

Berlin-Brandenburgische Akademie der Wissenschaften

# Mitteilungen

aus der Arbeit am

# WÖRTERBUCH DER ÄGYPTISCHEN SPRACHE

1

Herausgegeben von

Walter F. Reineke

unter Mitarbeit von

Adelheid Burkhardt, Elke Freier, Stefan Grunert und Ingelore Hafemann  
(Mitarbeiter am Akademienvorhaben Altägyptisches Wörterbuch)

Berlin 1993

Die Abkürzungen der ägyptologischen Literatur folgen allgemein den Regeln des *Lexikons der Ägyptologie*. Darauf basiert ein Abkürzungsverzeichnis zur numerischen Codierung aller Daten, das demnächst in der Reihe *Egyptology* des Internationalen Ägyptologen-Verbandes erscheint.

Als Abkürzung der *Mitteilungen aus der Arbeit am Wörterbuch der Ägyptischen Sprache* wird **MittWb** vorgeschlagen.

© 1993 by Altägyptisches Wörterbuch • Berlin  
Printed in Germany

Gedruckt mit Unterstützung des Internationalen Ägyptologen-Verbandes

# FRITZ HINTZE

Im Andenken  
an seine ständige  
Förderung und sachkundige Begleitung  
der Arbeiten für die Weiterführung des  
Wörterbuches  
der  
Ägyptischen  
Sprache



## Editorial

Die ersten zwei vorliegenden *Mitteilungen aus der Arbeit am Wörterbuch der ägyptischen Sprache* sind der Beginn einer in loser Folge erscheinenden Reihe von Berichten und Resultaten der real seit 1992 wiederaufgenommenen Arbeiten am neuen Wörterbuch-Projekt. Diese und auch die weiteren Veröffentlichungen verstehen sich nicht als eine weitere ägyptologische Zeitschrift: Als Verwalter eines sehr umfangreichen Textarchives ägyptischen Sprachgutes sehen wir es als eine für alle lohnende Aufgabe an, die zum Teil einzigartigen Materialien sowohl für die Zukunft international effektiver nutzbar zu machen, als auch als Basis bei der Fortführung der Sammlung ägyptischer Textzeugnisse anzuwenden.

Die gesamte Arbeit wird durch den Einsatz moderner Computertechnik unterstützt. Dabei erfordert der Aufbau einer lexikalischen Datenbank, die mit einer Textdatenbank und einem Textcorpusverwalter verbunden ist, eine Vielzahl von vorbereitenden Arbeiten, die als solche nicht nur für die Berliner Arbeiten an einem rechnergestützten Wörterbuch geeignete Hilfsmittel sind. Sie sollen für eine Veröffentlichung in den *Mitteilungen* aufbereitet und auf diesem Weg einer möglichst großen Zahl von Fachkollegen zur Nutzung angeboten werden. Der Charakter dieser gedruckten Dokumente wird ganz wesentlich durch deren ursprüngliche Funktion bestimmt - es sind tools für die Computerarbeit, die auf Grund ihrer Struktur und Informationsvielfalt ständiger Veränderung und Erweiterung unterworfen sind. Ihrer Menge nach sind sie nur bedingt in Papierformat umsetzbar. Daher können - über die ausgedruckten Hilfsmittel hinaus - auf Anfrage auch die primären Dateien in computergemäßer Form zur Verfügung gestellt werden.

## **Inhalt**

Editorial	5
J. Assmann, S. Grunert, W. F. Reineke Bericht über die Arbeitstagung "Die Zukunft des Altägyptischen Wörterbuchs" vom 3.-5. 9. 1992 in Berlin	7
Jan Assmann Folgerungen für die Zukunft	21
Nigel Strudwick The next generation of the Wörterbuch	25
Walter F. Reineke Das Wörterbuch der Ägyptischen Sprache. - Stand, Probleme, Perspektiven -	31
Silvia Köpstein Das Abklatscharchiv beim "Wörterbuch der ägyptischen Sprache" (Arbeitsbericht)	37
Ingelore Hafemann Lexikalische Datenbank contra Wörterbuch	43
Stefan Grunert Zur Berliner Wortliste	54

Bericht über die Arbeitstagung  
"Die Zukunft des Altägyptischen Wörterbuchs"  
vom 3.-5.9.1992 in Berlin  
J. Assmann, S. Grunert, W. F. Reineke

Erstes Ziel der Berliner Arbeitsgruppe Altägyptisches Wörterbuch, nach glücklich überstandener "Evaluierung" und gesicherter Weiterexistenz in den alten, neu ausgestatteten Räumen, war die Wiederherstellung der ursprünglichen, schon durch die Nazis beschnittenen und zu DDR-Zeiten dann vollends blockierten Internationalität des Projekts, das einmal das Herzstück der internationalen Ägyptologie gewesen war und für dessen Fortexistenz sich in der Phase der Abwicklung Kollegen aus allen Ländern eingesetzt hatten. Die internationale Arbeitstagung "Die Zukunft des Altägyptischen Wörterbuchs" sollte den ersten Schritt auf diesem Wege bilden. Ziel der Tagung war daher vor allem die Sondierung von Möglichkeiten internationaler Zusammenarbeit. Es galt, einen Überblick zu geben über die verschiedenen mit der Berliner Wörterbucharbeit verbundenen Einzelprojekte (Aufgaben, Zielsetzungen, Teilschritte) und Methoden (Programme) sowie einen Überblick zu gewinnen über vergleichbare Vorhaben und Erfahrungen in der internationalen Ägyptologie, um Überschneidungen zu vermeiden, rechtzeitige Abstimmung herzustellen und Möglichkeiten gegenseitigen Datenaustauschs zu eröffnen. Ferner galt es, sich über die Zielsetzungen und Methoden der Berliner Wörterbuch-Arbeit selbst zu verständigen. Was erwartet die internationale Ägyptologie von einer Neuausgabe des Wörterbuchs? Welche Möglichkeiten sieht sie, a) die vielerorts bereits gesammelten Ergänzungen zum Wb in diese Neuausgabe einzubringen, b) sich an den darüberhinaus notwendigen Arbeiten zu beteiligen?

In Anbetracht der begrenzten Mittel, der kurzfristigen Planung und des workshop-Charakters der Tagung war die Teilnehmerzahl bewußt kleingehalten worden. Es wurden vor allem solche Kollegen eingeladen, die mit der Entwicklung und/oder Anwendung von EDV-Programmen in der ägyptischen Lexikographie befaßt sind. Auf diesem Gebiet schien der größte Abstimmungsbedarf vorzuliegen. Demgegenüber wurde in Kauf genommen, daß Philologie und Sprachwissenschaft diesmal erheblich unterrepräsentiert bleiben mußten. Eine gewisse thematische Schwerpunktsetzung schien unabdingbar, um zu konkreten Ergebnissen zu kommen. Daher waren viele Kollegen bei der Tagung nicht anwesend, deren Forschungsschwerpunkte sich mit der allgemeinen Wörterbuchproblematik enger berühren als dies etwa bei den EDV-Spezialisten der Fall ist. Es ist aber vorgesehen, diese Versäumnisse in Zukunft nachzuholen. Die Tagung verstand sich als Auftakt einer Serie ähnlicher Konferenzen, die mit wechselnden Schwerpunktsetzungen auch wechselnde Teilnehmerkreise einbeziehen. Allerdings waren darüberhinaus auch

manche der diesmal in Berlin fehlenden Kollegen wie etwa E. Hornung, J. Osing und W. Westendorf verhindert. Anwesend waren aus Deutschland:

Horst Beinlich (Würzburg)	Frank Kammerzell (Göttingen)
Elke Blumenthal (Leipzig)	Steffen Kirchner (Berlin, HU)
Andrea Gnirs (Göttingen/Heidelberg)	Wolfgang Schenkel (Tübingen)
Jochen Hallof (Trier/Berlin)	Stephan J. Seidlmaier (Berlin, FU)
Friedrich Junge (Göttingen)	Pawel M. Wolf (Berlin, HU)

sowie die Mitglieder der Berliner Arbeitsgruppe Altägyptisches Wörterbuch  
Adelheid Burkhardt Stefan Grunert

Elke Freier Ingelore Hafemann  
Silvia Köpstein (ABM; Abklatschkatalogisierung)

Aus dem Ausland:

Hans van den Berg	(Niederlande, Utrecht)
Leonard H. Lesko	(USA, Providence)
Nicolas C. Grimal (Kairo)	(Frankreich, Institut Français d'Archéologie Orientale,
Jean Winand	(Belgien, Liège)
Nigel Strudwick	(Großbritannien, Cambridge)
Fayza Haikal	(Ägypten, American University Cairo)
Dirk van der Plas	(Niederlande, Utrecht)

Als Veranstalter:

Jan Assmann	(Heidelberg, als interimistischer Betreuer des Wörterbuchprojekts im Auftrag der Konferenz der Akademien)
Fritz Hintze †	(Berlin, als Leiter der ehemaligen Wörterbuch-Kommission)
Walter F. Reineke	(Berlin, als Leiter der Arbeitsgruppe Altägyptisches Wörterbuch)
Dietrich Wildung	(Berlin, als Präsident des IÄV und Direktor des Ägyptischen Museums, dem Ort der Tagung)

Nach der Begrüßung der Teilnehmer am 1. Tagungstag (3. September; 9.00 - 18.00 Uhr) durch D. WILDUNG und J. ASSMANN eröffnete F. HINTZE die Arbeitstagung mit einer kurzen Einführung in die Geschichte des seit 1897 laufenden Projektes "Wörterbuch der Ägyptischen Sprache" und erläuterte, wie es zur Einstellung aller Forschungsarbeit zwischen 1947 und 1985 kam.

Die Geschichte des 'Wörterbuches' gliedert sich in mehrere Phasen. Die 1., "klassische" Phase umfaßt die 50 Jahre von 1897-1947. Nur sie war der Arbeit an einem ägyptischen Gesamtwörterbuch gewidmet. Entscheidend war die Tatsache, daß diese Arbeit mit dem Erscheinen des ersten Bandes 1928 nicht

eingestellt, sondern in Form systematischer Verzettelung neuer Texte konsequent weitergeführt wurde. Das Wörterbuch war also von Anfang an auf Ergänzung und das heißt: auf ein Gesamtcorpus aller ägyptischen Texte als Basis künftiger Neuausgaben hin angelegt. Diese Konzeption wurde aber 1947 aufgegeben - und damit die Nachverzettelung neuer Texte eingestellt - zugunsten der Erstellung von Spezialwörterbüchern, wie sie von Gardiner im Vorwort zu den Onomastica 1947 angeregt worden waren (Medizin, Mathematik, Mineralogie, Äthiopentexte). Dies war die 2. Phase, die mit der Eingliederung der Wörterbuchgruppe in das Zentralinstitut für Alte Geschichte und Archäologie im Jahre 1969 endete. Die Wörterbucharbeit mußte zugunsten anderer Projekte eingestellt werden und blieb auf die Pflege der Archive und Überlegungen zur Konzeption einer Weiterführung beschränkt (3. Phase). Die 4. Phase begann mit der Erarbeitung einer 1. Konzeption zur Weiterführung der Wörterbucharbeiten (1971). Erst 1985 aber wurde die Arbeitsgruppe wieder dem Projekt Altägyptisches Wörterbuch rückgewidmet und eine Kommission eingesetzt, der F. Hintze vorstand. Wichtigstes Arbeitsziel war die Prüfung von EDV-Programmen, die für die Wiederaufnahme der Neuverzettelung geeignet wären. Die 5. Phase begann 1989, schon vor der Wende, durch die auf Betreiben von D. Wildung erfolgte Eingliederung in den "Berliner Ägyptologischen Datenverbund". Seitdem konnte durch eine entsprechende apparative Ausstattung, die inzwischen (ab 1992) auf ein angemessenes Niveau aufgestockt worden ist, mit konkreten Vorarbeiten zur Texterfassung begonnen werden. Die 6. Phase steht allerdings noch bevor: sie beginnt, wenn die neue Berliner Akademie gegründet, eine neue Wörterbuch-Kommission bestellt und auf internationaler Basis die eigentliche Arbeit an den Hauptprojekten der Arbeitsgruppe in Angriff genommen sein wird. Die Arbeitstagung diente der Vorbereitung dieser Phase, vor allem durch Herstellung der internationalen Basis.

## I. Fragen, Erwartungen und Forderungen an ein neues Wörterbuch.

Im Beitrag von W.-F. REINEKE wurden aktuelle Fragestellungen im Zusammenhang mit modernen Wörterbüchern und der Notwendigkeit der Erstellung eines umfassenden Textcorpus erörtert. Dabei ging es vor allem um 4 Probleme:

1. Das Problem der verschiedenen Sprachstadien des Ägyptischen. Das alte Wörterbuch hatte sich auf das Alt-, Mittel- und Neuägyptische beschränkt und Demotisch und Koptisch nur in gelegentlichen Vergleichen einbezogen. Die Neuausgabe des Wörterbuchs müßte sich dagegen konsequent auf sämtliche Sprachstadien des Ägyptischen beziehen. In diesem Zusammenhang gilt es auch, das Problem kanonisierter Sprachstadien (in heiligen, literarischen und sonstigen "kulturellen" Texten) und der darauf beruhenden,

- sich ebenfalls historisch verändernden "Polysystemik" (I. Evan-Zohar) des Ägyptischen zu berücksichtigen.
2. Das Problem der Belegstellen. Bei toten Sprachen und den ihnen gewidmeten Historischen Wörterbüchern muß die nicht nachprüfbare Akzeptabilität durch Belegbarkeit ersetzt werden. Daher ist sowohl Belegstellenvollständigkeit als auch eine kritische Würdigung des Belegstatus (wieviele Belege und Belege welcher Art müssen mindestens vorliegen, damit eine Zeichenfolge als "Lexem" anerkannt und ins Wb aufgenommen werden kann?) unabdingbar. Die Erstausgabe des Wörterbuches ist aber hinsichtlich der Belegstellenangabe äußerst selektiv verfahren, ganz abgesehen davon, daß bis 1926 überhaupt nur ca. 40% des heute verfügbaren Textmaterials erfaßt werden konnte.
  3. Möglichkeiten archäologisch-onomasiologischer Erschließung eines historischen Wortschatzes: "Benennungspflicht". Gefordert wird eine archäologisch-historische Grundlagenforschung. Alles kulturgeschichtlich Bezeugte muß eine sprachliche Bezeichnung gehabt haben. Von 83 archäologisch nachweisbaren Pflanzennamen z.B. sind aber nur 28 lexikalisch identifiziert. Für 366 Lexeme findet sich im Wörterbuch nur der Eintrag "Art Pflanze".
  4. Möglichkeiten philologisch-semasiologischer Erschließung: z.B. Wortfeldforschung. Auch hier bestehen in der Ägyptologie erhebliche Desiderate.

Der anschließende Beitrag von W. SCHENKEL "Wörterbuch versus Textcorpus" (erscheint in ZÄS) bezog eine genaue Gegenposition. Was die Ägyptologie braucht, ist nicht das belegstellenvollständige integrierte Thesaurus-Wörterbuch, das im Hinblick auf die derzeitige Situation der Ägyptologie als eine Utopie angesehen werden muß, sondern zwei verschiedene und sorgfältig auseinanderzuhaltende Arbeitsinstrumente: 1. ein aktuelles Handwörterbuch der ägyptischen Sprache, und 2. eine Textdatenbank mit Wortkonkordanz. Die beiden Instrumente dienen verschiedenen Zwecken. Das erste bedient auch Fachnachbarn, Sprachwissenschaftler, Semitisten sowie Studienanfänger usw.; das zweite dagegen dient nur dem professionellen, philologisch arbeitenden Ägyptologen als reines Forschungsinstrument. Es erfüllt ganz andere Ansprüche als (1). Erst wenn man die verschiedenen Ansprüche trennt und auf zwei getrennte Projekte verteilt, werden die Teilziele erreichbar. Schon die erste Ausgabe des Wörterbuches litt unter der Kombination mit einem Textcorpus. "Die Anhäufung des Materials erwies sich als leichter als seine Auswertung" (R. Borger). Die Umsetzung der anderthalb Millionen gesammelter Belege in lexikalische Information nahm 18 Jahre in Anspruch. Da EDV zwar das Sammeln, aber nicht die Umsetzung erleichtert, außerdem ein Genie vom Range Adolf Ermans nicht zur Verfügung steht, ist leicht abzusehen, daß der Weg vom Textcorpus zum Wörterbuch heute überhaupt nicht mehr gangbar ist. Daher empfiehlt sich die strikte Trennung in Handwörterbuch und Textcorpus.

Das Handwörterbuch soll das gesamte lexikalische Wissen bieten. Es soll aktuell, umfassend und kompakt sein. Es soll Angaben enthalten zu semito-hamitischen Etymologien, zu Rektion, Wortbildung, Derivation, Belegzeitraum, evtl. textsortenspezifische Belegpräferenzen, soll Demotisch und Koptisch integrieren, Hinweise auf wichtige Sekundärliteratur enthalten. Als Modell kann das Akkadische Handwörterbuch von W. v. Soden dienen.

Das Textcorpus soll den gesamten Textbestand in ägyptischer Sprache umfassen. Es kann sich auf Transkription beschränken. Es bietet die Möglichkeit des Zugriffs auf sämtliche Vorkommen sämtlicher Lexeme in ägyptischen Texten, überläßt aber die Relevanz- und Selektionskriterien dem Benutzer und seinen spezifischen Fragestellungen. Das Textcorpus braucht bzw. soll gar nicht in ausgedruckter Form herausgegeben werden, sondern z.B. als CD-ROM. Spezialwörterbücher nach dem Modell Gardiner von 1947 (Corpus- bzw. Sachgebiet-spezifische Wörterbücher) sind unnötig, da sie die Arbeit der Selektion und Zusammenfassung nur aufschieben. Was hier gebraucht wird, sind vielmehr

- ein Vokalisationswörterbuch,
- ein Lexikon hamito-semitischer Etymologien,
- ein Valenzwörterbuch,
- ein Graphienwörterbuch,
- ein Wörterbuch der semitischen Fremdwörter,
- Spezialwörterbücher in Gebieten wie z.B. Botanik und Zoologie, Mineralogie und Metallurgie,
- eine Zusammenstellung der Wortdiskussion in der Sekundärliteratur.

Als praktische Folgerungen für die Arbeit am Akademien-Projekt wurde von W. Schenkel hervorgehoben, daß vorrangig das Textcorpus inklusive eines Registers zu den erfaßten Textstellen (sog. TCN-Register) erarbeitet werden muß, nicht aber ein modernes Wörterbuch der ägyptischen Sprache. Seiner Ansicht nach müssen die Arbeiten zur Pflege und Betreuung der Archive vollständig hiervon abgetrennt werden; sowohl die erwähnten Spezialwörterbücher als auch die Datenbank der Texte sollten in Zusammenarbeit der verschiedenen interessierten Institutionen erfolgen.

