

DER COMPUTER ALS HILFSMITTEL FÜR DIE
LEXIKALISCHE UND GRAMMATISCHE BESCHREIBUNG
DES ALTÄGYPTISCHEN.
MÖGLICHKEITEN UND GRENZEN

VON WOLFGANG SCHENKEL, DARMSTADT

1 Elektronische Datenverarbeitung in Philologie und Linguistik

1.1 Elektronische Datenverarbeitungsanlagen (Computer) sind in den letzten Jahren mehr und mehr in Bereichen eingesetzt worden, die man den Geisteswissenschaften zuzuteilen pflegt. Gerade auf dem Gebiet der Linguistik, Philologie und Literaturwissenschaft haben solche Apparate Wunder gewirkt. Erinnerung sei an die Impulse, die von der bis heute nicht befriedigend gelösten automatischen Sprachübersetzung auf die linguistische Theorienbildung ausgingen. Die Durchdringung des gewaltigen uns vorliegenden Sprachmaterials wurde einmal durch den Zwang zu logisch einwandfreien Beschreibungen des Faktischen gefördert, andererseits dadurch, daß erst die elektronische Datenverarbeitungsanlage die systematische Erprobung einer Theorie in großem Umfang ermöglicht. Für die Gebiete der Philologie und der Literaturwissenschaft sei auf die Konkordanzen verwiesen, die erst durch den Einsatz von Datenverarbeitungsanlagen den hohen Grad, die Geschwindigkeit und die Wirtschaftlichkeit der Erschließung eines Textkorpus erreichen, die man gegenwärtig anstrebt¹.

1.2 Es läßt sich heute nicht allzu genau abschätzen, welche Möglichkeiten in den elektronischen Datenverarbeitungsanlagen noch stecken. Jedenfalls beginnt man jetzt überhaupt erst richtig, das Gebiet auf seine „nichtnumerischen“ Möglichkeiten systematisch zu erproben². Man hat indes schon so viel Erfahrung gesammelt, daß es müßig wäre, die Anwendbarkeit von Verfahren der elektronischen Datenverarbeitung für ein spezielles Anwendungsgebiet wie die Erschließung altägyptischen Sprachmaterials nachzuweisen.

¹ Vgl. z. B. die in Zusammenarbeit mit dem Deutschen Rechenzentrum hergestellte Shakespeare-Konkordanz: M. SPEVACK, *A Complete and Systematic Concordance to the Works of Shakespeare*, Band 1, Hildesheim 1968.

² S. Computers and the Humanities, herausgegeben vom Queens College of the City University of New York, New York 1966 ff.; darin die Betrachtungen von L. T. MILIC, „*The Next Step*“, I (1966), 3-6.

Die Frage kann nur lauten: wie setzt man eine elektronische Datenverarbeitungsanlage sinnvoll ein; oder: welche Komplexe lohnen sich für eine solche Bearbeitung.

2 Methoden und Verfahren der maschinellen Texterschließung

Die lexikalische und grammatische Erschließung altägyptischer Texte mit Hilfe einer elektronischen Datenverarbeitungsanlage ist unter zwei Aspekten zu betrachten: dem Aspekt der Methode und dem Aspekt des Verfahrens.

2.1 Die Methoden sind keine anderen als die philologischen und linguistischen, Methoden also, deren sich der Ägyptologe auch in konventionellem Arbeitsverfahren bedienen müßte³. Für die elektronische Datenverarbeitung müssen die Methoden lediglich solcherart sein, daß sie eine formal-logische Darstellung des jeweiligen Problemkomplexes in seiner Gänze gestatten; in methodisch nicht lösbaren Grenzfällen sind darüber hinaus pragmatische Lösungen statthaft. Die Auswahl der Methoden und die Konzeption ihrer Anwendung bleibt hier wie dort der Intuition des Forschers überlassen.

2.2 Das Verfahren besteht in einem für eine elektronische Datenverarbeitungsanlage geschriebenen Programmsystem. Die Hauptstufen der Verarbeitung sind die Eingabe der Daten (hier: der altägyptischen Texte) in der festgelegten Form, die eigentliche Verarbeitung der Daten und die Ausgabe der Ergebnisse.

