

### Die altägyptischen Sargtexte: Verfahrensschritte eines großen und komplizierten Textcorpus

Schenkel, Wolfgang; Buroh, Knut

Veröffentlichungsversion / Published Version  
Zeitschriftenartikel / journal article

Zur Verfügung gestellt in Kooperation mit / provided in cooperation with:  
GESIS - Leibniz-Institut für Sozialwissenschaften

#### Empfohlene Zitierung / Suggested Citation:

Schenkel, W., & Buroh, K. (1995). Die altägyptischen Sargtexte: Verfahrensschritte eines großen und komplizierten Textcorpus. *Historical Social Research*, 20(3), 144-149. <https://doi.org/10.12759/hsr.20.1995.3.144-149>

#### Nutzungsbedingungen:

Dieser Text wird unter einer CC BY Lizenz (Namensnennung) zur Verfügung gestellt. Nähere Auskünfte zu den CC-Lizenzen finden Sie hier:  
<https://creativecommons.org/licenses/by/4.0/deed.de>

#### Terms of use:

This document is made available under a CC BY Licence (Attribution). For more information see:  
<https://creativecommons.org/licenses/by/4.0>

beitet haben, zur Durchsicht gegeben. Ergeben sich bei der Durchsicht noch sachliche Änderungen, so werden diese wiederum in die zentrale Arbeitsdatei eingetragen. Die korrigierte Datei geht nochmals zum Satz.

Nachbesserungen, die nur die Typographie betreffen, werden erst nach Abschluß aller sachlichen Änderungen ausgeführt. Sie tangieren die zentrale Arbeitsdatei nicht mehr.

Die Ergebnisse der Umstellung haben wir schon 1992 auf der Tagung der Arbeitsgemeinschaft Philosophische Editionen in Berlin vorgetragen. Die meiste Zeit, die ich seither an dem Projekt gearbeitet habe, habe ich damit verbracht die Programme den Wünschen des Verlegers anzupassen, um z. B. Editions-konventionen oder aus Bleisatzzeit herübergerettete Konventionen wieder zu übernehmen. Diese Arbeiten haben nichts mit der editorischen Arbeit zu tun. Sie haben die vorgesehene Datenstruktur weder erweitert noch umgestoßen.

Auch aus dieser Erfahrung heraus plädiere ich für eine Trennung von wissenschaftlicher editorischer Arbeit und technischer Herstellung der Edition. Die Satzaufbereitung kostet viel Zeit und bringt der inhaltlichen Arbeit nichts.

Ein Arbeitsmittel wie TUSTEP, das sowohl die wissenschaftlichen Arbeitsschritte als auch die mehr technischen Arbeitsgänge von Satzherstellung und Aufbereitung für andere Publikationsformen gleichermaßen unterstützt, ohne daß ein Systemwechsel notwendig wird, erleichtert die reibungslose Integration aller Arbeitsschritte in einen Gesamtablauf, trotz der hier erfolgreich praktizierten Arbeitsteilung.

## Die altägyptischen Sargtexte Verfahrensschritte bei der Erschließung eines großen und komplizierten Textcorpus

*Wolfgang Schenkel, Knut Buroh\**

### 1. Das ägyptologische Projekt

Die altägyptischen Sargtexte sind Texte, die zwischen dem Ende des 3. und der Mitte des 2. vorchristlichen Jahrtausends meist auf den Innenwänden von kastenförmigen Särgen betuchter Personen angebracht wurden. Sie sind in der Hauptsache gedacht als Wissensvorrat für den Verstorbenen zum Gebrauch im Jenseits. Dem Ägyptologen sind sie eine wahre Fundgrube für religionshistorische und linguistische Fragen.

---

\* Protokoll des 62. Kolloquiums über die Anwendung der EDV in den Geisteswissenschaften an der Universität Tübingen am 26. November 1994.

Die Texte stehen in einer lebenden Texttradierung und -Weiterentwicklung zwischen den Pyramidentexten aus der Mitte des 3. vorchristlichen Jahrtausends und dem Totenbuch, das seit der Mitte des 2. Jahrtausends v. Chr. die Sargtexte ablöst. Dem Umfang nach sind die Sargtexte die größte der drei Textgruppen. Sie umfassen - soweit ediert - Texte im Gesamtumfang von ca. 140 000 laufenden Wörtern. Die Texte sind - soweit ediert - im Schnitt 3,8 Mal bezeugt, so daß das Textcorpus insgesamt mehr als eine halbe Million laufende Wörter umfaßt. Da die Textzeugen nicht unbeträchtlich voneinander abweichen, ist letztere Zahl die für das Projekt maßgebende.