F. HAikal stellte in ihrem Plädoyer für eine Neuausgabe des Wörterbuches einige der dringendsten Desiderate heraus, die aus der Sicht der internationalen und besonders der ägyptischen Ägyptologie an ein solches Vorhaben anzumelden sind und stellte vorab klar, daß das vorliegende Wörterbuch (1. Auflage 1926-1963) als nicht länger zufriedenstellend anzusehen ist. Die Neuausgabe des Wörterbuchs soll alle Sprachstadien des Ägyptischen umfassen und nicht etwa aufgespalten werden in eine entsprechend den historischen Zeitabschnitten (AR, MR, NR usw.) geordnete Reihe von Wörterbüchern. Es soll alle Schreibungen eines Lexems nachweisen sowie alle Belegstellen ange-

ben. Es soll die afroasiatischen Etymologien berücksichtigen, vor allem auch die zahlreichen Fälle der in heutigen arabischen, besonders ägyptischen Lokaldialekten überdauernden ägyptischen Wörter. Schließlich erachtete sie es als notwendig, auch verstärkt Angaben zum metaphorischen Sprachgebrauch einzubeziehen.

Auf den ersten Teil des Programms, und aus diesem Grunde hier in den Bericht eingesetzt, bezog sich auch die spätere Intervention von N. STRUDWICK, der erste Gedanken zur äußeren Form eines altägyptischen Wörterbuches äußerte. Sein wichtigster Punkt war, daß es nicht in gedruckter, sondern in einer elektronischen Form vorgelegt werden sollte. Wenngleich die Fertigstellung eines solchen Wörterbuches noch längere Zeit andauern würde, so müßte bereits jetzt Klarheit darüber erzielt werden, um so den Anforderungen, die sich mit der Anwendung modernster Technik ergeben, zu entsprechen.

## II. Aus den Erfahrungen mit anderen Projekten

In seinem schriftlich vorgelegten und verteilten Beitrag "High Tech Projects for Research and Distribution" (erscheint in ZÄS) berichtete L. H. LESKO über seine Erfahrungen bei der Bearbeitung eines "Glossary of the Late Ramesside Letters" und des "Dictionary of Late Egyptian", die er in ihren Einzelbereichen als eine Art Vorstufe für ein komplexes Wörterbuch verstand. Ein wesentlicher Aspekt war dabei die Verarbeitung und der Druck codierter Hieroglyphen. Aus den Erfahrungen eines vergleichbaren Projektes für Altgriechisch empfahl er die Anwendung der CD-ROM - Technik als wohl gegenwärtig geeignetstes Speichermedium, das sich gleichzeitig für eine Verteilung unter potentiellen Nutzern eignen würde. Als wesentlichen Bestandteil gespeicherter Daten bezeichnete er von Farbbildern eingescante Darstellungen der Originalquellen. Als weitere Notwendigkeiten einer Datenbank des Textcorpus nannte er ein Identifizierungsregister (dem entspricht das Teilprojekt TCN-Register) sowie eine Aufnahme der Texte in Transkription, die so eng wie möglich an das Original gebunden ist, wobei als primäre Informationsquelle die Hieroglyphen in codierter Form unverzichtbar sind. An technischer Ausrüstung wird vom Redner zur Anwendung von 386er bzw. 486er PCs mit einem Arbeitsspeicher von 4-10 MB und einem Festspeicher von 200-500 MB bei einer Verarbeitungsfrequenz von 25 bis 50 MHz geraten, wobei die Geräte unter LAN verbunden sein sollten.

Der Beitrag von A. GNIRS behandelte Perspektiven und Methoden textsortenbezogener Textdatenbanken und beschäftigte sich mit der gattungsspezifischen Behandlung von Texten aus lexikologischer Sicht. In Bezug auf das Wörterbuch-Projekt sollten sich die Kriterien für die Textcorpus-Aufnahme

nach den Zwecken des Wörterbuches richten, d. h. bereits bei der Korpusbearbeitung sollten lexikologische Fragen wie:

- linguistisch-sprachgeschichtliche Dimensionen,
- sozio-kulturelle Dimensionen bzw.
- pragmatische Dimensionen

eine Rolle spielen. Dabei wies sie auf Arbeiten von Erman zur Wortforschung hin. Obwohl allgemein in der Lexikologie eine repräsentative Auswahl von Texten für Wörterbuch-Fragen als ausreichend angesehen wird, empfahl sie, in der Ägyptologie alle verfügbaren Texte in das Textcorpus aufzunehmen. Die Erfassung der Texte sollte unbedingt nach Gattungen erfolgen, um sowohl die sozio-kulturellen als auch pragmatischen Dimensionen eines Wortes sowie seine semantischen und morphologischen Eigenheiten innerhalb einer Gattung zu berücksichtigen. Die Gattung, d.h. die Gesamtmenge der durch eine gemeinsame Verwendungssituation verbundenen Texte, bietet den einzigen Zugang zur pragmatischen Dimension des Wortschatzes sowie zur Wortfeldforschung und liefert auch die entscheidenden Hinweise auf die von F. Haikal angesprochene Problematik metaphorischen Wortgebrauchs. So hat z.B. das Wort "Herz" in biographischen Texten eine andere Bedeutung als in medizinischen. Die Ausnutzung des von S. Seidlmaier entwickelten Programms PPTES hat sich hinsichtlich einiger dieser Forderungen als sehr geeignet erwiesen. Es gibt hier Möglichkeiten zur morphologischen Analyse, zur Erstellung von Wortkonkordanzen sowie zum syntaktischen als auch syntagmatischen Verhalten von Worten.

In seinem 2. Beitrag berichtete W. SCHENKEL (erscheint in ZÄS) über den gegenwärtigen Stand der seit 25 Jahren laufenden Arbeiten zur maschinellen Analyse der altägyptischen Sargtexte, die ursprünglich mit dem System M.A.A.T. begonnen wurden und gegenwärtig mittels des integrierten Programmpaketes TUSTEP weitergeführt werden. Aus der langen Erarbeitungszeit ergibt sich, daß die Zielvorstellungen des Projektes Änderungen unterworfen waren. Eine Verarbeitung der in codierter Form ursprünglich eingegebenen Graphien ist daher nicht möglich, auch im Bereich der Transliteration führte die Entwicklung des Wissenstandes zu wiederholten Veränderungen im gesamten Datenmaterial. Gegenwärtig besteht die Möglichkeit, für ca. 30-40000 Graphien nach Identifizierung Codes einzugeben, die dann sowohl unter dem Aspekt der Graphien als auch unter Berücksichtigung der Flektionsformen sortierbar sind. Die Notation der Graphien erfolgt auf 6 Ebenen (Lautwerte; Inversionen und phonetische Determinative; 4 x textkritische Zeichen [Zerstörungen bzw. Fragliches; Duplikierungen; Verweise; Rubren]). Aus seinen Erfahrungen ist für das Wörterbuch-Projekt eine schnelle Bearbeitung des gesamten Materials notwendig, um möglichen, durch den Zeitfaktor bedingten Änderungen zu begegnen.

In seinem Beitrag "L'encodage du Néo-Égyptien" definierte J. WINAND das neuägyptische Corpus als eine in sich nicht einheitliche, sondern auf der Ebene verschiedener 'Register' zu unterteilende Gesamtmenge neuägyptischer Texte. Er verwies auf die bekannten Probleme bei der Transliteration dieser in ihrer Zeitstellung und Schriftweise besonderen Gruppe, die er durch verschiedene Thesauri (z.B. Form-Thesaurus, Morphem-Thesaurus, Ausdrucksregister) in gewisser Weise zu lösen vermochte.

Aus der Sicht seines Projekts einer Internationalen Ägyptologischen Datenbank erläuterte D. VAN DER PLAS seine Überlegungen zu Fragen eines altägyptischen Wörterbuches, das mit einem Textcorpus verbunden ist. Er bezog damit die genaue Gegenposition zu W. Schenkel, der in seinem ersten Beitrag für eine strikte Trennung der Projekte Wörterbuch und Textcorpus plädiert hatte. Van der Plas sprach sich für einen völligen Neuanfang, nach dem Prinzip "reculer pour mieux sauter" aus. Am Anfang hat die Erstellung eines umfassenden Textcorpus zu stehen: "ohne Textcorpus kein neues Wörterbuch". Das Textcorpus soll alle ägyptischen Texte mit codierten Hieroglyphen, Umschrift, Interlinearübersetzung umfassen. Ein solches Textcorpus ist nur in internationaler Zusammenarbeit erreichbar. In diesem Zusammenhang wurden folgende Vorschläge bzw. Forderungen gemacht:

- Dieses Textcorpus soll als Bestandteil der Internationalen Ägyptologischen Datenbank (IÄD) eines Netzwerkes dezentraler Datenbanken erstellt werden.
- Der IÄD und ihren Teilprojekten soll ein einziges verbindliches System zugrundegelegt werden.
- Eine Liste soll alle kooperierenden Kollegen und die von ihnen zu bearbeitenden Textgruppen erfassen.
- Die Mitarbeiter sollen sich in Spezialseminaren über Methoden und Probleme der kollektiven Textaufbereitung verständigen.
- Die Entwickler der Computerprogramme THOT und PPTES sollen ein gemeinsames Softwareprodukt erarbeiten, das als integriertes Standardpaket von Texterfassungsprogrammen dann im internationalen Maßstab als verbindlich vorzugeben sei.
- Die verbindlichen Regeln der Texteingabe sollen in einem Leitfaden zusammengestellt werden.
- Eine internationale Wörterbuch-Kommission sollte durch den IÄV gegründet werden.
- Diese hätte sich vor allem um die Außendarstellung des Projekts (public relations) zum Zwecke der Einwerbung finanzieller Unterstützung zu kümmern.

Die Geschichte des Wörterbuchprojekts, die Notwendigkeit der Neuauflage, die Beschreibung des Vorhabens und seiner internationalen Form, der Computertechnik, der innovativen Zielsetzungen sollte in einer mehrsprachigen Broschüre dargestellt werden.

Für die Einbeziehung in die Wörterbuchgestaltung unterbreitete er ein umfangreiches Angebot von codiert vorliegenden Texten, das sich allerdings auf Nachfrage als noch in der Anfangsphase befindlich erwies. Insgesamt verursachte der Beitrag von D. van der Plas eine umfangreiche Diskussion, in der das Für und Wider von international verbindlichen Standards und einer postulierten Notwendigkeit der durchgängigen Hieroglyphen-Codierung altägyptischer Texte näher erörtert wurden.

Im Ganzen neigte man mehrheitlich zu der Ansicht, daß die grundsätzliche Vorordnung des Textcorpus vor die Neuauflage des Wörterbuchs diese ins Unabsehbare verzögern würde. Der Weg vom Textcorpus zum Wörterbuch ist nach wie vor unklar. Anstatt die gesamte internationale Ägyptologie auf ein einziges allgemeinverbindliches System festzulegen, sollte man lieber Verfahren entwickeln, die größere Flexibilität und Konvertibilität ermöglichen.

Anschließend wurden zu diesem Sachbereich in Demonstrationen am Rechner durch die Entwickler H. VAN DEN BERG, J. HALLOF und S. J. SEIDL MAYER die Programme WINGLYPH, THOT und PPTES vorgestellt. Grundsätzlich machten sie darauf aufmerksam, daß diese Programme nicht für das Wörterbuch-Projekt entwickelt wurden. Ausgehend von der Bedeutung des Projektes sei es dringend notwendig, in Zusammenarbeit mit einem Programmierer und unter Anwendung der Erfahrungen aus der Nutzung der verschiedenen Programme eigene, wörterbuch-spezifische Software-Vorstellungen zu entwickeln.

### III. Die Projekte der Berliner Arbeitsstelle Altägyptisches Wörterbuch

Eingangs erläuterte P. M. WOLF (Berlin HU) die Strukturen der Berliner Ägyptologischen Datenbank, die sich aus einem IBM-Studienprojekt unter Anwendung des Betriebssystems OS/2 entwickelt hatte. Im Vergleich mit den Vorstellungen zu Beginn des Studienprojektes (Ende 1989), die auf einer Arbeitskonzeption von 1985 aufbauten, zeigen sich die substantiellen Veränderungen, die sich während der Projektlaufzeit (bis Ende 1991) insbesondere aus einer sich vertiefenden Kenntnis der inneren Strukturen und der vielfältigen Möglichkeiten von OS/2 ergaben. Durch die Entwicklung der neuen Version von OS/2 werden einerseits die Vorteile der Listenverarbeitung, andererseits aber auch deren Grenzen trotz Kompatibilität und bei Nutzung der bereits eingegebenen Daten deutlich. Für die Zukunft sollte die Möglichkeit einer Anwendung objektorientierter Datenbankprogramme geprüft werden, zu deren

Anwendung es in vergleichbaren Unternehmungen bereits entsprechend positive Erfahrungen gibt.

Ausgehend von der database-Struktur unter OS/2 erläuterte S. **GRUNERT** die Veränderungen innerhalb des seinerzeit als Arbeitsvoraussetzung geplanten computerinternen Wörterbuches (CiWb), an dessen Stelle jetzt eine sog. Berliner Wortliste gesetzt wurde. Diese Wortliste wird durch Kompilation der einschlägigen Literatur zum Wortbestand des publizierten Wörterbuches erstellt, wobei einer laufenden Nummer, die in verschiedenen Listen auftaucht und durch sog. Schlüssel untereinander verbunden ist, sich jeweils ergänzende Informationen (Hieroglyphencodierung, Normübersetzung, Codierung hieroglyph-graphischer sowie sich aus der historischen Entwicklung ergebender Schreibvarianten, Transkriptionscodierung, Datierungscodierung, Grammatikcodierung, Sachgebietscodierung usw.) zugeordnet sind. Diese Informationen können komplex verbunden innerhalb eines Reports abgefragt werden. Insbesondere hinsichtlich der Belegerfassung ist die Notwendigkeit einer Textcorpusnummern-Datei (TCN) gegeben, in der alle relevanten Daten zu einem altägyptischen Text erfaßt werden. Zu einzelnen Datenbereichen (z.B. Datierung, Herkunft, Textgattungsbestimmung) ist die Erarbeitung verbindlicher Thesauri zur Optimierung der Dateneingabe - insbesondere hinsichtlich der Daten-gleichartigkeit und der Datenreduzierung durch numerische Codes - notwendig.

Über Fragen einer vorhergehenden Identifizierung des Textmaterials sprach I. **HAFEMANN** und stellte mehrere Konkordanzen vor, durch die sowohl den Bearbeitern als auch den Nutzern des publizierten Wörterbuches eine Verifi-zierung der zitierten Textquellen ermöglicht wird.

Ein ähnliches Hilfsmittel ist der von E. **FREIER** vorgelegte Index der Museums-Objekte zu den Bänden IV-VII der Topographical Bibliography of Ancient Egyptian Hieroglyphic Texts, Reliefs, and Paintings.

Für bislang nicht verarbeitete Texte und deren Prioritäten bei der Bearbeitung stellte A. **BURKHARDT** die Berliner Überlegungen zur kurz-, mittel- und langfristigen Planung vor, wie sie sich aus dem Aspekt eines möglichst hohen Nutzens für den im Wörterbuch erfaßten Wortbestand ergibt.

Zum bislang bekannten Wortbestand gibt es innerhalb der wissenschaftlichen Literatur umfangreiche Diskussionen vielfältigen Inhalts, über deren Erfassung innerhalb des Projektes "Wortdiskussion" und den damit verbundenen Problemen E. **FREIER** berichtete.

Der Problemkreis zum Projekt "Schreibungswörterbuch", den I. HAFEMANN vorzustellen beabsichtigte und zugunsten einer vertiefenden Diskussion der vorerwähnten Sachbereiche zur Disposition stellte, wurde von verschiedenen Standpunkten aus diskutiert.

Insgesamt zeichnete die Präsentation der Berliner Arbeitsergebnisse nach knapp einjähriger Laufzeit des Akademien-Vorhabens "Altagyptisches Wörterbuch" ein sehr vielfältiges Bild hierfür notwendiger Arbeitsaufgaben, die sich nicht auf eine mechanische Eingabe vorher codierter hieroglyphischer Texte in den Computer beschränken lassen.

#### IV. Abschlußdiskussion

Die Notwendigkeit und die Abfolge einzelner Arbeitsschritte standen im Zentrum der von J. ASSMANN geleiteten Gesamtdiskussion am 3. Tagungstag (9 - 14.30 Uhr). Zentrale Fragestellung war dabei das Verhältnis von Wörterbuch und Textcorpus bzw. die Frage: Wie kommt man in vernünftiger Zeit zu einem abgeschlossenen Ergebnis? Allgemeiner Konsens bestand in der von W. Schenkel und anderen erhobenen Forderung, die Projekte Wörterbuch und Textcorpus grundsätzlich zu entkoppeln und die Bemühungen der Arbeitsgruppe vordringlich auf das Textcorpus-Projekt zu konzentrieren. Eine Neuausgabe des Wörterbuchs ist frühzeitig bis mittelfristig nicht zu erreichen. Nach Maßgabe der Möglichkeiten der Berliner Arbeitsgruppe und des vordringlich Wünschbaren und Erreichbaren sollte im Vorfeld der langfristig anzustrebenden Wörterbuch-Neuausgabe mittelfristig ein Textcorpus erstellt werden. Dieses sollte in möglichst einfacher und praktikabler Form alle Texte umfassen.

W. SCHENKEL plädierte nochmals im Hinblick auf die Abschließbarkeit der Arbeit für die Beschränkung auf die Transkription (ohne Graphematik) und für die Benutzung des Programms PPTES. Die Frage der Texterfassung - Transkription oder Hieroglyphen - wurde kontrovers diskutiert.

N. GRIMAL plädierte für Einbeziehung von soviel Information wie möglich und für die Privilegierung der hieroglyphischen Repräsentation. Transkription sei ein ägyptologisches Konstrukt; "ägyptisch" seien allein die Hieroglyphen.

S. SEIDL MAYER hielt dem die Unverzichtbarkeit der Transkription entgegen. Die Texterfassung braucht ein Rückgrat, das die Ordnung vorgibt und an das alle möglichen und nötigen Informationen angegliedert sein können. Für diese Rückgratfunktion kommt allein die Umschrift in Frage. Nur an die phonematische Repräsentation des ägyptischen Lexems können Zusatzinformationen graphematischer (Hieroglyphen), grammatischer, morphologischer und sonstiger Art angegliedert werden. Zu diesen Zusatzinformationen

können vor allem auch Hinweise im Hinblick auf Wörterbuch-relevante Befunde gehören. Schon bei der Erstellung von Textdateien können neue (im alten Wb nicht berücksichtigte) Wörter, interessante Schreibungen, ungewöhnliche Bedeutungen, Besonderheiten der Rektion u.a.m. hervorgehoben werden.

