

WOLFGANG SCHENKEL

Wörterbuch vs. Textkorpus

oder:

Wie und ob man überhaupt ein Wörterbuch machen kann*

Was die Ägyptologie braucht, ist ein Wörterbuch und ein Textkorpus. Um es deutlicher zu sagen: Was die Ägyptologie braucht, ist ein Handwörterbuch und eine Textdatenbank. Beides ist notwendig, aber es kann nicht beides in einem Zug produziert werden.

Unter Handwörterbuch verstehe ich eine Kompilation, in der das gesamte lexikalische Wissen der heutigen Ägyptologie gebucht ist, und zwar in einer so kompakten Weise, daß man eine solche Kompilation bequem zur Hand nehmen kann. Es ist ein Referenzwerk, das einem fortgeschrittenen Studenten zugänglich sein soll und dem Allgemein-Ägyptologen den Alltag erleichtert, das aber auch dem erfahrenen Forscher immer noch Wesentliches zu bieten hat. Es sollte aktuellere und mehr Information als der Erman/Grapow enthalten, wäre aber dadurch handlich zu halten, daß man nicht so viel Papier unbedruckt läßt. Ich stelle mir so etwas vor wie von Sodens „Akkadisches Handwörterbuch“. Dieses Handwörterbuch muß alles Wesentliche enthalten über

- die Lautformen der Lexeme, einschließlich ihrer Flexionsformen
 - die Grundbedeutungen und ihre Verzweigungen (Wortbildung)
 - die hamitosemitischen Etymologien
 - die demotischen Äquivalente (sofern sich das Wörterbuch weiterhin auf die hieroglyphische und hieratische Überlieferung beschränkt)
 - die koptischen Äquivalente
 - die hieroglyphischen Schreibungen, auch die Besonderheiten der Schreibungen der Flexionsformen
 - die syntaktischen Parameter wie namentlich die Rektionen
 - den Belegzeitraum
 - die Angabe der Textsorten, in denen das Lexem gebraucht wird
 - fallweise solcherlei veranschaulicht an einer guten Auswahl signifikanter Text-Zitate
 - nicht zuletzt die Angabe solcher Sekundärliteratur, die einen wesentlichen Beitrag zur Erhellung des lexikalischen Befundes geleistet hat
- (nicht weniger, allenfalls mehr).

Demgegenüber ist eine Textdatenbank, ein reines Forschungsinstrument, ein Instrument des spezialisierten Forschers. Sie leistet die Erfassung und Aufbereitung von Texten, letzten Endes aller überlieferten Texte mit allen Textzeugen, und ist so organisiert, daß der Forscher zur Klärung von Spezialfragen gezielt darin recherchieren kann. Ein solches Instrument kann in praktisch beliebigen Stufen der Vollkommenheit Nutzen bringen. Es kann bereits Nutzen erbringen, wenn nur ausgewählte Textkorpora erfaßt werden. Es kann bereits Nutzen bringen, wenn nur ausgewählte Aspekte der Texte erschlossen werden, z.B. bereits dann, wenn nur Lexeme in Transkription abfragbar sind. Die Arbeit mit einer Textdatenbank erfordert dann

* Für die Schriftform leicht adaptierte Fassung eines Vortrags auf der Internationalen Arbeitstagung „Ägyptisches Wörterbuch“, Berlin 3.–5. 9. 1992.

aber die Findigkeit eines erfahrenen Forschers. Er muß, ohne das Textkorpus wirklich überblicken zu können, Fragen an das Textkorpus zu formulieren imstande sein und aus der Textdatenbank Tatbestände herauszaubern können, die tatsächlich signifikant sind. Eine Textdatenbank ist nicht in usum delphini bestimmt.