- (1) Als Eingabe für eine Textverarbeitung werden philologisch einwandfrei fixierte Texte verlangt. Das Verfahren der elektronischen Datenverarbeitung kann methodische und andere Mängel, die sich der Philologe bei der Textaufnahme zuschulden kommen läßt, nur in begrenztem, jeweils in den Programmen vorgeplantem Umfang beseitigen. Darüber hinaus kann es den Philologen durch augenfällige Mängel in den Ergebnissen zu gründlichen Überlegungen auffordern. Die gediegene Arbeit eines guten Philologen mit allen seinen geistigen Fähigkeiten, seinem Assoziationsvermögen und seiner umfassenden Erfahrung, ist hier noch weniger entbehrlich als bei konventionellem Verfahren.
- (2) Die Verarbeitung der Daten ist der Kern des maschinellen Verfahrens. Sie umfaßt organisatorische Maßnahmen zur Steuerung des Datenflusses und, was den Ägyptologen unmittelbar interessiert, den Prozeß der Merkmalfindung. Merkmalfindung heißt in unserem Fall: Erkennen und

³ Die elektronische Datenverarbeitung hat selbstverständlich ihre eigenen Methoden; diese stehen hier, wo es nur um die ägyptologische Anwendung der elektronischen Datenverarbeitung geht, nicht in Rede.

Fixieren von lexikalischen und grammatischen Merkmalen, die in den Eingabedaten nicht explizit angegeben sind. Die stupide Arbeit der Fixierung immer wiederkehrender Merkmale und der Umgruppierung der Daten gemäß diesen und den bereits bei der Dateneingabe explizit angegebenen Merkmalen kann eine elektronische Datenverarbeitungsanlage vorzüglich ausführen.

Der Apparat kann allerdings nicht mit Hilfe eines Knopfdrucks etwa für eine lexikalische oder grammatische Analyse altägyptischer Texte in Gang gesetzt werden, sondern muß für die Erledigung dieser Aufgabe mit Hilfe maschinenverständlicher Anweisungen, eines Programms, gesteuert werden⁴. Da ein solches Programm jeden Analyseschritt explizit vorschreiben muß, ist seine Herstellung oft ein langwieriges Unternehmen. Das Programmieren kann aber ungemein lehrreich sein, weil die durch das Programm gelieferten Ergebnisse den Ägyptologen gegebenenfalls von der Irrigkeit seiner Formalisierungsbemühungen überzeugen können. Insofern liegt in der Programmierung einer elektronischen Datenverarbeitungsanlage und in der Konversation mit ihr beim Testen des Programms eine Quelle der Erkenntnis, eine Anregung zu neuen Lösungsversuchen.

- (3) Die Ergebnisse der Verarbeitung müssen in lesbarer und übersichtlicher Form ausgedruckt werden. Aus Gründen der Wirtschaftlichkeit wird der Ägyptologe dabei einige ihm ungewohnte, aber leicht zu memorierende Besonderheiten eines Maschinenausdrucks unter Umständen in Kauf nehmen müssen, insbesondere den, daß die Schnelldrucker nicht über die konventionellen Transkriptionszeichen verfügen. An sich ließe sich technisch ohne weiteres Abhilfe schaffen, doch ist die Entwicklung von Ausgabegeräten noch derart im Fluß, daß sich besondere finanzielle Anstrengungen von ägyptologischer Seite vorerst nicht lohnen.

3 Illustration eines Verfahrens der maschinellen Texterschließung

Ein Abschnitt des im Deutschen Rechenzentrum entwickelten Systems M.A.A.T. (Maschinelle Analyse Altägyptischer Texte) soll die Möglichkeiten der elektronischen Datenverarbeitung veranschaulichen. Mit Hilfe der bereits programmierten Systemteile kann man auf der Grundlage von philologisch aufbereiteten Texten ein detailliertes lexikalisch-grammatisches Verzeichnis über die Wörter der Texte produzieren, das Eigenheiten einer Konkordanz mit denen eines Wörterbuchs verbindet.

3.1 Abb. 1 zeigt einen Ausschnitt aus einer systemgerechten Textauf-

⁴ Die Programme des unten vorgeführten Verfahrens sind in den Programmiersprachen FORTRAN II und FAP für die Datenverarbeitungsanlage IBM 7090 programmiert.

