

Überlegungen zu den Quellen für ein historisches Wörterbuch an der Schwelle zum digitalen Zeitalter

1. Die Qualität von Wörterbüchern hängt von der Ergiebigkeit der ausgewerteten Quellen ab

Wenn man über Quellen für ein Wörterbuch der Zukunft redet, denkt man heute sicher zunächst an tendenziell unbegrenzte Möglichkeiten, digitale Korpora mithilfe von ausgefeilten Werkzeugen zu einem ganz neuartigen Auskunftsmittel zu formen. Auch wenn tatsächlich intensiv an dieser Vision geforscht und gearbeitet wird, weicht die Realität selbst der modernsten historisch arbeitenden Wörterbücher in vielerlei Hinsicht von einem Idealzustand ab, nicht zuletzt deshalb, weil die benötigten finanziellen Mittel weniger reichlich fließen als Ideen und Pläne.

Folgen wir deshalb ganz pragmatisch dem Gedanken, wie die Arbeit an einem historisch arbeitenden Wörterbuch aussehen könnte, wenn sich die Korpusituation, so wie sie sich im Moment darstellt, nicht sofort in einen Idealzustand versetzen lassen würde – und schreiben dazu die am Deutschen Wörterbuch bereits laufende Arbeit in die nahe Zukunft fort. Ein historisches Wörterbuch – und historisch verstehe ich vor allem als Methode unabhängig vom tatsächlichen Alter der behandelten Erscheinungen – wird nur so gut sein wie die Ressourcen, die ihm zur Verfügung stehen, und mit deren Entwicklung verschiedene Qualitätssprünge erleben.

Derartige Brüche, die auch aus einer veränderten Quellenlage folgten, sind für die Vergangenheit des Deutschen Wörterbuches bereits gut dokumentiert: „ihrer natur nach können bücher dieser art erst gut werden bei zweiter auflage“¹, formuliert bekanntermaßen Jacob Grimm, und seit ihrer Wiederentdeckung vor einigen Jahren kennen wir seine Sammlung an Nachträgen und Zusätzen in eigens für die Brüder Grimm mit einem breiten Korrekturrand versehenen Handexemplaren ihres Wörterbuches.²

1 Jacob Grimm am 21. Dezember 1861 an Karl Weigand, in: Kirkness, Alan (Hg.). 2010. Briefwechsel der Brüder Jacob und Wilhelm Grimm mit Rudolf Hildebrand, Matthias Lexer und Karl Weigand, S. 384.

2 Dazu Kirkness, Alan. 2006. Die Grimmschen Handexemplare des Deutschen Wörterbuchs A–F. In: Zeitschrift für Germanistische Linguistik 34, 207–209.

Die erste Auflage des Deutschen Wörterbuches ist bekannterweise bemerkenswert heterogen, und das hat im wesentlichen zwei Ursachen: Zum einen die sich über den langen Entstehungszeitraum hinweg stark wandelnden Vorlieben bei der Präsentation lexikographischer Ergebnisse – und zum anderen die Qualität des verwendeten Korpus: Erst im Laufe der Zeit hat sich durch immer neue und kritische Editionen älterer Quellen, durch breitere Streuung der Exzerpte und die wachsende Menge des gesammelten Materials die Möglichkeit eines zuverlässigen Überblicks über den Wortschatz eröffnet.

Manche Lücke in der Stichwortreihe der ersten Auflage des Deutschen Wörterbuches steht freilich für eine überlegte Auswahl von Stichwörtern. So sucht, wie wir wissen, Jacob Grimm etwa in der Fremdwortfrage einen Kompromiß: „dieser ausländerei und sprachmengung soll das wörterbuch keinen vorschub, sondern will ihr allen redlichen abbruch thun, geflissentlich aber auch die abwege meiden, auf welche von unberufenen sprachreinigern gelenkt worden ist“.³ Das erklärt, warum wir beispielsweise mit genauen Ausführungen zu einem so entlegenen Wort wie *Mandelkäse*⁴ versorgt werden, dessen Bestandteile zwar entlehnt sind, deren Herkunft aber längst phonetisch und morphologisch verschleiert war, dagegen aber die bloße Erwähnung von zum damaligen Zeitpunkt ähnlich lange in der deutschen Sprache heimischen, aber deutlich frequenteren Wörtern wie *Marmelade*⁵ vermissen.