In Berlin hat man sich einmal dazu entschieden, ein Wörterbuch der ägyptischen Sprache auf der Basis eines Textkorpus zu erarbeiten, Textkorpus- und Wörterbucharbeit miteinander zu verbinden. Wie wenig realistisch die an dieses Verfahren geknüpften Erwartungen waren, ist bekannt. Um mich an Rykle Borger's „Altorientalische Lexikographie“¹ anzulehnen: „Die Anhäufung des Materials erwies sich als leichter denn die Umsetzung dieses Materials in ein Wörterbuchmanuskript.“² Es war klar, daß, während die vorgängige Sammelarbeit auf breiter Personalbasis durchgeführt werden konnte, „diese Umsetzung nur von den eigentlichen Trägern des Unternehmens durchgeführt werden konnte, und so wurde ein bottle-neck unvermeidlich“³. Jetzt nämlich war „die für den Lexikographen so wichtige Kunst des Auswählens und des wohl erwogenen Weglassens“⁴ gefragt. Und, um es in Worten John Tait's (in: *Approaches to Demotic Lexicography*)⁵ zu sagen: „A proper dictionary entry cannot be achieved by the mere collecting together of the translations of a word that are most appropriate in individual contexts.“⁶ Man braucht die negativen Erfahrungen, die die Arbeit am alten Berliner Wörterbuch zeitigte, nicht im einzelnen auszuführen; man kennt sie aus Hermann Grapow's abschließendem Bericht oder kann sie dort nachlesen⁷.

Ich will hier keine Überlegungen darüber anstellen, ob es 1897 eine echte Alternative zum damals eingeschlagenen Verfahren gab. Zum mindesten im nachhinein stellt sich aber doch die Frage – im folgenden kommt wieder Borger zu Wort –, „ob der durch das mechanische Verzettelungsverfahren angestrebte Zeitgewinn wirklich erzielt worden ist. Die ungeheure Menge letztlich unnützer Zettel [fast 90 %] muß die Ausarbeitung des Wörterbuchs doch sehr aufgehalten haben. Das Verfahren wäre für einen Thesaurus ideal gewesen, war es aber kaum für ein eklektisches Wörterbuch.“⁸ Es fragt sich, ob der Lexikograph nicht besser fährt, wenn er „von vornherein nur die Stellen mit kurzem Kontext notiert, deren Aufnahme in das Wörterbuch ihm erforderlich erscheint“⁹, oder, ob man nicht besser daran täte – eine Zwischenlösung – „nur eine verhältnismäßig kleine Selektion aus dem Textmaterial mechanisch zu verzetteln, und aus der großen Masse der Texte nur solche Stellen auszuziehen, die wirklich für Aufnahme in das Wörterbuch geeignet zu sein scheinen“¹⁰. Selektion, so effektiv sie im Ergebnis ist, hat natürlich ihren Preis: „Selektives Exzerpieren von Texten setzt ... bessere Vertrautheit mit der Materie voraus als mehr oder minder mechanisches Verzetteln und läßt sich weniger gut durch Hilfskräfte erledigen.“¹¹ – Womit ich die Qualitäten, die man von

¹ Rykle Borger, *Altorientalische Lexikographie, Geschichte und Probleme*, NAWG, phil.-hist. Kl. 1984, S. 69–114.

² Borger, S. 76.

³ Borger, S. 76.

⁴ Borger, S. 82.

⁵ W. J. Tait, *Approaches to Demotic Lexicography*, in: S. P. Vleeming (ed.), *Aspects of Demotic Lexicography*, Löwen 1987, S. 95–108.

⁶ Tait, S. 96.

⁷ Adolf Erman †/Hermann Grapow, *Das Wörterbuch der ägyptischen Sprache*, DAW, Vorträge und Schriften 51, Berlin 1953.

⁸ Borger, S. 76.

⁹ Borger, S. 77, Anm. 7.

¹⁰ Borger, S. 95.

¹¹ Borger, S. 95.

einem prospektiven Lexikographen erwarten muß, streife, zu denen gleich noch ein Wort zu sagen ist.

Ich will, wie gesagt, nicht mit Erman 1897 rechten. Sicher bin ich mir nur in einem: Unter heutigen Bedingungen ist ein anderes Vorgehen angebracht. Unser Wörterbuch ist – ähnlich äußert sich Borger zu den Wörterbüchern des Akkadischen¹² – trotz allem so gut, daß man nicht erneut ein mechanisches Verzettelungsverfahren in Gang setzen muß; es genügt, das für das Wörterbuch tatsächlich benötigte zusätzliche Material selektiv zu erheben, allenfalls streng ausgewählte Teil-Textkorpora als Kontrollmaßnahme oder zur Evaluierung des Lexemverhaltens ergänzend zu verzetteln. (Ich spreche weiterhin von Verzetteln, denke natürlich an den Computer.)

