des GWb vergegenwärtigt diesen Impuls der Projektväter. ULRICH PRETZEL, seit 1947 Leiter der Hamburger Arbeitsstelle, war der Bruder von Raimund Pretzel alias SEBASTIAN HAFFNER (1907-1999). Das Pseudonym des Emigranten und Heimkehrers wurde zum Inbegriff der Kontinuität eines anderen, kritisch-intellektuellen Deutschland und seiner Sprache – als Heimat eines Johann Sebastian Bach und Wolfgang Amadeus Mozart.<sup>15</sup>

Heute sind solche historischen Projektionen nur noch bedingt auf den Wortschatz eines Klassikers an der Schwelle zum modernen Informations- und Kommunikationszeitalter übertragbar. Denn trotz auflebender Überlegungen zur Kanonbildung in Bildungswesen und Publizistik: Der gegenwärtige Zeitgeist scheint eher beherrscht von dem Wunsch nach enthierarchisierter Allverfügbarkeit kulturellen und historischen Wissens.<sup>16</sup> Auch die Begriffe von ‘Klassik’ und ‘Klassizität’ scheinen wissenschaftlich weitgehend historisiert und nivelliert – damit aber auch im Diesseits der kanonkritischen Weltsicht angelangt. Ihre Berechtigung wahrt die kanonisierende Wertschätzung dennoch, zumindest im Sinne einer anthropologischen Konstante. Denn jede Epoche oder (Hoch-)Kultur kannte bislang ein personal oder textuell manifestes Heiliges oder Höchstes, eben ‘Klassisches’ – sei es eine Verkörperung durch ein – heroisches, auch vergöttertes – Individuum oder das Textkorpus eines berühmten Weisen oder Dichters.

### *III. Goethe: einziger Neunzigtausender der deutschen Literatur?*

In ihrer monumentalen Masse und phänomenalen Bedeutungsfülle bleibt die Sprachwelt Goethes historisch unerreicht. Goethes Wortgebrauch, sein dichterisches Schöpferturn erscheinen einzigartig: Kein bekannt gewordenes individuelles Sprachvermächtnis

---

<sup>14</sup> WOLFGANG SCHADEWALDT: Das Goethe-Wörterbuch. Eine Denkschrift. In: Goethe. Neue Folge des Jahrbuchs der Goethe-Gesellschaft 11 (1949) 294.

<sup>15</sup> Zur institutionellen Einbettung der Goethe-Forschung und ihrer bedeutendsten Köpfe vgl. RUDOLF BENTZINGER: Das Institut für deutsche Sprache und Literatur bei der Deutschen Akademie der Wissenschaften zu Berlin (1952-1969) – Ort gesamtdeutscher Germanistik. In: Mitteldeutsches Jahrbuch für Kultur und Geschichte 11 (2004) 151-54; 161f.; 165-67 sowie die Kurzporträts von WILHELM WISSMANN (ebd. 169); WOLFGANG SCHADEWALDT (170) und ERNST GRUMACH (173).

<sup>16</sup> Vgl. dazu HELMUT FUHRMANN: Klassiker oder Computer? In: Goethe-Jahrbuch 113 (1996) 259f.

vereint eine derartige Verwendungsvielfalt von Gemein-, Fach- und Poesiesprache. Dies manifestiert sich in einem Gesamtwortschatz von rund 92.000 einzelnen Wörtern. Dabei handelt es sich neben Wörtern wie *Urphänomen*, *Wahlverwandtschaft* oder *Weltliteratur* um weitaus häufiger belegte Goethesche ‘Natur- und Wesenswörter’, aber auch zahlreiche Begriffe aus der Amts- und Rechtssprache oder naturwissenschaftliche Terminologie sowie Einmalbildungen und poetische Neologismen.<sup>17</sup> Goethes Wortschatz ragt aber nicht nur qualitativ in unerreichte Höhen. Denn die reiche Wortwelt seines schriftlichen Gesamtwerks, eine enorme Menge von Gedichten, dramatischen und Prosatexten, Privatbriefen, Amtsschreiben und wissenschaftlichen Schriften – mit erhaltenen Tagebüchern und (aufgezeichneten) Gesprächen aus vielen Jahrzehnten – bildet den bisher wohl einzig bekanntgewordenen ‘Neunzigtausender’ der deutschen Literatur. Zum Vergleich: Für das projektierte Schiller-Wörterbuch sind rund 70.000 Lemmata veranschlagt.<sup>18</sup> Die Aussagekraft rein zahlenmäßiger Betrachtungen ist allerdings eher begrenzt. Die Größe eines erschlossenen Wortschatzes allein erlaubt noch keine fundierten Rückschlüsse. Schließlich bleibt der Fokus eines jeden Ansetzungsrahmens fließend. So kannte die Goethezeit noch keine allgemein verbindliche Getrennt- und Zusammenschreibung. Ernsthafte orthographische Normierungsversuche datieren erst aus nachgoethischer Zeit oder setzten sich erst nach dem Tode des Dichters durch.<sup>19</sup> Singuläre Qualität und Quantität – dieses Merkmal von Klassizität ist auch für das GWb konstitutiv, indem es eine der letzten universalen und künstlerisch nahezu unerschöpflichen individuellen Sprachleistungen nach dem Thesaurusprinzip vollständig zu beschreiben anstrebt, und zwar nach dem Prinzip: Darbietung aller Wörter *und* Darstellung aller Bedeutungen eines jeden Worts bei Goethe.

#### *IV. Lexikographie als Meta-System*

An den Akademien des deutschsprachigen Raums bilden neben einer Fülle wichtiger, oft international einmaliger Forschungsprojekte eigene Archive und Bibliotheken ein unverrückbares Fundament. Eckpfeiler der geisteswissenschaftlichen Grundlagenforschung sind unter anderem Großvorhaben vom Typ Bibliographie, Edition, Lexikon oder Wörterbuch.<sup>20</sup> Mit diesen Elementen kultivieren die Akademien

---

<sup>17</sup> Vgl. dazu OTTO PNIOWER: Goethe als Wortschöpfer. In: *Euphorion* 31 (1930) 362-83.

<sup>18</sup> Laut Selbstdarstellung unter <http://grimm.adw-goettingen.gwdg.de/wb-liste/?29> [Stand bei Drucklegung].

<sup>19</sup> Vgl. etwa HEYSE 1833ff., aber auch ADELUNG <sup>2</sup>1835.

<sup>20</sup> Diese Grobeinteilung erfaßt die enorme Vielfalt akademiegestützter Forschung keineswegs restlos. Zu



gewissermaßen die Ordnungs- und Zugriffssysteme der modernen Medien- und Wissenswelt. Im technischen Sinn entspricht dies den Paradigmen ‘Inhalt/Speicher’ (content/memory) und ‘Zugriff/Abwurf’ (retrieval).