Zwingend notwendige Voraussetzung für ein Textcorpus ist die Schaffung einer Datei der Textcorpusnummern, in der, im Sinne eines Adressbuches, alle relevanten, zu den Texten gehörige Daten, nicht aber die Texte selber, erfaßt sind. Auf der Grundlage dieses Registers soll dann eine Liste der zu erfassenden Texte erstellt werden, wobei neue Texte prioritär aufgenommen werden sollen, letztlich aber der gesamte Bestand erfaßt werden soll.

Die Gruppe einigte sich auf folgende Punkte, die im Sinne von Empfehlungen an die Arbeitsgruppe und ihren Träger, die Konferenz der deutschen Akademien, weitergegeben werden sollen:

1. Die Erstellung einer Datenbank sämtlicher altägyptischer Texte soll vorrangiges Ziel sein.

Dafür müssen so bald wie möglich folgende Vorarbeiten abgeschlossen werden:

- a) das Textnummerncorpus
- b) das Eingabesystem.

Hierfür wird allgemein die Einstellung eines professionellen Programmierers in die Arbeitsgruppe für ein Jahr als unabdingbar erachtet. Die Diskussion der lauffähig vorhandenen und an begrenzter Textmenge überprüften Programme THOT und PPTES ergab, daß keines der beiden Systeme als ausreichend ausgereift, flexibel und vielseitig gelten kann, um zur verbindlichen Grundlage langfristiger und internationaler Arbeit gemacht zu werden. Besonders S. Seidlmayer wies wiederholt darauf hin, daß er PPTES nicht als internationale Norm zur Verfügung stellen kann. Eine internationale Norm kann aufgrund der damit verbundenen Verantwortung nur von professionellen Programmierern in engster Zusammenarbeit mit der Arbeitsgruppe erarbeitet werden. Die Forderung nach befristeter Einstellung eines Programmierers wurde von allen Anwesenden nachdrücklich unterstützt. Bevor nicht ein als internationale Norm konsensfähiges System zur Verfügung steht, kann mit der Arbeit an der Textdatenbank nicht sinnvoll begonnen werden.

2. Die Erstellung einer ägyptischen Textdatenbank übersteigt die Möglichkeiten der Berliner Arbeitsgruppe und kann nur in Form internationaler Kooperation geschehen. Die Einbeziehung in das Dach-Projekt einer Internationalen Ägyptologischen Datenbank (wie von D. v. d. Plas gefordert) erscheint aber nicht wünschenswert. Die Integration sämtlicher Datenbankprojekte unter ei-

nen Hut würde die Arbeit am Wörterbuch unnötig verzögern. Dafür wird aber die Wiederausweitung des Projekts auf mehrere kooperierende Akademien begrüßt (an dem ursprünglichen Projekt waren außer Berlin auch Leipzig, Göttingen und München beteiligt). An eine funktionierende Kooperation können dann leicht weitere Arbeitsstellen im Ausland angeschlossen werden.

3. Die Berliner Arbeitsstelle soll durch studentische Hilfskräfte und Stipendiaten in ihrer Textcorpusarbeit unterstützt werden. Die Einrichtung von Forschungsstipendien soll Ägyptologen aus aller Welt nach Abschluß ihres Studiums einen längeren Aufenthalt an der Arbeitsstelle ermöglichen, der einerseits deren eigenen Forschungen etwa im Zusammenhang von Dissertationen und Habilitationen zugute kommen soll und andererseits als Teilzeitarbeit für das Wörterbuch die Berliner Projekte fördert.
4. An die internationale Ägyptologie soll appelliert werden, der Berliner Arbeitsstelle Materialien zur Verfügung zu stellen, die als Ergänzungen zum alten Wb von Bedeutung sind, um die Arbeit an einem aktualisierten Wörterbuch möglichst zu beschleunigen.
5. Die Berliner Arbeitsgruppe wird auch in Zukunft Arbeitstagungen durchführen, die Spezialproblemen der Arbeit am Textcorpus sowie am Wörterbuch und Spezialwörterbüchern (etymologisches, Graphien-, Valenz-Wörterbuch usw.) gewidmet sind. Dabei sollen auch Vertreter von Nachbarfächern und theoretischen Disziplinen herangezogen werden.
6. Die Berliner Arbeitsgruppe wird eine Art Newsletter herausgeben, der als Publikationsorgan der Tagungsergebnisse, der laufenden Diskussion zu Einzelproblemen an den Projekten der Arbeitsgruppe(n) sowie als Plattform für den internationalen Austausch von Informationen und Materialien dienen soll.
7. Zum Zwecke der Finanzierung des Projektes, insbesondere hinsichtlich anteiliger Arbeitsübernahme durch Stipendiaten, der Entwicklung wörterbuchspezifischer Software und der Beschaffung notwendiger Hardware, soll gezielt nach Sponsoren auch im internationalen Rahmen gesucht werden.

Nach einer ersten Auswertung im Kreise der Arbeitsgruppe (W.-F. Reineke, A. Burkhardt, E. Freier, S. Grunert, I. Hafemann) zusammen mit dem Betreuer des Projektes (J. Assmann) wurden folgende Schlußfolgerungen für die weiteren Arbeiten bis zum Ende des Jahres 1993 gezogen:

1. Der Aufbau der Berliner Wortliste ist - da sie als Hilfsmittel bei der Erfassung neuer Texte dient - vorrangig abzuschließen. Die Berliner Wortliste

(vgl. Beitrag von Grunert) umfaßt neben Angaben aus dem Wörterbuch die aus den Spezialwörterbüchern (Medizinisches Wörterbuch, Drogenwörterbuch) und aus Untersuchungen (Germer: Arzneimittelpflanzen; Harris: minerals; Jones: ships u.a.) sowie die Eintragungen von Meeks: Annuaire Lexicographique und die aus Wörterbüchern für bestimmte Sprachstufen des Ägyptischen (Faulkner, Lesko, Westendorf). Die Berliner Wortliste soll Interessierten zur Verfügung gestellt werden. Dies würde der Forderung nach einem möglichst bald verfügbaren neuen Wörterbuch zwar nicht entsprechen, jedoch wesentlich entgegenkommen.

2. Angesichts der unterschiedlichen Positionen zur Frage der zwingenden Eingabe der Hieroglyphen und zu ihrer Notierung bzw. Codierung und unter Berücksichtigung des Zeitfaktors wird ein Kompromiß praktiziert. Neue Texte werden, beginnend mit der TCN 100001, von den Bearbeitern grundsätzlich in Transkription eingegeben, wobei während der dazu notwendigen Vorarbeiten die Berliner Wortliste als Überprüfungsinstrument dient (Vorhandensein des Wortes und der Schreibung werden überprüft). Neue Wörter und Schreibungsvarianten werden sofort in die Berliner Wortliste aufgenommen.
3. Als Programme zur Textanalyse werden weiterhin sowohl PPTES als auch THOT verwendet; das in Tübingen genutzte Programmpaket TUSTEP soll auf seine Verwendbarkeit für das Wörterbuch überprüft werden. Grundsätzlich soll die zeitgleiche Anwendung der verschiedenen Programme dazu dienen, die Anforderungen an ein zu entwickelndes wörterbuchspezifisches Programm fixieren zu können.
4. Die Arbeiten an der Textcorpusnummern-Datei, die in Verbindung mit der Textidentifizierung stehen, sowie an der Literatur-Datei werden unter Mithilfe studentischer Hilfskräfte weitergeführt.
5. Es werden für die Wb-Newsletters Materialien vorbereitet, wobei die Publikation von Beiträgen zur Arbeitstagung in Absprache mit E. Blumenthal (ZÄS) vorgenommen wird. Weitere Bestandteile sind die Arbeitshilfsmittel
  - Berliner Wortliste - Aleph
  - PM-Objektindex
  - Museums-Konkordanzen
  - Listen der Berliner Thesauri
6. Die Aufteilung der zur Verfügung stehenden Arbeitszeit wird in Anlehnung an die auf der Arbeitstagung geäußerten Vorschläge zu zwei Dritteln auf die neuen Aufgaben und zu einem Drittel auf die sog. Altlasten verteilt.

## Folgerungen für die Zukunft

### Jan Assmann

Die Berliner Tagung hat deutlich gemacht, daß es bei der Wiederaufnahme der 1947 eingestellten Wörterbuch-Arbeit nicht nur darum geht, so schnell und effektiv wie möglich eine aktualisierte Neuausgabe des Ägyptischen Wörterbuchs anzustreben, sondern auch und vor allem darum, der internationalen Ägyptologie wieder eine Zentralstelle einzurichten, in der alle Texte für die lexikographische Auswertung gesammelt werden und von der auf der Basis dieser Dokumentation Auskünfte über sprachliche und textliche Befunde gegeben werden, sei es in Form der Korrespondenz, sei es durch Ermöglichung längerer oder kürzerer Forschungsaufenthalte an der Arbeitsstelle. Diese Funktion hat die Berliner Wörterbuch-Stelle einmal ausgeübt und sie hat damit nicht nur die Ägyptologie in philologischer Hinsicht entscheidend vorangebracht, ja geradezu revolutioniert, sondern auch der deutschen Wissenschaft im internationalen Kontext zu ganz besonderem Ansehen verholfen. Die Wörterbuch-Kommission der deutschen Akademien war ein von den vier damaligen Akademien der Wissenschaften in Deutschland (Berlin, Leipzig, Göttingen und München; Heidelberg wurde erst 1909 gegründet) gemeinsam getragenes Unternehmen, und es war von Anfang an auf eine vollständige Sammlung und lexikographische Erfassung sämtlicher altägyptischer Texte (auch der unveröffentlichten) hin angelegt. Daher war der ursprüngliche Auftrag der Kommission mit Erscheinen des Wörterbuchs (ab 1926) auch keineswegs erfüllt, sondern es wurde konsequent weiter gesammelt und weiter verzettelt.

An diese ursprüngliche und erst 1947, also nach genau 50jähriger Laufzeit aufgegebene Konzeption gilt es wieder anzuknüpfen. Zwar stellt sich die Situation heute so dar, daß aufgrund der inzwischen neu gefundenen und veröffentlichten Texte das Verhältnis der erfaßten zu den noch zu erfassenden Texten etwa als 50:50 bis 40:60 einzustufen ist. Aber dafür sind durch die Möglichkeiten der elektronischen Datenverarbeitung inzwischen auch ungleich schnellere und effektivere Verfahren der Texterfassung erschlossen worden.

Als bedeutsamstes Fazit der Berliner Tagung ist jedoch festzuhalten, daß - anders als bei der ursprünglichen Berliner Konzeption - zwischen Textcorpus und Wörterbuch methodisch strikt unterschieden werden muß. Das Textcorpus ist nicht lediglich als Unterbau des Wörterbuchs anzusehen, sondern als Resultat eines eigenständigen (Rahmen-)Projekts anzustreben. Außerdem sind die Gewichte umzuverteilen. Das Textcorpus entsteht nicht als Nebenprodukt im Umfeld der Wörterbucharbeit, sondern umgekehrt ergeben sich aus der gezielten lexikographischen Auswertung bei der Arbeit am Textcorpus die entschei-

denden Informationen, die in eine Neuausgabe eingehen müssen. Das Textcorpus ist also als das Zentrum des Gesamtprojekts anzusehen, und es ist auch heute wie schon um 1900 nur in der Form einer Zusammenarbeit mehrerer Akademien und darüberhinaus auch ausländischer freier Mitarbeiter zu verwirklichen.

In der heutigen Situation ist klar, daß ein solches Gesamttextcorpus sinnvollerweise nur noch in Form einer Erstellung einzelner Textdatenbanken erfolgen kann, die später einmal zu einer Gesamt-Datenbank altägyptischer Texte zusammenwachsen sollen. Die Existenz derartiger Textdatenbanken für den Gesamtumfang der griechischen und lateinischen Literatur beweist, daß dieser Weg schon jetzt gangbar ist. Die ganz anderen Schwierigkeiten, mit denen eine Datenbank altägyptischer Texte zu rechnen hat, zwingen allerdings 1. zu erheblichen Vorarbeiten, 2. zu einer Aufteilung der Arbeit und 3. zu einer langfristigen Planung. Das Gesamtprojekt dürfte in weniger als 20 Jahren nicht abzuschließen sein.

Um die Berliner Arbeitsstelle für ihre Hauptaufgaben zu entlasten, müssen an anderen Akademien weitere Arbeitsstellen geschaffen werden, an die sinnvoll abtrennbare Teilprojekte ausgelagert werden. Die Hauptaufgaben der Berliner Arbeitsstelle sind

a) kurzfristig:

1. Ausarbeitung geeigneter Programme zur Texteingabe für Textdatenbanken, unbedingt in (befristeter) Zusammenarbeit mit einem professionellen Programmierer;
2. Fertigstellung des Textnummerncorpus als Grundlage für die Erstellung einzelner Textdateien.
3. Bewältigung von "Altlästen"
4. Verwaltung der bisherigen und zukünftigen Dokumentation

b) mittel- bis langfristig:

1. Beteiligung an der Erstellung von Textdateien
2. Koordination der Zusammenarbeit mit den Außenstellen und mit der internationalen Ägyptologie
3. Lexikographische Auswertung der Textdateien
4. Arbeit an einem aktualisierten "Handwörterbuch"

Die Wiederaufnahme systematischer, koordinierter und international verteilter Arbeit an einer Gesamterfassung und philologisch-lexikographischen Auswertung altägyptischer Texte wird die Erforschung der altägyptischen Kultur, soviel ist mit Sicherheit vorauszusagen, in ähnlicher Weise voranbrin-

gen, wie dies in der ersten Phase des Wörterbuch-Unternehmens von 1897-1947 der Fall gewesen war. Vorbedingung dieses Erfolges ist allerdings, daß die Erstellung der Textdateien mit einer philologischen Bearbeitung der Quellen Hand in Hand geht. Die wenigsten Gattungen der ägyptischen Literatur liegen in heutigen Ansprüchen genügenden Textausgaben, Übersetzungen und philologischen Bearbeitungen vor. Für weiteste Bereiche des altägyptischen Schrifttums ist diese Arbeit allererst zu leisten. Unter der Voraussetzung einer solchen Verbindung von philologischer Bearbeitung und elektronischer Datenerfassung bedeutet dann allerdings auch schon jeder Teilschritt des Gesamtprojekts für sich einen erheblichen Fortschritt. Das nur langfristig zu realisierende Gesamtvorhaben einer umfassenden ägyptischen Textdatenbank ist daher als Rahmenprojekt verschiedener Einzelprojekte vorzustellen. Das Rahmenprojekt gehört in jedem Falle in den Zuständigkeitsbereich der Berliner Arbeitsstelle. Die Einzelprojekte dagegen sollten zu einem erheblichen Teil nach Maßgabe fachlicher Kompetenz ausgelagert werden, und zwar einerseits auf einzurichtende Außenstellen an deutschen Akademien, andererseits auf sonstige in- und ausländische Institute und Wissenschaftler, die sich um Mitarbeit an dem Gesamtprojekt bewerben und bereit sind, sich auf die gemeinsamen Verfahren und Prinzipien der Texterfassung zu einigen.

Der wissenschaftliche Gewinn einer solchen Verbindung von zentralem Wörterbuchprojekt und dezentralen gattungsspezifischen Textcorpusprojekten ist gar nicht hoch genug zu veranschlagen, zumal die Projekte auf bestehenden Forschungsschwerpunkten aufzubauen und sich langjährige Erfahrungen auf den entsprechenden Gebieten zunutze machen können. Hier könnte mit verhältnismäßig geringen Mitteln eine Forschungsarbeit von internationaler Bedeutung in Angriff genommen werden, die das Fach an einem zentralen Punkt entscheidend fördert und Wissenschaftler der alten und neuen Bundesländer in einem gemeinsamen Projekt verbindet. Die internationale Fachtagung hat deutlich gemacht, daß nach wie vor die Erwartungen der internationalen Ägyptologie sich in diesem Punkt auf Berlin und die deutsche Ägyptologie richten, daß konkurrierende Projekte nicht bestehen und daß, falls das Projekt in dem geplanten Umfang in Angriff genommen werden sollte, mit internationaler Zusammenarbeit gerechnet werden kann. Auf längere Sicht wäre also, über die innerdeutsche Integrationswirkung hinaus, mit einer europäischen Integrationswirkung zu rechnen.



## The next generation of the Wörterbuch

Nigel Strudwick

My thoughts on the future of the Wörterbuch are restricted mainly to some ideas of how it should be distributed when its second edition is ready. I would only offer one short piece of general advice to those who are actively compiling the work. I would advise against making the Wörterbuch in whatever final form it takes all things to all men, this will only make it unwieldy and unusable. I have no fundamental objection to the input of information to the computer such as grammatical analysis, but it should not be a prerequisite of the form, if nothing else because scholars cannot agree on the details. I realise that people cannot also agree on reading and meaning, but that can easily be handled. A dictionary is an alphabetical word-list, with references, which are absolutely essential in the case of Ancient Egyptian. I would certainly advocate that even more references be included in the future, since this is a fundamental role of the Wörterbuch. I do not feel that it is essential that the references be included in hieroglyphic form; although interesting and in an ideal world a very good idea, I would argue that it would waste time that the compilers could better spend on getting the basic dictionary functions together. I would say that this was clearly realised by the original compilers of the Wörterbuch.

### Aspects of publication and dissemination

As one who would not consider himself a philologist, I would prefer to keep the bulk of my remarks to suggestions as to additional forms of publication and dissemination of the finished work, namely those using electronic media. I do not wish to see the demise of the printed book, and a further printed version of the Wörterbuch would be extremely useful. Nonetheless, printing costs are rising, and with the present reprints of the dictionary costing in the region of 900.- DM, the likely cost a new Wörterbuch is impossible to estimate.

In the course of the Berlin meeting in September 1992, at which this paper was originally presented, I felt most strongly that the question of the final form of the publication was being ignored, or at least not being given enough consideration.

There is indeed a case for considering at the present only the issues of encoding and treatment of texts, and of the form in which the material should exist in the computer database. A strong argument against over-concern about the final form is that, if media additional to the printed book are to be considered for the publication, it is pointless to speculate on what might be possible, given the pace of change, since that publication is a number of years away.

I accept this up to a point. That point is very simple: unless we have some concept of the sort of material we wish to make available to readers in

the new Wörterbuch, data cannot be collected and arranged in the most useful fashion. This is my only real concern in relation to the present project.

I should now like to turn to some individual subjects. Dictionaries are the ultimate reference work, and it is absolutely essential that they can be accessed in the course of other work. This should be kept in mind when considering ways in which the Wörterbuch can be electronically published.

There are a number of possibilities. The most popular one today is to publish the Wörterbuch on CD-ROM, accompanied by small management programs which can allow the material to be accessed. This would mean that the Wörterbuch can effectively be re-published on a regular basis, depending on how much new material is brought in; CDs are now relatively inexpensive to make, and the whole database can be placed onto a CD and disseminated much more quickly than if it were going to a printer. A very recent development is the appearance of CD ROM production drives which can be attached to present office computer equipment, and which allow the user to make his own discs, rather than having to send them away to a mastering factory. In this way, once the initial investment for the drive has been made, the Wörterbuch project will not have to estimate how many discs to be made at one time. CD ROM is more cost-effective, since no-one is going to pay for a new printed version every 2/3 years, but a CD service could be set up for a subscription. The Wörterbuch could also be kept 'in print' this way much more easily, and hopefully might just make this version cheaper than the original, but that of course depends so much on financing and other commercial factors.