SPALTE	10	20	30	40	50	60	70	80
0	003X101V,184-5,A,M8C							BAANA010
R'	I	NA	I		I			BAANA020
/-	I	NI	CI		ID21L/,Z1/			BAANA030
N.I	I	INI	I		I			BAANA040
PR.T	I	PR	I		ID54			BAANA050
M	I	NA 2	I		I			BAANA060
HRW.W	I	PR	I		IN5L,Z1/			BAANA070
M	I	NI 2	I	NA	I			BAANA080
CR.T	-	NA	I		I			BAANA090
NYR	I	NA	I		I			BAANA100
0	006X101V,184-5,B,M8C				I			BAANA110
XPR	I	SK	I		I			BAANA120
MD.T	I	NAW	I		I			BAANA130
NNK	I	PE	I		I			BAANA140
I+TH.W	I	NA	ID		I			BAANA150
0	015X101V,186-7,A,M8C				I			BAANA160
MNN	=	SKG	I		I			BAANA170
...	I	PE	I		I			BAANA180
MI.KW	I	PSI	I		I			BAANA190
...	I		I		I			

Abb. 1 Korrekturausdruck einer Textaufnahme

nahme des Spruches 335a der altägyptischen Sargtexte (CT IV 184–191, Sarg M8C)⁵. Zur leichteren Orientierung ist der hieroglyphische Text hinzugefügt; die zwischengeschobene Zählung ist eine in der Textaufnahme in den mit „0“ beginnenden Zeilen stehende künstliche Abschnittsnumerierung.

Jede Wortform beginnt in einer neuen Zeile. Zur Wortform gehören folgende Elemente der Beschreibung:

- (1) Spalte 1 bis 12: Wortform in genommter Transkription⁶ (zum Beispiel „MD.T“), darunter gegebenenfalls eine zeichenweise den Transkriptionszeichen zugeordnete Textkritik (zum Beispiel „/“ unter dem zerstörten „T“ von „MD.T“).
- (2) Spalte 25 bis 38: morphologisch-lexikalische Indizes, durch die die Wortform meist eindeutig in ihrer lexikalischen und grammatischen Zugehörigkeit identifiziert werden kann. So steht etwa in Spalte 25 bis 26 die Angabe der Wortart (zum Beispiel „NA“ für Nomen Appellativum); in Spalte 27 der Wurzeltyp (zum Beispiel „I“ für ult. *ǰ*); in Spalte 28 bis 29 ein Index, der homokonsonantische Wurzeln unterscheidet (zum Beispiel *hrw*₁ „zufrieden sein“ und *hrw*₂ zu *hrw.w* „Tag“)⁷; in Spalte 33 ein Namenindex (zum Beispiel „D“ für „Deus/Dea“); in Spalte 35 bis 36 eine Wortartangabe für Zusammensetzungen (zum Beispiel „NA“ bei „CR.T-NYR“, *hr.t-ntr*).
- (3) Spalte 48: Interpunktion (zum Beispiel „.“ für Satzende).
- (4) Spalte 49 bis 71: Notation von Graphemen⁸. Kern der Notation sind die Nummern der Zeichen in Hieroglyphenlisten, hier der Gardiner-Liste (zum Beispiel „A2“ für den sitzenden Mann mit der Hand am Mund).
- (5) Spalte 73 bis 80: Kennung der Lochkarte, ohne ägyptologisches Interesse.

Die lexikalisch-grammatische Identifikation durch Indizes ist stark redundant, was eine bequeme, schematische Notation erlaubt und das Analyseverfahren flexibel hält.

3.2 Abb. 2 zeigt ein Verarbeitungsprotokoll, wie es während der Hauptstufe der maschinellen Verarbeitung produziert wird. Der Maschinenaus-

⁵ Abgebildet ist nicht die originale Textaufnahme, sondern ein etwas übersichtlicherer Korrekturausdruck, in dem die Textkritik aus Spalte 13 bis 24 unter die Wortform in Spalte 1 bis 12 gesetzt ist. Die durchgehenden Kolumnen aus „I“s sind lediglich Orientierungshilfen.

⁶ Vgl. vorläufig R. GUNDLACH und W. SCHENKEL, „M.A.A.T., Ein System zur lexikalischen und grammatischen Erschließung altägyptischer Texte mit Hilfe einer elektronischen Datenverarbeitungsanlage“, *Chronique d'Égypte* XLII, 83 (1967), 41–64.

⁷ Der Index „0“ oder „1“ braucht nicht geschrieben zu werden.

⁸ In der Probe sind weder Phonogramme noch rein kalligraphische Zeichen notiert.