Umgekehrt hat die technische Medienwelt diese epistemischen Ordnungen verinnerlicht. Denn jeder Inhalt verlangt in der Welt der neuen Medien nach einer Verarbeitung im Sinne letztlich wissenschaftlicher Verfahren. Der Inhalt (nicht nur, aber in der Regel ein Text) verlangt elektronisch *ediert*, also in eine komplexe Struktur von Meta-Informationen eingebettet und formatiert bzw. annotiert zu werden. Für Kataloge oder (kommentierte) Linklisten muß er *bibliographierbar* und für eine Suchmaschine indexiert, d. h. im Sinne eines virtuellen *Wörterbuchs* der Suchbegriffe oder *Lexikons* der Schlagworte *lemmatisiert* werden. Schließlich werden Text und Metatext im wissenschaftlich-technischen Verarbeitungsprozeß wie jedes beliebige andere Datenobjekt in einem standardisierten Format *archiviert*.

Die Frage: Gehört ein (sprach)historisches oder begriffsgeschichtliches Wörterbuch ins Internet? verdient deshalb vor diesem Hintergrund reverse Betrachtung: Trägt die Webwelt nicht eher umgekehrt die Signaturen der (Wörter-)Buchwelt von klassischer Bildungskultur und Wissenschaft? Aus der Sicht der historischen Lexikographie kann man dies zweifellos bejahen. Zwar ist die Mehrzahl der retrodigitalisierten historischen Wörterbücher noch weitgehend von der Formensprache und den Nutzungsvorstellungen des 19. Jahrhunderts bestimmt.<sup>21</sup> Es zeigen sich aber bemerkenswerte Strukturparallelen mit der technologischen Morphologie des 21. Jahrhunderts. Das verdeutlicht eine sprechende Gegenüberstellung der Module eines Goethe-Wortartikels mit den typisierten Elementen des Fundstellen-Eintrags einer Suchmaschine (z. B. Google)<sup>22</sup>. Die Bausteine eines Wortartikels sind in der linken Spalte vereinfachend benannt:<sup>23</sup>

---

den Interdisziplinären Arbeitsgruppen als innovatives Forschungsinstrument der Berlin-Brandenburgischen Akademie der Wissenschaften vgl. DIETER SIMON: Akademie der Wissenschaften. Das Berliner Projekt. Ein Brevier (Berlin 1999) 100-04 bzw. 159-61.

<sup>21</sup> Vgl. MICHAEL SCHLAEFER: Standards retrodigitaler Wörterbücher. In: Lexikos (Afrilex-Reeks) 10 (2000) 165.

<sup>22</sup> Gemäß einer Selbstdarstellung des Produktes ‘Google’ ergibt sich der Name der zur Zeit erfolgreichsten Suchmaschine aus einem mathematischen Begriff, nämlich „Googol, einer 1 mit einhundert Nullen“, vgl. NORBERT THOMMA: Wie denkt das größte Hirn der Welt? In: Der Tagesspiegel Nr. 17871 vom 1.9.2002.

<sup>23</sup> Zum Gebrauch vgl. GWb Bd.1 (1978) Sp. 1\*-14\* bzw. Sp.17\*-20\*. Zur aktuellen Orientierung, auch im Hinblick auf eine geplante elektronische Nutzung vgl. GOETHE-WÖRTERBUCH. Hg. von der Berlin Brandenburgischen Akademie der Wissenschaften; Reihe Informationsbroschüren der Akademienvorhaben (Berlin 2004) 12-15 sowie 22ff. [Online-Dokumentation: Hyperlink ‘Artikelaufbau’ unter [http://bibliothek.bbaw.de/Goethe/gw/fr\\_haupt.htm](http://bibliothek.bbaw.de/Goethe/gw/fr_haupt.htm)]

Lemma	Such-, Schlag-, Stichwort
Vorbemerkung	Meta-Information (Fundort, Seitenstatistik)
Leitbemerkung	Trefferkommentar
Zitat	referenziertes Dokument (Bild, Grafik, Ton)
Stellenangabe	virtuelle Ortsreferenz (URL)
Verweise (zur Wortbildung)	Hyperlinks
Synonyme	Angabe ähnlicher Seiten
Lemmaliste	Index

Die Parallelität in der Mikrostruktur von Wörterbuchartikel und Fundeintrag gilt hier nur für die semantische Funktion der einzelnen Module. Denn es gilt: Bleibt der (gedruckte) Wörterbuchtext stets einer statisch-hierarchischen Dimension verhaftet, so sind die Datenbankfelder einer Fundstellenausgabe stets dynamisch und seriell organisiert<sup>24</sup>.

Internet und neue Medien haben die nutzerspezifischen Anforderungen und technischen Möglichkeiten der sprachhistorischen Lexikographie zweifellos revolutioniert. Allerdings erwiesen sich die Schlüsselbegriffe historischer Wörterbuchkonzepte umgekehrt auch als unverzichtbare evolutionäre Konstituenten des weltweiten Wissensnetzes. Beinahe mühelos lassen sich die ‘grammatischen’ Elemente der lexikographischen Beschreibungssprache als technische Morphologie der Internetwelt verstehen: Begriffsp parallelen wie Lemma/Suchwort, Textkorpus/Volltext oder Verweis/Hyperlink, aber auch statthafte Analogisierungen wie Wortschatz = Volltextlemmatisierung oder semantische Bestimmung = kommentierter Wortindex belegen: Viele Schlüsselbegriffe aus der Welt der Search engines oder Web crawler übersetzen Wörterbuchprinzipien in die Terminologie und Taxonomie der Suchmaschinen. Damit besetzen lexikalisch-thesaurische Muster eine sehr wichtige Schnittstelle zwischen semantisch-natürlichsprachlicher und logisch-maschinensprachlicher Welt.

#### *V. Konvergenz von Kulturwissenschaft und Worldwide Web*

Die Affinität zwischen moderner Web- und historischer Wörterbuchwelt erlaubt drei kurze Ausblicke über den engeren Rahmen der lexikographischen Betrachtung hinaus.