Another method which is less technically sophisticated but which allows for more rapid updates of material is the use of the FTP archive. FTP stands for FILE TRANSFER PROTOCOL, and is a means of accessing read-only electronic archives using international academic and commercial networks. If the Wörterbuch project were to be connected to the central computer resources at the Humboldt-Universität, an archive could be kept into which data files relating to certain aspects of the Wörterbuch material could be kept and updated very regularly. It would be incumbent on scholars who use the archive to find the best way of using it; it ought to be possible to keep information here which is not available either in the printed version of the book or in the CD ROM.<sup>1</sup>

### An example

As well as electronic media through which any future Wörterbuch should be disseminated, software has to be provided to make the material in the database

<sup>1</sup> For an explanation of FTP, see my article 'Electronic communication—possibilities for Egyptology' presented at the 1993 Geneva 'Informatique et Egyptologie' meeting. Until the proceedings of that conference are published, a copy will be available via anonymous FTP at the site newton.newton.ac.uk in the directory pub/ancient/egypt under the name Strudwick.Geneva.txt.

accessible to users. With the caveat mentioned above about how quickly software goes out of date, I developed very rapidly for the Berlin meeting a Hypercard stack which shows one way in which the Wörterbuch could be put out. This is in no way intended to be a formal suggestion, but just one way in which the material could be used in a user-friendly manner.

Hypercard is a pioneering program in "Hypermedia" which handles data in many different ways. It is written for the Macintosh and comes with each new machine, but there are comparable programs which run under MS DOS. Hypercard uses as its basic analogy for each file a set of "cards", which in Hypercard terminology is called a "Stack". The program does not use a traditional data structure in the manner of xbase programs; the concept is that it parallels the way in which scholars use index cards. Control buttons can be added to cards, and these can be programmed in Hypercard's own language, which includes various special effects as well as more conventional programming constructions. Graphics can also be added to cards, and it is also possible for the program to display photographs. Externally programmed commands can be added to the stacks to extend the range of capabilities.

The advantage of hypermedia is that one does not have to proceed linearly through the cards. Specific links can be made from one card to another, or the programming language can be used to jump dynamically from one card to another. Thus one can easily go where the reader/user wants to, rather than having to follow pre-set paths. These facilities which make Hypercard and its comparable applications so useful as a way of getting through a disparate mass of material.<sup>2</sup>

Some cards from the sample stack are shown in figures 1-4.

Fig. 1 shows the entry for the enclitic particle *ȝ*

Fig. 2 shows the Belegstellen for *ȝ*

Fig. 3 shows the index of words in Egyptian order in the stack

Fig. 4 shows the index of words in English translation in the stack

The buttons at the bottom of the cards allow the user to navigate to different options. In the word index lists, clicking on a word will take the user to that entry. The emphasis is that there is no fixed approach to the data, but rather that the user can move where he wishes, when he wishes. It is also quite simple to add personal notes to a card, rather like writing in the margins.<sup>3</sup>

The advantage of a hypermedia (or similar) approach as a way of distributing the Wörterbuch is that it can be operated in a manner akin to that in which most Egyptologists think: that is, we browse around in information, jump forward a couple of pages as ideas strike us, and then go back to where

<sup>2</sup> Some further examples of use of hypermedia may be found in my paper 'Data-Management in Research and Teaching', *Informatique et Egyptologie* 7 (1990), 137-46.

<sup>3</sup> A copy of this stack, under the name Woerterbuch.hqx, is available via anonymous FTP at the site newton.newton.ac.uk in the directory pub/ancient/egypt.

we were. A more conventional database approach, although undoubtedly more efficient in certain circumstances, such as searching a database for all examples of a particular combination of circumstances, does not really reflect the way in which we work with a dictionary. It is also possible with a computer to incorporate bookmarks into such a system as I am describing; versions of the Wörterbuch in languages other than German are likewise relatively easy to produce.

I do not feel I can stress this approach too strongly - we must not, if we are going to publish the Wörterbuch in any method other than paper, restrict ourselves to computer approaches which are conventional, and by that I mean the regular database. Hypermedia approaches offer us a wealth of possibilities to begin to emulate the manner in which we approach the information which can be stored in a computer. I believe that the main way in which we are going to be able to convince Egyptologists to make greater uses of computers is for the information on them to be presented in ways in which they do not appear to require a complete change of approach. Once users can be convinced that we do not have to change our ways dramatically, then I am sure that progress will be more rapid.

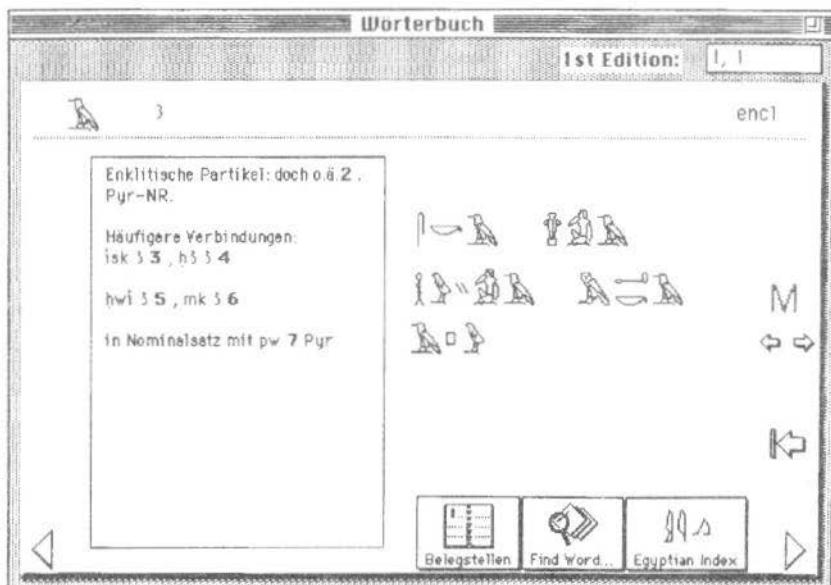


Fig. 1 Card for the enclitic particle 3



Fig. 2 Belegstellen for 3

**Wörterbuch**

**Egyptian Index**

}	(nm)	Egyptian vulture	↑
}	(encl)		
}	(vt)	Go into a place, Tread	
ḥr̥	(vt)	Do	
ḥry	(nm)	Supervisor	
pr	(nm)	House, Estate	
nswt	(nm)	King	
ḥtp	(nm)	Offering	

M  
↔ ↔

Rückläufiges Wb. English Index Update List

Fig. 3 Index of words in Egyptian order in the stack

**Wörterbuch**

**English Index**

Do	ḥr̥	↑
Egyptian vulture	}	
Estate	pr	
Go into a place	}	
House	pr	
King	nswt	
Offering	ḥtp	
Supervisor	ḥry	
Tread	}	

M  
↔ ↔

Egyptian Index Rückläufiges Wb. Update List

Fig. 4 Index of words in English translation in the stack

## Das Wörterbuch der Ägyptischen Sprache

- Stand, Probleme, Perspektiven -

Walter F. Reineke

Als im Jahre 1985 mit der Gründung der "Wörterbuchkommission" an der Berliner Akademie unter aktiver Teilnahme des vor kurzem tragisch verunglückten Fritz Hintze Überlegungen und Vorarbeiten einer computergestützten Version des Wörterbuchs am damaligen Zentralinstitut für Alte Geschichte und Archäologie begonnen hatten, waren sich die Mitglieder dieser Kommission darin einig, daß die künftige Wörterbuch-Arbeit folgende Schwerpunkte haben sollte:

- Aufarbeitung aller im Wörterbuch-Archiv vorhandenen Materialien
- Schaffung einer einheitlichen Kartei aller verzettelten Texte, d.h. Einordnung der sog. Nachverzettelung (1927 - 1947 nur nach dem Lemma sortiert) in die Hauptverzettelung (1897 - 1926, fein sortiert)
- Ausarbeitung einer Textcorpus-Nummern-Datei, in der unter der jeweiligen laufenden Nummer alle Angaben zum Text zu finden sind (z.B. Fundort, Datierung, Aufbewahrungsort, Schriftträger, Schriftart, Textgattung u.a.m.)
- Ausarbeitung eines computerinternen Wörterbuchs (CIWB), das alle wesentlichen Angaben des Wörterbuchs und neuere Erkenntnisse der ägyptischen Wortforschung enthalten sollte. (Im Laufe der Arbeiten zur Schaffung des CIWB zeigte es sich, daß dieser strategische Ansatz für weitere Wörterbuch-Arbeiten nicht optimal war. Vielmehr wird als Analyse- und Eingabeinstrument eine Wortliste benötigt, die neben den Grundangaben des Wörterbuchs und publizierter Glossare bzw. Spezialwörterbücher eine große Zahl anderer Angaben zum Lemma enthält).
- Weiterführung des vorhandenen umfangreichen Corpus ägyptischer Texte.

Dieser Orientierung entsprechend haben die Mitarbeiter des Wörterbuch-Teams ihre Arbeiten - neben anderen, damals akademiespezifischen - seit 1988 organisiert. Relativ problemlos, wenn auch zeitaufwendig, ist die Zusammenordnung von Haupt- und Nachverzettelung. Mehr als drei Fünftel des vorhandenen Materials sind inzwischen eingearbeitet. Ergebnis ist eine Kartei, nach dem Muster der alten Wörterbuch-Kartei angeordnet, in der jeder Belegzettel eindeutig einer Wörterbuch-Stelle zugeordnet ist. Für eine große Anzahl ägyptischer Wörter ergaben sich bei der Zusammenordnung neue, bislang nicht aufgeführte Schreibungsvarianten; die im Wörterbuch angegebenen Belegzeiträume können in vielen Fällen entscheidend verändert werden. Einige neue Wörter tauchten auf; bei manchen Wörterbuch-Einträgen ist die Tilgung notwendig, weil neue Belege aus der sog. Nachverzettelung die Zuordnung zu einem anderen Lemma geraten erscheinen lassen. Die Zusammenord-

nung der beiden Zettelarchive (1.75 Mill. + 0.75 Mill. Zettel) wird zügig weitergeführt. Damit steht der geplanten Computerversion des Wörterbuchs eine einheitliche, in sich geschlossene und mit allen notwendigen Referenzen versehene Zettelkartei für den Altbestand zur Verfügung.

Die Ausarbeitung der Textcorpus-Nummern-Datei (TCN) ist zu Beginn der Arbeiten zur Weiterführung des Wörterbuchs mit großer Intensität betrieben worden und für alle Berliner Texte, die im Rahmen des Datenverbundes der Berliner Ägyptologie zu bearbeiten waren, zu Ende geführt worden. Die TCN-Datei gilt als dringendes Desiderat für jede weitere Wörterbucharbeit, schafft doch nur sie eindeutig Klarheit über die im Wörterbuch-Archiv verarbeiteten Texte und macht auch neues Material zweifelsfrei ansprechbar. Schon bei der Konzipierung der Architektur für die TCN-Datei war deutlich geworden, daß verschiedene Angaben immer wieder auftauchen und deshalb gut geeignet sind, nur durch Nummerncodes mittels einer Hilfsdatei, die zu diesen Codes die notwendigen Angaben enthält, in der Hauptdatei TCN vertreten zu sein. Das betrifft insbesondere

- Fundort
- Aufbewahrungsort
- Datierung
- Schriftart
- Schrifträger
- Textgattung
- Literatur zum Text.

Neben codierten Angaben zum Text werden in der TCN-Datei auch Referenzen aus dem LÄ, bei PM, LD sowie zu den Inventarnummern der Museen gegeben.

In der Anfangsphase dieser Arbeit wurden alle diese Informationen uncodiert eingegeben, doch zeigte sich bald, daß hier dringend Änderungen in Richtung der erwähnten Hilfsdateien geschaffen werden mußten. Einige dieser Angaben sind identisch mit denen, die für die Beschreibung ägyptischer Denkmäler in der Datenbank der Museen in Berlin erforderlich sind. Deshalb wurden auch von der Wörterbuch-Arbeitsstelle Thesauruslisten erarbeitet, die - wie es die Verzahnung innerhalb des Datenverbundes der Berliner Ägyptologie notwendig macht - von allen Beteiligten dieses Verbundes genutzt werden können (so für die Datierung, Schriftarten, Textgattungen und für Ortsnamen als Fund- bzw. Aufbewahrungsort von Aegyptiaca). Die TCN-Datei verzeichnet intern nur die numerischen Codes für die Einzelpositionen in den verschiedenen Thesauruslisten; die Eingabe erfolgt unter Nutzung der Hilfsdateien in für den Bearbeiter sichtbarer Form.

Dies gilt auch für die Eingabe aller notwendigen Literaturangaben, die anfangs zwangsläufig eine Häufung gleichförmiger Informationen ergaben (z.B. Berliner Inschriften..., CG..., RAD... usw.). Hier wurde entschieden, eine

Hilfsdatei "Ägyptologische Literatur" nach einem numerischen Codesystem aufzubauen, aus der dann jeweils die geforderte Referenz für TCN abgerufen wird. Für die Ausarbeitung dieser Hilfsdatei, die inzwischen zur wohl größten Sammlung ägyptologischer Titel angewachsen sein dürfte, standen uns nach Absprachen mit den jeweiligen Autoren in gegenseitigem Einvernehmen deren Computerdateien zur Verfügung.

- Christine Beinlich-Seeber, Ägyptologische Bibliographie 1822 - 1946, Würzburg 1991;
- Daten der AEB, Leiden 1947 ff., auf der Grundlage einer Datei von Erhard Graefe (bis 1960);

Eingearbeitet sind durch eigene Aktivitäten die AEB-Daten von 1961 bis 1970 sowie von 1983 und 1984; Daten aus einem PEB-File wurden bisher nicht aufgenommen, da deren Zählungsstrukturen nicht der potentiellen AEB-Zählung entsprechen; die Berliner Literaturdatenbank ist weitestgehend um eine Kompatibilität mit den AEB-Daten bemüht. Dies gilt vor allem so lange, wie die Datenaufnahme der vorliegenden AEB-Bände nicht abgeschlossen ist. Als Ergänzung zu den genannten Bibliographien stehen uns jetzt ferner Materialien zur Verfügung, die uns Wolfgang Helck kurz vor seinem unerwarteten Tod zugesagt hatte. Es handelt sich dabei um Literaturangaben aus den Jahren 1939 bis 1945, die vielfach in keiner Bibliographie aufgenommen sind.

Die Ausgangsdaten - sofern wir sie als Dateien erhalten haben - wurden so weit manipuliert, daß eine Anpassung an die Datenstruktur des Berliner Datenverbundes gegeben ist. Tragendes Prinzip ist, daß jede Publikation Monographie, Artikel, Katalogteil, CAA-Blatt) eine Code-Nummer erhält, bei der die ersten vier Ziffern einer achtstelligen Zahl das Erscheinungsjahr repräsentieren; die verbleibenden vier Ziffern entsprechen nur für den Zeitraum der AEB-Datei weitgehend der dort publizierten Zählung. Ansonsten wird eine laufende Zählung innerhalb eines Jahrgangs nach Eingang der Daten vorgenommen.

Beginnt man mit der Aufarbeitung eines derart umfangreichen Materials wie der ägyptologischen Fachliteratur, so ist es bald deutlich, daß auch innerhalb der Literaturangaben einige Items immer wieder auftauchen, so z.B. die Abkürzungen für Zeitschriften und Reihen. So erschien es ökonomisch, zur Hilfsdatei "Ägyptologische Literatur-Datenbank" auch noch eine der gängigen Abkürzungen zu erarbeiten. Diese Abkürzungsliste, die weit über die Listen von AEB und LÄ hinausgeht, wird demnächst in der Reihe *Egyptology* durch den Internationalen Ägyptologen-Verband veröffentlicht werden.

Unabdingbar für die Analyse und Eingabe neuer Texte mittels Computer, die die eigentliche Hauptaufgabe des Berliner Wörterbuch-Teams darstellt, ist - entsprechend der Strategie des Wörterbuchs - eine maschinenlesbare Liste, in der alle bekannten ägyptischen Wörter der unterschiedlichen Zeit- und

Sprachstufen verzeichnet und mit numerischen Codes, der WortCorpusNummer, versehen sind. Über den numerischen Code sind die verschiedensten Angaben zum entsprechenden Wort abrufbar, wie beispielsweise dessen semantisch-grammatische Zuordnung, der Belegzeitraum und die Schreibungsvarianten. Seit der ersten Vorstellung einer solchen Liste auf der Wörterbuch-Konferenz in Berlin (September 1992) sind im Ergebnis interner und externer Diskussionen eine Vielzahl von Veränderungen und Verbesserungen notwendig geworden. Im vorliegenden Heft der Mitteilungen wird ein Teil des Datenbank-Materials zur Liste der Worteinträge unter Aleph allgemein zugänglich gemacht. Hilfreich für die Bearbeitung der ägyptischen Lexik war eine Computerdatei, die Horst Beinlich anhand des gedruckten Wörterbuchs erarbeitet und uns zur Verfügung gestellt hatte.

Die Datenquellen, die für die Berliner Wortliste (BWL) ausgewertet wurden, sind außer dem Wörterbuch und WbZ

- Andreu - Cauville, *Vocabulaire absent au Wörterbuch*
- Cerný, *Coptic Etymological Dictionary*
- Erichsen, *Demotisches Glossar*
- Faulkner, *Ancient Egyptian Coffin Texts*
- Faulkner, *Dictionary of Middle Egyptian*
- Gardiner, *Onomastica*
- Germer, *Flora und Arzneimittelpflanzen*
- Grundriß der Medizin der Alten Ägypter
- Jones, *Glossary (Ships, Parts of Ships)*
- Lesko, *Dictionary of Late Egyptian*
- Meeks, *Annuaire Lexicographique*
- Osing, *Nominalbildung*
- Pusch, *Register zu Junker, Giza I - XII*
- Schenkel, *Sargtextkonkordanz*
- Sethe, *Pyramidentexte*
- Westendorf, *Koptisches Handwörterbuch*

sowie eine Reihe anderer Referenzlisten zur ägyptischen Lexik.