Von derartigen Fällen abgesehen kann aber eben nicht ausgeschlossen werden, daß so manche Erscheinung auch aufgrund der schlechten Beleglage in den hauptsächlich literarischen Quellen der ersten Auflage keine Diskussion gefunden hat. Bis zum Aufkommen geeigneter Datenbanken ist im übrigen auch bei der Neubearbeitung des Deutschen Wörterbuches weitgehend nur beschrieben worden, was im eigenen Archiv ausreichend belegt werden konnte. Niemand hätte anders verfahren können; Lücken mußten notgedrungen in Kauf genommen werden. Nachdem große digitale Korpora zunehmend auch für historische Sprachstufen zur Verfügung stehen, läßt sich allerdings sagen: Ein modernes Wörterbuch ist heute ohne digitale Hilfsmittel undenkbar geworden.

3 ¹DWB 1,XXVIII.

4 ¹DWB 6,1536.

5 Vgl. Jones, William Jervis. 1976. A Lexicon of French Borrowings in the German Vocabulary (1575–1648), S. 425.

2. Ein modernes historisch arbeitendes Wörterbuch ist ohne digitale Hilfsmittel undenkbar

Lassen Sie mich das für die Neubearbeitung an drei Beispielen illustrieren:

Beispiel 1: Arbeitsvermögen

Die Neubearbeitung des Deutschen Wörterbuches stützt sich für die ersten Buchstaben des Alphabets auf ein in seiner Art vom Umfang her wahrscheinlich beispielloses Korpus: Trotzdem, und das ist eine häufig gemachte Erfahrung, ist es zu klein, um Zweifelsfälle aller Art sorgfältig klären und erklären zu können. Ist ein Wort etwa spät und spärlich⁶ belegt und scheint die Bildung unproblematisch zu sein, wird es oft schlicht nicht aufgenommen, wie seinerzeit das *Arbeitsvermögen*. Zwar ist das Wort im Belegarchiv dokumentiert, und zwar mit einigen Nachweisen seit 1894⁷. Allerdings hat das Wort, wie die Suche in den digitalen Textsammlungen von Google Books⁸ und Deutschem Textarchiv⁹ zeigt, eine im alles umstürzenden 19. Jahrhundert bereits viel längere Geschichte hinter sich, tatsächlich sogar bis in das 18. Jahrhundert hinein, eine Begriffsgeschichte, die sich bei so wichtigen Autoren wie Helmholtz oder Marx widerspiegelt. *Arbeitsvermögen* bezeichnet schon früh ganz unterschiedliche Phänomene im Bereich der Physik¹⁰ und der Ökonomie¹¹. Kaum etwas von den terminologischen Unsicherheiten und Überschneidungen des 19. Jahrhunderts findet sich allerdings in den Sammlungen des Deutschen Wörterbuches: Die Entwicklung des Wortes war für die Bearbeiterin mithin weitgehend unsichtbar.

Beispiel 2: Bärlauch

Anders etwa zehn Jahre später bei *Bärlauch*. Die Pflanze findet sich gegenwärtig in vielen Lebensmitteln – nichtsdestotrotz ist für das Deutsche Wörterbuch seinerzeit kein einziger Beleg dafür gesammelt worden. Offenkundig liegt der Grund dafür in der Quellenauswahl, die eben keinen Schwerpunkt in den Gebrauchstexten besitzt, und auch das Korpus des Deutschen Textarchivs

6 Gemäß einem Straffungskonzept werden in der Berliner Arbeitsstelle in der Regel nur Wörter aufgenommen, die nach der Stichwortliste ca. 30 Belege von unterschiedlichen Autoren aufweisen.

7 Bei Lueger, Otto. 1894–99. Lexicon der gesamten Technik und ihrer Hilfswissenschaften 1, S. 440.

8 <http://books.google.de/>

9 <http://www.deutschestextarchiv.de/>

10 Vgl. Online-Lexikon der Physik (Spektrum Akademischer Verlag) unter <http://www.wissenschaft-online.de/abo/lexikon/physik/729>

11 Vgl. Gabler Wirtschaftslexikon zum Stichwort *Humankapital* unter <http://wirtschaftslexikon.gabler.de/Archiv/54658/humankapital-v3.html>

bleibt uns wahrscheinlich aus demselben Grund einen Nachweis schuldig. Glücklicherweise sind die Hilfsmittel inzwischen so weit gediehen, daß die Geschichte des Wortes dennoch schnell offenkundig: Man findet es als Fortsetzer von *Bärenknoblauch* zuerst im 18. Jahrhundert, wo es weiträumig vor allem das ältere *Ramse* abzulösen beginnt. Ohne die fortgeschrittene Digitalisierungstätigkeit von Google und der beteiligten Bibliotheken¹² wahrscheinlich unbekannt gebliebene Belege konnten nun extrahiert werden, um überhaupt erst einen Artikel anlegen zu können.