Erstens: Lexikalische und lexikographische Techniken im medialen Umgang mit

---

<sup>24</sup> Vgl. dazu MARC MEYER: Korpusbasierte Erstellung eines Wörterbuchs des Deutschen. Chancen und Schwierigkeiten (Marburg 2003) 57-59. [zugl. unter <http://www.marc-meyer.de/korpus/text.pdf>]

Wissen spiegeln auch eine Verwissenschaftlichung der Öffentlichkeit wider. Ordnungsbegriffe wie Archiv, Enzyklopädie, Katalog, Index, Register u. a. m. sind jedem Netznutzer unabhängig von seinem Bildungsstand geläufig. Man könnte sagen: Das Internet hat die epistemische Formensprache in ähnlicher Weise popularisiert und demokratisiert wie Nomenklatur und Einheitensysteme der Naturwissenschaften und der Technik („Megahertz“, „Kilobit“ u. ä.). Selbst usuelle Metaphorik des digitalen Zeitalters ist in weiten Bereichen der Kultur- und Wissensgeschichte entnommen. Man bedenke nur die Vorstellung des „Scrollens“, die auf die antike Schriftrolle zurückgeht,<sup>25</sup> das virtuelle Werkzeug des „Editors“ oder die heuristische Vorstellung des Desktops als „Explorer“ – in Anlehnung an den Begriff der wissenschaftlichen Entdeckerfahrt. Weitere Leitbegriffe der Web- und Medienwelt säkularisieren Elemente der Wissenschaftskultur: Fachwörter wie *Software-Archiv*, *Digitale Bibliothek*, elektronisches *Buch*, elektronische *Edition*, *Viren-Enzyklopädie*, *Volltext-Index* oder *Internet-Katalog* sind wahre Termini technici der neuen Medienwelt. Sie enthalten jeweils ein semantisches Zweitglied, das die jahrhundertealte Strukturwelt des Wissens aktualisiert.

Zweitens: Auch im Blick auf die Standardisierung von Datenformaten, Programmiersprachen und Auszeichnungssystemen hat die Internetwelt die Wissenschaft um eines ihrer höchsten historischen Verdienste beerbt. Denn erst die akademische Konventionalisierung des Lateinischen oder Englischen als wissenschaftliche Verkehrssprache ermöglichten die Entstehung einer Scientific community. Dieses Prinzip der (sprachlichen) Standardisierung für die technisch schnellebigen Windungen des kulturellen Gedächtnisses gewinnt in der digitalen Wissenswelt täglich an Bedeutung. Spezielle bibliotheks- und geisteswissenschaftliche Standards wie z. B. die Vorgaben der Open Archive Initiative (OAI) oder der Text Encoding Initiative (TEI) bedienen sich, zumindest was die semantische Seite technischer Normierung betrifft – altbewährter szientistischer Konventionalisierung. Klassifikationssysteme wie die sog. Dewey Decimal Classification (DDC) oder der Versuch einer Universalkodierung aller Schriftzeichen der Menschheit durch UNICODE gehören ebenfalls in diesen Zusammenhang.

Drittens: Der populärwissenschaftliche Umgang mit Internet und neuen Medien läßt

---

<sup>25</sup> „Der in unsere Sprache herübergewanderte Ausdruck ‘scrollen’ macht es deutlich: Wir sind dabei, die kaiserzeitliche Entwicklung von der Buchrolle zum Codex wieder umzukehren.“ Vgl. CHRISTOPH MARKSCHIES: Digitalisierung antiker Texte. In: *Gegenworte* 8 (2001) 36.

historische Erfahrungsmuster als neue mediale Mythologeme wieder auf- und fortleben. So hat der Internetbeauftragte des Vatikans in Gestalt des verkündenden Boten- und Erzengels Gabriel einen Schutzheiligen für das weltweite Wissensnetz vorgeschlagen.<sup>26</sup> Auch Software-Entwickler bedienen sich gerne biblischer oder mythologischer Figuren und Vorstellungen, um ihre Produkte in einen bedeutenden Namen zu bannen. In diesem Sinne konstituiert das datenbankgestützte Autoren- und Redaktionssystem FAUST zwar keinen näheren Bezug zu seinem Namenspatron. Dafür empfiehlt sich die Lese-Software zum *Thesaurus Linguae Graecae* (TLG) mit ihrem sprechenden Namen in sehr ambivalenter Weise als (Büchse der) PANDORA, die mit ihren Treffer- und Textmyriaden auch ein 'Übel des Zuviel' über den Nutzer ausgießt.<sup>27</sup> Und der Korpusbrowser der Online-Bibliothek des Archimedes-Projektes, entwickelt am Berliner Max Planck-Institut für Wissenschaftsgeschichte, heißt nach der Geleiterin des Odysseus im Reiche des Phäakenkönigs (<http://NAUSIKAA2.rz-berlin.de>). Schließlich waltet die Muse der Geschichte gleich mehrmals als Namensgeberin für wissenschaftliche Online-Projekte im Bereich der Geschichtswissenschaften (als CLIO-online.de, „Fachportal für die Geschichtswissenschaften“, oder als Datenbank KLEIO, eine Entwicklung des Instituts für Historisch-Kulturwissenschaftliche Informationsverarbeitung an der Universität Köln, s. unter [www.hki.uni-koeln.de](http://www.hki.uni-koeln.de)).

#### *VI. Zur elektronischen Struktur des Goethe-Wörterbuchs*

Ein kurzer geschichtlicher Rückblick auf die Forschungs- und Methodengeschichte der Goethe-Lexikographie verspricht in diesem Kontext weiteren Erkenntnisgewinn. So erscheint das bewährte Neben- und Miteinander von thesaurischem und indexierendem Arbeitsprinzip in der germanistischen Literatur- und Sprachwissenschaft für zukünftige technische und konzeptionelle Entwicklungen in der Wörterbuchlandschaft aufschlußreich.

Zunächst möchte ich am Beispiel der Literaturwissenschaft die Intentionen bestimmen, die die Anwendung linguistisch *indexierender* Verfahren auf ein literarisches Korpus motivieren können. Die literaturwissenschaftliche Korpusanalyse dient in der Regel:

---

<sup>26</sup> Vgl. DANIEL WILLIAMS: Das Internet sucht seinen Schutzpatron. Der Vatikan fahndet nach dem perfekten Heiligen für das Netz. In: Der Tagesspiegel Nr. 18022 vom 3.2.2003 sowie unter [www.santiebeati.it](http://www.santiebeati.it)

<sup>27</sup> CHR. MARKSCHIES: Digitalisierung antiker Texte, a. a. 0. [Anm. 25] 35.

- der auktorialen Zuschreibung eines Textes oder Korpus;
- der Ermittlung von Form-, Stil- oder Motivaffinitäten oder
- der strukturellen Inhaltsanalyse.

