

# QUEST 2 – Computergestützte Philologie und Exegese

*Wolf-Dieter Syring (Greifswald)*

Nach den in mühevoller, oft jahrzehntelanger Detailarbeit und Niederschrift erstellten Konkordanzen von Mandelkern, Lisowsky und Even-Shoshan<sup>1</sup> sind bereits einige Konkordanzen zum hebräischen Alten Testament erschienen, die die schnellen Suchmöglichkeiten moderner Computer nutzen, um semantische und syntaktische Erscheinungen aus den über 400000 Wörtern des AT herauszufiltern.<sup>2</sup> Dabei bieten die „Rechenknechte“ insbesondere bei komplexen und wurzelunabhängigen Suchvorgängen einige Möglichkeiten, die bei der Verwendung gedruckter Ausgaben nicht oder nur unter unvertretbarem Mehraufwand zugänglich sind.

Eine komfortable und effektive Arbeit am Text des Alten Testaments setzt eine breite und zuverlässige Datenbasis voraus und benötigt ein einfach zu beherrschendes Programm mit einer unkomplizierten Übernahme von Ergebnissen in die vom Benutzer verwendeten Textprogramme. Inhalt dieses Artikels ist eine kurzgehaltene Vorstellung des Programms Quest, das in seiner zweiten Version voraussichtlich 1999 im wissenschaftlichen Verlagsprogramm der Deutschen Bibelgesellschaft erscheinen wird.

## *1. Die Datenbasis*

Die Güte einer Textdatenbank wird bestimmt von Umfang und Qualität der enthaltenen Texte. Quest 2 enthält neben den Ausgaben der BHS und des NT Graece auch die Texte der Septuaginta sowie der Vulgata.<sup>3</sup> Für die Septuaginta stehen die am CCAT<sup>4</sup> erarbeiteten morphologischen Daten zur Verfügung, für die BHS die unter d) kurz beschriebene WIT-Datenbank.<sup>5</sup>

<sup>1</sup> S. Mandelkern, *Veteris Testamenti Concordantiae Hebraicae atque Chaldaicae* (Graz 1955); G. Lisowsky, *Konkordanz zum Hebräischen Alten Testament* (Stuttgart 1958); A. Even-Shoshan, *A New Concordance of the Bible* (Jerusalem 1989).

<sup>2</sup> Quest 1.0 (Haarlem 1992); Bible Windows 5.1 (Cedar Hill, Texas, 1997); Bible Works 3.5 (Seattle, Washington, 1996); Accordance 2.1 (Vancouver 1997). Die aufgeführten Programme haben ähnliche Zielvorstellungen, wenn auch die Anzahl der beigegebenen Texte, deren Qualität und Kodierungstiefe z.T. erhebliche Unterschiede aufweisen.

<sup>3</sup> *Biblia Hebraica Stuttgartensia* (Stuttgart 1990<sup>4</sup>); *Septuaginta* (Stuttgart 1935); *Biblia Sacra iuxta Vulgatam Versionem* (Stuttgart 1983<sup>3</sup>); *Novum Testamentum Graece* (Stuttgart 1993<sup>27</sup>). Eine Reihe neuerer Übersetzungen stehen ebenfalls zur Verfügung, z.B. in deutscher Sprache: Luther-Übersetzung (Revision 1984), Einheitsübersetzung und Elberfelder Übersetzung (Revision 1993).

<sup>4</sup> Center for the Computer Analysis of Texts, University of Pennsylvania, Philadelphia, Pennsylvania. Die morphologische Analyse der Septuaginta wurde unter der Leitung von R. Kraft erstellt.

<sup>5</sup> Werkgroep Informatica Text. Die „Werkgroep Informatica“ wurde 1977 gegründet und hat sich – unter der Leitung von E. Talstra – nach der Erstellung und Korrektur eines maschinenlesbaren, morphologisch analysierten Textes des AT mit den theoretischen Voraussetzungen einer



a) Text: Nachdem der Herausgeberkreis der Biblia Hebraica editio quinta die Entscheidung getroffen hat, den Codex Leningradensis weiterhin als Grundlage zu verwenden,<sup>6</sup> ist es sinnvoll, bei der Herausgabe einer Computerkonkordanz zum AT ebenso zu verfahren. Der Michigan-Claremont-Text, eine maschinenlesbare Ausgabe der BHS, der als Referenztext für alle weiteren maschinenlesbaren Ausgaben diente, wurde für die Arbeit der „Wergroep Informatica“ an der theologischen Fakultät der Freien Universität Amsterdam überprüft und – in Zusammenarbeit mit der Deutschen Bibelgesellschaft – von Druck- bzw. Datenfehlern weitestgehend befreit. Der von Quest verwendete Text enthält somit den vollständigen Konsonanten- und Vokal-Text sowie die Akzentzeichen der BHS.