Die Erarbeitung eines Wörterbuchs, wie ich es verstehe, ist ein so schwieriges Geschäft, daß man sich überlegen muß, wie man den Entscheidungsprozeß der Wörterbucharbeit entlasten kann, wie man Fragenkomplexe abtrennen und gesondert behandeln kann. Ein probates Verfahren ist die vorgängige oder auch parallele Erarbeitung von Spezialwörterbüchern. Damit meine ich allerdings gerade nicht das, was Gardiner 1947 als Modell für die ägyptologische Lexikographie angeboten hat¹³. Ich denke also nicht an korpusbezogene Spezial-Wörterbücher. Denn, um wieder mit John Tait zu sprechen: „If they [die Spezialwörterbücher] are intended to be preliminary steps towards the compilation of a dictionary, then the bulk of the lexicographical work remains to be done at a later stage.“¹⁴ Diese Arbeit ist nämlich: zu selektieren und zusammenzufassen, kurz: sich einen lexikographischen Reim auf das Material zu machen.

Was ich mit Spezialwörterbüchern meine, ist z. B. folgendes:

- ein Vokalisationswörterbuch (übrigens von Jürgen Zeidler in Angriff genommen und schon weit gefördert)
- ein Lexikon der hamitosemitischen Etymologien
- ein Wörterbuch der semitischen Fremdwörter
- ein Wörterbuch der hieroglyphischen Graphien
- ein Valenz-Wörterbuch
- Sachwörterbücher zu Gebieten wie z. B. Botanik und Zoologie, Mineralogie und Metallurgie, deren Erarbeitung besondere Kompetenzen verlangt, über die selbst der philologisch beste Lexikograph nicht ohne weiteres verfügt
- eine Zusammenstellung der Wortdiskussion in der Sekundärliteratur, d. h. eine Zusammenstellung der substantiellen Beiträge zur Wortdiskussion, nicht eine Ausgrabung des unvermeidlichen Schrottes der wissenschaftlichen Produktion¹⁵.

In ähnlicher Weise könnte man – dies nur am Rande gesagt – die Arbeit am Textkorpus entlasten durch

- eine Textbibliographie, mit dem Nachweis der sinnvollen Textinterpretationen, die sich in der Sekundärliteratur finden unter Eliminierung der Falsch- und Füllzitate¹⁶.

Schließlich wäre für die Wörterbuch-Arbeit, wofür wiederum Borger plädiert¹⁷, zweckmäßig

- ein Register derjenigen Textstellen, die bereits in das Wörterbuch eingegangen sind, als Entscheidungshilfe bei der erneuten Exzerpierung der bereits einmal für die Zwecke des

¹² Borger, S. 113.

¹³ Alan H. Gardiner, *Ancient Egyptian Onomastica*, London 1947, I, S. XIX f.

¹⁴ Tait, S. 105.

¹⁵ Zur Wortdiskussion vgl. Gardiner, S. XXI.

¹⁶ Zur Textbibliographie – ein Herzensanliegen Borgers – vgl. Borger, S. 73.

¹⁷ Borger, S. 78, auch S. 83 und S. 113.

Wörterbuchs verzettelten Texte (man wird ja auch die alten Texte wieder durcharbeiten müssen, da sich das Verständnis inzwischen vielfach geändert hat).

„Auch bei lexikographischer Teamarbeit [ich bin bereits wieder mitten in Borger-Zitaten] kommt es ... gewaltig auf das Wissen und den Arbeitseinsatz der ganz wenigen wirklich hervorragenden Fachkräfte an.“¹⁸ „Zwar ist möglich, viele Mitarbeiter Material sammeln zu lassen, aber die Ausarbeitung des ... Wörterbuchs muß{te} doch von nur ganz wenigen Gelehrten vorgenommen werden.“¹⁹ „Leider kann man nicht mehrere schwache Wissenschaftler zu einer einzigen Spitzenkraft addieren.“²⁰ „Schwache Mitarbeiter können mehr Unheil anrichten, als u. U. von den besten Projektleitern behoben werden kann.“²¹

„Welche Voraussetzungen muß{te} nun ein prospektiver ... Lexikograph ... erfüllen ...?“

Er muß{te} über eine enorme Belesenheit verfügen ...

Er muß{te} die Fähigkeit haben, bei seiner Textlektüre die lexikalisch ergiebigen Stellen auszuwählen, und sich bereits sehr viele Belegstellen auf diesem Wege notiert haben.