Die erste, gleichsam ‘forensische’ Motivation korpuslinguistischer Arbeitsweisen ermöglicht mit Blick auf einen Text u. a. Antworten auf die Frage: Wer ist der Verfasser? Mittels entsprechender Wortschatz- und Stilanalysen konnte die ältere Forschung z. B. die Autorschaft Goethes für die Literatur-Rezensionen aus den *Frankfurter Gelehrten Anzeigen* von 1772 im einzelnen bestimmen.<sup>28</sup>

Die zweite intertextuelle Untersuchungsweise erlaubt vergleichende Aussagen über Parallelen, Ähnlichkeiten oder Übereinstimmungen zwischen verschiedenen Autoren. Die dritte beschränkt sich auf die textimmanente Betrachtung, z. B. zur Verwendungshäufigkeit oder Schreibweise eines Wortes oder einer Wortgruppe innerhalb eines Textes. Die Kulturinformatikerin Martina Schwanke bemerkt dazu in ihrem „Historischen Überblick über Versuche einer computergestützten Aufbereitung literarischer Texte“:

„Schon Ende der vierziger Jahre [des 20. Jahrhunderts, R. C.] wurde die Möglichkeit entdeckt, mit den neuen Rechenanlagen literaturwissenschaftliche Aufgaben zu lösen, von denen man wegen des immensen Personal- und Zeitaufwandes sonst zurückgeschreckt wäre. Damals begann man mit der automatischen Erstellung der ersten Werkindices und -konkordanzen, die wegen der noch geringen Leistungsfähigkeit der Anlagen zunächst einen nur bescheidenen Umfang haben konnten.“<sup>29</sup>

Die technischen Möglichkeiten der Verzettelung und „Textzerlegung“ beflügelte auch die Goethe-Lexikographie. Das belegt eine Reihe entsprechender Publikationen und Projekte.<sup>30</sup> Für die Exzerptionsphase des GWb waren großrechnergestützte Verfahren der Korpusanalyse unverzichtbar. So bilanzierten GERHARD STICKEL und MANFRED GRÄFE bereits im Erscheinungsjahr der ersten Lieferung des GWb die Debatte um „Natürliche Sprachen und elektronische Datenverarbeitung“ mit ihrem nüchternen Bericht über Möglichkeiten und Mühsal der „[a]utomatische[n]

---

<sup>28</sup> Vgl. Wörterregister zu Goethes Rezensionen in den *Frankfurter Gelehrten Anzeigen*, hg. von der Deutschen Akademie der Wissenschaften zu Berlin (Leipzig 1959) I-VII [masch. Manuskript.]

<sup>29</sup> MARTINA SCHWANKE: *Namen und Namengebung bei Goethe. Computergestützte Studien zu epischen Werken* (Heidelberg 1992) 92.

<sup>30</sup> Vgl. HOHLFELD/JOOS/TWADDELL 1940; KITTEL/FUERST 1946; SCHMIDT 1970; MOMMSEN 1986.

Textzerlegung und Herstellung von Zettelregistern für das Goethe-Wörterbuch”.<sup>31</sup> Die Werkzeuge aus dieser Frühzeit beliefen sich dabei noch auf Flußdiagramme in der Programmiersprache FORTRAN, umgesetzt auf den Lochkarten einer Rechenanlage vom Typ IBM 7090.<sup>32</sup> Noch lange vor Einführung des Personal Computer wurde das Goethesche Werk mittels „Verlochung” einzelner Textkorpora erstmals vollständig ausgewertet und durchlemmatisiert. Da aber noch keine komplette elektronische Fassung des gesamten Œuvres vorlag, ergänzte die maschinelle Exzerption lediglich die manuelle. Auch bei den automatisierten Analysen blieb eine „useful man/machine interaction” unverzichtbar.<sup>33</sup> Die technikgeschichtlich bedingte Struktur der Arbeitsgrundlagen des GWb durch automatische Indexierung konstituiert auch die heutige Datenbasis des Wörterbuchs.

Ein Beispiel: Die angesetzte Lemmaform „Goethit” verweist ebenso auf den Gesamtwortindex wie die Synonymen-Verweise am Artikelende. Die zahlreichen Abkürzungen, Werk- und Namenssiglen referenzieren die indexierten Listen mit den Auflösungen der Textkonventionen sowie der Werktitel und Namen der Briefempfänger:

---

<sup>31</sup> Vgl. GERHARD STICKEL/M. GRÄFE: Automatische Textzerlegung und Herstellung von Zettelregistern für das Goethe-Wörterbuch. In: Sprache im technischen Zeitalter 19 (1966) 247.

<sup>32</sup> G. STICKEL/M. GRÄFE: Automatische Textzerlegung, a. a. O. [Anm. 31] 251; 255-57.

<sup>33</sup> G. STICKEL/M. GRÄFE: Automatische Textzerlegung, a. a. O. [Anm. 31] 254.

**Goethit** von JGLenz 1806 in die mineralog Terminologie aufgenommener Name für Nadeleisenerz<sup>1)</sup>

Wohlwollende Männer aus dem Westerwald entdecken ein schönes Mineral und nennen es mir zu Lieb und Ehren G.; denen Herrn Cramer und Achenbach bin ich dafür noch vielen Dank schuldig, obgleich diese Benennung auch schnell aus der Oryctognosie verschwand. <sup>2)</sup> Es hieß auch Rubinglimmer, gegenwärtig kennt man es unter der Bezeichnung Pyrosiderit N6,162,18 MetamPfl Nachtr B27,217,22 Cramer [6.11.16] K uö Syn Pyrrhosiderit Rubinglimmer

<sup>1)</sup> vgl HFranke/VWahl, Zur Entstehung des Mineralnamens Göthit, GJb95, 1978,241ff <sup>2)</sup> erst nach G-s Tod wurde der Name zu einem festen Bestandteil der Terminologie (ebd 241)

Die Wörterbuchartikel verstehen sich somit nicht nur als Erträge semantisch bestimmender, hermeneutischer Arbeit, sondern sie bündeln auch eine Fülle von Informationen, die allein aus analytischen Verfahren zu schöpfen sind. Bei einem historischen Belegwörterbuch wie dem GWb gehören dazu statistische Aussagen über die Verwendungsbereiche und die Gebrauchsgewichtung der einzelnen Wörter, aber auch Angaben zu (seltenen) Schreibungen oder Fehlesungen. Letztlich speisen sich fast alle Module des einzelnen Wortartikels aus einem Index als universeller ‘Text hinter dem Text’.

Der lexikographische Arbeitsprozeß reflektiert damit die Genese der semantisch-thesaurischen Intensivform aus der automatisch indexierten Extensivform. Das technische Layout visualisiert den Entstehungsprozeß des Wörterbuchs als Aufstieg vom Index zum Thesaurus.

