

# STANDARDISIERUNG VON TRANSLITERATIONEN

Peter Damerow — Berlin

## 1. Prinzipien

Nachstehend werden für die weitere Diskussion die Ergebnisse eines Workshops zur Vorbereitung einer elektronischen Erfassung der mathematischen Keilschrifttexte kurz zusammengestellt. Der Workshop fand vom 7. bis zum 9. September 1994 im Max-Planck-Institut für Wissenschaftsgeschichte statt. An dem Workshop nahmen teil (in alphabetischer Reihenfolge) Peter Damerow (Max-Planck-Institut für Wissenschaftsgeschichte, Berlin); Robert Englund (Freie Universität Berlin); Jöran Friberg (Universität Göteborg, Schweden); Jens Høyrup (Universität Roskilde, Dänemark); Jörg Kantel (Max-Planck-Institut für Wissenschaftsgeschichte, Berlin); Stefan Maul (Universität Heidelberg); Gebhard Selz (Universität Freiburg); Marcel Sigrist (École Biblique, Jerusalem, Israel).

Der Workshop hatte zum Ziel, Mindestanforderungen an die Transliteration zu formulieren, ohne die technische Realisierung bereits vorwegzunehmen. Die Ergebnisse sollen die Grundlage für die Erarbeitung einer technischen Konzeption bilden. Dem formulierten Vorschlag liegen nachstehende Prinzipien zugrunde.

### 1.1. *Hard- und Softwareunabhängigkeit des Datenformats*

Die Transliterationen sollen in einem Datenformat gespeichert werden, das es ermöglicht, sie auf einfache Weise zwischen verschiedenen Rechnersystemen auszutauschen und die jeweils verfügbare Standardsoftware für die Verarbeitung von strukturiertem Text (Editoren, Datenbanken, Indizierungsprogramme etc.) zu verwenden. Das Datenformat soll sich zu diesem Zweck in ein Format konvertieren lassen, das nur die Zeichen ASCII 32 bis 127 enthält. Die Transliterationen sollen nur ein Minimum an Markup enthalten, so daß sie in einem einfachen Editor noch leicht lesbar sind und notfalls ohne spezielle Softwareunterstützung direkt verarbeitet werden können. Aus diesen Gründen sollen die Transliterationen nur den Minimalanforderungen genügen, die unter philologischen Gesichtspunkten erforderlich sind.

### 1.2. *Abbildungsgenauigkeit der Transliteration*

Die Regeln der Transliteration in dem zu definierenden Datenformat sollen so beschaffen sein, daß sich aus der Transliteration die tatsächlich auf der Tafel befindlichen Zeichen eindeutig rekonstruieren lassen.

### 1.3. *Konvertierbarkeit in Standardformate*

Um das Datenformat einfach zu halten, wird es notwendig sein, für die verschiedenen Computersysteme spezielle Formate zu definieren, in denen die Transliterationen einfach verarbeitet werden können. Durch Konvertierungsprogramme soll jedoch sichergestellt werden, daß solche speziellen Formate jederzeit sowohl in ein universelles Austauschformat (möglichst ein SGML-Format) als auch in Textverarbeitungsformate für eine hochwertige Druckausgabe (RTF, TeX etc.) konvertiert werden können.

## 2. **Standardisierung der Zeichen**

Die bei der Transliteration zulässigen Zeichen werden genau definiert. Sie werden vier verschiedene Gruppen zugeordnet:

- Zeichen, die Wörter trennen;
- Zeichen, die Keilschriftzeichen, aber keine Wörter trennen;
- Wortbildende Zeichen;
- Wortbildende, aber bei der Verarbeitung zu ignorierende Zeichen.

Da die weitere Verarbeitung der Transliterationen und insbesondere die Entwicklung spezieller Software für die Unterstützung der Verarbeitung auf dieser Unterscheidung beruhen soll, ist die eindeutige Verwendung der Zeichen im Sinne dieser Unterscheidung strikt einzuhalten.

### 2.1. *Worttrennende Zeichen*

- Leerraum (ASCII 32)
- Zeilenende und Linefeed (ASCII 13 und ASCII 12)
- Tabulator (ASCII 9)

(Sinnvollerweise sollte diese Gruppe vielleicht auf weitere ASCII-Zeichen zwischen ASCII 0 und ASCII 31 ausgedehnt werden)

### 2.2. *Zeichen, die innerhalb der Wörter Keilschriftzeichen trennen*

- Bindestrich: —
- Punkt: .
- Pluszeichen: +
- Malzeichen: × (eigenes Zeichen, nicht der Buchstabe x)
- Doppelpunkt: :
- Zeichen für den Anfang eines Determinativs/Komplements
- Zeichen für das Ende eines Determinativs/Komplements

### 2.3. Wortbildende Zeichen

- a bis z (x ist reserviert für unlesbare Zeichen);
- A bis Z;
- 0 bis 9;
- Zeichen für den Index x (eigenes Zeichen, nicht der Buchstabe x);
- Unbedingt notwendige Sonderzeichen: Š š ʹ T t S s;
- Optionale Sonderzeichen (Liste erweiterbar): ˘ ħ Ĝ ĝ ř â ê î û ā ē ī ū;
- s-akut (groß und klein);
- zd und mb;
- Nichttrennendes Leerzeichen für die Kennzeichnung abgebrochener Wort- und Textteile;
- / für die Schreibung von Brüchen;
- ( und ) für die Identifizierung von Zeichen mit Index x sowie für Kommentare in der Transliteration.