TXT01C CT,SP.335A

B0335A

MFT06F

IN DER FLEXIONSTABELLE
FEHLT DER STAMM
MIT DEM WURZELINDEX.
ODER VOR EINEM SUFFIX FEHLT

FPE
I
1

SYS21C

WEGEN DER VORSTEHENDEN BEANSTANDUNGEN)
MURDE DER FOLGENDE SATZ VON DER AUSWERTUNG ZURUECKGESTELLT

CT,SP.335A

B0335A

\$\$\$\$\$

0 132X10IV,202-3,C,M8C

0 135X10IV,202-3,D,M8C

IM		SA
RJ1.N	=	SKI
I		PE
\$W.TI	--	NA
I		PE
M		PR
TP	=	NA
I		PE
AAAAAA		

D37L
ALL
2H6L
ALL
DIL,Z1
.ALL

B0335A
B0335A
B0335A
B0335A
B0335A
B0335A
B0335A
B0335A

TXT02C CT,SP.335A
BEEDET

B0335A

Abb. 2 Analyseprotokoll

druck spricht für sich selbst; nur so viel sei zur Erläuterung hinzugefügt: Während des Analyseprozesses können formale Fehler der Textaufnahme erkannt werden. Das System gibt auch diese Fehler im Verarbeitungsprotokoll bekannt. Fehlerhafte Passagen werden dann in einer Weise von der Bearbeitung zurückgestellt, daß sie nach Korrektur der angezeigten Fehler wieder in den Analyseprozeß geleitet werden können. Zur Zeit können im Verarbeitungsprotokoll ca. 40 Feststellungen verschiedenen Typs auftreten. Der Diagnostizier- und Nachtragsmechanismus ist ein Kernstück des Analyseverfahrens, weil sich formale Fehler erfahrungsgemäß bei der Textaufnahme immer einschleichen⁹.

3.3 Abb. 3 zeigt einen Ausschnitt aus einer lexikalischen Liste zum Spruch 335a der altägyptischen Sargtexte, Sarg M8C (Liste LL/CT.335A. M8C-0). Zur Erläuterung folgende Hinweise:

- (1) Wurzeln stehen mit ihren Indizes in Stern-Kästchen. Den Wurzeln sind Angaben über den Wurzeltyp beigelegt.
- (2) Wörter mit ihren Indizes und Wortartangaben stehen in Strich-Kästchen. Den Wörtern sind aus einem maschinenverfügbaren Glossar Übersetzungen und Verweise auf Sekundärliteratur oder Textstellen beigelegt, soweit solche im Glossar zum Zeitpunkt der Produktion enthalten waren.
- (3) Die Belege zu einem Wort sind vor allem nach ihrer grammatischen (morphologischen) Form durchgeordnet.
- (4) Als Beleg ist jeweils der Belegsatz mit parallel darunter laufender Textkritik gegeben und die Referenz der Textstelle. Die zu belegenden Wortform ist durch Sperrdruck hervorgehoben. Grapheme sind der besseren Lesbarkeit wegen im wesentlichen nur bei der zu belegenden Wortform wiedergegeben.
- (5) Zusammengesetzte Wörter werden nur einmal in die Liste vollständig aufgenommen. Zusätzlich sind die Einzelelemente ausnahmslos aufgeführt und durch Verweise mit der Stelle verknüpft, an der die komplette Information steht.
- (6) Die Anzahl der Belege ist notiert.

4 Möglichkeiten und Grenzen

Die Überlegungen des Abschnitts 2 und die Skizze des im System M.A.A.T. praktizierten Verfahrens in Abschnitt 3 gestatten uns, einige Merkmale der elektronischen Datenverarbeitung zu bestimmen, die für das Anwendungsgebiet symptomatisch sind:

⁹ Formal richtige, aber sachlich falsche Aufnahmen werden vom System dagegen ohne Vorbehalt akzeptiert.