### *VII. Von der historischen zur korpusbasierten Lexikographie*

Ein Blick in die Geschichte der deutschen Wortschatzforschung ist in diesem Zusammenhang ebenfalls aufschlußreich. Historisch betrachtet treten den älteren ‘thesaurischen’ Wörterbüchern des Deutschen heute nämlich zahlreiche moderne lexikographische Informationssysteme vom Indextyp gegenüber. Die Vertreter der traditionellen Lexikographie setzten in der Regel den Namen ihrer großen Pioniere ein Denkmal gesetzt: JOHANN CHRISTOPH ADELUNG, JOACHIM HEINRICH CAMPE, JACOB und WILHELM GRIMM – aber auch MORIZ HEYNE oder HERMANN PAUL. Die Wörterbuchsysteme vom Indextyp dagegen sprechen mit mehr oder weniger klingenden

Akronymen wie CISLEX, COSMAS, DWDS, GERMANET, LIMAS oder NEGRA.<sup>34</sup> Sie zeugen damit von ihrer Herkunft aus den anglo-amerikanischen Corpus linguistics. Anders als klassische Printwörterbücher rationalisieren diese digitalen Projekte den lexikographischen Produktions- und Rezeptionsprozess, und zwar von der elektronischen Textquellenerschließung über die Segmentierung und Annotierung ihrer Korpora bis zur Extraktion ihres ‘Wissens über Wörter’. Kontrastiv läßt sich sich folgender Gegensatz von Thesaurus- und Indexprinzip konturieren:

- a) Wörterbücher vom Thesaurustyp bestimmen Wortbedeutungen im Sinne der historischen Semantik. Sie gründen in der Regel – nach beendeter Exzerptionsphase – auf einem genau quantifizierten Korpus (im Falle der gemeinsprachlichen Wörterbücher) bzw. auf einem geschlossenen Wortschatz (Autorenwörterbücher) und lemmatisieren von A bis Z. Ihre Redaktoren gleichen semantischen Archäologen, die ihre Worterklärungen vor allem an philologischer und sprachwissenschaftlicher Hermeneutik orientieren. Auf der Ebene der Wortartikel dominieren absolute Angabetypen, wie semantische, phonetische, morphologische, syntaktische oder pragmatische Aussagen, so z. B. der Verweis auf ein Hapax legomenon oder eine Erstdatierung. Es dominiert ein hierarchisierter Artikelaufbau. Medium dieses Typs ist das Print- oder (retro)digitalisierte Online-Wörterbuch.
- b) Wörterbuchsysteme vom Indextyp fokussieren weniger die ‘Bedeutung’ als die ‘Eigenschaften’ von Wörtern. Das betrifft vor allem das Verhalten von Wörtern im repräsentativen oder kontextabhängigen Massenverband. Sie basieren auf großen, in der Regel dynamisch erweiterbaren elektronischen Textkorpora, teilweise mit Hunderttausenden (oder gar Millionen, Milliarden) von Textwörtern und akkumulieren zumeist offene Wortschätze. Hier dominieren relative Angaben, basierend auf der seriellen Feldstruktur der zugrundeliegenden Datenbanken. Komplexe Retrievalfunktionen ermöglichen hauptsächlich objektivierte Wissensextraktion und favorisieren statistische, systematische oder heuristische Fragestellungen: Treten Wörter musterbildend miteinander in Verbindung? Haben sie gesetzmäßige Beziehungen zu bestimmten Textsorten? Medium dieses Typs sind Volltextdatenbanken mit mächtigen Benutzerschnittstellen und virtuelle Wörterbuchsysteme mit verteilten Online-Ressourcen.

---

<sup>34</sup> Vgl. M. MEYER: Korpusbasierte Erstellung eines Wörterbuchs, a. a. O. [Anm. 24] 26-30; 77-79.



Allerdings erhellen aus dieser Gegenüberstellung von ‘Thesaurus’ und ‘Index’ auch divergierende Tendenzen. So laufen nämlich die traditionellen Wörterbücher gelegentlich Gefahr, lediglich isolierte „Einzelwortschicksale“ in arbiträrer Abfolge aufzulisten oder auf Artikelebene in unüberschaubare Wortmonographien auszufern.<sup>35</sup> Auf der anderen Seite verkümmert die dynamische Datenausgabe lexikographischer Datenressourcen nicht selten zum bloßen Formular für Zahlen und Zeichen. Das Besondere und Interessante droht dabei in Myriaden von Textwörtern zu versinken.

Auch ist die Antinomie von Thesaurus- und Indextyp durch die technische Entwicklung in Bewegung geraten. Zum einen hat man einige historische Wörterbücher inzwischen zuverlässig (retro)digitalisiert, wie z. B. Adelung, Lexer oder Grimm. Soweit sie dabei im Volltext restrukturiert und annotiert vorliegen, werden sie dadurch selbst zum (Teil eines) Korpus. Dies dokumentieren Wörterbuchverbände wie das Trierer Projekt „Mittelhochdeutsche Wörterbücher auf CD-ROM und im Internet“ (<http://gaer27.uni-trier.de/MWV-online/MWV-online.html>) oder das „Wörterbuchportal“ der Berliner und Heidelberger Akademien der Wissenschaften ([www.woerterbuch-portal.de](http://www.woerterbuch-portal.de)). Andererseits haben korpuslinguistische Techniken im lexikographischen Arbeitsprozeß historischer Wörterbücher längst Einzug gehalten. So wird das Belegarchiv des GWb (mit rund 3,2 Millionen Textbelegen) laufend ergänzt durch die Arbeit mit der Datenbank *Goethes Werke auf CD-ROM*.<sup>36</sup> Es handelt sich dabei um eine SGML-kompatible Volltexterschließung des Gesamtwerkes nach der Großen Weimarer oder Sophien-Ausgabe in 143 Bänden, Weimar 1887-1919. Die Auszeichnungstiefe dieser Datenbank ermöglicht u. a. eine Kommandozeilensuche, die auch korpuslinguistischen Verfahren und Abfragen entgegenkommt.<sup>37</sup> Auf diese Weise wachsen traditionelle und korpusbasierte Lexikographie zusammen. Mit der universellen Verfügbarkeit digitaler Werkzeuge konvergieren auch Produktions- und Präsentationsverfahren, und zwar vom elektronischen Volltextkorpus auf CD-ROM bis zum Portal für alle Wörterbuchtypen.

Diese Konvergenz von Thesaurus- und Indexprinzip illustriert ein letzter Blick auf die elektronische Struktur des GWb. Innerhalb der Präsentation von ‘GWb online’ auf