### 2.4. Wortbildende, aber zu ignorierende Zeichen

- [ und ] für Ergänzungen des Textes;
- ˘ und ˙ für Beschädigungen von Zeichen;
- < und > für Weglassungen im Originaltext;
- << und >> für zu Tilgendes im Originaltext;
- ein Sonderzeichen für den Beginn von akkadischem Text (Kursivschrift);
- ein Sonderzeichen für das Ende von akkadischem Text (Kursivschrift);
- ? für fragliche Zeichen (direkt angefügt);
- ! zur Betonung der Richtigkeit (direkt angefügt);
- <sup>d</sup> zur Kennzeichnung kollationierter Zeichen);
- ein Sonderzeichen für die Fortführung eines Wortes in einer neuen Zeile der Tafel.

## 3. Standardisierung der Transliteration

### 3.1. Schreibung der Keilschriftzeichen

**Diakritische Zeichen** werden durch die Indexziffern 2 und 3 ersetzt.

**Indexziffern** werden als normale Ziffern geschrieben. Die Indexziffern werden später konvertiert. Dazu muß bei der Transliteration darauf geachtet werden, daß die Regel strikt eingehalten wird, daß Indexziffern stets auf einen Buchstaben folgen und stets von einem Zeichen gefolgt werden, das Keilschriftzeichen trennt.

**Index x** wird durch ein besonderes Zeichen dargestellt. Die Zeichenform wird in runden Klammern direkt angefügt; dabei dürfen keine trennenden Zeichen verwendet werden.

**Determinative und Komplemente** werden wie alle anderen Zeichen geschrieben, aber am Anfang und Ende durch spezielle Zeichen gekennzeichnet.

**Numerische Notierungen** werden vorerst nicht vereinheitlicht. Sie müssen aber stets abbildungsgenau die geschriebenen Zeichen zu rekonstruieren gestatten. Die Zeichenform wird in runden Klammern direkt angefügt. Zulässig sind auch standardisierte, aber eindeutige Kurzschreibweisen für gängige numerische oder metrologische Systeme.

### 3.2. *Standardisierung des Textformats*

Jede **Zeile** der Tafel besteht aus einem kurzen, standardisierten Zeilecode zur Kennzeichnung des Ortes auf der Tafel und der durch ein Tabulatorzeichen abgetrennten Transliteration.

Der **Zeilecode** einer Zeile besteht aus einem Code zur Identifizierung der *Tafel*, einem Code für die *Seite* (Vorderseite, Rückseite, Rand, Siegel etc.); einer zweistelligen Nummer der *Spalte* und einer zweistelligen Nummer der *Zeile*. Die Codes werden so gewählt, daß bei einer ASCII-Sortierung alle Zeilen in eine korrekte Reihenfolge gebracht werden.

**Kurze Kommentare**, z.B.: (Rasur), können in die Transliteration eingefügt werden. Sie werden generell in runde Klammern gesetzt, ohne worttrennende Zeichen geschrieben und stets durch Spacecharacters vom restlichen Text getrennt.

Auf der Tafel hervorgehobene **Trennstriche** oder **Leerzeilen** werden als Kommentare geschrieben, z.B.: (3-Leerzeilen). Ihr Zeilecode wird so definiert, daß sie durch ASCII-Sortierung korrekt einsortiert werden.

Spezielle **Wortklassen** können gekennzeichnet werden. Dies geschieht durch den Wörtern vorgesetzte, in runde Klammern ohne Trennzeichen geschriebene Codes, die mit einem zeichentrennenden Zeichen (z.B.<sup>s</sup> dem Bindestrich) mit dem Wort verbunden werden.

Die Klammern für **Beschädigungen** etc. werden stets paarweise verwendet und nur vor oder hinter, nie in ein Keilschriftzeichen geschrieben, bei numerischen Zeichen aber auch um jede Ziffer.

### 3.3. *Vereinheitlichung der Transliterationen*

Während des Scannens der vorhandenen Umschriften werden bereits **Formalkorrekturen** vorgenommen, insbesondere die Beseitigung der diakritischen Zeichen, die Vereinheitlichung der Kennzeichnung der Beschädigungen und die Umsetzung aller nach diesen Richtlinien nicht zulässigen Zeichen in ihre Entsprechungen (Sonderzeichen, Fragezeichen, Klammerformen etc.).

Die Vereinheitlichung der Umschriften nach **philologischen Kriterien** geschieht erst nachträglich.

Alle nach philologischen Gesichtspunkten vorgenommenen Änderungen werden als **Vereinheitslichungsregeln** explizit formuliert, damit spätere Bearbeiter die Arbeit konsistent fortführen können.

Die Zeichennamen der **nicht-lesbaren Zeichen** werden in Großbuchstaben geschrieben und nach Borger/Ellermeier vereinheitlicht.

Die **Übersetzungen** werden eingescannt und mit dem Zeilecode der Transliteration versehen.