Beispiel 3: Befindlichkeit

Mit den sich entwickelnden digitalen Hilfsmitteln können wir nicht nur früheste Vorkommen aufspüren, sondern auch irritierende Beleglücken schließen und die Geschichte des Wortes durch neue Erkenntnisse bereichern. Bisher war beispielsweise wenig bekannt, daß das in heutigen Diskursen so beliebte *Befindlichkeit* zwischen dem Gebrauch bei den spätmittelalterlichen Mystikern, ihren späteren Nachfolgern im 17. Jahrhundert wie Jakob Böhme und erst wieder viel später bei Martin Heidegger überhaupt eine Geschichte besitzt. Johannes Erben, der zu einem lexikologischen Beitrag¹³ die Belege des Deutschen Wörterbuchs verwendet hatte¹⁴, kann hier deshalb auch nur eine Lücke annehmen. Nach einer Suche in den heute zugänglichen digitalen Sammlungen läßt sich konstatieren, daß weder der schon bei Erben beschriebene und für unseren heutigen Gebrauch grundlegende Gebrauch bei Heidegger¹⁵ im Archiv des Deutschen Wörterbuchs belegt ist, noch die Tatsache, daß für die Jahrhunderte davor sogar verschiedene Bedeutungen des Wortes überliefert sind¹⁶, auch wenn das Wort nicht ausgesprochen häufig ist. Wir finden deutliche Hinweise darauf, daß – und daher stammt wahrscheinlich auch die entscheidende Anregung für Heidegger – die Beschäftigung mit der Sprache der Mystiker auch im 19. Jahrhundert noch lebendig gewesen ist¹⁷.

Insbesondere eine große Sammlung wie Google Books bietet also Erkenntnismöglichkeiten, auf die kaum noch ernsthaft verzichtet werden kann: Ein modernes historisches Wörterbuch ist ohne digitale Hilfsmittel nicht mehr denkbar.

12 Der älteste ermittelte Beleg stammt aus einem Exemplar der Bayerischen Staatsbibliothek. Trotz dieser Recherchemöglichkeiten bleibt freilich die lexikographische Aufbereitung des Themas wie hier etwa durch Heinrich Marzell (Wörterbuch der deutschen Pflanzennamen, Leipzig, 1943–1979) eine unschätzbare und einstweilen nicht ersetzbare Hilfe.

13 Erben, Johannes. 2000. *Befindlichkeit* – Anmerkungen zur Wortbildung, Semantik und pragmatischen Sprachgeschichte. In: Mechthild Habermann, Peter O. Müller und Bernd Naumann (Hgg.): Wortschatz und Orthographie in Geschichte und Gegenwart, S. 171–179, insbes. S. 176 f.

14 Ders., S. 177.

15 Heidegger, Martin. 1927. *Sein und Zeit*; vgl. etwa S. 134.

16 Gut belegt ist insbesondere eine lokale *Befindlichkeit an, bei* usw.; vgl. z. B. Matthies, Konrad Stephan. 1836. *Propädeutik der Neutestamentlichen Theologie*, S. 142.

17 Vgl. Mundt, Theodor. 1837. *Die Kunst der deutschen Prosa*, S. 170.

3. Ein modernes historisch arbeitendes Wörterbuch kann von nichtelektronischen Quellen profitieren

Werfen wir einen Blick auf den umgekehrten Fall: Im Moment darf man noch davon ausgehen, daß ein modernes Wörterbuch bei allem Vorzug digitaler Quellen nach wie vor von nichtelektronischen Quellen profitiert. Um ein zuverlässiges Nachschlagewerk zu schreiben, empfiehlt es sich nämlich, alle verlässlichen Quellen zu nutzen, die erhältlich sind und die sinnvoll verarbeitet werden können. Solange bestimmte Quellen noch nicht in elektronischer Form zur Verfügung stehen oder die Werkzeuge zu deren Nutzung noch nicht für lexikographische Zwecke optimiert sind, liegt es nahe, auch traditionelle Quellen zu verwenden.

Zum einen wird man immer wieder feststellen, daß auch bei der Nutzung von großen Korpora, wenn diese nicht ständig mit aktuellen Texten versorgt werden, gerade auf die neuesten Belege verzichtet werden müßte. Das ist vor allem dann bedauerlich, wenn man damit auch zentral wichtige Texte unausgewertet lassen müßte, die beispielsweise aufgrund der Rechtelage einerseits und der stets angespannten finanziellen Situation von Wörterbüchern andererseits für eine digitale Verarbeitung auf absehbare Zeit nicht zur Verfügung stehen werden.