b) Apparat: Ketib/Qere sowie der Apparat der BHQ (und des NT Graece) sind in Quest enthalten und von jeder betroffenen Textstelle aus zugänglich. Die Schriftzeichen, die nicht dem hebräischen, griechischen oder lateinischen Alphabet entstammen, werden in Umschrift wiedergegeben.

c) Zeichenkodierung: Die bei bisherigen Computersystemen vorhandene Beschränkung auf ca. 220 Zeichen je Schriftart hat dazu geführt, daß verschiedene Abbildungen des hebräischen Alphabets und der Vokale auf den ASCII-Zeichensatz verwendet werden.<sup>7</sup> Die mit Unicode 2.0 vorhandene Standardkodierung der hebräischen Zeichen ist im Blick auf den benötigten hebräischen Zeichenvorrat vollständig und wird von Quest verwendet.<sup>8</sup> Die zur Eingabe hebräischer Zeichen verwendete Zuordnung zur Tastatur kann den Gewohnheiten des Benutzers angepaßt werden. Ein dem Druckbild der BHS entsprechender Zeichensatz gehört zum Lieferumfang von Quest.

d) Grammatische Kodierung: Die Entwicklung einer grammatischen Datenbank stellt einen Neuanatz der Forschung innerhalb der Althebraistik dar.<sup>9</sup> Der in der

---

grammatischen Textdatenbank zum AT befaßt sowie mit der Entwicklung linguistischer Programme zum Aufbau dieser Datenbank, vgl. E. Talstra und F. Postma, *On Texts and Tools. A short History of the Wergroep Informatica*, in: *Computer Assisted Analysis of Biblical Texts* (Amsterdam 1989) 9-28.

<sup>6</sup> Vgl. A. Schenker, Eine Neuausgabe der Biblia Hebraica, ZAH 9 (1996) 58-61. Die erste Lieferung der BHQ (Megilloth) mit ihrem neu erarbeiteten und gestalteten Apparat wird mit Quest 2 ausgeliefert.

<sup>7</sup> So wird z.B. א mit „a“ oder „>“ kodiert, ו mit „w“ oder „\$“, die Vokale stehen vor oder nach den zugehörigen Konsonanten (אָ = „egr“ oder „ger“), ebenso die Akzentzeichen; mit Dagesch oder Mappiq versehene Konsonanten werden entweder als Großbuchstaben kodiert (א = „b“; אַ = „B“) oder als Folge von Konsonant und Punkt (אָ = „b.“ oder „.b“).

<sup>8</sup> Die Bemühungen des Unicode-Konsortiums, die Schriftzeichen aller Sprachen auf ca. 65000 Zeichen abzubilden und ihre Verwendung zu standardisieren, haben 1996 einen vorläufigen Abschluß gefunden: *The Unicode Standard, Version 2.0* (Reading 1996), bes. 6.20-6.21; 7.60-7.62. Da die verbreiteten Textprogramme bisher keine Unicodezeichen verarbeiten können, werden bei der Übernahme hebräischer und griechischer Texte die benötigten Konvertierungen automatisch vorgenommen.

<sup>9</sup> Vgl. C. Hardmeier, Elektronische Datenverarbeitung als Instrument der Analyse von hebräischen Texten des Alten Testaments, in: *Actes du premier Colloque International Bible et Informatique: Le Texte* (Paris/Genève 1985) 203-212; E. Talstra, *An hierarchically structured Database of Biblical Hebrew Texts. The relationship of grammar and encoding*, ebd. 335-349; E. Talstra, *The production of a syntactically orientated concordance of Biblical Hebrew Texts*, in: *Actes du*



„Werkgroep Informatica“ beschrittene Weg einer von der Wortebene ausgehenden, aufsteigenden Kodierung und die Unterscheidung distributionaler und funktionaler Kategorien gewährleisten ein hohes Maß an Transparenz und Überprüfbarkeit.<sup>10</sup>