Die Arbeit an den Worteinträgen der BWL zu Aleph hat allerdings und bedauerlicherweise gezeigt, daß die Wunschvorstellung eines geprüften Verzeichnisses aller ägyptischen Wörter mit den dazugehörigen Informationen wegen der notwendigen umfangreichen Recherchen unmöglich in solch einer Zeit zu realisieren ist, die uns recht bald die Eingabe neuer Texte gestattet. Da bei der Aufnahme neuer Texte ohnehin die Einträge der BWL nicht nur konsultiert sondern auch durch die Worteinträge der neuen Texte ergänzt werden, ergibt sich bei einer langfristigen Anwendung der BWL automatisch deren Verifikation. Daher haben wir uns inzwischen dazu entschlossen, zunächst nur eine grob bereinigte und geringfügig durch neuere Literatur ohne Tiefenprüfung vervollständigte Liste der Worteinträge zu erarbeiten. Mit der Analyse

und Eingabe neuer Texte auf der Grundlage der vorläufigen BWL wird diese ständig verbessert, sie wird angereichert durch neue Wörter, neue Übersetzungen, neu nachweisbare unumgängliche Streichungen, neue Belegzeiträume u.a.m.

Die enge Verbindung der BWL mit der Eingabe neuer Texte resultiert aus Erfahrungen, die mit dem Eingabe- und Analyseprogramm PPTES von Stephan J. Seidlmaier gemacht wurden. Diesem umschrift-orientierten Programm hinterliegt ebenfalls eine Wortliste, die zur Generierung lemmatisierter Formen ständig befragt und - falls erforderlich - verbessert werden kann. Das Programmpaket THOT von Jochen Hallof und Hans van den Berg wurde ebenfalls zur Texteingabe getestet. Ergebnis dieser Tests war die Einsicht, daß unter Nutzung von Ideen und Komponenten beider Verfahren ein eigenes Wörterbuch-Programm geschaffen werden muß. Diese Einsicht schien sich schon auf der Wörterbuch-Konferenz herauszukristallisieren. Es wurde damals angelegt, der Wörterbuch-Gruppe für ein Jahr einen professionellen Programmierer beizugeben, der in ständigem Kontakt mit den Ägyptologen dieses Programm erarbeiten soll. Es bleibt Aufgabe der nächsten Zeit, hierfür die entsprechenden Lösungen zu finden, denn ein spezielles Wörterbuch-Programm ist unverzichtbar.

Doch damit ist die Wörterbuch-Arbeit keineswegs erschöpft, insbesondere nicht die Bewältigung der "Altlästen". Ein großes Problem besteht in der Ansprache bzw. Benennung der Texte. Die ägyptologischen Vulgarbezeichnungen, wie Bauer, Traumstele, Chetaschlacht u.ä. werden über TCN sukzessive sowohl in wissenschaftlich vertretbare Bezeichnungen (nach Museumsnummern bzw. nach Standort entsprechend PM) als auch in numerische Codes umgesetzt. Große Schwierigkeiten ergeben sich aber aus der z.T. mehrfachen Umstellung von Museumsnummernsystemen. Hier sind im Wörterbuch-Material und dementsprechend natürlich auch in den Belegstellen jeweils die zum Zeitpunkt der Verzettelung gültigen Nummern verzeichnet. Viele Texte besaßen zur Zeit ihrer Aufnahme zudem noch keine Museums-Nummern und erhielten daher eine interne Wörterbuch-Zählung. Die Identität dieser Stücke ist zu ermitteln. Es besteht die dringende Notwendigkeit, Konkordanzen der damaligen, im Wörterbuch gegebenen Nummern mit den heute gültigen zu erarbeiten. Dies ist zeitaufwendig und - falls nicht kooperationsbereite Kollegen wie in Wien, London und Paris unsere Ausgangslisten be- und überarbeiten - ohne Arbeit am Ort nicht zu bewerkstelligen. Die hierzu notwendigen Recherchen wurden für eine Reihe von Sammlungen durchgeführt; zu nennen sind z.B. italienische, englische und französische. Für einen Teil der im Wörterbuch benutzten Texte aus Kairo konnten entsprechende Angaben durch das Entgegenkommen der ägyptischen AltertümERVERWALTUNG in den verschiedenen Verzeichnissen (CG, JE, Temporary Register) ausfindig gemacht werden.

Auf der Wörterbuchkonferenz wurde auf die Dringlichkeit dieser Arbeiten mehrfach hingewiesen und die Veröffentlichung der Rechercheergebnisse gefordert. Diesem wird nachgekommen, und in Kürze werden in den Mitteilungen neben weiteren Teilen der BWL erste Konkordanzlisten publiziert werden. Dies gilt auch für Kataloge zu den einzigartigen Sammlungen, die sich im Wörterbuch-Archiv befinden. Neben den Originalzeichnungen und den Papierabklatschen der Lepsius-Expedition handelt es sich dabei um ca. 10000 Abklatschserien, die seit der Mitte des vorigen Jahrhunderts von ägyptischen Altertümern vorzugsweise in europäischen Museen, von bedeutenden Denkmälern in Ägypten oder von Stücken aus dem Handel genommen wurden. Diese Serien bestehen zumeist aus mehreren Einzelblättern, können aber - wie bei den Pyramidentexten, den Philä-Inschriften und denen der nubischen Tempel - auch sehr umfangreich sein. Nur ein geringer Teil dieses Bestandes war seinerzeit im Rahmen der Wörterbucharbeit zu Beginn unsers Jahrhunderts inventarisiert worden.

Ein wichtiges Desiderat, das in der Berliner Wörterbuch-Arbeit zu allen Zeiten kaum eine Rolle gespielt hat, ist die Aufnahme von Belegen zur Wortsdiskussion, d.h. die Durchsicht der gesamten einschlägigen ägyptologischen Literatur auf Hinweise zu neuen Wörtern bzw. bei den schon vorhandenen zu neuen Schreibungen, Bedeutungen usw. Für einen bestimmten Zeitraum sind die Resultate einer solchen Arbeit bei Meeks gesammelt. Für den Fortgang der erforderlichen Recherchen muß jedoch wesentlich mehr Material verarbeitet, d.h. durchgesehen werden - mindestens alle Veröffentlichungen von 1925 an unter der Annahme, daß bis zur Ausarbeitung des ersten Bandes des Wörterbuchs die Autoren (Erman, Sethe und Grapow) alles Einschlägige parat hatten, es somit im Wörterbuch verarbeitet ist. Die bei dieser Arbeit neu gefundenen Wörter werden sofort der BWL zugefügt, ebenso Präzisierungen der Wortbedeutung.

Die Aufgaben für die Schaffung eines neuen Corpus ägyptischer Texte mit Hilfe von Computern als Grundlage jeder weiteren lexikographischen Arbeit sowie die Schaffung aller genannten Analysehilfsmittel sind so vielschichtig und bedürfen der Durcharbeitung eines derart immensen Materials, daß - wie in der Anfangsphase der Wörterbuch-Arbeit - die Zusammenarbeit mit anderen deutschen Akademien notwendig ist; Vorstellungen hierzu bestehen bereits. Auch wird die internationale Kooperation zur Bewältigung der Aufgaben unumgänglich sein.

# Das Abklatscharchiv beim "Wörterbuch der ägyptischen Sprache" (Arbeitsbericht)

Silvia Köpstein

Das Langzeitprojekt "Wörterbuch der ägyptischen Sprache" an der Berlin Brandenburgischen Akademie der Wissenschaften verfügt über ein in der Welt einzigartiges Archiv von Papierabdrucken altägyptischer Denkmäler. Das Archiv umfaßt 550 Kästen mit ca. 10000 Abklatschserien, die zum großen Teil zwischen 1833 und 1910 von Karl Richard Lepsius, Heinrich Brugsch, Auguste Mariette, Edouard Naville, Wilhelm Spiegelberg, Karl August Reinhardt und Hermann Junker genommen wurden und Stücke aus Ägyptischen Museen in Europa, Amerika und Ägypten sowie seinerzeit noch stehende Denkmäler in Ägypten und Nubien dokumentieren. Sie stellen einen Schatz von musealen Wert dar, da viele der damals aufgenommenen Denkmäler heute verloren oder nicht mehr in dem damaligen Zustand erhalten sind (Überflutung nach Staudammbauten, Zerstörung durch Erosion, Raubgrabungen, Bauarbeiten bzw. Abbruch der Bauwerke zur Gewinnung von Baumaterial in Ägypten, Kriegseinwirkungen in europäischen Museen).

Ein kleiner Teil davon (1800 Serien) war bis zum Beginn dieses Jahrhunderts inventarisiert worden. Nach dem 2. Weltkrieg wurde in den 60er und 70er Jahren der Bestand allgemein gesichtet, neu aufgestellt und ein erste Katalogisierungsversuch unternommen. Das Material, das Karl Richard Lepsius zwischen 1842 und 1845 in Ägypten sammelte (560 Serien), ist dann in den 80er Jahren mit den aktuellen Bestimmungen versehen und reinventarisiert worden. 1991 wurde für die Aufarbeitung des Archivs eine ABM-Stelle für 2 Jahre (1.11.1991-31.10.1993) eingerichtet.

Während dieser zwei Jahre konnten die Abklatsche aus europäischen Museen sowie aus dem Antikenhandel vollständig aufgearbeitet und für die anderen Archivbestände ein Überblick gewonnen werden. Nur durch den Einsatz eines Computers ließen sich die Arbeiten in dieser Zeit bewältigen. Die computerinterne Erfassung der Bestände ermöglicht nun den schnellen Zugriff auf die Bestände unter unterschiedlichen Fragestellungen, sowohl bei dem aufgearbeiteten Material als auch bei dem noch nicht bearbeiteten, und wird die Fortsetzung der Arbeiten erleichtern. Darüber hinaus ist sie die Grundlage für die geplante Publikation des Bestandskatalogs.

Zu den bearbeiteten Abklatschserien aus europäischen Museen gehören 1300 Serien, die zum großen Teil von Karl Richard Lepsius 1833-1866 und von Heinrich Brugsch 1848-1853 angefertigt wurden. Die ca. 200 Serien von Stücken aus dem Antikenhandel stammen von Karl August Reinhardt, der sie während seiner Dolmetschertätigkeit am Kaiserlichen Deutschen Generalkonsulat Kairo zwischen 1894 und 1899 nahm.

Die Blätter wurden identifiziert und - soweit möglich - die aktuellen Museumsinventarnummern der Vorlagen ermittelt. Da sich die Inventarnummern in den Museen innerhalb der letzten 100 Jahre vielfach geändert und die Museen ihre Bestände in unterschiedlichem Maße aufgearbeitet haben, waren z.T. langwierige Sucharbeiten über nur deskriptive Kataloge (ohne Abbildungen) notwendig. Zusätzlich zu den aktuellen bibliographischen Angaben wurden die Vollständigkeit der Serien sowie Urheber und Fertigungszeit der Blätter vermerkt. Darüber hinaus wurde angegeben, ob das Stück für das "Ägyptische Wörterbuch" verzettelt worden ist. Die Abklatschserien wurden inventarisiert und per Computer in einer Datenbank unter dBase IV, Version 2.0 erfaßt (s. Anlage 1).

Es wurden die im Archiv vorhandenen Serien von Stücken aus den Museen von Berlin, Bologna, Braunschweig, Budapest, Durham, Florenz, Grenoble, Hamburg, Hannover, London, Marseille, Neapel, Oxford, Paris, Petersburg, Pisa, Rom, Turin und Wien aufgearbeitet.

Bei den Serien aus dem Antikenhandel war die Ermittlung des aktuellen Standorts vergleichsweise schwieriger, da für eine sichere Identifikation Abbildungen erforderlich sind. Aufgrund von Porter/Moss sowie entsprechender Museumspublikationen ließen sich für die Hälfte dieser Serien die aktuellen Angaben ermitteln. Bei einem weiteren Teil konnte aufgrund stilistischer Vergleiche die Zugehörigkeit zu bestimmten Gräbern gesichert werden, ohne daß Porter/Moss die im Abklatsch dokumentierten Blöcke verzeichnet hätte. Bei weiteren Serien ließ sich die Herkunft nicht eruieren, da die in Frage kommenden Gräber bislang unpubliziert sind.

Mit der Dokumentation von Stücken aus dem Antikenhandel vom Ende des vorigen Jahrhunderts besitzt das Archiv einzigartiges Material, besonders hinsichtlich der inzwischen verschollenen Grabteile, aber auch bei denen, die nachweislich in Museen gekommen sind. Denn wie der Vergleich der Abklatsche mit den Museumspublikationen ergab, haben die Originale sehr häufig durch den Transport bzw. die Bedingungen in den Museen gelitten, wodurch die Abklatsche einzige Zeugen für den ursprünglichen besseren Erhaltungszustand sind.

Die durch Abklatsche dokumentierten Stücke aus dem Antikenhandel befinden sich heute in den Museen von Berlin, Chicago, Frankfurt/Main, Hamburg, Hannover, Heidelberg, Hildesheim, Karlsruhe, London, München und Straßburg.

Neben der systematischen Aufnahme der Museumsbestände und zugleich als Vorrarbeit dafür habe ich das gesamte Archiv gesichtet und ein aktuelles Verzeichnis der Kästen per Computer aufgestellt (s. Anlage 2). Für die große Gruppe von Abklatschen aus Philae wurde ein separates Verzeichnis angelegt, das bereits die ausführliche Benennung des Urhebers enthält. Damit wurde der Zugang zu dieser Materialgruppe vereinfacht (s. Anlage 3).

Da die Bestände während des Krieges in relativ kurzer Zeit ausgelager werden mußten und dafür Hilfskräfte herangezogen wurden, waren sie teilweise durcheinandergeraten und selbst zusammengehörige Serien über mehrere Kästen verteilt. Sofern möglich, habe ich diese Serien wieder zusammengeführt und darüber hinaus ein computergestütztes Verzeichnis angelegt, das die Zuweisung zu inventarisierten Serien und ihre systematische Aufarbeitung erleichtert wird. (MUS\_REST.DBF, gleiche Struktur wie bei Anlage 1). Auf diese Weise konnten 700 Serien aus ägyptischen Museen bzw. unklarer Herkunft für die Bearbeitung vorbereitet werden. Darüber hinaus wurden die noch offenen Probleme beim Material der Lepsius-Expedition 1842-1845 zum großen Teil geklärt.

Daneben habe ich alle verfügbaren Unterlagen im Archiv des "Ägyptischen Wörterbuchs" und der Staatlichen Museen zu Berlin, die jährlichen Arbeitsberichte des "Wörterbuchs" und des "Ägyptischen Museums" sowie Biographien und Publikationen der betreffenden Texte auf Hinweise zu Provenienz der Abklatsche gesichtet, da es beim "Wörterbuch der ägyptischen Sprache" keine Verzeichnisse gibt, die den Zuwachs der Sammlung dokumentieren. Darüber hinaus halfen Handschriftenvergleiche der Notizen auf den Blättern bei der Verifizierung des Urhebers und damit meist auch der Fertigungszeit.

Die Geschichte des Abklatscharchivs soll 1994 in den "Mitteilungen aus der Arbeit am Wörterbuch der ägyptischen Sprache" zusammen mit dem Katalog der während der ABM-Zeit bearbeiteten Serien aus europäischen Museen erscheinen.

Bisher nicht aufgearbeitet sind die Abklatsche aus den ägyptischen Museen Kairo, Alexandria, Ismailija (ca. 500 Serien), die Serien mit unbekannten Standort (ca. 200 Serien) sowie die von den ortsfesten Denkmälern in Ägypten und Nubien. Dieses Material umfaßt rund 400 Kästen mit ca. 20 Serien pro Kasten (s. Anlage 4). Für die Fortsetzung der umfangreichen Arbeiten mit den Teilbereich Nubien konnten Drittmittel gewonnen werden.

## Anlagen

### 1. Datensatzformat der dB-Datei CAT\INV\_KK.DBF

Feldname	Typ	Länge	Bemerkung
ABKLNR	Zeichen	12	aktuelle Inv.-Nr. des Abkl.
KASTEN_NEU	Zeichen	8	aktuelle Kasten-Nr.
BLATT_Z	Numerisch	4	Anzahl der Blätter der Serie
BENENNUNG	Zeichen	15	Art des Denkmals
SPEZIFIZ	Zeichen	50	Erläuterung zu Feld 4
NAME	Zeichen	50	Besitzer des Denkmals
BESCHR	Memo	10	Beschreibung des Abkl.
DATIERUNG	Zeichen	17	Datierung des Denkmals
FUNDORT	Zeichen	15	Fundort des Denkmals
MUS_AKT	Zeichen	40	aktueller Standort des Denkmals
INVNR_AKT	Zeichen	20	aktuelle Inv.-Nr. des Denkmals
INV_ALT	Zeichen	20	alte Inv.-Nr. des Denkmals
URHEBER	Zeichen	20	Urheber des Abklatsches
URH_ZEIT	Zeichen	20	Fertigungszeit des Abkl.
MUS_LT_URH	Zeichen	32	Standort lt. Urheber
INVNR_URH	Zeichen	15	Inv.-Nr. des Denkmals lt. Urheber
ABKLNR_ALT	Zeichen	15	Abklatsch-Nr. des Urhebers
BEZ_URH	Memo	10	Beschriftung des Abkl. durch den Urheber
MAPPE	Zeichen	30	Wb.-Verzettelung
PM_BAND	Zeichen	2	Band-Zahl bei PM
PM_SEITE	Zeichen	7	Seitenzahl bei PM
PM_SZENE	Zeichen	15	Szenenkennung bei PM
PUBLIK_1	Zeichen	80	Publikation des Denkmals
ERG_LITER	Memo	10	weitere Publikation
VOLLSTKEIT	Zeichen	20	Angabe zur Vollständigkeit der Serie
BEMERKUNG	Memo	10	zusätzliche Bemerkungen zur Serie
FEHLT	Logisch	1	Angabe, ob die ganze Serie fehlt

## 2. Datensatzformat der dB-Datei CAT\KASTEN.DBF

Feldname	Typ	Länge	Bemerkung
KASTEN_NEU	Zeichen	8	
KASTEN_ALT	Zeichen	10	alte Kasten-Nr.
BENENNUNG	Zeichen	15	
SPEZIFIZ	Zeichen	50	
NAME	Zeichen	40	
DATIERUNG	Zeichen	15	
FUNDORT	Zeichen	15	
URHEBER	Zeichen	15	
URH_ZEIT	Zeichen	15	
MUS_AKT	Zeichen	25	
ABKLR	Zeichen	12	
ABKLR_NEU	Zeichen	10	evtl. Abkl.-Inv.-Nr.
ABKLR_ALT	Zeichen	15	
PM_BAND	Zeichen	2	
PM_SEITE	Zeichen	7	
PUBLIK_1	Zeichen	80	
VOLLSTKEIT	Zeichen	20	
KONTROLLE	Logisch	1	
BEMERKUNG	Memo	10	

## 3. Datensatzformat der dB-Datei CAT\PHILAE.DBF

Feldname	Typ	Länge	Bemerkung
KASTEN_NEU	Zeichen	8	
KASTEN_ALT	Zeichen	2	
SERIE	Zeichen	8	alte Seriennummer
BENENNUNG	Zeichen	15	
SPEZIFIZ	Zeichen	80	
DATIERUNG	Zeichen	10	
PM_BAND	Zeichen	2	
PM_SEITE	Zeichen	7	
PM_SZENE	Zeichen	10	
PUBLIK_1	Zeichen	80	
VOLLSTKEIT	Zeichen	15	
KONTROLLE	Logisch	1	
BEMERKUNG	Memo	10	

4. Abklatsche von Denkmälern in Ägypten und Nubien  
 (nach Urheber geordnet)

Urheber	Zeit d. Abkl.	Ort	Kastenzahl
Borchardt	[nach 1898]	Nubien	6
	[nach 1898]	Karnak	5
	1902-1908	Abusir	7
Ebers/Stern	1872-1873	Felsinschriften El Kab	5
Fraas	1897	Hammamat	1
Heintze, Borchardt	1897-1898	Pyramidentexte	33
Jebens	[vor 1886]	Sakkara, Privatgräber	1
Junker, Schäfer	1908-1910	Philae	168
		Edfu	3
		Dakka	7
		Biggeh	3
		Dendur	3
		Debod	3
		Kalabscha	13
		noch unbestimmte nubische Tempel	7
Lepsius	1866	Delta	2
		Memphis	
Mariette	1859	Dendera	37
	[1858-1859]	Abydos	6
Möller	1907	Hatnub	4
Naville	1869	Theben: Königsgräber, Privatgräber	8
		Abu Simbel	
		Philae	
		Abydos	
		Dendera	
Naville	1886-1889	Edfu	
Reisner?	[1920-1923]	Bubastis	15
Roeder	1907-1910	Meroë	1
		Dakka	35
		Beit el-Wali	
		Wadi Gahr	
		Gnaui Scheima	
		Gamli	
		Abisko	
Schliemann	1852	Debod	
Spiegelberg	1895-1896	bei Kalabscha	
		Mittel- und Oberägypten	4
		Graffiti in Theben	7
		El Kab	
		Assuan	
		Debod	
		Tafe	
		Kertassi	
Stubel	[vor 1891]	Beit el-Wali	
?	1890	Kalabscha	
?	?	Wadi el-Arisch	1
		Nimrud Dagh	4
		Assuan, Privatgräber	5
		Summe:	394

## Lexikalische Datenbank contra Wörterbuch Ingelore Hafemann

Die Arbeitsstelle "Ägyptisches Wörterbuch" an der neugegründeten Berlin-Brandenburgischen Akademie der Wissenschaften hat gegenwärtig innerhalb des Wörterbuchprojektes drei Subprojekte.<sup>1</sup>

1. Sammlung von ägyptischen Texten, die nach Abschluß der Wörterbucharbeiten publiziert bzw. bekannt geworden sind. Ziel ist der Aufbau eines Text-Corpus innerhalb einer elektronischen DATENBANK (DB) mit allen Informationen zu den Texten sowie den Texten selbst.