Zum anderen werden sich bestimmte Textsorten, die auch in großen Korpora und selbst im Internet unterrepräsentiert sind, ohne traditionelle Mittel – oder jedenfalls ohne Hilfe von außen – nur eingeschränkt auswerten lassen. Das gilt zum Beispiel für Texte aus der Werbung, auf Verpackungen, für Rechnungen, Fahrkarten usw., aber auch für solche, die man für eine elektronische Exzerption vielleicht nie in Betracht gezogen hätte, wie etwa entlegene Lokalblätter, die aber für die Entstehung eines Sprachgebrauches hochinteressante Auskünfte liefern könnten.

Gerade hier finden sich übrigens Ansatzpunkte, mit einer interessierten Öffentlichkeit in Kontakt zu kommen und auch ihre Hilfe in Anspruch zu nehmen: De facto verzichten große Wörterbücher ohnehin nicht darauf, neue Belege nachzuexzerpieren. Die einen erledigen diese Arbeit mit Hilfe von Praktikanten und studentischen Hilfskräften, die anderen beziehen interessierte Außenstehende ein. Letztlich ist auch ein großes Korpus nur ein begrenzt nutzbares Hilfsmittel, wenn man noch nicht weiß, wonach zu suchen ist, und auch der Hinweis auf eine interessante Entwicklung ist nur wenig wert, wenn diese sich nicht sicher nachvollziehen läßt.

Vor diesem Hintergrund erstaunt es auch nicht zu sehen, daß das mit dem größten Aufwand produzierte historisch arbeitende Wörterbuch, das Oxford English Dictionary, zu den Wörterbüchern gehört, bei denen ein Zettelarchiv trotz aller zur Verfügung stehenden digitalen Hilfsmittel noch eine wichtige Rolle spielt.¹⁸ Tatsächlich sind Arbeitsweise und Arbeitsorganisation in den verschiedenen großen Wörterbüchern nicht fundamental verschieden – was sich beobachten läßt, sind lediglich verschiedene Stadien einer natürlichen Entwicklung hin zu immer stärkerer Nutzung digitaler Möglichkeiten. Letztlich geht es überall darum, einen immer möglichen Konflikt zwischen zwei Zielen aufzulösen: dem Ziel einerseits nämlich, sein Wörterbuch nicht nur elektronisch zu präsentieren, was weitgehend für selbstverständlich gehalten wird, sondern dieses auch ausschließlich elektronisch zu erarbeiten – und dem Ziel andererseits, zugleich auch ein zuverlässiges Nachschlagewerk zu sein und zu bleiben. Will man beides in einem Gleichgewicht halten, wird man pragmatische Entscheidungen treffen müssen, jedenfalls aber keine Quellenart von vornherein ausschließen wollen, sei sie digital oder nichtdigital. In jedem Fall werden flexible Optionen in einem Redaktionssystem¹⁹ erforderlich sein, die im übrigen rechtzeitig angelegt werden müssen, um später nicht zur Improvisation gezwungen zu sein.

4. Ein modernes historisches Wörterbuch bleibt für die früheren Sprachstadien teilweise noch auf traditionelle Hilfsmittel angewiesen

Die Arbeit an einem historischen Wörterbuch wird gegenwärtig dadurch erschwert, daß uns bisher große, unbeschränkt zugängliche und leicht auswertbare elektronische Korpora für die älteren Sprachstufen noch nicht zur Verfügung stehen, ein Problem, das dadurch verstärkt wird, daß die großen Sprachstadienwörterbücher, das Althochdeutsche²⁰, das Mittelhochdeutsche²¹ und das Frühneuhochdeutsche Wörterbuch²² sämtlich noch nicht abgeschlossen sind und mittelfristig auch nicht abgeschlossen sein werden. Noch in den nächsten Jahren wird deshalb lexikographische Arbeit wenigstens zur Zeit vor dem 16. Jahrhundert notwendigerweise auch nichtelektronische Quellen einschließen müssen.