Die von den graphischen Einheiten des Textes ausgehende morphologische Analyse liefert als Grundeinheiten der weiteren Arbeit die Wörter des AT sowie deren Eigenschaften.<sup>11</sup> In einem iterativen und rekursiven Strukturierungsprozess (parsing process)<sup>12</sup> werden zunächst Wortfügungen (phrases) definiert und durch fortschreitende Differenzierung und Ergänzung präzisiert. Dieser Vorgang wird anschließend auf Sätze (clauses) und Satzfügungen (clause connections) angewendet, im Blick auf eine funktionale Textgrammatik des hebräischen AT.<sup>13</sup> Der Strukturierungsprozess arbeitet mit rein deskriptiven Kategorien, die aus den morpho-syntaktischen Merkmalen der jeweiligen Untereinheiten gewonnen werden, und orientiert sich an reproduzierbaren und damit nachprüfbareren Vergleichen mit Mustersammlungen (pattern matching), die sukzessiv – unter Einbeziehung aller Texte des AT – vervollständigt werden. Dabei wird auf ein vorgegebenes grammatisches Regelwerk verzichtet, das stets Einzelentscheidungen in den verbleibenden Ausnahmefällen erfordert.

Neben diesem an Aufbau und Verteilung grammatischer Phänomene orientierten Weg zur deskriptiven Erfassung „distributionaler“ Daten wird ein zweiter Ansatz verfolgt, der die Relationen resp. Funktionen grammatischer Einheiten berücksichtigt.<sup>14</sup> Als Beispiele seien die Bestimmung der Satzteile (clause constituents), der Relationen zwischen Phrasen (z.B. attributiv) und zwischen Sätzen (z.B. parallel) genannt.

Die in Quest 2 verwendete Textdatenbank bietet distributionale und funktionale Daten zur Wort-, Wortfügungs- und Satzebene, in Teilbereichen auch zur Satzfü-

deuxieme Colloque International Bible et Informatique: Méthodes, outils, résultats (Paris/Genève 1989) 563-580; C. Hardmeier und E. Talstra, Sprachgestalt und Sinngehalt. Wege zu neuen Instrumenten der computergestützten Textwahrnehmung, ZAW 101 (1989) 408-428.

<sup>10</sup> Folgende Ebenen werden verwendet: Wortebene, Wortfügungsebene, Satzebene, Satzfügungsebene. Vgl. E. Talstra, Phrases, Clauses and Clause connections in the Hebrew database of the Werkgroep Informatica. Computer-assisted production of syntactically parsed textual data (Werkgroep Informatica, Faculty of Theology, Vrije Universiteit, Amsterdam 1996, unveröffentlicht).

<sup>11</sup> Diese Analyseebene wird von den meisten Konkordanz-Programmen zum AT zur Verfügung gestellt, vgl. die in Anm. 2 genannten Programme.

<sup>12</sup> Vgl. E. Talstra (1996) 2: „The process of syntactic parsing is, first, a process of text *segmentation* and the calculation of grammatical and lexical *features* of the segments produced. Second, it is a process of the combination of these segments by establishing their grammatical *relations*.“

<sup>13</sup> E. Talstra (1996) 1: „This research implies both the construction of a text database and the search for a text grammar of classical Hebrew.“

<sup>14</sup> E. Talstra (1996) 3: „The production of more abstract or functional grammatical categories will have to be done after the inventory of linguistic forms has been made extensive und consistent. ... From the study of distributional patterns the production is possible of functional labels for linguistic elements at all levels of grammatical analysis.“ So fungiert z.B. in Hiob 1,1 ein Personalpronomen mit Artikel in attributiver Relation zu einem vorangehenden, ebenfalls determinierten Nomen (הַיְהוָה אֱלֹהֵי יְהוָה) als Demonstrativum. Diese (nach dem Muster Artikel-Nomen-Artikel-Personalpronomen aufgebaute) Nominalphrase fungiert als Subjekt in einem Verbalsatz: ... וְהָיָה הָאֱלֹהִים יְהוָה יֶחֱדָם.



gungsebene. Die in der Datenbank vorliegenden Festlegungen sind nicht veränderbar, es besteht jedoch die Möglichkeit, eine Nutzer-Datenbank mit Alternativen aufzubauen und in die Suchvorgänge einzubeziehen.