Grundlage ist dabei der Textbestand des Zettelarchivs. Jeder seinerzeit verzeichnete Text erhält eine TEXTCORPUSNUMMER (TCN), wird exakt identifiziert und durch Informationen über Standort, Inventarnummer, Publikationsstand, Datierung, Textgattungsbestimmung etc. erläutert. Er findet damit Eingang in einen TEXTCORPUSVERWALTER (TCV-ALT). Eine Neuaufnahme dieser Texte selbst erfolgt vorerst allerdings nicht. Als Basisarchiv soll der TCV-ALT lediglich eine moderne Referenzgrundlage für die Belegstellen des Ägyptischen Wörterbuchs sein.

2. Aufnahme der Texte selbst, die bisher nicht für die Wörterbucharbeit erschlossen sind, mit einem geeigneten Texteingabe- und Analyseprogramm.
3. Aufbau einer lexikalischen DB. Diese soll möglichst das gesamte lexikalische Wissen aufnehmen, über das die Ägyptologie bereits verfügt. Grundlage ist das ausgedruckte Wörterbuch von Erman-Grapow, das ergänzt wird (incl. Streichungen) durch die Wörterbücher und Wortsammlungen der vergangenen Jahrzehnte wie Faulkner, Meeks, etc. Diese DB wird in ihrem ersten Arbeitsstadium in Gestalt einer Wortliste, der sog. BERLINER WORTLISTE (BWL), realisiert.

Die BWL soll zwei Funktionen erfüllen: So stellt sie einerseits den Kern der geplanten lexikalischen DB dar, wird folglich selbst schon als relationale DB aufgebaut, und ist andererseits als ein Hilfsinstrument bei der Neuaufnahme von Texten konzipiert.

Indem die DB interaktiv mit einem Textanalyseprogramm gekoppelt wird<sup>2</sup>

<sup>1</sup> Vgl. die bereits erschienenen Beiträge zum Berliner Projekt: S. GRUNERT, W. F. REINEKE, Das Berliner ägyptische Wörterbuch, in: GM 129 (1992), 9-20; P. M. WOLF, Ägyptologie und Informatik in Berlin, in: GM 128 (1992), 43-67, S. GRUNERT, I. HAFEMANN, Zum Berliner Projekt eines rechnergestützten Wörterbuches: Problemstellung und Lösungsvarianten, in: Informatique et Egyptologie 8 (1993), 32-41, P. M. WOLF, Standards für den Datenaustausch zwischen Datenbanken in der Ägyptologie, in: Informatique et Egyptologie 8 (1993), 127-136.

<sup>2</sup> Zu diesem Ansatz jetzt auch S. J. SEIDLAYER, Aus der Arbeit an einem philologisch-prosopographischen Texterschließungssystem, in: Informatique et Egyptologie 8 (1993), 104.

und damit beide Funktionen verbunden sind, wird gewährleistet, daß bei der Textanalyse immer das gesamte lexikalische Wissen dieser DB als Hilfsmittel genutzt werden kann. Außerdem wird durch die Analyse neuer Texte mittels der lexikalischen DB diese sowohl um Neueinträge als auch um morphologische und syntaktisch-semantische Varianten schon bestehender Einträge erweitert. Gleichzeitig wächst bei bestehenden Worteinträgen deren Belegzahl. Die konkret technische Verknüpfung beider Funktionen stellt z. Zt. noch ein Problem dar, das durch einen Programmierer zu lösen ist. Im Folgenden soll auf die Vielfältigkeit dieser lexikalischen DB eingegangen werden.

Der konzeptionelle Ansatz ist folgender: Texte, die in Abschnitte und Sätze untergliedert sind, werden in Phrasen und Wörter zerlegt, welche in der lexikalischen DB bereits erfaßt sind. Jeder sprachliche Ausdruck hat mehrere Eigenschaften, die er mit anderen Ausdrücken auf verschiedene Weise teilt. Im wesentlichen gibt es drei Analysedimensionen:

- I. Lexikalische Bedeutung,
- II. Morphologisch-syntaktische Form,
- III. Syntagmatische Funktion.

Im Mittelpunkt unserer Arbeit steht die lexikalische Analyse, d. h. die lexikalische Bedeutung der Wörter, auf deren Grundlage sich semantische Analysen vornehmen lassen. Semantisches Wissen kann repräsentiert werden durch semantische Netze, semantische Relationen (Sinnrelationen), Begriffssysteme oder Wissensbasen. Entscheidend ist die Art der Wissensrepräsentation, d.h. wie stellt man Wissen für die Weiterverarbeitung, konkret für maschinelle Weiterverarbeitung, dar?

In den letzten Jahren sind hierfür unterschiedliche Arten maschineller Wörterbücher erarbeitet worden.<sup>3</sup>

Die einfachste Variante ist das in maschinenlesbarer Form abgespeicherte, ursprünglich gedruckte Wörterbuch.

Weitergehend sind maschinelle Wörterbücher, deren Einträge mit codierten Angaben zur Grammatik, wie Wortart, morphologische Kategorie oder auch Valenzen u.ä. angereichert sind. Diese Arten maschineller Wörterbücher können aus verschiedenen Wörterbüchern zusammengestellt sein. Sie werden z.B. zur halbautomatischen Standardisierung (Lemmatisierung) bei literarischer Textbearbeitung oder Datenaufnahme verwendet. Ein solches Wörterbuch benutzt auch die neue Version des Programms PPTES (vgl. Anm. 2).

Eine lexikalische DB enthält strukturierte und formalisierte Informationen zu den einzelnen Einträgen auf verschiedenen Ebenen und sichern so einen wirklich effektiven Gebrauch der sich ständig entwickelnden Computer-

<sup>3</sup>Vgl. N. CALZOLARI, Computer-Aided Lexicography: Dictionaries and Word Data Bases, in: Handbuch für Sprach- und Kommunikationswissenschaft 4 (1989), 510-519 [ed. by I. S. Batori, W. Lenders, W. Putschke, Computational Linguistics].

technologie. Diese Präsentation lexikalischen Wissens kann völlig neue Ansätze und Anwendungsbereiche für die moderne Lexikographie und auch Lexikologie bieten.

Schließlich gibt es noch computerlinguistische Wörterbücher, die für Parser, für Anwendungen in der künstlichen Intelligenz oder für maschinelle Übersetzungsprogramme angelegt wurden. Sie sind äußerst spezifiziert und je nach Anwendung und Zielstellung sehr verschieden strukturiert. So können sie morphologische Informationen, Angaben zur Subkategorisierung, Strukturen der Prädikatenlogik, semantische Relationen etc. enthalten. Solcherart Wörterbücher sind ganz speziell auf den Parser bezogen, für den sie Daten zur Verfügung zu stellen haben, und sie sehen selbstredend für jede natürliche Sprache anders aus. Sie sind in starkem Maße von linguistischen Theorien abhängig und arbeiten mit entsprechend verschiedenen Formalisierungen.<sup>4</sup>

Tatsächlich sind die vorhandenen Wörterbücher als Resultat lexikographischer Arbeit ja nicht nur Wortlisten, sondern eine Hauptquelle unserer Kenntnisse über Weltwissen und Weltverständnis einer Kultur und ihrer Träger. Bei den Überlegungen, dieses Wissen aus den verschiedenen lexikalischen Quellen möglichst umfassend auszubeuten und dabei Wörterbücher verschiedenen Typs miteinander zu verbinden, stellt sich das Konzept einer lexikalischen DB als höchst fruchtbar dar. Diese ist multifunktional und kann größtenteils unabhängig von Systemtheorien sein.

Elementare Vorüberlegungen sollte man auf das Zusammenspiel von Zielstellung, Methoden der Speicherung und Suchroutinen richten. Man muß eine "Verbindungsmöglichkeit" zu den an sich sehr verschiedenen Typen von Informationen schaffen, um diese dann für eine Systemanalyse einer natürlichen Sprache zur Verfügung stellen zu können. Es handelt sich um morphologische, syntaktische, semantische und pragmatische Informationen, von denen jede verschiedene Anforderungen an Inhalt und Struktur des Eintrages stellt. So scheint uns der effektivste Weg, die o. g. sehr unterschiedlichen und in sich stark strukturierten Informationen mit Methoden zu verarbeiten, die typisch für *database-management-systems* (DBMS) sind. Die Verwendung von DBMS ermöglicht es, die Daten den Notwendigkeiten entsprechend zu strukturieren, ohne bestimmte Nachteile (wie z.B. Redundanz) in Kauf nehmen zu müssen. Es ist möglich, dem Nutzer Zugang zu den Daten zu verschaffen, ohne daß er mit der internen oder physischen Organisation der Daten vertraut sein muß, indem man noch ein vernünftiges Query-User-Interface auf die Datenstruktur programmiert. Es wird mit einer *query-language* gearbeitet, die der natürlichen Sprache recht ähnlich sein kann. Die *search*-Methoden ermöglichen einen direkten bzw. punktuellen Zugriff auf die Daten, im Gegensatz zum

<sup>4</sup> Vgl. ausführlich dazu P. HELLWIG, Parsing natürlicher Sprachen: Grundlagen, sowie ders., Parsing natürlicher Sprachen: Realisierungen, in: Handbuch für Sprach- und Kommunikationswissenschaft 8 (1989), 348-432.

*sequential-reading* anderer Systeme. Allerdings sind bisherige kommerzielle DBMS nicht so gut für unsere oben umrissene Aufgabenstellung geeignet.<sup>5</sup>

Zur Entwicklung einer solchen DB, also zur logischen Organisierung der Daten ist eine Art Modell der lexikalischen Welt zu entwerfen. Dazu muß folgendes festgelegt werden:

1. Objekt-Typ oder *entity-set* (*entity* ist ein Element, das durch einen Namen und bestimmte Eigenschaften charakterisiert ist).
2. Mögliche Beziehungen und Assoziationen zwischen den Eigenschaften des gleichen *sets* oder den Entitäten verschiedener *sets*.

Hierbei ist das relationale Datenbankmodell den Eigenschaften der lexikalischen Welt angemessen, wobei allerdings zukünftige Lösungen wohl durch objektorientierte Datenbankmodelle bestimmt sein werden.<sup>6</sup>

Das Design einer relationalen DB besteht aus der Bestimmung eines *sets* von Relationen, die Fakten beschreiben, mit denen wir umgehen. In unserem Fall ist das *set* der Lemmata die grundlegende Entität, die mit allen anderen Relationen in Beziehung steht. Mit der Vergabe von *primary* und *secondary keys* (Kupplungsglieder zwischen unterschiedlichen Tabellen, deren Werte übereinstimmen) werden potentiell die bestehenden physischen und virtuellen Verknüpfungen hergestellt. Im Moment der Anfrage können die in sich unterschiedlichen Daten nicht nur einzeln abgerufen, sondern auch entsprechend ihren Strukturen beliebig zusammengestellt werden, und zwar jeweils nach den definierten Attributen einer Entität. Dabei ergibt sich die Möglichkeit sowohl von 'eins-zu-eins'-Relationen, wie auch 'eins-zu-viele', 'viele-zu-eins' oder 'viele-zu-viele'. So können beliebige Verbindungen hergestellt werden, die bei einem alphabetisch aufgearbeiteten Wörterbuch selbst dann nicht möglich sind, wenn die Daten im Rechner abgespeichert wurden.

Die lexikalische DB sollte so angelegt sein, daß sie verschiedene linguistisch-lexikalische Fragestellungen befriedigen und klären helfen kann, d.h. verschiedene Nutzer der gleichen Daten auch verschiedenen Zugang zu ihnen bekommen können. Die unterschiedlichen *views* einer DB existieren dabei aber nur virtuell; tatsächlich sind sie definiert in Begriffen der *database-retrieval-functions*, d. h. es erfolgt die Projektion der Daten zu einem *view* bei der Abfrage. Daher können alle Daten ohne physische Änderungen mit großer Freizügigkeit automatisch so umgeordnet werden, wie sie für einen Report entsprechend einer ausgewählten Fragestellung erwünscht sind. Diese Ordnung kann z. B. identisch sein mit der Wortart- oder Satzglied-Bestimmung, die zum Such-Kriterium definiert wurde. So könnte man z. B. alle Adjektive zusammenstellen bzw. alle intransitiven Verben oder alle Substantive, die in Lokalbestimmungen verwendet werden.

<sup>5</sup> Dazu vgl. jetzt auch P. M. WOLF, a. a. O. (1993).

<sup>6</sup> Ders., a. a. O., 128.

Einerseits ist der Zugang zu einem Wort das Wichtigste, und andererseits muß die Computer-Organisation die verschiedenen Ebenen der linguistischen Informationen widerspiegeln. Durch diese Verknüpfung der verschiedenen Ebenen wird ein Wörterbuch, das in Form einer DB organisiert ist, die Grenzen eines normalen, sequentiell alphabetisch geordneten Wörterbuchs sprengen. Voraussetzung ist die exakte Analyse eines lexikalischen Eintrages und die Festlegung seiner Basis-Elemente, d.h. für die Formalisierung eines lexikalischen Eintrages nach den relevanten Kategorien in der DB muß vorher eine Arbeitshypothese erarbeitet werden. Einige solcher Kategorien findet man bereits in ausgedruckten Wörterbüchern codiert wie Wortartbestimmung (Subst., Verb, Adj.), Grammatikangaben (mask., fem., pl., dat., akk. etc.), oder Sprachgebrauchsebene (vulg., dial., poet. etc.) und Sachgebietsangaben (zool., techn., astron. etc.). Die interessante Beziehung zwischen lexikalischen und syntaktischen Aspekten sowie der Semantik und Pragmatik der Wörter könnte in Zukunft effektiver in einer lexikalischen DB abgefragt werden.

Die von uns konzipierte lexikalische DB für das Ägyptische soll vorerst folgende Datenfelder erhalten, die in verschiedenen Listen zusammengestellt werden können :

- |                                 |                                |
|---------------------------------|--------------------------------|
| 1. Wortnummer,                  | 2. Numerische Ordnungsangabe,  |
| 3. Derivationsangabe,           | 4. Transkription,              |
| 5. Neue Transkription,          | 6. Hieroglyphische Schreibung, |
| 7. Wortensprechung/Übersetzung, | 8. Äquivalenzzählung           |
| 9. Sachgebiet,                  | 10. Grammatik,                 |
| 11. Belegstelle,                | 12. Literatur                  |
| 13. Variantennummer             | 14. Datierung/Belegzeitraum,   |
| 15. Kommentar.                  |                                |

1. Die Wortnummer ist der nur einmal vorhandene, eindeutige "Name" einer Entität. Sie stellt als rein numerischer Ersatzcode für den einzelnen Worteintrag eine mathematisch ganze Zahl vom Datentyp *Integer* (Bereich zwischen - 2147483648 bis + 2147483647) dar, die zur Festlegung referentieller Integritätsbedingungen zwischen den verschiedenen Tabellen (Datenlisten) der Datenbank durch unterschiedliche Schlüsselvergabe (*primary und foreign key*) notwendig ist. Da sich alle Informationen in den verschiedenen Tabellen jeweils auf einen ganz konkreten Worteintrag beziehen, muß dieses Datenfeld in den Datenlisten, die einen direkten Bezug zur Wortnummer haben, als Verbindungsglied auftreten.
2. Die numerische Ordnungsangabe ist ein vorläufiges Hilfsmittel zur Ordnung nach der Lesung eines Worteintrags und wird als *numerus currens* vergeben. Der Datentyp entspricht dem der Wortnummer.
3. Die numerische Derivationsangabe, der sog. Owner, verweist auf die eingetragene Wortnummer eines Lexems, von dem das durch eine eigene Wort-

nummer benannte Lexem abgeleitet ist. Dadurch werden die im Grammatikfeld (s. u. 10.) codierten allgemeinen Ableitungstypen für jeden Worteintrag konkretisiert: Als deverbale Nomina werden beispielsweise die Einträge *jʒ Dieb* zu *jʒ fortnehmen* oder *zʒ Schutz* zu *zʒw schützen* gestellt. Hier können wichtige Forschungen der letzten Jahre in die DB einfließen, so die zur Nominalbildung im Ägyptischen, und auf effektive Art erreich- und verarbeitbar gemacht werden. Der Datentyp entspricht dem der Wortnummer.