Für die Verarbeitung im Computer wurden die Texte nicht in der Hieroglyphenzeichenfolge erfaßt, die die Edition bietet - eine solche linguistisch zu interpretieren, wäre hochkompliziert -, sondern in einer linguistischen Transkription, die mit graphematischen Informationen angereichert ist, mithin in zweiter Linie auch die Hieroglyphenzeichenfolge abbildet.

Ursprünglich war Ziel, eine Konkordanz zu den Sargtexten zu erstellen, d. h. die Belege für die einzelnen Lexeme mit ihrem Kotext wiederzugeben. Dieses Ziel wurde zwar nicht aufgegeben, aber einstweilen noch vertagt, da eine mechanische Gewinnung der Kotexte entweder, wenn z. B. als Kotexte einheitlich sinnvolle Textabschnitte wie Sätze oder Teilsätze gewählt würden, diese die Konkordanz zu voluminös machten, oder, wenn man mechanisch kürzere Kotexte ohne Rücksicht auf den Sinnzusammenhang abgrenzen würde, die Konkordanz nur schwer zu benutzen wäre, da ägyptische Wörtformen - in Hieroglyphen oder in Transkription - in weit höherem Maße mehrdeutig sind als etwa deutsche. Statt dessen werden, um auch einmal ein Ergebnis zu haben, Register produziert, lexikalische, morphologische, graphematische. Danach sollen die Texte syntaktisch analysiert werden und die syntaktischen Strukturen bei der Abgrenzung von Kotexten für die Zwecke einer echten Konkordanz genutzt werden

Im übrigen werden übergroße Belegmengen für ein und dasselbe Textelement - Lexeme sind bis gegen 30 000 Mal belegt - mit Hilfe manueller Eingriffe derart reduziert oder strukturiert, daß ein Leser die Aufstellungen auch tatsächlich benutzen kann (statt z. B. 10 000 Belege nachzuschlagen, würde man besser die Edition durchlesen).

## 2. Die philologische Prozedur

Die Aufnahme der Texte (in Transkription) ist so kompliziert und damit fehleranfällig, daß sie nicht in einem Zug durchgeführt werden kann. Im übrigen stand weder von Anfang an noch steht heute fest, was man notiert, und was nicht; die Ansprüche wuchsen und wachsen mit dem Projekt. Anfangs bestand die Textaufnahme vor allem aus der Notation des Wortlauts und textkritischen Zusätzen, später kamen Angaben zu den Graphien hinzu, schließlich soll künftig einmal eine syntaktische Analyse zugeordnet werden - und dergleichen

mehr. Die Erstellung der Texte für die Verarbeitung im Computer erfolgt also zyklisch oder besser: spiralförmig. Die Textaufnahmen werden in Etappen »aufgeschaukelt«. Dabei liefern frühere Etappen oft mehr als den Textbestand selbst: Die mehrfach bezugten Texte wurden derart gewonnen, daß ein - beliebiger - erster Textzeuge vervielfältigt wurde und durch Korrektur in die anderen Textzeugen verwandelt wurde. Oder Für die Zufügung der graphematischen Informationen zu den zuvor hergestellten Transkriptionen des Wortlauts wurde ein graphematisches Wörterbuch produziert, das jede mögliche Graphie *einmal* enthielt, und aus diesem wurden dann die Graphien der einzelnen Wörtformbelege ausgewählt.

Für das »Aufschaukeln« der Textaufnahme ist von wesentlichem Belang, daß zwischen den Zyklen Register erstellt werden können - z. B. sehr früh schon ein lexikalisches Register -, mit deren Hilfe man die Textaufnahmen auf Konsistenz hin überprüfen kann.

### 3. Die TUSTEP-Prozedur

Es gibt stets *eine maßgebende* Fassung des Textcorpus. Diese Fassung bietet die Texteinheiten in *einer* Reihenfolge, z. B. in der Folge des Textes oder in der lexikalischen Folge. Will man die Elemente in anderer Reihenfolge haben, muß diese »Maßgebende Textfassung« des Corpus entsprechend umsortiert werden. Es spricht allerdings nichts dagegen, neben der »Maßgebenden Textfassung« des Corpus in einer bestimmten Sortierung auch nicht-maßgebende Fassungen in anderer Sortierung zu haben. Abstrakt überlegt, wäre die Arbeit mit einer Datenbank praktischer, in der Praxis dagegen ist das gewählte Verfahren ohne Datenbank durchaus zweckmäßig, da man sich im allgemeinen in der Folge der Zyklen sehr lange auf eine einzige Reihenfolge konzentriert. Die »Maßgebende Textfassung« muß drei Bedingungen erfüllen:

- Sie muß am Bildschirm lesbar und auf der Tastatur (bzw. mit der Maus) schreibbar und korrigierbar sein.
- Sie muß nach unterschiedlichen Gesichtspunkten intern sortierbar sein.
- Sie muß als ~~Schnittstelle für~~ **Schnittstelle für** den Druck dienen können, also alle Informationen zur Steuerung des Druckbildes implizit enthalten (zur Umsetzung in ein Druckbild siehe unten § 4).

Wesentliche TUSTEP-Eigenschaften - neben den allgemeinen Qualitäten sind für unsere Zwecke vor allem diese:

- Die Möglichkeit, den Zeichenvorrat durch Zeichenkombinationen zu erweitern, dann aber diese Zeichenkombinationen bei der internen Weiterverarbeitung (Sortierung) auf einfache Weise in Einzelzeichen umcodieren zu können;
- die Möglichkeit, Daten sozusagen **zweidimensional zu organisieren und umzuorganisieren** (1. Dimension: Zeilen, 2. Dimension: »Spalten«, d. h. durch Trennzeichen gekennzeichnete Abschnitte in der Zeile), m. a. W. lange und

komplizierte Datensätze, die intern als Einheit manipuliert werden (Sortierung), am Bildschirm durch Zerlegung in Abschnitte und Aufteilung auf mehrere Zeilen überschaubar zu machen.

Den immensen positiven Qualitäten des Werkzeugs TUSTEP stehen ein paar Nachteile gegenüber, etwa diese:

- Um praktisch arbeiten zu können, muß das Textcorpus in Dateien von je etwa einem Hundertstel zerlegt werden. Andernfalls wäre das Zerstörungsrisiko bei (fehlerhaften) Pauschalkorrekturen zu hoch und die Wartezeit am Bildschirm beim Suchen - und Korrigieren - zu groß. Anderfalls würden auch die Zeilennummern zu groß oder unpraktisch groß. Im übrigen müssen, wenn durch eine Korrektur die Sortierfolge geändert würde, u. U. Daten aus einem Hundertstel in ein anderes manuell verschoben werden.
- Daten, die eliminiert werden (z. B. nur ergänzte Wörtformen) müssen manuell in Hilfsdateien ausgegliedert werden, da sie nicht einfach unsichtbar gemacht werden können und damit den Überblick über das Material am Bildschirm behindern würden.
- Die Transkriptionen haben eine Struktur, die mit TUSTEP-Mitteln allein nicht in einen Sortierschlüssel umgesetzt werden kann (Zerlegung in diverse Hauptinformationen und vielfältige Indizes hierzu); für solche Hauptinformationen und vor allem Indizes werden je eigene Sortieralphabete benötigt, die mit den zugelassenen drei TUSTEP-Sortieralphabeten nur durch trickreiche, schwer durchschaubare Überlagerung verschiedener Alphabete realisiert werden können.

Die genannten Problem-Punkte sind für die Zwecke des Projekts praktisch gelöst. Es steht zu vermuten, daß die Probleme weitestgehend mit den spezifischen Besonderheiten der benutzten Transkription und mit der Eigenart des Textcorpus zusammenhängen, Speziallösungen also gerechtfertigt sind.

#### 4. Beispiel einer Druckausgabe: der Druck der lexikalischen Konkordanz

Da ein wesentliches Ziel des Projekts ist, die wichtigsten Ergebnisse in der klassischen Form eines gedruckten Buches zu veröffentlichen, das nach wie vor für den Ägyptologen das meistgebrauchte Arbeitsmittel darstellt, war es von einem gewissen Grad der »Aufschauklung« und Sortierung der Daten an sinnvoll und erwünscht, diese schon als Teil- oder Zwischenergebnis in einer der endgültigen Druckgestalt nahekommenden, wenn nicht sogar identischen Form zu produzieren. Als Schnittstelle zu diesem Produktions schritt dient die oben § 3 eingeführte »Maßgebende Textfassung«, die die Daten in ihrer jeweils aktuellsten Form enthält

Die einzelnen Sätze der »Maßgebenden Textfassung« werden hierzu durch eine Kaskade von TUSTEP-Kopiere- und TUSTEP-Satz-Programmen geschickt in deren Verlauf die Daten verglichen und reduziert, interne Codes in

explizite TUSTEP-Satz-Codes umgewandelt sowie für den Satz und für das Druckbild nötige Ergänzungen vorgenommen werden. Am Ende des Durchlaufs steht ein TUSTEP-Satz-Programm, das dann die endgültige Ausgabe, letztlich auf Papier besorgt.