## 2. Das Programm

Die derzeit an einem interdisziplinären Forschungsschwerpunkt der Universität Greifswald<sup>15</sup> entwickelte zweite Version von Quest bietet – bei einer gegenüber der ersten Version weitreichenden Vereinfachung und Verbesserung der Bedienung – folgende Möglichkeiten:

Einfache Abfragen, die der Arbeit mit gedruckten Konkordanzen entsprechen, können im Hauptfenster des Programms ausgeführt werden. Als zusätzliche Möglichkeit steht neben der üblichen alphabetischen Wortliste auch eine Sortierung nach Häufigkeit zur Verfügung; beide Listen können auf Teile des AT (z.B. Pentateuch oder Dodekapropheten) oder Teile einzelner Bücher (z.B. Deuteronesaja) sowie auf selbstdefinierte Einheiten (z.B. die Rahmenerzählung des Hiobbuches: Hiob 1,1-2,13; 42,7-17 oder die weisheitliche Poesie des AT: Hiob 3,1-42,6; Prov; Ps 1; 37; 73 ...) beschränkt werden.

Einzelne grammatische Erscheinungen können in einem Abfragedialog definiert (z.B. die wurzel-unabhängige Suche nach den Verbformen der 1. und 2. Person im Hiobbuch oder die Suche nach allen außer den proklitischen Präpositionen im Dodekapropheten) und zu komplexen Abfragen zusammengestellt werden (z.B. die Suche nach absoluten Infinitiven, die mit einer finiten Form gleicher Wurzel verbunden sind, oder nach Nominalphrasen, die ein determiniertes Partizip enthalten). Es besteht die Möglichkeit, Zwischenergebnisse (z.B. die als Subjekt fungierenden Phrasen im Deuteronomium) für eine spätere Bearbeitung zu speichern oder durch zusätzliche Angaben (z.B. Subjekt in Verbalsätzen mit einem Narrativ als Verbform) einzuzugrenzen.<sup>16</sup>

Ergebnisse werden als Fundstellenliste im Hauptfenster ausgegeben. Die einzelnen Fundstellen können in verschiedenen Zusammenhängen (z.B. im Vers- oder Satzkontext, im Kontext einer definierbaren Anzahl von Phrasen vor bzw. nach der gefundenen Phrase) auf dem Bildschirm dargestellt, in ein Textverarbeitungsprogramm übernommen oder direkt gedruckt werden, wahlweise mit zusätzlichen grammatischen Informationen.<sup>17</sup>

---

<sup>15</sup> Der Forschungsschwerpunkt „Computergestützte Philologie und Bilderschließung“ wurde im Jahr 1996 auf Initiative und unter der Leitung von C. Hardmeier an der Universität Greifswald eingerichtet und ist der theologischen Fakultät zugeordnet. Neben Quest werden Programme für andere philologische Forschungsvorhaben (aus der Nordistik, Baltistik, Germanistik) entwickelt.

<sup>16</sup> Viele der genannten Möglichkeiten bestanden bereits in der ersten Version, waren aber aufgrund einer komplexen Abfragesprache nur schwer zugänglich. In der zweiten Version wurde der Datenbestand erheblich ausgeweitet und die Vielfalt der Abfragemöglichkeiten erweitert.

<sup>17</sup> Die Such- und Ausgabemöglichkeiten können hier aus Platzgründen nur auszugsweise genannt werden. Weitere Informationen können bei der am Ende des Artikels genannten Adresse angefordert werden.



### 3. Der Stand der Arbeit

Eine Vorversion des Programms und ein umfangreicher Teil der Datenbasis wurden während der CeBIT 1998 in Hannover vorgestellt, die letzte Testphase (Betatest) beginnt im Oktober 1998. Derzeit werden die Übersetzungen und Wörterbücher (Hebräisch-Deutsch, Hebräisch-Englisch, Griechisch-Englisch) schrittweise in das Programm eingebunden, und die Programmbedienung wird für deutsch-, englisch- und französisch-sprachige Nutzer angepaßt.

Es besteht bis September 1998 die Möglichkeit, zusätzliche, eventuell für spezielle Forschungsvorhaben benötigte Funktionen zu implementieren. Vorschläge dieser Art sind an den Verfasser zu richten und können bei vermutetem Interesse für einen größeren Nutzerkreis in die Endversion übernommen werden.

#### *Anschrift des Autors:*

*W.-D. Syring, Forschungsschwerpunkt Computergestützte Philologie und Bilderschließung, Theologische Fakultät der Universität Greifswald, 17487 Greifswald, E-mail: syring@rz.uni-greifswald.de*