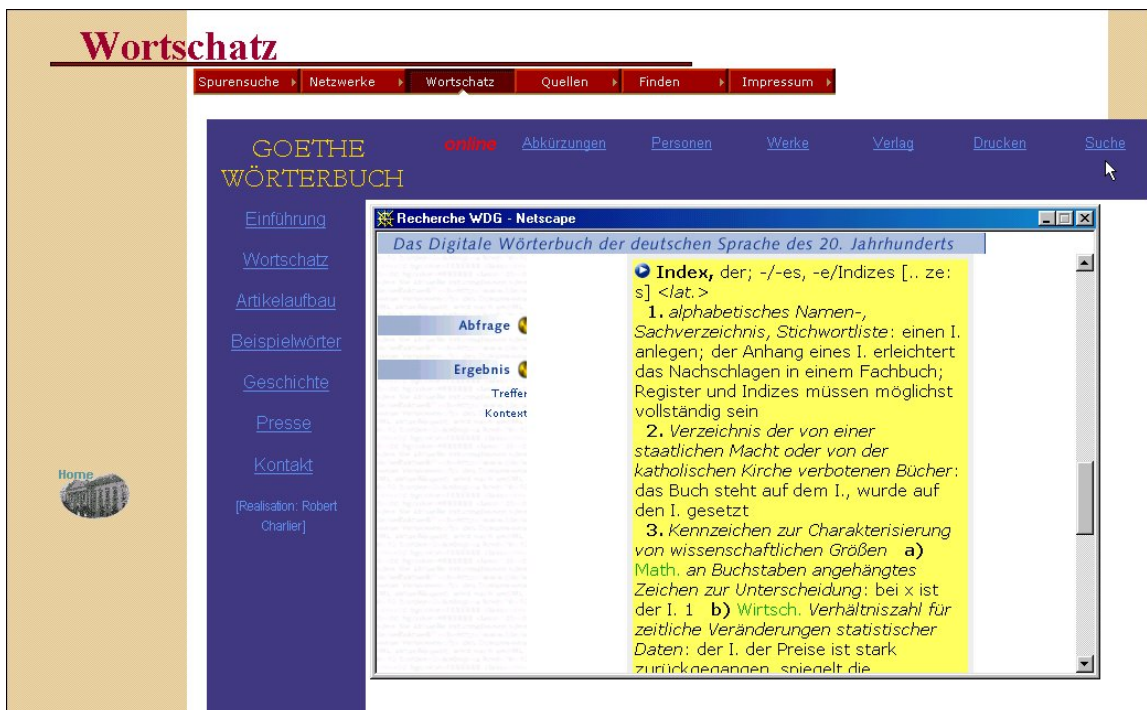
---

<sup>35</sup> Vgl. THOMAS GLONING/R. WELTER: Wortschatzarchitektur und elektronische Wörterbücher: Goethes Wortschatz und das Goethe-Wörterbuch. In: Chancen und Perspektiven computergestützter Lexikographie. Hypertext, Internet und SGML/XML für die Produktion und Publikation digitaler Wörterbücher, hg. von INGRID LEMBERG, B. SCHRÖDER und A. STORRER (Tübingen 2001) 118.

<sup>36</sup> Vgl. CHADWYCK-HEALEY 1995

<sup>37</sup> Vgl. dazu FOTIS JANNIDIS: Goethes Werke auf CD-ROM. In: Arbitrium 16 (1998) 192-201.

dem Goethe-Server der Berliner Akademiebibliothek läßt sich über den Hyperlink ‘Suche’ in der oberen Menüleiste (s. Cursorposition, Abbildung im folgenden) das *Digitale Wörterbuch der deutschen Sprache des 20. Jahrhunderts* (DWDS) ansteuern und der digitale Volltext des *Wörterbuchs der deutschen Gegenwartssprache* (WDG) virtuell ‘aufschlagen’. Das enge Ineinander von Thesaurus und Index verdeutlicht dabei der Aufruf des Artikels ‘Index’ aus dem WDG. Dieses vollständige Wörterbuch des Deutschen, von 1952 bis 1977 erarbeitet an der Deutschen Akademie der Wissenschaften zu Berlin, findet sich im Volltext restrukturiert im Korpusverbund des DWDS:



[http://bibliothek.bbaw.de/Goethe/my\\_html/wortschatz.htm](http://bibliothek.bbaw.de/Goethe/my_html/wortschatz.htm)

Als Repräsentation eines Thesaurustyp-Wörterbuchs eröffnet ‘GWb online’ zugleich den Zugang zu einer korpusbasierten lexikographischen Ressource wie dem DWDS, das ein enormes Textvolumen für die anspruchsvolle Analyse aufbereitet. Das moderne ‘historische’ WDG sorgt wiederum für vollständige semantische Ausdifferenzierung in Form eines geschlossenen, von A bis Z durchlemmatisierten Thesaurus. Dabei handelt es sich um den Wortschatz der modernen Gegenwartssprache der 1950er bis 1970er Jahre aus dem Blickwinkel der ostdeutschen Sprachwissenschaft. Das thesaurische WDG mündet damit in ein Referenzkorpus des Deutschen unter dem Dach des Indextyp-Wörterbuchs DWDS, das neben schöner und Gebrauchs-Literatur, Publizistik und

Fachtexten ganze historische Wörterbücher in seinen Vorrat von über 100 Millionen laufenden Textwörtern inkorporiert (s. unter [www.dwdsorpus.de](http://www.dwdsorpus.de)). Die Verschmelzung von Thesaurus- und Index-Wörterbuch (WDG bzw. DWDS) mit einem digital verfügbaren Referenzkorpus des Deutschen (DWDS CORPUS) verspricht, eine sinnvolle Vorform der großen lexikographischen Allsynthese von semantischer, struktureller und referentieller Worterschließung zu modellieren.

Diese Konvergenz von Thesaurus, Index und Korpus überwindet die Grenzen gedruckter Wörterbücher und rückt zumindest theoretisch die Verwirklichung eines ‘großen’, kumulativen ‘Thesaurus’ des Deutschen in den Bereich des technisch Denk- oder sogar Machbaren. Auch die Vertreter von Fachlexikographie, Begriffsgeschichte und allgemeiner Enzyklopädie erscheinen in absehbarer Zeit als enge Vernetzungspartner im Rahmen einer universellen Plattform für Historische Semantik (vermutlich als Portallösung) unverzichtbar.

Die Dimensionen dieser konvergenten Entwicklung vermag ein engerer Rückblick auf die Geschichte der historischen Lexikographie an der Berliner Wissenschaftsakademie zu erhellen. Um 1900 hatten die ersten Erfolge des *Thesaurus linguae Latinae* (TLL) eine analoge Diskussion an der Preußischen Akademie der Wissenschaften um einen ‘Thesaurus linguae Germanicae’ entzündet. Kein anderer als Herman Grimm entwickelte diesen ‘Thesaurus-Plan’ im Jahre 1893 als Erweiterung und Korrektiv für das Deutsche Wörterbuch – immerhin das Lebenswerk seines Vaters und Onkels. Das Projekt, noch 1903 von dem Altphilologen Hermann Diels postuliert, fußte bei aller Kontroverse auf einem Konsens: Es sollte sich nicht um ein gedrucktes Wörterbuch handeln.<sup>38</sup> Auch eingefleischten Verfechtern des Thesaurus-Projekts schauderte nämlich vor der „Hektometerzahl, die die Bände des Ungetüms nebeneinander gereiht füllen könnten“,<sup>39</sup> so Gustav Roethe, Vorsitzender der Deutschen Kommission, im Jahr 1913. An die Stelle des gedruckten Thesaurus trat vielmehr die Vorstellung eines großen historischen Material- und Wortarchivs, das der Forschung frei zugänglich sein sollte.<sup>40</sup> Dies allerdings eher im Sinne eines offenen historischen Wortschatzarchivs, denn auch die Textmenge von (Referenz-)Korpora verlangt

---

<sup>38</sup> Vgl. ULRICH SCHRÖTER: Von Moriz Heyne zur Deutschen Kommission. Zur Bearbeitung des Deutschen Wörterbuchs von 1867 bis 1908. In: Das Grimmsche Wörterbuch. Untersuchungen zur lexikographischen Methodologie, hg. von JOACHIM DÜCKERT (Stuttgart 1987) 116f.