4. Im Datenfeld "Transkription" erfolgt die Dateneingabe nach den Prinzipien des Wörterbuchs von Erman/Grapow und den Codierungsvorschriften von WINGLYPH. Der Datentyp ist *Varchar*, es handelt sich also um Zeichendaten (ASCII-Daten) mit variabler Länge, die bei der Definition des Datenfeldes allerdings auf 50 Zeichen begrenzt wurde. Die Umsetzung der ASCII-Daten in den Transkriptionszeichensatz von WINGLYPH erfolgt gegenwärtig nur außerhalb der Datenbank.
5. Das optionale Datenfeld "Neue Transkription" ist vorgesehen, um einem - im Rahmen einer Arbeitstagung 1994 noch zu erarbeitenden - moderneren Transkriptionssystem den nötigen Platz einzuräumen. Vorläufig wurde als Datentyp der der Transkription definiert.
6. Die hieroglyphische Schreibung wird durch einen Code, der dem "Leitfaden zur Verschlüsselung hieroglyphischer Texte" folgt, erfaßt.<sup>7</sup> Im multi-tasking-Betrieb (unter Nutzung der Clipboard-Technik) ermöglicht die Verwendung von WINGLYPH eine sofortige Sichtbarmachung der Hieroglyphen am Bildschirm; was die Überprüfung von Schreibungen in einem aktuellen Text mit den bereits in der Datenbank gespeicherten Varianten wesentlich beschleunigt. Alle Schreibungsvarianten tragen einen numerischen Code, der mit der Wortnummer verbunden den Bezug auf einen Worteintrag ermöglicht (s. u. 13.). Der alphanumerische Datencharakter bedingt *Varchar* als Datentyp; es handelt sich also wieder um ASCII-Daten in einem Datenfeldes, das auf 250 Zeichen begrenzt wurde. Die Umsetzung der ASCII-Daten in den hieroglyphischen Zeichensatz von WINGLYPH erfolgt gegenwärtig nur außerhalb der Datenbank.
7. Die Übersetzung wird, soweit vorhanden, dem Wörterbuch oder bei neuen Wörtern aus der Sprache entnommen, in der ein Wort oder Text erstmals publiziert wurde, vor allem, um nicht durch die Übertragung in eine dritte

<sup>7</sup> Für die Codierung der hieroglyphischen Schreibung unter WINGLYPH, das jetzt immerhin über 5000 Zeichen zur Verfügung stellt, kann in den Fenstern der Zeichenliste (Gardiner-Zeichensatz) und zusätzlich einem Handbuch geblättert werden. Als Eingabehilfe wurde auch ein unter WINGLYPH arbeitendes Graphiktablett mit frei definierten Makros entwickelt, das zwar weniger Hieroglyphen, diese aber weit übersichtlicher angeordnet bietet (maximal 1320 auf einer Ebene liegende Zeichen). Vgl. H. VAN DEN BERG, Glyph for Windows, in: Informatique et Egyptologie 8 (1993), 113-120; I. HAFEMANN, Eine Eingabehilfe bei der Hieroglyphencodierung, in: GM 129 (1992), 21-25.

Sprache Bedeutungsnuancen zu verwischen. Der rein alphabetische Datencharakter bedingt *Varchar* als Datentyp; es handelt sich also wieder um ASCII-Daten auf der Basis der Zeichenumsetztabellen 437 und 850 in einem Datenfeld, das auf 250 Zeichen begrenzt wurde. Dadurch können Wortentsprechungen der ägyptologischen Standardsprachen (Englisch, Französisch, Deutsch und Italienisch) problemlos übernommen werden. Es ist geplant, innerhalb der Wortentsprechungen separate Datenfelder für diese Sprachen unter Mithilfe von Muttersprachlern aufzubauen.

8. Der diachrone Charakter der Wortdatenbank führt zur Aufnahme semantischer (aber auch morphologischer und grammatischer; s.u.) Varianten, einschließlich des bildlichen Gebrauchs eines Worteintrages. Bei den Wortentsprechungen werden diese Einträge durch die Äquivalenzzählung erfaßt und mit dem Ausgangseintrag (Wortnummer) verbunden. Sie wird als *numerus currens* vergeben und stellt eine mathematisch ganze Zahl vom Datentyp *Smallint* (Bereich zwischen - 32768 bis + 32767) dar.
9. Für die Zuordnung eines Lexems zu einem Sachgebietes kann aus Begriffen eines verbindlichen Thesaurus gewählt werden. Dem Begriffsthesaurus können im Bedarfsfall anderssprachige Synonymlisten zugefügt werden. Als Verbindungsglied dient ein numerischer Code vom Datentyp der Äquivalenzzählung. Nur der numerische Code, der durch einen *foreign key* mit dem Begriffsthesaurus verbunden ist und daher jederzeit durch den eigentlichen Begriff ersetzt werden kann, wird in der DB gespeichert.

Die Erarbeitung des Begriffsthesaurus folgt der modernen lexikographischen Klassifikation und ist durch Termini der altägyptischen Begriffswelt ergänzt. Hierbei kann auf die Klassifizierung des 6. Bandes von Erman/Grapow (S. 195-221) zurückgegriffen werden.<sup>8</sup>

Da jedes Wort Teil eines Textes ist wird die Sprachgebrauchsebene, die durch einen Textgattungs-Thesaurus klassifiziert und innerhalb des TCV bei der Textbeschreibung angegeben ist, ebenfalls für jedes Wort feststellbar sein.<sup>9</sup> Das gilt allerdings nur für neu eingegebene Texte, da für das ausgedruckte Wörterbuch nicht alle einzelnen im WbZ vorhandenen, sondern nur die dort indizierten Lemma-Belege notiert werden.

Durch die so mögliche Kopplung von Sachgebiet und Sprachgebrauchsebene kann die Beziehung Semantik-Pragmatik erfaßt werden. Problematisch ist die Zuordnung eines Textes zu einer Gattung bei den Fällen, in denen ein Text aus mehreren Teilen unterschiedlicher Gattungen besteht, also z. B. Grabinschriften mit Opferformeln, Hymnen, Listen etc.

<sup>8</sup> Der Sachgebietsthesaurus ist noch nicht abgeschlossen.

<sup>9</sup> Zu den unterschiedlichen Auffassungen über die Anforderungen an einen Textgattungsthesaurus vgl. die Standpunkte von D. VAN DER PLAS und S. GRUNERT auf der letzten Arbeitstagung der Gruppe Informatique et Egyptologie 1993 in Genf (erscheint in: Informatique et Egyptologie 9, 1994).

10. Für die Grammatikangaben bei einem Lexems kann ebenfalls aus Begriffen eines verbindlichen Thesaurus gewählt werden. Dem Begriffsthesaurus können im Bedarfsfall anderssprachige Synonymlisten zugefügt werden. Als Verbindungsglied dient ein numerischer Code vom Datentyp der Äquivalenzzählung. Nur der numerische Code, der durch einen *foreign key* mit dem Begriffsthesaurus verbunden ist und daher jederzeit durch den eigentlichen Begriff ersetzt werden kann, wird in der DB gespeichert. Die Begriffe beziehen sich gegenwärtig ausschließlich auf die Ebene der Wortartbestimmung und beinhalten bedingt auch Genus und Verbkategorie (transitiv/intransitiv). Konkrete Derivationsangaben zu einem speziellen Eintrag finden sich als Nummerncode in einem eigenen Datenfeld (s.o. 3.). Als Ergebnis der Arbeit mit dem Textanalyseprogramm PPTES ist ferner geplant, bei Verben u.a. auch die Verbalklasse und bei anderen Wortarten auch Bildungsart und Ableitungstyp (Substantive als Nisbebildungen, deverbale Substantive von Partizipien, Präpositionaladverbien etc.) zu codieren.

Da die grammatische Bestimmung eines Beleges eindeutig sein muß, werden auch bestimmte semantische Varianten oder einige häufige Transpositionen auf der Ebene des Lexikons definiert, m.a.W. lexikalisiert. Somit werden Verben, von denen transitive Belege neben intransitiven bekannt sind, zweimal aufgenommen, desgleichen werden Eigenschaftsverben, als Verb und Adjektiv aufgenommen oder Zeitsubstantive auch als Adverbien - dies jeweils unter Angabe ihres Ableitungswortes durch den numerischen Derivationscode.<sup>10</sup>

11. Einerseits beinhaltet die Belegstelle bei neu eingegebenen Texten die für ein durch die Wortnummer spezifiziertes Lexem jeweils gültige Textcorpusnummer (TCN).<sup>11</sup> Bei dieser handelt es sich um einen rein numerischen Ersatzcode für jeden einzelnen Text, der eine mathematisch ganze Zahl vom Datentyp *Integer* (s.o. 1.) darstellt. Andererseits werden der TCN in eignen Datenfelder interne Zählungen (Zeilennummer+Wortnummer) hinzugefügt, die vom Datentyp her mit der Äquivalenzzählung identisch sind.
12. Innerhalb der Literatur-Angaben wird bei der Ausarbeitung der BWL im allgemeinen auf die Datenquellen verwiesen, die für die Berliner Wortliste ausgewertet wurden (vgl. Beitrag Reineke). Im Interesse einer sofortigen Überprüfbarkeit durch den Nutzer wurde hier auf die numerische Codie-

<sup>10</sup> Bei der Entscheidung, welche Transpositionen lexikalisiert werden, lassen wir uns von pragmatischen Gesichtspunkten im ägyptischen Sprachgebrauch leiten; hier sind aber jederzeit Einschübe möglich durch die Nutzung der numerischen Ordnungsangabe und der Wortnummer als Owner.

<sup>11</sup> Über die TCN ist der Zugriff zum TCV mit allen interessanten Informationen zu einem Text möglich, z. B. zu einer vollständigen Bibliographie. Ein bibliographisches Nummerncodesystem ist bereits erarbeitet und findet bei dem Aufbau der Literaturdatenbank Verwendung.

rung entsprechend dem vorliegenden Nummerncodesystem (vgl. Anm. 11) verzichtet und auf die gängigen ägyptologischen Abkürzungen zurückgegriffen. Der in aller Regel alphanumerische Datencharakter bedingt *Varchar* als Datentyp; es handelt sich also wieder um ASCII-Daten in einem Datenfeld, das auf 50 Zeichen begrenzt wurde.

13. Die Variantenummer - vom Datentyp her mit der Äquivalenzzählung identisch - dient der internen Unterscheidung zwischen unterschiedlichen Varianten in der hieroglyphischen Schreibung eines Lexems.
14. Die Datierung, also der Belegzeitpunkt bzw. Belegzeitraum der einzelnen spezifischen Schreibung eines Lexems folgt ebenfalls wieder den Prinzipien der Eingabe mittels Thesaurus (s.o. 9.), wobei ein verbindlicher Datierungsthesaurus vorliegt, der auch im TCV verwendet wird. Die dabei vorhandene eins-zu-viele-Beziehung (gleichartige Schreibungen eines Lexems können zu den unterschiedlichsten Zeiten auftreten) wurden durch den n-dimensionalen Charakter der aus insgesamt 4 numerischen Datenfelder (*Smallintr*) bestehenden Datierungscodierung erreicht (WCN-Variantennummer-interne Datierungszählung-Datierungscode).

Da für die Verarbeitung der Lexeme neben semantischen Aspekten auch grammatische und morphologische berücksichtigt werden, ergab sich, wie bereits ausgeführt, zwangsläufig die Aufspaltung eines Worteintrages in zwei oder in mehrere Einträge, wenn beispielsweise neben einer maskulinen auch eine feminine, neben einer intransitiven auch eine transitive Form bekannt ist und sich gegebenenfalls daraus Derivate entwickelten. Ausgangspunkt hierfür war die Überlegung, den Lexemen der BWL jeweils eindeutig Worte aus einzelnen Texten zuordnen zu können. Gerade für diesen Bereich scheint für die Zukunft ein Datenfeld interessant, das lemmaspezifische lexikalische Ergänzungen und ihre morpho-syntaktische Charakterisierung enthält.

Für jede der genannten Ebenen mußte eine eigene Datenstruktur entwickelt werden, d.h. das Datenfeld 'Hieroglyphische Schreibungen' ist anders aufgebaut als das Datenfeld 'Sachgebietsangabe' oder 'Datierung', letztere sind z.B. thesaurusabhängig. Die Ausarbeitung der entsprechenden Thesauri stellte und stellt eine nicht zu unterschätzende Vorarbeit dar, bei der Fachwissen eingebracht und mit praktischen Erwägungen verbunden werden muß. So darf ein Textgattungs-Thesaurus neuen Forschungen in der ägyptologischen Literaturwissenschaft nicht gerade widersprechen, kann aber nicht alle Feinheiten der dort vorgenommenen Klassifizierung bieten, zudem diese auch in der Diskussion sind.

Abhängig von den Fragestellungen kann eine Reorganisation der Daten vorgenommen werden, d.h. der Computer kann bei solcherart Ausnutzung den Pluralismus und das Dynamische der Daten deutlicher machen als jede herkömmliche Art Wörterbuch. Eine echte lexikalische DB kann als ein multifunktionales Objekt charakterisiert werden und ist in gewisser Hinsicht auch

ein multidimensionales Wörterbuch, in dem nach verschiedenen Aspekten eines Wortes gesucht werden kann, wenn man verschiedene Pfade innerhalb der DB verfolgt. Eine DB, die originäre Daten unter verschiedenen Aspekten speichert, hat die Eigenschaft, dieselben Informationen zu multiplizieren. Der Multiplikationsfaktor wird durch die verschiedenen *views* oder Suchstrategien gegeben. Neben der Nutzung einer solchen DB für die unterschiedlichsten Abfragen besteht auch die Möglichkeit daraus Sublexika zu erstellen, die ausgewählte Teile des gesamten Lexikons beinhalten, wie terminologische Fachlexika oder Sammlungen von Ortsnamen, Personennamen u.ä.. In unserem Fall könnte man durch die Derivations- und Owner-Angaben z.B. Vorarbeiten für ein Wurzelwörterbuch oder ein etymologisches Wörterbuch leisten, Untersuchungen zu Bildungstypen von Wortarten unter Berücksichtigung semantischer oder historischer Aspekte, oder zur Verwendung bestimmter Termini oder hieroglyphischer Schreibungen in bestimmten Textgattungen zu bestimmten Zeiten etc. anstellen. Ergebnisse können ausgedruckt, aber auch auf die neuen Speichermedien übertragen und als dynamische Teilwörterbücher verwendet werden. Damit würde z. B. auch die Trennung zwischen den verschiedenen Typen von Wörter- und Spezialwörterbüchern aufgehoben werden.

Selbstredend sind aber für die Bereitstellung einer solchen DB eine Reihe von Vorarbeiten nötig. Jegliche Operationen sind abhängig von den vorher definierten Strukturen. Da es hier vorrangig um lexikographische Arbeiten geht, stehen beispielsweise die syntaktischen Strukturen des Ägyptischen nicht im Vordergrund. So kann z. B. nicht die konkrete Subjekt-Objekt-Beziehung in einem Satz eines konkreten Textes abgefragt werden, oder die hieroglyphische Schreibung bestimmter Flexionsformen in einem Text oder einer Gruppe von Texten. Hierfür müssen spezielle Programme erarbeitet werden, und das für die verschiedenen Sprachstufen des Ägyptischen.<sup>12</sup>

Dieses Projekt folgt konzeptionell den neuen Ansätzen in der Computerlexikographie. Selbstverständlich wird es kaum mehr möglich sein, die Datenfülle einer solchen lexikalischen DB in einem Buch, also einem neuen *Wörterbuch* unterzubringen. Dennoch ist geplant, Teilergebnisse zu publizieren, vorerst in Form einer sich ständig erweiternden Wortliste.

Vielversprechend scheint uns der Ansatz zu sein, diese computerisierte Wortbasis als Lexikon zu benutzen und ihre strukturierten Informationen auf

<sup>12</sup> Vgl. das Textanalyseprogramm THOT, welches die drei Ebenen Transkription - Übersetzung - hieroglyphische Schreibung verknüpft analysiert - dazu J. HALLOF, THOT - Zum Konzept eines Programms zur Analyse ägyptischer Texte, in: Informatique et Egyptologie 8 (1993), 42-46.

Um auch syntaktische Erscheinungen in der oben beschriebenen DB besser zu erfassen, könnte man bspw. bestimmte Flexionsformen, die als Affixe oder Infixe vom Grundwort abzutrennen und separat abzulegen wären, lexikalisieren, oder aber eine Liste der Wortform-Schreibungen anlegen. Die Diskussionen hierüber sind nicht abgeschlossen.

einen laufenden, aktuell aufzunehmenden Text zu spiegeln, der dann vorher nicht indiziert werden muß. Das würde, natürlich immer entsprechend den vorher ausgewählten Kriterien, eine syntaktisch-semantische Lemmatisierung gestatten. Man kann so verschiedene Bedeutungsschattierungen ermitteln und Vergleiche mit anderen Begriffen des gleichen semantischen Feldes anstellen, Kollokationen oder idiomatischen Gebrauch sowie Häufigkeiten etc. feststellen. Andererseits kann man durch Befragung der Textdatei Wörter in ihrem Kontext studieren, neue Entscheidungen in der lexikalischen DB speichern und so für lexikologische Untersuchungen bzw. als Übersetzungshilfen bei der Bearbeitung weiterer Texte konsultieren.<sup>13</sup>

Abschließend muß darauf verwiesen werden, daß das Problem der Verknüpfung von lexikalischer DB und Text-DB durch ein integriertes Textanalyseprogramm noch nicht gelöst ist, da dies die Zusammenarbeit mit einem professionellen Programmierer erfordert. Gegenwärtig konzentrieren sich alle Vorarbeiten auf die Fertigstellung der "Berliner Wortliste" mit den oben beschriebenen verschiedenen Datenfeldern. Die Ausführungen sollten sowohl den Umfang der notwendigen Vorarbeiten verdeutlichen als auch die Notwendigkeit weiterer Diskussionen wie zu Fragen der Thesauri, zur Transkription u. a.m. aufzeigen.

---

<sup>13</sup>Zu diesem Ansatz vgl. H. H. ZIMMERMANN, Multifunctional Dictionaries, *Linguistica Computazionale III* (1983), 279-288, bes. 286. In der Ägyptologie jetzt auch S. J. SEIDL-MAYER, a. a. O., 104, der auf Transkriptionsbasis mit einer Wortliste von standardisierten Primäreinträgen arbeitet, indem er die in den Texten vorkommenden Wortformen diesen Primäreinträgen zuordnet, also die Textwörter standardisiert bzw. lemmatisiert. Den Primärwerten der Wortliste sind morphologisch-syntaktische Eigenschaften zugeordnet.

Zur Berliner Wortliste  
Stefan Grunert

Die Berliner Wortliste versteht sich als ein Hilfswerkzeug bei der Eingabe und Analyse neu aufzunehmender Texte in das Textcorpus am Berliner Wörterbuch der ägyptischen Sprache; sie will und kann keine überarbeitete oder erweiterte Fassung des von Adolf Erman und Hermann Grapow im Auftrag der deutschen Akademien herausgegebenen und in den Jahren von 1926 bis 1963 gedruckten Exemplares sein. Allen, die dessen zunehmende Unzulänglichkeiten beklagen, ist bewußt, daß die von Erman und Grapow geleitete Arbeit von mehr als 50 Ägyptologen bisher nicht ersetzt werden konnte; nach wie vor ist es als Standardwerk in der Ägyptologie unabdingbar.