Derzeit sind erste Grundlagen für annotierte Korpora von Texten der einzelnen Sprachstufen bereits

18 Instruktiv dazu beispielsweise der für das OED produzierte Kurzfilm *Writing the OED* unter <http://www.youtube.com/user/OUPAcademic#p/u/9/v9AgtJTVuuM>

19 Vgl. dazu den Beitrag von Elke Gehweiler.

20 Althochdeutsches Wörterbuch. 1968 ff. Berlin: Akademie-Verlag.

21 Mittelhochdeutsches Wörterbuch. 2006 ff. Stuttgart: Hirzel.

22 Frühneuhochdeutsches Wörterbuch. 1989 ff. Berlin [u. a.]: de Gruyter.

gelegt: Zu den schon seit längerem verfügbaren ausschnittshaften Textsammlungen und etwa der umfangreichen indogermanistischen Datenbank des Titus-Projektes²³ werden sich in den nächsten Jahren Referenzkorpora zum Althochdeutschen²⁴, Mittelhochdeutschen²⁵ und Frühneuhochdeutschen²⁶ gesellen. Schon heute ist ein in Nachbarschaft zum Mittelhochdeutschen Wörterbuch²⁷ als Mittelhochdeutsches Textarchiv²⁸ aufgearbeitetes umfangreiches Korpus mittelhochdeutscher Texte ebenso verfügbar wie ein kleiner Teil der im Bonner Frühneuhochdeutsch-Korpus versammelten Texte²⁹.

Derzeit noch verbleibende Lücken ließen sich auf unterschiedliche Weise mehr oder weniger gut schließen: Zum einen befinden sich im Archiv des Deutschen Wörterbuches umfangreiche Sammlungen zu den Buchstaben A – F, die vor allem das bisher besonders schlecht elektronisch erreichbare Frühneuhochdeutsche relativ dicht erfassen und sich darüber hinaus bis in das 11. Jahrhundert erstrecken. Bei der Exzerption einiger wichtiger Quellen sind außerdem Stichwörter aus dem gesamten Alphabet erfaßt worden, so etwa für eine große Zahl von Luther-Schriften. Solange diese Werke nicht digital bearbeitet werden können, wird das Archiv also nach wie vor eine ernstzunehmende Arbeitsgrundlage darstellen. Zum anderen ist die Sprache der wichtigsten Schriftsteller der frühesten Zeit durch Glossare relativ gut erschlossen. Insbesondere ist der Wortschatz des Althochdeutschen über Glossare³⁰ und natürlich über Wörterbücher auch außerhalb des an der Sächsischen Akademie der Wissenschaften bearbeiteten Althochdeutschen Wörterbuches nahezu vollständig zu überschauen.

Würde man die Arbeit des Deutschen Wörterbuches also in die nächste Zukunft fortschreiben, wäre kein plötzlicher Bruch zu erwarten, sondern ein kontinuierliches Voranschreiten hin zu einem immer stärkeren Einsatz digitaler Mittel – so wie es verschiedene vergleichbare Wörterbücher zeigen, die jeweils ganz unterschiedliche Positionen in diesem Kontinuum besetzen. Wenn auch perspektivisch der Nutzen eines Zettelarchivs und gedruckter Wörterbücher zurückgehen wird, behalten die traditionellen Mittel der historischen Lexikographie für eine Übergangszeit durchaus ihre Funktion.

23 <http://titus.uni-frankfurt.de/indexe.htm>

24 <http://clients.designato.de/humboldt/ddd/home/index.html>

25 http://www.linguistics.ruhr-uni-bochum.de/~dipper/project_ddd.html

26 http://www.germanistik.uni-halle.de/forschung/altgermanistik/referenzkorpus_fruhneuhochdeuts/

27 <http://www.mhdwb-online.de/>

28 http://mhgta.uni-trier.de/katalogsuche.php?ses_id=83139e0bf22ddfa6d33f836c5c0c2935

29 <http://www.korpora.org/FnhdSuche/>; eine Übersicht über alle in diesem Projekt gesammelten Texte bei Hoffmann, Walter und Wetter, Friedrich (Bearb.). 1987. Bibliographie frühneuhochdeutscher Quellen. 2. Aufl. Frankfurt [u. a.]: Lang

30 Beispielhaft dafür seien genannt: Schützeichel, Rudolf (Hg.). 2004. Althochdeutscher und Altsächsischer Glossenwortschatz. Tübingen: Niemeyer.; Sehrt, Edward und Starck, Taylor. 1955. Notker-Wortschatz. Halle/S.: Niemeyer.

Die Schaffung digitaler Quellenwerke muß dem Wandel der Arbeitsmethoden vorangehen. Wichtigstes Ziel auch beim Wechsel der Medien bleibt der Zuwachs an wissenschaftlicher Erkenntnis.