<sup>39</sup> Zit. nach KARL STACKMANN: Historische Lexikographie. In: ders.: Philologie und Lexikographie. Kleine Schriften II. Hg. von J. HAUSTEIN (Göttingen 1998) 140.

<sup>40</sup> Vgl. K. STACKMANN: Historische Lexikographie, a. a. O. [Anm. 39] 138-42

aufgrund struktureller und arithmetischer Gesetzmäßigkeiten nach sinnvoller Kappung. Schließlich werden Größe und Struktur eines Textkorpus in der Regel durch die sog. Type-Token-Ratio (TTR) bestimmt. Es handelt sich dabei um den Quotienten aus der „Anzahl der unterschiedlichen Wortformen“ geteilt durch die „Anzahl sämtlicher Wörter“ eines Korpus.<sup>41</sup> Die zu lemmatisierende Grundform eines Wortes (‘Type’) liegt in der Regel in einer Vielzahl von morphologischen Erscheinungsformen vor (‘Tokens’). Wächst die Textmenge, werden neue Tokens inkorporiert, und die Anzahl neuer Types nimmt ab. Insofern konvergieren Textkorpora gegen einen kritischen Wert.<sup>42</sup>

Das beschriebene Konvergenzphänomen möchte ich abschließend durch ein Thesenbild verlebendigen. Es handelt sich dabei um die Übertragung eines naturgeschichtlichen Bildes in den Technikbereich. Als ‘evolutionsgeschichtlich’ ältere Daseinsform verkörpern demgemäß die historischen Wörterbücher die fiktive Spezies der ‘Thesaurier’. Digitale lexikographische Informationssysteme dagegen übernehmen gleichsam die Rolle imaginärer ‘Indechsen’. Während das thesaurische Prinzip die vernetzte Wissenswelt in ihrem Funktionskern modelliert, profitieren historische Korpora von den immer mächtigeren Möglichkeiten technischer Indexierung, und zwar, wie gezeigt, vom Mark-up bis zum Retrieval. Der (teil)automatisierte Index benötigt also semantische Erläuterung, der komplex kommentierte *Wortschatz* verlangt dagegen nach Umwandlung in einen abfragbaren *Datenschatz*. Thesaurier und Indechsen stehen in einem ko-evolutionären Verhältnis zueinander und optimieren sich gegenseitig.

---

<sup>41</sup> M. MEYER: Korpusbasierte Erstellung eines Wörterbuchs, a. a. O. [Anm. 24] 15.

<sup>42</sup> M. MEYER: Korpusbasierte Erstellung eines Wörterbuchs, a. a. O. [Anm. 24] 16.

## VII. Quellen

### A. allgemeine Lexikographie

#### 1. Wörterbücher, neuere deutsche Wortschatzgeschichte, Thesaurustyp

CHRISTOPH ERNST STEINBACH: Vollständiges Deutsches Wörter-Buch, 2 Bde, Breslau 1734

JOHANN LEONHARD FRISCH: Teutsch-Lateinisches Wörter-Buch, Berlin 1741

JOHANN CHRISTOPH ADELUNG: Grammatisch-kritisches Wörterbuch der hochdeutschen Mundart mit beständiger Vergleichung der übrigen Mundarten, besonders aber der Oberdeutschen. 5 Bde, Leipzig <sup>2</sup>1793-1801 [1774-86]

JOACHIM HEINRICH CAMPE: Wörterbuch der deutschen Sprache. 5 Bde, Braunschweig 1807-11

JOACHIM HEINRICH CAMPE: Wörterbuch zur Erklärung und Verdeutschung der unserer Sprache aufgedrungenen fremden Ausdrücke. Ein Ergänzungsband zu Adelung's und Campe's Wörterbüchern. Neue stark vermehrte und durchgängig verbesserte Ausgabe. Braunschweig <sup>2</sup>1813 [1801]

JOHANN CHRISTIAN AUGUST HEYSE: Handwörterbuch der deutschen Sprache mit Hinsicht auf Rechtschreibung, Abstammung und Bildung, Biegung und Fügung der Wörter, sowie auf Sinnverwandtschaft. Nach den Grundsätzen einer Sprachlehre angelegt. 3 Bde, Magdeburg 1833-49

JOHANN CHRISTOPH ADELUNG: Vollständige Anweisung zur deutschen Orthographie, nebst einem kleinen Wörterbuch. 2 Teile, Leipzig <sup>2</sup>1835 [1788]

Kleines Wörterbuch der deutschen Sprache für die Aussprache, Rechtschreibung, Beugung und Ableitung, nebst der Erklärung aller grammatischen Benennungen und der Verdeutschung vieler fremder Wörter. Von J. Ch. ADELUNG. Sechste, nicht bloß vermehrte und berichtigte, sondern ganz umgearbeitete und als ein völlig neues Werk erscheinende, Ausgabe, verfaßt von K. B. Schade, Leipzig 1834

Deutsches Wörterbuch VON JACOB GRIMM und WILHELM GRIMM. 16 Bde, Leipzig 1854-1971 [DWb]

Deutsches Wörterbuch von J. und W. GRIMM. Neubearbeitung. Hg. von der Akademie der Wissenschaften der DDR (ab 1993 [Bd. 2, Lfg. 4 bzw. Bd. 7, Lfg. 8] von der

Berlin-Brandenburgischen Akademie der Wissenschaften) in Zusammenarbeit mit der Akademie der Wissenschaften zu Göttingen. Leipzig 1983ff. [zugleich als partieller digitaler Prototyp unter: [www.dwb-digital.adw-goettingen.gwdg.de/WebDwb/index.html](http://www.dwb-digital.adw-goettingen.gwdg.de/WebDwb/index.html)]

DANIEL SANDERS, Handwörterbuch der deutschen Sprache, 2 in 3 Bdn, Leipzig 1869

MORIZ HEYNE, Deutsches Wörterbuch. 3 Bde, Leipzig 1890-95

HERMANN PAUL, Deutsches Wörterbuch, Halle 1897

Wörterbuch der deutschen Gegenwartssprache von RUTH KLAPPENBACH und WOLFGANG STEINITZ, hg. von der Deutschen Akademie der Wissenschaften zu Berlin [seit 1972: Akademie der Wissenschaften der DDR], 6 Bde, Berlin 1961-77 [WDG]

## 2. *Online-Ressourcen, (korpusbasierter) Indextyp*

CISLEX (Projekt eines vollständigen elektronischen Wörterbuchs des Deutschen):

<http://www.cis.uni-muenchen.de/projects/CISLEX.html>

COSMAS I UND II (s. Institut für deutsche Sprache)