JJJ) 1 - JR 1	78	LL/CT.335A.M8C-0
-----+ I JJJ.T 1 A, NA I -----+	BEHOERDE, KOLLEGIUM WB 5/528, 1-529,20	BELEGE 2 BELEGE 2
SG.		
1. (REDE) (ZEILE IV,252-3,C.M8C) I+NJ.W VR=YN -- (ZEILE IV,254-5,A.M8C) NB.W M) (.T J) J) . T (GR..N33,AA8,A1,3N33) V).T WSIR (ZEILE IV,254-5,B.M8C) JJ.IW \$(.WT DS.W M ISF.TIW (ZEILE IV,256-7,A.M8C) IM.IW-XT VTP=S-XWI=S. CT,SP.335A, IV,254-5,A.M8C		
2. (ZEILE IV,266-7,A.M8C) IR VR.I N.I J) J) . T (GR..AA8,N33,3Z1) TN, (ZEILE IV,266-7,B.M8C) }IR-WR RN=F. + CT,SP.335A, IV,226-7,A.M8C		
***** * J W I 1 * *****	ULT. I	BELEGE 1
-----+ I JW.T 1 A, NA I -----+	SCHLECHTES, BOESES WB 5/547,11-548,17	BELEGE 1 BELEGE 1
SG.		
1. (REDE) (ZEILE IV,256-7,C.M8C) DR=YN J W . T (GR..G37/) IR.T=I (ZEILE IV,258-9,A.M8C) MI NW IRI.N=YN N IX 8 IPW I.N.IW-\$MS.W N.IW NB SP).WT (ZEILE IV,258-9,B.M8C) IRI.N INP.W S.WT=SN (ZEILE IV,260-1,A.M8C) HRW.W PF N.I MI-IR=K-IM=I. CT,SP.335A, IV,256-7,C.M8C		
***** * J B (1 * *****		BELEGE 1
-----+ I JBI 1 A, NA I -----+	FINGER WB 5/562,11-565,9 WMT1000-1003	BELEGE 1 BELEGE 1
PL.		
1. (ZEILE IV,236-7,C.M8C) IN GR.T JVWTI IRI NN (ZEILE IV,238-9,A.M8C) M J B (. W (GR..3D50L) = F. CT,SP.335A, IV,238-9,A.M8C		
***** * J R 1 * *****		
-----+ I JR 1 A, NA I -----+	ENDE, GRENZE, BEREICH WB 5/585,11-591,10	
SG.		
X. SIEHE NB 2, NB 1 A, NB-R-JR		

Abb. 3 Probe aus einer lexikalisch-grammatischen Liste

- (1) Die elektronische Datenverarbeitungsanlage ist kein Roboter; sie ist nicht in der Lage, auf Knopfdruck intelligente Arbeit zu verrichten. Sie verarbeitet lediglich formal-logisch einwandfreie Daten (hier: Texte) nach bestimmten, in Programmen festgelegten Algorithmen (hier zum Beispiel: grammatischen Regeln). Sie ist lediglich Partner des Ägyptologen, dem sie bestimmte Routinearbeitsgänge abnehmen kann. In jeder Etappe eines maschinellen Verfahrens bestimmt der Ägyptologe, was die Maschine tun soll. Daneben bleiben die philologische Fixierung des zu verarbeitenden Textes und die Auswertung der Ergebnisse des maschinellen Verfahrens ihm allein überlassen.
- (2) Sowohl die Entwicklung eines Verfahrens für die elektronische Datenverarbeitung als auch die philologische Textaufnahme sind Tätigkeiten, die nur der Mensch, der Ägyptologe und Datenverarbeitungsfachmann, vollbringen kann. Wollte man ohne präzise Vorstellungen „mal eben etwas durch die Maschine jagen“, so wäre sowohl die ägyptologische als auch die Programmier-Investition zu hoch im Vergleich zu einem bei übereilter Produktion erreichbaren Ergebnis. Es verbietet sich zumal, konventionelle Problemstellungen unbesehen in eine Bearbeitung mit der elektronischen Datenverarbeitungsanlage zu überführen. Im allgemeinen führt eine gründliche Problemanalyse im Hinblick auf die Realisierung mit einer elektronischen Datenverarbeitungsanlage zu einer Präzisierung der Problemstellung selbst.
- (3) Der Einsatz der elektronischen Datenverarbeitung lohnt sich für methodisch oder ihrem Umfang nach anspruchsvolle Texterschließung (Methoden- und Modellerprobung; Erfassung großer Datenmengen). Optimalen Nutzen erreicht man durch eine integrierte Texterschließung, d. h. wenn eine Vielzahl verschiedener Gesichtspunkte an ein und dasselbe Textkorpus herangetragen wird. So sollen zum Beispiel Textaufnahmen im System M.A.A.T. lexikalisch, morphologisch, syntaktisch, graphematisch und nach anderen Gesichtspunkten auswertbar sein.
- (4) Der Mensch ist Initiator und Mittelpunkt im Verfahren der elektronischen Datenverarbeitung; er allein trifft geistige Entscheidungen. Die elektronische Datenverarbeitung ist allerdings insofern inhuman, als sie formal-logisch einwandfreie Konzeptionen verlangt, und gerade diese Forderung von Disziplinen, die mehr sach- als methodenorientiert sind, naturgemäß eine stärkere Betonung solcher Gesichtspunkte verlangt.
- (5) Gewisse Verlegenheitslösungen im Druckbild der Ergebnisse sind als Scheingrenze einzustufen. Man kann hier ein leichtes Nachgeben des Ägyptologen erwarten, da nicht Geistiges, sondern lediglich gewisse Gewohnheiten betroffen sind.