Unter den neuen technischen Bedingungen ist es möglich geworden, den ursprünglichen Gedanken der Schaffung eines zentralen Gesamtcorpus altägyptischer Texte nach mehr als 40 Jahren wieder aufzunehmen und umzusetzen. In der Vergangenheit erfolgte die Sammlung der Texte nach dem Muster des "Thesaurus Linguae Latinac", und in einem über fast zwei Jahrzehnte dauernden Prozeß wurde - das Material auswertend - auf dieser Basis die Lexik der altägyptischen Sprache erforscht. Durch den Einsatz der Rechentechnik muß vom Prinzip des gewohnten, in zigfacher Ausfertigung vorliegenden Wörterbuchzettels abgewichen werden.

Durch das handschriftliche Abschreiben der Hieroglyphen erfolgte früher die Erfassung der Texte relativ originalgetreu - hieratischen Texten, die man in Hieroglyphenschrift umsetzte, wurde in Zweifelsfällen das originale Schriftbild als Kopie zur späteren Entscheidung beigegeben. Eine solche Form der Textspeicherung ist heute qualitativ besser möglich: Die neu zu erfassenden Texte werden - abhängig von ihrer Publikationsform - in ihrem Erscheinungsbild mittels Scanner erfaßt und in dieser, dem Original adäquateren Form in einer separaten Bild-Datenbank unter ihrer Textcorpusnummer abgespeichert.

Zur eigentlichen Erfassung durchlaufen die Texte einen umfangreichen Analyseprozeß. Bisher wurde dieser Ablauf unter Nutzung von zwei unterschiedlichen sog. fullscreen Programmen (PPTES und THOT) getestet, die aus speziellen, nicht mit dem Wörterbuch verbundenen Aufgabenanforderungen entstanden. Das Analyseprogramm THOT verlangt eine durchgängige programmspezifische alphanumerische Hieroglyphencodierung sowie zusätzlich die Angabe der Transkription. Beide Bereiche müssen sich in der Lexemzahl entsprechen. Eine Übersetzung kann angemerkt werden. Umfangreichere Texte können in einzelne Records, vergleichbar den früheren Wörterbuchzetteln, untergliedert werden. In der grammatischen Textanalyse arbeitet THOT mit einem starren System, daß auf der mittelägyptischen Sprachstufe basiert. Jeder Eintrag muß zwangsläufig einer durch das System vorgegebenen gram-

matischen Erscheinung zugeordnet werden. Bezuglich der Lexeme verfügt THOT über kein Grundwissen. Erst aus dem verarbeiteten Material wird eine Wortliste erstellt. Dadurch bietet es die Möglichkeit, aus einer bestimmten Textgruppe einen Spezialthesaurus zu erstellen.

Das Analyseprogramm PPTES wurde zur lexikographischen, philologisch-phraseologischen sowie prosopographischen Erschließung von Texten entwickelt. Es dient also primär der Erstellung von Wordindices, Wortkonkordanzen sowie Indices von Spezialworteinträgen wie Personennamen, Götternamen, Geographica. Der Textanalyse liegt eine interne Wortliste in Transkription zugrunde, die aus den Einträgen des Wörterbuches von Erman&Grapow entstand. Anhand dieser linearen Wortliste werden alle Wörter eines Textes lemmatisiert, d.h. standardisiert, indem sie einem Lexem der Wortliste zugeordnet und damit identifiziert werden. Der zu analysierende Text wird nur in einer systemspezifischen Transkription erfaßt, auf der auch die Analyse beruht. Durch eine interne Umsetzung kann dabei zwischen einer englischen und zwei deutschen Transkriptionen gewählt werden. Besondere Schreibungen (wie z.B. von Flexionsendungen) werden in einigen Fällen vom Programm selbst erkannt und dem richtigen Lexem zugeordnet. Meist aber erfolgt die Zuordnung der speziellen Wortform zum Lexem interaktiv. Jedoch kann in den Konkordanzen der Wortkontext mit der aktuellen Wortformschreibung ausgedruckt werden. Allen Einträgen der Wortliste sind bestimmte morphologisch-syntaktische Eigenschaften zugeordnet. Hieroglyphische Schreibungen in alphanumerischer Codierung werden z.Zt. nicht aufgenommen, können aber in einem optionalen Feld ergänzt werden.

Bereits in der ersten Phase der Vorbereitungen zur Wiederaufnahme der Wörterbucharbeit in Berlin war davon ausgegangen worden, das bisherige Wissen zur ägyptischen Lexik im weiteren Arbeitsprozeß produktiv in Form eines sog. Computerinternen Wörterbuches zu nutzen. Die Erfahrungen beim Aufbau einer Datenbank unter dem Betriebssystem OS/2 sowie die späteren bei der Textbearbeitung mittels der erwähnten Analyseprogramme führten zu der Entscheidung, eine lineare Liste der altägyptischen Lexeme aufzubauen, die relational mit einer Vielzahl anderer Informationen zum einzelnen Lexem verbunden ist. Die Diskussionen auf der Berliner Wörterbuchkonferenz (3.-5. September 1992), auf der eine deutliche Polarisierung zur Frage der Notwendigkeit einer alphanumerischen Codierung des hieroglyphischen Erscheinungsbildes vorherrschte, führten schließlich zu dem theoretisch angedachten Kompromiß der Berliner Wortliste. Ausgangspunkt hierfür waren drei grundsätzliche Überlegungen: Einerseits sollte der aufwendige und in der Mehrzahl wenig ertragreiche Prozeß einer vollständigen Codierung der hieroglyphischen Schreibung aller neu zu erfassenden Texte umgangen werden, um entsprechend der mehrheitlich geäußerten Forderung möglichst rasch die Texte erfassen und auswerten zu können. Andererseits sollte wie bisher über die

Transkription stets der Zugang zu den einzig originären Informationen, der hieroglyphischen Schreibung, gewährleistet sein - eine Erfahrung, die sich aus der Jahrzehntelangen Praxis in der Arbeit mit dem Wörterbuch-Archiv und der Beantwortung Hunderter von Anfragen ergab. Schließlich blieb ein technisches Problem zu lösen, daß auch in der Archivarbeit bereits eine Rolle gespielt hat. Bei einer Vielzahl von Lexemen kam es zu Änderungen bei der Lesung, jedoch konnte und wurde das Archiv-Material nicht permanent umsortiert. Ähnlich problematisch würde es werden, wenn auch im Bereich eines rechnergestützen Wörterbuchs oder einer Textdatenbank im Computer ausschließlich die Transkription als Primärinformation zu verarbeiten wäre. Änderungen in der Lesung hieroglyphischer Lexeme würden - bei unterschiedlichen Zeitpunkten ihrer Eingabe - zu einer Nicht-Identität gleicher Einträge führen bzw. das Textcorpus müßte permanent aktualisiert werden, was ab einer gewissen Textmenge kaum mehr realisierbar ist.

Da die Entwicklung der ägyptischen Sprache historisch abgeschlossen ist, kann deren ehemaliger Wortschatz insgesamt als eine endliche Menge betrachtet werden. Dabei hat das einzelne Lexem einen historischen Entwicklungsprozeß durchlaufen, kann also in dieser Menge mehrfach in unterschiedlicher Form und mit verschiedenen Inhalten vertreten sein. Daneben gibt es einen weiteren dynamischen Charakterzug dieser toten Sprache, der sich aus der Verteilung der einzelnen Textzeugnisse über drei Jahrtausende ergibt.

Um diese Aspekte in ihrer Gesamtheit zu berücksichtigen, wurde beschlossen, jedes einzelne Lexem durch einen numerischen Code, die **WORT-CORPUS-NUMMER**, zu repräsentieren. So wie über die Textcorpusnummer (TCN) sowohl die bildhafte Information zum Text als auch die zu dessen Publikationen aus den Datenbanken abrufbar sind, können zur WCN die unterschiedlichsten Informationen separat abgespeichert und gleichfalls separat oder aber auch komplex ausgeworfen werden. Die Abruf-Form wird bestimmt durch den Nutzer; da die einzelnen Informationen in ihrer Gesamtheit relational miteinander verknüpft sind, ergeben sich die vielfältigsten Möglichkeiten.

Texte, die unter Nutzung der WCN und einer durch das konkrete Erscheinungsbild spezifizierten Zusatzinformation numerisch codiert wurden, bleiben auch dann in ihrer physischen (ausschließlich numerischen) Existenz unverändert, wenn sich zu einem Lexem einzelne Informationsinhalte wie beispielsweise Transkription, grammatische Zuordnung oder Übersetzung ändern. Änderungen des einzelnen hieroglyphischen Erscheinungsbildes eines Lexems in einem einzelnen Text, der letztlich einzig originären Information, sind bedingt möglich; allerdings kann sich hier die Notwendigkeit ergeben, bei späteren Änderungen den gesamten Textcorpus auf das Erscheinen eines solchen Beleges überprüfen zu müssen, da er theoretisch mehrfach auftreten kann. Es kann jedoch davon ausgegangen werden, daß ein einmal festgestelltes Erscheinungsbild eine relativ hohe Stabilität besitzt; ohnehin sind interpretative Ver-

änderungen wie beispielsweise Emendationen in der Regel textbezogen und werden separat erfaßt.

Für die Eingabe der Texte sowie deren gleichzeitige Analyse existiert bisher keine aufgabenspezifische Software; grundsätzlich sind die dazu notwendigen Arbeitsprozesse aber im Multitasking unter Window-Technik simulierbar. Da allerdings gegenwärtig die Berliner Wortliste ohnehin noch nicht vollständig ist und zudem bei einer derartigen Eingabe eine 100%ige Dateneintegrität nicht gewährleistet wäre, wird vorrangig die Fertigstellung der Berliner Wortliste sowie die Software-Entwicklung betrieben. Nach Fertigstellung beider Komponenten wird die Texteingabe in verschiedenen Arbeitsschritten erfolgen. Die von THOT und PPTES vorgenommene direkte Einbindung aller Daten zur Textbeschreibung in den jeweiligen Text wird am Berliner Wörterbuch in der separaten TCN-Datei getrennt vom Text erfaßt. Der eigentliche Text wird unter seiner TCN in einem interaktiven Prozeß eingegeben. Dabei sind mittels Window-Technik auf dem Bildschirm mehrere Bereiche sichtbar: Im Sinne eines Computer-Thesaurus (d.h. eine Liste feststehender Begriffe, die mit einer Liste von Synonymen verbunden ist) wird die Berliner Wortliste in den dafür notwendigen Bestandteilen (Transkription, Hieroglyphencode, Grammatik, Übersetzung) bei deren optischer Präsenz über sensitive Eingabemasken konsultiert. Der Bearbeiter hat die Möglichkeit, sich anhand der aufbereiteten Vorlage bei der Eingabe für den Bereich der Transkription oder aber den der hieroglyphischen Schreibung zu entscheiden, wobei der Eingabeprozeß bei frühzeitiger Übereinstimmung zwischen einzugebendem Lexem und dem vorhandenen Eintrag der BWL durch dessen Übernahme abgebrochen werden kann. Nach einer derartigen Zuordnung wird das einzelne Lexem intern in den entsprechenden Nummerncode umgesetzt. Gleichzeitig erfolgt im Bereich der Wortformen-Datenbank, einer modifizierten Form der Berliner Wortliste, die Hinzufügung der jeweiligen TCN. Lexeme, die noch nicht oder nicht in der entsprechenden Variante vorhanden sind, müssen vollständig eingetragen werden, um die notwendigen internen Übernahmen in die Wortformen-Datenbank und in die BWL zu ermöglichen. Damit wird deren stete Verbesserung und Aktualität gewährleistet. Vor allem durch den hohen Variationsgrad in der hieroglyphischen Schreibung ergeben sich hier anfangs umfangreichere Codierungsarbeiten, die sich jedoch im Laufe der Zeit zunehmend reduzieren. Gerade dieser Umstand hat das hier vorgelegte Erscheinungsbild der Listen-Einträge unter Aleph entscheidend geprägt.

Zunächst war davon ausgegangen worden, anhand des archivierten Berliner Materials die dort vorhandene gesamte Variationsbreite in den Erscheinungsbildern eines jeden Lexem-Eintrages wiederzugeben (bis WCN 47), was neben der strikten Überprüfung aller Einträge auf deren Tragfähigkeit zu einem auf Dauer nicht vertretbaren Zeitverzug führte. Für den Rest der Einträge wurde ein selektives und interpretatives Verfahren gewählt, bei dem einerseits

die häufigsten, andererseits aber auch die untypischsten Vertreter für die Aufnahme in die BWL ausgesondert wurden. Während bei den Einträgen bis WCN 47 und ab 217 bis 358 grundsätzlich auch die Stellung der Hieroglyphen untereinander notiert wurde, ist diese für die Einträge WCN 48 bis 216 normiert worden, d.h. es wurde lediglich auf das Vorhandensein einzelner Hieroglyphen in einer bestimmten Reihenfolge geachtet, um die jeweilige Schreibung einem konkreten Eintrag zuordnen zu können. Dieses Verfahren hat insoweit seine Berechtigung, da ein großer Teil des Archivmaterials aus hieratischen Textzeugnissen durch die damaligen Bearbeiter unterschiedlich in Hieroglyphen umgesetzt worden war. Von Bedeutung war ferner der Umstand, daß die Wortform das konkrete Erscheinungsbild eines Beleges bestimmt. Während der Numerus bei Substantiva im allgemeinen in seinen differenzierteren Erscheinungsbildern ohne besondere Hervorhebung übernommen wurde, ist beim Genus strikt unterschieden worden, was beim Vorliegen unterschiedlicher Genera zu unterschiedlichen Haupteinträgen in der BWL führte (vgl. WCN 74 und 75). Außer wortbildenden Präfixen werden Affixe und Infixe für die Haupteinträge in der BWL abgetrennt. Die konkreten Erscheinungsbilder hierzu werden in einer separaten Datei erfaßt werden und sind später über einen in Arbeit befindlichen grammatischen Thesaurus ansteuerbar.

Bei allen Einträgen ist die Möglichkeit einer Angabe zur Derivation vorhanden. Ausgehend von Untersuchungen zum Gebrauch des Verbs *3wj* und seiner Komposita (WCN 49-54) wurde beschlossen, bei Vorhandensein von transitivem und intransitivem Gebrauch separate Einträge vorzunehmen und diese auf eine sog. Stammform, die im eigentlichen Sinne nicht belegbar ist, zurückzuführen. Dieses Verfahren soll klarere Angaben zu weiteren Derivationen ermöglichen. Beispielsweise entwickelt sich aus dem intransitiven Gebrauch von *3wj* in Verbindung mit dem Nomen *ib* ("weit sein an Herz") sowohl eine intransitive als auch transitive Verwendung: "froh sein" und "erfreuen" (WCN 53). Aus dem sog. Stammeintrag können sich - ohne Möglichkeiten einer differenzierteren Zuordnung - weitere Derivationen ergeben (vgl. WCN 41: ein Raum, der weit ist bzw. weit gebaut wurde). Für den vorliegenden Ausdruck der Aleph-Einträge werden die Derivations-Angaben aus drucktechnischen Gründen separat gegeben. Insgesamt ergeben sich Schwierigkeiten, alle zur Berliner Wortliste in Form einer Datenbank erfaßten Informationen in ausgedruckter Form präsentieren zu können. Um dies zu verdeutlichen, soll zumindest deren Struktur hier kurz vorgestellt werden:

## Struktur der Berliner Wortliste

1. Wortnummer "238"	(BWLWCN; LI; Data required)	"1" - Worteinträge, die nicht-synonyme, unterschiedliche Entsprechungen besitzen, werden durch Variantenzählung unterschieden. - Die Wortnummer wird als <i>numerus currens</i> vergeben.
2. Numerische Ordnungsangabe "23800"	(SEQUENCE; LJ; Data required)	9. Sachgebiets-Code "7329" - Der Sachgebiets-Code folgt den Prinzipien der Eingabe mittels Thesaurus (im Stadium der Ausarbeitung).
- Die numerische Ordnungsangabe ist als manuelles Hilfsmittel zur Ordnung nach der Transkription der einzelnen Worteinträge zu verstehen.		
3. Numerische Derivationsangabe "228"	(OWNER; LJ)	10. Grammatik-Code "1122" - Der Grammatik-Code folgt den Prinzipien der Eingabe mittels Thesaurus (im Stadium der Ausarbeitung).
- Die numerische Derivationsangabe ist als optionaler Eintrag zu verstehen und verweist auf die unter BWLWCN eingetragene Wortnummer.		
4. Traditionelle Transkription "Ax.ij" (= sh.ij)	(TTRANSCHR; 50 CHAR; Data required)	11. Belegstelle I "1000000" - Die Belegstelle ist in der Anfangsphase identisch mit einem Eintrag aus den Datenquellen (Wörterbuch u.a.); erst im späteren Verlauf werden die Belege mit der realen TCN gekoppelt.
- Die traditionelle Transkription folgt den Prinzipien des Wörterbuchs von Erman/Grapow.		
5. Neue Transkription "(noch zu bestimmen)"	(NTRANSCHR; 50 CHAR)	12. Belegstelle II "Wb 1,18.11" - Referenz-Angabe für Einträge, die nicht aus realen Texten übernommen sind.
- Die neue Transkription folgt Prinzipien, die sich aus noch zu fixierenden Festlegungen am Berliner Wörterbuch ergeben.		
6. Hieroglyphen-Code "H2:x-t;Z4-M2:Z2"	(HIEROCOD; 250 CHAR; Data required)	13. Variantennummer "1" - Die Wortvariantennummer wird als <i>numerus currens</i> vergeben und dient der numerischen Codierung unterschiedlicher Schreibungen eines Wortes
- Der Hieroglyphen-Code folgt den Prinzipien des Leitfadens zur Verschlüsselung hieroglyphischer Texte.		
7. Wortentsprechung "Ackerpflanzen"	(TRANSLAT; 250 CHAR)	14. Datierungs-Code "19" - Der Datierungs-Code folgt den Prinzipien der Eingabe mittels Thesaurus; ein verbindliches neues Codierungssystem zur Datierung liegt vor.
- Die optionale Angabe der Wortentsprechung wird aus der Sprache übernommen, in der der Text publiziert wurde; es ist beabsichtigt, mehrere gleichartige Datenfelder für unterschiedliche Zielsprachen (Englisch, Französisch, Italienisch, Arabisch) aufzubauen.		
8. Äquivalenzzählung (AEQVN; SI; Data required)		15. Kommentarfeld "(BWLCOMM; 2500 CHAR) "keine konkrete Nutzpflanze sondern allgemeiner Sammelbegriff" - Das Kommentarfeld ist inhaltlich offen gestaltet.