DEUTSCHES WÖRTERBUCH VON JACOB UND WILHELM GRIMM AUF CD-ROM UND IM

INTERNET (Retrodigitalisierung der Erstbearbeitung mit Gliederungsansicht):

[www.dwb.uni-trier.de/index.html](http://www.dwb.uni-trier.de/index.html) (s. auch DER DIGITALE GRIMM)

DER DIGITALE GRIMM. Deutsches Wörterbuch. Elektronische Ausgabe der Erstbearbeitung von Jacob Grimm und Wilhelm Grimm. Hg. vom Kompetenzzentrum für elektronische Erschließungs- und Publikationsverfahren in den Geisteswissenschaften an der Universität Trier in Verbindung mit der Berlin-Brandenburgischen Akademie der Wissenschaften. 2 CD-ROMs sowie Begleit- und Benutzerhandbuch. Frankfurt/M. (Zweitausendeins) 2004:

[www.Zweitausendeins.de/Grimm-Forum](http://www.Zweitausendeins.de/Grimm-Forum)

DWDS (= Digitales Wörterbuch der deutschen Sprache des 20. Jahrhunderts; Korpus, Online-Wörterbuch, Kollokationsprojekt): <http://www.dwds.de>

DWDS CORPUS (Textbasis des DWDS, Referenzkorpus des Deutschen):

<http://www.dwds.corpus.de>

ELEXIKO – WISSEN ÜBER WÖRTER (lexikographisches Informationssystem des IDS

Mannheim, s. auch INSTITUT FÜR DEUTSCHE SPRACHE): <http://www.ids-mannheim.de/elexiko/>

GERMANET (lexikalisch-semantisches Netzwerk): <http://www.sfs.nphil.uni-tuebingen.de/lsd>

INSTITUT FÜR DEUTSCHE SPRACHE, Mannheim (diverse Korpora, Analysemodule):  
<http://corpora.ids-mannheim.de>

KLASSIKERWORTSCHATZ, Freiburg (korpusbasiertes historisches Wörterbuchprojekt als Differenzwortschatz): <http://www.klassikerwortschatz.uni-freiburg.de>

LIMAS (Quellenkorpus): <http://linux-s.ikp.uni-bonn.de/Limas/index.htm>

NEGRA CORPUS VERSION 2 (Textsortenkorpus): [www.coli.uni-sb.de/sfb378/negra-corpus/negra-corpus.html](http://www.coli.uni-sb.de/sfb378/negra-corpus/negra-corpus.html)

WÖRTERBUCHPORTAL der Berliner und Heidelberger Akademien der Wissenschaften (Korpus-, Online-Wörterbuchverbund): <http://www.woerterbuch-portal.de>

WORTSCHATZLEXIKON (laufend aktualisierte Wortschatz-Datenbank, Online-Wörterbuch): <http://wortschatz.uni-leipzig.de>

## B. Goethe-Lexikographie

### *1. Thesaurustyp*

DILL, CHRISTA: Wörterbuch zu Goethes West-östlichem Divan. Tübingen 1987

FISCHER, PAUL: Goethe-Wortschatz. 2 Teile. Leipzig 1929

GOETHE-WÖRTERBUCH. Hg. von der Berlin-Brandenburgischen Akademie der Wissenschaften [bis Bd. 1, 6. Lfg.: Deutsche Akademie der Wissenschaften zu Berlin; bis Bd. 3, 4. Lfg.: Akademie der Wissenschaften der DDR], der Akademie der Wissenschaften in Göttingen und der Heidelberger Akademie der Wissenschaften. 3 Bde. Stuttgart; Berlin; Köln [Mainz]: 1966ff. [bis Band 4, Lieferung 11: Hinzutritt-Hosenknopf, 2004] [GWb]

MERKER, ERNA (u. a.): Wörterbuch zu Goethes Werther, begründet von E. Merker in Zusammenarbeit mit J. Graefe und F. Merbach, fortgeführt und vollendet von I. Engel, J. Graefe, E. Linke, J. Mattausch, F. Merbach. Berlin 1966

NEUENDORFF-FÜRSTENAU, Jutta: Wörterbuch zu Goethes Götz von Berlichingen. Lfg. 1 und 2 (Aal-Glück). Berlin 1958-63

STREHLKE, FRIEDRICH: Wörterbuch zu Goethe's Faust. Stuttgart; Leipzig; Berlin; Wien 1891

### *2. Indextyp*

- HOHLFELD, A. R. HOHLFELD/M. JOOS/W. F. TWADDELL: Wortindex zu Goethes Faust. Madison (Wisconsin) 1940
- KITTEL, PAULA M./N. FUERST: Der Wortschatz der Bühnenprosa in Goethes 'Faust'. Ein Nachtrag zum 'Wortindex zu Goethes Faust'. Madison (Wisconsin) 1946
- MOMMSEN, KATHARINA (Hg.): Verskonkordanz zu Goethes 'Faust, Erster Teil'. Bearbeitet von Steven P. Sondrup und David Chisholm (= Indices zur deutschen Literatur; 18). Tübingen 1986
- SCHMIDT, PETER: Der Wortschatz von Goethes 'Iphigenie'. Analyse der Werk und Personensprache mit EDV-Hilfe. Mit Wortindex, Häufigkeitwörterbuch und Wortgruppentabellen (= Indices zur deutschen Literatur). Frankfurt/M.; Bonn 1970
- Wörterbuch zu Friedrich Hölderlin [...] auf der [Text-]Grundlage der Großen Stuttgarter Ausgabe. I. Teil: Die Gedichte. Bearbeitet von H.-M. DANNHAUSER und K. SCHUFFELS.; II. Teil: Hyperion. Bearbeitet von H. O. HORCH, K. SCHUFFELS und M. KAMMER (= Indices zur deutschen Literatur; 10/11 bzw. 19). Tübingen 1983 bzw. 1992
- WÖRTERREGISTER zu Goethes Rezensionen in den Frankfurter Gelehrten Anzeigen, hg. von der Deutschen Akademie der Wissenschaften zu Berlin. [masch. Ms.] Leipzig 1959, S. I-VII

### C. Digitale Goethe-Korpora

- KARL EIBL/F. JANNIDIS/ M. WILLEMS (Hgg.): Der junge Goethe in seiner Zeit. Texte und Kontexte in 2 Bdn. und einer CD-ROM. Frankfurt a. M.; Leipzig 1998
- Goethes Werke (Weimarer Ausgabe) auf CD-ROM. Vollständige elektronische Edition im Sinne eines Nachdruckes der sog. Sophien-Ausgabe der Werke Johann Wolfgang Goethes. Cambridge 1995 [CHADWYCK-HEALEY]
- FOTIS JANNIDIS: Goethes Werke auf CD-ROM [= Rezension von CHADWYCK-HEALEY]. In: Arbitrium 16 (1998), S. 192-201