Er braucht{e} ein hervorragendes Gedächtnis und große Kombinationsfähigkeit. ...

Er braucht{e} Vertrautheit mit der Lexikographie der <vergleichbaren> Sprachen. ...

Er braucht{e} grammatische Begabung.

Er braucht{e} epigraphische und paläographische Erfahrung.

Er muß{te} bereit sein, auf manches zu verzichten, das schnelleren Ruhm verspricht; er <kann> sich kaum mehr ein gründlich studiertes Spezialgebiet leisten.

Er muß{te} ein gereifter Gelehrter sein, aber zugleich noch so jung, daß – Deo volente – gute Aussicht auf Abschluß der Arbeit <besteht>.“²²

1897 und die Folge war eine Sternstunde der ägyptologischen Lexikographie, die sich so nicht durch Wollen und Planen einfach wiederholen läßt. „Erman, der ‚Schöpfer der neuen Ägyptologie‘, war <damals> gewiß der einzige Ägyptologe, der sich der Durchführung eines großen Wörterbuchunternehmens mit Aussicht auf Erfolg unterziehen konnte. Seine führende Rolle war praktisch unumstritten. Er war der Lehrer einer ganzen Generation von Ägyptologen, ein Meister klaren Denkens und Formulierens, ein guter Organisator und fähig, Freunde zu gewinnen.“²³ Oder, in Grapows Worten: „[Seine] unbestrittene Führung und Autorität ... zog immer neue Kräfte heran und sicherte die einheitliche Lenkung dieser Kräfte auf das große gemeinsame Ziel wie eine Selbstverständlichkeit.“²⁴

Ich komme zu praktischen Folgerungen: Ich empfehle zunächst und vor allem, die Arbeit am Textkorpus, an der Textdatenbank, die ein großes Material mit zunächst geringen, nur nach und nach wachsenden Ansprüchen erfaßt und erschließt, strikt zu trennen von der Wörterbucharbeit, die den hohen Anspruch der Selektion der lexikalisch relevanten Befunde einlösen muß. Was die Berliner Situation angeht, würde ich weiter empfehlen, die Pflege und Erschließung der Altbestände, das heißt die Archivarbeit, organisatorisch abzutrennen. Schließlich sind alle Spezialwörterbücher u. dgl. organisatorisch abzutrennen. Nur auf diese

¹⁸ Borger, S. 94.

¹⁹ Borger, S. 94.

²⁰ Borger, S. 94.

²¹ Borger, S. 94.

²² Borger, S. 102f.

²³ Borger, S. 71.

²⁴ Erman †/Grapow, S. 14

Weise vermeidet man eine Verzettelung der Energien. Was man macht, sollte man ganz – als eigenständiges Projekt – machen, oder man soll es gar nicht machen.

Für realisierbar halte ich Spezialwörterbücher – sofern sich Köpfe finden, die sie redigieren –, für realisierbar halte ich eine Textdatenbank – eventuell in Kooperation zwischen mehreren Arbeitsgruppen, die sich jeweils geeignete Teilkorpora aus dem Gesamt-Textkorpus aussuchen –, sofern man zu einem *modus procedendi* kommt, der nicht durch blinde vorgängige Normierung einfach die linguistischen Probleme negiert. Was das Wörterbuch angeht, bin ich eher skeptisch, und zugleich empfinde ich die Situation hier geradezu als fatal: Was die große Mehrzahl der Ägyptologen braucht und erwartet, ist in der Tat das Handwörterbuch, nicht etwa das Forschungsinstrument der Textdatenbank, um hier nur wieder diese beiden Instrumente gegeneinander zu setzen. Entweder also muß man den Kollegen reinen Wein einschenken, ihnen sagen, daß ein Wörterbuch nicht in näherer Zukunft entsteht, oder aber man muß die Arbeit vollständig umkrepeln: einen Kopf suchen, der, unterstützt durch ein Team, hier und jetzt echte Wörterbucharbeit macht und sonst nichts anderes, also auch keine umfassende Datenbank in Angriff nimmt.