Die Schritte im einzelnen:

- Ein VORPROGRAMM (TUSTEP-Kopiere, bestehend aus 42 Durchgängen mit 282 Parameterzeilen) eliminiert aus jedem Datensatz der »Maßgebenden Textfassung« die Datenteile, die für das in diesem Programmablauf angestrebte Ergebnis (Druckseite) nicht auszuwerten sind. Des weiteren werden hier implizit eindeutige Grenzmarkierungen zwischen den hierarchischen Teilen eines Datensatzes für die weitere Abarbeitung in von TUSTEP leichter zu erkennende explizite Markierungen ausgetauscht. Als letztes werden hier sog. »Flags« gesetzt, d. h. es werden Datensatzteile aus ihrem festen Platz in der »Maßgebenden Textfassung« ausgeschnitten, und an einer nur für die Druckfassung benötigten anderen Stelle durch Markierungen ersetzt.
- Das PROGRAMM (TUSTEP-Kopiere, bestehend aus 96 Durchgängen mit 2058 Parameterzeilen) erledigt als zweite Komponente die Hauptarbeit des Druck-Paketes. Hier werden die einzelnen Komponenten der Datensätze der »Maßgebenden Textfassung« hierarchisch miteinander verglichen und je nach Ergebnis des Vergleichs im Programm weitergereicht. Da die Eingabedaten eine alphabetisch sortierte lexikalische Liste darstellen, bei der jeder Datensatz die vollständige grammatische und orthographische Beschreibung eines Wortes der Sargtexte enthält, werden hier die Angaben zu Wortwurzel, Wortform, Orthographie etc. verglichen, bei Wiederholung eliminiert und dem vorangegangenen Datensatz hierarchisch zugeordnet (Parameter av-ev, ak-ek, axx-exx).
- In einem NACHPROGRAMM (TUSTEP-Kopiere, bestehend aus 23 Durchgängen mit 490 Parameterzeilen) werden in den vorherigen Teilprogrammen zurückgestellte oder noch nicht lösbare Sonderfälle (wie z. B. Komposita) analog zu PROGRAMM verarbeitet.
- Den Abschluß der TUSTEP-Kopiere Reihe bildet ein SATZVORBEREITUNGS-Programm (TUSTEP-Kopiere, bestehend aus 28 Durchgängen mit 302 Parameterzeilen), das im wesentlichen alle verbliebenen projekt-spezifischen Codes in TUSTEP-Satz-Codes austauscht und z. B. lebende Kolummentitel aus den aktuellen Daten ermittelt und ergänzt.
- Den Druck schließlich besorgt ein einfaches TUSTEP-Satzprogramm, in dem nur noch datenunabhängige Pauschalanweisungen gesetzt werden, z. B. Schriftgrößen oder Überschriftengestaltung.

Die relativ hohe Zahl von Durchgängen und Parameterzeilen läßt die Frage aufkommen, ob eine Zusammenfassung von Verarbeitungsschritten in TUSTEP-Kopiere nicht möglich und wünschenswert wäre. Zwar ist dies theoretisch ohne weiteres denkbar, in der Praxis ginge dies aber auf Kosten der Nachvollziehbarkeit der einzelnen Programmschritte und ihrer Wirkungen im

komplexen Daten- und Codebestand. Durch die Verteilung auf vier, in ihren Aufgaben klar voneinander getrennte Teilprogramme, wird die Fehlersuche und Fehlerbeseitigung wesentlich vereinfacht. Manuelle »Schönung« der TUSTEP-Satz-Ausgabedatei ist nicht vorgesehen. Datenfehler, die in den Zwischenausdrucken festgestellt werden, werden in der »Maßgebenden Textfassung«, also in den Quelldaten korrigiert, Programmfehler im Programm abgefangen. Wie bereits erwähnt, ist es unter den gegenwärtigen Bedingungen jederzeit möglich, einen Ausdruck zu erstellen, der der endgültigen zu druckenden Form weitestgehend entspricht. Im Wesentlichen hängt es vom angestrebten Perfektionsgrad der »Maßgebenden Textfassung« ab, wann der Druck erfolgen kann.