Schließlich ein letztes: Zu glauben, man könne in absehbarer Zukunft ein Computerprogramm herzaubern, das die Brücke zwischen Textkorpus und Wörterbuch schlägt, halte ich für reine Utopie. Wollte man, um diese Brücke zu schlagen, sich zunächst auf Lexikologie konzentrieren, wäre die Lexikographie in unabsehbare Zukunft vertagt, wenn nicht endgültig begraben. Vielleicht liefert hierzu konkretes Anschauungsmaterial der Bericht über meine Arbeit an den Sargtexten, der an anderer Stelle dieses Heftes abgedruckt ist (S. 142–153).

POSTSCRIPTUM

In der Abschlußdiskussion habe ich in Abwägung der Hauptfaktoren der institutionellen und personellen Situation folgenden Kompromiß-Vorschlag unterbreitet:

1. Die Arbeitsstelle „Altägyptisches Wörterbuch“ sollte sich der Erstellung eines Textkorpus widmen, nicht allerdings – wie sie das in Fortschreibung des alten Wörterbuch-Verfahrens selbst sieht – als einem ersten Schritt in Richtung auf eine Neuauflage des Erman/Grapow, sondern – im Sinne meines Exposés – als einer eigenständigen Aufgabe.

2. Die Texte sollten in Transkription (plus Übersetzung) erfaßt werden, nicht – wie die Arbeitsstelle dies selbst plant – in Hieroglyphen, da die Erfassung der Hieroglyphen im vorgesehenen Verfahren (Stichwort: GLYPH) viel zu arbeitsaufwendig ist, um das Gesamt-Textkorpus in absehbarer Zukunft zu erfassen. – Die hieroglyphische Textgestalt sollte man zu einem späteren Zeitpunkt hinzufügen, und zwar mit Hilfe eines Scanners. Scannen könnte man Fotos, Zeichnungen, handschriftliche Texteditionen oder Hieroglyphen-Satz. (In der philologischen Praxis sind Faksimiles von höchstem Wert, ein zusätzliches Foto wäre indes nicht zu verachten.) Über die Stellenangaben wären die genannten Bilder mit den Transkriptionen zu vernetzen. Wie Leonard Lesko, der in seinem Referat in ähnliche Richtung denkt, halte ich die Speicherung der zu erwartenden Datenmenge für realistisch möglich (die Speicherungstechnik wird sich eher schneller entwickeln als der sinnvolle ägyptologische Gebrauch derselben).

3. Um die Wörterbuch-Arbeit nicht ganz aus den Augen zu verlieren, sollte bei der Aufnahme der Texte für die Zwecke des Textkorpus das lexikalisch interessante Material gekennzeichnet werden. Im Notfall genügt die Setzung eines „Ausrufezeichens“, zweckmäßig wäre jedoch, auch den Gesichtspunkt anzugeben, unter dem eine Textstelle lexikographisch

von Interesse ist. Für Standardfälle wäre ein System von Kennzeichnungen zu entwickeln, in Sonderfällen könnte man sich mit freien Kommentaren behelfen. Derart gekennzeichnetes Material könnte aus dem Textkorpus bei Bedarf selektiert und einer weitergehenden lexikographischen Bearbeitung zugeführt werden.

4. Akzeptable Zeitperspektiven für die Erfassung des Gesamt-Textkorpus sind z. B. 10 oder 20 Jahre, was bedeutet, daß bei einem Gesamt-Textkorpus von 5 oder 10 Millionen laufenden Wörtern pro Jahr Textmengen im Umfang etwa zwischen 250 000 und 1 Million laufenden Wörtern zu bewältigen sind, und zwar in philologisch stichhaltiger Weise mehr oder abschließend zu bewältigen sind – was einen mehrmaligen Durchgang durch die Texte erfordert (alle Zahlen grob geschätzt, in der Diskussion habe ich andere, aber in der Größenordnung in etwa entsprechende Zahlen genannt).

¹ Zu dessen 60sten und 75sten Geburtstag schon Ursula Hintze chronologisch geordnete Schriftenverzeichnisse zusammengestellt hat, s. u. Kap. VI: Verzeichnis der Schriften von Fritz Hintze, in: E. Endesfelder, K.-H. Priese, W.-F. Reineke und S. Wenig [Hrsg.] 1977: 507–512 und Schriftenverzeichnis Fritz Hintze, in: D. Apelt, E. Endesfelder und S. Wenig [Hrsg.] 1990: 385–390. S. auch E. Endesfelder 1988: 73–76.